Molecular Genetics and the Conservation of Plants: Two Case Studies

A thesis submitted to the University of Glasgow for the degree of Doctor of Philosophy

Christina T. Oliver

Division of Environmental and Evolutionary Biology

May 2004

MISSING PAGES ARE UNAVAILABLE

Declaration

I hereby declare that this thesis is composed of work carried out by myself unless otherwise acknowledged and cited and that this thesis is of my own composition. Research towards this thesis was carried out between January 2000 to January 2003. No part of this dissertation has been previously submitted for any other degree.

Abstract

The Convention on Biodiversity (CBD) signed at the 1992 Earth summit in Rio formally recognized biodiversity at the habitat, species, and genetic levels. For species and habitat biodiversity there is a well-established set of frameworks under which conservation programmes are constructed and delivered. From a genetic biodiversity perspective, however, there is no clear consensus on how best it should be measured, or how conservation programmes should be implemented.

The major reasons for conserving intra-specific genetic biodiversity can be summed up under two inter-related themes, (1) Protecting a broad spectrum of genetic biodiversity, and (2) Maintaining evolutionary fitness and adaptive variation. This thesis takes a case-study approach and explores the issues surrounding these themes for conservation strategies in two angiosperm species: *Saxifraga hirculus* and *Lathyrus japonicus*.

1) Protecting a broad spectrum of genetic biodiversity: This section of the thesis considered the evidence for major intra-specific genetic races in *Saxifraga hirculus* and the spatial distribution of its genetic biodiversity. Variation in *Saxifraga hirculus* chloroplast DNA was assessed in order to gain information on the biogeography of the British populations in the context of the wider European gene pool, and also to compare this with populations from Alaska and Colorado. In a European context, British populations have a high level of chloroplast diversity (three haplotypes) and contain a highly divergent lineage that was previously unsuspected. Seven haplotypes were found in total from 17 populations in Europe with marked inter-population differentiation (F_{ST} = 0.92). Higher diversity and lower population differentiation was detected in Alaska (33 haplotypes / 12 populations; F_{ST} = 0.46). Since most populations in Europe had unique haplotypes it is not possible to track migration routes or pinpoint refugia for the European populations, but the much higher diversity in Alaska compared to Europe indicates that the Beringia region may have acted as a refugium for this species throughout the Pleistocene. This highlights the importance of Alaska for the conservation of intra-specific genetic biodiversity in this species.

(2) Evolutionary fitness and adaptive variation: To assess the relationship between population size, genetic variation, morphological variation and fitness, genetic studies were undertaken on populations of *Lathyrus japonicus*. Eleven populations of *L. japonicus* were examined for variation using nine microsatellite loci. The populations show genetic isolation by distance across the distribution of the species in Britain, although isolation by distance breaks down when only the range centre populations are considered. There was no relationship between population size or isolation and genetic variation, with some small and/or isolated populations having high diversity, and large and/or range centre populations having low diversity. There was, however, a significant difference in the inbreeding coefficient of adult versus seedling plants. The heterozygosity of adult plants sampled in the field was significantly higher than seedlings grown in cultivation, indicating a survival advantage for heterozygotes.

Significant differences were found between populations for seed weight, number of seeds per pod, number of pods per cluster, and leaf shape of *L. japonicus* individuals in the field. For seedlings grown in common conditions significant differences were found in leaf shape, pigmentation, and dry weight after two season's growth. Morphological and genetic differentiation were well matched in this species, and gave similar signals. Seedlings from Carnoustie (Scotland) grew much more vigorously in cultivation in Edinburgh than seedlings sourced from English populations, indicating local adaptation. However no significant relationship was found between any fitness associated traits or morphological variation with genetic variation, in spite of the heterozygote advantage revealed by the genetic data.

The results from both research themes are discussed highlighting the difficulties in equating patterns of genetic marker variation to traits likely to be of evolutionary and ecological relevance.

Acknowledgements

To Supervisors Dr. Pete Hollingsworth, Chris Sydes, Dr. Richard Griffiths.

To Dr. Pete Hollingsworth for help setting up and planning this PhD project, and for helping to keep it on track.

To Dr. Pete Hollingsworth, Dr. Richard Gornall, Chris Sydes for editing and suggestions towards papers.

To Dr. Michelle Hollingsworth and Alex Ponge for support in the laboratories, and to Dr. Jane Squirrell for help and training for microsatellite development.

To Pete Brownless, Ross Irvine, and all the staff at the nursery for allowing space for, and looking after, 400 *Lathyrus japonicus* seedlings.

To Dr. Richard Ennos for statistical advice.

To Dr. David Welsh for help with *S. hirculus* fieldwork.

To Duncan Reddish for computer crisis support.

To the RBGE Library staff for help finding some obscure publications, and tips for working the copy machine.

To SNH for funding towards research costs, and the Louise Home Award for funding towards field work in Colorado, and everyone at RBGE who provided botanical illustration work.

To everyone at the RBGE for contributing to a pleasant and interesting working environment.

Table of contents

Table of figures

Table of tables

Chapter 1. Molecular genetics and the conservation of plants

1.1 Introduction

The 1992 Earth Summit in Rio led to the production of the Convention on Biodiversity (CBD) which was signed by 150 countries (www.biodiv.org). The CBD has subsequently had a major international impact on conservation as it provides a formalisation of global conservation efforts and has triggered a series of national and international conservation programmes and strategies, each with specified targets.

The CBD explicitly recognizes biodiversity at three different levels: Biodiversity at the habitat level, biodiversity at the species level, and biodiversity at the genetic level. This recognition of biodiversity at the genetic level as an integral component of biodiversity conservation programmes represented a major shift in conservation strategies, as previously conservation programmes rarely recognized this type of biodiversity.
Although the importance of biodiversity at the genetic level has now been established, conservation programmes that make efficient use of available resources to protect biodiversity at the genetic level are still in the minority. This is reflected in the documentation for many major conservation initiatives, which have genetic biodiversity described in their introductory strategic overviews, but lack any form of implementation measures for conservation action (HMG, 1995).

A contributing factor to this lack of clarity is confusion over the use of the term 'conservation genetics' which is often viewed as synonymous with 'the conservation of genetic biodiversity'. This is unfortunate, however, as Conservation Genetics is a broader subject which relates to the use of genetic data to contribute towards conservation programmes. This is encompasses several different research areas.

1.2 Conservation genetics

1.2.1 Conservation genetics: Taxonomy

Genetic markers can be used as tools to assist with the identification of species in taxonomically complex groups and also serve to identify cryptic species (Soltis & Gitzendanner, 1999; Hollingsworth, 2003). This enables basic decisions to be made regarding which species exist in the first place and where they occur - an essential prerequisite to knowing whether species are endangered or not.

This approach involves some level of species delimitation work and diagnostics. By using multiple genetic markers, the genetic integrity of morphologically defined units can be tested, and the resulting data can be used to contribute towards the decision as to which taxon/taxa a group of populations is best ascribed to. As costs of high throughput genetic analyses continue to fall, there is the opportunity to undertake widespread screening of populations for the presence of various diagnostic markers to allow rapid assignation of these samples to the defined and delimited taxa (Blaxter & Floyd, 2003; Ronquist & Gardenfors, 2003). Once the range and abundance of a given species has been clarified, threats can be identified and appropriate conservation strategies employed.

While this approach uses genetic markers, and is certainly well within the field of 'conservation genetics' it is essentially just a modern form of taxonomy, and is tackling conservation biology at the 'species level' of biodiversity, not the genetic level. The approach is just providing a new type of data to address long running taxonomic and evolutionary biology questions.

1.2.2 Conservation genetics: Reproductive biology

Successful reproduction and dispersal is essential for the survival of a species, and yet for many rare species little is known about their reproductive biology, and more importantly, what limits their successful reproduction and dispersal (e.g. Forrest *et al.*, 2004). Genetic

markers are thus often used in conservation programmes to increase the understanding of reproductive ecology (Frankham *et al.*, 2002).

In this approach, genetic markers are used to screen many individuals from a population to gain insights into their reproductive biology beyond what would be possible from field observations alone. Genetic markers can reveal a record of rare reproductive events which can be very important for a species, and which the chances of observing in the field are notoriously small. Likewise they can be used to determine whether reproduction is primarily sexual or asexual in species that can undertake both of these modes of reproduction. Also the relative proportions and importance of out-crossing events can be difficult, if not impossible, to determine from field observations alone. Thus key aspects of reproductive biology of a species may remain unknown unless genetic markers are employed. Genetic markers are ideally suited to these questions and are hence often employed in conservation programmes where some insights into reproductive biology are required for species management. However, it should again be stressed that the scientific driving force behind many of these investigations is to gain an insight into how the species is functioning and reproducing, rather that being driven by genetic biodiversity issues per se. As such, these types of reproductive ecology studies fall more within the discipline of managing biodiversity at the species-level, rather than them being studies concerning the conservation of biodiversity at the genetic level.

1.2.3 Conservation genetics: Conserving and managing genetic biodiversity

This strand of conservation genetics research is concerned with assessing the amounts and partitioning of genetic biodiversity and factors resulting in its loss, maintenance or enhancement. It is this subdivision of conservation genetics (rather than the whole discipline of conservation genetics) that is primarily relevant for the conservation of biodiversity at the genetic level, and it is this topic that is the subject of this thesis.

The rational underlying this branch of conservation is the recognition that species are not homogeneous panmictic units with even levels of genetic diversity, evenly distributed across their ranges. Species instead consist of a series of populations, which in turn may belong to local networks of populations that experience more genetic inter-change with each other, than with other population networks in the species. Individual populations can be large or small, genetically variable or depauperate. They can be geographically close to other populations, or they can be isolated. And of course the interplay of these factors is important, as being close to a small, genetically depauperate population, may be different to being close to a large genetically variable population, in terms of the likely influx of migrants and genes.

Post-Rio, post-CBD, there is now a clear recognition of the conservation importance of the sometimes 'hidden' level of biodiversity at the genetic level, and that this diversity is unlikely to be perfectly evenly distributed across a species range.

1.3 The importance of genetic biodiversity

The major reasons for conserving intra-specific genetic biodiversity can be summed up under two inter-related broad themes, which will in turn be the two major themes of this thesis: (1) Protecting a broad spectrum of genetic biodiversity (Regional conservation - protecting the most critical genetic biodiversity), and (2) Evolutionary fitness and adaptive variation. Each of these themes will be explored in turn below.

1.3.1 THEME 1: Protecting a broad spectrum of genetic biodiversity

Conservation programmes are typically constructed and delivered at the national level. However, it is useful to consider the broader genetic/geographical structure of species and

to consider them in an international perspective. Most plant species have gone through a series of expansions and contractions of their ranges as temperatures rose and fell through the ice ages. This means populations with very different histories occur in different parts of a species' range, and can result in different levels of genetic diversity and adaptive potential present in a species in different regions.

One of the major determinants of the distribution of range-wide genetic variation in plant species is major fluctuations in the earth's temperature. During periods of glaciation, the range of many plant species contract, and populations survive the ice ages in smaller refugial areas (Hewitt, 1996). As the climate warms, plants expand their distributions and recolonise the ranges they occupy during inter-glacial periods. Such massive changes in the distribution of species in turn impacts on their population genetic structure. The size and location of refugial populations will be an important determinant of genetic structure, as will be the mixing of lineages from different refugia and the speed and mode of colonization (e.g. a gradual advancing front versus a stepping stone colonization via long distance dispersal; Ferris *et al.*, 1999). Understanding the historical processes that have given rise to the present day distributions of plant species is important as it can serve as a framework in which to interpret the current observed distributions of genetic biodiversity. In addition these investigations into the historical movement of plant species in the face of environmental change may help with forward projections of the behaviour of species under future climate change.

Range edge populations are often small and isolated. When species are at the edge of their range in a particular country, they often are considered threatened and receive conservation attention, although in other regions the species may be widespread. Many range edge populations are the result of long distance dispersal, and have as a result experienced a severe bottleneck. Even when range edge populations are relicts from a time when the species was locally widespread, a prolonged small populations size is likely to have led to decreased genetic variation due to genetic drift. So for many species range edge populations contain only a small percent of the genetic diversity of the species, and are far less important to the survival of the species than range centre populations. In a species'

range centre, however, it can be quite widespread, and so receive little conservation attention. This sets up a conservation dilemma. What is the most important genetic diversity to conserve, and can generalizations be made about where it will be located?

1.3.1.1 Importance of refugial and range centre populations

Glacial refugia are important when considering the protection of range-wide genetic diversity. These refugia are areas that may contain a high proportion of unique divergent lineages, and often a large proportion of the genetic diversity of a species. Several studies have detected higher levels of genetic diversity in regions which are known to have been the sites of glacial refugia based on fossil evidence (Bennett *et al.*, 1991; Huntley & Birks, 1983) and climatic data (Bradley, 1985), with populations outside of refugial areas showing declining sub-sets of this variation (Hewitt, 1996; Ferris *et al.*, 1999). For species for which fossil evidence is not available, the concentration of high levels of genetic diversity in specific geographical areas is now sometimes used as evidence of regions being potential sites of glacial refugia. However areas of high genetic diversity can occur in 'melting pot' regions (Petit *et al.*, 2003) where migrants from several different refugia meet (Petit *et al.*, 2003; Abbott & Brochmann, 2003), so high diversity alone, without the presence of unique divergent lineages may not be a reliable indication of refugial areas.

From the perspective of conservation in the long-term, refugial areas are important. As the earth experiences future climate change in the form of glacial cycles, populations in these areas may again become key to the species' survival. If the species are lost from refugial areas, there will be no source populations to ensure inter-glacial recolonisation. This argument has been used to explain the current natural absence of hemlock (*Tsuga*) and spruce (*Picea*) in the UK, despite evidence of their natural colonization during previous inter-glacial periods (Ferris *et al.*, 1999).

1.3.1.2 Importance of 'melting pot' populations

It seems unlikely that populations are able to migrate back into ecologically filled (climax forest) refugial areas as climates cool (Ferris *et al.*, 1999), so the genetically diverse 'melting pot' populations are less important than refugial populations at a longer time scale. However knowledge of melting pot areas can also be useful to conservation for example, as diverse sources for sampling for *ex-situ* conservation.

1.3.1.3 Importance of local and divergent populations

The same forces of isolation and genetic drift that reduce the genetic diversity of many range edge populations, can leave some of these populations genetically and morphologically distinct from the range centre, and these populations can be candidates for incipient speciation as selection takes place on new forms. The range edge populations that occur in different ecological environments from the range centre can also acquire differing suites of adaptive genes through selection. These populations can thus be argued to have a conservation value in their own right (Lesica & Allendorf, 1995).

1.3.1.4 Molecular approaches for studying broad scale genetic structure (plant phylogeography)

To investigate broad scale genetic biodiversity within species and gain insights into historical migration patterns, genetic markers that are efficient at detecting population structure are required. These markers should have an appropriate mutation rate for tracking historic relationships in the context of modern geographic distributions and, ideally, be phylogenetically orderable (Avise, 2000).

The organelle genomes are the most popular source of genetic markers for phylogeographic studies. Organelle markers are effectively haploid (Birky *et al.*, 1989,) and so have half the effective population size of nuclear genes, making them more susceptible to drift (Ennos *et al.*, 1999; Schall *et al.*, 1998). An increased rate of genetic drift can result in population structure where nuclear markers might show no differentiation. Selective sweeps, the tendency of a strong selective advantage for a mutation in one region of the organelle genome to cause the entire genome to become fixed in a population, also contribute to higher population differentiation. Also, because organelle markers are non-recombinant, relationships between haplotypes are not obscured by the movement of genetic material from other sources.

In animals the rapid rate of nucleotide substitution in mitochondrial DNA (mtDNA) has led to powerful insights into patterns of range-wide genetic structure (Avise, 2000). Synonymous mutation rates are five to ten times higher in animals mitochondrial genomes than nuclear genomes (Wolfe *et al.*, 1989) and this allows for the development of geographic/phylogenetic structure even over relatively short periods of evolutionary time. In plants, however, the mutation rate of mtDNA is three to five times slower than cpDNA which mutates at half to one third the rate of the nuclear genome (the nuclear genome mutates at an equivalent rate in plants and animals) (Wolfe *et al.*, 1987). However different regions of the mitochondrial genome have differences in mutation rate, with the control region having a higher rate than other regions. Another trait of plant mitochondrial DNA that hampers its utility as a marker for phylogeographic studies is frequent intra-molecular recombination (Ennos *et al.*, 1999). Because of these characteristics of plant mtDNA, cpDNA is the marker of choice for phylogenetic studies in plants.

Plant cpDNA has a slow rate of mutation, one half to one third the rate of the nuclear genome (Wolfe *et al.*, 1987), and because of this slow rate of mutation, most variation is ancient, most likely from before the quaternary cycles of glaciations (Ferris *et al.*, 1999), and this reduces the relevance of phylogenetic approaches to plant phylogeography. Instead the frequency and distributions of chloroplast variants are often considered under the infinite allele model in which haplotypes are either considered the same or different,

rather than the degree of difference being qualified (Ennos *et al.*, 1999). Inheritance of the chloroplast is usually maternal in angiosperms (in conifers cpDNA is paternally inherited, while the mitochondrial DNA is maternally inherited). Because of this inheritance pattern cpDNA is theoretically moved by seed only, and can reveal historic distribution patterns that are not affected by more recent gene flow through pollen. CpDNA has been found to show more population structure than isozymes in many cases where both markers were studied in a single species of forest tree (Newton *et al.*, 1999), and this matches the theoretical expectations for a genome dispersed by seed rather than by pollen (Ennos, 1994). The combination of drift differentiating populations isolated in refugia, and low mutation rates not obscuring ancestral types, make this genome ideal for phylogeograhic studies.

There are, however, some disadvantages to organelle markers. Firstly the non-recombinant uni-parentally inherited nature of the markers means there is no replication possible across loci. All chloroplast genes are inherited *en masse* and hence data from the entire genome should be treated as data from a single locus (Ennos *et al.*, 1999). Secondly, the data is derived from the maternal lineage, and there is the potential that one is literally only getting 'half the story', and as a further complication the chloroplast can, at a very low rate, also be inherited through pollen (Wang *et al.*, 2004). Thirdly, the slow mutation rates sometimes make it difficult to detect any chloroplast variation at all making investigations into species history extremely difficult (Provan *et al.*, 2001). Finally, chloroplast molecules are known to move between species by chloroplast capture (perhaps most appropriately termed pollen swamping) if there are several generations of uni-directional hybridization (Potts & Read, 1998). It may thus be the case that what appears to be a potential hotspot of diversity for a given species, actually just reflects an area where it has hybridized with another species, and what is being seen is the product of hybridization rather than a rich resource of intra-specific biodiversity. However, despite these limitations to the use of cpDNA, it remains the mainstream approach for the study of plant phylogeography simply because there are no other technically simple alternatives. In the long term, large amounts of sequence data from the nuclear genome may prove informative, but the costs and time required to characterize,

isolate and sequence numerous single copy (and often heterozygous) genomic regions is currently too great to prevent the widespread application of this approach.

1.3.1.5 Generalizations that have emerged from phylogeographic studies to date

The largest body of phylogeographic studies using chloroplast DNA involves European trees, and their refugial areas and migration routes are now becoming fairly well understood. Species studied include *Quercus* spp. (Ferris *et al.*, 1993, 1995, 1998; Dumolin-Lapegue *et al.*, 1997; Petit *et al.*, 1993), *Alnus glutinosa* (King & Ferris, 1998), *Fagus sylvatica* (Demesure *et al.*, 1996), *Calluna vulgaris* (Rendell & Ennos, 2002), *Ilex aquifolium* (Rendell & Ennos, 2003), *Hedra* spp. (Grivet & Petit, 2002), *Prunus avium* (Mohanty *et al.*, 2001), *Prunus spinosa* (Mohanty *et al.*, 2002), *Sorbus aucuparia* (Raspe *et al.*, 2000), and *Salix caprea* (Palme *et al.*, 2003). A comparative genetic survey of 22 woody species from the same 25 European forests, showed the emergence of some common patterns. For these species, levels of genetic diversity were generally lower in the north of Europe and this was considered to be due to progressive population bottlenecks as part of the colonization process (Petit *et al.*, 2003). The major refugial regions identified on the basis of fossil data correspond well with the genetic data, and suggest that the most important refugial areas were the Iberia Peninsula, Italy, Corsica and the Balkans (Petit *et al.*, 2003). More cold tolerant species in Europe, such as *Pinus sylvestris* (Sinclair *et al.*, 1999), and *Betula* (Huntley & Birks, 1983), do not seem to follow these patterns (Ferris *et al.*, 1999).

The presence of these large comparative data sets, combined with fossil evidence, and simulation modeling of the genetic footprints of varying migratory patterns, has led to some powerful insights into the historical movements of plant species. However, outside of these European temperate tree species there are relatively few comparative studies available on plant phylogeography. Based on more limited sampling there is evidence that some plant species in the Pacific Northwest of North America show north/south partitioning of

chloroplast DNA (Soltis & Gitzendanner, 1999). Geographic structure has also been detected in North American *Liriodendron* (Sewell *et al.*, 1996). A geographic divide was found in *Nothofagus nervosa* in Argentina (Marchelli *et al.*, 1998). Phylogeographic structure was also found in the Aragan tree of Morocco (El Mousakik & Petit, 1996), *Aucoumea klaineana* in Gabon (Muloko-Ntoutoume *et al.*, 2000); and *Cunninghamia konishii* in Taiwan (Lu *et al.*, 2001).

From the perspective of UK biodiversity conservation, there is considerable interest in the conservation status of species with an arctic and alpine distribution, as it is populations of these species that are likely to be particularly sensitive to any future global warming which may greatly reduce or eliminate their habitat in the UK (Lusby & Wright, 1996). However compared to less cold tolerant species in Europe, there are far fewer studies in this area. Studies carried out to date include investigations into the phylogeographic structure of *Dryas integrifolia* (Tremblay & Schoen, 1999), *Saxifraga oppositifolia* (Abbott *et al.*, 2000) and *Silene acaulis* (Abbott *et al.*, 1995). The outcome of these studies has been a mixture of congruence and conflict. Some studies have produced convincing evidence for the persistence of species in high arctic refugia. For instance, large parts of Alaska are considered to have remained ice-free during the last ice age and this has led to the suggestion of the presence of the Beringial refugia (Abbott & Brochmann, 2003). Conversely, other studies have revealed evidence of long distance dispersal in many species (even those without specialized dispersal mechanisms) which indicates the possibility of complex and differing histories of post-glacial colonization in different species (Abbot & Brochmann, 2003).

Clearly, the identification of any general patterns in the distribution of genetic biodiversity in a given element of the flora would be useful as it would allow the design of regional conservation programmes aimed to conserve the range of diversity in several species simultaneously. However, while our understanding of the phylogeographic history of temperate European forest trees is becoming increasingly clear, there are simply not enough data to reach general conclusions regarding the arctic-alpine floristic element, and additional data are required to move this subject area forward.

1.3.2 THEME 2: Evolutionary and ecological fitness

1.3.2.1 Keys aspects of evolutionary and ecological fitness

Local adaptation

Plant populations are often not genetically equivalent to one another; they do not have equivalent levels of fitness. Their sessile nature is likely to lead to differential selection in different areas as plants become adapted to their local environment. This local adaptation can occur over a range of spatial scales from <1m to broad latitudinal clines (Galen *et al.*, 1991; Sork *et al.*, 1993; Kindell *et al.*, 1996). Understanding the differential adaptation of plant populations to different conditions is important for conservation biology. Preserving one population does not necessarily capture the potential of a species to exploit its full range of habitats, as different populations may contain different suites of adaptive genes.

Local adaptation is also important from the perspective of active conservation programmes such as population establishment, supplementation and re-introductions (Hufford & Mazer, 2003). For active conservation programmes to be successful there needs to be a good ecological match between donor populations and recipient populations, and the introduction of maladapted plants into a small existing population can lead to negative, rather than positive, conservation outcomes (Hufford & Mazer, 2003). However, despite a series of classical ecological genetic studies spanning several decades (Lowe *et al.*, 2004), studies of adaptation and quantitative differences between populations have been scarce in the literature in recent years. This decline is primarily attributable to the strong molecular biology focus of current genetics research, but this has come at the expense of ongoing and developing research into adaptation in natural populations.

Maximising the ability to evolve in a changing environment

Genetic variation is the raw material that natural selection acts on to lead to evolutionary change. Without genetic variation, there is no evolution. Thus the maintenance of genetic variation is important for the fitness of organisms and the persistence of species in a changing environment. This is particularly true for sessile organisms that lack the ability as individuals to actively modify their ranges to adjust to any environmental change.

The maintenance of genetic variation is dependent upon several factors including the size of a population and the amount of migration between populations (Barrett & Kohn, 1991). For a sessile species occurring in a changing environment, clearly the maintenance of some levels of genetic variation to allow for adaptation to changing conditions is likely to be beneficial. Small, isolated populations might be expected to be at risk due to genetic drift and low levels of migrants limiting the amount of genotypic combinations upon which selection can act (Ellstrand & Elam, 1993). In contrast, large populations existing as part of an inter-connected network of populations in a matrix of suitable habitat might be expected to be at a lower risk as drift is less severe in large populations and migration can allow a mix of genotypic combinations to be produced, on which selection can operate allowing populations to be continually evolving.

The avoidance of inbreeding depression

A potential negative consequence of a loss of genetic biodiversity is inbreeding depression (Barrett & Kohn, 1991). If a population experiences a reduction in size down to a small number of individuals then this can lead to mating among relatives. Individuals resulting from sib-matings can have lower levels of fitness than those resulting from mating events between more distantly related individuals. Thus if a plant population becomes small and isolated there is the possibility that subsequent mating events will lead to individuals of reduced fitness (Charlesworth & Charlesworth, 1987). The theoretical basis of inbreeding depression is well established, although there are still relatively few demonstrations of its importance for conservation biology under field situations. This is, at least in part, due to

the technical challenges of measuring firstly the pedigree of mating events, and secondly assessing the fitness of resulting offspring under field conditions.

1.3.2.2 Defining and measuring fitness

In considering evolutionary and ecological fitness, one immediately runs into a challenging set of questions. What exactly is fitness and how should it be measured (quantified)? Fitness itself can be a rather diffuse concept. Fitness can be defined in an ecological and evolutionary context as 'reproductive success – the capacity to pass genes on to the next generation'. This is a very important concept for anyone working in conservation. However fitness is not an easy trait to quantify. Since fitness is defined as reproductive success, it could be argued that seed characteristics could be a direct method of measuring fitness (e.g. seed weight, number and germinability). Measures of vegetative characters can also be used as indications of fitness, but perhaps it is more accurate to say these are indicators of vigor. Thus a large plant often produces many flowers, which in turn sets many seeds, but a vigorous plant does not necessarily pass its genes on to the next generation. For example when plants occur in cooler climates on the edges of their ranges, they may grow well, but never flower. Also, a plant that allocates all of its resources into growth will most likely be larger than a neighboring plant that allocates half of its resources into flowering, though the second will pass on more of its genes to the next generation.

Fitness in relation to context is another difficult issue. The idea of bringing plants from various populations into cultivation in order to compare their relative fitness in an environment that is free of uneven environmental stresses is intuitively appealing. Yet at the same time, the measurement of fitness that is obtained in cultivation may or may not have any relevance to fitness in the field. It is possible that under favorable conditions some individuals will grow and flower well while others do not, but those that do not grow well in cultivation (and hence appear 'less fit') could have a resistance to drought (or some other environmental factor) that gives them an advantage in the field. It is necessary to decide if

it is some inherent genetic fitness that one wants to measure, or a fitness of a particular taxon at a particular site at a particular time.

1.3.2.3 Factors likely to be important in determining fitness

Population genetic theory associates a small population size with low genetic diversity, and potentially inbreeding depression and the resulting decline in fitness and increase in likelihood of extinction (Saccheri *et al.,* 1998). Yet when empirical data on population size, molecular variation, and fitness are examined, a correlation between these variables does not seem to be consistently present.

Historic population size may be one reason why much of the existing empirical data does not fit neatly with the theories on the relationship between population size, molecular variation, and fitness. Most of the plants that receive conservation attention are plants whose ranges have recently become smaller, causing concern. Genetic drift has been documented to occur much more quickly in small populations (Hartl & Clark, 1997), but is still dependent on generation time. If a plant has not gone through several generations since the reduction in population size, or population fragmentation, there will not have been time for negative consequences of a lack of genetic variation to act. However if the species is an annual out-crosser, dependent on insect pollination, and the remaining population does not create enough of a display to attract many pollinators, very few seeds could be set, and the resulting fitness could rapidly become very low. On the other hand, if the plant in question evolved (speciated) in a small isolated population, perhaps in novel environmental conditions, its genetic load may have been purged of deleterious alleles and the plant may show very little genetic variation, yet have no ill effects and a high fitness, at least in the short term.

The breeding system is also likely to be correlated closely with levels and partitioning of molecular variation (Hamrick *et al.*, 1991). Species with different breeding systems may differ in the amounts of genetic variation they contain, and also differ in their susceptibility

to inbreeding depression (Barrett & Kohn, 1991). If inbreeding depression is caused by deleterious recessives, then selfing might be expected to purge a population of genetic load (Keller & Waller, 2002). In contrast, if the population is small and drift overrides selection, inbreeding depression may still occur in selfing populations. In addition, if over dominance is a cause of inbreeding depression, and heterozygotes are fitter than homozygotes, no amount of selfing will eliminate it (Keller & Waller, 2002).

Ecology and life habit may also affect the relationship between population size, molecular variation, and fitness. Plants that occupy different niches and have different life history strategies may have different levels of genetic diversity (Hamrick *et al.*, 1991). It follows that they will have different optimal levels of genetic diversity, and what would pose a conservation crisis for one taxon would not for another. A long-lived tree, with a lifespan of hundreds of years is at an advantage if its offspring have a wide variation of genotypes, since the environment that the tree is reproducing in is quite likely to be different from the one it germinated in, and the environment its offspring will be reproducing in is likely to be different again. An annual weed with a minimal seed bank will be producing seeds for the next season, and wide genetic variability in these seeds could even be detrimental, since the plants that have survived to reproduction are the successful ones, and it is to their advantage to produce offspring that are genetically similar to themselves.

1.3.2.4 The relationship between marker variability and fitness

Molecular techniques represent important tools for plant conservation studies (Falk & Holsinger, 1991; Frankham *et al.*, 2002). However, an important (but commonly overlooked) point with regards to studies examining the partitioning of population genetic diversity, is that the mainstream molecular ecological techniques reveal patterns of genetic variability that is most likely to be neutral (i.e. the distribution of genetic variants are considered neither beneficial nor detrimental to the individuals that possess them). These markers are uncoupled from the genes causing adaptive and fitness differences between individuals and populations. Thus a gulf remains between the observation that two

populations differ in the amounts or types of genetic marker variability, and the interpretation of this difference.

In a crisis-orientated discipline like conservation biology, it is appealing to search for simple and rapid solutions to problems, and approaches where general trends allow safe extrapolation to a wide range of situations. Population genetic surveys by techniques such as allozymes (Soltis *et al.*, 1989), RAPDs (random amplified DNA; Williams *et al.*, 1990), AFLPs (amplified fragment length polymorphism; Vos *et al.*, 1995), and even SSRs (simple sequence repeats; Edwards *et al.*, 1996) are becoming increasingly cheap and easy to carry out. There are conservation genetic studies that use marker variability as a surrogate measure of fitness, but the extent to which this is justified is currently unclear. The relationship between population size, neutral molecular variation, and fitness has been discussed extensively in the theoretical literature (Lynch *et al.*, 1995; Hedrick & Kalinowski, 2000), but in plants there is a rapidly increasing, but contradictory body of empirical research around the subject. Several molecular markers failed to distinguish between populations of *Pinus sylvestris* in Finland, that showed high differentiation in important adaptive traits (Karhu *et al.*, 1996). This is probably attributable to high gene flow and strong selective differences, causing the variation in neutral markers to poorly represent the variation in adaptive genes (Hedrick, 2001). Likewise in *Primula scotica*, virtually no genetic variation was recovered using RAPDs and allozymes (Glover & Abbott, 1996) leading the authors to conclude that populations of the species were genetically equivalent from a conservation perspective. However, significant differences in ecologically relevant heritable quantitative characters were detected when morphological variation was measured in a common garden experiment (Ennos *et al.*, 1997). In contrast there was an association between heterozygosity for neutral markers and fitness associated traits in *Salvia pratensis* (Ouborg & Treuren, 1994). Several reviews of the subject have been carried out, combining data from plants and animals (Hansson & Westerberg, 2002; Keller & Waller, 2002), as well as a meta-analysis (Reed & Frankham, 2003), but the conclusions from the empirical studies are not clear-cut. This creates a conservation dilemma. Should marker variability be used as an indicator of likely fitness problems? When should we become concerned about population reduction and fragmentation? What

population sizes and levels of genetic variation are sustainable and where can we best allocate conservation resources? There is thus a knowledge gap that acts as an impediment for conservation action.

1.4 Thesis aims

In this thesis I use a case-study based approach to explore the conservation biology issues surrounding the two research themes identified in the introduction.

A single species has been investigated for each of these themes, and the resulting data will then be discussed in the context of other studies to assess the merits and challenges facing the conservation of genetic biodiversity in relation to these topics.

The questions I specifically wish to address are:

Theme 1: The protection of a broad spectrum of genetic biodiversity
For arctic alpine species in Britain, how does the diversity present in the UK relate to the diversity in other countries, and is there any clear evidence for sites of glacial refugia. The study taxon selected for this investigation is the Marsh Saxifrage, *Saxifraga hirculus* (Chapter 2).

Theme 2: Evolutionary and ecological fitness
Is there a relationship between population size and isolation, with molecular variation, and with fitness? The study taxon selected for this investigation is the Sea Pea, *Lathryus japonicus* (Chapters 3, 4 and 5).

Chapter 2. Phylogeographic structure of an arctic plant, *Saxifraga hirculus*

2.1 Introduction

Conservation programmes are typically constructed and delivered at the national level (e.g. the UK Biodiversity Action Plans; H.M.G. 1995). However, in conserving the genetic biodiversity of a taxon, it is necessary to consider species in a broader international context to identify any major intra-specific genetic races and to establish how the genetic diversity of a species is geographically structured. The distribution of intra-specific genetic biodiversity is likely to be affected by historical events such as plant range expansions and contractions due to glacial cycles (Hewitt, 1996). This means that populations with very different histories can occur in different parts of a species' range, and this can result in different levels of genetic diversity being present in a species in different regions. Range edge populations are often small and isolated, and for many species these populations contain only a small percentage of the total intra-specific genetic diversity (Lesica & Allendorf, 1995). These populations thus might be considered less important to the survival of the species than range centre populations or populations in refugial areas. Range centre populations could be argued to be of higher conservation value. Range centre populations, however, can be locally widespread and hence receive little conservation attention. It should also be noted that some range edge populations can have a conservation value in their own right as these isolated populations can be candidates for incipient speciation due to local origins of new morphs and ecotypes (Hunter & Hutchinson, 1994; Lesica & Allendorf, 1995).

Molecular genetic studies are useful for gaining insights into the distribution of broad scale genetic variation within species (Newton *et al.* 1999; Avise, 2000). Furthermore, as such studies can shed light on the historical movement of plant species in the face of past environmental change, they may also help with assessing how species may respond to future climatic change. The chloroplast genome is ideal for broad scale genetic structure

studies because it evolves more slowly on average than nuclear markers, leaving a track of historic relationships, and its haploid behaviour makes it a good indicator of population differentiation caused by drift due to the smaller effective population size(Ennos *et al.*, 1999).

Several recent studies have examined the patterns of cpDNA variation in plant species with arctic distributions (Tremblay & Schoen, 1999; Abbott *et al.*, 2000). These studies have produced evidence bearing on the hypothetical existence of northern refugia, areas in which species may have survived the last glaciations *in situ*. In the past, attention has been directed at the Nordic region, where the debate over glacial survival versus the alternative tabula rasa hypothesis has been particularly fierce (Dahl, 1987; Borgen, 1987; Nordal, 1987; Birks, 1994). Patterns of nuclear genetic variation in this area in a range of species, including the out-breeding *Saxifraga oppositifolia* (Gabrielsen *et al.*, 1998) and the inbreeding *S. cespitosa* (Tollesfrud *et al.*, 1998), show very little evidence of geographical race formation or of deep geographical structure (Brochman *et al.*, 1996). Rather, studies have concluded that levels of gene flow have been so high that any evidence of centres of genetic diversity, such as might be expected to characterise refugia, is likely to have been obliterated; thus glacial refugia, if they existed in the Nordic region, are irrelevant to present day patterns of variation (Brochman *et al.*, 1996). On the other hand, at a larger spatial scale, a survey of chloroplast DNA restriction fragment-length polymorphism (RFLP) variation in *Saxifraga oppositifolia* across the entire arctic revealed both significant structure and a centre of diversity, indicating the presence of two principal lineages that may have originated from a refugium in western Beringia (Abbott *et al.*, 2000). Indeed, a growing body of evidence suggests that the Beringia region of Alaska may have acted as a refugium for many arctic plants throughout the ice ages (Abbott & Brochmann, 2003).

The aim of this chapter is to describe the genetic structure of the British populations of another arctic-montane species of *Saxifraga*, *S. hirculus* L. (Saxifragaceae), setting it in the context of its wider European gene pool, and comparing this with populations from Alaska and Colorado in North America. Although fossil evidence indicates that British and Irish populations of *S. hirculus* may have survived south of the limit of the Weichselian

glaciation (Godwin, 1975), it is nonetheless of interest to determine the patterns of genetic variation that have been established following the retreat of the ice and to see whether or not there is any evidence of extensive gene flow between populations or regions. The approach has been to study RFLP and sequence variation in chloroplast DNA, which is inherited maternally in the Saxifragaceae (Soltis *et al.*, 1990).

2.1.1 The study species: Saxifraga hirculus

Saxifraga hirculus is a loosely tufted, rhizomatous, perennial herb with a circumpolar, arctic-montane distribution that extends southwards distjunctly to the Rocky Mountains of Colorado, the Caucasus, Central Asia and the outer ranges of the Himalaya (Figure 2.1).

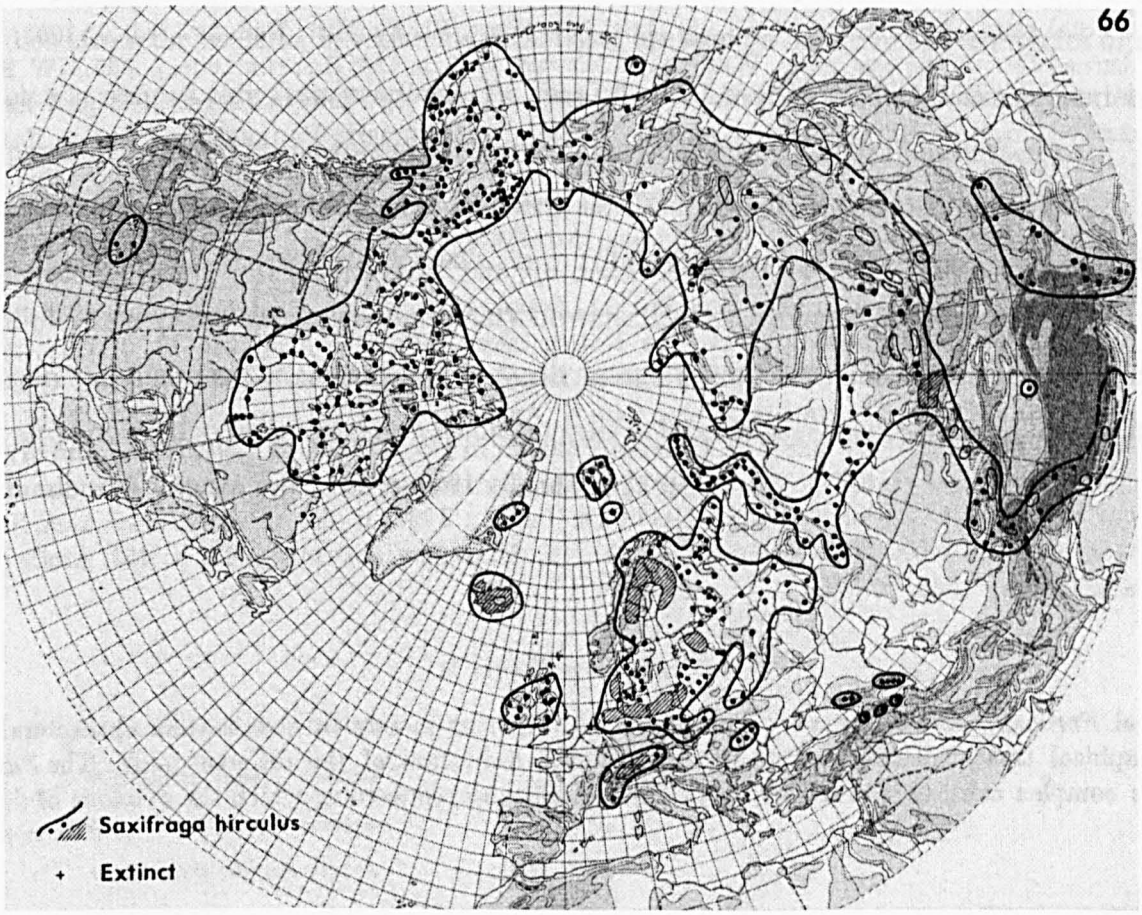

Figure 2.1. Distribution map of *S. hirculus*.
From Hultén E (1971)

The species is polymorphic, prompting Engler (1916) to recognise eleven infra-specific taxa. The morphological variation appears to be correlated to some extent with differences in chromosome number, such that diploids (2n = 16) and tetraploids (2n = 32) are often associated with particular phenotypes. Occasional triploids also occur. In the most recent taxonomic treatment of the complex, Hedberg (1992) recognised four subspecies. It seems that an imperfect distinction can be made between largely circumboreal-montane populations, in which the flowering stems are tall, bear many leaves and at least two flowers, and whose sepals are reflexed at anthesis (ssp. *hirculus*, 2n = 32) and predominantly circumpolar, arctic populations whose flowering stems are shorter, bear fewer leaves and solitary flowers, with sepals erect or spreading at anthesis. The arctic populations can be divided into those primarily from the palaearctic, in which the base of the petals is auriculate to truncate with a short claw (ssp. *compacta* O. Hedb., 2n = 32) and those from the nearctic, in which the petal bases are tapered but do not have auricles or a clearly defined claw (ssp. *propinqua* (R. Br.) Löve & Löve, 2n = 16). Alaskan material is particularly variable and contains all three of these subspecies, as well as numerous intermediates (Hultén, 1968: 568; Hedberg, 1992). The outlying populations from Colorado (ssp. *coloradensis* O. Hedb., 2n = 16) differ from ssp. *hirculus* not only in their diploid status but also in having solitary flowers. Before leaving this taxonomic summary/introduction, however, it is worth noting that at least some of the morphological variation seen in the species as a whole appears to be environmentally induced, not least the number of flowers produced per stem (Hedberg, 1992).

Adopting the taxonomy outlined above, the European gene pool contains the boreal-montane ssp. *hirculus* and the Palaearctic ssp. *compacta*. According to Hedberg (1992), these two taxa are sympatric in Iceland, where they can be found along with intermediates. However, when mapping the distribution of the species in Europe, Jalas (1999) recognised all material from Iceland and Svalbard as the palaearctic variant, using the name ssp. *alpina* (Engl.) A. Löve, and distinguishing it from the material found in the British Isles and mainland Europe which was treated as ssp. *hirculus*. (Since the type of ssp. *alpina* was described from Sikkim, it is preferable at this stage to retain Hedberg's nomenclature in the European context and call the palaearctic variant ssp. *compacta*).

In much of Europe *Saxifraga hirculus* usually inhabits base rich flushes and mires but in the arctic it occurs mainly in water-saturated moss-tundra. It extends discontinuously southwards to Switzerland and central Romania. During the 19th and 20th centuries, however, populations in the southern part of its range suffered a serious decline, so much so that the species is now extinct in Austria, the Czech republic and the Netherlands, and is severely depleted in southern Sweden, Germany, the Alps, Poland and the Baltic states, where it is registered as endangered or vulnerable (Welch, 2002; Jalas, 1999). It has also declined in the British Isles, where it is now restricted to the northern Pennines, the Pentland Hills, the Grampians and north-eastern Scotland and a few localities in Ireland (Preston *et al.*, 2002). It is a priority BAP (Biodiversity Action Plan) species listed in Annexes II and IV of the EU Habitats and Species Directive, and there are plans for reintroductions as part of the UK BAP for the species (H.M.Government, 1995)

Owing to concerns regarding its conservation there has been renewed interest in the reproductive biology of *Saxifraga hirculus*. *S. hirculus* reproduces by seed and also by basal axillary shoots. Because of this combination of sexual and vegetative reproduction, it is not clear exactly how many individuals occur in a population, or how they are distributed spatially. Research has been carried out estimating the number and position of genets in a population of *S. hirculus* in Denmark (Olesen & Warncke, 1990). Using isozyme data, anthesis, petal size, and local peaks of flower abundance, this study estimated 10 – 17

genets, or groups of closely related plants, in a population covering about 30 m. The population structure in Britain is not known. The flowers of *S. hirculus* are protandrous, with distinct male and female phases, though the plants are self-fertile (Olesen & Warncke, 1989b). Syrphid flies are the predominant pollinators, with different species being important in different regions (Warncke *et al.*, 1993). Olesen & Warncke (1989a) have studied pollinator flight patterns and pollen movement in detail. The seeds of *S. hirculus* have no specialized dispersal mechanism, and are dropped near the parent plant when shaken from the capsules by wind or rain. Average seed dispersal distance is estimated to be 13 cm (Olesen & Warncke, 1989a). *S. hirculus* seems to require disturbance for seedling recruitment. It has been speculated that large herbivores, such as deer, might move seeds that have become embedded in mud from place to place (Olesen & Warncke, 1990).

2.2 Materials and Methods

2.2.1 Sampling

To assess the phylogeographic structure of *Saxifraga hirculus,* leaf samples from a total of 488 individuals were collected from 30 populations from England, Scotland, Ireland, Iceland, Svalbard, Denmark, Switzerland, Alaska and Colorado. Full details about sample locality are provided in Table 2.1.

Samples were collected from a distance of at least a half metre apart, or a metre where possible, to decrease the likelihood of sampling twice from a single clone. Where possible, samples were collected evenly spaced from across the extent of the population. Samples from Denmark were collected along transects separated by 9-45m. However samples from Svalbard, Denmark, and Switzerland were collected by different people, and so less information is available on sampling technique.

Region	Population	Lat/Long	Code	Ni	Haplotype no.
Scotland, Grampians	Silverford	57°19'N, 2°57'W	SF	19	2
Scotland, Grampians	Buck of Cabrach	57°19'N, 2°59'W	BC	18	2
Scotland, Grampians	Slug Burn	57°19'N, 2°58'W	SB	20	2
Scotland, Pentlands	Craigengar, The Pike	55°46'N, 3°29'W	CNG	20	1
England, Pennines	Sally Grain Head	54°45'N, 2°19'W	SGH	20	3
England, Pennines	Yad Moss	54°44'N, 2°21'W	YM	19	3
England, Pennines	Great Shunner Fell	54°22'N, 2°14'W	GSF	13	2
England, Pennines	Knock Ore Gill	54°25'N, 3°26'W	KOG	16	2
Northern Ireland		54°45'N, 6°0'W	NI	5	3
Svalbard		78°N, 15°W	SH	20	4,5
Denmark	Rosborg	56°25'N, 9°13'E	DEN	20	3,6
Switzerland		43°33'N, 60°14'E	SWZ	8	6
Iceland	Southwest	64°10'N, 21°45'W	ISW	20	3
Iceland	Northwest	64°58'N, 21°8'W	INW	15	3
Iceland	North	65°35'N, 19°45'W	IN	15	3
Iceland	South 1	60°35'N, 19°50'W	IS1	16	3
Iceland	South 2	60°35'N, 19°45'W	IS2	20	3,7
Alaska, N. of Anchorage	Summit Lake	61°27'N, 149°27'W	SUM	22	8
Alaska, Denali Natl. Park	Primrose Ridge	63°45'N, 149°22'W	PRM	21	22,23
Alaska, Denali Natl. Park	Sable Pass	63°33'N, 149°36'W	SAB	19	9-15
Alaska, Denali Natl. Park	Polychrome Pass	63°31'N, 149°50'W	PCH	16	13,15-17
Alaska, Denali Natl. Park	Highway Pass	63°28'N, 150°10'W	HWY	19	13,15,18-21
Alaska, Richardson Hwy	Fielding Lake	63°6'N, 145°24'W	FLD	14	13,15-17,24-26
Alaska, Nome	Kougarok Road	64°30'N, 165°8'W	KOU	8	13,22,27
Alaska, Nome/Teller Rd	Penny River	64°36'N, 165°40'W	PEN	12	13,22,28-30
Alaska, Nome/Teller Rd	Kigluaik Mtns.	64°48'N, 165°48'W	KIG	13	13,22,30,31
Alaska, Nome/Teller Rd	Mile 56	64°54'N, 166°0'W	M56	7	13,17,38
Alaska, Nome/Teller Rd	Mile 62	65°0'N, 166°12'W	M62	4	17
Alaska, Teller	S of Teller	65°18'N, 166°24'W	TEL	13	17,22,32-37
Colorado	Summit Lake	39°35'N, 105°39'W	CSL	16	39
Colorado	River valley	39°19'N, 105°35'W	CRI	19	40

Table 2.1. Collection details of accessions of *Saxifraga hirculus* used in the study (Ni = no. individuals sampled). The haplotypes column refers to the chloroplast DNA type, see Results section.

2.2.2 Molecular methods

DNA extractions were carried out using the CTAB method of Doyle & Doyle (1990). Seventeen regions of chloroplast DNA were examined : *atp*B-*rbc*L, *trn*L-F, *trn*C-*trn*D, *trn*F-*trn*Vr, *trn*D-*trn*T, *trn*S-*trn*G, *trn*K2-*trn*Qr, *trn*Q-*trn*Rr, *trn*K2-*trn*K1, *trn*H-*trn*K, *trn*T-*psb*Cr, *rpo*C1-*trn*Cr, *trn*H-*psb*A, *psb*Ar-*trn*Fm, *rpl*20-*rps*12, *psb*B-*psb*F, and *trn*V-*rbc*Lr. These regions were PCR amplified using universal primers designed by Hamilton (1999), Chiang *et al.* (1998), Taberlet *et al.* (1991), Demesure *et al.* (1995), and Dumolin-Lapegue *et al.* (1997). The products were then blind cut with the restriction enzymes Alu I, Dra I, Hha I, Hinc II, Hinf I, Mse I, Msp I, Rsa I and Tru I. In order to avoid scoring the same mutation twice, two polymorphisms from the same region were used only if they separated different individuals. For a sub-set of samples, the regions *trn*L-F, *trn*H-*psb*A, *atp*B-*rbc*L, *rsp*20-*rpl*12, and *trn*S-*trn*G (Appendix 2) were sequenced on an ABI 377 sequencer and aligned using the program Sequence Navigator. When sequencing revealed polymorphisms, restriction enzyme assays were designed (where possible) using the program Webcutter (http://www.firstmarket.com/cgi-bin/cutter) and the assays extended to the full sample set. A small inversion detected in the region *trn*T-*psb*A and several point mutations present outside restriction sites in other sequenced regions were screened for one individual of each haplotype from each of the European populations (Appendix 2), since the low level of variation that was detected from these populations made it reasonable to assume that these samples represent their source populations. This assumption did not hold for the much more variable Alaskan populations, and since every individual could not be sequenced due to time and money constraints, these mutations were left out of the analysis. The sequence-only data was also left out of the F_{ST} analysis. A minimum-spanning network for European and North American haplotypes was constructed using the RFLP data only (Appendix 1).

The enzyme digests were carried out on 10µl of PCR product, to which 0.2µl of enzyme, 0.2 µl of BSA, 2µl of 10X buffer and 7.6µl H_2O were added following the manufacturers instructions (Gibco BRL). They were incubated at 37°C in a water bath for two hours.

The restriction digests were either separated on agarose or polyacrylamide gels and visualised using ethidium bromide and ultra violet light. In many instances, particularly where intra-population variation occurred, duplicate DNA extractions were carried out, and the haplotypes re-confirmed.

2.2.3 Data analysis

The results were analysed using the program Arlequin ver. 2000 (Schneider *et al.*, 2000) to obtain F_{ST} estimates for the European, the Alaskan, and combined samples. Arlequin was also used to construct minimum spanning networks between haplotypes. The program DISTON (Petit, 2000) was used to calculate G_{ST} and N_{ST}. G_{ST} is analogous to F_{ST} and considered interchangeable. The only reason both of these measures are used in this study is that Arlequin calculates one, and DISTON, the other. The measures G_{ST} and F_{ST} provide estimates of population subdivision under the assumption that drift is the primary cause of population differentiation. N_{ST} provides an estimate of population differentiation that takes into account not just the frequency of haplotypes, but also the number of mutations separating these haplotype by incorporating a coefficient of gene differentiation. Because DISTON could accept a maximum of only 25 haplotypes the G_{ST} and N_{ST} analysis was carried out using a subset of the markers, and thus collapsing groups of similar haplotypes. The subset was obtained by preferentially eliminating the RFLP loci that were not designed from sequence knowledge (e.g. the loci for which we had least knowledge regarding the homology/homoplasy of co-migrating fragments). Wilcoxon signed ranks tests to test the difference between N_{ST} and G_{ST} were performed on SPSS (SPSS Inc, 2001). Measures of genetic differentiation of populations were plotted against geographical distances between populations to assess the extent of a correlation between genetic and geographical distances.

2.3 Results

Of the seventeen regions of chloroplast DNA screened by means of RFLP analysis, seven regions contained usable polymorphisms (*atp*B-*rbc*L, *trn*L-F, *trn*C-*trn*D, *trn*F-*trn*Vr, *trn*D-*trn*T, *trn*S-*trn*G, *trn*K2-*trn*Q), which were scored in all 488 individuals.

This data revealed 40 haplotypes. No haplotypes were shared between the regions of Europe, Alaska or Colorado. Seven haplotypes occur in Europe (Table 2.2).

Hap	Craigengar (Pentlands)	Slug Burn (Grampions)	Buck of Cabrach (Grampions)	Silverford (Grampions)	Yad Moss (Pennines)	Sally Grain Head (Pennines)	Knock Ore Gill (Pennines)	Great Shunner Fell (Pennines)	Northern Ireland	Svalbard	Switzerland	Denmark DEN	Iceland Southwest	Iceland Northwest	Iceland North	Iceland South 1	Iceland South 2	Total
1	20																	20
2		20	18	19			16	13										86
3					19	20			5			10	20	15	15	16	19	139
4										18								18
5										2								2
6											8	10						18
7																	1	1
Total	20	20	18	19	19	20	16	13	5	20	8	20	20	15	15	16	20	284
Hap D	0.05	0.05	0.06	0.05	0.05	0.05	0.06	0.08	0.2	0.1	0.13	0.1	0.05	0.07	0.07	0.06	0.1	0.02

Table 2.2. Variation in chloroplast RFLP haplotypes among European populations of *Saxifraga hirculus*.

Columns represent populations and rows represent haplotypes. Cells show the number of individuals with a given haplotype. Hap = haplotype number. Hap D stands for haplotype diversity, or haplotypes present in a population divided by individuals sampled.

Thirty-one haplotypes, haplotypes 8 through 38, occur in Alaska (Table 2.3). Haplotypes 39 and 40 are from Colorado (Table 2.3).

Hap	Summit Lake	Sable Pass	Polychrome Pass	Highway Pass	Primrose Ridge	Fielding Lake	Kougarok Road	Penny River	Kigluaik Mtns	South of Teller	Mile 56 Nome/Teller Rd	Mile 62 Nome/Teller Rd	Summit Lake (Colorado)	River valley (Colorado)	Total
8	22														22
9		2													2
10		5													5
11		7													7
12		1													1
13		2	10	13		1	5	6	3		2				42
14		1													1
15		1	2	1		1									5
16			2			3									5
17			2			4				1	4	4			15
18				1											1
19				2											2
20				1											1
21				1											1
22					18		1	2	2	5					28
23					3										3
24						2									2
25						2									2
26						1									1
27							2								2
28								1							1
29								2							2
30								1	2						3
31									6						6
32										1					1
33										2					2
34										1					1
35										1					1
36										1					1
37										1					1
38											1				1
39													16		16
40														19	19
Total	22	19	16	19	21	14	8	12	13	13	7	4	16	19	203
Hap D	0.05	0.37	0.25	0.32	0.1	0.5	0.38	0.42	0.31	0.62	0.43	0.25	0.06	0.05	0.16

Table 2.3. Variation in chloroplast RFLP haplotypes among North American populations of *Saxifraga hirculus*.

Columns represent populations and rows represent haplotypes. Cells show the number of individuals with a given haplotype. Hap = haplotype number. Hap D stands for haplotype diversity, or haplotypes present in a population divided by individuals sampled. All populations are from Alaska except the two populations from Colorado.

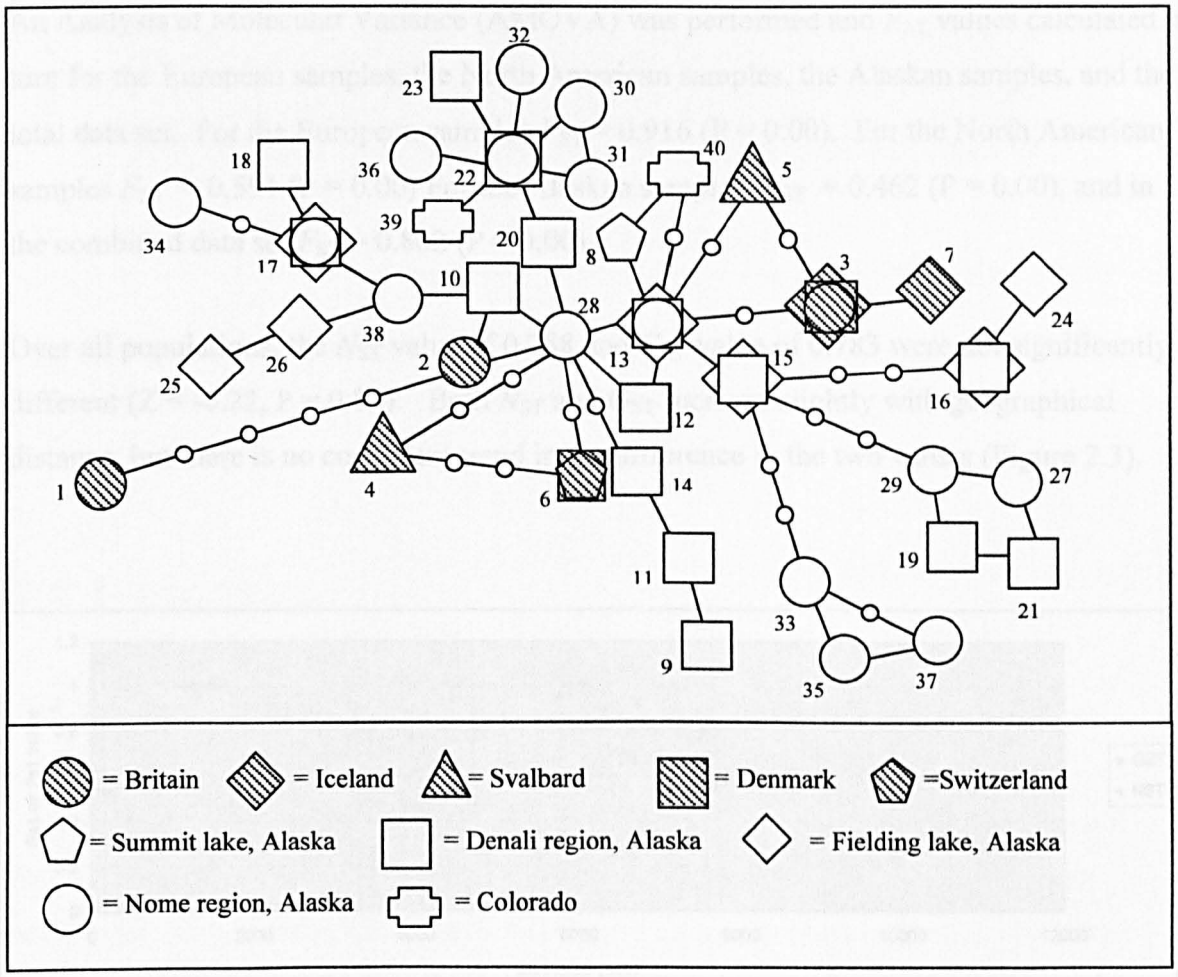

Figure 2.2. Minimum spanning network for European and North American chloroplast haplotypes of *S. hirculus*.

Numbers represent haplotypes with their regions of origin denoted by symbols. Superimposed symbols indicate the co-occurrence of haplotypes in a region. Shaded symbols are from Europe, unshaded are from North America. Small circles denote hypothetical intermediate haplotypes, each differing by one mutational step. For details of mutational differences among haplotypes see Appendix 1.

When a minimum spanning network is constructed, using RFLP data (Figure 2.2), it shows that the Alaskan haplotypes fill many of the intermediate haplotype states between European haplotypes. Haplotype 1 from the Pentlands in Scotland is situated on a long branch, separated by 5 unique mutations from haplotype two from the Grampians. Haplotype 4 is connected to haplotype 6 in this spanning network by three mutations, however mutations revealed by sequencing differentiated these two haplotypes further(Appendix 2).

An Analysis of Molecular Variance (AMOVA) was performed and F_{ST} values calculated in turn for the European samples, the North American samples, the Alaskan samples, and the total data set. For the European samples F_{ST} = 0.916 (P = 0.00). For the North American samples F_{ST} = 0.591 (P = 0.00) For the Alaskan samples, F_{ST} = 0.462 (P = 0.00), and in the combined data set F_{ST} = 0.802 (P = 0.00).

Over all populations, the N_{ST} value of 0.788 and G_{ST} value of 0.783 were not significantly different (Z = -0.22, P = 0.83). Both N_{ST} and G_{ST} increase slightly with geographical distance, but there is no consistent trend in the difference in the two values (Figure 2.3).

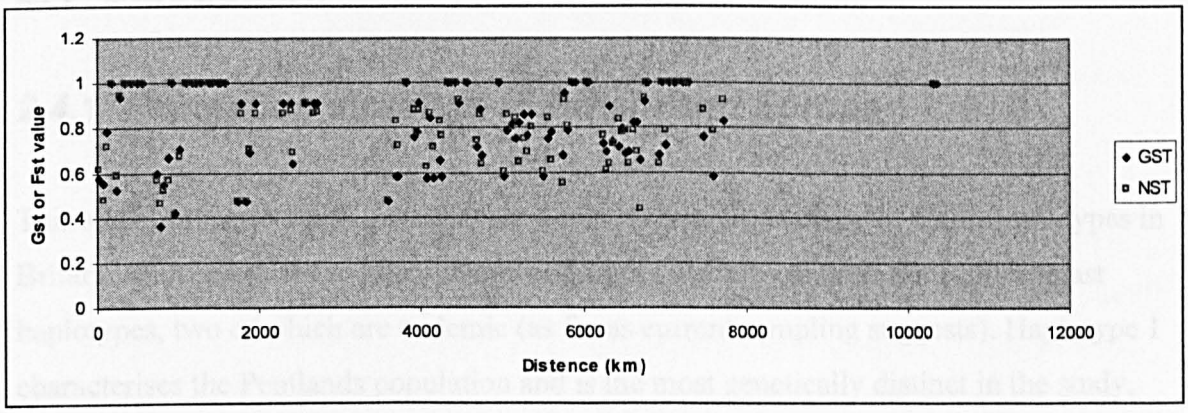

Figure 2.3. Pairwise estimates of G_{ST} and N_{ST}

Plotted against pairwise measures of geographical distance between populations

Two distance groups appear in the graphs of N_{ST} and G_{ST} against distance, one group making up populations under 2700 km apart and another over 3600 km apart. More variation is visible in the difference between G_{ST} and N_{ST} in the populations over 3600km apart, with variance= 0.008 in the over 3600km group, and variance = 0.001 in the under 2700 km group (Figure 2.4). However there is no significant difference detected between these broad population groups by ANOVA (F = 1.072, sig = 0.303).

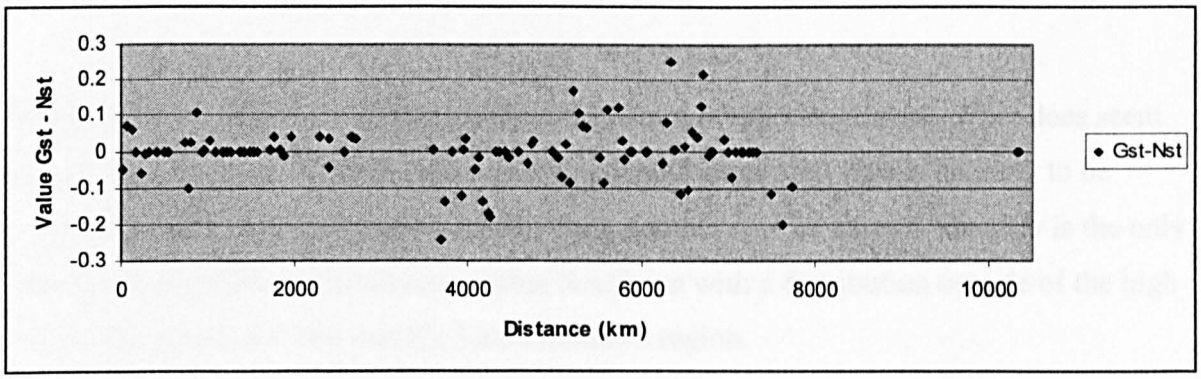

Figure 2.4. Pairwise values of G_{ST} – N_{ST} in 50km distance classes
Plotted against pairwise geographical distances between populations.

2.4 Discussion

2.4.1 Population structure in Britain and Europe

This study indicates that *S. hirculus* has a relatively high diversity of chloroplast types in Britain when compared to other places in Europe. Britain contains three chloroplast haplotypes, two of which are endemic (as far as current sampling suggests). Haplotype 1 characterises the Pentlands population and is the most genetically distinct in the study, separated from haplotype 2 by five mutations (Figure 2.2). Haplotype 2, found only in the Grampians and two populations in the Pennines, is separated from two Alaskan haplotypes by only one mutation (Figure 2.2). Haplotype 3 occurs in two of the Pennine populations as well as in Iceland and Denmark. The genetic distinctness of *S. hirculus* in the Pentlands is unexpected, and could have conservation implications. The other British populations of *Saxifraga hirculus* have greater haplotypic similarity to Alaskan populations than they do the Pentlands population. This raises the possibility that there is a previously unrecognised divergent lineage present in the UK. The question then arises as to whether this Pentlands *S. hirculus* population is a product of dispersal from somewhere different from other populations in the study, or if it is a relic. The answer to this, however, remains unclear, as despite the extensive sampling in this study, there is the eastern arctic part of the distribution of *Saxifraga hirculus* that remains unsampled. It may be that the Pentlands haplotypes would be recovered from this part of the species range (and if so, suggest

dispersal, rather than a local relic as being the origin of this population). What does seem safe to say, however, is that the divergent Pentlands chloroplast type is unlikely to be simply a case of local chloroplast capture from another species since *S. hirculus* is the only member of the section *Hirculoides* within *Saxifraga* with a distribution outside of the high mountains of central Asia and the Sino-Himalaya region.

The higher chloroplast diversity in Britain in comparison with other countries of Europe could, in part, be a sampling artefact since eight populations were sampled in Britain, whereas only one population was sampled from each of Switzerland, Denmark and Svalbard. However, the sampling reflects the species abundance in Denmark and Switzerland, and there is little opportunity for adding to this sample. Two haplotypes were detected from a single population from Svalbard, and should further sampling be undertaken there, additional chloroplast haplotypes might be found on this high arctic island.

In contrast to Britain, Iceland has low chloroplast diversity, with one haplotype being present in 86 samples from five well-separated populations. This homogeneity would be expected in a place that was colonized recently from a single refugium, while perhaps Britain was colonized from several refugia, or perhaps maintained populations surviving throughout the ice ages.

It is noteworthy that populations in the Pennines, as close as 8 km, are fixed for different chloroplast haplotypes. This indicates that no seed dispersal is occurring over distances as short as 8 km. In Denmark the situation appears to be even more marked. Here individuals from quadrats 1-4 contain one haplotype (no.3), whereas those from quadrats 5-7 contain another (no.6). The two sets of quadrats are separated from each other by metres (9-45m) rather than kilometres. Seed dispersal is thus perhaps even more restricted than would have been deduced from a study of population structure in the Pennines. The limiting factor, however, may not be dispersal itself, but rather seedling germination and establishment following dispersal. It is known that so-called 'priority effects' play a major role in controlling the genetic consequences of any dispersal, whereby immigrants fail to establish owing to the unavailability of suitable niches or fierce competition from the plants, pests

and pathogens that are already there. *Saxifraga hirculus* may need bare ground in order to become established.

As far as population differentiation is concerned, most variation is held between populations of *S. hirculus* in Europe, with $F_{ST} = 0.92$. This high level of F_{ST} indicates virtually no within population variation and high between population differentiation. Thus it appears that existence in restricted populations in a specific and local habitat type, and production of seeds without specialized dispersal mechanisms, has led to high inter-population differentiation in this species. Most European haplotypes are genetically distant from each other, and generally when two haplotypes occur in the same population, as in Denmark and Svalbard, they are not the genetically most similar haplotypes (Figure 2.2). This pattern is characteristic of old lineages that have become geographically dispersed since the time they arose by mutation (Pons & Petit, 1996). In contrast the single individual in Iceland with a different haplotype (haplotype 7) differs from the common Icelandic haplotype (haplotype 3) by only one mutation (Figure 2.2, Appendix 1), and this haplotype could have a more recent origin.

Most mutations found in the European material are autapomorphies - few are shared between populations. This makes a phylogenetic analysis of the data problematic, due to lack of resolution. The minimum-spanning network shows genetic similarity, but does not necessarily imply genealogical relationships. Because of the nature of this data, this hampers inferences about the evolutionary relationships between populations, and conclusions are not presently possible about clear migration routes or hypothetical refugial areas for European *S. hirculus*. The data, do, however, indicate clear differentiation between populations at a range of scales from the local to inter-country, and particularly in the case of the localised population structure, the differences were higher than might have been suspected in the absence of the data.

2.4.2 Population structure in North America

In contrast to the situation in Europe (where most variation is held between populations of *S. hirculus,* $F_{ST} = 0.92$), in Alaska most variation is held within populations with $F_{ST} = 0.46$. Six of the 31 haplotypes were present in more than one population, and 5 of these six haplotypes were present in more than one region. Also, in Alaska 31 haplotypes were present in 12 populations, compared with 7 haplotypes in 17 populations in Europe. There is thus greater diversity and more evenly distributed diversity, in Alaskan, versus European, populations. This indicates that many populations in Alaska may have been large for a long period of time, and also experienced greater levels of genetic exchange.

The Nome region is the only region sampled in Alaska that is clearly within the putative refugium of Beringia. This region does appear to have the highest haplotype diversity, with 15 haplotypes present in 57 individuals in 6 populations. Furthermore, the Teller population in this region (on the northern coast of the peninsula) has the highest haplotype diversity of any population, with 8 haplotypes present in 13 sampled individuals. However the Denali region is not far behind with 15 haplotypes in 97 individuals in 4 populations, and the single population at Fielding Lake has seven haplotypes in the 14 individuals sampled, which is second only to Teller. Fielding Lake, however, has fewer private haplotypes (43%), compared with Denali and Nome, which have 69% and 87%, respectively. There are no haplotypes shared between the Nome region and Fielding Lake that are not also present in Denali. The Denali populations occur in an area that appears in climatic reconstructions to be on the borderline between glaciated and unglaciated regions, on the northern side of the Alaska Range. The Denali region has a population structure most similar to the putative refugial area around Nome. However if Denali were glaciated, perhaps it was near enough to a large refugial area to be re-colonized quickly by many chloroplast lineages. In contrast, the Summit Lake population near Anchorage has only one haplotype in 22 sampled individuals, showing a similar population structure to the populations in other glaciated regions such as Colorado, which has two haplotypes among 35 sampled individuals in two populations.

2.4.3 Biogeography

In spite of the proven usefulness of cpDNA RFLPs for tracking the re-colonization history of species since the ice ages, and for pinpointing potential refugial areas (Ennos *et al.*, 1999; Petit *et al.*, 2003; Abbott *et al.*, 2000), this study is not able to resolve migration routes or identify glacial refugia in which European *S. hirculus* may have survived during the Pleistocene, because many populations of *S. hirculus* in Europe seem to have unique haplotypes. This could be due to a lack of sampling in northern Europe, and particularly in the regions of the Eurasian arctic and also in high mountains further south in central Asia, such as the Altai, where *S. hirculus* is known to occur. It could be that if these regions were sampled, common haplotypes could be found, along with more intermediate haplotypes. Or it could be that many populations have died out, causing intermediate haplotypes to be lost, and obscuring the relationships between populations.

The high cpDNA diversity of *S. hirculus* in Alaska with 31 haplotypes compared with 7 in European populations, along with the different population structures in the two regions, supports theories that the Beringian region acted as a glacial refugium for Arctic plants during the ice ages (Abbott & Brochmann, 2003). Beringia was first proposed as a refugial area by Hultén (1937). Recently molecular evidence has been presented that *Dryas integrifolia* (Tremblay & Schoen, 1999), and *Saxifraga oppositifolia* (Abbott *et al.*, 2000) survived the Quaternary glaciations in this region. Fossil evidence shows that Beringia was covered by various tundra types in this period (Ritchie & Cwynar, 1982; Edwards *et al.*, 2000), and macrofossils of *S. oppositifolia* have been found on the northern part of the Seward Peninsula of Alaska during the last full-glacial (Goetcheus & Birks, 2001).

While no haplotypes are shared between populations of *S. hirculus* in Europe and Alaska, there is no molecular evidence for two distinct lineages as there was in the case of *S. oppositifolia* (Abbott *et al.*, 2000). In general, European *S. hirculus* haplotypes are as likely to be more similar to an Alaskan haplotype as they are to be similar to another European haplotype. There is no relationship between geography and genetic similarity in the

European samples, indicating that the haplotypes arose before they dispersed to their present day locations. This presents something of a paradox because the seeds have no specialized dispersal mechanisms, and populations in the Pennines in Britain only 8km apart and those in Denmark only metres apart, are fixed for different cpDNA haplotypes. *Saxifraga hirculus* does not seem to be a species that can move very easily. However, it should be remembered that there would have been much bare ground and habitat with little competition at the end of the Pleistocene, such that priority effects would have been almost non-existent. Under these conditions dispersal is not so surprising.

2.4.4 Taxonomic implications

The molecular data lend little support to the morphological recognition of four subspecies of *S. hirculus* in the geographical area studied here. Thus, in Europe, material from Iceland and Svalbard is recognised as ssp. *compacta*, distinguished by its short, one-flowered stems from ssp. *hirculus* (taller, at least two-flowered stems) from localities to the south. This distinction is not reflected by the chloroplast haplotype data, where Iceland at least is dominated by a haplotype that occurs also in England, N. Ireland and Denmark. Similarly it is impossible to make a distinction on haplotype evidence between the North American ssp. *propinqua* and the primarily European ssp. *hirculus*, so intermixed are they in the minimum spanning network (Fig. 2.2). It should be admitted, however, that comparison here is confounded by the occurrence in Alaska of specimens that are identifiable as ssp. *hirculus* and the morphology of specimens from which the Alaskan chloroplast haplotypes were taken is unknown. In light of all the above, it is doubtful, therefore, whether much importance should be attached to the finding that the Colorado populations (ssp. *coloradensis*) cluster with other North American material but not with each other in this network (Fig. 2.2).

2.5 Conclusions

Genetic biodiversity is not evenly distributed in *Saxifraga hirculus.* There are clear differences in the amounts and partitioning of genetic diversity at both local and

international scales. In Britain, the divergence of the chloroplast type of the Pentlands population is a complete surprise. Its genetic difference would not have been suspected based on morphological observations alone. The lack of congruence between the genetic data and the sub-specific classification (e.g. morphology) highlights the difficulties of predicting the distribution of genetic biodiversity in the absence of genetic data. At an international scale, there is a clear difference in patterns of population genetic structure of the species in Alaska and Europe and this highlights the importance that historical events can have in shaping the present day distribution of genetic biodiversity. It also highlights the importance of Alaska as a centre of intra-specific genetic biodiversity for Arctic plants.

Chapter 3. Isolation of nine polymorphic microsatellite loci from the Sea Pea (*Lathyrus japonicus*)

3.1 Introduction

Lathyrus japonicus is an insect-pollinated perennial shrub capable of forming extensive clonal mats (Brightmore & White, 1963). It has a circumpolar north arctic distribution and it is exclusively confined to coastal habitats (Hultén, 1971). In Britain it occurs on shingle beaches, or more rarely sand dunes and its range centre in the UK is the south east of England, although there are records of outlying populations scattered around the UK coastline (Akeroyd, 1994). Recently, many populations have been lost or are declining, due predominantly to human disturbance and trampling. The notable decline in the species has led to conservation concern and intervention. In terms of assessing the genetic consequences of population isolation and bottlenecks, the clearly defined narrow coastal habitat of the species facilitates quantification of the occurrence of populations, and the distances between them, and thus this species can be used as a model to investigate the relationships between population size and isolation with genetic variation and fitness.

3.2 Materials, Methods, and Results

To provide a set of polymorphic neutral genetic markers, a set of microsatellite primers were isolated following the modified procedures of Edwards *et al.* (1996), and Squirrell & Wolff (2001) as described by Hughes *et al.* (2002). An enriched genomic DNA library was created by hybridizing the restricted, ligated, and amplified genomic DNA to Hybond (Amersham Pharmacia Biotech) N+ membrane, to which the oligonucleotides $(GA)_{13}$, $(CA)_{13}$ and $(AGG)_8$ had been fixed. The enriched DNA was then cloned using a PCR-Script™ Amp Cloning Kit (Stratagene). Plasmid DNA was extracted using either Quiagen minipreps, or TempliPhi kits (Amersham Biosciences) and the inserts were

sequenced using M13 primers and a Thermosequenase II dye terminator cycle sequencing kit (Amersham) and an ABI 377 sequencer.

Two hundred plasmids were sequenced, and approximately 95 contained microsatellites for which primers could be designed using the program PRIMER-3 (Rozen & Skaletsky, 2000) (http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi). These 95 primer pairs yielded 43 loci that produced reliable single or double bands in the appropriate size range. These loci were labelled with florescent dNTPs, and tested on two samples from each of 10 British populations to establish whether they were polymorphic or not. Nine primer pairs produced polymorphic apparently single locus products and these were fluorescently labelled using the dyes Joe or Fam (MWG-Biotech AG). These labelled primer pairs were then tested on a sample of up to 30 individuals from each of ten British populations of *Lathyrus japonicus* (Table 3.1).

The primer pairs revealed between 3 and 10 alleles per locus with gene diversity values (expected heterozygosity, H_E) ranging from 0.16-0.55. The observed heterozygosity values (H_O) ranged from 0.11-0.36. For all loci the observed heterozygosity was significantly lower than the expected (tested using randomization tests in Fstat; Goudet, 2001). Over all loci there was a deficit of observed heterozygosity leading to a globally significant F_{IS} estimate (F_{IS} = 0.270). The difference between the observed and expected heterozygosities may be attributable to null alleles although given its consistency over loci; or it may alternatively be due to some level of self-pollination. In terms of assessing the presence of null alleles, there were no cases of samples amplifying well across several loci (indicating good DNA quality) but persistently failing for others (indicating the presence of homozygous nulls). This suggests that null alleles are not present at a very high frequency in the data set. These primer pairs thus represent a potentially useful set of microsatellite markers for assessing the amounts and partitioning of genetic diversity in *Lathyrus japonicus*.

Table 3.1. Locus	Repeat motif	Primer sequence	T_{an}	Size range (bp)	No. alleles	H_O	H_E	N
L38x	$(TCT)_{18}TTT(TCT)_3$	L5'-GCATAAGCATTGATTGTCAAAGT-3'	55	186-201bp	5	0.36 *	0.55	~300
		R5'-AAAGAATCACATATCCCTGCAC-3'						
L29t	$(TG)_3CGG(TG)_{22}(CG)_5(TG)_3CG(TG)_5$	L5'-TGTCGCGAGTTCTACCCTATTG-3'	50	200-229bp	8	0.21 *	0.25	~300
		R5'-CTGGACTCACTAATTGTGGTTAAATGT-3'						
L13	$(TG)_{16}(CG)_2(TG)_4CGTGCATG(TG)_8$	L5'-TTCTGATTCAACTGACGCATCT-3'	56	253-285bp	10	0.18 *	0.27	~300
		R5'-CGGAGTTTGAAAAAGAGGAAGC-3;						
N76	$(CA)_4TA(CA)_{12}$	L5'-GAAAGACAAGAGGTGTGAAAACG-3'	55	155-169bp	5	0.37 *	0.41	~300
		R5'-AGAGGCTTTTCAAAGGGCTAAA3'						
N77	$(TA)_7$.....$(TG)_8TC(TG)_3TC$	L5'-GTGAGGAAACTGAGCAACATGA-3;	55	190-227	10	0.35 *	0.46	~300
		R5'-CTTGAGAAAGCACCCATCAACT-3'						
N81	$(CA)_2CGCACG(CA)_3CG(CA)_{13}CTAAAACT(TC)_6$	L5'-GAACGATTGTAAGGCAAAAGGA3'	56	160-179bp	7	0.19 *	0.28	~300
		R5'-TTTTCTCACAAAAGCACTTAGGC-3'						
N96	$(CA)_{12}TA(CA)_6TA(CA)_4CAAA(CA)_{19}$	L5'-TCGCATCTGAGTTATTGGTGTT-3'	57	230-242bp	6	0.25 *	0.30	~300
		R5'-TGATTCATCTGACTAGGCTCCA-3'						
N99	$(TG)_{10}$	L5'-TAAGGTGGGCATCATTTTACTG-3'	57	154-160	4	0.29 *	0.40	~300
		R5'-ACACTGTCATACAGGGTTCTCG-3'						
N164	$(CA)_{11}$	L5'-GTGAAAGCTCGTTTGATCATGT-3'	57	135-143	3	0.11 *	0.16	~300
		R5'-AAAAGTGAGAGGCTTCTTAAGAGTTTT-3'						

Table 3.1. *Lathyrus japonicus* microsatellite characteristics.

Primer sequences, annealing temperature and preliminary population genetic statistics from 10 British populations. T_{an} = annealing temperature, H_O = observed heterozygosity, H_E = expected heterozygosity, n = sample size. *Indicates that the observed heterozygosity value was significantly different from the expected.

Chapter 4. Population genetic structure of *Lathyrus japonicus* in Britain

4.1 Introduction

The maintenance of genetic variation within populations is a major concern in conservation biology. If a population becomes small in size, population genetic theory predicts it will contain low levels of genetic variation. This in turn is predicted to lead to a loss of fitness due to inbreeding depression and an inability to evolve in the face of changing environmental conditions (Frankham *et al.*, 2002). This is considered by some authors to be clear-cut and a direct and pressing issue for conservation biologists (Frankham, 1996). Other authors, however, question the importance of genetic variation and inbreeding as major determinants of the survival of natural populations (Lande, 1988). And while there are an increasing number of papers tackling work in this field, the "search remains open for general patterns as to how inbreeding depression varies among taxa, environments and populations with different demographic and genetic histories" (Keller & Waller, 2002).

One of the potential reasons for the ambiguities surrounding this issue, is the multi-faceted nature of the problem itself, and it is worth recognizing that there are two related components to it. (1) Is there a correlation between population size and levels of genetic variation? (2) Is there a correlation between levels of genetic variation and the fitness of the individuals in question? The first of these questions will be addressed in this chapter, and the second in Chapter 5.

4.1.1 Is there a correlation between population size and levels of genetic variation?

A pre-requisite to the notion that small populations are at greater genetic risk than large populations is that there is a correlation between levels of genetic diversity and population size. Supporting this notion, Frankham (1996) summarized 23 studies of plants and animals and from these concluded that there was a positive correlation between allozyme variation and the logarithm of population size in 22/23 studies. However, in a separate review by Ellstrand & Elam (1993) several studies show non-significant relationships between population size and genetic variation in plants. Likewise, Oostermeijer *et al.* (1994, 1995a,b) studying *Gentiana pneumonanthe* and Schmidt & Jensen (2000) studying *Pedicularis palustris* found that levels of genetic variability were to some degree independent of population size.

4.1.1.1.Why might there be a limit to the correlation between population size and levels of genetic diversity?

There are several factors that will potentially influence the relationship between population size and levels of genetic diversity.

Firstly population size can be thought of as being a historical concept with the long-term historical effective population size (*N*e) being the key determinant of levels of genetic variation. Population size, however, is usually simply measured as the census count (*N*). This is a snap-shot estimate of the population size, and various studies have estimated that *N*e can be as little as 10% of *N* in some populations (Frankham *et al.*, 2002). Factors that can lead to *N*e<*N* include unequal sex ratios, variation in fertility, and age structured populations. Thus if the ratio *N*: *N*e varies among populations (which seems entirely reasonable and likely) then one should expect this to introduce variance into correlations between *N* and levels of genetic diversity.

Secondly, population size is only meaningful in the context of inter-population distances; one should not expect a clear relationship between population size and levels of genetic diversity if this is not considered. A small population adjacent to a large population will experience a greater influx of genetic variation via pollen and seed, compared to that experienced by a small isolated population. This issue of the importance of both size and isolation of populations has led to the development of the concept of 'biological proximity' (Ehlers & Olesen, 2003), effectively a combined measure of these two variables. Biological proximity of neighboring populations may well impact on the levels of intra-population diversity and hence any correlations between population genetic diversity and the census count of individuals (Hanski, 1994).

Thirdly, the correlation between population size and levels of genetic variation is of course sensitive to statistical power. If a small sample of loci that show low levels of allelic diversity are examined, or if a misrepresentative set of individuals are collected, a significant correlation may appear non-significant simply due to a lack of statistical power.

While estimating *Ne* in natural populations is notoriously difficult and time consuming, it is clear that assessing population size in the context of the size and isolation of neighboring populations, and comparing this with genetic variation measured with a powerful suite of genetic markers, is both feasible and achievable in non-model taxa from natural populations. There is thus the opportunity to undertake studies aiming to control for at least two of the three variables mentioned above to assess how demographic factors influence levels of genetic variation and to establish whether a clearer picture emerges.

4.1.2 The study organism: Lathyrus japonicus

The organism selected to examine the relationship between population size and isolation with genetic diversity is the Sea Pea, *Lathyrus japonicus*. *Lathyrus japonicus* is restricted to coastal habitats where it is found predominantly on shingle beaches (occasionally on sand) in a zone spanning from just beyond the high water mark to the point where organic

matter begins to accumulate and vegetation cover becomes dense. This restricted habitat (essentially a 1-dimensional linear distribution of populations around the coastline) simplifies the concept of population isolation and inter-population differences by eliminating the chance of unknown populations occurring inland, and increases confidence in knowing its exact distribution in the study area, as it is a conspicuous plant growing in a narrow well-worked habitat.

Lathyrus japonicus has its range centre in the UK in the south east of England where it is locally abundant, though it occurs at a few sites on the south coast as far west as Dorset, and also occurs in two populations in Scotland (Figure 4.1). Historically its range was larger; the species is extinct from sites in Cornwall and Norfolk. It has been reported to occur transiently on the west coast of Ireland and the Hebridean Islands (Nelson, 2000). Outside the UK *Lathyrus japonicus* has a predominantly northern circumpolar distribution (Hultén, 1971) although it has also has been recorded from southwest South America. In Europe it occurs in Iceland, Finland, Norway, Denmark, Sweden, Germany, Poland and the former USSR, but it is extinct in France (Tutin *et al.*, 1968).

Lathyrus japonicus is a low growing, long lived perennial herb, which dies back in the winter, and re-grows from an extensive root system in spring. It occurs in large clumps, sometimes several metres wide, and vegetative spread is undoubtedly important. *L. japonicus* is pollinated by bumblebees and does not set seed if pollinators are excluded from an inflorescence (Brightmore & White, 1963). This has led to suggestions that the species is self-incompatible (Akeroyd, 1994). However, although the flowers are protandrous, viable pollen is still present at the time the stigma is receptive (Asmussen, 1993). The seeds are large with a hard outer covering and can be dispersed long distances in the ocean. Seeds can retain their viability in the sea for up to 5 years (Brightmore & White, 1963). Drift seeds of *L. japonicus* are reported sometimes in large numbers on the west coast of Ireland, Cornwall, and the Hebrides in Scotland (Nelson, 2000), though plants are rare and transient in these areas.

Lathyrus japonicus is sensitive to disturbance and trampling during the growing season, and has recently disappeared from many of its historical sites in Britain. This has led to

direct conservation interest and action in the form of protection and restoration programmes. This species is also an interesting model taxon for conservation genetic studies, because while it has declined and has some small and some isolated populations, large healthy populations are still present, and the species has a range of populations of differing sizes and different levels of isolation.

4.1.3 Study aims

The aims of this chapter are to assess patterns of population genetic variation in the context of the size and isolation of populations of *Lathyrus japonicus*. Specifically this involves sampling populations that differ in size and levels of isolation and establishing the extent to which these variables impact on patterns of population genetic structure. To provide a dynamic assessment of the impacts of population size and isolation, population genetic structure has been assessed not only from adult plants sampled from wild populations, but also from seedling populations derived from seed collected from these adult plants.

Once the population genetic structure of these populations have been assessed using neutral markers in this chapter, these data will be used as a framework for comparisons with quantitative characters in Chapter 5.

4.2 Materials and Methods

4.2.1 Sampling strategy

The goal of the sampling strategy was to identify large and small populations in the range centre of the species in the south east of England, and then include both large and small populations isolated to varying spatial degrees outside of this range centre (heading west along the south coast, and heading north along the east coast). However, while conceptually

appealing, this sampling strategy proved difficult to implement. In the first year of the study a general survey of populations was undertaken with the main challenge being the identification of small and isolated populations. However, several of the outlying small populations for which recent records were available were not found during the course of the study. For instance, the population of *L. japonicus* on Shetland was not re-found during a field visit in March 2003, due to storms moving large amounts of rock and shingle. Two small isolated populations (Hurst castle, and Sizewell-Thorpeness), were visited during the early stages of the project in 2000, but when a second visit was undertaken for sampling in 2001 the populations had been lost due to movement of shingle at Hurst castle, and for unknown reasons on the stretch of beach from Sizewell to Thorpeness. This loss of populations hampered the material available for sampling. Furthermore, one population extant at Cley, in Norfolk, (Figure 4.1) was not sampled as correspondence with local botanists suggested that it was planted and is not a natural population (Ellis, personal communication, 2000).

In total, 11 populations were eventually chosen as sample sites. These populations are summarised in Table 4.1 below, and marked on Figure 4.1. The sampling strategy at each site was as follows.

At each of the eleven sites visited, a rough estimate of the population size and extent was made by walking along the beach in either direction, with the population (census, *N*) size estimated from the number of discrete clumps. To establish the extent to which visual clumps are made up of single genets, two clumps from Kessingland were sampled with 10 leaves taken distributed around the extent of the clump. One of the clumps was the typical circular spreading mat that *Lathyrus japonicus* often develops (Figure 4.2). The second was a more diffuse linear clump (Figure 4.3). Intensive within-clump sampling was undertaken for the single circular clump at Nigg Bay (see below).

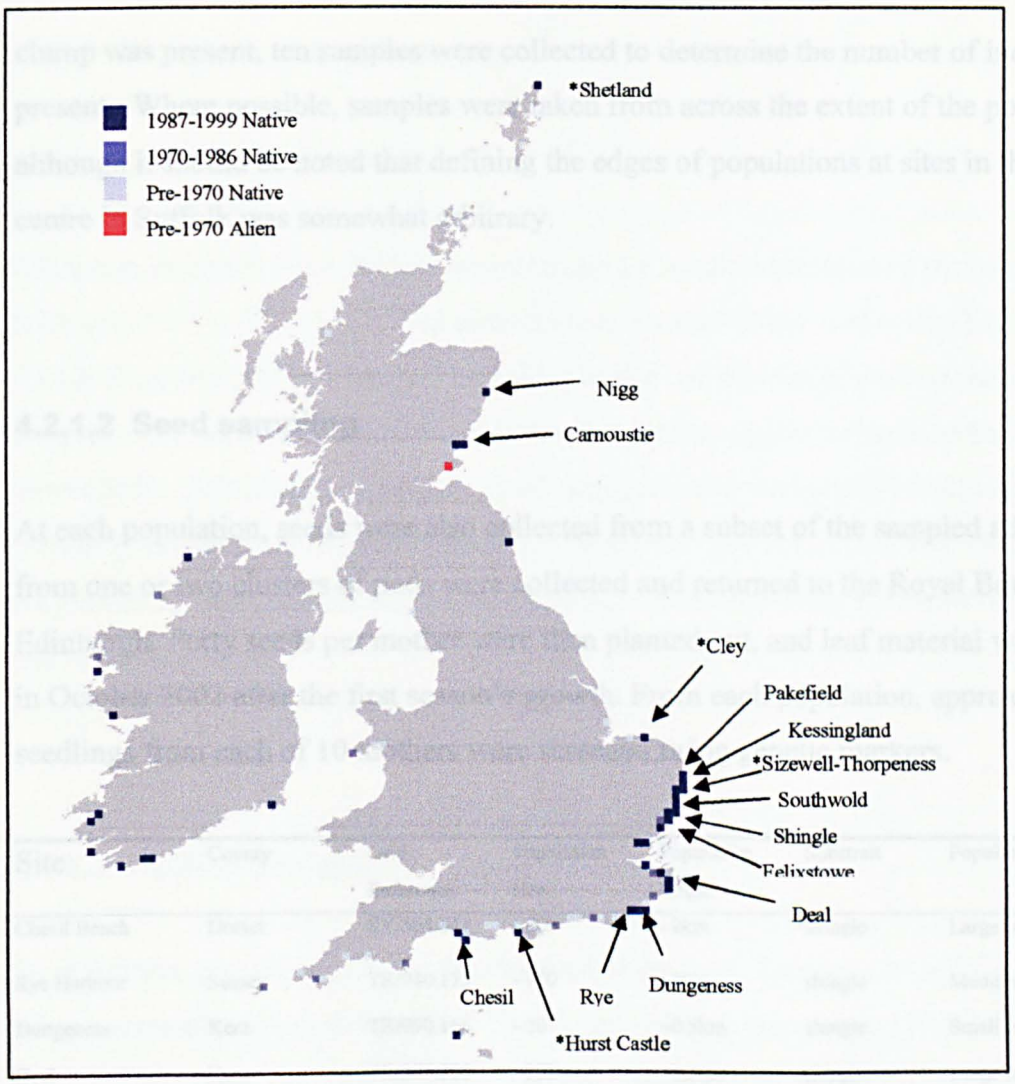

Figure 4.1. Distribution of *Lathyrus japonicus* in Britain, with the study locations indicated.
The distribution map is modified from Preston *et al.* (2002). * indicates populations that were not sampled but discussed in the sampling strategy text.

4.2.1.1 Adult sampling

From each of the populations, leaf samples were taken for DNA extraction from up to 30 distinct clumps, and immediately frozen in liquid nitrogen. If a population contained less than thirty clumps, every distinct clump present was sampled. At Nigg Bay, where only one

clump was present, ten samples were collected to determine the number of individuals present. Where possible, samples were taken from across the extent of the population, although it should be noted that defining the edges of populations at sites in the range centre in Suffolk was somewhat arbitrary.

4.2.1.2 Seed sampling

At each population, seeds were also collected from a subset of the sampled adults. Seeds from one or two clusters of pods were collected and returned to the Royal Botanic Garden Edinburgh. Forty seeds per mother were then planted out, and leaf material was harvested in October 2002 after the first season's growth. From each population, approximately 4 seedlings from each of 10 mothers were screened using genetic markers.

Site	County	Grid Reference	Population size	Population length	Substrait	Population type
Chesil Beach	Dorset	SY/564.841	>500	~2km	shingle	Large isolated
Rye Harbour	Sussex	TR/940.175	~150	~2km	shingle	Medium isolated
Dungeness	Kent	TR/090.166	~30	~0.5km	shingle	Small isolated
Deal	Kent	TR/380.500	~230	~1.5km	shingle	Large range centre
Felixstowe Ferry	Suffolk	TM/325.370	~85	~1km*	shingle	Medium range centre
Shingle Street	Suffolk	TM/365.423	>500?	~2km*	shingle	Large range centre
Southwold Denes	Suffolk	TM/507.752	~100	~1km	sand	Medium range centre
Kessingland	Suffolk	TM/537.857	~700	~2km	shingle	Large range centre
Pakefield Beach	Suffolk	TM/539.905	~260	~0.5km	sand/shingle	Large range centre
Carnoustie	Angus	NO/560.338	~100	~0.5km	sand	Medium isolated
Nigg Bay	Aberdeenshire	NJ/965.047	1 clump	~0.25km	sand/gravel	Small isolated

Table 4.1. Location and characteristics of study sites for *Lathyrus japonicus*.

* indicates populations that are part of a more or less continuous distribution in which neighbouring populations are separated by only a small (and potentially subjective) discontinuity. The cut-off for whether a population is large or small, isolated or not is also essentially somewhat arbitrary, but here small = <50 plants, medium = 51-200, large = >200 plants. Isolated populations are those populations away from the range centre of the south east of England.

4.2.2 Molecular approaches

4.2.2.1 DNA extraction

DNA was extracted from the leaf samples using a modified version of the protocol in Doyle & Doyle (1990). This was scaled down to use ~ 1 cm^2 of dry leaf with 400 µl of 2 X CTAB (cetyltrimethylammonium bromide) buffer and a pinch of acid washed sand. The samples were homogenised in an eppendorf tube using a glass rod attached to a domestic power drill. DNA samples were assessed alongside a HyperLadder I molecular weight marker (Bioline) by electrophoreses in 1.0 % agarose gels run in 1 x TBE (tris borate-ethylenediaminetetraacetic acid). The DNA was visualised using ethidium bromide staining and ultra-violet light.

4.2.2.2 CpDNA analysis

To search for genetic markers in the chloroplast genome, 8 regions were PCR amplified using universal primers designed by Hamilton (1999), Chiang *et al.* (1998), Taberlet *et al.* (1991), Demesure *et al.* (1995), and Dumolin-Lapegue *et al.* (1997). This was undertaken for a subset of the samples representing the full geographic range of the British populations, along with a single sample from the western USA. The regions *trn*L-*trn*F, *trn*S-*trn*G, and *psb*B-*psb*F were sequenced using an ABI 377 sequencer and aligned using Sequence Navigator. The regions ccmp4L-*atp*H, *rpo*C2f-*rpo*C2r, *psb*B-*psb*B, *trn*S-*trn*R, and *pet*B-*pet*D were each digested with four or five restriction enzymes. The enzyme digests were carried out on 10µl PCR products, to which 0.2µl of enzyme, 0.2 µl of BSA, 2µl of 10 X buffer and 7.6µl H_2O, were added. The enzyme digestions were carried out at 37°C in a water bath for 3 hrs. The products were visualized using polyacrylamide gels stained with ethidium bromide and ultra violet light.

4.2.2.3 Microsatellite analysis

Nine microsatellite primer pairs were developed for *Lathyrus japonicus* (Chapter 3). The population samples were then PCR (Polymerase Chain Reaction) screened for these microsatellite regions. The reactions were carried out in a volume of 10 µl with 1x *Taq* buffer (16mM $(NH_4)_2SO_4$, 67mM Tris-HCl (pH 8.8), 0.01% Tween-20), 2 mM $MgCl_2$, 100 µM dNTPs, 200 nM of forward and reverse primer (MWG Biotech) and 1 unit of *Taq* DNA polymerase (Bioline). The PCR cycle was 10 mins at 95 °C followed by 10 cycles of 94 °C for 15 sec, an annealing temperature of 50-58 °C for 15 sec and extension of 72 °C for 15 sec, then 20 cycles of 89 °C for 15 sec, 50-58 °C for 15 sec, and 72 °C for 15 sec. Then the samples were held at 72 °C for 30 min for a final extension. The products were visualized using an ABI 377 sequencer, and the resulting peaks analysed and sized using the computer program Genotyper (Applied Biosystems Inc.).

The adult samples were genotyped for all nine loci described in Chapter 2. These nine loci were used for assessing patterns of population genetic structure in the adults. However, as two of these loci (N99, N164) showed little variability, the seedlings were only screened for seven loci. To ensure comparisons between adult and seedling populations are meaningful, any direct comparisons between the adult and seedling data are based on the same 7 loci.

4.2.3 Data analysis

To check that the microsatellite loci were independent, a test for linkage disequilibrium was undertaken on the adult samples using Fstat (Goudet, 2001). No linkage disequilibrium was detected.

4.2.3.1 Intra-population genetic diversity

To assess levels of intra-population genetic diversity, the mean number of alleles per locus (*A*), the proportion of loci that are polymorphic (*P*), the mean gene diversity (expected heterozygosity, *He*), and the observed heterozygosity (*Ho*) were calculated using GDA (Lewis & Zaykin, 2001). To test for differences between observed and expected heterozygosities, F_{IS} was estimated using Fstat (Goudet, 2001) and the significance of deviations from Hardy-Weinberg equilibrium evaluated by permutation tests. $S = 2\ F_{IS}\ /1+\ F_{IS}$ was used as a rough approximation of the selfing rate (*S*), with the out-crossing rate (t) approximated from $t = 1\text{-}S$. This estimation of *t* from F_{IS} rests on a number of assumptions, including that the populations are at inbreeding equilibrium. All of these statistics were calculated for both the adult samples and the seedling samples. Generalised statistical comparisons between populations were undertaken using SPSS (SPSS Inc., 2001).

To formalise estimations of breeding behaviour, multi-locus out-crossing rates were then estimated from the seedling arrays (and maternal genotypes) from each of the populations using MLTR (Ritland, 2002). This estimation differs from direct inspection of the data because both the pollen and egg gene frequencies are considered, thus correcting for underestimation caused when out-crossing events pass on the same alleles present in the mother (Ritland, 2002).

4.2.3.2 Inter-population genetic diversity

Assessments of the amounts of differentiation between populations were made using Weir & Cockerham's (1984) estimators of Wright's (1978) F-statistics using Fstat (Goudet, 2001) with their significance assessed by permutation tests. To assess whether there were mutational differences (as opposed to drift-based differences) between populations, Goodnight's estimate of R_{ST} was calculated using Fstat (Goudet, 2001) following Rousset (1996), and Goodman (1997).

To test for a correlation between genetic distances and geographic distances, pairwise geographical distances (the shortest straight line distance around the coastline between populations) were compared against pairwise F_{ST} estimates using a Mantel test implemented in Arlequin (Schneider *et al.*, 2000). This was undertaken using the entire data set, as well as varying subsets of populations. To investigate the homogeneity of individuals within populations, an assignment test was carried out using DOH (Brzustowski, 2002; Paetkau *et al.*, 1995).

4.3 Results

4.3.1 CpDNA variability

In approximately 2200 bp of chloroplast DNA sequenced, no variation was detected. Of the 5 regions examined for RFLP variation, only one difference was detected. This was in the region *trn*S-*trn*R, cut with enzyme *Hinf* I. This difference, however, was between a sample of *L. japonicus* from the western USA and the British samples. As there was no useful variation detected in the British populations, the cpDNA data are not further considered.

4.3.2 Genetic diversity measures

4.3.2.1 Clonal diversity measures

The 10 samples taken from the intensively sampled circular patch at Kessingland (Figure 4.2) all possessed identical multi-locus genotypes consistent with this patch being a single clone. The 10 samples taken from the more diffuse linear patch at Kessingland (Figure 4.3) had one of two genotypes consistent with this patch being composed of two intermingled clones. There was no clear spatial aggregation of the genotypes within this patch. The 10

Figure 4.2. Circular patch of *Lathyrus japonicus* sampled to assess clonal diversity at Kessingland.

Figure 4.3. Linear patch of *Lathyrus japonicus* sampled the assess clonal diversity at Kessingland.

samples taken from the sole clump present at Nigg Bay all shared the same multi-locus genotype consistent with this clump being uni-clonal.

4.3.2.2 Population genetic diversity of adult plants

The percentage of the microsatellite loci that were polymorphic in each population was P = 100%, except for the populations from Deal, Shingle, Southwold, and Pakefield, where P = 89% (e.g. one out of the 9 loci was monomorphic in these populations), and Nigg where P = 22% (only two loci polymorphic) (Table 4.2). Excluding the uni-clonal Nigg population, the mean number of alleles per locus varied from A = 4.67 at Rye Bay, to A = 2.11 at Carnoustie, with an average of A = 2.9 (Table 4.2). The observed heterozygosity ranged from Hobs = 0.44 at Rye Bay to Hobs = 0.14 at Kessingland, with an average of Hobs = 0.26 (Table 4.2). The expected heterozygosites were consistently higher than the observed, and ranged from He = 0.52 at Dungeness to He = 0.23 at Pakefield, with an average of He =0.36 (Table 4.2).

Population	N	P	A	Ho	He	Population type	Location
Chesil	30	1.000	2.778	0.293	0.357	Large isolated	South coast
Rye	29	1.000	4.667	0.444	0.517	Medium isolated	South coast
Dungeness	29	1.000	4.222	0.360	0.521	Small isolated	South coast
Deal	30	0.889	2.556	0.222	0.358	Large range centre	South East coast
Felixstowe	28	1.000	2.556	0.206	0.317	Medium range centre	South East coast
Shingle	29	0.889	2.778	0.226	0.304	Large range centre	South East coast
Southwold	24	0.889	4.000	0.218	0.299	Medium range centre	South East coast
Kessingland	30	1.000	4.111	0.141	0.308	Large range centre	South East coast
Pakefield	29	0.889	2.778	0.169	0.233	Large range centre	South East coast
Carnoustie	30	1.000	2.111	0.315	0.384	Medium isolated	Scotland
Nigg	10 (1genet)	0.222	1.222	0.222	0.117	Small isolated	Scotland
Mean (Nigg excluded)		0.956	2.856	0.259	0.360		

Table 4.2. Intra-population measures of genetic diversity.

For each population, sample size (n), proportion of polymorphic loci (P), the mean number of alleles per locus (A), observed heterozygosity (Ho) and expected heterozygosity (He).

4.3.2.3 The relationship between population genetic diversity and population size and isolation

Figure 4.4 shows levels of population genetic diversity in the context of the size and isolation of the study populations. There is no clear relationship between the size, and/or isolation of populations with *A*, *He*, or *Ho*. The small isolated population at Nigg has the lowest number of alleles per locus (but note that only one genotype is present). The range centre populations do not consistently have higher levels of genetic variation than more isolated populations. The isolated Scottish populations have the lowest levels of allelic diversity, but the isolated south coast populations have, on average, higher levels of allelic diversity than the range centre populations, and the differences between these groups are significant (Kruskal-Wallis test, Chi-squared – 6.87, P = 0.032). Both expected and observed heterozygosity is consistently *lower* in the range-centre populations than the range-edge populations (Figure 4.5). Considering populations with more than one individual (e.g. all other than Nigg), the smallest population (Dungeness) has the highest levels of gene diversity (*He*) and the second highest level of observed heterozygosity (*Ho*) (Figure 4.5).

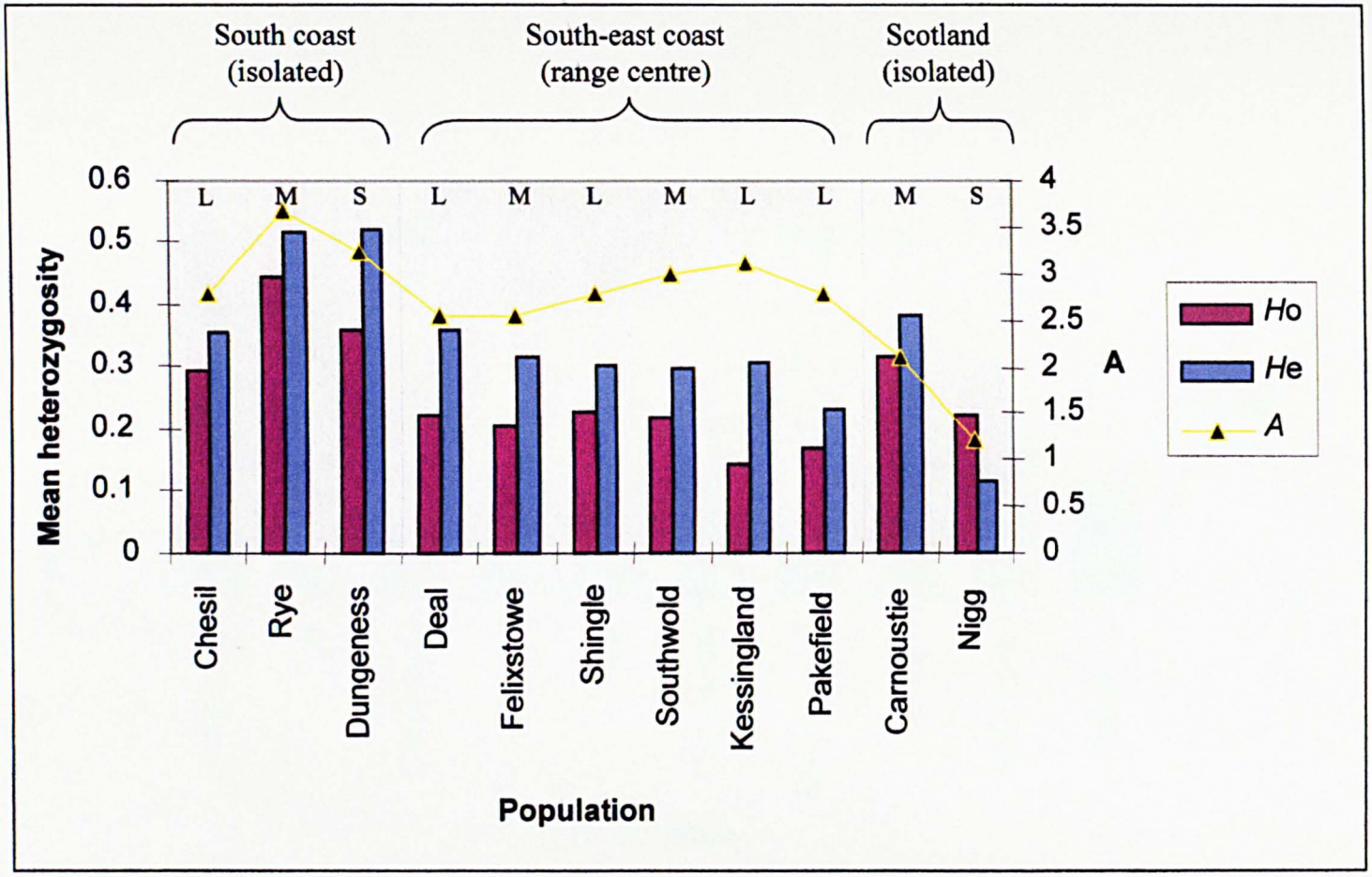

Figure 4.4. Allelic diversity (*A*) and expected (*He*) and observed (*Ho*) heterozygosity of British populations of *Lathyrus japonicus*.

L = large population, M = medium population, S = small population.

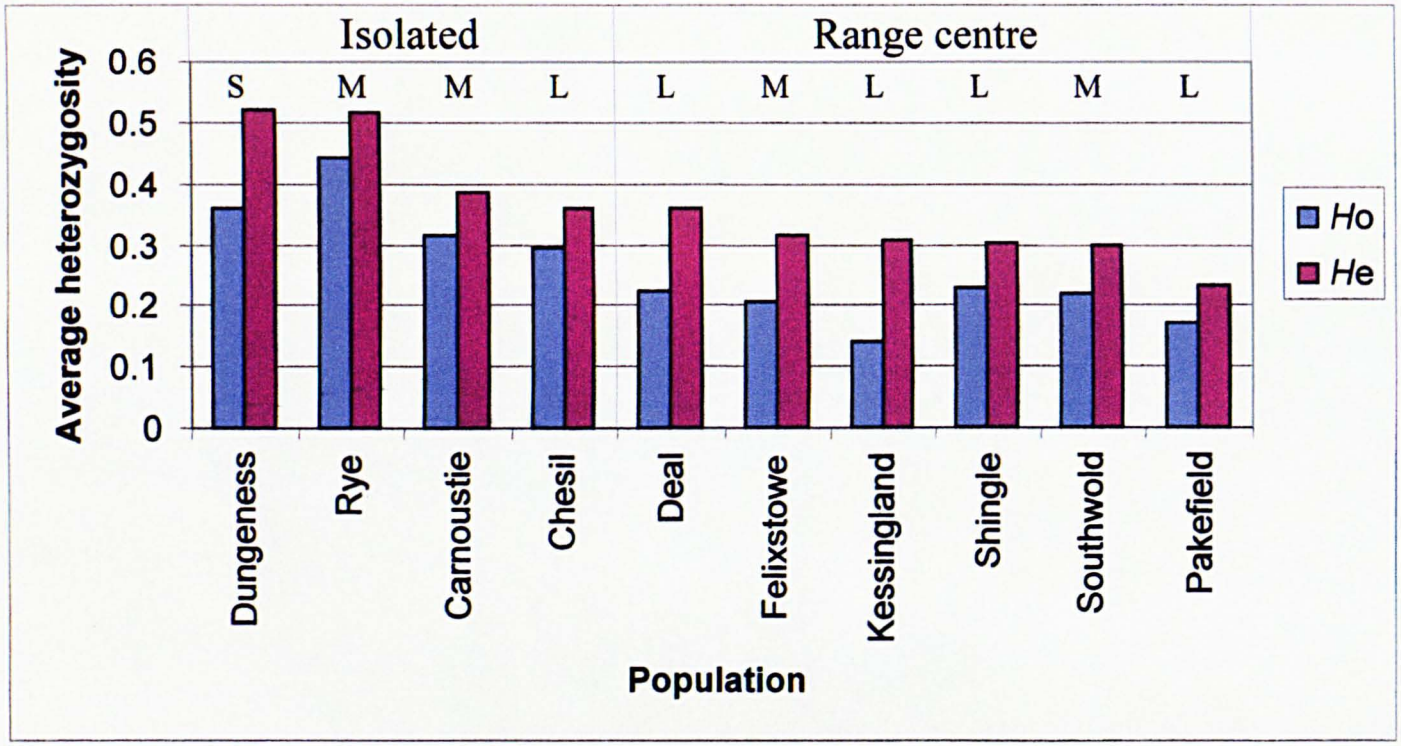

Figure 4.5. Expected (*He*) and observed (*Ho*) heterozygosity in relation to population isolation.

L = large population, M = medium population, S = small population.

4.3.3 Breeding behaviour

In field sampled adult *L. japonicus* the mean inbreeding coefficient over all loci and populations is $F_{IS} = 0.27$ which is significantly different from zero (Table 4.3). In all of the adult populations tested a deficit of heterozygotes was found, and in all populations except for Carnoustie the F_{IS} estimates were significantly different from zero (Table 4.4). At Carnoustie there was still a deficit of heterozygotes, but this was not significant (Table 4.4).

	Ho	*He*	*Ht*	F_{IT}	F_{IS}	F_{ST}	R_{ST}
L38x	0.357	0.545	0.757	0.512	0.341	0.258	0.345
L29t	0.210	0.245	0.540	0.629	0.141	0.568	0.558
N76	0.374	0.405	0.590	0.442	0.156	0.339	0.412
N77	0.345	0.457	0.674	0.498	0.244	0.335	0.761
N81	0.185	0.281	0.490	0.636	0.342	0.448	0.143
L13	0.178	0.274	0.492	0.652	0.352	0.463	0.618
N96	0.253	0.296	0.334	0.334	0.273	0.084	0.085
N99	0.289	0.399	0.473	0.379	0.275	0.145	0.149
N164	0.112	0.158	0.189	0.410	0.286	0.173	0.136
Total	0.256	0.340	0.504	0.516	0.270	0.337	0.3616
95% confidence interval				0.447-0.580	0.219-0.315	0.243-0.426	

Table 4.3. Estimates of the partitioning of microsatellite variation within and among individuals and populations of *Lathyrus japonicus*.

For each locus, and over all loci: Nei's estimation of observed heterozygotes (*Ho*), within sample gene diversity (*He*), overall gene diversity (*Ht*), Weir and Cockerham's estimators of Wright's F-statistics F_{IT}, F_{IS}, and F_{ST}, and Goodnight's estimator of R_{ST}.

Pop.	Chesil	Rye	Dungeness	Deal	Felixstowe	Shingle	Southwold	Kessingland	Pakefield	Carnoustie
Adult F_{IS}	0.182*	0.142*	0.312*	0.383*	0.353*	0.259*	0.276*	0.280*	0.547*	0.183
Seedling F_{IS}	0.356*	0.287*	0.545*	0.561*	0.357*	0.319*	0.534*	0.698*	0.583*	0.471*
Adult *t*	0.69	0.75	0.52	0.45	0.48	0.59	0.57	0.56	0.29	0.69
Seedling *t*	0.47	0.55	0.29	0.28	0.47	0.52	0.30	0.18	0.26	0.36

Table 4.4. Estimates of F_{IS} and estimates of out-crossing rates (*t*) derived from F_{IS} for British populations of *Lathyrus japonicus*.

The population from Nigg bay is excluded. * Indicates a significant F_{IS} estimate ($P < 0.01$).

For the seedling populations the mean inbreeding coefficient over all loci and populations was $F_{IS} = 0.47$ which is significantly different from zero. In all of the seedling populations tested, a deficit of heterozygotes was found, and all F_{IS} estimates were significantly different from zero (Table 4.4).

Levels of observed heterozygosity were significantly higher in the adult populations than in the seedling populations, but the expected heterozygosites of seedlings and adults were not significantly different (Table 4.5).

	Ho seed/*Ho* adult	*He* seed / *He* adult
Z	-2.366	-0.762
Asymp Sig.(2-tailed)	0.018	0.446

Table 4.5. The Wilcoxon Signed Ranks Test for heterozygosity in adult and seedlings.

The observed heterozygosity (*Ho*) of the seedling population in cultivation is significantly different from the adult population at the 5% level. However the expected heterozygosity (*He*) of the seedling population in cultivation is not significantly different from that of the adult population.

In all cases F_{IS} estimates from the seedlings were greater than from the adult populations (Table 4.4), and there was a significant difference between global adult and global seedling F_{IS} estimates ($P<0.01$). There was no clear relationship between isolation or population size with F_{IS} estimates (Figure 4.6), and indeed, for the adult populations, the largest F_{IS} estimates were from the range centre, rather than the more isolated populations. F_{IS} estimates for the seedling populations showed no relationship with either population size or isolation (Figure 4.6).

Estimates of out-crossing rates of adults derived from F_{IS} ranged from $t = 0.75$ at Rye Bay, to $t = 0.29$ at Pakefield, whilst F_{IS} derived estimates of seedling out-crossing rates were consistently lower and ranged from $t = 0.55$ at Rye Bay to $t = 0.18$ at Kessingland (Table 4.6).

Multi-locus out-crossing rates calculated from Mltr ranged from $tm = 0.37$ at Chesil, to $tm = 0.11$ at Southwold and Dungeness (Table 4.6).

Out-crossing rates calculated directly from the genotypes of seedlings (Table 4.6) had wide confidence intervals, generally ranging between 20 and 35 percent, due to small sample sizes. In all cases, however, the out-crossing rate that comprised the upper limit of the confidence interval was lower than the out-crossing rate calculated from the adult allele frequencies in the microsatellite data (Table 4.6). The direct measures of out-crossing rates were lower than the F_{IS} derived measures from seedlings (Figure 4.7), but the confidence limits of the direct measures overlapped the F_{IS} derived measures in several cases (Table 4.6).

Population	Out-crossing rate from F_{IS} Adult	Out-crossing rate from F_{IS} Seed	Out-crossing rate from Mltr	N	Direct out-crossing rate from seeds	95% confidence interval lower-upper bounds
Chesil	0.69	0.47	0.37	30	0.33	0.15 - 0.51
Rye	0.75	0.55	0.30	34	0.26	0.10 - 0.41
Dungeness	0.52	0.29	0.11	38	0.10	0.00 - 0.20
Deal	0.45	0.28	0.21	34	0.21	0.06 - 0.35
Felixstowe	0.48	0.47	0.33	31	0.29	0.12 - 0.46
Shingle	0.59	0.52	0.22	22	0.18	0.01 - 0.36
Southwold	0.57	0.30	0.11	35	0.09	-0.01 - 0.18
Kessingland	0.56	0.18	0.13	37	0.14	0.02 - 0.25
Pakefield	0.29	0.26	0.15	33	0.15	0.02 - 0.28
Carnoustie	0.69	0.36	0.18	37	0.16	0.04 - 0.29

Table 4.6 Out-crossing rates calculated directly from progeny arrays compared to those derived from F_{IS} and calculated in Mltr.

N = the sample size used for the direct estimate of out-crossing rates

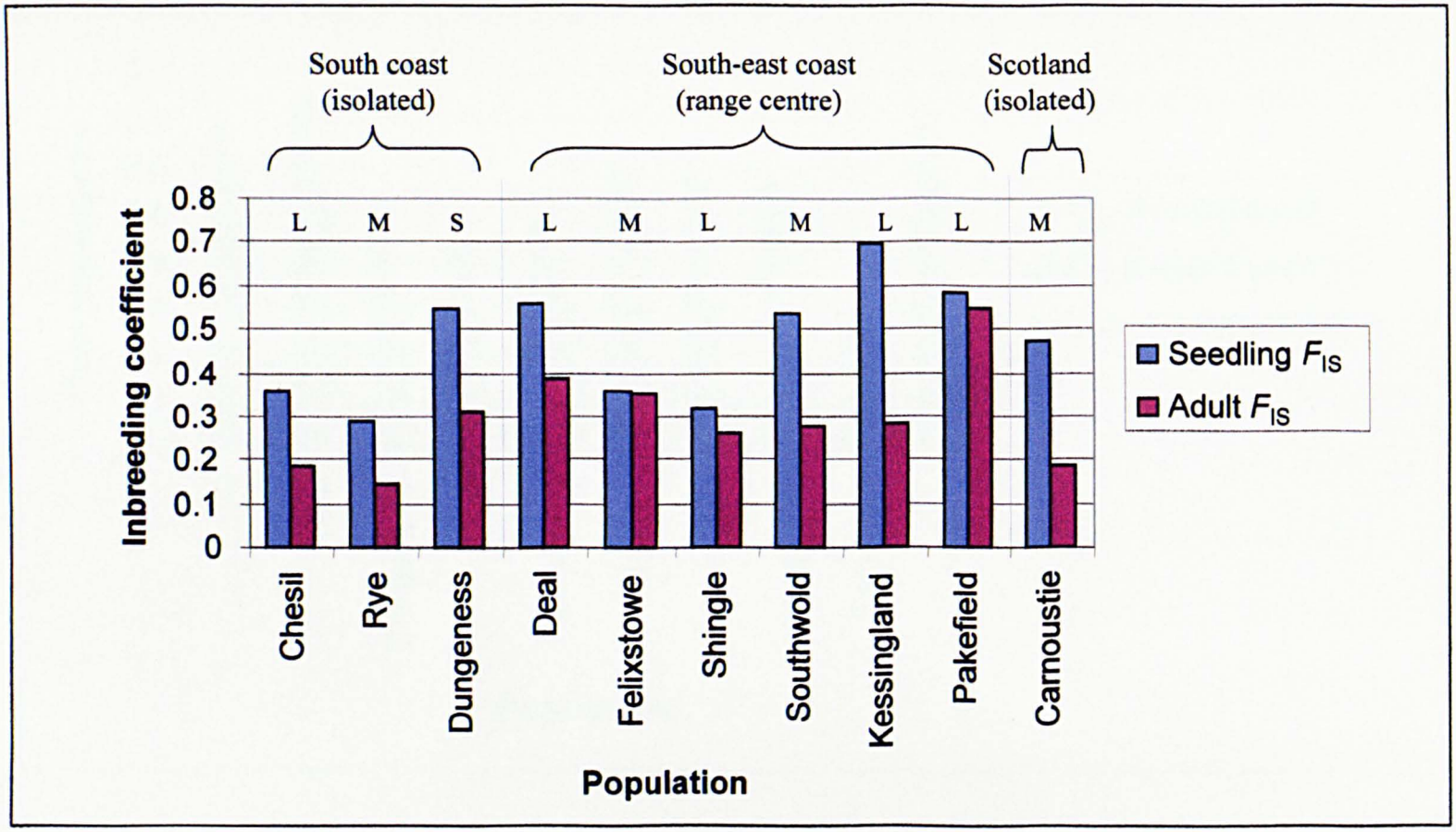

Figure 4.6. Mean F_{IS} estimates by population for adult plants and seedlings of *Lathyrus japonicus*.

L = large population, M = medium population, S = small population.

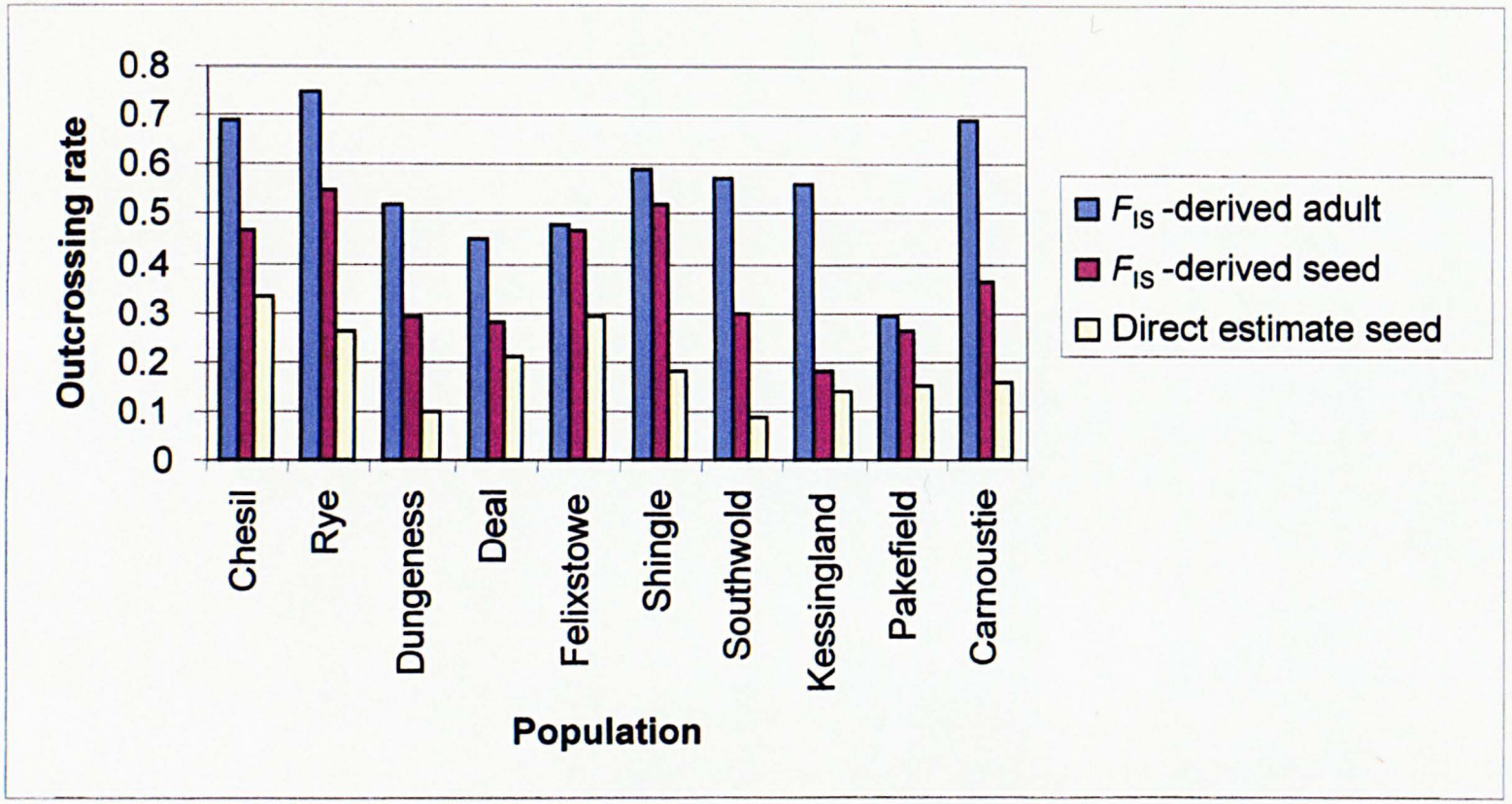

Figure 4.7. Outcrossing rate estimates for *Lathyrus japonicus* from F_{IS} and direct assessment of outcrossing via progeny arrays

4.3.4 Population differentiation

4.3.4.1 Global estimates of population structure

The global estimate of F_{ST} = 0.34, and was significantly different from zero (Table 4.3). A very similar estimate of population differentiation was obtained from R_{ST} = 0.36 (Table 4.3). These estimates of population differentiation were in turn very similar to that derived from the seedling populations (e.g. seedling F_{ST} = 0.39, not significantly different from the adult F_{ST}).

The distribution of private alleles (those only found in a single population) is given in Table 4.7. Of the 18 cases of private alleles, only 4 involved populations in the range centre (two from Kessingland, one from Deal, and one from Pakefield). Three of these were rare alleles present at a frequency of <0.05, and the remaining private allele from Deal was only at a frequency of 0.067. The other 14 private alleles in the data set were from the Scottish populations (Carnoustie has 5/18 of the private alleles) and the South Coast (Chesil has 5/18, and Rye has 4/18 of the private alleles). In contrast to the range-centre private alleles, those in isolated populations were sometimes at a high frequency with Carnoustie and Chesil in particular having some high frequency private alleles.

Locus	Allele	Frequency	Population restricted to
L38x	201	0.15	Chesil
L29t	210	0.583	Carnoustie
L29t	225	0.138	Rye
L29t	221	0.917	Chesil
N76	155	0.017	Kessingland
N76	167	0.017	Rye
N77	221	0.600	Carnoustie
N77	217	0.400	Carnoustie
N77	200	0.017	Rye
N77	227	0.617	Chesil
N81	160	0.300	Carnoustie
N81	173	0.033	Kessingland
N81	162	0.067	Deal
L13	279	0.017	Carnoustie
L13	271	0.017	Rye
L13	285	0.067	Chesil
L13	259	0.717	Chesil
N96	238	0.017	Pakefield

Table 4.7. Distribution of private alleles and their frequency in populations of *Lathyrus japonicus*.

4.3.4.2 Genetic/geographical structure

There is a significant correlation between pairwise estimates of F_{ST} and pairwise inter-population geographical distances ($P = 0.02$; Table 4.8). When the population (single individual) from Nigg Bay is excluded, the significance increases ($P < 0.01$) (Table 4.8). When only the Suffolk populations are included in the analysis the relationship between genetic and geographic distance breaks down ($P = 0.21$) (Table 4.8).

	Suffolk populations	Suffolk plus Deal	All populations except outliers	UK wide, Nigg excluded	UK wide, Nigg included
P	0.208860	0.084160	0.000200	0.00000	0.019840

Table 4.8. Significance of Mantel tests correlating genetic (F_{ST}) and geographic distance for different sets of populations.

Suffolk populations are Felixstowe, Shingle, Southwold, Kessingland and Pakefield. Outlying populations are Chesil, Carnoustie, and Nigg.

When the pairwise F_{ST} values for all *L. japonicus* populations are plotted against the geographic distances for each population the data can be described with a line ($r^2 = 0.40$), but it can be better described by a curve ($r^2 = 0.52$) (Figure 4.8)

When the single individual that makes up the population at Nigg Bay is excluded, the linear relationship becomes stronger ($r^2 = 0.60$), but the relationship is still best explained with a curve ($r^2 = 0.80$) (Figure 4.9).

The points making up the descent of the curve are pairwise distances involving the population from Carnoustie, indicated with cross-like symbols in Figure 4.10. The populations from the south coast (Chesil, Rye, and Dungeness) that are geographically farthest from Carnoustie are genetically more similar to Carnoustie than are the Suffolk populations (Figure 4.10).

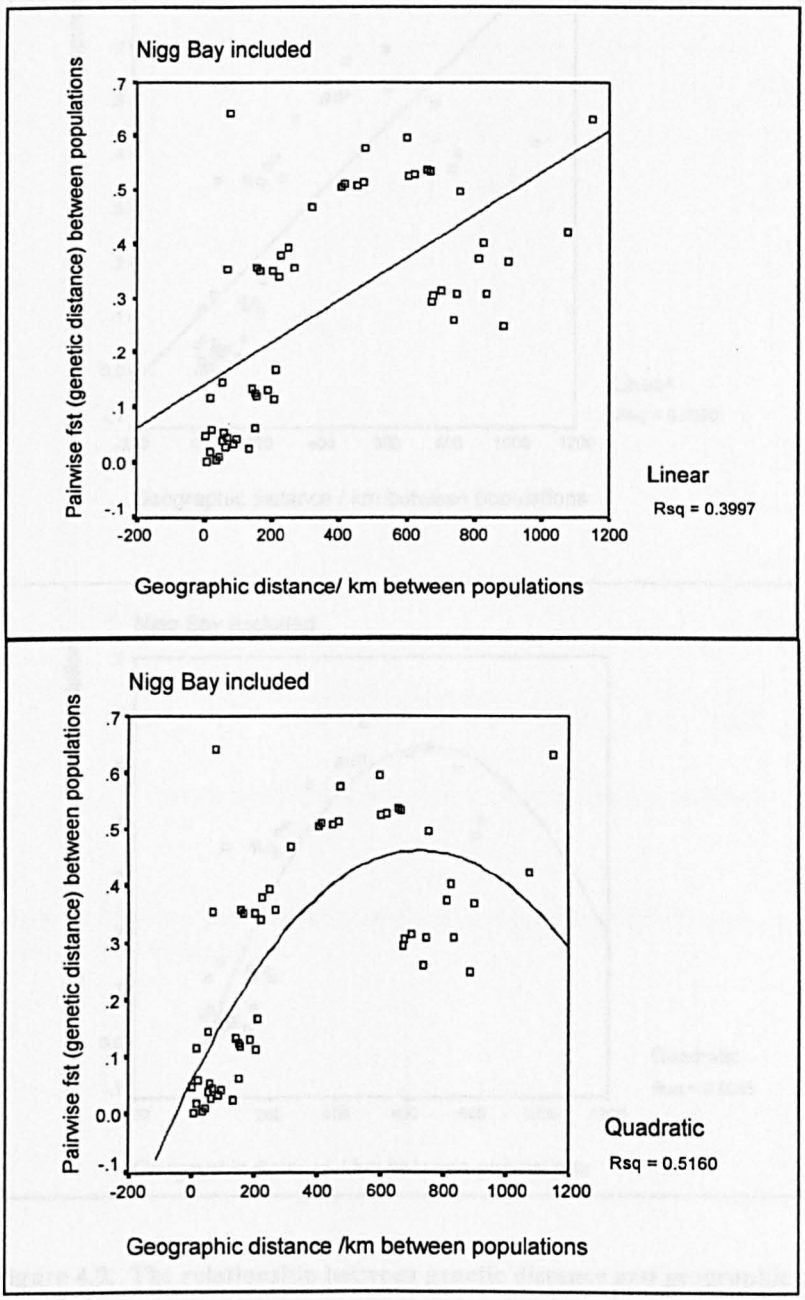

Figure 4.8. Relationship between genetic and geographic distance, Nigg Bay included.

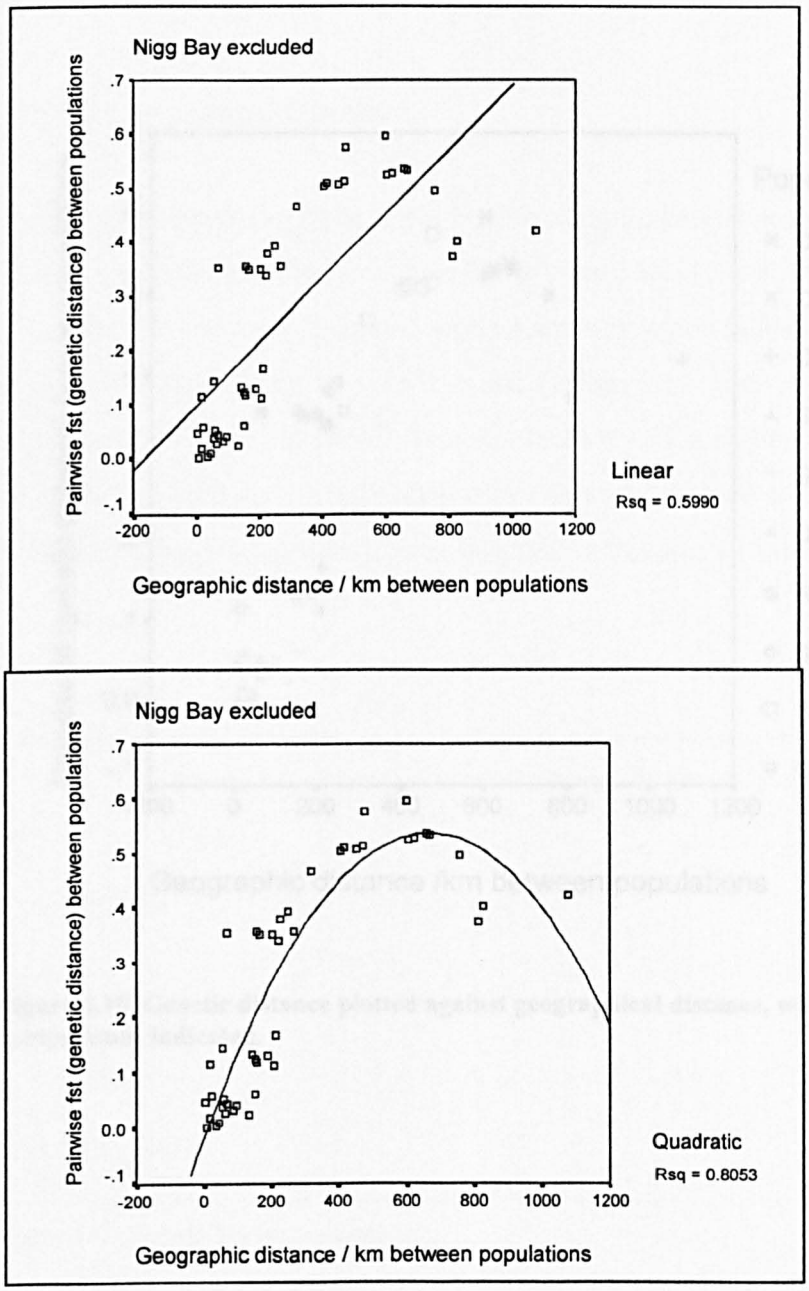

Figure 4.9. The relationship between genetic distance and geographic distance Nigg Bay population excluded.

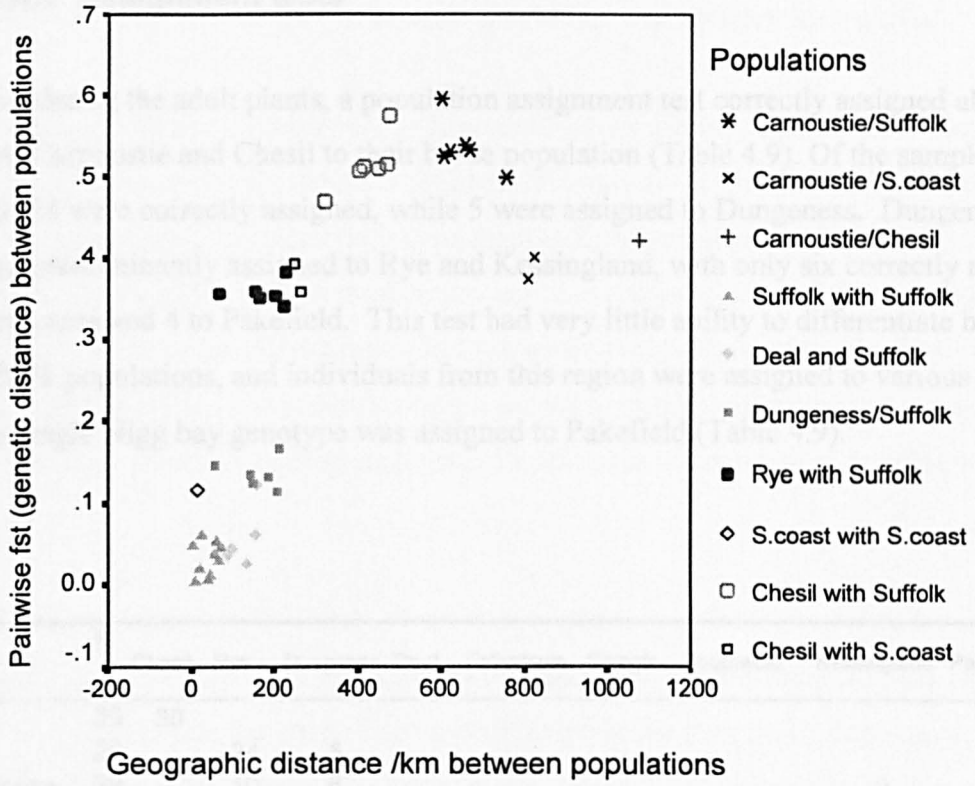

Figure 4.10. Genetic distance plotted against geographical distance, with the pairwise population comparisons indicated.

4.3.4.3 Assignment tests

Considering the adult plants, a population assignment test correctly assigned all 30 samples from Carnoustie and Chesil to their home population (Table 4.9). Of the samples from Rye Bay, 24 were correctly assigned, while 5 were assigned to Dungeness. Dungeness samples were predominantly assigned to Rye and Kessingland, with only six correctly assigned to Dungeness and 4 to Pakefield. This test had very little ability to differentiate between the Suffolk populations, and individuals from this region were assigned to various populations. The single Nigg bay genotype was assigned to Pakefield (Table 4.9).

	N	Chesil	Rye	Dungeness	Deal	Felixstowe	Shingle	Southwold	Kessingland	Pakefield	Carnoustie
Chesil	30	30									
Rye	29		24	5							
Dungeness	29		10	6					9	4	
Deal	30				18	2	2	2	3	3	
Felixstowe	28			2	1	8	5	6	1	5	
Shingle	29			2	2	9	8	3	3	2	
Southwold	24				1	2	3	9	4	5	
Kessingland	30			2	2	3	4	3	6	10	
Pakefield	29					3	2		3	21	
Carnoustie	30										30
Nigg Bay	1									1	

Table 4.9. Population assignment test.

Rows represent the true populations, and the columns show which population samples are assigned to. *N* represents the sample size from each population.

4.4 Discussion

4.4.1 Patterns of population genetic variation in adults

A simple prediction of the levels of genetic diversity within populations is that the size and isolation of a population should correlate with amounts of variation. Small isolated populations should be genetically depauperate. Large, range centre populations should be genetically variable. If these predictions can be considered generalities, then approximation of likely levels of genetic variation in a population might be achievable from demographic observations alone. This study, however, serves to highlight the complexity of the situation and the difficulties of making such generalizations in the absence of population genetic data.

In *Lathyrus japonicus* in Britain, there is no positive correlation between the sizes of populations or their proximity to other populations on the one hand, and the levels of genetic variation on the other. The only slight evidence for genetic depauperacy and isolation came from the isolated Scottish populations. The population of *Lathyrus japonicus* at Nigg Bay consisted of just a single multi-locus genotype, highlighting the precarious nature of the species at this site. However, this extremely small census count precludes meaningful comparisons of comparative levels of diversity with other populations. The population at Carnoustie, however, despite having a gene diversity estimate higher than many range centre populations, did show a reduced number of alleles per locus. It is known that *A* is a more sensitive indicator of genetic bottlenecks than *He*, as rare alleles are the initial casualties of any bottleneck. *He* can be relatively insensitive to population bottlenecks unless they are very severe and prolonged (Nei *et al.*, 1975).

The south coast populations showed no evidence whatsoever of genetic depauperacy associated with isolation or population size. Populations of *Lathyrus japonicus* on the south coast of England showed, on average, higher levels of genetic variation compared to populations in the centre of the species range in the UK (Table 4.2, Figure 4.4). The highest gene diversity in the sample set came from a small population at Dungeness (Figure 4.5)

that had less than 30 individuals present when the sampling was undertaken. Levels of gene diversity (*He*) for all of the south coast populations were higher than the average for all populations in the data set, and the mean number of alleles per locus (*A*) at Dungeness and Rye was higher than the average values, and the population at Chesil was only slightly lower than the average value (Figure 4.4).

4.4.1.1 Why do geographic isolation and small population sizes not lead to genetic depauperacy?

One potential reason why geographical isolation and population size might not correlate with levels of genetic diversity in the current study would be if gene flow is sufficient to overcome any apparent geographical isolation. However, this seems unlikely. There is clear genetic/geographical structure in the data set. Although in the UK range centre, there is lower differentiation among populations (Tables 4.8 & 4.9), there is clear evidence of population differentiation when greater distances are involved. This is clear in Figure 4.10 in which genetic isolation by distance (IBD) is evident. At the highest levels of geographical distances this relationship drops off to some extent due to some greater similarities between the south coast populations and the Scottish Carnoustie population than would be predicted by geographic distance alone. However, these populations are still strongly differentiated from one another with pairwise F_{ST} >3.5 (Figure 4.10) and the highest frequencies of private alleles of all populations (Table 4.7). (The population at Nigg bay is too small to make formal assessments of population genetic structure, but its lack of private alleles, and close similarity to the range-centre populations (Table 4.9) suggests it may be the result of a chance long distance dispersal event, or potentially even planted). But the general picture is that there is some restriction to gene flow at a UK wide level in the data set, and that these populations are certainly not at panmixia.

Given the genetic differentiation between the range-centre populations and both the south coast populations and Carnoustie, one possibility is that the genetic variability in these outlying populations is attributable to an influx of genes from other non-British populations. The prevailing currents through the English Channel run west to east, and the

prevailing North Sea surface currents circulate in a counterclockwise direction (Hill, 1971). This means that the outlying populations experience currents that flow to, rather than from, the range centre populations. Of course floating seeds are also likely to be affected by wind direction on the surface of the sea and this is more changeable and less predictable.

The south coast populations do not have another obvious proximal source to provide migrants. *Lathyrus japonicus* is not found elsewhere along the English Channel coastlines, and the only known French locality for this species is no longer extant. There is thus no obvious proximal source of migrants. What cannot be ruled out, however, is a general influx of seeds from a range of localities. The English Channel reflects a narrowing of water, and is connected to both the Atlantic Ocean and the North Sea. As such it represents a contact zone between two water bodies and populations here may receives some influx of variation from both. *Lathyrus japonicus* is globally widespread, and sea dispersed. Connecting channels between water bodies might be expected to be 'melting pots' of genetic diversity due to the potential for mixing different and distant gene pools.

The population at Carnoustie may potentially receive migrants from Scandinavian populations of *Lathyrus japonicus*. The species is relatively common in Scandinavia and occurs in Finland, Norway, Denmark, Sweden, and Germany, as well as Poland and the former USSR. However, the lower levels of allelic diversity in the Carnoustie population means that the genetic structure of this isolated population is 'less of a problem' to explain than the South Coast populations.

In seeking an explanation for the difference between the predicted levels of genetic diversity in this study, and the observed data, one has to accept that there are many potentially unknown variables that could be important. Shingle is occasionally moved around the coast-line to support coastal defenses, and the transport of large numbers of seeds is possible this way. There has, for instance, been some working of the shingle banks along the channel coast near Dungeness (O. Leyshon, personal communication, 2001) and this may have influenced patterns of genetic diversity at this site. Likewise, historical records suggest that *Lathyrus japonicus* has been used as a food source in times of famine in the centre of its range in the UK (Akeroyd, 1994). This may have led to some level of

population bottlenecks, although one would expect even extensive cropping of seeds for a few years to have minimal effects on levels of genetic diversity in a long lived, clonally spreading species.

Anthropogenic events aside, it is worth stressing the point that population genetic theory derived for sexual species is not directly transferable to species capable of extensive clonal growth. The genetic fingerprinting approach used in the current study detected large spreading clones, covering several meters. If plants persist for long periods of time vegetatively, this may act as an anchor against genetic drift and serve to maintain genetic variation by reduced generation cycling.

Regardless of the true biological reason underlying the observed data, this study serves to highlight clearly the difficulties of estimating levels of genetic variation from demographic data alone. Frankham *et al.* (2002) stated that there is "overwhelming evidence for associations between population size and genetic diversity." However, it is worth exploring the basis for this statement. The theory is clear, for completely isolated populations in which the effective population sizes match the census counts. But populations are often not completely isolated, and the census size rarely matches the evolutionary effective population size. Thus from these reasons alone, one should expect many exceptions to the statement of Frankham *et al.* (2002). Secondly, this is a topic likely to exhibit a large reporting bias. Where data sets have a clear correlation between population size and levels of genetic variation, this is likely to be reported, even if assessing this was not the objective of the study. In contrast, a lack of association between population size and levels of genetic variation is far more likely to go unreported.

4.4.2 *Reproductive biology of* Lathyrus japonicus

Lathyrus japonicus in Britain has been described as self-incompatible (Akeroyd, 1994). The data presented in this study, however, strongly contradict this assertion. Firstly, the Nigg bay 'plant' set seed, despite this population consisting of a single multi-locus

genotype. Secondly, F_{IS} estimates (inbreeding coefficient) were significantly different from zero in 9 out of 10 populations examined. Thirdly, the out-crossing rates measured directly from progeny arrays suggested high levels of selfing. While the sample sizes per family were too low for a precise estimate of the out-crossing rate (t) to be obtained, the consistency of the values across populations (Tables 4.4 & 4.6) suggest that self-pollination is important for this species, and the species should be described as having a mixed mating system, rather than being an obligate out-crosser. The extent to which this self-pollination is due to within-flower selfing, or geitonogamy is unknown. During field-work, bees were often observed visiting multiple flowers per plant, and between-flower within-individual matings may be an important source of this self-pollination.

4.4.3 Differences in inbreeding coefficients between adult and seedling populations

The population level inbreeding coefficients (F_{IS}) calculated from allele frequencies of wild collected adult plants are significantly smaller than the F_{IS} calculated from allele frequencies of seedlings in cultivation based on the same 7 loci (Table 4.4). Likewise, there was a marked difference between the out-crossing rates derived from adult F_{IS} and those estimated directly from progeny arrays (Table 4.6). The simplest explanation of this result is that there is preferential survival of out-crossed individuals, and that more out-crossed individuals survive to adulthood than inbred individuals. Thus the difference between the seedling F_{IS} and the adult F_{IS} can be attributed to selection. The populations are carrying some genetic load, this genetic load is unmasked to selection under inbreeding, and there is differential survival of individuals in relation to heterozygosity.

Keller & Waller (2002) noted that such selection can be either 'hard' or 'soft'. Hard selection is where the differential survival of individuals affects population dynamics by effectively limiting recruitment. In contrast, soft selection would involve selective deaths at a level that could be carried by a population with no consequences on the number of

individuals recruited to future generations. It is not clear as to whether these populations are experiencing hard or soft selection. The scale of differences between adult and seedling F_{IS} varies among populations, and one might predict the issue to be most severe when the discrepancy is at its largest (Figure 4.6). However, there are some issues regarding the interpretation of these data that should be considered before extensive biological conclusions are drawn.

4.4.3.1 Methodological issues

Approximately 40 seedlings from each population were used to estimate seedling F_{IS}. These seedlings originated from pooled seed collections from 5-10 pods from 1 or 2 pod clusters from each individual, with 10 separate individuals (families) sampled per population. Thus the seedling F_{IS} estimates are derived from 10 groups of 4 sibs in each population, whereas the adult F_{IS} was derived from ca 30 independent individuals per population. If there is any sub-structure within any sites, the correlated maternity in the seedling samples may elevate F_{IS} in individual populations.

The fact that the direct estimate of out-crossing for the seedlings, the F_{IS} derived out-crossing rates for seedlings, and Mltr derived out-crossing rates, were all significantly lower than the F_{IS} derived out-crossing rates for the adult plants, gives confidence in the difference between heterozygosity levels in seedlings versus adults (Figure 4.5 & 4.6). The consistent replication of this pattern across populations is also reassuring, as is the similarity in values between *tm* from Mltr and directly estimated out-crossing rates (Table 4.6). However, while there is confidence in there being a difference, quantifying the scale of the differences among populations is more difficult. The difference in inferred out-crossing rates between the two estimates from the seedlings suggests some caution is required before any further interpretation (Figure 4.7). To increase confidence in the direct assessments of out-crossing rates, sampling up to 30 individuals per mother would result in narrower confidence intervals around the mean (Ritland, 2002). All that can be said for now, is that there is good evidence for selection against inbred individuals surviving in wild populations, and that heterozygosity increases between seedling and adult life stages.

Chapter 5. The relationship between genetic variation and fitness in *Lathyrus japonicus*

5.1 Introduction

The theory of inbreeding depression, a reduction in fitness of progeny from selfed or closely related matings compared to out-crossed progeny, is well documented (Charlesworth & Charlesworth, 1987). Inbreeding reduces levels of heterozygosity and genetic variation, and can reduce survival and reproductive output. As the population structure of many species is altered due to changes in land use or climate, resulting in reduced population sizes and/or fragmentation, inbreeding depression is likely to become an increasing problem. As a result, conservation practitioners are becoming increasingly concerned with strategies to promote optimal levels of gene flow and the maintenance of genetic variation. But in natural systems, when knowledge of pedigree information and even breeding system are not always available, the effects of inbreeding can be difficult to determine. In these situations inbreeding is usually assessed using molecular markers to give a level of genetic variation, or heterozygosity.

However, there is little consensus between the results of studies on the correlation between neutral marker heterozygosity and fitness. Many studies have established a positive relationship between heterozygosity and fitness in plants (Hammerli & Reusch, 2003; Oostermeijer *et al.*, 1994; Paschke *et al.*, 2002; Stilwell *et al.*, 2003), while others have found no relationship (Jacquemyn *et al.*, 2003; Ouborg & Van Treuren, 1995). Several reviews of the subject exist, combining plants and animals (Hansson & Westerberg, 2002; Keller & Waller, 2002), a meta-analysis has been carried out (Reed & Frankham, 2003) and the overall consensus seems to be that a positive relationship exists, while warning that it is necessary to remember that there could be a bias due to the under-reporting of negative results.

When a positive correlation is present between heterozygosity and fitness, there are multiple theories for the mechanism underlying the correlation. One theory, called overdominance (Charlesworth & Charlesworth, 1987), or the direct effect hypothesis (Hansson & Westerberg, 2002), considers the fitness advantage to be due directly to the heterozygosity of the loci being screened. This hypothesis is not directly related to level of inbreeding. This hypothesis is a possibility with isozyme data, where possessing different alleles could give an advantage by each coding for slightly different versions of an enzyme, but this theory does not explain correlations present between heterozygosity of microsatellite markers, generally considered neutral, and fitness associated traits. Another theory, partial dominance (Charlesworth & Charlesworth, 1987), ascribes a decline in fitness in homozygotes to the fixation of deleterious alleles throughout the genome due to inbreeding. This hypothesis can be further broken down into two sub-hypothesis (Hansson &Westerberg, 2002). The local effects hypothesis considers the correlation between neutral marker heterozygosity and fitness to be a result of the effect of heterozygosity at closely linked fitness loci and requires phenomena such as a recent bottleneck to cause linkage disequilibria. The general effects hypothesis considers the correlation of increased fitness with neutral marker heterozygosity to be due to the effects of heterozygosity at a genome wide level.

There are many complicating factors that may obscure a signal between heterozygosity and fitness. The breeding system can influence this relationship, as predominantly selfing species may have been purged of recessive lethal alleles early in their history, and be much less likely to show a reduction in fitness than predominantly out-crossing species. The history of a population is also expected to effect whether a fitness advantage will be present in more heterozygous individuals. Populations with a long history of small population size and high levels of inbreeding will potentially have a much lower inbreeding load due to purging, than a historically large population that has only recently suffered a reduction in size. The strength of the relationship between heterozygosity and fitness can vary throughout the life span of an organism as well, as shown in a study of pitch pine where various directions and intensities of the relationship between heterozygosity and growth rate were explained by the age of the

stand, with the relationship only taking effect in older populations that had experienced competition (Ledig *et al.*, 1983).

Decreases in fitness may also be caused by factors other than genetics. In one of few studies comparing environmental quality, heterozygosity levels, and fitness, most effect on fitness was found to correlate with environmental quality and not with heterozygosity in *Primula vulgaris* (Jacquemyn *et al.*, 2003). The relationship between heterozygosity and fitness can vary over time in a natural system as well. A study on the mussel *Myrtilis edulus* demonstrated the significance and direction of the relationship varied between years (Gaffney, 1990).

Microsatellite markers have established a significantly greater proportion of heterozygotes in adult populations of *Lathyrus japonicus* compared with seedlings in cultivation, which indicates a positive selective advantage for heterozygotes throughout their lifetimes (Chapter 4). In this study fitness associated traits and levels of variation are measured in the field and in a common garden experiment to see if there are correlations between fitness associated traits and heterozygosity. Population differentiation in these traits is also examined in relation to patterns of population differentiation detected by the microsatellites.

5.2 Materials and Methods

5.2.1 Study sites

Each of the eleven sites shown in Table 1 were visited in 2001, the length of beach occupied by the populations were estimated, and the approximate number of clumps of *Lathyrus japonicus* were counted. At each site, thirty leaf samples were collected and the samples were screened for genetic variation using microsatellite markers (Chapter 4). These data were used for comparisons with fitness-associated traits. The measures of genetic diversity that were used were allelic richness (this is A standardized by sample size across populations), and H_O (Chapter 4).

Site	County	Grid Reference	Population size	Population Area	Substrait	Population type
Chesil Beach	Dorset	SY/564.841	>500	~2km	shingle	Large isolated
Rye Harbour	Sussex	TR/940.175	~150	~2km	shingle	Medium isolated
Dungeness	Kent	TR/090.166	~30	~0.5km	shingle	Small isolated
Deal	Kent	TR/380.500	~230	~1.5km	shingle	Large range centre
Felixstowe Ferry	Suffolk	TM/325.370	~85	~1km*	shingle	Medium range centre
Shingle Street	Suffolk	TM/365.423	>500?	~2km*	shingle	Large range centre
Southwold Denes	Suffolk	TM/507.752	~100	~1km	sand	Medium range centre
Kessingland	Suffolk	TM/537.857	~700	~2km	shingle	Large range centre
Pakefield Beach	Suffolk	TM/539.905	~260	~0.5km	sand/shingle	Large range centre
Carnoustie	Angus	NO/560.338	~100	~0.5km	sand	Medium isolated
Nigg Bay	Aberdeenshire	NJ/965.047	1 clump	~0.25km	sand/gravel	Small isolated

Table 5.1. Study sites for *Lathyrus japonicus* (from Chapter 4)

* = populations that are part of a more or less continuous distribution in which neighbouring populations are separated by only a small (and potentially subjective) discontinuity. The cut-off for whether a population is large or small, isolated or not is also essentially somewhat arbitrary, but here small = <50 plants, medium = 51-200, large = >200 plants. Isolated populations are those populations away from the range centre of the south east of England.

5.2.2 Field data 2001

For each sample included in the genetic analysis, measurements were taken of the widest and narrowest diameter of the clump from which it came, and the number of seedpods in fifteen pod clusters. Where pods were mature, the numbers of seeds in 20 pods were counted. The clumps from which the measurements were taken should approximate to ramets, but since neither genetic identity nor physical connectivity was individually determined, they may be multi-clonal.

For 10 of these samples chosen randomly from each population, one or two pod clusters from the same stem as the leaf sample (e.g. the pods were demonstrably from the same ramet as the sample in the genetic analysis) were collected for cultivation experiments.

The pods were taken back to the laboratory, and the seeds from each pod were counted and weighed to the nearest milligram. Before weighing, each seed was examined for predation by either beetles or wasps, this was recorded, and the predated seed was then excluded from the analysis to prevent eaten seed causing a downward bias in seed weights. A total of 5648 seeds were weighed from the eleven populations (range 80 - 800 unpredated seeds per population).

5.2.3 Cultivation experiment

For each population, 40 seeds were germinated (four taken at random from each of 10 individuals) at the RBGE nursery in Edinburgh. The seeds were nicked with a blade and soaked in water for two days. Then on the 25th January 2002, they were placed on the surface of a mix of half sand and half John Innes No.2 compost (a balanced fertilizer is incorporated in this compost) in individual 4-inch pots. Seeds were placed with the hilum facing upwards, so that germination would be visible without disturbing the seed. The pots were then placed in a randomised block, so that seeds from each population were distributed throughout the occupied space. The pots were located in an unheated greenhouse and bottom heat was provided. The seeds were kept uncovered and constantly moist until germination. As germination occurred (i.e. when the radicle was visible outside of the seed coat), this was recorded and the seed was covered in approximately 2 cm of the sand and compost mix. The date that the seedlings emerged from the sand and compost mix was also recorded. From this data, the rate and percentage of germination was calculated. The level of pigmentation in the first week after germination was also estimated using the Royal Horticultural Society (RHS) colour chart (Royal Horticultural Society, 1966). After a week of growth, as the plants developed leaves, the pigmentation differentiation became less obvious. Leaves were collected from the seedlings for genetic analysis in June 2002, as the seedlings were transplanted into larger pots and moved outdoors. Plants were grown outdoors, receiving fertilizer every two weeks until the autumn, when they were moved back into

an unheated greenhouse to prevent water logging during their winter dormancy. Further data was recorded on survival through winter, the number of stems present in June 2003, the dry weight of individuals in October 2003, and the length and width of five leaves for approximately 10 individuals per population.

5.2.4 Field data 2002

In 2002, the field sites were visited again and the length and width of five leaves from each of thirty individuals were measured, following an observation of variation in leaf shape between populations in the seedlings in cultivation. Pods were also collected from ten individuals at each site in 2002, to assess year-to-year variation in seed weight. Approximately 100 seeds per population were weighed in 2002. Data were not collected from the same thirty plants used in the genetic analysis since the plants were not marked in 2001.

5.2.5 Analysis

The data were recorded in Excel spreadsheets then analyzed using the computer program SPSS (SPSS Inc, 2001).

Each of the variables (allelic richness, observed heterozygosity, seed weight, clump area, leaf shape, pods per cluster and seeds per pod, dry weight and number of stems in cultivation) were checked for significant differences between populations, using the Kruskal Wallis test, and presented visually in the form of charts showing population means and 95% Confidence intervals.

The data was then screened for normality, and transformed where necessary. For the data set consisting of the mean of various measurements for each of the 10 populations, the majority of the distributions do not resemble the normal distribution, so the natural log, square root, and arcsine transformations were applied where appropriate, but none made substantial improvements, and the untransformed data was retained for the sake of simplicity. Because of deviations from normality, due most likely to the small sample

size of ten populations, the statistic Spearman's rho was used to obtain non-parametric correlations.

For individual level data, sample sizes were larger and the data more closely followed the normal distribution. Since heterozygosity data was in the form of a proportion it was arcsine transformed, and this created a distribution which approached normal. Adult clump size was natural log transformed to bring it to normality. Seedling dry weight was also natural log transformed. The number of seeds per pod, and the number of pods per cluster, and seed weight were not transformed, because they approximated a normal distribution without any transformations. Because the individual level data seems to be satisfactorily close to a normal distribution, parametric methods were used.

The population from Nigg Bay was not used in comparisons involving the genetic data to avoid comparing the results from one genetic individual with population samples of >20 plants from the other populations.

5.3 Results

5.3.1 Microsatellite data

The results of the microsatellite analyses are presented in Chapter 4 and not repeated here. However, one additional figure is presented here. The Kruskal Wallis Test (non-parametric version of independent samples t-test) showed that there are significant differences in observed heterozygosity among both the ten adult populations (Chi-Square 22.918, df = 9, p<0 .01), and the seedlings in cultivation (Chi-Square 57.572, df = 9, $p < 0.01$) (Figure 5.1).

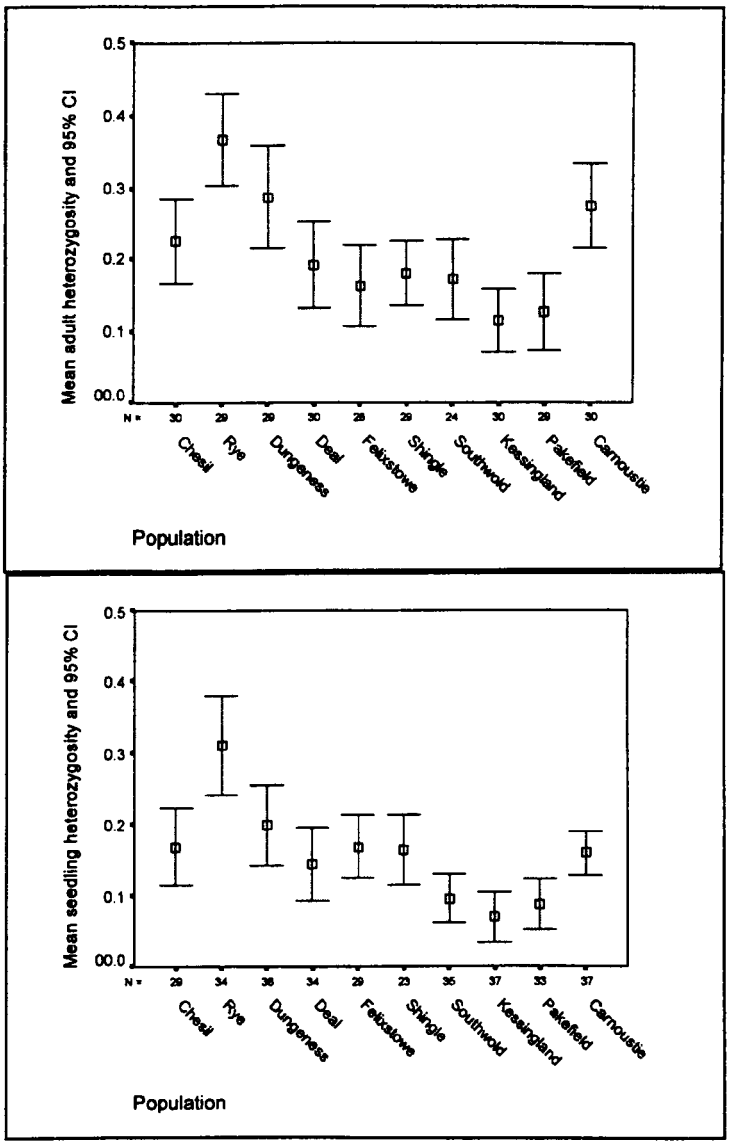

Figure 5.1. Mean adult and mean seedling heterozygosity of each population.

The data is based on the same seven microsatellite loci from adults and seedlings. CI = confidence interval.

5.3.2 *Differences between populations*

5.3.2.1 Germination and seedling survival

Germination of seeds from individual populations ranged from 68-100%, with a mean 86% (Table 5.2).

There was a correlation between the population level germination success, and the extent of predation damage ($r^2 = 0.57$, Spearman's rho = 0.766, $p < 0.01$) (Figure 5.2). Germination success was not related to seed weight (Spearman's rho = -0.105, p = 0.773).

The cumulative rate of germination of seeds in different populations reflected the overall success of germination, with the populations with the slowest initial rate of germination, being those that had the highest overall failure rate of germination (Shingle, Felixstowe) (Figure 5.3).

Site	Number of seeds germinated	Number of seeds planted	Percent germination
Chesil Beach	31	40	77.5
Rye Bay	34	40	85
Dungeness	38	40	95
Deal	32	40	80
Felixstowe	31	40	77.5
Shingle Street	27	40	67.5
Southwold Denes	38	40	95
Kessingland	38	40	95
Pakefield Beach	34	40	85
Carnoustie	40	40	100
Nigg Bay	34	40	85

Table 5.2. Percentage germination of seeds from each population.

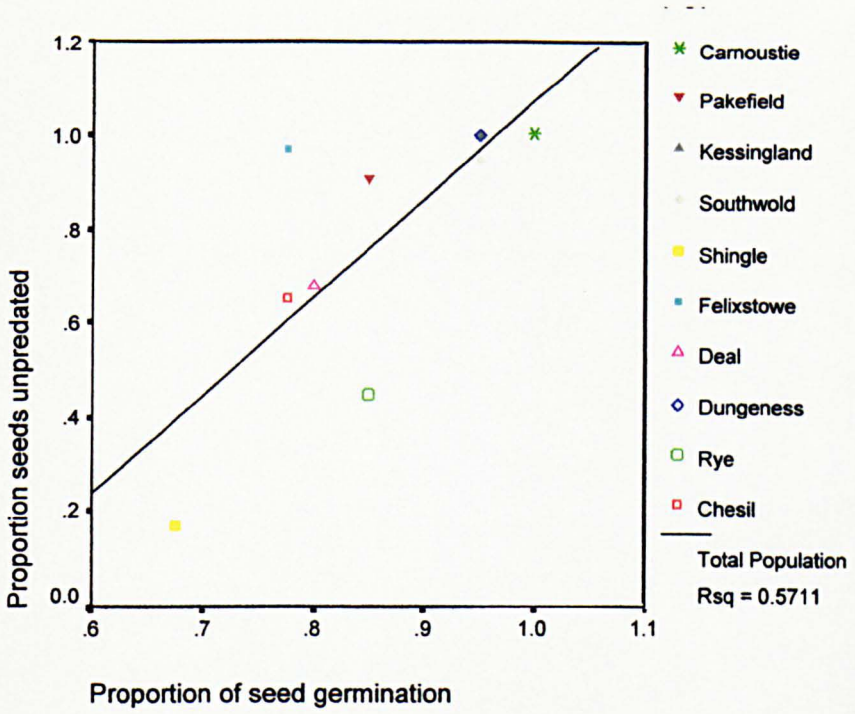

Figure 5.2. The relationship between levels of seed germination and the percentage of unpredated seeds.

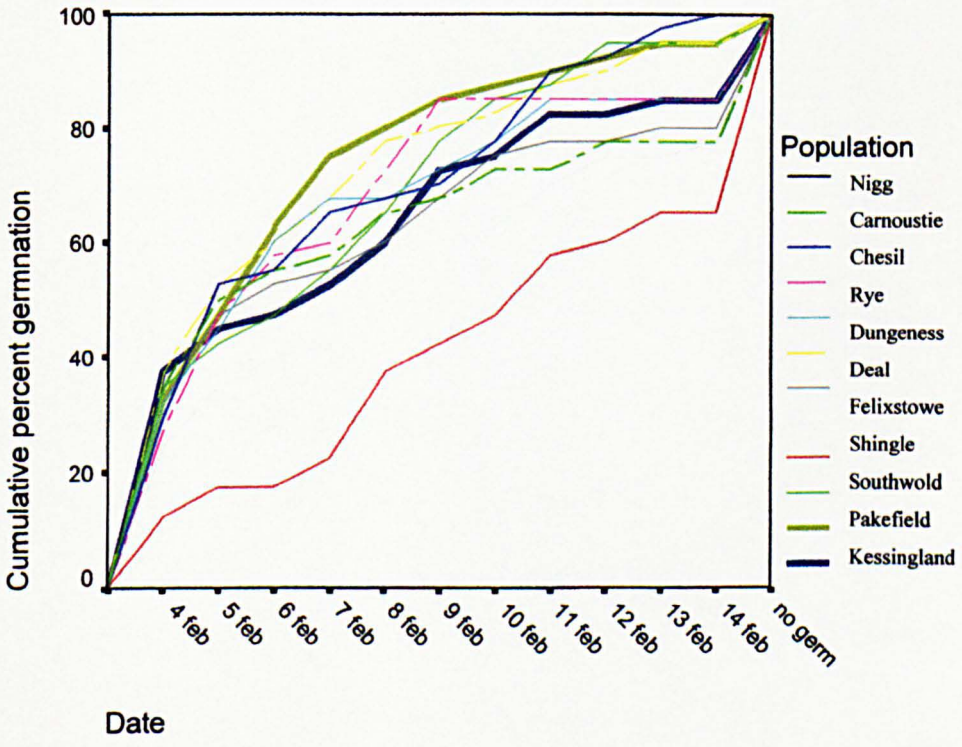

Figure 5.3. The cumulative percentage of seed germination in relation to time for cultivated populations of *Lathyrus japonicus*.

No germ = the remaining seedlings that did not germinate during the course of the study.

There were significant differences between populations in the numbers of seedlings surviving through the first winter (Kruskal-Wallace test, Chi-Square = 23.709 $P < 0.01$) (Figure 5.4). Four populations had seedling survival percentages of < 65%; these were all large populations (Figure 5.4, Table 5.1).

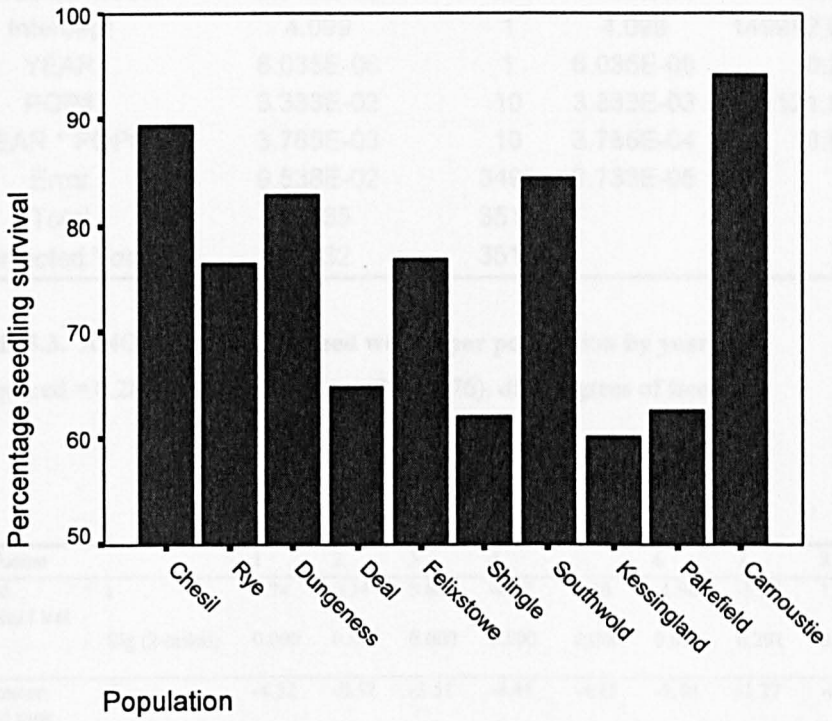

Figure 5.4. Percentage of seedlings surviving through the first winter for each population.

5.3.2.2 Seed weight

Significant differences were detected by ANOVA both in the weights of the seeds from different populations, and in different years (Table 5.3). Significant differences were present in many, but not all, of the populations between years. The individual significances of means between years for each population are shown in a paired sample t-test (Table 5.4). Assumptions for paired sample t-test are normal distribution with variances equal or unequal. Seed weights approximately follow the normal distribution, but a Wilcoxon signed rank test is presented as a non-parametric equivalent (Table 5.4).

The results of the two tests differ only in two borderline cases; Shingle (pop 6) where the near significance was lost and Nigg Bay (pop 11) where significance was gained in the non-parametric test.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3.712E-02	21	1.768E-03	64.677	0.000
Intercept	4.099	1	4.099	149982.071	0.000
YEAR	6.035E-06	1	6.035E-06	0.221	0.638
POP#	3.333E-02	10	3.333E-03	121.946	0.000
YEAR * POP#	3.785E-03	10	3.785E-04	13.850	0.000
Error	9.538E-02	3490	2.733E-05		
Total	5.065	3512			
Corrected Total	0.132	3511			

Table 5.3. ANOVA results for seed weight per population by year.

R Squared = 0.280 (Adjusted R Squared = 0.276). df = degrees of freedom.

Population		1	2	3	4	5	6	7	8	9	10	11
Paired samples t test	t	4.74	0.24	3.68	-4.12	4.80	-1.98	-1.28	1.50	-6.43	3.46	-1.78
	Sig (2-tailed)	0.000	0.815	0.000	0.000	0.000	0.051	0.201	0.136	0.000	0.001	0.079
Wilcoxcon signed rank test	Z	-4.52	-0.52	-3.51	-4.41	-4.25	-1.40	-1.27	-0.96	-7.16	-3.01	-2.79
	Sig (2-tailed)	0.000	0.605	0.000	0.000	0.000	0.163	0.205	0.337	0.000	0.003	0.005

Table 5.4. Paired-samples t-test and Wilcoxon signed rank test for differences in mean seed weight of populations in two years.

A significance of >0.05 means that the means for year one and two are not significantly different.

Population 1 is Chesil, 2 is Rye, 3 is Dungeness, 4 is Deal, 5 is Felixstowe, 6 is Shingle, 7 is Southwold, 8 is Kessingland, 9 is Pakefield, 10 is Carnoustie, and 11 is Nigg.

Populations that did not have significantly different mean seed weights in 2001 and 2002 are Rye, Southwold, and Kessingland, with Shingle being of borderline significance. It could be possible to explain some of the between year differences in seed weight by predation levels. In the course of weighting the seeds it was observed that in heavily predated pods, if some seeds escaped predation they were often the smallest seeds, and

this could bias the mean weights. Of the seven populations with the largest differences in seed weight between years, in five the direction of variation matches the direction of change in predation level. The changes in population mean seed weight over two years are shown in (Figure 5.5). This figure shows that in both years Chesil and Southwold have higher than average seed weights, while in both years Rye, Carnoustie, and Nigg have lower than average seed weights.

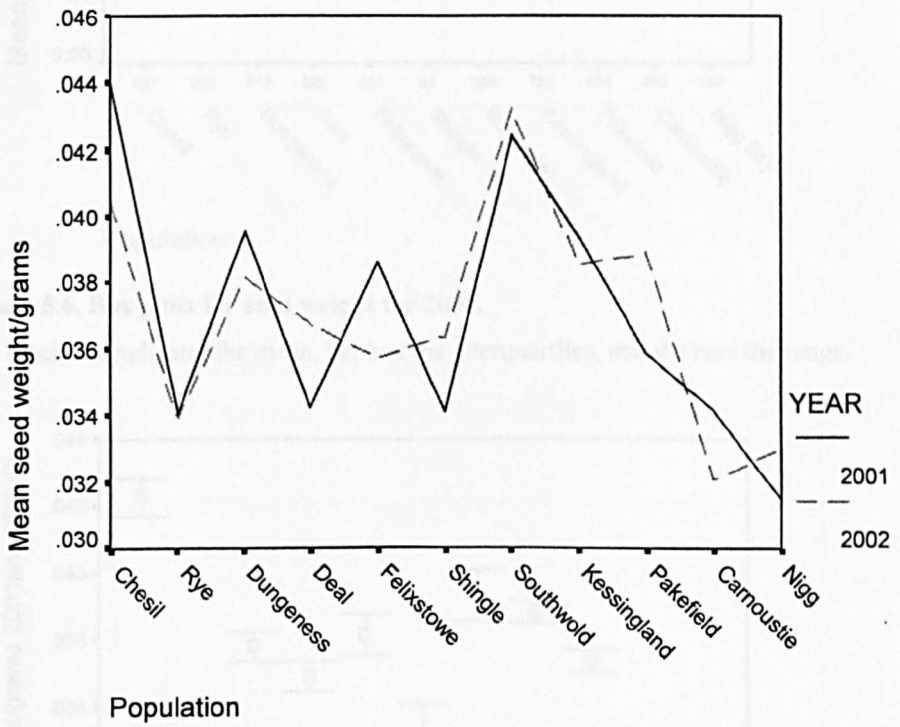

Figure 5.5. Mean seed weights for each population over two years, with sample sizes standardized between years for comparisons.

Figure 5.6 and Figure 5.8 show box plots with the mean and quartiles of seed weights for year 1 and year 2 respectively. Figure 5.7 and Figure 5.9 show means and 95% confidence intervals for the same data.

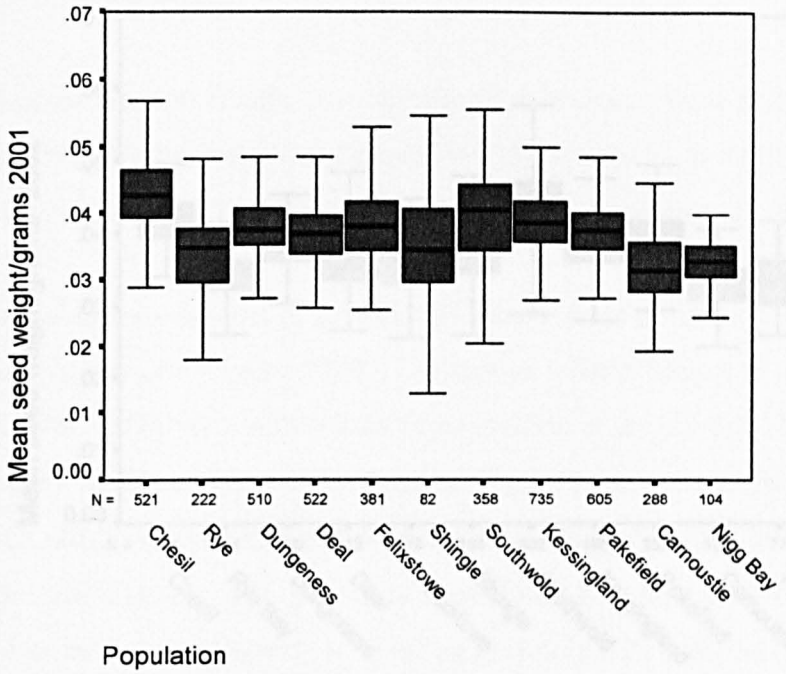

Figure 5.6. Box plots for seed weight for 2001.

The black bar indicates the mean, the box the interquartiles, and the bars the range.

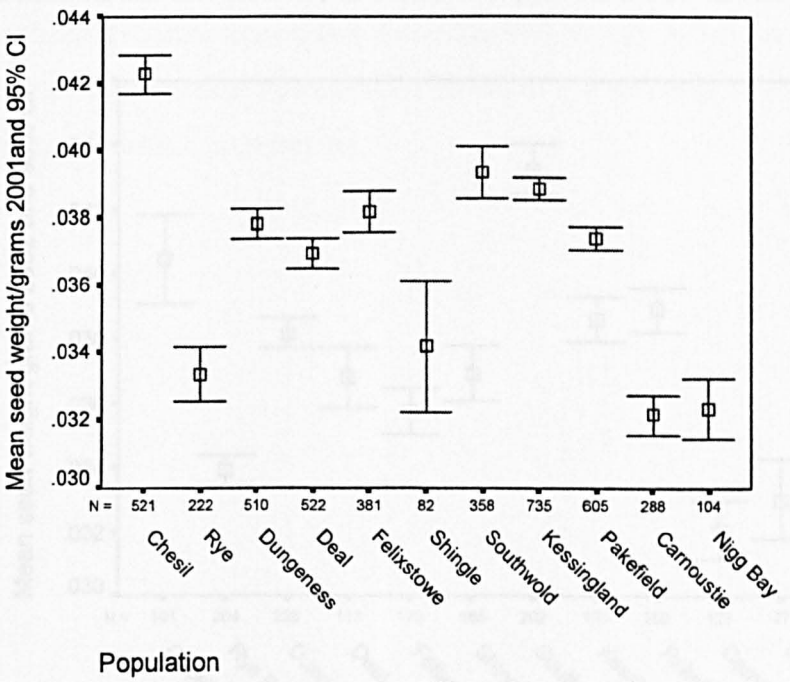

Figure 5.7. Mean and CI (confidence interval) for seed weight for 2001.

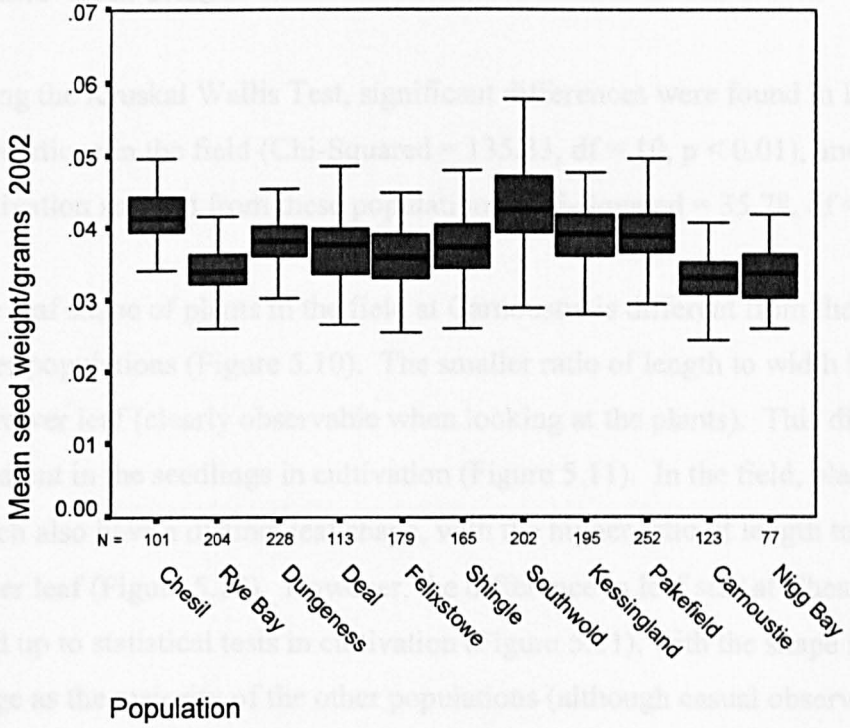

Figure 5.8. Boxplots for seed weight for 2002.

The black bar indicates the mean, the box the interquartiles, and the bars the range.

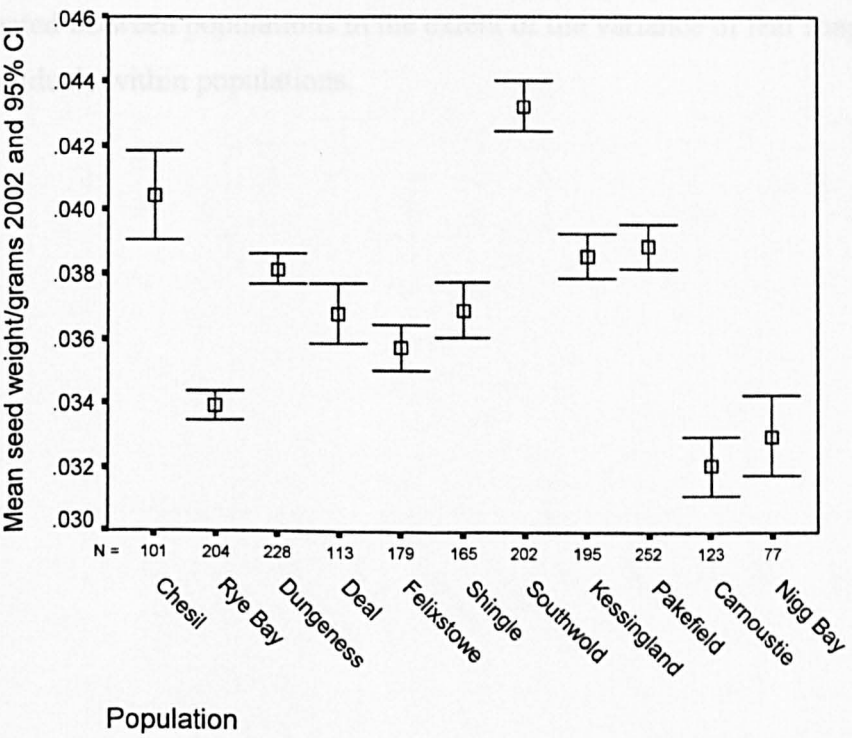

Figure 5.9. Mean and 95% confidence interval for seed weight for 2002.

5.3.2.3 Leaf shape

Using the Kruskal Wallis Test, significant differences were found in leaf shape of populations in the field (Chi-Squared = 135.33, df = 10, $p < 0.01$), and of seedlings in cultivation sourced from these populations (Chi-Squared = 35.78, df = 10, $p < 0.01$)

The leaf shape of plants in the field at Carnoustie is different from the leaf shape of any other populations (Figure 5.10). The smaller ratio of length to width indicates a much narrower leaf (clearly observable when looking at the plants). This difference is also apparent in the seedlings in cultivation (Figure 5.11). In the field, plants from Chesil beach also have a distinct leaf shape, with the higher ratio of length to width indicating a wider leaf (Figure 5.10). However, the difference in leaf size at Chesil beach did not hold up to statistical tests in cultivation (Figure 5.11), with the shape falling in the same range as the majority of the other populations (although casual observations could still distinguish the shape of these leaves). A correlation between the natural log dry weight of an individual seedling and the natural log variance in leaf shape of that seedling is significant (Pearson correlation = -0.308, $P<0.01$). However no differences were detected between populations in the extent of the variance of leaf shape among individuals within populations.

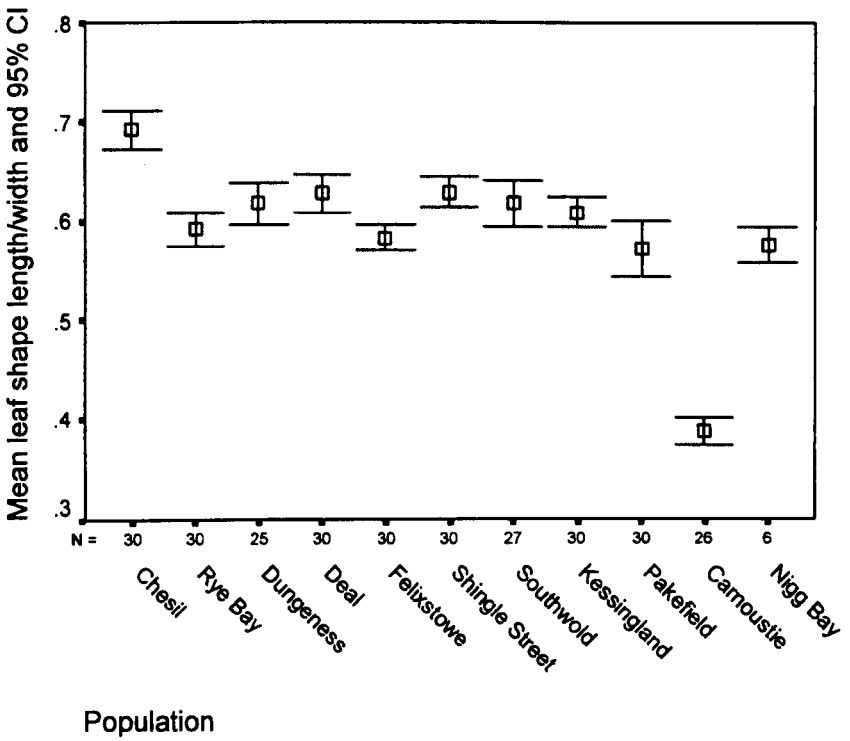

Figure 5.10. Mean and 95% confidence intervals (CI) for leaf shape per population of adults in the field.

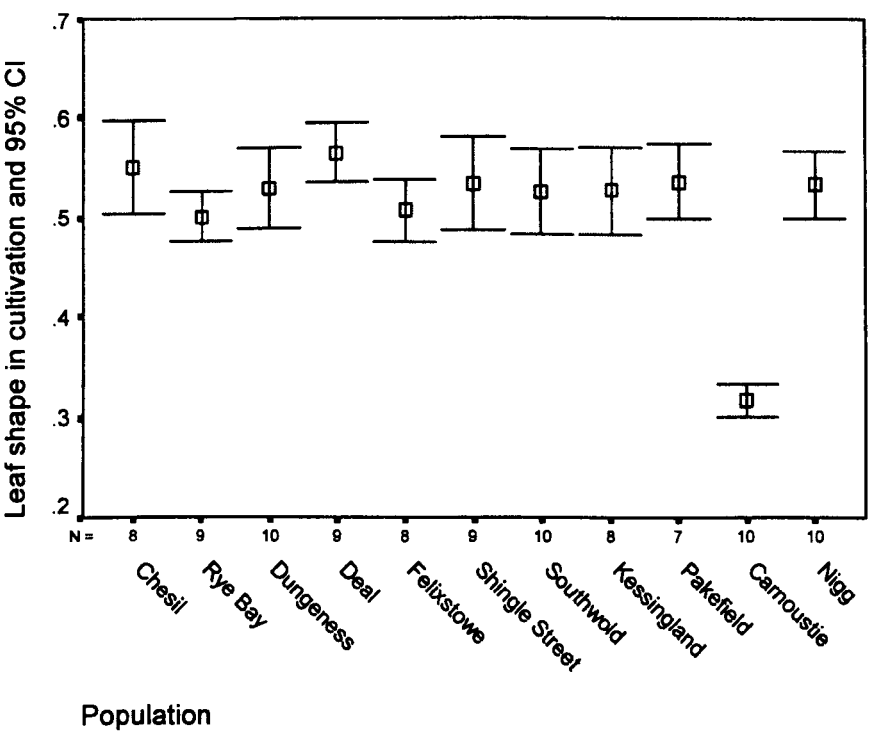

Figure 5.11. Mean and confidence intervals (CI) for leaf shape of seedlings in cultivation.

5.3.2.4 Pods per cluster and seeds per pod

The Kruskal Wallis test showed significant differences between populations in both the number of pods per pod cluster (Chi-Square = 81.47, df = 9, $p < 0.01$), and the number of seeds per pod (Chi-Squared = 55.04, df = 9, $p < 0.01$). The number of pods per cluster follows a geographic pattern, with the south coast populations (Chesil, Rye, Dungeness) and Deal having the highest numbers of pods per cluster, the Suffolk populations (Felixstowe, Shingle, Southwold, Kessingland, and Pakefield) having an intermediate number, and Carnoustie having the lowest number (Figure 5.12).

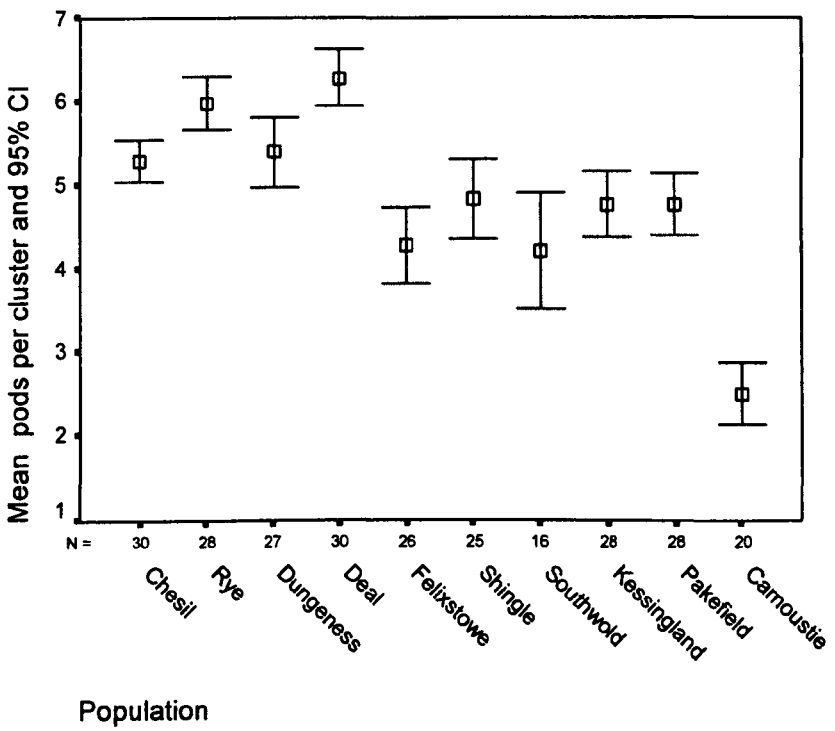

Figure 5.12 Mean and confidence intervals (CI) for number of pods per cluster for each population.

The number of seeds per pod, however, varies independently of geography (Figure 5.13). Instead there is an association with population size for this variable with the largest number of seeds being produced by the largest populations (Chesil, Deal, Shingle, Kessingland, Pakefield; Figure 5.13, Table 5.1). The two variables are independent of each other on a population mean level; with Deal having one of the

highest values in both, and Carnoustie one of the lowest in both, while other populations have inverse values for each variable, such as Rye, Dungeness and Shingle (high for one, low for another).

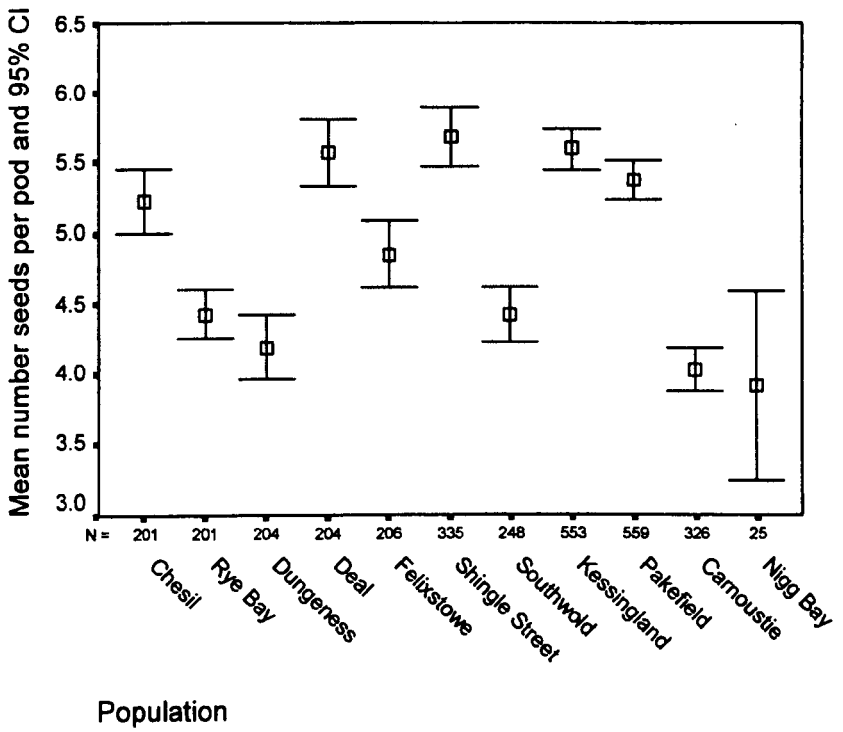

Figure 5.13. Mean and confidence intervals (CI) for number of seeds per pod.

However when the data is treated at an individual level, rather than a population level, a positive correlation exists between number of seeds per pod and number of pods per cluster (Pearson correlation = 0.233, P = <0.01 bonferroni corrected) and between number of pods per cluster and seed weight (Pearson correlation = 0.254, P = < 0.05 bonferroni corrected).

5.3.2.5 Dry weight and number of stems of seedlings in cultivation

Significant differences were found between the dry weights of seedlings sourced from different populations (Kruskal Wallis test, Chi-Squared = 92.67, df = 10, $p < 0.01$). Figure 5.14 shows mean seedling dry weight for each population after two summers growth. Seedlings from Carnoustie are larger than the English populations by twofold, and seedlings from Nigg Bay are somewhat smaller than the rest.

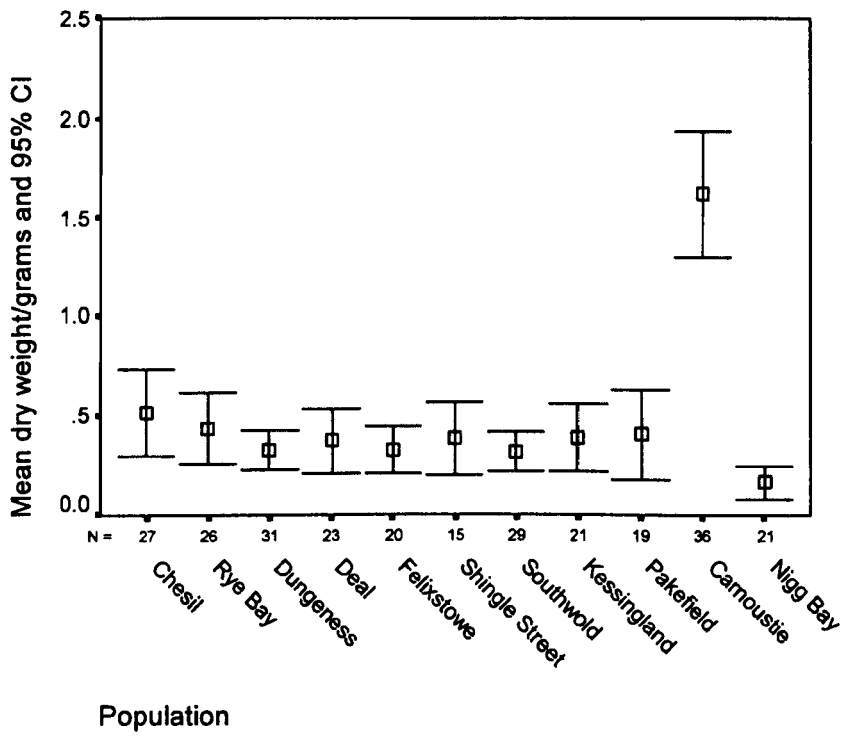

Figure 5.14. Mean and 95% confidence intervals (CI) of the dry weight of seedlings. Measurements were taken on seedlings that survived in cultivation after two summers growth.

There is a clear relationship between the 'dry weight' data and the 'number of stems' data shown in Figure 5.15, with most populations roughly the same, while Carnoustie is much higher in both variables. Because these two variables are giving the same signal and are not independent, the measure of the number of stems will be excluded from the rest of the study.

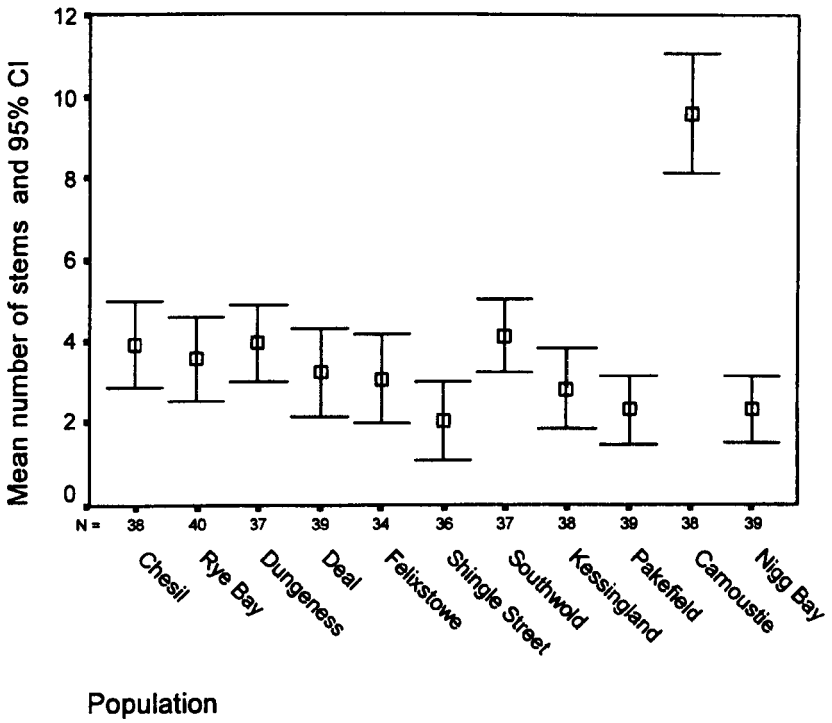

Figure 5.15. Mean and 95% confidence intervals (CI) of number of stems of seedlings. Measurements were taken on seedlings that survived in cultivation after two summers growth.

5.3.2.6 Pigmentation

Considerable differences in pigmentation between populations were visible in the seedlings in the first weeks after germination, before they developed leaves and extensive chlorophyll. The pigmentation forms present in each population are shown in Figure 5.16. Though most populations shared most of the pigmentation forms, there were differences in frequency of these forms. Over 50% of the seedlings from Chesil beach had no red pigmentation at all, over 90% were green or greenish, while less than 10% had enough pigmentation to appear either orange or red. Carnoustie at the other extreme, had no seedlings with greenish pigmentation, and nearly 60% of the seedlings could be described as dark orange or dark red. In general there was a trend of light to increasingly heavy pigmentation moving from south to north. The south coast populations held the highest proportions of green and greenish seedlings, the Suffolk

populations held only 7 to 15% green and greenish seedlings, and no green or greenish seedlings were present in Carnoustie in Scotland. An exception to the trend is Felixstowe – the most southerly of the Suffolk populations – with 50% seedlings dark orange or dark red, second only to Carnoustie.

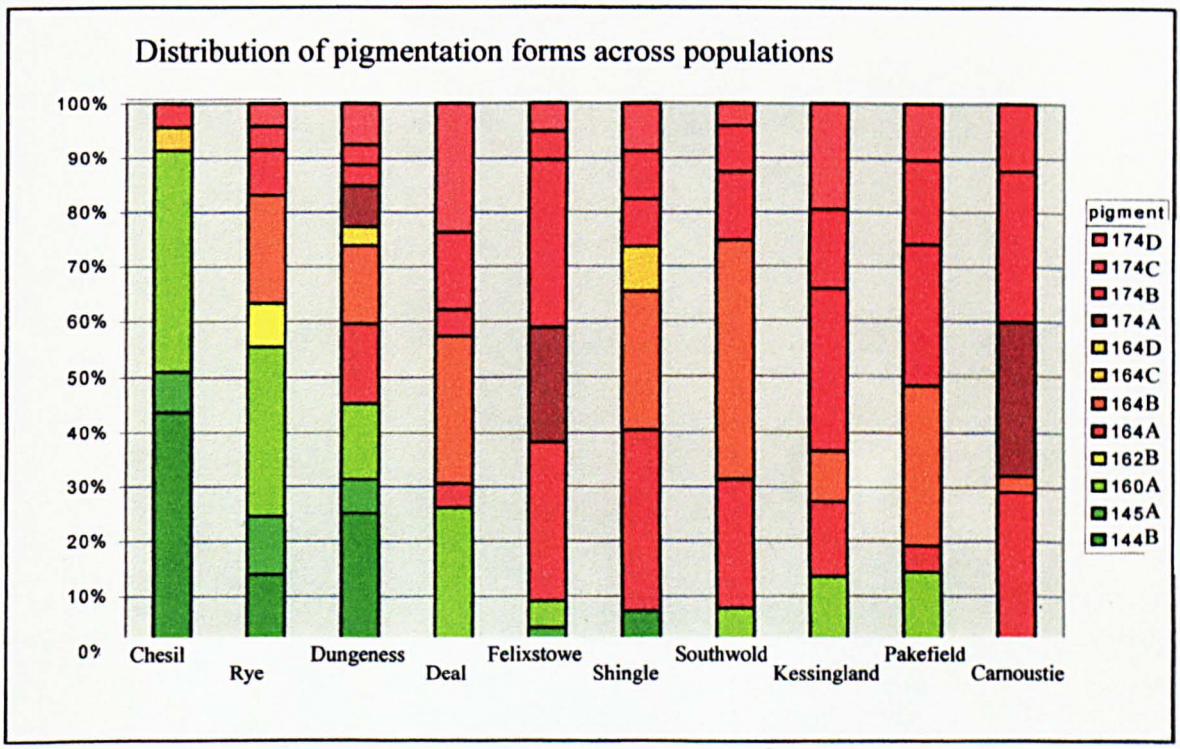

Figure 5.16. Pigmentation of seedlings in cultivation from each population.

The numerical codes next to the pigment colours represent the Royal Horticultural Society Pigment codes (Royal Horticultural Society, 1966).

5.3.3 Survival and heterozygosity

Thirty seven percent of seedlings did not survive through the first winter. Though the mean heterozygosity is slightly higher for seedlings that survived, the confidence intervals of the two means overlap substantially (Figure 5.17). There was no significant difference between the mean heterozygosites of surviving and dead seedlings (Mann-Whitney test, $Z = -0.529$, $p = 0.597$).

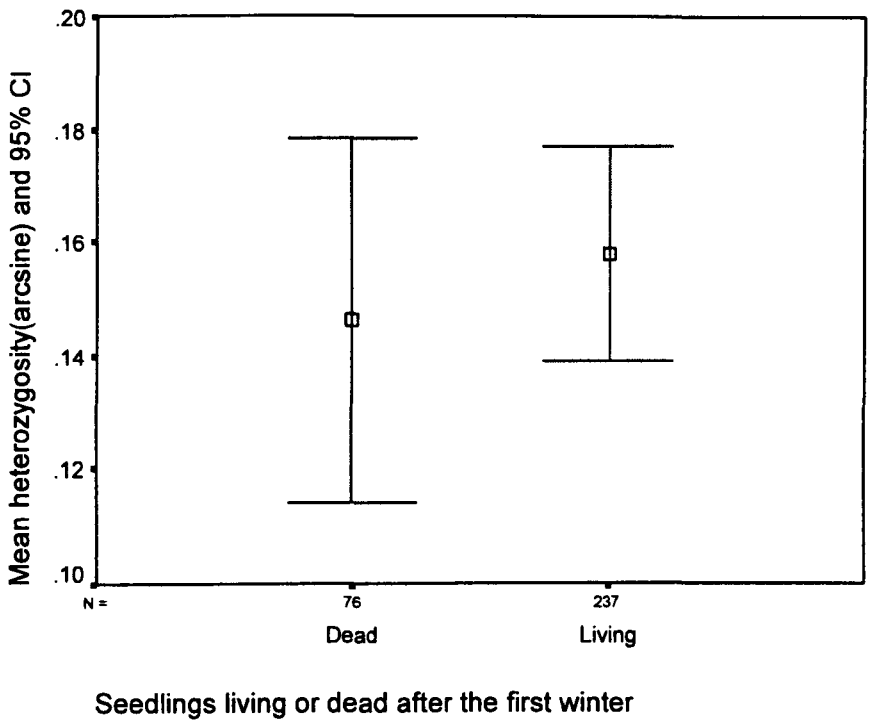

Figure 5.17. Mean heterozygosity and 95% confidence interval for seedlings that survived or died during the first winter.

5.3.4 Population means data -genetic diversity and fitness measures

No linear relationship involving population means has significance provided non-parametric methods are used.

Although population means did show some linear relationships, the directions of the strongest of these relationships were not consistent. The relationship between population average heterozygosity and percent seedling survival through their first winter appears to be positive and somewhat linear, but the relationship was not significant (Spearman's correlation coefficient = 0.515, P = 0.128).

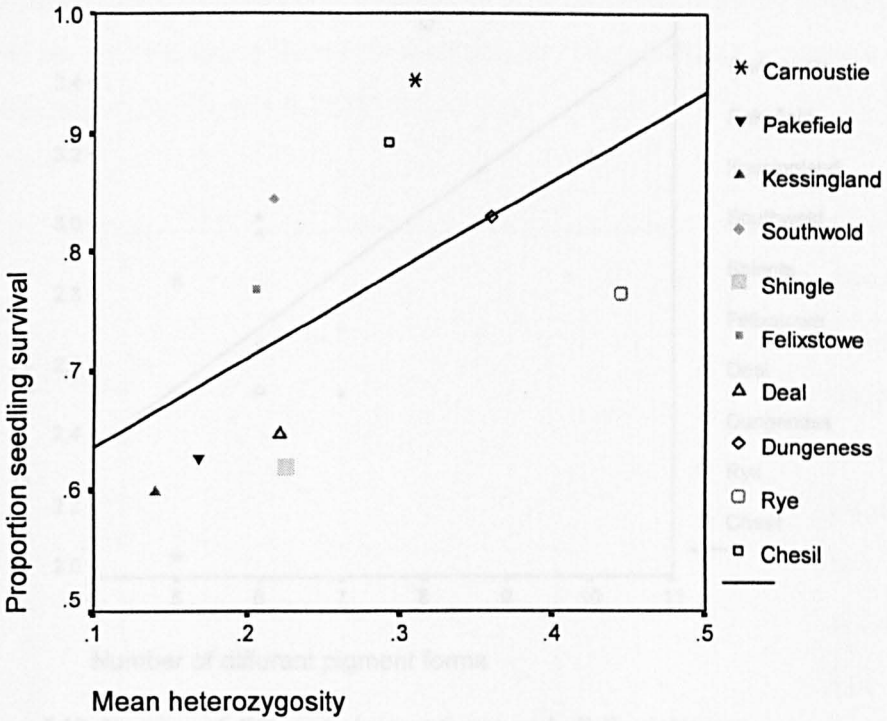

Figure 5.18. Seedling survival and mean heterozygosity.

Proportion of seedlings surviving the first winter and mean heterozygosity of the adult population from which they were sourced.

Allelic richness and number of different pigment forms appear to have a positive linear relationship, but the correlation was not significant (Spearman's correlation coefficient = 0.497, P = 0.144) (Figure 5.19). Allelic richness and number of seeds per pod also appear to have a positive linear relationship, but again, the correlation was not significant (Spearman's correlation coefficient = 0.467, P = 0.174)

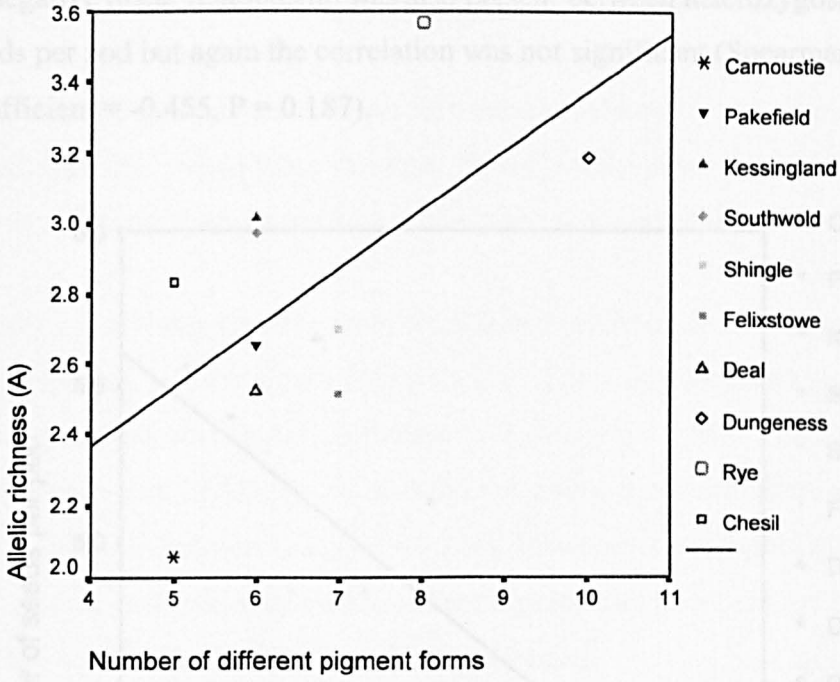

Figure 5.19. Number of different pigment forms and allelic richness.
Allelic richness is a population average (of adults in the field) and pigment forms are measured from the seedlings.

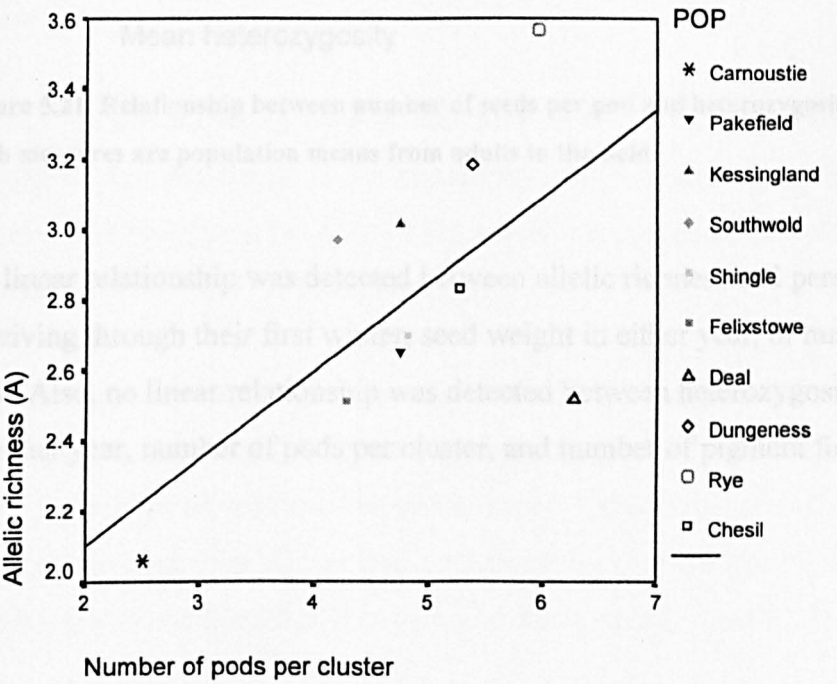

Figure 5.20. Relationship between number of pods per cluster and allelic richness.
Both measures are population means, for adults in the field.

A negative linear relationship was also present between heterozygosity and number of seeds per pod but again the correlation was not significant (Spearman's correlation coefficient = -0.455, P = 0.187).

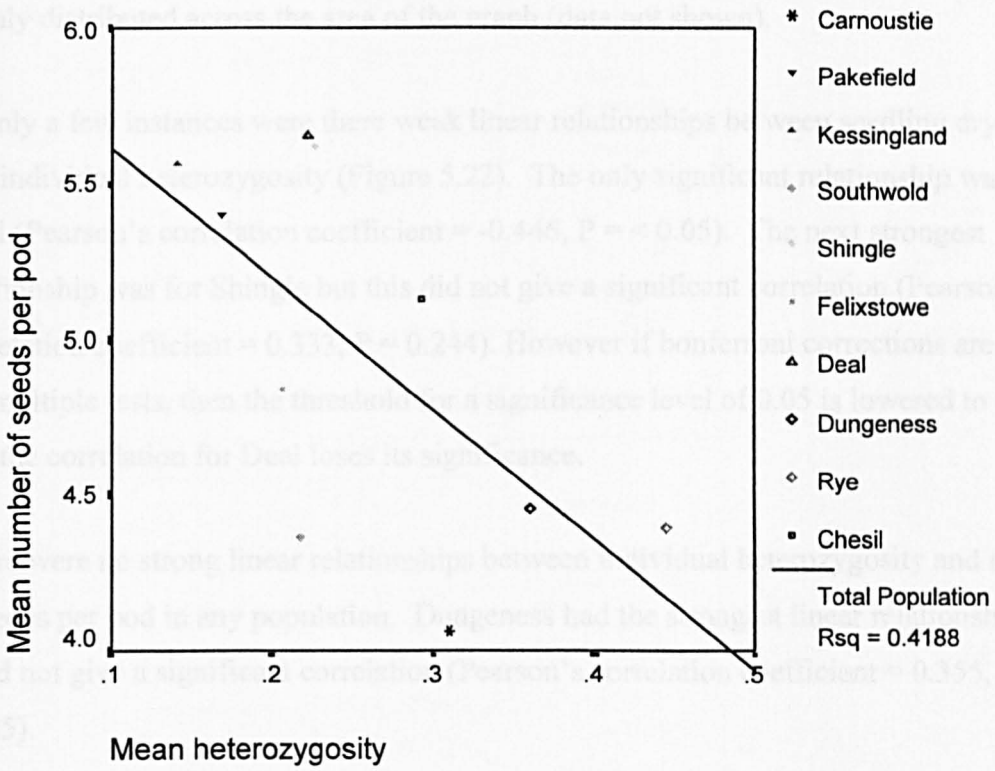

Figure 5.21. Relationship between number of seeds per pod and heterozygosity. Both measures are population means from adults in the field.

No linear relationship was detected between allelic richness and percent of seeds surviving through their first winter, seed weight in either year, or number of seeds per pod. Also, no linear relationship was detected between heterozygosity and seed weight in either year, number of pods per cluster, and number of pigment forms.

5.3.5 *Individual level heterozygosity and fitness measures*

In no population was there a linear relationship between clump size and individual heterozygosity. When these variables were plotted against one another, the points were evenly distributed across the area of the graph (data not shown).

In only a few instances were there weak linear relationships between seedling dry weight and individual heterozygosity (Figure 5.22). The only significant relationship was for Deal (Pearson's correlation coefficient = -0.446, P = < 0.05). The next strongest relationship was for Shingle but this did not give a significant correlation (Pearson's correlation coefficient = 0.333, P = 0.244). However if bonferroni corrections are made for multiple tests, then the threshold for a significance level of 0.05 is lowered to 0.025, and the correlation for Deal loses its significance.

There were no strong linear relationships between individual heterozygosity and number of seeds per pod in any population. Dungeness had the strongest linear relationship but it did not give a significant correlation (Pearson's correlation coefficient = 0.355, P = 0.125).

There were also no strong linear relationships between individual heterozygosity and number of pods per cluster in any population. The strongest linear relationship was present for Kessingland but again it did not give a significant correlation (Pearson's correlation coefficient = -0.331, P = 0.086).

Some linear relationships were present between seed weight and individual heterozygosity, although the strongest linear relationship, present in Carnoustie, did not give a significant correlation (Pearson's correlation coefficient = -0.548, P = 0.065). Other populations with weaker linear relationships (but not significant) between heterozygosity are Rye, Deal, Shingle, and Southwold (Figure 5.23).

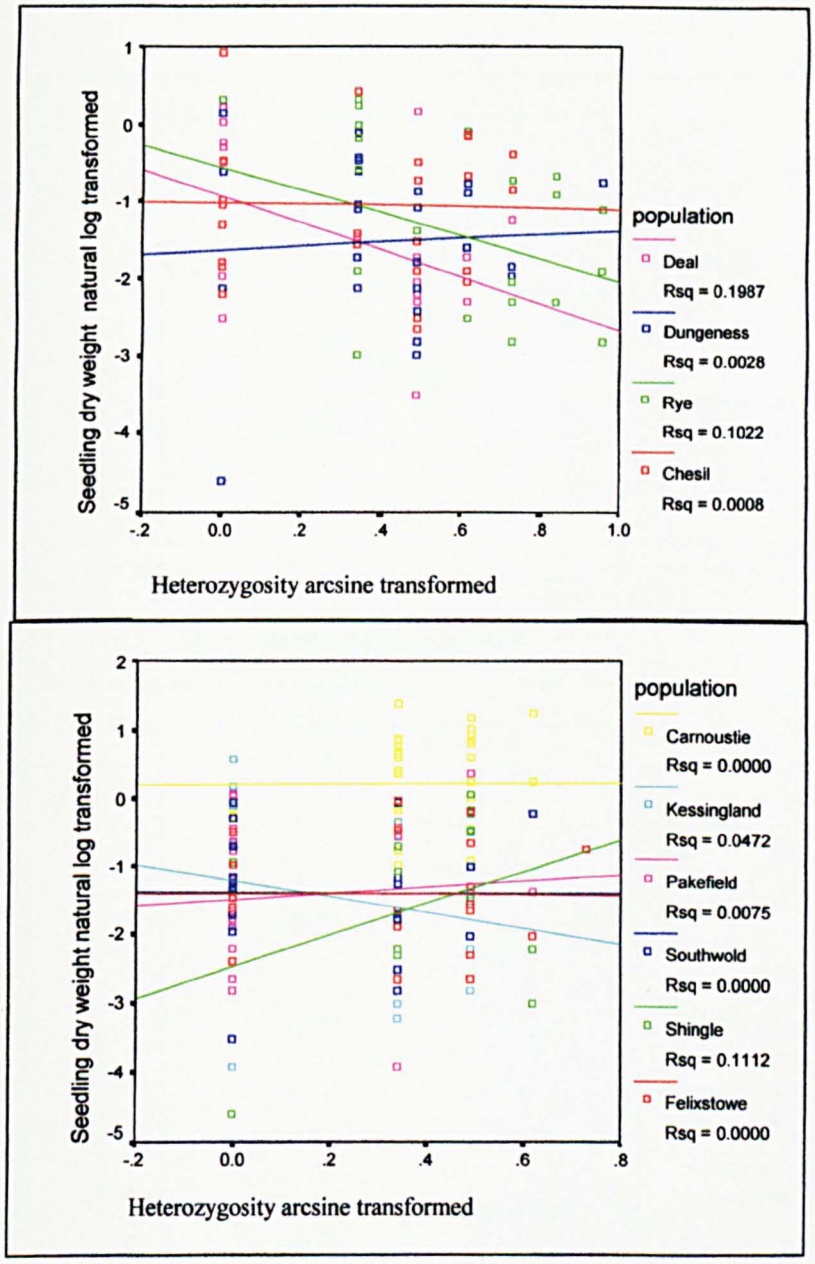

Figure 5.22. Scatter plot of heterozygosity against seedling dry weight in each population. Heterozygosity is of individual seedlings and is arcsine transformed. Seedling dry weight is natural log transformed.

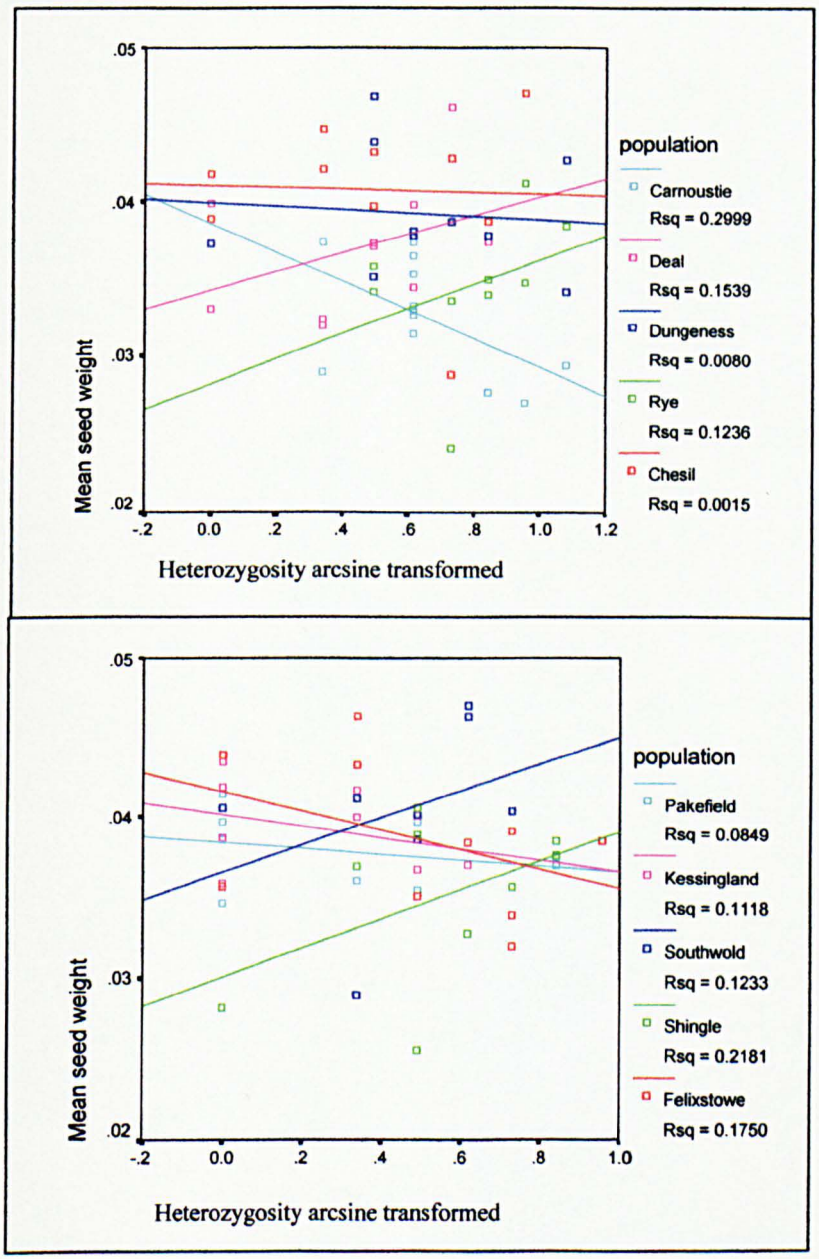

Figure 5.23. Scatter plot of heterozygosity against mean seed weight for each population.
Heterozygosity is arcsine transformed, and is for individual adults paired with seed weight from the same individual.

5.3.6 *Fitness measures at various heterozygosity levels*

Grouping individuals by level of heterozygosity, and looking at the mean and 95% confidence intervals of fitness-associated-measures offers another way of visualizing the data. Figure 5.24 shows mean seed weight for each heterozygosity class, Figure 5.25 shows seedling dry weight for each heterozygosity class, with individuals from Carnoustie removed due to the extreme size difference, and Figure 5.26 showing number of seeds per pod for each heterozygosity class. These graphs show clearly the lack of significant differences between means.

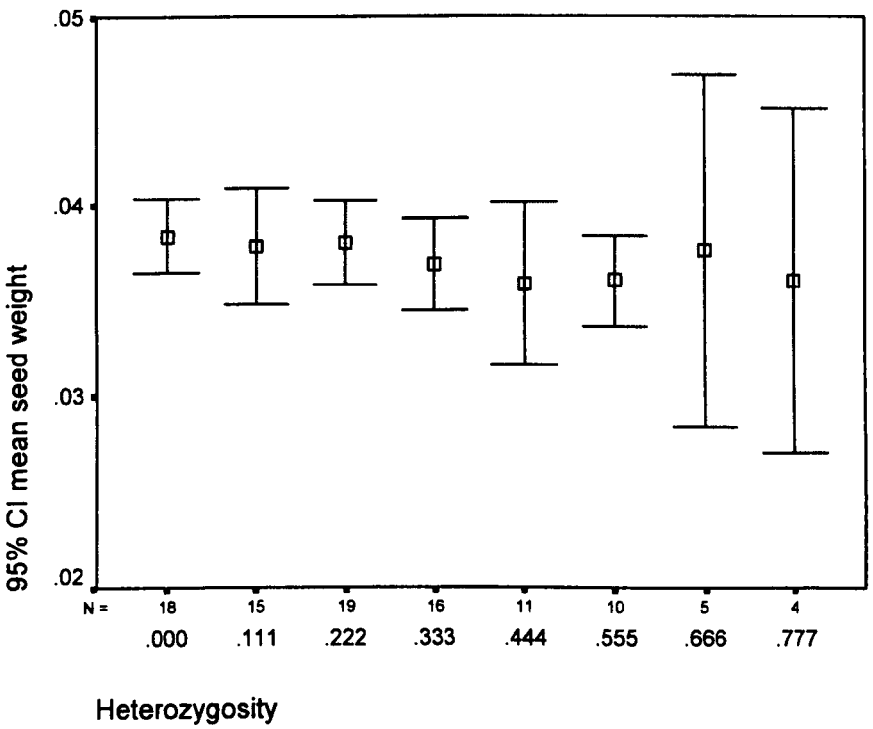

Figure 5.24. Mean seed weights grouped by level of heterozygosity.

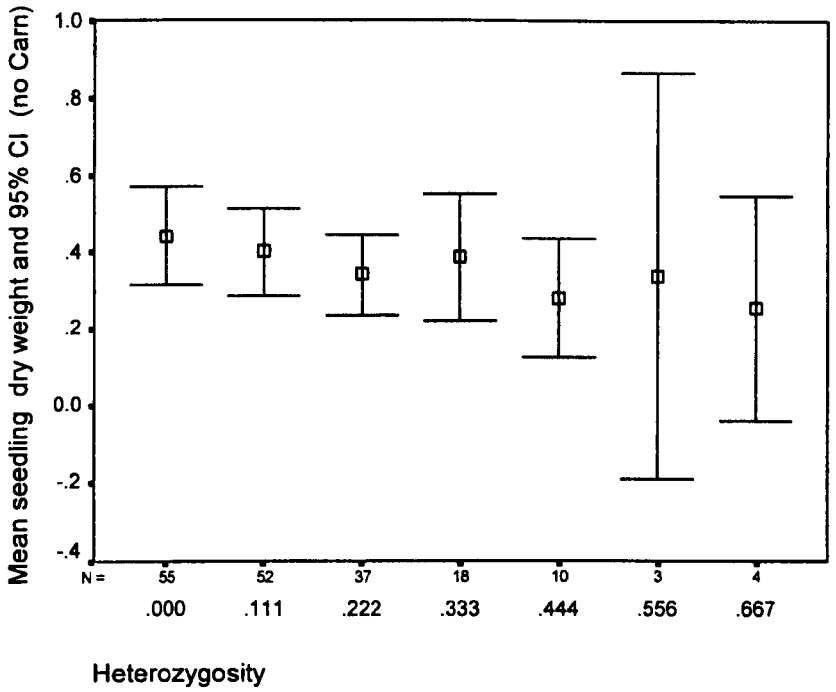

Figure 5.25. Dry weights of seedlings in cultivation grouped by level of heterozygosity. Heterozygosity is measured from the seedlings. Carnoustie is excluded because of large size of plants causing bias.

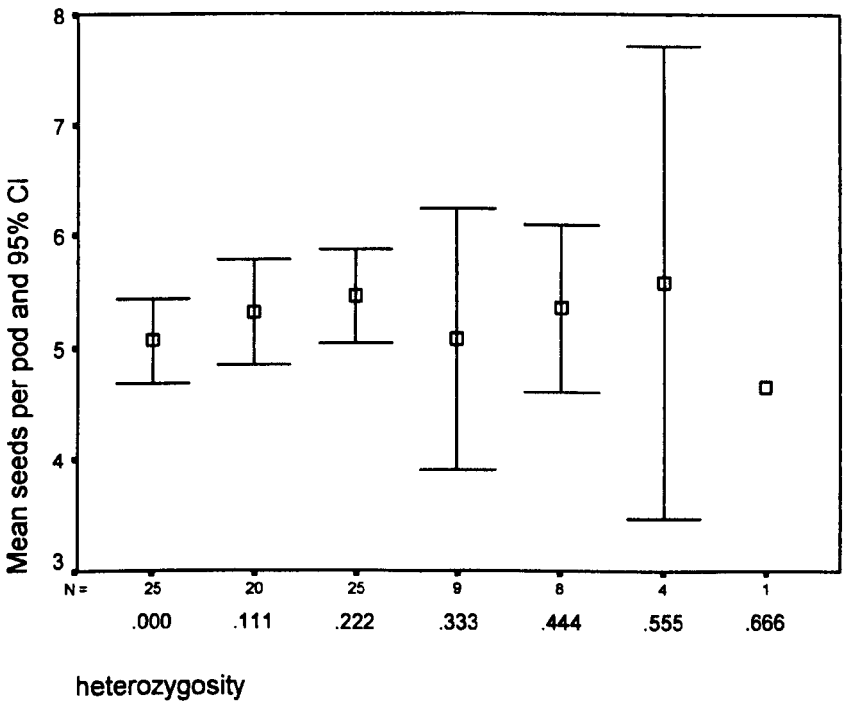

Figure 5.26. Number of seeds per pod grouped by mean heterozygosity.

5.4 Results summary

5.4.1 *Differences among populations in phenotypic characters*

- There was a significant negative correlation between seedling germination and the predation levels experienced by different populations.
- Significant differences were detected in the numbers of seedlings surviving the 1st winter (all of the low survival rates (<65%) are from range centre populations). The highest survival rates are from the most isolated populations (Carnoustie and Chesil).
- Significant differences were detected between seed weights of different populations. In both years seeds from Chesil and Southwold are larger than average, seeds from Rye, Carnoustie and Nigg are smaller than average. There is also some year-to-year variation in seed size.
- Significant differences were detected in leaf shape between different populations. Plants from Carnoustie have narrower leaves in both adult and seedling populations. Plants from Chesil have wider leaves, although this difference was not significant in the seedling population.
- A significant correlation was detected between the dry weight of individual seedlings and the variance in that individuals leaf shape.
- No significant differences were detected in the variance of leaf shapes in different populations.
- There is a significant difference in the number of pods per pod cluster between populations (south coast populations have the most, range centre populations have intermediate numbers, Carnoustie has the least).
- There is a significant difference in the mean number of seeds per pod in different populations. This does not vary with geography or genetic variation, but it does show some association with population size. All of the large populations have the most seeds per pod. At the individual level there is a correlation between the number of seeds per pod, and the number of pods per cluster.
- There were significant differences between the dry weight of seedlings from different populations. Seedlings from Carnoustie were much larger than those from other populations, those from Nigg are slightly smaller.
- There were clear differences in pigmentation of seedlings between populations. South coast populations had the least red pigmentation, range centre populations intermediate levels, and Carnoustie had heavy red pigmentation.

5.4.2 *Comparisons with genetic data*

5.4.2.1 Sample wide

- There was no linear relationship between heterozygosity and survival of individual seedlings through the first winter.

5.4.2.2 Population genetic comparisons

- There were no significant correlations between heterozygosity or allelic richness and any fitness surrogate measures either as population means or between individuals, although there were some linear relationships.
- There was no linear relationship between allelic richness and number of seeds per pod, or number of different pigment forms, or seed weight, or survival through the first winter.
- There was no linear relationship between heterozygosity and number of seeds per pod, number of pods per cluster, number of different pigment forms, seed weight, or survival through the first winter.

5.4.2.3 Individuals within populations

- There was no linear relationship between heterozygosity of individuals within populations and clump size, seedling weight, number of seeds per pod, number of pods per cluster or seed weight.

5.4.2.4 Grouping by heterozygosity

- There were no linear relationships between heterozygosity classes and seed weight, seedling weight, or number of seeds per pod.

5.5 Discussion

Levels of observed heterozygosity were significantly higher in the adult populations of *L. japonicus* than in the seedlings in cultivation, while expected heterozygosity was not significantly different (Chapter 4). This indicates that heterozygotes have been selected for throughout their lives in the field. This has been reported in several other studies and is usually associated with long-lived perennials and stressful conditions (Ledig *et al.*, 1983, Stilwell *et al.*, 2003). Having demonstrated a selective advantage for heterozygotes in *L. japonicus,* can this advantage be picked up by measuring fitness-associated traits such as reproductive output and vigor?

5.5.1 Heterozygosity, allelic richness, and fitness

In this study, significant differences in levels of heterozygosity were detected between populations in the field and their progeny in cultivation (Chapter 4), and significant differences in fitness-associated traits between populations were also detected (this chapter). Yet there were no correlations between fitness-associated traits and either heterozygosity or allelic richness that could not be adequately explained by chance variation alone.

Although population means data gave some linear relationships with genetic variability, the directions of even the strongest of these relationships were not consistent, with numbers of pods per cluster and the number of pigmentation forms having positive relationships with allelic richness (Figure 5.19, Figure 5.20), and number of seeds per pod having a negative relationship with heterozygosity (Figure 5.21).

When the data are examined at an individual level, there are no relationships between heterozygosity and fitness that cannot be explained by chance alone. Most fitness associated traits showed no linear relationships with heterozygosity. There were weak

linear relationships present between seedling dry weight and heterozygosity in only three of ten populations, and two of these were negative and one positive (Figure 5.22). There were linear relationships present between individual heterozygosity and seed weight in seven of the ten populations (Figure 5.23). Four of the seven relationships were positive, and three negative. In the data set there is a trend of linear relationships but no statistical significance, possibly because the sample size (10 individuals per population in seed weight, ~40 in seedling dry weight) for these measures was not large enough to give much statistical power.

In no population was there a linear relationship between clump size and individual heterozygosity. When these variables were plotted against one another, the points were evenly distributed across the area of the graph (data not shown). It is very possible that clump size is more a measure of age than vigor, as size was recorded in a natural population, with no knowledge of age of individuals.

There was also no significant correlation between genetic variability measures and phenotypic variability measures. Thus there was no significant correlation between the variance of leaf shape in different populations nor the number of different pigment forms, with levels of genetic variability.

So overall, in spite of the heterozygote advantage noted in adult versus seedling plants in *L. japonicus* populations in Britain, this study did not detect a relationship between fitness associated traits or variability with heterozygosity or allelic richness.

5.5.2 Why was no correlation between genetic variation and fitness traits detected?

There are several potential reasons for this discrepancy.

Firstly, it is possible that heterozygote advantage is present only in a limited stage of the life cycle, where more homozygotes than heterozygotes die, and in adulthood there is no difference in vigor or reproductive output of the surviving plants. In nature, each adult *L. japonicus* plant in a healthy population drops hundreds to thousands of seeds onto the shingle beneath it (this is easy to see by taking a quick count of the pod clusters, then extrapolating up using averages obtained when counting and weighting seeds), and yet microsatellites have shown that the clumps examined are made up of a single individual (or a small number of individuals) (Chapter 4). This means that a very low proportion of seeds actually germinate. And indeed, seedlings are a rare observation in the field. It has been reported that seedlings are very sensitive to drought before they become established (Brightmore & White, 1963). In cultivation I removed this limiting factor by nicking the seed coat and soaking the seeds for two days, then keeping the germinating seeds well watered, and the germination rates were uniformly high (80-90%) in most populations (Table 5.2). Where the germination rate was lower, it was correlated with predation damage (Figure 5.2). So it seems possible that in the field, seedling establishment could be a life history stage that could potentially eliminate individuals that are more homozygous, while those homozygous individuals that do survive this stage, may have a comparable fitness to more heterozygous individuals later on in life. However mortality at other life history stages could also limit survival to adulthood.

Secondly, it is possible that the measures of fitness used here are not actually measuring the fitness attributes of most importance to the plants in the field. However, an attempt was made to control for this problem by measuring multiple different fitness measures, both in the field and in cultivation. Seed weight is a commonly used fitness measure in studies comparing heterozygosity and fitness. In *Lathyrus japonicus*, seed weight from adult plants in the field showed some year-to-year variation in populations (Figure 5.5), which suggests caution should be exercised in the interpretation of single year fitness measures based on seed weight. Because seed weights were not consistent across populations across years, the differences may reflect a response to local ecological/climatic conditions rather than inherent genetic quality. Variation in seed

weight has been observed to be highly influenced by environmental conditions in other studies (Ouborg & Van Treuren, 1995). The number of seeds per pod, and number of pods per cluster were also measured, and are perhaps better indicators of fitness. And as they are not correlated with seed weight it is unlikely that the variation in these characters is solely related to allocation of resources. Ideally an entire lifetime reproductive output would be a useful measure (though almost impossible to obtain from a long lived perennial species) and indeed, if the cumulative effect of minor advantages are the reason for heterozygote success, this measure should correlate positively with heterozygosity, while perhaps reproductive output in any given year may not.

Thirdly, if the relationship between heterozygosity and fitness is present but very weak in any given year, with an additive effect throughout the lifetime of the plants, there may not have been sufficient statistical power to pick up very small fitness advantages. For the population means, with only ten populations, there was little statistical power in this study. Generally the linear relationships at the population level were stronger than those in the individual data, yet there were more statistically significant correlations in the individual data sets (which had considerably greater sample sizes). And when individuals were grouped into classes based on heterozygosity level, a lack of samples in the higher heterozygosity classes led to broad confidence intervals effectively limiting the power for comparisons with less heterozygous classes. There is a possibility that if sample sizes were larger, that a significant relationship might have been picked up. However, it should be noted that for some potentially useful characters (e.g. the number of seeds per pod and number of pods per cluster) data was available for 30 individuals from each population, and still no linear relationships were detected.

5.5.3 What do the fitness associated traits show?

Significant differences were found between populations for several characters. For seedlings in cultivation, significant differences were noted between populations for survival through the first winter (Figure 5.4), leaf shape (Figure 5.11), dry weight

(Figure 5.14), and pigmentation (Figure 5.16). For adults in the field, significant differences were noted between populations for numbers of seeds per pod (Figure 5.13), number of pods per cluster (Figure 5.12), and seed weight (Figure 5.7, 5.8, 5.9; but seed weight showed significant variation from year to year, and this is likely to reduce its suitability as a fitness measure; see above). Likewise, though there was variation in germination rate (Figure 5.3), this is correlated with the level of predation of the seeds, and is probably in this instance a measure of environmental stress, rather than fitness.

There were, however, some clear differences between populations that could be considered to be ecologically and environmentally relevant measures of performance differences. Seedlings grown from the Carnoustie population, in spite of coming from the lightest seeds, were considerably larger after two summers' growth than seedlings grown from all of the English populations (Figure 5.14). This better performance of the Carnoustie plants growing in Edinburgh indicates better adaptation to the local (Scottish) environmental conditions than plants from further south. This factor was clearly more important for the plants growth and survival than the low allelic diversity present in this population (Chapter 4; Table 4.2; this chapter, Figure 5.14).

The seedlings grown from the other Scottish population (Nigg Bay) were, in contrast, some of the least vigorous seedlings in the entire data set (Figure 5.14). These seedlings must be self-pollinated as only a single genotype is present at this site. It is tempting to attribute the poor performance of these seedlings to inbreeding depression, but without experimental replication this remains speculation. This population is genetically and morphologically more similar to the English populations than to Carnoustie (although the seed size of the two Scottish populations are similar; Figure 5.6 & 5.8). An alternative explanation for this result is that the Nigg Bay plants represent a recent planting/dispersal event and the lack of vigour is attributable to a lack of adaptation to local conditions.

There was a significant difference between populations in the number of pods per cluster, with a decline in the number of pods per cluster correlating with geography (Figure 5.12). This trend show a decreasing number of pods moving from the south

coast populations (Chesil, Rye, Dungeness and Deal), through the Suffolk range center populations northwards, with the isolated population at Carnoustie having the least number of pods per cluster. It is not clear why this relationship should exist and what the explanatory variable could be. The relationship is clear, even if the explanation is not. There is also a significant difference in the mean number of seeds per pod in the different populations (Figure 5.13). This does not show the same correlation with geography or with any measures of genetic variation. It does, however, show some association with population size. All of the large populations have the most seeds per pod (Table 5.1, Figure 5.13). The medium sized populations have intermediate numbers of seed per pod, and the smallest populations the fewest seeds per pod. This could potentially be attributable to pollinator behaviour, with larger populations offering greater floral displays and hence attracting greater pollinator activity. However, the number of seeds per pod is small (<10) and it is not clear whether pollen limitation is likely to operate on such a small number of ovules in an entomophilous flower, in which one would expect the number of pollen grains deposited per visit to be sufficient to fertilise all of the ovules.

In contrast to the number of seeds per pod (largest in large populations), survival through the first winter showed no clear association with population size (Table 5.1, Figure 5.4). Indeed the populations that experienced the lowest survival through the first winter were all large range centre populations. Smaller populations and more isolated populations showed greater survival rates. This is not predictable from population genetic expectations.

Although the intra-population microsatellite variation did not correlate with fitness measures in this study, the microsatellite differentiation between populations did reflect morphological differentiation in the range edge populations such as Chesil beach and particularly Carnoustie. The population at Carnoustie was genetically very distinct from any other population of *L. japonicus* in Britain based on microsatellites (Chapter 4, Figure 4.10). The population at Chesil Beach in Dorset was also genetically distinct (Chapter 4, Figure 4.10). These two populations, however, showed some greater

microsatellite similarities to each other, than they did to the range center populations, despite the large geographical distance between them (Chapter 4, Figure 4.10).

However, despite the 'greater than expected similarities' between these two populations based on microsatellite data, these populations were morphologically the most different from each other in the data set. Plants from Carnoustie had much narrower leaves than all of the other English populations (Figure 5.10). This was visible at a glance, and held up when seedlings were grown in cultivation (Figure 5.11). This difference in leaf shape was highly statistically significant. Chesil Beach, the population at the southwest UK range edge, had wider leaves than the other populations (Figure 5.10 & 5.11). This was not initially so obvious in the field, but was easily noticeable when seedlings were together in cultivation. However, ironically this difference in leaf shape at Chesil Beach was significantly different based on field measurement, but not significantly different based on measurements from the seedlings in cultivation (Figure 5.10 & 5.11). This type of result could indicate environmentally induced variation. However, given that a difference was still clearly visible in cultivation (even if not captured as significant by the test statistic) an alternative explanation requires consideration. Seedlings from Carnoustie were all fairly large, while seedlings from other populations varied widely in size. It could be that the lack of statistical significance is due to measurements from the smaller individuals (which had not yet developed the distinctive adult leaf proportions) affecting the results. In this respect it is noteworthy that the variance in leaf shape in individual seedlings shows a strong linear relationship with seedling weight (Pearson's correlation coefficient = -0.308, $P<0.01\%$).

Another clear geographically structured phenotypic difference among populations was seedling pigmentation (Figure 5.16). Although this character was only evident in very young seedlings, while it was expressed, the differences were striking. There was a marked gradation from west to south to north, with the two range edge populations being the most different. Thus the phenotypic differences between populations are better associated with geography than the microsatellite data. Both measures suggest that the populations from Chesil and Carnoustie are genetically different, but the affinity

between these two populations suggested by the microsatellites, is not supported by the phenotypic data.

The populations of narrow leaved *L. japonicus* on Shetland have been called subspecies *acutifolia*, and this study supports this distinction due to the high number of unique alleles present and the ecological differentiation, as well as the narrow-leaved morphology present in the Carnoustie population.

5.5.4 Summary

This study has revealed a mixture of congruence and incongruence of fitness results and the genetic data. In terms of levels of intra-population genetic diversity, there is virtually no association between this and fitness measures. This may be due to a genuine lack of an association, or a weakness in statistical power. The enhanced heterozygosity of adults compared to seedlings suggests that a fitness difference is present, but capturing this difference in an experiment was not achieved in the current study.

What is clear, however, from measuring fitness-associated traits is the extent of differences among populations. These populations show significant genetic differentiation for microsatellites, but they also show strong phenotypic and performance measure differences. The populations are not selectively equivalent to one another. There is clear evidence of local adaptation, and also phenotypic differentiation in relation to geography.

Chapter 6. Conclusions

6.1 THEME 1: Phylogeography

6.1.1 Identification of diversity hotspots

Genetic markers can contribute to conservation by aiding in the identification of diversity hotspots. The relatively slow mutation rate and maternal inheritance of organelle markers makes them well suited to tracking the broad scale distribution of genetic biodiversity (Ennos *et al.*, 1999). Populations that occur in regions that have been indicated to be ice free throughout the Pleistocene by fossil pollen and climate studies, often have the highest genetic divergence (Petit *et al.*, 2003). This means that the patterns of genetic diversity can be used at least to some extent to extrapolate refugial areas for species for which fossil pollen data is not available. The patterns of genetic diversity are well known for temperate European taxa (Ferris *et al.*, 1998; King & Ferris, 1998; Demesure *et al.*, 1996; Rendell & Ennos, 2002; Grivet & Petit, 2002; Mohanty *et al.*, 2002; Raspe *et al.*, 2000; Palme *et al.*, 2003), and a synthesis of this information has been carried out, allowing generalizations to be made (Petit *et al.*, 2003). This has led to a call for conservation prioritization of southern European populations: "Because most northern European populations are eliminated during glacials, the identification of the locations of southern long-term refugia should be a conservation priority" (Tzedakis *et al.*, 2002).

However diversity hotspots will be different for different elements of a flora. More cold tolerant species in Europe, such as *Pinus sylvestris* (Sinclair *et al.*, 1999), and *Betula* (Huntley & Birks, 1983), do not follow the same phylogeographic patterns identified for temperate species (Ferris *et al.*, 1999). Arctic plants, because of their different ecological requirements will respond differently to glacial cycles.

As there is something of a bias in the phylogeography literature towards temperate species in Europe, there is likely to be a bias in the areas being highlighted as the most important diversity hotspots, when actually what is being described are the most important diversity

hotspots for a specific set of organisms. There are simply not enough data available to make generalizations about other elements of the European flora. Yet regional conservation programmes need to reflect all element of a flora. Identification of refugial sites for temperate trees will not help in devising regional conservation programmes for arctic alpines.

Many plants threatened in the UK are northern in distribution, so it is important to understand the history of this element of the flora, as it is populations of these species that are likely to be particularly sensitive to any future global warming which may greatly reduce or eliminate their habitat in the UK (Lusby & Wright, 1996). There is now a growing body of evidence on the importance of high arctic refugia for some species of plants. There is fossil evidence that the Beringian region remained unglaciated and was covered by a mix of various types of tundra throughout the Pleistocene (Brochmann *et al.*, 2003). Phylogeographic studies on *Dryas integrifolia* (Tremblay & Schoen, 1999), *Saxifraga oppositifolia* (Abbott *et al.*, 2000) and *Silene acaulis* (Abbott *et al.*, 1995), show high diversity in this region. The *Saxifraga hirculus* data from Chapter 2 also supports the presence of a Beringian refugium, with populations from this region having more haplotypes and lower partitioning of genetic variation between populations.

6.1.2 Is there any evidence for cryptic northern refugia in the UK?

Accepting the importance of Alaska for arctic plants, it is also worth asking whether there is evidence for any similar diversity hotspots in Europe. There have, for instance, been claims of the presence of cryptic northern refugia occurring in the UK (Stewart & Lister, 2001; Stewart, 2003). The evidence supporting the presence of northern refugia is primarily related to assemblages of mammals, typically associated with deciduous woodland, being found further north than the occurrence of deciduous woodland during the late Pleistocene. Stewart & Lister (2001) also cite studies on population differentiation in western Scottish Scots Pine populations (Sinclair *et al.* 1999) as potential evidence of cryptic local refugia. However, it is worth qualifying these claims. Firstly mammals found in deciduous woodlands can also be found in open habitats and coniferous environments

2003), and there is the possibility of behavioral and adaptive changes during periods of environmental change. Secondly the references to Scots Pine rather misrepresent the inferences of the initial studies. The distinct lineages of Scots Pine in western Scotland were postulated to perhaps stem from a different (western) refugium to other Scots Pine, with southern Ireland being suggested as a possible source (Sinclair *et al.*, 1999), rather than suggesting that the species survived *in situ* in western Scotland as Stewart & Lister (2001) imply.

The population of *Saxifraga hirculus* from the Pentlands (Chapter 2) does possess the most genetically distinct haplotype detected in the study. This is an unexpected result. It may indicate the presence of an isolated and relic population that survived in or close to the UK. Alternatively it may involve post-glacial colonization from an as yet unsampled area (Chapter 2). Despite the large amount of data gathered from this study, it remains the case that more samples are required for *Saxifraga hirculus*, and more studies are required from other species before detailed insights can be gained into the phylogeography of arctic plants. The greatest difficulty in doing this will be undertaking the widespread sampling that is undoubtedly required, especially along the north coast of Russia. The detailed studies of refugial populations of temperate species in Europe have benefited by their coincidental co-occurrence with accessible regions in developed countries. Forming a picture of arctic phylogeography will present a much greater logistical challenge.

6.1.3 Does phylogeographic data only reflect diversity of chloroplast DNA or can it reflect general trends in intra-specific biodiversity?

It is worth stepping back from the data and asking a general question as to the extent to which organelle phylogeographies are representative of intra-specific genetic biodiversity. In angiosperms, cpDNA contains only a 100 or so genes, mtDNA contains only ca 40 genes (Palmer, 1987, 1992; Li, 1997). Should one use data from such small genomes to make inferences on the distribution of intra-specific genetic biodiversity? By far the largest

component of genes in an individual are stored in the nuclear genome in higher plants, and due to pollen flow, these nuclear genes may have independent distributions from the organelle genes.

The only study I am aware of to test this issue is in the European Oaks (Kremer *et al.*, 2002). A study of chloroplast variation, phenotypic variation and nuclear DNA markers was undertaken on samples from France, Britain and Germany. An association between organelle markers and nuclear markers was detected with evidence of cytonuclear disequilibrium (Kremer *et al.*, 2002). There was an association between both nuclear and organelle lineages, as well as comparative similarity in different levels of genetic variation from the different genomes. However, less association was detected with regards to phenotypic traits. This indicates that while the organelle DNA reflects broad scale genomic history, local selection and adaptation can influence individual phenotypic traits. Nevertheless, it is noteworthy that the organelle DNA did to some extent reflect broad scale genomic differentiation for the nuclear genome. This is especially pertinent given the reproductive biology of oaks. A pollen:seed flow ratio for oaks has been calculated at *ca.* 200:1 (Ennos *et al.* 1999). This high level of pollen flow (wind pollinated tree releasing pollen from a high canopy) would be expected to rapidly erode cytonuclear disequilibrium. If one considers insect pollinated herbs, one would expect the pollen:seed flow ratio to be much lower (Squirrell *et al.*, 2001), or at least that pollen flow would be less extensive (and cytonuclear disequilbria to be much higher). Under this scenario, it would seem even more likely that organelle phylogeography will reflect broad genome wide differences. Thus phylogeographic structure is likely to be more useful in this type of species for establishing the geographical distribution of intra-specific biodiversity.

A recent analysis of phylogeographic structure in European butterflies showed a correlation between levels of genetic variation and population demography (Schmitt & Hewitt, 2004). Populations from hypothesized refugial areas, with high levels of diversity, showed greater demographic stability than populations with less diversity that were distant from refugial areas. In contrast, species with no differentiation in levels of diversity across their ranges showed greater demographic stability across their range (Schmitt & Hewitt, 2004).

Although this is the only study of its type, it again shows an important association between phylogeographic data and conservation related issues. *Saxifraga hirculus* is declining in much of Europe, and Chapter 2 of this thesis indicates that European populations of this species have the low cpDNA diversity characteristic of populations far from refugial areas, and this echoes the patters found by Schmitt & Hewitt (2004). Also, a study by Dahlgaard & Warncke (1995) shows that *S. hirculus* plants in Denmark, resulting from cross pollination of individuals from different populations, grew and survived better than plants with parents from the same population, and hence the species is suffering from inbreeding depression. Further fitness/demographic studies set in a phylogeographic framework would be desirable.

6.2. THEME 2: Evolutionary fitness and adaptive variation.

6.2.1 The theory is well established, but in practice fitness is difficult to measure

The population genetic theory underlying the isolation and size of populations with respect to levels of neutral marker variability is straightforward (Ellstrand & Elam, 1993). Likewise, the theory underlying the potential for genetic depauperacy to impact on fitness is also relatively straightforward (Frankham *et al.*, 2002). However, as this study demonstrated (Chapter 4 & 5), the practical reality of uncovering these associations can be far from simple. As has been highlighted in the above discussion on phylogeography, historical events are likely to impact on patterns of population genetic diversity. Observing a set of present day extant populations hides a series of often complex historical events that have led to their present day distributions and patterns of genetic diversity.

One potential factor worth exploring in the future would be to assess whether there are improved correlations between demographic observations (size, isolation) with genetic

factors (levels of variation, fitness) in relation to species level F_{ST}. If gene flow can (even erratically) occur over long distances, even low levels of gene flow might impact on patterns of population genetic variation, particularly for loci under selection (Rieseberg *et al.*, 2003). In contrast, associations among demographic factors are likely to improve if 'isolated' populations are genuinely isolated, rather than being wedded by low levels of gene flow. But even here, population history will be important, and large populations may be genetically depauperate if historically they have not always been large. It is also noteworthy that in species like *Lathyrus japonicus* which can reproduce asexually as well as sexually, there is an additional level of complexity which may impact on the maintenance of genetic variation and also fitness associated traits.

6.2.3 Does our data source alter our understanding?

In the current study on *Lathyrus japonicus,* essentially three sets of data were available: demographic data (population sizes and distributions), genetic marker data, and data on fitness associated traits. If one considers within population measures of diversity, it is clear that we would not have been able to predict information about one type of data from the others (Chapters 4 & 5). This highlights the importance of multiple data sources and the difficulties of extrapolating from census/distribution data alone. However, there was a higher congruence between geographical distributions of populations and the distribution of genetic and phenotypic variation. This offers some reassurance that based on field/distributional observations alone, one would have been able to make some predictions on the organization of genetic biodiversity in this species. However, even this is not straightforward. The high divergence of the Pentlands *Saxifraga hirculus* population that was geographically relatively close to some other genetically very different populations illustrates how cryptic the organization of genetic biodiversity can be. Likewise the lack of association of infraspecific differences in *Saxifraga hirculus* with the genetic data also illustrates a potential uncoupling of data sources. At the risk of reiterating a point, it is clear that if there are ancient phylogeographic events overlaid on factors affecting

contemporary population structure, then obtaining generalities across data sources will be difficult. Probably the major challenge remaining is truly extensive sampling and experimentation on populations of a range of taxa with a range of demographic and life history characteristics for all three data sources. Until this is undertaken on a large scale, the chances of generalities emerging are slim.

6.3 Conclusions

There is a large body of demographic and neutral marker data that has been, and is being, collected. This will offer increased opportunities for understanding the relationships between demographic data and genetic marker data. However, at both a regional conservation level (Theme 1), and on a more localised conservation level (Theme 2), there are still few studies measuring adaptive differentiation and comparing this to levels of genetic diversity, phylogeographic differences and demography. With recent technological advances, gathering the molecular data is becoming increasingly straightforward; gathering the fitness data, in contrast, can be more challenging. But if these fitness data are not gathered, there will continue to be an important aspect of genetic biodiversity (perhaps *the* most important aspect of genetic biodiversity) that we do not understand or know how it is distributed.

References

Abbott RJ, Chapman HM, Crawford RMM, Forbes DG (1995) Molecular diversity and derivations of populations of *Silene acaulis* and *Saxifraga oppositifolia* from the high Arctic and more southerly latitudes. *Molecular Ecology*, **4**, 199-207.

Abbott RJ, Smith LC, Milne RI Crawford RMM, Wolff K, Balfour J (2000) Molecular analysis of plant migration and refugia in the Arctic. *Science*, **289**, 1343-1346.

Abbott RJ, Brochmann C (2003) History and evolution of the arctic flora: in the footsteps of Eric Hultén. *Molecular Ecology*, **12**, 299-313.

Akeroyd JR (1994) *Lathyrus japonicus*. In: *Scarce plants in Britain.* (eds. Stewart A, Pearman D, Preston CD), pp. 223. Joint Nature Conservation Committee, Peterborough.

Asmussen C (1993) Pollination biology of the Sea Pea, *Lathyrus japonicus*: floral characters and activity and flight patterns of bumblebees. *Flora,* **188**, 227-230.

Avise J (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Harvard.

Barrett SCH, Kohn JR (1991) Genetic and evolutionary consequences of small population size in plants: implications for conservation. In: *Genetics and conservation of rare plants* (eds. Falk DA, Holsinger KE), pp. 3-30. Oxford University Press, Oxford.

Bennett KD, Tzedakis PC, Willis KJ (1991) Quaternary refugia of north European trees. *Journal of Biogeography*, **18**, 103-115.

Blaxter M, Floyd R (2003) Molecular taxonomics for biodiversity surveys: already a reality. *Trends in Ecology and Evolution*, **18,** 268-269.

Borgen L (1987) Postglacial evolution in the Nordic flora a review. *Blyttia*, **45**, 147-168.

Bradley RS (1985) *Quaternary paleoclimatiology: methods of paleoclimatic reconstruction.* Allen and Unwin, Boston.

Brightmore E, White PHF (1963) *Lathyrus japonicus* Willd. Biological Flora of the British Isles, No. 94. *Journal of Ecology*, **51**, 795-801.

Brochmann C, Nilsson T, Gabrielsen TM (1996) A classic example of postglacial allopolyploid speciation re-examined using RAPD markers and nucleotide sequences: *Saxifraga osloensis* (Saxifragaceae). *Acta Universitatis Upsaliensis*

MISSING PAGES ARE UNAVAILABLE

Dumolin-Lapegue S, Demesure B, Fineschi S, Le Corre V, Petit RJ (1997) Phylogeographic sturcture of white oaks throughout the European continent. Genetics, **146**, 1475-1487.

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques*, **20**, 758-760.

Edwards ME, Anderson PM, Brubaker LB, Ager TA, Andreev AA, Bigelow NH, Cwynar LC, Eisner WR, Harrison SP, Hu FS, Jolly D, Lozhkin AV, MacDonald GM, Mock CJ, Ritchie JC, Sher AV, Spear RW, Williams JW, Yu G, (2000) Pollen-based biomes for Beringia 18,000, 6000 and 0 14C yr BP. *Journal of Biogeography*, **27**, 521-55.

Ehlers BK, Olesen JM (2003) Flower and fruit herbivory in a population of *Centaurea scabiosa* (Asteraceae): Importance of population size and isolation. *Ecoscience*, **10** 45-48.

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: Implications for plant conservation. In: *Annual Review of Ecology and Systematics*, **24**, 217-242.

El Mousadik A, Petit RJ (1996) Chloroplast DNA phylogeography of the argan tree of Morocco. *Molecular Ecology*, **5**, 547-555.

Engler A, Irmscher E (1916) Saxifragaceae – *Saxifraga* I. *Das Pflanzenreich 1, IV, 117, 1.* Leipzig.

Ennos RA (1994) Estimating the relative rates of pollen and seed migration among plant populations, *Heredity*, **72**, 250-259.

Ennos RA, Cowie NR, Legg CJ, Sydes C (1997) Which measures of genetic variation are relevant in plant conservation? A case study of *Primula scotica.* In: *The role of genetics in conserving small populations* (eds Tew TE, Crawford TJ, Spencer JW, Stevens DP, Usher MB, Warren J), pp 73-79. JNCC, Peterborough.

Ennos RA, Sinclair WT, Hu X-S, Langdon A (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. In: *Molecular systematics and plant evolution* (eds Hollingsworth PM, Bateman RM, Gornall RJ), pp.1-19. Taylor & Francis, London.

Falk DA, Holsinger KE (1991) *Genetics and conservation of rare plants*. Oxford

MISSING PAGES ARE UNAVAILABLE

Hedrick PW (2001) Conservation genetics: Where are we now? *Trends in Ecology and Evolution*, **16**, 629-636.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276.

Hill HW (1971) Currents and water masses. In: *North Sea Science* (ed. Goldberg ED). MIT Press, London.

HMG (Her Majesty's Government) (1995) *Biodiversity: the UK steering group report.* HMSO, London.

Hollingsworth PM (2003) Taxonomic complexity, population genetics, and plant conservation in Scotland. *Botanical Journal of Scotland*, **55**, 55-63.

Hufford KM, Mazer SJ (2003) Plant ecotypes: Genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution*, **18**, 147-155.

Hughes M, Hollingsworth PM, Squirrell J (2002) Isolation of polymorphic microsatellite markers for *Begoina sutherlandii* Hook. f. *Molecular Ecology Notes*, **2**, 185-186.

Hultén E (1937) *Outline of the history of arctic and boreal biota during the quaternary period.* Lehre J Cramer, New York.

Hultén E (1968) *Flora of Alaska and neighbouring territories*. Stanford University Press, California.

Hultén E (1971) *The circumpolar plants II: Dicotyledons*. Almquist & Wiksells, Uppsalla.

Hunter ML Jnr, Hutchinson A (1994) The virtues and shortcomings of parochialism: conserving species that are locally rare, but globally common. *Conservation Biology*, **8**, 1163-1165

Huntley B, Birks HJB (1983) *An atlas of past and present pollen maps for Europe: 0–13 000 years ago*. Cambridge University Press, Cambridge.

Jalas J, Suominen J, Lampienin R, Kurtto A (1999) *Atlas Florae Europaeae, Distribution of Vascular Plants in Europe 12; Resedaceae to Platinaceae*. The Committee for Mapping the Flora of Europe & Societus Biologica Fennica Vanamo, Helsinki.

Jacquemyn H, Van Rossum F, Brys R., Endels P, Hermy M, Triest L, De Blust G (2003) Effects of agricultural land use and fragmentation on genetics, demography, and population persistence of the rare *Primula vulgaris*, and implications for conservation. *Belgian Journal of Botany*, **136**, 5-22.

Karhu A, Hurme P, Karjalainen M, Karvonen P, Karkkainen K, Neale D, Savolainen O (1996) Do molecular markers reflect patterns of differentation in adaptive traits of conifers? *Theoretical and Applied Genetics*, **93**, 215-221.

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. *Trends in Ecology and Evolution*, **17**, 230-241.

Kindell CE, Winn AA, Miller TE (1996) The effects of surrounding vegetation and transplant age on the detection of local adaptation in the perennial grass *Aristida stricta*. *Journal of Ecology*, **84**, 745-754.

King RA, Ferris C (1998) Chloroplast DNA phyleogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151-1163.

Kremer A, Kleinschmit J, Cottrell J, Cundall EP, Deans JD, Ducousso A, König AO, Lowe A, Munro RC, Petit RJ, Stephan BR (2002) Is there a correlation between chloroplastic and nuclear divergence, or what are the roles of history and selection on genetic diversity in European oaks. *Forest Ecology and Management*, **156**, 75-87.

Lande R (1988) Genetics and demography in biological conservation. *Science*, **241**, 1455-1460.

Ledig FT, Guries RP, Bonefeld BA (1983) The relation of growth to heterozygosity in pitch pine. *Evolution*, **37**, 1227-1238.

Lewis PO, Zaykin D (2001) Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Available from: http://lewis.eeb.uconn.edu/lewishome/software.html.

Lesica P, Allendorf FW (1995) When are peripheral populations valuable for conservation? *Conservation Biology*, **9**, 753-760.

Li WH (1997) *Molecular Evolution*. Sinaeur, Sunderland, Massachusetts.

Lowe A, Herris SE, Ashton P (2004) *Ecological genetics*. Blackwell Publishers, Malden, MA.

Lu SY, Peng CI, Chen YP, Hong KH, Chiang TY (2001) Chloroplast DNA phylogeography of *Cunninghamia konishii* (Cupressaceae), an endemic conifer of Taiwan. *Genome*, **44**, 797-807.

Lusby PS, Wright J (1996) *Scottish Wild Pla*nts. The Stationery Office Limited, Edinburgh.

Lynch M, Conery J, Bürger R (1995) Mutation accumulation and the extinction of small populations. *American Naturalist,* **146**, 489-518.

Marchelli P, Gallo L, Scholz F, Zeigenhagen B (1998) Chloroplast DNA markers reveal a geographical divide across Argentinean southern beech *Nothofagus nervosa* (Phil.) Dim. et Mil. distribution area. *Theoretical and Applied Genetics*, **97**, 642-646.

Mohanty A, Martin JP, Aguinagalde I (2001) A population genetic analysis of chloroplast DNA in wild populations of *Prunus avium* L. in Europe. *Heredity*, **87**, 421-427.

Mohanty A, Martin JP, Aguinagalde I (2002) Population genetic analysis of European *Prunus spinosa* (Rosaeae) using chloroplast DNA markers. *American Journal of Botany*, **89**, 1223-1228.

Muloko-Ntoutoume N, Petit RJ, White L, Abernethy K (2000) Chloroplast DNA variation in a rainforest tree (*Aucoumea klaineana*, Burseraceae) in Gabon. *Molecular Ecology*, **9**, 359-363.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1-10.

Nelson CE (2000) *Sea beans and nickar nuts*. BSBI, London.

Newton AC, Allnutt TR, Gillies ACM, Lowe AJ, Ennos RA (1999) Molecular phylogeography, intraspecific variation and the conservation of tree species. *Trends in Ecology and Evolution*, **14**, 140-145.

Nordal I (1987) Tabula rasa after all? Botanical evidence for ice free refugia in Scandinavia reviewed. *Journal of Biogeography*, **14**, 377-388.

Olesen JM, Warncke E (1989a) Temporal changes in pollen flow and neighbourhood structure in a population of *Saxifraga hirculus* L., *Oecologia*, **79**, 205-211.

Olesen JM, Warncke E (1989b) Flowering and seasonal changes in flower sex ratio and frequency of flower visitors in a population of *Saxifraga hirculus*. *Holarctic Ecology*, **12**, 21-30.

Olesen JM, Warncke E (1990) Morphological, phenological and biochemical differentiation in relation to gene flow in a population of *Saxifraga hirculus*. *Sommerfeltia*, **11**, 159-172.

Oostermeijer JGB, Van Eijck MW, Den Nijs JCM (1994) Offspring fitness in relation to population size and genetic variation in the rare perennial plant species *Gentiana*

pneumonanthe (Gentianaceae). *Oecologia*, **97**, 289-296.

Oostermeijer JGB, Van Eijck MW, Van Leeuwen NC, Den Nijs JCM (1995a) Analysis of the relationship between allozyme heterozygosity and fitness in the rare *Gentiana pneumonanthe* L. *Journal of Evolutionary Biology*, **8**, 739-759.

Oostermeijer JGB, Altenburg RGM, Den Nijs HCM, (1995b) Effects of out-crossing distance and selfing on fitness components in the rare *Gentiana pneumonanthe* (Gentianaceae). *Acta Botanica Neerlandica*, **44**, 257-268.

Ouborg NJ, Van Treuren R (1994) The significance of genetic erosion in the process of extinction. IV. Inbreeding load and heterosis in relation to population size in the mint *Salvia pratensis*. *Evolution*, **48**, 996-1008.

Ouborg NJ, Van Treuren R (1995) Variation in fitness-related characters among small and large populations of *Salvia pretensis*. *Journal of Ecology*, **83**, 369-380.

Paetkau D, Calvert W, Sterling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, **4**, 347-354.

Palme AE, Semerikov V, Lascoux M (2003) Absence of geographical structure of chloroplast DNA in *Salix caprea* L. *Heredity*, **91**, 465-474.

Palmer JD (1987) Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *The American Naturalist*, **130**, S6-S29.

Palmer JD (1992) Mitochondrial DNA in plant systematics: applications and limitations. In: *Molecular systematics of plants*. (eds. Soltis PS, Soltis DE and Doyle JJ). Chapman and Hall, London.

Paschke M, Abs C, Schmid B (2002) Relationship between population size, allozyme variation, and plant performance in the narrow endemic *Cochlearia bavarica*. *Conservation Genetics*, **3**, 131-144.

Petit RJ, Kremer A, Wagner DB (1993) Geographic structure of chloroplast DNA polymorphisms in European oaks. *Theoretical and Applied Genetics*, **87**, 122-128.

Petit R (2000) TURBO PASCAL program (DistoN for Windows and Linux) yields pairwise estimates of diversity. Available at: http://www.pierroton.inra.fr/genetics/labo/Software/Diston/

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Muller-Starck G, Demesure-

Musch B, Palme A, Martin JP, Rendell S, Vendramin GG (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563-1565.

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

Potts BM, Reid JB (1998) Hybridization as a dispersal mechanism. *Evolution*, **42**, 1245-1255.

Preston CD, Pearman DA, Dines TD (2002) *New atlas of the British and Irish flora*. Oxford University Press, Oxford.

Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution*, **16**, 124-147.

Raspe O, Saumitou-Laprade P, Cuguen J, Jacquemart A-L (2000) Chloroplast DNA haplotype variation and population differentation in *Sorbus aucuparia* L. (Rosaceae: Maloideae). *Molecular Ecology*, **9**, 1113-1122.

Rendell S, Ennos RA (2002) Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe. *Molecular Ecology*, **11**, 69-78.

Rendell S, Ennos RA (2003) Chloroplast DNA diversity of the dioecious European tree *Ilex aquifolium* L. (English holly). *Molecular Ecology*, **12**, 2681-2688.

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. *Conservation Biology*, **17**, 230-237.

Rieseberg LH, Church SA, Morjan CL (2003) Integration of populations and differentiation of species. *New Phytologist*, **161**, 59-69.

Ritchie JC, Cwynar LC (1982) The late-Quaternary vegetation of the North Yukon. In: *Paleoecology of Beringia* (eds Hopkins DM, Matthews JV, Schweger CE, Young SB), pp. 113-126. Academic Press, New York.

Ritland K (2002) Extensions of models for the estimation of mating systems using *n* independent loci. *Heredity*, **88**, 221-228.

Ronquist F, Gardenfors U (2003) Taxonomy and biodiversity inventories: time to deliver. *Trends in Ecology and Evolution*, **18**, 269-270.

Rousset F (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics*, **142**, 1357-1362.

Royal Horticultural Society (1966) R.H.S. Colour Chart, Royal Horticultural Society, London.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, **132**, 365-386

Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I, (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature*, **392**, 491-494.

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

Schmidt K, Jensen K, (2000) Genetic structure and AFLP variation of remnant populations in the rare plant *Pedicularis palustris* (Scrophulariaceae) and its relation to population size and reproductive components. *American Journal of Botany*, **87**, 678-689.

Schmitt T, Hewitt GM (2004) The genetic pattern of population threat and loss: a case study of butterflies. *Molecular Ecology*, **13**, 21-31.

Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: A software for population genetics data analysis. Available from: http://lgb.unige.ch/arlequin/.

Sewell MM, Parks CR, Chase MW (1996) Intraspecific chloroplast DNA variation and biogeography of North American *Liriodendron* L. (Magnoliaceae). *Evolution*, **50**, 1147-1154.

Sinclair WT, Morman JD, Ennos RA (1999) The postglacial history of Scots pine (*Pinus sylvestris* L.) in Western Europe: evidence from mitochondrial DNA variation. *Molecular Ecology*, **8**, 83-88.

Soltis DE, Soltis PS (1989) *Isozymes in plant biology*. Portland OR, Dioscorides Press.

Soltis DE, Soltis PS, Ness BD (1990) Maternal inheritance of the chloroplast genome in *Heuchera* and *Tolmiea* (Saxifragaceae). *Journal of Heredity*, **81**, 168-170.

Soltis PS, Gitzendanner MA (1999) Molecular systematics and the conservation of rare species. *Conservation Biology*, **13**, 471-483.

Sork VL, Stowe KA, Hochwender C (1993) Evidence for local adaptation in closely adjacent subpopulations of northern red oak (*Quercus rubra* L.) expressed as resistances to leaf herbivores. *American Naturalist*, **142**, 928-936.

SPSS Inc. (2001) *SPSS base 11.0 User's Guide*, SPSS Inc. Chicago, IL.

Squirrell J, Wolff K (2001) Isolation of polymorphic microsatellites in *Plantago major* and *P. intermedia*. *Molecular Ecology Notes*, **1**, 179-181.

Squirrell J, Hollingsworth PM, Bateman RM, Dickson JH, Light MHS, MacConaill M, Tebbitt MC (2001) Partitioning and diversity of nuclear and organelle markers in native and introduced populations of *Epipactis helleborine* (Orchidaceae). *American Journal of Botany*, **88**, 1409-1418.

Stewart JR, and Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology and Evolution*, **16**, 608-613.

Stewart JR (2003) Comment on "Buffered tree population changes in a quaternary refugium: Evolutionary implications. *Science*, **299**, 825a.

Stilwell KL, Wilbur HM, Werth CR, Taylor DR (2003) Heterozygote advantage in the American chestnut, *Castanea dentate* (Fagaceae). *American Journal of Botany*, **90**, 207-213.

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

Tollefsrud MM, Bachmann K, Jakobsen KS, Brochmann C, (1998) Glacial survival does not matter - II: RAPD phylogeography of Nordic *Saxifraga cespitosa*. *Molecular Ecology*, **7**, 1217-1232.

Tremblay NO, Schoen DJ, (1999) Molecular phylogeography of *Dryas integrifolia*: Glacial refugia and postglacial recolonization. *Molecular Ecology*, **8**,1187-1198.

Tutin G, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA (1968) *Flora Europaea Volume 2: Rosaceae to Um*belliferae. Cambridge University Press, Cambridge.

Tzedakis PC, Lawson IT, Frogley MR, Hewitt GM, Preece RC (2002) Buffered tree population changes in a quaternary refugium: Evolutionary implications. *Science*, **297**, 2044-2047.

Tzedakis PC, Lawson IT, Frogley MR, Hewitt GM, Preece RC (2003) Response to comment on "Buffered tree population chances in a quaternary refugium: Evolutionary implications". *Science*, **299**, 825b.

Vos P, Hogers R, Bleeker M, Reijans M, van der Lee T, Hornes M, Frijters A, Pot J,

Peleman J, Kuiper M, Zabeau M (1995) A new technique for DNA-fingerprinting. *Nucleic Acids Research*, **23**, 4407-4414.

Wang T, Li Y, Shi Y, Reboud X, Darmency H, Gressel J (2004) Low frequency transmission of a plastid-encoded trait in *Setaria italica*. *Theoretical and Applied Genetics*, **108**, 315-320

Warncke E, Terndrup U, Michelsen V, Erhardt A (1993) Flower visitors to *Saxifraga hirculus* in Switzerland and Denmark, a comparative study. *Botanica Helvetica*, **103**, 141-147.

Weir and Cockerham (1984) Estimating F-statistics for analysis of population structure. *Evolution*, **38**, 1353-1370.

Welch D. 2002. The establishment of recovery sites for *Saxifraga hirculus* in NE Scotland. *Botanical Journal of Scotland*, **54**, 75-88.

West RG (1970) Pollen zones in the Pleistocene of Great Britian and their correlation. *New Phytologist*, **69**, 1179-1183.

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.

Wolfe KH, Li W, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. *Proceedings of the Natural Academy of Sciences USA*, **84**, 9054-9058.

Wolfe KH, Sharp PM, Li WS (1989) Rates of synonymous substitution in plant nuclear genes. *Journal of Molecular Evolution*, **29**, 208-211.

Wright S (1978) *Evolution and the genetics of populations vol 4 variability within and among natural populations*. University of Chicago Press, Chicago.

Appendix 1. Haplotype chart showing differences among haplotypes from blind and targeted RFLPs

Haplotype	atpB-rbcL cut with Hinc II	atpB-rbcL cut with Hinf I 1	atpB-rbcL cut with Hinf I ?	trnL C-F Hinf I 1	trnL C-F Hinf I 2	trnL C-F Hinf I 3	trnL C-F Hinf I 4	trnL C-D Hinf I indel 5	trnL C-D Tru I 1	trnL C-D Tru I 2	trnL C-D Tru I 3	trnL C-D Tru I 4	trnK2-Qr cut with Msp I	trnK2-Qr cut with Rsa I 1	trnK2-Qr cut with Rsa I 2	trnK2-Qr cut with Rsa I 3	trnK2-Qr cut with Rsa I 4	F-Vr 1	trnF-Vr cut with Rsa I N184	trnC-D cut with Rsa I 1	trnC-D cut with Rsa I 2	trnD-T cut with Mse I	trnD-T ind	trnS-G cut with Rsa I 1	trnS-G cut with Rsa I 2	trnS-G cut with Dra I
1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	1
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1
3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	?	0	0	0
4	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
5	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
6	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
7	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	?	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	1	2
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	1
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	2
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	1
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	3	0	0	0	0	0	1	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	1	0	0	0	0	0	0
19	0	0	0	0	0	1	0	0	0	0	?	1	0	0	0	0	3	0	0	0	0	1	?	0	1	0
20	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
21	0	0	0	0	0	1	0	0	0	0	?	1	0	0	0	0	3	0	5	0	0	1	?	0	1	0
22	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
23	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	3	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	4	0	0	0	0	0	1	0
26	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0
27	0	0	0	0	0	1	0	0	0	0	?	1	0	0	0	0	0	0	5	0	0	1	?	0	1	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
29	0	0	0	0	0	1	0	0	0	0	?	1	0	0	0	0	0	0	0	0	0	1	?	0	1	0
30	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	1
31	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
32	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
33	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	3	1	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	1	0	0	0	0	1	0
35	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	3	1	0	0	0	3	1	0
36	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
37	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	5	0	0	0	0	3	1	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0
39	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1
40	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0

Appendix 2

Haplotype chart for European populations for RFLPs.

	N	atpb-rbcl Hinc II c/a subs	atpb-rbcl Hinf I indel	trnl C-F cut with Hinf I	trnS-G Mse I t/a subs bp235 nexus	trnS-G Rsa I indel	trn C-D cut with Rsa I	trnC-D cut with Rsa I lower band	trnK2-Qr Msp I	trnK2-Qr Rsa I	trnD-T cut with Mse I	trnFm-Vr cut with Rsa I	n	haplo-type
CNG	20	A s3	B s3	B s1	B s1	B s1	A r	B....r1	B r	A r	B r	C r	20	1
SB	20	B s1	B s1	C sr	B s1	B s1	B r	B....r1	B r	B r	B r	A r	20	2
BC	18	B s1	B s1	C sr	B s1	B s1	B r	B....r1	B r	B r	B r	A r	18	2
SF	19	B s1	B s1	C s1	B s1	B s1	B r	B....r1	B r	B r	B r	A r	19	2
YM	19	B s1	B s1	B sr	A s1	B s1	B r	B....r1	B r	B r	C r	B r	19	3
SGH	20	B s1	B s1	B sr	A s1	B s1	B r	B....r1	B r	B r	C r	B r	20	3
KOG	16	B s1	B s1	C s1	B s1	B s1	B r	B....r1	B r	B r	B r	A r	16	2
GSF	13	B s2	B s2	C sr	B s1	B s1	B r	B....r1	B r	B r	B r	A r	13	2
SH	20	B s6	A s6	B s1	B s4	B s4	B r	B....r1	A r	B r	B r	B r	18	4
		B s2	B s2	B sr	A s2	A s2	B r	B....r1	B r	B r	B r	B r	2	5
O	8	B s3	B s3	A s1	B s1	B s1	B r	B....r1	B r	B r	A r	B r	8	6
D2-4	10	B s2	B s2	B s1	A s1	B s1	B r	B....r1	B r	B r	C r	B r	10	3
D5-7	10	B s1	B s1	A s2	B s1	B s1	B r	B....r1	B r	B r	A r	B r	10	6
SW	20	B s1	B s1	B sr	A s1	B s1	B r	B....r1	B r	B r	C r	B r	20	3
NW	15	B s1	B s1	B s1	A s1	B s1	B r	B....r1	B r	B r	C r	B r	15	3
N	15	B s1	B s1	B sr	A s1	B s1	B r	B....r1	B r	B r	C r	B r	15	3
S1	16	B s1	B s1	B sr	A s1	B s1	B r	B....r1	B r	B r	C r	B r	16	3
S2	20	B s1	B s1	B s1	A s1	B s1	B r	B....r1	B r	B r	C r	B r	19	3
		B sr	B sr	B sr	A s1	B s1	B r	B....r1	B r	B r	C r	C r	1	7

s=sequence data from this population (the number indicates the number of individuals sequenced)
r=pcr-rflp data (no sequence knowledge)
sr = pcr-rflp data (based on sequence from other populations)

Haplotype chart for European populations for sequenced regions.

	N	atpb-rbcl g/t subs bp208 A=G B=T	atpb-rbcl t/c subs bp271 A=T B=C	atpb-rbcl inde -A	atpb-rbcl bp137 A=A B=C	atpb-rbcl bp285 A= B=G	atpb-rbcl bp327 A=A B=G	trnS-G g/t subs bp481 A=T B=G	trnS-G c/a subs bp351 A=A B=C	trnS-G c/t subs bp190 A=T B=C	trnS-G bp236 A=T B=A	trnS-G bp462 A=T B=G	trnS-G bp543 A=T B=A	trnL indel bp237 A=pres B=abs	trnL indel bp348 A=pres B=abs	trnL indel bp368 A=pres B=abs	trnH-psbA bp36 A=A B=C	trnH-Ps A invers A=ATA B=TAT bp62	haplo-type
CNG	20	B s3	A s3	B s3	B s3	B s3	A s3	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	1
SB	20	A s1	A s1	B s1	B s1	B s1	A s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	B s1	2
BC	18	A s1	A s1	B s1	B s1	B s1	A s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	B s1	2
SF	19	A s1	A s1	B s1	B s1	B s1	A s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	B s1	2
YM	19	A s1	A s1	B s1	B s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	3
SGH	20	A s1	A s1	B s1	B s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	3
KOG	16	A s1	A s1	B s1	B s1	B s1	A s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	B s1	2
GSF	13	A s1	A s1	B s1	B s1	B s1	A s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	B s1	2
SH	20	A s6	B s6	C s6	B s6	B s6	A s6	A s4	B s4	A s4	B s4	B s4	B s1	B s1	B s1	B s1	A s1	B s1	4
		A s2	A s2	A s2	B s2	B s2	A s2	B s2	B s2	B s2	B s2	B s2	B s1	B s1	B s1	B s1	A s1	A s1	5
O	8	A s3	A s3	B s3	B s3	B s3	A s3	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	6
D2-4	10	A s2	A s2	B s2	B s2	B s2	A s2	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	3
D5-7	10	A s1	A s1	B s1	B s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	6
SW	20	A s1	A s1	B s1	B s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	3
NW	15	A s1	A s1	B s1	B s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	3
N	15	A s1	A s1	B s1	B s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	3
S1	16	A s1	A s1	B s1	B s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	3
S2	20	A s1	A s1	B s1	B s1	B s1	A s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	B s1	A s1	A s1	3

APPENDIX 3

Non-technical report for SNH on a possible re-introduction programme of

LATHYRUS JAPONICUS AT ELLIOT LINKS, ANGUS.

Christina Oliver
Royal Botanic Garden Edinburgh

Contents table

Table of Figures

SUMMARY

RE-INTRODUCTION OF *LATHYRUS JAPONICUS* AT ELLIOT LINKS

Background

Lathyrus japonicus is a low growing, long lived perennial herb with a circumpolar distribution. In Britain, it occurs predominantly as an early colonist of shingle beaches though it can also occur on dunes or sand and gravel. It dies back in the winter, and regrows from an extensive root system in spring, and can form clumps several meters wide. *L. japonicus* seems sensitive to disturbance and trampling during the growing season, and has recently disappeared from some of its historical sites in Britain. Elliot links a site where *L. japonicus* has occurred in the past, but for some reason has become extinct.

This report aims to give information on the ecological preferences of *L. japonicus*, assess the habitat suitability at Elliot Links, and give recommendations on seed sourcing and management if the reintroduction were to be carried out.

Main Findings

- Small areas of potentially suitable habitat for *L. japonicus* occur at Elliot Links. Specifically a patch of relatively open shingle at the northern edge of the site, near the outlet of the river, and a narrow band of sand and shingle just above the high water line, stretching approximately 20 meters south from the outlet of the river.

- However Elliot Links has considerable recreational use, and since *L. japonicus* can not tolerate heavy disturbance, reintroduction would probably not be successful unless fencing were used to protect the plants from being unknowingly walked upon.

- The population of *L. japonicus* at Carnoustie is markedly genetically distinct from other populations in Britain, and has low levels of genetic variation when compared to the other populations in Britain.

- Seedlings sourced from Carnoustie show local adaptation to the Scottish climate when grown in Edinburgh in common garden conditions with seedlings sourced from English populations. Carnoustie seedlings grew more vigorously and held their leaves longer into the winter.

- If reintroduction were attempted it could be the best tactic to introduce young plants rather than seeds for several reasons including, issues of seed predation, rareness of occurrence of germination in the field, and the necessity of protection from disturbance.

- If reintroduction were attempted it would probably be the safest option to source the introduction from Carnoustie, because the advantages from local adaptation probably outweigh the risk of inbreeding depression.

INTRODUCTION

Lathyrus japonicus is a low growing, long lived perennial herb with a circumpolar distribution. In Britain, it occurs predominantly as an early colonist of shingle beaches though it can also occur on dunes or sand and gravel. It dies back in the winter, and regrows from an extensive root system in spring, and can form clumps several meters wide. *L. japonicus* seems sensitive to disturbance and trampling during the growing season, and has recently disappeared from some of its historical sites in Britain. *L. japonicus* is pollinated by bumblebees. It does not set seed if pollinators are excluded from an inflorescence (Brightmore and White, 1963), although ample seed set by the single isolated plant at Nigg Bay indicates that *L. japonicus* is self-fertile. The flowers are protandrous[1], but viable pollen is still present at the time the stigma is receptive (Asmussen, 1993). The seeds are large with a hard outer covering and can be dispersed long distances in the ocean. Seeds can retain their viability in the sea for up to 5 years (Brightmore and White, 1963). Drift seeds of *L. japonicus* are reported sometimes in large numbers on the west cost of Ireland, Cornwall, and the Hebrides in Scotland (Nelson, 2000), though plants are rare and transient in these areas.

PHENOLOGY

The onset of growth in *Lathyrus japonicus* in spring is variable, but generally growth begins to show above ground by April (Brightmore and White, 1963). Flowering begins in May, and extends into September. The majority of seeds mature in August and September, and are shed onto the substrate beneath the parent plant when the pods dehisce. Flowering and fruiting densities at a given time can vary considerably between adjacent sites, and are probably dependent on local weather conditions. The above ground growth can remain until December or January. Germination of the seeds takes place in the following April and May (Brightmore and White, 1963).

ECOLOGICAL PREFERENCES

Lathyrus japonicus predominantly occurs in Britain as an early colonist of shingle beaches. In some of the more extensive sites a progression is visible from pure stands of *L. japonicus* near the ocean, to clumps of *L. japonicus* slightly more inland being colonized by other species, and finally to scraggly plants of *L. japonicus* remaining in overgrown

[1] The term protandrous means that the anthers of a particular flower mature and release their pollen before the stigma becomes receptive. This is a mechanism to promote out-crossing, and minimise self-fertilization. But, as in this case, it does not always completely eliminate self-fertilization.

areas towards the back of the beach. The species is not sensitive to frost, so temperature is probably not a limiting factor. Though the adult plants are not sensitive to drought, and indeed require well-drained habitats, seedlings cannot tolerate drying out. *L. japonicus* can occur on sand dunes, although where other vegetation such as marram grass is dense the plants become scraggly and flowering and fruiting is reduced.

FEASIBILITY OF RE-ESTABLISHMENT

Habitat Suitability

The habitat needs of *L. japonicus* seem not too far off what appears to exist in a few patches at Elliot Links. However the area of habitat that seems suitable is small. On the northern edge of the site, near the outlet of the river, there is a small area of bare shingle that appears to be above the high water line and relatively stable. Also, moving south, there is a narrow band of mixed sand and shingle covered with marram grass, immediately above the high water line. *L. japonicus* most commonly occurs on open shingle, as one of the first colonist species, but in sites such as the Southwold Denes (Sussex) and Carnoustie (Angus) it occurs on sandy substrates, with denser vegetation cover. Plants on such sites have a sparser growth habit, and produce fewer flowers and fruits, but they do persist. Interestingly, in Alaska, *L. japonicus* is widespread, and appears to occur typically on sandy coastal areas. So the sand and marram grass present at Elliot links in itself should not be enough to deter establishment.

Climate

It is possible that *Lathyrus japonicus* disappeared from Elliot links due to climate change, and if this is the case, re-introduction could be problematic. *L. japonicus* is not sensitive to frost, either in the above or below ground parts, and is widespread in more Arctic environments such as Alaska, Greenland, and Iceland, but also thrives in the south of England where climates are milder than Scotland, so temperature is probably not a limiting factor at Elliot Links. However *L. japonicus* requires well drained habitats, and if the area where plants occurs becomes waterlogged, either because of a build-up of organic matter, or a series of particularly wet winters, the plants could probably not tolerate this. In the south of England the plants occur predominantly on shingle, which holds very little water; and in more northerly habitats such as Alaska, where it occurs commonly on sand, much of the winter moisture would be tied up as ice and the environment could be very dry. It follows that one could speculate that a gradual change in Scotland towards a

milder and wetter climate could cause *L. japonicus* to require a narrower set of more well drained habitats than it might have previously needed.

Disturbance

Another perhaps more immediately serious problem that *L. japonicus* would face were it to be re-introduced at Elliot links would be human disturbance. In England, the main reason attributed to the recent decline in *L. japonicus* is disturbance. It does not seem to be able to tolerate being trodden on, particularly early in the growing season. This has been demonstrated at many sites in the south by the setting up of enclosures. And at some sites the plant now remains only inside the enclosures. Unfortunately the ecological requirements of *L. japonicus* seem to match closely with the holiday and recreational requirements of humans! Elliot Links appears to be a popular place for people to go walking, and to walk their dogs. Heavy traffic is apparent from the state of the footpaths. Even during my site visit, I observed nearly a dozen people arrive, most with several dogs. And a footpath seems to stretch along the point just above the high water line where the vegetation is sparsest and the environment most promising for *L. japonicus*. Realistically, given the small size of the suitable habitat, and the heavy recreational use of the site, I do not think *L. japonicus* would have a chance of becoming re-established unless fences were put in place to prevent the plants from being unknowingly walked on. This brings up a question of priorities, and what the local people and government want from the site. If the people who use the site were interested in the re-introduction project, it could be a simple enough procedure to put up small fences around a few individual plants or groups of plants in areas of suitable habitat, and include information signs explain the enclosures. But without the willingness to set up enclosures, perhaps it would be better to seek a less heavily used site for potential reintroduction.

TACTICS IF RE-ESTABLISHMENT IS CARRIED OUT

Young Plants vs. seed

There are several reasons that make the introduction of young plants rather than seed seem like the best option. Firstly, a large clump of *L. japonicus* can produce hundreds of pods, and thousands of seeds in a season, which are dropped onto the substrate under the parent plant in the autumn. Recent molecular research using microsatellite markers indicates that a single visual clump tends to comprise of a single

genetic individual. This means that the thousands of seeds dropped under the parent plants each year have not germinated, and that germination is in fact a rare occurrence. When I started seeds in cultivation, I chipped the outer shell with a sharp knife, and soaked them in water until they sank. This took 2-3 days, and in that time the seeds doubled in size. 80 –90 percent germination then occurred in about a week. The seeds have hard outer coverings, and it is possible that for germination to occur, a very particular series of events need to happen, such as a sea journey which would allow for the absorption of water, or abrasion of the testa by the action of waves, or ideal climatic conditions that allow for slow breakdown of this outer covering, and the absorption of water without the occurrence of rot. This would help to explain why germination is a rare occurrence in the natural habitat, while 80 to 100 percent germination occurs in cultivation. It is reported that untreated seeds shed in the autumn germinate in the following spring due to breakdown of the testa by soil organisms when sown in potting soil (Brightmore and White, 1963), and the same article reports that seeds sown in pure shingle failed to germinate unless they were chipped.

It might appear to be the least labour intensive method to collect seed and broadcast it over potential habitats, but I could see this easily resulting in no recruitment of seedlings in most years. Most likely the introduction of seed would have to be repeated for years, in spite of seeds having a viability of several years (Brightmore and White, 1963), in order to have enough seeds present when the climate conditions were right for establishment. And if seedlings were to germinate, there would be a period of time when they would be very vulnerable to drying out which older plants with established root systems would not be effected by. Finally, if seedlings did germinate they could be quite difficult to locate and protect with fencing before they were trodden upon!

A more sure technique, though perhaps more labour intensive in the beginning, would be to set out one or two year old seedlings in the early spring, perhaps even as few as half a dozen, protect them with fencing, at least until they become established, and wait for the year or two until they reach flowering size and start producing crops of seed. At this point, if the habitat is satisfactory, there should be ample time to wait for the chance occurrence of years with suitable conditions for germination, while the plants themselves broadcast the seeds each year.

Seed Sourcing

It is important when re-introducing a plant to consider how to source the introduction material. The introduced plants should be as ecologically well adapted to the site as possible, moving pests and disease with the introduced plants should be avoided, and if genetic information is available it should be used to guide the decision in a way to minimize the potential of inbreeding or outbreeding depression.

Genetic issues

A morphological difference in leaf width in *Lathyrus japonicus* has been noted in the past, with the plants from Carnoustie having narrow leaves, as do the plants from Shetland, and plants from England having wider leaves. This morphological difference remained in seedlings cultivated in common conditions in Edinburgh in 2001, and is mirrored by the genetic differences described below. Interestingly, the single plant present at Nigg Bay near Aberdeen has both the wider leaves and similar genetic makeup to the English plants.

Genetic research using microsatalite markers indicates that the population of *L. japonicus* at Carnoustie is genetically distinct from the other populations of *L. japonicus* in Britain, and has a large number of alleles[2] not found in other British populations.

The Carnoustie population also has a relatively low number of alleles. The average number of Alleles per locus among the English populations studied is 3.4, while in Carnoustie the average is only 2.1. So Carnoustie seems to have at least some evidence for comparatively low genetic variation. However it is not known whether this impacts the health of the population, and in light of the ecological differences discussed below, I would guess the low genetic variation present at Carnoustie is not a serious concern.

[2] The term allele refers to a copy of a gene. Most genes, also called loci, used in conservation research are neutral – that is they give no advantage or disadvantage to the organisms that possess them, and do not undergo selection. Each individual has two alleles at each locus, one inherited from each parent. In a population many different alleles can exist for a single locus, with each individual possessing two copies, either identical or different. An individual with different alleles at a particular locus is referred to as heterozygous for that locus, while an individual with identical alleles at a particular locus is referred to as a homozygous. When selfing or inbreeding occurs, homozygosity increases, and genetic variation declines. It is often speculated that higher heterozygosity results in higher fitness in a population, although the relationship between neutral genetic markers and fitness is not clear-cut.

Ecological considerations

In January of 2002, 40 seeds sourced from each of 11 populations of *Lathyrus japonicus* throughout Britain were germinated at the RBGE Edinburgh. They were cultivated in an unheated greenhouse until spring, and then grown outside. At the end of the summer, it was clearly observable that the plants sourced from Carnoustie had produced more vegetative growth than the seedlings sourced from any other populations. Also, the seedlings from Carnoustie held their leaves longer into the winter than the other seedlings. This suggests that a significant amount of local adaptation for the particular climate in Scotland is present in the Carnoustie seedlings. In simplest terms, the plants sourced from Carnoustie grew better in Edinburgh than the plants from England, or the plant from Nigg Bay (which is genetically most similar to the English plants). This has important implications for any reintroduction that might take place. Although the Carnoustie plants might have low genetic variation, they appear to be locally adapted to the Scottish climate, and grow more vigorously in Edinburgh than plants from England. Although the climate in Edinburgh is not the same as that at Elliot Links, it is definitely closer to it than the Southern English climate. So, although it might seem desirable to introduce seeds from a wide range of sites in order to increase genetic variation, and potential adaptability, in this case it would probably not be the best tactic. Most likely English plants would not compete in a situation where mixed sourcing was carried out, but there is a danger that if they did survive to flowering that they could introduce less well adapted genetic material to the local type plants, and result in an overall lowering of fitness.

Seed predators

English populations of *L. japonicus* have high levels of predation by beetle larva, which develop inside the seeds. A parasitoid wasp also occurs, which presumably feeds on the beetle larvae. Seeds which contain beetle larva are mostly hollow at the time the beetle has matured, although seeds that contain wasps are usually only damaged somewhat, and some still germinate. Levels of seed predation were measured in late summer of 2001, and the results are given in Table X.1. These numbers are probably conservative because an effort was made to collect groups of pods that were not heavily predated since the original objective was to obtain weights of undamaged seeds. So it is likely that levels of predation are actually higher than this table reflects. These levels of predation, where they are highest, must have a considerable impact on the reproduction of

L. japonicus, particularly at Shingle Street where only 16 percent of the seed crop was undamaged.

Table X.1. Numbers of seeds from 11 L. japonicus sites predated by beetle larva, including seeds found to contain parasitoid wasps, which presumably fed on the beetle larva. Data taken from seeds collected in Sept 2001.

	Seeds containing wasps	Seeds predated by beetle	Total damaged seeds	Total undamaged seeds	Total seeds examined	Proportion of undamaged seeds
Chesil Beach	49	223	272	512	784	0.653061
Rye Harbour	121	172	293	239	532	0.449248
Dungeness	1	0	1	508	509	0.998035
Deal	145	100	245	522	767	0.680574
Felixstowe	12	0	12	382	394	0.969543
Shingle Street	261	142	403	81	484	0.167355
Southwold Denes	21	0	21	359	380	0.944737
Kessingland	2	1	3	735	738	0.995935
Pakefield Beach	59	4	63	601	664	0.90512
Carnoustie	0	0	0	292	292	1
Nigg Bay	0	0	0	104	104	1

It is notable that neither the beetle nor the wasp has been found in Scotland. Perhaps there is a climate limitation, but perhaps it is a chance distribution, and care should be taken not to introduce this seed predator to the Scottish populations of *L. japonicus*. This is another reason why it might be better to introduce young plants rather than a large number of seeds from any site. And introducing seed from any English population should be avoided, as the beetles would stand a high chance of being introduced with the seed.

MANAGEMENT

The main management concern if *L. japonicus* were to be reintroduced at Elliot links would be how to prevent the plant from being trodden on by recreational users of the site. This would probably involve fencing, as it has been demonstrated in various sites in England that this species cannot tolerate heavy traffic. Without this bit of management, reintroduction would probably not be worth attempting.

Another potential area where management might be considered is in the removal of marram grass. The marram grass present at Elliot Links is very similar to the habitat at Southwold Denes, where *L. japonicus* persists, but has lowered flowering and seed production. Any procedure that would thin the marram would give *L. japonicus* a better chance of thriving, however this would only be worth while in the narrow band of already potentially suitable habitat just above the high water line as farther inland, where organic matter has built up, probably no amount of clearing could make the habitat suitable. However thinning the marram could be highly labour intensive, and would only probably be a temporary solution. And since the problem posed by the marram grass is perhaps not critically serious, and *L. japonicus* does persist in sites with a similar type and level of vegetation, it might not be worth the time and money.

REFERENCES

Asmussen C 1993. Pollination biology of the Sea Pea, *Lathyrus japonicus*: floral characters and activity and flight patterns of bumblebees. *Flora* 188. 227 – 23

Brightmore E, White PHF. 1963. *Lathyrus japonicus* Willd. Biological Flora of the British Isles, No. 94. *Journal of Ecology* **51**: 795-801.

Nelson CE. 2000. *Sea beans and nickar nuts.* London: BSBI.