THE UNIVERSITY *of* EDINBURGH

The evolution and conservation of tetraploid *Euphrasia* L. in Britain

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

Harriet Stone

PhD Thesis

The University of Edinburgh

2012

Abstract

In the UK, nearly half of the plants short listed for high conservation priority in the UK Biodiversity Action Plan are found in taxonomically complex groups. It is thought that a shift from species- to process-based conservation strategies, aimed at conserving the processes that generate diversity as opposed to simply the end product of these dynamic interactions, may benefit these groups. One group for which this strategy has been proposed is tetraploid *Euphrasia*. The underlying taxonomic complexity in this group is hypothesised to arise via breeding systems, hybridisation and local ecotypic adaptation. The goal of this thesis is to use morphological, ecological and molecular marker data to examine taxon limits and evolutionary processes in order to further understand the mechanisms involved in maintaining species boundaries and generating taxonomic complexity in tetraploid *Euphrasia*. This will not only make conservation in this group more effective, but will also provide a broader insight into some of the processes involved in plant speciation. A detailed study of two widespread, small flowered, tetraploid taxa, *E. micrantha* and *E. scottica*, showed that offspring are almost exclusively the result of self-fertilization. These taxa maintain distinctive morphologies, habitat preferences and chloroplast DNA variation throughout their range, suggesting that they represent coherent lineages within Scotland. As in other widespread inbreeding species, there are high levels of microsatellite differentiation among different populations of the same species. In northwest Scotland three complex populations of tetraploid *Euphrasia* were identified which comprised an array of many different morphs (recognised species, and putative hybrids). Analysis of chloroplast and microsatellite markers suggests that these different morphs represent distinct genetic groups. Within each site there is evidence both for habitat heterogeneity, and for association of morphs with this habitat variation. Intermediate morphs were not simple F1 hybrids, but are likely to have originated via hybridisation and subsequent selfing, surviving as independent recombinant lines, perhaps specialised for habitat types different from that of their progenitor parents. These stable morphs of hybrid origin could represent groups with adaptive potential that may result in the origin of a novel *Euphrasia* species. It will be important to further examine the processes involved in generating novel diversity in *Euphrasia*. For the time being, these complex populations must be recognised as sites requiring special protection within the context of a process-based conservation strategy.

Declaration

I hereby declare that this thesis is composed of work carried out by myself unless otherwise acknowledged and that this thesis is of my own composition. The research was carried out in the period of September 2008 to March 2012. This thesis has not in whole or in part been previously presented for any other degree.

Harriet Stone

June 2012

Acknowledgements

Completing this PhD has felt like something of a personal *tour de force* and would certainly not have been possible without a great deal of help, guidance and support from a number of people. First, I owe a huge debt of thanks to my three supervisors. I wish to thank both Richard Ennos and Pete Hollingsworth who have assisted me in numerous ways, accompanying me on field trips, advising me when planning experiments, helping with my analysis, providing a sounding board for ideas and generally being fantastic mentors, steering me in the right direction. I also want to thank Alan Silverside without whom, I think it's fair to say, this study would not have been possible, particularly when it came to searching out field sites, and identifying separate morphs in the complex sites. Without this detailed identification it would have been extremely difficult to interpret the molecular data from these sites. His extensive knowledge of British *Euphrasia* is frankly awe inspiring.

I have also received practical help from many people in lots of smaller ways. Sabrina Renoux, Marcus Ruhsam, Roberta Kovaks, Michelle Hollingsworth, and Jane Squirrell introduced me to new laboratory techniques and guided me with my experimental work. Jane also planned and performed the complex population collections in Betty Hill. Alan Forrest, Rhiannon Crichton, Markus Ruhsam and Frieda Christie assisted me with the use of the many software programs I have used in my analyses. For all their help I am extremely grateful. I should also thank the Edinburgh Gene Pool for their excellent bioinformatics support, aligning the *Euphrasia* EST library and mining for microsatellites extremely efficiently.

This PhD was funded by NERC, and additionally supported by the Royal Botanic Garden Edinburgh. Many thanks must go to them for making these studies possible.

During my time in Edinburgh, I have been privileged to meet and befriend many people who have provided wonderful moral support during my PhD. For fun and friendship, board games, dinner and wine, thanks go to Matty, Sabrina, Lynne, Tobias, Lorna, Lydia, Heather and David. A special mention must go towards my fellow thesis writing team, who have been a tower of

strength throughout my write up, buoying me up when feeling low, keeping me motivated and being there to celebrate achievements; so thank you to Rebekah, Fiona and Johanna for generally being amazing.

Of course none of this would be possible without my family, particularly my parents who have provided me with an excellent education, always encouraged me to follow my heart and have allowed me the freedom to make my own choices in life. Thank you also to my superb proof readers: my cousin Katie, the wonderful Grahame, and my loving boyfriend Hervé, who proof read my thesis, even though he could hardly understand it. Whilst they may be rather far away, I always know that I have the love and support of my family: parents and grandparents, sisters, brother and cousins, aunts and uncles, close friends and extended family. I am extremely lucky to have such an excellent and inspiring network of people behind me who will always be there for me, good and bad, who encourage me to take risks and who will catch me if I fall. I love them all dearly.

Table of Contents

List of Figures

List of Tables

List of Abbreviation

A	number of alleles per locus
AFLP	amplified fragment length polymorphism
AMOVA	analysis of molecular variance
ANOVA	analysis of variance
bp	base pairs
CIE	Commission international de l'éclairage
cpDNA	chloroplast DNA
CTAB	hexadecyltrimenthylammonium bromide
d.f.	degrees of freedom
DNA	deoxycytosine acid
dNTP	deoxynucleotide triphosphate
EDTA	ethylenediaminetetraacetic acid
EST	expressed sequence tag
F_{is}	Wright's inbreeding coefficient
F_{st}	pairwise population differentiation
G_{st}	level of differentiation among populations
H_E	mean expected heterozygosity
h_T	total diversity
h_S	average intrapopulation diversity
ICUN	international union for the conservation of nature
K	number of structural units
mRNA	messenger ribonucleic acid
MS	mean square
N	number
N_e	effective population size
N_{st}	level of differentiation among populations
NS	not significant
PC1	principal component 1
PCA	principal component analysis
PCO	principal coordinate analysis
PCO1 or 2	principal coordinate 1 or 2
PCR	polymerase chain reaction

PVPP	polyvinylpolypyrrolidine
RAPD	randomly amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
RHS	Royal Horticultural Society
RNA	ribonucleic acid
S	selfing rate
SE	standard error
SNP	single nucleotide polymorphism
SSq	sums of squares
SSR	simple sequence repeat
SSSI	special site of scientific interest
T_M	annealing temperature
t_m	multilocus outcrossing rate
t_s	single locus outcrossing rate
TE	Tris-EDTA buffer
UK	United Kingdom

Chapter 1: General introduction

1.1. Introduction

The conservation of biodiversity involves both the prevention of extinction, and the maintenance of conditions under which novel biodiversity can evolve. Species-based conservation plans account for the prevention of biodiversity loss and are suitable for the majority of plant genera. Some groups, however, are not so easily classed into discrete taxa, and different lines do not correspond to reproductively isolated lineages. For these taxonomically complex groups it has been proposed that process-based conservation strategies may be more appropriate (Ennos et al. 2012)

Taxonomically complex groups are often characterised by recent evolution of novel biodiversity (Ennos et al. 2012). Reticulate evolution which involves processes such as polyploidy, uniparental reproduction (self-fertilisation, apomixis, gynogenesis and hybridogenesis), and hybridisation, produces entities whose morphological and reproductive boundaries are blurred. Many endemic species in countries whose flora have colonised since the retreat of the last ice sheet approximately 10,000 years ago are found in taxonomically complex groups as a result of active diversification. (Brochmann et al. 2003). Since endemic species in these countries are rare, they are often accorded a high conservation value. In the UK, almost half of all the species prioritised for conservation under the UK Biodiversity Action Plan are from taxonomically complex groups (HMSO 1995).

Setting conservation strategies for species in taxonomically complex groups can be hugely problematic, not least because it is difficult for non-specialists to identify individual species. Hybridisation is a particular problem as it leads to the blurring of species boundaries making taxonomic identification difficult. It is a process common to many taxonomically complex plant taxa in Britain, and can take two forms: (1) allopolyploid hybridisation, in which the hybrid's chromosome complement is doubled, and (2) homoploid hybridisation, where there is no ploidal change. Polyploid complexes have been recognised in many taxonomically complex groups in

Britain, such as *Alchimella* L. (Walters 1949), *Cochlearia* L.(Rich 1991), *Dactylorhiza* Necker (Bateman & Denholm 1983; 1985; 1989), *Euphrasia* L. (Yeo 1978) and *Limonium* Mill. (Ingrouille & Stace 1986). Allopolyploidy is recognised to be a prominent mode of speciation in flowering plants, and allopolyploid offspring are often highly reproductively isolated from their progenitor parents (Soltis & Soltis 1993). Homoploid hybridisation is becoming increasingly recognised as a force that may promote speciation (Rieseberg 1997; Gross & Rieseberg 2005), and is also a process that occurs in many taxonomically complex species. Extensive hybridisation between morphologically distinct taxa results in the increased complexity of taxa such as *Calamagrostis* Adans. (Stewart 1997), *Dactylorhiza* Necker (Bateman & Denholm 1983; 1985; 1989), *Euphrasia* L. (Yeo 1978) and *Gentianella* Moench (Pritchard 1959; 1960a; 1960b).

Reproductive isolation is an important factor in maintaining species distinctions. In many taxonomically complex groups, reproductive isolation is achieved as a result of uniparental reproduction. This can add to issues of taxonomic complexity as it is hard to define species limits in groups composed of selfing or apomictic lines (Squirrell et al. 2002; Hollingsworth 2003). In fact, many of the groups which contain the most endemics in the UK have an agamospermous life history. *Hieracium* L., for example, has over 250 named taxa in Britain, many of which are recognised as locally rare endemics (Stace 2010). Other taxonomically complex groups with a high conservation status that contain species with an agamospermous life history include *Alchemilla* L. (Walters 1949), *Limonium* Mill. (Ingrouille & Stace 1986), and *Sorbus* L. (Warburg 1962). Uniparental reproduction via self-fertilisation is also common in some taxonomically complex groups, such as *Euphrasia* L. (Yeo 1978) and *Epipactis* Zinn. (Richards 1997).

It is likely that a combination of the above processes has resulted in present day taxonomic complexity. The identification of discrete taxa in taxonomically complex groups is an acute problem and there are often disagreements as to whether taxonomic units warrant species status or should be recognised at lower taxonomic levels, such as subspecies, or microspecies. This distinction is important as it can have consequences in light of conservation legislation since species tend to be given a higher priority than other taxonomic levels, and may therefore be

included on lists of conservation importance whereas other taxonomic entities are not (Hollingsworth 2003).

Since these groups evolved relatively recently in the UK, protection of future adaptive potential and the processes that generate taxonomic diversity may be of greater importance than the protection of the currently defined products of these dynamic processes (Ennos et al. 2012). Also, the large number of species identified for conservation priority in taxonomically complex groups complicates the already difficult process of allocating limited resources by conservation practitioners (Milligan et al. 1994). With conservation strategies that concentrate on the processes underlying the generation of novel diversity, rather than separate species, these resources may be more effectively targeted. However, for such conservation to occur successfully, it is vital not only to fully understand the processes by which present day diversity came about, but also to establish whether these processes are currently active and spatially localised (Ennos et al. 2012). In this way a process-based conservation plan may be developed which could aid the continued survival of *Euphrasia* in the light of global climate change (Hollingsworth 2003; Ennos et al. 2005; Ennos et al. 2012).

Process-based conservation has been advocated for a number of taxonomically complex taxa, including *Sorbus, Epipactis, Euphrasia* and *Dactylorhiza* (Ennos et al. 2012). The generation of novel taxonomic diversity is particularly well understood in *Sorbus* where rare outcrossing events between outcrossing and autogamous species result in the production of new taxonomic variation (Ennos et al. 2012). Sites have been identified in which these processes currently occur and conservation strategies have been set up for their protection. There is a great deal of evidence to suggest that diversification in European *Dactylorhiza* is generated by hybridisation and introgression between progenitor diploid and tetraploid lineages (Shipunov et al. 2005; Pillon et al. 2007; Nordstrom & Hedren 2009b; Paun et al. 2010; De Hert et al. 2012). These processes are particularly active in the base-rich fens of Greece and northern Sweden (Nordstrom & Hedren 2009a), and conservation must therefore be targeted in these areas. In *Epipactis* the recurrent transition from outcrossing to selfing breeding strategies have created novel multilocus genotypes and generated much of the present day diversity (Hollingsworth et al. 2006).

It is likely that diversity in tetraploid British *Euphrasia* come about as a result of a combination of processes, much like the above examples. Hybridisation is likely to have led to the initial creation of diversity, much like we see in *Sorbus* and *Dactylorhiza*. It is thought that high levels of selfing have been important, providing reproductive isolation from progenitor species and also generating novel multilocus genotypes, as we find in *Epipactis*. It is important to provide some evidence for this theory and to identify particular sites where conservation efforts may be targeted in order to maintain the processes that generate diversity in *Euphrasia*.

This research project aims to enhance the process-based conservation strategy advocated for *Euphrasia* in Britain by looking further into how diversity is generated in this genus.

1.2. The taxonomically complex plant group, *Euphrasia*

Euphrasia L. (Orobanchaceae, formerly a subfamily within Scrophulariaceae) comprises approximately 350 perennial and annual hemi-parasitic plant species and has a bipolar distribution (Gussarova et al. 2008). The northern hemisphere contains the largest number of *Euphrasia* species, which are renowned for being taxonomically complex (Gussarova et al. 2008). The European *Euphrasia* are all hemiparasitic annuals which reproduce sexually and are capable of self-fertilisation (Pugsley 1930; Yeo 1978). The most recent monograph of European *Euphrasia* defined 48 species (Yeo 1978). Within Europe, the highest concentration of *Euphrasia* species can be found in Britain. Of the 19 currently recognised species, 11 are endemic or predominantly restricted to Britain and 6 of these are on the "short list" for conservation priority according to the UK Biodiversity Action Plan for vascular plants (Yeo 1978; HMSO. 1995). Individual species within the genus *Euphrasia* are difficult to classify into discrete and unambiguous entities because of their large intraspecific variability and relatively small interspecific differentiation (Yeo 1968). There is also a great deal of intermediate variation in the form of hybrids, with 64 accepted hybrid combinations recognised (Silverside 1991c; Stace 2010). This is an issue when it comes to conservation, as only a few specialists are able to distinguish the species which may occur within a swarm of unclassifiable variation (Ennos et al. 2005).

It is thought that this species-complex has formed as a result of recent evolution which has involved processes such as hybridisation, polyploidy and segregation of selfing lines. A recent study by French et al. (2008) has helped to clarify some of the diversity to be found within British *Euphrasia*. This paper indicates that the diploid *Euphrasia* species in the British Isles represent morphologically and genetically meaningful biological units. Tetraploid *Euphrasia*, however, show varying degrees of morphological distinctiveness. They do not appear to represent genetically distinct clusters, suggesting that they are not reproductively isolated lineages (French et al. 2008). As a result, while species-based conservation may be appropriate for diploid *Euphrasia*, it is not the best option for conserving the overall diversity of the genus (Ennos et al. 2005). A process-based conservation plan is needed.

Some areas of Britain, such as the north and west Scottish coasts, boast a diverse array of tetraploid morphs (Silverside 1991b). A dynamic set of genetic interactions are thought to be at the root of this diversity. In order to construct an effective conservation plan it is necessary to fully understand these interactions. Understanding the evolutionary processes that generate and maintain diversity in these hotspots will facilitate the formulation of a process-based conservation plan. Under this strategy the conditions that preserve current diversity along with those that generate novel diversity will be maintained (French et al. 2008). In this chapter I will present what is currently known about the diversity and diversification processes in *Euphrasia*. I will outline areas that warrant further study and describe specific aims and approaches of the current research

1.3. Diversification in *Euphrasia*

There are a number of hypothesised mechanisms by which the observed taxonomic complexity in *Euphrasia* may have arisen.

There is evidence to suggest that differences in ploidy level represent a strong barrier for reproduction. Controlled crosses between diploid and tetraploid *Euphrasia* result in hybrids with very low fertility (Yeo 1956, 1966; Liebst & Schneller 2005). French et al. (2008) confirmed that the two ploidy levels are genetically distinct with support from both chloroplast and nuclear-gene

genome analysis. Polyploidy may therefore have had a large influence on the diversity of *Euphrasia* species witnessed in Britain, by creating novel entities that were unable to cross back with progenitor species.

Within ploidal groups species distinctions are less clear. Current molecular evidence suggests that diploid *Euphrasia* represent genetically as well as morphologically discrete entities (French et al. 2008). Tetraploid *Euphrasia*, however, do not appear to represent genetically isolated clusters. Of the 19 recognised *Euphrasia* in the UK, 16 are tetraploid, and 5 of these are endemic (Table 1:1). Much of the taxonomic complexity in British *Euphrasia* resides, therefore, within the tetraploids. Tetraploid *Euphrasia* in the UK colonise areas much further north than any of the diploid species, and most of the endemics have a markedly restricted northern British distribution (Yeo 1978). It will therefore be important to further analyse the extent of partitioning of genetic variation within the tetraploid group.

1.3.1. Breeding System

Breeding system differences also seem to play a role in the partitioning of genetic variation in *Euphrasia*. French *et al.* (2005) demonstrated that there is a strong relationship between breeding system and corolla size within the diploid taxa, with small flowered plants showing high levels of selfing, whilst large flowered plants tend to be predominantly outcrossing. In Britain, 11 of the 19 named species have corolla sizes that are consistent with a high level of self-fertilisation (French et al. 2005). Many of these small flowered species are tetraploid and it will be necessary to find out whether the trend continues as predicted in tetraploid taxa.

French et al. (2008) performed a geographically paired sample of two widespread outcrossing tetraploid species, *E. arctica* and *E. nemorosa*. This study demonstrated that a significant proportion of genetic variability could be accounted for by geographic location, but not by taxa (French et al. 2008). Even though morphological differences have been maintained, there is evidence of widespread genetic exchange between these species, so much so that they may be regarded as an outcrossing complex. It is generally believed that selfing species should display more genetic differentiation as the selfing habit means that opportunities for genetic exchange

Table 1:1. Summary of 19 British *Euphrasia* including their ploidy level, flower size, breeding system, distribution and conservation status. Conservation status is from Cheffings and Farrell (2005)

Taxon	Series	Ploidy	Breeding System	Range	Habitat	Conservation Status
E. officionalis subsp. *anglica*	*Euphrasia*	Diploid	Mixed	South England	Grazed pastures and heathland	Near endemic Endangered
E. officionalis subsp. *rostkoviana*	*Euphrasia*	Diploid	Mixed/ Outcrossing	Northern Britain	Damp fertile meadows	Vulnerable
E. vigursii	*Euphrasia*	Diploid	Mixed	South West England	*Agrostis curtisii-Ulex gallii* heaths	Endemic Endangered
E. rivularis	*Euphrasia*	Diploid	Mixed	North West England & Wales	Montane flushes	Endemic Endangered
E. arctica subsp. *borealis*	*Boreales*	Tetraploid	Mixed/ Outcrossing	Throughout Britain	Damp meadows and pastures	Data Deficient
E. nemorosa	*Nemorosae*	Tetraploid	Mixed/ Outcrossing	Throughout Britain	Grassland, heathland, coastal calcareous	Least concern
E. pseudokerneri	*Nemorosae*	Tetraploid	Outcrossing	South East England and Wales	Calcareous grassland and damp fen	Near endemic Endangered
E. confusa	*Nemorosae*	Tetraploid	Inbreeding	Throughout Britain	Turf of heathland and moorland	Data deficient
E. tetraquetra	*Nemorosae*	Tetraploid	Inbreeding	Throuhout; rare North Scotland	Turf of cliffs and dunes	Data deficient
E. foulaensis	*Parviflorae*	Tetraploid	Inbreeding	North Scotland	Turf of cliffs	Data deficient
E. micrantha	*Parviflorae*	Tetraploid	Inbreeding	Throughout Britain	Dry heathland	Data deficient
E. scottica	*Parviflorae*	Tetraploid	Inbreeding	Northern Britain and Wales	Wed moorland and acidic hill flushes	Least concern
E. ostenfeldii	*Parviflorae*	Tetraploid	Inbreeding	Northern Britain and Wales	Dry limestone and ultrabasic ledges	Data deficient
E. frigid	*Parviflorae*	Tetraploid	Inbreeding	Northern Britain	Wet, basic montane ledges	Data deficient
E. marshallii	*Parviflorae*	Tetraploid	Mixed	North Scotland coast	Damp, basic cliff turf	Endemic Endangered
E. campbelliae	*Parviflorae*	Tetraploid	Inbreeding	North West Scotland	Coastal heath turf	Endemic Data deficient
'E. fharaidensis'	*Parviflorae*	Tetraploid	Inbreeding	North Scotland	Base rich coastal flushes	Endemic awaiting description and assessment
E. heslop-harisonii	*Parviflorae*	Tetraploid	Inbreeding	North Scotland	Salt marshes	Endemic Least concern
E. cambrica	*Parviflorae*	Tetraploid	Inbreeding	Wales	Base poor montane turf	Endemic Vulnerable

between different species should be reduced (Schemske & Lande 1985). An assessment of widespread inbreeding species should allow us to distinguish if they represent coherent taxonomic groups across their range, or if genetic variability is accounted for by geographic location as with the outcrossing tetraploids.

1.3.2. Hybridisation

Hybridisation is assumed to be a very important factor in creating diversity in *Euphrasia* and may result in the evolution of new species (Yeo 1978; Vitek 1986). However, it is also a process that can erode genetic differentiation (Levin et al. 1996; Rhymer & Simberloff 1996). With 64 recorded *Euphrasia* hybrid combinations in Britain it is probable that this is a significant force generating diversity within the genus (Stace 1991). In the hotspots of *Euphrasia* diversity, such as in northern Scotland, many species co-occur and hybridisation could be important in driving the generation of novel diversity. An assessment of the levels of genetic diversity within these hotspot sites may give an indication of the importance of reproductive isolation and hybridisation in the generation of distinct morphologies or genetic entities. It could also point towards certain species that regularly produce hybrids and are possibly important in driving diversification.

It has been proposed that cross pollination of small-flowered *Euphrasia* could be facilitated by other, more showy, flowers growing in close proximity, which would attract pollinators (Yeo 1968). Hybridisation between widespread outcrossing and widespread inbreeding taxa may therefore be driving diversification in British *Euphrasia* (French 2004). However, small flowered *Euphrasia* may also hybridise with each other as witnessed by hypothesised hybrids found in the field (Yeo 1978). An experiment growing artificially established mixed populations of *Euphrasia* demonstrated that cross-pollination occurred even within populations of inbreeding species, suggesting that interspecific cross pollination in selfing species may occur more often than originally thought (Liebst 2008). It is possible that this takes place in mixed populations of *Euphrasia*, where plants with small and large corollas can be found together. The study of the level of genetic isolation between different morphs within complex populations may allow us to identify those crosses that result in independent lineages.

1.3.3. Habitat Preference

Ecological differentiation, such as habitat specialisation or differential flowering time, is also hypothesised to result in reproductive isolation of *Euphrasia* species (Yeo 1978). In fact, the habitat in which a species is found can be key to the plant's identification (Yeo 1978; Silverside 1990). Selection can be especially strong in annual species (Vitek 1998) such that it is likely that specialisations for particular environments may evolve relatively quickly in *Euphrasia*. Nevertheless, it is unclear whether species found over a large geographic range are coherent groups, or are in fact the result of convergent evolution with 'species' having homoplastic phenotypic similarities. Indeed, the results of the paired population study (French et al. 2008) could be explained by the independent evolution of morphology under similar selection pressures or differential selection on a small number of genes responsible for those morphological traits.

Different species may well be specially adapted for survival in particular environments. Since *Euphrasia* seeds are relatively large and have no obvious adaptations for dispersal by animals, it is thought that gravity is their major form of dispersal. Therefore, even if interspecific crosses occur in nature, hybrid seed are unlikely to be competitive in their maternal parent's highly selective environment. It may be that only the con-specific crosses are able to survive to adulthood, thereby maintaining the phenotypic and genotypic integrity of the species.

1.4. Aims for the study of diversification in *Euphrasia*

Since tetraploid British *Euphrasia* have been shown to harbour the most taxonomic complexity, this thesis will concentrate on examining the processes involved in generating diversity in this group. Findings will be used to enhance the process-based conservation action plan advocated for *Euphrasia*. In particular this thesis aims to look into:

1. The degree of inbreeding within small flowered tetraploid *Euphrasia.* Do tetraploid taxa follow the trend indicated in the diploid taxa?
2. Whether widespread self-fertilising taxa represent coherent morphological and genetic groups in line with current taxonomy. Do they represent genetically differentiated

progenitor groups suitable for creating novel hybrid diversity or is genetic marker variation completely unrelated to morphology and dependent on geography as with the outcrossing tetraploids?

3. Does morphological variation in complex sites correspond with genetic variation in complex *Euphrasia* populations?
4. Are morphological and/or genetic entities within complex populations located in ecologically distinct habitats?

1.5. Approach

1.5.1. Overview of sampling strategy

For the first major study of the partitioning of molecular variation in British *Euphrasia* a sampling strategy that maximised the number of populations sampled was used, rather than including large numbers of individuals from few populations (French 2004). This study concentrated on examining recognised species, excluding populations with character combinations from differing species. It was decided, therefore, to apply a more focussed assessment that allowed the analysis of within population variation, as well as between population variation.

In the study of widespread inbreeding tetraploid *Euphrasia*, two species were chosen for comparison, *E. micrantha* and *E. scottica*. The taxonomic identity of each population collected was confirmed by British *Euphrasia* expert, Alan Silverside. Samples were collected from four regions, and within each region at least 30 individuals from at least 3 putatively pure populations of each species were collected.

A more stratified, but equally focused approach was used for the sampling of complex populations. Complex populations are populations in which two or more *Euphrasia* species are found growing amongst one another, along with individuals of putative hybrid origin. Samples were collected from within quadrats along transects, or within a grid, placed in order to encompass the majority of morphological variation within sites. This allowed a very detailed

assessment of within population variation. Again, species identifications were confirmed by Alan Silverside.

1.5.2. Overview of morphological analysis

Time was the major limiting factor when collecting morphological data in the field. The majority of collections were made on two, week long field trips, where most of our time was spent looking for and identifying suitable collection sites. Since *Euphrasia* exhibit large intraspecific variability and relatively small interspecific differentiation (Yeo 1968), species identification relies on a combination of character traits. Within complex sites, as morphological variation was so subtle, each species and hybrid identification served to provide an indication of different morphological groups. For the analysis of widespread inbreeding tetraploid species, we were looking to see if taxa retained coherent morphologies across their range. It was decided that between *E. scottica* and *E. micrantha* key components to morphological identification lie in the shape and colour of flowers and leaves. Therefore, for each individual, the final flowering node leaf and a fully developed flower were collected from each individual sampled in the field, for later morphological analysis. This will be described in more detail in Chapter 4.

1.5.3. Overview of habitat analysis

As with the morphological analysis, time was a major limiting factor when it came to assessing the habitats in which different species of *Euphrasia* grew. Previous studies of ecological specialisation in *Euphrasia* have used associated species as an indicator of habitat differentiation (Bobear 1969; Karlsson 1984). From associated species, Ellenberg environmental indicator values can be used to allow relatively quick, easy and sensitive measures of environment as they are highly correlated with environmental variables (Ellenberg 1979; Dzwonko 2001). From sites which have only one species or hybrid of *Euphrasia* present, associated species have been collected as one sample for the whole population. Within complex sites, associated species were identified within individual quadrats, in order to gauge the level of habitat heterogeneity within the sites, and between individual species and hybrids.

1.5.4. Molecular markers used in this study

With the increasing development and improvement of population genetic approaches the study of ecology, evolution and conservation has become ever more approachable and applicable for biologists studying natural populations (Manel et al. 2005). It is important to determine which combination of markers can most optimally be used to answer the questions posed by a study. The molecular markers in this study will be used to three different ends. First, they will be used to examine the degree of inbreeding of small flowered tetraploid taxa. Second, they will be used to examine and compare partitioning of genetic variation among and within populations of two widespread inbreeding tetraploid *Euphrasia* species, and finally to assess within population variation in complex populations of *Euphrasia*.

Tetraploid *Euphrasia* display low levels of genetic integrity (French et al. 2008). Markers with high levels of resolution are therefore needed so any differentiation present can be detected. Many tetraploid *Euphrasia* species are assumed to be highly selfing due to their small flower size (French et al. 2005). Consequently, these species may have very low levels of heterozygosity at many loci. Overall, it is necessary to identify markers with a large number of loci, or markers with a few loci that are highly polymorphic to provide the high taxonomic resolution necessary for this study.

Marker types can fall into two distinct groups, nuclear and organelle. Each has specific properties which affect the properties of the markers or sequence data collected from them.

1.5.4.1. Nuclear markers

Nuclear markers are designed from DNA contained within the nucleus of a eukaryotic organism. Nuclear DNA is sexually inherited and undergoes recombination. This means each locus will have more than one allele (e.g. two alleles at each locus in a diploid organism, four in a tetraploid etc.) and individual loci are independent of one another. As a result nuclear loci can evolve at a much higher rate than organelle loci and hence show greater levels of polymorphism.

In the past allozyme variation has been used to study *Euphrasia* populations (Oliver 1999). Allozymes are variant forms of an enzyme that are coded for by different alleles at the same

locus. They may display different shapes, sizes or level of charge. Electrophoresis of the enzymes results in banding patterns which can be used to assess genetic variation. However, allozymes show low levels of polymorphism compared with other markers available today (microsatellites, SNPs) and are now considered to be of historical, but not practical, importance. Since a limited amount of information can be gained from their analysis, they are unsuitable in the context of this study.

PCR-based techniques, such as randomly amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) can also be used to detect partitioning of genetic variation among taxa, using a phenetic approach to identify individual genotypes. Both of these techniques have been used in the past to assess molecular variation in *Euphrasia*. RAPDs have been used to look at the genetic variation of four Scottish *Euphrasia* species (Oliver, 1999) and hybridisation in Alpine *Euphrasia* (Liebst & Schneller 2005). French et al. (2008) used AFLPs with some success in order to assess the partitioning of genetic variation between the British species of *Euphrasia*. Both techniques are relatively cheap, but whereas the analysis of RAPDs is relatively easy, the production of AFLPs requires a high degree of technical skill. With both techniques, it is difficult to establish reproducible results and there are uncertainties as to the independence of loci (Sweigart & Willis 2003; Lowe et al. 2004). In addition, both of these techniques produce dominant markers, meaning that much larger samples are required in order to estimate allele frequency in comparison to co-dominant markers. Furthermore allele frequency can only be estimated if the population is assumed to be in Hardy-Weinberg equilibrium, a rather critical assumption (Lowe et al. 2004). Hardy-Weinberg equilibrium is, in fact, not anticipated to occur in *Euphrasia* as these species show significant levels of self-fertilisation.

Assessment of the mating system in small flowered tetraploid *Euphrasia* will give us a greater understanding of the partitioning of variation within and between widespread, presumably inbreeding species. The analysis of progeny genotype arrays is commonly used for estimating the level of outcrossing within a population (Ritland 1990b). This requires the use of co-dominant markers with high levels of within population variation. Allozymes, RAPDs and AFLPs are therefore unsuitable in this context.

Single nucleotide polymorphisms (SNP) are co-dominant markers that provide a high degree of resolution between and within populations. Whilst individual loci are generally biallelic, and therefore not highly polymorphic, large numbers of loci can be isolated and can therefore achieve statistical power comparable to that of microsatellites (Morin et al. 2004). However, they are not suitable for the study of tetraploid *Euphrasia* as they are presumed allopolyploids (Yeo 1968), and there is therefore a high risk of homoplasy.

Microsatellite markers, which are co-dominant, highly polymorphic, highly reproducible and relatively easy to score (Jarne & Lagoda 1996) are much more suitable. Their highly polymorphic nature provides a high degree of resolution that will also allow the detection of any variation present within populations of presumed highly inbreeding *Euphrasia* species. Microsatellite markers have been used in the past to measure the inbreeding coefficient in diploid *Euphrasia* (French et al. 2005). They have also been developed for use in population genetic analysis of Taiwanese *Euphrasia* (Wang et al. 2009). However, unique primers must be developed for each microsatellite loci, and they do not transfer easily between species. Of the 5 SSRs developed from the British diploid, *E. nemorosa* (French et al. 2003) and the 9 SSRs developed from Taiwanese *Euphrasia* (Wang et al. 2009), only two of the French *et al*. primers were found to amplify in a polymorphic manner in tetraploid *Euphrasia*. Since tetraploid *Euphrasia* are thought to be allotetraploids (Yeo 1968) it may be possible to develop loci that amplify in a diploid manner despite the tetraploid genome. This trait will be useful for estimating outcrossing rates.

Given the advantages of microsatellite markers over other nuclear genetic markers, they were chosen to estimate outcrossing rates and inbreeding coefficients in tetraploid populations. They will also be useful for examining the partitioning of genetic variation within and between populations of tetraploid *Euphrasia*. Part of this study therefore involved the isolation of new microsatellite loci.

1.5.4.2. Organelle markers

When integrated with nuclear markers, markers developed from organelle genomes can greatly enhance data on genetic variation and can be used to great effect when studying seed and pollen flow, hybridisation and migration (Ennos et al. 1999). This is because the organelle genomes are non-recombining, haploid and of uniparental inheritance, so male and female migration can be decoupled (Ennos et al. 1999). Each organelle, however, behaves as a single gene and they evolve slowly, resulting in low resolution and only one piece of data per sample, hence the need to study them in combination with nuclear markers (Ennos et al. 1999).

Plants have two potential organelle genomes to study: chloroplast and mitochondria. Chloroplasts have been used to a greater extent in plant population studies. This is because chloroplast DNA has a conserved structure across taxa whilst mitochondrial DNA frequently undergoes intrachromosomal recombination in plants (Atlan & Couvet 1993; Ennos et al. 1999). The mitochondrial genome in plants is very large and has a slower rate of mutation than the chloroplast, making it less suitable for population level studies (Provan et al. 2001). Since high levels of chloroplast variation have been observed in other species within Orobanchaceae (dePamphilis et al. 1997), it is possible that chloroplast genome variation will be useful not only for studies of genetic variation between species, but also for intra-specific studies.

As the chloroplast genome is relatively well conserved (Ennos et al. 1999) a number of universal primers have been developed that amplify across broad taxonomic groups. French et al. (2008) used primer pairs designed to detect sequence variation by RFLPs (Demesure et al. 1995) to look at partitioning of variance in *Euphrasia*. Universal primers providing chloroplast sequence data have also been used successfully by Gussarova et al. (2008) to study the relationship between *Euphrasia* across the globe, and by Wu et al. (2005) to study the genetic structure within the *E. transmorrisonensis* complex. Screening chloroplast genomes with universal primers for noncoding chloroplast DNA sequences provides a relatively quick method for detecting sequence variation. Therefore, in addition to the use of microsatellite markers, chloroplast sequence data will be used to assess the partitioning of genetic variation within and between British tetraploid *Euphrasia.*

1.6. Layout of the thesis

First, the isolation of eight microsatellite loci by sequencing of an expressed-sequence tag library, is described in Chapter 2. These loci, along with two previously developed *Euphrasia* microsatellite loci were used to examine the level of inbreeding in very small flowered tetraploid *Euphrasia*, and this is the topic of Chapter 3. Chapter 4 examines the partitioning of morphological, ecological and molecular (chloroplast sequence and microsatellite) variation within two widespread inbreeding tetraploid *Euphrasia*, and looks at its level of congruence with current taxonomy. In Chapter 5, a similar analysis of the partition of morphological, ecological and molecular variation is used at a much finer scale, examining variation within complex populations and its level of congruence with current taxonomy.

The final chapter, Chapter 6, examines the findings of this study. This chapter explores how our findings have developed our understanding of the delimitation of widespread inbreeding tetraploid *Euphrasia* as well as our understanding of the evolutionary processes involved in generating taxonomic diversity in British *Euphrasia*. This knowledge will be used to enhance the process-based strategy developed for the conservation of British *Euphrasia.*

Chapter 2: Isolation and characterisation of 8 EST-derived microsatellite markers for cross utility in tetraploid British *Euphrasia*

Abstract Tetraploid British *Euphrasia* are taxonomically complex. Previous genetic studies have failed to elucidate any clear structure between described taxa. There is therefore a need for a high resolution marker to further clarify the relationships between taxa within this group. The sequencing of an expressed sequence tag (EST) library produced 9517 cDNA contigs, from which 419 simple sequence repeats (SSRs) were identified in 390 sequences. Primer design was possible for 187 of the SSRs. Forty-three of these were selected for testing. Nineteen EST-SSRs were found to amplify successfully and display polymorphism. Of these 19 loci, 8 were further developed as they could be multiplexed. These 8 loci were screened on 959 individuals from 28 natural populations and 6 different species and a putative hybrid revealing between 5 and 18 alleles per locus. These EST-SSR loci, with their polymorphism across a wide number of species, will be a valuable genetic resource permitting fine-scale analysis of *Euphrasia* populations.

2.1. Introduction

Tetraploid British *Euphrasia* are a taxonomically complex group in which it has been difficult to place species boundaries. Reticulate evolution, involving hybridisation, polyploidy and breeding system transitions, is thought to be generating taxonomic complexity in this group (French et al. 2008). There are currently 15 recognised tetraploid species in Britain (Yeo 1978), 5 of which are endemic and 4 of which are classified with a high conservation status in the UK Biodiversity Action Plan (HMSO. 1995). Whilst there is some evidence that genetic variation within the group can be accounted for by taxon, it has not been possible to come to any clear conclusions about the relationships between described taxa from cpDNA or AFLP markers (French et al. 2008). It is therefore necessary to develop markers with a high resolution for further analysis of this group.

Simple sequence repeats (SSRs) are effective markers that are widely utilised in molecular ecology. In plants, higher frequencies of SSRs are detected in the transcribed regions of the genome (Morgante et al. 2002) and an estimated 2-5% of all plant-derived expressed sequence

tags (ESTs) are thought to contain SSRs (Kantety et al. 2002; Pashley et al. 2006; Ellis & Burke 2007). The majority of these SSRs have been found to be polymorphic, making them useful as population genetic markers (Pashley et al. 2006; Ellis & Burke 2007). EST-SSRs are significantly more transferable across taxonomic boundaries than SSRs obtained through other methods (Decroocq et al. 2003; Thiel et al. 2003; Woodhead et al. 2003; Bandopadhyay et al. 2004; Saha et al. 2004; Varshney et al. 2005; Pashley et al. 2006; Ellis & Burke 2007; Kim et al. 2008), have a reduced occurrence of null alleles (Kim et al. 2008) and neutrality tests have shown that loci are generally not under positive selection (Kim et al. 2008). It is relatively simple to generate large quantities of single-pass sequences of ESTs utilising next-generation sequencing technologies. EST-SSR libraries can therefore be constructed without the need for cloning. Taking into account all these factors, EST-SSRs are the ideal marker to develop for use in the study of tetraploid British *Euphrasia*.

2.2. Materials and methods

2.2.1. cDNA library construction

We extracted and sequenced *Euphrasia* mRNA to create an EST library. *E. nemorosa* flower buds were collected from the north eastern bank of the Bonaly Reservoir and a footpath leading up to it in the Pentland Hills (Grid Ref: NT 21105 66204). The flower buds were frozen in liquid nitrogen for RNA isolation, in the field. Total RNA was extracted from 300 mg of plant material using the Qiagen RNeasy Midi Kit according to the manufacturer's protocol. One additional step was made to the protocol: after the initial centrifugation, the supernatant was passed through a filtered syringe to remove any debris that might clog up the Qiagen column. From the total RNA, mRNA was isolated according to the instructions of the mRNA Isolation Kit (Roche). In total 8 µg of mRNA were extracted. This was converted to cDNA and sequenced using the 454 pyrosequencing platform.

An eighth of a plate was sequenced using Roche 454 Titanium pyrosequencing technology. Contigs were produced from the original sequencing output using ClcBio assembly software. EST-SSR containing sequences were found using MISA (http://pgrc.ipk-gatersleben.de/misa), a

Perl program. Mononucleotide repeats were excluded from the study. Parameters were set for the identification of perfect di-, tri-, tetra- and pentanucleotide motifs with a minimum of 6, 5, 5 and 5 repeats respectively. When a gap of over 25 bp was present between two SSRs found in one EST, they were described as separate SSRs rather than compound SSRs. Primer3 software (Rozen & Skaletsky 2000) was used to find potential primer sites surrounding the SSR region with parameters set as follows: primer length between 18 and 24 bases with an optimum length of 20 bases, PCR product size range of 100 to 300 base pairs, optimum annealing temperature of 60^{o}C, and GC content between 40 and 60%. Each locus had an output of 10 primer pair pairs which were then analysed using NetPrimer (Premier Biosoft International) checking for absence of primer dimers and hairpins. The sequencing, contig production and SSR mining were performed at the University of Edinburgh's Gene Pool facility.

2.2.2. Estimation of polymorphism and diversity

Initial testing was performed using DNA from *E. micrantha* and *E. scottica*. Screening for polymorphism in the loci was performed across *E. micrantha*, *E. scottica*, *E. fharaidensis*, *E. nemorosa*, *E. foulaensis* and *E. marshallii*, as well as a number of hybrids derivatives of presumed *E. micrantha x scottica x fharaidensis* origin. Total DNA was isolated from silica dried leaf material according to the hexadecyltrimethylammonium bromide (CTAB) method (Doyle & Doyle 1987) with the inclusion of 0.1% insoluble polyvinul-poly-pyrrolidine (PVPP) and 0.2% 2-mercaptoethanol to the 2x CTAB buffer (1.4 M NaCl, 100 mM Tris-HCl, pH 8.0, 20 mM EDTA, 2% CTAB). Two choloroform/isoamyl alcohol extractions were performed before addition of isopropanol (at -20^{o}C). The DNA pellet produced was then suspended in 150 µl of Tris-EDTA (TE) buffer at pH 8.

Initial PCR amplifications were carried out in a final volume of 10 µL with DNA diluted to approximately 4 µg µl^{-1} from the initial suspension, 1 unit of *Taq* buffer [16 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl (pH 8), 0.01%Tween-20], 2 mM $MgCl_2$, 100 µM dNTPs, 100 nM forward and reverse primer and 1 unit of *Taq* DNA polymerase (Bioline) under the following cycling conditions: 80^{o}C for 5 minutes, then 30 cycles of 94^{o}C for 1 minute, 60^{o}C for 1 minute and 72^{o}C for 1 minute followed by 10 minutes at 72^{o}C with a final hold at 4^{o}C. Primers that did not

amplify successfully (product the wrong size, double bands, no amplification) were subject to a range of $MgCl_2$ concentrations (30, 40, 50 and 70 mM) and annealing temperatures (56 to 62^{o}C) in an attempt to optimise the output. Alleles were separated by electrophoresis on a 2% agarose gel.

Primers which amplified successfully in both species and were of the predicted locus size were tested for polymorphism using forward primers with 5' M13 tails (CACGACGTTGTAAAACGA) for the purpose of genotyping. The M13 method uses a two-part primer sequence where the tail corresponds to a standard M13 primer that is fluorescently labelled. Amplification is performed with three primers, the forward tailed primer, an M13 fluorescently labelled primer, and the corresponding reverse primer (Oetting et al. 1995). Initial rounds of amplification use the forward and reverse primers, but in subsequent rounds of amplification, the amplified product has the M13 sequence incorporated and hence the M13 labelled primer takes the place of the forward primer.

PCR reactions remained the same except 100 nM of M13 primer were used, and 50 nM of forward primer. The thermal cycling conditions for M13 reactions were as follows: 80^{o}C for 5 minutes, then 10 cycles of 94^{o}C for 1 minute, optimum primer annealing temperature for 1 minute and 72^{o}C for 1 minute followed by 30 cycles of 94^{o}C for 1 minute, 57^{o}C for 1 minute, 72^{o}C for 1 minute, then 10 minutes at 72^{o}C with a final hold at 4^{o}C. PCR products were analysed on an ABI 3730 sequencer. The genotyping output was analysed using GENEMAPPER (v4.0) software (PE Applied Biosystems).

2.3. Results

A total of 106,543 sequence reads were obtained (GC content 44.17%; mean read length 310 bp). From these sequence reads 9517 contigs were produced. SSR mining found 419 SSRs (7 in compound formation) contained within 390 (4.10%) sequences (Table 2:1.). The most frequent repeats found were trinucleotide repeats (54.42%) and the least frequent were pentanucleotides (1.43%). It was possible to design primers for 187 of these SSRs. Primer pairs were chosen for further development if they contained a dinucleotide motif of over 8 repeats, or a tri- or

Table 2:1. Frequency of di-, tri-, tetra- and pentanucleotide repeat motifs in the EST-SSR library of *E. nemorosa*

Repeat motif (including mirror motifs e.g. AC/GT = TG/CA)	**Repeat number**							**total**	**%**
	5	**6**	**7**	**8**	**9**	**10**	**10+**		
AC/GT	-	19	11	4	5	2	0	41	9.79
AG/CT	-	33	12	9	8	1	3	66	15.75
AT/AT	-	30	10	7	3	1	3	54	12.89
AAC/GTT	7	4	5	2	0	0	0	18	4.3
AAG/CTT	33	6	5	1	1	2	1	49	11.69
AAT/ATT	30	11	6	2	3	0	0	52	12.41
ACC/GGT	13	3	3	1	0	0	0	20	4.77
ACG/CTG	8	3	1	0	0	0	0	12	2.86
ACT/ATG	14	8	2	1	1	1	0	27	6.44
AGC/CGT	5	2	0	0	0	0	0	7	1.67
AGG/CCT	9	8	0	1	0	0	0	18	4.3
AGT/ATC	10	5	1	0	0	0	0	16	3.82
CCG/CGG	6	1	2	0	0	0	0	9	2.15
AAAC/GTTT	2	1	0	0	0	0	0	3	0.72
AAAT/ATTT	7	0	0	0	0	0	0	7	1.67
AAGT/ATTC	1	0	0	0	0	0	0	1	0.24
AATG/ACTT	1	1	0	0	0	0	0	2	0.48
AATT/AATT	4	2	0	0	0	0	0	6	1.43
ACAT/ATGT	1	1	0	0	0	0	0	2	0.48
ACTC/AGTG	0	0	1	0	0	0	0	1	0.24
AGAT/ATCT	1	0	0	0	0	0	0	1	0.24
AGCT/ATCG	1	0	0	0	0	0	0	1	0.24
AAAAG/CTTTT	2	0	0	0	0	0	0	2	0.48
AAAAT/ATTTT	0	1	0	0	0	0	0	1	0.24
AAAGC/CGTTT	1	0	0	0	0	0	0	1	0.24
AACGT/ATTGC	2	0	0	0	0	0	0	2	0.48

Table 2:2. Summary data for 8 microsatellite loci developed for cross amplification in British tetraploid *Euphrasia*

Locus	Repeat Motif	Primer Sequence (5'-3')	T_M (^oC)	Size (bp)	*A*
Earc1	$(TAA)_7$	F: CGACGACGACTCTTCATCAC R: TTGGGCCTTGGAAATTACAG	60	299-328	9
Earc2	$(TTG)_8$	F: GGCATAGTTGTCACACCCAA R: CCCACCTCACCAATCTCTTC	60	245-267	9
Earc3	$(AGA)_7$	F: CGATAAGGAATTGTCCGTGG R: TCAACAATGGTTTCAGCAGC	60	301-328	7
Earc4	$(GGC)_7$	F: GTTTGCACCGACCACTCTTT R: GGCATAATCGCAGCAATCTT	60	144-154	8
Earc5	$(AC)_{10}$	F: TGACGAGTGCAGAACAGACA R: TGACTGACTGACCTTTGCCA	60	201-214	9
Earc6	$(AGA)_{10}$	F: GGATATCAAGGCAGCTCCAA R: TCCTCAACCCGTCGATTATT	60	218-230	5
Earc7	$(TA)_9$	F: TGAAATCTCTGGCTGCCTCT R: TGACTGCATAACATTCTGTCCA	60	394-416	18
Earc8	$(ATT)_7$	F: CTCCTCGCTTTCATTCTCGTT R: GATGGGAGTAATTCGGGTGA	60	186-207	7

Forward primers have an M13 5' tail for the purpose of genotyping; T_M = annealing temperature; *A* = number of alleles

Table 2:3. Cross-species amplification of 8 microsatellite loci within 6 tetraploid *Euphrasia* taxa and a group of putative hybrid origin, *E. micrantha x scottica x fharaidensis.*

	Earc1		*Earc2*		*Earc3*		*Earc4*		*Earc5*		*Earc6*		*Earc7*		*Earc8*	
	N	*A* (sizes)	N	*A* (sizes)	N	*A* (sizes)	N	*A* (sizes)	N	*A* (sizes)	N	*A* (sizes)	N	*A* (sizes)	N	*A* (sizes)
E. micrantha	348	5 (313-328)	362	9 (245-267)	355	4 (310-319)	367	5 (144-152)	363	8 (201-214)	364	5 (218-230)	346	9 (394-406)	362	7 (186-207)
E. scottica	332	6 (299-325)	397	7 (248-267)	376	6 (307-328)	391	7 (144-154)	398	9 (201-214)	411	4 (218-227)	388	16 (394-416)	416	6 (186-201)
E. fharaidensis	68	6 (299-325)	68	3 (258-261)	68	5 (301-319)	68	2 (150-152)	68	5 (201-210)	68	4 (218-227)	68	9 (396-412)	68	5 (186-198)
E. nemorosa	32	5 (299-314)	31	4 (248-261)	32	3 (307-313)	32	2 (150-152)	32	4 (201-210)	32	2 (218-221)	32	5 (396-405)	32	4 (186-198)
E. foulaensis	19	2 (313-325)	19	3 (258-261)	19	5 (307-319)	19	2 (150-152)	19	4 (201-210)	19	2 (218-224)	19	4 (396-406)	19	4 (186-195)
E. marshallii	6	2 (299-325)	6	3 (258-261)	6	3 (307-313)	6	3 (144-152)	6	3 (203-210)	6	1 (218)	6	5 (396-415)	6	3 (186-195)
E. micrantha x scottica x fharaidensis	167	4 (299-325)	168	5 (248-264)	168	5 (307-319)	168	5 (144-146)	168	6 (201-210)	168	4 (218-227)	168	12 (396-415)	168	6 (186-207)

N = number of individuals; *A* = number of alleles

tetranucleotide repeat of over 7 repeats, so long as they were not redundant. This resulted in the testing of 43 loci. Of these 43, 19 were found to amplify well in *E. micrantha* and *E. scottica* and displayed polymorphism. A selection of 8 of these were chosen for further development as they could be multiplexed in groups of two or three primers and diluted together for genotyping (Table 2:2.). All of these primers successfully amplified across all tetraploid *Euphrasia* tested, and were shown to be polymorphic (Table 2:3.).

2.4. Conclusions

This study shows that EST derived SSR primers are indeed highly transferable across species boundaries, and remain polymorphic among species. The EST-SSR primers developed in this study represent useful tools for population genetic studies in tetraploid *Euphrasia*. They will make the investigation of breeding systems in tetraploids possible. Their transferability will allow us to make direct comparisons of different tetraploid *Euphrasia* species as well as allowing us to look into whether hybridisation has acted as a mechanism to enhance biodiversity within this group. These topics will be explored in the following chapters.

Chapter 3: Assessing breeding systems in natural populations of tetraploid British *Euphrasia*

Abstract: Tetraploid *Euphrasia* in Britain are a taxonomically complex group of morphologically similar hemi-parasitic annual species that have a high conservation value. It has been suggested that many tetraploids have high levels of self-fertilisation due to their small flower size. The objective of this chapter is to determine what the breeding system of small-flowered tetraploid *Euphrasia* is. Five populations were selected for genetic assessment of mating system. Between three and six microsatellite loci were used to estimate outcrossing rates in progeny arrays, and inbreeding coefficients, F_{is}, in adult and progeny populations. Outcrossing rates (t_m) varied from 0.001 to 0.571, and F_{is} values varied from 0.827 to 1.000. These results show that there are very high levels of self-fertilisation in small-flowered tetraploid *Euphrasia*. The potential factors influencing breeding systems are discussed along with the implications they have for taxonomic complexity and lineage differentiation within the genus.

3.1. Introduction

The breeding system of a species can have a profound effect on patterns of intraspecific genetic diversity. In particular, high levels of self-fertilisation can alter levels of evolutionary diversification as gene flow within and among populations is reduced, limiting their ability to respond to selection (Jarne 1995a; Charlesworth 2003). It is known that selfing reduces within-population diversity by a factor of $(2 – S)/2$ at equilibrium, where S is the selfing rate (Pollak 1987; Nordborg & Donnelly 1997). If S approaches 100%, the effective within population diversity will be reduced by ½ in effect halving the effective population size (N_e) and increasing the rate of genetic drift (Pollak 1987; Nordborg & Donnelly 1997). Inbreeding also reduces the efficiency of recombination (Nordborg 2000), resulting in further reduction of neutral polymorphisms which are associated with hitchhiking on selective sweeps (Smith & Haigh 1974; Barton 2000) and selection against deleterious mutations (Charlesworth et al. 1993). Self-fertilisation is a common trend in angiosperm evolution (Holsinger 1992; Schoen et al. 1997; Barrett 1998) and is thought to provide reproductive assurance when opportunities for outcrossing are low (Barrett 1998; Aarssen 2000; Holsinger 2000).

Selfing breeding systems can result in the evolution of distinct multilocus genotypes that may result in diverse combinations of character states that are maintained unchanged from one generation to the next (French et al. 2005). This can make defining species limits in groups composed of selfing lines difficult. Selfing entities can vary from being assigned as full species, to unnamed intraspecific variants creating taxonomically complex groups (Squirrell et al. 2002; Hollingsworth 2003). Excessive splitting of a genus into a high number of species can increase their regional endemic flora, which can impact the group's conservation prioritisation (Hollingsworth 2003). Conversely, representing the group as a single species complex would reduce conservation attention with fewer endemic taxa being recognised within the flora.

The genus *Euphrasia* (Orobanchaceae) in Britain is such a case of a genus showing taxonomic complexity with 19 recognised species, 7 of which are endemic along with 64 accepted hybrid combinations (Silverside 1991c; Stace 2010). Two ploidy levels are present in this genus, diploid and tetraploid (Yeo 1978), with most of the complexity occurring within the tetraploid group (French et al. 2008). Experimental evidence shows that this genus represents a self-compatible (Darwin 1876; Yeo 1966), but not apomictic group (Yeo 1966). High levels of self-pollination have been invoked as a potential mechanism for generating the high levels of taxonomic complexity in this group, by facilitating lineage differentiation (French et al. 2005).

At present, quantitative data that estimate outcrossing rates in tetraploid *Euphrasia* are lacking. Measurements of the equilibrium inbreeding coefficient, F_{is}, were used to infer the selfing rate in four diploid species, and these data showed a significant negative correlation with flower size (French et al. 2005). Extrapolation from these inferred selfing rates, and taking into account flower size, suggests that self-pollination predominates in British *Euphrasia* species with 11 of the 19 named species predicted to have selfing rates of > 80% (French et al. 2005). However, these estimates make the assumption that the populations being studied are at equilibrium for a fixed selfing rate and that there is no subdivision occurring within the population. Only one of the 11 species with small flowers is diploid and it may be inappropriate to extrapolate from diploids to tetraploids. Moreover, tetraploid *Euphrasia* in Britain tend to have much smaller

flowers than diploids. Therefore, a linear extrapolation outside the range of currently measured flower sizes may not provide good estimates for levels of selfing within the tetraploid group.

Hybrid populations identified in the field have been accounted for by the suggested outcrossing of small-flowered tetraploid *Euphrasia* (Yeo 1978), suggesting that small flower size does not completely prevent hybridisation. In fact, a higher outcrossing rate has been proposed for the small flowered *E. micrantha*, resulting from the hypothesised mimicry of *E. micrantha* to flowers of *Calluna vulgaris* (L.) Hull (Yeo 1968). Artificially established populations of mixed *Euphrasia* species also demonstrate that cross-pollination can occur within populations of inbreeding species, suggesting outcrossing may occur at a higher rate than originally thought (Liebst 2008).

The primary objective of this chapter is to estimate the rate at which selfing occurs in small-flowered (lower corolla lip area < 2 mm^2) tetraploid *Euphrasia*. In order to do this we shall be using both isolated populations and populations of inbreeding *Euphrasia* found growing in close proximity to one another, in an attempt to replicate the mixed *Euphrasia* populations of Liebst (2008). This should allow us to detect outcrossing events between *E. scottica* and *E. micrantha* if they occur and provide more clarity on outcrossing rates in natural populations of small flowered tetraploid *Euphrasia*. We also wish to determine if there is any evidence of inbreeding depression or outcrossing vigour in this group. Data from these studies should also allow us to observe the mode of inheritance in tetraploid taxa. This will increase our understanding of the dynamics of gene flow in this group which in turn may elucidate mechanisms that generate taxonomic complexity in this genus. This is of a high priority as tetraploid *Euphrasia* constitute approximately 12% of the vascular plant species on the priority 'Short List' for conservation in the UK Biodiversity Action Plan (HMSO. 1995).

To investigate the mating systems of small-flowered tetraploid *Euphrasia*, molecular marker analysis of open pollinated progeny arrays were used to estimate outcrossing rates (Ritland 1990b). This involved fitting the mixed mating model in which maternal plants self-fertilise at rate *S* or outcross with an outcross pollen pool at rate *t*. To determine if there was any evidence of inbreeding depression, or outcrossing vigour, we took the indirect approach of assessing

whether populations of adult plants from the standing population have been subject to selection *in situ*. One would expect populations with inbreeding depression or outcrossing vigour to have lower values of F_{is} than progeny arrays from seed collected from the same population, as highly inbred progeny may be purged from natural populations which are more competitive (Crow 1970; Lande & Schemske 1985; Charlesworth et al. 1990a). The mode of inheritance in tetraploid *Euphrasia* was investigated by comparing molecular marker loci from an *E. micrantha* mother-progeny array.

Microsatellite markers have been chosen for this study as they are codominant. This allows for the identification of heterozygous individuals. Microsatellite markers also have high levels of allelic diversity due to their high mutation rate, which increases the precision of estimates of genetic diversity (Jarne & Lagoda 1996). This is especially important when investigating species which are likely to have high levels of inbreeding as they are usually less polymorphic than outcrossing species (Jarne 1995b; Hamrick & Godt 1996; Charlesworth 2003). We shall implement both progeny arrays and population structures to look into breeding systems as these two methods can provide complementary information (Jarne & David 2008). In this study we shall estimate outcrossing rates in two small-flowered tetraploid *Euphrasia*, namely *E. micrantha* and *E. scottica*, as well as a specimen of putative complex hybrid origin with a morphology representing a cross of *E. micrantha x scottica x fharaidensis*.

3.2. Materials and methods

3.2.1. Study system

E. micrantha and *E. scottica*, two widespread tetraploid *Euphrasia* species, and a putative complex hybrid form the object of this study. Both *E. micrantha* and *E. scottica* grow throughout Scotland, and indeed have a Europe wide distribution (Yeo 1978). Both are small-flowered and thought to be highly inbreeding. *E. micrantha* are purple flowered *Euphrasia* with many branches and purple tinged leaves and stems, and are found on dry heath land. *E. scottica* have small white flowers and tend to have only one branch if any. They have green foliage and are found in wet moorland and hill flushes. Neither species is endemic, though they are thought to

play a key role in the creation of endemic species in Britain through rare hybridisation events with widespread outcrossing tetraploid *Euphrasia* (French 2004).

3.2.2. Sampling of plant material

From locations in Scotland, populations of *E. micrantha* (BEm_a) and *E. scottica* (BEs_c) were collected from putatively pure and isolated populations (no other *Euphrasia* species present within 500m), along with putatively pure populations of *E. micrantha* and *E. scottica* found growing directly adjacent to one another (PEms_a and PEsm_a respectively) (Table 3:1). In the adjacent populations, *E. micrantha* was found growing on dry heath land and *E. scottica* in a hill flush that cut through the heath land (Figure 3:1). In addition, one population of assumed hybrid *E. micrantha x scottica x fharaidensis* (LDM_Track) was identified for collection (Table 3:1). Identifications were confirmed by the British *Euphrasia* expert, Alan Silverside.

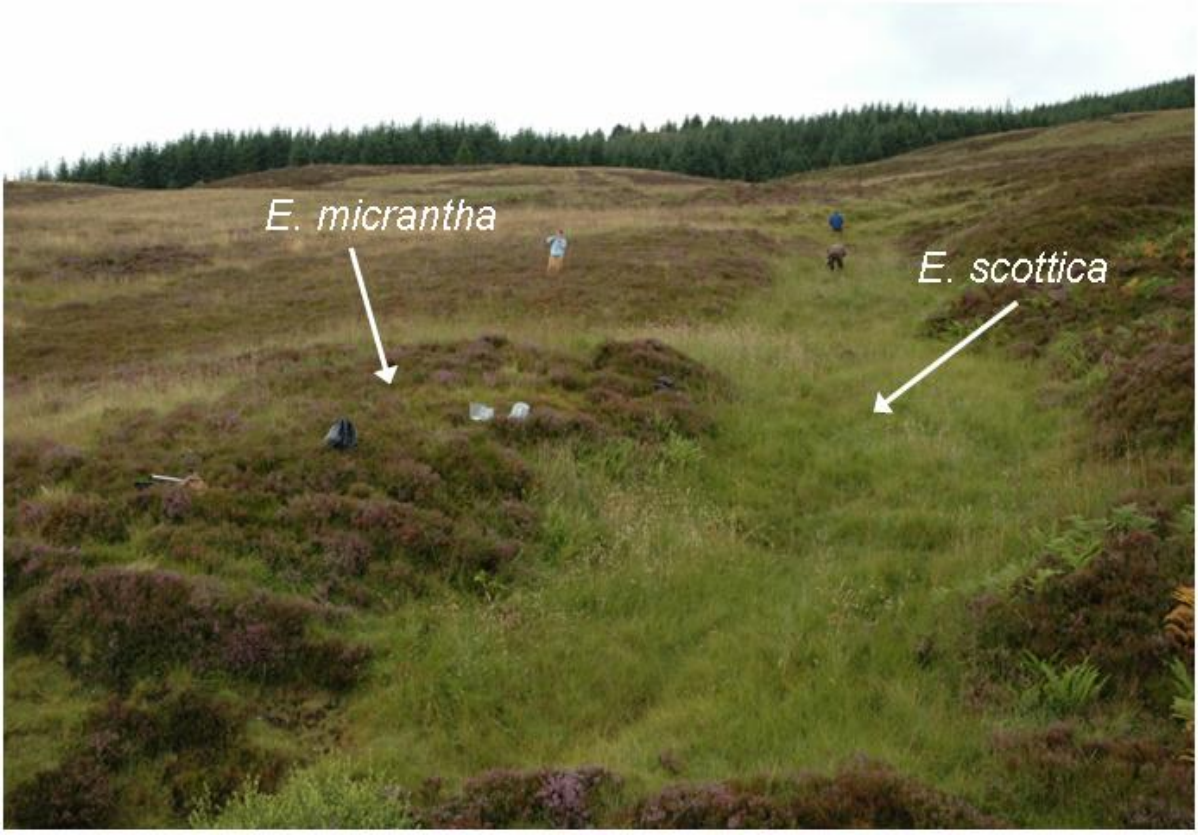


Figure 3:1. Perthshire site where *E. micrantha* and *E. scottica* grow adjacent to one another.

3.2.3. Estimation of outcrossing rate and inbreeding coefficients

To establish the rate of outcrossing in *E. micrantha, E. scottica* and the hybrid population of *E. micrantha x scottica x fharaidensis*, 20 families of open-pollinated seeds were collected from each of the populations previously described (Table 3:1). Samples of the maternal parents were collected from the *E. micrantha* Berwickshire, Twinlaw Ford population and stored for DNA extraction in silica gel. It was not possible to collect the maternal parents in the other populations they were mostly dead by the time they became suitable for seed collection (multiple ripe seed pods on one plant). Plant collections of additional individuals from all populations were made for calculation of inbreeding coefficient (F_{is}) in the adult generation (BEm_a, N = 31; PEms_a, N = 32; BEs_c, N = 29; PEsm_a, N = 32; LDM_Track, N = 64). At least 20 seeds from a single maternal plant (seed families) were sown in individual pots with no host in late December 2010 and left outside over winter. In late February 2011 the families were brought into an unheated greenhouse. A pilot study in December 2009 showed that no planting *Euphrasia* seed with no host was the most effective and easy way of cultivating *Euphrasia*. After germination, up to five seedlings per family were chosen randomly and stored in silica gel for DNA extraction. Poor germination and survival rates meant that for some families it was not possible to collect five seedlings for DNA extraction, and as many plants as were present were collected.

Table 3:2. Site data for *Euphrasia* seed family collections across Scotland (maps in Chapters 4 and 5).

Species	Population Location	Population Code	Altitude (m)	UK Grid Reference
E. micrantha	Berwickshire, Twinlaw Ford	BEm_a	343	NT6482255960
E. micrantha	Perthshire, Coire Screabaig	PEms_a	461	NN7807754186
E. scottica	Berwickshire, Hell's Cleugh	BEs_c	318	NT7404955056
E. scottica	Perthshire, Coire Screabaig	PEsm_a	461	NN7807754186
E. micrantha x scottica x fharaidensis	Sutherland, Loch an Daimh Mor	LDM_Track	96	NC1659742971

3.2.4. Molecular data

3.2.4.1. DNA extraction

DNA was extracted from 6 to 12 silica dried leaves using the hexadecyltrimethyl-ammonium bromide (CTAB) method, as described by Doyle and Doyle (1987), with the inclusion of 0.1% insoluble polyvinyl-poly-pyrrolidine (PVPP) to the 2x CTAB buffer (1.4 M NaCl, 100 mM Tris-HCl, pH 8.0, 20 mM EDTA, 2% CTAB). Two chloroform/isoamyl alcohol extractions were performed after which DNA was precipitated using isoproponol at -20 ^{o}C. Samples were washed with 70% ethanol, dried and suspended in 150µL of Tris-EDTA, ph 8 (TE buffer) and stored at -20 ^{o}C.

3.2.4.2. Microsatellites

All microsatellite markers developed in chapter 2 plus the microsatellite markers *Ene1* and *Ene5*, developed by French *et al.* (2003), were tested on each family to establish their suitability for estimating outcrossing rate. Loci were deemed suitable and chosen for use in analysis if they amplified a maximum number of two alleles per individual with no evidence of fixed heterozygosity (all family members heterozygous for the same two alleles across a locus). This is because we needed the loci to be behaving in a diploid manner for this analysis. In total, seven loci were chosen to analyse the *E. micrantha* populations (BEm_a and PEms_a), six for the complex hybrid population (LDM_Track) and the *E. scottica* population BEs_a, and five for the *E. scottica* population, PEsm_a (Table 3:2).

3.2.4.3. Microsatellite genotyping

The forward primers were tagged at the 5' end with an M13 sequence (CACGACGTTGTAAAA GCA). PCR reactions were performed in 10 µL volumes using the following protocol: 3.95 µL purified water, 1 µL 10 × buffer (Bioline), 3 mM $MgCl_2$, 0.1 mM dNPTs, 0.25 µM forward primer, 0.5 µM reverse primer, 0.1 µM labelled M13 primer (6-FAM_M13 for *Earc3* and *Earc7*, VIC_M13 for *Earc2, Earc5* and *Ene5*, PET_M13 for *Earc1* and *Ene1*), 2 µg of bovine serum albumin (Promega), 1 unit of Bio*Taq* (Bioline), 0.5 µL of DNA (approximately 5-50 ng μL^{-1}).

The cycling conditions were as follows: 80°C for 5 min; 10 cycles of 1 min denaturing at 95°C, 1 min annealing at 60°C, 1 min extension at 72°C; 30 cycles of 1 min denaturing at 95°C, 1 min annealing at 56°C, 1 min annealing at 72°C, ending with 4 mins at 72°C to complete extension followed by storage at 4°C. The PCR products were run out on an ABI 377 DNA sequencer, followed by allele size scoring using GENEMAPPER (v4.0) software (PE Applied Biosystems).

Table 3:3. Summary of microsatellite primers used for breeding system analysis of two *E. micrantha* and *E. scottica* populations, and one hybrid *Euphrasia* population. T_M = annealing temperature; *Primers developed by French et al. (2003).

Locus	Population	Alleles detected		Primer sequence	T_M
Earc1	BEm_a	313, 325	2	F: CGACGACGACTCTTCATCAC	60
	PEms_a	297, 325	2	R: TTGGGCCTTGGAAATTACAG	
	BEs_a	313, 314, 325	3		
	PEsm_a	313, 314, 319	3		
	LDM_Track	309, 313, 314, 325	4		
Earc2	BEm_a	261, 264, 267	3	F: GGCATAGTTGTCACACCCAA	60
	PEms_a	248, 254, 258, 261, 264	5	R: CCCACCTCACCAATCTCTTC	
	BEs_a	254, 258, 261, 264	4		
	LDM_Track	248, 258, 261, 264, 268	5		
Earc3	BEm_a	310, 313, 316	3	F: CGATAAGGAATTGTCCGTGG	60
	PEms_a	313, 316	2	R: TCAACAATGGTTTCAGCAGC	
	BEs_a	313	1		
	PEsm_a	310, 313, 319	3		
	LDM_Track	313, 316, 319	3		
Earc5	BEm_a	206, 210	2	F: TGACGAGTGCAGAACAGACA	60
	PEms_a	207, 210, 222	3	R: TGACTGACTGACCTTTGCCA	
	BEs_a	210, 212	2		
	PEsm_a	208,210	2		
	LDM_Track	206, 208, 210	3		
Earc7	BEm_a	396, 402	3	F: TGAAATCTCTGGCTGCCTCT	60
	PEms_a	396, 402	4	R: TGACTGCATAACATTCTGTCCA	
*Ene1**	BEm_a	169, 171	2	F: AATTCCATCACTGCCAGAAAGAA	60
	PEms_a	169	1	R: TGCACAAACACTCCCTAAGTTTG	
	BEs_a	169	1		
	PEsm_a	165, 169	2		
	LDM_Track	169	1		
*Ene5**	BEm_a	131	1	F: CCGTCACATACTCACATTACACA	60
	PEms_a	131, 135	2	R: CCATTGACTTCGATTTGAAGATT	
	BEs_a	131	1		
	PEsm_a	131, 135, 147	3		
	LDM_Track	131, 135	2		

3.2.5. Analysis of molecular data

3.2.5.1. Progeny-array estimates of breeding system (direct estimate)

In order to estimate multilocus (t_m) and single locus (t_s) outcrossing rates, populations of progeny were analysed individually by fitting the Ritland & Jain (1981) and Ritland (1990b) mixed mating model using the software MLTR (Ritland 2002) (progeny-array estimate of breeding system). In general, default settings were used, except that the "pollen=ovule gene frequencies" option was unchecked. 1000 bootstaps were used to calculate standard errors for mating system parameters t_m and t_s.

3.2.5.2. Population structure estimates of inbreeding coefficients (indirect estimate)

To estimate the inbreeding coefficient of the population F_{is} (population structure estimate of breeding system) in samples of adults and progeny, the software program Fstat v2.9.3 (Goudet 1995) was used. At inbreeding equilibrium, the values would be expected to be equal. If there is either outcrossing vigour or inbreeding depression, F_{is} values are expected to be significantly higher in the offspring than in the parents (Ritland 1990a). Two sample *t*-tests were conducted to determine if there was a significant difference in F_{is} values between parent and offspring generations.

3.2.5.3. STRUCTURE analysis of adjacent *E. micrantha* and *E. scottica* populations

A STRUCTURE analysis (Falush et al. 2007) was performed using version 2.3 of the STRUCTURE software in order to see if there was any evidence of hybridisation or introgression between *E. micrantha* and *E. scottica* when the populations are found growing adjacent to one another. This was performed separately for progeny arrays and adult population collections using the five microsatellite loci that are found amplifying in a common diploid manner in both populations (*Earc1, Earc3, Earc5, Ene1* and *Ene5*). K (the number of structural units) was set to 2, to represent the two species and the analysis was run with a burnin period of 40,000 with 40,000 MCMC repeats after burnin.

3.3. Results

3.3.1. Estimation of outcrossing rates

Euphrasia micrantha populations BEm_a (20 families, 68 progeny), and PEms_a (20 families, 88 progeny) were polymorphic at 6 microsatellite loci (*Earc1*, *Earc2*, *Earc3*, *Earc5*, *Earc7* and *Ene1*, and *Earc1*, *Earc2*, *Earc3*, *Earc5*, *Earc7* and *Ene5* respectively; Table 3:2). Estimated outcrossing rates were not significantly different from $t = 0$ in both populations, with outcrossing rates of 0.001 ± 0.001 and 0.001 ± 0.163 in BEm_a and PEms_a respectively (Table 3:3). Estimates for levels of bi-parental inbreeding, assumed by ($t_m - t_s$) were also very low at 0.000 ± 0.001 for BEm_a and 0.017 ± 0.160 for PEms_a (Table 3:3), and not significantly different from zero.

Similar patterns were found in populations of *Euphrasia scottica*. Fewer loci were found to be polymorphic with populations BEs_c (19 families, 71 progeny) and PEsm_a (14 families, 55 progeny) polymorphic at 3 loci (*Earc1*, *Earc2* and *Earc5*) and 4 loci (*Earc1*, *Earc3*, *Earc5* and *Ene1*) respectively (Table 3:2). Total multilocus outcrossing rates were not significantly different from zero in *E. scottica* populations, much like *E. micrantha*, at 0.047 ± 0.038 in BEs_c and 0.571 ± 0.347 in PEsm_a. Again, levels of bi-parental inbreeding ($t_m - t_s$) were not significantly different from zero at 0.004 ± 0.007 and 0.542 ± 0.340 in BEs_c and PEms_a respectively. The small number of loci used in the analysis of PEsm_a populations affected the accuracy of the estimation of t_m resulting in a large standard error and a t_m value that was an order of magnitude greater than any other results. When pollen gene frequencies were constrained to equal ovule frequencies in the MLTR analysis, a t_m of 0.019 ± 0.018 was observed. Of the 14 seed families tested in PEsm_a, only 2 showed any polymorphism. In the first family, one individual of three was heterozygous at one allele (*Ene5*), and in the second family all three individuals showed heterozygosity suggesting low levels of outcrossing within the population.

The stable population of hybrids, LDM_Track (20 families, 93 progeny), acted in much the same way as its progenitor species with very low level of estimated outcrossing (0.001 ± 0.121) and

extremely low levels of bi-parental inbreeding ($t_m - t_s$) (-0.013 ± 0.119) (Table 3:3). Again, neither was significantly different from zero.

3.3.2. Difference in F_{is} between adult and progeny populations

A two sample *t*-test revealed that in all populations (BEm_a, PEms_a, BEs_c, PEsm_a and LDM_Track), inbreeding coefficients (F_{is}) in the adult (0.984, 0.827, 1.000, 0.895 and 0.874) and the progeny generation (1.000, 0.874, 0.914, 0.919 and 0.924 respectively) were not significantly different. Within species, there was no significant difference in F_{is} detected between *E. scottica* populations. This was also the case in *E. micrantha* however, inbreeding coefficients in Perthshire were significantly different from 1.000 (adult: 0.827; progeny: 0.884).

Table 3:4. Mating system parameters for two pure *E. micrantha* (BEm_a and PEsm_a) and *E. scottica* (BEs_c and PEsm_a) populations and one *E. micrantha x scottica x fharaidensis* (LDM_Track) population estimated using MLTR (Ritland 2002). t_m = multilocus outcrossing rate, t_s = single locus outcrossing rate, t_m-t_s = biparental inbreeding rate; F_{is} = parental inbreeding coefficient; SE = standard error.

Population Code	Number of families per population	Number of progeny per population	Number of loci used	t_m (SE)	t_s (SE)	t_m-t_s (SE)	F_{is}
Euphrasia micrantha							
BEm_a	20	68	7	0.001 (0.001)	0.001 (0.001)	0.000 (0.001)	0.984 (0.013)
PEms_a	20	88	7	0.001 (0.163)	0.018 (0.163)	0.017 (0.160)	0.884 (0.041)
Euphrasia scottica							
BEs_c	19	71	5	0.047 (0.038)	0.043 (0.034)	0.004 (0.007)	0.999 (0.031)
PEsm_a	14	55	5	0.571 (0.347)	0.029 (0.027)	0.542 (0.340)	0.915 (0.031)
Euphrasia micrantha x scottica x fharaidensis							
LDM_Track	20	93	6	0.001 (0.121)	0.014 (0.012)	-0.013 (0.119)	0.882 (0.025)

Table 3:5. Population genetic parameters from microsatellite analysis of adult and progeny generations in two *E. micrantha* and *E. scottica* populations and a putative hybrid population. *N* = number of individuals sampled; *A* = average number of alleles per locus; H_E = average expected heterozygosity over all loci; F_{is} = inbreeding coefficient within populations; 95% (F_{is}) = 95% confidence interval for F_{is} estimated using Fstat v2.9.3 (Goudet, 2001).

Species/population		*N*	Number of loci screened	*A*	H_E	F_{is}	95% (F_{is})
Euphrasia micrantha							
BEm_a	Adult	51	7	2.4	0.352	0.984	$0.955 \leq F_{is} \leq 1.000$
	Offspring	68	7	2.4	0.309	1.000	$1.000 \leq F_{is} \leq 1.000$
PEms_a	Adult	31	7	2.1	0.336	0.827	$0.710 \leq F_{is} \leq 0.982$
	Offspring	88	7	2.1	0.129	0.878	$0.551 \leq F_{is} \leq 0.968$
Euphrasia scottica							
BEs_c	Adult	30	5	1.4	0.073	1.000	$1.000 \leq F_{is} \leq 1.000$
	Offspring	71	5	1.4	0.073	0.914	$0.894 \leq F_{is} \leq 1.000$
PEsm_a	Adult	34	5	2.2	0.284	0.895	$0.727 \leq F_{is} \leq 0.931$
	Offspring	55	5	2	0.241	0.919	$0.647 \leq F_{is} \leq 1.000$
Euphrasia micrantha x scottica x fharaidensis							
LDM_Track	Adult	63	6	2.3	0.392	0.874	$0.745 \leq F_{is} \leq 0.970$
	Offspring	93	6	2.5	0.241	0.924	$0.840 \leq F_{is} \leq 1.000$

3.3.3. STRUCTURE analysis of adjacent *E. micrantha* and *E. scottica* populations

STRUCTURE analysis (Falush et al. 2007) showed that, in general, individuals from the adjacent *E. micrantha* and *E. scottica* populations in Perthshire were highly differentiated, even though they are found growing adjacent to one another. In the progeny array, there is some evidence of outcrossing among species in one family of *E. micrantha*. In this case we have detected potential pollen flow from *E. scottica* to *E. micrantha*, though the sample size is too small to be able to say if this is typicl (Figure 3:2, A). Those individuals in the *E. micrantha* progeny array that appear to share some alleles in common with *E. scottica* are not F1 hybrids, as they are homozygous for all but one allele (*Earc 5*, all heterozygous for 210, 222) and thus may have been present in the population for a number of generations as a result of a historical hybridisation event. There no evidence of any contemporary hybridisation between *E. micrantha* and *E. scottica* (Figure 3:2). This suggests that even though hybridisation can occur between adjacent natural populations, it happens only very rarely.

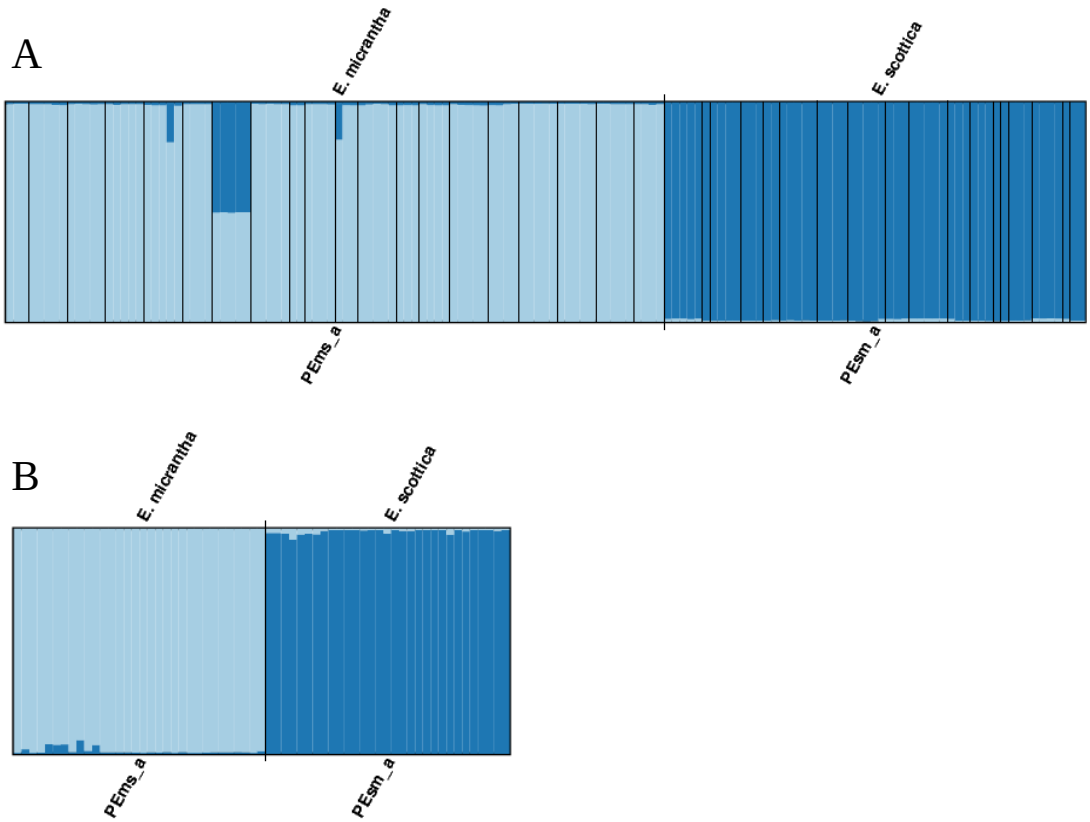


Figure 3:2. STRUCTURE analyses of microsatellite multilocus families for (A) seed families from adjacent *E. micrantha* and *E. scottica* families (each family is separated by a black line) and (B) standing populations of adjacent *E. micrantha* and *E. scottica* populations in Perthshire.

3.3.4. Mode of inheritance

Allotetraploids can exhibit two distinct modes of inheritance: disomic and tetrasomic. Homologous chromosomes are expected to pair faithfully during meiosis, leading to disomic inheritance. However, if homeologous chromosomes from progenitor parents are nearly identical, then tetrasomic inheritance may occur. In this case, four chromosomes pair at meiosis instead of two, resulting in inter-genomic recombination.

Fixed heterozygosity at three loci (*Earc4*, *Earc6* and *Earc8*) between mother and offspring in the BEm_a population indicate that the mode of inheritance in these allopolyploids is disomic, as homeologous chromosomes are not able to recombine and break up heterozygous alleles.

3.4. Discussion

3.4.1. Breeding system

In this chapter, we examined the extent of inbreeding in two species of small flowered, tetraploid *Euphrasia*. The direct estimates of low outcrossing rates, reported for the first time in the present study, revealed that these species are almost exclusively self-fertilising. *E. micrantha, E. scottica* and a putative hybrid population all showed a low rate of outcrossing not significantly different from zero. Indirect estimates of F_{is} values calculated for progeny and adult populations were correspondingly high, with a range of 0.827 to 1.000. These findings are in line with the negative correlation between F_{is} and flower size demonstrated in diploid *Euphrasia* (French et al. 2005). There is also no evidence to suggest that *E. micrantha* show higher levels of outcrossing than other small flowered tetraploid *Euphrasia*. The extremely low levels of outcrossing are in line with features of these taxa that indicate high levels of selfing, such as low levels of herkogamy and dichogamy, as well as poorly developed nectaries (Schulz, and reported by Knuth (1909)). This is especially notable in *E. scottica*, where the stigma is often found curled around against the anthers in a flower.

E. micrantha plants typically have more flowers open at once than *E. scottica* (Yeo 1978). This leads to a higher likelihood of geitonogamous pollination in *E. micrantha*. It may therefore be expected that *E. micrantha* may have higher levels of inbreeding than *E. scottica*. If this were the case, then populations of adult *E. micrantha* would have higher F_{is} estimates. This was not the case in our study, as adult *E. micrantha* had estimated F_{is} values of 0.846 and 0.984 compared to estimates of 0.889 and 1.000 for *E. scottica*. This gives some indication that geitonogamy does not have a great effect on inbreeding coefficients within *Euphrasia.* However, in this instance geitonogamy is difficult to detect as there is only likely to be a significant difference detected if there is outcrossing, and outcrossing rates in this instance are extremely low.

Based on evidence from studies of artificially established mixed *Euphrasia* populations, Liebst (2008) proposed that outcrossing within inbreeding species may occur at a higher rate than previously thought. In Liebst's experiments, a number of hybrid progeny were found growing in

artificial populations of mixed small-flowered tetraploid *Euphrasia*, *E. minima* and *E. salisburgensis*, grown in the Swiss Alps. It may be that the detected outcrossing rates are artificially lowered in natural populations, as outcrossing events between sister lines cannot be distinguished from selfing events. There is no evidence in our study of biparental inbreeding in any population, as t_m - t_s calculations were never significantly different from zero. In any case it would be difficult to detect as individuals from the same family are likely to have the same allele combinations, since the species have such a high rate of inbreeding. In Perthshire, the two *Euphrasia* species were growing so close to one another that they could be considered to be growing as one population. STRUCTURE analysis indicates that whilst outcrossing can occur between the two species, it does not happen often and stable individuals of hybrid origin are extremely rare in reciprocal habitats. There is no evidence to suggest that outcrossing occurs at a higher rate in mixed populations of *Euphrasia* in Britain.

High levels of self-fertilisation in *Euphrasia* may be due to poor pollinator availability, and this could indeed change year on year. Occasional "good" years for pollinators may act to increase outcrossing rates and increase the overall genetic diversity within a population. Yeo (1966) observed poor pollinator service in *Euphrasia*. Moreover, a study of small-flowered *E. willkommii* populations in Spain showed that, despite low levels of insect visitation (3%), approximately 80% of ovules were pollinated, suggesting that reproductive assurance may be an important adaptation for *Euphrasia* species (Gomez 2002).

3.4.2. Inbreeding depression

Inbreeding depression, a result of an individual's genetic load, is expected to be lower in selfing taxa as recessive alleles are purged through selection (Lande & Schemske 1985; Charlesworth & Charlesworth 1990; Husband & Schemske 1996). In particular, annuals are more likely to show purging than perennial plants (Byers & Waller 1999) and biparental inbreeding in annuals is thought to play an important role in reducing genetic load and increasing the relative advantage of selfing genes (Molau 1993).

Even though in four of the five populations lower F_{is} values were seen in parental populations (BEm_a, PEms_a, PEsm_a and LDM_Track), there is no evidence to suggest that there is any fitness advantage conferred to outcrossed progeny. This is because parental and progeny F_{is} values did not significantly differ in any of the populations.

It is possible that mildly recessive alleles may not be purged and hence even highly inbred populations may have some level of inbreeding depression (Charlesworth et al. 1990b; Charlesworth et al. 1991). One possible explanation for the lack of evidence for inbreeding depression may be the poor germination and low survival of *Euphrasia* seedlings in a glasshouse situation (Yeo 1961), which could result in failed germination of highly inbred lines. This holds especially for *E. scottica*, where germination rates were low and conditions in the greenhouse differed more significantly from their natural habitat. Selection, therefore, may have already taken place in seedlings before they were scored. Another explanation for the lack of evidence for inbreeding depression may be the limited power of the test due to small sample sizes and large standard errors and also because there is likely to be very little genetic variation within families.

3.4.3. Evolutionary implications

The high levels of self-fertilisation observed among small-flowered tetraploid species helps our understanding of diversification mechanisms in the taxonomically complex genus *Euphrasia*. A combination of high selfing rates and rare outcrossing events could result in unique recombinant inbred lines and these novel selfing lines may acquire novel multilocus character combinations. The isolation brought about by high levels of inbreeding may contribute to the high diversity of species of *Euphrasia* in Britain, with 11 of the 19 taxa having an inferred selfing rate of > 80% (French et al. 2005). A similar process is thought to be involved in lineage diversification within the taxonomically complex genus *Epipactis* (Squirrell et al. 2002; Hollingsworth et al. 2006).

Further examination of how high levels of self-fertilisation affect gene diversity within and between widespread inbreeding species may help towards understanding processes involved in the evolutionary diversification of *Euphrasia*. It would also be interesting to see how inbreeding

species of *Euphrasia* interact when found in the presence of an outcrossing species. It is possible that rare outcrossing events could generate new recombinant inbreeding lines. Deeper understanding of the processes involved in the diversification of *Euphrasia* may influence conservation programmes in this group. This influence could manifest itself through the creation of programmes that are aimed at preserving the evolutionary processes that create diversity, rather than merely the products of evolution (Hollingsworth 2003; Ennos et al. 2005; Ennos et al. 2012).

Chapter 4: Partitioning of variation within widespread self-fertilising tetraploid *Euphrasia* in Scotland. Do these species represent coherent lineages in line with current taxonomy?

Abstract: Tetraploid *Euphrasia* in Britain are a taxonomically complex group that harbour a large number of endemic taxa, resulting in a high conservation interest. It has been suggested that a process-based strategy that promotes the generation of novel biodiversity would be most suitable for the conservation of this group. In order to implement such a plan it is important to have a good understanding of the mechanisms that generate biodiversity. Hybridisation between widespread outcrossing and widespread inbreeding tetraploids has been hypothesised as a mechanism for generating novel selfing lines and hence novel diversity. A detailed analysis of morphological, habitat and genetic variation within and between two widespread tetraploid *Euphrasia*, *E. micrantha* and *E. scottica*, was performed to assess whether these taxa represent coherent species or regional ecotypes. Consistent differences between species morphology and habitat preference were witnessed across Scotland. Analysis of cpDNA suggests distinct taxonomic groupings, with a degree of regional variation. Analysis of microsatellite haplotypes is less clear. High levels of variation within and between populations of the same species make it difficult to detect differences between the two species. At the regional level, differences between the two species become more obvious. On the strength of distinctive phenotype, habitat preference and chloroplast distributions, the current taxonomic delimitation can be considered consistent with selfing lineages that are widespread throughout Scotland, and not with regional ecotypes.

4.1. Introduction

Taxonomically complex groups, characterised by recent evolution of novel biodiversity, contribute one-third of the total flowering plant species diversity within the British Flora (Rich et al. 2008). Taxonomic complexity is thought to occur as a result of reticulate evolution involving polyploidy, hybridisation and breeding system changes, particularly with regard to the high level of uniparental reproduction (self-fertilisation, apomixis, gynogenesis and hybridogenesis) found within these groups. Taxonomically complex groups in the UK are associated with high levels of endemism. These endemics are often afforded a high conservation status. Indeed, in the UK approximately half of the taxa included in the 'shortlist' of endangered vascular plants are present in taxonomically complex groups (HMSO. 1995).

It is generally agreed that conservation should aim to aid ecological and evolutionary processes in order to maintain the ability of a taxon to respond adaptively to environmental change (Crandall et al. 2000; Moritz 2002; Latta 2008). However, in taxonomically complex groups, it is the evolutionary processes between rather than within taxa that have greater consequences for conservation (Ennos et al. 2005). Process based conservation strategies have therefore been invoked for the conservation of taxonomically complex groups, to encourage the generation of novel taxonomic diversity rather than attempting to preserve the individual entities that the dynamic evolution produces (Hollingsworth 2003; Ennos et al. 2005; Ennos et al. 2012). Before such a plan can be drawn up it is critically important to understand the evolutionary processes that generate biodiversity in the taxonomically complex group (Ennos *et al*., 2012).

The genus *Euphrasia* is a classic example of a group where active diversification has lead to taxonomic complexity and subsequently problems for conservation (Pugsley 1930; Yeo 1978; Ennos et al. 2005; French et al. 2008). Of the currently recognised 19 British species, seven are endemic, presumably the result of post-glacial evolution. Five of the 19 *Euphrasia* species are classified by the IUCN as 'endangered' and three as 'vulnerable' (Cheffings & Farrell 2005). Difficulties arise as species cannot be easily delineated, and thus species based conservation plans cannot be implemented. Consequently, understanding how biodiversity is generated in this group is of great importance for the successful implementation of a process based conservation strategy aiming to conserve the generation of new diversity.

A recent study of genetic variation between *Euphrasia* species in the UK has gone some way in unravelling the taxonomic complexity in this group (French et al. 2008). It was concluded from this analysis that, of the two ploidal levels in *Euphrasia* (diploid and tetraploid), the diploid group represent coherent and distinct genetic clusters (French et al. 2008). Taxonomic complexity therefore resides mostly within the tetraploid *Euphrasia*. One proposed mechanism driving this taxonomic complexity is the recurrent hybridisation between widespread outcrossing and widespread inbreeding species (French et al. 2008). In this scenario, novel character combinations could be brought together and propagated through new inbreeding lines. A paired population analysis of the widespread outcrossing taxa, *E. arctica* and *E. nemorosa*, showed that

there are few barriers to genetic exchange, suggesting they are more of an interbreeding outcrossing complex rather than separate species (French et al. 2008). There has been no such in depth study of widespread inbreeding tetraploid *Euphrasia* in Scotland.

In this chapter we assess variation within and among two widespread inbreeding tetraploids *E. micrantha* and *E. scottica*. These species are widespread across Europe with *E. micrantha* occurring on dry heathland and *E. scottica* in wet moorland and acidic hill flushes (Silverside 1991a). They have been suggested as important progenitor groups for the formation of novel diversity in *Euphrasia* (Biodiversity Action Plan for Eyebrights, 2008). The distinctness of these taxa is in some dispute with Stace (2010) suggesting that *E. scottica* is "close to *E. micrantha* and perhaps not distinct". This view tends to suggest that the two currently recognised species are ecotypes and that *E. micrantha* and *E scottica* do not represent discrete genetic lineages but morphs that have arisen on numerous independent occasions from different progenitor populations. Common garden experiments conducted by Yeo (1962) show that tetraploid *Euphrasia* maintain their morphological differences outside of their natural habitats, suggesting a genetic and not ecological basis to morphological distinctness. Current evidence from AFLP analysis also shows some level of support for discrete genetic lineages (French et al. 2008). Within widespread inbreeding tetraploid *Euphrasia* a principal coordinate analysis of AFLP data showed a general correspondence between named taxa and genetic similarities, though the individual taxa do not appear to form genetically distinct groups (French et al. 2008). It is important to examine the differentiation among and within populations of these inbreeding tetraploid *Euphrasia* species in order to construct a broader picture of the diversity found within and between widespread inbreeding tetraploids and to further our understanding of their propensity to generate novel diversity within *Euphrasia*.

The objective of this chapter is to use a combined analysis of variation in morphology, habitat, chloroplast and DNA markers across a wide geographic range to explore the status of the widespread inbreeding taxa *E. micrantha* and *E. scottica*.

We shall be asking the following questions:

1. Are there coherent morphological entities corresponding to *E. micrantha* and *E. scottica*?
2. Do different entities occupy different habitats?
3. Is there any evidence that these morphological entities represent independent genetic lineages or is genetic marker variation unrelated to morphology, but dependent on geography (i.e. within regions the two morphs are more similar than between regions)?

Plant samples were collected from 25 populations (12 x *E. micrantha* and 13 x *E. scottica*) from four regions across Scotland. Sites for each species were chosen according to the morphological identification of 'good species' by *Euphrasia* specialist, Alan Silverside. Various measurements were taken to assess morphological variation, and associated species lists were used to characterise the habitat of each site. Chloroplast sequence data and microsatellite markers were used to investigate genetic diversity in these populations and evidence of genetic lineages. A multiple regression of geographic distance and habitat differentiation against genetic distance among these populations was also analysed in order to investigate the relative importance of geographical structuring and habitat differentiation on genetic diversity within and between morphs.

4.2. Materials and methods

4.2.1. Study species

Two widespread inbreeding tetraploid *Euphrasia* species were included in this study: *E. micrantha* and *E. scottica*. These species can be found growing throughout Scotland and have distributions that spread across Europe (Yeo 1978) (Figures 4:1 & 4:2). Both species have distinct morphologies and are believed to have different habitat preferences (Silverside 1991a), though they can be found growing in close proximity to one another. *E. micrantha* (Figure 4:1) and *E. scottica* (Figure 4:2) can be distinguished from one another by colour and shape differences in the leaf and flower and also by their differential branching patterns.

E. micrantha can have many branches which are generally straight and rigid. They have purple tinged leaves, especially on the upper surface. The corolla tends to be purple with an elongate lower lip (Yeo 1978). *E. micrantha* is most likely to be found in areas of dry heathland with a heather/grass mix.

E. scottica on the other hand is a slender, erect and commonly unbranched plant though it may have a single, short, arcuate-ascending branch from the middle of the stem (Yeo 1978). The corolla of *E. scottica* is usually white, though the upper lip is sometimes pale lilac. The lower corolla lip is hardly longer than the upper lip. *E. scottica* is most likely to be found growing on wet moorland, and in hill flushes. *E. micrantha* and *E. scottica* are most likely to be found growing close to one another where wet flushes run through drier moorland.

4.2.2. Population sampling

Population samples of between 30 and 37 individuals of each of the tetraploid inbreeding species, *E. micrantha* and *E. scottica* were collected from three to four sites within each of four geographic areas in mainland Scotland (Figure 4:3, Table 4:1). In each region we aimed to sample at least three populations of each species. A mixture of isolated sites where populations grew alone, and mixed sites where the two species grew adjacent to one another were collected from sites around Scotland. Populations were considered isolated if there was a minimum separation of 400m and adjacent if two species were found growing no more than 20m apart. Populations of species growing adjacent to one another were only found in Perthshire and Sutherland. Collections of populations in the Cairngorms/Grampians region are more widely scattered geographically than in other regions due to difficulties in finding sites. However all sites share a similar geology, different from that of Sutherland, Perthshire or Berwickshire. The regional collections allowed the testing of the null hypothesis that two species represent morphologically, ecologically and genetically distinct groups.

Figure 4:1. A typical *E. micrantha* plant and its UK distribution. Map from Preston et al. (2002).

Figure 4:2. A typical *E. scottica* plant and its UK distribution. Map from Preston et al. (2002).

Figure 4:3. Collections for the populations analysis were taken from four regions, Sutherland (A), Cairngorms/Grampians (B), Perthshire (C) and Berwickshire (D). A Google version of this map can be found here.

4.2.3. Morphological analysis

One entire flower (including corolla and sepal) and the pair of leaves associated with that flower were collected from each individual sampled and mounted on card using transparent adhesive tape within 6-12 hours of collection (Figure 4:4). In total, 811 *Euphrasia* plants were used for the morphological analysis: 390 *E. micrantha* and 421 *E. scottica*. All plant species were identified

according to Yeo's monograph (Yeo 1978), with identifications confirmed by the British *Euphrasia* expert, Alan Silverside.

Leaf (upper surface) and flower (upper and lower lip) colour were recorded using Royal Horticultural Society (RHS) colour cards (Anon. 1995). These colours were converted into three coordinates (x, y and Y) according to the 'Commission internationale de l'éclairage' (CIE) colour space using the conversion chart supplied with the colour cards. The coordinates x and y refer to the CIE 1931 colour space chromaticity diagram, and Y to the level of colour saturation. Shape measurements were taken as shown in Figure 4:9. Colour measures (y and Y for flower colour, and Y for leaf colour) and shape measures were combined in a principal component analysis (PCA). In total 12 measures were used in the PCA. The first principal component (PC1) was then used in an analysis of variance (ANOVA) to test the null hypotheses that (1) there is no difference in morphology between regions, (2) there is no difference between species and (3) there are no differences between populations of the same species.

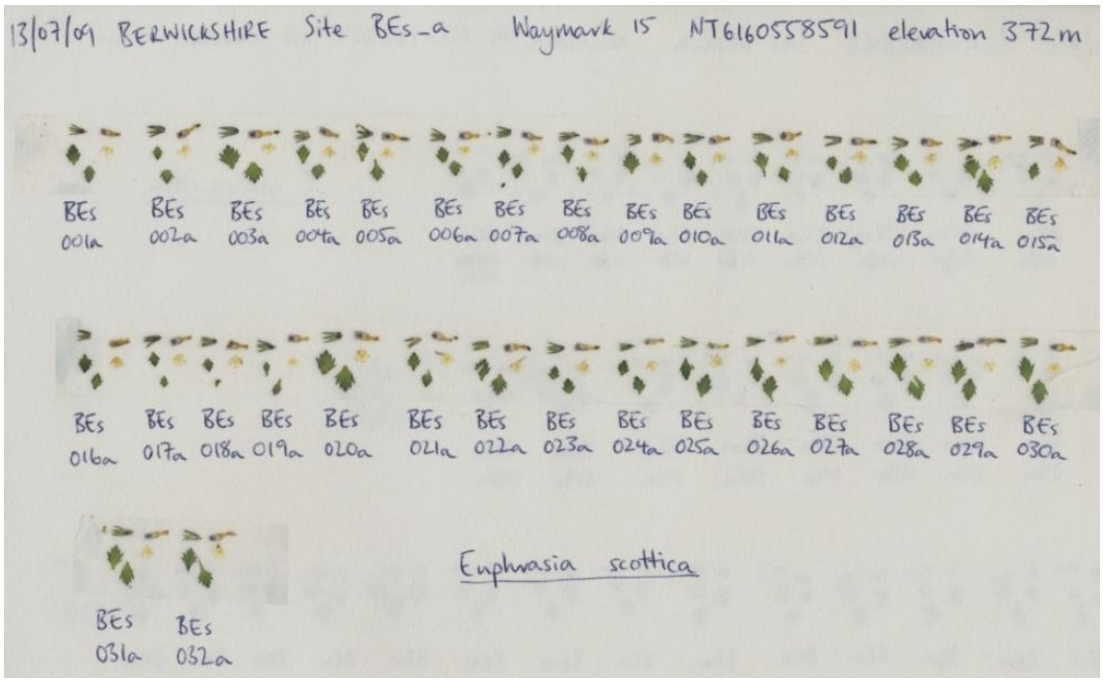
4.2.4. Habitat analysis

Associated plant species were recorded in all populations. Mean Ellenberg indicator values (Ellenberg et al. 1991) for light (L), moisture (F), reaction (R), nitrogen (N) and salt tolerance (S) were compiled from associated species lists using PLANTATT (Hill et al. 2004). Tables of these values compiled at the population, region and species level were then used in an ANOVA to test the null hypothesis of no habitat difference between and within species and regions.

A principal component analysis of mean population Ellenberg indicator values for light, moisture, reaction, nitrogen and salt tolerance provided a single value indicative of a site's ecology in each population (when considering only the first principal component). Pairwise comparisons of a population's first principal component allowed calculations of ecological distance between populations.

4.2.5. Common garden experiment

In order to establish whether the morphological differences are genetically or environmentally determined a common garden experiment was conducted. Seed collections were made between



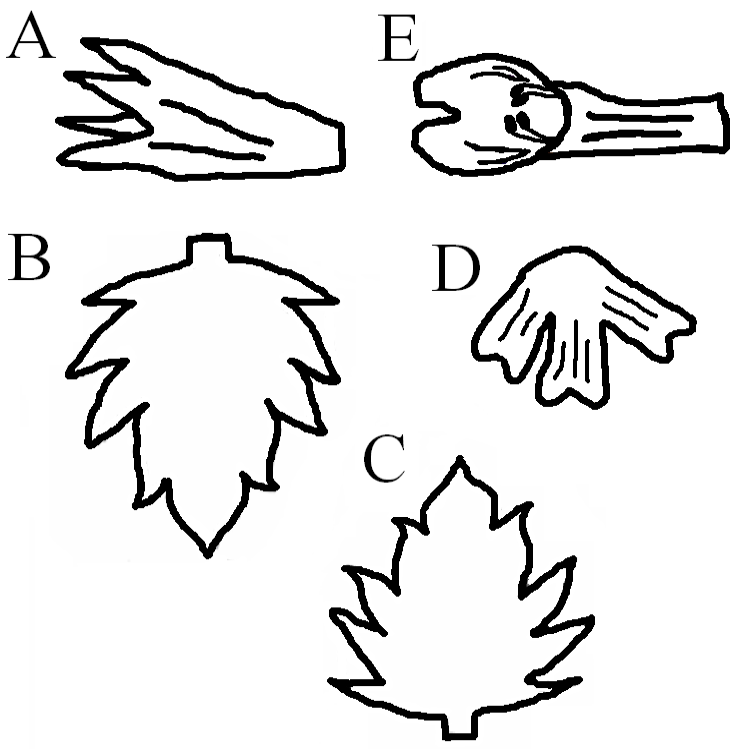


Figure 4:4. An example of mounted flowers and leaves taken from a population of *E. scottica* in Berwickshire. For each individual the (A) sepal, (B) upper face of the final flowering node leaf, (C) lower face of the final flowering node leaf, (D) lower corolla lip and (E) upper corolla lip were mounted.

Table 4:1. Regional sampling of *E. scottica* and *E. micrantha* within Scotland. Adjacent populations have the same site name.

Site Name	Site Code	Collection Area	Species sampled	UK Grid Reference	Elevation (m)	Date visited	No. Individuals Sampled	Approx. Size of site (m)
Brock's Cleugh	BEs_a	Berwickshire	*E. scottica*	NT6160558591	372	13.07.09	31	10 x 15
Upper Stot Cleugh	BEs_b	Berwickshire	*E. scottica*	NT6343359371	424	13.07.09	32	10 x 10
Hell's Cleugh	BEs_c	Berwickshire	*E. scottica*	NT7404955056	318	15.07.09	32	10 x 10
Twinlaw Ford	BEm_a	Berwickshire	*E. micrantha*	NT6482255960	343	13.07.09	32	20 x 20
Herron Scar	BEm_b	Berwickshire	*E. micrantha*	NT6087258733	395	15.07.09	32	10 x 20
Lower Stot Cleugh	BEm_c	Berwickshire	*E. micrantha*	NT6376258124	326	15.07.09	32	1 x 20
Schiehallion	PEs_a	Perthshrie	*E. scottica*	NN7511955203	419	10.08.09	37	50 x 50
Meall Odhar Quarry	PEm_a	Perthshire	*E. micrantha*	NN7771054180	437	10.08.09	37	5 x 1 (x3)
Coire Screabaig	PEsm_a	Perthshire	*E. scottica*	NN7807754186	461	11.08.09	32	20 x 5
Coire Screabaig	PEms_a	Perthshire	*E. micrantha*	NN7807754186	461	11.08.09	32	20 x 5
Ben Lawres	PEsm_b	Perthshire	*E. scottica*	NN5871641776	580	12.08.09	34	40 x 10
Ben Lawres	PEms_b	Perthshire	*E. micrantha*	NN5871641776	580	12.08.09	32	10 x 10
Upper Howbog	CEs_a	Grampians	*E. scottica*	NJ4047626255	456	14.08.09	32	20 x 30
Loch An Eilein	CEs_b	Cairngorms	*E. scottica*	NH 9038 0775		14.08.09	32	20 x 30
A9, Lay by 92	CEs_c	Cairngorms	*E. scottica*	NN6482385893	415	07.08.10	32	5 x 5
Creag Meagaidh	CEs_d	Cairngorms	*E. scottica*	NN4691588929	536	13.08.10	32	30 x 30
Leids Hill	CEm_a	Grampians	*E. micrantha*	NJ4146726143	467	13.08.09	32	2 x 15
Suie Wood	CEm_b	Grampians	*E. micrantha*	NJ5470223202	443	13.08.09	32	1 x 300
Creag Meagaidh	CEm_c	Cairngorms	*E. micrantha*	NN4704188688	536	13.08.10	32	2 x 30
Gleann Leireag	SEs_a	Sutherland	*E. scottica*	NC1583630836	120	12.08.10	32	2x5 & 2x10
Rhiconich	SEm_a	Sutherland	*E. micrantha*	NC2558052620	124	12.08.10	32	2 x 20
Newton	SEsm_a	Sutherland	*E. scottica*	NC2318531426	128	09.08.10	32	30 x 30
Newton	SEms_a	Sutherland	*E. micrantha*	NC2318531426	128	09.08.10	32	1 x 50
Loch Merkland	SEsm_b	Sutherland	*E. scottica*	NC3739133160	175	11.08.10	32	30 x 30
Lock Merkland	SEms_b	Sutherland	*E. micrantha*	NC3739133160	175	11.08.10	32	3 x 20

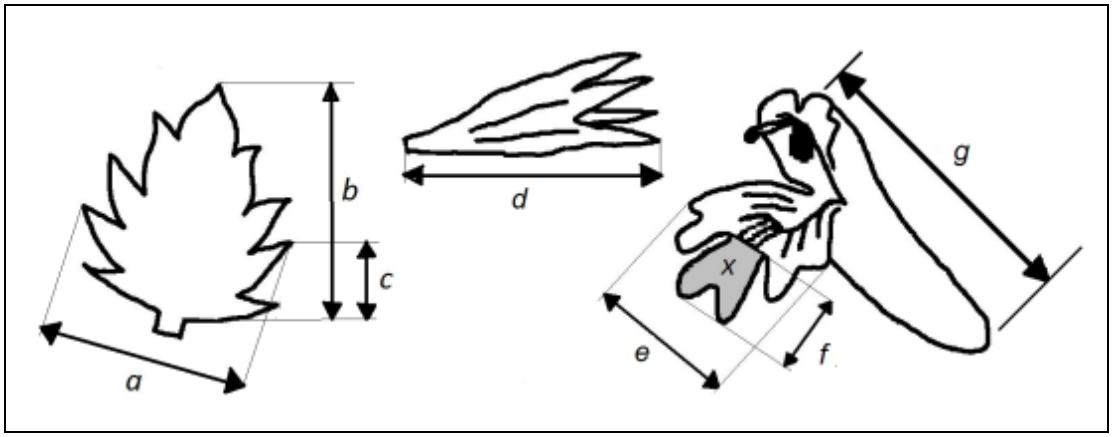


Figure 4:5. Measurements used in PCA of morphology. Relative size indices were used for the following measures: c/b, b/a, e/f and g/f. Area of lower corolla lip is indicated by x. Angle of leaf tip also included as a shape measure. Measurements taken using AxioVision Rel. 4.8.

August and September 2009 with a total of 90 *E. micrantha* and 60 *E. scottica* seeds collected. *E. micrantha* seed were collected from populations BEm_b (40) and BEm_c (40) in Berwickshire and PEms_b (10) in Perthshire. *E. scottica* seeds were collected from the population BEs_a (60) in Berwickshire. Five seeds were sown to a pot in late December 2009 and left to winter outside. The pots were brought into an unheated greenhouse in February 2010. Flower and leaf samples were taken for morphological analysis according to the procedure detailed from population collections in early July 2010. In total 14 *E. micrantha* (5 from BEm_b, 6 from BEm_c and 3 from PEms_b) were collected and pooled for morphological analysis under the population name Mx. Only 4 *E. scottica* were collected due to low levels of germination and survival to adulthood and were labelled collectively as Sx. These morphological data were used only for presentation in the graph of the principal component analysis and not in the ANOVA. As in Chapter 3 no hosts were used in this common garden experiment as pilot studies showed that this was the best method for cultivation.

4.2.6. Molecular analysis

4.2.6.1. DNA extraction

DNA was extracted from between 6 and 12 silica dried leaves using the method as described in Chapter 3.

4.2.6.2. Chloroplast sequencing

Between 4 and 10 individuals from each population were sequenced at the trnL$^{(UAG)}$-rp132 locus (Shaw et al. 2005). Approximately 4ng of DNA was used for the PCR of the trnL-UAG region. Amplification reactions were performed in 10 µL, with 1 µL 10 × *Taq* buffer (Bioline: 16 mM $(NH_4)_2SO_4$,), 3 mM $MgCl_2$, 0.2 mM dNTPs, 0.1 µM of forward and reverse primer, 2 µg bovine serum albumin (Promega) and 1 unit of *Taq* DNA polymerase (Bioline). Thirty thermal cycles were carried out for amplification with an annealing temperature of 50^{o}C for 30 s, and an extension temperature of 72^{o}C for 45 s. PCR products were run out on 2% agarose (Bioline) and visualised under ultra-violet light after staining with safeview (NBS Biologicals). PCR products were purified using ExoSap and sequenced with a Sanger sequencer. Fragments were sequenced

only in the reverse direction as the primers provided a good quality read along the whole of the sequence length.

4.2.6.3. Microsatellite genotyping

Eight microsatellite loci (Earc1, Earc2, Earc3, Earc4, Earc5, Earc6, Earc7 and Earc8) were amplified using primers and PCR conditions as described in Chapter 2, along with two other microsatellite loci *Ene1* and *Ene5* as described by French et al. (2003) (Table 4:2). The same PCR conditions as described in Chapter 2 were used for the French et al. (2003) loci. The PCR products were run out on an ABI 377 DNA sequencer, followed by allele size scoring using GENEMAPPER software (v 4.0) (PE Applied Biosystems).

Table 4:2. Summary of French et al. (2003) primers used Forward primers have an M13 5' tail for the purpose of genotyping; T_M = annealing temperature; *A* = number of alleles

Locus	Repeat Motif	Primer Sequence (5'-3')	T_M (^oC)	Size (bp)	*A*
Ene1	$(GT)_7$	F: AATTCCATCACTGCCAGAAAGAA R: TGCACAAACACTCCCTAAGTTTG	60	165 – 173	4
Ene5	$(CA)_7$	F: CCGTCACATACTCACATTACACA R: CCATTGACTTCGATTTGAAGATT	60	129 – 147	7

4.2.7. Analysis of molecular data

4.2.7.1. Chloroplast analysis

Sequences were aligned, and polymorphisms detected using the software Sequencher (Gene Codes Inc., Ann Arbor, MI, USA). Sequence editing and haplotype formation were performed manually. The overall diversity of chloroplast haplotypes within and between populations of each species were estimated by calculating total diversity (h_T) and average intrapopulation diversity (h_S) following the methods described by Pons and Petit (1995), using the computer software Haplodiv. Measures of population differentiation G_{st} and N_{st} were calculated using the software Haplonst (Pons & Petit 1996). To test for the presence of phylogenetic structure between populations, comparisons of G_{st} and N_{st} were made. An analysis of molecular variance (AMOVA) was performed to test the null hypotheses of no difference in chloroplast sequence between populations, regions and species using the computer software program ARLEQUIN v3.1

(Excoffier et al. 2005). The same program was used to create a minimum spanning network indicating the relatedness between haplotypes.

4.2.7.2. Microsatellite analysis

Since these two tetraploid species have been shown to be highly inbreeding and have disomic inheritance with overlapping isoloci it was not possible to use currently available software for analysing polyploid microsatellites. Estimation of allele copy number and hence frequency is difficult or impossible to determine in polyploids because inheritance patterns are complex (Bruvo et al. 2004; De Silva et al. 2005; Clark & Jasieniuk 2011). Disomic inheritance in polyploids presents a computational challenge as scored alleles can represent two or more independent loci (Clark & Jasieniuk 2011). Software has been created that can estimate allele frequency in allotetraploids such as TETRASAT (Markwith et al. 2006), TETRA (Liao et al. 2008), and ATETRA (Van Puyvelde *et al.*), and population genetic parameters can then be calculated using the software POLYSAT (Clark & Jasieniuk 2011). This is only appropriate, however, for allopolyploids where isoloci have no alleles in common. An added complication is that the French et al. (2003) primers seem to amplify in a diploid manner across all populations and would therefore have to be analysed separately.

It was decided, therefore, to score the genotyping output of the microsatellites as RFLP-type scores coding each allele as a dominant marker. Every allele detected at each locus was scored as either present (1) or absent (0), creating a microsatellite "multilocus phenotype". Obviously within a locus there was not complete independence of allele scores as there are necessarily a maximum of four allele scores for each individual. This approach is therefore not ideal and much information is lost, but it is the best solution for this analysis. These RFLP-type scores, or microsatellite haplotypes, were then analysed using the software ARLEQUIN v3.1 to perform an AMOVA testing the null hypothesis of no difference between populations, regions and species. ARLEQUIN v3.1 also provides pairwise measures of genetic distance between populations (F_{st}). A multiple regression of pairwise geographical distance and ecological distance against pairwise F_{st} was performed using the statistical package R in order to determine if there was any evidence for different ecotypes within species or any geographic structuring of genetic variation.

A principal coordinate analysis was performed using the software PAST (Hammer et al. 2001) along with analysis using the Bayesian population assignment software STRUCTURE Version 2.3 (Falush et al. 2007) to look for a pattern of genetic structuring between and within species and regions from two fundamentally different perspectives. For the STRUCTURE analysis, estimates of K (the number of structural units) represent values that most accurately capture the major structure in our collections. That is, the largest value of K, beyond which further increases resulted in the random subdivision of previously defined structural units. Between two and four independent runs with a burnin period of 40000 and with 40000 MCMC repeats after burnin were performed for each value of K in order to confirm that the STRUCTURE output remained consistent.

4.3. Results

4.3.1. Morphological analysis of flower and leaf

In the principal component analysis of the morphological data, 37% of the variation was explained by principal component 1 (PC1), falling to 19% and 10% in principal components 2 and 3 respectively. It was therefore decided to further analyse PC1 only. Analysis of variance (ANOVA) showed significant partitioning PC1 variation between species explaining 77.52% of overall variation (Table 4:3). Grouping by region explained a significant, but much smaller proportion of partitioning of PC1 variation explaining only 1.51% (Table 4:3). A significant partitioning of PC1 variance was also detected among populations, accounting for 5.12% of the total variation (Table 4:3). *E. micrantha* (Mx) and *E. scottica* (Sx) grown in a common environment retain their distinct morphological appearance (Figure 4:6).

The most influential factors in the principal component analysis were upper and lower corolla flower colour (Table 4:4). Flower corolla lip area, the ratio of upper to lower corolla lip lengths and the ratio of lower corolla lip width and lower corolla lip length were also influential (Table 4:4). *E. micrantha* tended to have a much larger (both wider and longer) lower corolla lip than flowers of *E. scottica*. Leaf morphology did not have a large effect on PC1, with the most influential leaf factor being leaf tip angle (Table 4:4).

Table 4:3. Analysis of variance (ANOVA) of PC1 for morphology among *E. micrantha* and *E. scottica* and among the four collection regions: Berwickshire, Perthshire, the Cairngorms and Sutherland. d.f. = degrees of freedom; SSq = sums of squares (**p < 0.001).

Source of Variation	d.f.	SSq	Variance components	% total Variation
Among species	1	2011.16	2011.16	77.52**
Among regions	3	85.7	28.57	1.51**
Among populations	20	226.4	11.32	5.12**
Within populations	677	418.22	0.62	15.86
Total	701	2741.48	2051.67	100

Table 4:4. Weighting of morphological measurements included in principal component 1 (PC1) of the principal component analysis. Colour coordinates Y and y refer to the 'Commission internationale de l'éclairage' (CIE) colour space. ':' indicates a ratio.

Rotation	PC1
Upper Leaf Colour (Y)	0.05335808
Upper Corolla Lip Colour (y)	-0.406616654
Upper Corolla Lip Colour (Y)	-0.411275255
Lower Corolla Lip Colour (y)	-0.459339891
Lower Corolla Lip Colour (Y)	-0.465364367
Leaf length:Leaf width	0.027040445
Leaf length:Leaf length to widest point	0.009642227
Leaf Tip Angle	-0.101692785
Sepal Length	-0.001320837
Lower Corolla Lip Area	0.301187352
Upper Corrola Lip Length:Lower Corolla Lip Length	-0.305192113
Lower Corolla Lip Width:Lower Corolla Lip Length	-0.206376487

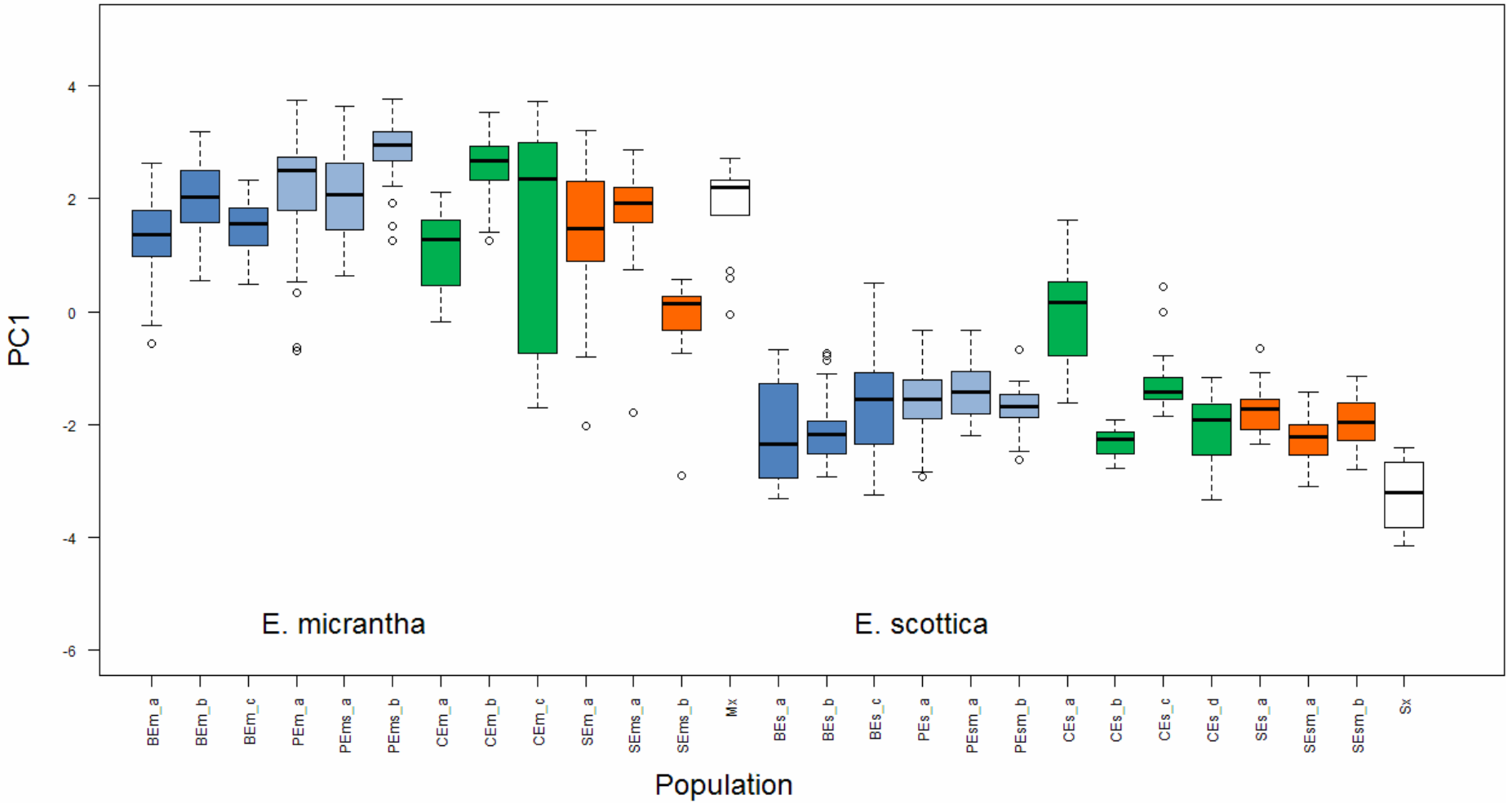


Figure 4:6. Box plots of principal component 1 (PC1) values based on morphological characters for individuals from the 25 populations collected. Population names correspond to the code in Table 4:3 (methods section). Two further populations, one *E. micrantha* (Mx) and one *E scottica* (Sx) are included in this graph to illustrate the morphology of species when grown in a common greenhouse environment. The colour of a box plot represents the region in which the population was collected: Berwickshire (dark blue), Perthshire (light blue), Cairngorms (green) and Sutherland (orange).

4.3.2. Habitat analysis

The sites occupied by *E. micrantha* and *E. scottica* differed significantly in mean Ellenberg values according to three environmental variables (Figure 4:7), light, moisture and nitrogen values (Light: F-value of 36.48, $p < 0.001$; Moisture: F-value of 107.47, $p < 0.001$; Reaction: F-value of 2.14, $p > 0.05$; Nitrogen: F-value of 10.74, $p < 0.05$; and Salt: F-value of 3.23, $p > 0.05$) (Table 4:5). Regions differed significantly in mean nitrogen and moisture values (Light: F-value of 2.51, $p > 0.05$; Moisture: F-value of 5.81, $p < 0.05$; Reaction: F-value of 1.37, $p > 0.05$; Nitrogen: F-value of 4.84, $p < 0.05$; and Salt: F-value of 1.57, $p > 0.05$) (Table 4:5). There was no evidence of a species region interaction (Table 4:5). Whilst there were no significant differences among *E. micrantha* populations between regions, the *E. scottica* sites differed significantly among regions in mean moisture, reaction and nitrogen values (Light: F-value of 0.9807, $p > 0.05$; Moisture: F-value of 5.7314, $p < 0.05$; Reaction: F-value of 4.6719, $p < 0.05$; Nitrogen: F-value of 5.5863, $p < 0.05$; Salt: F-value of 0.6509, $p > 0.05$). *E. micrantha* sites tended to have partial shade with relatively infertile, acid soils of average dampness. *E. scottica* sites, however, were more well lit with infertile, acid and wet soils. Berwickshire sites were on average more fertile and drier than sites in Perthshire, the Cairngorms or Sutherland. Across the regions *E. scottica* sites were drier in Berwickshire, and became less acidic and less fertile further north.

4.3.3. Molecular analysis

4.3.3.1. Chloroplast sequence analysis

Sequencing of the trnL$^{(UAG)}$-rp132 locus resulted in the detection of 20 sequence differences which included point mutations (polymorphic sites 1, 44, 87, 89, 119, 163, 199, 235, 291, 302, 433, 481, 589 and 600), indels (polymorphic sites 239, 339, 393, 422 and 628) and an inversion (polymorphic site 516) (Table 4:6). In total 18 haplotypes were identified across the 25 populations.

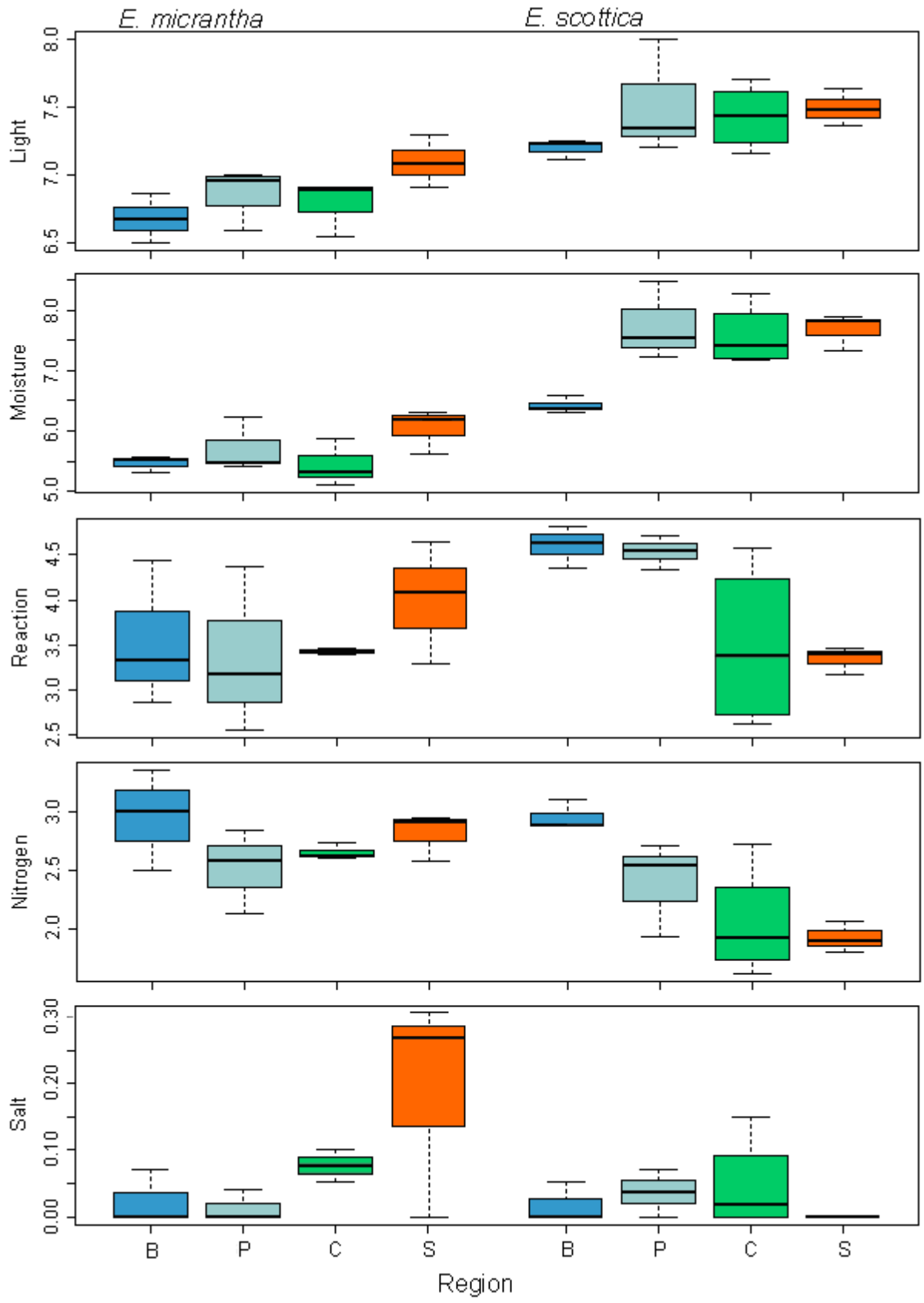


Figure 4:7. Mean Ellenberg values for five environmental variables, calculated from the associated species of each collection site. Mean Ellenberg values are shown for *E. micrantha* and *E. scottica* populations grouped within thier respective regions, Berwickshire (B), Perthshire (P), the Cairngorms (C) and Sutherland (S).

Table 4:5. Analysis of variance (ANOVA) of mean Ellenberg values for five environmental variables, among and within *E. micrantha* and *E. scottica* and among the four collection regions: Berwickshire, Perthshire, the Cairngorms and Sutherland. d.f. = degrees of freedom; MS = mean square (**p < 0.001; *p < 0.05, NS = not significant).

Source of Variation	d.f.	MS Light	MS Moisture	MS Reaction	MS Nitrogen	MS Salt
All Populations across all regions						
Among species	1	1.996**	18.6600**	0.08372^{NS}	1.14083*	0.015281^{NS}
Among regions	3	0.13538^{NS}	1.0086*	0.5339^{NS}	0.51355*	0.00744 ^{NS}
Species:Region interaction	3	0.02438^{NS}	0.4443^{NS}	1.1346^{NS}	0.26449^{NS}	0.013752^{NS}
Residuals	17	0.0539	0.1735	0.3904	0.10619	0.004738
***E. micrantha* only**						
Among regions	3	0.09591^{NS}	0. 24145^{NS}	0. 2475^{NS}	0.10913^{NS}	0.019779^{NS}
Residuals	8	0.04130	0.12082	0.4875	0.09094	0.007623
***E. scottica* only**						
Among regions	3	0.06384^{NS}	1.2115*	1.4210*	0.66891*	0.001414^{NS}
Residuals	9	0.06510	0.2114	0.3042	0.11974	0.002173

Table 4:6. Polymorphic sites detected in the trnL$^{(UAG)}$-rp132 chloroplast locus sequence.

Haplotype	Relative codon position																			
	1	44	87	89	119	163	199	235	239	291	302	339	393	422	433	481	516	589	600	628
h01	C	C	G	G	C	T	A	T	:	T	G	:	G	:	T	C	TTACT	G	G	T
h02	C	A	G	G	C	T	A	T	:	T	G	:	G	:	T	C	AGTAA	G	G	T
h03	T	A	G	G	C	T	C	C	:	T	T	:	G	:	G	C	TTACT	A	G	T
h04	T	A	G	G	C	T	C	C	:	T	T	:	G	:	G	C	AGTAA	A	G	T
h05	C	A	G	G	C	T	A	T	:	T	G	:	G	:	T	C	AGTAA	G	T	T
h06	C	A	G	G	C	T	A	T	:	T	G	:	G	:	T	C	AGTAA	G	G	:
h07	C	A	G	G	C	T	A	T	:	T	G	T	G	:	T	C	TTACT	G	G	T
h08	C	A	G	G	C	T	A	T	:	T	G	:	G	:	T	C	TTACT	G	G	T
h09	C	A	G	G	C	T	A	T	:	G	G	:	G	:	T	C	AGTAA	G	G	T
h10	T	A	G	G	C	T	C	C	:	T	T	:	G	:	G	C	AGTAA	G	T	T
h11	C	A	G	G	C	T	A	C	:	T	G	:	:	:	T	A	AGTAA	G	G	T
h12	C	A	A	A	C	T	A	C	:	T	G	:	:	A	T	C	AGTAA	G	G	T
h13	C	A	G	G	C	G	A	T	:	T	G	:	G	:	T	C	AGTAA	G	G	T
h14	T	A	G	G	C	T	C	C	A	T	T	:	G	:	G	C	TTACT	A	G	T
h15	T	A	G	G	T	T	C	C	:	T	T	:	G	:	G	C	TTACT	A	G	T
h16	T	A	G	G	T	T	C	C	:	T	T	:	G	:	G	C	AGTAA	A	G	T
h17	C	A	G	A	C	T	A	C	:	T	G	:	:	:	T	C	AGTAA	G	G	T
h18	C	A	G	G	C	T	A	T	:	T	G	T	G	:	T	C	AGTAA	G	G	T

In general each population was predominantly fixed for one chloroplast haplotype with only 14.95% of total chloroplast variation witnessed within populations, compared to 58.44% of the variation partitioned between populations (Tables 4:7 & 4:8). Four haplotypes were shared between species. Haplotypes h05 and h11 were present in different species in different regions (h05: *E. scottica* in the Cairngorms and *E. micrantha* in Berwickshire; h011: *E. scottica* in the Cairngorms and *E. micrantha* in Perthshire). Haplotypes h03 and h15 were found in both species in Sutherland only. The remaining 14 haplotypes were species specific, with 9 haplotypes present only in *E. scottica* and 5 only in *E. micrantha*. Six haplotypes were distributed across more than one region, though only one (h03) was found in all four regions and mostly in *E. micrantha* but also in 1 *E. scottica*. Twelve haplotypes showed a geographically restricted distribution. Haplotype h01 was restricted to Berwickshire; h06, h07, h09, h12 and h18 restricted to Perthshire; h13 restricted to the Cairngorms and h14, h15, h16 and h17 restricted to Sutherland. All Sutherland *E. scottica* haplotypes seem to be more closely related to other *E. micrantha* haplotypes than the *E. scottica* from the other regions (Figure 4:8).

The total chloroplast diversity (h_T) among *E. scottica* populations is significantly higher than that of *E. micrantha* populations. *E. scottica* also have significantly higher intrapopulation diversity (h_S) than *E. micrantha*. Comparisons of G_{st} and N_{st} were made at three different levels, (i) among populations across all regions, (ii) among *E. micrantha* only and (iii) among *E. scottica* only. Significantly higher values of N_{st} than G_{st} indicate phylogenetic structure between populations. No significant difference was found between G_{st} and N_{st} in any of the comparisons (Table 4:8). It is interesting to note, however, that G_{st} is consistently lower than N_{st}.

In light of unique chloroplast types *E. scottica* found only in Sutherland, four analyses of molecular variance (AMOVA) were performed: (i) all populations in all regions (Table 4:9); (ii) all populations from Berwickshire, Perthshire and the Cairngorms (not Sutherland) (Table 4:10); (iii) *E. micrantha* populations only and (iv) *E. scottica* populations only (Table 4:11). An AMOVA across all populations in all regions showed that there is significant partitioning of genetic variation among species, explaining 26.61% of haplotype variation (Table 4:9).

Table 4:7. The number of chloroplast haplotypes detected by sequence differences in the trnL$^{(UAG)}$-rp132 chloroplast locus

Region	Species	Population	h01	h02	h03	h04	h05	h06	h07	h08	h09	h10	h11	h12	h13	h14	h15	h16	h17	h18	No. diff. haplotypes
Berwickshire	*E. micrantha*	BEm_a			5																1
Berwickshire	*E. micrantha*	BEm_b				5															1
Berwickshire	*E. micrantha*	BEm_c					5														1
Perthshire	*E. micrantha*	PEm_a			3							1	4								3
Perthshire	*E. micrantha*	PEms_a											4								1
Perthshire	*E. micrantha*	PEms_b												5							1
Cairngorms	*E. micrantha*	CEm_a			4																1
Cairngorms	*E. micrantha*	CEm_b			5																1
Cairngorms	*E. micrantha*	CEm_c			4											1					2
Sutherland	*E. micrantha*	SEm_a				4													1		2
Sutherland	*E. micrantha*	SEms_a			1	3											1				3
Sutherland	*E. micrantha*	SEms_b			5																1
Berwickshire	*E. scottica*	BEs_a	5																		1
Berwickshire	*E. scottica*	BEs_b	5																		1
Berwickshire	*E. scottica*	BEs_c		5																	1
Perthshire	*E. scottica*	PEs_a		3				4	2	1											4
Perthshire	*E. scottica*	PEsm_a								5											1
Perthshire	*E. scottica*	PEsm_b		1				1		6	1									1	5
Cairngorms	*E. scottica*	CEs_a		2			6														2
Cairngorms	*E. scottica*	CEs_b								5											1
Cairngorms	*E. scottica*	CEs_c											5								1
Cairngorms	*E. scottica*	CEs_d		4											1						2
Sutherland	*E. scottica*	SEs_a											1				4				2
Sutherland	*E. scottica*	SEsm_a			1												4				2
Sutherland	*E. scottica*	SEsm_b															3	2			2

Table 4:8. Parameters of chloroplast diversity among populations, total diversity (h_T) and intrapopulation diversity (h_S) calculated using the software Haplodiv (Pons & Petit 1995) and differentiation among populations (G_{st}) and (N_{st}) calculated using the software Haplonst (Pons & Petit 1996).

Source of Variation	h_T (se)	h_S (se)	G_{st} (se)	N_{st} (se)
Among all populations	0.914 (0.0226)	0.234 (0.0574)	0.744 (0.0623)	0.864 (0.0499)
Among *E. micrantha* populations	0.782 (0.0865)	0.181 (0.0814)	0.785 (0.1064)	0.823 (0.1099)
Among *E. scottica* populations	0.917 (0.0160)	0.283 (0.0815)	0.692 (0.0903)	0.855 (0.0590)

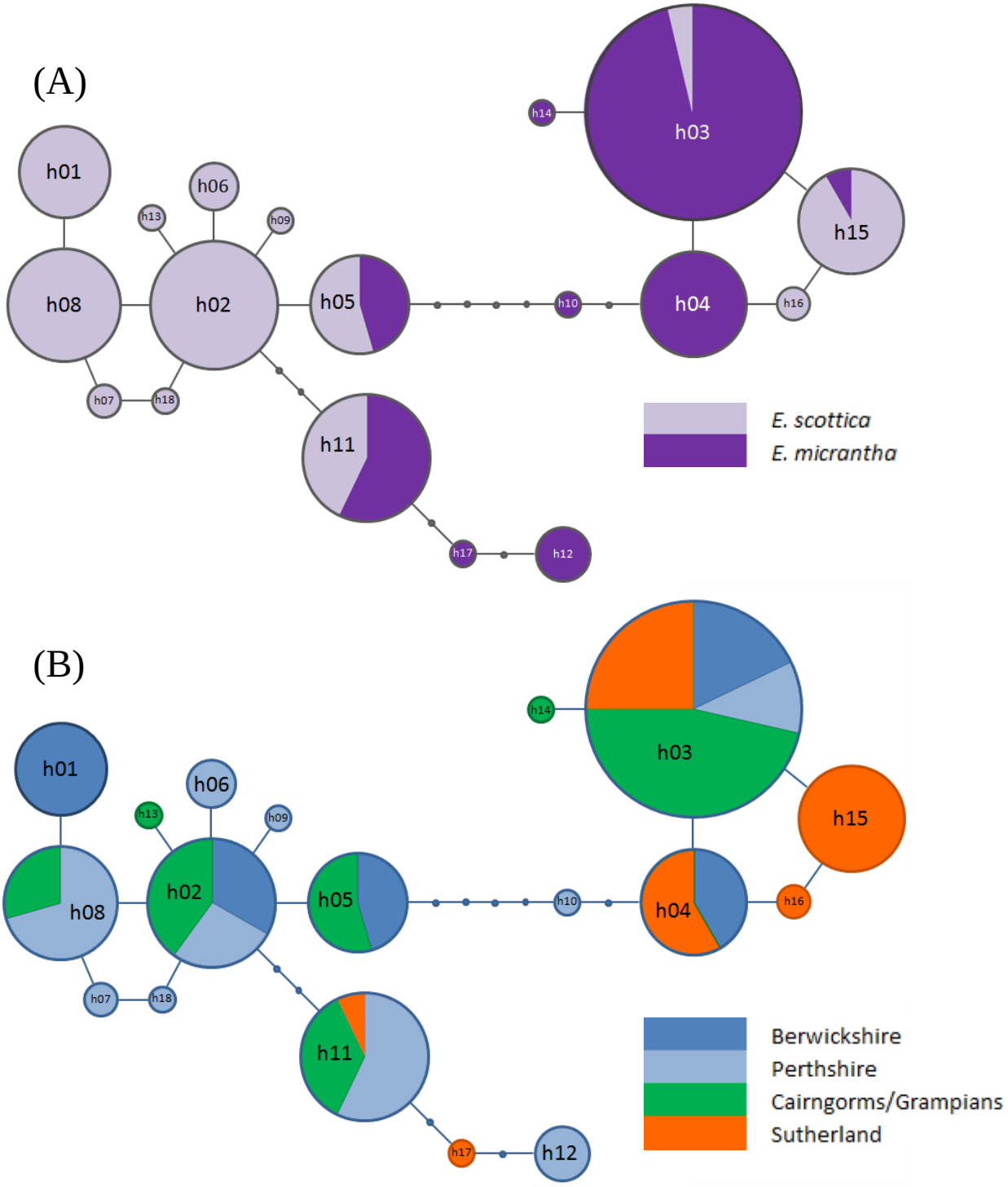


Figure 4:8. Minimum spanning network (MSN) representing chloroplast haplotype groups. Haplotypes are described in Table 4:6. Dots along the line indicate missing haplotypes in the network. The area of the circle represents the number of individuals scored with each haplotype. MSN (A) shows the differences in haplotype between *E. micrantha* and *E. scottica*. MSN (B) shows the region in which the haplotype was found.

Grouping populations by region also showed similar partitioning, explaining 22.46% of total haplotype variation. When Sutherland is removed from the analysis there is no longer any evidence of partitioning between the regions, but species boundaries still explain a significant and greater amount of partitioning in haplotype variation at 40.55% (Table 4:10). An individual AMOVA of each species shows that there is significant partitioning at the regional level in both species, but that it accounts for a smaller proportion of haplotype variation in *E. micrantha* than *E. scottica* at 30.23% and 61.95% respectively (Table 4:11). Accordingly there is a greater proportion of variation explained by differences between populations within a region in *E. micrantha* than *E. scottica* with 47.67% and 22.34% of variation explained respectively.

4.3.3.2. Microsatellite analysis

All microsatellite loci were polymorphic in both *E. micrantha* and *E. scottica* populations, with a total of 75 alleles detected (7 from *Earc1*, 9 from *Earc2*, 5 from *Earc3*, 7 from *Earc4*, 10 from *Earc5*, 4 from *Earc6*, 16 from *Earc7*, 7 from *Earc8*, 3 from *Ene1* and 7 from *Ene5*). Seven alleles were found to be limited to *E. micrantha* and 16 to *E. scottica* (Table 4:12). Most species limited alleles occurred at a very low frequency ($N \leq 10$) and were also found to be population limited. One allele specific to *E. micrantha* (*Earc4*, 149) was found across three populations in two different regions. Six alleles specific to *E. scottica* were found across more than one population: two were specific to *E. scottica* in Sutherland (*Earc1*, 299 and *Earc5*, 205), one specific to *E. scottica* in the Cairngorms (*Ene5*, 137) and three specific to *E. scottica* across four different populations and two regions (*Earc7*, 397, 407 and 412).

Genotyping of microsatellites resulted in the detection of 378 different molecular phenotypes. The vast majority of these were population restricted. Within population diversity was high. In one population (CEm_b) every individual had a different molecular phenotype. However in most populations there was one multilocus phenotype that occurred at high frequency along with a number of poorly represented multilocus phenotypes (Table 4:13). In general, however, *E. micrantha* showed lower levels of diversity in molecular phenotype than *E. scottica*. This is particularly obvious in Sutherland where all *E. micrantha* populations have fewer than 10

Table 4:9. Analysis of molecular variance (AMOVA) for the trnL$^{(UAG)}$-rp132 chloroplast locus including all populations sampled from all regions. d.f. = degrees of freedom; SSq = sums of squares (**p < 0.001; *p < 0.05).

	d.f.	SSq	Variance components	% total Variation
Populations grouped among species				
Among species	1	60.985	1.75162	26.61*
Among populations within species	23	218.094	1.65106	58.44**
Within populations	113	47.725	0.42235	14.95
Total	137	326.804	2.82502	100.00
Populations grouped among regions				
Among regions	3	86.989	0.57445	22.46*
Among populations within regions	21	186.479	1.54501	60.41**
Within populations	113	49.525	0.43827	17.13
Total	137	322.993	2.55774	100.00

Table 4:10. Analysis of molecular variance (AMOVA) for the trnL$^{(UAG)}$-rp132 chloroplast locus including populations sampled from Berwickshire, Perthshire and the Cairngorms only. d.f. = degrees of freedom; SSq – sums of squares (**p < 0.001; *p < 0.05).

	d.f.	SSq	Variance components	% total Variation
Populations grouped among species				
Among species	1	66.125	1.10926	40.55**
Among populations within species	17	126.780	1.258859	46.01**
Within populations	89	32.725	0.36770	13.44
Total	107	225.630	2.73555	100.00
Populations grouped among regions				
Among regions	2	20.552	0.00676	0.29
Among populations within regions	13	131.236	2.01472	87.64**
Within populations	62	17.200	0.27742	12.07
Total	77	17.200	2.29889	100.00

Table 4:11. Analysis of molecular variance (AMOVA) for the trnL$^{(UAG)}$-rp132 chloroplast locus for *E. micrantha* and *E. scottica* separately. d.f. = degrees of freedom; SSq – sums of squars (**p < 0.001; *p < 0.05).

	d.f.	SSq	Variance components	% total Variation
***E. micrantha* populations grouped among regions**				
Among regions	3	50.898	0.71302	30.23*
Among populations within regions	8	48.461	1.12425	47.67**
Within populations	48	25.025	0.52135	22.10
Total	59	124.383	2.35862	100.00
***E. scottica* populations grouped among regions**				
Among regions	3	89.299	1.37699	61.95**
Among populations within regions	9	29.437	0.49652	22.34**
Within populations	65	22.700	0.34923	15.71
Total	77	141.436	2.22274	100.00

Table 4:12. Species specific microsatellite alleles across all *E. micrantha* and *E. scottica* populations. N = number of individuals

Species	Locus	Species Limited Allele (allele size)	N	Population(s)
E. micrantha	*Earc1*	328	1	PEm_a
	Earc2	245	2	BEm_b
		249	9	CEm_c
	Earc4	149	37	CEm_b, CEm_c, SEms_b
	Earc6	230	1	CEm_b
	Earc8	207	10	PEm_a
	Ene5	139	1	SEm_a
E. scottica	*Earc1*	299	85	SEs_a, SEsm_a, SEsm_b
		315	2	CEs_a
	Earc3	328	5	SEsm_b
	Earc4	146	10	SEs_a
		154	2	BEs_c
	Earc5	205	2	SEsm_a, SEsm_b
	Earc7	397	1	PEs_a
		399	53	CEs_c, CEs_d, PEs_a, PEsm_a
		407	13	CEs_b, SEs_a, SEsm_a, SEsm_b
		408	6	BEs_c
		409	1	SEsm_a
		412	93	CEs_b, SEs_a, SEsm_a, SEsm_b
		416	4	SEs_a
	Ene5	129	1	BEs_c
		133	8	CEs_d
		137	18	CEs_a, CEs_b

different molecular phenotypes (Table 4:13). In general each molecular phenotype was only encountered once (mode = 1) (Table 4:13). This indicates a high level of genetic diversity within populations.

An AMOVA of the microsatellite multilocus phenotypes across all populations in all regions showed that there is significant partitioning of genetic variation among species, explaining 4.62% of haplotype variation (Table 4:14). This is very small but considering the amount of variation found within species, any significant difference found between species is pertinent. Grouping populations by region also showed significant partitioning explaining similar levels of haplotype variation at 6.86%. In accordance with the chloroplast data, an AMOVA was conducted without Sutherland. When Sutherland is removed from the analysis there is no longer any convincing evidence of a difference between regions or species (Table 4:15). An individual AMOVA of each species shows that there is significant partitioning at the regional level in both species, but that it accounts for a smaller proportion of haplotype variation in *E. micrantha* than *E. scottica* at 8.39% and 18.87% respectively (Table 4:16). Differences between populations account for the greatest proportion of variation explained by differences in microsatellite haplotypes at 52.99% among populations within species and 50.44% among populations within regions.

A significant ($p < 0.001$) sample regression coefficient was found between pairwise population differentiation (F_{st}) and geographic distance (km) between all populations across all regions, though it was a very weak signal, explaining only 5% of the total variation (Figure 4:9, A). No relationship was found between ecological distance and pairwise population differentiation (F_{st}) (Figure 4:10, A).

Between all populations of *E. micrantha* a significant sample regression coefficient was found between population differentiation and both geographical distance ($p < 0.05$) and ecological distance ($p < 0.01$), together explaining 22% of the observed variation in F_{st} (Figures 4:9, B and 4:10, B respectively). Within *E. micrantha*, if habitats are very similar, smaller values of F_{st} are observed than between highly heterogeneous habitats, even between regions, although there are a lot of scatter in the residuals. This is not observed in *E. scottica* where there is no significant relationship between population differentiation and geographical distance (Figure 4:9, C) and a

Table 4:13. Summary of the diversity of microsatellite multilocus phenotypes detected in each population. Highest freq. refers to the number of individuals in which the most common phenotype was found.

				Number of molecular phenotypes				Populations with shared molecular phenotypes (Number shared)
Region	Species	Population	No. individuals genotyped	Total	Highest freq.	Median	Mode	
Berwickshire	*E. micrantha*	BEm_a	25	11	8	1	1	PEs_a (2), PEsm_b (5)
Berwickshire	*E. micrantha*	BEm_b	32	9	18	1	1	none
Berwickshire	*E. micrantha*	BEm_c	31	8	17	1.5	1	none
Perthshire	*E. micrantha*	PEm_a	29	21	6	1	1	PEsm_b (6)
Perthshire	*E. micrantha*	PEms_a	27	17	4	1	1	none
Perthshire	*E. micrantha*	PEms_b	31	18	8	1	1	none
Cairngorms	*E. micrantha*	CEm_a	28	15	9	1	1	none
Cairngorms	*E. micrantha*	CEm_b	23	23	1	1	1	none
Cairngorms	*E. micrantha*	CEm_c	27	21	3	1	1	none
Sutherland	*E. micrantha*	SEm_a	25	5	18	1	1	none
Sutherland	*E. micrantha*	SEms_a	20	9	5	1	1	none
Sutherland	*E. micrantha*	SEms_b	24	9	10	2	1	none
Berwickshire	*E. scottica*	BEs_a	21	10	5	1	1	BEs_b (1)
Berwickshire	*E. scottica*	BEs_b	26	15	6	1	1	BEs_a (1)
Berwickshire	*E. scottica*	BEs_c	21	19	2	1	1	none
Perthshire	*E. scottica*	PEs_a	27	19	4	1	1	PEsm_b (3), BEm_a (2)
Perthshire	*E. scottica*	PEsm_a	29	19	8	1	1	none
Perthshire	*E. scottica*	PEsm_b	28	17	6	1	1	PEm_a (6), BEm_a (5), PEs_a (3)
Cairngorms	*E. scottica*	CEs_a	32	28	3	1	1	none
Cairngorms	*E. scottica*	CEs_b	22	20	2	1	1	none
Cairngorms	*E. scottica*	CEs_c	25	12	10	1	1	none
Cairngorms	*E. scottica*	CEs_d	25	9	10	2	1	none
Sutherland	*E. scottica*	SEs_a	30	16	4	1	1	none
Sutherland	*E. scottica*	SEsm_a	32	21	6	1	1	none
Sutherland	*E. scottica*	SEsm_b	31	23	7	1	1	none

Table 4:14. Analysis of molecular variance (AMOVA) for microsatellite multilocus phenotypes including all populations sampled from all regions. d.f. = degrees of freedom; SSq – sums of squars (**p < 0.001; *p < 0.05; *p < 0.1).

	d.f.	SSq	Variance components	% total Variation
Populations grouped among species				
Among species	1	216.684	0.32957	4.62**
Among populations within species	23	2398.699	3.77901	52.99**
Within populations	646	1953.076	3.02334	42.39**
Total	670	4568.459	7.13192	100.00
Populations grouped among regions				
Among Regions	3	546.020	0.48546	6.86*
Among populations within regions	21	2069.363	3.57139	50.44**
Within populations	646	1953.076	3.02334	42.70**
Total	670	4568.459	7.08019	100.00

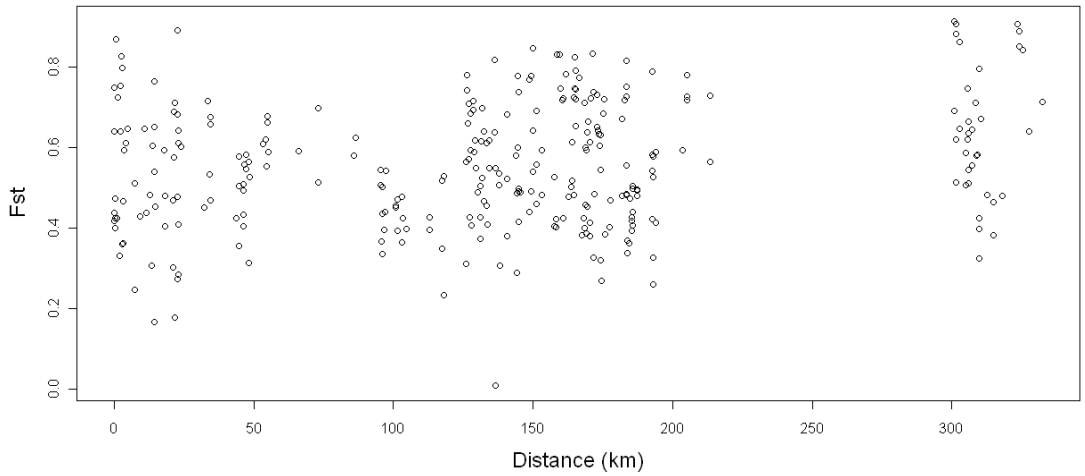
Table 4:15. Analysis of molecular variance (AMOVA) for microsatellite multilocus phenotypes including populations sampled from Berwickshire, Perthshire and the Cairngorms only. d.f. = degrees of freedom; SSq – sums of squares (**p < 0.001; *p < 0.05; $^{+}$p < 0.1).

	d.f.	SSq	Variance components	% total Variation
Populations grouped among species				
Among species	1	141.427	0.18792	2.93^{+}
Among populations within species	17	1563.432	3.33155	52.01**
Within populations	490	1414.182	2.88609	45.06**
Total	590	3119.041	6.40555	100.00
Populations grouped among regions				
Among Regions	2	248.306	0.18559	2.91^{+}
Among populations within regions	16	1456.553	3.30035	51.79**
Within populations	490	1414.182	2.88609	45.29**
Total	508	17.200	2.29889	100.00

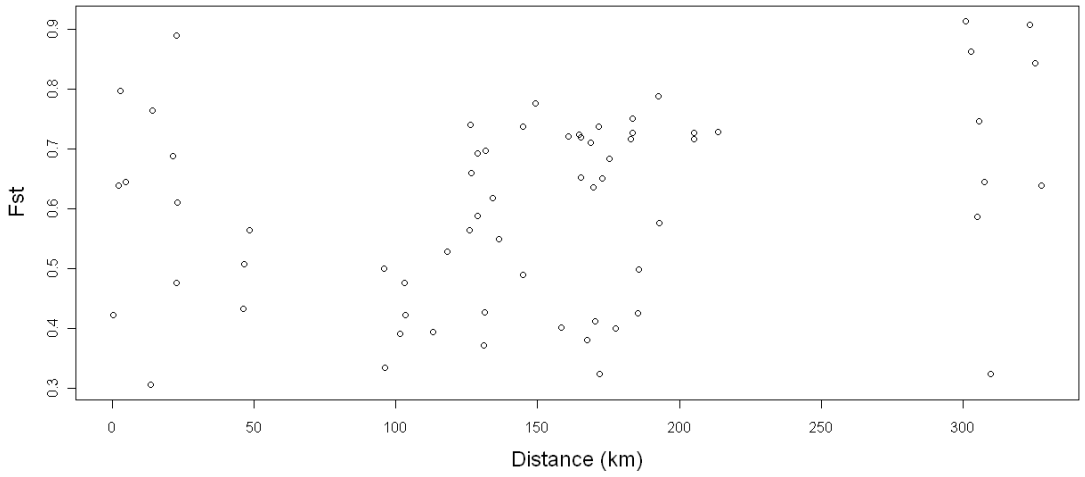
Table 4:16. Analysis of molecular variance (AMOVA) for microsatellite multilocus phenotypes for *E. micrantha* and *E. scottica* separately. d.f. = degrees of freedom; SSq – sums of squars (**p < 0.001; *p < 0.05; $^{+}$p < 0.1).

	d.f.	SSq	Variance components	% total Variation
***E. micrantha* populations grouped among regions**				
Among regions	3	453.369	0.57848	8.39*
Among populations within regions	8	830.996	3.79094	54.97**
Within populations	310	2067.534	2.51645	36.64**
Total	321	2067.543	6.89587	100.00
***E. scottica* populations grouped among regions**				
Among regions	3	540.147	1.33453	18.87**
Among populations within regions	9	574.217	2.25708	31.91**
Within populations	336	1169.887	3.48178	49.22**
Total	348	2284.241	7.07339	100.00

(A)



(B)



(C)

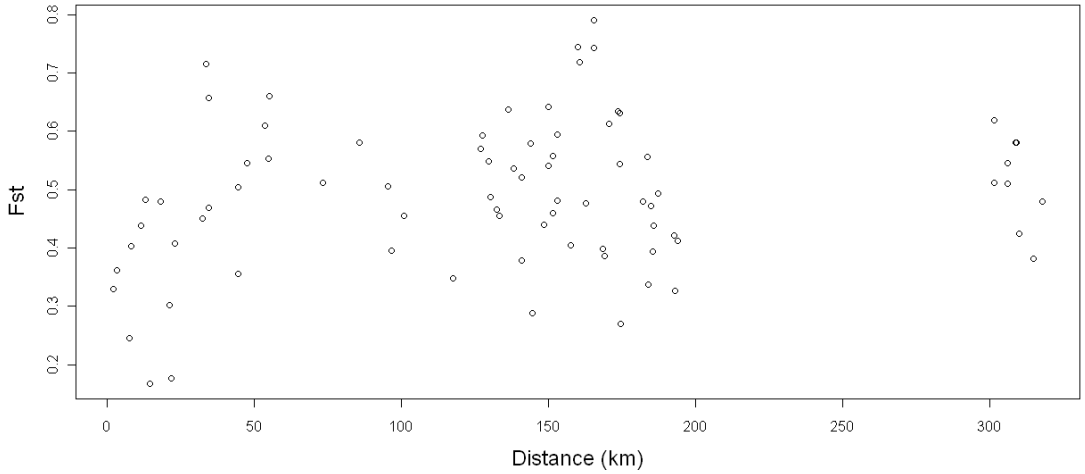


Figure 4:9. Scatter plot showing the regression between pairwise F_{st} and distance (km) between (A) all populations across all regions, (B) all *E. micrantha* populations across all regions and (C) all *E. scottica* populations across all regions.

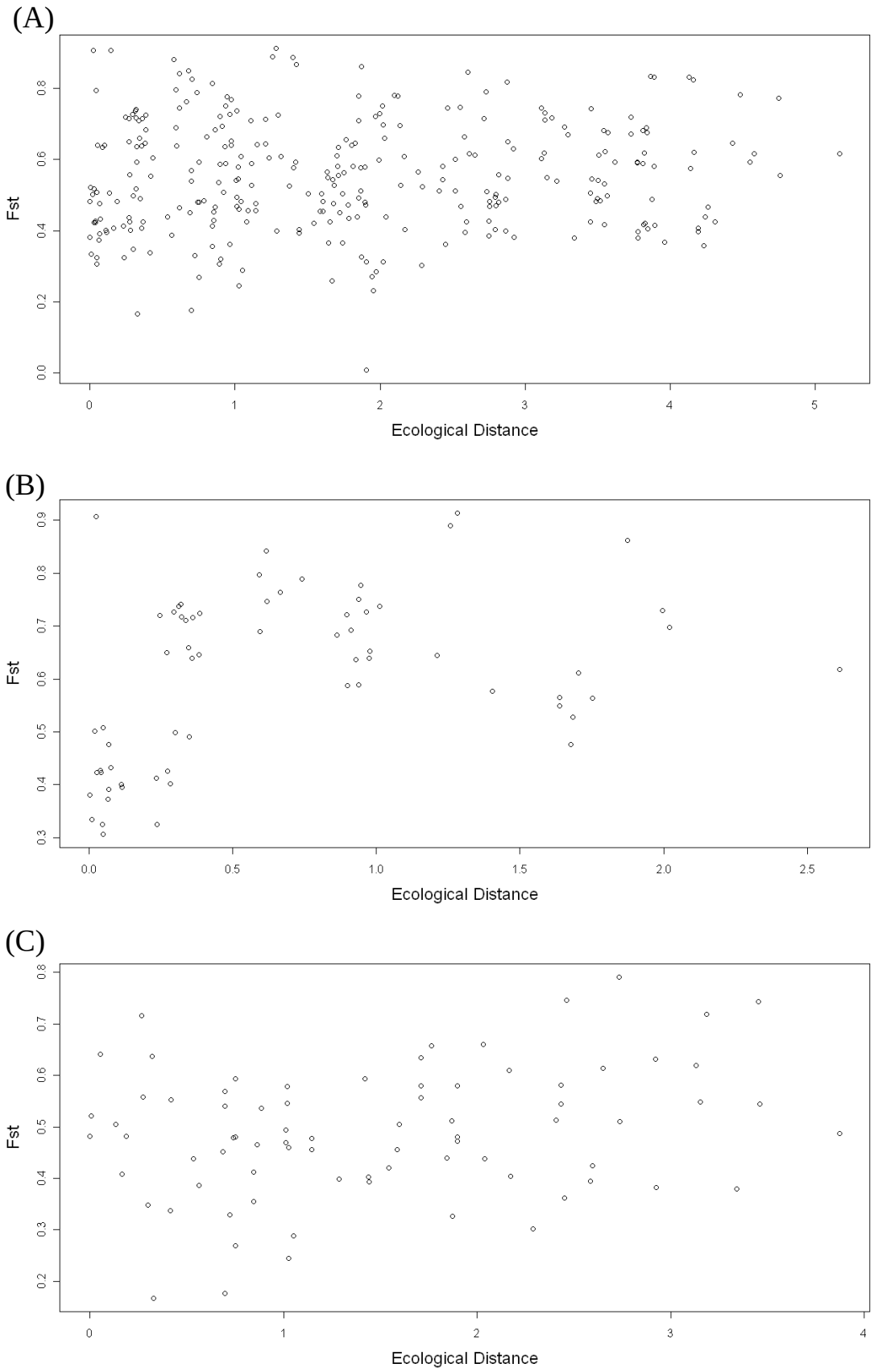


Figure 4:10. Scatter plot showing the regression between pairwise F_{st} and the pairwise ecological distance for (A) all populations across all regions, (B) all *E. micrantha* populations across all regions and (C) all *E. scottica* populations across all regions.

barely significant difference ($p < 0.08$) between ecological distance and population differentiation which explains only 8% of total variation (Figure 4:10, C). This is even though the range of ecological distance is larger in *E. scottica* than *E. micrantha* (Figure 4:10, B&C).

In the principal coordinate analysis looking at genetic variation in all populations across all regions neither species nor region groupings were discriminated along the 1st or 2nd coordinate (Figure 4:11), the first principal coordinate (PCO1) explaining 11.96% or the genetic variation and the second (PCO2) 9.78%.

Generally a smaller genetic distance was found between individuals within a species(region) group than between either different species within the same region, or between the same species from different regions. In other words, each species(region) group occurred in a distinctive but largely overlapping phenetic space (Figure 4:11). Within Berwickshire and Sutherland *E. micrantha* and *E. scottica* are largely non-overlapping and do not share a great deal of phenetic space. *E. scottica* from Perthshire stretches down into more *E. micrantha*-like phenetic space. *E. micrantha* and *E. scottica* from the Cairngorms however, have a much greater proportion of overlapping phenetic space with *E. micrantha* from the Cairngorms clustering in a much more *E. scottica*-like area.

Principal coordinate analyses of separate regions (Figure 4:12) show the same general pattern with discrete differences in phenetic space found between *E. scottica* and *E. micrantha* in Berwickshire and Sutherland (PCO1 = 28.00%; PCO2 = 20.98% and PCO1 = 28.57%; PCO2 = 15.36% respectively) (Figure 4:12). In both cases, there seem to be two separate groups of *E. micrantha*. In Perthshire, separate species group into distinct but largely overlapping phenetic space (PCO1 = 16.87; PCO2 = 15.49). In the Cairngorms *E. micrantha* clusters within the *E. scottica* which are widely distributed across the phenetic space (PCO1 = Figure 4:12).

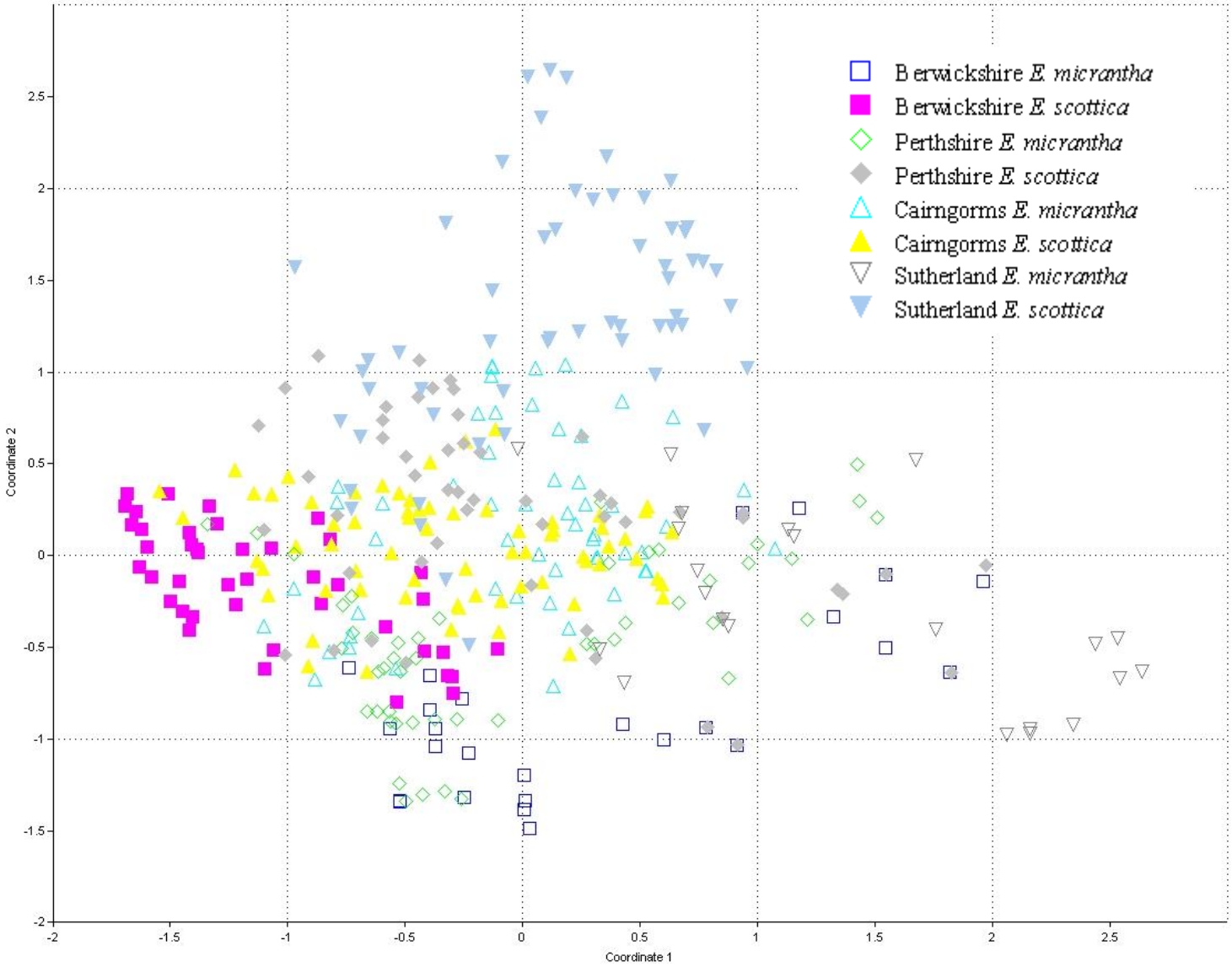


Figure 4:11. Principal coordinate analysis of genetic distance between *E. micrantha* and *E. scottica*, constructed from microsatellite multilocus phenotypes of all samples across all regions. The x axis (principal coordinate 1) explained 11.96% and the y axis (principal coordinate 2) 9.78% of the total variation. Regions are represented by shapes (square = Berwickshire, diamond = Perthshire, triangle = Cairngorms, inverted triangle = Sutherland), species by hollow (*E. micrantha*) or filled (*E. scottica*) points.

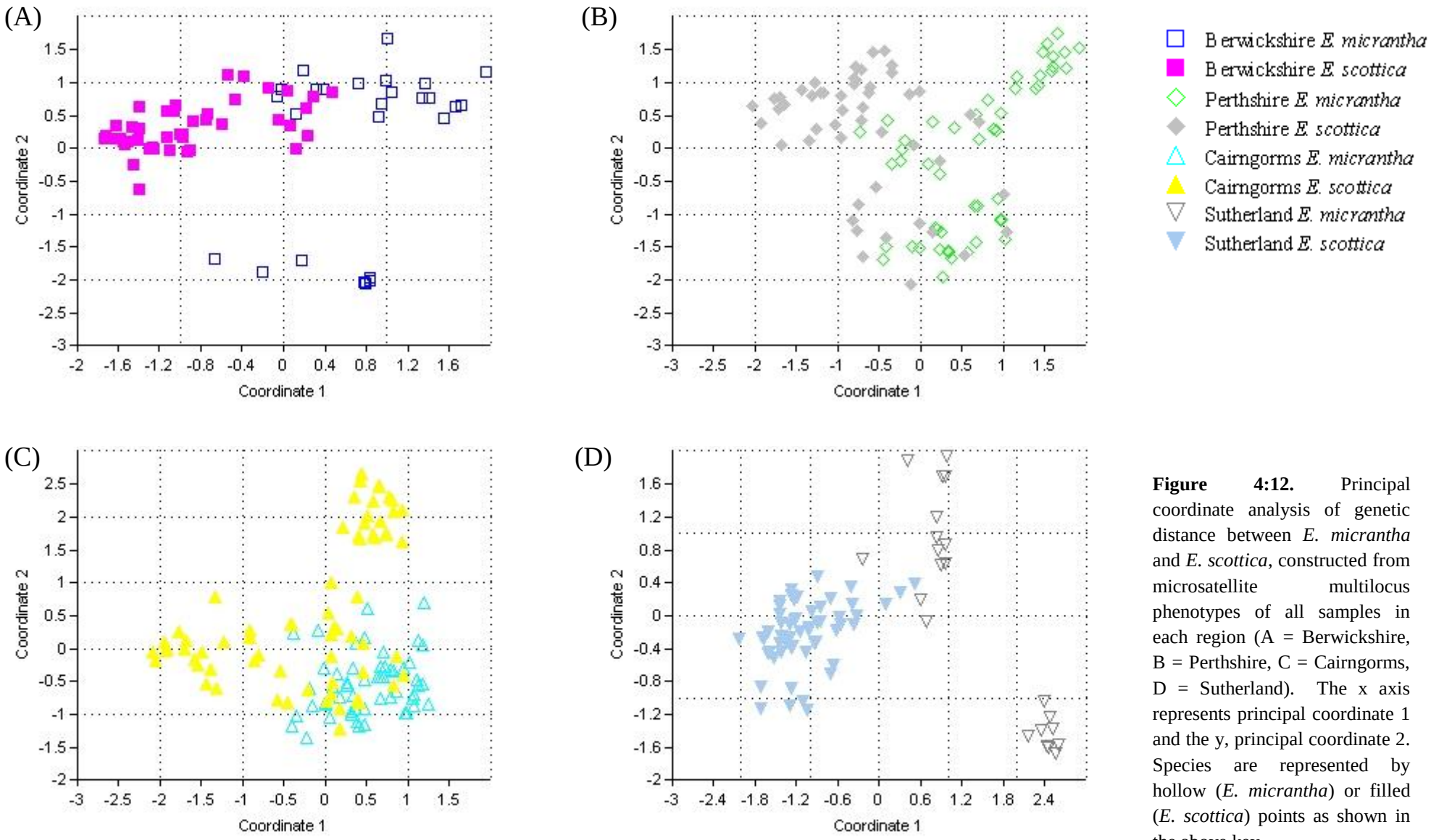
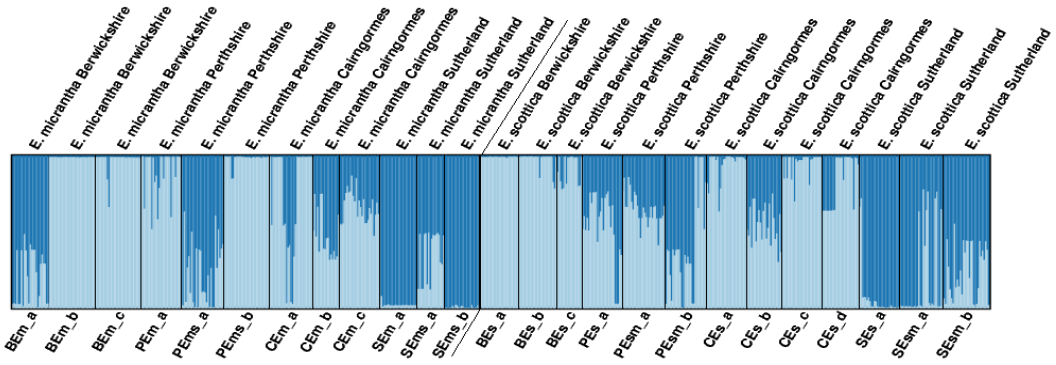


Figure 4:12. Principal coordinate analysis of genetic distance between *E. micrantha* and *E. scottica*, constructed from microsatellite multilocus phenotypes of all samples in each region (A = Berwickshire, B = Perthshire, C = Cairngorms, D = Sutherland). The x axis represents principal coordinate 1 and the y, principal coordinate 2. Species are represented by hollow (*E. micrantha*) or filled (*E. scottica*) points as shown in the above key.

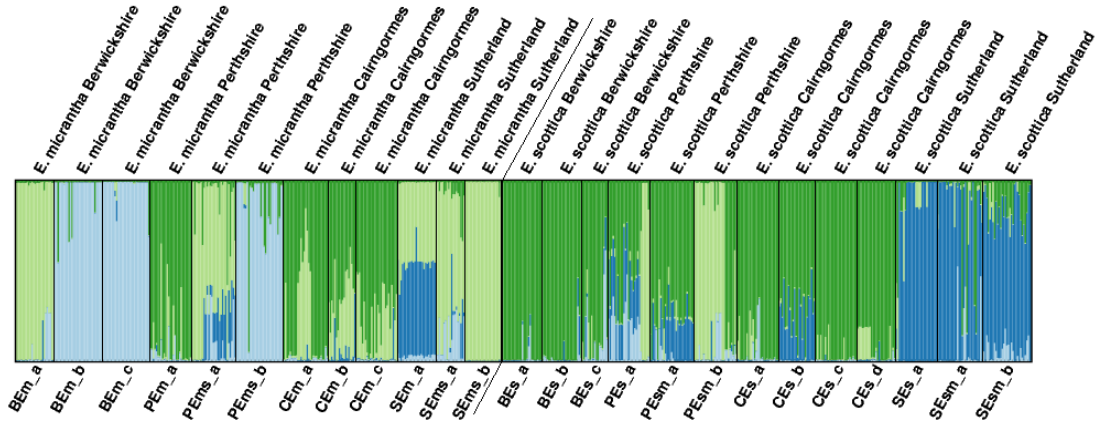
A STRUCTURE analysis of all populations across all regions does not distinguish between different species, regions or species within regions (Figure 4:13) with a best fit of K = 16. Greater structure appears to be distinguished at the population and not the species level, with separate populations coming out as separate entities. If two populations are found in the same structure group (cluster) they are more often populations of the same species from within the same region. For example, *E. scottica* from both Sutherland and Berwickshire appear to form a coherent cluster and there is some clustering of *E. micrantha* populations within Sutherland and Berwickshire. There appears to be less distinction between the populations in Perthshire and the Cairngorms. The STRUCTURE analysis therefore comes to a conclusion similar to that of the principal coordinate analysis (PCO).

The pattern determined by STRUCTURE analysis of separate regions also concurs with that found under PCO analysis (Figure 4:14). This is most obvious in Sutherland where *E. scottica* and *E. micrantha* separate neatly into two distinct clusters. In Berwickshire *E. scottica* forms one cluster whereas the *E. micrantha* are split into three separate entities (Figure 4:14). In Berwickshire however, one of the *E. micrantha* clusters (L Green, BEm_a) is less diverged from the *E. scottica* cluster than other *E. micrantha* groups (Table 4:17). In the Cairngorms, *E. micrantha* forms one cluster, similar to the tight clustering witnessed in the Principal Coordinate analysis. *E. scottica* are more distinct from one another, with one population of *E. scottica* clustering with the Cairngorms *E. micrantha* (Figure 4:14). The dark blue cluster, corresponding to *E. scottica* from CEs_c and some individuals in CEs_d, is the most divergent group in the Cairngorms (Table 4:17). In Perthshire, all *E. micrantha* populations come out as distinctly different from one another, and there is one *E. scottica* population that appears to cluster with *E. micrantha* better than with other *E. scottica* within the region (Figure 4:14).

Predicted K = 2



Predicted K = 4



Predicted K = 8

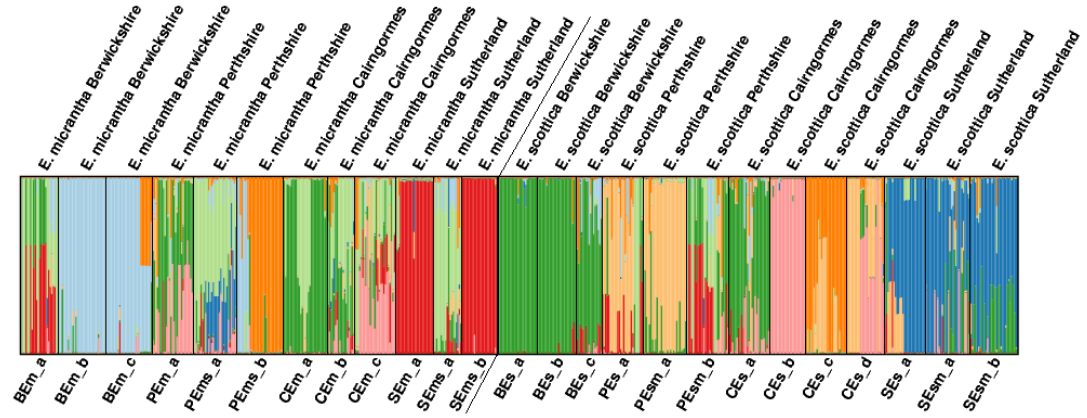
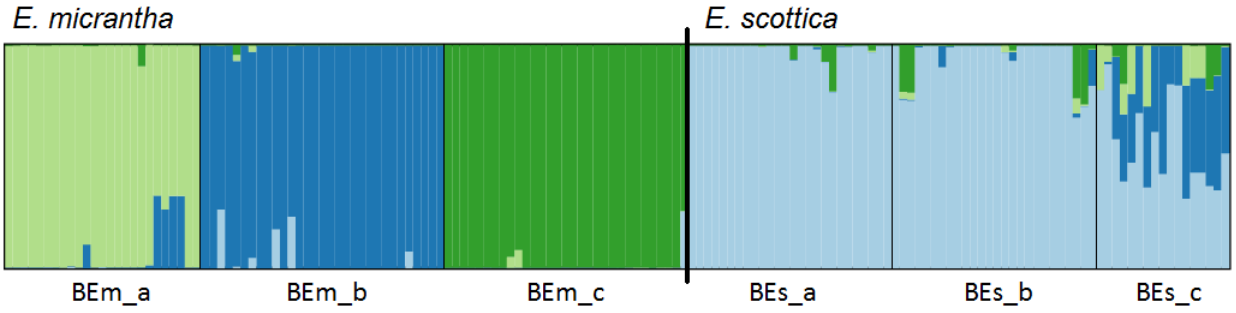
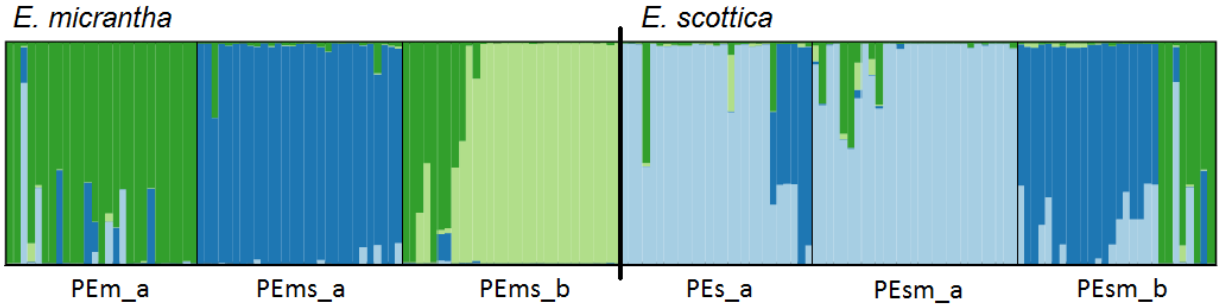


Figure 4:13. STRUCTURE analyses of all populations of *Euphrasia* across all regions. K = 2 corresponds to the prediction that variation in microsatellite multilocus phenotypes can be explained by separate *E. micrantha* and *E. scottica* species, K = 4 variation predicted to be within regions and K = 8 variation predicted to be in separate species within separate regions.

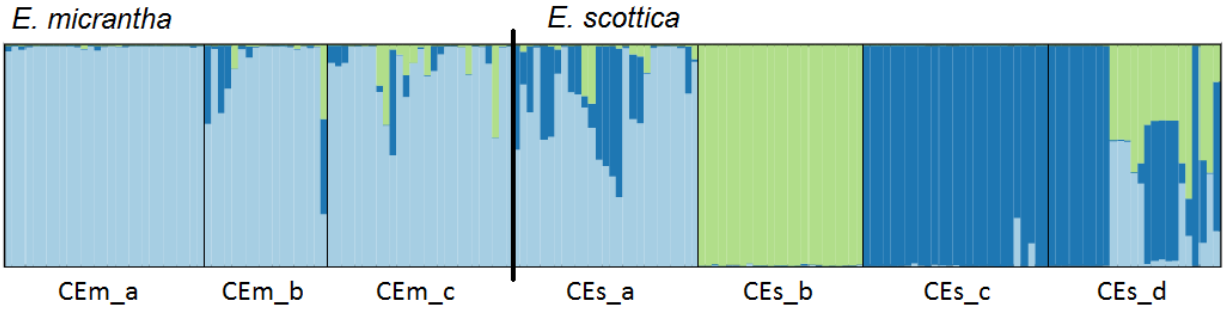
Berwickshire K = 4



Perthshire K = 4



Cairngorms K = 3



Sutherland K = 2

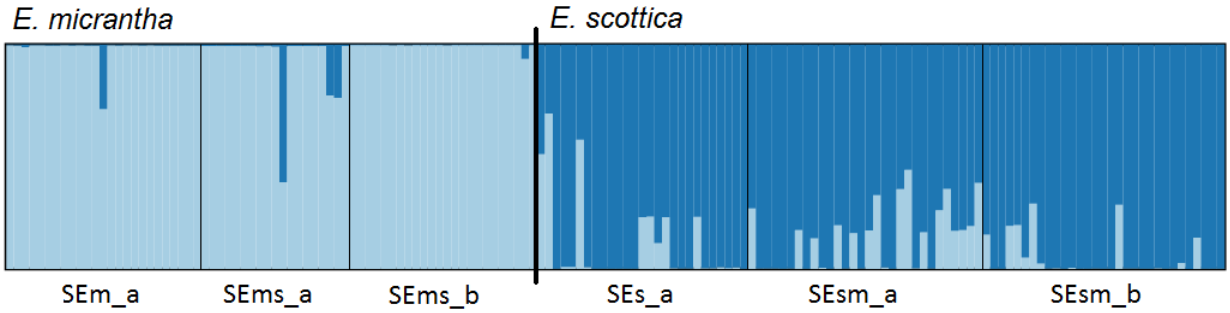


Figure 4:14. STRUCTURE analyses of all populations of *Euphrasia* within Berwickshire, Perthshire, the Cairngorms and Sutherland. The value of K corresponds to the most likely number of structural groups.

Table 4:17. Allele frequency divergence (net nucleotide distance) among STRUCTURE clusters grouped within regions. STRUCTURE groups are indicated by their colour.

Region	Net nucleotide difference			
Berwickshire		Dark Blue	Light Green	Dark Green
	Light Blue	0.1366	0.1146	0.1221
	Dark Blue		0.1268	0.1245
	Light Green			0.1476
Perthshire		Dark Blue	Light Green	Dark Green
	Light Blue	0.1065	0.1368	0.0751
	Dark Blue		0.1398	0.0719
	Light Green			0.1042
Cairngorms		Dark Blue	Light Green	
	Light Blue	0.0991	0.0763	
	Dark Blue		0.1111	
Sutherland		Dark Blue		
	Light Blue	0.1136		

4.4. Discussion

4.4.1. Natural morphology and habitats of *E. micrantha* and *E. scottica* in Scotland

Both *E. micrantha* and *E. scottica* are widely distributed throughout Scotland, and have overlapping distributions. The two species have been found to have very distinct morphologies that are largely consistent across their distribution. Species differences account for 78% of observed variation in morphology and regional differences only 2%. Even when individuals from the different species were planted in the greenhouse under common garden conditions, they maintained their morphological distinctness, a result that has previously been observed by Yeo (1962). It can be concluded, therefore, that *E. micrantha* and *E. scottica* form coherent morphological groups.

It was also shown that the morphology of individual plants varied greatly within populations, accounting for 16% of overall variation. It has been well recorded that the morphology of an adult plant in this hemi-parasitic group is strongly affected by the quality of alternative host plants (Yeo 1962; Seel & Press 1993, 1994; Matthies 1998; Svensson et al. 2001). Therefore, it is proposed that variation within populations may be explained largely by differential host

parasitism. Variation among populations accounted for 5% of observed variation. This may be due to differences in the ecology or availability of suitable hosts at the different sites.

Analysis of environmental variability showed that *E. micrantha* and *E. scottica* inhabit clearly distinct habitats, most notably with respect to levels of moisture and light. In sites where the two species are found growing adjacent to one another, neither species was found growing in the habitat of the other, even though seed dispersal into neighbouring sites is not improbable. This suggests that there is strong selection against germination and survival to adulthood for either species growing in reciprocal habitats. Across Scotland there was no significant difference between *E. micrantha* sites. *E. scottica* sites, however, did vary across their range, becoming more nutrient rich, acidic and dry the more southerly they were found.

As *Euphrasia* taxonomy is largely based upon morphological and ecological variation, these results conform to our expectations (Pugsley 1930; Yeo 1978). The fact that these species maintain their morphological differences in a common garden suggests that these differences are not a result of growth under different environmental conditions. Therefore, the fact that these species are restricted to very distinct habitats, even when growing directly adjacent to one another, suggests that there is some genetic component to their differentiation, resulting in specialisation for these specific habitats.

4.4.2. Chloroplast variation

There was a much higher level of variation in chloroplast haplotype within populations than was expected for a self-fertilising species with seeds that are dispersed by gravity. More than one chloroplast haplotype was found in eleven of the 25 populations. Of these, seven had haplotypes that differed by only one mutational step, and four (two in each of *E. micrantha* and *E. scottica*) had chloroplasts that differed by more than one mutational step. The plastid genome in hemiparasites has been shown to be under relaxed functional constraints as they are not completely self-reliant for photosynthesis (dePamphilis et al. 1997). The relaxed constraints lead to a subsequent high rate of substitution and indel formation making high levels of within species variation likely. This is not, however, thought to greatly affect within population

variation. If populations have been stable for very long periods of time, then one might expect some variation to be caused by mutation within a population, though this is unlikely. The lack of significant difference between G_{st} and N_{st} estimates suggests that this is not the main cause of variation within populations in either *E. micrantha* or *E. scottica*. It may be that there was a large pool of chloroplast variation at the initial colonisation of the site, or that separate historical migration events added new chloroplast haplotypes hence increasing variation within populations. Since these species are known to hybridise, it cannot be discounted that hybridisation and introgression may have also contributed to high levels of within population chloroplast variation.

4.4.2.1. Within species chloroplast variation

Chloroplast variation within both species is very high, both within and among regions. As chloroplast sequences exhibit a higher rate of mutation in hemiparasites (dePamphilis et al. 1997) this is not unexpected. As the species expanded across their present range it might be expected that high mutation rates could result in some of the chloroplast variation witnessed within species. However, chloroplast variation within species is not completely homogenous. The minimum spanning network shows that both species have disparate haplotype groups, suggesting mutation is not the only source of variation. Multiple allopolyploid events between the same two diploid progenitors at the origin of these species may have resulted in an initially large and diverse pool of chloroplast variation. Drift may have caused the pattern of present day chloroplast variation within and among regions of the same species. This may explain how the disparate *E. micrantha* haplotype groups are found in multiple regions. Another theory for separate haplotype groups appearing within species might be explained by hybridisation followed by introgression of a different chloroplast haplotype which may then stabilise in the species. The only occasion where a chloroplast haplotype is found in both species within a region is in Sutherland with haplotypes h03 (normally *E. micrantha* though here found in one *E. scottica* individual) and h15 (normally *E. scottica* but found in one *E. micrantha* individual). These are present in the adjacent populations SEms_a and SEsm_a. This supplies some support for the theory of hybridisation and introgression in Sutherland.

4.4.2.2. Between species chloroplast variation

Most importantly for this study, the chloroplast haplotypes have been found significantly different among species, accounting for 27% of total variation. Whilst some haplotypes are shared among species, this is not usually within the same region, with the exception of the SEms_a and SEsm_a populations. The majority of chloroplast haplotypes are species specific and occur in the same species across multiple regions. This pattern can only be explained by a long period of separation between these species, providing strong evidence to suggest that *E. micrantha* and *E. scottica* represent distinct genetic entities.

4.4.3. Microsatellite variation

4.4.3.1. Within population microsatellite variation

Typically, selfing species are expected to show low levels of genetic diversity within populations and have much greater levels of differentiation among populations than outcrossing species (Allard et al. 1968; Loveless & Hamrick 1984; Hamrick & Godt 1996). However, similar to chloroplast variation, neutral variation within populations has been found to be extremely high in both species explaining 37% and 49% of overall nuclear genetic variation in *E. micrantha* and *E. scottica* respectively. High levels of within population genetic variation have been found in other wide-spread inbreeding plant species such as *Arabidopsis thaliana* (Jorgensen & Mauricio 2004; Le Corre 2005; Stenoien et al. 2005; Bakker et al. 2006; He et al. 2007; Bomblies et al. 2010), *Bromus tectorum* (Ramakrishnan et al. 2006) and *Medicago trucatula* (Bonnin et al. 1996; Bonnin et al. 2001; Siol et al. 2008). This has largely been explained by rare outcrossing events resulting in divergent recombinant inbred lines (Bonnin et al. 2001; Ramakrishnan et al. 2006; Siol et al. 2008; Bomblies et al. 2010). Even though outcrossing rates are low in these species, it may play a major role in the organisation of genetic variation within populations. However, effective population sizes would have to be relatively large, or at least larger in the past, for this to have an effect, as outcrossing events would have to occur between genetically distinct individuals for there to be any noticeable effect of recombinant inbred lines.

Another explanation for the large amount of microsatellite variation is the high mutation rate at microsatellite loci. Microsatellites are well known to have high mutation rates, and are thus expected to reveal high levels of polymorphism when used at the species level (Thuillet et al. 2005). Recent estimates have suggested a mutation rate in plants of approximately 5×10^{-3} per locus per generation (Udupa & Baum 2001; Thuillet et al. 2002; Vigouroux et al. 2003; Cieslarova et al. 2011). Since the effective population size of a selfing species is expected to be small as individuals are expected to have very similar genotypes, a high mutation rate is likely to result in divergence among populations with low polymorphism within a population (Crow & Kimura 1970). If a population is stable over a long period of time, a high mutation rate would cause polymorphism within a local population, unless the effective population is very small.

Ultimately, high levels on variation can only be maintained if populations have large effective populations. Therefore, it is necessary to consider the ecology of *Euphrasia*. These species are not early successional annuals that occupy disturbed habitats, but occur in habitats that are thought to be stable over long periods of time. Sites identified a number of years ago have been re-visited during this study and found to be still present and much the same. This suggests that populations can be stable over long periods. Population sizes can also be large, even though they do not cover large areas. *Euphrasia* plants are very small and it is not uncommon to find as many as 150 individuals per m^2. Each population may therefore contain many thousands of individuals. Seed banks are also present in *Euphrasia* and can last from three to five years (Yeo 1961; Vitek & Kiehn 1998; Lammi et al. 1999; Liebst & Schneller 2008). Whilst this is not a very long dormancy, a persistent seed bank can increase the effective population size by retaining genetic variants in populations and hence increasing the effective generation time (Vitalis et al. 2004; Lundemo et al. 2009). It may therefore be supposed that *Euphrasia* populations may have large effective population sizes.

Since outcrossing rates have been shown to be extremely low in these two species it seems likely that a high mutation rate in microsatellites may have more of an influence on current levels of within population diversity. However, low levels of outcrossing cannot be completely

discounted and I would suggest that a combination of these two processes could be contributing to the high levels of neutral diversity witnessed within populations.

4.4.3.2. Among population, within species microsatellite variation

The results described in this chapter indicate significant differentiation among populations of the same species both within and among regions. Variation among populations of the same species within regions accounted for a much greater proportion of neutral genetic variation than among regions. This suggests that low levels of gene flow between populations within regions accounts for a greater level of differentiation than large scale geographic isolation. Such differentiation corresponds to what is usually observed among populations of selfing plant species (Schoen & Brown 1991; Green et al. 2001; Bakker et al. 2006; Ronfort et al. 2006; Bomblies et al. 2010). This phenomenon is expected as theory suggests that strong linkage disequilibria are likely to accumulate in selfing species with strong drift and lack of inter-population gene flow (Allard 1990).

In *E. micrantha* a weak signal of isolation by distance was detected, but most surprisingly there was a significant positive correlation between ecological distance and pairwise F_{st}. This perhaps suggests the presence of ecotypes within *E. micrantha* and some kind of habitat specialisation, even though there is not a large ecological distance between any populations. No such pattern of isolation by distance is detected among *E. scottica* populations though there is a very weak signal of isolation by ecological distance suggesting that there may be some kind of ecotypic differentiation within *E. scottica*.

4.4.3.3. Among species microsatellite variation

Among species, there is a weak but significant indication of differentiation. Across both species in all regions there was a weak signal of isolation by distance, but no signal for isolation by ecological distance. The principal component analysis showed that whilst there was a great deal of overlap between *E. micrantha* and *E. scottica* across all regions, they did cluster in distinctive groups within the phenetic space. STRUCTURE analysis among all species and regions showed that neutral variation tends to cluster at the population level, and not the species or region level.

This indicates that non-identical genotypes within a population were on average much more similar to one another than genotypes sampled from different populations. There was more clear evidence of differentiation among species within the context of separate regions, though patterns changed within different regions. In Sutherland and Berwickshire the two species seem to be most distinct from one another, but genetic signals in Perthshire and the Cairngorms are more blurred. These patterns are consistent across STRUCTURE and principal coordinate analyses.

Given the high mutation rate of microsatellites and the large number of alleles shared between the species (more than cpDNA, possibly also homoplasious), in hindsight, it is perhaps not surprising that a large differentiation was not found between *E. micrantha* and *E. scottica*. Indeed, in the diploid *Euphrasia*, significant differences were found among species using AFLP markers, but not microsatellite markers (French 2004). It may be that the high level of differentiation within and among populations of the same species reduces the likelihood of detecting any significant differences between the two species. Therefore, even though differentiation between species is only small, explaining 5% of total variation, it does serve to suggest that these two species do represent different entities. Unfortunately no similar studies comparing two widespread inbreeding species have been found in the literature, but there may be such studies in the pipeline involving inbreeding sister species in *Medicago* (Ronfort et al. 2006) and *Capsella* (Caullet et al. 2011). It will be interesting to compare these results with ours to see if they share the same general patterns.

4.4.4. Summary

Where the two species, *E. micrantha* and *E. scottica*, are found growing adjacent to one another there is no evidence of extensive genetic exchange between the two groups. Only in Sutherland, where chloroplasts are most similar between species, is the same chloroplast haplotype found in both species when populations are growing adjacent to one another. Adjacent populations tend to cluster in very distinct STRUCTURE groups suggesting little to no genetic exchange between them. This strongly suggests that each species is specialised for a particular habitat. This specialisation is highly likely to be a result of differential gene expression. Patterns of

chloroplast DNA variation strongly support the hypothesis of separate species, but neutral microsatellite variation shows limited but nonetheless significant differentiation between the taxa. This is probably due to the nature of microsatellites in selfing species, where high mutation rates coupled with longstanding genetic isolation will tend to blur differences between species.

From the evidence put forward, I would suggest that *E. micrantha* and *E. scottica* represent two distinct, widespread inbreeding species. They can therefore be seen as two separate pools of genetic variation that could potentially hybridise, perhaps with widespread outcrossing *Euphrasia* tetraploids to produce novel hybrids.

CHAPTER 5: Partitioning of variation within complex populations of tetraploid *Euphrasia* in Scotland. Do different morphological entities represent coherent groups within a population?

Abstract: Taxonomic uncertainty in tetraploid British *Euphrasia* has hindered conservation efforts in this critical group. A process-based species action plan approach that encourages the generation of novel diversity as opposed to trying to conserve present diversity has been suggested as most appropriate for their conservation. To implement such action plans it is essential to identify complex sites of high morphological diversity, to establish whether there is a genetic basis to this morphological variation, and to identify the processes generating and maintaining this diversity. In this chapter analyses of habitat and genetic variation were performed at three *Euphrasia* 'hotspot' sites of increasing morphological complexity. The objective was to assess whether the different morphs present in a site represented genetically distinct groups, to determine whether there was significant habitat heterogeneity within sites, and to test whether morphological/genetic groups were associated with distinct habitats. Whilst many chloroplast haplotypes were shared among different morphs within sites, some haplotypes were significantly associated with specific morphs. STRUCTURE analysis of microsatellite variation showed that large flowered, outcrossing, morphs tended to cluster together genetically within a site, suggesting an outcrossing complex. However, each small flowered, inbreeding morph on the whole represented a separate independent genetic lineage. The genotype of these lineages was consistent with a hybrid origin. Significant habitat heterogeneity was found within sites, and this increased with increasing diversity of morphs. Morphs were also shown to inhabit significantly different habitats. The results suggest that the morphological diversity at complex *Euphrasia* sites has a genetic basis, and that many of the morphs in complex sites have a hybrid origin. They also suggest that the maintenance of morphological diversity may be associated with the presence of a diversity of habitats to which the morphs are differentially adapted.

5.1. Introduction

Taxonomically complex groups pose a problem for conservation, with their high levels of endemism and a range of morphological diversity that often defies classification by any but the most experienced taxonomists. Diversity in these groups is hypothesised to be generated by a combination of evolutionary processes such as hybridisation, polyploidy, ecological

specialisation and breeding system transitions. These processes generate novel lineages that may be genetically isolated from their progenitor species and have unique adaptive potential. Strategies that focus not only on preserving individual species within these complex groups, but also the processes that generate novel diversity are considered most appropriate for their conservation (Hollingsworth 2003; Ennos et al. 2005; Ennos et al. 2012). For the successful implementation of a process-based conservation strategy an essential first step is to investigate the underlying processes that are generating taxonomic biodiversity (Ennos et al. 2005; Ennos et al. 2012).

Tetraploid *Euphrasia* in the UK are a taxonomically complex group in which conservation has been hindered due to difficulties in identifying critical taxa (Ennos et al. 2012). Approximately 12% of the vascular plant species on the priority 'Short List' for conservation in the UK Biodiversity Action Plan (HMSO., 1995), are tetraploid members of *Euphrasia*. In order to more effectively conserve the diversity within *Euphrasia* it is crucial to study the mechanisms that are involved in generating this complexity so that conservation resources may be pooled and a process-based conservation plan may be implemented.

Diversity in tetraploid *Euphrasia* can be broadly split into three categories. First, there are well defined species that are widespread throughout Britain, and are also present in continental Europe. These species include the widespread inbreeding species such as *E. micrantha, E. scottica* and *E. frigida,* and widespread outcrossing species such as *E. nemorosa* and *E. arctica*. Second, there are a number of endemic tetraploid taxa. It is likely that these endemics have evolved recently, within the last 10,000 years, since the end of the last glaciation. A feature common to all five endemic tetraploid taxa generated in this period is their very strict habitat requirements. The two relatively large flowered endemics *E. marshallii* (Pugsley 1933, 1936) and *E. rotundifolia* (Pugsley 1929) are restricted to damp basic cliff top turf and are found growing along the north Scottish coast (Silverside 1991b). Also found on the northern and western Scottish coast are the small flowered *E. campbelliae* (Pugsley 1940) and *E. heslop-harisonii* (Pugsley 1945) which are restricted to coastal heathy turf and grazed estuarine saltmarshes respectively (Silverside 1991b, c). Finally, *Euphrasia cambrica* (Pugsley 1929,

1936) is a small flowered *Euphrasia* restricted to poor montane grazed truf and can be found growing only in north Wales (Silverside 1991c). Whatever the origin of these taxa, their novel habitat requirements are likely to have facilitated their ability to diversify. Hybridisation may also have played a key role, as four of these endemics (*E. marshallii, E. rotundifolia, E. campbelliae* and *E. heslop-harisonii*) are of putatively hybrid origin (Yeo, 1978). Finally, there are sites in which "novel" variation occurs, much of which is unclassifiable into easily recognisable taxa.

These sites are found to occur within "hotspot" areas of British *Euphrasia* diversification. There are areas where a great deal of morphological variation in *Euphrasia* exists. These "hotspots" have been identified in the western and northern oceanic regions of the UK and together contain six species with a high conservation status (Wigginton 1999; Preston et al. 2002; French 2004). Inland and coastal heathland habitats in south-west England, montane habitats of north-west England and north Wales, and coastal habitats of western and northern Scotland have been identified as sites with high levels of *Euphrasia* diversity (French 2004). "Hotspot" areas are characterised by heterogeneous habitats and large numbers of different *Euphrasia* species, including endemics. *Euphrasia* species within these areas are not always found in isolated populations but can be found co-occurring in a complex manner.

These complex populations contain a mix of morphological variation ranging from easily classifiable widespread species, to unclassifiable taxa. It is possible that novel *Euphrasia* are being generated within these complex populations, and in order to develop a process-based conservation strategy it will be important to investigate whether this range of morphological diversity is associated with partitioning of genetic diversity, and to try to explain how this diversity has arisen. It is possible that there is sufficient environmental heterogeneity at these sites that different widespread species can coexist. It is hypothesised that hybridisation between these widespread species could produce the novel morphological diversity witnessed (French, 2004). In these sites, the environmental heterogeneity is thought to be such that hybrid offspring may have the potential to survive. It is thought that the inheritance of a predominantly inbreeding mating system may lead to the segregation of novel inbreeding lines that are genetically isolated,

and morphologically different from their progenitor taxa (French, 2004). These inbreeding lines may have unique adaptive potential and therefore natural complex sites, with a mixture of 'pure' species and putative hybrid morphs, may merit conservation status within the context of a process-based conservation strategy.

No in depth analysis has been performed within these complex populations of *Euphrasia*. Therefore, the objective of this chapter is to explore the status of different morphs within complex populations of mixed *Euphrasia* species in order to gather evidence to support our hypothesis of hybridisation resulting in novel *Euphrasia* lineages. In order to determine if this idea is sound we shall conduct a detailed investigation of a number of these complex sites. Using a combined analysis of variation in chloroplast DNA, microsatellite markers and habitat differences we shall ask whether:

1. The morphological variation that we see within sites represents different genetic lineages and hybrids between these lineages.
2. It is possible that novel morphs are the result of hybridisation.
3. Sites are environmentally heterogeneous.
4. There is an association between genetic lineages and environmental variation within the sites.

If the above can be shown to be true, these results would be consistent with the hypothesis that hybridisation takes place, and that environmental heterogeneity allows the coexistence of multiple genetic entities.

Plant samples were collected from three different complex populations in the "hotspot" coastal area of north-west Scotland. Morphological identifications were carried out by *Euphrasia* specialist, Alan Silverside. Chloroplast sequence data and microsatellite markers were used to investigate genetic diversity within these populations and to assess evidence of genetic partitioning between morphs. Associated species lists were used to characterise environmental heterogeneity and to assess whether different morphs were associated with different environments. If data from these sites is consistent with our assumptions of how diversity is

generated within these sites, then it will be worth pursuing the hypothesis with further observations of offspring, and experimental analysis.

5.2. Materials and methods

5.2.1. Study sites

Complex sites are sites in which a number of distinct *Euphrasia* morphs are found growing amongst one another. These morphs can be representative of "pure" species identified in the UK, or are morphs that combine characters of more than one *Euphrasia* species. These may either be recent hybrids, or stable inbreeding lines of hypothesised hybrid origin. Complex *Euphrasia* populations can be found along the north and west coasts of Scotland, and three study sites that represent increasing levels of complexity were identified in Sutherland (Figure 5:1). At these complex sites all morphological identifications were performed by the *Euphrasia* specialist, Alan Silverside.

Figure 5:1. Sites selected for complex population analysis: Loch an Diamh Mor (A), Faraid Head (B) and Betty Hill (C). A digital version of this map can be found here.

5.2.1.1. Loch an Daimh Mor

The first site, Loch an Daimh Mor (Figure 5:1, A) represents a situation where complexity is relatively low. This site represents an area of wet heathland, similar to the habitats in which *E. scottica* is usually found and shows a low level of patchy environmental heterogeneity. It should also be noted that at just over a kilometre from the coast, this site was furthest from the sea. In this site two morphs, *E. scottica* and *E. fharaidensis*, and their putative hybrid, *E. scottica x fharaidensis*, were identified. *E. fharaidensis* is a taxon that is yet to be formally described, and is not present in Yeo's (1978) monograph of *Euphrasia* in Europe. It is a small flowered endemic taxon that is restricted to moderately base rich coastal flushes in north Scotland (Silverside 1991c, 1998). A description of *E. scottica* can be found in Chapter 4. No large flowered tetraploids were present at this site.

5.2.1.2. Faraid Head

The second site, Faraid Head (Figure 5:1, B), was chosen as here, different habitat types merged with one another, and it represents a site of potentially increased environmental heterogeneity compared to the Loch an Daimh Mor site. A number of *Euphrasia* morphs were identified in this site, including a mix of outcrossing and inbreeding taxa. Three putatively pure species were identified, the medium flowered *E. nemorosa* and the small flowered *E. fharaidensis* and *E. scottica* along with putative hybrids *E. nemorosa x fharaidensis*, *E. scottica x* and *E. scottica x fharaidensis.* *E. scottica x* represents a morph that is similar in many respects to *E. scottica* but with some indication of introgression from another, unidentified *Euphrasia* species. *E. nemorosa* (Pugsley 1929) has a wide distribution across Europe and in Britain is found in grassland and lowland heathland, but in the north and west is restricted to coastal calcareous soils (Silverside 1991a).

5.2.1.3. Betty Hill

Betty Hill (Figure 5:1, C) represents the site with the highest degree of complexity with many different tetraploid *Euphrasia* found across the site, again including a mix of outcrossing and inbreeding taxa. As with Faraid Head, three putatively pure species were identified, the large

flowered *E. marshallii* and the small flowered *E. foulaensis* and *E. fharaidensis*. However, there were many more hybrid combinations found, including hybrids that did not have a progenitor species present: *E. marshallii x, E. marshallii x arctica, E. marshallii x foulaensis, E. marshallii x fharaidensis, E. electa, E. electa x* and *E. electa x fharaidensis*. As previously described, *E. marshallii* (Pugsley 1933, 1936) is a large flowered *Euphrasia* that is restricted to damp basic cliff top turf in north Scotland. *E. electa* is a putatively stable *E. micrantha x scottica* hybrid that is found growing on damp moorland. *E. electa x* and *E. marshallii x* are morphs that show many morphological characteristics of *E. electa* and *E. marshallii* respectively, but are not quite the same, indicating some form of introgression from an unidentified *Euphrasia* species.

Each species and hybrid identification serves to provide an indication of the different morphological groups in each site that will be used for analysing these populations. It was thought that these identifications, performed by an experienced taxonomist, were able to capture a greater amount of morphological information than a few measurements taken from field samples. Thus these samples were able to best represent the variation present in each site.

5.2.2. Population sampling

Loch an Daim Mor is a much smaller site than either Faraid Head or Betty Hill. Samples from Loch an Dim Mor were therefore collected slightly differently.

Table 5:1. Complex *Euphrasia* sites within Scotland

Site Name	Site Code	No. Transects	UK Grid Reference	Date visited	No. Individuals Sampled	No. morphs
Loch an Daimh Mor	LDM	1	NC16584298	10.08.2010	24	3
Faraid Head	FH	2	NC16554287	08.08.10	100	6
Betty Hill	BH	3	NC70436276	02.08.2006 and 03.08.2006	137	10

5.2.2.1. Loch an Daimh Mor

In Loch an Daim Mor a two dimensional grid of 30 by 28 meters was formed for collection from the complex population of *E. scottica, E. foulaensis* and their putative hybrid. Quadrats of $1m^2$

were placed at randomly selected coordinates. These quadrats were separated into 25 (20cm x 20cm) partitions. Within each quadrat one plant from four randomly selected partitions was collected. In total, six quadrats were sampled and 24 individuals collected (Table 5:1).

5.2.2.2. Faraid Head

In Faraid Head, two 50m transects that encompassed a variety of morphological diversity were selected for collections (Figure 5:2). These two transects were approximately 400m apart from one another. Quadrats of $1m^2$, separated into 25 (20cm x 20cm) partitions, were placed at 5m intervals along the transect. One plant from five randomly selected partitions was collected from each quadrat. In total, 50 plants were collected from each transect, resulting in the collection of 100 individuals in total at the Faraid Head site (Table 5:1).

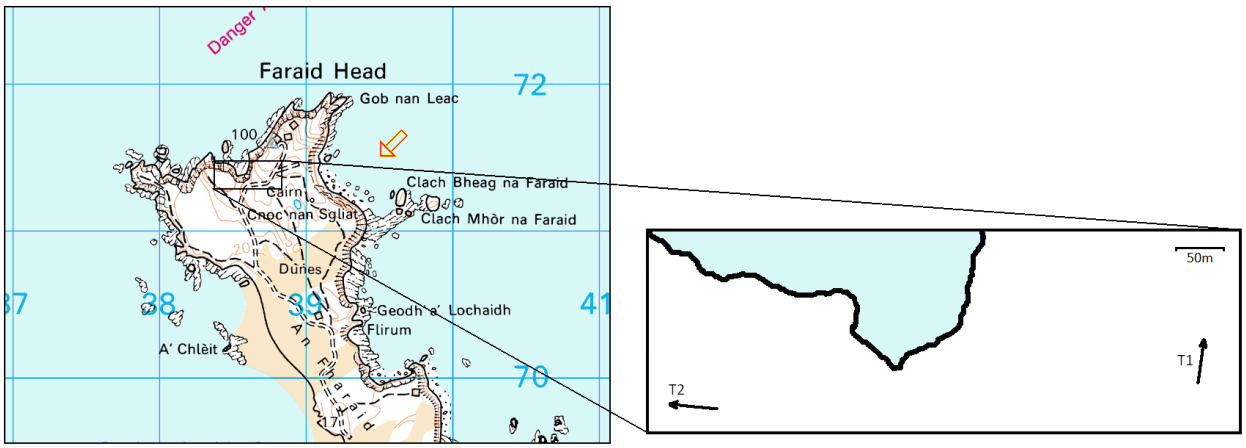


Figure 5:2. Diagram of transect placement at the Faraid Head site. Map on the left taken from www.streetmap.co.uk; each grid square represents 1 km^2. T1 = Transect 1; T2 = Transect 2.

5.2.2.3. Betty Hill

In Betty Hill three 50m transects encompassing a variety of morphological diversity were selected for collections. These transects were taken at right angles from the shore, with the three parallel transects separated by approximately 100m (Figure 5:3). Samples along the quadrat were taken in the same way as at Faraid Head, with up to 5 individuals collected from each quadrat. From each transect 47, 44 and 46 individuals were collected respectively with a grand total of 137 individuals collected across the whole Betty Hill site (Table 5:1).

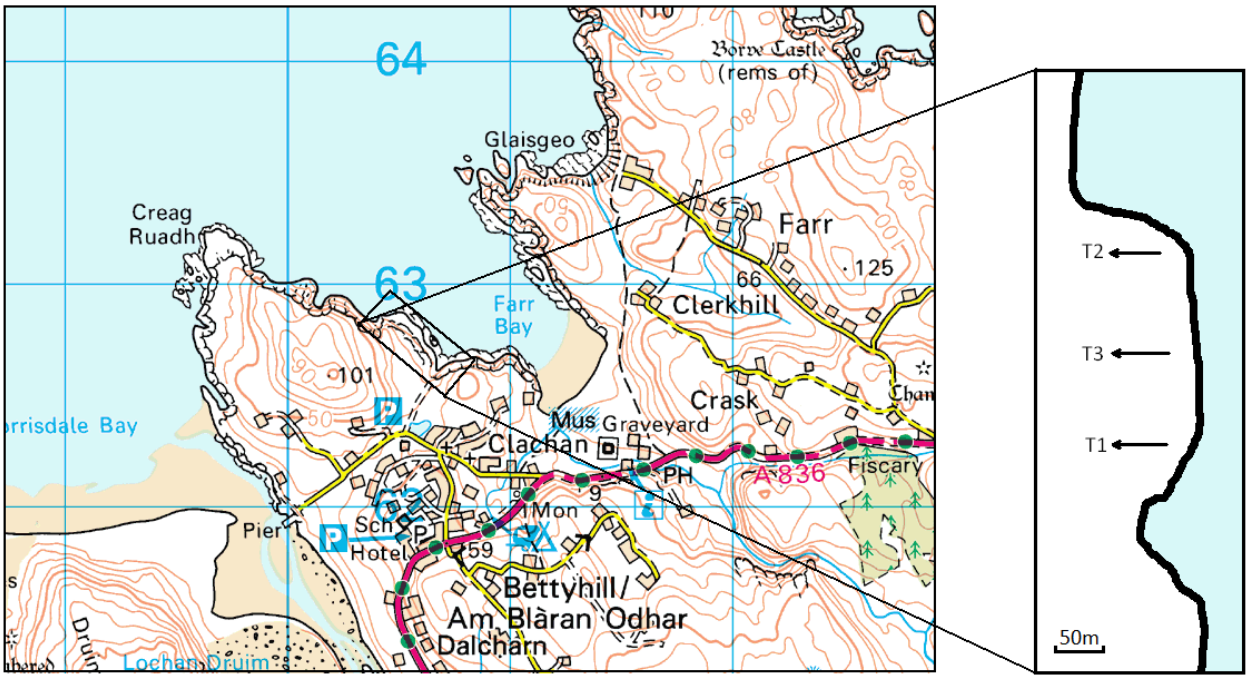


Figure 5:3. Diagram of transect placement at the Betty Hill site. Map on the left taken from www.streetmap.co.uk; each grid square represents 1 km^2. T1 = Transect 1; T2 = Transect 2; T3 = Transect 3.

5.2.3. DNA Extraction and molecular marker analysis

DNA extraction was performed as described in Chapter 3. Molecular diversity was assessed in both the chloroplast and nuclear genome. In Loch an Daim Mor, all individuals were sequenced at the trnL$^{(UAG)}$-rp132 locus (Shaw et al. 2005). In Faraid Head and Betty Hill one individual of each morph identified was sequenced per quadrat. Genotypes were scored in every individual collected from the amplification of ten microsatellite loci (*Earc1*, *Earc2*, *Earc3*, *Earc4*, *Earc5*, *Earc6*, *Earc7*, *Earc8*, *Ene1* and *Ene5*). Both techniques were performed and scored as described in Chapter 4.

5.2.4. Molecular data analysis

5.2.4.1. Chloroplast analysis

Sequences were aligned and polymorphisms detected using the software Sequencher (Gene Codes Inc., Ann Arbor, MI, USA). Sequence editing and haplotype formation were performed manually. Estimation of the partitioning of genetic variation within each site was performed using analyses of molecular variance (AMOVA) to test the null hypothesis of no difference in

chloroplast sequence between morphs using the computer software program ARLEQUIN v3.1 (Excoffier et al. 2005). Depending on the site, in order to make sense of the diversity, different groupings of morphs were tested, for example: pure species only, large flowered morphs vs small flowered morphs, hybrid groups and individual morphs. Minimum spanning networks indicating the relatedness between haplotypes were created using ARLEQUIN v3.1. A randomisation test for association was also performed using the program Resampling Stats for Excel, in order to test the null hypothesis that individual chloroplast haplotypes were not associated to specific morphs or hybrid groups in each site.

5.2.4.2. Microsatellite analysis

As in Chapter 4, genotype scores for individuals were converted into an RFLP-type haplotype. The Bayesian population assignment software STRUCTURE Version 2.3 (Falush et al. 2007) was used to look for a pattern of genetic structuring between and within transects and morphological groups within each site separately. Estimates of K (the number of structural units) represent values that most accurately capture the major structure at each site. As in Chapter 4, this meant that the largest value of K, beyond which further increases resulted in the random subdivision of previously well defined structural units (clusters). Between two and four independent runs with a burnin period of 40000 and with 40000 MCMC repeats after burnin were performed for each value of K in order to confirm that the STRUCTURE output was consistent.

5.2.5. Habitat Analysis

Associated plant species were recorded from within the quadrats in the Loch an Daimh Mor grid and the Faraid Head and Betty Hill transects. Each associated species acted as an independent measure of five Ellenburg environmental indicator values: light (L), moisture (F), reaction (R), nitrogen (N) and salt tolerance (S). Associated species lists for each quadrat and transect were compiled, resulting in tables of environmental indicator values which were taken from PLANTATT, as in Chapter 4. These values were used to perform an ANOVA to first test the null hypotheses of (i) no habitat difference between transects within sites and (ii) no habitat different between quadrats along a transect or within the grid. Associated species for the

different morphs were also complied and consisted of those species found within the same quadrat. An ANOVA was then performed to test the null hypothesis that there was no association between morph and environment within a site.

5.3. Results

At each of the sites I performed the same set of analyses. First, the diversity and distribution of individual morphs within each site will be described. Then the extent to which the variation at cpDNA and microsatellites can be accounted for by the difference in morphs will be examined. A simple analysis will then be conducted to see whether there is environmental variation among the quadrats sampled and we shall look into whether there is an association between morphs and environment.

5.3.1. Loch an Daim Mor

5.3.1.1 Diversity and distribution of morphs

An overview of collections within the grid at Loch an Daimh Mor can be seen in Figure 5:4. Three morphs were identified in this site, *E. fharaidensis, E. scottica* and their putative hybrid, *E. fharaidensis x scottica*. Of the three morphs identified within the Loch and Daim Mor grid, *E. fharaidensis* is the most common, being identified in four of the six quadrats. Figure 5:4 also shows that individual morphs tend to cluster together within quadrats. However, separate morphs can be found growing in close proximity, as is shown between quadrats Q4 and Q2, which contain *E. fharaidensis* and *E. scottica* respectively.

5.3.1.2. Genetic variation among morphs

From the overview of Loch an Daimh Mor (Figure 5:4) we can see that whilst morphs may cluster together within a quadrat, they do not necessarily come from the same family, as different chloroplast groups are found within the same morph within one quadrat. The putative hybrid

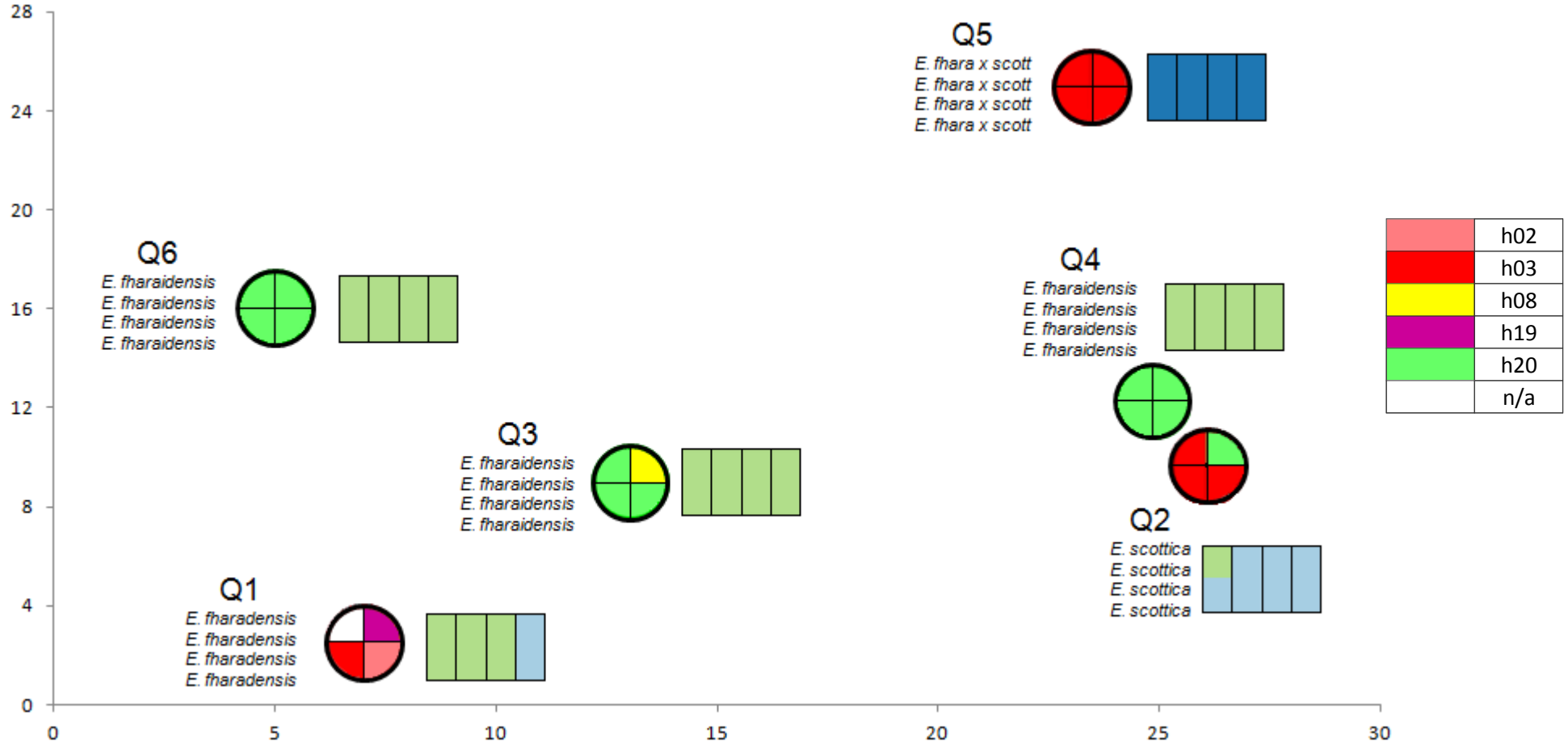


Figure 5:4. Overview of sampling at the Loch an Daimh Mur site in the grid. Numbers along the axis represent meters and indicate the boundary of the grid. The approximate position of each quadrat (Q1 – Q6) is indicated by the placement of the circle quadrant. Circle quadrants indicate chloroplast haplotypes identified for each individual (individuals going down the list from each quadrat are represented clockwise from the top right hand quarter). Chloroplast haplotypes are represented by the colours shown in the key to the right of the grid. The rectangular graphs indicate STRUCTURE groups for each individual (individuals moving down the list are indicated from left to right along the graph). Morphs identified are *E. fharaidensis*, *E. scottica* and *E. fharaidensis x scottica* (*E. fhara x scott*).

E. fharaidensis x scottica does share a chloroplast haplotype with *E. scottica* which is an observation consistent with their proposed hybrid origin. It also looks as though chloroplast haplotypes and STRUCTURE groups may be associated with different morphs. The analysis of chloroplast and microsatellite data shall now be approached in further detail.

Sequencing of the trnL$^{(UAG)}$-rp132 chloroplast locus resulted in the detection of 11 sequence differences across the site which included point mutations (polymorphic sites 1, 89, 119, 163, 235, 302, 433, 589 and 628), an indel (polymorphic site 393) and an inversion (polymorphic site 516) (Table 5:2). In total, 5 haplotypes were identified across the site (h02, h03, h08, h19 and h20) (Figure 5:5).

Table 5:2. Polymorphic sites detected in the trnL$^{(UAG)}$-rp132 chloroplast locus sequence at the Loch an Diamh Mor site.

Haplotype	Polymorphic Sites (relative position in sequence)										
	1	89	119	163	235	302	393	433	516	589	628
h02	C	G	G	A	T	G	G	T	AGTAA	G	T
h03	T	G	G	C	C	T	G	G	TTACT	A	T
h08	C	G	G	A	T	G	G	T	TTACT	G	T
h19	C	A	A	A	C	G	:	T	AGTAA	G	T
h20	C	G	G	A	T	G	G	T	TTACT	G	G

Separate AMOVAs were performed to compare the difference between morphs of putatively pure origin (*E. fharaidensis* and *E. scottica*) and between all morphs detected in the grid. Differences between all morphs accounted for 58% of total chloroplast variation within the grid, although there is no significant difference detected when *E. fharaidensis x scottica* is not included in the analysis (Table 5:3). A randomisation test performed using the program Resampling Stats for Excel showed that haplotype h20 is significantly associated with *E. fharaidensis* ($p < 0.01$) and h03 is significantly associated with *E. fharaidensis x scottica* ($p < 0.01$). Since there are only 23 individual chloroplast samples in the grid of Loch an Daim Mur, p-values of < 0.1 are considered significant.

Table 5:3. Analysis of molecular variance (AMOVA) for the trnL$^{(UAG)}$-rp132 chloroplast locus for *Euphrasia* across the Loch an Daimh Mor grid. d.f. = degrees of freedom; SSq – sums of squares (*p < 0.05; $^{+}$p < 0.1; ^{NS}Not Significant).

	d.f.	SSq	Variance components	% total Variation
Group = pure species only^{++}				
Among Group	1	9.410	1.13759	47.88^{NS}
Among quadrats within species	3	6.450	0.33355	14.04*
Within quadrats	14	12.667	0.90476	38.08
Total	18	28.526	2.37589	100
Group = all morphs^{+++}				
Among Group	2	22.883	1.55663	58.13^{+}
Among quadrats within morph	3	6.450	0.37631	14.05*
Within quadrat	17	12.667	0.74510	27.82
Total	22	42.000	2.67804	100

$^{++}$*E. scottica* and *E. fharaidensis;* $^{+++}$*E. scottica, E. fharaidensis, E. scottica x fharaidensis*

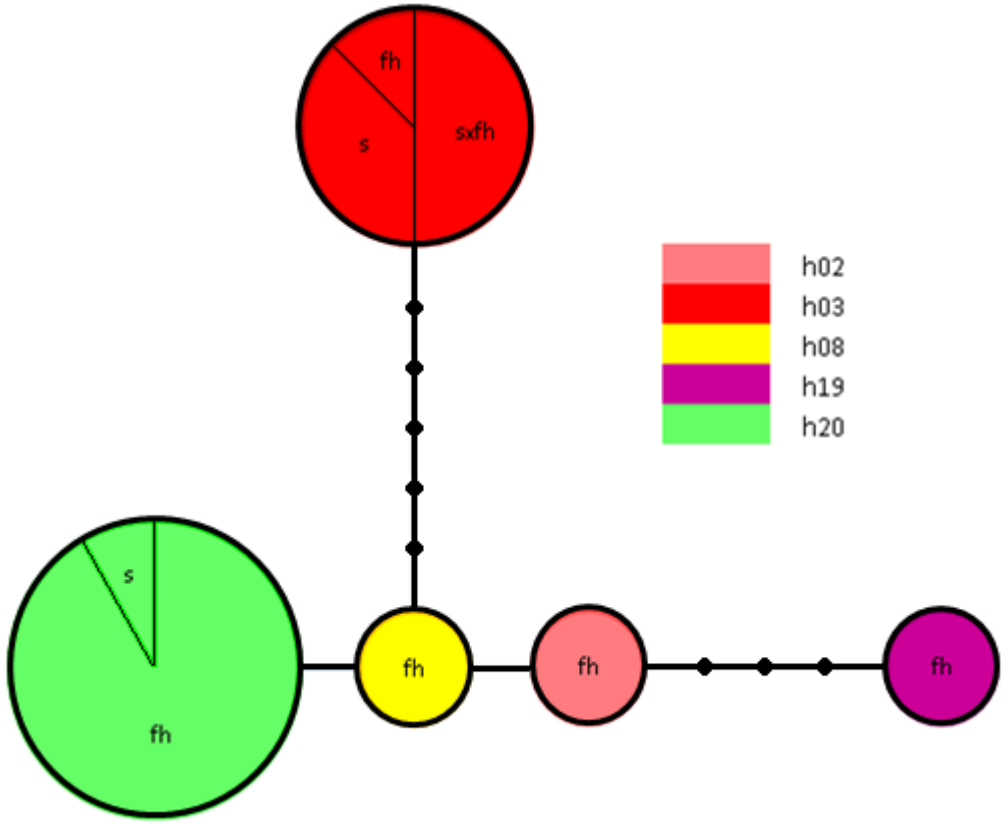


Figure 5:5. Minimum spanning network (MSN) representing chloroplast haplotype groups from the Loch an Daimh Mor grid. Haplotypes are represented by colours as shown in the key. Dots along the line indicate missing haplotypes. The area of the circle represents the number of individuals with each haplotype. Species identifications are represented as follows: *E. fharaidensis* (fh); *E. scottica* (s); *E. scottica x fharaidensis* (sxfh).

STRUCTURE analysis of the microsatellite data (Falush et al. 2007) in Loch an Daim Mor shows that within the grid different morphological entities are clearly distinguished as genetically isolated groups (Figure 5:6). Individuals identified as *E. fharaidensis x scottica* do not appear to be F1 hybrids of *E. fharaidensis* and *E. scottica*, but an entity separate from both *E. scottica* and *E. fharaidensis*. This is likely to be the result of historical hybridisation producing an individual of intermediate morphology, propagated by self-fertilisation and thus producing a novel inbred line. If K=2, then *E. fharaidensis x scottica* clusters with *E. fharaidensis* suggesting it is more closely related to *E. fharaidensis*. There is some evidence for more modern introgression of genes between *E. scottica* and *E. fharaidensis* in one *E. scottica* individual (Figure 5:6). One *E. fharaidensis* individual comes out under STRUCTURE analysis as most similar to other *E. scottica* (Figure 5:6). An AMOVA of microsatellite haplotypes shows only weak evidence for partitioning of genetic variation among morphs, though it accounts for a large proportion (63%) of the total genetic variation within the grid (Table 5:4).

Table 5:4. Analysis of molecular variance (AMOVA) for microsatellite multilocus phenotypes from the Loch an Daimh Mor grid. d.f. = degrees of freedom; SSq = sums of squares (**p < 0.001; $^{+}$p < 0.1).

	d.f.	SSq	Variance components	% total Variation
Group = morph				
Among group	2	45.333	3.17361	63.12^{+}
Among quadrats within groups	3	10.875	0.59028	11.74**
Within groups	18	22.750	1.26389	25.14
Total	23	78.958	5.02778	100

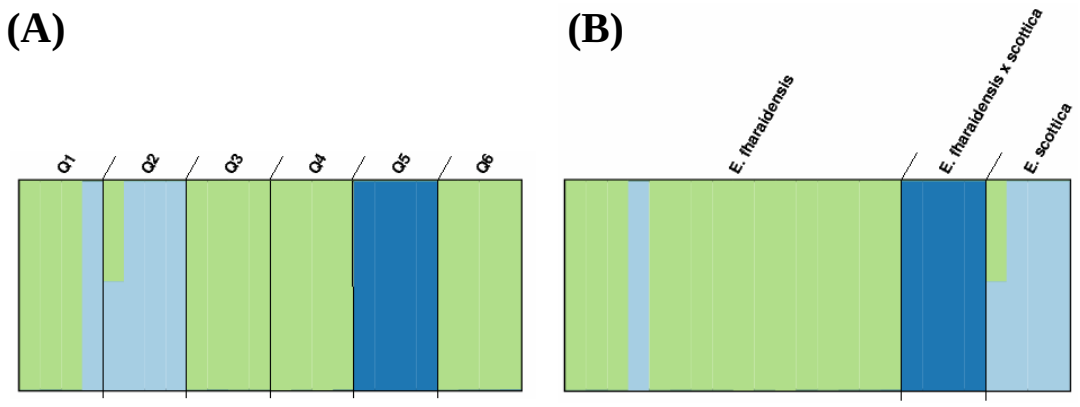


Figure 5:6. STRUCTURE analysis of *Euphrasia* morphs detected in the Loch an Daimh Mor grid. K (the most likely number of structural groups) is 3. (A) = results ordered according to quadrat. (B) = results ordered according to morph.

5.3.1.3. Environmental heterogeneity within sites

Associated species were recorded within each quadrat within the grid. Lists of their environmental indicator values were then used to test the null hypothesis of no habitat difference between quadrats within the grid. This analysis shows that quadrats differed significantly in levels of moisture (Light: F-value of 1.283, $p > 0.1$; Moisture: F-value of 2.4086, $p < 0.05$; Reaction: F-value of 1.3185, $p > 0.1$; Nitrogen: F-value of 1.0663, $p > 0.1$; and Salt: F-value of 0.8721, $p > 0.1$) (Table 5:5).

Table 5:5. Analysis of variance (ANOVA) of mean Ellenberg values among quadrats for five environmental variables at the Loch an Daim Mur site. d.f. = degrees of freedom; MS = mean square (*$p < 0.05$).

Source of Variation	d.f.	MS Light	MS Moisture	MS Reaction	MS Nitrogen	MS Salt
Grid						
Among quadrats	5	0.3973	4.136*	2.533	0.496	0.0506
Residuals	50	24.7695	1.717	1.921	0.465	0.058

5.3.1.4. Association between genotypes and environment

Lists of species associated to each morph were collated from the associated species recorded within the quadrats in which that morph was found. Environmental indicator values were then gathered to test the null hypothesis of no habitat difference between morphs within the Loch an Daimh Mor site. There was less evidence for difference in habitat preference between different morphs than environmental heterogeneity within the site as a whole, with barely significant differences detected between levels of moisture (Light: F-value of 2.1492, $p > 0.1$; Moisture: F-value of 2.5687, $p = 0.085$; Reaction: F-value of 1.6935, $p > 0.1$; Nitrogen: F-value of 1.4993, $p > 0.1$; and Salt: F-value of 1.1685, $p > 0.1$) (Table 5:6 and Figure 5:7). The box plots show that *E. fharaidensis x scottica* morphs inhabit environments that are less moist than those inhabited by *E. fharaidensis* and *E. scottica*. Since each morph has been shown to have a different genotype there is weak evidence to suggest that genotypes are associated with different environments.

Table 5:6. Analysis of variance (ANOVA) of mean Ellenberg values among morphs for five environmental variables at the Loch an Daim Mur site. d.f. = degrees of freedom; MS = mean square ($^{+}p < 0.1$).

Source of Variation	d.f.	MS Light	MS Moisture	MS Reaction	MS Nitrogen	MS Salt
Morph^{++}						
Among Morphs	2	0.6890	4.771^{+}	3.692	0.8179	0.0955
Residuals	57	0.3206	1.857	2.108	0.5456	0.0817

$^{++}$*E. fharaidensis, E. scottica, E. fharaidensis x scottica*

5.3.2. Faraid Head

5.3.2.1. Diversity and distribution of morphs

Figure 5:8 shows an overview of the collections made within each quadrat and each transect within the Faraid Head site. Five morphs were identified in this site, *E. nemorosa, E. fharaidensis* and *E. scottica* and the putative hybrids, *E. nemorosa x fharaidensis* and *E. fharaidensis x scottica.* Within this site, *E. nemorosa* and *E. fharaidensis* were the most commonly identified morphs. *E. nemorosa x fharaidensis* and *E. fharaidensis x scottica* were also quite prevalent within the site and *E. scottica* had a very low level presence, with only three individuals identified. In 9 of the 20 quadrats sampled, only one morph was found to be present. Consequently 11 of the 20 quadrats contained a mix of different *Euphrasia* morphs. This is most obvious in Transect 2 where the different *Euphrasia* morphs are commonly found growing amongst one another, particularly between quadrats Q2 and Q7. In Transect 1, morphs can generally be seen to gradate from *E. nemorosa* to *E. fharaidensis x scottica* to *E. fharaidensis* between quadrats Q1 and Q10. In Transect 2, there is also a gradual change in the morphs detected with a change from *E. nemorosa* to *E. nemorosa x fharaidensis* to *E. fharaidensis* between quadrats Q1 and Q10.

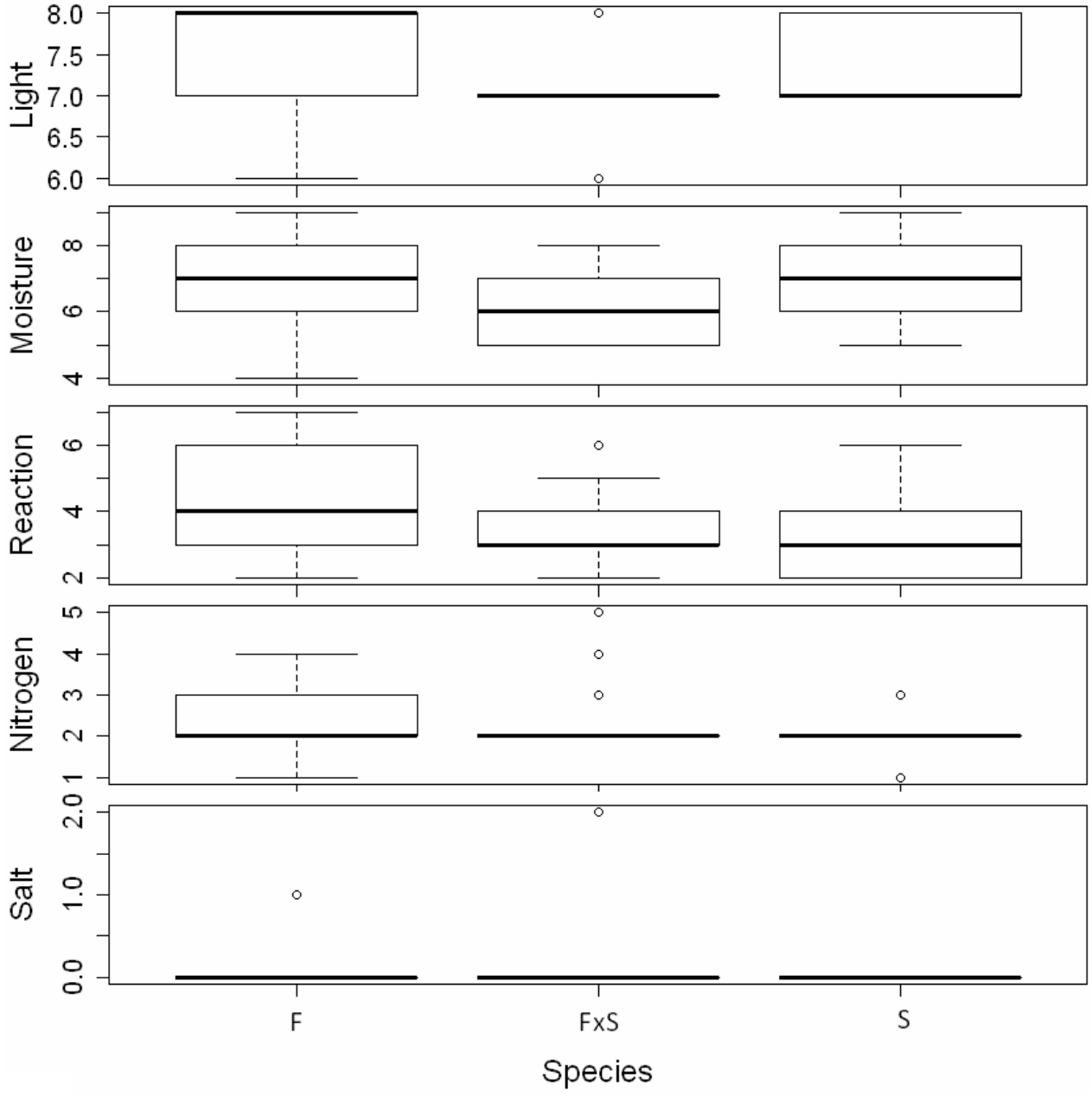


Figure 5:7. Ellenberg values for five environmental variables, calculated from associated species within quadrats. Ellenberg values are shown for each morph present in the Loch an Daim Mor site: *E. fharaidensis* (F), *E. fharaidensis x scottica* (FxS) and *E. scottica* (S).

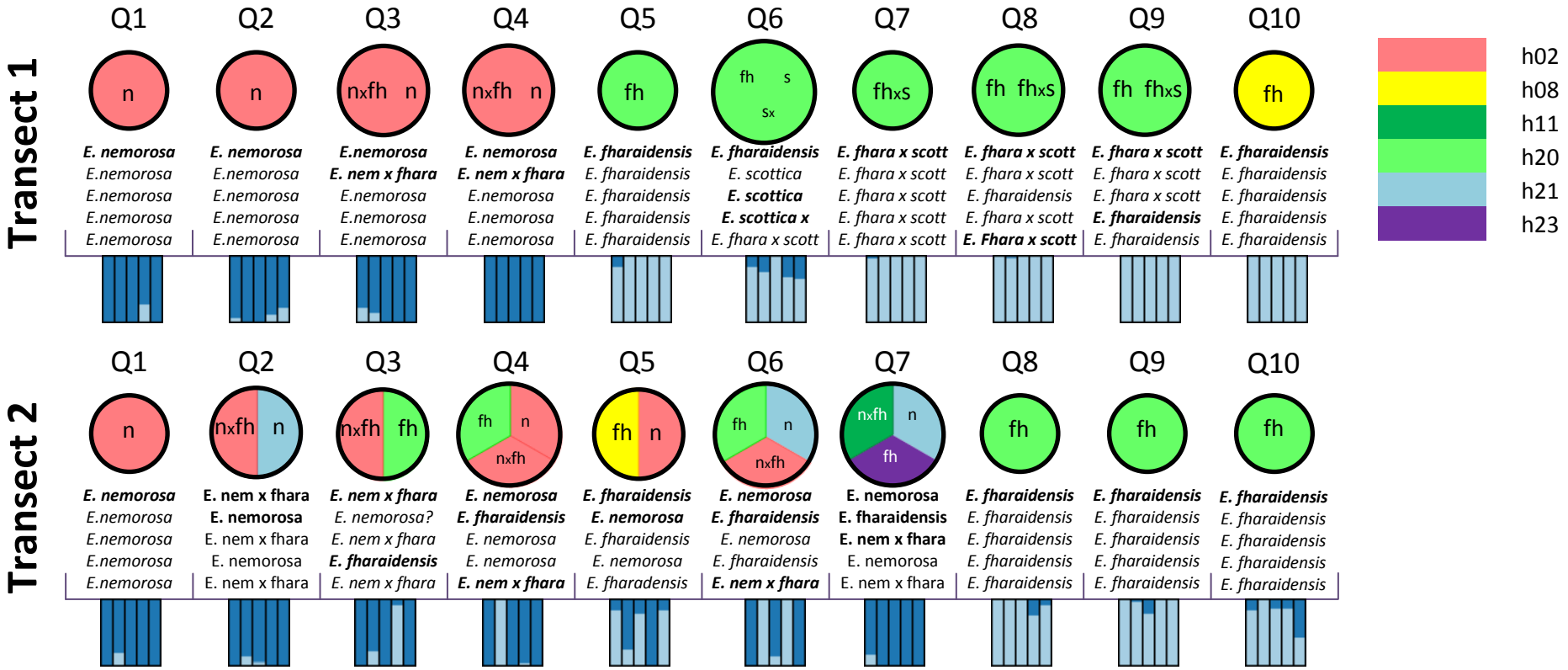


Figure 5:8. Overview of sampling at the Faraid Head site. Quadrats (Q1 – Q10) along each 50m transect were placed at 5m intervals. Individual morphs collected are listed for each quadrat. From each quadrat, one individual of each species was sequenced at the trnL$^{(UAG)}$-rp132 chloroplast locus (indicated in bold from the list). Chloroplast haplotypes are represented in the circles by the colours shown in the key in the top right. The letters within the circles indicate the morph in which the haplotype was detected (n = *E. nemorosa*; nxfh = *E. nemorosa x fharaidensis*; fh = *E. fharaidensis*; fhxs = *E. fharaidensis x scottica*; s = *E. scottica* and sx = *E. scottica x*). The rectangular graphs below the list of sampled morphs indicates the STRUCTURE group of each individual (individuals moving down the list are indicated from left to right along the graph).

5.3.2.2. Genetic variation among morphs

Sequencing of the trnL$^{(UAG)}$-rp132 chloroplast locus resulted in the detection of 9 sequence differences across the site including point mutations (polymorphic sites 44, 165, 481, 513 and 628), indels (polymorphic site 239, 510, 515 and 516) and an inversion (polymorphic site 516) (Table 5:7). In total 6 haplotypes were identified across the Faraid Head site (Figure 5:9).

The overview of the Faraid Head site (Figure 5:8) shows that the gradation of morphological forms along transects are replicated in the change of chloroplast haplotypes and STRUCTURE groups. All but one of the proposed hybrid individuals (*E. nemorosa x fharaidensis* from Transect 2, Q7) shares its chloroplast haplotype with one of their proposed progenitor species, consistent with their hypothesised hybrid origins. It looks as though chloroplast and STRUCTURE groups are broadly related to morphological groups, and this warrants further analysis.

Table 5:7. Polymorphic sites detected in the trnL$^{(UAG)}$-rp132 chloroplast locus sequence at the Faraid Head site.

Haplotype	Polymorphic Sites								
	44	165	239	481	510	513	515	516	628
h02	A	T	G	C	GT	G	G	AGTAA	T
h08	A	T	G	C	GT	G	G	TTACT	T
h11	A	C	:	A	GT	G	G	AGTAA	T
h20	A	T	G	C	GT	G	G	TTACT	G
h21	C	C	G	C	:	C	:	:	T
h23	A	C	:	A	:	G	G	TTACT	T

Three different levels of morphological variation were tested for variation in chloroplast haplotypes: (i) between morphs of putatively pure species (*E. nemorsa*, *E. fharaidensis* and *E. scottica*); (ii) between *E. nemorosa* plus its hybrid morph and all other morphs ([*E. nemorosa* and *E. nemorosa x fharaidensis*] and [*E. fharaidensis*, *E. fharaidensis x scottica* and *E. scottica*]), and (iii) between all separate morphs. Significant partitioning of variation between morphs was found at each level of analysis (Table 5:8), though the largest proportion of chloroplast sequence variation (57%) was explained by differences between *E. nemorosa* and its hybrid morph, and all

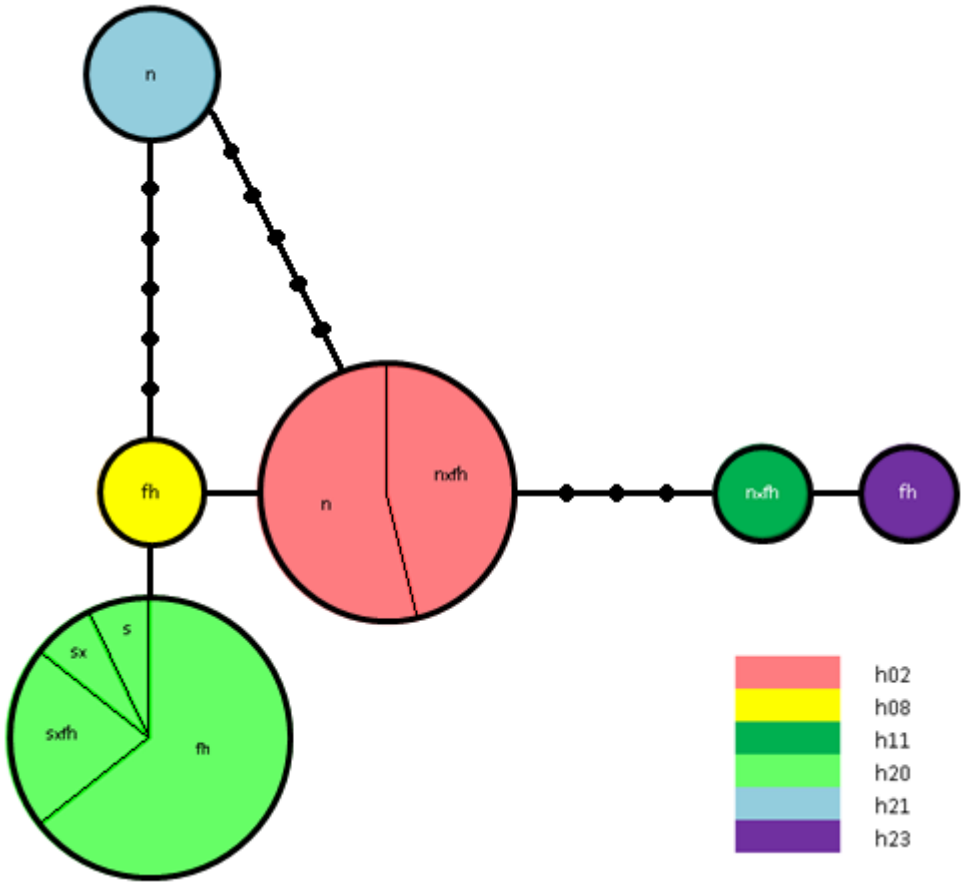


Figure 5:9. Minimum spanning network (MSN) representing chloroplast haplotye groups from Faraid Head. Haplotypes are represented by colours as shown in the key. Dots along the line indicate missing haplotypes. The area of the circle represents the number of individuals with each haplotypes. Species identifications are represented as follows: *E. nemorosa* (n); *E. nemorosa x fharaidensis* (nxf); *E. fharaidensis* (fh); *E. scottica* (s); *E. scottica x* (sx); *E. scottica x fharaidensis* (sxfh).

Table 5:8. Analysis of molecular variance (AMOVA) for the trnL$^{(UAG)}$-rp132 chloroplast locus for *Euphrasia* across the Faraid Head site. d.f. = degrees of freedom; SSq = sums of squars (**p < 0.01; *p < 0.05).

	d.f.	SSq	Variance components	% total Variation
Group = pure species morphs only				
Among Group	2	12.285	0.57222	39.67*
Among transects within species	2	3.937	0.24526	17.00*
Within populations	21	13.125	0.625	43.33
Total	25	29.346	1.44248	100
Group = *E. nemorosa* and its hybrid morphs vs everything else				
Among Group	1	16.759	0.85567	56.87**
Among transect within species	7	6.419	0.09457	6.29*
Within populations	28	15.525	0.55446	36.85
Total	36	38.703	1.50471	100
Group = all separate morphs				
Among Group	5	19.070	0.46957	39.10*
Among transects within species	3	4.108	0.17700	14.76*
Within transects	28	15.525	0.55446	46.14
Total	36	38.703	1.20103	100

other morphs. Significant differences in chloroplast haplotype within morphs at different transects was also detected, though this accounted for a much smaller proportion of variation (17% between pure species morphs; 6% between *E. nemorosa* and its hybrid compared with *E. fharaidensis*, *E. scottica* and *E. fharaidensis x scottica* and 15% between separate morphs) (Table 5:8). A randomisation test performed using the program Resampling Stats for Excel showed that haplotype h20 was significantly associated with *E. fharaidensis* ($p < 0.01$), h02 and h21 were significantly associated with *E. nemorosa* ($p < 0.05$) and haplotype h02 was significantly associated with *E. nemorosa x fharaidensis* ($p < 0.01$). Within the two groups, h02 is found to be significantly associated with [*E. nemorosa* and *E. nemorosa x fharaidensis*] ($p < 0.001$) and haplotype h20 to be significantly associated with [*E. fharaidensis*, *E. fharaidensis x scottica* and *E. scottica*] ($p < 0.01$). Therefore, it can be concluded that chloroplast haplotypes are not randomly associated with different morphs.

In Faraid Head, microsatellite variation seems to be partitioned between large (*E. nemorosa* and *E. nemorosa x fharaidensis*) and small flowered (*E. fharaidensis*, *E. fharaidensis x scottica* and *E. scottica*) morphs, and not among individually identified morphs (Figure 5:10). The difference between these morphs accounts for 40% of the overall genetic variation, and there is no evidence for spatial genetic structuring between the two transects (Table 5:9). Unlike in the Loch an Daimh Mor grid, *E. fharaidensis* and *E. scottica* are not represented by different STRUCTURE groups. In this case, (as compared to the Loch an Daim Mor grid) each separate morph is not prepresented by a separate STRUCTURE group.

Table 5:9. Analysis of molecular variance (AMOVA) for microsatellite multilocus phenotypes from the Faraid Head site. d.f. = degrees of freedom; SSq – sums of squares (**p < 0.001; *p < 0.05; ^{NS} = Not Significant).

	d.f.	SSq	Variance components	% total Variation
Group = morph				
Among group	5	154.601	2.11678	40.44*
Among transects within groups	3	14.368	0.15443	2.95*
Within groups	81	240.044	2.96351	56.61**
Total	89	409.022	5.23472	100
Group = transect				
Among group	1	3.334	-0.01118	-0.98^{NS}
Among species within groups	7	19.843	0.59415	52.24**
Within groups	28	15.525	0.55446	48.75**
Total	36	38.703	1.13744	100

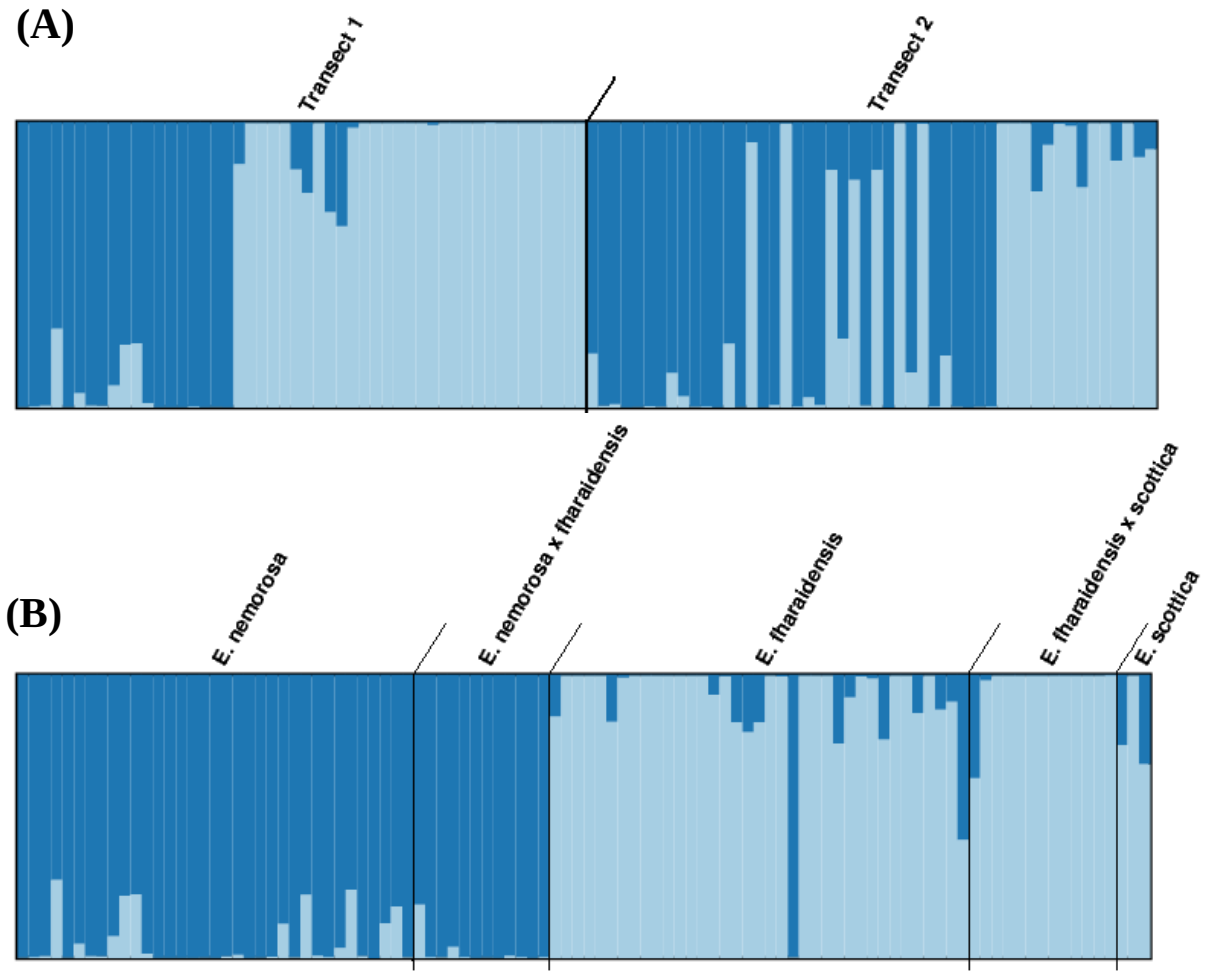


Figure 5:10. STRUCTURE analyses of *Euphrasia* morphs detected at the Faraid Head site. K (the most likely number of structural groups) is 2. (A) = results ordered according to position along separate transects. (B) = results ordered according to morph.

5.3.2.3. Environmental heterogeneity within sites

Whilst there is no significant difference between any of the environmental variables among the two transects, within transects there is a significant difference in the moisture levels along Transect 1 (Light: F-value of 0.3765, $p > 0.1$; Moisture: F-value of 3.8506, $p < 0.001$; Reaction: F-value of 1.3, $p > 0.1$; Nitrogen: F-value of 1.6571, $p = 0.096$; and Salt: F-value of 0.9553, $p > 0.1$), and levels of nitrogen in Transect 2 (Light: F-value of 0.1958, $p > 0.1$; Moisture: F-value of 1.3112, $p > 0.1$; Reaction: F-value of 1.3497, $p > 0.1$; Nitrogen: F-value of 3.7431, $p < 0.001$; and Salt: F-value of 0.4109, $p > 0.1$) (Table 5:10). It can therefore be seen that there is indeed environmental heterogeneity within the Faraid Head site.

Table 5:10. Analysis of variance (ANOVA) of mean Ellenberg values for five environmental variables at the Faraid Head site within among transects. d.f. = degrees of freedom; MS = mean square (***$p < 0.001$, $^{+}p < 0.1$).

Source of Variation	d.f.	MS Light	MS Moisture	MS Reaction	MS Nitrogen	MS Salt
Transect 1						
Among quadrats	10	0.13	5.825***	1.852	2.608^{+}	0.296
Residuals	151	0.346	1.513	1.424	1.574	0.309
Transect 2						
Among quadrats	9	0.068	2.317	1.963	4.872***	0.218
Residuals	153	0.345	1.767	1.454	1.302	0.53
Among Transects						
Among Transects	1	0.233	3.74	1.001763	1.5	0.583
Residuals	323	0.331	1.79	1.47	1.57	0.411

5.3.2.4. Association between genotypes and environment

Morphs were tested at three different levels to look for differences in habitat preference: (i) between morphs of putatively pure species (*E. nemorosa*, *E. fharaidensis* and *E. scottica*); (ii) between *E. nemorosa* and its hybrid morph and all other morphs ([*E. nemorosa* and *E. nemorosa x fharaidensis*] and [*E. fharaidensis, E. fharaidensis x scottica* and *E. scottica*]), and (iii) all separate morphs (Table 5:11). A significant difference in the level of moisture was found between *E. nemorosa* and its hybrid morph and the *E. fharaidensis/E. scottica* morphs (Light: F-value of 0.1556, $p > 0.1$; Moisture: F-value of 7.5244, $p < 0.01$; Reaction: F-value of 1.8123, $p > 0.1$; Nitrogen: F-value of 0.7805, $p > 0.1$; and Salt: F-value of 1.033, $p > 0.1$) (Table 5:11). *E. fharaidensis, E. fharaidensis x scottica* and *E. scottica* are found in locations with higher moisture levels than *E. nemorosa* and *E. nemorosa x fharaidensis* (Figure 5:11). At no other level of analysis was there any significant difference between the habitats of each morphological group (Table 5:11). However this is also the level at which different morphs split into different genotypes (Figure 5:10), suggesting that there may be an association between genotype and environment.

Table 5:11. Analysis of variance (ANOVA) of mean Ellenberg values for five environmental variables at the Faraid Head site among morphs. d.f. = degrees of freedom; MS = mean square (***p < 0.001, $^{+}$p < 0.1).

Source of Variation	d.f.	MS Light	MS Moisture	MS Reaction	MS Nitrogen	MS Salt
Group = pure 'speices'						
Among groups	2	0.0165	5.045	4.06^{+}	2.698	0.11
Residuals	100	0.3074	2.164	1.503	1.574	0.427
Group = *E. nemorosa* and its hybrid morph vs all other morphs						
Among groups	1	0.047	15.67**	2.796	1.277	0.4
Residuals	168	0.302	2.08	1.543	1.636	0.388
Group = all separate morphs						
Among groups	4	0.02	4.95^{+}	2.303	1.454	0.184
Residuals	165	0.308	2.1	1.532	1.638	0.393

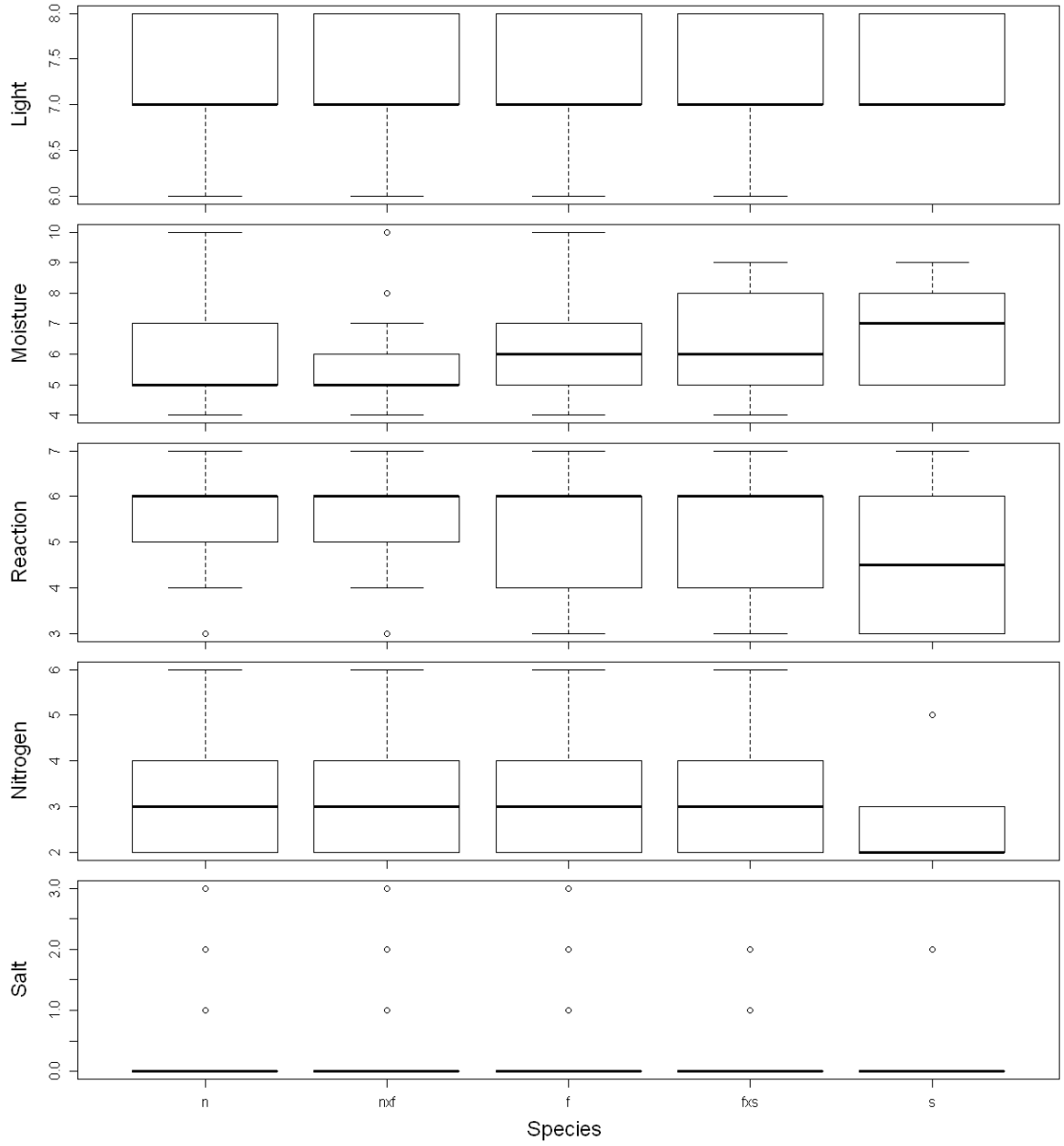


Figure 5:11. Ellenberg values for five environmental variables, calculated from associated species within quadrats. Ellenberg values are shown for each morph present in the Faraid Head site: *E. nemorosa* (n), *E. nemorosa x fharaidensis* (nxf), *E. fharaidensis* (f) *E.fharaidensis x scottica* (fxs) and *E. scottica* (s).

5.3.3. Betty Hill

5.3.3.1. Diversity and distribution of morphs

The overview of collections from Betty Hill (Figure 5:12) shows that there is a great deal of morphological variation among and within transects. Betty Hill shows the greatest degree of morphological variation of all the sites, with 11 different morphs being detected (*E. marshallii, E. marshallii x, E. marshallii x arctica, E. marshallii x foulaensis, E. marshallii x fharaidensis, E. fharaidensis, E. electa x fharaidensis, E. electa x, E. electa* and *E. foulaensis*) along with 9 individuals of unknown taxonomic origin. The most common morphs found at the Betty Hill site were *E. electa x, E. foulaensis, E. fharaidensis* and *E. marshallii x fharaidensis* all represented by more than 10 individuals. Whilst individual quadrats may contain a number of different morphs, as in the Faraid Head site, they generally seem to be dominated by one morphological group. Different morphs were found to be present at different frequencies within each transect. In Transect 1, putative hybrids of *E. marshallii* (*E. marshallii x* and *E. marshallii x arctica*) and *E. electa x* dominate, although *E. fharaidensis* is also present. In the first four quadrats of Transect 2, only *E. foulaensis* and *E. marshallii* are found. Between quadrats Q5 and Q11 a higher diversity of morphological variation is detected, with morphs changing quickly from *E. electa x fharaidensis* in Q5 to a mix of *E marshallii x foulaensis* and *E. electa* between Q6 and Q10 and then to *E. fharaidensis* in Q11. Transect 3 is different again. Unlike transects 1 and 2, no *E. electa* or hybrid with *E. electa* was detected. Morphs of *E. foulaensis* are identified in the first three quadrats, and between quadrats Q4 and Q11 morphs of *E. fharaidensis* and *E. marshallii x fharaidensis* dominate. Transect 3 displays the lowest level of morphological diversity, with only 5 different morphs being detected and transect 2 the highest, with 7 different morphs identified.

5.3.3.2. Genetic variation among morphs

At the Betty Hill site, increased morphological diversity does appear to coincide with an increase in genetic diversity (Figure 5:12). The overview in Figure 5:12 seems to suggest a general correspondence between chloroplast type and morph, and also between STRUCTURE group and

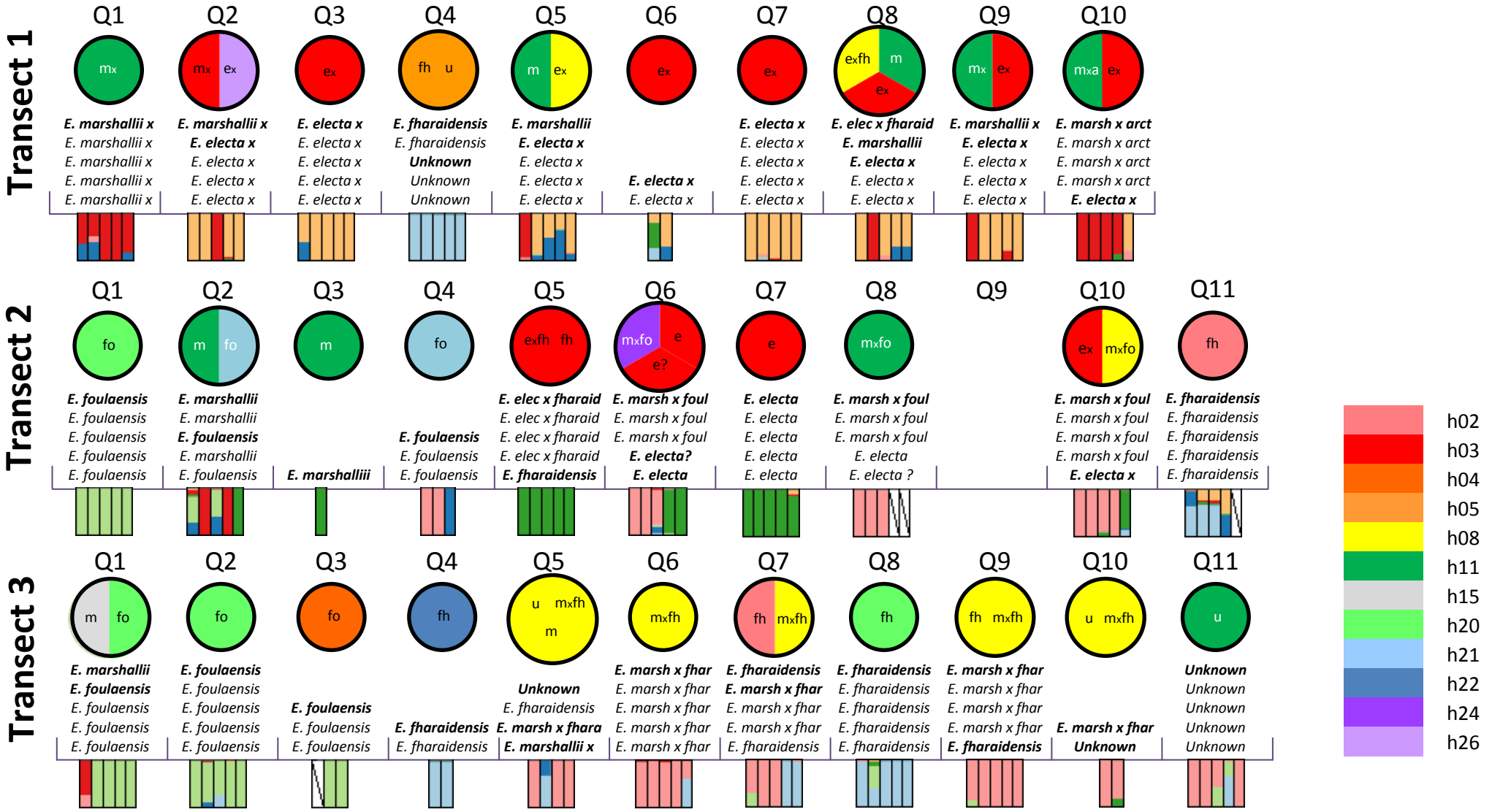


Figure 5:12. Overview of sampling at the Betty Hill site. Quadrats were placed at 5m intervals along 50m transects (Q). Individual morphs collected are listed for each quadrat. From each quadrat, one individual of each species was sequenced at the the trnL$^{(UAG)}$-rp132 chloroplast locus (indicated in bold from the list). Chloroplast haplotypes are represented in the circles by the colours shown in the key at the right. The letters within the circles indicate the morph in which the haplotype was detected (*E. marshallii* (m); *E. marshallii hybrid* (mx); *E. marshallii x arctica* (mxa); *E. marshallii x foulaensis* (mxa); *E. marshallii x fharaidensis* (mxa); *E. foulaensis* (fo); *E. fharaidensis* (fh); *E. electa* (e); *E. electa?* (e?) *E. electa hybrid* (ex); *E. electa x fharaidensis* (exfh)). The rectangular graphs below the list of sampled morphs indicates the STRUCTURE group of each individual (individuals moving down the list are indicated from left to right along the graph).

different morph. This is consistent across transects. For proposed hybrids in which two progenitor parents have been identified, individuals are generally found to share a chloroplast haplotype with one of these putative progenitor parents with the exception of *E. marshallii x foulaensis* in Q6 which has a unique chloroplast haplotype (h24). A more in depth analysis of chloroplast sequence data and microsatellite loci follows.

Sequencing of the trnL$^{(UAG)}$-rp132 chloroplast locus resulted in the detection of 14 sequence differences across the site including point mutations (polymorphic sites 1, 165, 235, 291, 303, 433, 481, 489, 600 and 628), indels (polymorphic site 339, 393, 422 and 516) and an inversion (polymorphic site 516) (Table 5:12). In total 12 haplotypes were identified across the Betty Hill site (Figure 5:13).

Table 5:12. Polymorphic sites detected in the trnL$^{(UAG)}$-rp132 chloroplast locus sequence at the Betty Hill site.

Haplotype	Polymorphic Sites (relative position in sequence)													
	1	165	235	291	303	339	393	422	433	481	516	589	600	628
h02	C	C	A	T	G	:	G	:	T	C	AGTTA	G	G	T
h03	T	C	C	C	T	:	G	:	G	C	TTACT	A	G	T
h04	T	C	C	C	T	:	G	:	G	C	AGTTA	A	G	T
h05	C	C	A	T	G	:	G	:	T	C	AGTTA	G	T	T
h08	C	C	A	T	G	:	G	:	T	C	TTACT	G	G	T
h11	C	C	A	C	G	:	:	:	T	A	AGTTA	G	G	T
h15	T	T	C	C	T	:	G	:	G	C	TTACT	A	G	T
h20	C	C	A	T	G	:	G	:	T	C	TTACT	G	G	G
h21	C	C	C	C	G	:	G	:	T	C	:	G	G	T
h22	C	C	A	T	G	:	G	:	T	C	AGTTA	G	T	T
h24	C	C	A	C	G	:	G	:	T	A	TTACT	G	G	T
h26	C	C	A	T	G	T	G	T	T	C	AGTTA	G	G	T

Again, morphological variation was split into three separate categories in order to perform an AMOVA testing for evidence of variation in chloroplast haplotypes among morphs. These categories were: (i) between morphs of putatively pure species (*E. marshallii*, *E. fharaidensis* and *E. foulaeneis*); (ii) between hybrid groups ([*E. marshallii*, *E. marshalii x*, *E. marshallii x arctica*, *E. marshallii x fharaidensis*], [*E. fharaidensis*], [*E. electa x fharaidensis*, *E. electa x*, *E. electa*] and [*E. foulaensis*]) and (iii) all separate morphs (Table 5:13).

Significant partitioning of chloroplast sequence variation was detected among pure species (15%) and among *E. marshallii* and its hybrid morphs when compared against other morphological

variation (12%) (Table 5:13). Significant levels of partitioning of variation were also detected among transects within species groups (21% among pure species, 9% among the *E. marshallii* group and all other morphs and 58% among all separate morphs). No significant partitioning of chloroplast haplotype variation was found among separate morphs. However, a randomisation test performed among separate morphs using the program Resampling Stats for Excel showed that several chloroplast haplotypes were significantly associated with specific *Euphrasia* morphs: h11 was associated with *E. marshallii* ($p < 0.01$); h08 with *E. marshallii x fharaidensis* ($p < 0.01$); h03 with *E. electa x* ($p < 0.001$); h02 with *E. fharaidensis* ($p < 0.05$) and h04, h21 and h22 with *E. foulaensis* (all $p < 0.05$). This shows that even though there is a great deal of variation in chloroplast haplotypes at the Betty Hill site among and within species, there is some pattern to their distribution that can be accounted for by morphological and not spatial variation.

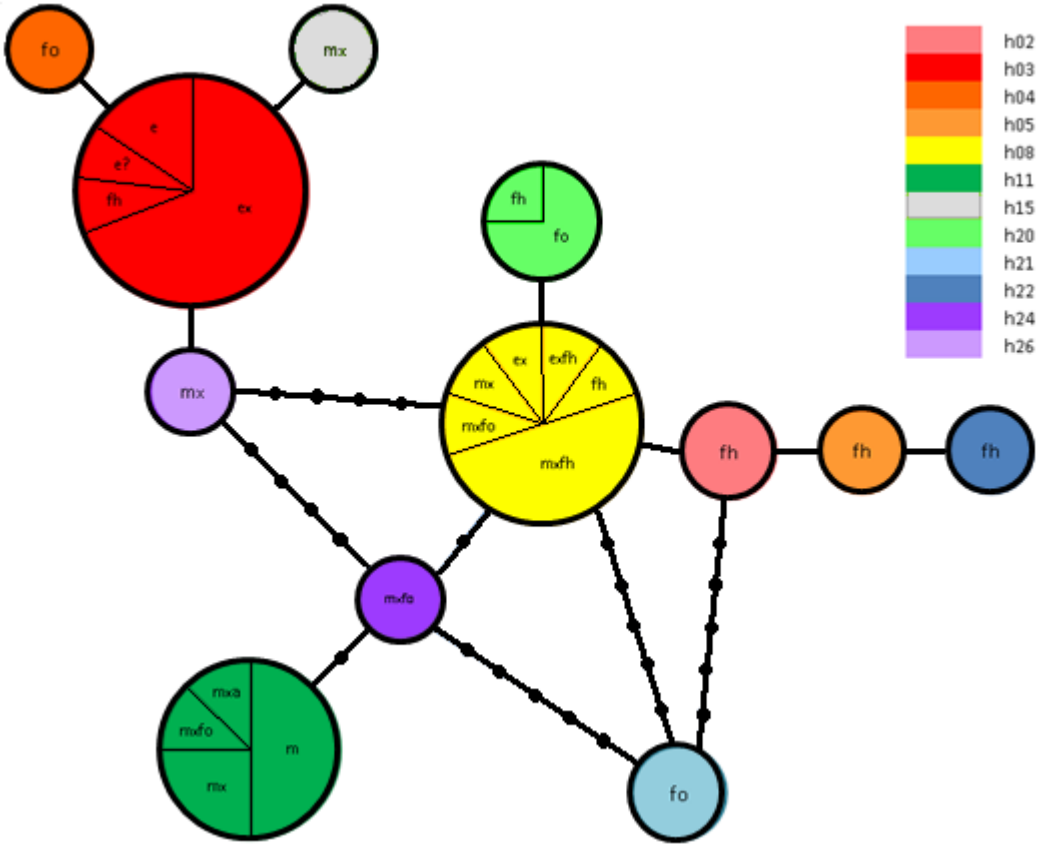


Figure 5:13. Minimum spanning network (MSN) representing chloroplast haplotype groups from Betty Hill. Haplotypes are represented by colours as shown in the key. Dots along the line indicate missing haplotypes. The area of the circle represents the number of individuals with each haplotypes. Species identifications are represented as follows: *E. marshallii* (m); *E. marshallii hybrid* (mx); *E. marshallii x arctica* (mxa); *E. marshalli x foulaensis* (mxa); *E. marshallii x fharaidensis* (mxa); *E. foulaensis* (fo); *E. fharaidensis* (fh); *E. electa* (e); *E. electa?* (e?) *E. electa hybrid* (ex); *E. electa x fharaidensis* (exfh).

Table 5:13. Analysis of molecular variance (AMOVA) for the trnL$^{(UAG)}$-rp132 chloroplast locus for *Euphrasia* across the Betty Hill site. d.f. = degrees of freedom; SSq = sums of squars (**p < 0.01, *p < 0.05).

	d.f.	SSq	Variance components	% total Variation
Group = pure species morphs				
Among Groups	2	10.856	0.38826	14.78*
Among transects within species	5	14.110	0.56291	21.43*
Within transects	10	16.750	1.67500	63.78
Total	17	41.722	2.62617	100
Group = *E. marshallii* and its hybrid morphs vs everything else				
Among Groups	1	8.1610	0.25434	12.41*
Among transects within species	7	17.283	0.18697	9.12*
Within transects	35	56.283	1.60810	78.47
Total	43	81.727	2.04941	100
Group = all separate morphs				
Among Groups	9	27.962	-0.30479	-15.15
Among transects within species	10	34.992	1.17887	58.60**
Within transects	30	34.125	1.12750	56.55
Total	49	97.080	2.01158	100

Variation among microsatellite multilocus phenotypes seems to be structured between morphological groups and not between transects (Figure 5:14). An AMOVA of the microsatellite data indicates that morphological grouping accounts for 31% of the total genetic variation (Table 5:14). A significant proportion of variation is detected among transects (Table 5:14), but this may be accounted for by the different levels of morph representation within each transect. A STRUCTURE analysis (Figure 5:14) of the microsatellite data indicate that genetic structuring is not at the spatial scale but that changes in structural groups within and between transects are associated with changes in the presence of morphological groups. There are fewer structural groups (K = 7, though only 6 groups are widely represented) than separate morphological identifications (10), with some groups clustering together within structural units. *E. marshallii*, *E. marshallii x* and *E. marshallii x arctica*, all large flowered, and thus outcrossing *Euphrasia* can be seen to cluster together; *E. marshallii x foulaensis* and *E. marshallii x fharaidensis* which are medium-sized flower morphs of a hypothesised intermediate levels of outcrossing, cluster together as do *E. electa* and *E. electa x fharaidensis*. The small flowered *E. fharaidensis*, *E*

electa x and *E. foulaensis* all constitute independent morphological and genetic groups with *E. fharaidensis* and *E. foulaensis* maintaining their genetic identity to a large extent, even when found growing in different transects (Figure 5:12).

Table 5:14. Analysis of molecular variance (AMOVA) for microsatellite multilocus phenotypes from the Betty Hill site. d.f. = degrees of freedom; SSq = sums of squares (**p < 0.001; *p < 0.05).

	d.f.	SSq	Variance components	% total Variation
Group = morph				
Among group	9	391.090	2.00877	30.79*
Among transects within groups	10	117.273	2.23453	34.25**
Within groups	104	237.112	2.27993	34.95
Total	123	745.476	6.52322	100
Group = transect				
Among group	2	201.672	1.61457	23.60*
Among species within groups	17	306.691	2.94586	43.07**
Within groups	104	237.112	2.27993	33.33
Total	123	745.476	6.84036	100

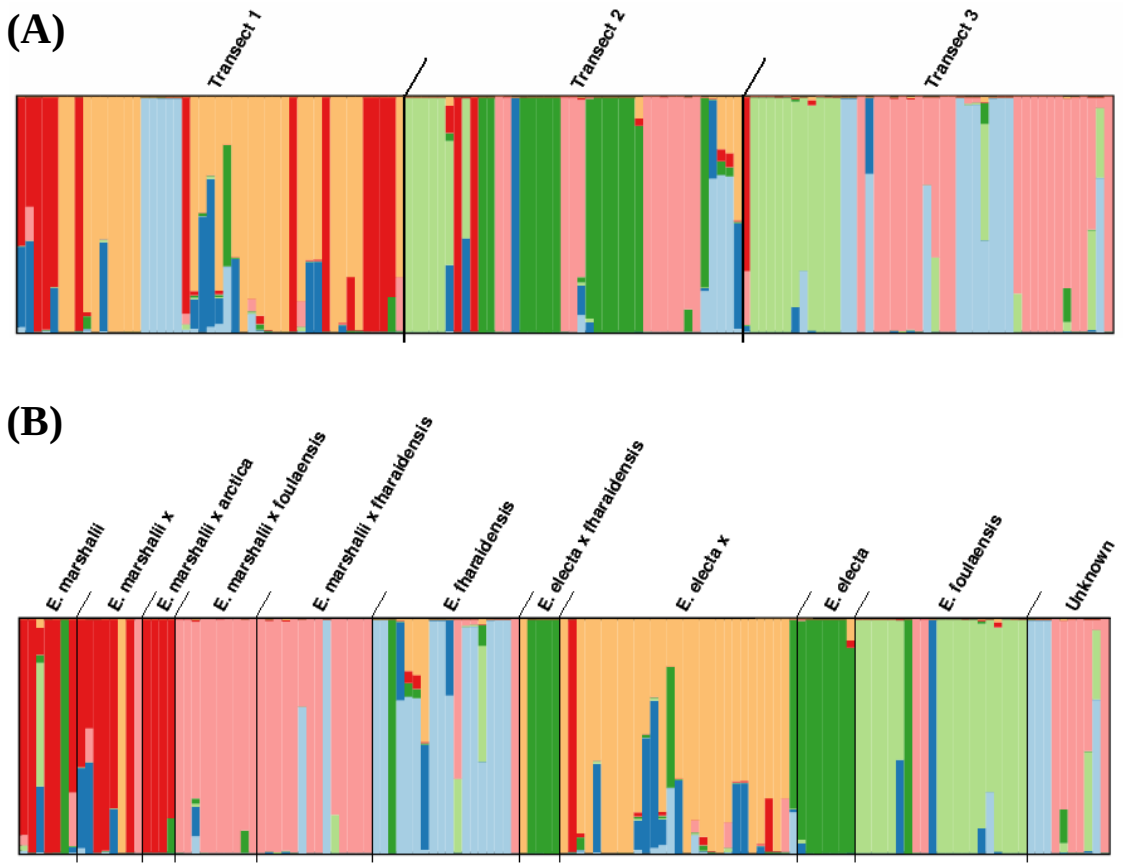


Figure 5:14. STRUCTURE analyses of *Euphrasia* morphs detected in the Betty Hill site. K (the most likely number of structural groups) is 7. (A) = results ordered according to position along transects. (B) = results ordered according to morph.

5.3.3.3. Environmental heterogeneity within sites

There is a great deal of habitat heterogeneity within the Betty Hill site. This is apparent both within and between transects, with the exception of within transect variation among quadrats in Transect 2 (Table 5:15). Within Transect 1, there is significant variation among quadrats with regard to moisture and nitrogen levels (Light: F-value of 0.6542, $p > 0.1$; Moisture: F-value of 1.9747, $p < 0.05$; Reaction: F-value of 1.5661, $p > 0.1$; Nitrogen: F-value of 2.4871, $p < 0.05$; and Salt: F-value of 0.8459, $p > 0.1$). Within Transect 3, significant variation amongst quadrats is found with regard to moisture, nitrogen and salt levels, and to a certain extent, levels of acidity (reaction) (Light: F-value of 0.5376, $p > 0.1$; Moisture: F-value of 6.0348, $p < 0.001$; Reaction: F-value of 1.6867, $p = 0.089$; Nitrogen: F-value of 2.1963, $p < 0.05$; and Salt: F-value of 2.0541, $p < 0.05$) (Table 5:15). There are significant differences between every environmental variable when transects are compared among one another (Light: F-value of 6.1949, $p < 0.05$; Moisture: F-value of 10.911, $p < 0.001$; Reaction: F-value of 5.6228, $p < 0.05$; Nitrogen: F-value of 4.0476, $p < 0.05$; and Salt: F-value of 5.2354, $p > 0.01$) (Table 5:15).

Table 5:15. Analysis of variance (ANOVA) of mean Ellenberg values for five environmental variables at the Betty Hill site. Within transect and among transect variation has been tested. d.f. = degrees of freedom; MS = mean square (***$p < 0.001$; **$p < 0.01$, *$p < 0.05$, $^{+}p < 0.1$).

Source of Variation	d.f.	MS Light	MS Moisture	MS Reaction	MS Nitrogen	MS Salt
Transect 1						
Among quadrats	9	0.096	2.641*	3.152	2.148*	0.36
Residuals	153	0.272	1.338	2.012	0.864	0.425
Transect 2						
Among quadrats	10	0.265	1.003	3.25	1.086	0.644
Residuals	131	0.474	1.727	2.77	1.359	1
Transect 3						
Among quadrats	10	0.185	7.913***	3.229^{+}	1.328*	1.247*
Residuals	139	0.344	1.311	1.914	0.605	0.607
Among Transects						
Among Transects	2	2.149**	17.53***	12.80**	3.92*	3.506**
Residuals	425	0.347	1.61	2.28	0.97	0.67

5.3.3.3. Association between genotypes and environment

As with the different morphs found within the Faraid Head site, morphs detected in the Betty Hill site were tested at three different levels to look for differences in habitat preference: (i) between morphs of putatively pure species (*E. marshallii*, *E. fharaidensis* and *E. foulaeneis*); (ii) between hybrid groups ([*E. marshallii*, *E. marsahlii x*, *E. marshallii x arctica*, *E. marshallii x foulaensis* and *E. marshallii x fharaidensis*], [*E. fharaidensis*], [*E. electa x fharaidensis*, *E. electa x*, *E. electa*] and [*E. foulaensis*]), and (iii) all separate morphs (Table 5:16). The most highly significant level of habitat variation among species was detected in the analysis when all morphs were compared as separate entities, with significant differences in moisture being detected (Light: F-value of 0.3765, $p > 0.1$; Moisture: F-value of 3.8506, $p < 0.001$; Reaction: F-value of 1.3, $p > 0.1$; Nitrogen: F-value of 1.6571, $p = 0.096$; and Salt: F-value of 0.9553, $p > 0.1$) (Table 5:16). *E. marshallii x fharaidensis* is found growing in quadrats that tend to be drier than average, and *E. marshallii x foulaensis*, *E. electa x fharaidensis* and *E. foulaensis* tend to be found in slightly damper sites (Figure 5:15).

Table 5:16. Analysis of variance (ANOVA) of mean Ellenberg values for five environmental variables at the Betty Hill site among morphs. d.f. = degrees of freedom; MS = mean square (***$p < 0.001$, $^{+}p < 0.1$).

Source of Variation	d.f.	MS Light	MS Moisture	MS Reaction	MS Nitrogen	MS Salt
Among pure speices only						
Among morphs	2	0.012	5.278^{+}	2.086	0.125	0.095
Residuals	104	0.437	2.119	2.678	1.081	0.618
Among hybrid groups						
Among quadrats	3	0.175	4.87*	2.53	0.176	0.17
Residuals	266	0.395	1.85	2.52	1.119	0.623
Among all separate morphs						
Among Transects	9	0.183	6.06***	1.65	1.063^{+}	0.376
Residuals	260	0.4	1.74	2.55	1.11	0.627

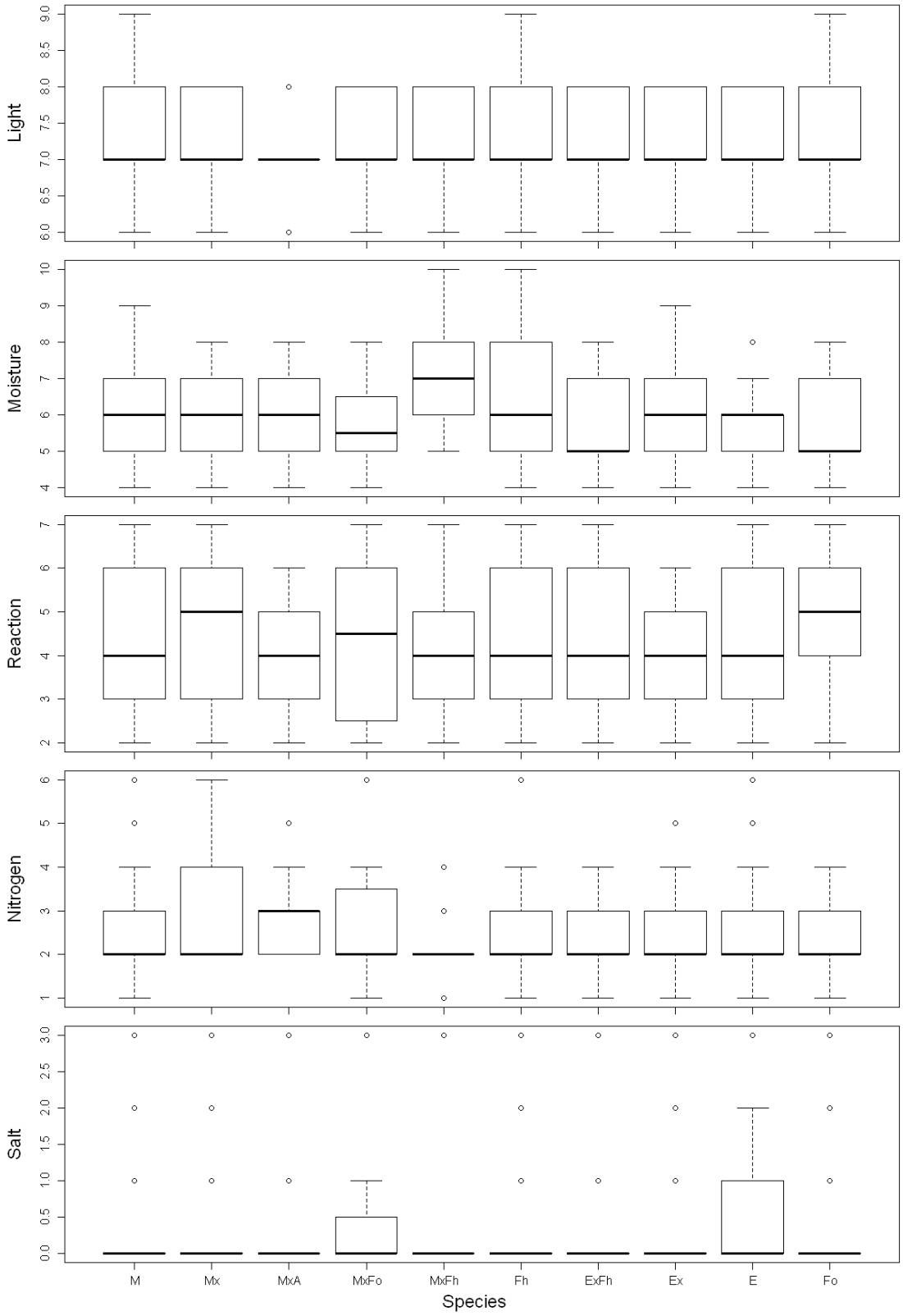


Figure 5:15 Ellenberg values for five environmental variables, calculated from associated species within quadrats. Ellenberg values are shown for each morph present in the Betty Hill site: *E. marshallii* (M), *E. marsahlii x* (Mx), *E. marshallii x arctica* (MxA), *E. marshallii x fharaidensis* (MxFh) *E. fharaidensis* (Fh), *E. electa x fharaidensis* (ExF), *E. electa x* (Ex), *E. electa* (E) and *E. foulaensis* (Fo).

5.4. Discussion

5.4.1. Genetic heterogeneity among morphs

Within all sites, the analysis of molecular markers from both the chloroplast and nuclear genomes provides clear evidence that individual morphs or groups of morphs are genetically distinct. This shows that they do indeed correspond to discrete genetic entities. The complex population studies show a large number of such discrete genetic entities. Thus at Loch an Daimh Mor there were 3 different genetic entities by STRUCTURE analysis, corresponding to the three morphological entities, at Faraid Head 2 structural units were detected amongst the 5 morphological entities, and in Betty Hill, the 10 morphological entities detected were split into 6 broad genetic groups. These complex populations also showed a high degree of chloroplast variation, with 5, 6 and 12 different chloroplast haplotypes being detected within Loch and Daimh Mor, Faraid Head and Betty Hill respectively. This is in stark contrast to what was witnessed in sites in other parts of the *Euphrasia* range of widespread species, where only one species is present. In these populations microsatellite data tends to cluster into one structural group and only one or two chloroplast haplotypes are present (Chapter 4). The high mutation rate of microsatellites has served this study well as they have allowed for the detection of genetic differences between morphs within a population. At no site is molecular variance completely homogenous among all morphs. It should also be noted that morphs recognised as intermediates between two 'pure' species, did not appear as F1 hybrids in the STRUCTURE analysis, as their name would suggest, but most likely inbreeding segregants from such hybrids. Whilst genetic evidence does not prove the hybrid origin of these morphs, shared chloroplast haplotypes with suggested progenitor parents does indicate that historical hybridisation could have resulted in the present-day morphological entities identified.

5.4.1.1. Loch an Daimh Mor

Within the Loch an Diamh Mor grid, chloroplast groups were generally subdivided between h20 in *E. fharaidensis* and h03 in *E. fharaidensis x scottica* and *E. scottica*. This suggests that *E. scottica* could be the maternal progenitor of the *E. fharaidensis x scottica* morph. If hybridisation

occurred, it is likely to be historical, as *E. fharaidensis x scottica* clusters independently under STRUCTURE analysis, and does not appear to be an F1 hybrid. These presumed ancient hybrids are however all based on individuals from the same 1 m^2 quadrat and perhaps cluster together because they are in the same family. Even though there is genetic variation within morphs in the grid at Loch an Daimh Mor, each morph clusters into discrete groups that correspond to separate inbred lines. In this case, a historical hybridisation event may have resulted in a novel hybrid that has propagated through self fertilisation to result in a genetically isolated and novel recombinant inbred line.

5.4.1.2. Faraid Head

Within Faraid Head, unlike at Loch an Daimh Mor, *E. fharaidensis*, *E. fharaidensis x scottica* and *E. scottica* were all found to cluster together in the STRUCTURE analysis and all shared the h20 chloroplast haplotype. Faraid Head was chosen as a site because it was known that *Euphrasia* in this site display morphological heterogeneity, and in particular because *E. scottica* was known to be present. Unfortunately it was a very dry year and few *E. scottica* (3 specimens), which prefer moist sites (Chapter 4), were collected. It may be that in this year, conditions were not suitable for the germination and survival to adulthood of *E. scottica* in Faraid Head, and those individuals that did survive to adulthood had, at some point in the past, experienced a certain level of introgression with *E. fharaidensis* that allowed them to survive in this dry year. This could also explain the presence of the same chloroplast group (h20) in all *E. fharaidensis*, *E. fharaidensis x scottica* and *E. scottica* morphs.

The large flowered and presumed outcrossing morphs, *E. nemorosa* and *E. nemorosa x fharaidensis* also clustered together in the STRUCTURE analysis. The presence of the h02 chloroplast haplotype, significantly associated with these morphs, suggests that *E. nemorosa* is the original maternal progenitor of the *E. nemorosa x fharaidensis* morph. French et al. (2008) showed that there was extensive genetic exchange between outcrossing species suggesting that they represent more of an outcrossing complex than separate species. This theory is in accord with our results at Faraid Head.

5.4.1.3. Betty Hill

At Betty Hill, there is a striking correspondence between genetic variation and morphological identification, both at the chloroplast and nuclear level, despite the huge mix of morphological variation across the population. This is particularly noticeable at the nuclear genetic level where STRUCTURE groups correspond well with separate morphs or morphological groups. Again, morphs with an intermediate phenotype of two putatively 'pure' species appear to be stable genetic lines, and have not resulted from recent hybridisation events. As with the French et al. (2008) study of paired populations of outcrossing tetraploid *Euphrasia* and the Faraid Head site, the large flowered and presumed outcrossing morphs of Betty Hill, *E marshallii*, *E. marshallii x* and *E. marshallii x arctica*, cluster into the same STRUCTURE group, suggesting an outcrossing complex. Presumed hybrids between outcrossing and inbreeding species, *E. marshallii x fharaidensis* and *E. marshallii x foulaensis* also appear to cluster into the same STRUCTURE group. They perhaps represent a group that experience an intermediate level of outcrossing which has led to them becoming homogenised at a neutral genetic level but remain separate from other more highly outcrossing morphs. Small flowered, inbreeding morphs tend to cluster into separate STRUCTURE groups suggesting that their self fertilising breeding system is sufficient to create a strong barrier to genetic exchange between morphs. This allows small flowered morphs to retain a distinct genetic identity within a sea of morphological variation.

5.4.2. Habitat heterogeneity within sites and its association with morph presence

The ANOVA of different environmental variables showed clear evidence for the presence of a wide range of different habitat types within sites. This was most notable in the Betty Hill site where there were significant differences between all five environmental variables tested among transects, and where significant differences were detected in more than one environmental variable within two of the three transects. This increase in habitat heterogeneity at the Betty Hill site also corresponds with an increase in observed morphological and genetic variation. In the Loch an Daimh Mor grid and within the Faraid Head transects, there was a relatively low level of habitat heterogeneity, corresponding to the lower levels of morphological and genetic complexity.

The distribution of *Euphrasia* morphs was found to be significantly associated with specific habitats within sites in Faraid Head and Betty Hill. This is not entirely unexpected as one of the major factors influencing taxonomy in *Euphrasia* is ecological variation (Pugsley 1930; Yeo 1978). Generally speaking different levels of moisture within a quadrat were most likely to affect *Euphrasia* morph presence. Nitrogen levels were also shown to have an effect on the distribution of *Euphrasia* morphs, though to a lesser extent. Whilst there was evidence of differential morph presence in different habitats, this analysis did not manage to separate out each morph into separate preferred habitats. This may be because different morphs are able to inhabit multiple habitat types, but it may also be that Ellenberg indicator values are not sensitive enough to detect fine scale differences in habitat among different morphs where they exist. The associated species assigned to a particular morph were those that were present within the quadrats in which that morph was found. Sometimes more than one morph was present in a quadrat. These morphs would therefore share the same associated species and hence environmental indicator values, even if habitats varied within the quadrat. To get a more accurate estimate of the habitats in which particular morphs were found, different methods of acquiring environmental indicator values should be used, such as the use of pH meters, taking soil samples and restricting the range around a plant in which associated species can be recorded. Nevertheless, associated species lists did indicate that there were ecological differences amongst morphs, even using this relatively inaccurate system. This indicates that there is a possibility that different morphs are associated with different habitats.

5.4.3. Hybridisation, genetic isolation and habitat heterogeneity

It has been predicted that hybridisation between widespread outcrossing and inbreeding taxa may play an important role in *Euphrasia* diversification in Britain (French et al. 2008). It certainly seems that in these populations intermediate morphs are of hybrid origin as they tend to share a chloroplast haplotype with one of their proposed progenitor species. These hybridisation events appear to result in genetically distinct hybrid morphs that are likely to be isolated due to their highly inbreeding mating system. Our observations would also suggest that hybridisation between separate inbreeding species may also play an important role in the production of novel

inbreeding lines. Since outcrossing in *Euphrasia* occurs via insect pollination (Pugsley 1930; Yeo 1978), it is possible that the presence of outcrossing species within a population of mixed *Euphrasia* morphs enhances the presence of suitable pollinating insects, and thus encourages the pollination of inbreeding species. It may also be that habitat heterogeneity increases the likelihood of co-occurrence of inbreeding species, heightening the likelihood of a hybridisation event. The presence of intermediate habitats within a site may also play a role in allowing those hybrid offspring to survive and propagate.

As separate morphs, including the hybrid morphs, represent independent genetic lineages, their association with specific habitats may be attributed to their separate lineages, and not simply to morphological variation occurring as a result of growth under different environmental conditions, a trait that has been shown in common garden experiments between different *Euphrasia* species (Chapter 4, Yeo 1962). This suggests that different morphs are specialised to particular habitat types.

It has long been suggested that for hybrids to form stable populations, the availability of a suitable habitat that is different from that of its progenitor parents is required (Anderson 1948; Templeton 1981; Rieseberg 1997; Rieseberg et al. 2003). In fact, all species of confirmed homoploid hybrid origin are ecologically diverged and display some degree of ecogeographic isolation (Gross & Rieseberg 2005). Many hybrids appear to be found in habitats that form extremes of parental environments, as is evident in many *Helianthus* and *Gossypium* species of ancient hybrid origin (Wendel et al. 1991; Welch & Rieseberg 2002; Rieseberg et al. 2003).

Studies of recombinant inbred hybrid lines of different ecotypes of the grass species, *Avena barbata* (Johansen-Morris & Latta 2006; Latta et al. 2007; Johansen-Morris & Latta 2008) showed that F_6 generation hybrids were capable of outperforming parental ecotypes in novel environments. Whilst, on average, recombinant inbred lines were significantly less fit than their progenitor parents, particularly in the parental habitats, they had a much greater range of fitness values (Johansen-Morris & Latta 2006; Latta et al. 2007; Johansen-Morris & Latta 2008). These experiments showed results consistent with the theory presented by Barton (2001). Whilst most hybrid lines will perform poorly, rare lines will be competitive, especially when presented with a

novel environment. According to these results, for novel hybrid lines to be successful and competitive, a range of habitat types other than those of parental species must be available within a site. This theory has also been demonstrated by Buerkle et al. (2000; 2003) in models that featured (i) habitats for both parental species, and a third habitat that diverged ecologically and (2) only the two parental habitats. In the model containing a unique habitat, it was found that in that habitat hybrids were most successful, and resultant speciation became more likely as the strength of ecological selection increased. Hybrids were shown to be much less successful when there was no unoccupied habitat. Populations that did manage to form within a parental habitat were small and appeared to represent more of a transitory state eventually ending up in the genetic assimilation of one of the parental species.

This theory may help to explain why hybrids were rarely detected between adjacent sites of *E. micrantha* and *E. scottica* (Chapters 3 & 4), where their habitats are very distinct from one another with no gentle gradation between habitats, but are present in abundance in complex sites which have a great deal of habitat heterogeneity. Of course, in complex *Euphrasia* sites, we are dealing with hybridisation between separate species and not the ecotypes described by the *Avena barbata* papers (Johansen-Morris & Latta 2006; Latta et al. 2007; Johansen-Morris & Latta 2008). We are also not suggesting that novel hybrid recombinant inbred lines represent new species as is suggested by the Buerkle et al. papers (2000; 2003), but that these hybrid groups represent a pool of novel genetic variation that potentially have speciation potential.

5.4.4. Summary

From analyses of these separate sites, it has been shown that morphological differences have an underlying genetic basis and there is some evidence to suggest that different morphs have significant associations with different habitats. Generally speaking, morphological variation is associated with changing levels of moisture, although it can also be seen to a lesser extent with changes in nitrogen levels across a site. Sites with higher levels of habitat heterogeneity correspond with sites that hold the greatest levels of morphological diversity. Different morphs within complex sites broadly correspond with discrete genetic lineages. Our analyses of microsatellite data show that inbreeding species tend to cluster separately, and that larger

flowered, outcrossing species tend to cluster together. These results are consistent with the hypothesis that sites with high levels of morphological diversity are actively generating taxonomic biodiversity. They should therefore form the focus for process-based conservation action plans for British tetraploid *Euphrasia*.

Chapter 6: General discussion

Taxonomically complex groups, where species distinctions cannot be easily defined, are characterised by recent and ongoing evolution. In countries with recently introduced flora, such as those that have only been available for colonisation since the retreat of the last ice sheet, many of the endemic taxa belong to taxonomically complex groups (Brochmann et al. 2003; Hollingsworth 2003). Whilst identifying and classifying species within these groups can be problematic, especially with regard to conservation legislation, they can provide unique opportunities for studying speciation. Understanding the processes that create diversity in these groups will not only allow conservationists to conserve taxonomically complex groups more effectively through the implementation of process-based strategies, but will also aid our general understanding of the evolutionary processes involved in speciation.

6.1. What is a plant species?

Classifying groups of plants at the species level has been a major problem for botanists (Mishler & Donoghue 1982). Widely accepted species concepts, such as the biological (Mayr 1940), evolutionary (Simpson 1961; Wiley 1978; Grant 1981; Wiley 1981), cohesion (Templeton 1989) and phylogenetic (Rosen 1978; Dequeiroz & Donoghue 1988; McKitrick & Zink 1988) concepts have not been wholeheartedly accepted by the botanical community (Mishler & Donoghue 1982). This is because there is a great deal of discontinuity in ecological, morphological and genetic variation with frequent reports of interspecific hybrids (Ehrlich & Raven 1969; Mishler & Donoghue 1982; Arnold 1997). However, defining species is important for at least two reasons: First, from a conservation perspective, they tend to be the main focus of conservation effort (Hollingsworth 2003). Secondly, species definitions provide a good basis for the study of the mechanisms involved in speciation, and indeed defining speciation itself.

Reproductive isolation is often invoked as an important criterion for delimiting species. Although the majority of plant species are held together by gene flow and correspond well with reproductively isolated lineages, there are exceptions (Rieseberg 2001). Classifying species in terms of sexual isolation is difficult in species that reproduce asexually, as there is no gene flow

between individuals. Moreover, many species are self-fertilising. Low levels of gene flow between populations means that populations, and not species, represent reproductively isolated units (Ehrlich & Raven 1969; Mishler & Donoghue 1982). The groups that cause perhaps the most difficulty for species delimitation are those in which reproductive isolation between separate entities is not complete. This is particularly common in groups in which active speciation events are currently occurring and in groups with long generation times (Rieseberg & Willis 2007). For these groups, in which species cannot be assigned readily, it has been suggested that species should be defined according to the important patterns and processes that account for the observed variation (Mishler & Donoghue 1982). Many flowering plant groups have recently undergone multiple speciation events, and some effort should be made to gain an understanding of the species in these situations as this could be particularly interesting for studies that look into how speciation occurs.

6.2. Taxonomic treatment of tetraploid *Euphrasia* in Britain

Tetraploid British *Euphrasia* are a prime example of a taxonomically complex group in which reproductive isolation between species is not complete. Species within this group have largely been defined according to morphological and ecological differences (Pugsley 1930; Yeo 1978). This method for delimiting species is controversial, especially in the light of genetic analysis of the group. Large-flowered tetraploid *Euphrasia* have been shown to cluster together genetically, and genetic distinctions between small-flowered tetraploids are not discrete (French et al. 2008). If reproductive isolation was the only criteria for defining species in tetraploid *Euphrasia*, they could be considered to be one polymorphic species. However, this would not do justice to the diversity witnessed in this group. The species concept that comes closest to doing justice to the diversity found in *Euphrasia* is Wu's genic species concept (2001). Here different species are defined by the genes that are responsible for differential adaptation to different natural or sexual environments (2001). This may be very important for defining *Euphrasia* species, but without detailed knowledge of their genetics it is very difficult to identify these adaptive genes, and therefore we have had to concentrate on analysing neutral genetics and trying to recognise potential for adaptive differences between the species that may have a genetic basis. The aim of

this thesis is (i) to look more closely at the genetic integrity of inbreeding tetraploid species and (ii) to gain an understanding of how currently defined species interact to generate novel diversity in tetraploid British *Euphrasia*.

6.3 Maintenance of species distinctions in tetraploid British *Euphrasia*

An initial genetic analysis of tetraploid *Euphrasia* in Britain showed that there is little to no reproductive isolation in the large-flowered tetraploid *Euphrasia* (French et al. 2008). Despite there being no significant reproductive barriers (Yeo 1966), and despite extensive reticulation between tetraploid *Euphrasia*, it was shown that there was a general correspondence between small-flowered species and genetic similarity (French et al. 2008). This is further supported by a detailed analysis of two widespread inbreeding *Euphrasia*, where chloroplast haplotypes were not randomly distributed among species and different species maintained their morphological and ecological integrity across their range (Chapter 4). This suggests that whilst hybridisation between small-flowered taxa is possible, it is not extensive enough to completely merge genetic variation between species within the tetraploid group.

It is likely that a combination of habitat specialisation (Chapter 4) and high levels of inbreeding in the small-flowered *Euphrasia* (Chapter 3) present a large barrier to gene flow among tetraploid progenitor taxa. Whilst there is little to no gene flow among populations, as is evident from analysis of neutral nuclear loci (Chapter 4), their adaptation to specific environments, morphological integrity and associated chloroplast types serve to differentiate progenitor taxa from one another. It is highly likely that adaptation to specific environments is under genetic control and that these adaptive traits are heritable. Therefore, separate inbreeding tetraploid species can be thought of as having separately adapted gene pools. If separate species meet and hybridise, these hybrids will have a mix of gene pools. This is likely to give them an adaptive potential that is different from both the progenitors.

6.4. Generation of diversity in tetraploid British *Euphrasia*

British tetraploid *Euphrasia* are extremely diverse with 15 recognised species, 5 of which are endemic. In fact, these classifications oversimplify the diversity witnessed among tetraploid

Euphrasia, as there is also a swarm of unclassifiable variation which is mainly thought to be generated through hybridisation (Yeo 1978). The diversity present in tetraploid *Euphrasia* is thought to have occurred very recently and is likely to be in a state of flux, with new discrete entities evolving and becoming extinct during the diversification process. Thus, tetraploid *Euphrasia* represent a group in which there is active formation of novel morphs. Many of these populations are impossible to define, with individuals that have an intermediate morphology to that of recognised taxa. Small isolated populations of indistinct morphology may be relatively ephemeral and therefore do not warrant naming under current *Euphrasia* taxonomic concepts. However, if some have managed to persist and form distinct morphological entities with discrete ecological habitats, spreading across a wide geographic area, then they may be defined as separate species. It is likely that diversification in British *Euphrasia* involves a complex interaction of a number of processes combining the effects of ecology, breeding systems and hybridisation between species.

Studies into homoploid hybrid speciation have highlighted the importance of reproductive isolation and ecological specialisation in the maintenance of new hybrid populations (Barton 2001; Gross & Rieseberg 2005). The conditions under which progenitor populations meet and hybridise will affect the fitness of the hybrid genotype and ultimately determine whether a hybrid is able to establish (Barton 2001).

To determine whether hybridisation is a likely route to speciation, it must first be shown that hybridisation is theoretically possible and that there is evidence for hybridisation in nature. Yeo (1966) showed that artificial crosses between separate tetraploid taxa were able to produce fully fertile F1 hybrids. The morphology of these hybrids varied from plants that exhibited either intermediate forms or combinations of characters from both progenitor parents, to hybrids that showed morphological features outside of the range of both parents (Yeo 1966). Successive generations of self-fertilised hybrids resulted in lines that either reverted to the morphology of one of the parent progenitors or to a large extent maintained their original hybrid form (Yeo 1966). Hybridisation has also been shown to occur readily in nature, as fully fertile hybrids were observed in artificial populations of mixed tetraploid *Euphrasia* (Yeo 1966; Liebst 2008).

Therefore, it seems likely that populations of *Euphrasia* that appear to have a combination of morphological features commonly found in currently defined *Euphrasia* taxa, such as those identified in Chapter 5, are of hybrid origin.

Second, one must establish how this hybridisation might lead to the establishment of new species. Reproductive isolation is vital for a new hybrid population to become established and differentiated from its progenitor parents. Reproductive isolation in homoploid hybrids is thought to occur through ecological divergence, spatial isolation and/or rapid chromosomal reorganisation (Grant 1981; Rieseberg 1997; Buerkle et al. 2000). Much attention has been paid to the chromosomal model, and karyotypic divergence appears to have played an important component in roughly half of the well-studied species of homoploid hybrid origin (Gross & Rieseberg 2005). It is unlikely, however, that rapid chromosome rearrangements have played much of a role in reproductive isolation between tetraploid *Euphrasia*, as there do not appear to be strong genetic barriers to hybridisation within the group. More likely, inheritance of a self-fertilising life history is sufficient to provide reproductive isolation. Certainly, from analysis of the partitioning of genetic variation in complex *Euphrasia* populations, small-flowered and presumably highly-selfing morphs formed a greater number of separate genetic clusters than did larger-flowered morphs (Chapter 5). This indicates a higher degree of reproductive isolation. It was also shown in Chapter 3 that small flowered *Euphrasia* of putative hybrid origin displayed very low rates of outcrossing, not significantly different from zero. Predominant self-fertilisation has in fact been advocated as the most important factor maintaining reproductive isolation in other homoploid hybrids, such as the new hybrid species *Senecio eboacensis* (Lowe et al. 2004).

Whilst high selfing rates may maintain the genetic integrity of a new hybrid, it is thought that ecological divergence plays a critical role in their ultimate establishment. Computer simulations indicate that hybrid speciation is unlikely to occur in the absence of niche separation (Grant 1981; Greig et al. 2002). In fact, all clearly defined examples of homoploid hybrid species display a degree of ecological divergence from their progenitor parents, and some even occur in more extreme habitats (Rieseberg 1991; Abbott 1992; Rieseberg 1997). Traits that allow hybrids to occupy new environments may be inherited from their progenitor species (Ellstrand &

Schierenbeck 2000; Barton 2001; Schwarzbach et al. 2001). Hybrids with intermediate trait values, combinations of traits form both parents or extreme trait values may have high fitness in new environments (Gross & Rieseberg 2005). Endemic tetraploid *Euphrasia*, presumably of hybrid origin, seem to exhibit habitat segregation in Scotland. Most occupy sites that are not dissimilar from those occupied by progenitor species. *E. marshallii* (Pugsley 1933, 1936) and *E. rotundifolia* (Pugsley 1929) are restricted to damp basic cliff top turf and *E. campbelliae* (Pugsley 1940) to coastal heathy turf along the northern Scottish coast (Silverside 1991b). However, *E. heslop-harrisonii* (Pugsley 1945) is restricted to grazed estuarine salt marshes, an extreme environment for *Euphrasia* (Silverside 1991c). This transition would require the inheritance of extreme traits, which could occur as a result of transgressive segregation after hybridisation (Rieseberg et al. 1999; Schwarzbach et al. 2001; Lexer et al. 2003a; Lexer et al. 2003b; Gross et al. 2004).

Environmental variation also appears to be an important factor in the establishment of new *Euphrasia* hybrids. In our study, complex populations with the highest degree of environmental variation also exhibited the highest number of presumed hybrid morphs (Chapter 5). Different morphs are also found to be significantly associated with specific habitat types (Chapter 5). Further evidence to suggest of the importance of ecological specialisation comes from adjacent populations of *Euphrasia* species in which only parental habitats are available. In these populations, the proximity of the different species suggests that, whilst outcrossing is rare, some outcrossing between species is likely to occur. However, very little evidence of hybridisation between species was found in these situations (Chapters 3 and 4). It may be that large flowered *Euphrasia* increase the level of outcrossing in complex populations by attracting suitable pollinators, which would make hybridisation in these populations more likely than between adjacent populations of small flowered *Euphrasia.* Post-zygotic selection is likely to play a large role in reducing the number of hybrids detected in adjacent populations. Hybrids may have reduced fitness in parental habitats, and those that do survive are likely to be transitory (Buerkle et al. 2003). It has been shown that hybridisation is far more likely to result in introgression than speciation (Buerkle et al. 2003). This implies that the conditions suitable for supporting stable hybrids are rarely met (Rieseberg & Wendel 1993; Arnold 1997; Buerkle et al. 2000). Therefore,

complex sites with intergrading habitat types that support large numbers of stable hybrids may be of key importance, both in the context of process-based conservation in *Euphrasia* and for the study of homoploid hybrid speciation.

Initial hybrid establishment must necessarily occur in sympatry with both parent species if seeds are not widely dispersed, as is the case in *Euphrasia*. It has been argued that hybrids are unlikely to speciate unless they colonise a new locality that is ecologically and/or spatially isolated (Charlesworth 1995). Thus, whilst complex populations of *Euphrasia* may represent hotspots of hybrid diversity, for a new hybrid to form a new species, it must somehow escape and form an independent population. There is thus more work to be done to fully understand the mechanisms involved in homoploid hybrid speciation in *Euphrasia*.

6.5. Future work

There is clear evidence to suggest that tetraploid *Euphrasia* have extremely porous genomes, as all species seem to be able to produce fertile hybrids readily (Yeo 1966). Further work is needed to clarify the extent to which breeding systems and ecological specialisation can maintain species boundaries. The further study of endemic species could also increase our understanding of evolution and speciation within *Euphrasia*.

In order to further understand the role of breeding systems in maintaining species boundaries and generating diversity within complex populations, it will be important to collect seed and perform a progeny array analysis to estimate outcrossing rates, as in Chapter 3. This is for at least three reasons. First, confirmation of the breeding systems of individuals with differently sized flowers would allow a greater understanding of the patterns of genetic diversity witnessed within the populations. Second, it would allow us to test the hypothesis that outcrossing rates in small-flowered *Euphrasia* are increased in the presence of *Euphrasia* with larger flowers, due to the increased presence of suitable pollinators. Under this hypothesis, outcrossing rates for small-flowered individuals in such populations would be higher than those witnessed in populations of purely small-flowered *Euphrasia*. Finally, it could also increase our understanding of the level of hybridisation between species. If F1 hybrids are found to be more common in progeny arrays

than in natural populations, it would suggest that there is a degree of selection against hybrid morphs.

To test the importance of ecology on the maintenance of species boundaries, reciprocal transplantations of different species could be performed. For example, seed collected from *E. micrantha* and *E. scottica* could be sown artificially in reciprocal habitats. This could be performed in populations where the two species grow adjacent to one another. An assessment of the sites the following year would show whether the reciprocally sown seed were able to germinate and survive to adulthood in environments other than those in which they are commonly found. For more complex interactions, such as of those between persistent hybrids and their progenitor parents, a different approach may be taken. Small populations in which both progenitor parent and hybrid offspring are found could be purged (over several years to account for variation found within the seed bank) of their *Euphrasia* by removing plants from the population, preferably before they set seed. Seed collected from adults found at this site may then be re-sown at random across the area in which the *Euphrasia* were found. Patterns of growth should then be recorded over a number of years. Should the different morphs return to the same distribution pattern, it would be suggestive of a selective advantage to changing environments across the site.

From the perspective of hybrid speciation, it could also be interesting to explore the genetic structuring of currently described endemic taxa in north Scotland. It may be possible to establish whether single or multiple hybridisation events led to their establishment. As chloroplast haplotypes are largely associated with progenitor groups, it may also be possible to identify the 'mother' progenitor species. Assessment of chloroplast haplotypes may also indicate whether hybridisation in reciprocal directions can lead to the formation of a hybrid species. The location of sites with high levels of diversity could then be used to determine the area in which the original hybrid population was formed.

Detailed studies of homoploid speciation have been performed mainly on "good" species of ancient hybrid origin (e.g. *Heliantus*, Rieseberg et al. 2003). The study of diversification and speciation in British tetraploid *Euphrasia* would provide a new perspective. This is because the

genus contains not only young endemics of presumed hybrid origin, but also very young hybrid populations that may have the adaptive potential to become new endemics in the future. Together, the dynamic processes creating novel hybrids with unique adaptive potential are available for investigation. Greater insight into the evolutionary processes involved in the diversification of *Euphrasia* is, therefore, not only useful in the context of creating a process-based conservation plan, but also presents an interesting opportunity for the study of homoploid hybrid speciation.

6.6. Process-based conservation of tetraploid British *Euphrasia*

In order to produce an effective process-based conservation plan for the preservation of the diversification in taxonomically complex groups, it is important to consider a number of factors. An understanding of how diversity is generated and established must be developed and it must be shown that these processes are currently active (Ennos et al. 2012). For any conservation plan to be implemented, important evolutionary processes generating diversity must be shown to be spatially localised so that management may be targeted effectively (Ennos et al. 2012).

The current conservation plan for British *Euphrasia* emphasises the protection of common species, such as *E. micrantha, E. scottica, E. nemorosa, E. arctica* and *E. frigida*, which represent progenitor groups for hybridisation. Another objective of the action plan is to survey "hotspots" of *Euphrasia* diversity, such as those with a western and northern oceanic distribution, in order to clarify their protection status. In this context, I would recommend conservation designation of complex sites, such as Betty Hill, and the development of habitat management plans to ensure the continued evolution of novel *Euphrasia* diversity within them. Where this has not already been done, it may be worth considering notifying these sites as SSSIs. These sites represent hotspots for the creation of novel diversity in *Euphrasia,* occurring against a background of intergrading ecological diversity. In these sites, it will not be necessary to assign all individuals to discrete and unambiguous taxa. Importance should be given to maintaining populations of progenitor taxa, either in the site itself or in the surrounding area, and maintaining the diversity of morphological entities within these complex sites (Ennos et al. 2012). Furthermore, it is important to increase our understanding of currently identified endemic species and their habitats

in order to gain a greater understanding of (i) the overall process of diversification in *Euphrasia* and (ii) the role played by said endemic species in this process.

The concept behind this conservation action plan can be paralleled in other actively diversifying taxonomically complex groups such as *Sorbus*, *Epipactis* and *Dactylorhiza* (Ennos et al. 2012). A detailed conservation action plan has been proposed for *Sorbus*. Some of the diversity in this genus is initially generated via hybridisation events for between the sexual *S. aucuparia* and different apomicts, a process that is initiated by the crossing between *S. aucuparia* and *S. rupicola* (Robertson et al. 2004b, a). *S. aucuparia* must be within pollination distance of apomictic taxa for this hybridisation to take place. Conservation action therefore aims to preserve *S. aucuparia* in sites where active diversification is thought to occur (Ennos et al. 2012). In *Epipactis*, recurrent breeding system transitions from cross-pollination to self-pollination have been important in generating novel diversity in the genus (Hollingsworth et al. 2006). It has been recommended that hotspots of diversity be identified for the targeting of management plans that encourage evolutionary processes rather than the end products of evolution (Ennos et al. 2012). Allopolyploid hybridisation, hybridisation between diploid and tetraploid progenitor lineages, and subsequent back-crossing, combined with ecological divergence, are thought to be the main mechanisms creating an array of diversity within European *Dactylorhiza* (Shipunov et al. 2005; Pillon et al. 2007; Nordstrom & Hedren 2009a; Paun et al. 2010; Paun et al. 2011; De Hert et al. 2012). It has been proposed that conservation efforts should focus on preserving active diversification of *Dactylorhiza* in certain habitats, such as base-rich fens and important geographic areas, such as Greece and northern Sweden, as opposed to conserving individual taxa (Nordstrom & Hedren 2009b).

Whilst these taxonomically complex groups may have unifying similarities with regard to the generation of diversity, such as breeding system transitions, ecological divergence and hybridisation, the role each aspect plays is different in each group. Future studies of these taxonomically complex groups may highlight processes that are key to the generation of diversity. They will also allow the study of speciation in action, providing a different perspective in a field that is dominated by studies of "good" species in groups that are no longer seen to be

actively speciating. It will also provide the detailed information that is necessary to develop process based conservation strategies, which should help these groups to diversify in the wake of climate change.

6.7. Summary

High levels of inbreeding combined with ecological specialisation have likely played an important role in maintaining species distinctions among tetraploid *Euphrasia* despite their extremely porous genomes. Rare hybridisation events produce fertile hybrids. These hybrids are only likely to survive and propagate, however, if they are able to grow in a habitat that is different from that of both of the progenitor species. In this way, new hybrid entities with novel multilocus genotypes that are adapted to different ecological niches represent groups of *Euphrasia* with adaptive potential. These entities may be transient, but they could go on to form new endemic taxa. Populations of *Euphrasia* that display high levels of hybridisation and occupy a range of ecological habitats must be protected under conservation legislation in order to encourage the generation of these novel adaptive entities. Further study will increase our understanding of the diversification process in *Euphrasia* and may also broaden our appreciation of the role which homoploid hybridisation plays in speciation.

References

Aarssen LW (2000) Why are most selfers annuals? A new hypothesis for the fitness benefit of selfing. *Oikos*, **89**, 606-612.

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

Allard RW (1990) Future direction in population genetics, evolution, and breeding In: *Plant population genetics, breeding, and genetic resources* (eds. Brown AHD, Clegg MT, Kahler AL, Weir BS), pp. 1-19. Sinauer, Sunderland, MA.

Allard RW, Jain SK, Workman PL (1968) Genetics of inbreeding populations. *Advances in Genetics Incorporating Molecular Genetic Medicine*, **14**, 55-&.

Anderson E (1948) Hybridization of the habitat. *Evolution*, **2**, 1-9.

Anon. (1995a) Royal Horticultural Society Colour Chart.

Arnold ML (1997) Natural hybridization and evolution. In: *Oxford Series in Ecology and Evolution*. Oxford University press, New York.

Atlan A, Couvet D (1993) A model simulating the dynamics of plant mitochondrial genomes. *Genetics*, **135**, 213-222.

Bakker EG, Toomajian C, Kreitman M, Bergelson J (2006) A genome-wide survey of R gene polymorphisms in *Arabidopsis*. *Plant Cell*, **18**, 1803-1818.

Bandopadhyay R, Sharma S, Rustgi S, Singh R, Kumar A, Balyan HS, Gupta PK (2004) DNA polymorphism among 18 species of *Triticum-Aegilops* complex using wheat EST-SSRs. *Plant Sci.*, **166**, 349-356.

Barrett SCH (1998) The evolution of mating strategies in flowering plants. *Trends in Plant Science*, **3**, 335-341.

Barton NH (2000) Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **355**, 1553-1562.

Barton NH (2001) The role of hybridization in evolution. *Molecular Ecology*, **10**, 551-568.

Bateman RM, Denholm I (1983) A reappraisal of the British and Irish dactylorchids 1. The tetraploid marsh-orchids. *Watsonia*, **14**, 347-376.

Bateman RM, Denholm I (1985) A reappraisal of the British and Irish dactylorchids 2. The diploid marsh orchids. *Watsonia*, **15**, 321-356.

Bateman RM, Denholm I (1989) A reappraisal of the British and Irish dactylorchids 3. The spotted orchids. *Watsonia*, **17**, 319-350.

Bobear JB (1969) An analysis of populations of Irish *Euphrasia* D. *Watsonia*, **7**, 68-91.

Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, Fitz J, Weigel D (2010) Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. *Plos Genetics*, **6**.

Bonnin I, Huguet T, Gherardi M, Prosperi JM, Olivieri I (1996) High level of polymorphism and spatial structure in a selfing plant species, *Medicago truncatula* (Leguminosae), shown using RAPD markers. *American Journal of Botany*, **83**, 843-855.

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

Brochmann C, Gabrielsen TM, Nordal I, Landvik JY, Elven R (2003) Glacial survival or tabula rasa? The history of North Atlantic biota revisited. *Taxon*, **52**, 417-450.

Bruvo R, Michiels NK, D'Souza TG, Schulenburg H (2004) A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology*, **13**, 2101-2106.

Buerkle CA, Morris RJ, Asmussen MA, Rieseberg LH (2000) The likelihood of homoploid hybrid speciation. *Heredity*, **84**, 441-451.

Buerkle CA, Wolf DE, Rieseberg LH (2003) The origin and extinction of species through hybridization. In: *Population viability in plants: Conservation, management, and modeling of rare plants*, pp. 117-141. Springer-Verlag New York Inc., 175 Fifth Avenue, New York, NY, 10010-7858, USA; Springer-Verlag GmbH & Co. KG, Heidelberger Platz 3, D-14197, Berlin, Germany.

Byers DL, Waller DM (1999) Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. *Annual Review of Ecology and Systematics*, **30**, 479-513.

Caullet CML, Pernin F, Poncet C, Le Corre V (2011) Development of microsatellite markers in *Capsella rubella* and *Capsella bursa-pastoris* (Brassicaceae). *American Journal of Botany*, **98**, E176-E179.

Charlesworth B, Charlesworth D, Morgan MT (1990) Genetic loads and estimates of mutation rates in highly inbred plant populations. *Nature*, **347**, 380-382.

Charlesworth B, Morgan MT, Charlesworth D (1991) Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. *Genetical Research*, **57**, 177-194.

Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics*, **134**, 1289-1303.

Charlesworth D (1995) Hybrid speciation - evolution under the microscope. *Current Biology*, **5**, 835-836.

Charlesworth D (2003) Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **358**, 1051-1070.

Charlesworth D, Charlesworth B (1990) Inbreeding depression with heterozygote advantage and its effect on selection for modifiers changing the outcrossing rate. *Evolution*, **44**, 870-888.

Charlesworth D, Morgan MT, Charlesworth B (1990b) Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. *Evolution*, **44**, 1469-1489.

Cheffings CM, Farrell L (2005) *The vascular plant red data list for Great Britain*. JNCC, , Peterborough.

Cieslarova J, Hanacek P, Fialova E, Hybl M, Smykal P (2011) Estimation of pea (*Pisum sativum* L.) microsatellite mutation rate based on pedigree and single-seed descent analyses. *Journal of Applied Genetics*, **52**, 391-401.

Clark LV, Jasieniuk M (2011) POLYSAT: an R package for polyploid microsatellite analysis. *Molecular Ecology Resources*, **11**, 562-566.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*, **15**, 290-295.

Crow JF (1970) Genetic loads and the cost of natural selection. In: *Mathematical Topics in Population Genetics* (ed. Kojima KI), pp. 128-177. Springer, Berlin.

Crow JF, Kimura M (1970) An introduction to popualtion genetics theory. In: *An Introduction to Population Genetics Theory*, p. 591.

Darwin C (1876) *The effects of cross and self fertilisation in the vegetable kingdom*. Murray, London.

De Hert K, Jacquemyn H, Van Glabeke S, Roldan-Ruiz I, Vandepitte K, Leus L, Honnay O (2012) Reproductive isolation and hybridization in sympatric populations of three *Dactylorhiza* species (Orchidaceae) with different ploidy levels. *Annals of Botany*, **109**, 709-720.

De Silva HN, Hall AJ, Rikkerink E, McNeilage MA, Fraser LG (2005) Estimation of allele frequencies in polyploids under certain patterns of inheritance. *Heredity*, **95**, 327-334.

Decroocq V, Fave MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. *Theoretical and Applied Genetics*, **106**, 912-922.

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

dePamphilis CW, Young ND, Wolfe AD (1997) Evolution of plastid gene rps2 in a lineage of hemiparasitic and holoparasitic plants: Many losses of photosynthesis and complex patterns of rate variation. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 7367-7372.

Dequeiroz K, Donoghue MJ (1988) Phylogenetic systematics and the species problem. *Cladistics-the International Journal of the Willi Hennig Society*, **4**, 317-338.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull. Bot. Soc. Am.*, **19**, 11–15.

Dzwonko Z (2001) Assessment of light and soil conditions in ancient and recent woodlands by Ellenberg indicator values. *Journal of Applied Ecology*, **38**, 942-951.

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

Ellenberg H (1979) Indicator values of vascular plants in central Europe. *Scripta Geobotanica*, **9**, 1-122.

Ellenberg H, Weber HE, Dull F, Wirth V, Werner W, Paulissen D (1991) Zeigerwerte von Pflanzen in Mitteleuropa. *Scripta Geobotanica*, **18**, 1-248.

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. *Heredity*, **99**, 125-132.

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

Ennos RA, French GC, Hollingsworth PM (2005) Conserving taxonomic complexity. *Trends in Ecology & Evolution*, **20**, 164-168.

Ennos RA, Sinclair WT, Hu XS, Langdon A (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. *Molecular Systematics and Plant Evolution*, **57**, 1-19.

Ennos RA, Whitlock R, Fay MF, Jones B, Neaves LE, Payne R, Taylor I, De Vere N, Hollingsworth PM (2012) Process-Based Species Action Plans: an approach to conserve contemporary evolutionary processes that sustain diversity in taxonomically complex groups. *Botanical Journal of the Linnean Society*, **168**, 194-203.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, **1**, 47-50.

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

French GC (2004) Conservation genetics of the critical plant genus *Euphrasia* L. in Britain. University of Edinburgh.

French GC, Ennos RA, Silverside AJ, Hollingsworth PM (2005) The relationship between flower size, inbreeding coefficient and inferred selfing rate in British *Euphrasia* species. *Heredity*, **94**, 44-51.

French GC, Hollingsworth PM, Ennos RA (2003) Isolation of polymorphic microsatellite markers for British *Euphrasia* L. *Molecular Ecology Notes*, **3**, 626-628.

French GC, Hollingsworth PM, Silverside AJ, Ennos RA (2008) Genetics, taxonomy and the conservation of British *Euphrasia*. *Conservation Genetics*, **9**, 1547-1562.

Gomez JM (2002) Self-pollination in *Euphrasia willkommii* Freyn (Scrophulariaceae), an endemic species from the alpine of the Sierra Nevada (Spain). *Plant Systematics and Evolution*, **232**, 63-71.

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

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

Green JM, Barker JHA, Marshall EJP, Froud-Williams RJ, Peters NCB, Arnold GM, Dawson K, Karp A (2001) Microsatellite analysis of the inbreeding grass weed Barren Brome (*Anisantha sterilis*) reveals genetic diversity at the within- and between-farm scales. *Molecular Ecology*, **10**, 1035-1045.

Greig D, Louis EJ, Borts RH, Travisano M (2002) Hybrid speciation in experimental populations of yeast. *Science*, **298**, 1773-1775.

Gross BL, Kane NC, Lexer C, Ludwig F, Rosenthal DM, Donovan LA, Rieseberg LH (2004) Reconstructing the origin of *Helianthus deserticola*: Survival and selection on the desert floor. *American Naturalist*, **164**, 145-156.

Gross BL, Rieseberg LH (2005) The ecological genetics of homoploid hybrid speciation. *J. Hered.*, **96**, 241-252.

Gussarova G, Popp M, Vitek E, Brochmann C (2008) Molecular phylogeny and biogeography of the bipolar *Euphrasia* (Orobanchaceae): Recent radiations in an old genus. *Molecular Phylogenetics and Evolution*, **48**, 444-460.

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica* **4**, 9.

Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **351**, 1291-1298.

He F, Kang D, Ren Y, Qu LJ, Zhen Y, Gu H (2007) Genetic diversity of the natural populations of *Arabidopsis thaliana* in China. *Heredity*, **99**, 423-431.

HMSO (Her Majesty's Stationery Office). (1995) *UK Biodiversity Steering Group*.

Hill MO, Preston CD, Roy DB (2004) *PLANTATT. Attributes of British and Irish Plants: Status, Size, Life History, Geography and Habitats*. Centre for Ecology & Hydrology.

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

Hollingsworth PM, Squirrell J, Hollingsworth ML, Richards AJ, Bateman RM (2006) Taxonomic complexity, conservation and recurrent origins of self-pollination in *Epipactis* (Orchidaceae). In: *Current taxonomic research on the British and European Flora.* (eds. Bailey JP, Ellis RG), pp. 27–44. BSBI, London.

Holsinger KE (1992) Ecological models of plant mating systems and the evolutionary stability of mixed mating systems. In: *Ecology and Evolution of Plant Reproduction*, pp. 169-191.

Holsinger KE (2000) Reproductive systems and evolution in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 7037-7042.

Husband BC, Schemske DW (1996) Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution*, **50**, 54-70.

Ingrouille MJ, Stace CA (1986) The *Limonium binervosum* aggregate (Plumbaginaceae) in the Brithish Isles. *Botanical Journal of the Linnean Society*, **92**, 177-217.

Jarne P (1995a) Mating system, bottlenecks and genetic polymorphism in hermaphroditic animals. *Genetical Research*, **65**, 193-207.

Jarne P (1995b) Mating system, bottlenecks and genetic polymorphism in hermaphroditic animals. *Genetical Research*, **65**, 193-207.

Jarne P, David P (2008) Quantifying inbreeding in natural populations of hermaphroditic organisms. *Heredity*, **100**, 431-439.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution*, **11**, 424-429.

Johansen-Morris AD, Latta RG (2006) Fitness consequences of hybridization between ecotypes of *Avena barbata*: Hybrid breakdown, hybrid vigor, and transgressive segregation. *Evolution*, **60**, 1585-1595.

Johansen-Morris AD, Latta RG (2008) Genotype by environment interactions for fitness in hybrid genotypes of *Avena barbata*. *Evolution*, **62**, 573-585.

Jorgensen 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.

Kantety RV, La Rota M, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology*, **48**, 501-510.

Karlsson T (1984) Early-flowering taxa of *Euphrasia* (Scrophylariaceae) on Gotland, Sweden. *Nordic Journal of Botany*, **4**, 303-326.

Kim KS, Ratcliffe ST, French BW, Liu L, Sappington TW (2008) Utility of EST-Derived SSRs as population genetics markers in a beetle. *Journal of Heredity*, **99**, 112-124.

Knuth P (1909) *Handbook of Flower Pollination 3*. Clarendon Press, Oxford.

Lammi A, Siikamaki P, Salonen V (1999) The role of local adaptation in the relationship between an endangered root hemiparasite *Euphrasia rostkoviana*, and its host, *Agrostis capillaris*. *Ecography*, **22**, 145-152.

Lande R, Schemske DW (1985) The evolution of self-fertilization and inbreeding depression in plants .1. Genetic Models. *Evolution*, **39**, 24-40.

Latta RG (2008) Conservation genetics as applied evolution: from genetic pattern to evolutionary process. *Evolutionary Applications*, **1**, 84-94.

Latta RG, Gardner KM, Johansen-Morris AD (2007) Hybridization, recombination, and the genetic basis of fitness variation across environments in *Avena barbata*. *Genetica*, **129**, 167-177.

Le Corre V (2005) Variation at two flowering time genes within and among populations of *Arabidopsis thaliana*: comparison with markers and traits. *Molecular Ecology*, **14**, 4181-4192.

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

Lexer C, Welch ME, Durphy JL, Rieseberg LH (2003a) Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: Implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. *Molecular Ecology*, **12**, 1225-1235.

Lexer C, Welch ME, Raymond O, Rieseberg LH (2003b) The origin of ecological divergence in *Helianthus paradoxus* (Asteraceae): Selection on transgressive characters in a novel hybrid habitat. *Evolution*, **57**, 1989-2000.

Liao W-J, Zhu B-R, Zeng Y-F, Zhang D-Y (2008) TETRA: an improved program for population genetic analysis of allotetraploid microsatellite data. *Molecular Ecology Resources*, **8**, 1260-1262.

Liebst B (2008) Do they really hybridize? A field study in artificially established mixed populations of *Euphrasia minima* and *E. salisburgensis* (Orobanchaceae) in the Swiss Alps. *Plant Systematics and Evolution*, **273**, 179-189.

Liebst B, Schneller J (2005) How selfing and intra- and interspecific crossing influence seed set, morphology and ploidy level in *Euphrasia*: An experimental study of species occurring in the Alps of Switzerland. *Plant Systematics and Evolution*, **255**, 193-214.

Liebst B, Schneller J (2008) Seed dormancy and germination behaviour in two *Euphrasia* species (Orobanchaceae) occurring in the Swiss Alps. *Botanical Journal of the Linnean Society*, **156**, 649-656.

Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, **15**, 65-95.

Lowe A, S. H, Ashton P (2004) *Ecological genetics: design, analysis, and application*. Blackwell Publishing, Oxford, UK.

Lundemo S, Falahati-Anbaran M, Stenoien HK (2009) Seed banks cause elevated generation times and effective population sizes of *Arabidopsis thaliana* in northern Europe. *Molecular Ecology*, **18**, 2798-2811.

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions techniques with appropriate. *Trends in Ecology & Evolution*, **20**, 136-142.

Markwith SH, Stewart DJ, Dyer JL (2006) TETRASAT: a program for the population analysis of allotetraploid microsatellite data. *Molecular Ecology Notes*, **6**, 586-589.

Matthies D (1998) Influence of the host on growth and biomass allocation in the two facultative root hemiparasites *Odontites vulgaris* and *Euphrasia minima*. *Flora*, **193**, 187-193.

Mayr E (1940) Speciation phenomena in birds. *American Naturalist*, **74**, 249-278.

McKitrick MC, Zink RM (1988) Species concepts in ornithology. *Condor*, **90**, 1-14.

Milligan BG, Leebensmack J, Strand AE (1994) Conservation genetics - beyond the maintenance of marker diversity. *Molecular Ecology*, **3**, 423-435.

Mishler BD, Donoghue MJ (1982) Species concepts - a case for pluralism. *Systematic Zoology*, **31**, 491-503.

Molau U (1993) Relationships between flowering phenology and life history strategies in tundra plants. *Arctic and Alpine Research*, **25**, 391-402.

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics*, **30**, 194-200.

Morin PA, Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, **19**, 208-216.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238-254.

Nordborg M (2000) Linkage disequilibrium, gene trees and selfing: An ancestral recombination graph with partial self-fertilization. *Genetics*, **154**, 923-929.

Nordborg M, Donnelly P (1997) The coalescent process with selfing. *Genetics*, **146**, 1185-1195.

Nordstrom S, Hedren M (2009a) Evolution, phylogeography and taxonomy of allopolyploid *Dactylorhiza* (Orchidaceae) and its implications for conservation. *Nordic Journal of Botany*, **27**, 548-556.

Nordstrom S, Hedren M (2009b) Genetic diversity and differentiation of allopolyploid *Dactylorhiza* (Orchidaceae) with particular focus on the *Dactylorhiza majalis* ssp *traunsteineri/lapponica* complex. *Biological Journal of the Linnean Society*, **97**, 52-67.

Oetting WS, Lee HK, Flanders DJ, Wiesner GL, Sellers TA, King RA (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics*, **30**, 450-458.

Oliver C (1999) Conservation genetics of *Euphrasia*. *MSc thesis. University of Edinburgh and Royal Botanic Garden Edinburgh*.

Pashley CH, Ellis JR, McCauley DE, Burke JM (2006) EST databases as a source for molecular markers: Lessons from Helianthus. *Journal of Heredity*, **97**, 381-388.

Paun O, Bateman RM, Fay MF, Hedren M, Civeyrel L, Chase MW (2010) Stable Epigenetic Effects Impact Adaptation in Allopolyploid Orchids (*Dactylorhiza*: Orchidaceae). *Molecular Biology and Evolution*, **27**, 2465-2473.

Paun O, Bateman RM, Fay MF, Luna JA, Moat J, Hedren M, Chase MW (2011) Altered gene expression and ecological divergence in sibling allopolyploids of *Dactylorhiza* (Orchidaceae). *Bmc Evolutionary Biology*, **11**.

Pillon Y, Fay MF, Hedren M, Bateman RM, Devey DS, Shipunov AB, van der Bank M, Chase MW (2007) Evolution and temporal diversification of western European polyploid species complexes in *Dactylorhiza* (Orchidaceae). *Taxon*, **56**, 1185-1208.

Pollak E (1987) On the theory of partially inbreeding finite populations .1. Partial selfing. *Genetics*, **117**, 353-360.

Pons O, Petit RJ (1995) Estimation, variance and optimal sampling of gene diversity .1. Haploid Locus. *Theoretical and Applied Genetics*, **90**, 462-470.

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

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

Pritchard NM (1959) *Gentianella* in Britain I. *G. amarella*, *G. anglica* and *G. uliginosa*. *Watsonia*, **4**, 169-193.

Pritchard NM (1960a) *Gentianella* in Britain II. *Gentianella septrentrionalis* (Druce) E. F. Warb. *Watsonia*, **4**, 218-237.

Pritchard NM (1960b) *Gentianella* in Britain III: *Gentianella germanica* (Willd.) Böner. *Watsonia*, **4**, 290-303.

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

Pugsley HW (1929) New British Species of *Euphrasia*. *Journal of Botany*, **67**, 224-225.

Pugsley HW (1930) A revision of the British *Euphrasiae*. *Journal of the Linnean Scoiety of London Botany*, **48**, 467-544.

Pugsley HW (1933) The *Euphrasias* of Iceland and the Faeroes. *Journal of the Linnean Society of London, Botany*, **48**, 467-544.

Pugsley HW (1936) Notes on the British *Euphrasias* - IV. *Journal of Botany*, **74**, 71-75.

Pugsley HW (1940) Notes of British *Euphrasias* - VI. *Journal of Botany*, **78**.

Pugsley HW (1945) The Eyebrights of Rhum. *The Naturalist*, **April/June**, 41-44.

Ramakrishnan AP, Meyer SE, Fairbanks DJ, Coleman CE (2006) Ecological significance of microsatellite variation in western North American populations of *Bromus tectorum*. *Plant Species Biology*, **21**, 61-73.

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

Rich TCG (1991) *Crucifers of Great Britain and Ireland*. Botanical Society of the British Isles, London.

Rich TCG, McDonnell EJ, Lledo MD (2008) Conservation of Britain's biodiversity: the case of *Hieracium cyathis* (Asteraceae) and its relation to other apomictic taxa. *Botanical Journal of the Linnean Society*, **156**, 669-680.

Richards AJ (1997) *Epipactis*. In: *Plant Crib* (eds. Rich TCG, Jeremy AC), pp. 378-379. Botanical Society of the British Isles, London.

Rieseberg LH (1991) Homoploid reticulate evolution in *Helianthus* (Asteraceae): evidence from riobsomal genes. *American Journal of Botany*, **78**, 1218-1237.

Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology and Systematics*, **28**, 359-389.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution*, **16**, 351-358.

Rieseberg LH, Archer MA, Wayne RK (1999) Transgressive segregation, adaptation and speciation. *Heredity*, **83**, 363-372.

Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science*, **301**, 1211-1216.

Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In: *Hybrid Zones and the Evolutionary Process*, pp. 70-114.

Rieseberg LH, Willis JH (2007) Plant speciation. *Science*, **317**, 910-914.

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

Ritland K (1990b) A series of fortran computer programs for estimating plant mating systems. *Journal of Heredity*, **81**, 235-237.

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

Ritland K, Jain S (1981) A model for the estimation of outcrossing rate and gene frequencies using N independent loci. *Heredity*, **47**, 35-52.

Robertson A, Newton AC, Ennos RA (2004a) Breeding systems and continuing evolution in the endemic *Sorbus* taxa on Arran. *Heredity*, **93**, 487-495.

Robertson A, Newton AC, Ennos RA (2004b) Multiple hybrid origins, genetic diversity and population genetic structure of two endemic *Sorbus* taxa on the Isle of Arran, Scotland. *Molecular Ecology*, **13**, 123-134.

Ronfort J, Bataillon T, Santoni S, Delalande M, David JL, Prosperi J-M (2006) Microsatellite diversity and broad scale geographic structure in a model legume: building a set of nested core collection for studying naturally occurring variation in *Medicago truncatula*. *Bmc Plant Biology*, **6**.

Rosen DE (1978) Vicariant patterns and historical explanation in biogeography. *Systematic Zoology*, **27**, 159-188.

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

Saha MC, Mian MAR, Eujayl I, Zwonitzer JC, Wang LJ, May GD (2004) Tall fescue EST-SSR markers with transferability across several grass species. *Theoretical and Applied Genetics*, **109**, 783-791.

Schemske DW, Lande R (1985) The evolution of self-fertilization and inbreeding depression in Plants .2. Empirical observations. *Evolution*, **39**, 41-52.

Schoen DJ, Brown AHD (1991) Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 4494-4497.

Schoen DJ, Johnston MO, Lheureux AM, Marsolais JV (1997) Evolutionary history of the mating system in *Amsinckia* (Boraginaceae). *Evolution*, **51**, 1090-1099.

Schwarzbach AE, Donovan LA, Rieseberg LH (2001) Transgressive character expression in a hybrid sunflower species. *American Journal of Botany*, **88**, 270-277.

Seel WE, Press MC (1993) Influence of the host on 3 sub-Arctic annual facultative root hemiparasites .1. Growth, mineral accumulation and above groud dry matter partitioning. *New Phytologist*, **125**, 131-138.

Seel WE, Press MC (1994) Influence of the host on 3 sub-Arctic annual facultative root hemiparasites .2. Gas exchange characteristics and resource use efficiency. *New Phytologist*, **127**, 37-44.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu WS, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142-166.

Shipunov AB, Fay MF, Chase MW (2005) Evolution of *Dactylorhiza baltica* (Orchidaceae) in European Russia: evidence from molecular markers and morphology. *Botanical Journal of the Linnean Society*, **147**, 257-274.

Silverside AJ (1990) A guide to eyebrights (*Euphrasia*) - I. Introduction: *E. nemorosa, E. confusa* and *E. arctica*. *Wild Flower Magazine* 30-34.

Silverside AJ (1991a) A guide to eyebrights (*Euphrasia*) - III. Species of restricted habitats. *Wild Flower Magazine*, 31-35.

Silverside AJ (1991b) A guide to eyebrights (*Euphrasia*) - IV. On northern coasts and western hills. *Wild Flower Magazine*, 29-33.

Silverside AJ (1991c) A guide to eyebrights (*Euphrasia*) - V. Local endemics and concluding remarks. *Wild Flower Magazine*, 32-36.

Silverside AJ (1998) *Euphrasia*. In: *Plant Crib* (eds. Rich TCG, Jeremy AC), pp. 269-272. Botanical Society of the British Isles, London.

Simpson GG (1961) Principles of animal taxonomy (Columbia Biological Series No. XX.). In: *Principles of Animal Taxonomy*.

Siol M, Prosperi JM, Bonnin I, Ronfort J (2008) How multilocus genotypic pattern helps to understand the history of selfing populations: a case study in *Medicago truncatula*. *Heredity*, **100**, 517-525.

Smith JM, Haigh J (1974) Hitch-hiking effect of a favorable gene. *Genetical Research*, **23**, 23-35.

Soltis DE, Soltis PS (1993) Molecular-data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences*, **12**, 243-273.

Squirrell J, Hollingsworth PM, Bateman RM, Tebbitt MC, Hollingsworth ML (2002) Taxonomic complexity and breeding system transitions: conservation genetics of the *Epipactis leptochila* complex (Orchidaceae). *Molecular Ecology*, **11**, 1957-1964.

Stace C (1991) New Flora of the British Isles. *Stace, C. New Flora of the British Isles. Xxx+1226p. Cambridge University Press: Cambridge, England, Uk; New York, New York, USA. Illus. Maps*, XXX+1226P.

Stace CA (2010) *New Flora of the British Isles*, 3 edn. Cambridge University Press. Cambridge.

Stenoien HK, Fenster CB, Tonteri A, Savolainen O (2005) Genetic variability in natural populations of *Arabidopsis thaliana* in northern Europe. *Molecular Ecology*, **14**, 137-148.

Stewart OM (1997) *Calamagrostis*. In: *Plant Crib.* (eds. Rich TCG, Jeremy AC), pp. 363-364. Botanical Society of the British Isles, London.

Svensson BM, Seel WE, Nilsson CH, Carlsson BA (2001) Roles played by riming of seedling development and host identity in determining fitness of an annual, subarctic, hemiparasitic plant. *Arctic Antarctic and Alpine Research*, **33**, 299-305.

Sweigart AL, Willis JH (2003) Patterns of nucleotide diversity in two species of *Mimulus* are affected by mating system and asymmetric introgression. *Evolution*, **57**, 2490-2506.

Teacher AGF, Griffiths DJ (2011) HapStar: Automated haplotype network layout and visualisation. *Molecular Ecology Resources*, **11(1)**, 151-153.

Templeton AR (1981) Mechanisms of speciation - a population genetic approach. *Annual Review of Ecology and Systematics*, **12**, 23-48.

Templeton AR (1989) The meaning of species and speciation - A genetic perspective. In: *Otte, D. and J. a. Endler*, pp. 3-27.

Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.*, **106**, 411-422.

Thuillet AC, Bataillon T, Poirier S, Santoni S, David JL (2005) Estimation of long-term effective population sizes through the history of durum wheat using microsatellite data. *Genetics*, **169**, 1589-1599.

Thuillet AC, Bru D, David J, Roumet P, Santomi S, Sourdille P, Bataillon T (2002) Direct estimation of mutation rate for 10 microsatellite loci in durum wheat, *Triticum turgidum* (L.) Thell. ssp *durum* desf. *Molecular Biology and Evolution*, **19**, 122-125.

Udupa SM, Baum M (2001) High mutation rate and mutational bias at (TAA)(n) microsatellite loci in chickpea (*Cicer arietinum* L.). *Molecular Genetics and Genomics*, **265**, 1097-1103.

Van Puyvelde K, Van Geert A, Triest L ATETRA, a new software program to analyse tetraploid microsatellite data: comparison with TETRA and TETRASAT. *Mol. Ecol. Resour.*, **10**, 331-334.

Varshney RK, Sigmund R, Borner A, Korzun V, Stein N, Sorrells ME, Langridge P, Graner A (2005) Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. *Plant Sci.*, **168**, 195-202.

Vigouroux Y, Matsuoka Y, Doebley J (2003) Directional evolution for microsatellite size in maize. *Molecular Biology and Evolution*, **20**, 1480-1483.

Vitalis R, Glemin S, Olivieri I (2004) When genes go to sleep: The population genetic consequences of seed dormancy and monocarpic perenniality. *American Naturalist*, **163**, 295-311.

Vitek E (1986) Evolution in Alpine populations of *Euphrasia* (Scrophulariaceae) - the tetraploid *E. minima*. *Plant Systematics and Evolution*, **151**, 241-269.

Vitek E (1998) Are the taxonomic concepts of agamospermous genera useful for autogamous groups - A critical discussion using the example of *Euphrasia* (Scrophulariaceae). *Folia Geobotanica*, **33**, 349-352.

Vitek E, Kiehn M (1998) *Euphrasia salisburgensis* var. *stiriaca* (Scrophulariaceae): A relic taxon from the eastern Alps. *Annalen des Naturhistorischen Museums in Wien Serie B Botanik und Zoologie*, **100**, 715-723.

Walters SM (1949) *Alchemilla vulgaris* L. agg. *Watsonia*, **1**, 6-18.

Wang K, Wu M, Chiang T, Chou C (2009) Isolation and characterization of polumorphic microsatellite DNA markers for *Euphrasia nankotaizanensis* (Orobanchaceae) and cross amplification in another *Euphrasia* L. *Conservation Genetics*, **10(4)**, 1163-1165.

Warburg EF (1962) *Sorbus* L. . In: *Flora of the British Isles* (eds. Clapham AR, Tutin TG, Warburg EF), pp. 423-237. Cambridge University Press, Cambridge.

Welch ME, Rieseberg LH (2002) Habitat divergence between a homoploid hybrid sunflower species, *Helianthus paradoxus* (Asteraceae), and its progenitors. *American Journal of Botany*, **89**, 472-478.

Wendel JF, Stewart JM, Rettig JH (1991) Molecular evidence for homoploid reticulate evolution among Australian species of *Gossypium*. *Evolution*, **45**, 694-711.

Wigginton MJ (1999) *British Red Data Books 1: Vascular Plants*. Joint Nature Conservation Committee, Peterborough.

Wiley EO (1978) Evolutionary species concept reconsidered. *Systematic Zoology*, **27**, 17-26.

Wiley EO (1981) Phylogenetics: The theory and practice of phylogenetic systematics. John Wiley and Sons, Inc., New York, N.Y., USA.

Woodhead M, Russell J, Squirrell J, Hollingsworth M, Cardle L, Ramsay L, Gibby M, Powell W (2003) Development of EST-SSRs from the alpine lady-fern, *Athyrium distentifolium*. *Molecular Ecology Notes*, **3**, 287-290.

Wu CI (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851-865

Wu MJ, Huang SF, Huang TC, Lee PF, Lin TP (2005) Evolution of the *Euphrasia transmorrisonensis* complex (Orobanchaceae) in alpine areas of Taiwan. *Journal of Biogeography*, **32**, 1921-1929.

Yeo PF (1956) Hybridisation between diploid and tetraploid species of *Euphrasia*. *Watsonia*, **3**, 253-269.

Yeo PF (1961) Germination, seedlings, and the formation of haustoria in *Euphrasia*. *Watsonia*, **5**, 11-22.

Yeo PF (1962) A study of variation in *Euphrasia* by means of outdoor cultivation. *Watsonia*, **5**, 224-235.

Yeo PF (1966) The breeding relationships of some European *Euphrasiae*. *Watsonia*, **6**, 216-245.

Yeo PF (1968) Evolutionary significance of speciation of *Euphrasia* in Europe. *Evolution*, **22**, 736-&.

Yeo PF (1978) Taxonomic revision of *Euphrasia* in Europe. *Botanical Journal of the Linnean Society*, **77**, 223-334.

Appendix I: Site dossier

Berwickshire

BEm_a – Twinlaw Ford

Collection date: 13.07.09
Waymark: 14
Elevation: 343m
Grid reference: NT6482255960

Site notes: *E. micrantha* amongst the heather on the other side of the valley from the approaching path.

Photographs:

BEm_b – Henon Scar

Collection date: 15.07.09
Waymark: 20
Elevation: 395m
Grid reference: NT6087258733

Site notes: Along the roadside on both sides.

Photographs: n/a

BEm_c – Lower Stot Cleugh

Collection date: 15.07.09
Waymark: 21
Elevation: 326m
Grid reference: NT6376258125

Site notes: Along the track on a steep slope above a few trees, among the heather and in a huge stand where nothing much else is growing.

Photographs: n/a

BEs_a – Brock's Cleugh

Collection date: 13.07.09
Waymark: 15
Elevation: 372m
Grid reference: NT6160558591

Site notes: Just off the road, *E. scottica* just by bridge/higher bit of road, and also on the other side of a bush nearer the stream

Photographs:

BEs_b – Upper Stot Cleugh

Collection date: 13.07.09
Waymark: 19
Elevation: 424m
Grid reference: NT6343359371

Site notes: Right up at the top of the stream, there is an obvious patch where the habitat changes – keep going!

Photographs: n/a

BEs_c – Hell's Cleugh

Collection date: 15.07.09
Waymark: 28
Elevation: 318
Grid reference: NT7404955056

Site notes: On a relatively bare patch of earth which is on a slope (unusual for *E. scottica*), but with water constantly flowing through it.

Photographs: n/a

Perthshire

PEm_a – Meall Odhar Quany

Collection date: 10.08.09
Waymark: 31
Elevation: 437m
Grid reference: NN7771054180

Site notes: Population spread in three patches, (1) on the opposite bank of the stream from the entrance to the site plants 1-5 (first photo); (2) directly opposite the first patch, slightly further up on the other (near) side of the stream, plant 6 from the bottom of the gully, and 7-11from the top of the gully; (3) to the left of the old farm track as you walk up from the road (second photo) plants 12-37.

Photographs:

PEm33a

PEm34a

PEm35a

PEm36a

PEm37a

PEms_a – Coire Screabaid

Collection date: 11.08.09
Waymark: 32
Elevation: 461m
Grid reference: NN7807754186

Site notes: Wet flush passing through heather moorland, *E. scottica* in the flush, *E. micrantha* in the heather out towards the electricity pylon. Plants were sampled from the top to the bottom of the site (going downhill) and there were four patches of heather (*E. micrantha*) with a flush running between. First patch of *E. micrantha* plants 1,2&4; second patch 3, 5-7, 9-24, 31; third patch 8, 25-29; fourth patch 30&31 sampled.

Photographs:

PEms08a

PEms31a

PEms18a

PEms_b – Ben Lawers

Collection date: 12.08.09
Waymark: 34
Elevation: 580m
Grid reference: NN5871641776

Site notes: Site just off the road with the *E. micrantha* growing on the slope of short heather. This site was evenly sampled across the 10x10m area. There is *E. arctica* in the short grassy bits in the verge between both sites. A small way down the hill is a "stable" hybrid population. N.B. the bracts on one of the plants look as though they might be minutely hairy – possible hybrid!?

Photographs:



PEms01b


PEms04b


PEms24b


PEms24b (bracts)


PEms32b

PEs_a – Schiehallion

Collection date: 10.08.09
Waymark: 29
Elevation: 419
Grid reference: NN7511955203

Site notes: Large population extending over 50x50m – also an extra 5 *E. scottica* found from close to the site where it is slightly more acidic – 100m away (NN7536855033)

Photographs: n/a

PEsm_a – Coire Screabaig

Collection date: 11.08.09
Waymark: 32
Elevation: 461m
Grid reference: NN7807754186

Site notes: Wet flush passing through heather moorland, *E. scottica* in the flush, *E. micrantha* in the heather out beyond the electricity pylon. *E. scottica* were sampled from the top to the bottom of the flush (going downhill). PEsm27a was slightly *E. micrantha* like. N.B. these *E. scottica* plants look heterogenous and are probably hybridised.

Photographs:

PEsm13a PEsm20a PEsm31a PEsm27a

PEsm_b – Ben Lawers

Collection date: 12.08.09
Waymark: 34
Elevation: 580m
Grid reference: NN5871641776

Site notes: Site just off the road with *E. scottica* either side of the burn. It's on the other side of the road from the micrantha population. This site was evenly sampled across the 40x10m area. There is *E. arctica* in the short grassy bits in the verge between both sites. This *E. scottica* is not perfect – they may have some arctica in them as the plants are very robust looking. This *E. scottica* population has rather aristate foliage, possible introgression from *E. arctica* (but the name '*E. scottica*' was based on similar plants, i.e. this population resembles the type specimens. A small way down the hill is a "stable" hybrid population.

Photographs:

PEsm31b

PEsm32b

PEsm33b

PEsm34b

Cairngorms

CEm_a – Leids Hill

Collection date: 13.08.09
Waymark: 35
Elevation: 467m
Grid reference: NJ4146726143

Site notes: *E. micrantha* are found growing alongside the road in the slightly more sparse patch of ground before the heather really gets started. Some *E. electa* (*E. micrantha x scottica*) were found nearby.

Photographs:






CEm_b – Sui Wood

Collection date: 13.08.09
Waymark: 36
Elevation: 443m
Grid reference: NJ5470223202

Site notes: This site is in a ditch along the road and is a very long thin site: approx 300m. Samples start from further down the hill (the first couple were infected with mildew) and plants were collected every 10 paces moving up the hill.

Photographs:






CEm_c – Creag Meagaidh

Collection date: 13.08.10
Waymark: 70
Elevation: 536m
Grid reference: NN4704188688

Site notes: Just as you leave the trees there is a small population of very good *E. micrantha* on bare peat (way: 72; elev: 463m; Coord: NN4737188208), 5 of which were collected. Below that on the other side of the wooded area are, some much more hybrid looking (*E. nemorosa*???) *E. micrantha* can be found along the path.

Photographs: n/a

CEs_a – Upper Howbog

Collection date: 14.08.09
Waymark: 42
Elevation: 456m
Grid reference: NJ4047626255

Site notes: This was an unusual site as the site itself was surrounded by very unsuitable *Euphrasia* habitat making it very isolated. There were two main patches, one 20m^2 patch (collection 1-19 & 24) and another 10m stretch about 30m from the first patch (20-23 & 25-32). **After inspecting the photographs Alan actually defines this as a putative hybrid – *E. electa* (*E. scottica x micrantha*)**

Photographs:

CEs_b – Loch an Eilein

Collection date: 14.08.09
Waymark: n/a
Elevation: 216m
Grid reference: NH 9038 0775

Site notes: 20x20m site next to the lake a short way off the path and down the hill where the stream meets the lake (when facing the lake, take a left from the car park down that track)

Photographs:




CEs29b


CEs30b


CEs31b


CEs32b

CEs_c – A9, Lay-by 92

Collection date: 07/08/10
Waymark: 61
Elevation: 415m
Grid reference: NN6482385893

Site notes: This *E. scottica* site is just over the fence if you stop at lay-by 92 on the A9. It's 20m from the road in and around a small flush. Samples were taken from across the whole site. Main site was 5x5m with scottica also growing further down the stream more sporadically. Collection 1-24 main site and 25-32 as you go further down the stream.

Photographs:




CEs_d – Creag Megaidh

Collection date: 13/08/10
Waymark: 70
Elevation: 536m
Grid reference: NN4691588929

Site notes: Just to the left off the Craig Meghidgh path about 200m from after you leave the wooded area of the.

Photographs: n/a

Sutherland

SEm_a

Collection date: 12.08.10
Waymark: 68
Elevation: 124m
Grid reference: NC2558052620

Site notes: Just off the Kinlochbervie turning from Riconich in lay by, along the roadside in bank just further on from the layby.

Photographs: n/a

SEms_a

Collection date: 09/08/2010
Waymark: 64
Elevation: 128
Grid reference: NC2318531426

Site notes: Just in the corner of the Drumbeg turning with easy parking (from Unapool). *E. micrantha* are to be found along the roadside in the Scourie direction. The site was sampled over about 50m. The *E. micrantha* are mixed with an *E. arctica x nemorosa* all around the roadside. Intermediates also collected (Ex01-07)

Photographs: n/a

SEms_b

Collection date: 11.08.10
Waymark: 66
Elevation: 175m
Grid reference: NC3739133160

Site notes: *E. micrantha* was found around the parking space; much smaller population, also not entirely *E. micrantha* like.

Photographs: n/a

SEs_a

Collection date: 12.08.10
Waymark: 67
Elevation: 120m
Grid reference: NC1583630836

Site notes: *E. scottica* along the footpath, where the footpath comes closer to the stream and also by the stream – almost at an oxbow. First 16 collected by the stream and second 16 by the path.

Photographs: n/a

SEsm_a

Collection date: 09/08/2010
Waymark: 64
Elevation: 128
Grid reference: NC2318531426

Site notes: Just in the corner of the Drumbeg turning with easy parking (from Unapool). There is a very large *E. scottica* population with plants in abundance, in the peaty flushes 20 yards from the road. It's a mosaic community with no obvious end to the population. Samples were taken over a 20-30m^2 area. Midges found in abundance.

Photographs: n/a

SEsm_b

Collection date: 11.08.10
Waymark: 66
Elevation: 175m
Coordinates: NC3739133160

Site notes: *E. scottica* found to the left of the parking spot (quarry), facing away from the road. Most *E. scottica* were sampled in a large population 30x30m with a few (6) from approx 50m away at the bottom of a flush further from the quarry.

Photographs: n/a

Loch an Daim Mor Complex Site

Collection date: 10.08.10
Waymark: 65
Elevation: 140m
Grid reference: NC1655742878

Site notes: Population just up the path through the gate. *E. fharidensis x scottica*, *E. fharaidensis* and *E. scottica* population in the damper area to the right of the track and *E. micrantha x scottica x fharaidensis* along the track.

Sampling strategy: A $1m^2$ quadrate was placed within 6 consecutive 5x30m blocks. Within the 10cm grid of the quadrate 5 squares were chosen at random from which the most central *Euphrasia* plant was taken. If there were no *Euphrasia* in that square, we just picked another random square until 5 samples resulted [Bags E0810_101-150; Q1 7m/2.5m *E. fraidensis* type quadrat; Q2 26m/10m classic *E. scottica*; Q3 13m/9m *E. fraidensis* type quadrat; Q4 25m/12m *E. fraidensis* type quadrat; Q5 23.5/25 *E. fraidensis x scottica* quadrat; Q5 5m/16m *E. fraidensis* quadrat]. Second set was of white and purple *E. micrantha x scottica x fharaidensis* and one of each colour was picked every two meters.

Seed collection: 10.08.10 collection of the *E. micrantha x scottica x fharaidensis* mix.

Photographs:

Faraid Head complex site

Collection date: 08.08.10
Waymark: n/a
Elevation: 20 m
Grid reference: NC16554287

Site notes: These sites were chosen to transect both nemorosa (drier patches) and *E. fharidensis/scottica x* in the flushes. This year it was very dry so there was a much lower *E. scottica* presence than expected. Site 1 was nearer the MOD base and transect 2 was closer to the top L hand side of the headland.

Sampling strategy: $1m^2$ quadrats were placed at 5m intervals along a 50m transect. Each quadrat was split into a 20 x 20cm grid and 5 squares were chosen at random from which the most central Euphrasia plant was taken. If there were no *Euphrasia* in that square, we just picked another random square until 5 samples resulted.

Photographs:

Betty Hill complex site

Collection date: 02.08.2006 and 03.08.2006
Waymark: n/a
Elevation: 10m
Grid reference: NC70436276

Site notes: These sites were as there was a very complex mix of different *Euphrasia* morphs. Three transects were sampled from.

Sampling strategy: $1m^2$ quadrats were placed at 5m intervals along a 50m transect. Each quadrat was split into a 20 x 20cm grid and 5 squares were chosen at random from which the most central *Euphrasia* plant was taken. If there were no *Euphrasia* in that square, we just picked another random square until 5 samples resulted.

Photographs:

Appendix II: Associated species

Table AII:1. Associated species lists for all *E. micrantha* populations sampled in Chapter 4. Populations are referred to by their code, as described in Table 4:1 p48.

Associated Species	Population											
	BEm_a	**BEm_b**	**BEm_c**	**PEm_a**	**PEms_a**	**PEms_b**	**CEm_a**	**CEm_b**	**CEm_c**	**SEm_a**	**SEms_a**	**SEms_b**
Agrostis canina											+	
Agrostis capillaris	+	+	+	+	+			+	+	+	+	
Agrostis vinealis				+	+	+	+	+		+		+
Aira caryophyllea										+		
Alchemilla glabra				+								
Anthoxanthum odoratum	+	+	+	+		+	+	+		+	+	+
Blechnum spicant									+			+
Calluna vulgaris	+	+	+	+	+	+	+	+	+	+	+	+
Campanula rotundifolia				+								
Carex binervis						+						
Carex echinata											+	
Carex flacca				+							+	
Carex leporina (ovalis)											+	
Carex nigra											+	
Carex pilulifera				+		+				+		
Carex pulicaris											+	
Cirisium arvense			+									
Cirsium vulgare		+										
Cynosurus cristatus											+	
Cytisus scoparius							+	+				
Dactylorhiza maculata												+
Danthonia decumbens				+						+	+	
Deschampsia cespitosa												
Deschampsia flexuosa	+	+	+		+	+	+	+				+
Empetrum nigrum	+											
Erica cinerea		+	+	+	+	+	+	+	+	+	+	
Erica tetralix										+	+	+
Eriophorum angustifolium											+	
Euphrasia arctica var. borealis				+								
Euphrasia arctica										+		
Euphrasia arctica x nemorosa											+	
Euphrasia scottica											+	
Festuca filiformis					+		+	+				
Festuca ovina	+	+	+	+	+	+	+	+	+	+		
Festuca rubra										+	+	
Festuca vivipara						+				+	+	+
Fragaria vesca			+									
Galium saxatile	+	+										+
Gentianella campestris											+	
Holcus lanatus			+		+			+		+	+	+
Holcus mollis								+				
Hypericum pulchrum											+	
Hypochaeris radicata								+			+	

Table AII:1. ***Continued***

Associated Species	Population											
	BEm_a	**BEm_b**	**BEm_c**	**PEm_a**	**PEms_a**	**PEms_b**	**CEm_a**	**CEm_b**	**CEm_c**	**SEm_a**	**SEms_a**	**SEms_b**
Juncus articulatus											+	
Juncus squarrosus						+					+	+
Leontodon saxatilis											+	
Linum catharticum				+							+	
Lotus corniculatus			+	+			+	+	+	+	+	
Luzula multiflora				+		+				+	+	+
Luzula pilosa					+							
Luzula sylvatica							+					
Lycopodium clavatum								+				+
Molinia caerulea									+		+	
Nardus stricta				+		+		+	+	+	+	+
Pedicularis sylvatica												+
Picea sitchensis (seedlings)								+				
Pinguicula vulgaris											+	
Plantago maritima										+	+	
Plantago media			+									
Plantago lanceolata				+								
Poa annua										+		
Poa humilis											+	
Potentilla erecta ssp. erecta		+	+	+	+	+			+	+	+	
Potentilla erecta ssp. strictissima												+
Prunella vulgaris				+								
Pteridium aquilinum			+	+					+	+	+	
Ranunculus acris											+	
Rhinanthus minor											+	
Rumex acetosella												+
Scorzoneroides autumnalis											+	
Senecio jacobaea												+
Silene flos-cuculi											+	
Sorbus aucuparia									+			
Succisa pratensis									+	+		+
Thymus polytrichus				+								
Trichophorum germanicum						+						
Trifolium repens			+	+							+	
Ulex europaeus								+				
Vaccinium myrtillus					+	+						
Vaccinium vitis-idaea	+				+	+		+				
Veronica officinalis				+				+	+			
Viola palustris												+
Viola riviniana				+						+	+	+

Table AII:2. Associated species lists for all *E. scottica* populations sampled in Chapter 4. Populations are referred to by their code, as described in Table 4:1 p48.

Associated Species	**Population**												
	BEs_a	**BEs_b**	**BEs_c**	**PEs_a**	**PEsm_a**	**PEsm_b**	**CEs_a**	**CEs_b**	**CEs_c**	**CEs_d**	**SEs_a**	**SEsm_a**	**SEsm_b**
Achillea ptarmica					+								
Agrostis canina					+	+	+		+		+	+	+
Alchemilla glabra							+						
Anthoxanthum odoratum	+	+			+		+						
Blechnum spicant										+			
Briza media	+	+	+		+		+						
Calluna vulgaris	+	+	+			+	+	+	+	+	+	+	+
Cardamine pratensis						+							
Carex dioica	+	+	+										
Carex echinata					+	+	+		+		+	+	+
Carex hostiana				+	+		+		+		+	+	
Carex nigra					+	+	+	+				+	
Carex panicea	+			+		+	+	+			+	+	+
Carex pauciflora								+					
Carex pulicaris						+	+		+		+		+
Cirsium palustre						+							
Cynosurus cristatus		+	+										
Dactylorhiza maculata	+									+			+
Danthonia decumbens											+	+	
Deschampsia cespitosa		+	+				+						
Deschampsia flexuosa													+
Drosera anglica												+	
Drosera rotundifolia				+				+			+	+	+
Eleocharis multicaulis												+	
Eleocharis quinqueflora				+					+				
Empetrum nigrum		+											
Epilobium palustre	+				+	+							
Equisetum palustre			+			+	+						
Erica cinerea										+	+		
Erica tetralix	+			+	+	+		+	+	+	+	+	+
Eriophorum angustifolium							+	+	+	+		+	+
Eriophorum latifolium				+									
Eriophorum vaginatum													+
Festuca ovina					+	+					+		
Festuca vivipara						+						+	
Galium palustre					+		+						
Galium uliginosum		+											
Holcus lanatus	+	+	+		+		+						

Table AII:2. ***Continued***

Associated Species	Population												
	BEs_a	**BEs_b**	**BEs_c**	**PEs_a**	**PEsm_a**	**PEsm_b**	**CEs_a**	**CEs_b**	**CEs_c**	**CEs_d**	**SEs_a**	**SEsm_a**	**SEsm_b**
Juniperus communis							+						
Juncus articulatus				+		+	+		+				
Juncus bulbosus					+				+				
Juncus conglomeratus			+		+				+			+	
Juncus effusus	+	+											
Juncus squarrosus		+	+				+			+	+		+
Linum catharticum			+			+							
Lotus corniculatus		+											
Luzula campestris	+							+					
Luzula multiflora		+				+	+						
Lysimachia nemorum					+								
Molinia caerulea	+	+	+	+	+	+		+	+	+	+	+	+
Myrica gale				+					+	+	+	+	+
Nardus stricta	+		+			+	+	+			+		+
Narthecium ossifragum								+	+	+	+	+	+
Pedicularis palustris		+					+						
Pinguicula lusitanica												+	
Pinguicula vulgaris			+	+	+	+	+		+		+		+
Polygala serpylifolia								+					
Potamogeton polygonifolius									+				+
Potentilla erecta	+	+	+			+	+	+	+	+	+	+	+
Prunella vulgaris	+	+				+	+						
Pteridium aquilinum										+			
Ranunculus acris		+				+	+						
Ranunculus flammula					+	+	+				+		+
Rumex acetosella	+												
Salix aurita	+										+		
Saxifraga aizoides				+	+								
Schoenus nigricans												+	
Succisa pratensis				+	+	+	+				+	+	+
Thalictrum alpinum						+							
Thymus serpyllifolia			+										
Tofieldia pusilla				+									
Trichophorum cespitosum										+			
Trichophorum germanicum				+					+		+	+	+
Trifolium repens	+	+	+										
Triglochin palustris									+				
Viola palustris ssp. palustris						+	+		+				+

Table AII:3. Associated species lists for the Loch an Daimh Mor quadrats within the grid, Chapter 5. Q = Quadrat

Associated Species	**Grid**					
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
Agrostis canina	+	+	+		+	+
Agrostis capillaris				+		
Agrostis vinealis					+	
Anthoxanthum odoratum	+					
Calluna vulgaris		+			+	+
Carex demissa			+	+		+
Carex echinata	+	+	+	+		
Carex flacca	+	+			+	
Carex hostiana	+		+	+		+
Carex panicea	+		+	+		+
Carex pulicaris	+			+		
Dactylorhiza maculata		+				
Danthonia decumbens	+	+	+	+	+	
Drosera rotundifolia			+			
Epilobium palustre (young)		+				
Erica cinerea		+			+	
Erica tetralix	+	+	+	+		
Festuca filiformis						
Festuca ovina	+	+	+	+	+	
Festuca rubra ssp. rubra					+	
Hypericum pulchrum			+			
Hypochaeris radicata	+					
Juncus articulatus	+					
Linum catharticum	+					
Molinia caerulea	+	+		+	+	
Myrica gale			+	+		
Nardus stricta	+				+	+
Narthecium ossifragum		+		+		+
Pinguicula lusitanica	+	+		+		
Potentilla erecta	+	+	+	+	+	
Ranunculus flammula				+		+
Salix repens	+					
Schoenus nigricans			+			
Selaginella selaginoides		+		+		
Silene flos-cuculi	+					
Succisa pratensis	+	+	+			
Trichophorum germanicum		+			+	+
Viola riviniana					+	

Table AII:4. Associated species lists for all quadrats in transects 1 and 2 of the Faraid Head site, Chapter 5. Q = Quadrat

Associated Species	**Transect 1**										**Transect 2**									
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**
Agrostis canina						+														+
Agrostis capillaris	+	+	+		+		+	+		+	+	+	+	+	+		+	+		
Angelica sylvestris				+			+		+				+			+				
Anthoxanthum odoratum	+	+	+	+		+	+	+		+	+		+		+	+	+	+		
Avenula pratensis					+															
Bellis perennis		+			+								+				+			
Cardamine pratensis	+																			
Campanula rotundifolia											+									
Carex binervis																			+	
Carex flacca	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Carex hostiana						+	+		+	+					+			+	+	+
Carex nigra								+		+										
Carex panicea				+				+		+										
Carex pulicaris		+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+
Centaurea nigra															+	+				
Cerastium fontanum											+	+				+	+			
Cynosurus cristatus												+	+	+						
Dactylis glomerata																	+			
Dactylorhiza maculata						+			+					+		+			+	+
Danthonia decumbens							+			+	+			+				+	+	
Daucus carota	+																			
Festuca ovina						+			+	+		+	+			+		+	+	+
Festuca rubra	+	+		+	+	+		+	+		+	+	+	+	+	+	+	+		+

Table AII:4. ***Continued***

Associated Species	Transect 1										Transect 2									
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Filipendula ulmaria						+	+													
Galium verum		+													+	+				
Gentianella campestris		+	+										+			+				
Holcus lanatus	+	+	+	+			+	+	+			+	+	+	+					
Hypochaeris radicata	+	+							+	+	+	+	+	+	+	+	+			
Juncus articulatus								+										+		
Juncus bufolius s.s.																	+			
Koeleria macrantha	+	+	+		+				+		+		+	+	+	+	+	+		+
Linum catharticum		+				+	+			+	+	+	+	+	+			+	+	
Lotus corniculatus		+	+		+						+	+		+		+		+		
Luzula campestris	+	+		+								+								
Molinia caerulea				+		+	+	+	+		+	+		+	+	+		+	+	+
Parnassia palustris								+												
Pedicularis sylvatica						+														
Persicaria vivipara														+						
Pinguicula vulgaris																			+	
Plantago lanceolata		+	+				+						+	+	+		+			
Platanthera bifolia																				
Plantago maritima													+	+	+	+	+			
Poa humilis		+	+		+												+			
Potentilla erecta			+	+	+	+	+		+	+	+	+		+	+	+		+		+
Prunella vulgaris	+	+	+		+								+		+					
Ranunculus acris	+	+	+	+			+		+				+				+			

Table AII:4. ***Continued***

Associated Species	**Transect 1**										Transect 2									
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**
Rumex acetosa	+	+																		
Scorzoneroides autumnalis	+				+								+		+		+			
Succisa pratensis						+	+		+	+	+	+		+		+		+		+
Thymus polytrichus		+														+				
Trifolium pratense								+					+			+				
Trifolium repens	+	+	+	+	+		+					+	+		+		+			
Vicia cracca							+	+		+										
Viola riviniana	+			+	+										+	+				

Table AII:5. Associated species lists for all quadrats in transects 1, 2 and 3 of the Betty Hill site, Chapter 5. Q = Quadrat

Associated Species	**Transect 1**										**Transect 2**											**Transect 3**										
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Agrostis canina					+	+		+	+																	+		+	+			
Agrostis capillaris	+	+	+	+		+	+	+	+	+	+			+	+	+			+	+	+	+	+	+							
Agrostis stolonifera																+	+	+														
Anthoxanthum odoratum	+	+	+	+	+	+	+	+	+	+																		+				
Bellis perennis														+																		
Calluna vulgaris	+	+	+		+	+	+	+	+	+		+			+	+		+	+	+		+	+	+	+	+	+	+		+		+
Campanula rotundifolia														+																		
Carex binervis				+																						+	+	+		+		
Carex caryophoyllea												+												+								
Carex demissa													+	+			+								+						+	
Carex echinata				+	+																											
Carex flacca		+	+		+	+		+			+										+			+							+
Carex hostiana																																+
Carex nigra																			+						+			+	+			
Carex panicea								+	+	+		+	+		+			+	+	+					+	+	+	+	+	+		+
Carex pilulifera		+	+	+	+		+																									
Carex pulicaris						+			+						+	+		+		+				+	+	+	+	+	+		+	+
Carex viridula																															+	
Cerastium fontanum	+									+																						
Dactylorhiza maculata		+			+	+	+		+	+									+												
Danthonia decumbens		+		+	+	+		+	+			+	+		+		+			+			+		+	+	+	+	+	+	
Deschampsia flexuosa																		+	+													
Empetrum nigrum					+						+				+	+	+			+	+						+					
Erica cinerea	+	+	+	+	+	+	+	+	+	+		+			+		+	+							+	+		+				+

Table AII:5. ***Continued***

Associated Species	**Transect 1**										**Transect 2**										**Transect 3**									
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Erica tetralix			+	+	+							+	+		+		+	+		+							+	+	+	+	+	
Eriophorum angustifolium																														+	+	+
Festuca ovina		+	+	+	+	+	+	+		+			+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+
Festuca rubra	+	+															+															
Gymnadenia borealis	+																															
Holcus lanata							+	+		+																						
Hypochaeris radicata										+													+	+								
Juncus articulatus				+																												
Juncus bulbosus				+																											+	
Koeleria macrantha	+										+				+	+					+			+								
Leontodon autumnalis														+				+							+							
Linum cartharticum																		+														
Lotus corniculatus	+	+	+		+	+	+		+	+	+	+			+		+	+	+	+	+	+	+	+								
Luzula multiflora		+	+	+			+		+	+							+															
Molinia caerulea				+																											+	
Nardus stricta				+	+				+		+	+									+				+	+	+			+	+	+
Narthecium ossifragum				+	+																											
Parnassia palustris											+	+		+							+				+			+			+	
Pedicularis palustris																												+				
Pedicularis sylvatica				+	+				+																		+			+	+	+
Pinguicula lusitanica																																+
Pinguicula vulgaris																											+					
Plantago coronopus											+										+	+										

Plantago lanceolata	+										+	+	+	+	+						+	+	+	+	+							

Table AII:5. ***Continued***

Associated Species	**Transect 1**										**Transect 2**											**Transect 3**										
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Plantago maritima	+	+		+		+			+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+			+	+	+
Poa humilis	+																															
Polygala serpylifolium													+															+				
Potentilla erecta	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+			+	+	+	+		+	+	+	+	+	+	+	+	+	+
Primula scotica											+	+									+											
Prunella vulgaris				+								+																				
Pteridium aquilinum				+																												
Ranunculus flammula																														+	+	
Rhinanthus minor	+						+																									
Salix repens		+	+	+	+	+	+	+		+																						
Schoenus nigricans																																+
Scilla verna											+	+			+		+	+			+	+	+									
Selaginella selaginoides												+	+													+		+			+	+
Succisa pratensis	+	+	+	+	+	+	+	+	+	+									+				+	+	+	+	+	+	+	+	+	+
Thymus polytrichus											+	+		+	+	+	+		+		+	+	+	+						+		
Trifolium repens	+			+									+	+																		
Ulex europaea										+																						
Viola riviniana	+												+	+	+					+			+	+				+	+		+	

Appendix III: trnL$^{(UAG)}$-rp132 contig

Below is the contig for the *Euphrasia* chloroplast locus trnL$^{(UAG)}$-rp132. Polymorphic sites are in bold.

TAATAGC**C**TATGAATAAGCAATCGTCTAGATTTAAGAATGAAATCGGGTG**TAA**

TAGTTTTTAGGAAATATTCAAATTGATGAATAAAAATTTT**A**G**G**TATTTGAAGGT

TCATTATTTTAATTTCAG**G**CTAGGGTCATAATTTTATTGTGTATTTTCTATTACTA

CTTTAT**C**G**T**GATTTGCACTTTTTTTGTAAAACCATACAGTCT**A**TGGGTTAGAATC

GCCTCAAAAAAATGGCTGTTTGC**C**TTT**:**TTGACTTATTTAAAATTTCCACATATTT

ATTCAGAAAACACCTTAACTAAG**T**AAGTATTTTT**G**CGTAGATTAATGCTGATATA

TCTAGAAGGGGTCTAT**:**ATAGTAGATTAGCAAAGAAAGAAAGTTATACAAAAAC

TAAAATTTAGTATCTT**G**TTTTTTTGTCAACAAGACGGACGTCTTT**:**TCTTTTTATT**T**

TTAATTCGGAAAGGTAATATAGAAAAATTCTTATTCTAAATTCTATT**C**TAAAATC

CAAATTATCGAATTCCATTTC**GT**A**G**T**GAGTAA**CACTACGAAATGGAATTCGAT

AATTTGGATTTTATATTTGATAGGAAATGACTTCAAAAAAGTGAGTC**G**AGTCTAT

TTA**G**ATTATTCAAACATGTTATGTTTTAATA**T**TTTTTTTTA

Appendix IV: Full trnL$^{(UAG)}$-rp132 minimum spanning network for

Below is the minimum spanning network for the chloroplast locus trnL$^{(UAG)}$-rp132 in all the *Euphrasia* sampled. This network was produced using the software HapStar (Teacher and Griffthis 2011)