ECOLOGICAL GENETICS OF THE MOSS

PHYSCOMITRELLA PATENS

Erica Jane Hooper

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty of Biological Sciences

September 2008

Acknowledgements

Firstly I would like to thank my supervisors Andy Cuming and Bill Kunin, for all their support, guidance and encouragement. This has been immensely valuable to me whilst conducting and completing this project, and without which I would have been unable to finish. I am also extremely grateful to Yasuko Kamisugi, for all her help and advice with laboratory work, particularly the AFLP analysis. Thanks also go to Claire Smith, for her help with plant tissue culture.

I am very grateful to all the bryologists who have assisted in finding bryophyte populations, without which I would have been completely unable to undertake this research. Particular thanks go to Sam Bosanquet, David Chamberlain, Gorden Rothero, John Blackburn, Hans Blom and Rachel O'Hara for accompanying me on my hunt for populations. I am also very grateful to Chris Preston, David Long, David Holyoak, Mark Pool, Paul King, Tom Blockeel, Ron Porley, Angela Newton and Mark Lawley, for supplying details of populations. Thanks go to all the landowners and wardens who gave their permission for plant collections to take place.

Many thanks go to Jeff Duckett at Queen Mary, University of London, for his help with identifying *Physcomitrella readeri*, without whose expertise the moss would remain a mystery. Thanks also go to both Jeff Duckett and Sylvia Pressel, for supplying beautiful SEM images of the *P. readeri* plants.

I would like to thank my colleagues at the University of Leeds, for providing both their friendship and support, along with a stimulating environment in which to work. Lastly I will thank my family for their enthusiasm and endless encouragement, and for accompanying me on cold, wet, field excursions. I particularly thank my partner

for all his crucial support and understanding, and for spotting the odd sporophyte when I was about to give up.

Abstract

Molecular genetic studies using the model bryophyte *Physcomitrella patens* have advanced the body of knowledge surrounding plant functional genetics and molecular biology, yet very little is known about the ecology and population genetics of this species. Although the bryophytes are the second largest group of plants, there is little information regarding the population genetics of bryophytes in general. To address these issues I have conducted the first study into the population genetics of *P. patens*, where plants from eight populations in Britain have been collected. Sampling within these populations was conducted according to a hierarchical scale, so as to assess not only the level of genetic variation within populations, but how this is structured spatially. Analysis of the plants collected was conducted using amplified fragment length polymorphism (AFLP) analysis. No spatial genetic structure was found within populations of *P. patens*, and it is hypothesised that the nature of the ephemeral aquatic habitats that *P. patens* occupies may account for this finding. In this thesis a novel method for studying the mating systems operating within bryophyte populations has been proposed, which exploits the dominant haploid stage of the bryophyte life cycle. This methodology has been applied to natural populations of *P. patens*, and evidence of mixed mating has been observed. Bryophytes are often overlooked or under-recorded in their natural environment, and distribution data within Great Britain is likely to be inaccurate for a large number of species. This issue is highlighted in this thesis, as a bryophyte species new to Europe has been discovered.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations Used

AFLP	Amplified fragment length polymorphism
AMOVA	Analysis of molecular variance
CI	Confidence interval
C.I.	Consistency index
bp	Base pairs
cDNA	Complementary DNA
EST	Expressed sequence tag
FAM	Blue fluorescent dye (Applied Biosystems)
GPS	Geographical positioning system
HEX	Green fluorescent dye (Applied Biosystems)
IBD	Isolation by distance
ITS	Internal transcribed spacer
NED	Yellow fluorescent dye (Applied Biosystems)
PCR	Polymerase chain reaction
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic DNA
R.I.	Retention index
SCALP	Sequence characterised amplified length polymorphism
SEM	Scanning electron microscope

Chapter 1 General Introduction

1.1 Introduction to project

The bryophytes are a diverse and ecologically interesting group of plants, comprising 25,000 species occupying a wide variety of habitats (During & van Tooren, 1987). They are the second largest group of land plants in terms of species number and in many ecosystems they are the main primary producers (During & van Tooren, 1987; Mishler, 1988). Despite the diversity and ecological importance of this group they are often overlooked in population studies, and little information regarding the evolutionary genetics of bryophytes has been published (Innes, 1990). Bryophytes are often missed or incorrectly recorded in their natural environment, so for many groups of mosses little information is available on their distribution, ecology and population biology. During the course of this project, a moss species new to Europe has been found. This species has been verified as *Physcomitrella readeri* (C. Müell.) Stone & Scott, a species previously only recorded in the U.S.A, Japan, Australia and New Zealand.

Knowledge surrounding the molecular genetics of bryophytes has increased dramatically in recent years, with the development of the moss *Physcomitrella patens* (Hedw.) Bruch & Schimp. as a model organism. The *Physcomitrella* Genome Programme has generated over 300,000 expressed sequence tags (ESTs) and has recently published the first draft genome sequence assembly (Rensing *et al.*, 2008) Other model organisms and their relatives, such as *Arabidopsis thaliana*, *Drosophila melanogaster* and *Escherichia coli*, have increasingly been utilised in ecological and evolutionary studies to investigate aspects such as adaptation, genetic diversity and

reproductive strategies in natural populations (as reviewed by Mitchell-Olds 2001; Shimizu, 2002; Travis, 2006).

One area of current focus in plant population ecology is that of the spatial genetic structure within plant species, and how this relates to features of the habitat they occupy. This field has received much attention in vascular plants throughout the last 20 years, with many studies focussing explicitly on how genetic variation is partitioned spatially within a species. There has been an increasing trend towards using mapped individuals as the operational taxonomic unit in population genetic studies, rather than randomly sampling within populations, which are often arbitrarily and inaccurately defined (see Manel *et al.*, 2003 for a review). However, this is an approach that has only very recently been applied to bryophyte populations (e.g. Snäll *et al.*, 2004), and little work into population structure in general has been undertaken outside commonly studied moss genera such as *Sphagnum* (e.g. Gunnarsson *et al.*, 2005; Stenøien & Såstad, 1999; Thingsgaard, 2001). Additionally, little has been done to clarify reproductive processes in bryophyte populations, yet this aspect of organismal life history has one of the most significant impacts on levels of diversity, genetic structure and population behaviour (Shaw, 2000).

Model organisms, such as *P. patens* are a useful starting point when addressing questions surrounding the genetic processes in evolutionary ecology as a wealth of information and tools is available for these organisms. This investigation aims to be the first study into the population genetics of *P. patens*, thus allowing the detailed molecular information available for this species to be related to its ecology.

1.2 Genetic analysis of population structure and evolutionary processes

1.2.1 Fine-scale spatial structure

Advances in the accuracy, resolution and scale of the molecular tools used for population genetic studies have allowed utilisation of molecular markers in fine-scale population studies (Manel *et al.*, 2003). When using these tools to investigate the structure within populations at differing spatial scales, the distance over which genetic variation within populations is distributed and the levels of gene flow within a population can both be determined. Plant populations are often highly spatially structured, which can be due to a number of reasons (reviewed in Ennos, 2001). Plants are generally static, and genetic dispersal is by means of pollen and seeds, or in the case of bryophytes by spores, sperm and vegetative fragments. Thus gene flow can be limited and genetic isolation by distance (IBD) occurs within populations (Vekemans & Hardy, 2004; Wright, 1943). Many plants also exhibit vegetative dispersal, which will lead to a patchy distribution of genotypes. In addition, spatial structure within populations could be observed due to sampling events upon colonisation of new areas or regeneration of populations (Ennos, 2001).

How gene flow operates within species greatly affects not only the spatial structure of populations, but also the ability of a population to adapt to local conditions, with a subsequent possibility of independent evolution of those populations resulting in speciation (Slatkin, 1985). When spatially dependent processes, such as IBD, have been identified in plant populations, they can be correlated with landscape and population features, such as population size or fragmentation (Manel *et al.*, 2003). This information can be of use in conservation biology to predict the effect that

changes in habitat could have on population genetic processes (Escudero *et al.*, 2003; Storfer *et al.*, 2007).

There are a number of different methods for assessing the spatial genetic structure within plant populations (see reviews in Escudero *et al.*, 2003; Manel *et al.*, 2003; and Storfer *et al.*, 2007). Commonly used population analyses, such as Analysis of Molecular Variance, or AMOVA (Excoffier *et al.*, 1992) and estimates of genetic differentiation and distance between populations, have long been used to infer the hierarchical structure of genetic diversity within and between populations (Escudero *et al.*, 2003). The discrete mapping of sampled individuals and subsequent genetic profiling, allows far more sophisticated statistical analysis to be conducted. Spatial autocorrelation, that is the association of one variable at one location with the same variable at another location (Schabenberger & Gotway, 2005), can therefore be examined. A widely used method of investigating spatial autocorrelation is a Mantel test (Mantel, 1967). Here spatial clustering of variables is examined by the construction of two distance matrices, one with genetic distance between individuals and one with geographic distance between individuals. The correlation between these two matrices is examined, and a randomisation procedure is used to determine if clustering based on genetic similarity can be explained by spatial position (Escudero *et al.*, 2003; Schabenberger & Gotway, 2005). The disadvantage of this approach is that spatial genetic autocorrelation does not necessarily follow a linear pattern (Escudero *et al.*, 2003). An enhancement of a Mantel test is to partition measurements of geographic distance into distance classes. Spatial autocorrelation analysis can then be conducted based on these classes rather than linear distance, with data often presented as a correlogram (Escudero *et al.*, 2003).

1.2.2 Mating systems within populations

The mating patterns in plant species, and specifically the mating systems operating within populations, are a significant feature of a species' ecology. The mating system of any species or population is not static, but can vary in response to changing environments and evolutionary forces. Changes in the mating system can have important consequences for the genetic structure and behaviour of that population (Barrett & Eckert, 1990). Although some species, such as dioecious species with separate male and female plants, are obligate out-crossers, many plant species exhibit levels of inbreeding. Classic population genetic theory dictates, particularly in animal populations, that out-crossing is evolutionarily favourable to inbreeding. This is because inbreeding may result in inbreeding depression, which is measured by the fitness of selfed offspring compared to those that are the product of out-crossing (Ritland, 1990). Inbred individuals are thought to be reduced in fitness compared to out-crossed individuals due to two reasons; that heterozygosity is superior to homozygosity at any one locus, and that inbreeding may result in an accumulation of deleterious alleles in a population (Charlesworth & Charlesworth, 1987). However, due to the propensity of plants for self-fertilization, predominant inbreeding has been found to be a stable state in many plant populations (Lande & Schemske, 1985). Studies into the mating systems operating within plant populations have found that mixed mating, or the presence of both inbreeding and out-crossing in a population, also frequently occurs (Barrett & Eckert, 1990; Goodwillie *et al.*, 2005; Vogler & Kalisz, 2001; Walter, 1986). Despite the importance of mating systems to the ecology and evolution of a species, knowledge is limited with regards to within-population variation of mating systems in plants (Cruzan, 1998).

1.2.2.1 The use of AFLPs in population genetic analysis

The use of neutral molecular markers in plant diversity studies has greatly enhanced the range of genetic studies that can be conducted (Cruzan, 1998), and amplified fragment length polymorphism (AFLP) analysis is increasingly being favoured as the marker of choice. The technique involves restriction digestion of DNA followed by selective amplification of the resulting fragments (see Figure 1.1), which can then be viewed on a gel or automated sequencer (Blears *et al.*, 1998; Ritland & Ritland, 2000; Vos *et al.*, 1995). The main advantage of using AFLPs in population studies is that the number of loci scored and polymorphisms identified in each reaction is high (Blears *et al.*, 1998). This is important when studying diversity within populations as the genetic variation between individuals may be low. AFLP is also favoured as it is a robust and reliable genetic marker, which is highly reproducible due to the specificity of restriction enzymes. It is more reliable than similar techniques such as RAPD, which is more sensitive to the reaction conditions (Blears *et al.*, 1998). AFLPs provide dominant markers, and in diploid tissue heterozygotes cannot be distinguished. Ideally in population genetic studies co-dominant markers such as microsatellites would be used to study genetic structure and the inheritance of alleles in most species (Ennos, 2001). However, when using microsatellites far fewer loci can be scrutinised, and in addition they are expensive and time consuming to develop. In species where the dominant stage of the life cycle is haploid, such as in mosses, the disadvantages of using a dominant marker are negated, and AFLPs provide an ideal marker system.

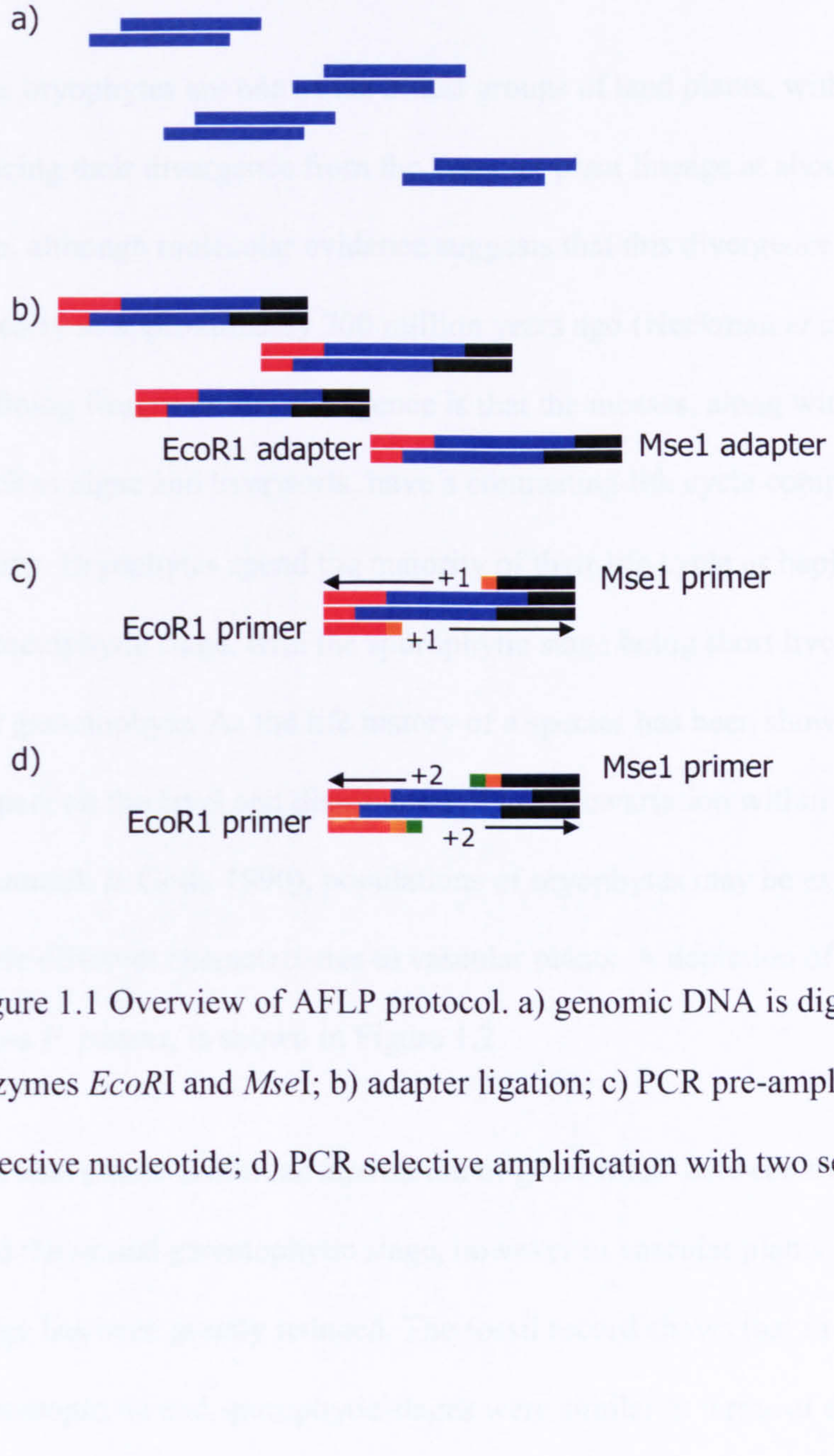


Figure 1.1 Overview of AFLP protocol. a) genomic DNA is digested with restriction enzymes *EcoR*I and *Mse*I; b) adapter ligation; c) PCR pre-amplification with one selective nucleotide; d) PCR selective amplification with two selective nucleotides.

1.3 Bryophyte biology and life history

The bryophytes are one of the oldest groups of land plants, with the fossil record placing their divergence from the vascular plant lineage at about 450 million years ago, although molecular evidence suggests that this divergence may have occurred as early as approximately 700 million years ago (Heckman *et al.*, 2001). The defining feature of this divergence is that the mosses, along with other plant groups such as algae and liverworts, have a contrasting life cycle compared to vascular plants. Bryophytes spend the majority of their life cycle as haploids in the gametophytic stage, with the sporophytic stage being short lived and dependent on the gametophyte. As the life history of a species has been shown to have a profound impact on the level and distribution of genetic variation within natural populations (Hamrick & Godt, 1990), populations of bryophytes may be expected to exhibit quite different characteristics to vascular plants. A depiction of this life cycle in the moss *P. patens*, is shown in Figure 1.2.

All land plants exhibit an alternation of generations between the sporophytic stage and the sexual gametophytic stage, however in vascular plants the gametophytic stage has been greatly reduced. The fossil record shows that in early land plants the gametophytic and sporophytic stages were similar in terms of complexity, with the reduction of the gametophyte in vascular plants developing after the transition to land (Kenrick & Crane, 1997; Taylor *et al.*, 2005). The bryophytes represent the earliest group of plants to diverge in the land plant phylogeny, so they therefore provide an opportunity to study the evolutionary processes associated with the conquest of land (Rensing *et al.*, 2008; Renzaglia *et al.*, 2007).

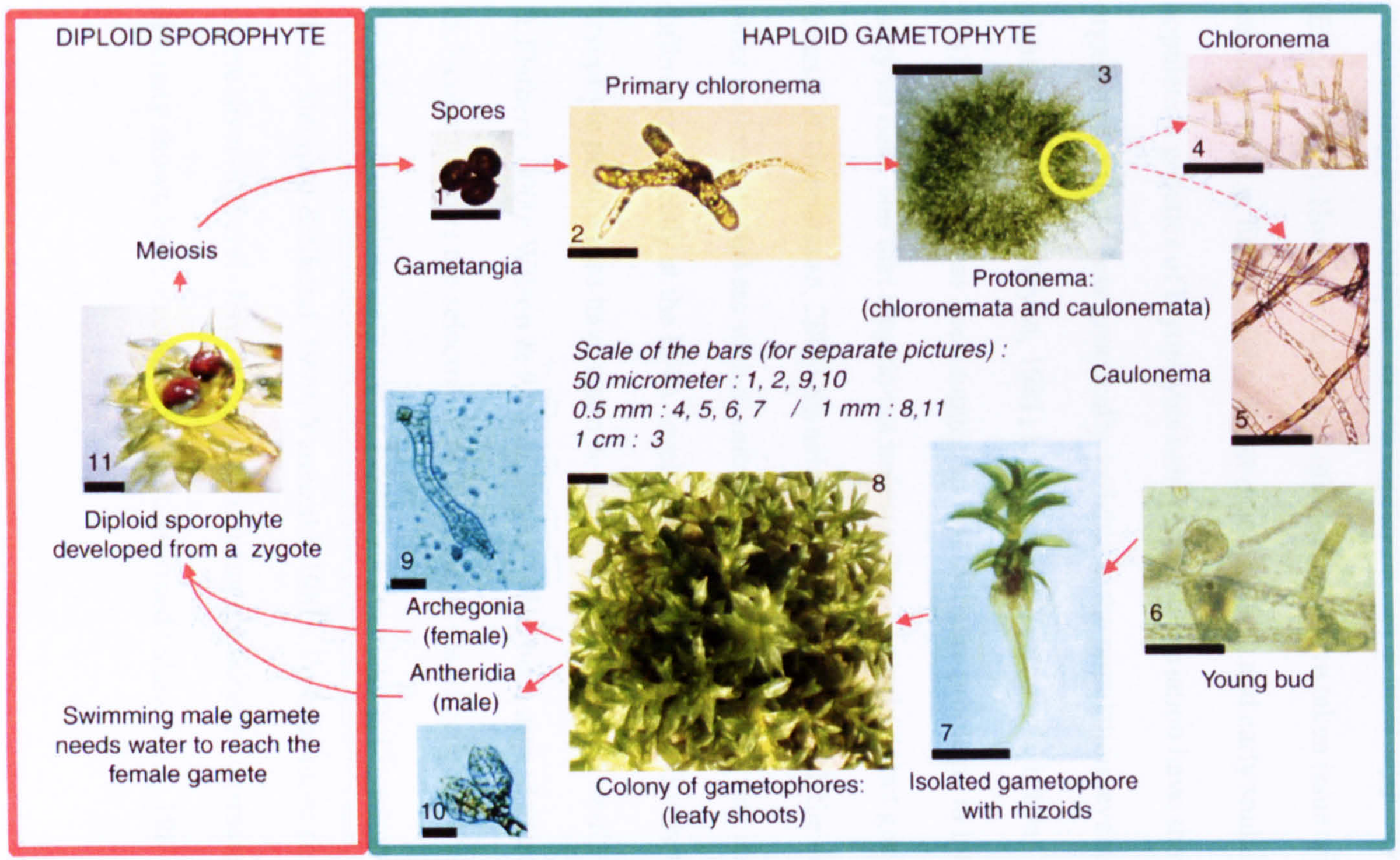


Figure 1.2 Diagram depicting the life cycle of the moss *P. patens* Taken from Schaefer & Zrÿd (2001).

1.3.1 Genetic diversity in bryophyte populations

The consequences of a dominant haploid life stage in bryophytes has long been thought to have significant implications for levels of diversity in mosses, and until recently there was a strongly held view that bryophytes lacked the "evolutionary potential" (Wyatt *et al.*, 1989b) of vascular plants. There was general acceptance that diploid organisms possess greater genetic variation, and thus a greater evolutionary potential than haploid organisms due to heterozygosity in diploids (Ennos, 1990). However, a number of reviewers have taken issue with this assumption (e.g. Ennos, 1990; Wyatt *et al.*, 1989b), and early studies of the population genetics of bryophytes using isozyme variation have shown haploid organisms to be no less genetically variable at the population level than vascular plants (Cummins & Wyatt, 1981; Innes, 1990; Wyatt *et al.*, 1989a; Yamazaki, 1981). However, it has been suggested that there are problems in interpreting isozyme data, and that they do not necessarily accurately detect genetic variation (Hassel & Gunnarsson, 2003; Stenøien & Såstad, 1999). With the development of other molecular markers such as microsatellites and RAPDs, that record genetic differences directly at the DNA level, high levels of genetic diversity within bryophyte populations have again been found (e.g. Skotnicki *et al.*, 1999; Stenøien & Flatberg, 2000; Wilson & Provan, 2003). It has been argued that as most molecular markers are selectively neutral, the levels of genetic variability found do not reflect the evolutionary rate of bryophyte populations (Stenøien & Flatberg, 2000; Stenøien & Såstad, 1999; Yamazaki, 1981). Perhaps more convincingly, where diversity levels have been studied using quantitative characters, the levels of diversity shown by molecular data are confirmed (Shaw *et al.*, 1987).

1.3.2 Bryophyte mating systems

In bryophytes there is a wide range of mating systems operating within species, owing to the unique life history of these plants (Shaw, 2000). Bryophyte mating systems fall into two general categories; dioicous species with separate male and female gametophytes, and monoicous species with both sexes on one gametophyte. Monoicous bryophytes can also be divided into a number of different forms; with archegonia and antheridia either in the same inflorescence (synoecious), the same florescence but separated by bracts (paroecious), or on the same gametophyte but different stems (autoecious) (Shaw, 2000). The proximity of antheridia and archegonia on the same gametophyte will therefore have consequences for the likelihood of the occurrence of inbreeding in a species. It is likely that all monoicous bryophytes have at least an element, if not to a significant degree, of inbreeding within populations, and no bryophyte species have been found to be self-incompatible (Shaw, 2000).

Bryophytes can not only undergo "true" self-fertilisation *i.e.* the mating of gametes produced from the same individual (intragametophytic selfing), but also intergametophytic selfing; the mating of gametes produced from haploid individuals produced from the same diploid parent (Klekowski, 1972). Intergametophytic selfing is equivalent to self-fertilisation in seed plants, which results in a reduction in heterozygosity in the diploid stage of 50% in one generation, and a reduction of 90% in 10 generations (Hedrick, 1987a; Hedrick, 1987b; Shaw, 2000). In contrast to this, intragametophytic selfing results in a complete loss of heterozygosity in one generation (Hedrick, 1987a; Hedrick, 1987b; Shaw, 2000). It is intuitive therefore to conclude that monoicous bryophytes should be less genetically diverse than dioicous species (Wyatt *et al.*, 1989b). This has not been found to be the case however when

comparisons have been conducted (Cronberg *et al.*, 2006). It is thought that this is because sexual reproduction is rare in dioicous bryophytes (Cronberg, 2002), with male and female gametophytes often spatially separated (Longton, 1976). This is supported by evidence that sporophytes are frequently produced in monoicous species, but rarely in dioicous species (Longton & Schuster, 1983).

Along with the effect of diversity levels within bryophyte populations, the mating systems in operation will also have an effect on gene flow within a population and thus the scale at which spatial structure is observed. In monoicous species, where inbreeding is frequent, small scale spatial structuring would be expected (Ennos, 2001). If true inbreeding does not occur, then crosses with siblings are likely as gamete dispersal has been shown to be restricted to distances of tens of centimetres in bryophyte populations (Anderson & Lemmon, 1974; Longton, 1976). However, as monoicous bryophytes produce sporophytes more profusely than dioicous species, gene flow mediated by spores, which could travel tens of metres or further, may be more common.

1.4 Population genetics of model species

Population and ecological genetic studies tend to focus on ecologically interesting organisms or organisms that are potentially useful, such as crop wild relatives. The focuses of evolutionary studies are, however, beginning to change. As well as identifying and analysing the evolutionary processes that occur within natural populations, researchers also aspire to identify the genes that affect these evolutionary processes (for a review see Feder & Mitchell-Olds, 2003). This marriage of ecology and molecular genetics has been termed "evolutionary and ecological functional genomics" (Feder & Mitchell-Olds, 2003). To conduct

molecular studies requires a considerable financial investment, with well equipped laboratories and skilled researchers required. Therefore to investigate the functional genomics of the wide number of organisms that have been the focus of prior ecological studies would be untenable, especially since many such species have been poorly characterised in terms of their molecular biology. By contrast, to focus ecological and evolutionary studies on appropriate model organisms, which have been the targets of molecular genetic studies, would have many benefits. Gene function in model organisms has classically been studied by creating mutants in a limited number of laboratory "wild-type" strains, and identifying genes through traditional "forward" and "reverse" genetics. The number of mutants that can be created is limited by the phenotype of the wild type strains, so in order to apply such an analytical approach it is necessary to obtain a wide variety of accessions from various habitats and locations which would enable the identification of ecologically relevant phenotypic traits (Alonso-Blanco & Koornneef, 2000).

The ecological genetics of a number of model species and their close relatives have been studied, however the diversity of model organisms available for ecological studies is low, particularly among plants, in which the majority of studies have investigated the population genetics of *A. thaliana* (Jørgensen & Mauricio, 2004; Miyashita *et al.*, 1999; Sharbel *et al.*, 2000) and its close relatives. The genetic variation found in plants sampled from natural populations has been used to investigate evolutionary interesting genes (Johanson *et al.*, 2000), and quantitative trait loci (QTL) maps have been produced (see Alonso-Blanco & Koornneef, 2000; Tonsor *et al.*, 2005 for reviews). Work on animal species, particularly *D. melanogaster*, has been far more prolific, with a long historical body of work to support this (for a review see Mousette & Derome, 2004), and organisms that are

easily manipulated in the laboratory, such as *Saccharomyces cerevisiae* also have the potential to be useful evolutionary models (Landry *et al.*, 2006).

If investigations into the molecular ecology of model species are to be extrapolated generally, a wide diversity of species need to be studied (Travis, 2006). In plants, genetic studies have mainly been restricted to *A. thaliana* or crop plants, which provide limited coverage of the diversity of plant life. The model moss, *P. patens*, therefore provides an alternative valuable model in which to combine ecological and evolutionary studies with the many molecular tools available for this species. *P. patens* is a useful comparative model to *A. thaliana* and other flowering plants, as it offers the potential to explore characteristics fundamental to the evolution of land plants (Nishiyama *et al.*, 2003; Waters, 2003; Zimmer *et al.*, 2007).

1.4.1 *P. patens* – a useful ecological model

P. patens has grown in popularity as a model species due to the experimentally amenable characteristics of its life cycle, biology and genetics. Haploid protonemal and gametophytic tissue is easily cultured under laboratory conditions, and the dominance of this haploid stage allows for easy identification of recessive mutants (Cove *et al.*, 1997). Gametophytic tissue can be regenerated both from spores and fragments of protonemal tissue, and many plants can be cultured with relative speed and little space requirements, particularly if cultured on agar. The wealth of knowledge on the culture of *P. patens* in the laboratory, and the ability to easily manipulate the resultant gametophytes, should help make the development of laboratory based ecological studies more straightforward.

The life cycle of the most commonly used strain, named "Gransden" after its site of collection near Gransden Wood in Cambridgeshire, takes around three to four

months to complete (Cove, 2005), and years of culture in the laboratory have resulted in a loss of fertility in this strain. However, antheridia and archegonia can be produced in the laboratory, and the Gransden strain is still able to undergo both self- and cross-fertilisation. During the course of this work, wild accessions have been shown to be far more fertile than Gransden, with the production of gametangia occurring several weeks earlier than in the Gransden strain.

P. patens is also an excellent model for "reverse genetic" studies as targeted transgene insertion by homologous recombination occurs with high frequency, in contrast to most other plant species (Schaefer & Zrÿd, 1997). Therefore, interesting genes can be more easily manipulated for functional investigation than in any other plant species.

1.4.1.1 Biology of *P. patens*

P. patens inhabits ephemeral aquatic habitats; namely ponds, lakes, ditches and streams in the northern hemisphere. Populations are found in Europe, from north of the Mediterranean to Scandinavia, Siberia, and North America (Hill *et al.*, 1994; Smith, 2004; Tan, 1979). For a detailed description of *P. patens* see Appendix II, which is taken from Smith (2004). Plants develop in early summer, but their appearance is dependent on the populations being submerged in winter, and on water levels subsequently falling in summer (Cove, 2005). Sporophytes are produced from late summer to early winter. *P. patens* is classified as occasional in Smith (2004), but it is likely to be a relatively under-recorded moss (S. Bosanquet, *pers. comm.*). Further information regarding the habitats that this species occupies is required in order to utilise fully its potential as a model species, however its

distribution in Europe, Japan and the USA facilitates study by global research groups.

P. patens is monoicous, with antheridia and archegonia on the same gametophore, where self-fertilisation can occur (Cove, 1983). The production of sporophytes in laboratory strains of *P. patens* has been shown to be dependent on a day length of about eight hours (Hohe *et al.*, 2002) and a temperature range of 15-19°C (Engel, 1968). In these conditions each sporophyte produces between three and four thousand spores (Engel, 1968). No study has investigated the biology of *P. patens* in natural populations or in an environment other than the laboratory, so very little is known about its ecology.

1.4.1.2 Phylogenetics and taxonomy

P. patens is a member of the Funariaceae, which are in the order Funariales in the Bryopsida. This includes the mosses with a diplolepideous-alternate peristome, which describes two alternate concentric rings of toothlike structures at the capsule opening (Cox *et al.*, 2000). A phylogeny of bryophytes, indicating the position of the Funariales, is shown in Figure 1.3. No phylogeny of the Funariales or Funariaceae has been produced. The Funariaceae comprise six genera; *Physcomitrella* Bruch & Schimp., *Aphanorrhegma* Sull., *Entosthodon* Schwäger., *Funaria* Hedw., *Physcomitrium* (Brid.) Brid. and *Pyramidula* Brid. *Physcomitrella* comprises only two species, *P. patens* and *P. readeri*. Together with *P. patens*, the Funariaceae includes a second well studied species, *Funaria hygrometrica*.

The Funariaceae have posed problems for taxonomists due to a large variability in sporophyte morphology, which is unusual in mosses. The family is characterised by gametophytic characters, including thin-walled obovate leaves with inflated terminal

cells. Shared sporophytic characters include stomata enclosed by a single guard cell (Fife, 1985). Members of this family also readily hybridise, with *P. patens* forming hybrids in natural populations with *Aphanorrhegma* and *Physcomitrium*, which has led to suggestions that these are the same genus (Bryan, 1957). There is still controversy over the naming of *P. patens*, and it is classified in The Moss Flora of Britain and Ireland (Smith, 2004) as *Aphanorrhegma patens* (Hedw.) Bruch & Schimp. This is after the re-interpretation of the Funariaceae by Lindberg (1864), which included *Physcomitrella* within *Aphanorrhegma* (in Fife, 1985). *Physcomitrella*, as first described by Bruch, Schimper and Gümbel (1836-1855; in Fife, 1985), remains as a separate genus in many other regional floras, including the Flora of the USA (Goffinet, 2007), and important floras from Japan (Ochi, 1968) and New Zealand (Scott & Stone, 1976; Stone & Scott, 1973). This nomenclature has been retained for the purposes of this study as it is most commonly recognised by the scientific community as *Physcomitrella*.

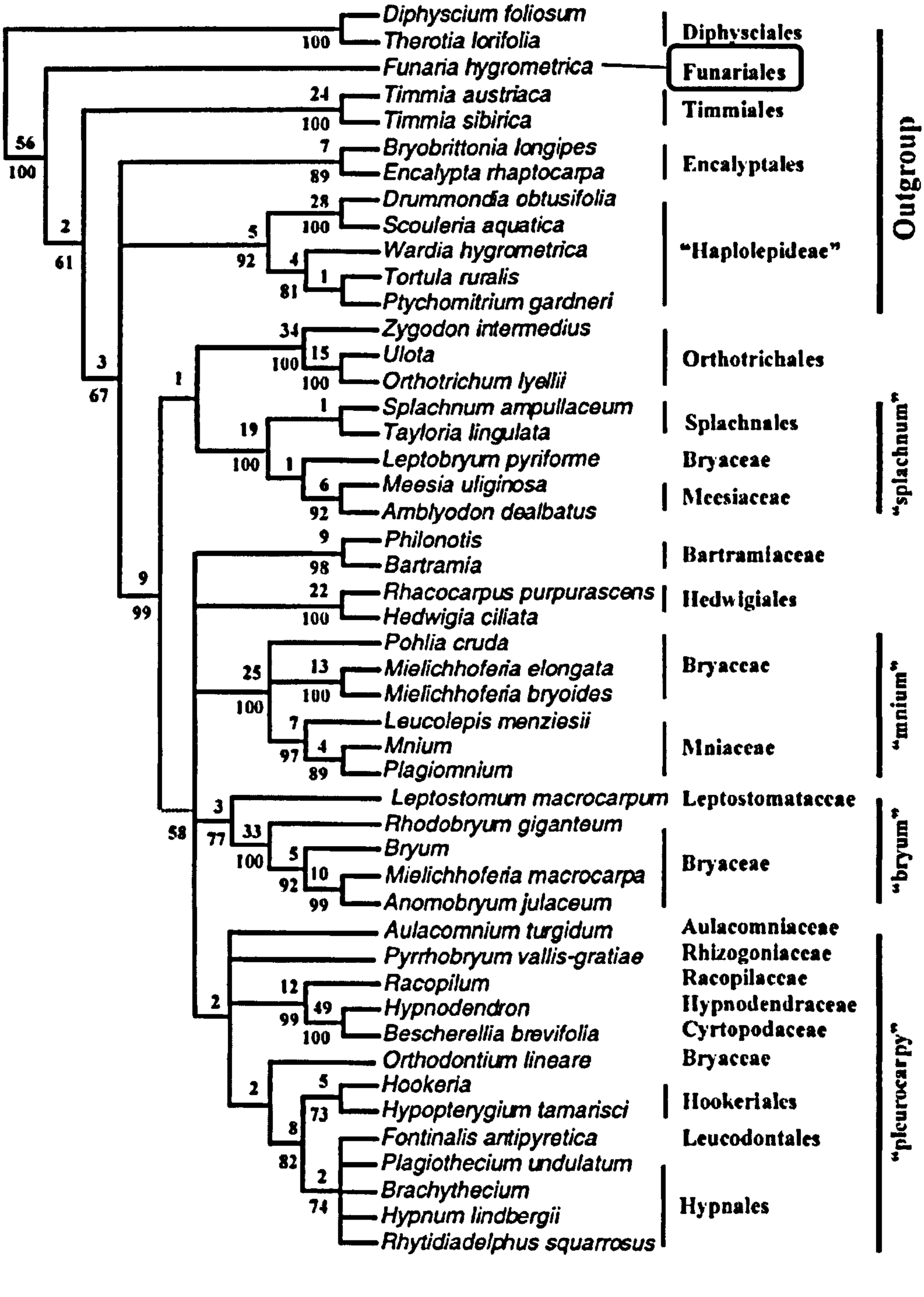


Figure 1.3 Phylogeny of the diplolepideous-alternate mosses, based on 18S rDNA, *trn*L-F, *rbc*L, and *rps*4 sequences, taken from Cox *et al.,* (2000). The position of the Funariales, as represented by *F. hygrometrica*, is indicated.

1.4.1.3 Historical perspective

Mosses have long been the focus of genetic and developmental studies, with mosses utilised as important model systems as early as the 1940's (see Cove *et al.*, 1997 for a review). The first significant genetic studies with *P. patens* were conducted by Engel (1968), who produced mutants of *P. patens* using chemical and X-ray mutagenesis. Further significant studies did not come until the 1970's, when *P. patens* was adopted as a model by David Cove at the University of Leeds, and Wolfgang Abel at the University of Hamburg (Schaefer & Zrÿd, 2001). Important developments during this period included the regeneration of plants from protoplasts and the resultant ability to conduct complementation studies using protoplast fusion (Grimsley *et al.*, 1977a; Grimsley *et al.*, 1977b). Shortly after this, self-sterile but cross-fertile mutants were developed, which allowed test crosses to be conducted (Courtice *et al.*, 1978). Further studies of *P. patens* focussed on its developmental biology, with a number of developmental mutants identified (see Cove & Knight, 1993 for a review).

P. patens became the focus of molecular genetic experimentation in the late 1990s, with the discovery that when transformation was attempted via polyethyleneglycol transformation of protoplasts, DNA containing homology with endogenous sequences integrated preferentially at the cognate loci. This "gene targeting" occurred by homologous recombination, a feature absent in other model plants. This allows far greater versatility in *P. patens* than other plant species, as specific genes can be targeted for very precise mutagenesis, revealing insights into gene function (see Reski, 1999; and Schaefer & Zrÿd, 2001 for reviews).

1.4.1.4 Genomics

P. patens has been found to possess 27 chromosomes in the wild type strain (Reski *et al.*, 1994), and following the assembly of a first-draft complete genome sequence, its genome has been estimated to contain over 35,000 gene models (Rensing *et al.*, 2008). An important resource that could be utilised in evolutionary studies of bryophytes, is the *P. patens* expressed sequence tag (EST) database, which has been used to annotate the first draft of the genome sequence assembly (Rensing *et al.*, 2008). ESTs provide an insight into possibly evolutionarily significant genes, as they provide information about the genes expressed in particular tissues, at particular times and under particular conditions (Bouck & Vision, 2007). Along with information about gene function, ESTs can also be used to produce molecular markers that could be utilised in population genetics studies, and a number of microsatellite markers have recently been developed using ESTs from *P. patens* (von Stackelberg *et al.*, 2006).

The function of a number of interesting genes has been discovered in *P. patens*, including genes that could have been important in the transitions to life on land. *P. patens* is tolerant to a number of environmental stresses, and the *P. patens* EST database contains sequences that are homologous to stress-related genes in other species, including those that are up-regulated in response to dehydration (Frank *et al.*, 2005). *P. patens* is a dehydration tolerant species, and plants have been shown to have the ability to regenerate after losing 92% of their fresh weight after dehydration treatments (Frank *et al.*, 2005). Conversely, desiccation tolerant species are able to regenerate after losing over 99% of their fresh weight, and they possess specific mechanisms that reduce the effect of desiccation and repair the cellular damage caused by it (Oldenhof *et al.*, 2006; Oliver *et al.*, 2000). To fully explore the

mechanisms that allowed plants to colonise land, desiccation tolerant species should be studied, and bryophytes such as *Tortula ruralis* are being developed as models for this purpose (Oliver *et al.*, 2000; Oliver *et al.*, 2005).

Although the functions of some genes have been inferred through interrogation of the genome sequence and EST database, the function of the majority of genes in *P. patens* remains unknown (Rensing *et al.*, 2008). A significant step on the road to relating the genome sequence to gene function has been the production of a high density linkage map (Kamisugi *et al.*, 2008). This linkage map was produced using both AFLP and microsatellite markers, and the AFLP markers used have been anchored to the genome sequence by converting the AFLP loci into sequence characterised amplified length polymorphisms (SCALPs). The development of these AFLP markers provides a potentially significant resource for ecological and evolutionary genetics. Any AFLP loci that are found to be of interest through genetic association with evolutionarily important traits could potentially be used to identify sequenced genes subsequently found to be present as markers in the linkage map.

1.5 Aims of the project

The model moss *P. patens* is one of the most well resourced model plant species and is certainly the most well-characterised bryophyte, in terms of its molecular biology. Nevertheless, there is virtually nothing known of the ecology and population genetics of this plant. To extend the variety of studies that can be conducted and increase the utility of the molecular knowledge that has been gleaned from this model, more information about the diversity of this species in its natural environment must be obtained. The laboratory strains of *P. patens* that are currently studied include a number of *P. patens* accessions, termed "Physcotypes". They comprise isolates from widely different parts of the world, but only a few sporophytes have been collected from each locality. Although probably unintentional, these Physcotypes are likely to have been selected based on their phenotypes or suitability for laboratory culture (Alonso-Blanco & Koornneef, 2000). Consequently, a larger collection of plants representative of the genetic variation in different populations is required, if the molecular tools available for genetic analysis of *Physcomitrella* are to be effectively mobilised in the study of its population genetics. This requirement has underpinned the investigation described in this thesis: the first study of the genetic diversity of *P. patens* in a variety of natural populations, using multiple samples derived from a number of sites.

P. patens is an interesting model in which to answer evolutionary questions. Being a haploid, inbreeding, and ephemeral bryophyte, it has a life history in contrast to most well studied plant species. In the chapters ahead I will examine the level of genetic variation within natural populations of *P. patens* and how this diversity is partitioned spatially within and between species. I will also address how the mating systems operating within populations of bryophytes can be investigated, by trialling

a novel method of analysis, and applying this to natural populations of *P. patens.* I have also identified a population of *P. readeri* in West Yorkshire, which is a bryophyte new to Europe. By conducting this study I hope to characterise the population genetic processes occurring within natural populations of *P. patens*, so that this may further the understanding of the ecology and evolution of ephemeral plants and bryophytes in general, about which little is currently known.

Chapter 2 General Materials and Methods

2.1 Study sites

Between August 2005 and November 2006, mature sporophytes of *P .patens* were collected and population data were recorded from nine populations in Great Britain. Populations are generally defined by the limits of the pool, lake, ditch or other such habitat in the vicinity that they occupy, so populations are therefore discrete but hugely variable in size. Populations were chosen according to a hierarchical spatial scale protocol, with regions and populations separated by distances of 5-20km, 50-100km and 100km-400km. Population details are given in Table 2.1with their distribution in the UK shown in Figure 2.1. Photographs showing examples of the habitats that *P. patens* occupies are given in Figure 2.2.

When identifying possible target populations, a total of 42 locations were examined, based on previous records of *P. patens* at these locations. Of these only 11 locations were found to contain *P. patens* plants. The details of these locations and the historical records of *P. patens* are given in Appendix I. The success of discovery of populations was generally dependant on the age of the record, and often plants were not found in locations where the records were over 10 years old. However, Papercourt Marshes in Surrey was found to contain a large population of *P. patens*, yet the record was dated from 1973. In many populations that had only recently been recorded, evidence of *P. patens* was still not found. In certain cases, such as at Bosherton Lakes in Pembrokeshire and several locations in Cambridgeshire, exact GPS grid references were provided by bryologists who had found *P. patens* only a few years previously. In these instances it is likely that the water levels in the lakes,

reservoirs and streams in which the populations inhabit were too high, and that the populations were submerged. Other locations were found to contain established plant communities, with little evidence of aquatic areas, indicating that the area that the population had previously occupied had since become too dry for *P. patens*. It can be concluded that *P. patens* is very much dependent on a regular cycle of inundation and exposure for the persistence of populations.

Figure 2.1 Location of populations of *P. patens* (•) sampled. The distribution of populations within Cambridgeshire is shown in greater detail in b).

Table 2.1 Site details of *P. patens* populations

Population ID	Population/ Location	County	Country	Description	OS grid reference
A	Pant-y-llyn	Carmarthenshire	Wales	Pond	SN6063116683
B	Llanfihangel Gobion	Monmouthshire	Wales	Lake/cattle trampled ground	SO3454309362
C	Stokesay	Shropshire	England	Pond	SO4353681633
D	Papercourt Marshes	Surrey	England	Lake	TQ0347556403
E	Ouse Washes	Cambridgeshire	England	Ditch	TL4718784999
F	Wicken Fen	Cambridgeshire	England	Shallow pool	TL5610969842
G	Swavesey	Cambridgeshire	England	Ditch/cattle trampled ground	TL3640169733
H	Mepal	Cambridgeshire	England	Ditch	TL4363181163
I	Lindley Wood Reservoir	West Yorkshire	England	Reservoir	SE2110349868

B

D

E

F

Figure 2.2 Photographs of example *P. patens* populations. B = Llanfihangel Gobion, D = Papercourt Marshes, E = Ouse Washes, F = Wicken Fen.

Sporophytes were collected according to a hierarchical spatial scale protocol within populations, with the sporophytes collected that were separated by distances in the region of 10cm – 50cm, 1m-5m and 20m – 50m within each site. An example of the sampling structure is shown in Figure 2.3. Within-site distances were measured between consecutive samples and bearings were taken between samples, so the points where samples were collected from could be converted to x and y coordinates, to produce a plan of each site and enable distances between each pair of samples to be calculated. Mature, undehisced spore capsules were collected and stored in 1.5ml microcentrifuge tubes. Population area was recorded by walking around the perimeter of the population with a geographical positioning system and recording the perimeter position at regular intervals. The area within this perimeter was then calculated using ArcGIS 9.1 (ESRI, Redlands, California, USA).

The density of *P. patens* around each individual collected was also measured, using quadrats of 10x10 cm and 1x1m. The quadrats were made up of a 7 x 7 grid comprising 49 cross-over points, and the occurrence of a *P. patens* plant beneath each point was recorded, with the sample collected as the centre point.

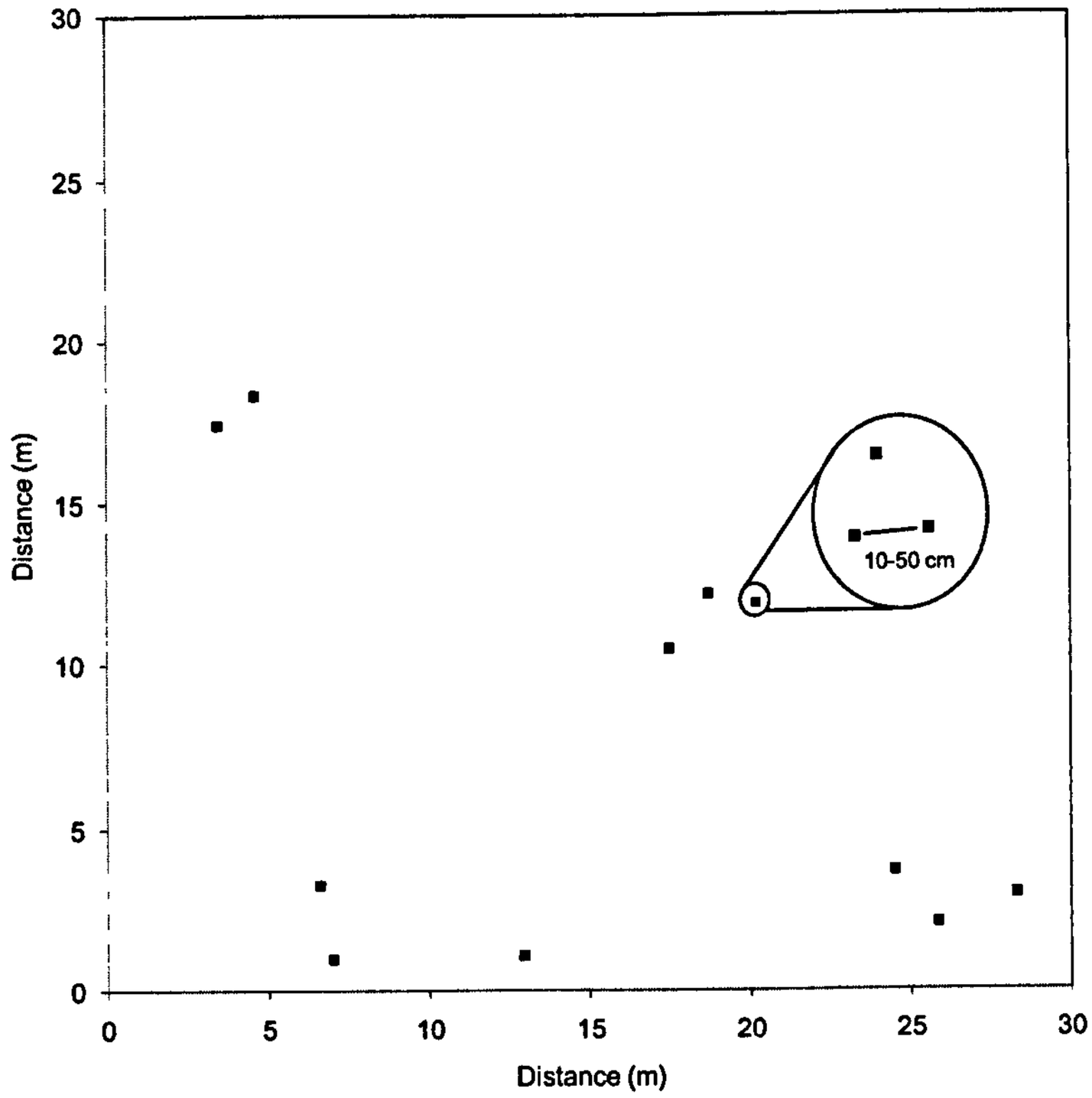


Figure 2.3 Positions of *P. patens* sporophytes that were collected from Pant-y-llyn in Carmarthenshire. Each labelled point indicates clusters of 1-3 sporophytes separated by distances of less than 50cm.

2.2 Sample culture and DNA extraction

Intact spore capsules were rinsed in sterile distilled water, and then transferred to domestic bleach (*ca.* 5% free Cl_2) for five minutes in a sterile 1.5ml microcentrifuge tube. Capsules were rinsed twice in sterile distilled water, transferred to a clean tube and crushed in 1ml sterile distilled water, to release spores. Spores were germinated on BCD medium (Knight *et al.*, 2002) containing NH_4 (5mM ammonium tartrate) and 10mM $CaCl_2$ under constant illumination at 25°C. Following germination, individual plants were transferred to BCD medium containing NH_4 and 1mM $CaCl_2$, and grown under constant illumination at 25°C. For DNA extraction, individual plants were grown on medium overlaid with cellophane for isolation of DNA as described by Knight *et al.* (2002).

2.3 AFLP analysis

AFLP analysis (Vos *et al.*, 1995) of each plant was conducted as described by Myburg and Remington (2000). Approximately 100-200ng DNA was digested with *EcoR*I and *Mse*I and ligated with the adapter sequences described by Myburg & Remington (2000).

AFLP pre-amplification reactions were conducted using *EcoR*I and *Mse*I primers with a single selective nucleotide (*EcoR*I: GACTGCGTACCAATTCN; *Mse*I: GATGAGTCCTGAGTAAN) and selective amplification reactions were conducted with either one or all of the following primer combinations (see Chapters 3 and 5 for specific details): *EcoR*I+AC/*Mse*I+CG, *EcoR*I+AG/*Mse*I+CT and *EcoR*I+CG/*Mse*I+TA. The primer combinations selected were chosen as they produced easily scorable traces with polymorphic loci, based on a trial of AFLP

primer combinations by Kamisugi *et al.* (2008) for production of the *P. patens* genetic linkage map. These *EcoR*I primers were fluorescently labelled with FAM, HEX and NED respectively, so that the length of fragments could be scored automatically using a capillary sequencer. PCR reactions were conducted as follows according to a modified method of Berres (2001). Each PCR reaction (15µl) contained PCR buffer, 0.2mM dNTP mix, 18ng *EcoR*I primer, 90ng *Mse*I primer, 0.15µl *Taq* polymerase, and 1.5µl pre-amplification product (diluted 1:50). Cycle parameters were 12 cycles of denaturation at 94°C for 10 seconds, annealing at 65°C for 30 seconds and extension at 72°C for 60 seconds – 0.7°C/cycle, followed by 24 cycles of denaturation at 94°C for 10 seconds, annealing at 56°C for 30 seconds and extension at 72°C for 60 seconds + 1 second/cycle, with a final extension at 72°C for 2 minutes. Amplified AFLP fragments were then electrophoretically resolved using an ABI 3130 Genetic Analyser (Applied Biosystems, Foster City, USA).

Chapter 3 Genetic structure within and between populations of *P. patens*

3.1 Introduction

P. patens provides an interesting model in which to study evolutionary questions as it is an ephemeral, inbreeding species, which has a number of implications for the population genetics of this organism. *P. patens* resides in aquatic habitats such as reservoirs and ponds (Smith, 2004), with receding water levels exposing populations (Cove, 2005). Thus *P. patens* populations could be expected to exhibit features such as low levels of genetic diversity and high fecundity, as observed in ephemeral flowering plant species (Barrett & Husband, 1997).

An important aspect of the population genetics of plant species is the spatial structure of plant genetic variation within and between populations, yet this has been little studied in bryophyte populations. Where this has been investigated it has been studied in dioicous (Cronberg, 2002; Cronberg *et al.*, 1997; Gunnarsson *et al.*, 2005) or epiphytic bryophytes (Snäll *et al.*, 2004), which would be expected to exhibit very different population biology to a monoicous ephemeral such as *P. patens*.

The spatial genetic structure of individual populations can be studied in plant populations using spatial autocorrelation analysis whereby a measure of genetic similarity between pairs of individuals, such as the kinship coefficient, is correlated with the position of individuals in a mapped population (Heywood, 1991; Vekemans & Hardy, 2004). There are a number of different methods for estimating the genetic relationship between a pair of individuals when using spatial autocorrelation

analysis, such as those by Ritland (1996), Queller and Goodnight (1989), Li *et al.* (1993), Hardy and Vekemans (1999), Lynch and Ritland (1999) and Wang (2002). However, Loiselle *et al.* (1995) has developed an estimator based on that of Hardy and Vekemans (1999), which is itself based on Moran's *I* statistic, that performs well in highly selfing species (Vekemans & Hardy, 2004), and it has been used successfully in spatial autocorrelation analysis of a population of an epyphytic bryophyte by Snäll (2004).

In this study I investigated the level of genetic diversity within populations of *P. patens*, along with the spatial genetic structure both within and between populations. I have used amplified fragment length polymorphism (AFLP) analysis, as it is a robust and reliable genetic marker which is highly reproducible, with many loci amplified in individual polymerase chain reactions (Blears *et al.*, 1998; Ritland & Ritland, 2000). Maximising the number of loci to be sampled is important when studying diversity within populations as the genetic variation between individuals may be low. I also investigated the spatial partitioning of genetic diversity between populations by structuring the distance between populations at differing spatial scales as well as by sampling within populations at multiple spatial scales.

3.2 Materials and methods

For a description of the populations of *P. patens* sampled, see Chapter 2, Table 2.1. For details of the sampling procedure, sample culture, DNA extraction and AFLP analysis, see the relevant sections in Chapter 2. The number of sporophytes sampled from each population studied in this chapter is given in Table 3.1.

Table 3.1 Populations sampled, including the number of sporophytes collected at each population.

Population ID	Population/ Location	No. of sporophytes
A	Pant-y-llyn	19
B	Llanfihangel Gobion	34
C	Stokesay	20
D	Papercourt Marshes	32
E	Ouse Washes	27
F	Wicken Fen	24
G	Swavesey	15
H	Mepal	3
I	Lindley Wood Reservoir	37

3.2.1 AFLP profile scoring and error-rate testing

AFLP fragments to be scored were initially filtered using Genemapper® Software v3.7 (Applied Biosystems). Only unambiguous fragments that were between 70 and 500 bp in length were included in the analysis. The fragment fluorescence intensity, or peak height data that was produced by Genemapper® was then converted into a genotype table, where AFLP loci were scored as either present (1) or absent (0) in each individual, using the free r-script based software, AFLPScore version 1.3 (Whitlock *et al.*, 2008). AFLPScore normalises peak height data between

individuals, and scores loci as either present or absent by determining the optimum scoring conditions based on error-rate testing. The error rate testing procedure determines the loci to be scored based on the average peak height at each locus, or the locus-selection threshold, and the relative height at which fragments are scored as present or absent, or the genotype-calling threshold. The locus-selection threshold and the genotype-calling threshold were chosen based on the % mismatch error rate, and the number of loci retained. A binary matrix of AFLP profiles was then produced by AFLPScore based on these criteria.

To estimate the rate of error within the AFLP procedure, each step of the process was conducted on two replicated samples from 24 individuals. Each of these 24 samples were sub-divided into two separate tissue samples, with separate DNA extractions, AFLP reactions and profile scoring was then conducted independently on each set of 24 samples.

3.2.2 Data analysis

3.2.2.1 Genetic diversity and linkage disequilibrium

Within-population diversity estimates based on allele frequencies were calculated using AFLP-SURV version 1.0 (Vekemans *et al.*, 2002), based on methods used by Lynch and Milligan (1994). The proportion of polymorphic loci at the 5% level, and Nei's gene diversity index (Nei, 1973) were calculated for each population. The correlation between the Nei's gene diversity index within each population and population area was tested by calculating the Spearman rank correlation coefficient using XLSTAT (Addinsoft, 2008). The presence of linkage disequilibrium, or the non-independence of loci, was tested for using the programme LIAN (Haubold & Hudson, 2000). LIAN tests for statistically independent assortment of alleles by first

computing the number of loci at which each pair of samples differs. LIAN then calculates the variance of the mismatch values (V_D), and compares this to the variance expected under linkage equilibrium (V_e). The null hypothesis $H_0 : V_D = V_e$, was tested by a Monte Carlo simulation test using 10,000 re-samplings of the data (see Haubold & Hudson, 2000). A measure of linkage, the standardised index of association (I_A^s), was also calculated in LIAN as:

$$I_A^s = \frac{1}{l-1}\left(\frac{V_D}{V_e} - 1\right)$$

Where l is the number of loci analysed.

3.2.2.2 Genetic structure within and between populations

To analyse genetic differentiation within and among populations, Wright's F_{st} (Wright, 1978) was calculated and an Analysis of Molecular Variance (AMOVA, Excoffier *et al.*, 1992) was performed in ARLEQUIN (Schneider *et al.*, 2000). The significance of variance components and F_{st} was tested using a non-parametric permutation method, as described in Excoffier *et al.* (1992), using 2000 permutations of the data. Pairwise F_{st} values between populations were calculated using the free statistical software package SPAGeDi (Hardy & Vekemans, 2002). The relationship between pairwise F_{st} and geographical distance between populations was tested by conducting a Mantel test using XLSTAT.

The kinship coefficient between pairs of individuals within and among populations was calculated according to Loiselle *et al.* (1995) using SPAGeDi (Hardy & Vekemans, 2002). Kinship coefficients, or coancestry coefficients, are based on the

probability of identity of two homologous alleles, p_i and p_j, sampled randomly within two mapped individuals, *i* and *j* (Hardy & Vekemans, 2002), or as defined in Snäll *et al.*(2004) for haploid bryophyte individuals as;

$$\hat{\rho}_{ij} = \frac{\sum_{ij}(p_i - \tilde{p})(p_i - \tilde{p})}{k\tilde{p}(1-\tilde{p})} + \frac{n}{n+1}$$

where the first term is the expected value of ρ_{ij}, $\tilde{p}$ is the average frequency of the allele in the sample and k =n(n − 1)/2 is the total number of possible pairwise connections between n samples. The second term adjusts for bias attributable to finite size, and results in ρ_{ij} having an expected value of zero for a population in Hardy–Weinberg equilibrium. The fine-scale genetic structure amongst all individuals and within populations was estimated using spatial autocorrelation analysis (see Vekemans & Hardy, 2004), which was also calculated in SPAGeDi. All individuals within the dataset were first pooled and the kinship coefficients of all pairs of individuals were binned into physical mapped distance classes of <1m, 1-10m, 10-50m, 50-500m, 500m – 100km and 100-400km. SPAGeDi calculated the mean kinship coefficient for each distance class, and the significance of this kinship coefficient (significantly different from 0) was tested by performing 2000 permutations of the data. The observed mean kinship coefficient was then compared to the 95% confidence intervals (CIs) produced from the permuted kinship coefficients. This procedure is analogous to a Mantel test (Hardy & Vekemans, 2002). This spatial autocorrelation analysis was then repeated within populations, excluding population H from Mepal in Cambridgeshire, as only three sporophytes were collected from this population. Distance classes of <10m, 10-50m, 50-100m, 100-200m and 200-500m, were used in these analyses, however not every distance

class was used in all analyses, as the power of this procedure is dependent on the number of pairs within each distance class. Only distance classes with at least 30 pairs in each were used for each within-population analysis.

3.3 Results

After initial analysis of a subset of sporophytes collected from all populations studied, one population, at Lindley Wood Reservoir in West Yorkshire, was found to have very different AFLP profiles to those obtained from samples from all other populations. A preliminary analysis of pairwise F_{st} between populations (calculated using SPAGeDi), and geographic distance (Figure 3.1), showed that the inclusion of plants from the population at Lindley Wood distorts the data. It has since been confirmed that the samples from Lindley Wood were not infact *P. patens*, but *P. readeri*, and further details of this are discussed in Chapter 4.

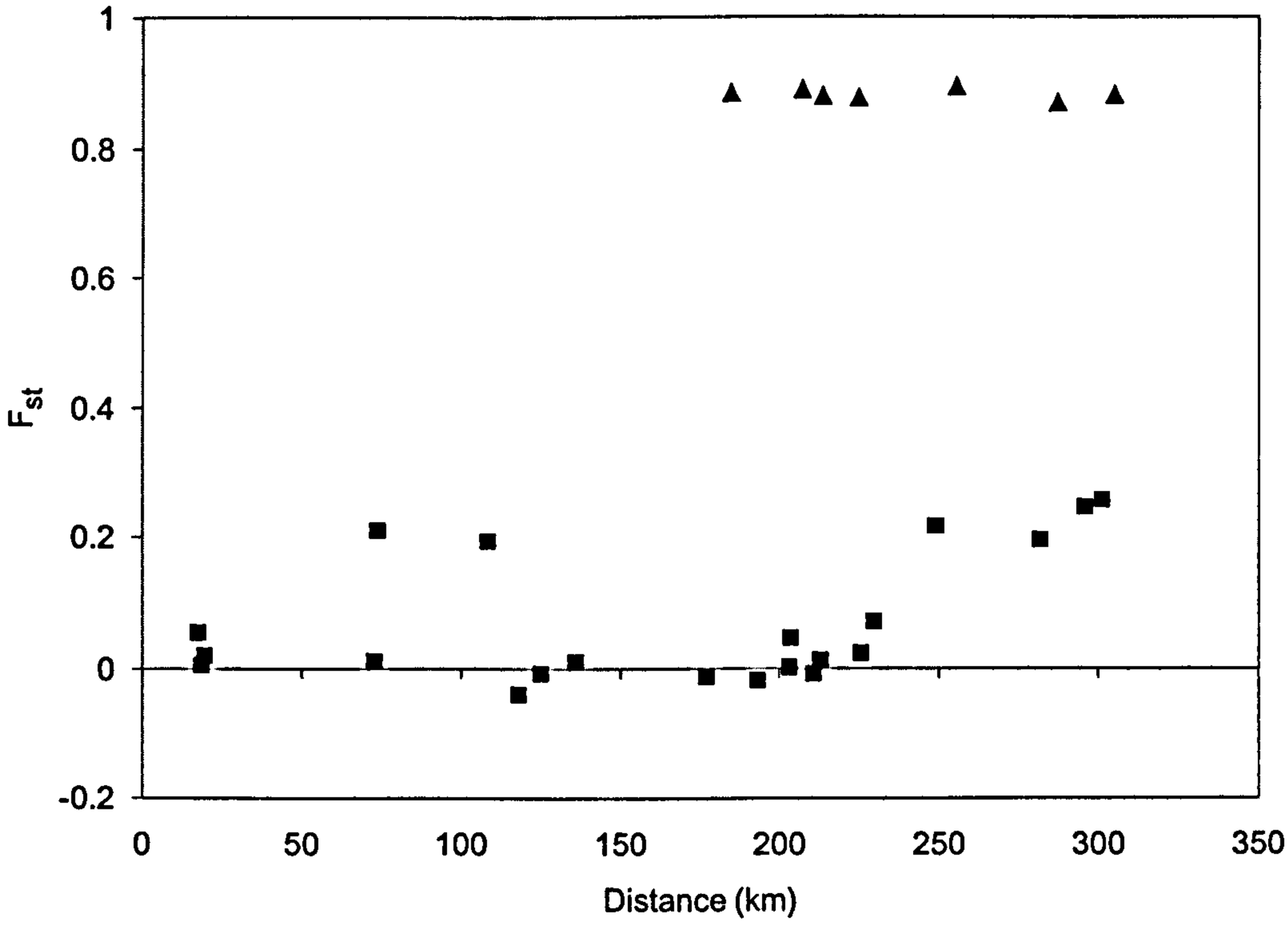


Figure 3.1 The relationship between pairwise F_{st} between populations sampled and geographic distance between populations, calculated using a subset of samples in a preliminary analysis. F_{st} between plants from Lindley Wood Reservoir and all other populations = ▲, and calculations between all other populations = ■.

3.3.1 Genetic diversity and linkage disequilibrium

From a total of 233 loci that were scored using the three primer combinations, 152 of these were found to be polymorphic (65.5%). The AFLP analysis error rate, averaged over all loci, was 3.00%, which is within reasonable bounds for AFLP analysis in plant populations (Bonin *et al.*, 2004). The proportion of polymorphic loci, expressed as a percentage, in each population is shown in Table 3.2 along with estimates of Nei's gene diversity. These estimates were plotted against population area (Figure 3.2). There is a negative correlation between population area and the

gene diversity index within populations (correlation coefficient, *r*, = -0.716), however this correlation is not significant (P = 0.088). There is no significant difference between the gene diversity indices in populations B to F, however, the level of diversity within population A from Pant-y-llyn in Carmarthenshire, was higher than that found in other populations. The gene diversity index within this population was shown to be 0.11875, compared with an average gene diversity index of 0.04536 for all other populations.

Significant linkage disequilibrium was found in the seven populations analysed (P<0.001 in all cases), and the standardised index of association was similar in all, ranging from 0.0122 in Ouse Washes to 0.0540 in Pant-y-llyn (Table 3.2).

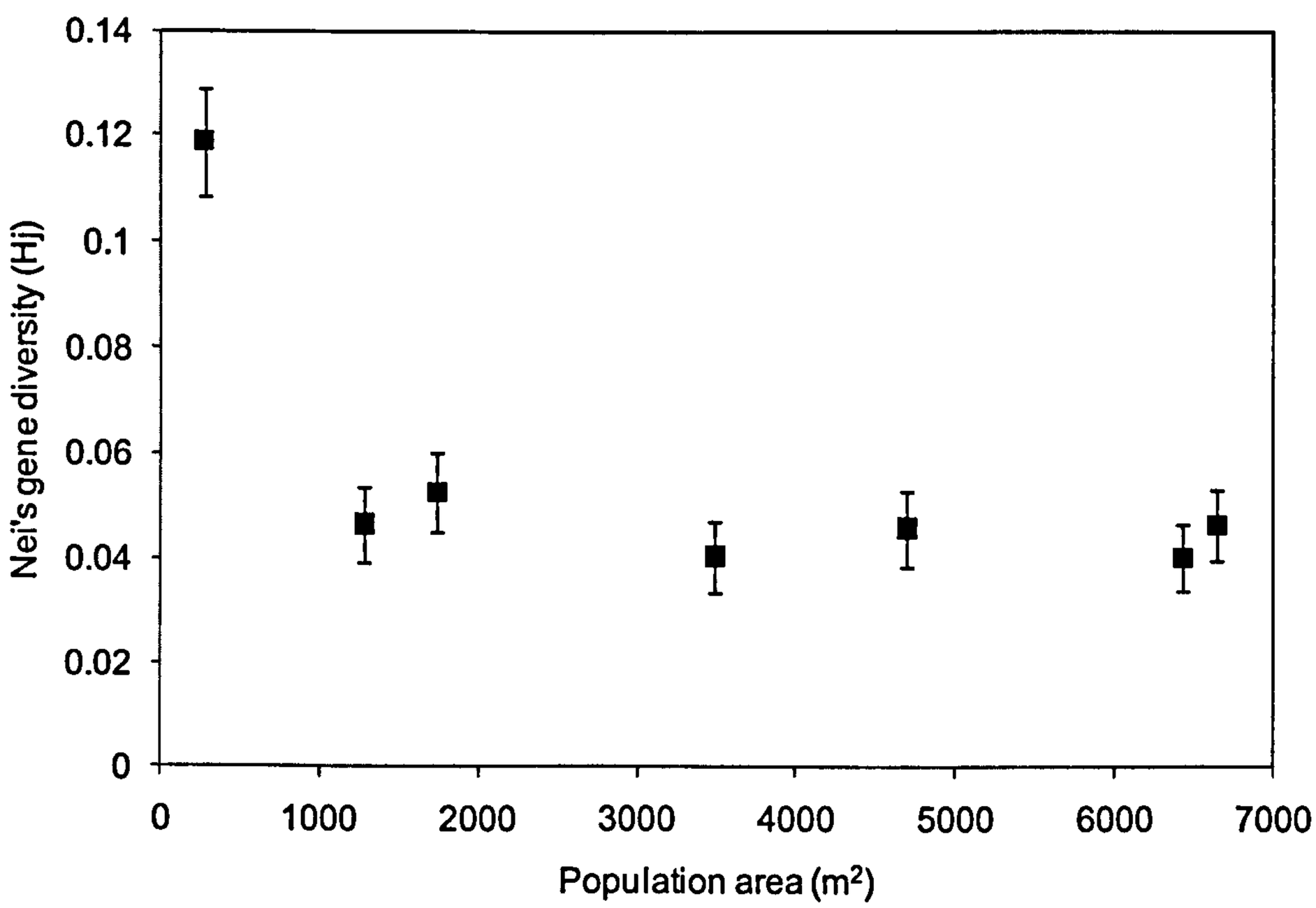


Figure 3.2 Correlation between Nei's gene diversity within populations of *P. patens* (excluding population H) and population area. *r* = -0.714, P = 0.088.

Table 3.2 Proportion of polymorphic loci, Nei's gene diversity (Hj) and the standardised index of association (I^s_A) within *P. patens* populations.

Pop ID	Population/Location	No. of sporophytes	% Polymorphic loci	Hj (S.E)	I^s_A
A	Pant-y-llyn	19	51.3	0.11875 (0.01031)	0.0540
B	Llanfihangel Gobion	34	12.9	0.04023 (0.04023)	0.0270
C	Stokesay	20	19.0	0.04033 (0.00654)	0.0134
D	Papercourt Marshes	32	14.7	0.04651 (0.00702)	0.0186
E	Ouse Washes	27	10.3	0.04640 (0.00686)	0.0122
F	Wicken Fen	24	12.9	0.04562 (0.00707)	0.0153
G	Swavesey	15	22.8	0.05248 (0.00757)	0.0138
H	Mepal	3	7.8	0.04598 (0.01043)	-

3.3.2 Genetic structure within and between populations

The results of the AMOVA analysis, which was used to test for genetic differentiation between populations of *P. patens*, are shown in Table 3.3. These results indicate that most of the variation was partitioned within populations, at 93.18%, compared to only 6.82% amongst populations. A significant fixation index (F_{st}) of 0.062 indicates that there is only moderate genetic differentiation between populations (Wright, 1978).

The relationship between pairwise F_{st} and geographical distance between populations is shown in Figure 3.3. Analysis of this relationship using a Mantel test indicated that there was no significant relationship between pairwise F_{st} and geographical distance ($r = 0.396$, $P = 0.081$).

Graphical representations of the kinship coefficient spatial autocorrelation analysis are shown in Figure 3.4. When all individuals from all populations were pooled, spatial structure within and between populations was detected. For distance classes of 1-10m, 10-50m, and 50-500m, significant positive kinship coefficients were observed. A significant negative kinship coefficient was also observed in the distance class 100-400km. For within population analysis of kinship coefficients alone, only population B, at Llanfihangel Gobion in Monmouthshire, contained a distance class, at 10-50m, with a kinship coefficient significantly different to zero.

Table 3.3 Analysis of molecular variance (AMOVA)

Source of variation	d.f	Sum of squares	Variance	%	*P*
Among Populations	6	101.292	0.45122	6.82	<0.001
Within Populations	161	992.952	6.16740	93.18	
Total	167	1094.244	6.61862		

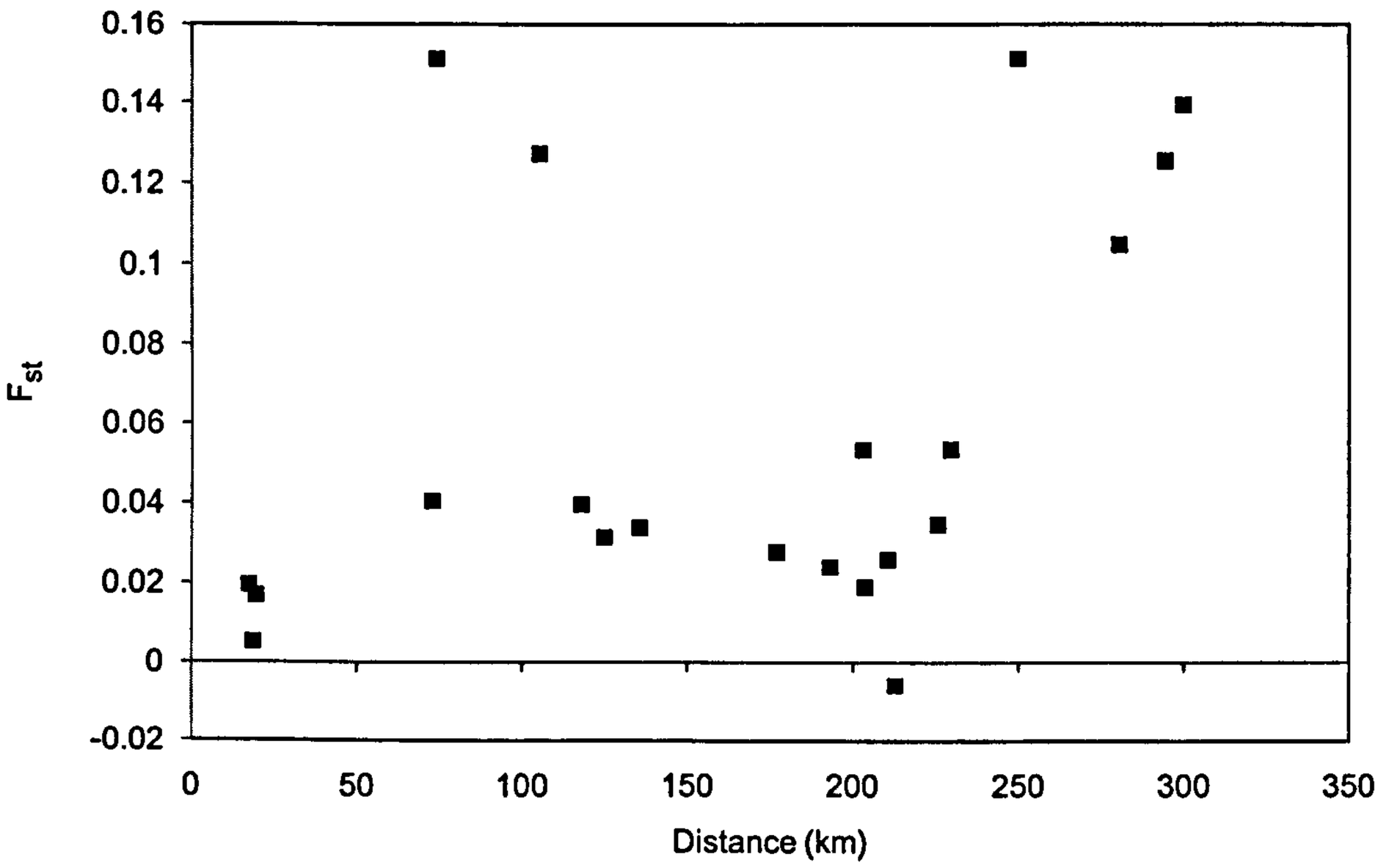


Figure 3.3 The relationship between pairwise F_{st} between populations sampled and geographic distance between populations. $r = 0.396$, P = 0.081.

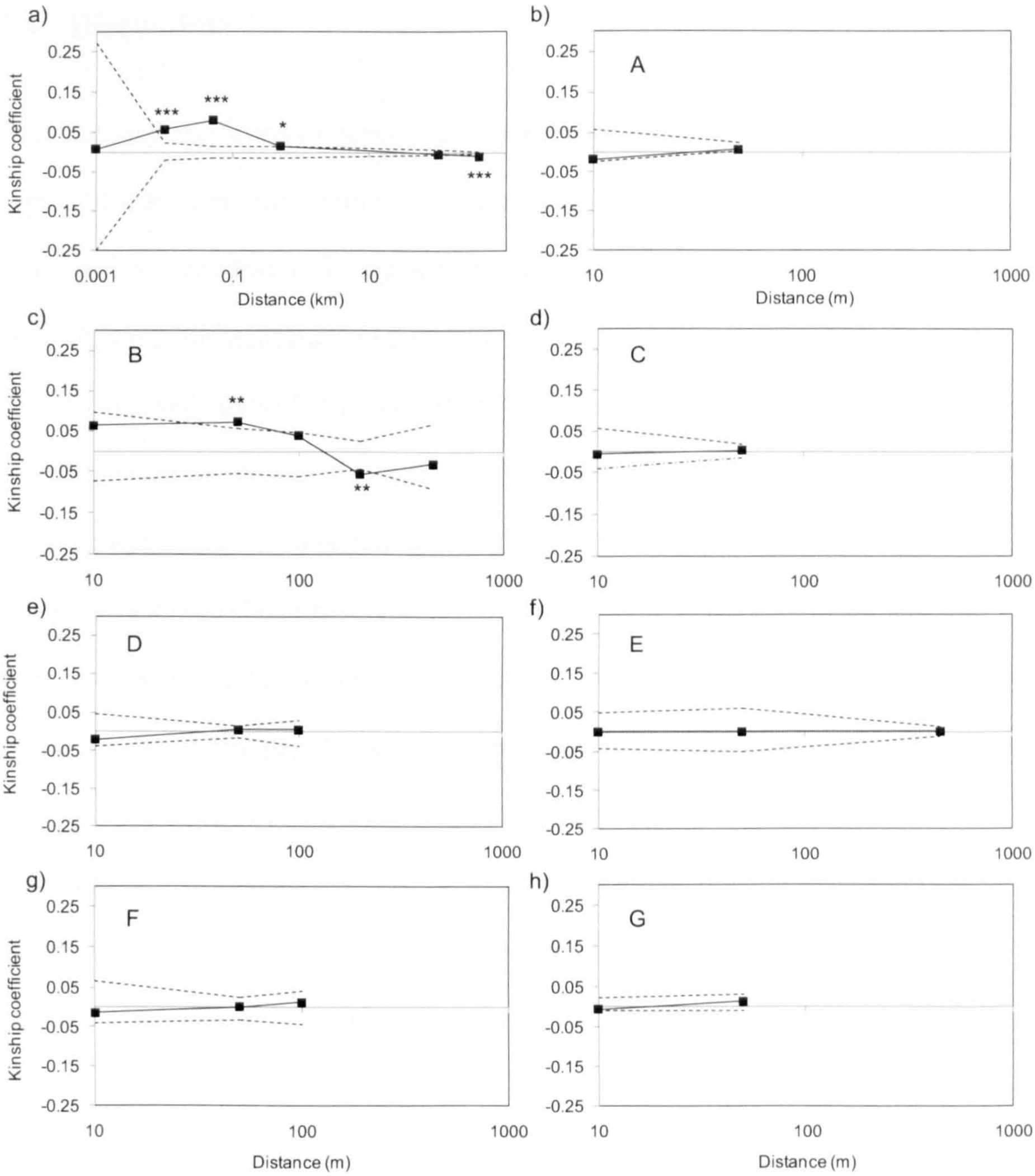


Figure 3.4 Kinship coefficient autocorrelation at a) over all sites, and b) to h) within populations A to G. Mean kinship coefficient of each distance class is plotted against maximum class distance. Significant coefficients are indicated as $P < 0.05$ = *, $P < 0.01$ = ** and $P < 0.001$ = ***, based on the null hypothesis of kinship coefficient = 0. The upper and lower 95% confidence intervals are shown by dashed lines.

3.4 Discussion

The relatively low levels of genetic variation observed in most of the *P. patens* populations are consistent with the levels of variation observed in other ephemeral plant species (reviewed in Barrett & Shore, 1989). Although low levels of genetic diversity are often associated with small population size, this does not appear to be the case in populations of *P. patens*. Barrett and Husband (1997) found that population size had no significant effect on diversity when using allozyme loci to measure heterozygosity in populations of *Eichhornia paniculata*, an aquatic plant occupying ephemeral habitats. However, the authors did find a significant positive correlation between the harmonic mean population size over a three year period and heterozygosity within populations. As populations studied were not static over this time but variable in size, any measure of heterozygosity taken in a single year would have been affected by the size of the populations in previous years, particularly if the population was much smaller. The authors of this study also found a significant positive correlation between population size and diversity in morphological characters. This may be due to the fact that allozymes, like other molecular markers such as AFLPs, are selectively neutral, and analysing morphological characters under strong selection pressure may yield different results.

The relatively higher level of genetic variation observed within population A in Pant-y-llyn, was, however, unexpected, as this was also the smallest population studied. This could reflect the stability of this population, in terms of population age, along with the frequency and extent of flooding in this habitat, when compared with the other populations. An alternative explanation, as discussed above, is that the current size of populations of *P. patens* may not be representative of population

sizes in previous years. Studies that assess the impact of population age on levels of diversity in bryophyte populations are rare due to the complexities in accurately determining the time of population establishment. However this has been investigated by Cronberg (2002), who assessed levels of diversity in populations of *Hylocomium splendens* in an archipelago with islands of known age. In this study genetic diversity was found to increase with increasing age, which could be explained by the constant recruitment of new alleles by populations.

Although no data are available to corroborate these theories, it is likely that where an ephemeral population is relatively stable, and in the case of this *P .patens* population, where flooding in the habitat is infrequent, that a higher level of genetic variability may be accumulated or maintained. To date there have been very few studies into the dynamics of genetic structure within ephemeral plant populations. A number of studies have investigated the level of genetic diversity within species that occupy different successional stages (reviewed in Barrett & Shore, 1989), and significantly lower level of within population diversity have been found in species that inhabit early successional stages, compared to those occupying mature habitats. However, as these studies generally involve different species at different successional stages, it is difficult to extrapolate these results.

When examining the spatial structure of the pooled populations, a significant kinship coefficient was only observed within distance classes of less than 500m, or within populations, with no spatial structure between populations. This indicates a strong restriction to gene flow within populations. A similar pattern in terms of the distances at which isolation by distance occurs in bryophyte populations was also found by Snäll (2004), with a significantly positive kinship coefficient found at

distances between samples of up to 300m. However, in analyses of individual populations, the only population to exhibit a significant kinship coefficient within a distance class was population B at Llanfihangel Gobion. This suggests that populations are not genetically structured within populations, or more accurately, within the pools they occupy. The significant kinship coefficient in population B from 10-50m further corroborates this, as this population was actually divided into two subpopulations, separated by dry land that is not flooded, and by distances greater than 50m. A significant kinship coefficient at distances up to 50m but not in classes of greater distances, suggests that these sub-populations are genetically distinct.

In pooled population spatial autocorrelation analysis, a negative kinship coefficient was found in distances between individuals of greater than 100km, as well as within population B at Llanfihangel Gobion, at distances of 100-200m. A negative kinship coefficient, which suggests that the kinship coefficient is lower than that expected between any pair of individuals picked at random from a sample, seems unlikely, and this finding may be an artefact of the methodology. However, there may be strong isolation by distance, particularly at large distance classes relative to the analysed sample, which could result in negative kinship coefficients relative to the whole sample from all populations. A negative kinship coefficient at distance classes of over 300m was observed by Snäll *et al.*(2004), in an epiphytic bryophyte population, as well as by Bonnin *et al.* (2001) in *Medicago* populations.

Linkage disequilibrium was found in all populations of *P. patens* analysed, which is as expected in a selfing species. The level of linkage disequilibrium found was either within or slightly higher than that found in other monoicous bryophyte species

(Gunnarsson *et al.*, 2005; Snäll *et al.*, 2004), however it was generally less than that of dioicous bryophytes (Grundmann *et al.*, 2007; Grundmann *et al.*, 2008). As the increased level of linkage disequilibrium found in dioicous bryophytes can generally be explained by the clonal nature of these species, the comparatively lower linkage disequilibrium, along with the lack of spatial structure found within populations, suggests that asexual dispersal is not common in *P. patens*.

The level of genetic differentiation between populations of *P. patens*, where most genetic variation is partitioned within, rather than between populations is consistent with levels observed in other bryophyte species (Gunnarsson *et al.*, 2005; Zartman *et al.*, 2006). However, the lack of spatial structure found within populations is more unusual. As *P. patens* is known to readily self fertilise in the laboratory, and as gene flow through gamete dispersal is thought to be severely restricted in bryophytes in general (Shaw, 2000) it may be expected that natural populations may be highly spatially structured with patchy genotypic distribution (Bonnin *et al.*, 2001). However, few studies have investigated spatial autocorrelation analysis in either selfing species or bryophyte populations. Most analyses of fine-scale spatial structure within plant populations have focussed on largely out-crossing species, occupying stable habitats with gene dispersal by seed. As *P. patens* occupies very different habitats and exhibits a very different life history to these well studied species, the trend observed in populations of *P. patens* could be expected to be very different. As the ephemeral habitats that this species occupies can be extremely variable in both size and existence from one year to the next, it could be expected that taxa with such ecology would show little within-population spatial structure. There have been no other investigations into the fine-scale spatial structure within ephemeral plant populations, but Bonnin *et al.* (2001) has used kinship

autocorrelation analysis within populations of an inbreeding *Medicago* species. Here, a high level of spatial structure was only found in the sub-population that was subjected to the least disturbance, and no spatial genetic structure was found in a sub-population that was ploughed regularly. If the populations of *P. patens* within the ponds and lakes that they occupy are periodically submerged throughout the fruiting season, this action could act as disturbance in a similar way. Although gamete dispersal may be restricted in mosses, spores have the potential to be carried in air or water over many metres or kilometres. Thus spore dispersal could account for much of the lack of genetic structure within populations, as alleles could be dispersed almost at random after sufficient rainfall.

This study is the first investigation into the population genetics of *P. patens*, and one of few into the population of ephemeral bryophytes. Populations of *P. patens* have been shown to exhibit genetic structure that is symptomatic of its ephemeral nature, which may provide further avenues for investigation. AFLPs were used successfully in this study to investigate both the level of genetic diversity and the structure of this diversity within populations of *P. patens*. There is a degree of error within this analysis that is not so present in molecular markers such as simple sequence repeats and isozymes, that may cause a low level of spatial structuring within populations to be missed. However, this is negated by the efficiency of AFLPs given that they provide a large number of loci for utilisation. Also the fact that spatial genetic structure was found between pools a few hundred metres apart, but not within the pools, further confirms the role of pool mixing in breaking up spatial genetic structure within populations of *P. patens* when they occupy continuous habitats. The lack of spatial structure found within populations of *P. patens* is consistent with that

found in other species of bryophytes analysed by using AFLP analysis (Snäll *et al.*, 2004), which suggests that the results found in this study are robust.

Chapter 4 *P. readeri* in Europe

4.1 Introduction

The moss genus *Physcomitrella* Bruch & Schimp. currently consists of one species recorded in Europe, *P. patens* (Hedw.) Bruch & Schimp. *P. patens* has a geographical distribution within the northern hemisphere, extending from the U.S.A to Europe (excluding the Mediterranean), Scandinavia and Siberia (Smith, 2004; Tan, 1979). The only other accepted species in the genus, *P. readeri* (C. Müell.) Stone & Scott, has up until now been recorded only from California, U.S.A, Australia, New Zealand and Japan (Dixon, 1926; Goffinet, 2007; Ochi, 1968; Scott & Stone, 1976). The taxonomic status of *P. readeri* is contentious; it has been classified as a subspecies of *P. patens*. Tan (1979) argued for the subdivision of *P. patens* into four subspecies, based mostly on geographical distribution. These were; subsp. *patens*, comprising most plants from the Northern hemisphere, subsp. *readeri* from Australia, subsp. *californica* from the U.S.A and Japan and subsp. *magdalenae* from Rwanda. *P. readeri* is now accepted as a distinct species comprising the following synonyms; *Ephemerella readeri* C. Müell., *Physcomitridium readeri* (C. Müell.). Roth, *P. austro-patens* Broth. ex Roth, *P. patens* ssp. *californica* (Crum & And.) B. C. Tan and *P. californica* Crum & And. (Goffinet, 2007; Stone & Scott, 1973). *P. readeri* is mainly distinguished from *P. patens* by a shorter leaf midrib, or costa, and arguments for its designation as a separate species are partly based on the geographical distribution of plants with this feature (Goffinet, 2007).

Where natural populations of *P. readeri* have been found, they occupy the inundation zones of reservoirs and lakes, on bare, exposed mud after water levels

have fallen, and they have been found with *P. patens* in the U.S.A (Goffinet, 2007). During collection of plants from a population of *P. patens* in West Yorkshire in November 2006, I collected plants with sporophytes that appeared to differ from typical *P. patens*. Further analysis has shown the plants sampled to not be *P. patens*, but *P. readeri*, therefore recording the first population of this species in Europe. The area visited, at Lindley Wood reservoir in the Washburn Valley, was initially visited as *P. patens* had been recorded at the reservoir in 2001 (Hodgetts, N., *pers. comm.*). However, with further analysis of the plants, differences between those collected from Lindley Wood reservoir and other samples of *P. patens* in the collection were discovered.

AFLP profile analysis (see Chapter 3) was conducted on plants cultured from spores from the collected sporophytes. Kinship coefficients were calculated between all plants analysed, including a minimum of 13 plants from each of the 7 populations of *P. patens* in the UK and 14 plants cultured from the sporophytes collected at Lindley Wood reservoir. The AFLP profiles from plants collected in this population were found to differ greatly from the profiles of plants from the other 7 populations of *P. patens* sampled in the UK, as well as accessions from Germany, France, Switzerland and the Ukraine.

Upon culturing the Lindley Wood samples for production of gametophytic tissue and sporophytes, further morphological difference between these plants and *P. patens* became evident. To clarify the identity of the plants from Lindley Wood reservoir, analysis of the internal transcribed sequence (ITS) region of ribosomal DNA was conducted on an individual plant cultured from sporophytes collected at Lindley Wood Reservoir, along with confirmed *P. patens* plants and 4 other species of bryophyte within the Funariaceae. This chapter details the findings of this

analysis, which along with morphological observations, confirm the discovery of a species of bryophyte new to Europe.

4.2 Materials and methods

4.2.1 Population details and sample collection

Sporophytes of *P. readeri* were collected from a population at Lindley Wood reservoir in West Yorkshire, UK in November 2006 (OS grid reference SE210498). For the sampling procedure followed, see Chapter 2. Figure 4.1 shows the location of the population in the UK, and a photograph depicting the site is given in Figure 4.2.

4.2.2 Sample culture

For details of spore culture from intact sporophytes, see Chapter 2. When mature plants had been obtained, individual *P. readeri* plants were cultured on BCD medium (Knight *et al.*, 2002) containing 1mM $CaCl_2$, and grown at 15°C for 7 hours light over a 24 hour period. Reducing the temperature and day length in the growth chamber is routinely used to induce sporophyte production in *P. patens* (Cove, 2005), and this was trialled in *P. readeri*, as sporophytes greatly aid morphological analysis. Individual *P. patens* plants were also cultured as above, to enable a morphological comparison with the plants collected from Lindley Wood reservoir. For ITS sequence analysis, the following plants were cultured for DNA extraction, as detailed in Chapter 2: *P. readeri* from Australia and Japan, *P. patens* from the UK and France, *Physcomitrium pyriforme* (Hedw.) Hampe and *Entosthodon obtusus* (Hedw.) Lindb. Provenance data for these samples can be found in Table 4.1. Two plants collected from the Villersexel population in France were included in this

analysis, as the two lab strains propagated from these plants have been shown to be quite different. One strain, termed "K4" has been shown to be very similar to Gransden, whilst the second strain, "K3" is the most genetically divergent strain of the *P. patens* physcotypes available (see Kamisugi *et al.*, 2008).

Table 4.1 Provenance data for samples used in phylogeny reconstruction

Species	Population/ Location	County/Region	Country
P. readeri	Lindley Wood Reservoir	West Yorkshire	England
P. readeri	Melton Reservoir	Victoria	Australia
P. readeri	Masu pond	Okayama-shi	Japan
P. patens	Gransden	Cambridgeshire	England
P. patens	Villersexel	Franche-Comté	France
P. pyriforme	Carbondale	Illinois	U.S.A
E. obtusus	St Agnes	Cornwall	England

Figure 4.1 Location of *P. readeri* population in West Yorkshire, UK.

Figure 4.2 Panoramic photograph of Lindley Woods Reservoir, where *P. readeri* was found in November 2006.

4.2.3 Nuclear ITS region analysis

4.2.3.1 Amplification and sequencing

DNA was extracted from plant tissue as described in Chapter 2. The ITS1 and ITS2 regions of the nuclear genome were first amplified from each sample using PCR as follows. The primer sequences used to amplify the entire ITS region were; TCGTAACAAGGTTTCCGTAGGTG at the 3' end of the 18S rDNA gene and ATTTTCAAGCTGGGCTCTTTCC at the 5' end of the 26S rDNA gene.

Each PCR reaction (50µl) contained PCR buffer, 0.2mM dNTP mix, 0.3µM each of forward and reverse primer, 1µl Taq polymerase, and 2µl pre-amplification product (maximum concentration 20ng/µl). The PCRs were then carried out with the following conditions over 35 cycles; pre-melt for 2 minutes at 94°C, denaturation at 94°C for 20 seconds, annealing at 58°C for 40 seconds and extension at 72°C for 60 seconds, with a final extension at 72°C for 10 minutes. PCR products were purified using a QIAquick PCR purification kit (QIAGEN Ltd, Crawley, West Sussex U.K). The purified products were than cycle-sequenced by the Integrated Genomic and Gene Expression Analysis Facility at the University of Leeds, with a separate reaction for each of the amplification primers. Products of the cycle-sequencing were run on an ABI 3130 Genetic Analyser (Applied Biosystems, Foster City, U.S.A).

The sequence obtained from the Villersexel "K3" strain was found to be a mixed trace containing more than one sequence. The PCR product produced from this strain was then cloned into the *EcoR*V site of pBluescript by Dr Y. Kamisugi, and two separate sequences were provided for this strain.

4.2.3.2 Sequence alignment and analysis

The sequences were viewed and assembled using the ABI trace viewer in BioEdit, version 7.0.9.0 (copyright© 1999-2005 Tom Hall), with the ITS1 and ITS2 regions treated separately. Successfully amplified regions were first automatically aligned in a matrix using ClustalW version 2 (Ramu *et al.*, 2003). The resulting alignments were then manually modified in the Sequence Alignment Editor in BioEdit (copyright© 1997-2007 Tom Hall) The ITS region sequence from *F. hygrometrica* Hedw. was obtained from GenBank (Benson *et al.*, 2006) and included in the alignment matrix.

After sequence alignment, most parsimonious phylogenetic trees were produced using the phylogenetic analysis package TNT (Goloboff *et al.*, 2008), which has been made freely available by the Willi Hennig Society. This was done by performing a branch-and-bound search using the implicit numeration option in TNT, which guarantees to find all trees. To test support for each clade, bootstrap analysis was performed in TNT, using 1000 replicates.

Maximum likelihood analysis was also conducted, using the free software package TREEFINDER (Jobb, 2008). Optimal substitution models for each ITS region were automatically selected using the model proposer dialog in TREEFINDER. Of the 32 models proposed, an HKY+G model of evolution (Hasegawa *et al.*, 1985) was selected for each ITS region, which was then used to reconstruct phylogenies in TREEFINDER. Bootstrap analysis was performed using 1000 replicates.

4.2.4 Morphologcial analysis

Plants cultured from the sporophytes collected at Lindley Wood reservoir were compared with the *P. patens* plants currently in culture in the laboratory. Scanning Electron Micrographs (SEM) were taken of the Lindley Wood plants using methodology described in Pressel *et al.*, (2007) by J. Duckett. Leaf specimens were photographed using a digital camera under an Olympus SZX12 microscope. Comparisons were made between the two species with regards to sporophyte and leaf morphology.

4.3 Results

4.3.1 ITS sequence analysis

Figures 4.3 and 4.4 show the maximum likelihood phylograms produced using the ITS1 and ITS2 sequences from the 10 samples analysed. The most parsimonious trees are not shown as the topologies obtained were identical using both methods.

Both trees place the sample collected from Lindley Wood reservoir with *P. readeri* from Japan and Australia in separate clades to the *P. patens* plants analysed from the UK and France. Figure 4.5 also depicts the relevant section of sequence alignment between the known *P. readeri* sample from Australia, and the sample from Lindley Wood, produced using BLASTN (Altschul *et al.*, 1990). This figure shows that within these two regions there is a single-base mismatch between these samples, in ITS1. This result is in contrast with the percentage identity between the Lindley Wood sample and the *P. patens* sample from the UK, where the percentage identity is only 86% (averaged over all characters in BLAST alignment). The ITS1 and ITS2 region sequences from the Gransden and Villersexel K4 strains of *P. patens* were

identical, indicating that there is little variation within species. However, the K3 strain differed, with a percentage identity between the K4 strain and the two sequences derived from the K3 strains of 99%. This strain has been shown to be the most genetically distinct when compared to all other samples of *P. patens* in laboratory culture, and has thus been used to produce the *P. patens* genetic linkage map (Kamisugi *et al.*, 2008). These results corroborate the identification of the Lindley Wood plants as *P. readeri*, as opposed to *P. patens*.

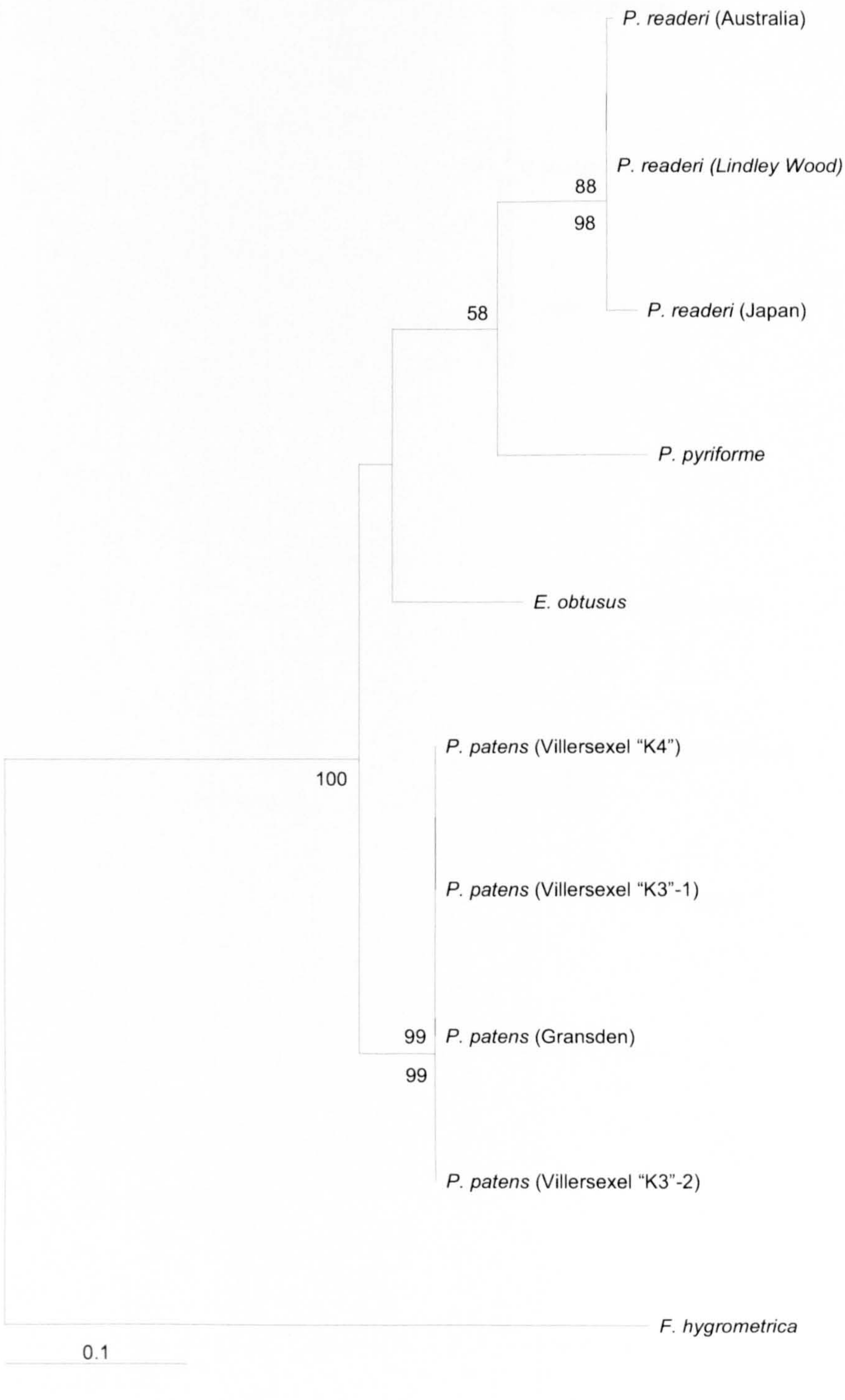


Figure 4.3 ITS1 maximum likelihood phylogram. Likelihood = -1132. Bootstrap values from the maximum likelihood analysis are shown above branches, with bootstrap values from the parsimony analysis shown below (1 tree, 159 steps, C.I. = 0.86 and R.I. = 0.80).

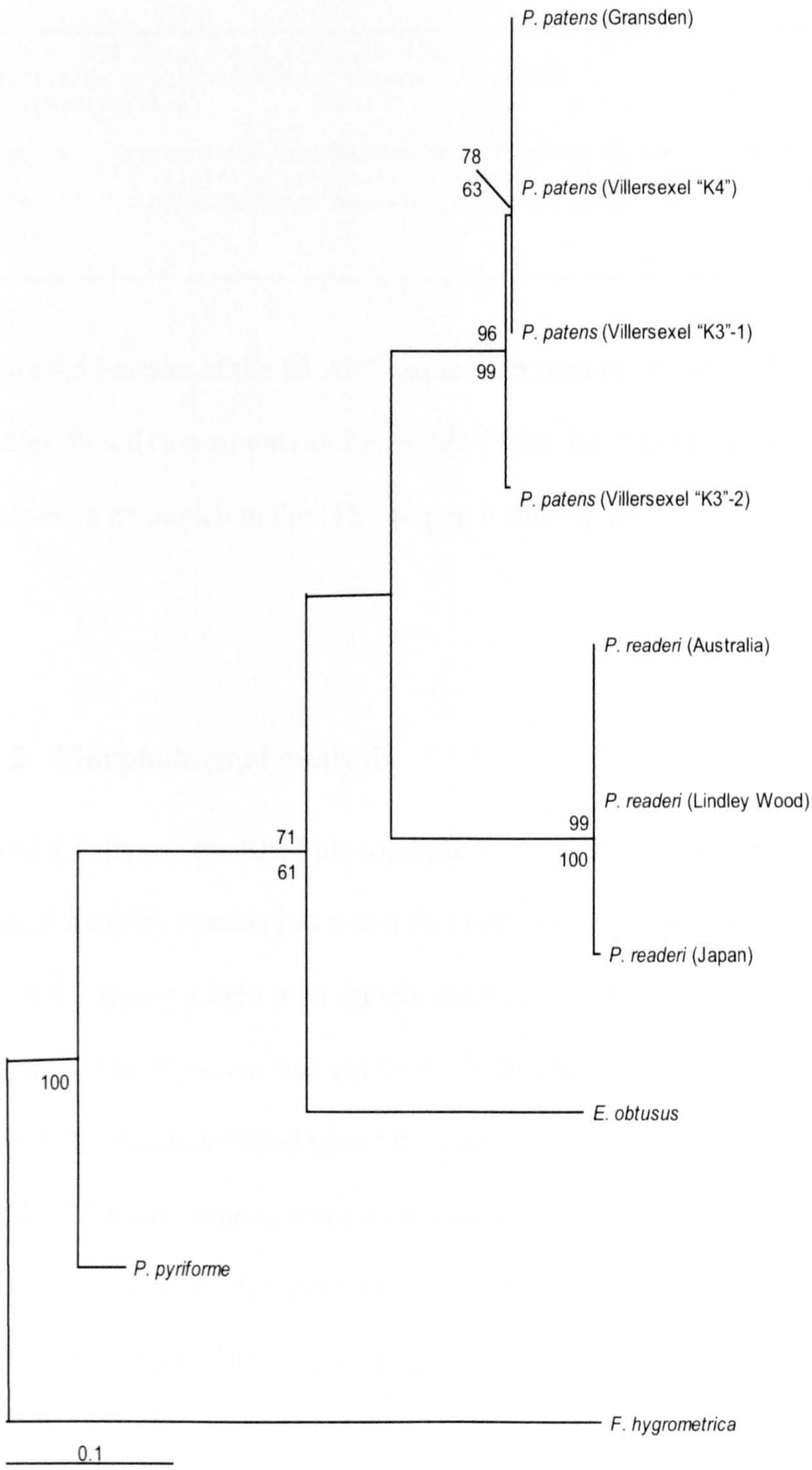


Figure 4.4 ITS2 maximum likelihood phylogram. Likelihood = -1225. Bootstrap values from the maximum likelihood analysis are shown above branches, with bootstrap values from the parsimony analysis shown below. (1 tree, 167 steps, C.I. = 0.89 and R.I. = 0.84)

```
Score =  852 bits (443),  Expect = 0.0
 Identities = 555/556 (99%), Gaps = 0/556 (0%)
 Strand=Plus/Plus

Query  1    CACACACAAAAACTGCAAACAACCCCCTTTTTGAGAACTATATCTTATTTGTACAAAACC  60
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct  1    CACACACAAAAACTGCAAACAACCCCCTTTTTGAGAATTATATCTTATTTGTACAAAACC  60
```

Figure 4.5 Section of the BLAST sequence alignment between *P. readeri* from Lindley Wood (top strand) and *P. readeri* from Australia (bottom strand). The single-base mismatch in the ITS1 region is highlighted.

4.3.2 Morphological analysis

Figure 4.6 depicts the SEM photographs of the Lindley Wood plants. The sporophyte of *P. readeri* has a seta that is much longer than those of *P. patens*. Figure 4.7 shows a light micrograph of a leaf from both a Lindley Wood plant and *P. patens*. The *P. patens* leaf exhibits a well defined costa, extending to the apex, whereas the Lindley Wood plants have a much shorter costa, extending only 2/3 the length of the leaf. This short costa is a diagnostic feature of *P. readeri* (Goffinet, 2007). For a detailed description of *P. readeri* see Appendix III, which is adapted from Scott & Stone (1976) and Smith (2004).

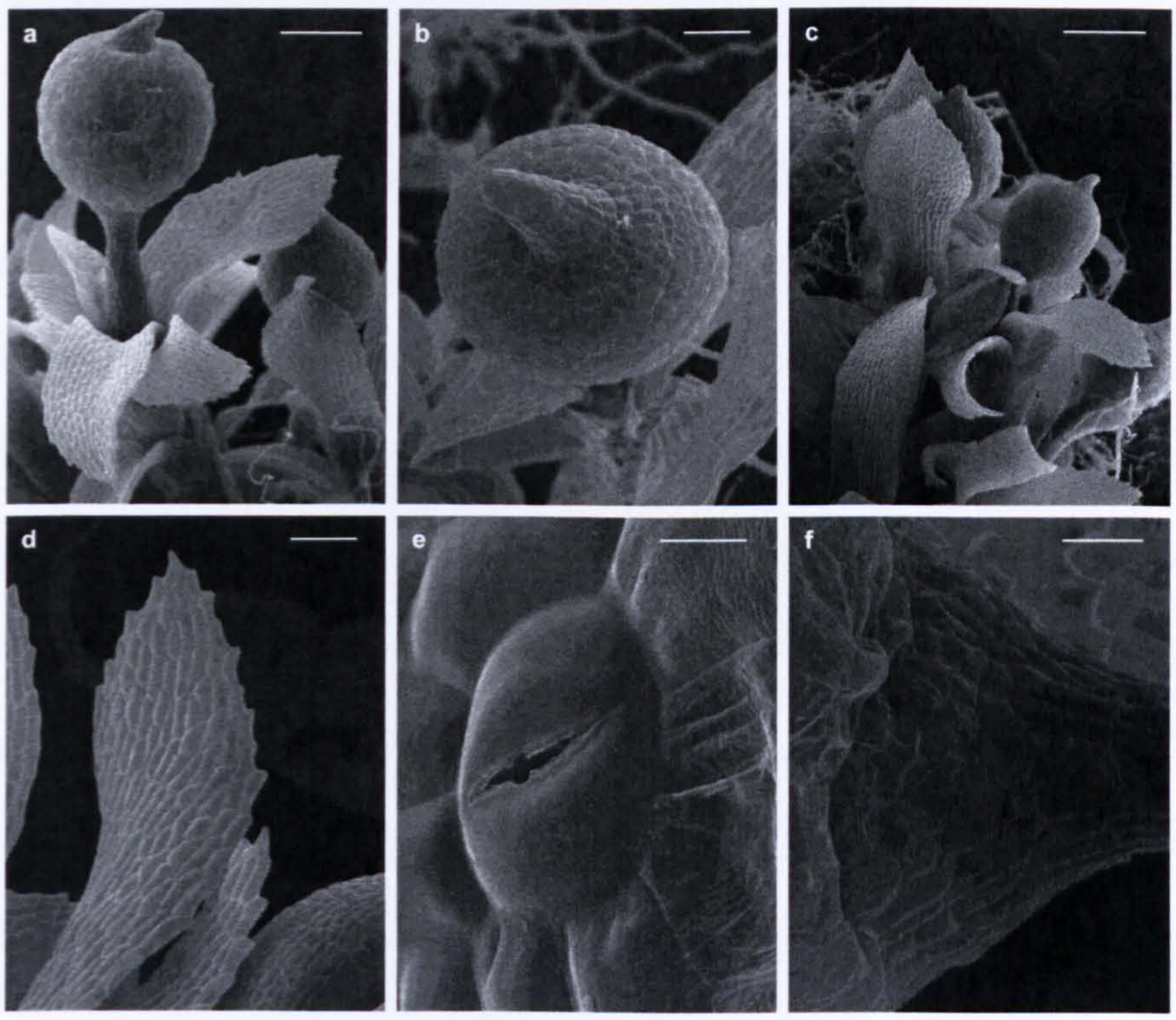


Figure 4.6 a-f. SEM images of *P. readeri* from Lindley Wood Reservoir. a-c) sporophyte, d) leaf, e) stoma f) stomata at base of seta. Scale bars: a, c = 400µm, b, d = 100µm, e = 10µm, f = 30µm.

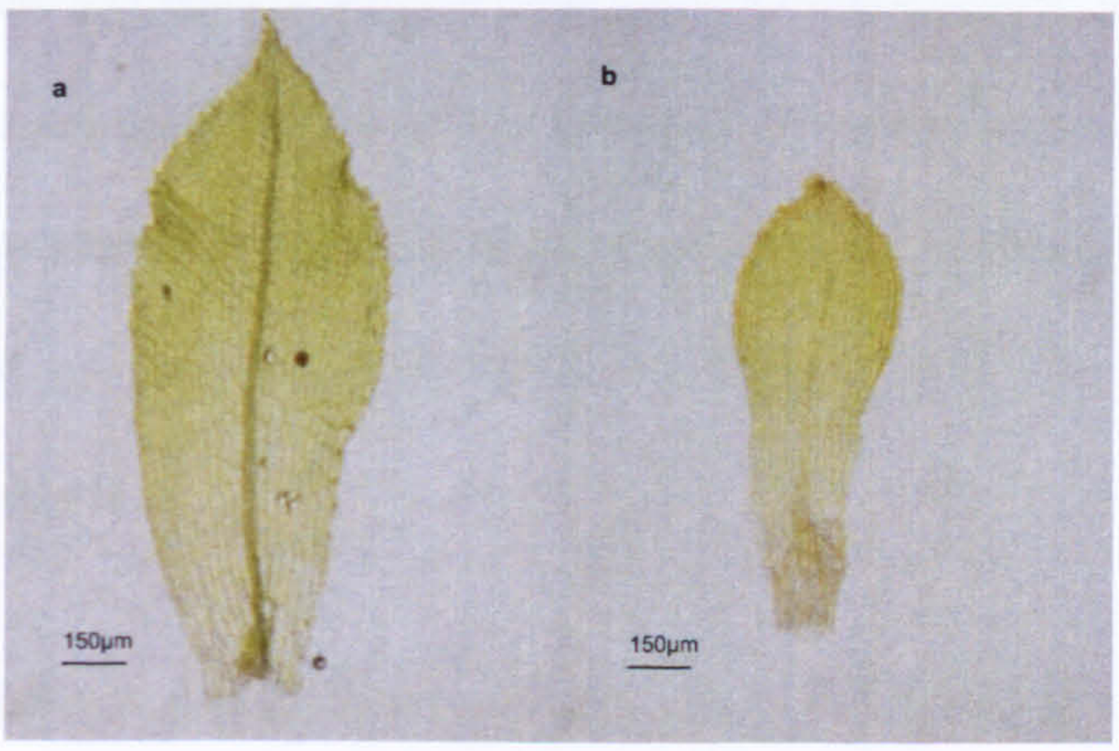


Figure 4.7 a-b. Light micrograph showing entire leaves from a) *P. patens* and b) *P. readeri* from Lindley Wood Reservoir.

ephemeral ecology, where gametophytic and subsequently sporophytic material appears in autumn following the dropping of water levels in aquatic habitats. This winter inundation has been found to be essential for the re-appearance of gametophytes in species of bryophytes and pteridophytes that occupy the same habitats, (Duckett & Duckett, 1980; Hill *et al.*, 1994; Holyoak & Bryan, 2005), and it is likely that this is also the case for *P. readeri*. Many of the bryophytes that occupy reservoir inundation zones also have a very restricted distribution in the UK, for example *Micromitrium tenerum*, *Physcomitrium eurystomum*, *Physcomitrium sphaericum* and species of *Ephemerum* have been recorded from only a few locations (Hill *et al.*, 1994; Holyoak & Bryan, 2005). It is therefore likely, as with the fore-mentioned taxa, that *P. readeri* is native to the UK, but under-recorded. As *P. readeri* occupies habitats that are for large periods of time submerged and in-accessible, it is likely that populations of this species have been excluded from species surveys. Whilst looking for populations of *P. patens* in autumn 2006, many habitats that may also have been suitable for *P. readeri* were visited, yet no other populations of this species were discovered. It is therefore likely that *P. readeri* is rare in Britain.

Although the population of *P. readeri* at Lindley Wood reservoir was the only population found during the year, it was a particularly large population, and it occupied the entire extent of bare mud in the reservoir, which covered an area of 65,581m^2. Further discoveries of *P. readeri* may therefore be expected with greater exploration of reservoirs and lakes during the summer months. It is likely to be found in other reservoirs nearby in the Washburn Valley, such as Swinsty and Fewston Reservoirs, but water levels in these reservoirs were very high in autumn 2006, which may account for the absence of *P. readeri* during this period. Like other

rare bryophytes occupying similar habitats, *P. readeri* is especially vulnerable, particularly if water management strategies keep water levels high, which would result in the disappearance of this species.

Chapter 5 A method for estimation of the mating systems operating within populations of *P. patens*

5.1 Introduction

As stated in previous chapters, there have been relatively few studies into the molecular genetics of bryophyte populations, and among these the investigation of bryophyte mating systems has been particularly neglected. However, the mating system of any species, especially that operating within natural populations, has important implications for the evolutionary ecology of these populations (Barrett & Eckert, 1990; Seymour *et al.*, 2005).

In diploid species, Wright's fixation index (F_{is}) has commonly been used to estimate the levels of selfing within a population. This approach has also been used to estimate the out-crossing rate exhibited in the diploid stage in bryophytes (Eppley *et al.*, 2007). Eppley (2007) estimated the rate of selfing in a number of populations of both dioicous and monoicous moss species, by calculating inbreeding coefficients using allozyme markers. In this study, the authors not only characterised "selfing" as the mating of gametes produced from the same individual (intragametophytic selfing), but also as the mating of gametes produced from haploid progeny from the same diploid parent (intergametophytic selfing). The authors found that selfing rates were far higher in monoicous species than in dioicous species, but that there was evidence of mixed mating systems in some populations of dioicous species, with greater than expected levels of intergametophytic selfing. Eppley *et al.* (2007) concluded that mixed mating occurs in populations of both monoicous and dioicous bryophytes, suggesting that further studies into the mating systems of bryophytes are

necessary, as the level of inbreeding and out-crossing in different species is largely unknown. However, the methods used in this study are not as applicable to bryophytes as to seed plants, particularly where genetic analysis is attempted as diploid sporophyte tissue is not readily cultured.

Innes (1990) studied the level of genotypic diversity within populations of the dioicous moss *Polytrichum juniperinum*, and confirmed that sexual reproduction was present. The author also utilised both the haploid gametophyte and the diploid sporophyte to further study the mating systems within these populations. In this investigation the sporophytic tissue from plants sampled in the populations studied were genotyped using isozymes, along with the attached maternal gametophytic tissue and the male gametophytic tissue from within the populations. As the genotype of the maternal parent is easily identified as it is attached to the sporophyte, the genotype of the male parent was inferred from the residual loci. Using this methodology Innes (1990) was able to determine that female gametophytes were usually fertilised by male gametophytes that were close by, and at most within a few metres of the female plant. Roads & Longton (2003) also compared the genotypes of maternal gametophytes to their attached sporophytes from populations of the monoicous mosses *Phascum cuspidatum* and *Pottia truncata*. No variation was found between sporophytes and maternal gametophytes, suggesting a high degree of inbreeding in these populations.

The predominance of the haploid stage in bryophytes can assist in identifying the breeding systems found in bryophyte populations, as haploid siblings can easily be isolated from the diploid sporophyte. Differing degrees of inbreeding within a population will result in differing numbers of sporophytes that contain either homozygous or heterozygous haploid siblings, with a true inbreeding event

producing homozygous sporophytes and genetically identical gametophytes (Shaw, 2000). This approach has been utilised to study the mating systems within natural populations of fungi (e.g. Ennos & Swales, 1987; Free *et al.*, 1996; Marra *et al.*, 2004; Milgroom *et al.*, 1993; Milgroom *et al.*, 1992) and lichens (e.g. Murtagh *et al.*, 2000; Seymour *et al.*, 2005).

The purpose of this study was to investigate a method for estimating the mating systems that are present in natural populations of *P. patens*. As *P. patens* is a monoicous bryophyte (Reski, 1998), with both archegonia and antheridia on the same gametophyte, there is the possibility that selfing can occur, and *P. patens* has been shown to produce homozygous offspring due to selfing (Cove, 1983). This was done by first analysing the genotypes of the gametophytes released from a sporophyte that was the result of a known cross in the laboratory. The same methodology was then applied to natural populations of *P. patens*, to determine the level of out-crossing or inbreeding within these populations.

As gamete dispersal has been found to be locally restricted within most bryophyte populations, with estimates suggesting that where cross-fertilisation does occur that it is between plants separated by distances of on average no more than 10cm (Longton, 1976), it can be assumed that plant density would have a pronounced affect on the probability of cross fertilization. Studies investigating the effect of density on out-crossing rates in angiosperms have found that increased density both increased (Wolff *et al.*, 1988) as well as decreased (Ellstrand *et al.*, 1978) out-crossing rate. The effect of density was therefore also investigated in this study.

5.2 Materials and methods

5.2.1 Trial using known progeny

To determine if the methodology to be used is effective at identifying progeny that are the result of an out-crossing event, the method was trialled using the progeny of a known cross. The gametophytes analysed in this trial were the progeny of a cross between two laboratory strains of *P. patens*: the standard laboratory strain, "Gransden" and a genotypically distinct strain "Villersexel K3". These progeny have been used to produce a genetic linkage map of *P. patens* (see Kamisugi *et al.*, 2008 for details of the crossing procedure), and they will from henceforth be referred to as the mapping population. A sample of 22 gametophytes from this population was randomly selected, and the following protocol was used. Individual plants were cultured from laboratory stocks as described in Chapter 2.

The methodology to be trialled involved pooling progeny to be genotyped, and identifying the presence of heterozygous individuals within that pool by comparing the pooled AFLP profile to that of a reference individual from the same sporophyte. Using this methodology, any additional loci not found in the reference individual immediately identifies that sporophyte as the result of a cross, or a "discernible" mating, as described by Shaw *et al.*, (1981). Where profiles of the pooled sample and reference sample are found to be identical, this is termed a non-discernible mating, as the sporophyte could be the product of either a cross with a genetically identical individual, or a self fertilisation event.

A similar methodology has been trialled by Marra *et al.*, (2004), where the authors were investigating the mating systems within populations of the chestnut blight fungus, *Cryphonectria parasitica*. In this study the ascospores from a tree canker

were pooled, with the DNA fingerprint of the pooled ascospores compared to the maternal fingerprint, to determine if the progeny segregate at the DNA fingerprint loci. The authors found this method to be successful, and there is therefore potential for this methodology to be applied to bryophyte systems, with a great reduction in the experimental time and expense that is involved with genotyping progeny individually.

5.2.1.1 DNA extraction and AFLP analysis of progeny

The 22 plants selected were divided into two replicate groups of 11 individuals, with 10 plants pooled and one reference plant in each group. 10 plants were estimated to be a sufficient number of individuals to analyse in order to identify segregating alleles, as sporophytes that are the result of an out-crossing event will comprise progeny with a genotype ratio of 1:1. Therefore the probability of identifying any one segregating allele from 10 individuals will be $1-(0.5)^{10} = 0.999$. DNA was extracted from the pooled plants and the reference individual as described in Chapter 2.

AFLP analysis was conducted on each of the pooled samples and reference samples, with each group analysed in a separate AFLP reaction. AFLP analysis was conducted as described in Chapter 2, using the *EcoR*I+AC/*Mse*I+CG FAM labelled primer combination. This primer combination was chosen as it displayed the lowest error rate of the three combinations, at 1.92%, as calculated in Chapter 3. AFLP analysis was also conducted on the parental strains, Gransden and Villersexel, so that the source of any segregating alleles could be identified.

5.2.2 AFLP profile scoring

AFLP profiles from the pooled and reference samples were visualised using Genemapper® Software v3.7 (Applied Biosystems). The AFLP profiles of each pooled sample were compared to the reference samples. Polymorphic loci were scored as those present in the pooled sample but absent from the reference, indicating a segregating allele within the sporophyte. Only unambiguous fragments that were between 70 and 500 bp in length were included in the analysis.

5.3 Analysis of natural populations

To estimate the mating systems present in natural populations of *P. patens*, the method above was applied to 56 sporophytes from three populations. The populations were selected as they encompassed contrasting areas of both high and low density *P. patens* plants. The populations chosen were; Llanfihangel Gobion (Population B), Papercourt Marshes (Population D) and Ouse Washes (Population E), as detailed in Chapter 2. In Llanfihangel Gobion and Papercourt Marshes 10 sporophytes were selected randomly from within both the high and low density areas. From Ouse Washes only 8 sporophytes were available from each area, so all were used.

The AFLP genotype analysis was conducted as above, with 10 gametophytes per sporophyte cultured and pooled, and a single reference gametophyte cultured. DNA extraction and AFLP analysis was conducted as above.

5.4 Results

5.5 Trial using known progeny

Loci with segregating alleles were identified in both replicates of pooled gametophytes from the mapping population. From a total of 148 scorable loci, 6 polymorphic loci were identified when the pooled sample was compared to the reference samples, and both replicates of pooled gametophytes exhibited the same polymorphisms. A screen shot depicting an example of these polymorphic loci is given in Figure 5.1, with screen shots depicting all polymorphic loci given in Appendix IV. The presence of the alleles in the AFLP profile of either Gransden or Villersexel, *i.e.* the source of these segregating alleles, is detailed in Table 5.1.

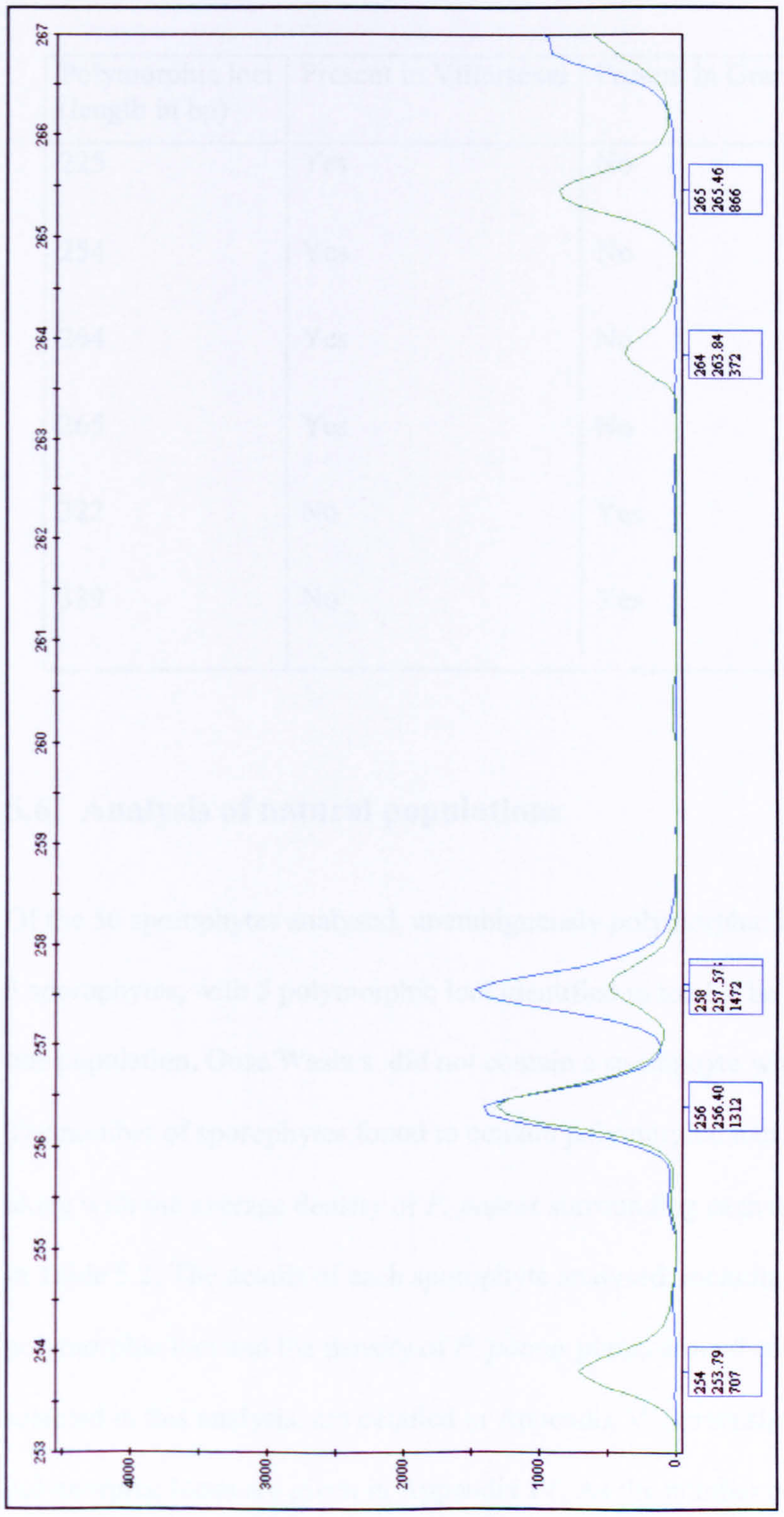


Figure 5.1 Screenshot depicting an example of polymorphisms between the pooled and reference gametophytes produced from a cross between the Villersexel and Gransden strains of *P. patens*. Polymorphic loci are named using the size of the AFLP fragments. Villersexel cross. Three polymorphic loci correspond to AFLP fragments of size 254, 264 and 265 bp.

Table 5.1 Table detailing the polymorphic loci observed within a sporophyte produced from a cross between the Villersexel and Gransden strains of *P. patens*. Polymorphic loci are named using the size of the AFLP fragment.

Polymorphic loci (length in bp)	Present in Villersexel	Present in Gransden
225	Yes	No
254	Yes	No
264	Yes	No
265	Yes	No
322	No	Yes
339	No	Yes

5.6 Analysis of natural populations

Of the 56 sporophytes analysed, unambiguously polymorphic loci were identified in 3 sporophytes, with 5 polymorphic loci identified in total. The sample derived from one population, Ouse Washes, did not contain a sporophyte with polymorphic loci. The number of sporophytes found to contain polymorphic loci in each population, along with the average density of *P. patens* surrounding each sporophyte, is shown in Table 5.2. The details of each sporophyte analysed, including the number of polymorphic loci and the density of *P. patens* plants around each sporophyte selected in this analysis, are detailed in Appendix V. Screenshots showing each polymorphic locus are given in Appendix IV. As the number of sporophytes

containing polymorphic loci was so low, any effect of density could not be measured.

Table 5.2 The number of sporophytes collected in each of the three populations that contained segregating loci, and the density of *P. patens* plants surrounding the sporophytes sampled.

Population/Density designation	Number of sporophytes with no segregation	Number of sporophytes with polymorphic loci	Number of polymorphic loci	Average density (fraction fill) 1x1m	Average density (fraction fill) 10x10cm
Papercourt Marshes/High[1]	10	0		0.328	0.565
Papercourt Marshes/Low[2]	9	1	1	0.155	0.151
Ouse Washes/High[1]	10	0		0.176	0.374
Ouse Washes/Low[2]	10	0		0.034	0.172
Llanfihangel Gobion/High[1]	10	0		0.177	0.273
Llanfihangel Gobion/Low[2]	8	2	4	0.045	0.155

[1]Sporophytes within a high density area
[2]Sporophytes within a low density area

5.7 Discussion

The results of the trial using a known cross between two strains of *P. patens* indicate that this methodology can be used to identify sporophytes that are the product of an out-crossing event. The methodology has also been found to be repeatable, as identical polymorphisms were found between both the replicates analysed and the reference samples. When natural populations were analysed, a low level of variation within the sporophytes sampled was observed, with only three sporophytes found to contain polymorphic loci. However, this does indicate that out-crossing does occur in natural populations of *P. patens*, and this suggests that mixed-mating systems operate within them.

The sample collected from the Ouse Washes population contained sporophytes with no polymorphic loci. A low level of sporophytes containing polymorphic loci would be expected in populations of *P. patens*, as self-fertilisation is known to occur with high frequency under laboratory conditions. In selfing bryophytes, not only can true inbreeding occur, but a cross between siblings from the same sporophyte (intergametophytic selfing) would likely be frequent. This has the potential to greatly reduce heterozygosity within a sporophyte, even if no true inbreeding occurs (Hedrick, 1987a; Hedrick, 1987b).

A homozygous, inbreeding population, such as is probable in Ouse Washes, maybe stable in bryophyte species such as *P. patens*. Taylor *et al.* (2007) studied the effects of inbreeding on fitness using the dioicous moss *Ceratodon purpureus* and the monoicous moss *F. hygrometrica*. By conducting experimental crosses of these two species, the authors found that inbreeding did have an effect on the fitness of *C. purpureus* but not *F. hygrometrica*. In some species, inbreeding may even be

advantageous, by purging the population of deleterious recessive alleles (Lande & Schemske, 1985). *F. hygrometrica*, like *P. patens*, is a weedy bryophyte in the Funariaceae, which colonises disturbed environments (Taylor *et al.*, 2007). Due to the nature of the habitats that both these species occupy, selfing may be favoured so as to allow re-colonisation. Intragametophytic selfing would then purge the population of deleterious alleles in one generation (Hedrick, 1987a). Ephemeral populations would be expected to experience genetic bottlenecks, which would also reduce deleterious alleles in the population (Taylor *et al.*, 2007; Kirkpatrick & Jarne 2000).

Although no relationship between density and the mating system in populations of *P. patens* could be ascertained in this study, future studies into the breeding systems of bryophytes and plants in general should aim to investigate the impact of both environmental and genetic influences. Lande & Schemske (1985), characterised plants species as either predominantly selfing or predominantly out-crossing. However, plant species show great variability in out-crossing rates both within and between populations (reviewed in Walter, 1986), so species cannot easily be classed as either selfing or out-crossing. Many different factors, including genetic factors such as flower size and colour, and ecological and environmental factors such as density, mode of pollination and altitude, have been shown to have a significant impact on the mating systems within populations (reviewed in Barrett & Eckert, 1990). Therefore care should be taken over extrapolating findings from single population studies.

Wolff *et al.* (1988), found that the timing of analysis had a significant impact on the out-crossing rates observed in a population of *Plantago coronopus*. The authors sampled over a period of 6 weeks, and they found a significant decrease in out-

crossing rates in the plants sampled later in the study, mainly due to a decrease in flowering, which is analogous to a decrease in density. Therefore in bryophytes, particular note should be taken of the growth stage of the plants in the population, and in particular the presence of mature antheridia and archegonia. In practice this could be very time consuming, particularly when taking into account the number of populations and repeated sampling that would be required to investigate the effect of this on out-crossing rates. In a study such as this, a more readily recordable variable maybe the number of mature sporophytes in the focal area. Studies investigating how ecological and environmental factors impact on mating systems in natural populations have the potential to be extremely complex (Barrett & Eckert, 1990). Therefore, *P. patens*, as a well established experimental system, has the potential to assist with advancement in this area through laboratory based studies and controlled field experiments.

The low genetic diversity levels within populations of *P. patens* may pose problems for further studies into the mating systems of this species, as the number of polymorphic loci that can be observed is low. As AFLPs provide a large number of loci for scrutiny, these markers are most appropriate for such an investigation. The methodology used in this study could be most applicable in an initial screen to identify populations of interest, as pooling a number of gametophytic progeny from a sporophyte before conducting AFLP analysis is much quicker and cheaper than genotyping the progeny individually. As the number of populations with either no within-sporophyte polymorphic loci or low levels is likely to be high, this method could be used to identify only the populations where mixed mating can be observed, which maybe of evolutionary interest. Full genotyping could then be conducted on individual progeny from sporophytes from these populations.

Chapter 6 General Discussion

Model species provide ideal systems in which to explore evolutionary questions, as there are often numerous molecular resources available for the study of such species, along with established laboratories and culture techniques. General evolutionary questions can therefore be answered with greater efficiency in model species than in other organisms. It is however important for a wide variety of model species to be available for such studies, with the diversity of natural systems fully represented, in order to extrapolate any findings (Travis, 2006). The starting point in any evolutionary genetic study using any system must be to investigate the level of genetic variation within that system, along with an estimation of how this is distributed. The level of diversity found within populations of a species reflects its past ability to respond to changing environments and evolutionary pressures, and likewise it will also determine the ability of that species to respond to these changes in the future. An estimation of how gene flow operates within natural populations, and therefore how this diversity is structured spatially, will determine the way in which populations will respond (Jørgensen & Mauricio, 2004). Therefore, if aspects of a species' biology that are of evolutionary interest are to be studied, the genetic processes that affect these aspects must be determined.

6.1 Genetic diversity and spatial genetic structure

I have investigated the levels of diversity and spatial genetic structure in natural populations of *P. patens* in the UK. There was no correlation between the level of diversity in the population and size of the area the population occupies, which contradicts population genetic theory. However, these theories have largely been developed for organisms with very different life histories to *P. patens*. As *P. patens* occupies ephemeral habitats, where the size of the population area is likely to alter from one year to the next, the size of the populations at the time they were sampled may not be representative of historical populations sizes. Therefore, an extension of this study could be to measure both population area and genetic diversity levels within these populations in future years, as levels of diversity have been shown to correlate with the harmonic mean of population size (Barrett & Husband, 1997).

No spatial genetic structure was found within populations of *P. patens*, whereas theory suggests that populations of bryophytes such as *P. patens* should exhibit a degree of spatial structure due to an expected degree of inbreeding and restricted gamete dispersal in bryophytes (Anderson & Lemmon, 1974; Longton, 1976). The lack of spatial structure found could be due to limitations of the AFLP technique, where the error rate observed could restrict the resolution of the genetic differentiation that this marker can detect. However, a similar lack of spatial structure has been found in other bryophyte species (e.g. Snäll *et al.*, 2004) and this finding could be related to the ecology of *P. patens*. If gene flow in populations of *P. patens* is primarily mediated by spores, then periodically rising water levels in the habitats that *P. patens* occupies could disperse spores over a great distance, leading to a lack of spatial structure. This may also account for the finding that genetic diversity is partitioned within rather than between populations, as if *P. patens*

populations are connected to a degree by waterways, gene flow between populations could be common. As very little is known about the dispersal of spores in bryophytes, or the regeneration of bryophyte populations from spores versus gametophytic tissue, then this aspect of the ecology of *P. patens* could provide interesting avenues for research (Innes, 1990; Thingsgaard, 2001).

One limitation of this study was the sample sizes available for within-population spatial autocorrelation analysis. When all the samples analysed were pooled, this provided a large data set on which to conduct spatial autocorrelation analysis. However, to derive conclusions about the population structure in individual populations is less robust than that with the pooled sample, given the sample sizes used. After conducting this study I would ideally have re-sampled from a single population in 2007, to enable a within-population study with a larger sample size. However, extremely high rainfall in summer 2007 made this impossible as water levels in the populations did not recede.

6.2 Mating systems in *P. patens*

The mating systems operating within three populations of *P. patens* have been investigated, and a novel procedure for estimation of this proposed. This procedure was found to be proficient in identifying progeny that are the result of an out-crossing event, when tested using progeny from a known cross. Haploid species such as bryophytes provide convenient models in which to investigate plant mating systems, as the haploid progeny from diploid sporophytes can be screened for segregating alleles, thus identifying sporophytes that are the result of an out-crossing event. Although the progeny from only three *P. patens* sporophytes were found to contain segregating alleles, this indicates that there is at least a degree of out-

crossing in populations of *P. patens*. It is likely that more than three sporophytes analysed in this chapter were the result of out-crossing events, as only those crosses between genetically different individuals as determined by the AFLP loci can be detected.

As little is known of the mating systems in bryophytes, *P. patens* can be used as a model to investigate the levels of inbreeding and out-crossing in bryophytes given differing environmental conditions or population characteristics, as the findings of this study suggest that mixed-mating systems operate in populations of *P. patens*. The procedure trialled provides a relatively inexpensive and quick method of conducting initial analysis and identifying populations of interest, before proceeding with more in depth analysis involving screening individual progeny. As *P. patens* can easily be cultured in the laboratory, artificial populations of *P. patens* could be set up and mating system investigations conducted. For example, gamete dispersal distance could be estimated by setting up populations of *P. patens* with plants of known genotype at known spatial positions, and then screening the progeny of any sporophytes produced for segregating alleles at target loci.

6.3 *P. readeri* in Europe

Possibly the most unexpected finding in this project was the discovery of a moss new to the bryophyte flora of Europe, with *P. readeri* found in a reservoir in West Yorkshire. The subsequent molecular analysis of *P. readeri* and *P. patens* indicates that these species should possibly be placed in separate genera, but further analysis using a representative set of species from the Funariaceae is required to confirm this.

The finding of *P. readeri* in West Yorkshire highlights how little is known about the distribution of many bryophytes within the UK. Species that occupy habitats such as *P. readeri* and *P. patens* are particularly vulnerable to being missed in surveys and thus under-recorded. Whilst looking for populations of *P. patens* in the UK, I have found that populations cannot always be found in the same locality that they have been found previously, primarily due to the water levels in the habitats where they are found. It is likely that this is partly the reason why little is known of the ecology and evolution of the plants that occupy such habitats, and why many are classed as rare. Thus *P. patens* offers a potential model for investigations into the ecology and evolution of other ephemeral bryophytes.

6.4 Combining molecular genetics and ecology

This study has been the first investigation into the population genetics of *P. patens*. AFLP analysis was used successfully, with allele frequencies identified for a number of polymorphic loci. The AFLP primer combinations that were used in this study were also used to construct the genetic linkage map of *P. patens* (Kamisugi *et al.*, 2008). Further work could involve scrutiny of those loci that were found to be polymorphic in natural populations, which also correspond to loci with identical characteristics (*i.e.* primer combination used and fragment length) that are polymorphic loci in the mapping population. The polymorphic AFLP loci that were used to produce the linkage map correspond to DNA fragments of known length, but not known DNA sequence. However, the DNA sequence of each locus used has been predicted by analysis of the sequenced *P. patens* genome by Kamisugi *et al.*(2008) using the SCALPHunter software. Therefore, the DNA sequence corresponding to the AFLP loci used in this study can also be predicted with possible inference of gene function based on sequence homology to known genes. An example of such an exercise is described below.

For all polymorphic loci that were characterised in Chapter 3 of this project, candidates were searched for in the *P. patens* linkage map. From those loci identified as having identical characteristics to a locus on the linkage map, the most frequently polymorphic locus, with an allele frequency among the pooled samples of 0.33, was locus EAGMCT229 (*EcoR*I+AG, *Mse*I+CT, 229 base pairs in length). An electropherogram depicting this polymorphic locus is given in Appendix VII. SCALPHunter identified this locus as a candidate sequence on sequence Scaffold 38 in linkage group 1 of the *P. patens* linkage map. Using this predicted sequence, a BLAST search of the *P. patens* genome was conducted. From this search a predicted

gene model was identified using the *Physcomitrella* genome browser (http://genome.jgi-psf.org/Phypa1_1/Phypa1_1.home.html). Inspection of this gene model suggested that the computationally defined model was incorrect, in that the existing EST clones showed additional exons. Sequence alignment of the predicted gene model and the possible homologous ESTs was conducted, with a new gene model predicted from cDNAs in the *P. patens* EST database. This gene model was found to contain a putative protein coding sequence. This polypeptide is a member of a small (2 genes) multigene family in *P. patens*. This family of proteins is highly conserved in evolution, and encodes a "lipid raft" protein involved in membrane structure and modification.

To identify genes that maybe of evolutionary significance in *P. patens*, further sampling of a wide variety of populations from differing environments and habitat types is required. If this is achieved, possible future work that combines what is known of the molecular genetics of *P. patens* with its ecology could be possible. If a large number of plants are sampled within many populations, analyses with molecular markers such as AFLPs can be used to more accurately identify interesting loci (Luikart *et al.*, 2003). With a large number of both samples and loci, "outlier loci" can be identified. Outlier loci are loci that exhibit behaviour that deviates from behaviour predicted in simulations (Luikart *et al.*, 2003). For example, these outlier loci could be identified based on simulations of expected F_{st} values (e.g. in Wilding *et al.*, 2001). In an inbreeding species such as *P. patens*, where the level of diversity, and thus the number of polymorphic loci, within a population is low, a great many loci would be required in order to identify outlier loci. Such a study would best be conducted on a population previously identified as containing a high level of genetic variation. Based on current knowledge of diversity within

populations of *P. patens*, the Villersexel population in France may be a suitable candidate for such a study. At present only two accessions have been collected from this location, however these accessions have been shown to be highly divergent based on AFLP and ITS sequence analysis (see Chapter 4 and Kamisugi *et al.*, 2008). Further analysis into the genetics of the population at Villersexel may therefore provide a diversity of evolutionarily interesting genes. As two ITS sequence traces were found in the K3 strain of Villersexel, suggesting hybridisation has taken place, this could potentially provide a means of studying hybridisation and introgression in *P. patens* populations.

P. patens has the potential to be utilised as a model for both molecular genetic and ecological and evolutionary studies, as is the case with other model organisms such as *A. thaliana* and *D. melanogaster*. The study of natural variation within *P. patens* combined with molecular genetic investigations may provide an insight into the early evolution of land plants. However, *P. patens* is not only useful as a model bryophyte, as considering the ephemeral habitats that it occupies, it also represents a group of ecologically interesting plants. Little is known of the population genetics of both ephemeral plants and bryophytes, yet only by studying the full breadth of plant diversity can we fully describe the evolutionary processes that occur within natural populations.

Literature cited

Addinsoft (2008) XLSTAT Pro version 2008.7.03 available at http://www.xlstat.com.

Alonso-Blanco C, Koornneef M (2000) Naturally occuring variation in *Arabidopsis*: An underexploited resource for plant genetics. *Trends in Plant Science* **5**, 22-29.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Anderson LE, Lemmon BE (1974) Gene flow distances in the moss *Weissia controversa* Hedw. *Journal of the Hattori Botanical Laboratory* **38**, 67-90.

Barrett SCH, Eckert CG (1990) Variation and evolution of mating systems in seed plants. In: *Biological Approaches and Evolutionary Trends in Plants* (ed. Kawano S), pp. 229-254. Academic Press, London.

Barrett SCH, Husband BC (1997) Ecology and genetics of ephemeral plant populations: *Eichhornia paniculata* (Pontederiaceae) in northeast Brazil. *Journal of Heredity* **88**, 277-284.

Barrett SCH, Shore JS (1989) Isozyme variation in colonizing plants. In: *Isozymes in plant biology* (eds. Soltls DE, Soltls PS), pp. 106-126. Dloscorldes Press, Portland, Oregon.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2006) GenBank. *Nucleic Acids Research* **34**, D16-D20.

Berres ME (2001) General fluorescent AFLP (fAFLP) protocols. University of Wisconsin Department of Zoology, Madison, Wisconsin. Available at http://ravel.zoology.wisc.edu/sgaap/AFLP_html/fAFLP_protocols.htm. Accessed 02/03/2006.

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

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

Bonnin I, Ronfort J, Wozniak F, Olivieri I (2001) Spatial effects and rare outcrossing events in *Medicago truncatula* (Fabaceae). *Molecular Ecology* **10**, 1371-1383.

Bouck A, Vision T (2007) The molecular ecologist's guide to expressed sequence tags. *Molecular Ecology* **16**, 907-924.

Bruch P, Schimper WP, Gümbel T (1836-1855) *Bryologia Europaea seu genera Muscorum Europaeorum monographice illustrata* E. Scweizerbart, Stuttgart.

Bryan VS (1957) Cytotaxonomic studies in the Ephemeraceae and Funariaceae. *The Bryologist* **60**, 103-126.

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

Courtice GRM, Ashton NW, Cove DJ (1978) Evidence for the restricted passage of metabolites into the sporophyte of the moss *Physcomitrella patens* (Hedw.) Br. Eur. . *Journal of Bryology* **10**, 191-198.

Cove D (1983) Genetics of bryophyta. In: *New Manual of Bryology* (ed. Schuster RM), pp. 222-231. Hattori Botanical Laboratory, Nichinan.

Cove D (2005) The moss *Physcomitrella patens*. *Annual Review of Genetics* **39**, 339-358.

Cove DJ, Knight CD (1993) The moss *Physcomitrella patens*, a model system with potential for the study of plant reproduction. *The Plant Cell* **5**, 1483-1488.

Cove DJ, Knight CD, Lamparter T (1997) Mosses as model systems. *Trends in Plant Science* **2**, 99-105.

Cox CJ, Goffinet B, Newton AE, Shaw AJ, Hedderson TAJ (2000) Phylogenetic relationships among the diplolepideous-alternate mosses (Bryidae) inferred from nuclear and chloroplast DNA sequences. *The Bryologist* **103**, 224-241.

Cronberg N (2002) Colonization dynamics of the clonal moss *Hylocomium splendens* on islands in a Baltic land uplift area: Reproduction, genet distribution and genetic variation. *Journal of Ecology* **90**, 925-935.

Cronberg N, Molau U, Sonesson M (1997) Genetic variation in the clonal bryophyte *Hylocomium splendens* at hierarchical geographical scales in Scandinavia. *Heredity* **78**, 293-301.

Cronberg N, Rydgren K, Økland RH (2006) Clonal structure and genet-level sex ratios suggest different roles of vegetative and sexual reproduction in the clonal moss *Hylocomium splendens*. *Ecography* **29**, 95-103.

Cruzan MB (1998) Genetic markers in plant evolutionary ecology. *Ecology* **79**, 400-412.

Cummins H, Wyatt R (1981) Genetic variability in natural populations of the moss *Atrichum angustatum*. *The Bryologist* **84**, 30-38.

Dixon HN (1926) Funariaceae. In: *Studies in the bryology of New Zealand, with special reference to the herbarium of Robert Brown. New Zealand Institute. Bulletin No.3, Part IV*, pp. 191-192. New Zealand Institute, Wellington.

Duckett JG, Duckett AR (1980) Reproductive biology and population dynamics of wild gametophytes of *Equisetum*. *Botanical Journal of the Linnean Society* **80**, 1-40.

During HJ, van Tooren BF (1987) Recent developments in bryophyte population ecology. *Trends in Ecology and Evolution* **2**, 89-93.

Ellstrand NC, Torres AM, Levin DA (1978) Density and the rate of apparent outcrossing in *Helianthus annuus* (Asteraceae). *Systematic Botany* **3**, 403-407.

Engel PP (1968) The induction of biochemical and morphological mutants in the moss *Physcomitrella patens*. *American Journal of Botany* **55**, 438-446.

Ennos RA (1990) Population genetics of bryophytes. *Trends in Ecology & Evolution* **5**, 38-39.

Ennos RA (2001) Inferences about spatial processes in plant populations from the analysis of molecular markers. In: *Integrating Ecology and Evolution in a Spatial Context* (eds. Silvertown J, Antonovics J), pp. 45-71. Blackwell Science, Oxford.

Ennos RA, Swales KW (1987) Estimation of the mating system in a fungal pathogen *Crumenulopsis sororia* (Karst.) Groves using isozyme markers. *Heredity* **59**, 423-430.

Eppley SM, Taylor PJ, Jesson LK (2007) Self-fertilization in mosses: A comparison of heterozygote deficiency between species with combined versus separate sexes. *Heredity* **98**, 38-44.

Escudero A, Iriondo JM, Torres ME (2003) Spatial analysis of genetic diversity as a tool for plant conservation. *Biological Conservation* **113**, 351-365.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Feder ME, Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Nature Reviews Genetics* **4**, 649-655.

Fife AJ (1985) A generic revision of the Funariaceae (Bryophyta: Musci) Part I. *Journal of the Hattori Botanical Laboratory* **58**, 149-196.

Frank W, Ratnadewi D, Reski R (2005) *Physcomitrella patens* is highly tolerant against drought, salt and osmotic stress. *Planta* **220**, 384-394.

Free SJ, Holtz BA, Michailides TJ (1996) Mating behaviour in field populations of *Monilinia fructicola*. *Mycologia* **88**, 208-211.

Goffinet B (2007) *Physcomitrella*. In: *Flora of North America* (ed. Commitee FoNAE), pp. 194-196. Oxford University Press, New York and Oxford.

Goloboff PA, Farris JS, Nixon KC (2008) TNT, a free program for phylogenetic analysis. *Cladistics* **24**, 1-13.

Goodwillie C, Kalisz S, Eckert CG (2005) The evolutionary enigma of mixed mating systems in plants: Occurrence, theoretical explanations and empirical evidence. *Annual Review of Ecology, Evolution and Systematics* **36**, 47-79.

Grimsley NH, Ashton NW, Cove DJ (1977a) Complementation analysis of auxotrophic mutants of the moss, *Physcomitrella patens*, using protoplast fusion. *Molecular and General Genetics* **155**, 103-107.

Grimsley NH, Ashton NW, Cove DJ (1977b) The production of somatic hybrids by protoplast fusion in the moss, *Physcomitrella patens*. *Molecular and General Genetics* **154**, 97-100.

Grundmann M, Ansell SW, Russell SJ, Koch MA, Vogel JC (2007) Genetic structure of the widespread and common Mediterranean bryophyte *Pleurochaete squarrosa* (Brid.) Lindb. (Pottiaceae) - evidence from nuclear and plastidic DNA sequence variation and allozymes. *Molecular Ecology* **16**, 709-722.

Grundmann M, Ansell SW, Russell SJ, Koch MA, Vogel JC (2008) Hotspots of diversity in a clonal world - the Mediterranean moss *Pleurochaete squarrosa* in Central Europe. *Molecular Ecology* **17**, 825-838.

Gunnarsson U, Hassel K, Söderström L (2005) Genetic structure of the endangered peat moss *Sphagnum angermanicum* in Sweden: A result of historic or contemporary processes? *The Bryologist* **108**, 194-203.

Hamrick JL, Godt MJW (1990) Allozyme diversity in plant species. In: *Plant population genetics, breeding and genetic resources* (eds. Brown AHD, Clegg MT, Kahler AL, Weir BS), pp. 43-63. Sinauer Associates, Sunderland, M.A.

Hardy OJ, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation and population genetics models. *Heredity* **83**, 145-154.

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**, 618-620.

Hasegawa M, Kishino H, Yano K (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**, 160-174.

Hassel K, Gunnarsson U (2003) The use of inter simple sequence repeats (ISSR) in bryophyte population studies. *Lindbergia* **28**, 152-157.

Haubold B, Hudson RR (2000) LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics* **16**, 847-848.

Heckman DS, Geiser DM, Eidell BR, *et al.* (2001) Molecular evidence for the early colonization of land by fungi and plants. *Science* **293**, 1129-1133.

Hedrick PW (1987a) Genetic load and the mating system in homosporous ferns. *Evolution* **41**, 1282-1289.

Hedrick PW (1987b) Population genetics of intragametophytic selfing. *Evolution* **41**, 137-144.

Heywood JS (1991) Spatial analysis of genetic variation in plant populations. *Annual Review of Ecology and Systematics* **22**, 335-355.

Hill MO, Preston CD, Smith AJE (1994) *Atlas of the Bryophytes of Britain and Ireland.* Harley Books, Colchester.

Hohe A, Rensing SA, Mildner M, Lang D, Reski R (2002) Day length and temperature strongly influence sexual reproduction and expression of a novel MADS-box gene in the moss *Physcomitrella patens*. *Plant Biology* **4**, 595-602.

Holyoak DT, Bryan V (2005) *Ephemerum hibernicum* sp. nov. (Bryopsida: Ephemeraceae) from Ireland. *Journal of Bryology* **27**, 89-95.

Innes DJ (1990) Microgeographic genetic variation in the haploid and diploid stages of the moss *Polytrichum juniperinum* Hedw. *Heredity* **64**, 331-340.

Jobb G (2008) TREEFINDER version of March 2008. Munich, Germany. Distributed by the author at www.treefinder.de.

Johanson U, West J, Lister C, *et al.* (2000) Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* **290**, 344-347.

Jørgensen S, Mauricio R (2004) Neutral genetic variation among wild North American populations of the weedy plant *Arabidopsis thaliana* is not geographically structured. *Molecular Ecology* **13**, 3403-3413.

Kamisugi Y, von Stackelberg M, Lang D, *et al.* (2008) A sequence-anchored genetic linkage map for the moss *Physcomitrella patens*. *The Plant Journal* **56**, 855-866.

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. *Nature* **389**, 33-39.

Kirkpatrick M, Jarne P (2000) The effects of a bottleneck on inbreeding depression and the genetic load. *The American Naturalist* **155**, 154-167.

Klekowski EJ (1972) Genetical features of ferns as contrasted to seed plants. *Annals of the Missouri Botanical Garden* **59**, 138-151.

Knight CD, Cove DJ, Cuming AC, Quatrano RS (2002) Moss gene technology. In: *Molecular Plant Biology* (eds. Gilmartin PM, Bowler C), pp. 285-301. Oxford University Press, Oxford.

Lande R, Schemske DW (1985) The evolution of self-fertilization and inbreeding depression in plants I. Genetic models. *Evolution* **39**, 24-40.

Landry CR, Townsend JP, Hartl DL, Cavalieri D (2006) Ecological and evolutionary genomics of *Saccharomyces cerevisiae*. *Molecular Ecology* **15**, 575-591.

Li CC, Weeks DE, Chakaravarti A (1993) Similarity of DNA finger-prints due to chance and relatedness. *Human Heredity* **43**, 45-52.

Lindberg SO (1864) Uppställning af familjen Funariaceae. *Öfvers. Af K. Vet.-Akad. Foerh.* **21**, 589-600.

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* **82**, 1420-1425.

Longton RE (1976) Reproductive biology and evolutionary potential in bryophytes. *Journal of the Hattori Botanical Laboratory* **41**, 205-223.

Longton RE, Schuster RM (1983) Reproductive biology. In: *New Manual of Bryology* (ed. Schuster RM), pp. 386-462. Hattori Botanical Laboratory, Nichinan.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics* **4**, 981-994.

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

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. *Genetics* **152**, 1753-1766.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology & Evolution* **18**, 189-197.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209-220.

Marra RE, Cortesi P, Bissegger M, Milgroom MG (2004) Mixed mating in natural populations of the chestnut blight fungus, *Cryphonectria parasitica*. *Heredity* **93**, 189-195.

Milgroom MG, Lipari SE, Ennos RA, Liu Y-C (1993) Estimation of the outcrossing rate in the chestnut blight fungus, *Cryphonectria parasitica*. *Heredity* **70**, 385-392.

Milgroom MG, Lipari SE, Powell WA (1992) DNA fingerprinting and analysis of population structure in the chestnut blight fungus, *Cryphonectria parasitica*. *Genetics* **131**, 297-306.

Mishler D (1988) Reproductive ecology of bryophytes. In: *Plant Reproductive Ecology: Patterns and Strategies* (eds. Lovett Doust J, Lovett Doust L), pp. 285-306. Oxford University Press, Oxford.

Mitchell-Olds T (2001) *Arabidopsis thaliana* and its wild relatives: A model system for ecology and evolution. *Trends in Ecology and Evolution* **16**, 693-700.

Miyashita NT, Kawabe A, Innan H (1999) DNA variation in the wild plant *Arabidopsis thaliana* revealed by amplified fragment length polymorphism *Genetics* **152**, 1723-1731.

Mousette S, Derome N (2004) Molecular polymorphism in *Drosophila melanogaster* and *D. simulans*: What have we learned from recent studies. *Genetica* **120**, 79-86.

Murtagh GJ, Dyer PS, Crittenden PD (2000) Sex and the single lichen. *Nature* **404**, 564.

Myburg AA, Remington DL (2000) Protocol for High-Throughput AFLP analysis using LI-COR IR2 Automated Sequencers. Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria. Available at http://www.up.ac.za/academic/fabi/eucgenomics/euc_mapping/AFLP_protocol.pdf. Accessed 02/03/2006.

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

Nishiyama T, Fujita T, Shin T, *et al.* (2003) Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: Implication for land plant evolution. *Proceedings of the National Academy of Sciences* **100**, 8007-8012.

Ochi H (1968) A revision of the family Funariaceae (Musci) in Japan. *Japanese Journal of Botany* **20**, 1-34.

Oldenhof H, Wolkers WF, Bowman JL, Tablin F, Crowe JH (2006) Freezing and desiccation tolerance in the moss *Physcomitrella patens*: An *in situ* Fourier transform infrared spectroscopic study. *Biochimica et Biophysica Acta* **1760**, 1226-1234.

Oliver MJ, Velten J, Wood AJ (2000) Bryophytes as experimental models for the study of environmental stress tolerance: *Tortula ruralis* and desiccation-tolerance in mosses. *Plant Ecology* **151**, 73-84.

Oliver MJ, Velten J, Mishler BD (2005) Desiccation tolerance in bryophytes: A reflection of the primitive strategy for plant survival in dehydrating habitats? *Integrative and Comparative Biology* **45**, 788-799.

Pressel S, Matcham HW, Duckett JG (2007) Studies of protonemal morphogenesis in mosses. XI. *Bryum* and allied genera: a plethora of propagules. *Journal of Bryology* **29**, 241-258.

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution* **43**, 258-275.

Ramu C, Hideaki S, Tadashi K, *et al.* (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Research* **31**, 3497-3500.

Rensing SA, Lang D, Zimmer AD, *et al.* (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **319**, 64-49.

Renzaglia KS, Schuette S, Duff RJ, *et al.* (2007) Bryophyte phylogeny: Advancing the molecular and morphological frontiers. *The Bryologist* **110**, 179-213.

Reski R (1998) Development, genetics and molecular biology of mosses. *Botanica Acta* **111**, 1-15.

Reski R (1999) Molecular genetics of *Physcomitrella*. *Planta* **208**, 301-309.

Reski R, Faust M, Wang X-H (1994) Genome analysis of the moss *Physcomitrella patens* (Hedw.) B.S.G. *Molecular and General Genetics* **244**, 352-359.

Ritland C, Ritland K (2000) DNA-fragment markers in plants. In: *Molecular Methods in Ecology*. (ed. Baker AJ), pp. 208-234. Blackwell Science Ltd., Oxford.

Ritland K (1990) Inferences about inbreeding depression based on changes of the inbreeding cooefficient. *Evolution* **44**, 1230-1241.

Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetical Research* **67**, 175-185.

Roads E, Longton RE (2003) Reproductive biology and population studies in two annual shuttle mosses. *Journal of the Hattori Botanical Laboratory* **93**, 305-318.

Schabenberger O, Gotway CA (2005) *Statistical Methods for Spatial Data Analysis* Chapman & Hall, Boca Raton.

Schaefer DG, Zrÿd J-P (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *The Plant Journal* **11**, 1195-1206.

Schaefer DG, Zrÿd J-P (2001) The moss *Physcomitrella patens*, now and then. *Plant Physiology* **127**, 1430-1438.

Schneider S, Roessli D, Excoffier L (2000) Arlequin: A software for population genetics data analysis. Ver 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Scott GAM, Stone IG (1976) *Physcomitrella* B.S.G. In: *The mosses of southern Australia*, pp. 260-262. Academic Press, London, New York, San Francsico.

Seymour FA, Crittenden PD, Dickenson MJ, *et al.* (2005) Breeding systems in the lichen-forming fungal genus *Cladonia*. *Fungal Genetics and Biology* **42**, 554-563.

Sharbel TF, Haubold B, Mitchell-Olds T (2000) Genetic isolation by distance in *Arabidopsis thaliana*: Biogeography and postglacial colonisation of europe. *Molecular Ecology* **9**, 2109-2118.

Shaw AJ (2000) Population ecology, population genetics, and microevolution. In: *Bryophyte Biology* (eds. Shaw AJ, Goffinet B), pp. 369-402. Cambridge University Press, Cambridge.

Shaw DV, Kahler AL, Allard RW (1981) A multilocus estimator of mating parameters in plant populations. *Proceedings of the National Academy of Sciences, U.S.A.* **78**, 1298-1302.

Shaw J, Antonovics J, Anderson LE (1987) Inter- and intraspecific variation of mosses in tolerance to copper and and zinc. *Evolution* **41**, 1312-1325.

Shimizu KK (2002) Ecology meets molecular genetics in *Arabidopsis*. *Population ecology* **44**, 221-233.

Skotnicki ML, Selkirk PM, Ninham JA (1999) RAPD analysis of genetic variation and dispersal of the moss *Bryum argenteum* in Ross Island and Victoria Land, Antarctica. *Polar Biology* **21**, 417-422.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics* **16**, 393-430.

Smith AJE (2004) *The moss flora of Britain and Ireland*, second edn. Cambridge University Press, Cambridge.

Snäll T, Fogelqvist J, Ribeiro PJJ, Lascoux M (2004) Spatial genetic structure in two congeneric epiphytes with different dispersal strategies analysed by three different methods. *Molecular Ecology* **13**, 2109-2119.

Stenøien HK, Flatberg KI (2000) Genetic variability in the rare norwegian peat moss *Sphagnum troendelagicum*. *The Bryologist* **103**, 974-801.

Stenøien HK, Såstad SM (1999) Genetic structure in three haploid peat mosses (*Sphagnum*). *Heredity* **82**, 391-400.

Stone IG, Scott GAM (1973) Name changes in Australian mosses. *Journal of Bryology* **7**, 603-605.

Storfer A, Murphy MA, Evans JS, *et al.* (2007) Putting the 'landscape' in landscape genetics. *Heredity* **98**, 128-142.

Tan BC (1979) A new classification for the genus *Physcomitrella* B.S.G. *Journal of the Hattori Botanical Laboratory* **46**, 327-336.

Taylor PJ, Eppley SM, Jesson LK (2007) Sporophytic inbreeding depression in mosses occurs in a species with separate sexes but not in a species with combined sexes. *American Journal of Botany* **94**, 1853-1859.

Taylor TN, Kerp H, Hass H (2005) Life history biology of early land plants: deciphering the gametophyte phase. *Proceedings of the National Academy of Sciences* **102**, 5892-5897.

Thingsgaard K (2001) Population structure and genetic diversity of the amphiatlantic haploid peatmoss *Sphagnum affine* (Sphagnopsida). *Heredity* **87**, 485-496.

Tonsor SJ, Alonso-Blanco C, Koornneef M (2005) Gene function beyond the single trait: natural variation, gene effects, and evolutionary ecology in *Arabidopsis thaliana*. *Plant, Cell and Environment* **28**, 2-20.

Travis J (2006) Is it what we know or who we know? Choice of organism and robustness of inference in ecology and evolutionary biology. *The American Naturalist* **167**, 303-314.

Vekemans X, Beauwens T, Lemaire M, Roldan-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology* **11**, 139-151.

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analysis in plant populations. *Molecular Ecology* **13**, 921-935.

Vogler DW, Kalisz S (2001) Sex among the flowers: The distribution of plant mating systems. *Evolution* **55**, 202-204.

von Stackelberg M, Rensing SA, Reski R (2006) Identification of genic moss SSR markers and a comparative analysis of twenty-four algal and plant gene indices reveal species-specific rather than group-specific characteristics of microsatellites. *BMC Plant Biology* **6**, 9.

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

Walter DM (1986) Is there disruptive selection for self-fertilization? *The American Naturalist* **128**, 421-426.

Wang J (2002) An estimator for pairwise relatedness using molecular markers. *Genetics* **160**, 1203-1215.

Waters ER (2003) Molecular adaptation and the origin of land plants. *Molecular Phylogenetics and Evolution* **29**, 456-463.

Whitlock R, Hipperson H, Mannarelli M, Butlin RK, Burke T (2008) An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources*, In press.

Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *Journal of Environmental Biology* **14**, 611-619.

Wilson PJ, Provan J (2003) Effect of habitat fragmentation on levels and patterns of genetic diversity in natural populations of the peat moss *Polytrichum commune*. *Proceedings of the Royal Society of London B.* **270**, 881-886.

Wolff K, Friso B, Damme JMM (1988) Outcrossing rates and male sterility in natural populations of *Plantago coronopus*. *Theoretical and Applied Genetics* **76**, 190-196.

Wright S (1943) Isolation by distance. *Genetics*, 114-138.

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

Wyatt R, Odrzykoski IJ, Stoneburner A (1989a) High levels of genetic variability in the haploid moss *Plagiomnium ciliare. Evolution* **43**, 1085-1096.

Wyatt R, Stoneburner A, Ordrzykoski IJ (1989b) Bryophyte isozymes: Systematic and evolutionary implications. In: *Isozymes in Plant Biology* (eds. Soltis DE, Soltis PS), pp. 221-239. Chapman & Hall, London.

Yamazaki T (1981) Genic variabilities in natural population of haploid plant, Conocephalum conicum. I. The amount of heterozygosity. *Japanese Journal of Genetics* **56**, 373-383.

Zartman CE, McDaniel SF, Shaw AJ (2006) Experimental habitat fragmentation increases linkage disequilibrium but does not affect genetic diversity or population structure in the Amazonian liverwort *Radula flaccida. Molecular Ecology* **15**, 2305-2315.

Zimmer A, Lang D, Richardt S, *et al.* (2007) Dating the early evolution of plants: Detection and molecular clock analyses of orthologs. *Molecular Genetics and Genomics* **278**, 393-402.

Appendix I Table detailing potential populations of *P. patens* that were visited.

Location	County	Date recorded*	OS grid reference	Date visited	Found?
Harpenden Reservoir	Midlothian	Unknown	NT090605	12/11/2005	No
Threipmuir Reservoir	Midlothian	14/09/2003	NT187641	12/11/2005	No
Clubbiedean Reservoir	Midlothian	25/03/1973	NT197667	12/11/2005	No
Todd Moor	Devon	Unknown	SX632540	04/09/2006	Very small number of plants
Seaborough	Dorset	1967	ST430058	4/09/2006	No
Prime Coppices	Dorset	8/1971	SY388973	4/09/2006	No

Cheddar Reservoir	Somerset	7/10/1989	ST438542	04/09/2006	No
Blagdon Lake	Somerset	7/10/1989	ST506603	05/09/2006	No
Chew Valley Lake	Somerset	7/09/1986	ST573615	05/09/2006	No
Chew Valley Waterworks	Somerset	13/07/1986	ST573617	05/09/2006	Small number of plants - no sporophytes
Coed-y-fedw	Monmouthshire	25/09/2004	SO435089	08/09/2006	No
Whitehall	Monmouthshire	27/09/2004	SO444094	08/09/2006	No
Llanfihangel Gobion	Monmouthshire	10/09/2006	SO345093/ SO343093	11/10/2006	Yes
Bosherston Lakes	Pembrokeshire	22/08/2005	SR977947	07/09/2006	No

Llandegfedd Reservoir	Monmouthshire	30/11/2003	SO331004	22/10/2006	No
Pant-y-cos Reservoir	Monmouthshire	Unknown	ST256915	22/10/2006	No
Leintwardine	Herefordshire	07/2004	SO424718	25/10/2006	No
Pant-y-llyn	Carmarthenshire	14/08/2004	SN606166	29/08/2005	Yes
Stokesay Pool	Shropshire	09/2002	SO434817	24/10/2006	Yes
Wicken Fen	Cambridgeshire	15/10/2005	TL561698	18/10/2006	Yes
Mepal	Cambridgeshire	02/08/2006	TL436811	18/10/2006	Yes
Ouse Washes	Cambridgeshire	06/10/2002	TL472849	28/09/2006	Yes

Chippenham Park	Cambridgeshire	04/09/2006	TL663691	17/10/2006	No
Swavesey	Cambridgeshire	19/09/2004	TL366697	17/10/2006	Yes
Little Wilbraham	Cambridgeshire	12/11/2005	TL533581	19/10/2006	No
Great Wilbraham Common	Cambridgeshire	12/11/2005	TL534578	19/10/2006	No
Gransden wood	Cambridgeshire	Unknown	TL2655	19/10/2006	No
Nene Washes	Cambridgeshire	12/10/2003	TF319-000	19/10/2006	No
Langham Pond	Surrey	16/08/1980	TQ0072	25/09/2006	No
Papercourt marshes	Surrey	23/10/1973	TQ0356	25/09/2006	Yes

Cock Marsh	Berkshire	Unknown	SU880868	27/09/2006	Yes - submerged
Langley Reservoir	Cheshire	22/09/1994	SJ940710	30/10/2006	No
Damflask Reservoir	South Yorkshire	08/12/2001	SK270915	7/11/2006	No
Underbank Reservoir	South Yorkshire	07/12/2001	SK239997	8/11/2006	No
More Hall Reservoir	South Yorkshire	07/12/2001	SK274959	8/11/2006	No
Toddbrook Reservoir	Derbyshire	04/10/1986	SJ9980	9/10/2006	No
Linacre Upper Reservoir	Derbyshire	16/10/1994	SK336727	9/10/2006	No
Swinsty reservoir	West Yorkshire	Unknown	SE1853	16/11/2006	No

Lindley Wood Reservoir	West Yorkshire	05/11/2001	SE210495	09/11/2006	No
Fewston reservoir	West Yorkshire	Unknown	SE1754	16/11/2006	No
Twywell Hills and Dales	Northamptonshire	12/09/2002	SP945774	16/10/2006	No
Ashby Folville	Leicestershire	Unknown	SK713135	18/09/2006	No

*Date recorded = previous record of population

Appendix II Description of *P. patens*

(Taken from Smith, 2004)

Physcomitrella patens (Hedw.) Bruch & Schimp.

Ephemeral green patches or scattered plants, to 2.5mm high. Leaves erect or erect-patent when moist, ovate to lanceolate, shortly acuminate; margins plane, bluntly dentate from about the middle; costa ending in or below apex; cells narrowly rectangular below, shorter above, 15-20µm wide in mid-leaf, 1-2 marginal rows narrower. Setae very short, *c.* 100 µm long; capsules immersed, cleistocarpous, globose, with short blunt beak; spores 26-32µm. By streams, ditches, ponds, lakes, reservoirs, bottoms of dried-out ponds, wet tracks and woodland rides. Lowland. Occasional in England, rare in Wales, very rare in Scotland, extending north to Dunbarton and Kintyre, rare in Ireland. Eurosiberian Temperate. Europe, except the Mediterranean region, north to southern Scandinavia, Yenesei, N. America.

Appendix III Description of *P. readeri*

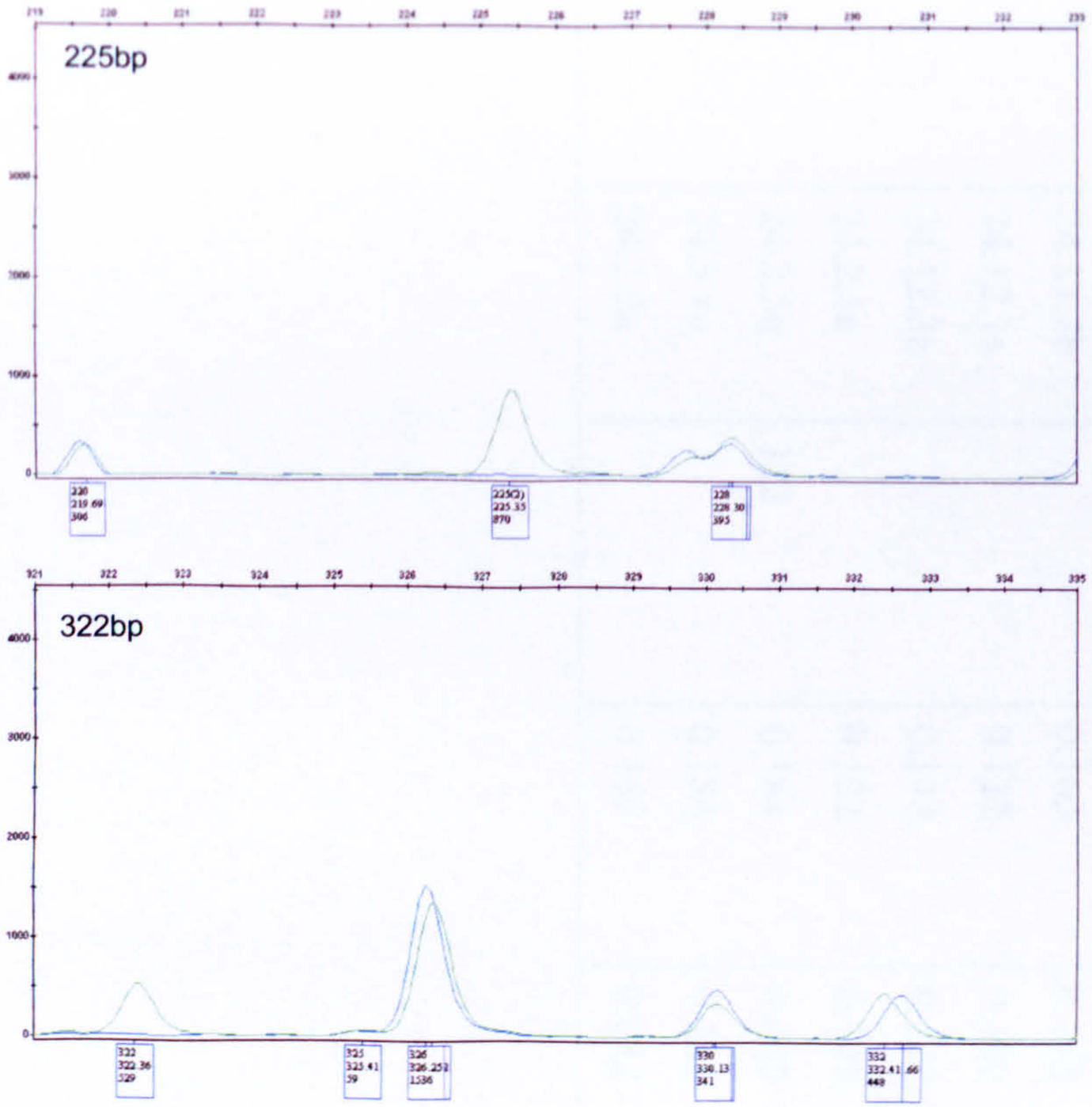
(Adapted from Scott & Stone, 1976, and Smith, 2004)

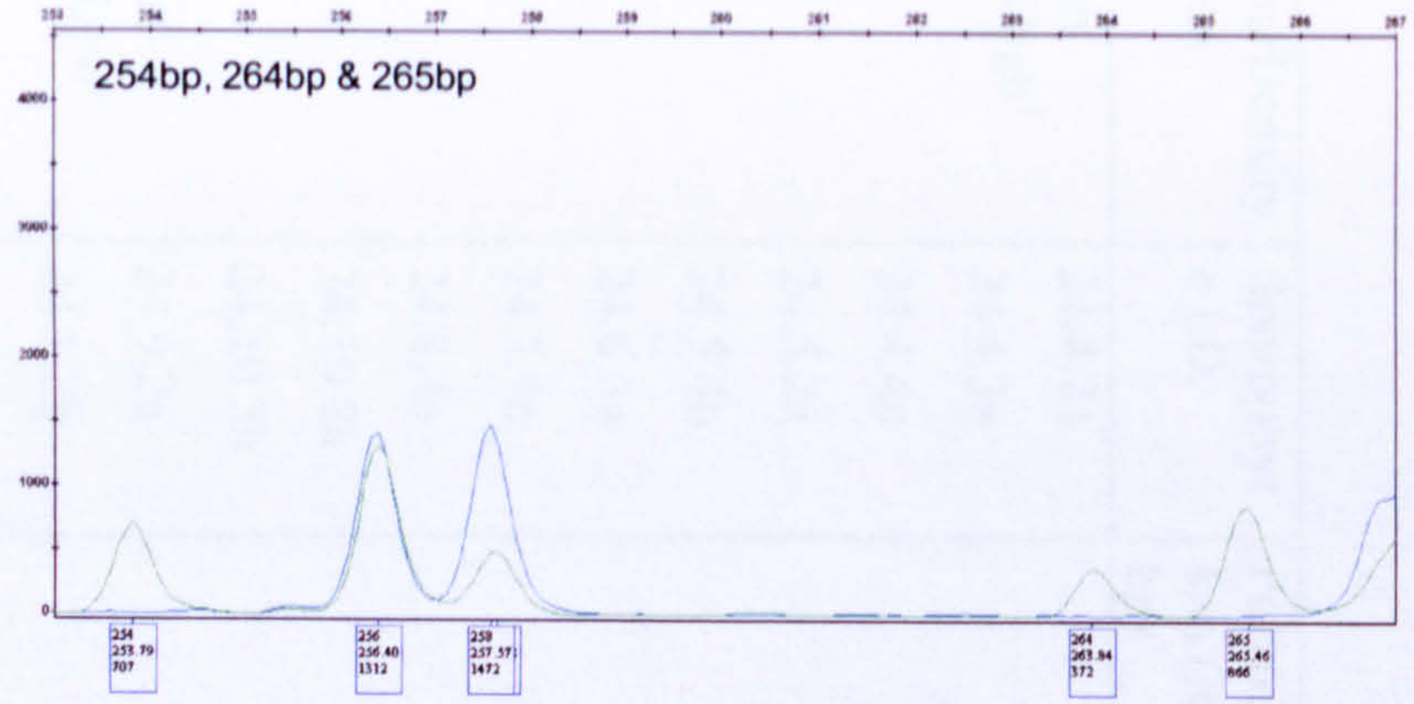
Physcomitrella readeri (C. Müell.) Stone & Scott

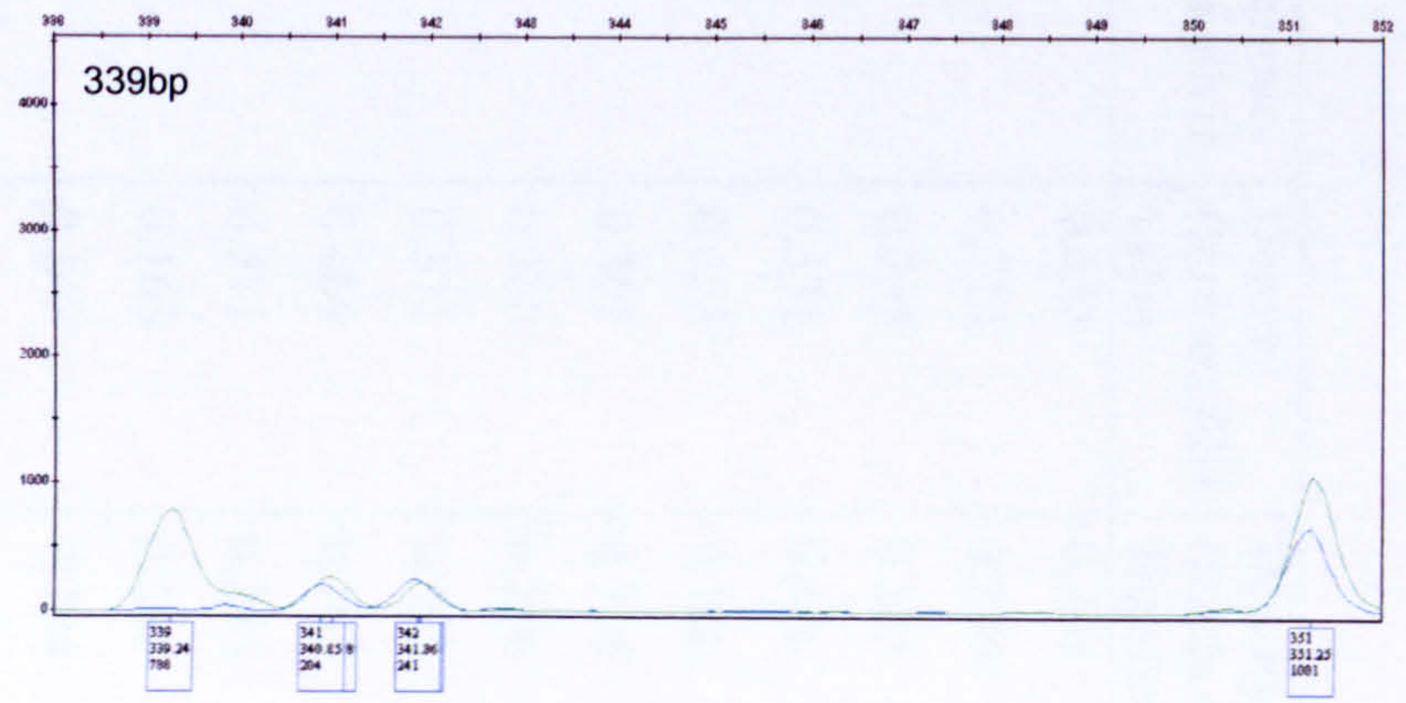
Ephemeral green plants, 0.5 - 4mm high. Leaves erect when moist, oblanceolate, spathulate or obovate with a short apiculus, margins plane and entire near the base, serrate above; costa ending far short of the apex; cells large and thin-walled, rectangular or rhomboidal to six-sided, 20-30µm wide and mostly 2-3 X 1 but sometimes much larger in the lower part of the leaf and adjoining the costa. Setae very short, usually half the length of the cleistocarpous capsule; capsules very variable in size, globular or oblong, apiculate, dark brown to black at maturity. Stomata at base of the capsule not immersed, no line of dehiscence on the capsule. Spores dark rusty brown, 30-45µm, spinulose with slightly hooked spines. Paroecious antheridia in clusters of 4-5 in the leaf axils just below the archegonia at the apex.

Growing in wide patches on mud or silt in recently flooded areas. UK population in West Yorkshire, also Australia, New Zealand, U.S.A. and Japan.

Appendix IV Screenshots depicting polymorphisms between the pooled and reference gametophytes produced from the Gransden X Villersexel cross. Polymorphic fragment sizes are indicated.







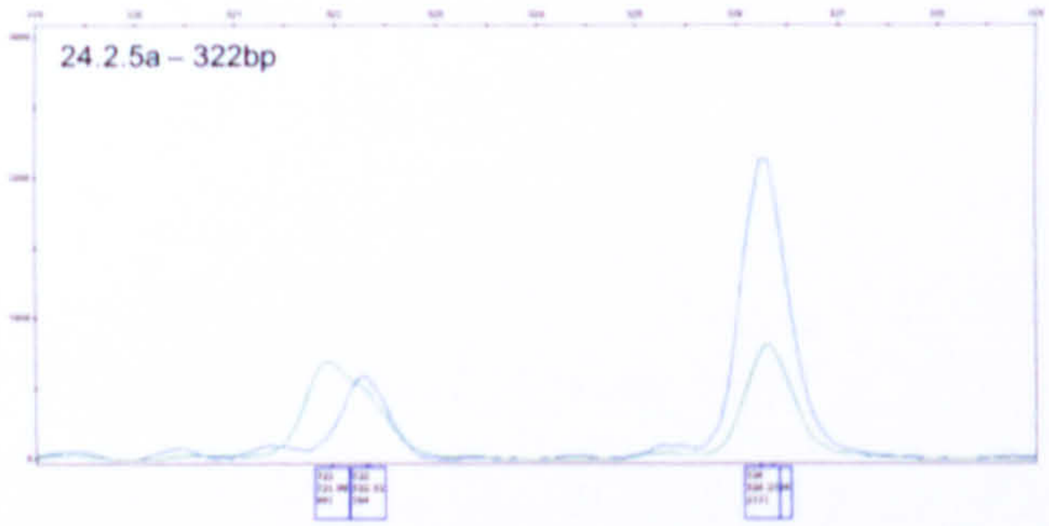
Appendix V Table detaling each sporophyte analysed in Chapter 5, including the number of polymorphic loci and the density of *P. patens* plants around each sporophyte.

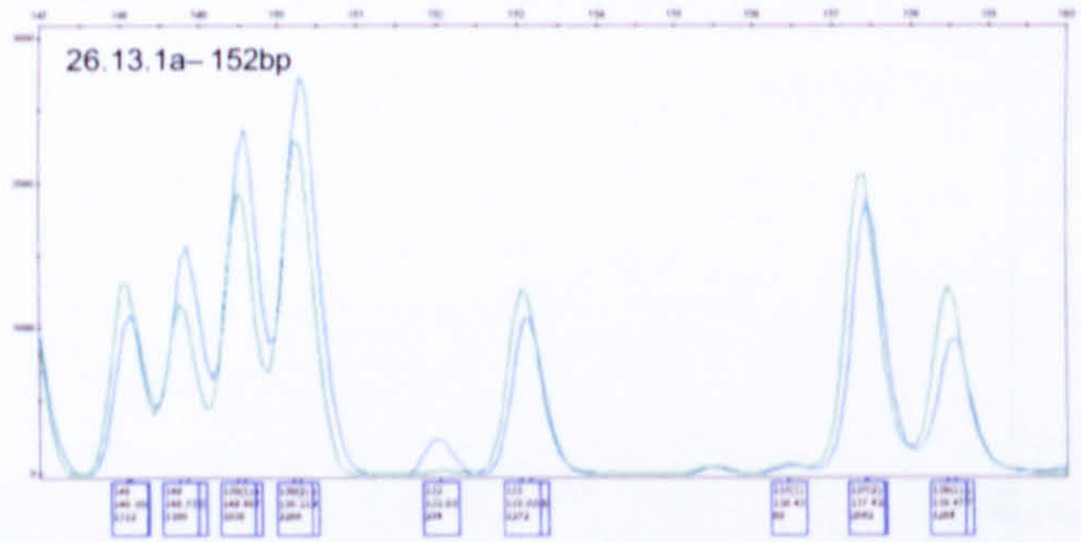
Population/Density designation	Sporophyte I.D.	Polymorphic loci (length in bp)	Density (fraction fill) 1x1m	Density (fraction fill) 10x10cm
Papercourt marshes/High[1]	24.4.2a		0.653	0.408
	24.4.3a		0.714	0.286
	24.4.4a		0.735	0.265
	24.5.2a		0.531	0.449
	24.5.4b		0.653	0.306
	24.6.3a		0.367	0.776
	24.7.1a		0.388	0.286
	24.8.4a		0.571	0.102
	24.10.2a		0.469	0.163
	24.10.3b		0.571	0.245
Papercourt marshes/Low[2]	24.9.2a		0.163	0.265
	24.1.2b		0.143	0.204
	24.1.4a		0.184	0.163
	24.11.2b		0.163	0.163
	24.12.1a		0.122	0.102
	24.12.3a		0.143	0.122
	24.2.1a		0.122	0.245
	24.2.5a	322	0.184	0.082
	24.3.1a		0.184	0.082
	24.3.2a		0.143	0.082

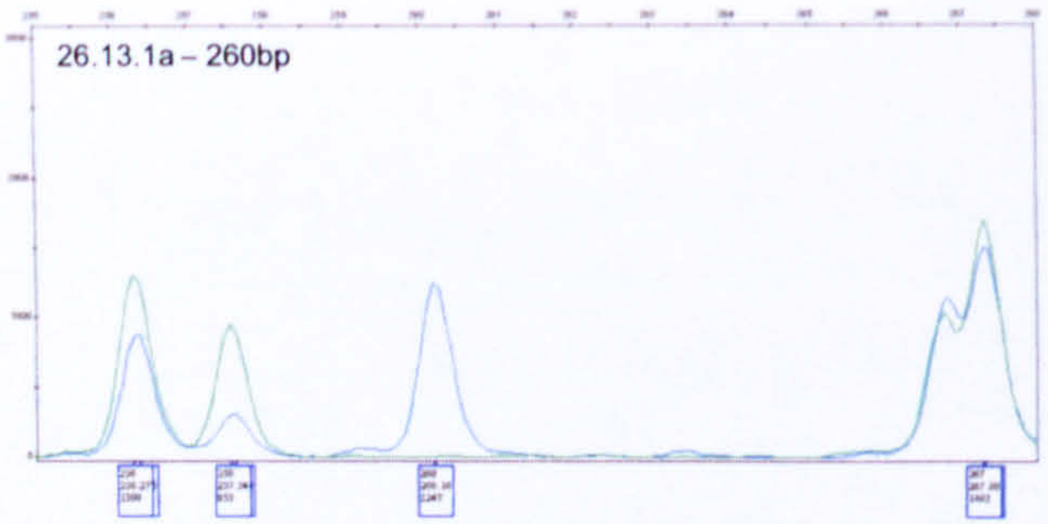
Ouse washes/High[1]	25.1.3a		0.204	0.143
	25.12.3a		0.245	0.245
	25.2.1b		0.163	0.531
	25.2.2a		0.184	0.612
	25.2.3b		0.143	0.449
	25.3.1a		0.122	0.245
	25.5.1b		0.163	0.510
	25.6.1b		0.184	0.306
	25.6.2a		0.184	0.327
Ouse washes/Low[2]	25.8.1a		0.020	0.102
	25.9.1a		0.020	0.286
	25.11.2		0.041	0.122
	25.11.3a		0.041	0.122
	25.12.1b		0.041	0.286
	25.13.2b		0.020	0.102
	25.14.2a		0.041	0.204
	25.14.3a		0.041	0.163
	25.7.1a		0.041	0.163
Llanfihangel Gobion/ High[1]	26.3.2a		0.204	0.306
	26.4.1a		0.204	0.204
	26.4.2a		0.245	0.367
	26.4.3a		0.408	0.265
	26.5.1a		0.102	0.327
	26.6.3		0.122	0.143
	26.8.1a		0.122	0.143
	26.8.3a		0.122	0.245
	26.10.1b		0.122	0.367
	26.13.1a	152, 260, 279	0.122	0.367

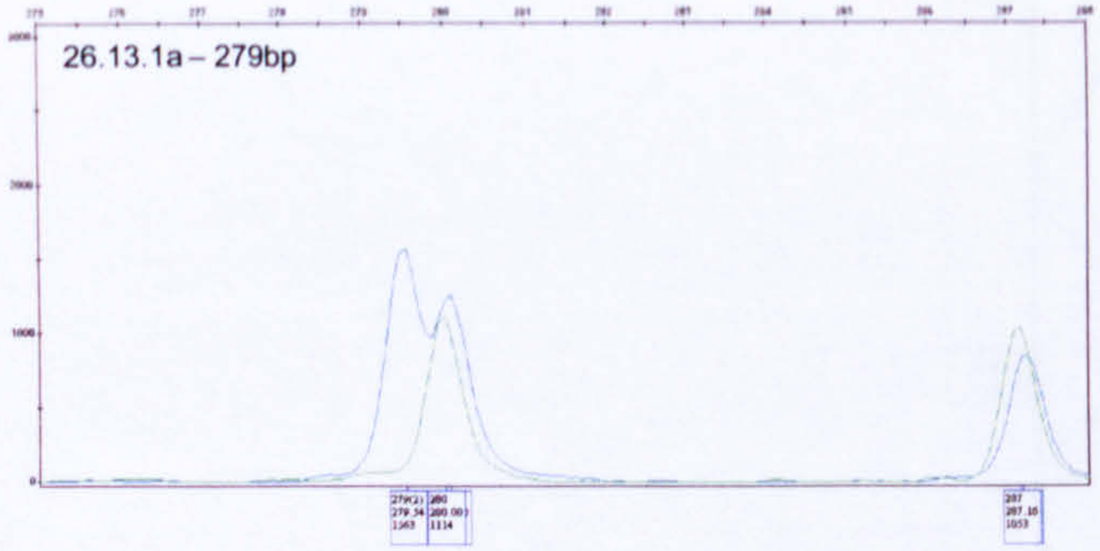
Llanfihangel Gobion/Low[2]	26.12.1a		0.041	0.163
	26.12.2b		0.041	0.122
	26.12.3b		0.061	0.082
	26.13.2a		0.082	0.184
	26.14.1		0.061	0.204
	26.14.2a		0.041	0.122
	26.14.3a		0.061	0.286
	26.15.1c		0.040	0.122
	26.15.3b	199	0.020	0.204
	26.9.3b		0.041	0.061

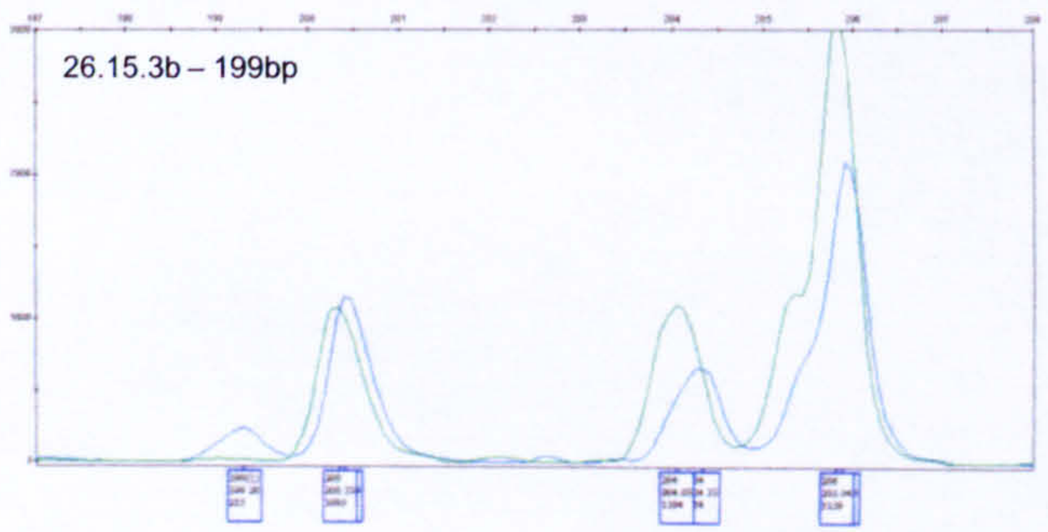
[1]Sporophytes within a high density area
[2]Sporophytes within a low density area











Appendix VI Screenshots depicting polymorphisms between the pooled and reference gametophytes from sporophytes collected from natural populations of *P. patens*, as detailed in Chapater 5. Sporophyte I.D's and polymorphic fragment sizes are given.

Appendix VII Screenshot depicting polymorphic locus EAGMCT229.