Effects of Chronic Habitat Fragmentation on Population Genetic Processes in Temperate Tree Species

The example of rowan and ash in a deforested landscape and implications for native woodland restoration in southern Scotland

Cécile Fanny Emilie Bâcles

A thesis submitted to the University of Edinburgh in fulfilment of the degree of Doctor of Philosophy

Institute of Cell, Animal and Population Biology
The University of Edinburgh

2004

ACKNOWLEDGEMENTS

I would like to thank immensely my supervisors Richard Ennos and Andrew Lowe for their valuable guidance throughout the Ph.D.. Richard, I am indebted to you for sharing your infinite knowledge of population genetics with me and for unconditionally trusting my decisions in the running of this project whilst being always available for advice when I needed it. Andy, thank you so much for bringing your dynamism and experience to this project, it would not have been possible to achieve it without your help.

Very special thanks are due to Sarah Rendell for having been my unofficial third supervisor, not only guiding my first steps in the lab four year ago and today correcting thesis drafts from Oxford but also listening to the ups and downs of Ph.D. life in front of a glass of wine.

I would like to thank the Wildwood Group of the Borders Forest Trust and in particular, Philip Ashmole and Adrian Newton whose initiative developed into this project. I hope this thesis will provide useful guidance in the accomplishment of their brilliant enterprise. I am very grateful to Hugh Chalmers for granting my requests, to Tilhill Forestry, John Barker and Christopher Moffat for allowing me to visit their land, especially during the very difficult time of the foot-&-mouth outbreak.

Financial support from the Scottish Forestry Trust and from the Natural Environment Research Council is gratefully acknowledged. I also would like to thank the Entente Cordiale Scholarship Trust and the British Council for their early contribution in the form of a research scholarship which allowed me to make this dream come true. Thanks are also due to Roger Arditi and André Gallais at the Institut National Agronomique Paris-Grignon for their inspirational teaching.

Thank you to everyone in the biosystem management section at CEH Edinburgh and to Joan Cottrell at NRS for a very warm welcome in their lab for the mammoth task of microsatellite screening. I had a brilliant time during my year spent at CEH. Andy, thank you again for giving me this opportunity. Stephen, thank you for being Andy when Andy was not around, and Bob, thank you for your cheerfulness despite my tantrums! I am looking forward to working with you all again soon.

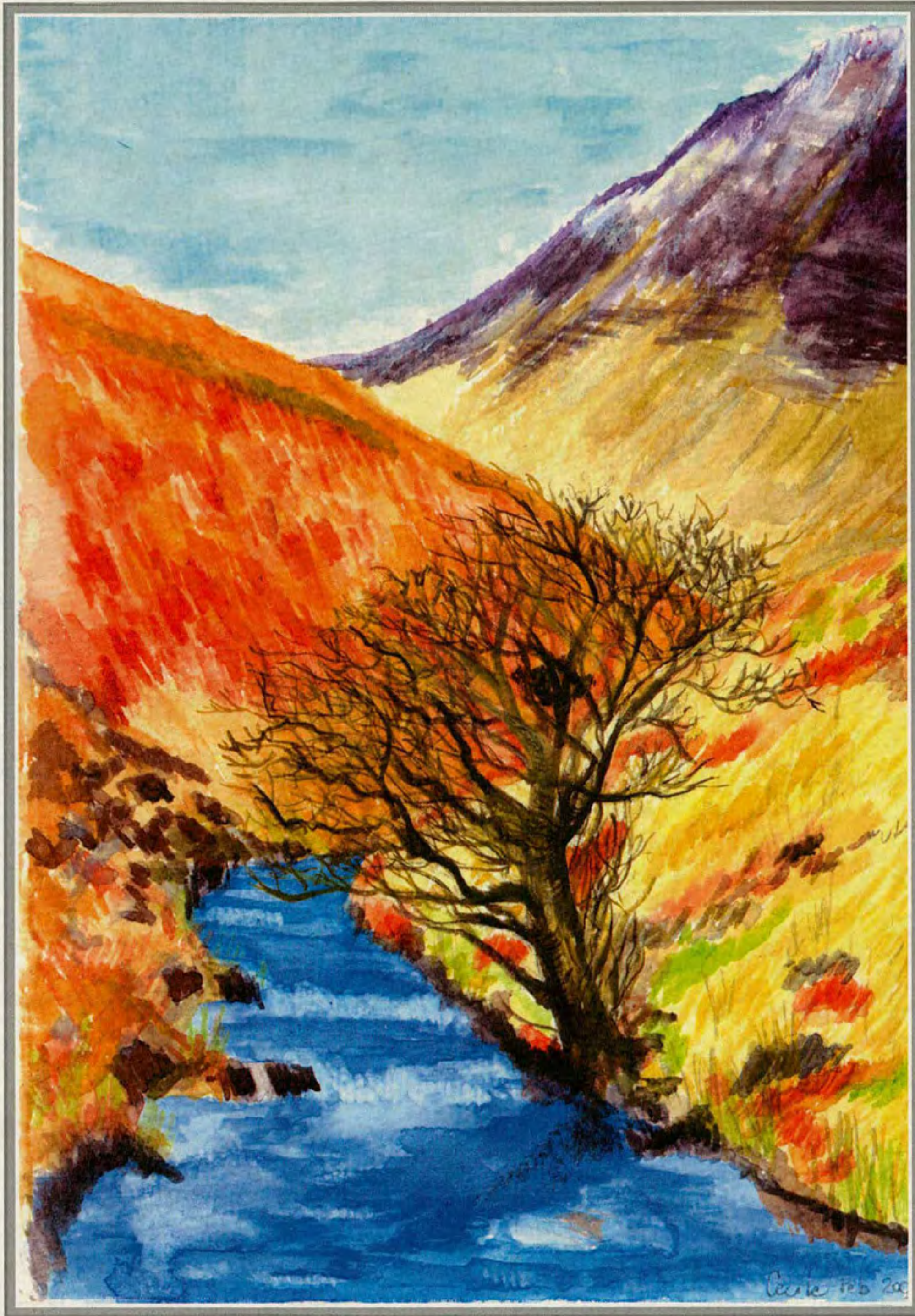
Many thanks are due to students and staff at the University of Edinburgh and elsewhere. In particular, I would like to thank John Morman for his help in the lab, Shiela Wilson and Connie Fox for knowing everything and staff at the Darwin library and the Multimedia Centre of the university for their very efficient help and great resources. Several people have contributed to improving the content of this thesis by debugging a programme or sharing thoughts and ideas when meeting at a conference, and I am indebted to them for their time and interest in the project. Thank you especially to Sophie Gerber, Olivier Hardy, Eva Wallander, Stephen Harris for offering solutions to isozyme extraction in Fraxinus excelsior, and Oliver Raspé for extensive advice on Sorbus aucuparia, isozymes, seed stratification and more! Thank you to Sylvie Oddou-Muratorio for sharing her S. torminalis microsatellite primers with me and sending me her outstanding Ph.D. thesis, to Jarek Burzyck for letting me play with NEIGHBOR and to Jim Wright who saved me from a divorce with ARCVIEW and GIS analysis. Many thanks are also due to Peter Levy and Fabiola Lopez-Barberra for last minute thesis proof-reading.

Thank you to Becky, Chris, Sofia, Sylvia, and Matt (Matt, will I ever be able to thank you enough?) for climbing, getting wet, getting cold, getting lost, and even collecting samples. Thank you to Ashley, Sarah B., Sarah R. again, Claudia, Fabiola, Simoneta and Graham for advice, sympathy and generally good atmosphere in the lab. You have all become great friends.

The Ph.D. experience was certainly not 'all work no fun' and many people, old friends or new encounters, have contributed to making my time here very special. Edinburgh University Sailing Club has to be thanked, of course, for reminding me what student life is truly about: going to the pub three nights a week, sailing with a hangover and getting dressed up in impossible fancy costumes, but also for teaching me how to sail and sharing the challenge of running a really good university club. Thank you Allan, David, Heather, Jenny, Matt, Ross, Tommy G. and Viv. I am also grateful to Rodolfo and Tumi at Dancebase for transforming stress into physical exhaustion and positive energy, to George for being the best flatmate, to Chiara, Daniela and Tim for the drawing group and others at CEH for memorable outings, and to Sazini for everything. Two people have managed to stay close friends despite leaving Scotland so many years ago, Aidan and Friederike, thank you so much for your long-distance, long-lasting friendship.

Un mot de français pour remercier tous ceux que j'ai laissés en France et à qui j'ai consacré trop peu de temps. Les amis fidèles, Thierry et Stéphanie, mon chef de chœur préféré Jean-Michel, et la

famille éparpillée aux quatre coins de la France, merci pour vos coups de fil, emails, cartes de vœux et d'anniversaire. Cette année a été difficile puisque certains nous ont quittés, mon oncle Michel, ma presque cousine Virginie et mon grand-père de cœur Tonton Pierre. Je pense à vous. Enfin, je n'aurai pu être épanouie et heureuse comme je le suis aujourd'hui sans le soutien et l'amour de ma sœur et de mes parents. Ils m'ont toujours supportée dans mes choix, même les plus extravagants et ambitieux et je leur en suis à jamais reconnaissante.



'THE END SPROUTS FROM THE MEANS, AS THE TREE GERMINATES FROM THE SEED'

Mohandas Karamchad Gandhi

DEDICATION

I dedicate this thesis to my grand-parents, **Pierre et Thérèse Brunel**, who taught me the natural world. They gave me the will to understand it better and to help finding a balance between preserving the Earth resources and satisfying human needs.

Je dédie cette thèse à mes grands-parents de cœur, ***Pierre et Thérèse Brunel****, qui m'ont appris la Nature. Ils m'ont donné la volonté de mieux la comprendre et de contribuer à établir une relation d'équilibre entre sauvegarder la Planète et satisfaire les besoins des hommes.*

ABSTRACT

Concerns have been expressed regarding the viability of forest remnants due to detrimental genetic consequences of habitat fragmentation. However, empirical studies conducted so far suggest that population genetic processes respond in more varied ways than expected, thus highlighting the need for evidence from a wider range of species and in situations where fragmentation is long-standing.

In southern Scotland, human-mediated deforestation for pasture since the Neolithic has dramatically altered the landscape. A single catchment (Moffat Dale) was intensively surveyed for severely fragmented populations of *Sorbus aucuparia* L., an insect pollinated bird dispersed species, and *Fraxinus excelsior* L., which is wind pollinated and wind dispersed. These remnants are being considered for seed collection in a native woodland restoration programme currently being implemented.

Quantifying genetic variation at isozyme and chloroplast DNA markers in *S. aucuparia* remnants revealed that high levels of genetic diversity are maintained. However, genetic differentiation among remnants was detected for both types of marker and the estimated ratio of pollen flow to seed flow between fragments is close to one (r=1.36) suggesting reduced historical pollen-mediated gene flow but efficient seed dispersal.

Similarly, *F. excelsior* remnants maintain high levels of genetic diversity at nuclear microsatellite markers and low interpopulation differentiation (Θ=0.080). Using the neighbourhood model, it was estimated from open-pollinated progeny arrays that contemporary pollen flow is extensive and that effective pollen dispersal distance within the catchment averages 328 m. A detailed paternity analysis conducted on progeny arrays confirmed these results. Although pollen flow is an important component of realised gene flow, a parentage analysis showed that it is not predominant as 56.6% of the seedlings that recently established in Moffat Dale immigrated into the catchment.

S. aucuparia and *F. excelsior* remnants in a severely deforested landscape are part of a wide reproductive network. Genetic diversity within remnants and gene exchange among them have been maintained by efficient long distance seed and pollen-mediated dispersal, making remnants an appropriate seed source for planting stock.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS i

DEDICATION v

DECLARATION vii

ABSTRACT ix

TABLE OF CONTENTS xi

Chapter 1: GENETIC CONSEQUENCES OF HABITAT FRAGMENTATION IN PLANT POPULATIONS: A REVIEW OF EXPECTATIONS AND EVIDENCE 1

I Introduction 1
II Theoretical population genetic consequences of habitat fragmentation 3
1 Organisation of genetic variation 3
2 Loss of genetic variation 4
3 Increased inbreeding and inbreeding depression 5
4 Reduced migration 7
III Empirical evidence of genetic effects in naturally fragmented populations of plants 10
1 Loss of genetic variation 10
2 Alteration of mating patterns 11
3 Interpopulation differentiation and gene flow 13
4 Conclusions 15
IV The southern Uplands of Scotland: a deforested landscape 17
1 Study site: Moffat Dale 17
2 The Carrifran Wildwood Project and reforesting southern Scotland 18
V Study species: same habitat, contrasting pollen and seed dispersal 21
1 *Sorbus aucuparia* L., rowan or mountain ash 21
2 *Fraxinus excelsior* L., common ash 25
VI Objectives 30

Chapter 2: USING MOLECULAR MARKERS TO IDENTIFY HISTORICAL AND CONTEMPORARY POPULATION GENETICS PROCESSES 33

I Molecular markers: polymorphism and neutrality 33
1 Isozymes 33
2 Microsatellites 34
3 Chloroplast DNA (CpDNA) markers 34
II Population structure and historical gene flow 35
1 Hardy-Weinberg principle 35
2 Measuring polymorphism 35
3 Inbreeding coefficient and *F*-statistics 36
4 Effective number of migrants 37
III Estimating contemporary patterns of mating and dispersal 38
1 Paternity analysis 38
2 Parentage analysis 41

Chapter 3: GENETIC EFFECTS OF CHRONIC HABITAT FRAGMENTATION ON TREE SPECIES: THE CASE OF *SORBUS AUCUPARIA* L. IN A DEFORESTED SCOTTISH LANDSCAPE 43

Abstract 43
I Introduction 44
II Material and methods 47
1 Study site and populations 47
2 Plant material 48
3 Protein extraction and electrophoretic procedure 49
4 DNA isolation and PCR-RFLP analysis of cpDNA 49
5 Statistical analysis 50
III Results 53
1 Genetic diversity at isozyme loci 53
2 Characterisation of cpDNA haplotypes and diversity 55
3 Population differentiation in *S. aucuparia* remnants 56
4 Pollen versus seed-mediated gene flow 58
IV Discussion 59
V Conclusions 65

Chapter 4: GENETIC STRUCTURE AND MATING PATTERNS OF *FRAXINUS EXCELSIOR* L. REMNANTS IN A DEFORESTED SCOTTISH LANDSCAPE 67

Abstract 67
I Introduction 68
II Material and methods 71
1 Study site and populations 71
2 Mapping 72
3 Plant material 73
4 DNA isolation 73
5 Microsatellite amplification and genotyping 74
6 Statistical analysis 75
III Results 79
1 Independence of microsatellite loci and choice of markers 79
2 Genetic diversity within remnants 79
3 Departure from random mating 79
4 Genetic differentiation among *F. excelsior* remnants 81
5 Historical gene flow among remnants 82
6 Mating patterns of *F. excelsior* within Moffat Dale 82
IV Discussion 86
1 Genetic structure and historical gene flow among forest fragments 86
2 Contemporary mating patterns within a deforested landscape 89
Appendix 4.1 : Effective pollen dispersal distance and male neighbourhood size for 19 *F. excelsior* mother trees sampled in three remnants of Moffat Dale. 92

Chapter 5: LANDSCAPE APPROACH TO CONTEMPORARY GENE DISPERSAL OF *FRAXINUS EXCELSIOR* L. I. PATERNITY ANALYSIS OF POLLEN FLOW IN FOREST FRAGMENTS 93

Abstract 93
I Introduction 94
II Material and methods 97
1 Study site and sampling strategy 97
2 Tree gender and sample collection 99
3 DNA isolation 99

4 Microsatelllite amplification and genotyping 100
5 Data analysis 100
6 Paternity analysis 100
III Results 105
1 Choice of markers and introduction of error rate 105
2 Paternity analysis 107
IV Discussion 116
1 Methodology of paternity analysis 116
2 Contemporary pollen flow for *F. excelsior* across a fragmented landscape 118
V Conclusions 122
Appendix 5.1: Paternal contribution to seed sample of *F. excelsior* mature individuals sampled in Moffat Dale 123

Chapter 6: LANDSCAPE APPROACH TO CONTEMPORARY GENE DISPERSAL OF *FRAXINUS EXCELSIOR* L. II. PARENTAGE ANALYSIS REVEALS EXTENSIVE SEED FLOW IN FOREST FRAGMENTS 125

Abstract 125
I Introduction 126
II Material and methods 129
1 Study site 129
2 Sampling 129
3 DNA isolation and microsatellite genotyping 131
4 Genetic diversity 131
5 Parentage analysis 132
III Results 136
1 Microsatellite diversity in adult and seedling samples 136
2 Parentage analysis 137
IV Discussion 142
V Conclusions 147

Chapter 7: IMPLICATIONS OF THE DYNAMIC GENETIC BEHAVIOUR OF FOREST FRAGMENTS FOR NATIVE WOODLAND RESTORATION IN SCOTLAND 149

I Introduction 149
II Summary of results 150
III Genetic management of forest restoration 154
IV Conclusions and perspectives 157

REFERENCES 159

APPENDICES 175

APPENDIX A: DESCRIPTION OF GENOTYPING PROCEDURE FOR EIGHT *FRAXINUS EXCELSIOR* MICROSATELLITE LOCI: ADDRESSING THE PROBLEM OF NULL ALLLELES AND GENOTYPING ERROR 177

APPENDIX B: PUBLICATIONS 203

Chapter 1 : GENETIC CONSEQUENCES OF HABITAT FRAGMENTATION IN PLANT POPULATIONS: A REVIEW OF EXPECTATIONS AND EVIDENCE

I Introduction

Habitat fragmentation occurs in natural ecosystems, notably through disturbance by fire, wind and flooding (Andren 1994). However, human activities, such as deforestation for commercial wood products and subsequent land-use for cultivation and cattle grazing, are the most important, large-scale causes of fragmentation and loss of natural habitat. For the last 20 years, a great deal of attention has been given to the sudden deforestation of the tropics which is occurring at an alarming rate of 100 000 to 200 000 km^2 per year (Katzman and Cale 1990; White *et al.* 1999), threatening the existence of the world's reservoir of biodiversity (Sih *et al.* 2000). Less commonly considered is the much more ancient and continuing deforestation of the temperate zone which has also greatly reduced forest cover (Wilcove *et al.* 1986). In Great Britain, for instance, forest fragmentation began over 5000 years ago with permanent clearances by Neolithic farmers and was well advanced by the 11th century (Wilcove *et al.* 1986).

It is widely accepted that habitat fragmentation associated with anthropogenic disturbances threatens the maintenance of biological diversity (Wilcox and Murphy, 1985), leading to population extinction and species rarefaction (Wilcove *et al.* 1986). Habitat fragmentation also affects communities (Robinson *et al.* 1992; Didham *et al.* 1996) and physical processes (Young and Mitchell 1994). Research so far has primarily focused on assessing the effects of habitat fragmentation on ecosystem processes (e.g. Andren, 1994; Didham *et al.* 1996), although Robinson *et al.* (1992) pointed out that population analyses are fundamental in shedding light over the changes occurring in fragmented habitats. Nevertheless, the need for a more integrative approach, combining ecology and population genetics, has been recently acknowledged (Sih *et al.* 2000; Young and Clarke 2000).

Habitat fragmentation reduces the size and increases the isolation of suitable sites for a species (Wilcox and Murphy 1985). Three major component processes can be distinguished (Andren 1994): loss of original habitat (pure habitat loss), reduction in habitat patch size and increased isolation between habitat patches. Such 'insularisation' (Wilcove *et al.* 1986) of habitats produces a patchwork-like landscape, in a mosaic, where small patches of similar habitats are isolated from one another by a matrix of different habitats.

Theoretical predictions are that habitat fragmentation will be accompanied by dramatic changes in population genetics, affecting population viability in the short and in the long term (Young *et al.* 1996). For plant populations, these changes are likely to be complicated by their sessile habit once established (Loveless and Hamrick 1984), and therefore, their prefragmentation abundance. They will also be affected by their wide variety of reproductive systems (Cunningham 2000), the possibility of gene dispersal by both pollen and seed, the presence of seed banks (McCue and Holtsford 1998; Mahy *et al.* 1999) and interactions with pollinators and seed dispersers which may also be affected by fragmentation (Didham *et al.* 1996; Dick *et al.* 2003).

In this general introduction, I review the genetic effects predicted by population genetics theory of habitat fragmentation on populations of plants. I then present the empirical evidence for such effects that can be drawn from the existing literature in naturally fragmented plant populations, focusing mainly on forest tree species. Finally, I introduce the motivations for undertaking the research described herein and the main objectives of the project. A detailed description of the study site and species is also given.

II Theoretical population genetic consequences of habitat fragmentation

1 Organisation of genetic variation

The amount and distribution of genetic variation within and among populations of a species has been a major subject of population genetics research for 200 years or so because of its ecological and evolutionary significance (Linhart and Grant 1996). Genetic variation within a species is an important factor affecting its long-term survival (Avise 1994). It allows adaptation to environmental change in the short term and defines the evolutionary potential of a species in the long term (Vida 1994).

Both theoretical and empirical studies have now established that genetic variation, whether quantitative or discrete, is not randomly distributed within a species (Loveless and Hamrick 1984). Spatial and temporal genetic structure within and among populations results from the opposing action of evolutionary forces tending to produce local genetic differentiation and forces tending to produce genetic homogeneity (Slatkin 1987). While mutation is the ultimate source of genetic variation, it the distribution of the latter among regions, populations and individuals within populations depends on the relative importance of natural selection, migration and random genetic drift in the context of a species' biology and ecology (Loveless and Hamrick 1984; Barrett and Kohn 1991). Mutation, genetic drift because of finite population size, and selection favouring adaptation to local environment produce genetic differentiation. Gene flow resulting from the dispersal of gametes (for plant species, pollen and seed) and from the migration of individuals counteract these effects (Slatkin 1987). Relatively low levels of gene flow will balance the effects of genetic drift and directional selection (Ellstrand 1992) and according to Wright's shifting balance theory (Wright 1931; 1982), gene flow may favour the spread of beneficial alleles more rapidly than would have been possible under the unique action of selection (Slatkin 1987).

Theoretical predictions are that habitat fragmentation will reduce genetic diversity within populations and increase interpopulation genetic differentiation (Young *et al.* 1996). In the long term, loss of genetic variation is thought to reduce the ability of

populations to respond to changing selection pressures (Young *et al.* 1996), and in the short term, increased susceptibility to pest and diseases (Barrett and Kohn, 1991) and reduced population viability associated with the fixation of deleterious alleles (Huenneke 1991) may also be expected. In more general terms, the prediction is thathabitat fragmentation enhances the effects of diverging evolutionary forces and limits the effects of homogenising forces.

For plant species, a reduction in habitat patch size is usually equivalent to a reduction in population size. Plant population remnants are therefore expected to undergo genetic effects associated with small population size. It is fair to say that no population genetic theory of habitat fragmentation has been developed *per se*, rather, expectations have been borrowed from well established theory of the genetic behaviour of populations of small, finite size and from subdivided populations and metapopulation theory. The approach is justifiable since the matrix of small islands of habitat created by habitat fragmentation (Wilcove *et al.* 1986) corresponds to assumptions underlying these models. Hereafter I review in detail each of the expected population genetic effects of habitat fragmentation.

2 Loss of genetic variation

As a consequence of habitat fragmentation, plant populations are sharply reduced in size. Only a sample of the original number of individuals will survive. Thus, it is likely that only a sample of the alleles of the original population will remain in the remnants. The loss of genetic variation associated with a sudden reduction in population size is referred to as a genetic bottleneck (Barrett and Kohn 1991). The intensity of such a bottleneck will depend on the severity of habitat fragmentation, that is to say, the severity in reduction in population size.

Populations may subsequently remain small for many generations and undergo increased random genetic drift (Ellstrand and Elam 1993). Small populations are especially sensitive to random genetic drift because gene frequencies can undergo large fluctuations in different generations (but only small fluctuations will occur in sufficiently large populations; (Barrett and Kohn 1991). Whilst the first alleles to be

lost in a genetic bottleneck and in any one generation of genetic drift are rare alleles, increased genetic drift acting over generations leads to additional loss of more common alleles. This results in severe depletion of genetic diversity within populations (Lande 1988). Population viability may be affected within a relatively short period, since random genetic drift acts equally on neutral and selected genes and deleterious alleles may become fixed by chance (Huenneke 1991). However, if the disturbance triggers a dynamic of frequent extinction and recolonisation events, such that the average time that a population persists in an area is less than the time it takes for genetic drift to fix neutral alleles, then little loss of genetic diversity (due to random genetic drift) is expected.

3 Increased inbreeding and inbreeding depression

Not only is a loss of genetic variation within a population expected to occur following reduced patch size and increased isolation between patches, but the structure of genetic variation among populations, among individuals within population and within individuals is also altered. In particular, increased inbreeding is expected to arise following changes in mating patterns within populations. In normally outcrossed species, increased inbreeding may arise in either of two ways: if the proportion of selfing is increased, or if the proportion of mating with close relatives is increased.

For instance, following a bottleneck event, individuals of a self-incompatible species may find it difficult to find a compatible mate (Ellstrand and Elam 1993)asalleles conferring self-incompatibility may be lost. Self-incompatibility systems require large numbers of incompatibility alleles to maintain high levels of cross-compatibility and loss of such alleles may lead toselfing being selected for (Reinhartz and Les 1994; Washitani 1996).

In a more general manner, as a result of increased isolation between populations (and between individuals within populations), gamete exchange with more distant individuals or populations is reduced and mating between spatially close individuals becomes more likely. As spatially close individuals are more likely to be relatives

(and therefore are more likely to share identical alleles at one particular locus), such mating patterns result in biparental inbreeding (Barrett and Kohn 1991).

Increased inbreeding translates into a reduction of individual heterozygosity. Theory predicts a reduction of individual fitness associated with increased homozygosity (Charlesworth and Charlesworth 1987). The phenomenon is commonly referred to as inbreeding depression which encompasses the loss in fitness in inbred offspring compared with offspring from random mating within the same population (Kirkpatrick and Jarne 2000). Although the actual relationship between reduced fitness and increased homozygosity is poorly understood (Hedrick 1994; Husband and Schemske 1997), two hypotheses prevail, the overdominance hypothesis and the partial dominance hypothesis (Charlesworth and Charlesworth 1987; Keller and Waller 2002). In the overdominance hypothesis, inbreeding depression is caused by a selective advantage of heterozygotes over homozygotes at individual loci affecting fitness. In the partial dominance hypothesis, inbreeding depression is caused by the fixation of recessive or partially recessive deleterious alleles in inbred offspring (Charlesworth and Charlesworth 1987; Hansson and Westerberg 2002).

Inbreeding depression will affect individual fecundity, establishment and survival and will ultimately affect population viability (Hedrick 1994; Keller and Waller 2002). These effects are likely to be greater for plant than animal species (Barrett and Kohn 1991). However, how population viability is affected depends on the composition of the genetic load (i.e. the population component of inbreeding depression), which in turn may be influenced by population size and history (Keller and Waller 2002) and the nature of mutations causing inbreeding depression (Bataillon and Kirkpatrick 2000).

Historically outcrossing populations are expected to harbour a high genetic load. A shift to mating patterns which increase inbreeding after fragmentation is therefore expected to have severe effects on fitness (Reinhartz and Les 1994), although Kirkpatrick and Jarne (2000) argue that following a genetic bottleneck, the inbreeding depression caused by deleterious mutations may initially decrease.

Chronically small populations are likely to exhibit lower levels of inbreeding depression if deleterious mutations have been purged by selection. The nature and degree of purging depends on the genetic basis of inbreeding depression (Crnokrak and Barrett 2003). Successive generations of inbreeding will purge lethal alleles while the genetic load resulting from mildly deleterious alleles will persist (Hedrick 1994; Bataillon and Kirkpatrick 2000).

Whilst a reduction of genetic variation will affect population viability in the longer term, changes in mating parameters respond more rapidly to habitat fragmentation. For plant species, the effects on mating parameters may be duplicated by association with pollinators and seed dispersers that may also be affected by fragmentation.

4 Reduced migration.

Colonisation. As habitat fragmentation reduces the extent of original habitat, the probability for a migrant, a seed in the case of plant species, to establish in a suitably colonisable site is lowered. When a species' dynamics are characterised by frequent extinction of local populations and colonisation of new sites (i.e. a metapopulation), the metapopulation is under threat of extinction from this purely demographic effect, if extinction cannot be balanced by recolonisation in a fragmented habitat (Lande 1988).

Nevertheless, when colonisation of a new site eventually occurs, it is likely that only a few original founders, that carry only a fraction of the total genetic variation of the parental population, succeed in establishing. Such a founder effect will have similar population genetic consequences to a bottleneck (Barrett and Kohn 1991). The founder effect may be especially strong in a fragmented habitat because the population source of migrants may itself have suffered the detrimental effects of increased drift and inbreeding.

Gene flow among extant remnants. Following the most commonly used definition of Endler (1977), gene flow refers to gene exchange among populations and

corresponds to the proportion of newly immigrant genes in a given population, while gene movement refers to intra-population phenomena.

Initial predictions are that habitat fragmentation, by increasing isolation between patches, reduces gene flow among populations. Such reduction in the proportion of immigrant genes, combined with the effects of increased genetic drift and inbreeding within subpopulations, increases genetic differentiation among fragmented remnants (Young *et al.* 1996). Ultimately, the action of selection favouring local adaptation is strengthened and the fitness of seed fertilised by immigrant pollen or the offspring of immigrant seed may be lowered. Such reduction of fitness may occur when the coadapted complex of genes are broken apart when genotypes from different locations are crossed (Barrett and Kohn 1991). The process called outbreeding depression renders gene flow between fragments more limited.

Gene flow is best known as a force homogenising the distribution of genetic variation among populations, with only limited levels of gene flow necessary to counteract genetic drift, moderate levels of directional selection and maintain high levels of genetic diversity (Slatkin 1987).

However, gene flow in plants is idiosyncratic, varying greatly between species, populations and seasons (Young *et al.* 1996). The ultimate effect of gene flow on local population diversity will also depend on the nature of genetic variation in the gene flow source relative to the sink population (Ellstrand and Elam 1993) and on how the source population is connected to other populations (Mills and Allendorf 1996; Couvet 2002) .Furthermore, the actual relationship between inbreeding depression and gene flow has been addressed theoretically only very recently (Couvet 2002). It appears that the viability of individuals is expected to be reduced in populations of small size when gene flow is limited. Levels of gene flow greater than one migrant per generation (sufficient to balance the effect of genetic drift) may be required to prevent the accumulation of deleterious mutations in small populations (Couvet 2002). Mills and Alllendorf (1996) suggest that, in natural populations, a

minimum of one and a maximum of 10 migrants per generation is an appropriate level of genetic connectivity between populations.

In species with historically high levels of gene flow, habitat fragmentation will most likely provoke a drop in the amount of gene flow (by increasing the distance between populations and by subsequently disturbing pollinator and seed disperser guilds). This enhances the detrimental effects associated with genetic drift and inbreeding associated with small population size. However, if the disturbance created by fragmentation reduces the size of a population so that the fraction of seed fertilised by immigrant pollen or the fraction of immigrant seed increases, gene flow is enhanced even if the absolute number of migrants may be reduced (Ellstrand and Elam 1993).

III Empirical evidence of genetic effects in naturally fragmented populations of plants

Theoretical population genetic consequences of habitat fragmentation are straightforward extensions of population genetics theory. Empirical validation of these predictions is limited, as few studies so far have investigated the effects of fragmentation in natural populations. The available results show that habitat fragmentation may have genetic effects in plant species more complex and varied than first expected.

1 Loss of genetic variation

Empirical data usually confirm that the amount of genetic variation maintained within a fragment is related to its population size. For instance, Prober and Brown (1994) found a significant positive logarithmic relationship between the population size of *Eucalyptus albens* fragments and both the percentage of polymorphic loci and allelic richness measured at isozyme loci. Similarly, Young *et al.* (1999) found a strong relationship between polymorphism and allelic richness measured at isozyme loci and the logarithm of the reproductive population size of fragments of the grassland daisy, *Rutidosis leptorhynchoides,* reflecting the loss of rare alleles in smaller populations.

Interestingly, Prober and Brown (1994) found that more isolated remnants of *E. albens* were genetically depauperate (when less isolated remnants of similar size were not), suggesting the possibility of a fragmentation threshold (Young *et al.* 1996). Similarly, Cruzan (2002) reported a fragmentation threshold at a scale of 8 km, where metapopulation size was positively associated with higher levels of genetic diversity in the herb, *Scrutellaria montana.* Andren, (1994) also raised the issue regarding the maintenance of species diversity.

The level of genetic variation maintained at microsatellite markers in adult and juvenile cohorts of continuous and fragmented stands have been compared in several neotropical tree species. In these recently fragmented forests, the level of genetic variation sampled in the adult cohort represents a prefragmentation abundance while

sapling and seedling cohorts should reflect the genetic effects of habitat fragmentation. Dayanandan *et al.* (1999) found that allelic richness was lower in saplings of isolated fragments of the tropical tree *Carapa guianensis*. For *Swietenia humilis,* another tropical tree species, White *et al.* (1999) also found that rare alleles present in the continuous stand were lost in fragments and that such allelic loss increased as population size decreases. However, Aldrich *et al.* (1998) found a similar number of alleles in adult and sapling cohorts of both continuous and fragmented stands of the tropical tree, *Symphonia globulifera* and increased genetic diversity was observed in fragments of the temperate tree, *Acer saccharum* compared to continuous stands (Young *et al.* 1993).

Whether a reduction of genetic diversity is observed in fragmented remnants depends on the measure considered. Indeed, in the same studies as mentioned above, expected heterozygosity was homogenous among cohorts and stands (Aldrich *et al.* 1998; White *et al.* 1999; Young *et al.* 1999). One possible explanation for such discrepancies between measures is that the populations of interest may have been fragmented for only a few generations (and maybe not more than one generation in neotropical tree species). Genetic drift is unlikely to have had time to have a significant effect and the loss of rare alleles observed therefore reflects genetic bottlenecks. While the loss of alleles through a genetic bottleneck is detected by a measure of allelic richness, expected heterozygosity will only decrease when alleles in medium frequencies are lost through genetic drift (Petit *et al.* 1998).

2 Alteration of mating patterns

Pollinator loss has been shown to result in strong selection for a self-fertile homostyle morph and therefore to lead to the loss of heterostyly in an isolated population of the endangered herb, *Primula sieboldii* (Washitani 1996). Pollinator loss also explains increased inbreeding (as measured by the inbreeding coefficient, F_{is}) in small populations of the herb *S. montana* (Cruzan, 2002). Franceschinelli and Bawa (2000) found that the outcrossing rate in populations of *Helicteres brevisp*era was directly correlated with plant density, which is low in forest areas and high in the savanna. Increased selfing rate in areas of low density may be explained by

pollinators tending to visit several flowers per plant (Franceschinelli and Bawa 2000). Similarly, geitonogamy is more likely to occur in isolated trees of the self-incompatible tropical tree *Samanea saman* resulting in higher levels of selfing associated with low population density and mass flowering (Cascante *et al.* 2002). In the tropical tree *Enterolobium cyclocarpum* contrasting results were reported as identical outcrossing rates were found in pasture and forest trees but the correlation of paternity among progeny was higher for trees in continuous forest suggesting that nearby trees producing abundant pollen contribute to fertilising several pods (Rocha and Aguilar 2001).

Nonetheless, a clear consensus among empirical studies is that significantly different mating behaviour is observed for trees in disturbed areas where trees occur in lower density (Rocha and Aguilar 2001). For self-incompatible species, alteration of mating patterns due to increased isolation and low density may be intensified as reproductive success may also be limited by lower transfer of compatible pollen as it is the case, for instance, for *Shorea siamens*, a partially self-incompatible tropical tree (Ghazoul *et al.* 1998).

Evidence for increased biparental inbreeding in fragmented populations has been reported for several species. In *R. leptorrhynchoides*, Young and Brown (1999) reported a paternal bottleneck (i.e. greater divergence between the allele frequencies in the populations and those in the pollen pool estimated from progeny arrays) and higher correlation of outcrossed paternity (i.e. production of more full-sibs) in small isolated populations. Similarly, Fuchs *et al.* (2003) found that the progeny of *Pachira quinata* trees from continuous stands, experienced lower levels of relatedness as a result of higher outcrossing rate or higher number of sires than the progeny of isolated trees.

Aldrich *et al.* (1998) compared the genetic structure at microsatellite loci of adults, saplings and seedlings cohorts of a continuous and several fragmented stands of *S. globulifera.* They found that a significant inbreeding coefficient was most often associated with the fragmented stands and the seedlings. However, in a similar

context, no significant inbreeding (as measured by F_{is}) was observed in either adult or sapling cohorts in fragmented populations of *C. guianensis* (Dayanandan *et al.* 1999).

A comprehensive picture of the extent and severity of inbreeding depression in natural populations is lacking because very few studies exist (Keller and Waller 2002), let alone in fragmented habitats. To my knowledge, inbreeding depression *per se* has never been measured in plant populations disturbed by habitat fragmentation but lower progeny vigour has been reported in progeny of pasture trees of *E. cyclocarpum* (Rocha and Aguilar 2001) and *S. saman* (Cascante *et al.* 2002). Interpretation of reduced viability or endangerment of fragmented populations usually relies on observation of increased inbreeding based either on estimates of outcrossing rates or inbreeding coefficients at neutral marker loci. Unfortunately, as laid out above, theoretical studies show that the correlation between heterozygosity and fitness is a complicated issue (Hansson and Westerberg 2002), especially since neutral and selected genomes may respond differently to small population size (Milligan *et al.* 1994). Further research in this area is clearly needed to assess the actual threat posed by inbreeding depression in naturally fragmented populations.

3 Interpopulation differentiation and gene flow

An impressive body of literature describes the genetic structure at neutral marker loci within and among plant populations using *F*-statistics. This relatively inexpensive approach has proven useful to assess genetic differentiation among populations and to infer the effective number of migrants among them from it (see Chapter 2). Restricting this review to forest tree species occurring in fragmented habitats, contrasting results are reported.

For instance, Hall *et al.* (1996) compared the genetic structure at isozyme loci of populations from a large reserve and several fragmented stands of the neotropical tree, *Pithecellobium elegans*. They found a positive relationship between the level of differentiation from the reserve population and the distance of the fragment from the reserve, thus suggesting that gene flow decreases as fragments become increasingly

isolated. Dayanandan *et al.* (1999) also found greater genetic differentiation among sapling cohorts than among adult cohorts in fragments of *C. guianensis* and concluded that this was a result of restricted gene flow as a consequence of fragmentation.

Interestingly, in *A. saccharum*, Fore *et al.* (1992) found that habitat fragmentation led to decreased genetic differentiation and increased gene flow in juvenile cohorts compared to the adult cohort. These results are in agreement with another study of *A. saccharum* (Young *et al.* 1993), suggesting that, in this species, gene flow has increased in forest remnants, as a result of logging allowing seed dispersal by wind to occur across greater distances (Fore *et al.* 1992; Young *et al.* 1993). Furthermore, significant inbreeding and genetic differentiation found in seedling cohorts of fragmented stands of *S. globulifera* is evidence of a genetic bottleneck (Aldrich *et al.* (1998). However, this genetic bottleneck is not the consequence of reduced gene flow into fragments. The reduction in effective population size results from increased gene flow from only a few sources in neighbouring pastures following fragmentation (Aldrich and Hamrick 1998). This was shown by a direct assessment of gene flow, performing a parentage analysis on seedlings established in forest fragments.

Studies that have attempted to measure contemporary gene flow directly are rare. This is however currently a very active area of research thanks to the advent of highly polymorphic markers and new statistical methods which have rendered the task more achievable (Sork *et al.* 1999). Interestingly, estimates from recently fragmented stands of several neotropical species suggest that contrary to theoretical expectations, isolated trees and fragments are part of a large reproductive network, for instance in *Enterolobium cyclocarpum* (Apsit *et al.* 2001). Gene flow by pollen occurs over very great distance in *P. elegans* (Chase *et al.* 1996), in *Spondias monbin* and in *Ficus* spp. (Nason and Hamrick 1997). Isolated trees of *S. humilis* (White *et al.* 2002) and *Dinizia excelsa* (Dick 2001; Dick *et al.* 2003) in pasture receive pollen gene flow from distant sources and are reproductively active. In *D. excelsa*, such long distance pollen-mediated gene flow is possible because an alien pollinator, African honeybees, visit isolated trees in pastures although native pollinators do not

(Dick 2001). Nevertheless, Sork *et al.* (2002) found that for the wind pollinated temperate species *Quercus lobata*, very few trees contribute to the pollen pool of individuals in a severely fragmented habitat.

4 Conclusions

Few studies have addressed the questions raised by theoretical population genetics concerning the consequences of habitat fragmentation for plant species. The literature is especially poor for trees species (Adams 1992). Although patterns of genetic diversity within and among populations have been well documented in many species, further empirical evidence is needed to understand the effects of habitat fragmentation on inbreeding depression and gene flow in remnant populations. All too often, interpretations of increased inbreeding depression or limited gene flow are drawn from the population genetic structure of fragmented populations, and rely on theoretical assumptions without actually testing them. One assumption often made is that increased homozygosity necessarily leads to reduced fitness in fragmented remnants. However, both theoretical and empirical evidence suggest that the reality is rather more complicated (Keller and Waller 2002). Since theoretical expectations are that reduced gene exchange between isolated remnants will lead to increased genetic differentiation among them, the latter is therefore often interpreted in terms of limited gene flow. However, recent empirical evidence suggests that this may not be true (Aldrich and Hamrick 1998). These issues need to be critically addressed as recommendations for conservation and management are being made from such interpretations.

Interestingly, the major conclusion that can be drawn from the existing literature is that genetic variation in plant populations responds to habitat fragmentation in more varied ways than first expected. No general pattern can be inferred from the current empirical evidence. Studies of natural populations of *A. saccharum* (Fore *et al.* 1992; Young *et al.* 1993) or *S. globulifera* (Aldrich and Hamrick 1998) gave insights into changes in gene flow dynamics more complex than expected or unpredicted by theory. Most recent direct estimations of pollen-mediated gene flow into recently fragmented stands of tropical tree species suggest that isolated trees and remnants are

part of a wide reproductive network. Considering the complexity and diversity of plant mating systems and the idiosyncrasies of gene flow in plant species, this stresses the need for the collection of empirical data for a wider range of species and in situations when fragmentation is of long standing.

IV The Southern Uplands of Scotland: a deforested landscape

1 Study site: Moffat Dale

The Southern Uplands of Scotland is one of the regions of Britain most lacking in natural woodlands and although areas have been planted with non-indigenous conifer species, it is highly deforested. For instance, in Peeblesshire, only 0.1% of the land-area has long-established natural woodlands (Newton and Ashmole 1998).

Moffat Dale is a wide valley, located in the central part of the Southern Uplands of Scotland (Figure 1.1), in Dumfries and Galloway (N55° 24' W3° 20') where the Moffat-Tweedmuir Hills rise to 800 m (Wildwood Group of the Borders Forest Trust 2000). Adjacent valleys to Moffat Dale, such as Black Hope and Carrifran whose streams drain into Moffat Water (Figure 1.1) have steep slopes that have been excavated during the last glacial period (Figure 1.2a). Carrifran is bordered on the north-east by White Coomb which, rising at 821 m, is one of the highest points in the south of Scotland.

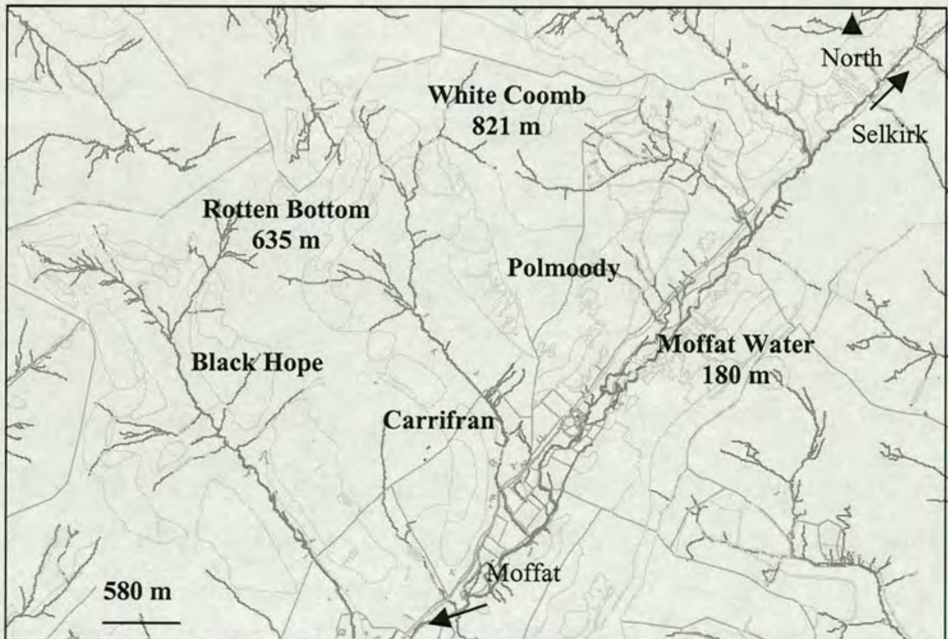


Figure 1.1: Location of Moffat Dale in the Southern Uplands of Scotland.
Moffat Dale at the bottom of which runs Moffat Water, is situated in Dumfries and Galloway, (N55° 24' W3° 20'). Streams of side valleys in Moffat Dale such as Carrifran and Black Hope drain into Moffat Water.

The landscape of Moffat Dale and its surroundings is a patchwork of bare rough grazings and dense conifer plantations (Figure 1.2b). Native woodland cover is very limited. native woodland remnants are very small (comprising between one and a few tens of individuals), confined to ravines which are inaccessible to grazing animals (Figures 1.2a and 1.2c). Trees commonly have a stunted, deformed shape, as a result of both intensive grazing and exposure to strong winds. In Carrifran, a recent survey indicates that the remaining woodland fragments are composed of only a few species, *Fraxinus excelsior*, *Betula pubescens*, *Sorbus aucuparia*, *Corylus avellana*, *Crataegus monogyna*, *Salix spp.*, *Prunus padus*, and *Ilex aquilifolium* (Wildwood Group of the Borders Forest Trust 2000).

Evidence from pollen records indicates that, originally, woodland cover in the area would have been extensive (Tipping 1998) but anthropogenic activities, which can be dated back 6000 years (Wildwood Group of the Borders Forest Trust 2000), have greatly altered the landscape. Indeed, an Early Neolithic flatbow dated at 4040-3640 BC was recently found, preserved in the peat at Rotten Bottom, at an elevation of 600 m asl north-west of Carrifran (Figure 1.1). Pollen analysis reveals that land use for grazing has been significant from very early times (Wildwood Group of the Borders Forest Trust 2000). Sheep farming was first introduced in the area in the 12th century, expanded rapidly between the 15th and 18th century, and has been intensive ever since. The last 60 years, following the second world war, have seen the establishment of non-indigenous dense conifer plantations. The Polmoody plantation (Figure 1.1) covers a large area of Moffat Dale, east of Carrifran. In Polmoody (and elsewhere), native woodland remnants struggle to persist in ravines which have been left clear of conifers (Figure 1.2d).

2 The Carrifran Wildwood Project and reforesting southern Scotland

Context. In recognition of the growing concerns caused by the current massive destruction of forest habitats world-wide, in particular of tropical rainforests, a group of residents of the Southern Uplands of Scotland initiated a local project to restore the original woodland. The initiative gave birth to the Wildwood Group of the

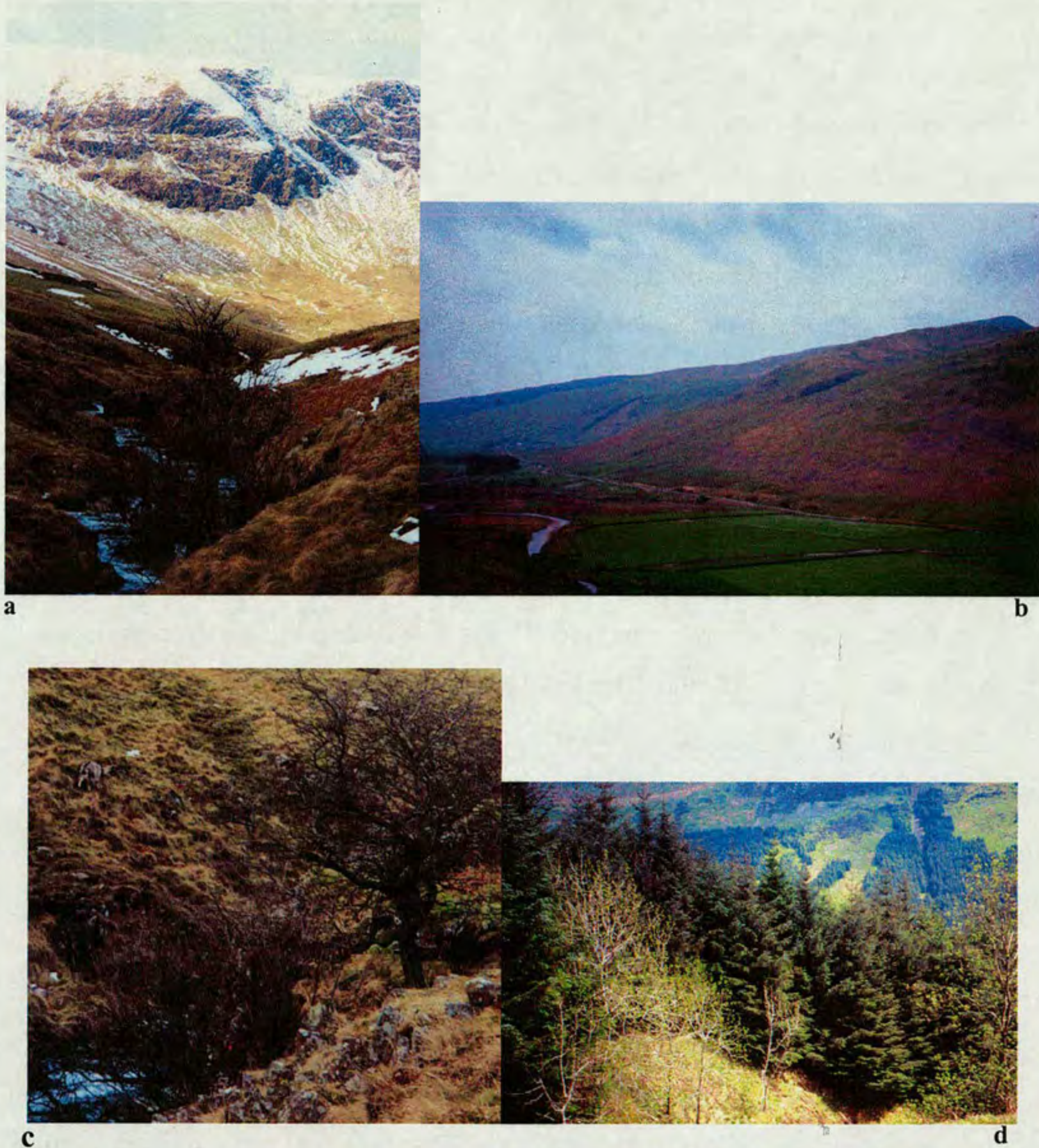


Figure 1.2: The deforested landscape of Moffat Dale.
a) Native woodland remnants at the bottom of the Carrifran valley. b) Moffat Dale looking down towards Moffat. c) *S. aucuparia* remnants in Carrifran. A feral goat is pictured grazing (top left corner). d) A remnant of 12 *F. excelsior* trees situated along a stream within Polmoody conifer plantation. In the background, view of the opposite side of Moffat Dale.

Borders Forest Trust, founded in 1995. The objectives of the Wildwood group are described in their mission statement (Wildwood Group of the Borders Forest Trust 2000):

> *"The Wildwood Project aims to re-create, in the Southern Uplands of Scotland, an extensive tract of mainly forested wilderness with most of the rich diversity of native species present in the area before human activities became dominant. The woodland will not be exploited commercially and human impact will be carefully managed. Access will*

be open to all, and it is hoped that the Wildwood will be used throughout the next millennium as an inspiration and an educational resource.,,

The main ecological objective of the Wildwood Group is to recreate a prehistoric woodland, an *original- natural* woodland, defined by Peterken (1996) as a woodland comprising all species (and only those) present prior to human impact (Ashmole 1998). After two years of fundraising, the Carrifran valley, off Moffat Dale, was purchased and forest restoration began in 2000. Such a forest restoration project is unusual in the way it originated from a local initiative. It is also one of the largest scale native woodland restoration programmes in Britain (a watershed of 650 ha) (Wildwood Group of the Borders Forest Trust 2000).

Forest restoration at Carrifran. How such prehistoric woodland may be restored at Carrifran has been described in great details in a management plan compiled by the Wildwood Group of the Borders Forest Trust (2000; available at URL: http://www.carrifran.org.uk). The species composition of the woodland has been defined using information provided by pollen analysis, historical records and vegetation surveys of the remaining fragments.

An immediate aim of the project is the establishment of native woodland cover on a large proportion of the valley, by direct seeding from local remnants and by encouraging natural regeneration. To such end, large areas of the valley have been fenced to prevent deer, goat and sheep grazing. However, considering the extent of deforestation, planting over ten years will be the favoured mean of reforestation. In order to obtain a planting stock best adapted to the local environmental conditions, seed have been collected as locally as possible and within 20 km of Carrifran (Wildwood Group of the Borders Forest Trust 2000). In particular, seed were collected from *S. aucuparia* and *F. excelsior* remnants in Carrifran or in the neighbouring valleys, where remnants are more extensive (>20 trees) than for most other native species.

V Study species: same habitat, contrasting pollen and seed dispersal

1 *Sorbus aucuparia* L., rowan or mountain ash

Taxonomy. S. aucuparia L has recently been described in the biological flora of the British Isles (Raspé *et al.* 2000b). It is a diploid species (2n=34), belonging to the Rosaceae, subfamily Maloidae which includes *Malus* species (Apples). The genus *Sorbus* includes more than 80 different species which differ in having pinnate (rowans) or simple (whitebeams) leaves (Mitchell 1994) although there are many hybrids between them. Indeed, *S. aucuparia*, *S. torminalis* Crantz, *S. chamaemespilus* Crantz and *S. aria* Crantz form an hybrid complex (Raspé *et al.* 2000b; Oddou-Muratorio 2002). In particular, hybridisation between *S. rupicola/S. aria* and *S aucuparia* has given rise to two microspecies, *S. arranensis* and *S. pseudofennica*, endemic to the Isle of Arran and to another taxon morphologically similar to *S. mennishii sensu lato* (Robertson 2000).

Morphology. S. aucuparia L. can be recognised by its silver grey smooth bark (Rameau *et al.* 1989). It is a small, insect pollinated, bird dispersed, tree, growing on average to 15 m and up to 20 m in height. The twigs are pubescent at first, then glabrous (Mitchell 1994; Raspé *et al.* 2000b) with long ovoid buds (Figure 1.3). The leaves are pinnate with pairs of leaflets (Raspé *et al.* 2000b). The leaflets are sessile, finely and densely pubescent beneath at first, but are later nearly glabrous (Mitchell 1994). The terminal leaflet more or less equals the lateral ones but is never larger (Figure 1.3).

The inflorescence is a compound corymb, dense, with usually approximately 250 small white hermaphrodite flowers (Figures 1.3 and 1.4). The fruits are fleshy pome-like berries that vary in colour from scarlet red to orange (Raspé *et al.* 2000b). The ripe fruit contains a variable number of rather soft small seeds (Figure 1.3). Usually one to five full sized seeds are produced per fruit but up to eight seeds in a 4-celled pome can be observed (Raspé 1998; pers. obs.).

Figure 1.3: Plate showing drawings of *Sorbus. aucuparia* by Thomé (1885) © 1999, Stüber online library (URL: http://www.biolib.de/).
A inflorescence (1-4 details of flower). B berries (5-6 section of a berry showing seeds).

Reproductive biology and phenology. Vegetative buds break in early spring and *S. aucuparia* flowers in May and early June (Raspé *et al.* 2000b). Flowers attract generalist insect pollinators, mainly diptera (Raspé 1998) but also beetles and bees, taking both pollen and nectar (Raspé *et al.* 2000b).

Figure 1.4: ***S. aucuparia*** **tree in blossoms.**

Although flowers are hermaphrodite, *S.aucuparia* is completely outcrossed thanks to a single locus gametophytic system of self-incompatibility also identified in other Rosaceae, for instance in *Crataegus monogyna* (Raspé and Kohn 2002). Experiments involving controlled pollination (Raspé 1998) showed that self-pollination results in virtually no fruit set and very low fruit initiation. However, selfing has been reported in Swedish populations of *S. aucuparia* (Sperens 1996, cited in Raspé *et al.* 2000a) suggesting a possible breakdown of the incompatibility system in the northern range of the species.

In Great Britain, fruits ripen in late August and are eaten by birds (Snow and Snow 1988) but also mammals (Grime *et al.* 1988; pers. obs.). Blackbirds are the main dispersers in towns and rural areas but migratory birds such as fieldfares and redwings will also disperse the seeds during their autumn migration (Mitchell 1994). *S. aucuparia* shows a large individual variation in fruit production between years and

also suffers from attack by the apple fruit moth (*Argyresthia conjugella*). As much as 89% of the fruits can be destroyed (Edland 1975).

Geographical distribution and ecology. S. aucuparia L. is a short-lived tree. It has a maximum life span of about150 years and seed bearing begins at about 15 years of age. *S aucuparia* is native to most of Europe and is widespread across the region (For map of distribution in Europe, see Raspé *et al.* 2000b). It can grow at altitudes nearly as high as 1000 m in Scotland, higher than any other tree (Mitchell 1994) and occurs at altitudes up to 2000 m in France (Rameau *et al.* 1989) and 800 m in Sweden (Raspé *et al.* 2000b). It is absent from Europe only in the Azores, Balearic Islands, Crete, Faroes, Sardinia, Spitzbergen and Turkey but occurs in Morocco and North Asia minor. It has been introduced in North America as an ornamental tree (Raspé *et al.* 2000b) and is commonly planted in parks and gardens all over Britain (Mitchell 1994). The distribution of *S. aucuparia* in Europe is thought to be limited by a combination of poor drought tolerance, adaptation to short growing seasons and cold requirement for bud burst (Raspé *et al.* 2000b).

S. aucuparia L. has a wide topographical range from flat lowland sites to rocky mountain slopes. It grows commonly on cliffs inaccessible to grazing livestock up to 650 m (Raspé *et al.* 2000b) but above that level is usually found as seedlings and small saplings displaying a stunted morphology. Seed may also establish in the trunk of other trees and may grow to maturity (pers. obs.). *S. aucuparia* regenerates entirely by seed and although no persistent seed bank has been detected within the Sheffield region (Grime *et al.* 1988), Hill (1979) suggested that seed have a considerable longevity in the soil. Seedlings and saplings are very tolerant to shade and are mainly restricted to wooded sites (Raspé *et al.* 2000b). Although *S. aucuparia* is a stress tolerant competitor, its establishment in woodlands appears to be adversely affected by the presence of grazing stock or game (Raspé *et al.* 2000b). In lowland habitats, *S. aucuparia* behaves as a hardy pioneer or post-pioneer species and is usually succeeded by late successional species (Raspé *et al.* 2000a). At higher altitude, however, it is one of the few species to maintain the tree habit where it forms part of the late successional vegetation.

2 *Fraxinus excelsior* L., common ash

Taxonomy. *F. excelsior* L. (2n=46), is among 50 species of the genus *Fraxinus*, of the family Oleaceae. The Oleaceae is a medium size family of approximately 600 species and 25 genera (Wallander and Albert 2000) containing many fragrant flowering shrubs such as lilac and jasmine (Mitchell 1994). A third of these are insect pollinated and occur in Asia while the remainder are wind pollinated and occur in North America and Eurasia (Wallander 2001). Within the genus *Fraxinus*, three sections are distinguished (*Diptaleae*, *Ornus* and *Fraxinus*). *F. excelsior* belongs to the section *Fraxinus*, which includes 23 wind pollinated species characterised by having apetalous flowers in inflorescences that emerge from lateral buds before the leaves of the terminal bud begin to expand (Figure 1.5; Wallander 2001).

Morphology. *F. excelsior* is a tall, wind pollinated, wind dispersed, tree. It can reach an extreme height of 43m, but is usually nearly 12-18m high (Wardle 1962). The bark is grey, smooth in younger trees, becoming furrowed in mature trees (Mitchell 1994). In winter, *F. excelsior* is easily recognised by its black dormant buds (Figure 1.5), up to 10 mm long. These open into pinnate opposite leaves, with up to 13 sessile opposite glabrous leaflets (Figure 1.5; Wardle 1962).

F. excelsior produces small hermaphrodite flowers that lack calyx and corolla (Figure 1.5). Hermaphrodite flowers consist of one ovary and two stamens but in some flowers one sex can be aborted. Flowers may be male, with only functional stamens, or female, with a functional pistil and sometimes rudimentary stamens (Wallander 2001); Figure 1.6a). There are two locules in a syncarpous ovary, with two ovules each (Figure 1.5) but the mature fruit is dry and elongated into a wing (i.e. samara, Figure 1.5), containing only one seed (Wallander 2001).

Figure 1.5: Plate showing drawings of *Fraxinus. excelsior* by Thomé (1885) © 1999, Stüber online library (URL: http://www.biolib.de/).
A male inflorescence (2 male flower). B hermaphrodite inflorescence (1 hermaphrodite flower, 3-4 sections of ovary). C fruits (5-6 seed in samara, 7 seed, 8-9 sections of seed showing embryo).

Figure 1.6: *F. excelsior* in Spring. a) Detail of a female inflorescence. b) Flushing of vegetative buds in May in Moffat Dale.

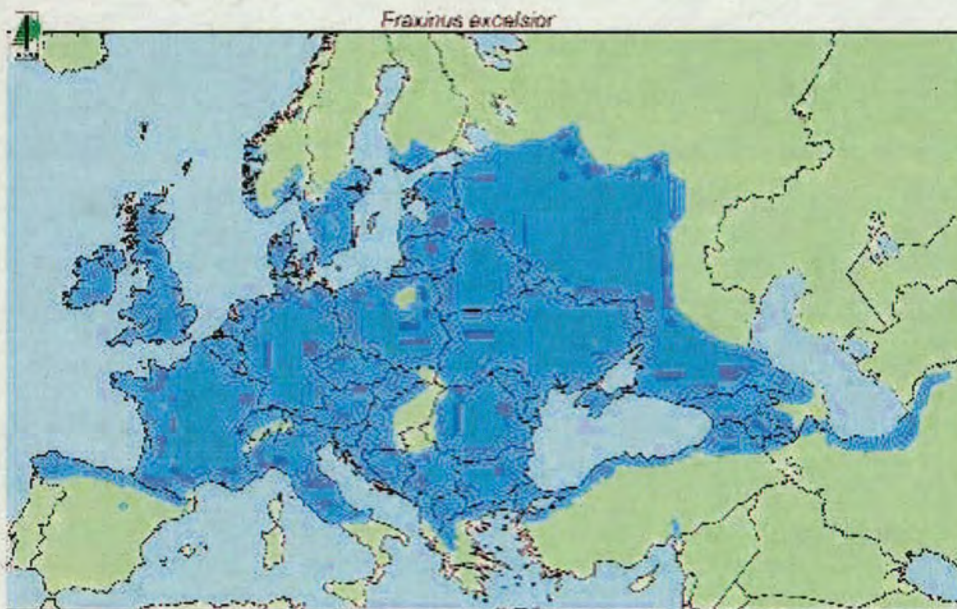


Figure 1.7: Distribution map of *F. excelsior* in Europe. Map obtainded from URL: http://www.ipgri.cgiar.org (EUFORGEN Noble Hardwood Network, Heuertz 2002).

Reproductive biology and phenology. F. excelsior has a complex polygamous sexual system (Wallander 2001). Sex expression varies at the floral, inflorescence and tree level as well as between years, and individuals may be classified into a continuum from purely male to purely female with a whole range of hermaphrodite intermediates (Wallander 2001). Inflorescence buds open by mid-April (Wardle

1962) or early May in Scotland (pers. obs.) and vegetative buds open soon after (Figure 1.6).

Flowers are wind pollinated (Wardle 1961) but attract insects (pers. obs.). Recent evidence from controlled crosses show that self-pollination in hermaphrodites is possible (Wallander 2001; Morand-Prieur *et al.* 2003) and that *F. excelsior* is at least partially self-compatible with a low selfing rate found in controlled experiments involving pollen competition (Morand-Prieur *et al.* 2003). However, selfing in hermaphrodite flowers may be avoided by protogyny (Wallander 2001).

F. excelsior is a prolific seed former, trees can bear up to 10kg of fruits (Wardle, 1962; Marigo *et al.* 2000). Fruits are fully grown by the beginning of the summer following pollination and fall from September onwards but most seeds may not be dispersed before the following spring (Wardle 1962). Irregular fruiting between years has been observed and synchrony among individuals suggests that it may be induced by some environmental components (Tapper 1996; Heuertz 2002).

Geographical distribution and ecology. F. excelsior has a life-span of normally 200 to 300 years, however some individuals may live up to 400 years (Marigo *et al.* 2000). Regular fruit bearing may be observed as early as 16 years of age (Hulden 1941). *F. excelsior* is widespread in temperate Europe (Marigo *et al.* 2000) and native throughout the British Isles (Figure 1.7). Generally present in the lowlands in Britain, it grows in the mountains in the Mediterranean region and at greatest altitude in central Europe (Wardle 1962).

Fraxinus species distribution and community patterns have developed in response to heterogeneous environment conditions and a wide range of historical factors including the complex histories of anthropogenic disturbances (Marigo *et al.* 2000). Pollen records suggest that the postglacial recolonisation of Europe from south-eastern refugia was fastest in the period up to 6000 BP, with migration rates of 200 m to 500 m per year (Heuertz 2002). The distribution limits of the species were reached between 6000 and 4000 BP and *F. excelsior* had fully colonised the British Isles

around 4500 BP (Heuertz 2002). The northern limits of *F. excelsior* are probably defined by intolerance of winter cold or lack of a hot summer (Wardle 1962). In southern Europe, the limits are more difficult to draw as *F. excelsior* is progressively replaced by *F. angustifolia* and hybridisation between the species occurs frequently (Marigo *et al.* 2000).

F. excelsior is a mesophilic species that usually thrives on alluvial soils and can also survive strong water deficit on hill slopes. It is a light demanding plant (Marigo *et al.* 2000) and is outcompeted in shaded environment, persisting as a minor component in beech dominated woods (Wardle 1962). *F. excelsior* is ubiquitous with respect to the major ecological gradients, has a wide distribution and is present in various communities (Marigo *et al.* 2000). It is an important pioneer (Wardle, 1962). Seedlings occur wherever the field layer is reduced in density, but are shade tolerant (Marigo *et al.* 2000). Seedlings are highly tolerant to partial submersion but are particularly sensitive to spring/summer flooding (Marigo *et al.* 2000). *F. excelsior* can also persist by vegetative reproduction in a site after disturbance but grazing by goat and sheep are particularly harmful to its regeneration and probably limited its expansion throughout the Middle Ages and 19th century (Marigo *et al.* 2000).

VI Objectives

As discussed in the first part of this introduction, empirical evidence of the detrimental effects that habitat fragmentation may have on population genetic processes is not well documented, especially for temperate tree species. Little research has addressed how habitat fragmentation may affect mating patterns and gene dispersal within and among populations, especially in cases where fragmentation is long-standing.

In Moffat Dale, *Sorbus aucuparia* and *Fraxinus excelsior* occur together in severely fragmented population remnants, that have existed for centuries, as a result of sustained anthropogenic activities. The heterogeneity of the Moffat Dale landscape is such that forest remnants vary in their population size, in their degree of physical isolation and distance to other remnants and in their exposure to grazing. This situation provides a unique opportunity to assess the effects on population genetics of chronic habitat fragmentation in two temperate tree species with contrasting dispersal. In particular, the relative importance of population size, isolation, exposure to grazing, or mode of pollen and seed dispersal and the possibility of a fragmentation threshold may be investigated. I wish to address specifically the following questions:

1- Are chronically fragmented forest remnants genetically depauperate?
2- Is there evidence of inbreeding and preferential mating with neighbouring individuals within forest remnants?
3- Are forest remnants genetically differentiated at selectively neutral loci?
4- What is the pattern of pollen and seed exchange among forest remnants?
5- What are the implications of these findings for the use of fragmented populations in the regeneration and re-establishment of native woodland?

This general introduction is followed by a chapter introducing the general methods used to investigate these issues (Chapter 2). I address mainly how neutral codominant genetic markers may be useful for identifying population genetic processes. The following four chapters, which form the basis of papers to be

submitted for publication, describe the results of the research that has been conducted over the past four years :

In chapter 3, an assessment is made of the genetic structure of mature *S. aucuparia* trees in population remnants using isozymes and chloroplast DNA (cpDNA) markers. This provides a quantification of neutral genetic variation within remnants and describes its distribution among them. The contrasting mode of inheritance of isozymes, coded by biparentally inherited nuclear genes, and the chloroplast genome, maternally inherited in *S. aucuparia* (Raspé 2001) is used to infer the relative contribution of pollen and seed mediated dispersal to historical gene flow among remnants.

As a result of a failure in optimisation of appropriate isozymes and cpDNA markers, nuclear microsatellite markers available for the species were used to describe the patterns of genetic variation within and among remnants of *F. excelsior*. They were also used to estimate current mating patterns for *F. excelsior* within Moffat Dale by fitting the neighbourhood model (Burczyk *et al.* 2002) to open-pollinated progeny arrays. This is the subject of chapter 4.

Chapters 5 and 6 deal with a detailed assessment of contemporary patterns of gene dispersal for *F. excelsior* in the fragmented landscape of Moffat Dale. In chapter 5, several methods of paternity analysis are applied to open-pollinated progeny arrays and compared to describe patterns of pollen-mediated gene flow into forest remnants. In chapter 6, a maximum likelihood-based parentage analysis of newly established seedlings in forest remnants is used to describe patterns of seed dispersal and to quantify pollen and seed-mediated gene flow into remnants.

In a final discussion chapter (Chapter 7), results from the different studies are brought together in a summary and discussed in the context of native woodland restoration in southern Scotland. Issues regarding, in particular, the suitability of collecting seed for planting stock from severely fragmented remnants and in general, the genetic component of forest management are discussed.

Chapter 2 : Using molecular markers to identify historical and contemporary population genetic processes

I Molecular markers: polymorphism and neutrality

The general term of molecular marker refers to discrete traits, genes or proteins coded by genes that follow Mendelian laws of heredity and for which several allelic forms may be identified in populations. Such polymorphism and the factors, evolutionary forces and mating patterns, that influence it may be investigated in the framework of population genetics models. Selectively neutral markers (i.e. genes for which the replacement of an allele by another does not affect individual fitness) are especially useful to assess the influence of migration and mating patterns on the distribution of genetic diversity within and among populations of a species. This chapter aims at introducing, in a rather brief and general manner, the types of marker and the theoretical background of the methods used in this thesis, which are detailed in subsequent chapters in the specific context of the research undertaken.

1 Isozymes

Isozymes are different active forms of a protein, usually enzymes involved in photosynthesis or respiration, that are segregated according to their electric charge by electrophoresis and are revealed by inducing a biochemical reaction catalysed by the enzyme. Isozymes may be encoded by one or more loci, but variation at each form may be equated to variation at a single protein-coding locus at which alleles are codominant (May 1998). Polymorphism at isozyme loci is usually moderately low in natural populations compared to more recently developed DNA markers and the assumption of selective neutrality at loci coding for isozymes has been questioned (Jarne and Lagoda 1996). However, this relatively fast and inexpensive technique has proven a useful tool for analysing genetic variation in natural populations for several decades (May 1998). Therefore, the significant literature available has made it possible to compare genetic variation between species with different life-history traits (e.g. Hamrick and Godt 1996).

2 Microsatellites

Microsatellites are tandemly repeated DNA sequences whose unit of repetition is between one and five base pairs (Jarne and Lagoda 1996). Genomic microsatellites have been found in a wide variety of eukaryotes, and in fact in all the species analysed so far (Jarne and Lagoda 1996; Li *et al.* 2002) but may also be found in the chloroplast genome of plant species. Genomic microsatellites are codominant and inherited in a Mendelian fashion, and are considered neutral (Jarne and Lagoda 1996).They are highly polymorphic in natural populations and as such have become a favoured choice of marker for the power they provide to statistical analyses (Hedrick 1999). High polymorphism at microsatellite loci is generated by exceptionally high mutation rates that are generally explained by polymerase slippage during DNA replication (Li e*t al.* 2002). This suggests that the occurrence of homoplasy (i.e. that alleles that share the same size or sequence are identical by state but are not identical by descent) may be a concern for population genetic analyses although it may not be a problem on short evolutionary timescales (Jarne and Lagoda 1996; Estoup e*t al.* 2002).

3 Chloroplast DNA (cpDNA) markers

The absence of recombination in the chloroplast –haploid- genome means that its structure is highly conserved among plant species. This characteristic has made it possible for consensus, universal primers to be developed (Demesure *et al.* 1995; Dumoulin-Lapègue *et al.* 1997). Although these primer sequences are located in highly conserved regions, they flank non-coding more variable regions (Demesure *et al.* 1995) in which polymorphism can be detected following amplification by polymerase chain reaction (PCR) by restriction fragment length polymorphism (RFLP) or sequencing. Since the chloroplast genome is uniparentally inherited in most plant species (maternally in angiosperms and paternally in gymnosperms), such polymorphism has been most useful in plant population studies to infer, when used in combination with nuclear markers, the relative importance of pollen and seed-mediated migration (Ennos 1994; McCauley 1995).

II Population structure and historical gene flow

1 Hardy-Weinberg principle

The Hardy-Weinberg principle states that (Henry and Gouyon 1998):

> 1-In a diploid, sexually reproducing population of infinite size, and if there is no mutation, selection, migration and random genetic drift, allele frequencies are constant between generations.
>
> 2-If mating occurs at random (i.e. panmixia), then genotype frequencies may be inferred directly from allele frequencies, and are also constant. When considering one locus with two alleles A and a in frequencies p and q in the population, then there is a proportion p^2 of individuals with genotype AA, $2pq$ with genotype Aa and q^2 with genotype aa, referred to as the Hardy-Weinberg equilibrium (for two alleles, Hartl 1988).

The Hardy-Weinberg principle describes a very simple model of population genetic structure, from which departure under the influence of mating patterns or acting evolutionary forces may be investigated in natural populations by estimating allele frequencies at marker loci.

2 Measuring polymorphism

Population genetic diversity is often measured in terms of mean expected heterozygosity per locus under random mating or gene diversity (H_e), which may be defined by the mean over loci of $1-\Sigma p_i^2$ where p_i is the frequency of the ith allele in the population and Σ stands for summation over all alleles (Nei 1978). Estimates of H_e are often compared to the observed heterozygosity (H_o), or proportion of heterozygotes sampled. Polymorphism may also be measured in terms of percentage of polymorphic loci ($P_{\%}$) or allelic richness (A) which may be standardised to a particular sample size (El Mousadik and Petit 1996).

3 Inbreeding coefficient and *F*-statistics

Departure from Hardy-Weinberg equilibrium in natural populations may occur in particular when the hypothesis of panmixia is not verified. Mating patterns influence genotype frequencies at a particular locus by determining whether alleles of homologous genes in the progeny are more likely to be identical by descent (inbreeding) or different (outbreeding). Departure from random mating can therefore be estimated by the probability of two alleles of a locus in an individual (in a given population) being identical by descent and corresponds to the inbreeding coefficient denoted F. This is easily translated in terms of departure from the expected frequency of heterozygotes at Hardy-Weinberg equilibrium in the population such as (Hartl 1988):

$$F=1-(H_o/H_e)$$

Equation 2.1

Where H_o is the observed frequency of heterozygotes in the population and H_e the expected frequency of heterozygotes at Hardy-Weinberg equilibrium.

Natural populations often occur in spatially subdivided units in which it may be difficult to identify reproductive boundaries. Therefore observed frequencies of heterozygotes may depart from those expected at Hardy-Weinberg equilibrium if more than one reproductive unit is considered as one population (even if there is panmixia in each subpopulation). The phenomenon is referred to as a Wahlund effect (Henry and Gouyon 1998).

To account for the spatial structuring of genetic diversity in subdivided populations, Wright (1951) defined a set of parameters (F-statistics) that correspond to the inbreeding coefficient taken at different levels:

F_{is} measures a departure from Hardy-Weinberg equilibrium within sub-populations and corresponds to the inbreeding coefficient F as defined above. It accounts for non-random mating patterns. F_{st} measures the deficit in heterozygotes observed due

to genetic differentiation between sub-populations, that is to say, the degree to which the total population is divided into distinct reproductive units (Wahlund effect) and F_{it} measures the global departure from Hardy-Weinberg equilibrium.

4 Effective number of migrants

Estimation of *F*-statistics only requires sampling of allele frequencies at marker loci (Weir and Cockerham 1984) and therefore they constitute an easily applicable tool to estimate the relative importance of mating patterns and isolation in the structuring of genetic diversity in natural populations. In particular, Wright (1951) has demonstrated that under the assumptions of an infinite island model of migration, at migration-drift equilibrium and when selection (s) and mutation (μ) are negligible before migration (m), F_{st} may be related to the migration rate among subpopulations of size (N) as follows:

$$F_{st} = 1/(1+4Nm)$$

Equation 2.2

The infinite island model of migration (Wright 1951) considers an infinite number of panmictic populations of equivalent size N and linked by a constant flow of migrants (m) equivalent among populations, where mutation brings new alleles into populations (Infinite allele mutation model). Because natural populations necessarily deviate from these unrealistic assumptions, it has been suggested that inferring the number of effective migrants per generation (Nm) from estimations of F_{st} may be misleading (Whitlock and McCauley 1999). However, it has also been suggested that estimates may be accurate within a factor of two (Slatkin and Barton 1989). Furthermore, such an indirect approach may be useful for estimating the cumulative effects of gene flow or for understanding the genetic structure of populations with a history of metapopulation dynamics (Sork *et al.* 1999).

III Estimating contemporary patterns of mating and dispersal

However, in order to investigate how habitat fragmentation may affect contemporary gene dispersal within and among population remnants, it is necessary to use methods that directly estimate gene flow and mating patterns (Sork *et al.* 1999).

Progeny analyses and related methods offer such a tool. Rather than considering averaged allele frequencies in a population, these methods are concerned with the comparison of genotypes between individuals in different generations and provide detailed estimation of individual reproductive success or pollen or seed-mediated dispersal.

1 Paternity analysis

In particular, paternity analyses of progeny arrays or from dispersed progeny of know mothers provide estimates of individual male reproductive success and of pollen movement within populations (Smouse and Meagher 1994), and also of the contribution of immigrant pollen (i.e. pollen-mediated gene flow) to the pollen pool when paternity of an offspring cannot be attributed to an individual within the population (Ellstrand and Marshall 1985). In essence, three approaches to paternity analysis may be distinguished, simple exclusion, maximum likelihood and fractional likelihood.

The simple exclusion approach is straightforward. Knowing the genotype of an offspring and of its mother for a given number of loci, potential male parents are excluded as father if they cannot have produced gametes bearing a haplotype compatible with the genotype of the progeny given the genotype of the mother (Ellstrand 1984). However, this approach has severe limitations (Schnabel 1998), it depends greatly on the level of polymorphism at marker loci as there is no way to discriminate between two male candidate that could have produced gametes bearing the same haplotype and gene flow events can only be inferred when none of the potential male parents within the target population could have sired the progeny. It may therefore critically underestimate the actual number of pollen immigration

events as a certain proportion of foreign gametes may share the same haplotype at the loci observed with locally produced gametes. Such hidden immigration events are referred to as cryptic gene flow (Devlin and Ellstrand 1990).

Likelihood-based methods have been developed to overcome the pitfall of simple exclusion and to allow assignment when several candidate fathers cannot be excluded. While the maximum-likelihood approach, first introduced by Meagher (1986) concentrates on individual reproductive success (categorical assignment), the fractional-likelihood approach concentrates on population phenomena (Devlin *et al.* 1988).

The principal of the maximum likelihood-based paternity analysis is derived from Bayes theorem (Meagher 1986) and consider the statistical likelihood that a given male is the paternal parent, given the genotypes of the offspring and the known maternal parent for a set of genetic markers, as follows:

Let R be the relationship between three individuals B, C, D, and g_i the genotype of individual I.

The likelihood of relationship R_1 that D is the father, given offspring B and known mother C is

$$L(R_1) = T(g_B / g_C, g_D)\, P(g_C)\, P(g_D)$$

where $P(g_i)$ is the population probability of genotype g_i,
and $T(g_B/g_C,g_D)$ is the Mendelian segregation probability from parent to offspring, or transition probability.

The likelihood of relationship R_2 in which C is mother of B but D is an unrelated male is:

$$L(R_2) = T(g_B / g_C)\, P(g_C)\, P(g_D)$$

Hence,

$$\mathbf{LOD}\ (R_1{:}R_2) = \log\ [T(g_B / g_C, g_D) / T(g_B / g_C)]$$

Equation 2.3

Attributing Log of likelihood ratio (LOD) scores to each male allows one to compare the likelihood of their paternity of a given offspring by considering their value relative to a base-point value of no relation. The possible male parent giving the highest LOD score is the most likely male parent for a given offspring and is accepted as the male parent (Meagher 1986).

This method has the advantage of requiring no prior assumption, for instance equal fertility, about the probability of possible males being the father of a given offspring ; it identifies mother-father-offspring triplets (Schnabel 1998) and provides measures of fertility variation within the population. Nevertheless, three major criticisms have been expressed. First, even if the assumptions of the model are met, there is a bias in respect to individuals homozygous at many loci, whose fertility is overestimated. Also, the individual with the highest LOD score is arbitrarily chosen as the likely parent with little regard to the relative likelihood of other possible parents. Finally, the model does not give any answer, in its original form, to tied LOD scores (Devlin *et al.* 1988; Schnabel 1998).

The fractional paternity assignment method (Devlin *et al.* 1988) is very similar to the maximum likelihood method of Meagher (1986) but, instead of attributing paternity to the most-likely male, the method assigns a fraction of the paternity of a progeny to all the non-excluded males, proportionally to a male's likelihood of paternity. Unlike the most-likely method, all progeny are assigned with a parent but some may not be assigned a single father (Devlin *et al.* 1988).

Using Bayes theorem, the conditional probability of male D* being the father of offspring B given mother C, where g_i is the genotype of individual I is :

$$P(MP=D^*/FP=C, O=B)= [T(g_B/g_C,g_{D^*})P(MP=D^*/FP=C)].[\Sigma T(g_B/g_C,g_D)P(MP=D/FP=C)]^{-1}$$

Equation 2.4

where $T(g_B/g_C,g_D)$ is the transition probability of offspring to parents and P(MP=D/FP=C) is the prior probability of paternity, encompassing all the ecological and genetic characteristics of the population that can make one individual more or less likely to be parent.

Under the assumption that P(MP=D/FP=C) is constant for all D (for justification, see Devlin *et al.* 1988), then the conditional probability of D* being the father of B given mother C is the likelihood ratio of paternity versus non paternity for male parent D, divided by the sum of all potential paternal likelihood ratios (Devlin *et al.*, 1988):

$$P(MP=D^*/FP=C, O=B)= T(g_B/g_C,g_{D^*}). [\Sigma T(g_C/g_C,g_D)]^{-1}$$

Equation 2.5

Although both methods did not originally account for the potential of cryptic gene flow, they have recently been refined (e.g. Marshall e*t al.* 1998; Nielsen *et al.* 2001) to estimate the statistical confidence in assignment and to account for the lack of reproductive isolation of natural populations.

2 Parentage or parent-pair analysis

Parentage analyses of dispersed progeny deal with the more complex situation of identifying parent-offspring relationships when neither the maternal parent nor the paternal parent are known (Meagher and Thompson 1987). They provide estimates of gene dispersal that are a composite of pollen and seed-mediated dispersal, and the two processes cannot be distinguished on a pure genealogical basis (Dow and Ashley 1996).

Although it is possible to reconstruct the parentage of progeny of unknown mother and father using a similar exclusion technique to the one described above (e.g. Dow and Ashley 1996), candidate parents must not only match the genotype of the considered offspring but must also match genotypes as a pair for assignment to be possible.

Meagher and Thompson (1986) have developed a maximum likelihood-based method of parentage assignment for natural populations based on the principle of attribution of LOD scores in which the emphasis is placed on first identifying most-likely mothers and fathers for a given individual and only then identifying the most-likely parent pairs (among most-likely parents). This is justifiable as there is a strong correlation between single-parent and parent-pair likelihoods for a given set of offspring and candidates (Meagher and Thompson 1986).

Such a maximum-likelihood approach to parentage assignment has recently been refined to account for non-isolation of the population (Gerber *et al.* 2003, approach detailed in Chapter 6). Methods of parentage analysis have rarely been applied to natural populations. One objective of this thesis will be to test how such methods can provide useful information on individual reproductive success and patterns of contemporary pollen and seed-mediated gene flow in fragmented populations of trees.

Chapter 3 [1]: GENETIC EFFECTS OF CHRONIC HABITAT FRAGMENTATION ON TREE SPECIES: THE CASE OF *SORBUS AUCUPARIA* L. IN A DEFORESTED SCOTTISH LANDSCAPE

Abstract

Sustainable forest restoration and management practices require a thorough understanding of the influence that habitat fragmentation has on the processes shaping genetic variation and its distribution in tree populations. Genetic variation at isozyme markers and chloroplast DNA (cpDNA), analysed by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP), was quantified in severely fragmented populations of *Sorbus aucuparia* (Rosaceae) in a single catchment (Moffat) in southern Scotland. Remnants maintain high levels of gene diversity (H_e) for isozymes (H_e=0.195) and cpDNA markers (H_e=0.490). Estimates are very similar to those from non-fragmented populations in continental Europe, even though the latter were sampled over a much larger spatial scale. Overall, no genetic bottleneck or departures from random mating were detected in the Moffat fragments. However, genetic differentiation among remnants was detected for both types of marker (isozymes Θ_n=0.043, cpDNA Θ_c=0.131; G-test, p-value<0.001). In this self-incompatible, insect pollinated, bird dispersed tree species, the estimated ratio of pollen flow to seed flow between remnants is close to one (r=1.36). Reduced pollen-mediated gene flow is a likely consequence of habitat fragmentation, but effective seed dispersal by birds is probably helping to maintain high levels of genetic diversity within remnants and reduce genetic differentiation between them.

Keywords: *Sorbus aucuparia* L., isozymes, cpDNA, *F*-statistics, pollen versus seed migration, fragmented landscape.

[1] Bacles, C.F.E., Lowe, A.J. and Ennos, R.A. (2004) Genetic effects of chronic habitat fragmentation on tree species: the case of *Sorbus aucuparia* in a deforested Scottish landscape. *Molecular Ecology* **13**, 573-584. presented in Appendix B.

I Introduction

Extensive anthropogenic use of natural resources to meet the needs of the ever expanding human population has led to growing concern among environmentalists about the detrimental effects of deforestation and habitat fragmentation. A substantial body of literature has documented the influence of deforestation on community and ecosystem processes (Wilcove *et al.* 1986; Robinson *et al.* 1992; Benitez-Malvido and Martinez-Ramos 2003). However, only a few studies have empirically addressed the genetic consequences of habitat fragmentation on population structure in, mainly tropical, forest tree species (e.g. Aldrich *et al.* 1998; White *et al.* 1999; Aguirre-Plantier *et al.* 2000; Dick, 2001). In addition, little effort has been made to integrate genetic factors into forest restoration and management programmes (Ennos *et al.* 1998).

Habitat fragmentation decreases the size of and potentially increases isolation between forest fragments (Andren 1994). It also reduces the availability of suitable colonisation sites (Wilcox and Murphy 1985). Theoretical predictions are that habitat fragmentation will reduce genetic variation within populations and increase interpopulation genetic differentiation, affecting population viability in the short and in the long term (Young *et al.* 1996). These effects are mainly due to an increase in the level of inbreeding and genetic drift in remnants associated with small census size and reduced gene flow between fragments (Couvet 2002). Loss of genetic variation (Dayanandan *et al.* 1999; White *et al.* 1999) and increased inbreeding (Fuchs *et al.* 2003) have been observed for several tropical tree species. However, there is growing empirical evidence for enhanced gene flow between isolated trees (Young *et al.* 1993; White *et al.* 2002; Dick *et al.* 2003). It appears that the effects of habitat fragmentation on tree population dynamics are more varied and complex than first expected (Aldrich and Hamrick 1998). Only addressing these issues for a wide range of species and ecological conditions will provide the information needed to formulate sustainable management practices, especially for temperate species, for which little data is available.

The objectives of the present study are to assess the extent of genetic structure between remnants of a temperate tree species, *Sorbus aucuparia* in the southern Uplands of Scotland. In particular, this chapter addresses the amount of genetic variation remaining within remnant populations, the extent of differentiation between remnants, and the inferred levels of historical gene flow occurring among remnants.

In southern Scotland, forest cover is extremely limited with, in some instances, less than one percent of native woodland remaining (Newton and Ashmole 1998). Major human impact in this area has been estimated to have begun as early as 6000 years ago (Wildwood Group of the Borders Forest Trust 2000). Ancient land use for pasture and much more recent implementation of conifer plantations have greatly fragmented the landscape. Many native tree species, including *S. aucuparia*, are now confined to steep riversides not accessible to grazing herds. Remnants are very small, comprising typically ten to 30 mature individuals with very little to no evidence of natural regeneration in grazed areas. Remnants are separated by at least hundreds of meters and may be isolated from any other remnant by a distance of more than 1 km. As a result of continuous grazing, trees are shrub-like, undersized, and display a convoluted shape. There is an immediate need to assess the population genetics of these forest fragments as they are being considered for use in seed collection to restore native woodlands in Scotland.

S. aucuparia L. (Rosaceae, Maloideae), rowan or mountain ash, is a small tree, native to most of Europe. In southern Scotland, it grows commonly along river beds on rocky mountain slopes and cliffs. At lower altitude, the species behaves as a hardy pioneer or post pioneer and is usually succeeded by late successional species (Raspé *et al.* 2000a). At high altitude, it is one of the few species to maintain the tree habit where it forms part of the late successional vegetation. It can grow at altitudes up to 1000 m in Scotland, higher than any other tree (Mitchell 1994). Seed bearing begins at 15 years of age and the maximum life span of *S. aucuparia* is 150 years. Small, hermaphrodite flowers, grouped into inflorescences are pollinated mainly by flies but also attract beetles and bees (Raspé *et al.* 2000b). Selfing is prevented by a gametophytic system of self-incompatibility (Raspé and Kohn 2002). The species

produces fleshy fruits eaten by birds, such as thrushes, blackbirds and starlings, which are its main seed dispersers (Snow and Snow 1988).

Previous studies have shown that European populations of *S. aucuparia* ranging from Finland to the Pyrenees maintain high gene diversity within population and low inter-population differentiation at isozyme markers compared to other temperate tree species (Raspé and Jacquemart 1998). CpDNA variation in populations sampled in several regions of France and in Belgium also demonstrate low differentiation and high diversity (Raspé *et al.* 2000a). However, populations of *S. aucuparia* separated by a few kilometres, in Belgium, display atypically high genetic differentiation at isozyme markers (Raspé and Jacquemart 1998) which was explained by a founder effect in these recently colonised populations (Raspé and Jacquemart 1998).

Isozymes and cpDNA markers are used to quantify neutral genetic variation and to assess interpopulation genetic differentiation between *S. aucuparia* remnants. The contrasting modes of inheritance of isozymes, coded by biparentally inherited nuclear genes, and the chloroplast genome, maternally inherited in *S. aucuparia* (Raspé 2001), is especially helpful to infer the relative importance of pollen and seed migration in historical gene flow processes (Ennos 1994). Previous studies describing the genetic structure of non-fragmented European populations will provide a useful comparison to assess the extent to which habitat fragmentation has altered patterns of genetic diversity in *S. aucuparia* remnants in southern Scotland.

II Material and methods

1 Study site and populations

The study site is located 80 km south of Edinburgh in Dumfries and Galloway (N55° 23' 51" W3° 19' 50"). The study focuses on a single catchment (Moffat Dale) that has been sampled for *Sorbus aucuparia*. Eight populations were defined within the Carrifran valley and its immediate surroundings (Figure 3.1). Remnants are typically small in size (10-30 individuals, Table 3.1), are separated by a distance of at least 300 m and are on average 1.5 km apart. Trees tend to grow on steep river sides, away from grazing herds, along with other native species such as *Fraxinus excelsior*, *Corylus avellana*, *Betula pubescens* and *Crataegus monogyna*. One remnant, in Swine Cleuchs, is a pure stand of ten trees, situated within a 30-year-old conifer plantation. In five of the eight populations, all identified mature trees were sampled. In the two larger remnants, Spoon Burn and Whitewells (Figure 3.1), a sample of 19 and 20 randomly distributed individuals was taken respectively. In Firthope Burn, tiny shrub-like individuals hang onto very steep cliffs in poor conditions and only two individuals (in a remnant of less than ten trees) were accessible for sampling, bringing the total sample to 123 individuals. The sampling strategy reflects the patchy nature of the landscape (Table 3.1). *S. aucuparia* remnants sampled vary in size, spatial isolation, altitude and accessibility to grazing, thus allowing testing of the effects of long-standing habitat fragmentation.

Table 3.1: location, abbreviated name, landscape characteristics, altitude and size of eight *S. aucuparia* remnants present in Moffat Dale.

Location	Abbreviation	Site characteristics	Altitude	Size
Spoon Burn	BMr	grazed riverside	262 m	50-70*
Carrifran (down)	CDr	grazed riverside	186 m	13
Carrifran (middle)	CMr	grazed riverside	215 m	21
Carrifran (top)	CTr	grazed riverside	332 m	27
Firthope Burn	FBr	grazed riverside	464 m	<10*
Bodesbeck	Kr	grazed pasture	215 m	12
Whitewells	Wr	protected riverbank	170 m	40-50*
Swine Cleuchs	SCr	grazed riverside in conifer plantation	384 m	10

Size expressed as the approximate number of individuals censed. Census sampling in all fragments but *19, 2 & 20 individuals were sampled in BMr, FBr & Wr respectively.

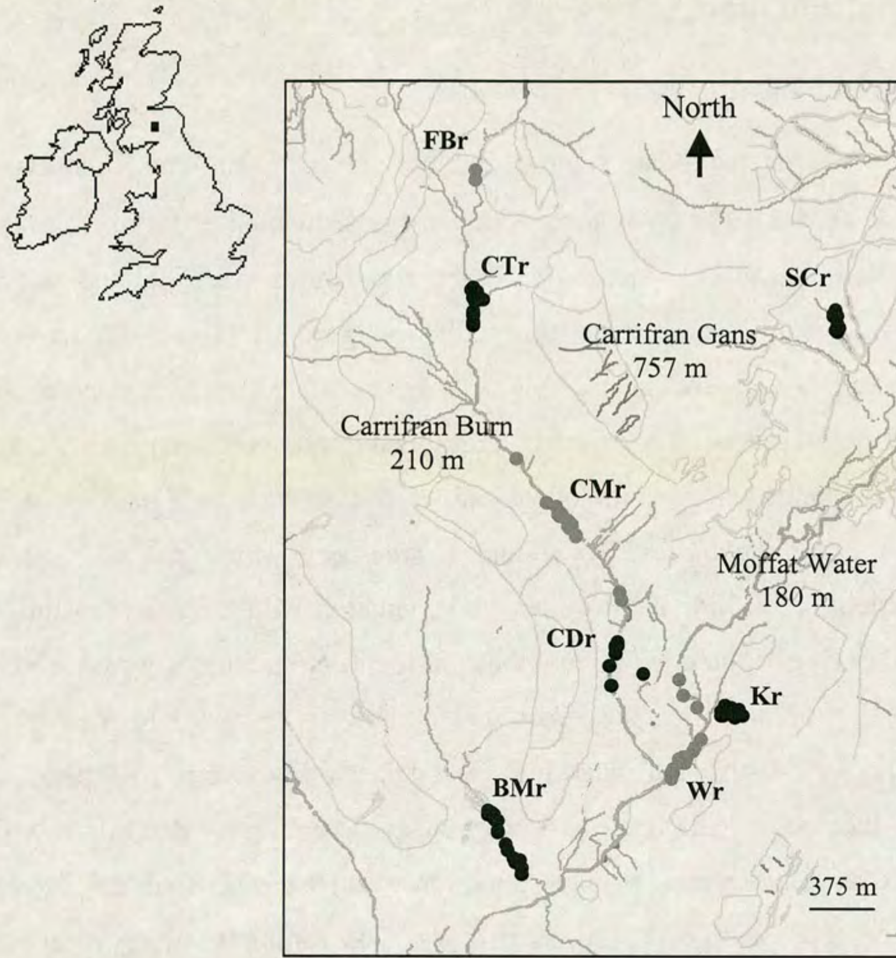


Figure 3.1: Distribution of *S. aucuparia* populations sampled in Moffat Dale.
Trees are grouped into populations as described in Table 3.1. Each dot represents a tree. The background map is a section of Ordnance Survey product Land-Line.Plus-nt11©Crown copyright Ordnance Survey. An EDINA digimap / JISC supplied service.

2 Plant material

Plant material collection was carried out in winter 1999 and spring 2000. Dormant buds were initially sampled to allow optimisation of molecular marker methods. In spring, a resampling was performed, collecting twigs with breaking buds or newly grown leaves from the 123 individuals. The latter were eventually retained for isozyme screening and provided the best material for optimising electrophoresis conditions.

3 Protein extraction and electrophoretic procedure

Protein extraction was performed on fresh material, immediately after collection. Small buds or young leaves were crushed on ice using extraction buffer (Cheliak and Pitel 1984) freshly prepared and stored at -20 °C. Immediately prior to extraction 3.2 µl/ml of 2-βmercaptoethanol and 0.03 g of polyvinyl-pyrrolidone 40 (PVP-40) were added to the extraction buffer. Extracted proteins were absorbed on filter paper wicks and stored at -80°C until electrophoresis. Electrophoresis was carried out at 4°C, on 13 % starch (Sigma) gels using available gel, electrode (Soltis *et al.* 1983; Wendel and Weeden 1989) and staining recipes (Cheliak and Pitel 1984). Choice of isozyme systems and electrophoretic conditions was guided by the developmental work of Raspé *et al.* (1998), based on scorability and reproducibility of results. Conditions were successfully optimised for eight enzyme systems: aspartate-aminotransferase (AAT, EC 2.6.1.1), leucine-aminopeptidase (LAP, EC 3.4.11.1), malate dehydrogenase (MDH, EC 1.1.1.37), menadione reductase, (MNR, EC 1.6.99.2), peroxidase (PRX, EC 1.11.1.7), 6-phosphogluconate dehydrogenase (6-PGD, EC 1.1.1.44), phosphoglucomutase (PGM, EC 2.7.5.1), shikimic acid dehydrogenase (SDH, EC 1.1.1.25). 15 putative loci were identified and phenotypes at AAT, MDH, PRX, 6-PGD and PGM were scored in accordance with zymograms and genetic analysis previously reported (Raspé *et al.* 1998). Eight loci encoding six enzymes found to be polymorphic either in this study (prx-1, pgm-1, pgm-2, pgm-3,sdh-1, aat-1) or previously ((Raspé and Jacquemart 1998), 6-pgd-1, 6-pgd-2), were retained for detailed analysis.

4 DNA isolation and PCR-RFLP analysis of cpDNA

DNA was isolated from 100 mg of frozen leaf or bud material following a standard extraction protocol adapted from Doyle and Doyle (1987) and described by Rendell and Ennos (2002). Extracted DNA was dissolved in 100 µl of Tris-ethylene-diamine-tetraacetic acid (TE, pH 7.6). All individuals were characterised by restriction analysis of three cpDNA fragments amplified by PCR with universal primers (Demesure *et al.* 1995; Table 3.2). PCR reactions were carried out in a GeneAmp® PCR system 9700 (PE Applied Biosystems) following the conditions of Raspé *et al.*

(2000a). PCR products of fragments AS, DT and HK digested by *Hinf*1 (Raspé *et al.* 2000a) were run on 1.8% agarose gels in 0.5X TBE buffer. 7.5 μl of each PCR product was loaded and 500 ng of 1kb ladder (Gibco-BRL Life Technologies) was included as a size marker. Agarose gels were stained with ethidium bromide and photographed with a DOC-IT gel documentation system (UVP) under ultra-violet light.

Table 3.2: CpDNA universal primers used for amplification and restriction enzyme used for digestion of PCR products of *S. aucuparia* samples.

Primer pair[†]	Abbreviation	Restriction enzyme[‡]
trnH-trnK1	HK	*HinfI*
trnD-trnT	DT	*HinfI*
psaA-trnS	AS	*HinfI*

[†]Primer sequences described in Demesure *et al.* (1995).[‡]For digestion conditions see Raspé *et al.* (2000a).

5 Statistical analysis

Genetic diversity. Gene diversity (H_e) per population and overall was estimated using an unbiased estimator for each isozyme and cpDNA locus in FSTAT 2.9.3 (Goudet 2001). Mean H_e estimates over eight isozyme loci at the population and species level as well as other classical estimates of genetic diversity (Percentage of loci polymorphic at the 95% level ($P_{95\%}$) and mean number of alleles observed per locus (A_o)) were obtained using GENETIX 4.01 (Belkhir *et al.* 1998). However, these measures are prone to bias due to variation in sample size (Leberg 2002). To correct for variation in sample size, the rarefaction method described by El Mousadik and Petit (1996) was used and total allelic richness (R_t) was estimated for each isozyme locus in FSTAT 2.9.3. Allelic richness over all isozyme loci per population (R_s) was also computed excluding populations FBr and BMr (because of missing values for one sample at one locus) and including all populations for chloroplast markers. It has been argued that the most relevant criteria for measuring diversity should be allelic richness (Petit *et al.* 1998). The statistic has proven useful for identifying historical processes such as bottlenecks and population admixture (Comps *et al.* 2001)

Departure from random mating. The inbreeding coefficient (F_{is}) was estimated over all isozyme loci for each population and departure from Hardy-Weinberg equilibrium was assessed using FSTAT 2.9.3. Tests for deficit in heterozygotes (*HD*) and excess in heterozygotes (*HE*) were performed by randomising alleles among individuals. Luikart *et al.* (1998) demonstrated that populations that have gone through a severe and recent genetic bottleneck show a transient excess of heterozygotes. To test whether fragmentation in *S. aucuparia* populations has caused a genetic bottleneck, a sign test (Cornuet and Luikart 1996) and Wilcoxon test (Luikart *et al.* 1998) were performed using BOTTLENECK (Cornuet and Luikart 1996). In the analysis, all individuals were pooled together as a single sampled population, i.e. the bottlenecked population, and all enzyme loci are assumed to fit an infinite allele model of mutation (IAM).

Genetic differentiation. Overall inbreeding (f) and genetic differentiation (Θ_n) coefficients were estimated following Weir and Cockerham (1984) for each isozyme locus and overall loci using FSTAT 2.9.3. The genetic differentiation coefficient was also estimated for the cpDNA data set (Θ_c). All estimations excluded population FBr (where only two individuals were sampled). Departure from Hardy-Weinberg equilibrium within samples was tested by comparing the distribution of the *F*is statistic (f) for the observed data set with its distribution for a randomised data set obtained after 10,000 permutations of alleles among individuals within samples. A *G*-test (Goudet *et al.* 1996), based on 10000 permutations of genotypes among samples, was performed to test for population differentiation at isozyme loci and for cpDNA markers. Whenever possible, pairwise tests of differentiation were performed for each population pair (excluding FBr) using the *G*-statistic (Goudet *et al.* 1996) by randomisation of multilocus genotypes for isozymes, and for cpDNA markers using FSTAT 2.9.3.

Independence of loci. Calculation of mean estimates of *F*-statistics assumes independence of loci. Genotypic linkage equilibrium between all polymorphic isozyme loci and between isozyme and cpDNA loci was tested using GENEPOP 3.2a

EX BIBL. UNIV. EDINBURGH

(Raymond and Rousset 1995). Furthermore, homogeneity of F-statistics estimates between isozyme loci was tested by jack-knifing in GENETIX 4.01.

Pollen versus seed migration ratio. Several authors have derived equations to infer the relative rate of pollen and seed mediated gene flow (Petit *et al.* 1993; Ennos 1994; Hamilton and Miller 2003). Ennos (1994) showed that under the assumption of an infinite island model of migration (Wright 1951), at migration-drift equilibrium, a ratio (r) of the amount of migration by pollen (m_p) by the amount of migration by seed (m_s) can be inferred from F-statistics estimated for biparental markers (F_{stb}) and for uniparentally inherited maternal markers (F_{stm}) as:

$$r = \frac{m_p}{m_s} = \frac{(\frac{1}{F_{stb}} - 1)(1 + F_{is}) - 2(\frac{1}{F_{stm}} - 1)}{\frac{1}{F_{stm}} - 1}$$

Equation 3.1: *r*-ratio.

The r-ratio was estimated for sampled populations of *S. aucuparia* by substituting in equation (1) F_{stb} with Θ_n and F_{is} with f estimated overall loci, while Θ_c provided an estimate of F_{stm} as maternal inheritance of the chloroplast genome has been demonstrated in *S. aucuparia* (Raspé 2001).

III Results

1 Genetic diversity at isozyme loci

Over the 123 individuals scored, six polymorphic isozyme loci were identified (prx-1, sdh-1, pgm-1, pgm-2, pgm-3, aat-1) and 15 alleles recorded (Table 3.4). Allele frequencies (*p*) per locus and population are detailed in Table 3.4. The number of alleles at polymorphic loci vary between two (sdh-1, pgm-3, aat-1) and three (prx-1, pgm-2, pgm-3). 6-pgd-1 and 6-pgd-2 were monomorphic in the sampled populations. Among polymorphic loci, H_e varies between 0.073 (aat-1) and 0.549 (pgm-2) whereas R_t varies between 1.145 (aat-1) and 2.409 (pgm-2). A_o varies between 1.200 for population FBr and 2.000 for Kr (Table 3.4). No population is monomorphic for all loci but no population includes all 15 alleles. Population Kr is most polymorphic ($P_{95\%}$=0.667) while FBr is least polymorphic ($P_{95\%}$=0.200). H_e and R_s are quite homogeneous across populations although CTr is most diverse when diversity is measured by H_e and Kr is most diverse when R is considered (Table 3.3). Values of $P_{95\%}$, A_o and H_e were 0.421, 1.392 and 0.163 when averaged over populations and 0.500, 2.000 and 0.195 overall. F_{is} ranges between -0.500 (FBr) and 0.344 (CDr) and a significant deficit of heterozygotes was detected solely for the latter suggesting a departure from random mating in this population. (Table 3.3). No genetic bottleneck was detected for the sampled *Sorbus aucuparia* fragmented populations either with a sign test (*p*-value =0.222) or with a one tailed-Wilcoxon test (*p*-value (*HE*)= 0.055).

Table 3.3: Mean over isozyme loci (including monomorphic loci) gene diversity (H_e), percentage of polymorphic loci ($P_{95\%}$) & number of alleles (A_o) per population & overall for *S. aucuparia* remnants of Moffat Dale.

Population	H_e	$P_{95\%}$	A_o	R_s	F_{is_est}	*p*-value HD	*p*-value HE
CDr	0.174	0.333	1.500	1.513	0.344	<0.01	ns
CMr	0.158	0.333	1.667	1.571	-0.062	ns	ns
SCr	0.143	0.333	1.500	1.556	-0.012	ns	ns
CTr	0.210	0.500	1.833	1.651	-0.081	ns	ns
FBr	0.075	0.200	1.200		-0.500	ns	ns
Kr	0.183	0.667	2.000	1.795	0.180	ns	ns
Wr	0.204	0.500	1.833	1.710	-0.047	ns	ns
BMr	0.153	0.400	1.600		0.203	ns	ns
All	0.163	0.421	1.392				
Overall	0.195	0.500	2.000				

Also shown are inbreeding coefficient estimates (F_{is_est}) & *p*-value for heterozygote deficit (*HD*) and excess (*HE*), allelic richness (R_s is based on a minimum sample size of 4), all estimated using FSTAT 2.9.3 (Goudet 2001). H_e, $P_{95\%}$ and A_o estimated using GENETIX 4.01 (Belkhir *et al.* 1998). Mean includeg loci monomorphic for the sampled populations.

Table 3.4: Allele frequencies (p(i)), gene diversity (H_e) and total allelic richness (R_t) per locus and population for isozyme and cpDNA markers for eight *S. aucuparia* remnants of Moffat Dale.

	Population								
locus	**CDr**	**CMr**	**SCr**	**CTr**	**FBr**	**Kr**	**Wr**	**BMr**	**Overall**
prx-1	N=13	20	10	27	2	10	20	19	121
p(1)	0.500	0.750	0.750	0.537	1.000	0.750	0.825	0.816	0.702
2	0.500	0.200	0.150	0.426	0.000	0.200	0.175	0.053	0.248
3	0.000	0.050	0.100	0.037	0.000	0.050	0.000	0.132	0.050
H_e	0.526	0.404	0.433	0.538	0.000	0.411	0.297	0.325	0.443
$R_t^{\dagger}$									1.860
sdh-1	13	18	5	11	2	7	20	19	95
1	0.731	0.639	0.500	0.682	0.500	0.714	0.500	0.711	0.637
2	0.269	0.361	0.500	0.318	0.500	0.286	0.500	0.289	0.363
H_e	0.404	0.467	0.550	0.455	0.500	0.452	0.511	0.424	0.463
R_t									1.822
pgm-1	9	14	4	26	0	8	19	12	92
1	0.833	0.786	0.750	0.846	NA	0.688	0.632	0.792	0.767
2	0.167	0.179	0.000	0.115		0.250	0.316	0.208	0.190
3	0.000	0.036	0.250	0.038		0.063	0.053	0.000	0.043
H_e	0.306	0.368	0.417	0.274		0.518	0.509	0.356	0.375
R_t									2.126
pgm-2	12	21	10	27	0	10	20	13	113
1	0.625	0.476	0.650	0.370	NA	0.750	0.550	0.654	0.540
2	0.375	0.524	0.350	0.574		0.150	0.200	0.346	0.394
3	0.000	0.000	0.000	0.056		0.100	0.250	0.000	0.066
H_e	0.508	0.512	0.478	0.538		0.422	0.611	0.468	0.549
R_t									2.409
pgm-3	13	17	8	22	2	8	19	0	89
1	0.962	0.971	1.000	0.886	0.750	0.938	0.974	NA	0.944
2	0.038	0.029	0.000	0.114	0.250	0.063	0.026		0.056
H_e	0.077	0.059	0.000	0.206	0.500	0.125	0.053		0.106
R_t									1.250
got-1	13	21	7	26	2	10	20	19	118
1	1.000	1.000	1.000	1.000	1.000	0.900	0.825	1.000	0.962
2	0.000	0.000	0.000	0.000	0.000	0.100	0.175	0.000	0.038
H_e	0.000	0.000	0.000	0.000	0.000	0.189	0.297	0.000	0.073
R_t									1.145
cpDNA	12	21	10	23	2	12	20	19	119
A	0.833	0.381	0.800	0.696	1.000	0.667	0.700	0.789	0.681
B	0.083	0.619	0.100	0.087	0.000	0.083	0.000	0.105	0.168
C	0.000	0.000	0.100	0.174	0.000	0.250	0.300	0.105	0.134
D	0.083	0.000	0.000	0.043	0.000	0.000	0.000	0.000	0.017
H_e	0.318	0.495	0.378	0.498	0.000	0.530	0.442	0.374	0.490
R_2	1.623	1.850	1.737	2.024	1.000	2.017	1.771	1.743	2.021

†Based on a min. sample size of 89, including monomorphic loci. All in FSTAT 2.9.3 (Goudet 2001).

2 Characterisation of cpDNA haplotypes and diversity

Three indels were detected using three primer/restriction enzyme combinations. Single locus haplotype nomenclature was attributed to polymorphic fragments according to Raspé *et al.* (2000a). Combining information from those three indels allowed the identification of four haplotypes in the sampled populations (Table 3.5).

Table 3.5: Description of four cpDNA halotypes found in eight *S. aucuparia* populations of Moffat Dale as a combination of polymorphism observed at three PCR-RFLP fragments.

	Polymorphism observed for each primer/enzyme pair†		
CpDNA haplotype	HK/*Hinf*I	DT/*Hinf*I	AS/*Hinf*I
A	**1**	**1**	**1**
B	**1**	**1**	**2**
C	**1**	**2**	**2**
D	**2**	**1**	**2**

†For polymorphic fragments nomenclature and size, see Raspé *et al.* (2000a).

Haplotype A was most frequent overall (p(A)=0.681, Table 3.3) whereas haplotype D was found in only two individuals, one in population CTr and one in CDr (Table 3.6). Haplotype B was more frequent in CMr than in any other population and haplotype C was not found in populations CMr and Wr (Figure 3.2). Estimates of allelic richness based on a minimum of two genes is highest for population CTr with a population mean of 2.021 (Table 3.4). Gene diversity varies between 0 (population FBr) and 0.530 (population K) and is high overall (H_e=0.490).

Table 3.6 : Distribution of four cpDNA haplotypes found in eight *S. aucuparia* remnants sampled in Moffat Dale.

	CpDNA haplotype				
Population	A	B	C	D	TOTAL
CDr	10	1		1	12
CMr	8	13			21
CTr	16	2	4	1	23
FBr	2				2
Kr	8	1	3		12
SCr	8	1	1		10
Wr	14		6		20
BMr	15	2	2		19
TOTAL	81	20	16	2	119

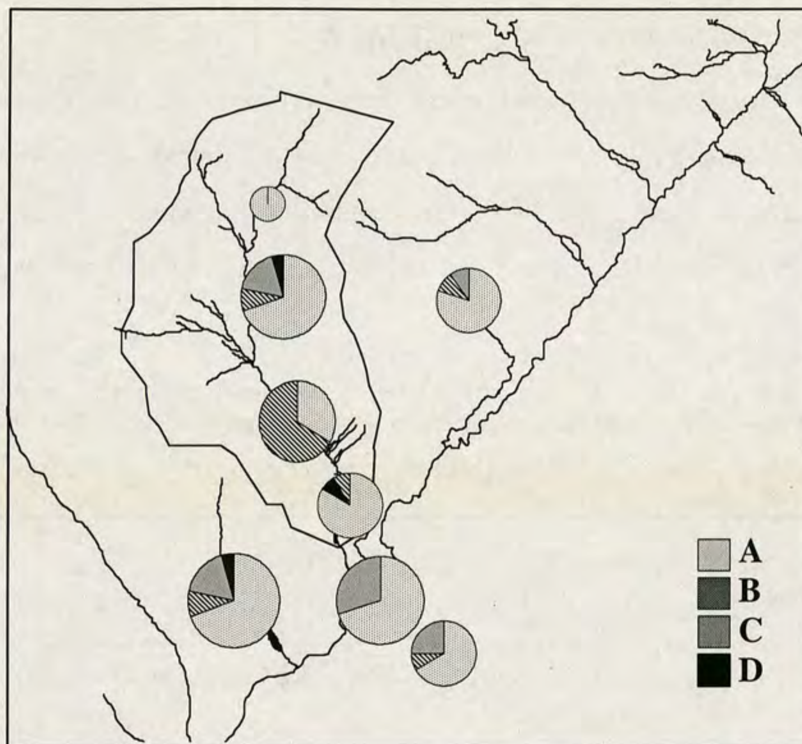


Figure 3.2: Map of the distribution of four cpDNA haplotypes in sampled populations of *S. aucuparia* according to their location in Moffat Dale.
For population names and characteristics, see Figure 3.1 and Table 3.1, haplotype nomenclature refers to Table 3.5. Pie chart diameter reflects relative sampling effort in each population.

3 Population differentiation in *S. aucuparia* remnants

Across isozyme loci, significant genotypic association was found for only one pair of loci across all populations (sdh-1 & pgm-2 ; *p*-value=0.013). Such results do not depart from random association and therefore suggest that loci are at linkage equilibrium. Homogeneity of estimates among isozyme loci was verified as jack-knifed estimates on isozyme loci for f and Θ are within a 95% confidence interval of the jack-knife mean (Table 3.7). Excluding one or other locus did not alter the trend observed in overall estimates and multilocus estimates reported include all six polymorphic isozyme loci. Estimates of f per locus vary between -0.146 (sdh-1) and 0.190 (pgm-1). Departure from Hardy-Weinberg equilibrium was detected at locus pgm-1 only (table 3.8). f estimated over all loci and populations is 0.024 and not significantly different from zero. On average, there is no tendency towards heterozygote excess or deficiency in the sampled populations and therefore no evidence of inbreeding.

Table 3.7: Jack-knifed estimates over isozyme loci of inbreeding coefficient (F_{is}) and of genetic differentiation (F_{st}) for *S aucuparia* samples.

without locus...	F_{is_est}	F_{st_est}
prx-1	0.015	0.034
sdh-1	0.077	0.054
pgm-1	-0.015	0.053
pgm-2	0.023	0.034
pgm-3	0.023	0.046
got-1	0.021	0.041
mean	**0.022**	**0.044**
standard error	**0.061**	**0.018**

Jack-knifing performed in GENETIX 4.01 (Belkhir *et al.* 1998).

Table 3.8: Estimates of *f* and Θ for each of six isozyme loci, for cpDNA markers and over all isozyme loci for *S. aucuparia* populations sampled in Moffat Dale.

locus	smallf (*f*)	test HW[†]	theta (Θ)	test PD[‡]
prx-1	0.060	ns	0.077	*p*-value<0.05
sdh-1	-0.146	ns	0.009	ns
pgm-1	0.190	*p*-value<0.05	0.004	ns
pgm-2	0.028	ns	0.069	*p*-value<0.001
pgm-3	-0.052	ns	0.004	ns
got-1	0.107	ns	0.115	*p*-value<0.01
overall	**0.024**	**ns**	**0.043**	***p*-value<0.001**
cpDNA	**--**	**--**	**0.131**	***p*-value<0.001**

[†]HW: Hardy Weinberg equilibrium. [‡]PD: population differentiation, test not assuming random mating. ns: non significant at 5% level. Estimates obtained and tests performed using FSTAT 2.9.3 (Goudet 2001).

Estimates of Θ are consistently low among loci but population differentiation was detected for loci prx-1, pgm-2 and aat-1 (Table 3.8), whether assuming Hardy Weinberg equilibrium or not. Overall, Θ_n is low (Θ_n=0.043) but significant population differentiation was detected for isozyme loci (*p*-value<0.001). For chloroplast markers, Θ_c is higher (Θ_c=0.131) and genetic differentiation was detected (*p*-value<0.001) among the populations sampled. Tests for pairwise differentiation

show that population SCr is the least differentiated at isozyme loci while differentiation for the chloroplast genome is significant only for pairs including CMr (Table 3.9). These results suggest that overall genetic structure detected at isozyme loci is the result of differences between most of the populations sampled whereas genetic structure detected for chloroplast markers is influenced mainly by population CMr.

Table 3.9: *p*-values for pairwise tests of differentiation between *S aucuparia* remnants sampled in the Moffat Dale.

Population	CDr	CMr	SCr	CTr	Kr	Wr	BMr
CDr		0.06905	0.17381	0.16905	0.02619	**0.00238***	**0.00238***
CMr	0.00476		0.51905	0.08095	0.00476	**0.00238***	0.18333
SCr	0.84048	0.00476		0.02857	0.10000	0.021430	0.70000
CTr	0.42143	**0.00238***	0.71905		**0.00238***	**0.00238***	**0.00238***
Kr	0.21905	0.00714	0.81667	0.88095		0.49048	0.01905
Wr	0.03333	**0.00238***	0.18810	0.20000	0.66667		**0.00238**
BMr	0.49524	**0.00238***	1.00000	0.81190	1.00000	0.20000	

Above diagonal: test based on isozyme loci prx-1, pgm-2 and aat-1. Below diagonal: test based on cpDNA data. *p*-values obtained after 420 permutations in FSTAT 2.9.3 (Goudet 2001). *indicative adjusted nominal level (5%) for multiple comparisons is: 0.002381.

4 Pollen versus seed mediated gene flow

Under the assumptions of the island model, levels of differentiation at biparentally (Θ_n=0.043) and maternally (Θ_c=0.131) inherited markers indicate that migration by pollen is of the same order of magnitude as migration by seed between the *S. aucuparia* remnants in Moffat Dale, r=1.36. The difference increases slightly when taking into account the estimated value for the inbreeding coefficient (f=0.024), r=1.44.

IV Discussion

Habitat fragmentation is expected to have a detrimental impact on the genetic diversity of plant species, but empirical evidence suggests that population responses to fragmentation are idiosyncratic (Young *et al.* 1996). Remnants of *Sorbus aucuparia* in the deforested landscape of the southern Uplands of Scotland have suffered severe long-standing fragmentation associated with anthropogenic land use. This provides an unique opportunity to empirically assess the extent to which fragmentation affects the amounts and patterns of gene diversity in a temperate tree species.

My findings show that fragmented populations of *S. aucuparia* in Moffat Dale maintain high levels of isozyme diversity both within populations (H_{ep}=0.163) and overall (H_{es}=0.195). Values are only marginally lower than those of Raspé and Jacquemart (1998) for unfragmented European populations of *S. aucuparia.* They found that the species maintains high levels of isozyme diversity, H_{ep}=0.212 within populations and H_{es}=0.229 at the species level, compared to species with similar ecological traits.

Estimates of diversity for the chloroplast genome can also be compared with those of unfragmented populations in France and Belgium (Raspé *et al.* 2000a). To obtain comparable estimators of variation, data provided by Raspé *et al.* (2000a) were used to redefine cpDNA haplotype distribution in each population using only information provided by primer/enzyme combination in common with the present study (polymorphic fragments DT1, AS2 and HK3). At the wider scale of France and Belgium, *S. aucuparia* displays higher levels of gene diversity (H_e=0.657) than in Moffat Dale (H_e=0.490). This is also the case for Belgian populations separated by several kilometres (H_e=0.671). Lower diversity is expected in Moffat Dale populations because of the smaller spatial scale and reduced size of individual populations. Nevertheless it is most remarkable that, on the basis of the same primer/enzyme combinations, just as many chloroplast haplotypes (i.e. four) are found within Moffat Dale as in much larger continental populations spread over a much greater area. It can be concluded that despite severe habitat fragmentation, *S.*

aucuparia remnants of Moffat Dale maintain substantial levels of genetic diversity in both nuclear and chloroplast genomes.

S. aucuparia is a self-incompatible species (Raspé and Kohn 2002) which makes it an obligate outcrosser. Populations of Moffat Dale typically have a low inbreeding coefficient (f=0.024) with no evidence for departure from Hardy-Weinberg equilibrium (p-value>0.05). In a fragmented habitat, a departure from random mating may be expected to arise in two ways. There is a theoretical expectation of increased spatial isolation between populations arising from fragmentation which makes mating between relatives (in normally outcrossed species) more likely (Barrett and Kohn 1991). As related individuals are more likely to be genetically similar, biparental inbreeding in a typically outcrossing species will lead to a reduction in heterozygosity (Charlesworth and Charlesworth 1987). A similar effect would also be detected if selfing had become possible. Autogamy can arise if a severe genetic bottleneck leads to a breakdown of self-incompatibility (Reinhartz and Les 1994). Furthermore, in a situation where few compatible mates are available, selection for self-compatibility may be strong (Stephenson *et al.* 2000). Self-compatibility has previously been reported in Swedish populations of *S. aucuparia* (Sperens 1996 cited in Raspé *et al.* (2000b). Based on the results of this study there is no evidence for biparental inbreeding or selfing in the fragmented populations in Moffat Dale. However, inbred progeny are very likely to suffer from inbreeding depression (Charlesworth and Charlesworth 1987) and may not have survived to maturity. If this were the case, inbred individuals and changes in the mating system of *S. aucuparia* would not be detected with the indirect approach taken here.

In order to investigate this issue further, I tested for a genetic bottleneck, but found no strong support for the hypothesis that a genetic bottleneck has accompanied habitat fragmentation in the studied populations of *S. aucuparia* (One tailed-Wilcoxon sign rank test, p-value=0.055). It is important to note that the test relies on the transient excess of heterozygotes that recently bottlenecked populations will exhibit at marker loci (Luikart *et al.* 1998). However, the tests for heterozygosity excess can only detect bottlenecks for a limited window of time after a bottleneck

has occurred (0.2-2.5 times $2N_e$, discussed by Cornuet and Luikart (1996). Therefore, only recent population declines are detectable. Caution must be taken when interpreting the lack of evidence for a genetic bottleneck in the studied populations of *S. aucuparia*. It is possible that populations have gone through a genetic bottleneck not recent enough to be detected. This interpretation would be supported by anecdotal evidence that habitat fragmentation in the area is ancient, since anthropogenic influence on the forest habitat can be dated back 6000 years (Wildwood Group of the Borders Forest Trust 2000).

Populations from Moffat Dale are slightly less differentiated (Θ_n=0.043) at isozyme markers than populations sampled within five different regions across Europe, ranging from Finland to the Pyrenees (G_{stn}=0.060, Raspé and Jacquemart 1998). However, in order to determine the effects of fragmentation on population genetic differentiation, it is most appropriate to compare the results of this study with those for unfragmented populations of the same species situated within a limited geographic region. Interestingly, Raspé and Jacquemart (1998) found that, within regions, unfragmented populations of *S aucuparia* separated by at least 20 km displayed much less genetic differentiation (G_{stn}~0.010) than populations in Moffat Dale. For the chloroplast genome, comparison with data reported by Raspé *et al.* (2000a, Tables 4 & 5) shows that genetic differentiation between populations sampled in France and Belgium (Θ_c=0.258) and between Belgian populations sampled within a single region (Θ_c=0.218) is higher than in Moffat Dale (Θ_c=0.131).

It is possible to conclude that the remnant *S. aucuparia* populations of Moffat Dale are more genetically differentiated at isozyme markers than would be expected in non-fragmented populations separated by less than a few kilometres and remarkably, within a single catchment, statistically significant differentiation is detected for both nuclear and cytoplasmic markers (*G*-test, p-value<0.01). It is therefore likely that the genetic structure observed has been influenced by habitat fragmentation.

Higher genetic structure for maternally inherited cytoplasmic markers than for nuclear markers has been described for numerous plants species (e.g. McCauley,

1997; Rendell and Ennos, 2002; Bonnin *et al.* 2002). Several factors can induce the differences observed between the two genomes. First, they can be under different selection regimes. It is unlikely that any selective force acts directly on the marker loci studied. However, diversity at neutral loci can be influenced by selection if they are linked to loci under selection (Charlesworth *et al.* 1997; Schierup *et al.* 2000). Background selection is likely in non recombining genomes, (e.g. chloroplast) and may, in certain conditions, contribute to an increase of F_{st} (Charlesworth *et al.* 1993; Charlesworth *et al.* 1997). Drift will also be enhanced in non-recombining genomes. Moreover, in hermaphrodite species, the effective population size of uniparentally inherited chloroplast genome is half the effective population size of biparentally inherited markers (McCauley 1995). Finally, the chloroplast genome is maternally inherited in *S. aucuparia* (Raspé 2001), thus it is only dispersed by seed while nuclear genes are dispersed by both pollen and seed. Pollen is usually considered the primary vector of gene movement (Dyer and Sork 2001). The differential migration of pollen and seed contributes to increase population differentiation of the chloroplast genome as compared to the nuclear genome (Ennos 1994).

Indeed, in tree species, a low level of gene flow by seed and extensive pollen-mediated gene flow are advocated to explain the contrasting strong differentiation observed in maternal markers, as in *Pinus ponderosa* (Latta *et al.* 1998) and in *Fagus sylvatica* (Demesure *et al.* 1996). In *Quercus* spp., one of the most extensively studied species complex (Streiff *et al.* 1998; Cottrell *et al.* 2003), a ratio of 196:1 is necessary to explain structure (Ennos 1994). Austerlitz *et al.* (2000) argue, however, that delayed reproduction and long generation time in tree species allow a large number of initial founders to establish before reproduction begins and result in high diversity but low differentiation for nuclear genes. Recent studies on *S. aucuparia* (Raspé *et al.* 2000a) and the related species *S. torminalis* (Oddou-Muratorio *et al.* 2001) report much lower pollen versus seed migration ratios. In fragmented populations of *S. aucuparia* of Moffat Dale, no predominance of pollen over seed migration was found. These results can be explained by its differences in life history traits. In contrast to *Pinus* and *Quercus* spp., *S aucuparia* is insect pollinated and its seed are dispersed by birds and mammals (Raspé *et al.* 2000b). The low population

structure observed for cpDNA markers reflects seed dispersal effective over long distance associated with bird dispersal.

It must, nevertheless, be remembered that the ratio does not provide a quantification of gene flow *per se*. The exceptionally low pollen flow to seed flow ratio may also be explained by reduced pollen flow among population remnants. Low pollen dispersal may arise if, as a result of increased isolation between remnants, pollinators tend to forage more locally and travel less between remnants. In the exposed situation of Moffat Dale long distance insect flights between fragments may well be rare and this is the most likely cause of the observed genetic differentiation. Rare long distance insect flights have been reported for tropical bee species (Powell and Powell 1987; Jennersten 1988) although other tropical studies suggest that this is not always the case (e.g. Roubik and Aluja 1983; Dick *et al.* 2003).

One important aspect to consider when estimating historical gene flow from *F*-statistics is that inferences assume an infinite island model of migration. Because natural populations necessarily deviate from the unrealistic assumptions of the model, it has been argued that indirect quantitative estimates of gene flow are at best not needed, at worst, misleading (Whitlock and McCauley 1999). Estimates can especially lack precision for species with long generation time, when disturbance has occurred and when pollen flow is not predominant (Ennos 1994) although they may be reasonable when the spatial scale is small (Whitlock and McCauley 1999). Furthermore, estimates should be accurate within a factor of two, even for population structure quite different from an island model (Slatkin and Barton 1989). Bearing these considerations in mind, the reported estimate of pollen versus seed mediated migration is probably a realistic figure, and is relevant for comparison with other species or ecological contexts (Ennos, 1994).

S aucuparia is often a pioneer that will later be replaced by late successional tree species (Raspé *et al.* 2000a). This extinction/recolonisation dynamic is likely to be enhanced in the fragmented landscape of Moffat Dale as small population size increases the risk of extinction. Pannell and Charlesworth (2000) have shown that as

a result of such population turnover, genetic differentiation is increased if colonisation occurs through a similar process to migration and particularly if colonisation follows a propagule-pool model of dispersal more closely. Several empirical studies (reviewed in Giles and Goudet 1997) have confirmed these theoretical predictions and found higher F_{st} values among younger populations than among older populations (Pannell and Charlesworth 2000). This is likely to also be the case for bird dispersed *S. aucuparia.*

Indeed, Raspé (unpublished data) described a metapopulation of *S. aucuparia* in Belgium where genetic diversity is lower in young populations and genetic differentiation is higher between recently founded populations for both nuclear and maternal markers. There are substantial peculiarities in the genetic structure of *S. aucuparia* remnants of Moffat Dale that suggest a similar dynamic. It may be hypothesised that populations of Moffat Dale which are most diverse (Kr and CTr, Tables 3.3 and 3.4) and least differentiated (SCr, Table 3.9) are older than populations which are least diverse (Wr), depart from random mating (CDr) or are distinct in their cpDNA haplotype composition (CMr, Figure 3.2 and Table 3.9). The latter three populations are situated at lower altitude where *S. aucuparia* is more likely to be outcompeted by late successional species and where population turnover may therefore be especially important. It is also conceivable that birds preferentially travel between remnants at higher altitude (CTr, SCr, BMr) or where the species occurs on its own (SCr, Kr) which would increase diversity and decrease differentiation between these populations. Further research is evidently needed to test these predictions.

V Conclusions

The overall conclusions of this study are that, despite being highly fragmented for many generations, the populations of *Sorbus aucuparia* in Moffat Dale have retained substantial genetic diversity for both isozyme and chloroplast DNA markers. In the extant population there is no evidence of inbred individuals. However significant genetic differentiation is present among the population remnants implying that gene flow among them has been restricted. A very low value for the ratio of pollen to seed flow among remnants indicates that it is reduced pollen flow that has been principally responsible for the increased genetic differentiation of remnants.

From the point of view of forest restoration these results are encouraging. They indicate that within Moffat Dale as a whole there has been no dramatic loss of genetic variation. However the data only provide information about the adult population of trees. They do not make it possible to predict the genetic quality of the seed that would be collected from these adult trees. In order to determine whether the seed would be suitable for including in a forest restoration scheme, an understanding the mating patterns of trees within and between remnants is needed. Only with such information will it be possible to determine whether inbreeding is occurring within the population, and whether the genetic variation present in the adult population is effectively transmitted to the seed generation. Unfortunately, failure to optimise suitably polymorphic markers for the species (data not shown) gave no conclusive evidence of the mating patterns of *S. aucuparia* in southern Scotland. However, it has been possible to address these issues for another tree species occurring in the same landscape, *Fraxinus excelsior*, which is also considered for forest restoration. The findings of which are the subject of the next three chapters.

Chapter 4 : GENETIC STRUCTURE AND MATING PATTERNS OF *FRAXINUS EXCELSIOR* L. REMNANTS IN A DEFORESTED SCOTTISH LANDSCAPE

Abstract

Empirical evidence of the detrimental impact that habitat fragmentation may have on the processes shaping genetic variation and its distribution in tree populations is needed for a wide range of species and ecological conditions in order to formulate sustainable management practices, especially for temperate tree species, for which little data is available. Genetic variation at microsatellite markers was quantified in severely fragmented populations of the wind pollinated, wind dispersed tree *Fraxinus excelsior* (Oleaceae) in a single catchment (Moffat) in southern Scotland. Remnants maintain high levels of genetic diversity, comparable to those reported for large non-fragmented populations occurring in south-eastern Europe, and low interpopulation differentiation (Θ=0.080) suggesting that historical gene exchange has not been limited (Nm=3.48). Using the neighbourhood model, it was estimated from open-pollinated progeny arrays collected from all trees producing fruits in three of five remnants, that *F. excelsior* is completely outcrossing in Moffat Dale (selfing rate <1%) and that contemporary pollen-mediated gene flow into the catchment is extensive (46%). Effective pollen dispersal distances within the catchment averaging 328 m are at least 2.5 times greater than reported for a non-fragmented Romanian population suggesting that habitat fragmentation, by opening the landscape, facilitates airborne pollen movement and thus enhances pollen-mediated gene exchange between remnants.

keywords: *Fraxinus excelsior* L., microsatellites, *F*-statistics, neighbourhood model, pollen dispersal, fragmented landscape.

I Introduction

In chapter 1, I have discussed how habitat fragmentation may influence genetic variation within tree populations and genetic differentiation between them, ultimately affecting population viability (Young *et al.* 1996). In particular, it was pointed out that empirical evidence of the population genetic consequences of habitat fragmentation is needed for a wider range of species, especially temperate ones, in order to formulate sustainable management practices.

Severely fragmented populations of an insect pollinated bird dispersed species, *Sorbus aucuparia*, were shown to have retained substantial genetic variation (Chapter 3) but significant genetic differentiation among remnants implies a reduction, to some extent, of pollen-mediated gene flow. The impact of habitat fragmentation on the population genetic structure of species with contrasting modes of dispersal may be very different. Therefore, in this chapter, I address the genetic consequences of chronic habitat fragmentation for a wind pollinated wind dispersed species, *Fraxinus excelsior*, which occurs in the deforested landscape previously described for *S. aucuparia* remnants.

In southern Scotland, forest cover is extremely limited with less than one percent of native woodland remaining in some areas (Newton and Ashmole 1998). Major human impact in this region has been estimated to have begun as early as 6000 years ago (Wildwood Group of the Borders Forest Trust 2000). Ancient land use for pasture and much more recent implementation of conifer plantations have greatly fragmented the landscape. Many native tree species are confined to steep riversides not accessible to grazing herds. Remnants are very small, comprising ten to 30 mature individuals with very little or no evidence of natural regeneration in grazed areas. Typically, remnants are separated by hundreds of meters although some forest patches can be isolated from others by more than 1 km. As a result of continuous exposure in a very windy climate and intensive grazing, trees are undersized, and display a convoluted shape. There is an immediate need to assess the population genetics of these forest fragments as they are being considered for native woodland restoration in Scotland. In particular, extensive planting using seed collected from *F.*

excelsior remnants is planned (Wildwood Group of the Borders Forest Trust 2000). It is therefore necessary to understand the mating patterns of trees within and between fragments to determine whether the genetic quality of such planting stock would be suitable for forest restoration.

In this study, the objectives are to assess the extent of genetic structure and to characterise mating system parameters for *F. excelsior* in chronically fragmented populations of the southern Uplands of Scotland. In particular, I consider the amount of genetic variation remaining within remnants, the extent of genetic differentiation between them and the inferred levels of historical gene flow among fragments. Detailed patterns of contemporary pollen-mediated gene dispersal are also described using the neighbourhood model of Burczyk *et al.* (2002) within the heterogeneous landscape where remnants occur.

F. excelsior L. (Oleaceae), common ash, is a postpioneer heliophilous tree species widespread in temperate Europe (Marigo *et al.* 2000) and native throughout the British Isles. In southern Scotland, individuals grow commonly along river beds in exposed pastures and rarely exceed 8-10 m height, which is smaller than the average 12-18 m (Wardle 1962). The species' polygamous mating system is complex and poorly known (Wallander 2001). Sex expression varies at the floral, inflorescence and tree level, as well as between years, and individuals may be classified phenotypically in a continuum from purely male to purely female with a whole range of hermaphroditic intermediates (Wallander 2001). Hermaphrodite flowers are protogynous and self-pollination appears to be physically possible (Wallander 2001), but the functional gender of hermaphrodites remains unclear (Morand-Prieur *et al.* 2003). Fruits are wind dispersed, regular fruit bearing begins around 20 years of age and the species displays a masting behaviour with irregular fruiting between years (Tapper 1996).

Difficulties with optimising isozyme markers (This thesis, data not shown; Heuertz *et al.* 2001) have hindered the characterisation of patterns of genetic variation in *F. excelsior* and until recently, little information was available. However, thanks to the

development of nuclear microsatellite markers (Brachet *et al.* 1999; Lefort *et al.* 1999) and chloroplast DNA (cpDNA) markers (Petit *et al.* 2003) for the species, new research on the genetic make up of *F. excelsior* has been initiated for populations across Europe, with projects such as GENFOR (Heuertz 2002), CYTOFOR (Petit *et al.* 2003) and FRAXIGEN (URL: www.fraxigen.net). In particular, Heuertz *et al.* (2001) have found significant inbreeding and low genetic differentiation among populations sampled in continuous stands across Bulgaria. Similar results were reported for French populations, in which high heterozygote deficiency was reported (Morand *et al.* 2002). However, for a continuous population, in Romania, no inbreeding was detected but spatial genetic structure fitting an isolation by distance model was attributed, in a simulation procedure, to restricted seed-mediated and moderate pollen-mediated gene flow (Heuertz *et al.* 2003).

In this chapter, microsatellite markers are used to quantify neutral genetic variation and to assess interpopulation differentiation from mature trees sampled in *F. excelsior* remnants. In addition, using a model-fitting approach (Burczyk *et al.* 2002), detailed patterns of contemporary pollen-mediated gene flow and effective pollen dispersal within the fragmented landscape are estimated from open-pollinated progeny arrays collected from known mother trees. Previous studies describing the genetic structure and mating patterns of non-fragmented Europeans populations provide a useful comparison to assess the extent to which habitat fragmentation has altered the population genetic dynamics of *F. excelsior* remnants in southern Scotland.

II Material and methods

1 Study site and populations

The study site is located 80 km south of Edinburgh in Dumfries and Galloway (N55° 23' 51" W3° 19' 50"). It is the same catchment previously described for *Sorbus aucuparia* remnants (Chapter 3). *Fraxinus excelsior* is only present in five forest remnants within the Carrifran valley and its immediate surroundings. There are no other ash remnant in Moffat Dale and a similar landscape is repeated throughout the southern Uplands of Scotland (Chapter 1). *F. excelsior* remnants are situated at least 600 m apart from one another and two trees (marked A & B, Figure 4.1) are isolated from the nearest individuals by a distance of 250 m (Figure 4.1).

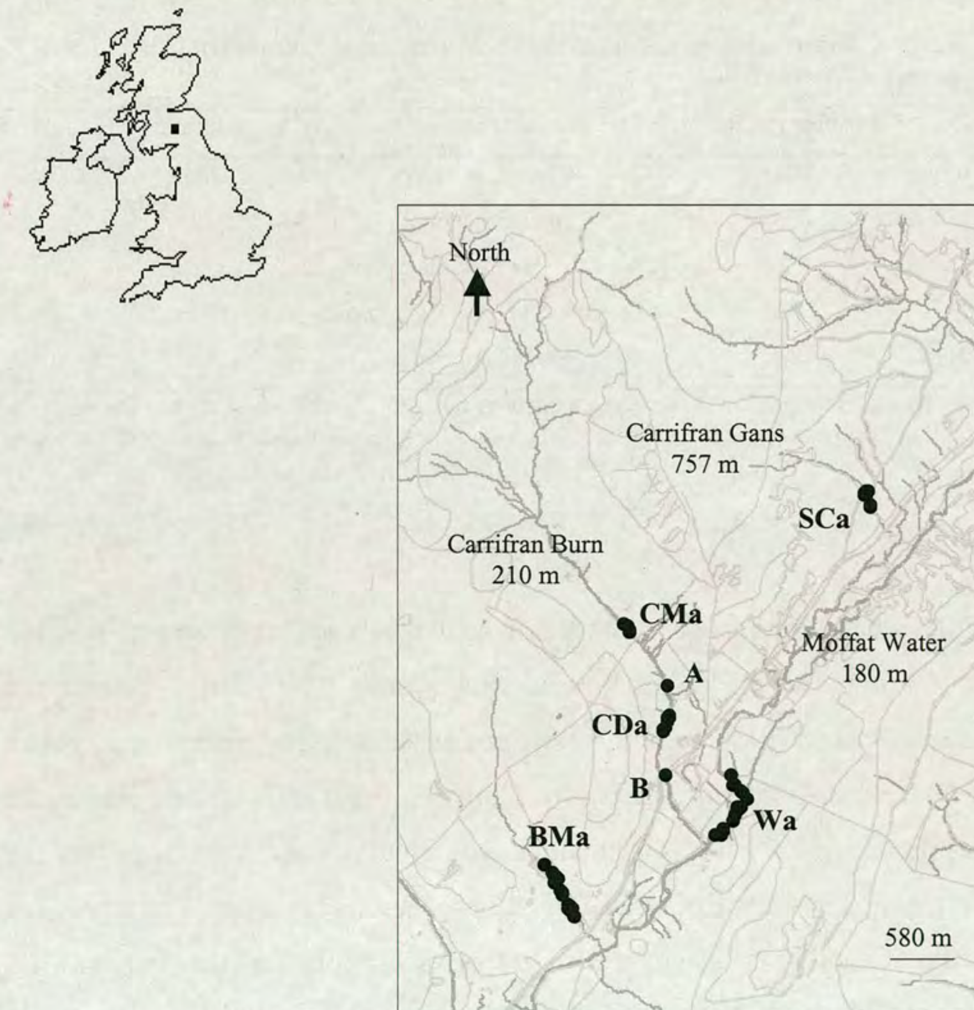


Figure 4.1: Distribution of *F. excelsior* populations sampled in Moffat Dale.
Trees are represented by a dot and grouped into populations as described in Table 4.1. The number of trees occurring in each population is also given in Table 4.1. The background map is a section of Ordnance Survey product Land-line.Plus-nt11 © Crown copyright Ordnance Survey. An EDINA digimap / JISC supplied service.

In four of the fragments identified, *F. excelsior* occurs together with *Sorbus aucuparia* and other native trees such as *Betula pubescens* and *Corylus avellana*. *F. excelsior* is the most abundant species in all but one remnant, maintaining a population size of more than 50 trees in Spoon Burn and Whitewells (Table 4.1). One fragment, in Swine Cleuchs, is a pure stand of 12 trees situated within a 30-year-old conifer plantation, located 1.7 km from the nearest identified population (Figure 4.1). In CDa, CMa and SCa, all identified mature trees were sampled. In each of two larger remnants (BMa and Wa), a sample of 20 randomly distributed individuals was taken (Table 4.1). The sampling strategy reflects the patchy nature of the landscape. Indeed, the sampled *F. excelsior* remnants vary in size, spatial isolation, altitude and accessibility to grazing (Table 4.1).

Table 4.1: location, abbreviated name, landscape characteristics, altitude and size of five *F. excelsior* remnants of Moffat Dale.

Location	Abbreviation	Site characteristics	Altitude	Size (N)
Carrifran (down)	CDa	grazed riverside	188m	30 + 2†
Carrifran (middle)	CMa	grazed riverside	212m	4
Swine Cleuchs	SCa	grazed riverside in conifer plantation	271m	12
Whitewells	Wa	riverbank protected from grazing	170m	~50*
Spoon Burn	BMa	grazed riverside	262m	50-70*

†Isolated trees A and B (Figure 4.1) were grouped with trees in population CDa for data analysis. *Census sampling in all fragments except in BMa and Wa where 20 individuals were randomly sampled.

2 Mapping

Sampled individuals were tagged with permanent labels and their spatial position was recorded using a handheld global positioning system (GPS) Etrex ™ set to the World geodetic system 1984 coordinate system (WGS84). For mapping purposes, coordinates were transformed to the United Kingdom (UK) national grid coordinate system. For Euclidean distance calculations, coordinates were transformed to the Cartesian projection system. Both transformations were made using a Projection and Transformation Calculations macro in Microsoft Excel supplied by the National GPS Network of the Ordnance Survey (URL: www.gps.gov.uk). Individual positions were plotted in Arcview 3.1 on an Ordnance Survey digital map made available through the EDINA digimap service (Figure 4.1).

3 Plant material

In Spring 2001, leaf material was collected from all mature individuals in remnants CDa, CMa and SCa (30, 4 and 12 individuals respectively), from the two isolated trees (marked A and B, Figure 4.1) and from a representative sample of 20 randomly distributed mature trees in each of fragments BMa and Wa, bringing the total sample to 88 individuals. Leaf material was stored at –80°C immediately after collection until DNA extraction.

In order to describe the mating patterns of *F. excelsior* remnants of Moffat Dale, in Autumn 2000, 50-200 fruits from the current year's reproduction were collected from all trees producing fruits (19 trees in total, see detailed description of sampling strategy in Chapter 5) in remnants CDa, CMa and SCa, chosen for their difference in size (Table 4.1) and in degree of spatial isolation (mean distance to other remnants 1093 m, 1296 m, 2210 m respectively). Whenever possible, fruits were collected from several branches throughout the canopy in an attempt to adequately reflect the pollen pool received by the tree in the sample. Fruits were dried and kept in a cool dry place until DNA extraction.

4 DNA isolation

For the adult trees, genomic DNA was isolated from 100 mg of frozen leaf following a standard extraction protocol adapted from Doyle and Doyle (1987) and described by Rendell and Ennos (2002). Pericarps were removed and seeds were soaked in sterile water for 4-6 hours to allow rehydration. Seeds were then dissected and the embryo was removed to a new Eppendorf tube immediately prior to DNA extraction. Genomic DNA was isolated using a Nucleon™ Phytopure plant extraction kit (Tepnel Life Sciences). The extraction protocol was adapted for samples weighing less than 0.1 g. Volumes of Phytopure reagents 1, 2 and 3 used were 400 µl, 120 µl and 50 µl respectively. Extracted DNA was dissolved in 100 µl of Tris-ethylene-diamine-tetraacetic acid (TE, pH 7.6) and stored at -20 °C.

5 Microsatellite amplification and genotyping

Eight microsatellite loci previously characterised for *F. excelsior* (Brachet *et al.* 1999; Lefort *et al.* 1999) were successfully amplified by polymerase chain reaction (PCR). For FEMSATL2, FEMSATL5, FEMSATL8, FEMSATL16, FEMSATL19 (Lefort *et al.* 1999) and for 1.19 (Brachet *et al.* 1999), PCR reactions (20 μl) contained autoclaved deionised water, 1X BIOTAQ NH_4 reaction buffer (Bioline), 1.5 mM $MgCl_2$ (Bioline), 0.2 mM deoxynucleotide mix (Sigma), 0.4 μM of each primer (MWG-Biotech), one unit of BIOTAQ DNA polymerase (Bioline) and 20-30 ng of template DNA. For M2-30B and 3.1 (Brachet *et al.* 1999), amplification of non-specific bands was reduced by using IMMOLASE hot start DNA polymerase (Bioline). PCR reactions were performed in a GeneAmp® PCR system 9700 (PE Applied Biosystems) or in a MJ Research™ PTC-100 Programmable Thermal Controller. PCR cycles and annealing temperatures were set as described for each primer by Brachet *et al.* (1999) and Lefort *et al.* (1999) except for M2-30B and 3.1 for which the initial denaturation step was modified to 96 °C for 7 min to allow for activation of the hot start DNA polymerase. The forward primer of each pair was labelled with a fluorescent dye at its 5' end (IRDye™ 700 or IRDye™ 800, MWG-Biotech). For FEMSATL5 and FEMSATL8, the yield of amplification reaction was increased by using both labelled and unlabelled forward primers in a ratio 2:3.

Amplified products were separated on a LI-COR® Long Read IR 4200 DNA Sequencer. Two PCR products, each labelled with one of two different fluorescent dyes (IRDye™ 700 or IRDye™ 800), were multiplexed and mixed with an equal volume of IR^2 formamide stop solution (LI-COR Biosciences). Reactions were denatured for 5 min at 95 °C prior to loading, then 0.6 to 0.8 μl of reaction was loaded onto a 6% denaturing polyacrylamide 25 cm gel (LongRanger™ Gel solution, FMC Bioproducts) prepared according to manufacturers instructions and including treatment with AG®501-X8 resin and filtration prior to polymerisation. Gels were run at 1200 V and 45 °C for 1.75 hours. On each gel, size standards (LI-COR® IR^2 size standards 50-350bp) were loaded at regular intervals and individuals were repeated between gels as positive controls. In addition, maternal parents were always

run with progeny arrays. Electrophoregrams were analysed independently by C. Bacles and A. Lowe using LI-COR® Gene ImagIR 3.59 genotyping software.

The 88 adult trees were screened for all eight microsatellite markers. A subset of five loci, chosen for their scorability and high level of polymorphism, was used to characterise progeny arrays. A sample of 30 seeds (or all seeds if the seed crop was less) for each of the 19 trees sampled for fruits, was screened for microsatellites 3.1, M2-30B, 1.19, FEMSATL2 and FEMSATL5.

6 Statistical analysis

Independence of loci. Calculation of mean estimates of F-statistics over loci and statistical estimation of mating system parameters using multilocus genotypic data assume independence of loci. Genotypic linkage equilibrium for each locus pair across populations was tested using GENEPOP 3.2a (Raymond and Rousset 1995).

Genetic diversity within remnants. For the five remnants sampled for mature trees in Moffat Dale, gene diversity (H_e) per population was estimated using an unbiased estimator for each microsatellite locus in FSTAT 2.9.3 (Goudet 2001). H_e estimates per locus and averaged over eight loci at the species level were obtained using GENETIX 4.01 (Belkhir *et al.* 1998). The number of alleles (k) sampled at each locus in each population and overall is also reported. Allelic richness (R) per locus in each population and overall was estimated using the rarefaction method of El Mousadik and Petit (1996) in FSTAT 2.9.3. For these analyses and subsequent population level analyses, isolated trees A and B were grouped with trees in the nearest remnant CDa (Figure 4.1).

Departure from random mating. The inbreeding coefficient (F_{is}) was estimated per microsatellite locus for each population and departure from Hardy-Weinberg equilibrium was assessed using FSTAT 2.9.3. Tests for deficit in heterozygotes (*HD*) and excess in heterozygotes (*HE*) significant at the 5% level were performed by randomising alleles among individuals.

Genetic differentiation among remnants. Overall inbreeding (*f*) and genetic differentiation (Θ) coefficients were estimated following Weir and Cockerham, (1984) for each locus and overall loci using FSTAT 2.9.3. Departure from Hardy-Weinberg equilibrium within samples was tested by comparing the distribution of the F_{is} statistic (*f*) for the observed data set with its distribution for a randomised data set obtained after 10,000 permutations of alleles among individuals within samples. A *G*-test (Goudet *et al.* 1996), based on 10 000 permutations of genotypes among samples (i.e. not assuming Hardy-Weinberg equilibrium within samples), was performed to test for population differentiation at microsatellite loci.

Effective number of migrants between remnants. Using the private allele method of Barton and Slatkin (1986), the effective number of migrants per generation (*Nm*) between the five *F. excelsior* remnants sampled in Moffat Dale was estimated in GENEPOP 3.2a assuming an infinite island model of migration (Wright 1951). The method is an interesting alternative to inferring *Nm* from the measure of genetic differentiation based on F_{st} (Wright 1951) because the relationship between the frequency of private allele ($p(1)$) and *Nm* depends only weakly on mutation rate (Slatkin and Barton 1989) which may not be assumed to be much smaller than migration when considering microsatellite loci (Hedrick 1999; Balloux and Lugon-Moulin 2002).

Comparison of levels of genetic diversity among cohorts. Mean H_e over three microsatellite loci (M2-30B, 1.19, FEMSATL2, see results section) was estimated for both adult and seed cohorts for each of the three remnants (CDa, CMa and SCa) where both adult trees and seed from all trees producing fruits were sampled, using an unbiased estimator (Nei 1978) in GENETIX 4.01 (Belkhir *et al.* 1998). Allelic richness per sample (R_s) was also estimated using FSTAT 2.9.3 (Goudet 2001) and averaged over the three microsatellite loci. R_s was estimated considering all six samples together (i.e. adult and seed samples for each of three remnants) in order not only to assess the extent of genetic diversity present in the adult population that is maintained in the seeds produced, but also to compare levels of genetic diversity

across cohorts between the three remnants which vary in size and degree of spatial isolation from other fragments.

Mating system. NEIGHBOR v.2.0 (Burczyk *et al.* 2002) was used to estimate a number of mating systems parameters for *F. excelsior* in Moffat Dale. The proportion of offspring resulting from selfing events (s), from pollen-mediated gene flow from unsampled trees (m) and the influence of the distance to mother tree (β) on male reproductive success of trees sampled within Moffat Dale were estimated by fitting the neighbourhood model (Burczyk *et al.* 2002) to genotypic arrays in the offspring of the 19 trees sampled for seeds (i.e. mother trees) in remnants CDa, CMa and SCa.

The neighbourhood model is described in details by Burczyk *et al.* (2002). The approach is particularly interesting as not only does it utilise the genetic information on paternity available from progeny arrays to describe mating systems in plant species but it also allows estimation of the influence of ecological factors on mating patterns. Furthermore, one distinct advantage of this model approach over traditional paternity analyses is that information on allele frequencies in the immigrant pollen pool is not required and may be estimated simultaneously with other parameters (Burczyk *et al.* 2002). The main interest of this study lies in describing mating patterns for *F. excelsior* at the catchment level (i.e. Moffat Dale). In addition, the model requires that mother trees have an equally sized neighbourhood when data across mother trees are combined to calculate 'catchment level' estimates (Burczyk *et al.* 2002) Therefore, I chose to consider a neighbourhood radius of 3000 m which is the smallest radius for which the 19 mother trees sampled have an equal number of neighbours. Since the actual functional gender of the trees sampled in Moffat Dale is unknown for the year of sampling (Chapter 5), all 88 mature trees were considered as potential pollen donors. Therefore, there are 87 potential pollen donors in the 3000 m radius neighbourhood of each of the 19 mother trees sampled (Figure 4.1). Genotypic data considered for analyses combine information across loci M2-30B, 1.19 and FEMSATL2 for which rare alleles were binned with more common alleles of nearest size in order to minimise genotyping error (Chapter 5, Appendix A). In total, 422

seeds were genotyped which is a large enough sample to obtain reliable parameter estimates (Burczyk *et al.* 2002).

Parameter estimates ($s_{_est}$, $m_{_est}$ and $\beta_{_est}$) were calculated using NEIGHBOR v.2.0 and tested for significance at the 5% level using numerical procedures based on maximum likelihood methods (Burczyk *et al.* 2002) and allele frequencies in the background pollen pool (i.e. immigrant) were estimated simultaneously with other parameters. The best-fitting model (maximising likelihood) was retained, including only parameters estimated to be significantly different from zero (p-value<0.05).

Given s, m and β estimated jointly for the 19 neighbourhoods, the male reproductive success of trees, and subsequently the mean pollen effective dispersal distance and effective male population size within the neighbourhood of each of the 19 mother trees sampled were inferred in NEIGHBOR v.2.0. The mean pollen effective dispersal distance and mean effective male population size over all 19 neighbourhoods are reported herein, thus providing an estimate of effective pollen dispersal for *F. excelsior* across Moffat Dale. Mean estimates of effective dispersal distance and effective male population size over all neighbourhoods within each of the three forest remnants where mother trees were sampled (CDa, CMa and SCa) are also reported to assess how population size and spatial isolation may affect contemporary mating patterns for *F. excelsior* in Moffat Dale.

III Results

1 Independence of microsatellite loci and choice of markers

Across eight microsatellite loci, highly significant association was found for two pairs involving four loci, across all populations, for 3.1 x FEMSATL2 (p-value<0.001) and for FEMSATL5 x FEMSATL8 (p-value<0.001). Furthermore, non-Mendelian segregation at loci 3.1 and FEMSATL5 was observed for open-pollinated progeny arrays (Chapter 5, Appendix A). Therefore, estimates of genetic diversity and inbreeding coefficient at loci 3.1 and FEMSATL5 per population are presented for information but subsequent analyses exclude these two loci.

2 Genetic diversity within remnants

Across 88 mature trees sampled in five *Fraxinus excelsior* remnants of Moffat Dale, 172 alleles were recorded over eight microsatellite loci. Allelic composition varies greatly between populations with only a fraction of the total number of alleles identified in each remnant (Table 4.2). In Moffat Dale, FEMSATL2 is most polymorphic with 40 alleles recorded overall and most diverse when considering the measure of allelic richness (R=5.217) while gene diversity is highest at locus M2-30B (H_e=0.940). FEMSATL16 is least diverse (H_e=0.649, R=2.933) with only six alleles recorded overall (Table 4.2). The most spatially isolated remnant, SCa (N=12), is least diverse for both the measure of allelic richness and gene diversity at all loci except FEMSATL8 and FEMSATL5 at which CMa (N=4), the smallest remnant (Figure 4.1) is least diverse. Wa (N=20 sampled, Table 4.1) is most diverse at all loci except 1.19 (CMa), M2-30B (BMa) and 3.1 at which BMa is most diverse on the measure of allelic richness (R=3.689) and gene diversity is highest for remnant CMa (H_e=0.833, Table 4.2).

3 Departure from random mating

Estimates of inbreeding coefficient ($F_{is\text{-}est}$) are heterogeneous among loci and populations but are high overall (Table 4.3). $F_{is\text{-}est}$ varies between –0.200 at locus 1.19 and FEMSATL16 and 1.000 at locus FEMSATL8 for CMa, although these extreme values may be due to the small size (N=4) of this remnant.

Table 4.2: Sample size (*N*), number of alleles sampled (*k*), gene diversity (H_e), and allelic richness (*R*) at eight microsatellite markers for mature individual sampled in each of five *F. excelsior* remnants of Moffat Dale.

		Population					
locus		**CDa**	**CMa**	**SCa**	**BMa**	**Wa**	**Overall**
3.1	*N*	30	4	12	16	12	74
	k	7	4	6	7	6	12
	H_e	0.797	**0.833**	**0.712**	0.779	0.754	0.859
	$R^{\dagger}$	3.674	3.464	**3.376**	**3.689**	3.309	4.283
M2-30B	*N*	32	4	12	19	20	87
	k	16	5	11	12	15	30
	H_e	0.893	0.875	**0.841**	**0.920**	0.908	**0.940**
	R	4.650	4.214	**4.248**	**4.825**	4.835	5.193
1.19	*N*	26	4	11	18	18	77
	k	8	5	3	6	7	10
	H_e	0.758	**0.833**	**0.264**	0.611	0.823	0.772
	R	3.668	**4.214**	**1.753**	2.976	3.925	3.705
FEMSATL2	*N*	29	3	12	20	20	84
	k	23	5	12	19	21	**40**
	H_e	0.903	0.917	**0.898**	0.947	**0.958**	0.938
	R	4.847	5.000	**4.787**	5.243	**5.379**	**5.217**
FEMSATL5	*N*	31	4	6	19	19	79
	k	15	3	6	12	16	34
	H_e	0.781	**0.750**	0.933	0.927	**0.942**	0.921
	R	3.837	**2.750**	4.288	4.826	**5.072**	5.021
FEMSATL8	*N*	27	4	11	17	18	77
	k	9	2	10	11	14	21
	H_e	0.801	**0.667**	0.845	0.871	**0.943**	0.888
	R	3.810	**2.000**	4.349	4.356	**5.042**	4.688
FEMSATL19	*N*	28	4	11	19	20	82
	k	14	5	8	11	12	19
	H_e	0.912	0.875	**0.750**	0.769	**0.921**	0.907
	R	4.746	4.214	**3.734**	3.855	**4.910**	4.811
FEMSATL16	*N*	20	4	12	17	15	68
	k	5	2	3	5	4	**6**
	H_e	0.676	0.417	**0.239**	0.627	**0.693**	**0.649**
	R	2.957	1.964	**1.696**	2.820	**2.990**	**2.933**

[†]Measure of allelic richness based on a minimum sample size of three diploid individuals. Estimations were performed in FSTAT 2.9.3 (Goudet 2001) and GENETIX 4.01 (Belkhir *et al.* 1998). In bold, lowest and highest estimates.

Heterozygote deficit significant at the 5% level was detected in three out of the five *F. excelsior* remnants sampled in Moffat Dale at loci 3.1 and FEMSATL5 (Table 4.3) indicating that non-Mendelian segregation observed at those loci may be the result of occurrence of null alleles (Heuertz *et al.* 2001; Morand *et al.* 2002). Excluding loci 3.1 and FEMSATL5, mean $F_{is\text{-}est}$ were found to be high overall, ranging from 0.074 in population SCa to 0.318 in BMa (Table 4.3). Significant heterozygote deficit was detected in populations CDa, BMa and Wa suggesting a departure from random mating in these populations.

Table 4.3: Inbreeding coefficient estimates ($F_{is\text{-}est}$) for each of five *F. excelsior* populations of Moffat Dale per locus and overall based on sampling of mature individuals.

locus	Population				
	CDa	**CMa**	**SCa**	**BMa**	**Wa**
3.1	**0.415***	0.700	**0.649***	**0.519***	0.779
M2-30B	0.125	0.143	0.009	**0.370***	0.229
1.19	-0.065	-0.200	0.655	0.182	0.259
FEMSATL2	0.046	-0.091	-0.021	0.208	0.113
FEMSATL5	0.245	0.667	**0.821***	**0.659***	**0.441***
FEMSATL8	0.307	1.000	-0.075	**0.595***	**0.470***
FEMSATL19	**0.608***	0.143	0.152	**0.316***	0.077
FEMSATL16	0.039	-0.200	0.302	0.155	-0.155
All(6)‡	**0.188***	0.127	0.074	**0.318***	**0.178***

*In bold, estimates that show a deficit in heterozygotes significant at the 5% level. In no instance was significant heterozygote excess detected (5% level). ‡$F_{is\text{-}est}$ averaged over six loci, excluding 3.1 and FEMSATL5 which show non-Mendelian segregation. Estimations and tests for heterozygote excess and deficit were performed using FSTAT 2.9.3 (Goudet 2001).

4 Genetic differentiation among *F. excelsior* remnants

Estimates of overall inbreeding (*f*) for five *F. excelsior* remnants sampled for mature trees in Moffat Dale are highly variable between loci, ranging from a low value of 0.032 at locus FEMSATL16 to a high one of 0.388 at locus FEMSATL8, and departure from random mating was detected at all markers except FEMSATL16 (Table 4.4). Mean *f* estimated over six microsatellite loci is high (f=0.202) and significantly different from zero (Table 4.4).

Table 4.4: Estimates of f and Θ at each of six microsatellite loci, and overall among five *F. excelsior* remnants sampled for mature trees in Moffat Dale.

locus	smallf (f)	test HW†	theta (Θ)	test PD‡
M2-30B	0.191	p-value<0.001	0.066	p-value<0.001
1.19	0.114	p-value<0.05	0.161	p-value<0.001
FEMSATL2	0.089	p-value<0.001	0.015	p-value<0.01
FEMSATL8	0.388	p-value<0.001	0.061	p-value<0.001
FEMSATL19	0.334	p-value<0.001	0.067	p-value<0.001
FEMSATL16	0.032	ns	0.138	p-value<0.001
All	**0.202**	**p-value<0.001**	**0.080**	**p-value<0.001**

†HW: Hardy Weinberg equilibrium. ‡PD: population differentiation, test not assuming random mating. ns: non significant at 5% level. Estimates obtained and tests performed using FSTAT 2.9.3 (Goudet 2001).

Estimates of Θ are consistently low between loci, ranging from 0.015 at locus FEMSATL2 to 0.138 at locus FEMSATL16 but statistically significant genetic differentiation among the five *F. excelsior* populations sampled in Moffat Dale was detected at all six markers (Table 4.4). Over all loci, Θ is low (Θ=0.080) but population differentiation significant at the 5% level, was detected among Moffat Dale remnants.

5 Historical gene flow among remnants.

Under the assumption of an infinite island model of migration, at migration-drift quasi-equilibrium (Barton and Slatkin 1986), the effective number of migrants per generation (*Nm*) among the five remnants sampled for mature trees in Moffat Dale (with a mean sample size of 15.83), was estimated over six microsatellite loci for which the mean frequency of private alleles ($p(1)$) is 4.85E^{-02} for the sampled populations. *Nm* among remnants within Moffat Dale was found equal to 3.48 after correction for sample size.

6 Mating patterns of *F. excelsior* within Moffat Dale

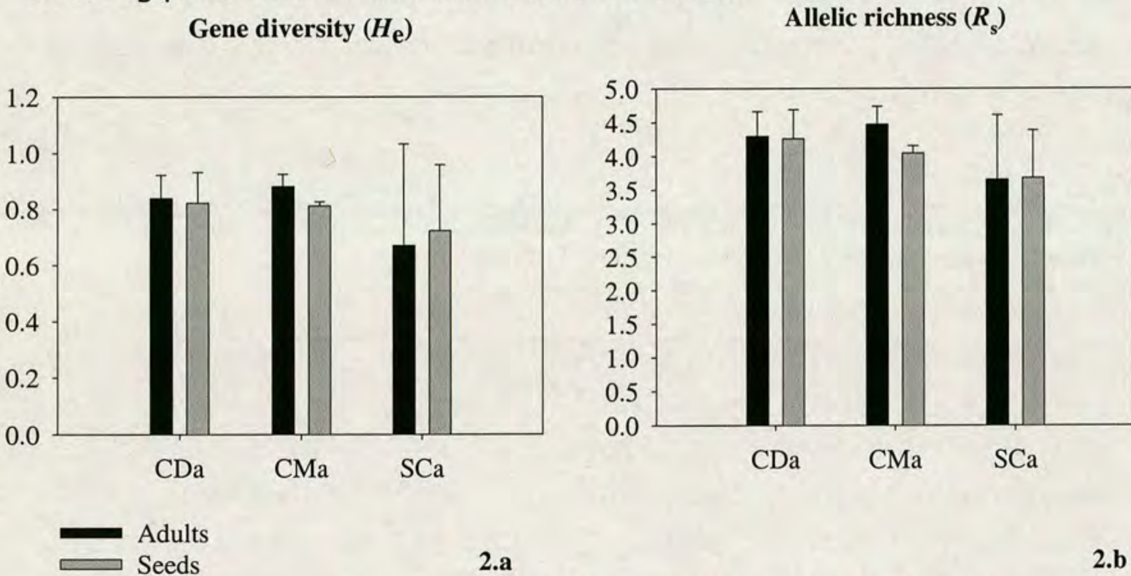


Figure 4.2: Comparison of gene diversity (H_e, 2.a) and allelic richness (R_s, 2.b) between adult and seed cohorts in each of three *F. excelsior* remnants of different size and degree of spatial isolation in Moffat Dale.
Mean H_e over three microsatellite loci estimated in GENETIX 4.01 (Belkir, *et al.* 1998), R_s estimated for a minimum sample size of three individuals in FSTAT 2.9.3 (Goudet 2001) and averaged over three microsatellite loci. Remnants CDa, CMa and SCa comprise 30, 4 and 12 trees and 282, 32, 108 seeds were collected from all trees producing fruits in each of them respectively. The mean distance to other remnants is 1093 m, 1286 m, and 2210 m respectively for CDa, CMa and SCa. Bars show standard errors.

Comparison of genetic diversity between cohorts. Comparison of gene diversity (*He*) and allelic richness (R_s) averaged over three microsatellite loci in adult and seed cohorts in three remnants of varying size and degree of spatial isolation shows that similar levels of genetic diversity are maintained in adult and seed cohorts across remnants (Figure 4.2). Although mean H_e and R_s estimates for the most isolated remnant (SCa) are lower than for other remnants, their variance is also greater (Figure 4.2) and overall, differences between cohorts and between remnants are not statistically significant.

Mating system. Mean estimates of mating system parameters for *F. excelsior* in Moffat Dale are reported in Table 4.5. From a sample of 422 seeds collected from 19 mother trees, the neighbourhood model estimates that the selfing rate is less than 1% ($s_{_est}$=0.00791±0.00783, *p*-value=0.001) and that over 45% of successful pollination is achieved by immigrant pollen ($m_{\text{-}est}$=0.45898±0.02747, *p*-value=0.001) suggesting not only that complete outcrossing occurs for *F. excelsior* in Moffat Dale but also that pollen flow into the catchment is extensive.

Furthermore, the distance between mother trees and potential pollen donors within a 3000 m - radius neighbourhood (β) has a significant effect (*p*-value=0.001) on the male reproductive success of the 88 mature trees sampled (Table 4.5). β is expected to be negative if male reproductive success decreases with increasing distance to the mother tree and vice versa (Burczyk *et al.* 2002). The estimated β is negative ($\beta_{_est}$=-0.00181±0.00021, Table 4.5) suggesting that, although pollen-mediated gene flow into the catchment is extensive, preferential mating with close neighbours occurs for *F. excelsior* within Moffat Dale.

Table 4.5: Mean estimates of selfing (*s*), immigrant pollen (*m*), and influence of distance between mates (β) on male reproductive success of trees within Moffat Dale for *F. excelsior* when mother trees have a 3000 m-radius neighbourhood.

sample size[†]	neighbourhood size[‡]	mating parameters[#]		
		$s_{\text{-est}}$ (se)	$m_{\text{-est}}$ (se)	β_{est}(se)
19/422	87	0.00791 (0.00783)	0.45898 (0.02747)	-0.00181 (0.00021)

Mating parameters estimated in NEIGHBOR v.2.0 (Burzyck *et al.* 2002) by fitting the neighbourhood model (Burzyck *et al.* 2002) to genotypic arrays of progeny from all trees producing fruits in three remnants of Moffat Dale. [†]Estimates are based on a sample of 422 seeds in total collected from 19 mother trees. [‡]87 potential pollen donors are located within a 3000 m neighbourhood of each mother tree (i.e. Moffat Dale, Figure 5.1). [#]Precision for all parameters is 0.001 based on likelihood ratio tests and standard errors (se) are derived from the variance-covariance matrix (Burzyck *et al.* 2002).

Effective pollen dispersal within Moffat Dale. When mating system parameters are estimated in the neighbourhood model across all neighbourhoods (19 mother trees, N=422), the effective pollen dispersal distance (m_p) inferred from the male mating success of mature trees sampled within a 3000 m-radius neighbourhood is, on average, 328 m (Table 4.6) which is considerably less than the mean distance between trees within Moffat Dale (1050 m, Table 4.6). Although pollen dispersal appears to be localised, nearly half the trees sampled contribute, on average, to the pollen pool of mother trees (N_e=47.75%, Table 4.6).

The 19 *F. excelsior* trees sampled for seeds are located in three forest remnants. 11, 2 and 6 mother trees are located in CDa, CMa and SCa respectively. The three remnants vary in size (30, 4, 12 mature trees respectively) and degree of spatial isolation from other fragments (mean distance to other fragments 1093 m, 1286 m, 2210 m respectively). Within remnants, estimates of effective pollen dispersal distance and male neighbourhood size are strikingly homogeneous among individuals (Appendix 4.1).

However, estimates vary greatly for mother trees among remnants. In CDa, m_p was estimated to be 245 m (Table 4.6) while the average distance between mother trees and potential pollen donors within Moffat Dale is nearly three-fold the mean effective pollen dispersal distance (667 m). Although m_p in CMa, the smallest remnant, is about three times as great as in CDa (656 m), a similar pattern is

observed in this remnant where mother trees are located, on average, 1023 m away from sampled pollen donors (Table 4.6). Interestingly, 61.63% of trees sampled contribute to the pollen pool of mother trees in CDa (Table 4.6). Similarly in CMa, N_e is only slightly lower than in CDa (N_e=56.81%) which suggests that a significant number of trees located in neighbouring fragments are reproductively active in CMa. However, in SCa, patterns of pollen dispersal are very different. The mean distance between mother trees and potential mates within the catchment is much greater (1762 m) but m_p is found to be limited to 370 m and N_e is about a third that of other remnants (N_e=19.37%, Table 4.6). Considering that the remnant nearest to SCa is located over 1600 m away (Figure 4.1) and that about 14% of the trees sampled are located in SCa, results suggest that mating (within the neighbourhood) occurs mainly between trees within this remnant.

Table 4.6: Mean effective pollen dispersal distance and male neighbourhood size for *F. excelsior* within Moffat Dale and comparison of pollen dispersal parameters between three remnants of varying size and degree of spatial isolation.

			effective pollen dispersal in Moffat Dale	
population	sample size†	mean distance between mates‡	mean effective pollen dispersal distance# m_p	mean effective male neighbourhood size N_e§ (%)
CDa	11/282	667 m	245 m	61.63%
CMa	2/32	1023 m	656 m	56.81%
SCa	6/108	1762 m	370 m	19.37%
Catchment	**19/422**	**1050 m**	**328 m**	**47.75%**

Effective pollen dispersal parameters obtained in NEIGHBOR v.2.0 (Burzyck *et al.* 2002) from male reproductive success of 88 trees sampled within Moffat Dale based on mating parameters estimated across all neighbourhoods for 19 trees sampled for seed (N=422, Table 4.5). Values for each neighbourhood given in Appendix 4.1. †Number of trees sampled for fruits in remnant/total number of seeds sampled. ‡Mean distance between mother tree and 87 potential pollen donor sampled within a 3000 m-radius neighbourhood (i.e. Moffat Dale) and averaged over all mother trees sampled for fruits in remnant. #Mean effective pollen dispersal distance within a 3000 m-radius neighbourhood. §Mean N_e expressed as the percentage of the 87 trees located within the neighbourhood of each mother tree actively contributing to the pollen pool of the sampled tree

IV Discussion

1 Genetic structure and historical gene flow among forest fragments

The genetic diversity assessed at eight microsatellite markers in five *Fraxinus excelsior* remnant populations in a severely deforested catchment, is high. Between six and 40 alleles were sampled at a particular marker locus and overall gene diversity ranges between 0.649 (FEMSATL16) and 0.940 (M2-30B) for a sample of only 88 mature trees. These figures are remarkably similar to values reported at five microsatellite markers by Heuertz *et al.* (2003) for a Romanian population. For this large population (152 individuals), occurring in a continuous mixed deciduous forest, 10 to 42 alleles were sampled overall and H_e varies between 0.663 (FEMSATL16) and 0.918 (M2-30; Heuertz *et al.* 2003). Furthermore, for continuous stands sampled across Bulgaria, Heuertz *et al.* (2001) observed between 10 and 59 alleles at five microsatellite loci for a much larger sample size (N=322) and overall gene diversity (He=0.793; Heuertz *et al.* 2001) was in the range of estimated values for *F. excelsior* in Moffat Dale. The comparison with such findings shows not only that populations of *F. excelsior* in Scotland harbour just as much diversity at neutral nuclear microsatellite markers as populations occurring in south-eastern Europe which has been identified as a hotspot of genetic diversity (Petit *et al.* 2003) but also that, despite small remnant size and severe isolation, chronic habitat fragmentation over 6000 years has not led to loss of genetic variation within the species.

Overall, the inbreeding coefficient estimated for populations sampled in Moffat Dale is high (f=0.202) and significant departure from random mating was detected (p-value<0.001). Nevertheless, estimates are heterogeneous among loci with no significant departure from random mating detected at FEMSATL16 (f=0.032, p-value>0.05). The highest value for the inbreeding coefficient is estimated at FEMSATL8 (f=0.388, p<0.001). Results in this study are in contrast with very low inbreeding coefficients reported for continuous stands in Bulgaria (F_{is}=0.029, p-value<0.001, Heuertz *et al.* 2001) and Romania (F_{is}=0.014, p-value>0.05, Heuertz *et al.* 2003). However, Morand *et al.* (2002) have reported high and significant heterozygote deficiency for *F. excelsior* populations sampled in continuous forests across France. In particular, they estimated from a sample of 60 mature trees in three

north-eastern French populations using microsatellite markers, an overall F_{is} of 0.292 (p-value<0.001) but with a great range of values for single-locus estimates, from 0.077 (p-value>0.05) to 0.450 (p-value<0.001) at locus M2-30 (Morand *et al.* 2002).

Although unusually high values of hetorozygote deficiency for an outcrossing tree species may be difficult to explain (Morand *et al.* 2002), findings for *F. excelsior* occurring in continuous stands in France show that the behaviour of the severely fragmented remnants presented in this study is not atypical and therefore suggest that habitat fragmentation has had little influence on mating patterns within remnants.

As Morand *et al.* (2002) have pointed out, there are two main, non exclusive, hypotheses that may explain the overall excess of homozygotes detected in *F. excelsior* remnants of Moffat Dale. Primarily, if selfing or mating between close relatives occurs, a reduction of heterozygosity is expected (Charlesworth and Charlesworth 1987). As detailed below, a selfing rate of less than 1% estimated from open-pollinated progeny arrays suggest that *F. excelsior* remnants sampled in Moffat Dale are completely outcrossed. Biparental inbreeding is most likely to occur when increased spatial isolation and reduced population size have decreased the number of potential mates (Barrett and Kohn 1991). However, significant deficit of heterozygotes was only detected for the largest and least isolated of the three remnants (Table 4.3) and is highest for the largest population (BMa, $F_{is\text{-}est}$=0.318, p-value<0.05).

Alternatively, mutation and amplification patterns of the microsatellite markers used in this study may affect estimates. The issue of non-amplifying alleles at microsatellite loci has been addressed by several authors (e.g. Pemberton *et al.* 1995). The occurrence of such null alleles would result in identifying true heterozygotes as homozygotes at a particular locus and therefore an excess of homozygotes would be detected. Non-Mendelian segregation that could be explained by the presence of null alleles has been observed in previous studies (Heuertz *et al.* 2001; Morand *et al.* 2002) for several of the microsatellite markers developed for *F. excelsior* (Brachet *et al.* 1999; Lefort *et al.* 1999). In this study, non-Mendelian

segregation was observed on open-pollinated progeny arrays at two loci (3.1 and FEMSATL5) which display high null allele frequencies (Chapter 5 and Appendix A). Furthermore, significant heterozygote deficiency detected in three remnants of Moffat Dale (Table 4.3) over six microsatellite markers, reflects significant departure from random mating at only one of two loci, FEMSATL8 for remnant CDa, FEMSATL19 for Wa, or both for BMa. The occurrence of non-amplifying alleles at these two further loci cannot be rejected as a possibility (Appendix A). To investigate the potential impact of non-amplifying alleles on estimation of F_{is}, estimates presented in tables 4.3 and 4.4 were calculated again when, at each marker locus, homozygote individuals were recoded as heterozygotes for null allele and non-amplifying individuals as homozygotes for null allele. Such approach certainly overestimates the true frequency of null alleles as true homozygotes are then identified as heterozygotes and as lack of amplification may be caused by other factors (e.g. poor quality of DNA extract is a most commonly reported problem) but it also allows to discuss the issue in a quantitative manner. When performing such transformation, departure from random mating is not detected (significant at the 5% level) at any of the eight markers in any of five *F. excelsior* remnants and nor overall (mean f = -0.056 p>0.05). Such results suggest that occurrence of non-amplifying alleles may inflate the observed F_{is} value.

However, taking into account null alleles at loci 3.1 and FEMSATL5 is not enough to explain non-Mendelian segregations in the progeny arrays (presented in Appendix A). For such highly polymorphic markers, constituted of large sequences of mono or dinucleotide repeats, patterns of mutation may also affect *F*-statistics estimates. In particular, size homoplasy may be high (Balloux and Lugon-Moulin 2002). However, Estoup *et al.* (2002) have argued that F_{is} is virtually insensitive to size homoplasy and other mutation processes because most coalescent events for pairs of genes within subpopulations occur before the occurrence of mutation.

Genetic differentiation across the five *F. excelsior* remnants sampled in Moffat Dale was moderate (Θ=0.080) but significant (*p*-value<0.001) which suggests that gene exchange has been, to some extent, restricted among fragments. However, caution

must be taken when interpreting statistical significance of genetic differentiation based on markers with high mutation rates (Hedrick 1999). High levels of polymorphism (generated by mutation) provides great statistical power to the exact *G*-test (Goudet *et al.* 1996) performed herein, and very fine differences in allele frequencies are detected but they may not be biologically meaningful (Balloux and Lugon-Moulin 2002). In fact, the effective number of migrants per generation estimated with the private allele method of Barton and Slatkin (1986) is greater than 1 (Nm=3.48 after correction for sample size) suggesting that historical gene exchange among remnants has been sufficient to counteract significant genetic drift (Wright 1951).

2 Contemporary mating patterns within a deforested landscape.

F. excelsior has a complex, polygamous mating system and the maintenance of diverse sexual forms in the species has been poorly documented (Wallander 2001). In particular, little is known about the ability of hermaphrodite individuals to self (Morand-Prieur *et al.* 2003). Recent evidence from controlled crosses show that self-fertilisation in hermaphrodites is possible (Wallander 2001; Morand-Prieur *et al.* 2003) and that *F. excelsior* is at least partially self-compatible with a low selfing rate found in controlled experiments involving pollen competition (Morand-Prieur *et al.* 2003). Estimates of selfing rates within natural populations reported herein corroborate these results. It was found that in *F. excelsior* populations of Moffat Dale, outcrossing is clearly predominant. The selfing rate estimated from open-pollinated progeny arrays was less than 1%. Although *F. excelsior* appears to be self-compatible, the very low incidence of selfing in populations of Moffat Dale may be explained by an advantage of male pollen over hermaphrodite pollen as suggested from controlled pollination investigating pollen competition (Wallander 2001; Morand-Prieur *et al.* 2003) or by very early acting inbreeding depression.

Estimates of contemporary pollen immigration into Moffat Dale corroborate the information inferred from spatial genetic structure of mature trees that historical gene exchange between remnants has not been limited (Nm>1). Contemporary pollen-mediated gene flow into the catchment is extensive suggesting that gene exchange

between remnants occur at a much larger scale than the sampled area (>20 km^2), thus maintaining genetic diversity in newly produced individuals across remnants.

Within Moffat Dale, nearly half of the 88 trees sampled contribute, on average, to the pollen pool of mother trees (N_e=48%) but their male mating success decreases with distance (β_{est}=-0.00181±0.00021, *p*-value<0.001) and effective pollen dispersal is limited, on average, to 328 m. The estimate of effective pollen dispersal distance in this study is at least 2.5 times greater than the distance inferred from the spatial genetic structure of *F. excelsior* trees in a population occurring in a continuous forest in Romania (70 m<σ_p<140 m; Heuertz *et al.* 2003). Such an increase in pollen dispersal distance in a deforested catchment may be the result of facilitation of airborne pollen movement in an open landscape as has been suggested for fragmented populations of the wind pollinated tree species, *Acer saccharum*, (Fore *et al.* 1992; Young *et al.* 1993).

Indeed, in a more spatially isolated remnant, SCa, distant from other fragments by 2210 m (Figure 4.1) and surrounded by a 30-year-old conifer plantation, the number of effective pollen donors from the catchment is reduced by two thirds (N_e=19%) and pollen dispersal is mainly restricted to trees within SCa (m_p= 370 m). Such estimates are strikingly lower than for the two remnants occurring in a bare and open landscape and would suggest further that pollen movement in Moffat Dale is greatly influenced by tree density within remnants and in the landscape surrounding them. A slight reduction in genetic diversity, for both adults and seed cohorts in this remnant, would suggest that when habitat fragmentation isolates small populations by producing a canopy closure, pollen-mediated gene flow may be limited, potentially creating a genetic bottleneck.

The neighbourhood model (Burczyk *et al.* 2002) has proved to be a useful framework for describing mating patterns for *F. excelsior* remnants in Moffat Dale. The contribution of self, immigrant pollen, and pollen from the 88 trees sampled within the catchment to fertilisation of progeny sampled from all trees producing fruits in the year of sampling in three of five forest fragments, were estimated

without requiring prior knowledge of the allelic composition of the immigrant pollen pool (Burczyk *et al.* 2002). The model - fitting approach adopted by NEIGHBOR v.2.0, provides robust estimates of average selfing rate, average values of pollen flow into the catchment, together with an estimate of the overall relationship between pollen dispersal and distance. An alternative method of analysing the microsatellite data from parents and their offspring is to conduct a paternity analysis in which the potential paternal parent of individual seeds is determined. By combining data from many such paternity assignments, a more detailed picture of the distribution of pollination events within remnants, between remnants, and into the catchment can be built up. In the next chapter, a paternity analysis has been applied to the seeds collected from mother trees of known genotype to build up this more detailed picture of the pollination process for *F. excelsior* in Moffat Dale.

Appendix 4.1: Effective pollen dispersal distance and male neighbourhood size for 19 *F. excelsior* mother tree sampled in three remnants of Moffat Dale.

		effective pollen dispersal in Moffat Dale		
mother tree	sample size[†]	mean distance between mates[‡] (m)	mean effective pollen dispersal distance[#] (m)	mean effective male neighbourhood size N_e[§] (%)
CDa102	22	671	242	60.99%
CDa105	29	665	243	61.61%
CDa108	23	665	238	61.22%
CDa111	29	665	240	61.41%
CDa112	24	666	239	61.29%
CDa118	27	666	237	61.06%
CDa123	23	669	239	60.97%
CDa125	30	669	273	63.60%
CDa126	29	667	265	63.15%
CDa129	26	666	237	61.64%
CDa130	21	670	240	60.98%
CMa24	20	1030	664	56.80%
CMa26	11	1016	648	56.81%
SCa33	22	1768	368	19.27%
SCa34	2	1773	368	19.21%
SCa35	28	1768	366	19.26%
SCa36	24	1757	372	19.43%
SCa38	3	1750	375	19.55%
SCa41	29	1753	373	16.97%

Effective pollen dispersal parameters obtained in NEIGHBOR v.2.0 (Burzyck *et al.* 2002) from male reproductive success of 87 trees sampled within a 3000 m- radius neighbourhood based on mating parameters estimated across all neighbourhoods for 19 trees sampled for seed (N=422, Table 5.5). [†]Number of seeds genotyped at three microsatellite loci. [‡]Mean distance between mother tree and 87 potential pollen donor sampled within a 3000 m-radius neighbourhood (i.e. Moffat Dale). [#]Effective pollen dispersal distance within a 3000 m-radius neighbourhood. [§] N_e expressed as the percentage of the 87 trees located within the neighbourhood actively contributing to the pollen pool of the mother tree.

Chapter 5 : Landscape approach to contemporary gene dispersal of *Fraxinus excelsior* L. I. Paternity analysis of pollen flow in forest fragments

Abstract

Estimation of contemporary gene dispersal in tree species is essential to determine whether habitat fragmentation results in genetic isolation of populations through limited gene flow. In particular, investigation of the influence that spatial isolation and remnant size have on population dynamic genetic processes of wind pollinated temperate tree species is required. Microsatellites markers were used and exclusion and likelihood-based paternity analyses were performed on progeny arrays in order to quantify pollen gene flow and to describe patterns of pollen dispersal in three remnants of *Fraxinus excelsior* (Oleaceae) in a heavily grazed catchment in southern Scotland. The applicability of such methods of paternity analysis to natural populations is discussed. For *F. excelsior*, at least 54.3% of pollen immigration was detected into the catchment and extensive pollen flow was found in all remnants representing up to 100% of effective pollination in isolated trees in open pastures. Pollen dispersal over a distance of 2.9 km was recorded. Population density rather than spatial isolation seems to affect quantitative estimates of gene flow at the landscape scale. Remnants appear to be part of a much wider network of reproductively active trees over tens of kilometres.

Keywords: *Fraxinus excelsior* L., microsatellites, paternity analysis, likelihood, pollen flow, fragmented landscape.

I Introduction

Understanding the dynamics of pollen and seed dispersal in tree species is essential for forest geneticists and conservation biologists to predict the impact of management and anthropogenic disturbances on population genetic diversity (Young *et al.* 1996). As Sork *et al.* (1999) have pointed out, this may be achieved by the study of gene flow processes considering an ecological (rather than evolutionary) time scale and integrating the features of the landscape in which they occur.

Traditional methods that indirectly infer the effective number of migrants from population structure only give an average figure of historical gene flow under the rather unrealistic assumptions of the island model (Whitlock and McCauley 1999; Sork *et al.* 1999). Alternatively, direct methods of estimation of pollen or seed movement using mark and recapture techniques (Campbell and Waser 1989; Wells and Young 2002) or tracking of pollinator movement (reviewed in Levin and Kerster 1974) present important limitations. These methods estimate dispersal from a source plant and as such reflect potential rather than effective gene dispersal (Cain *et al.* 2000) and may also greatly underestimate rare and long distance dispersal events (Campbell 1991). Fortunately, the availability of highly polymorphic molecular markers coupled with the advent of new statistical methods have recently revolutionised the study of contemporary gene flow in plants (Sork *et al.* 1999).

Paternity analysis of pollen dispersal in natural populations has been reported for several tree species (Schnabel and Hamrick 1995; Schuster and Mitton 2000; Konuma *et al.* 2000). High exclusion probabilities obtained from microsatellite markers have allowed the detailed patterns of pollen gene flow in *Quercus macrocarpa* (Dow and Ashley 1998), *Q. robur* and *Q. petrea* (Streiff *et al.* 1999) to be described by simple exclusion analysis of paternity. However, there has been increasing recognition that even when full exclusion is possible, paternity analysis of pollen flow by simple exclusion has major drawbacks (Oddou-Muratorio 2002; Hardy *et al.* submitted). In particular, the method does not resolve paternity when multiple males are not excluded but most importantly, it does not account for the lack of reproductive isolation of the sampled populations (Marshall *et al.* 1998; Nielsen *et*

al. 2001). To address these difficulties, refined methods of paternity analysis using a likelihood approach have been developed recently to assess the statistical confidence of paternity assignment in non-isolated populations (Marshall *et al.* 1998; Nielsen *et al.* 2001), and especially to estimate gene flow in plants (Gerber *et al.* 2000).

Classical views that gene flow in plants is limited (Levin and Kerster 1974) and that habitat fragmentation results in increased population isolation (Andren 1994) are being challenged by empirical evidence from insect-pollinated neotropical tree species (Nason and Hamrick 1997). In particular, spatially isolated pasture trees have been found to be reproductively dominant in *Symphonia globulifera* (Aldrich and Hamrick 1998) and to act as stepping stones for gene flow in fragmented populations of *Pithecellobium elegans* (Chase *et al.* 1996), *Dinizia excelsa* (Dick 2001) and *Swietenia humilis* (White *et al.* 2002). However, there is little evidence of how wind-pollinated tree species respond to habitat fragmentation (Kaufman *et al.* 1998) especially in temperate zones where interest has been so far mainly focused on continuous stands (Dow and Ashley 1998; Streiff *et al.* 1999 but see Young *et al.* 1993; Schuster and Mitton 2000; Sork *et al.* 1999; Sork *et al.* 2002).

The objective of this study is twofold. First, the applicability of paternity analysis methods to natural tree populations is considered. Second, the influence of habitat fragmentation and landscape heterogeneity on contemporary gene flow for a wind pollinated temperate tree species is addressed.

In southern Scotland, as part of a local initiative, large areas of upland pasture have recently been fenced to release grazing pressure from severely fragmented forest remnants (detailed description in Chapter 4) and a native woodland restoration programme involving extensive planting has been initiated (Wildwood Group of the Borders Forest Trust 2000). The programme intends to collect planting stock from local remnants. There is therefore an immediate need to assess the genetic quality of seed produced in such impoverished stands.

F. excelsior L. (Oleaceae), common ash, is a wind pollinated wind dispersed species with a complex polygamous mating system (Wallander 2001). Sex expression varies at the floral, inflorescence and tree level, as well as between years, and individuals may be classified phenotypically into a continuum from purely male to purely female with a whole range of hermaphroditic intermediates (Wallander 2001). Hermaphrodite flowers are protogynous, self-fertilisation seems possible (Wallander 2001) but the functional gender of hermaphrodites remains unclear (Morand-Prieur *et al.* 2003). Regular fruit bearing usually begins around 20 years of age and the species displays a masting behaviour with irregular fruiting between years (Tapper 1996).

In this study, microsatellite markers are used to screen open-pollinated progeny arrays collected from *F. excelsior* trees in a range of remnant size and isolation to describe detailed patterns of effective pollen dispersal and individual male reproductive success. In particular, I address how small population size, spatial isolation and features of the surrounding landscape may affect quantitative and qualitative aspects of gene flow by pollen in this wind pollinated temperate tree species.

II Material and methods

1 Study site and sampling strategy

The study site is the Moffat Dale, previously described for *Sorbus aucuparia* (Chapter 3) and for *Fraxinus excelsior* forest remnants (Chapter 4). This study focuses on three remnants (Figure 5.1), two in Carrifran (CDa and CMa) and one in Swine Cleuchs (SCa), which differ in their density, population size (Table 5.1) and degree of spatial isolation (minimum distance to other remnants 412 m, 497 m, 1613 m and mean distance to other remnants 1093 m, 1286 m, 2210 m respectively).

In these three remnants, all mature trees were mapped and sampled for leaf material. In addition, seeds were sampled from all trees bearing fruits in 2000 (Table 5.2). Two additional populations were surveyed, Whitewells (Wa) and Spoon Burn (BMa), and leaf material was taken from a representative sample of 20 randomly distributed trees as well as from two isolated trees, A and B (Figure 5.1). The latter two populations (BMa, Wa) are considered as potential sources of immigrant pollen for the studied remnants and will be referred to as 'source populations'. This sampling strategy was adopted to determine how two major components of habitat fragmentation (i.e. reduced population size and increased spatial isolation) affect estimates of the source and quantity of pollen flow.

Sampled individuals were tagged with permanent labels and their spatial position was recorded as described in chapter 4.

Table 5.1: location, abbreviated name, landscape characteristics, altitude and size of five *F. excelsior* remnants present in Moffat Dale.

Location	Abbreviation	Site characteristics	Altitude	Size (N)#
Carrifran (down)	CDa	grazed riverside	188m	30‡
Carrifran (middle)	CMa	grazed riverside	212m	4‡
Swine Cleuchs	SCa	grazed riverside in conifer plantation	271m	12‡
Whitewells	Wa	protected riverbank	170m	~50*
Spoon Burn	BMa	grazed riverside	262m	50-70*

#Size expressed as the approximate number of mature trees in the population. ‡Total number of mature trees in remnant. *Leaves collected from 20 individuals, sampled as pollen source for trees sampled for seed in CDa, CMa and SCa.

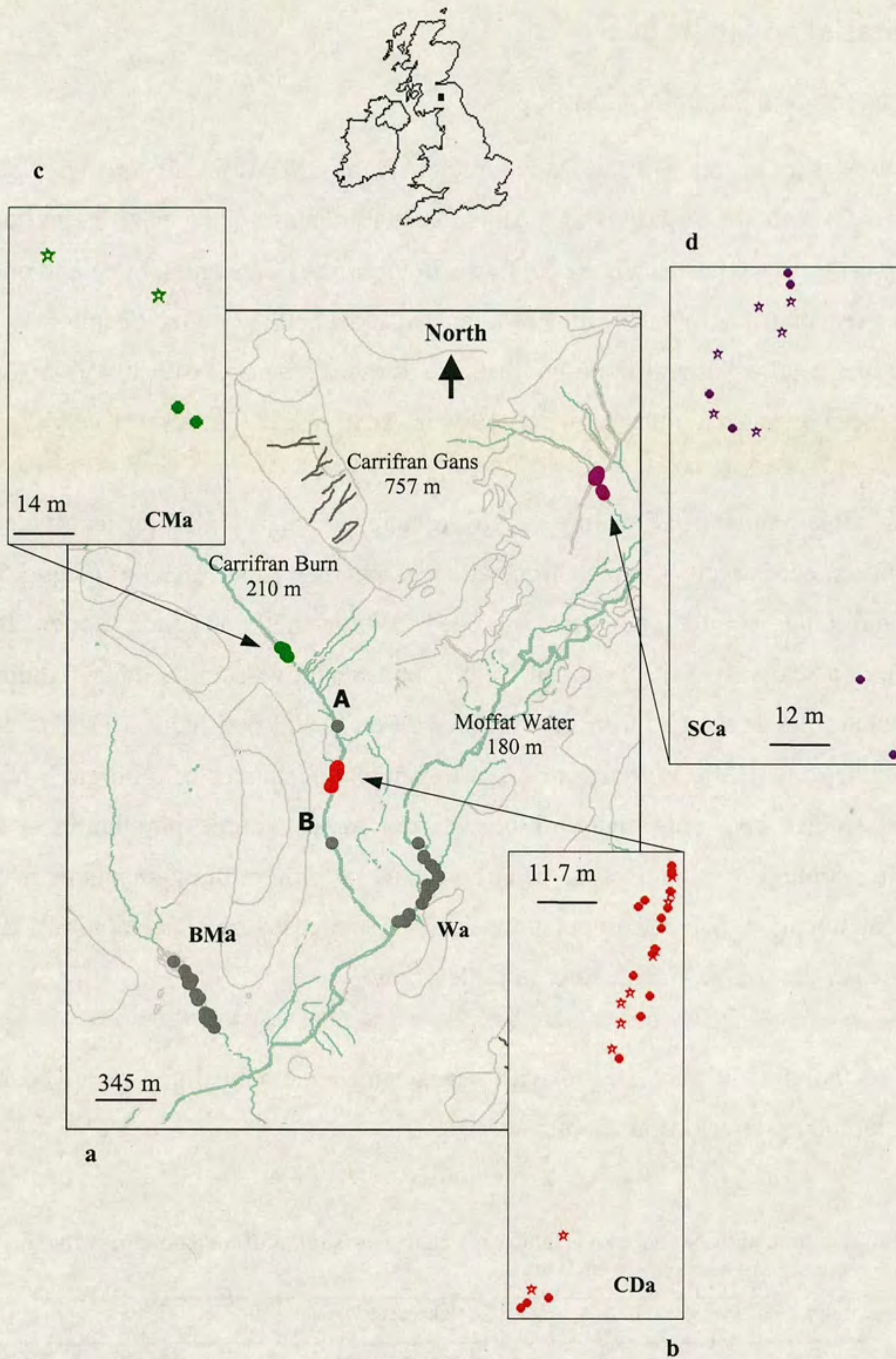


Figure 5.1: Distribution of *F. excelsior* individuals sampled in Moffat Dale remnants.
a) Trees grouped into populations as described in Table 5.1. Each dot represents a tree. The background map is a section of Ordnance Survey product Land-Line.Plus-nt11©Crown copyright Ordnance Survey. An EDINA digimap / JISC supplied service. b, c & d) Close up of spatial distribution of individuals in remnant CDa, CMa and SCa respectively. All individuals producing fruits in 2000 are represented by a star, 50-200 seeds were collected from each of them.

2 Tree gender and sample collection

In autumn 2000, the gender of all mature trees in CDa, CMa and SCa was crudely estimated during field observations as follows; trees bearing no fruits were recorded as male, those with only a few branches bearing fruits were recorded as predominantly male and those with a good seed crop were recorded as predominantly female. 50-200 fruits were collected from all trees producing fruits (Table 5.2). Whenever possible, fruits were collected from several, if not all, branches throughout the canopy in an attempt to adequately reflect the tree pollen pool in the sample. Phenotypic gender of all trees was also observed during flowering in spring 2001 but may not be representative of sex expression in 2000. Since there is a continuum of sex expression in the species (Wallander 2001), these observations were used only as an indication of individual investment in male and female functions. However, no tree was excluded as potential male parent on the basis of these observations. Fruits were dried and kept in a cool, dry place until DNA extraction. In spring 2001, young leaves were collected from all mature trees in CDa, CMa and SCa and from a sample of 20 trees in Wa and BMa and were immediately stored at -80 °C until DNA extraction.

Table 5.2: Number of individuals producing seed and total number of seeds sampled in 2000 in each remnant and overall.

Population	Size	Number of individuals producing seed in 2000	Seed sample
CDa	30	11	330
CMa	4	2	32
SCa	12	6	131
total	**46**	**19**	**493**

3 DNA isolation

Protocols for isolation of genomic DNA from frozen leaf for adult trees and from seed are described in chapter 4. In total, DNA was isolated from 88 adult trees of which 19 had also been sampled for seed (referred to as mother trees). For each mother tree, DNA was extracted from a sample of 30 seeds, or all seeds if the seed crop was less, bringing the total number of seed sampled to 493 (Table 5.2).

4 Microsatellite amplification and genotyping

Five microsatellite loci previously characterised for *F. excelsior,* 3.1, M2-30B, 1.19 (Brachet *et al.* 1999) and FEMSATL2, FEMSATL5 (Lefort *et al.* 1999) were successfully amplified by polymerase chain reaction (PCR) and chosen for their scorability and high level of polymorphism. Amplification conditions and genotyping analysis are as described in Chapter 4.

5 Data analysis

Diversity & independence of loci. The number of alleles observed (*k*) was recorded for each locus. Concern has been expressed regarding the occurrence of non-amplifying alleles at microsatellite loci (Pemberton *et al.* 1995). Therefore, null allele frequency was estimated, under the assumption of Hardy-Weinberg equilibrium, for each locus in CERVUS 2.0 (Marshall *et al.* 1998). Moreover, likelihood approaches to paternity analysis using multilocus genotypic data assume linkage equilibrium among loci. Previous analysis revealed no significant genotypic linkage disequilibrium between any pair of microsatellite loci for the sample (Chapter 4).

6 Paternity analyses

Exclusion probability. An exclusion probability is a population statistic that measures the power of a given data set to exclude a given relationship (Gerber *et al.* 2000). The importance of the exclusion probability to paternity assignment is that an increase in exclusion probability increases the probability of paternity among the set of non-excluded parents (Devlin *et al.* 1988). Paternity exclusion probabilities (i.e. probability of excluding an individual as the father of an offspring when the genotype of the mother is known) were computed for each locus following Jamieson and Taylor (1997) and Gerber *et al.* (2000) in FAMOZ (version released on 30th January 2003, Gerber *et al.* 2003).

Methods of assignment. Paternity analyses were undertaken for the seeds sampled from 19 mother trees in three *F. excelsior* remnants. All 88 mature individuals

sampled both inside and outside the remnants were considered as potential male parents (including self). Outcomes of paternity assignment may be that progeny remain unassigned or that at least one individual is identified as a father among the sampled individuals. Paternity was assigned following three different methods. 1) a simple exclusion analysis using FAMOZ, 2) a maximum likelihood approach to estimate the proportion of offspring fathered by non-sampled individuals using FAMOZ and 3) a Bayesian approach to fractionally assign paternity to a subgroup of non-excluded potential fathers using PATRI (Signorovitch and Nielsen 2002).

While simple exclusion is a useful starting point for paternity inference (Marshall *et al.* 1998), statistical approaches are necessary to assess the confidence in paternity assignment. In particular, it is virtually impossible to sample all adults contributing to reproduction in natural tree populations. It is therefore necessary to assess the risk of excluding a candidate as the true father on the sole grounds that it has not been sampled (Marshall *et al.* 1998; Nielsen *et al.* 2001). Genotyping error or mutation at marker loci may also result in erroneous paternity assignment. It is moreover necessary to resolve paternity in cases where are several possible candidate fathers.

In FAMOZ, determining confidence levels for rejecting a male with a compatible genotype as a potential father relies on a simulation procedure detailed in Gerber *et al.* (2000). The most-likely fathers were detected by means of 'log of the odds' ratios (LOD scores, Meagher, 1986; Gerber *et al.* 2000) using population allele frequencies estimated from seed and adult multilocus genotypic data pooled together (N=510). No significant genotypic association was detected among any pair of microsatellite loci for the sampled trees (Chapter 4), therefore, LOD scores over all loci were obtained by adding LOD scores calculated for each locus (Gerber *et al.* 2000).

The distribution of the LOD scores of the most-likely fathers of 10,000 randomly generated offspring with their father sampled (i.e. by randomly choosing a father among the genotyped trees) was compared to the distribution of the LOD scores of the most-likely fathers of 10000 offspring whose genotypes were randomly generated according to allele frequencies for the whole sample. The test threshold for

rejecting a candidate as the true father (TP) was chosen at the intersection of the two distributions of LOD scores to minimise both type I error of wrongly rejecting a candidate as the true father and type II error of wrongly accepting a candidate as the true father (Gerber *et al.* 2000).

Paternal inference in PATRI is an extension of the fractional likelihood assignment method first introduced by Devlin *et al.* (1988). However, the originality of the method is that isolation of the studied population (i.e. complete sampling of breeding males) is not a required assumption (Nielsen *et al.* 2001). PATRI estimates the posterior probability that a particular male has sired a particular offspring given the size of the total breeding male population (Nielsen *et al.* 2001; Signorovitch and Nielsen 2002).

Source of pollen. Results of paternity assignment by simple exclusion and by maximum likelihood analyses in FAMOZ were compared and used to describe patterns of pollen dispersal in the sampled remnants of *F. excelsior.* On the one hand, the proportion of unassigned offspring provided an estimate of apparent pollen flow in Moffat Dale (i.e. proportion of pollen originating from unsampled individuals). On the other hand, when offspring were assigned at least one father among the sampled trees, the following were determined: 1) the proportion of local pollen, immigrant pollen from unsampled individuals and immigrant pollen from trees in surrounding remnants contributing to individual pollen pools for each of the 19 mother trees sampled; 2) the maximum number of seed sired by each sampled individual; 3) the maximum number of sires identified among the sampled individuals for each mother tree.

Pollen dispersal distance. The actual number of pollination events was determined as a function of the distance between seed maternal parent and the identified pollen source and compared to the total number of potential male parents as a function of their distance to each mother tree.

Fractional paternity assignment to remnants. PATRI was used to fractionally assign paternity in the five *F. excelsior* remnants sampled in Moffat Dale. Adult trees sampled (including maternal parents) were categorically assigned to five male groups according to their remnant origin. Although the total breeding male population size is unknown, it seems reasonable to assume that up to 150 individuals can potentially contribute to the pollen pool as this is the approximate number of trees identified within Moffat Dale (Table 5.1). To account for the uncertainty around the actual size of the total male breeding population, I chose to model the effective male population size as a uniform function with a lower bound of 100 and an upper bound of 500. Results that assume an effective male population size of exactly 150 individuals are also discussed. I found this fractional approach most useful for studying the paternal contribution to offspring for *F. excelsior* forest remnants in Moffat Dale. Rather than identifying individual paternity *per se,* I wish to highlight population patterns of paternity (Devlin *et al.* 1988; Schnabel 1998) and especially differential contribution to effective pollen dispersal by the remnants of *F. excelsior* studied.

Comparison of gene flow between remnants. To assess whether habitat fragmentation affects estimates of the source and quantity of pollen flow, I estimated and compared the proportion of offspring not pollinated by local trees for each of the three remnants. The proportion of immigrant pollen attributable to identified trees in surrounding remnants was also estimated.

Estimation of cryptic gene flow. Pollen flow events may be undetected if gametes are the product of non sampled individuals that cannot be genotypically distinguished from gametes produced by male candidates. The confidence level of inference is already indicative of the probability of erroneously attributing paternity to a sampled male. However, a simulation based test undertaken in FAMOZ (Gerber *et al.* 2003) allows the estimation of the proportion of cryptic to total gene flow for a sample data set. The simulation estimates the actual proportion of offspring fathered with immigrant pollen (m_p) compared to apparent pollen flow (m_a) i.e. the proportion of offspring with no father identified among the sampled trees. Input data were chosen to resemble the actual sampling in the *F. excelsior* remnants of Moffat Dale.

Simulated data are based on the observed allele frequencies calculated for the sampled adults and seed. 25 offspring per mother were simulated for 19 mothers selected at random and 88 candidate male parents. The test requires input of an estimate of the effective male population size (N_p). N_p equal to 150 was used as a starting point for simulation (as has been used for PATRI) and verified that estimates of 'true gene flow' i.e. m_p, with this value for N_p (or others), were equal to the 'expected gene flow', i.e. $(N_p-88)/N_p$ the proportion of offspring expected to be fathered by the non-sampled individuals contributing to the pollen pool (Hardy *et al.* submitted). The test threshold for attributing paternity of simulated offspring to local or immigrant pollen was determined using the maximum likelihood approach described above. Simulations were repeated ten times in each case to crudely determine the variance of the estimate.

III Results

1 Choice of markers and introduction of error rate

Choice of markers. The number of alleles (k) observed at each locus was high. k varies between ten at locus 1.19 and 56 at locus FEMSATL2 (Table 5.3). Null allele frequencies estimated at Hardy-Weinberg equilibrium were especially high at loci 3.1 and FEMSATL5 (Table 5.3).

Table 5.3: Sample size (N), number of alleles sampled (k), gene diversity (H_e), and allelic richness (R) at five microsatellite markers including mature individual and progeny arrays sampled in each of five *F. excelsior* remnants of Moffat Dale.

		Population						
locus[a]		**CDa**	**CMa**	**SCa**	**BMa**	**Wa**	**Overall**	**Null frequency**[b]
3.1[‡]	N	338	30	115	16	12	511	**0.418**
	k	12	9	13	7	6	17	
	H_e	0.825	0.761	0.75	0.779	0.754	0.848	
	$R^{\dagger}$	6.266	6.685	5.931	6.772	6.000	7.447	
M2-30B[‡]	N	344	34	136	19	20	553	0.126
	k	27	19	27	12	15	35	
	H_e	0.877	0.840	0.862	0.92	0.908	0.929	
	R	9.876	10.608	10.393	10.766	12.198	13.112	
1.19[‡]	N	351	35	129	18	18	551	0.141
	k	10	8	8	6	7	10	
	H_e	0.727	0.804	0.45	0.611	0.825	0.782	
	R	7.058	6.863	4.113	5.514	6.543	7.125	
FEMSATL2[&]	N	357	35	143	20	20	575	0.048
	k	42	19	31	19	21	56	
	H_e	0.911	0.833	0.872	0.947	0.958	0.931	
	R	12.714	10.128	10.225	14.799	15.639	14.320	
FEMSATL5[&]	N	354	31	126	19	19	549	**0.332**
	k	30	17	24	12	16	46	
	H_e	0.810	0.816	0.869	0.927	0.942	0.899	
	R	9.612	10.547	10.576	10.562	12.981	12.717	

[†]Measure of allelic richness based on a minimum sample size of 12 diploid individuals. Estimations were performed in FSTAT 2.9.3 (Goudet 2001) and GENETIX 4.01 (Belkhir *et al.* 1998). [a]Microsatellite sequences described in [&]Lefort *et al.* (1999) and in [‡]Brachet *et al.* (1999). [b]Null allele frequency calculated on adult sample only (N=88) assuming Hardy-Weinberg equilibrium, in CERVUS 2.0 (Marshall *et al.* 1998).

The occurrence of a null allele at locus FEMSATL5 has already been suggested by Morand *et al.* (2002). Furthermore, non-Mendelian segregation was observed at loci 3.1 and FEMSATL5 even when taking into account null alleles at those loci, as detailed in Appendix A. Therefore, loci 3.1 and FEMSATL5 were discarded from further analysis. No genotypic association (tested significant at the 5% level) was established among any pair of loci for FEMSATL2, M2.30B and 1.19 (Chapter 4)

and detailed paternity analyses were accordingly carried out using multilocus data combining genotypes across the three lattermost loci.

Accounting for genotyping error. A number of progeny displayed a genotype not compatible with their maternal parent at one or more loci. The combined error rate (e) estimated in CERVUS was 0.0420 (Appendix A). Sources of error which may include sampling, genotyping and actual mutation, are discussed in detail in Appendix A. Progeny with incompatible genotypes were therefore discarded from the data set for further analysis.

However, to account for the possibility of genotyping error or mutation at marker loci, I chose to bin rare alleles with more common alleles of the nearest size. Alleles were deemed rare when they occurred at a frequency of less than 0.01 for M2.30B and FEMSATL2 (Appendix A). At locus 1.19, a procedure reflecting difficulties in gel scoring was used (Appendix A). The procedure reduced the number of alleles observed at loci FEMSATL2, M2.30B and 1.19 to 29, 21 and seven respectively (Table 5.4). Unbinned and binned data sets were considered for paternity analyses and results for both are described for comparison.

Table 5.4: Sample size (*N*) and number of alleles observed (*k*) before and after binning of alleles at the three loci used for paternity analysis of *F.excelsior* in Moffat Dale

Locus	*N*	*k* before binning	*k* after binning
FEMSATL2	505	54	29
M2.30B	486	35	21
1.19	484	10	7

Procedure for binning of alleles detailed in Appendix A.

Accounting for non-amplifying alleles. Considering this reduced data set, 35 seeds and 14 of the sampled adult trees failed to amplify at one or two of the three loci used. Lack of amplification may be caused by technical problems including failed PCR and poor quality of DNA extract (which is most likely when non-amplification is observed across several loci) but also by the occurrence of null alleles (Pemberton *et al.* 1995). To investigate how such null alleles may affect outcomes of paternity

analysis and estimates of pollen flow, analyses were also performed on the binned dataset modified so that homozygote individuals were recoded as heterozygote null and failed amplifications were recoded as homozygote null.

2 Paternity analysis

Choice of methods. The paternity exclusion probability (*PEP*) over the three loci FEMSATL2, M2.30B and 1.19 was high (>0.99) and remained high after binning alleles at all loci (>0.987, Table 5.5). This provided great confidence in paternal assignment and left only a small fraction of offspring, between 8% and 36% - depending on the method and dataset considered - with unresolved (i.e. multiple) paternity (Table 5.6). In PATRI, missing values are not permitted. To obtain comparable results, simple exclusion and likelihood analyses were also therefore performed on a reduced data set of 387 seeds and 74 candidate male parents. When including individuals with missing values, confidence in paternity assignment is weakened (13% to 36% unresolved paternity) as candidates may not be excluded as fathers on the only basis that information is missing and pollen flow may be greatly underestimated. However, it is then possible to account for the potential maximum paternal contribution of the 14 trees excluded from other analysis. Overall, comparison of apparent pollen flow estimates obtained with several approaches from least conservative (maximum likelihood excluding individual with incomplete genotypes) to most conservative (simple exclusion accounting for genotyping error (binned alleles) and maximum occurrence of null alleles) allows to quantify the uncertainty around apparent pollen flow estimates due to technical factors.

Table 5.5: Paternity exclusion probability (*PEP*) at each microsatellite locus and combined over loci for *F. excelsior* adults and seed sampled in Moffat Dale.

Locus	*PEP* without allele binning	*PEP* with allele binning‡	*PEP* with allele binning and null†
FEMSATL2	0.857075	0.851047	0.864281
M2.30B	0.851870	0.836928	0.844027
1.19	0.586746	0.505916	0.608445
Cumulated	**0.991251**	**0.987999**	**0.991711**

PEP calculated in FAMOZ (Gerber *et al.* 2003) using allele frequencies estimated when including adult and progeny multilocus genotypes. ‡Procedure for binning of alleles described in Appendix A. †Maximum estimate of null allele occurrence obtained by recoding all homozygote genotypes as heterozygote null and non-amplifications as homozygote null.

Apparent pollen flow into Moffat Dale. Gene flow by pollen of *F. excelsior* into Moffat Dale is remarkably extensive. Selfing was the only explanation for fertilisation outcome of 1 to 3 (depending on the estimation method considered) seed out of 387. The high exclusion probability of the dataset allowed to attribute at least 38.6% and up to 72.9% of paternity to pollen of unknown origin (Table 5.6).

Table 5.6: comparison of overall results of paternity assignment and percentage of apparent pollen flow for *F. excelsior* into Moffat Dale for different methods of analysis.

Method	N_{seed} tested	N_{ad} tested	TP	*p*-value	N_{seed} excluded	N_{seed} resolved	N_{seed} unresolved	Apparent pollen flow
SE[a],MV[b]	422	88	-	-	229	105	88	54.3%
SE	387	74	-	-	260	89	71	67.2%
SE,MV,BIN[c]	422	88	-	-	163	107	152	38.6%
SE,BIN	387	74	-	-	219	115	53	56.6%
SE,BIN,NUL[d]	422	88	-	-	239	114	69	56.6%
ML,MV	422	88	2.47	<0.05	268	96	58	63.5%
ML[e]	387	74	3.37	<0.01	282	74	31	72.9%
ML,MV,BIN	422	88	3.04	0.05	204	110	109	48.3%
ML,BIN	387	74	2.25	<0.01	221	114	52	57.1%
ML,BIN,NUL	422	88	2.43	<0.01	244	111	67	57.8%

Paternity assignments expressed as the number of seed (N_{seed}) for which no father among the adults sampled (N_{ad}) was found (excluded), or unique paternity (resolved) or multiple paternity (unresolved) were identified with each method. [a]SE simple exclusion analysis. [b]MV including individuals with missing values at one or two loci. [c]BIN binning alleles to account for the potential of genotyping error. [d]NUL accounting for maximum occurrence of null alleles. [e]ML maximum likelihood analysis. Confidence level (*p*-value) and threshold (TP) for assignment test determined using a simulation procedure in FAMOZ (Gerber *et al.* 2003).

Origin of pollen of unknown origin may be either unsampled trees in the source populations, BMa and Wa, or trees occurring outside the sampled area. When including individuals with incomplete genotypes a simple exclusion analysis provided an estimate of apparent pollen flow into the catchment of 54.3%, when alleles have been binned together, this estimate is greatly reduced (38.6%). With no missing values, 67.2% of pollination events were attributed to unsampled individuals by simple exclusion while allele binning only decreased estimates slightly (56.6%). A maximum likelihood analysis in FAMOZ assigned paternity with great confidence

(p-value<0.01) at the chosen threshold for paternity (TP=3.36) with apparent pollen flow into the catchment estimated to be as high as 72.9%. When binning alleles, the maximum likelihood analysis gave results nearly identical to the simple exclusion (apparent pollen flow=57.1%, TP=2.25, p-value<0.01) with 114 seed out of 387 assigned a unique father among the trees sampled allowing identification of exact distances between the maternal parent and the pollen source. Accounting for occurrence of null alleles do not modify estimates.

Although there is variation in individual assignments depending on the method chosen, trends in estimates are similar. Bearing this in mind and for the sake of clarity in the presentation, some further illustrations (Figures 5.2, 5.3, 5.4 and Table 5.7) are based on results obtained with the maximum likelihood analysis with allele binning and excluding individuals with incomplete genotypes.

Remnant contribution to individual pollen pools. When considering individual pollen pools for each of the 19 mother trees in the sampled populations of Moffat Dale, it was found that most fertilisation events could be attributed either to immigrant pollen of unknown origin (i.e. from trees outwith the remnant that were not genotyped) or to local pollen (i.e. from genotyped trees in the same remnant). However, for most mother trees, a few seeds were fathered by genotyped trees located in the neighbouring remnants, especially for CMa, where no seed was found to be fertilised locally (Figure 5.2).

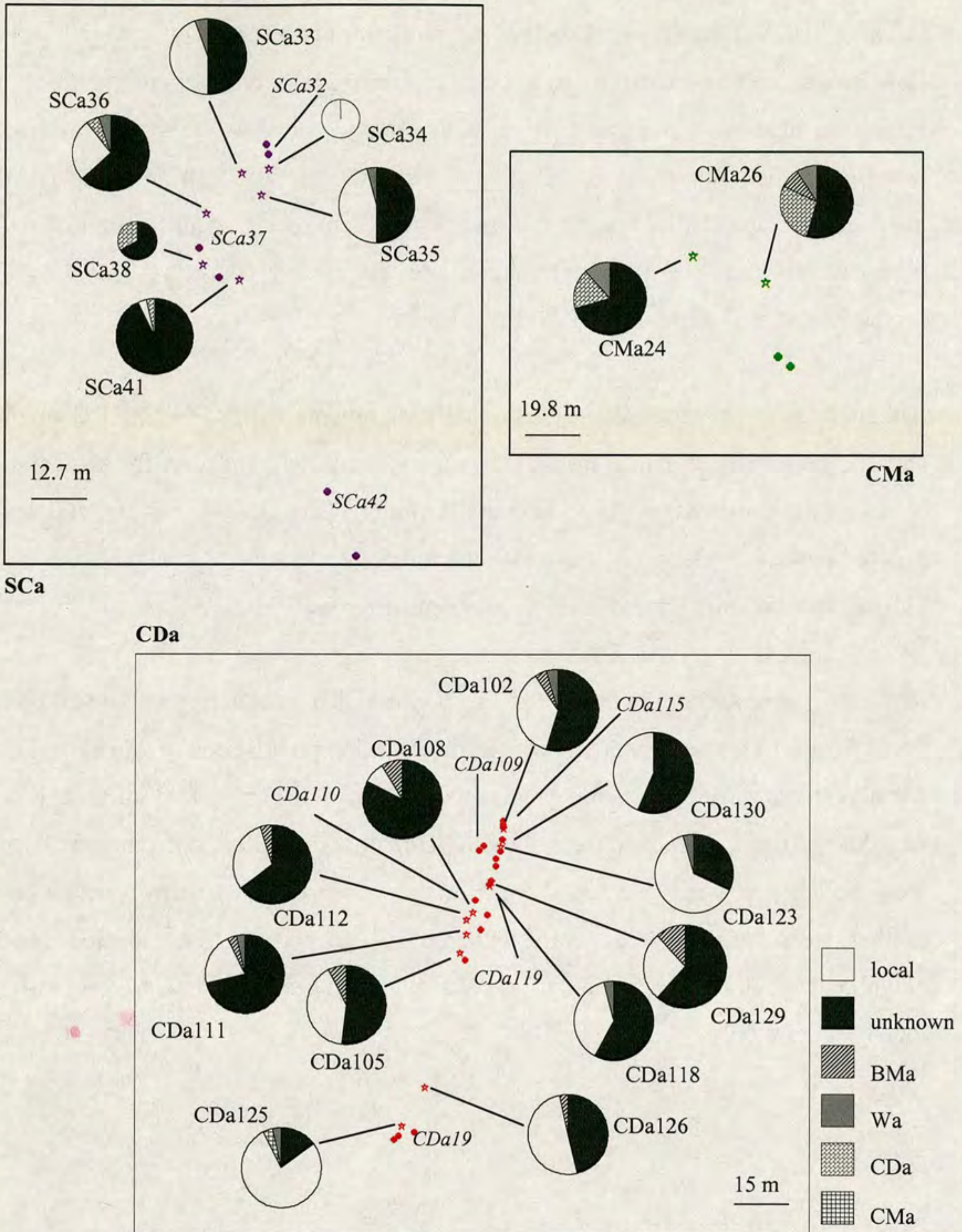


Figure 5.2: Remnant contribution to individual pollen pools according to the spatial position of *F. excelsior* mother trees sampled in remnants CDa, CMa and SCa.
Mother trees are labelled and represented by a star, other mature individuals are represented by a dot. Size of the pies is approximately proportional to the number of seeds sampled. Slices of pies represent the percentage of seed sired by pollen from different sources: pollen within the remnant, immigrant pollen of unknown origin, immigrant pollen from known origin in sampled remnants and source populations. Results of paternity assignment by maximum likelihood method in FAMOZ (Gerber *et al.* 2003) with allele binning excluding individuals with incomplete genotypes. For unresolved paternity (more than one candidate male non-excluded), paternity was attributed to the tree nearest to the mother tree. Seed crops were small for SCa38 (n=3) and SCa34 (n=2). Labels in italic indicate top ten sires as reported in Table 5.7.

Paternity assignment to trees within fragments. Interestingly, only a few of the mature trees in each fragment are responsible for the identified pollination events (Figure 5.3 and Appendix 5.1).

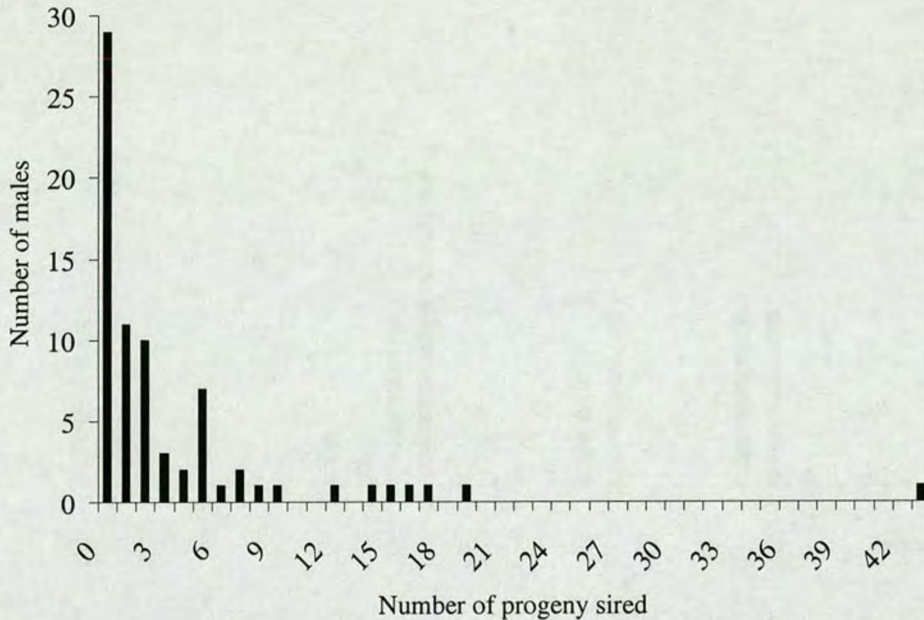


Figure 5.3: Distribution male contribution to paternity in the three *F. excelsior* remnants as a function of the number of seed they sired in Moffat Dale.
Results of paternity assignment using a maximum likelihood approach with allele binning and excluding individuals with incomplete genotypes in FAMOZ (Gerber *et al.* 2003).

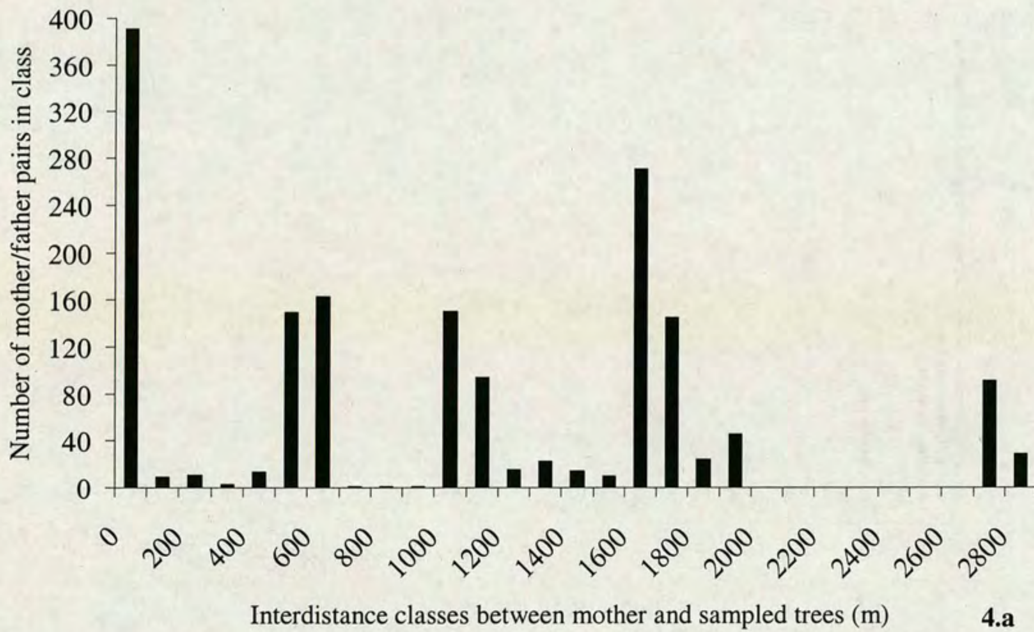
Table 5.7: Top ten trees among the *F. excelsior* mature individuals sampled which sired most progeny among the seed sampled in Moffat Dale

Rank	Tree	Number of progeny sired†
1	CDa19	43
2	CDa110	19
3	CDa130	17
4	CDa119	16
5	Sca37	15
6	Sca32	14
7	CDa112	12
8	CDa115	8
9	CDa109	7
10	Sca42	5

†Results of paternity assignment using a maximum likelihood approach with allele binning and excluding individuals with incomplete genotypes in FAMOZ (Gerber *et al.* 2003). For a full description of paternity assignment for each sampled tree in Moffat Dale see Appendix 5.1.

A majority of trees sired only one seed (Figure 5.3) and paternity of most seed can be attributed to five trees in remnant CDa and to two trees in remnant SCa (Table 5.7). Moreover, the few pollinations identified in the source populations Wa, and BMa may be attributed to five trees only among the 20 individuals sampled in each

remnant, namely, Wa11, Wa16, BMa6, BMa7 and BMa20 (Appendix 5.1). Isolated trees A and B (Figure 5.1) did not sire any of the analysed seed (Appendix 5.1).



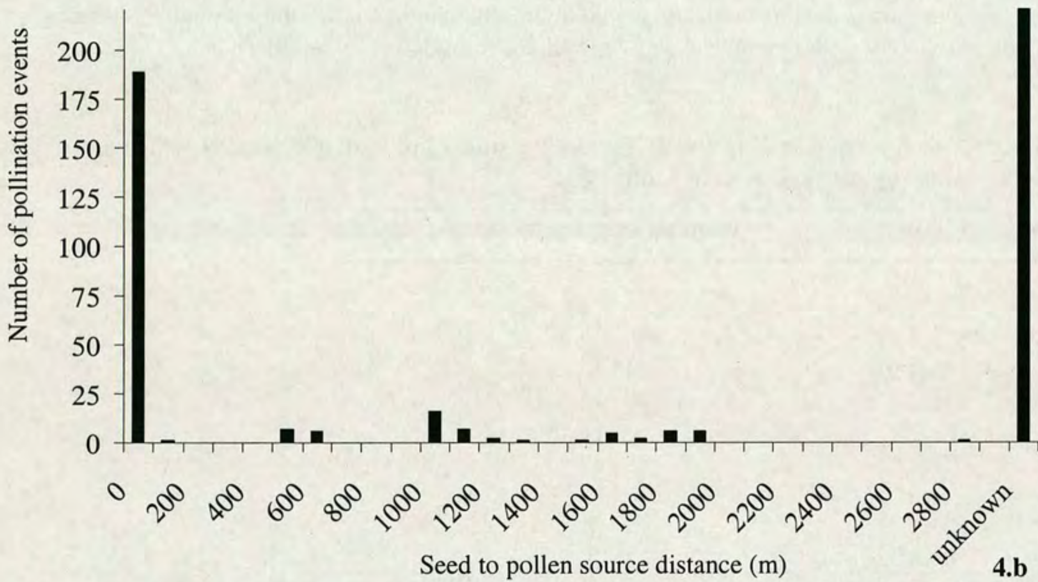


Figure 5.4: a) Distribution of distance between 88 potential *F. excelsior* male parents and 19 mother trees sampled in remnants of Moffat Dale. b) Distribution of actual distance of seed from pollen source for pollination events identified among the sampled individuals.
Maximum number of pollination events attributable to trees sampled within Moffat Dale by maximum likelihood with allele binning in FAMOZ (Gerber *et al.* 2003). Progeny may be assigned to several fathers if paternity is unresolved.

Pollen dispersal distance. The distance between mother trees and all potential male parents sampled in Moffat Dale are shown in figure 5.4a. The patchiness of the fragmented habitat results in an uneven distribution of trees. Many fathers are present within a 100 m neighbourhood of the mother trees but others also occur between 500 m and 700 m, between 1000 m and 2000 m and between 2700 m and 2900 m of the mother trees while no tree occurs between 2000 m and 2700 m (Figure 5.4a). This is reflected in the actual distance of pollen dispersal events observed within Moffat Dale. Most identified pollen parents can be found in close proximity to the mother tree (<100m) but a few events have been identified much further away, at distances between 500 m and 700 m, between 1000 m and 2000 m and in one instance at 2900 m (Figure 5.4b).

Fractional paternity assignment to remnants. The fractional likelihood assignment procedure in PATRI attributes over 50% of identified paternity to trees from remnant CDa and about 20% from remnant SCa, while the contribution of sampled trees in the source populations BMa and Wa is less than 10%, and in remnant CMa less than 1% (Table 5.8). Results differ little when binning alleles but the contribution of sampled trees in the source populations BMa and Wa is doubled when genotypes are modified to include null alleles. Furthermore, the expected number of offspring fathered by sampled trees is reduced by half when the male effective population size (N_{em}) is modelled as a uniform function with 100 and 500 as lower and upper limits respectively rather than when assuming an exact count of 150 individuals (Table 5.8).

Table 5.8: Fractional assignment of paternity to trees in the five remnants of *F. excelsior* sampled in Moffat Dale.

	Posterior expectations of the number of sampled offspring					
Population	no allele binning		allele binning		allele binning & null	
N_{em} ‡#	**150**	**[100-500]**	**150**	**[100-500]**	**150**	**[100-500]**
CDa	65.81	28.92	65.16	27.36	69.15	27.53
CMa	0.72	0.23	1.35	0.47	1.94	0.96
SCa	22.55	17.41	23.47	18.12	35.55	21.45
Wa	7.20	4.51	7.09	4.30	16.87	9.36
BMa	9.94	5.19	9.93	5.19	22.14	9.31
Total#	**106**	**56**	**107**	**55**	**146**	**69**

Posterior expectations of the number of offspring sampled in each remnant as estimated by fractional likelihood analysis in PATRI (Signorovitch and Nielsen 2002). ‡Male breeding population size. #When specifically accounting for the occurrence of null alleles (null), 422 seed and 88 trees are included in analysis, 387 seed and 74 trees were considered otherwise.

Note the expected fraction of offspring fathered by non sampled individuals is the pollen flow estimated by other methods (Table 5.6) suggesting that a male effective population of 150 individuals is a reasonable assumption.

Comparison between remnants. Pollen dispersal parameters were found to differ between remnants of varying size and degree of spatial isolation. In particular, the maximum (one seed may be attributed several paternal parents) number of male parents contributing to individual pollen pools, identified within or outwith a remnant, is generally greater in remnant CDa (Table 5.9) where more trees occur in high density (30 individuals) and for which the distance to surrounding remnants is the least, on average 1093 m (Figure 5.1).

Table 5.9: Maximum number of sires identified among the mature *F. excelsior* individuals sampled in five remnants of Moffat Dale for each mother tree studied.

	Maximum number of identified sires[a]			
Mother tree	**ML**	**ML,BIN**	**SE**	**SE,BIN**
CDa102	2	9	6	9
CDa105	9	14	14	14
CDa108	1	4	4	4
CDa111	4	10	8	9
CDa112	5	7	6	7
CDa118	11	12	10	12
CDa123	2	8	2	2
CDa125	6	9	7	6
CDa126	4	5	4	4
CDa129	7	7	7	7
CDa130	5	8	5	7
CDa all	*56*	*93*	*73*	*81*
CMa24	3	6	3	4
CMa26	1	4	1	1
CMa all	*4*	*10*	*4*	*5*
SCa33	3	7	3	4
SCa34	3	3	3	3
SCa35	2	3	2	3
SCa36	5	6	5	6
SCa38	0	1	0	1
SCa41	3	3	3	3
SCa all	*16*	*23*	*16*	*20*

[a]Progeny may be assigned several fathers. In the case of unresolved paternity, progeny attributed to all potential fathers. Maximum number of mates identified among the potential mates sampled for each mother tree as estimated by maximum likelihood (ML), simple exclusion (SE) and with allele binning (BIN) in FAMOZ (Gerber *et al.* 2003)

In addition, apparent pollen flow in the smallest remnant CMa (four individuals) is higher, with a value of 96%, 75% and 100% as estimated with binning of alleles by simple exclusion and maximum likelihood respectively (Table 5.10). Levels of apparent pollen flow in remnant SCa, which is isolated from other remnants by a distance of nearly 1 km more, vary between 69% (when binning alleles) and 75% (without binning of alleles) and are lower for CDa (at least 57% and up to 74%) which is the largest (30 individuals) and least isolated remnant (Table 5.10). Pollen immigration from surrounding sampled remnants and source populations is highest in CMa and estimated to represent at least 7% and up to 36% of fertilisation events in this remnant.

Table 5.10: Percentage of total apparent pollen flow and of apparent pollen flow from known[#] populations in the three *F. excelsior* remnants of Moffat Dale studied for their different size and degree of spatial isolation.

			total apparent pollen flow in %		pollen flow from known remnant in %		
population	size	seed sample	SE (BIN)[a]	ML (BIN)[b]	SE (BIN)	ML (BIN)	mean distance to other remnants
CDa	30	265	66 (58)	74 (57)	3 (5)	3 (6)	1093 m
CMa	4	28	96 (75)	96 (100)	11 (11)	7 (36)	1286 m
SCa	12	94	73 (69)	75 (69)	2 (5)	2 (6)	2210 m

[a]Results of paternity analysis by simple exclusion in FAMOZ (Gerber *et al.* 2003). [b]Results of paternity analysis by maximum likelihood in FAMOZ. In brackets, same analysis with allele binning (BIN). Population size estimated by the number of individuals censed, mean distance to other remnants expressed in meters. [#]Remnants other than the remnant of origin or source populations.

Cryptic pollen flow into Moffat Dale. Simulations of 25 seed per mother tree for 19 trees assuming a male effective population size of 150 individuals estimate that the proportion of true pollen flow (m_p) over apparent pollen flow (m_a) (at a test threshold of paternity of 3.37) is m_p/m_a=123.23% ± 5.33 and when binning alleles (at a test threshold of paternity of 2.25) that m_p/m_a=150.92% ± 5.10. This suggests that the level of apparent pollen flow detected for the data set, under the assumption that the actual male breeding population size for the sample is 150 individuals, underestimates actual pollen flow into the catchment by a quarter (when alleles are not binned) to a half (when alleles are binned). This provides an upper estimate of pollen flow from unsampled fathers for Moffat Dale.

IV Discussion

1 Methodology of paternity analysis

The paternity analysis of *Fraxinus excelsior* remnants of Moffat Dale reported in this study is one of the most detailed assessments of contemporary pollen flow that have so far been reported for a wind-pollinated temperate tree species. The sampling strategy adopted enabled me to address patterns of pollen dispersal at the individual and population levels. Not only were adult trees exhaustively sampled within remnants and a sample of trees was taken in surrounding remnants but seed were also sampled from all mother trees, thus avoiding the problem of describing merely the idiosyncratic behaviour of arbitrarily chosen individuals. Two likelihood-based methods of paternity analysis which assess the statistical confidence of an assignment and account for the non-isolation of sampled populations were used (Gerber *et al.* 2000; Nielsen *et al.* 2001). While both methods essentially provided identical individual assignment of paternity (comparative results not shown), they highlight different aspects of population patterns of paternity. The maximum likelihood approach developed in FAMOZ (Gerber *et al.* 2003) focuses on the estimation of apparent and cryptic gene flow from the unknown fraction of males contributing to the pollen pool while the fractional likelihood approach of PATRI (Signorovitch and Nielsen 2002) assesses the male reproductive success of trees in the defined remnants.

A number of difficulties must nonetheless be acknowledged. Although the sampling strategy sought to include all the males contributing to the pollen pool of *F. excelsior* remnants of Moffat Dale, the majority of seed (>54.3%) was fathered by unsampled males. Therefore, the actual male effective population size of the sampled remnants is unknown. This may be problematic as this is a key parameter in determining confidence in assignment of paternity (Nielsen *et al.* 2001; Oddou-Muratorio 2002). Indeed, modifying the effective population size in the fractional likelihood analysis in PATRI resulted in great quantitative variation in the expected number of offspring assigned to sampled males (Table 5.8). Direct estimates of cryptic gene flow by simulation in FAMOZ appears to be little affected by assumptions about the actual effective population size (Gerber *et al.* 2000). The proportion of cryptic gene flow

was estimated to be 23% to 50% of apparent gene flow for *F. excelsior* into Moffat Dale. In accordance with other authors, I feel that this may be an overestimation of the true value (Oddou-Muratorio 2002; Hardy *et al.* submitted). In particular, simulations in FAMOZ do not account for the possibility of heterogeneity in allele frequencies between the sampled and the unsampled fraction of males contributing to the pollen pool (Hardy *et al.* submitted). Furthermore it was chosen, for the purpose of LOD scores calculations and simulations in FAMOZ, to estimate allele frequencies of the pollen pool from seed and adult genotypic data pooled together. This procedure may greatly overestimates the frequency of maternal alleles of the pollen pool as the contribution of maternal alleles is accounting for not only in genotyped mothers but also in each of their sampled progeny. To avoid such bias, one can alternatively estimate allele frequencies for the pollen pool considering the paternal contribution to progeny genotypes only. Using such procedure barely affects most-likely paternity assignments (57.6% apparent pollen flow (ML, BIN), $p<0.01$; results not shown) but reduces cryptic gene flow estimates nearly by half (results not shown) which thus suggests that that estimates of cryptic gene flow in FAMOZ are affected by assumptions made on the allelic composition of the pollen pool.

Another factor that will affect the confidence in paternity assignment is the genotyping error (whether as a result of misreading or of actual mutation) associated with the use of microsatellite loci (Marshall *et al.* 1998). The error estimated for the data set was high (e=0.0420). Introducing such a mean blind error rate or a much lower one as advised by San Cristobal and Chevalet (1997; cited in Oddou-Muratorio 2002) increased stochasticity (results not shown). With such a high experimental error rate, it is nevertheless essential to take into account the risk of erroneously attributing paternity associated with the use of microsatellite markers. The decision to proceed by binning of rare alleles with more common alleles of the nearest size seems appropriate and it accounts for the higher probability of stepwise mutation and of misreading the size of rare alleles.

Another concern may be the occurrence of one or several non-amplifying alleles at marker loci. Estimations of null allele frequency in the tree sample with CERVUS

((Marshall *et al.* 1998) suggest that the frequency of null alleles may be as high as 14% at the marker loci retained for analysis although one may argue that using am method assuming Hardy-Weinberg equilibrium in such a fragmented population may not be suitable (discussed in more details in Chapter 4). One the one hand, the failure to amplify at one or two loci of a number of individuals lower the confidence in paternity assignment as candidates may not be excluded as fathers on the on the only basis their genotype cannot be compared to that of the progeny/mother pair (Appendix 5.1). This is a most conservative approach that may however underestimate true apparent gene flow (Table 5.6). On the other hand, candidate trees with incomplete genotypes may not be included for analysis in PATRI which reduce the size of the sample of candidates and may underestimates the contribution to paternity of remnants in which incompletely genotyped trees occur. Indeed, if one assumes that such amplification failures indicate the occurrence of null alleles, a fractional likelihood paternity assignment in PATRI results in a twofold increase in the posterior expectations of the number of sampled offspring in the source populations BMa and Wa (Table 5.8). However, making the assumption than amplification failures are indicative of null alleles does not affect estimates of apparent (Table 5.6) or cryptic gene flow (results not shown).

Fortunately, comparative results reported in this study (Table 5.6) show that, thanks to high paternity exclusion probabilities (*PEP*>0.99) and although individual assignments may vary depending on the method chosen and the factors considered, trends in population patterns of pollen dispersal are globally the same. Finally, estimates have only been considered for one year. This behaviour may vary from that of other years and generalisation to one-generation estimate of pollen dispersal for *F. excelsior* in Moffat Dale may not be justified (Irwin *et al.* 2003).

2 Contemporary pollen flow for *F. excelsior* across a fragmented landscape

Gene flow by pollen for *F. excelsior* into Moffat Dale is extensive. Apparent pollen flow into the catchment was estimated to be between 38.6% and 72.3%. These values are comparable to those reported for other wind-pollinated temperate species. Streiff *et al.* (1999) estimated from microsatellite data that levels of pollen flow into a

continuous stand were 65% and 69% for *Quercus robur* and *Q. petraea* respectively, while Dow and Ashley (1998) reported 62% of pollen flow into a population of *Q. macrocarpa*. However, it is important to note that these studies considered gene dispersal in sampling areas not exceeding 150 m radius, in dense habitats. In an population of the wind-pollinated conifer *Pinus flexilis* covering an area of nearly 400 m in radius, Schuster and Mitton (2000) estimated from isozyme markers that as much as 6.5% of pollen immigrated from populations located at least 2 km away. Similarly, pollen flow into Moffat Dale may have originated from unsampled trees occurring at least 1 km from the sampled mother trees, but is likely to be from much further away as most trees within a 3 km radius were sampled.

Results reported herein show that *F. excelsior* maintains extensive pollen exchange over much larger distances than would be anticipated in a landscape heavily deforested not only locally but also at the wider regional scale of the southern Uplands of Scotland (>50 km). This is in contrast to findings by Sork *et al.* (2002) from a TWOGENER analysis (Austerlitz and Smouse 2001) of *Q. lobata* that pollen movement is limited in a fragmented savannah population of this wind-pollinated species but agrees with evidence from insect-pollinated tropical species, that extensive long distance pollen flow is observed for isolated pasture trees (Dick 2001; White *et al.* 2002).

Pollen dispersal events in pasture trees of the tropical tree species, *Dinizia excelsa* were identified as far away as 3.2 km (Dick 2001) which is similar to the maximum distance of 2.9 km recorded for pollination events identified here. Gene flow events at distances over 4.5 km and 6-14 km were reported for *Swietenia humilis* (White *et al.* 2002) and for *Ficus* spp. (Nason and Hamrick 1997) respectively, and over 10 km for the wind-pollinated *Cecropia obtusifolia* (Kaufman *et al.* 1998). Other studies report much smaller distances of dispersal at 280 m for *Cordia alliodora* (Boshier *et al.* 1995), 350 m in *Pithecellobium elegans* (Chase *et al.* 1996) and 550 m in *P. flexilis* (Schuster and Mitton, 2000). As pointed out by Boshier *et al.* (1995), these figures are likely to greatly underestimate actual pollen dispersal distance because of the limit imposed by the sampling area.

Pollen dispersal has often been described to as following a leptokurtic distribution (Boshier *et al.* 1995; Streiff *et al.* 1999) although Schuster and Mitton (2000) found no excess of near-neighbour matings for *P. flexilis*. Most pollination events for *F. excelsior* within Moffat Dale were identified locally, originating from trees within a 100 m radius but rare events were also identified in other forest remnants throughout the sampling area and up to 2.9 km away. However, unidentified events, which are in the majority, may have originated at greater distances. This emphasises further the difficulty of estimating the 'fat-tail' of the dispersal curves in tree species (Ellstrand, 1992).

Several studies of allozyme variation in *Acer saccharum*, a wind-pollinated temperate tree species, have shown a reduction of genetic differentiation among fragmented patches compared to stands in continuous forest which suggested that habitat fragmentation enhanced gene flow by facilitating airborne pollen movement between fragments (Fore *et al.* 1992; Young *et al.* 1993). In the southern Uplands of Scotland, the climate is influenced by prevailing western Atlantic winds (Wildwood Group of the Borders Forest Trust 2000) but deforestation and use of land for pasture has greatly opened the landscape and high intensity wind is regularly observed in the area (for data on windiness for the southern Uplands of Scotland see associate figure 6B of the Carrifran Wildwood management plan, Wildwood Group of the Borders Forest Trust 2000). It is likely that habitat fragmentation has facilitated pollen movement and thus enhanced effective pollen flow for *F. excelsior* in Moffat Dale.

Interestingly, among the 88 sampled trees, only eight seem to make a large contribution to paternity in the seed sampled and others do not contribute at all (Appendix 5.1). Temporal variation in flowering between trees may affect mate availability, although 7 years of field observations (Wallander 2001) suggests that flowering is often synchronised at the population level for *F. excelsior.* It may also be that few trees within Moffat Dale were flowering or producing pollen in the year of study, therefore favouring effective long distance pollination.

Comparing male reproductive success and levels of pollen flow between the three remnants sampled in Moffat Dale, differing in population size, density and degree of spatial isolation to other remnants, sheds further light on the impact of landscape heterogeneity on tree population genetic dynamics. 68% of the seeds analysed were sampled in the largest remnant, CDa, and 24% in the most isolated remnant, SCa. A fractional assignment attributed over 50% of identified paternity to trees from remnant CDa and about 20% from remnant SCa, reflecting the fact that identified fathers were found primarily among local trees within these remnants. However, the male reproductive success of the four trees of remnant CMa was less than 1%, total pollen flow into CMa was highest (75% to 100%) and lowest for CDa (57% to 74%).

Similar results were found for fragments of *S. humilis* (White *et al.* 2002) challenging the preconception that isolated trees are 'living dead' (Dick, 2001; White *et al.* 2002). In fact, as Ellstrand and Elam (1993) suggested, although habitat fragmentation reduces the size of the local remnants, gene flow into smaller fragments is enhanced because the fraction of seed fertilised by immigrant pollen increases. Indeed, while pollen flow from different (known) sources in Moffat Dale is much higher for CMa (up to 36%), mating in CDa, where density is highest, tends to be more localised. Interestingly, the physical barrier provided by a conifer plantation coupled with a greater distance to other remnants (over 2 km) did not act as a barrier to gene flow into SCa, where it is higher than for CDa (69% to 75%).

V Conclusions

A detailed paternity analysis of seed collected from mother trees revealed extensive gene flow by pollen for *Fraxinus excelsior* and distances of effective pollen dispersal up to 2.9 km long were recorded in Moffat Dale. Not unlike insect-pollinated tropical tree species, fragmented populations of a wind-pollinated temperate tree species maintain contemporary gene exchange at a much larger scale than anticipated. Habitat fragmentation appears to enhance pollen flow by opening the landscape to wind dispersal and by increasing the fraction of seed fertilised by immigrant pollen in low density populations. This study provides further evidence that isolated trees are part of a wide reproduction network thus helping maintain genetic diversity in tree populations despite anthropogenic disturbances.

How such immigrant pollen will ultimately affect the genetic structure of *F. excelsior* remnants depends on how much of the pollen pool genetic diversity is effectively carried to successive generations by established seedlings that reach maturity. Natural regeneration in Moffat Dale is severely limited by continuous grazing pressure and pasture habitats may be unfavourable to seedling establishment. Thus, actual gene flow may be recruitment limited rather than dispersal limited (Imbert and Lefèvre 2003). A parentage analysis on established seedlings, presented in a companion paper (Chapter 6), will help one understand how seed dispersal and seedling recruitment affect the dynamics of *F. excelsior* remnants of Moffat Dale.

Appendix 5.1: Paternal contribution to seed sample of *F. excelsior* mature individuals sampled in Moffat Dale.

tree	gender	missing data	maximum number of progeny sired[a]			
			ML	ML, BIN	SE	SE, BIN
CDa019	male		29	43	40	44
CDa101	male	FEMSATL2	0	13	0	37
CDa102	female		0	3	2	3
CDa103	male	1.19	0	0	0	1
CDa104	male	1.19	2	3	2	3
CDa105	female		3	4	4	4
CDa107	male	FEMSATL2	6	6	18	32
CDa108	female		0	0	0	0
CDa109	male		4	7	6	8
CDa110	female		10	19	15	18
CDa111	female		3	3	3	3
CDa112	female		7	12	10	12
CDa113	male	FEMSATL2, 1.19	7	19	37	71
CDa114	male	1.19	11	13	14	14
CDa115	male		6	8	7	8
CDa116	male		5	5	5	5
CDa117	male	1.19	5	7	10	15
CDa118	female		1	2	2	2
CDa119	male		7	16	11	17
CDa120	male		0	3	0	0
CDa121	female		1	1	1	2
CDa122	female		1	2	2	2
CDa123	female		0	0	0	0
CDa124	male	1.19	7	7	11	12
CDa125	female		4	6	5	6
CDa126	female		1	2	2	2
CDa127	female		1	5	2	5
CDa128	female		2	2	2	2
CDa129	female		1	1	2	2
CDa130	female		10	17	12	17
B	male		0	0	0	0
CMa24	female		1	0	1	1
CMa26	female		1	2	2	2
CMa27	male	FEMSATL2	0	1	2	3
CMa28	male		0	2	0	0
A	male		0	0	0	0
SCa31	male		0	0	0	0
SCa32	male		10	14	11	15
SCa33	intermediate		5	5	5	5
SCa34	intermediate		0	0	0	0
SCa35	female		0	0	0	0
SCa36	intermediate		0	0	0	0
SCa37	male		13	15	13	16
SCa38	mostly male		2	2	2	2
SCa40	mostly male		1	1	1	1
SCa41	female		0	0	0	0
SCa42	male		5	5	5	5
SCa43	male	1.19	0	0	2	2

Appendix 5.1 (continued)

tree	gender	missing data	maximum number of progeny sired[a]			
			ML	ML,BIN	SE	SE,BIN
Wa01	not recorded		0	0	0	0
Wa02			0	0	0	0
Wa03			0	0	0	0
Wa04			0	0	0	0
Wa05			0	0	0	0
Wa06			0	1	0	0
Wa07			0	2	0	2
Wa08			0	0	0	0
Wa09		1.19	0	0	1	1
Wa10			0	0	0	0
Wa11			2	2	2	2
Wa12			0	5	0	5
Wa13		1.19	1	5	1	6
Wa14			0	1	0	1
Wa15			0	0	0	0
Wa16			2	5	3	5
Wa17			0	0	0	0
Wa18			0	0	0	1
Wa19			0	5	0	5
Wa20			1	1	1	1
BMa01			2	2	2	2
BMa02			0	0	0	1
BMa03		1.19	0	1	1	2
BMa04			0	1	0	1
BMa05			0	0	0	0
BMa06			0	9	1	9
BMa07			0	7	0	4
BMa08			0	0	0	0
BMa09			1	1	1	1
BMa10			1	0	0	0
BMa11			0	0	0	0
BMa12			1	1	1	1
BMa13			0	0	0	0
BMa14		M2.30B, 1.19	2	19	22	73
BMa15			0	0	0	0
BMa16			0	1	0	1
BMa17			0	0	0	0
BMa18			0	1	0	1
BMa19			0	0	0	0
BMa20			3	4	3	3

[a]Progeny may be assigned several fathers. In the case of unresolved paternity, progeny attributed to all potential fathers. Maximum number of progeny sired by each individual as estimated by maximum likelihood (ML), simple exclusion (SE) and with allele binning (BIN) in FAMOZ (Gerber *et al.* 2003). Gender estimated from observations of fruit production in 2000.

Chapter 6 : LANDSCAPE APPROACH TO CONTEMPORARY GENE DISPERSAL OF *FRAXINUS EXCELSIOR* L. II. PARENTAGE ANALYSIS REVEALS EXTENSIVE SEED FLOW IN FOREST FRAGMENTS

Abstract

Estimation of contemporary gene dispersal in tree species is essential to determine whether habitat fragmentation results in genetic isolation of populations through limited gene flow. In particular, seed dispersal and post-dispersal recruitment may be important factors in determining population and metapopulation viability. Microsatellite markers were used and a parentage analysis was performed on recently established seedlings of *Fraxinus excelsior* L. (Oleaceae) in severely fragmented remnants of a single catchment (Moffat) in southern Scotland. This approach enabled me to estimate the contribution of pollen and seed dispersal to realised gene flow and seed dispersal distances for this wind pollinated wind dispersed tree species. Realised gene flow into Moffat Dale is remarkably extensive with only 6.6% of seedlings pollinated and dispersed locally. Although pollen flow is an important component of effective gene flow, it is not predominant and 56.6% of the seedlings that established in Moffat Dale immigrated into the catchment. Seed dispersal within 20 m of the source was found to be prevalent (30% of analysed seed) but long-distance dispersal was recorded as far as 1.4 km away from the source and may be effective over tens of kilometres in the open landscape of Moffat Dale. Comparison between remnants suggests that landscape fragmentation may affect gene flow processes not only quantitatively but also qualitatively. In a smaller remnant, a greater number of sources contribute to realised gene flow while seed-mediated gene flow appears to be reduced and to originate from fewer sources in a more spatially isolated remnant. This study highlights the importance of seed dispersal in maintaining within remnant genetic diversity and gene exchange between them.

Keywords : *Fraxinus excelsior* L., microsatellites, parentage analysis, maximum likelihood, seed gene flow, landscape fragmentation.

I Introduction

Seed dispersal is an essential factor in determining a species' resilience to forest fragmentation. Widespread anthropogenic disturbances including deforestation and subsequent land use for pasture reduce the size of extant forest fragments and increases isolation between them (Andren 1994). Such disturbances also increase the isolation of suitable colonisation sites (Wilcox and Murphy 1985), often creating large areas of habitat unsuitable for recruitment (Cain *et al.* 2000). Consequently, severely fragmented tree populations may not only suffer an increased risk of extinction but also reduced genetic diversity and increased genetic differentiation if gene exchange is limited between them, ultimately affecting population viability (Young *et al.* 1996). Seed dispersal patterns will affect a species' ability to reach suitable colonisation sites and determine temporal and spatial genetic structure within and among populations (Nathan and Muller-Landau 2000).

Empirical evidence from direct measurements of seed dispersal from a point source suggest that almost invariably seeds disperse very close, within one to 20 m, to the source plant (mainly herbaceous species, Levin and Kerster 1974) and dispersal declines leptokurtically with distance (Ellstrand 1992). However, estimations often ignore long distance seed dispersal because of the inherent difficulties of sampling such dispersal events with these methods (Nathan and Muller-Landau 2000; Cain *et al.* 2000). Indirect approaches using genetic markers can help overcome these problems (Cain *et al.* 2000). In particular, recent developments of highly polymorphic microsatellite markers and refined statistical methods (Gerber *et al.* 2000; Austerlitz and Smouse, 2001; Nielsen *et al.* 2001; Burczyk *et al.* 2002) have revolutionised the study of contemporary gene dispersal processes in plants (Sork *et al.* 1999).

For the reasons aforementioned, pollen dispersal is believed to be predominant (Ellstrand 1992; Dyer and Sork 2001) especially in wind pollinated tree species for which pollen and seed mass may differ greatly. Therefore, attention has so far primarily focused on estimating the pollen dispersal component of gene flow. Many studies have reported paternity analysis of pollen dispersal using progeny arrays for

temperate (Dow and Ashley 1998; Kaufman *et al.* 1998; Streiff *et al.* 1999; Schuster and Mitton 2000) and tropical tree species (Chase *et al.* 1996; Kaufman *et al.*, 1998;Dick, 2001; White *et al.* 2002) but only a restricted number of studies have investigated pollen dispersal patterns in dispersed and established individuals, e.g. *Gleditsia triacanthos* (Schnabel and Hamrick 1995) or *Neobalanocarpus heimii* (Konuma *et al.* 2000). Furthermore, little interest has been devoted to seed-mediated gene flow. Parentage analysis of seed dispersal has been reported mainly for gravity or animal dispersed temperate species such as *Quercus macrocarpa* (Dow and Ashley 1996), *Gleditsia triacanthos* (Schnabel *et al.* 1998), *Magnolia obovata* (Isagi *et al.* 2000) and *Pinus pinaster* (Gonzalez-Martinez *et al.* 2002). To the best of my knowledge, only one study has reported patterns of both pollen and seed-mediated gene flow for seedlings established in fragmented forest remnants, being for the neotropical animal dispersed tree species, *Symphonia globulifera* (Aldrich and Hamrick 1998).

Evidently, information on gene dispersal patterns in established seedlings is needed for a wider range of species and ecological conditions. As some authors have pointed out, successful pollination by a foreign gamete may not be equivalent to successful establishment of that gamete's genes in the local gene pool (Schnabel and Hamrick 1995; Dow and Ashley 1996). Effective pollen flow may be substantially less than that observed from paternity analysis on progeny arrays if, for instance, immigrant genes are at selective disadvantage for establishment (Schnabel and Hamrick 1995). Conversely, in fragmented populations, immigrant seed may have a selective advantage if local seed suffer from inbreeding depression. Long distance seed dispersal and recruitment may therefore be determinant factors of the quality and quantity of effective gene flow.

This study describes contemporary patterns of realised pollen and seed mediated gene flow for *Fraxinus excelsior* in severely fragmented forest remnants in southern Scotland. The main objective of this study is to assess the population genetic response of a wind-pollinated, wind-dispersed temperate tree species to long-term habitat fragmentation and landscape heterogeneity.

In southern Scotland, as part of a local initiative for native woodland restoration (Wildwood Group of the Borders Forest Trust 2000), large areas of upland pasture have recently been fenced to release grazing pressure from severely fragmented forest remnants (detailed description in Chapter 4). There is therefore an immediate need to understand the population genetics dynamics of these forest remnants and to assess the long-term potential for natural regeneration in protected areas.

F. excelsior L. (Oleaceae), common ash, is a wind pollinated species with a complex polygamous sexual system (Wallander 2001). Sex expression varies at the floral, inflorescence and tree level, as well as between years, and individuals may be classified phenotypically into a continuum from purely male to purely female with a whole range of hermaphroditic intermediates (Wallander 2001). Hermaphrodite flowers are protogynous, but self-fertilisation seems possible (Wallander 2001). However, it was shown in a previous study that complete outcrossing occurs for *F. excelsior* forest remnants in southern Scotland (Chapter 4). Regular fruit bearing begins around 20 years of age and the species displays a masting behaviour with irregular fruiting between years (Tapper 1996). Fruits which are dry with a wing-shaped extension (i.e. samara) present special adaptations to wind dispersal (Horn *et al.* 2001).

In this study, microsatellite markers are used to perform a parentage analysis by maximum likelihood on recently established seedlings of *F. excelsior* and to describe patterns of realised gene flow and to estimate seed dispersal distances in forest remnants of this wind-dispersed tree species in southern Scotland. Furthermore, comparisons of levels of genetic diversity between cohorts and of potential gene flow estimated from progeny arrays (Chapter 5) with realised gene flow are useful to provide a comprehensive picture of the genetic dynamic of *F. excelsior* populations enduring long-term anthropogenic disturbance.

II Material and methods

1 Study site

The study site is the Moffat Dale, as previously described for *Sorbus aucuparia* (Chapter 3) and for *Fraxinus excelsior* remnants (Chapters 4 and 5). Full description of *F. excelsior* remnants sampled in Moffat Dale is given in chapter 4.

The study focuses on the same three forest remnants in which patterns of pollen dispersal were previously described (Chapter 4 and 5). Two of these remnants, CDa and CMa, are situated in a grazed pasture and one, SCa, is situated on a grazed riverside surrounded by a 30-year-old conifer plantation (Figure 6.1). CDa, CMa and SCa differ in their population size (30, 4 and 12 individuals respectively) and degree of spatial isolation (mean distance to other remnants 1093 m, 1286 m, 2210 m respectively). Two additional remnants were surveyed, Wa and BMa (Figure 6.1). The latter two were sampled as potential sources of immigrant pollen and seed for the three studied remnants and will be referred to as 'source populations'.

2 Sampling

In spring 2001, 20 newly established (first year of growth) seedlings were collected per site in CDa, CMa and SCa, distributed throughout each site (Figure 6.1). In each of the three remnants, leaf material was collected from all mature individuals (30, 4 and 12 individuals in CDa, CMa and SCa respectively). Two isolated trees were also sampled (A and B, Figure 6.1). In each of the source populations, BMa and Wa, leaf material was collected only from a representative sample of 20 randomly distributed mature trees. Plant material was stored at -80 °C immediately after collection until DNA extraction.

The spatial position of the 88 sampled trees was determined as reported in chapter 4. Spatial position of the 60 sampled seedlings was either recorded using a handheld global positioning system (GPS) Etrex™ as described in chapter 4, or estimated from the location of the nearest identified tree (which was estimated by GPS, Chapter 4) when sampled within the canopy area of the tree (precision < 10 m).

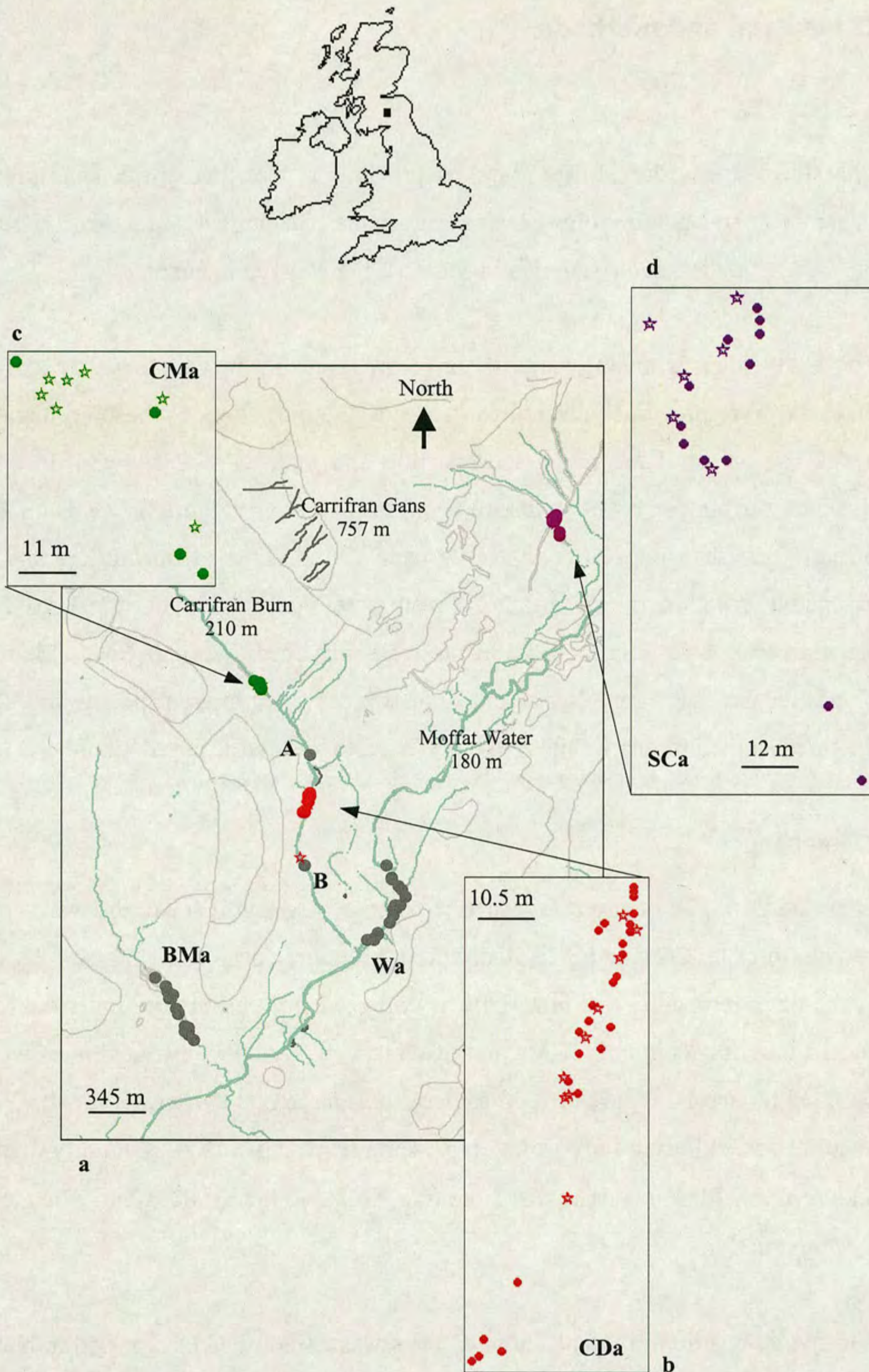


Figure 6.1: Distribution of *F. excelsior* trees and seedlings sampled in Moffat Dale remnants.
a) Each dot represents a tree. Trees are grouped in populations as described in chapter 4. Background map is a section of product Land-Line.Plus-nt11©Crown copyright Ordnance Survey. An EDINA digimap / JISC supplied service. b, c & d) Close up of spatial distribution of individuals in remnant CDa, CMa and SCa respectively. Stars represent approximate locations where one to five newly established seedlings were sampled.

3 DNA isolation and microsatellite genotyping

Genomic DNA was isolated from 100 mg of frozen leaf tissue for adult trees, and from 50 mg of frozen plant tissue for seedlings, following a standard extraction protocol adapted from Doyle and Doyle (1987) and described by Rendell and Ennos, (2002). Extracted DNA was dissolved in 100 µl of TE (pH 7.6) and stored at -20 °C.

Five microsatellite loci, FEMSATL2, FEMSATL8, FEMSATL19 (Lefort *et al.* 1999) M2-30B and 1.19 (Brachet *et al.* 1999) were used in this study for their scorability and high level of polymorphism (Table 6.1). Amplification conditions and genotyping analysis were as described in chapter 4.

Table 6.1: Initial sample size (*N*), number of alleles sampled (*k*), gene diversity (H_e), and allelic richness (*R*) at eight microsatellite markers for mature individual and established seedlings sampled in each of five *F. excelsior* remnants of Moffat Dale.

		Population						
locus		**CDa**	**CMa**	**SCa**	**BMa**	**Wa**	**Overall**	**Null frequency**[‡]
M2-30B	*N*	52	24	31	19	20	146	0.126
	k	22	16	13	12	15	34	
	H_e	0.899	0.887	0.812	0.920	0.908	0.944	
	R[†]	14.179	13.550	9.545	11.843	14.180	16.949	
1.19	*N*	46	21	30	18	18	133	0.141
	k	9	9	6	6	7	11	
	H_e	0.816	0.854	0.649	0.588	0.825	0.826	
	R	7.660	8.424	5.124	5.943	6.943	8.378	
FEMSATL2	*N*	49	23	31	20	20	143	0.048
	k	27	18	17	20	21	45	
	H_e	0.899	0.861	0.899	0.949	0.958	0.945	
	R	16.296	14.910	13.527	17.790	19.248	19.146	
FEMSATL8	*N*	47	24	30	17	18	136	0.262
	k	11	14	11	11	14	22	
	H_e	0.806	0.816	0.811	0.850	0.943	0.892	
	R	8.315	11.423	9.042	11.000	13.717	13.310	
FEMSATL19	*N*	48	24	30	19	20	141	0.224
	k	17	13	13	11	12	20	
	H_e	0.914	0.855	0.816	0.769	0.921	0.908	
	R	13.293	11.062	10.318	10.560	11.767	13.131	

[†]Measure of allelic richness based on a minimum sample size of 18 diploid individuals. Estimations were performed in FSTAT 2.9.3 (Goudet 2001) and GENETIX 4.01 (Belkhir *et al.* 1998). [‡]Null allele frequency calculated on adult sample only (*N*=88) assuming Hardy-Weinberg equilibrium, in CERVUS 2.0 (Marshall *et al.* 1998).

To account for the possibility of genotyping error or mutation at marker loci (Chapter 5), I chose to bin rare alleles with more common alleles of the nearest size. Alleles were generally deemed rare when they occurred in frequency less than 0.01 for all

markers except 1.19 for which a procedure reflecting difficulties in gel scoring was used (Appendix A). The procedure reduced the number of alleles observed from 45 to 27, from 22 to 18, from 22 to 17, from 34 to 24 and from 11 to 7 at locus FEMSATL2, FEMSATL8, FEMSATL19, M2-30B and 1.19 respectively (procedure detailed in Appendix A).

4 Genetic diversity

Mean gene diversity (H_e) was estimated over the five microsatellite loci for both adult and seedling cohorts in each of the three remnants CDa, CMa and SCa, using GENETIX 4.01 (Belkhir *et al.* 1998). Since the size of the adult sample varies greatly between remnants, allelic richness is probably the most relevant criteria for measuring diversity (Petit *et al.* 1998). Using the rarefaction method of El Mousadik and Petit (1996) to correct for variation in sample size, allelic richness per sample (R_s) was estimated using FSTAT 2.9.3 (Goudet 2001) and averaged over five microsatellite loci. R_s was estimated considering all six samples together (i.e. adult and seedling samples for each of three remnants) in order not only to assess the extent of genetic diversity present in the adult population that is maintained in a newly established generation, but also to compare levels of genetic diversity between the three remnants.

5 Parentage analysis

Method of assignment. A parentage analysis was undertaken to determine the parentage of the 60 seedlings sampled in the three *F. excelsior* remnants of Moffat Dale. A maximum likelihood approach to parentage assignment was taken using FAMOZ (version released on 30th January 2003, Gerber *et al.* 2003).

The most-likely parents and parent pairs were detected by means of 'log of the odds' ratios (LOD scores, (Meagher and Thompson 1986; Gerber *et al.* 2000) using population allele frequencies estimated from seedling and adult multilocus genotypic data pooled together (N=148). No significant genotypic association was detected among any pair of the five microsatellite loci for the sampled trees (Chapter 4),

therefore, LOD scores over all loci were obtained by adding LOD scores calculated for each locus (Meagher and Thompson 1986).

There are several reasons why the most-likely parents may not be the true parents of a tested offspring. In particular, it is virtually impossible to sample all adults contributing to reproduction in natural tree populations. It is therefore necessary to assess the risk of excluding a candidate as one of the true parents on the sole grounds that it has not been sampled (Marshall *et al.* 1998; Nielsen *et al.* 2001). Genotyping error or mutation at marker loci may also result in erroneous parentage assignment. It is moreover necessary to resolve parentage in case of tied-LOD scores between several candidate parents and parent-pairs.

One way to solve the problem of confidence in parentage assignment is by hypothesis testing (Marshall *et al.* 1998; Gerber *et al.* 2000). In FAMOZ, confidence levels for rejecting a parent or a parent pair as the true one relies on a simulation procedure detailed in Gerber *et al.* (2000) since LOD scores in these cases do not follow any particular statistical law (Gerber *et al.* 2003). The distribution of the LOD scores of the most-likely parents (or parent pairs) of 10000 randomly generated seedlings with both parents sampled (i.e. by randomly choosing pair of parents among the 88 sampled trees) was compared to the distribution of the LOD scores of the most-likely parents (or parent pairs) of 10000 seedlings whose genotype was randomly generated according to allele frequencies for the whole sample. The test thresholds for rejecting a parent as a true parent (TP) and for rejecting a parent pair as the true one (TC) were chosen at the intersection of the two distributions of LOD scores to minimise both type I error of wrongly rejecting a parent (or parent-pair) as the true one and type II error of wrongly accepting a parent (or parent pair) as the true one (Gerber *et al.* 2000).

The test was applied to determine parentage of the 60 seedlings sampled in the three *F. excelsior* remnants of Moffat Dale. All 88 trees sampled both within the remnants (CDa, CMa and SCa) and outside the remnants in surrounding 'source populations' (BMa, Wa and isolated trees) of Moffat Dale were considered as potential parents

(including self). Outcomes of parentage assignment may be: i) that no parent, ii) at least one parent (but no parent pair) or iii) at least one parent pair is identified for a given seedling among the sampled trees.

Apparent gene flow into Moffat Dale. The proportion of seedlings with no parent identified among the sampled trees provided an estimate of apparent gene flow by seed into Moffat Dale. To undertake further analysis on seedlings for which at least one parent was identified among the sampled trees, two assumptions were made following Dow and Ashley (1996) and Gonzalez-Martinez *et al.* (2002). 1) If only one parent was identified among the sampled trees, it was assumed to be the maternal parent. The proportion of seedlings with only one parent among the sampled trees, therefore provided an estimate of apparent gene flow by pollen into Moffat Dale. 2) If a parent pair was identified, the nearest parent was assumed to be the maternal parent and the other the paternal parent. It was also assumed, in case of unresolved parentage, that the nearest parent was the maternal parent. Studying spatial patterns of seed dispersal within Moffat Dale was allowed with these further assumptions

Cryptic gene flow into Moffat Dale. Gene flow events may be undetected if gametes are actually the product of non-sampled individuals that cannot be genotypically distinguished from gametes produced by sampled parents. To specifically estimate the proportion of such cryptic gene flow, a simulation based test (Gerber *et al.* 2003) is available in FAMOZ. The procedure estimates the proportion of cryptic gene flow (i.e. the number of simulated offspring that were attributed a parent among the genotyped sample when their actual parents were not sampled) over apparent gene flow (i.e. the number of simulated offspring with no parent among the sampled trees). Input data were chosen to resemble the actual sampling in the *F. excelsior* remnants of Moffat Dale: 60 offspring were randomly simulated by picking both parents either among the 88 sampled trees or among unsampled individuals (i.e. generated according to allele frequencies calculated for the whole data set) assuming an effective population size of 150. Although the actual effective population size is unknown, it seems reasonable to assume that 150 individuals contribute to reproduction as this is the approximate number of trees identified within Moffat Dale

(Chapter 4, Table 4.1). The test thresholds for attributing parentage of simulated offspring to sampled or unsampled individuals was determined using the maximum likelihood approach described previously Gerber *et al.* (2000) and levels of apparent and cryptic gene flow were estimated accordingly. Simulations were repeated ten times to crudely determine the variance of the estimate.

Female reproductive success of sampled trees. The female reproductive success of each of the 88 sampled trees of Moffat Dale was inferred from the number of seedlings whose maternity was attributable to this sampled individual by parentage analysis among the seedlings for which at least one parent could be identified.

Seed dispersal distance within Moffat Dale. The distance distribution of seedlings to their identified maternal parent was plotted for seedlings with at least one parent identified among the sampled trees of Moffat Dale by parentage analysis, and compared to the distance distribution of potential dispersal events between the 60 seedlings and 88 trees sampled in Moffat Dale.

Source and level of gene flow into remnants. 20 established seedlings were sampled in each of three remnants, CDa, CMa and SCa. The proportions of seed flow, pollen flow and local dispersal in each of the three remnants were inferred from the parentage analysis. The proportion of pollen and seed to total gene flow from known sources in Moffat Dale, i.e. either the source populations (BMa and Wa) or the other remnants (CDa, CMa and SCa), was also estimated. The comparison of levels of seed and pollen gene flow between remnants of different size and degree of spatial isolation to other populations allows one to assess how landscape fragmentation may affect estimates of the source and quantity of gene flow.

III Results

1 Microsatellite diversity in adult and seedling samples

Gene diversity. For the adult cohort, mean *He* over five microsatellite loci is highest in CDa (H_e=0.820±0.103, *N*=30, Figure 6.2a) and lowest in SCa (H_e=0.709±0.259, *N*=12) which is the most isolated remnant (mean distance to other remnants: 2210 m). In CMa, where only four trees are present, gene diversity was high (H_e=0.815±0.140).

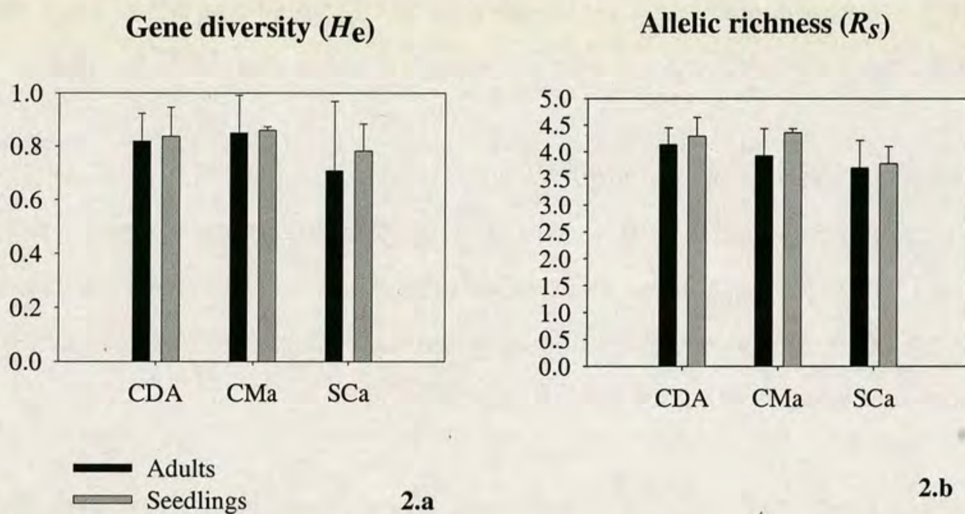


Figure 6.2: Comparison of gene diversity (H_e, 2.a) and allelic richness (R_s,, 2.b) between adult and seedling cohorts in each of three *F. excelsior* remnants of different size and degree of spatial isolation in Moffat Dale.
Mean H_e over five microsatellite loci estimated in GENETIX 4.01 (Belkir, *et al.* 1998), R_s estimated for a minimum sample size of three individuals in FSTAT 2.9.3 (Goudet 2001) and averaged over five microsatellite loci. Remnants CDa, CMa and SCa comprise 30, 4 and 12 trees respectively, 20 seedlings were sampled in each remnant. The mean distance to other remnants is 1093 m, 1286 m, and 2210 m respectively for CDa, CMa and SCa. Bars show standard errors of estimates.

For the seedling cohort (*N*=20 in each remnant) levels of gene diversity were found to be higher than for the adult cohort (Figure 6.2a). H_e is highest in CMa (H_e=0.860±0.015) and lowest in SCa (H_e=0.783±0.102).

Allelic richness. R_s averaged over five microsatellite loci and based on a minimum sample of four individuals, showed similar trends to H_e in the comparison of genetic diversity between the three *Fraxinus excelsior* remnants (Figure 6.2b). For the adult cohort, mean allelic richness was highest in CDa (R_s=4.136±0.310) and lowest in

SCa (R_s=3.702±0.510). For the seedling cohort, R_s is highest in CMa (R_s=4.357±0.082) and higher than for the adult cohort (R_s=3.928). In CDa and SCa, R_s is also higher (R_s=4.193±0.321 and R_s=3.781±0.321, respectively). However, differences in mean H_e and R_s estimates are not statistically significant (Figure 6.2). Overall, similar levels of genetic diversity and allelic richness are maintained in adult and seed cohorts across remnants.

2 Parentage analysis

Overall results. The single parent and parent pair exclusion probabilities (Jamieson and Taylor 1997; Gerber *et al.* 2000) combined over five microsatellite loci and estimated in FAMOZ were high (single-parent exclusion probability >99.59% and parent-pair exclusion probability >99.99%) providing great confidence in parentage assignment.

At the chosen thresholds for rejecting a parent as a true parent (TP=4.68, *p*-value<0.16) and for rejecting a parent pair as the true one (TC=5.40, *p*-value<0.0001), 21 of the 60 seedlings have at least one parent (but no parent pair) identified among the 88 trees sampled in Moffat Dale. For all but one seedling, one unique parent was identified (Table 6.2). In addition, another four seedlings have one parent pair among the 88 trees sampled in the catchment, but parentage is unresolved for three of them (Table 6.2). In total, at least one parent was found among the sampled trees for eight out of 20 seedlings sampled in remnant CDa and nine out of 20 seedlings sampled in each of CMa and SCa.

Table 6.2: Global results of parentage analysis performed on 60 seedlings established in three *F. excelsior* remnants in Moffat Dale.

population	number seedlings sampled	number of seedlings with one parent among sampled trees		number of seedlings with one parent pair among sampled trees	
		unique parent	unresolved	unique pair	unresolved
CDa	20	5	1	0	2
CMa	20	9	0	0	0
SCa	20	7	0	1	1
Total	**60**	**22**		**4**	

The parentage analysis was performed in FAMOZ (Gerber *et al.* 2003) using a maximum likelihood approach. Results of parentage assignment for test thresholds TP=4.68 (p-value<0.16) and TC=5.40 (*p*-value<0.001) to reject a parent and a parent pair respectively as the most-likely one based on 10000 simulations.

Total gene flow into Moffat Dale. The parentage analysis shows that among the 60 established seedlings sampled, 34 do not have any parent among the 88 *F. excelsior* trees sampled and must have immigrated into the catchment. Seed gene flow into Moffat Dale is therefore estimated to represent 56.7% of effective dispersal. A further 22 seedlings have only one parent (but no parent pair), assumed to be the maternal parent, among the sampled trees. Thus, at least 36.7% of pollen fertilising the seedlings sampled in Moffat Dale originated from unsampled trees and 84.6% of the established seedlings sampled that have dispersed from local trees, were fertilised by immigrant pollen. Apparent gene flow into Moffat Dale is remarkably extensive, with 93.4% of seedlings resulting from either (or both) pollen and seed immigration and only 6.6% of seedlings being pollinated and established locally.

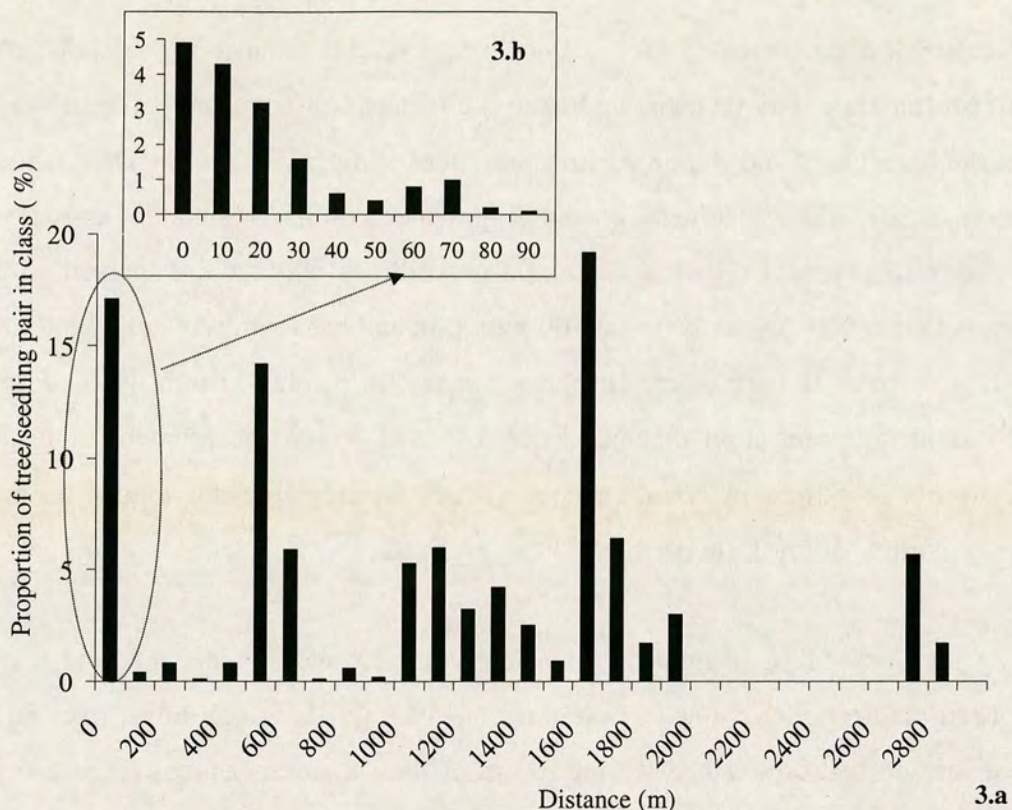
Gene flow may have originated from unsampled trees in the source populations (BMa and Wa) within Moffat Dale or from forest remnants outwith the catchment. The *F. excelsior* remnants of Moffat Dale are not reproductively isolated despite severe habitat fragmentation. Simulations in FAMOZ to account for the possibility of cryptic gene flow in such non-isolated populations, when assuming an effective population size of 150, estimate cryptic gene flow to be 7.5% (±2.4%) of apparent gene flow. Therefore, total gene flow into Moffat Dale may be underestimated by as much as 7% which would suggest that at most, all established seedlings sampled in Moffat Dale result from either pollen or seed immigration.

Female reproductive success of sampled trees of Moffat Dale. Of the 26 seedlings with at least one parent identified among the 88 trees sampled in the catchment, 18 mother trees are identified. Eight of these trees are located in remnant CDa, five in remnant SCa and two in CMa. Another two trees are located in the source population BMa and one in Wa. Only four individuals out of 18 contributed to maternity of more than one seedling. Four seeds (6.6%) have dispersed from tree 24 from CMa, three seeds (5%) have dispersed each from tree 36 and 37 in SCa and two (3.3%) from 103 in CDa.

Seed dispersal distance within Moffat Dale. The potential distance for seed dispersal within Moffat Dale may be estimated from the distance distribution between the 60 established seedlings and the 88 *F. excelsior* trees sampled in Moffat Dale (Figure 6.3a and 6.3b). The patchiness of the fragmented habitat results in an uneven distribution of potential dispersal. Essentially, seedlings may have dispersed within 100 m, between 500-700 m, between 1000-2000 m and between 2700 and 2900 m of their mother trees (Figure 6.3a). Dispersal under 100 m (and within 40 m, Figure 6.3b) would represent short distance local dispersal within the remnant of origin, while events identified in other distance classes would represent longer distance dispersal within Moffat Dale between forest remnants.

Among the 60 seedlings sampled, actual dispersal distance to the maternal parent have been precisely determined in 26 cases (~43%). Of these seedlings, over 80 % have dispersed and established within 100 m of their maternal parent (Figure 6.3c) and half of these (~22% of the total sample) have established at the foot of their maternal parent (<10 m, Figure 6.3d) reflecting very limited seed dispersal.

However, a proportion of seed dispersal events, ~ 20%, have been identified at great distance, at least 200 m and up to 1400 m away from the maternal parent (Figure 6.3c). Most interestingly, for 34 seedlings (~57%) a parent could not be found suggesting that they did not disperse from trees sampled in Moffat Dale and that a majority of seedlings have dispersed over distances greater than 3000 m. It is nevertheless possible that these seedlings dispersed within Moffat Dale from unsampled trees in the source populations (BMa and Wa). The minimum dispersal distance is, on average, 1660 m, but depends on the remnant of origin, and is less for seedlings established in CDa (mean distance CDa-Wa 660 m) and much more for seedlings established in SCa (mean distance SCa-Wa 2140 m).



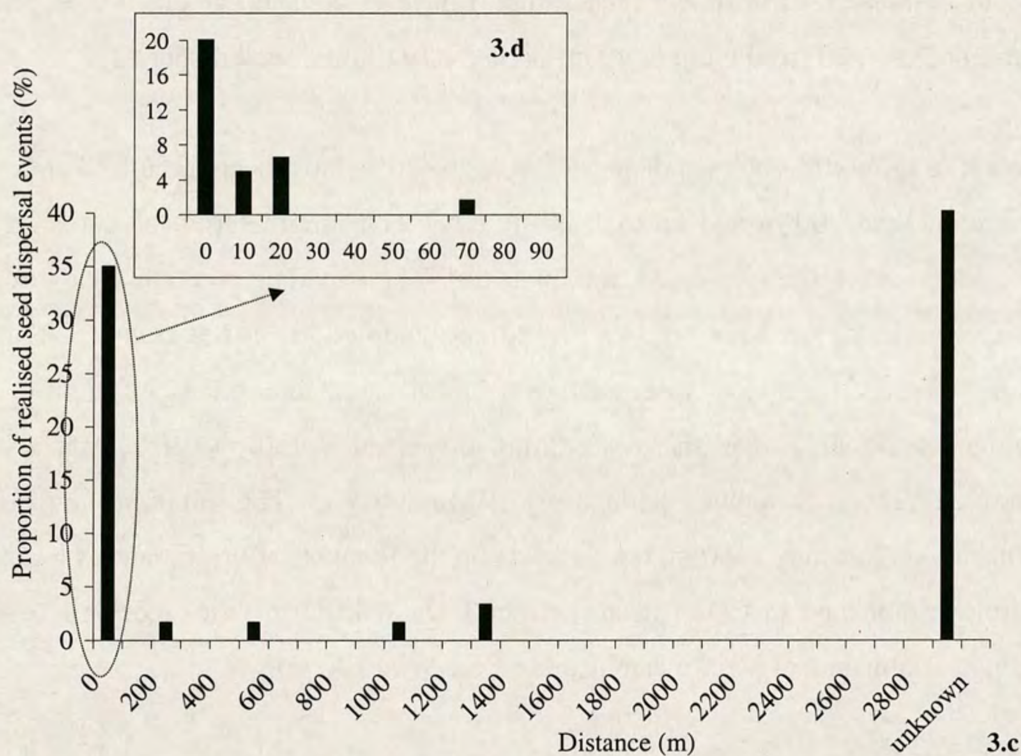


Figure 6.3: a & b) Frequency distribution of the distance between 88 *F. excelsior* trees and 60 established seedlings sampled in Moffat Dale. c & d) Frequency distribution of dispersal distance of 60 seedlings established in Moffat Dale.

Seed dispersal events identified by means of parentage analysis in FAMOZ (Gerber *et al.* 2003) for 60 seedlings sampled in Moffat Dale. When only one parent is identified, it is assumed to be the maternal parent. When a parent pair is identified, the nearest parent is assumed to be the maternal parent.

Source and level of gene flow into remnants. Levels of gene flow in CDa, CMa and SCa, which are remnants of different size (30, 4, 12 trees respectively) and degree of spatial isolation from other forest remnants within Moffat Dale (mean distance 1093 m, 1286 m, 2210 m respectively), are comparable and remarkably high (Table 6.3). Only 10% of the sampled seedlings have been pollinated and established locally in CDA and SCa while none did so in CMa, which is the smallest remnant.

Furthermore, the nature and origin of gene flow differ between the three remnants. Seed gene flow is highest (60%) into CDa, the largest and least isolated remnant and pollen gene flow highest (45%) into CMa, the smallest remnant (Table 6.3). Interestingly, in no instance, could seed gene flow from other remnants and source populations be identified except for CMa. In the latter remnant, 20% of seedlings dispersed from local trees and were pollinated by trees in remnant CDa and in the 'source populations' Wa and BMa within Moffat Dale.

Table 6.3: Comparison of the source and quantity of pollen and seed flow between *F. excelsior* remnants CDa, CMa and SCa of Moffat Dale studied for their different size and degree of spatial isolation to other known remnants.

	source of pollen and seed (%)				
		pollen flow‡	seed flow#		
population	local pollen/seed	local seed	known seed source†	unknown pollen/seed source	total gene flow
CDa	10	30	0	60	90
CMa	0	25	20	55	100
SCa	10	35	0	55	90

Results based on a parentage analysis performed in FAMOZ (Gerber *et al.* 2003). ‡Percentage of seedlings with at least one parent (but no parent pair) among the 88 sampled trees-parent identified within Moffat Dale assumed to be the maternal parent. #Percentage of seedlings with no parents among the 88 sampled trees. †known sources may be trees in remnants CDa, CMa and SCa or one of the 20 trees sampled in each of the source populations, BMa and Wa, the pollen source is unknown.

IV Discussion

The effects of habitat fragmentation on tree population genetic structure appear to be more complex and varied than expected from population genetic theory (Aldrich and Hamrick 1998). Direct estimation of contemporary pollen and seed mediated gene dispersal in relation to the landscape in which they occur is an essential step towards understanding how species respond to disturbances and implementing sustainable forest restoration practices (Sork *et al.* 1999). In this study, extensive seed gene flow is reported for severely fragmented populations of *Fraxinus excelsior* in Southern Scotland, suggesting that long distance seed dispersal is eminent in maintaining genetic diversity within forest remnants and gene exchange among them.

Adult and seedling cohorts maintain similar levels of genetic diversity at microsatellite makers in the three *F. excelsior* remnants sampled in Moffat Dale. Comparing genetic diversity between adults and recently established seedlings is indicative of the influence that habitat fragmentation may have on the population genetic dynamic of the remnants because the seedling cohort represents a generation submitted to longer term fragmentation. Mean gene diversity was higher for the seedling cohort than for the adult cohort in all three remnants (Figure 6.2a). However, since the number of adults present in remnant CMa (N=4) and SCa (N=12) is substantially less than the number of seedlings sampled in each of them (N=20), such results may be explained by differences in sample size (as suggests the greater variance for estimates adult cohorts in CMa and SCa) as well as by high rates of contemporary gene flow into the sampled remnants. Estimates of allelic richness (R_s), corrected for sample size, are not significantly higher in the seedling cohorts than in the adult cohorts of all remnants.

Contrary to theoretical expectations, the allelic composition of the seedling sample is not a simple subset of that found in *F. excelsior* trees present in the three remnants sampled. Similar results were found in forest fragments of the neotropical species, *Carapa guianensis* for which reduced allelic richness but increased genetic distance were observed in saplings cohorts (Dayanandan *et al.* 1999) and *Symphonia globulifera* (Aldrich *et al.* 1998). In *S. globulifera*, differences in seedling and adult

gene pools could be explained by high rates of gene flow into forest fragments which originated from nearby pastures but from only a few donors thus creating a genetic bottleneck (Aldrich and Hamrick 1998). In this study, remnant SCa is most isolated not only in terms of distance (mean distance to other remnants 2210 m) but also topographically because it is enclosed within a conifer plantation. I reported in previous chapters, substantial pollen flow into the remnant (on progeny arrays, Chapter 5) but from fewer pollen donors from Moffat Dale compared to other remnants (Chapter 4). It may be that the spatial isolation of remnant SCa also contributes to reducing the number of seed donors resulting in the (slight) reduction in allelic richness observed in both adult and seedling cohorts of SCa compared to other remnants. Nevertheless, differences between cohorts and remnants are not statistically significant and overall, levels of genetic diversity are maintained, if not increased, in the seedling cohorts, suggesting substantial gene flow into the sampled *F. excelsior* remnants.

Indeed, a parentage analysis of the 60 newly established seedlings sampled in Moffat Dale estimates that at least 56.6% of them have immigrated into the catchment. In contrast, Dow and Ashley (1996) reported only 6.6% of seed immigration into a continuous stand of the gravity-dispersed tree species, *Quercus macrocarpa,* although their sampling area was much smaller (~150 m in radius) but substantial seed gene flow (at least 40%) into a watershed was reported for *Magnolia obovata* (Isagi *et al.* 2000) which is bird dispersed. Although the sampled area of Moffat Dale is about three times as big as the watershed sampled for *M. obovata*, still greater seed immigration into severely fragmented populations of *F. excelsior* was detected in this study. Similarly to forest fragments of the bat-dispersed *S. globulifera* (Aldrich and Hamrick 1998) which produced 4.3% of the seedlings established in them, only 6.6 % of the established seedlings sampled in remnants of Moffat Dale were pollinated and dispersed locally. These findings are of fundamental importance suggesting that far from being limited, seed-mediated gene flow is a significant component of realised gene exchange in fragmented populations of the wind dispersed *F. excelsior* in Moffat Dale.

Both the barrenness of the landscape in Moffat Dale and *F. excelsior* dispersal characteristics may account for the high levels of seed gene flow reported in this study. Short distance seed dispersal events (within 20 m of the identified maternal parent) were prevalent (~30%) suggesting that seed dispersal of *F. excelsior* remnants of Moffat Dale is typically leptokurtic (Ellstrand 1992) but the tail of the distribution may stretch over many kilometres with seed dispersal events recorded as far as 1.4 km away from the source. Maximum seed dispersal distances reported for other tree species are much shorter, 165 m for *Q. macrocarpa* (Dow and Ashley 1996), 175 m for *Gleditsia triacanthos* (Schnabel *et al.* 1998) or 600 m for *M. obovata* (Isagi *et al.* 2000), but long-distance seed dispersal may be underestimated in these species because of the limits set by the sampling area.

The most striking result of this study is that long-distance dispersal (dispersal distance>100 m; Cain *et al.* 2000) represents 65% of effective dispersal and immigrant seed (56.6%) may have dispersed from forest remnants outwith Moffat Dale or from unsampled trees in the source populations (BMa and Wa which are at least several hundreds of meters away. Indeed, they may have immigrated from much further away since most mature trees within 3 km were sampled. Contrary to the species aforementioned, *F. excelsior* is wind-dispersed and its fruits are elongated into a wing (Wardle 1962). Evidence from ecological modelling shows that such winged-seed are better adapted to long-distance dispersal (Horn *et al.* 2001) although seeds must be lifted above the canopy by updrafts to have a chance of further dispersal in high velocity horizontal winds (Horn *et al.* 2001). In the bare landscape of southern Scotland where there is no canopy closure and where forest remnants are exposed to very windy climatic conditions, the ability for winged-seed of *F. excelsior* to disperse long distance may be further enhanced resulting in extensive effective long-distance dispersal for remnants of Moffat Dale.

Only 18 out of the 88 trees sampled in Moffat Dale have contributed to the production of seeds established locally. It may be that these trees were the only ones producing fruits in the year of production as *F. excelsior* trees may not flower or produce fruits every year (Wallander 2001). Interestingly, 10 of these trees were

among the 19 trees producing fruits in 2000 in the three remnants in which seed production was assessed that year (Chapter 5). Furthermore, the female reproductive success of these trees is low (<5%) and comparatively even, with all but four trees identified as the maternal parent of one unique seed (1.6%). This is very much in contrast to *G. triacanthos* (Schnabel *et al.* 1998) and *S. globulifera* (Aldrich and Hamrick 1998) for which distinct reproductive dominance of a few trees was reported. The actual establishment of seedlings in Moffat Dale may be recruitment rather than dispersal limited (Cain *et al.* 2000). A paternity analysis of progeny arrays (Chapter 5) shows no evidence for biparental inbreeding in the sampled remnants and suggests that local seeds are unlikely to suffer from inbreeding depression. However, biotic factors, such as grazing may be important in determining establishment. In grazed pastures, only two cohorts are observed, mature trees and newly established seedlings, but the latter only in early spring. It may be that the habitat has become unsuitable for establishment or that grazing *per se* creates a sampling effect and, by chance, seedlings established in the remnants of Moffat Dale have been produced by multiple sources from within and outwith the remnants. Measuring effective gene flow in established seedlings over several consecutive years would be necessary to assess the year to year stochasticity in patterns of effective gene dispersal.

The impact of habitat fragmentation on *F. excelsior* population genetic processes can be investigated further by comparing the level and origin of effective gene flow between the three remnants which differ in size, density and degree of spatial isolation. Total effective gene flow is highest in the smallest remnant, CMa (N=4), for which none of the established seedlings were pollinated and dispersed locally. Furthermore, only for remnant CMa was it possible to identify the source of immigrant seed (20%) in other remnants (CDa) and in the source populations (BMa and Wa) within Moffat Dale. Conversely, in the most isolated remnant, SCa (mean distance to other remnants, 2210 m) although total gene flow (90%) is just as high as in remnant CDa, which is least isolated and most dense, seed gene flow was the lowest (55%).

These results are remarkably similar to those reported for contemporary pollen gene flow into the three remnants estimated from progeny arrays (Chapters 4 and 5) thus reinforcing the hypothesis that the fragmentation of the landscape influences patterns of gene flow into forest fragments not only quantitatively but also qualitatively. In smaller remnants, effective gene flow is not only increased but also originates from a greater number of sources. This may be explained because there are fewer trees producing pollen and seed locally, creating more establishment opportunities for immigrant pollen or seed. In more spatially isolated remnants, seed but not pollen-mediated gene flow is reduced (Chapter 5) and fewer donors from the catchment contribute to pollen flow (Chapter 4), potentially creating a genetic bottleneck. The effect may be even stronger in species other than *F. excelsior* which do not have wind as an efficient vector of dispersal, for instance if they are insect pollinated or gravity dispersed.

V Conclusions

Seed dispersal is eminent in maintaining extensive gene flow between severely fragmented populations of the wind pollinated wind dispersed temperate tree species, *Fraxinus excelsior*. Although pollen gene flow is a significant component of contemporary effective gene flow between remnants, it does not prevail over seed flow. This study highlights the needs to measure both contemporary pollen and seed-mediated gene flow in relation to the landscape in which these processes occur. So far, the importance of seed-mediated gene flow may have been overlooked, especially in open landscapes where long-distance seed dispersal by wind may be effective over many kilometres.

The results of this study have important implications for native woodland restoration in southern Scotland. *F. excelsior* forest remnants of Moffat Dale are part of a wide reproductive network which maintains genetic diversity within remnants and gene exchange between them. Grazing has historically limited recruitment and natural regeneration in Moffat Dale. Removal of grazing pressure should ensure the restoration of a genetically diverse woodland. Ultimately, the genetic structure of the restored woodland will be determined by how canopy closure may limit both further recruitment and seed dispersal in the longer-term.

Chapter 7 : IMPLICATIONS OF THE DYNAMIC GENETIC BEHAVIOUR OF FOREST FRAGMENTS FOR NATIVE WOODLAND RESTORATION IN SCOTLAND

I Introduction

Although theoretical expectations are that habitat fragmentation will have significant detrimental effects on genetic processes in population remnants, literature reporting empirical validation or otherwise of such effects is scarce, especially for tree species (Chapter 1). In particular, how habitat fragmentation may affect mating patterns within and among population remnants has been little studied to date. However, this information is of crucial importance for forest restoration in severely fragmented areas (Ennos 1998).

In regions, such as southern Scotland, where native woodland cover is so limited and fragmented that native tree species are present in remnants comprising only a few isolated trees, the most practical approach to forest restoration is planting. Collecting seed from local remnants would provide the best adapted material to local climatic and edaphic conditions (Ennos 1998; Ennos *et al.* 1998). However, current approaches (Herbert *et al.* 1999) to seed collection from remnant woodlands (e.g. Carrifran Wildwood Project) are based on almost total lack of knowledge of the scale of local adaptation in trees and of how habitat fragmentation may have affected the amounts and patterns of genetic variation within and among remnants.

If theoretical expectations are verified, it should be anticipated that planting stock collected from such seed sources would be genetically depauperate and would suffer reduced fitness. Furthermore, potential for natural regeneration may be very limited if pollen and seed dispersal among remnants is so restricted that isolated trees are not reproductively active. Therefore, a thorough understanding of the dynamic genetic behaviour of native woodland remnants is fundamental to obtaining good quality seed.

Consequently, one objective of this thesis is to provide the missing information on the genetic processes that affect remnant woodlands and to suggest how this information may be used to ensure the long-term success of forest restoration programmes, with particular emphasis on the Carrifran Wildwood Project.

II Summary of results

The objective of assessing the genetic effects of long-standing and severe habitat fragmentation on forest remnants of *Sorbus aucuparia* and *Fraxinus excelsior* occurring in the same landscape in southern Scotland (Moffat Dale) is to address the concerns raised by population genetics theory. In particular, results reported in this thesis help in answering the following questions (Chapter 1):

1- Are chronically fragmented remnants genetically depauperate?
2- Is there evidence of inbreeding and preferential mating with neighbouring individuals within forest remnants?
3- Are forest remnants genetically differentiated?
4- Is pollen and seed exchange limited among forest remnants?

S. aucuparia and *F. excelsior* remnants of Moffat Dale are not genetically depauperate. In fact the amounts of genetic variation at molecular markers, sampled in mature trees, are similar to those reported for large, non-fragmented populations of continental Europe (Chapters 3 and 4).

Furthermore, there is no suggestion that mating patterns within remnants have been altered in such a way that inbreeding is favoured. In *S. aucuparia* remnants, no significant departure from random mating was detected overall (Chapter 3). Although a high and significant inbreeding coefficient (F_{is}) was found in *F. excelsior* remnants, comparison with non-fragmented French populations (Morand *et al.* 2002) shows that such findings are not atypical for the species and therefore suggests that the reported value does not reflect the influence of habitat fragmentation (Chapter 4).

Indirect estimation of inbreeding by measuring F_{is} in samples of mature individuals may be misleading if inbred progeny suffer from inbreeding depression and do not reach maturity. In addition, it does not reflect contemporary mating patterns among extant individuals, which may only be estimated from the seed they produce. Direct estimation of contemporary mating patterns from open-pollinated progeny arrays using a model-fitting approach (Burczyk *et al.* 2002) suggests that *F. excelsior*

remnants are completely outcrossed (Chapter 4). Although male reproductive success of trees within the catchment is negatively correlated to their distance to mother trees, nearly half of them contribute, on average, to the pollen pool, and mean pollen dispersal distance in Moffat Dale, 2.5 times greater than reported for a continuous stand (Heuertz e*t al.* 2003) suggests that habitat fragmentation may facilitate airborne pollen movement by opening the landscape (Chapter 4). Such results therefore indicate that preferential mating between spatially close individuals has not been occurring following habitat fragmentation.

Significant genetic differentiation was detected among forest remnants of both *S. aucuparia* and *F. excelsior* (Chapters 3 and 4). A remarkably low pollen versus seed migration ratio (r=1.36) and low genetic differentiation at chloroplast DNA (cpDNA) markers suggest efficient historical seed dispersal among *S. aucuparia* remnants and that significant genetic differentiation at both isozymes and cpDNA markers is most likely due to more restricted historical pollen-mediated gene flow among them (Chapter 3). Although moderate and significant genetic differentiation was detected at microsatellite markers among *F. excelsior* remnants, historical levels of gene flow appear to have been sufficient to counteract genetic drift (Chapter 4) when the effective number of migrants per generation (Nm) is estimated using the private allele method (Barton and Slatkin 1986).

The latter findings illustrate that genetic differentiation at marker loci is not necessarily equivalent to limited historical gene exchange among remnants. As pointed out in chapter 1, significant spatial genetic structure among fragmented populations is often interpreted in terms of restricted levels of historical of gene flow, leading to management recommendations aiming to enhance gene flow. Conclusions from chapter 4 stress that inferring mating patterns within and among remnants from *F*-statistics cannot substitute for estimating contemporary gene flow directly (Whitlock and McCauley 1999; Sork *et al.* 1999).

Indeed, direct measurements of effective pollen and seed dispersal using paternity analyses and related approaches show that contemporary gene exchange is extensive

among *F. excelsior* remnants and processes occur at a much wider scale than that of the sampled area, probably over tens of kilometres (Chapters 5 and 6).

A detailed paternity analysis performed on open-pollinated progeny arrays, collected from all trees producing fruits in three *F. excelsior* remnants of Moffat Dale, shows that pollen-mediated gene flow is very extensive. Furthermore, it has also been found that isolated trees in grazed pastures of Carrifran are not reproductively isolated and receive more immigrant pollen than trees less exposed in the landscape (Chapter 5).

However, a parentage analysis performed on *F. excelsior* seedlings newly established in the same three remnants, reveals that, although an important proportion of seed fertilised by immigrant pollen do establish in Carrifran, a majority of these seeds have immigrated into the remnants. Realised gene flow in *F. excelsior* remnants is extensive but both pollen and seed dispersal contribute –almost evenly- to gene exchange among remnants (Chapter 6).

This is an important finding for the dynamic of fragmented populations of trees. In a fragmented landscape, seed dispersal occurs over large enough distances that colonisation of suitable habitat and migration to extant remnants is possible. From a demographic perspective, such processes lower the risk of population or metapopulation extinction. From a genetic perspective, the results highlight the importance of seed-mediated gene flow in the structuring of genetic variation within and among forest fragments.

III Genetic management of forest restoration

Native woodland restoration projects in Scotland, such as the Carrifran Wildwood Project, have long-term objectives of creating a dynamic and expanding woodland, sustainable without human intervention (Wildwood Group of the Borders Forest Trust 2000). In such cases, forest management should aim at preserving (or creating) the processes that maintain adaptive diversity and evolutionary potential (Ennos 1998; Crandall *et al.* 2000). This may be achieved by carefully selecting seed sources for planting stock. In that respect, local forest remnants, when available, should be given priority because they provide the best adapted material to local climatic and edaphic conditions (Worrell 1992; Ennos *et al.* 1998; Worrell *et al.* 2000).

Findings of this thesis are encouraging for reforestation programmes using seed collected from trees in such forest remnants. In Carrifran, *Sorbus aucuparia* and *Fraxinus excelsior* remnants maintain high levels of genetic diversity in mature trees, in the seed they produce and in recently established individuals (Chapters 4 and 6 for *F. excelsior*, results not shown for *S. aucuparia*), thanks to extensive pollen and seed-mediated gene exchange at a much wider scale than the valley. Therefore, both natural regeneration and planting using seed collected from local forest fragments should ensure that high levels of genetic variation and low levels of inbreeding are maintained in the restored woodland.

However, a survey of the genetic diversity sampled in seedlings of *S. aucuparia* planted in Carrifran during the first two years of the wildwood restoration project was recently undertaken (Kettle 2001). Kettle (2001) found that the levels of genetic diversity at isozyme markers were reduced in the planting stock compared to that of Carrifran forest remnants. Furthermore, high inbreeding coefficients and significant deficit in heterozygotes were reported in the planting stock (Kettle 2001) whereas no departure from random mating was detected in Carrifran (This thesis, Chapter 3).

The planting stock of *S. aucuparia* collected for Carrifran and surveyed for diversity at isozyme markers originated from forest remnants located only a few kilometres north of Carrifran, in Talla Linn and Craigdilly (Kettle 2001). A complementary

survey of genetic variation at isozyme markers in mature individuals of these forest fragments (Ennos *et al.* 2002) shows that *S. aucuparia* forest fragments at Talla Linn and Craigdilly both maintain more genetic diversity (up to 25% more) than has been sampled in the planting stock and display very low, non-significant, inbreeding coefficients. This would suggest that the genetic potential of the seed sources has not been maximized in the planting stock. This may arise in either of two ways.

It is possible that only a few trees in remnants of Talla Linn and Craigdilly actively contribute to reproduction leading to biparental inbreeding. The very low inbreeding coefficient measured in the adult sample would then suggest that inbred progeny do not reach maturity as they suffer from inbreeding depression. This would have important consequences for the success of planting at Carrifran. However, preliminary results (not shown) of genetic analysis of *S. aucuparia* open-pollinated progeny arrays and seedlings established in forest remnants of Moffat Dale indicate that levels of genetic diversity are maintained across life stages and that at least moderate levels of pollen-mediated gene flow (>10%) occur in remnants. Considering the close proximity of Talla Linn and Craigdilly to Moffat Dale, it seems reasonable to assume similar population genetic processes in those populations. If this hypothesis can be confirmed, the reduced genetic quality of the planting stock can be explained by biased sampling strategy where seed are collected from only a fraction of the reproductively active trees in forest remnants.

This example illustrates that even in instances when forest remnants are a suitably diverse and dynamic seed source, restoration strategies must aim at preserving the population genetic dynamic of the forest remnants in the restored woodland. In practice, for planting, sampling most of the genetic diversity present in the mature trees may be best achieved by collecting a handful of seed from most trees within remnants. Collecting seeds from most remnants and planting seedlings at random in regard to their origin will also mimic the gene exchange naturally occurring among remnants.

Furthermore, findings for *F. excelsior* remnants in Moffat Dale of extensive realised gene flow resulting from long-distance seed dispersal suggests that the lack of natural regeneration in forest remnants is due to intensive grazing and not to reproductive isolation. Excluding grazing herds by fencing seems necessary to ensure seedling recruitment. However, there are also indications that extensive pollen and seed-mediated gene flow for *F. excelsior* in Moffat Dale result from a facilitation of airborne pollen and seed movement in a deforested landscape (Chapters 4 to 6). It may also be desirable to leave areas unplanted, to facilitate recruitment of seed and contribution to the pollen pool from distant sources.

IV Conclusions and perspectives

Results of the studies reported herein corroborate findings for a number of neotropical insect pollinated tree species that pollen-mediated gene flow for trees in deforested pasture is extensive (e.g. Dick, 2001; White *et al.* 2002), and that seed-mediated gene flow into forest fragments is not limited (Aldrich and Hamrick 1998). The most remarkable finding for *Fraxinus excelsior* in Moffat Dale is that, although most mature trees within a 3000 m radius were sampled, only a few contributed to reproduction in the forest fragments. It was, however expected that such a sampling strategy would have included –if not all- most reproductively active trees within the area.

The trend that seems to be emerging from recent studies which have assessed contemporary pollen and seed–mediated gene flow in forest remnants, including this one, is that habitat fragmentation, at its current scale, may not be as detrimental for tree species as predicted. It appears that human-induced habitat fragmentation, when it replaces forest patches with open pastures, enhances tree species' pollen and seed dispersal abilities, although in some cases, this may result in increased inbreeding because of reproductive dominance of a small number of trees (Aldrich and Hamrick 1998). Nonetheless, further research is needed to confirm these predictions. For instance, combining population genetic data to landscape and ecological modelling would help testing models that predict for gene flow processes among fragments from on landscape structure and species' dispersal characteristics. In the more practical context of forest restoration, such approach may also prove a useful decision tool by helping to predict the impact of management practices on population genetic structure (Oddou-Muratorio 2002)

From a purely methodological perspective, findings that gene flow processes in tree species occur at a much wider scale than it is practically possible to sample, only confirm the difficulty of accurately estimating realised (and potential) gene dispersal. In particular, the precision of paternity and parentage analyses is greatly affected not only by the degree of polymorphism at marker loci but also by the degree of sampling of the actual reproductive population (e.g. Nielsen *et al.* 2001). Although

current methods allow us to estimate the statistical precision of parentage assignment (e.g. Marshall *et al.* 1998; Gerber *et al.* 2000; Nielsen *et al.* 2001), such estimation relies on assumptions about the size and genetic make up of the (unknown) reproductive population. Recent methodological developments have been made to address these issues (Austerlitz and Smouse 2001; Burczyk *et al.* 2002). However, mating patterns, and in particular the importance of long-distance dispersal in shaping population genetic structure would be best described if complete sampling of the reproductive population was possible. The current challenge for population genetics research is to better describe processes at the scale on which they occur. This thesis also highlights methodological issues related to the use of highly polymorphic markers (Appendix A). It has become obvious as further empirical evidence is collected that, in the framework of current population genetics theory, interpretations in terms of species' reproductive ecology are limited by the influence of mutation patterns at these marker loci. It appears therefore necessary for a new theoretical framework, accounting for the mutational behaviour of microsatellite loci, to be developed (Hedrick 1999).

There is little evidence that levels of variation in quantitative, adaptative traits are correlated to those at variation at marker loci (Ennos *et al.* 1997; Milligan *et al.* 1994). Furthermore, gene flow processes may vary greatly between years (Irwin *et al.* 2003). Further research in *F. excelsior* and *S. aucuparia* remnants of Moffat Dale, involving the measurement of adaptative diversity and contemporary pollen and seed-mediated gene flow over several years would provide a more complete picture of population phenomena. However, as Milligan *et al.* (1994) have pointed out, beyond the quantification of genetic diversity in forest fragments, using molecular markers has made it possible to identify ecological processes of dispersal in tree species and has developed our understanding of the population genetic dynamics of forest fragments at the scale of the landscape in which they occur.

REFERENCES

Adams, W.T. (1992) Gene dispersal within forest tree populations. *New Forests* **6**, 217-240.

Aguirre-Plantier, E., Furnier, G.R. and Eguiarte, L.E. (2000) Low levels of genetic variation within and high levels of genetic differentiation among populations of species of *Abies* from southern Mexico and Guatemala. *American Journal of Botany* **87**, 365-371.

Aldrich, P.R., Hamrick, J.L., Chavarriaga, P. and Kochert, G. (1998) Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*. *Molecular Ecology* **7**, 933-944.

Aldrich, P.R. and Hamrick, J.L. (1998) Reproductive dominance of pasture trees in a fragmented tropical forest mosaic. *Science* **281**, 103-105.

Andren, H. (1994) Effects of habitat fragmentation on birds and mammals in landscape with different proportion of suitable habitat - a review. *Oikos* **71**, 355-366.

Apsit, V.J., Hamrick, J.L. and Nason, J.D. (2001) Breeding population size of a fragmented population of a Costa Rican dry forest tree species. *Journal of Heredity* **92**, 415-420.

Ashmole, P. (1998) Case Study - Carrifran Wildwood project. In *Native woodland restoration in southern Scotland: principles and practice*. (Eds. Newton A.C. and Ashmole, P.) pp.43-48. Occasional paper no. 2, Borders Forest Trust, Ancrum, Jedburgh, Scotland, UK.

Austerlitz, F., Mariette, S., Machon, N., Gouyon, P.-H. and Godelle, B. (2000) Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics* **154**, 1309-1321.

Austerlitz, F. and Smouse, P.E. (2001) Two-generation analysis of pollen flow across a landscape. II. Relation between ϕ_{ft}, pollen dispersal and interfemale distance. *Genetics* **157**, 851-857.

Avise, J.C. (1994) *Molecular markers, natural history and evolution*, New York: Chapman and Hall.

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

Barrett, S.C. and Kohn, J.R. (1991) Genetic and evolutionary consequences of small population size in plants: implications for conservation. In: *Genetics and*

Conservation of Rare Plants, (Eds. Falk, D.A. and Holsinger, K.E.) pp. 3-30. Oxford: Oxford University Press

Barton, N.H. and Slatkin, M. (1986) A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* **56**, 409-415.

Bataillon, T. and Kirkpatrick, M. (2000) Inbreeding depression due to midly deleterious mutations in finite populations : size does matter. *Genetical Research* **75**, 75-81.

Belkhir, K., Borsa, P., Goudet, J., Chikhi, L. and Bonhomme, F. (1998) GENETIX *logiciel sous Windows™ pour la génétique des populations*. Montpellier, France: Laboratoire Génome et Population, CNRS-UPR 9060.

Benitez-Malvido, J. and Martinez-Ramos, M. (2003) Impact of forest fragmentation on understory plant species richness in Amazonia. *Conservation Biology* **17**, 389-400.

Bonnin, I., Colas, B., Bacles, C., Holl, A.-C., *et al.* (2002) Population structure of an endangered species living in contrasted habitats: *Parnassia palustris* (Saxifragaceae). *Molecular Ecology* **11**, 979-990.

Boshier, D.H., Chase, M. and Bawa, K. (1995) Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 3. Gene flow, neighbourhood, and population substructure. *American Journal of Botany* **82**, 484-490.

Brachet, S., Jubier, M.F., Richard, M., Jung-Muller, B. and Frascaria-Lacoste, N. (1999) Rapid identification of microsatellite loci using 5'anchored PCR in the common ash *Fraxinus excelsior*. *Molecular Ecology* **8**, 157-168.

Burczyk, J., Adams, W.T., Moran, G.F. and Griffin, A.R. (2002) Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. *Molecular Ecology* **11**, 2379-2391.

Cain, M.L., Milligan, B.G. and Strand, A.E. (2000) Long-distance seed dispersal in plant populations. *American Journal of Botany* **87**, 1217-1227.

Campbell, D.R. and Waser, N.M. (1989) Variation in pollen flow within and among populations of *Ipomopsis aggregata*. *Evolution* **43**, 1444-1455.

Campbell, D.R. (1991) Comparing pollen dispersal and gene flow in a natural population. *Evolution* **45** , 1965-1968.

Cascante, A., Quesada, M., Lobo, J.J. and Fuchs, E.A. (2002) Effects of dry tropical forest fragmentation on the reproductive success and genetic structure if the tree Samanea saman. *Conservation Biology* **16**, 137-147.

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

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

Charlesworth, B., Nordborg, M. and Charlesworth, D. (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genetical Research* **70**, 155-174.

Chase, M.R., Moller, C., Kesseli, R. and Bawa, K. (1996) Distant gene flow in tropical trees. *Nature* **383**, 398-399.

Cheliak, W.M. and Pitel, J.A. (1984) *Techniques for starch gel electrophoresis of enzymes from forest tree species*, Petawa National Forest Institute: Canada.

Comps, B., Gömöry, D., Letouzey, J., Thiébaut, B. and Petit, R.J. (2001) Diverging trends between heterozygosity and allelic richness during postgracial colonisation in the European beech. *Genetics* **157**, 389-397.

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

Cottrell, J.E., Munro, R.C., Tabbener, H.E., Milner, A.D., Forrest, G.I. and Lowe, A.J. (2003) Comparison of fine-scale genetic structure within two British oakwoods using microsatellites; consequences of recolonisation dynamics and past management. *Forest Ecology and Management* **176**, 287-303.

Couvet, D. (2002) Deleterious effects of restricted gene flow in fragmented populations. *Conservation Biology* **16**, 369-376.

Crandall, K.A., Bininda-Edmonds, O.R.P., Mace, G.M. and Wayne, R.K. (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution* **15**, 290-295.

Crnokrak, P. and Barrett, S.C.H. (2003) Perspective: purging the genetic load: a review of the experimental evidence. *Evolution* **56**, 2347-2358.

Cruzan, M.B. (2002) Population size and fragmentation thresholds for the maintenance of genetic diversity in the herbaceous endemic *Scutellaria montana* (Lamiaceae). *Evolution* **55**, 1569-1580.

Cunningham, S. (2000) Effects of habitat fragmentation on the reproductive ecology of four plant species in Mallee Woodland. *Conservation Biology* **14**, 758-768.

Dayanandan, S., Dole, J., Bawa, K. and Kesseli, R. (1999) Population structure delineated with microsatellite markers in fragmented populations of a tropical tree, *Carapa guianensis* (Meliaceae). *Molecular Ecology* **8**, 1585-1592.

Demesure, B., Sodzi, N. and Petit, R.J. (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondria and chloroplast DNA in plants. *Molecular Ecology* **4**, 129-131.

Demesure, B., Comps, B. and Petit, R.J. (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution* **50**, 2515-2520.

Devlin, B., Roeder, K. and Ellstrand, N.C. (1988) Fractional paternity assignment: theoretical development and comparison to other methods. *Theoretical and Applied Genetics* **76**, 369-380.

Devlin, B. and Ellstrand, N.C. (1990) The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. *Evolution* **44**, 248-259.

Dick, C.W. (2001) Genetic rescue of remnant tropical trees by an alien pollinator. *Proceedings of the Royal Society London B* **268**, 2391-2396.

Dick, C.W., Etchelecu, G. and Austerlitz, F. (2003) Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. *Molecular Ecology* **12**, 753-764.

Didham, R.K., Ghazoul, J., Stork, N.E. and Davis, A.J. (1996) Insects in fragmented forests: a functional approach. *Trends in Ecology and Evolution* **11**, 255-261.

Dow, B.D. and Ashley, V. (1996) Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. *Molecular Ecology* **5**, 615-627.

Dow, B.D. and Ashley, V. (1998) High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. *Journal of Heredity* **89**, 62-70.

Doyle, J.J. and Doyle, J.L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**, 11-15.

Dumoulin-Lapègue, S., Pémonge, M.H. and Petit, R.J. (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology* **6**, 393-397.

Dyer, R.J. and Sork, V.L. (2001) Pollen pool heterogeneity in shortleaf pine, *Pinus echinata* Mill. *Molecular Ecology* **10**, 859-866.

Edland, T. (1975) Forecast of risk of attacks on apple by the apple fruit moth (*Argyresthia conjugella* Zell.) : results from 1974 and forecast for 1975. *Gartneryrket* **65**, 462-465.

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

Ellstrand, N.C. (1984) Multiple paternity within the fruits of the wild radish, *Raphanus sativus*. *The American Naturalist* **123**, 819-828.

Ellstrand, N.C. and Marshall, D.L. (1985) Interpopulation gene flow by pollen in wild radish, *Raphinus sativus*. *The American Naturalist* **126**, 606-616.

Ellstrand, N.C. (1992) Gene flow among seed plant population. *New Forests* **6**, 241-256.

Ellstrand, N.C. and Elam, D.R. (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics* **24**, 217-242.

Endler, J.A. (1977) *Geographic Variation, Speciation, and Clines*, Princeton, New Jersey: Princeton University Press.

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

Ennos, R.A. (1998) Genetic constraints on native woodland restoration. In *Native woodland restoration in southern Scotland: principles and practice*. (Eds. Newton A.C. and Ashmole, P.) pps.27-34. Occasional paper no. 2, Borders Forest Trust, Ancrum, Jedburgh, Scotland, UK.

Ennos, R.A., Cowie, N.R., Legg, C.J. and Sydes, C. (1997) Which measures of genetic variation are relevant in plant conservation? A case study of *Primula scotica*. In: *The role of genetics in conserving small populations*, (Eds. Tew, T.E., Crawford, T.J., Spencer, J.W., Usher, M.B. and Warren, J.) pp. 73-79. Peterborough: JNCC.

Ennos, R.A., Worrell, R. and Malcolm, D.C. (1998) The genetic management of native species in Scotland. *Forestry* **71**, 1-23.

Ennos, R.A., Bacles, C. and Kettle, C. (2002) *Genetics of fragmented populations of native trees*. 21 month progress report for Scottish Forestry Trust, 7 pp University of Edinburgh.

Estoup, A., Jarne, P. and Cornuet, J.-M. (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* **11**, 1591-1604.

Fore, S.A., Hickey, R.J., Vankat, J.L., Guttman, S.I. and Schaeffer, R.L. (1992) Genetic structure after forest fragmentation: a landscape ecology perspective on *Acer saccharum*. *Canadian Journal of Botany* **70**, 1659-1668.

Franceschinelli, E.V. and Bawa, K. (2000) The effect of ecological factors on the mating system of a south american shrub species (*Helicteres brevispira*). *Heredity* **84**, 116-123.

Fuchs, E.J., Lobo, J.A. and Quesada, M. (2003) Effects of forest fragmentation and flowering phenology on the reproductive success and mating patterns of the tropical dry forest tree *Pachira quinata*. *Conservation Biology* **17**, 149-157.

Gerber, S., Mariette, S., Streiff, R., Bodenes, C. and Kremer, A. (2000) Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Molecular Ecology* **9**, 1037-1048.

Gerber, S., Chabrier, P. and Kremer, A. (2003) FAMOZ: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. *Molecular Ecology Notes* **3**, 479-481.

Ghazoul, J., Liston, K.A. and Boyle, T.J.B. (1998) Distrubance-induced density-dependent seed set in *Shorea siamensis* (Dipterocarpaceae), a tropical forest tree. *Journal of Ecology* **86**, 462-473.

Giles, B.E. and Goudet, J. (1997) A case study of genetic structure in a plant metapopulation. In: *Metapopulation biology, ecology, genetics and evolution*, (Eds. Hanski, I. and Gilpin, M.) pp. 429-454. San Diego CA: Academic Press.

Gonzalez-Martinez, S.C., Gerber, S., Cervera, M.T., Martinez-Zapater, J.M., Gil, L. and Alia, R. (2002) Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. *Theoretical and Applied Genetic* **104**, 1290-1297.

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

Goudet, J. (1998) FSTAT (version 1.2): A computer program to calculate *F*-statistics. *Journal of Heredity* **86**, 485-486.

Goudet, J. (2001) FSTAT, *a program to estimate and test gene diversities and fixation indices (version 2.9.3)*. Updated from Goudet (1995). Available from http://www.unil.izea/softwares/fstat.html.

Grime, J.P., Hodgson, J.G. and Hunt, R. (1988) *Comparative plant ecology. A functional approach to common Britsh species*, London: Unwin Hyman.

Hall, P., Walker, S. and Bawa, K. (1996) Effects of forest fragmentation on genetic diversity and mating system in a tropical tree, *Pithecellobium elegans*. *Conservation Biology* **10**, 757-768.

Hamilton, M.B. and Miller, J.R. (2003) Comparing relative rates of pollen and seed gene flow in the island model using nuclear and organelle measures of population structure. *Genetics* **162**, 1897-1909.

Hamrick, J.L. and Godt, M.J.W. (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London-B* **321**, 1291-1298.

Hansson, B. and Westerberg, L. (2002) On the correlation between heterozygosity and fitness in natural populations. *Molecular Ecology* **11**, 2467-2474.

Hardy, O.J., Gonzalez-Martinez, S.C., Fréville, H., Boquiem, G., Mignot, A., Colas, B. and Olivieri, I. Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae) 1. Pattern of pollen dispersal. *Submitted to Journal of Evolutionary Biology*

Hartl, D.L. (1988) *A primer of population genetics. Second edition.* Sunderland, Mass: Sinauer associates.

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

Hedrick, P.W. (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**, 313-318.

Henry, J.-P. and Gouyon, P.-H. (1998) *Précis de génétique des populations*, Paris: Masson.

Herbert, R., Samuel, S. amd Patterson, G., (1999) *Using local stock for planting native trees and shrubs.* Practice Note 8. Forestry Commission,. Edinburgh.

Heuertz, M., Hausman, F., Tsvetkov, I., Frascaria-Lacoste, N. and Vekemans, X. (2001) Assessment of genetic structure within and among Bulgarian populations of the common ash (*Fraxinus excelsior* L.). *Molecular Ecology* **10**, 1615-1623.

Heuertz, M. (2002) *Population genetic structure in common ash: a focus on southeastern european genetic resources.* Ph.D. thesis. Université Libre de Bruxelles.

Heuertz, M., Vekemans, X., Hausman, F., Palada, M. and Hardy, O.J. (2003) Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology* **12**, 2483-2495.

Hill, M.O. (1979) The development of a flora in even-aged plantations. In: *The Ecology of Even-aged Forest Plantations*, (Eds. Ford, E.D., Malcolm, D.C. and Paterson, J.C.) pp. 175-192. Cambridge: NERC, Institute of Terrestrial Ecology

Horn, H.S., Nathan, R. and Kaplan, S.R. (2001) Long-distance dispersal of tree seeds by wind. *Ecological Research* **16**, 877-885.

Huelden, E. (1941) Studien uber *Fraxinus excelsior* L. *Acta Botanica Fennica* **28**, 1-25.

Huenneke, L.F. (1991) Ecological implication of genetic variation in plant population. In: *Genetics and Conservation of rare plants*, (Eds. Falk, D.A. and Holsinger, K.E.) Oxford : Oxford University press.

Husband, B.C. and Schemske, D.W. (1997) The effect of inbreeding in diploid and tetraploid populations of E*pilobium angustifolium* (Onagraceae): implications for the genetic basis of inbreeding depression. *Evolution* **51**, 737-746.

Imbert, E. and Lefèvre, F. (2003) Dispersal and gene flow of *Populus nigra* (Salicaceae) along a dynamic river system. *Journal of Ecology* **91**, 447-456.

Irwin, A.J., Hamrick, J.L., Godt, M.J.W. and Smouse, P.E. (2003) A multiyear estimate of the effective pollen donor pool for *Albizia julibrissin. Heredity* **90**, 187-194.

Isagi, Y., Kanazashi, T., Suzuki, W., Tanaka, H. and Abe, T. (2000) Microsatellite analysis of the regeneration process of *Magnolia obovata* Thunb. *Heredity* **84**, 143-151.

Jamieson, A. and Taylor, St.C.S. (1997) Parentage exclusion. *Animal Genetics* **28**, 397-400.

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

Jennersten, O. (1988) Pollination in *Dianthus deltoides* (Caryophyllaceae): Effects of habitat fragmentation on visitation and seed set. *Conservation Biology* **2**, 359-366.

Katzman, M.T. and Cale, W.G. (1990) Tropical forest preservation using economic incentives. *Bioscience* **40**, 827-833.

Kaufman, S.R., Smouse, P.E. and Alvarez-Buylla, E.R. (1998) Pollen-mediated gene flow and differential male reproductive success in a tropical pioneer tree, *Cecropia obtusifolia* Bertol. (Moraceae): a paternity analysis. *Heredity* **81**, 164-173.

Keller, L.F. and Waller, D.M. (2002) Inbreeding effects in wild populations. *Trends in Ecology and Evolution* **17**, 230-241.

Kettle, C. (2001) *Founder effects and genetic structure of Sorbus aucuparia (L.) planting stock of a native woodland restoration project.* B.Sc. Honours (Ecology) thesis. University of Edinburgh.

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

Konuma, A., Tsumura, Y., Ting Lee, C., Leong Lee, S. and Okuda, T. (2000) Estimation of gene flow in the tropical-rainforest tree *Neobalanocarpus heimii* (Dipterocarpaceae), inferred from paternity analysis. *Molecular Ecology* **9**, 1843-1852.

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

Latta, R.G., Linhart, Y.B., Fleck, D. and Elliot, M. (1998) Direct and indirect estimates of seed versus pollen movement within a population of *Ponderosa pine*. *Evolution* **52**, 61-67.

Leberg, P.L. (2002) Estimating allelic richness: effects of sample size and bottlenecks. *Molecular Ecology* **11**, 2445-2449.

Lefort, F., Brachet, S., Frascaria-Lacoste, N., Edwards, K.J. and Douglas, G.C. (1999) Identification and characterisation of microsatellite loci in ash (*Fraxinus excelsior* L.) and their conservation in the olive family (Oleaceae). *Molecular Ecology* **8**, 1075-1092.

Levin, D.A. and Kerster, H.W. (1974) Gene flow in seed plants. *Evolutionary Biology* **7**, 139-220.

Li, Y.-C., Korol, A.B., Fahima, T., Beiles, A. and Nevo, E. (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* **11**, 2453-2465.

Linhart, Y.B. and Grant, M.C. (1996) Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics* **27**, 237-277.

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

Luikart, G., Allendorf, F.W., Cornuet, J.-M. and Sherwin, W.B. (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* **89**, 238-247.

Mahy, G., Vekemans, X. and Jacquemart, A.-L. (1999) Patterns of allozymic variation within *Calluna vulgaris* populations at seed bank and adult stage. *Heredity* **82**, 432-440.

Marigo, G., Peltier, J.-P., Girel, J. and Pautou, G. (2000) Success in the demographic expansion of *Fraxinus excelsior L. Trees* **15**, 1-13.

Marshall, T.C., Slate, J., Kruuk, L.E.B. and Pemberton, J. (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**, 639-655.

May, B. (1998) Starch gel electrophoresis of allozymes. In *Molecular Genetic analysis of populations. A practical approach, second edition.* (Eds. Hoetzel, A.R.).

McCauley, D. (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology and Evolution* **10**, 198-202.

McCauley, D.E. (1997) The relative contribution of seed and pollen movement to the local genetic structure of *Silene alba. Journal of Heredity* **88**, 257-263.

McCue, K.A. and Holtsford, T.P. (1998) Seed bank influence on genetic diversity in the rare annual *Clarkia springvillensis* (Onagraceae). *American Journal of Botany* **85**, 30-36.

Meagher, T. (1986) Analysis of paternity within a natural population of *Chamelirium luteum.* I. Identification of most-likely parents. *American Naturalist* **128**, 199-215.

Meagher, T.R. and Thompson, E. (1986) The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. *Theoretical Population Biology* **29**, 87-106.

Meagher, T.R. and Thompson, E. (1987) Analysis of parentage for naturally established seedlings of *Chamaelirium luteum* (Lilliaceae). *Ecology* **68**, 803-812.

Milligan, B.G., Leebens-Mack, J. and Strand, A.E. (1994) Conservation genetics: beyond the maintenance of marker diversity. *Molecular Ecology* **3**, 423-435.

Mills, S. and Allendorf, F.W. (1996) The One-migrant-per-generation rule in conservation and management. *Conservation Biology* **10**, 1509-1518

Mitchell, A. (1994) *Trees of Britain and Northern Europe*, London: HarperCollins publishers.

Morand-Prieur, M.E., Raquin, C., Shykoff, J.A. and Frascaria-Lacoste, N. (2003) Males outcompete hermaphrodites for seed siring success in controlled crosses in the polygamous *Fraxinus excelsior* (Oleaceae). *American Journal of Botany* **90** , 949-953.

Morand, M.E., Brachet, S., Rossignol, P., Dufour, J. and Frascaria-Lacoste, N. (2002) A generalized heterozyogte deficiency assessed with microsatellites in French common ash populations. *Molecular Ecology* **11**, 377-385.

Nason, J.D. and Hamrick, J.L. (1997) Reproductive and genetic consequences of forest fragmentation: Two case studies of neotropical canopy trees. *Journal of Heredity* **88**, 264-276.

Nathan, R. and Muller-Landau, H.C. (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology and Evolution* **15**, 278-285.

Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583-590.

Newton, A.C. and Ashmole, P. (1998) How may native woodland be restored to southern Scotland?. *Scottish Forestry* **52**, 168-171.

Nielsen, R., Mattila, D.K., Clapham, P.J. and Palsbøll, P.J. (2001) Statistical approaches to paternity analysis in natural populations and applications to the north Atlantic Humpback whale. *Genetics* **157**, 1673-1682.

Oddou-Muratorio, S., Petit, R.J., Le Guerroué, B., Guesnet, D. and Demesure, B. (2001) Pollen versus seed-mediated gene flow in a scattered forest tree species. *Evolution* **55**, 1123-1135.

Oddou-Muratorio, S. (2002) I*mpact des Processus démographiques et spatiaux sur la diversité génétique des arbres forestiers. Le cas d'une espèce disseminée, l'Alisier torminal (*Sorbus torminalis *L. Crantz).* Ph.D. thesis. Ecole National Du Génie Rural, Des Eaux et Des Forêts.

Pannell, J.R. and Charlesworth, B. (2000) Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society of London-B* **355**, 1851-1864.

Pemberton, J., Slate, J., Bancroft, D.R. and Barret, J.A. (1995) Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* **4**, 249-252.

Peterken, G.F. (1996) *Natural woodland: Ecology and conservation in northern temperate regions*, Cambridge: Cambridge University Press.

Petit, R.J., Kremer, A. and Wagner, D.B. (1993) Finite island model for organelle and nuclear genes in plants. *Heredity* **71**, 630-641.

Petit, R.J., El Mousadik, A. and Pons, O. (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* **12**, 844-855.

Petit, R.J., Aguinagalde, I., de Beaulieu, J.-L., *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**, 1563-1565.

Powell, A.H. and Powell, G.V.N. (1987) Population dynamics of male englossine bees in Amazonian forest fragments. *Biotropica* **19**, 176-179.

Prober, S.M. and Brown, A.H.D. (1994) Conservation of the grassy white box woodlands : population genetics of *Eucalyptus albens*. *Conservation Biology* 1003-1013.

Rameau, J.C., Mansion, D. and Dume, G. (1989) *Flore Forestière Francaise*, Paris: I*nstitut pour le developement forestier.*

Raspé, O. and Jacquemart, A.-L. (1998) Allozyme diversiy and genetic structure of European populations of *Sorbus aucuparia* L.(Rosaceae: Maloideae). *Heredity* **81**, 537-545.

Raspé, O., Jacquemart, A.-L. and de Sloover, J.R. (1998) Isozymes in *Sorbus aucuparia* (Rosaceae: Maloideae): genetic analysis and evolutionary significance of zymograms. *International Journal of Plant Science* **159**, 627-636.

Raspé, O. (1998) *Biologie de la reproduction et variation génétique d'un arbre entomophile:* Sorbus aucuparia *(Rosaceae: Maloideae)*. Ph.D. thesis. Catholic University of Louvain-la-Neuve.

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

Raspé, O., Findlay, C. and Jacquemart, A.-L. (2000b) Biological flora of the British Isles NO. 214 *Sorbus aucuparia* L. *Journal of Ecology* **88**, 910-930.

Raspé, O. (2001) Inheritance of the chloroplast genome in *Sorbus aucuparia* L. (Rosaceae). *The Journal of Heredity* **92**, 507-509.

Raspé, O. and Kohn, J.R. (2002) S-allele diversity in *Sorbus aucuparia* and *Crataegus monogyna* (Rosaceae: Maloideae). *Heredity* **88**, 458-465.

Raymond, M. and Rousset, F. (1995) GENEPOP : population genetics software for exact test and eucumenism. *Journal of Heredity* **86**, 249

Reinhartz, J.A. and Les, D.H. (1994) Bottleneck induced dissolution of self-incompatibility and breeding systems consequences in *Aster furcatus* (Asteraceae). *American Journal of Botany* **81**, 446-455.

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

Robertson, A. (2000) *The origin, evolution and conservation of the Arran Sorbus microspecies*. Ph.D. thesis. University of Edinburgh.

Robinson, G.R., Holt, R.D., Gaines, M.S., Hamburg, S.P., Johnson, M.L., Fitch, H.S. and Martinko, E.A. (1992) Diverse and Contrasting Effects of Habitat Fragmentation. *Science* **257**, 524-526.

Rocha, O. J. and Aguilar, G. (2001) Reproductive biology of the dry forest tree *Enterolobium cyclocarpum* (Guanacaste) in Costa Rica: A comparison

between trees left in pasture and trees in continuous forest. *American Journal of Botany* **88**, 1607-1614.

Roubik, D.W. and Aluja, M. (1983) Flight ranges of *Melipona* and *Trigona* in tropical forests. *Journal of the Kansas Entomological Society* **56**, 217-222.

San Cristobal, M. and Chevalet, C. (1997) Error tolerant parent identification from a finite set of individuals. *Genetical Research* **70**, 53-62.

Schierup, M.H., Charlesworth, D. and Vekemans, X. (2000) The effect of hitch-hiking on genes linked to a balanced polymorphism in a subdivided population. *Genetical Research* **76**, 63-73.

Schnabel, A. and Hamrick, J.L. (1995) Understanding the population genetic structure of *Gleditsia triacanthos* L.: The scale and pattern of pollen gene flow. *Evolution* **49**, 921-931.

Schnabel, A. (1998) Parentage analysis in plants: Mating systems, gene flow, and relative fertilities. In: *Advances in Molecular Ecology*, (Ed. Carvalho, G.R.) pp. 173-189. I.O.S. Press

Schnabel, A., Nason, J.D. and Hamrick, J.L. (1998) Understanding the population genetic structure of *Gleditsia triacanthos* L.: seed dispersal and variation in female reproductive success. *Molecular Ecology* **7**, 819-832.

Schuster, W.S.F. and Mitton, J.B. (2000) Paternity and gene dispersal in limber pine (*Pinus flexiis* James). *Heredity* **84**, 348-361.

Signorovitch, J. and Nielsen, R. (2002) PATRI-paternity inference using genetic data. *Bioinformatics* **18**, 341-342.

Sih, A., Jonsson, B.G. and Luikart, G. (2000) Habitat loss: ecological, evolutionary and genetic consequences. *Trends in Ecology and Evolution* **15**, 132-134.

Slatkin, M. (1987) Gene flow and the geographic structure of natural populations. *Science* **236**, 792

Slatkin, M. and Barton, N.H. (1989) A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* **43**, 1349-1368.

Smouse, P.E. and Meagher, T. (1994) Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) Gray (Liliaceae). *Genetics* **136**, 313-322.

Snow, B. and Snow, D. (1988) *Birds and Berries* , London: Poyser.

Soltis, D., Haufler, C., Darrow, D. and Gastony, G. (1983) Starch gel electrophoresis of ferns : a compilation of grinding buffers, gel and electrode buffers. *American Fern Journal* **73**, 410-418.

Sork, V.L., Nason, J.D., Campbell, D.R. and Fernandez, J.F. (1999) Landscape approaches to historical and contemporary gene flow in plants. *Trends in Ecology and Evolution* **14**, 219-224.

Sork, V.L., Davis, F.W., Smouse, P.E., *et al.* (2002) Pollen movement in declining populations of California Valley oak, *Quercus lobata*: where have all the fathers gone? *Molecular Ecology* **11** , 1657-1668.

Sperens, U. (1996) Is fruit and seed production in *Sorbus aucuparia* L. (Rosaceae) pollen limited? *Ecoscience* **3**, 325-329.

Stephenson, A.G., Good, S.V. and Vogler, D.W. (2000) Interrelationships among inbreeding depression, plasticity in the self-incompatibility system, and the breeding system of *Campanula rapunculoides* L -(Campanulaceae). *Annals of Botany* **85**, 211-219.

Streiff, R., Labbé, T., Bacilieri, R., Steinkellner, H., Gloessl, J. and Kremer, A. (1998) Within-population genetic structure in *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Molecular Ecology* **7**, 317-328.

Streiff, R., Ducousso, A., Lexer, C., Steinkellner, H., Gloessl, J. and Kremer, A. (1999) Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. *Molecular Ecology* **8**, 831-841.

Tapper, P.-G. (1996) Long-term survival of mast fruiting in *Fraxinus excelsior*. *Ecology* **77**, 2567-2572.

Thomé, O.W. (1885) *Flora von Deutshland, Osterreich und der Schweiz*,

Tipping, R. (1998) The application of paleoecology to native woodland restoration: Carrifran as a case-study. In *Native woodland restoration in southern Scotland: principles and practice*. (Eds. Newton A.C. and Ashmole, P.) pp.9-21. Occasional paper no. 2, Borders Forest Trust, Ancrum, Jedburgh, Scotland, UK.

Vida, G. (1994) Global issues of genetic diversity. In: *Conservation genetics*, (Eds. loeshcke, V., Tolmiuk, J. and Jain, S.K.,) Basel: Birkhauser.

Wallander, E. and Albert, V.A. (2000) Phylogeny and classification of Oleaceae based on rps16 and trnL-F sequence data. *American Journal of Botany* **87**, 1827-1841.

Wallander, E. (2001) *Evolution of wind-pollination in Fraxinus (Oleaceae) - an ecophylogenetic approach*. Ph.D. thesis. Botanical Institute, Gotebörg University.

Wardle, P. (1962) Biological flora of the British Isles NO. 1362 *Fraxinus excelsior* L. *Journal of Ecology* 739-751.

Washitani, I. (1996) Predicted genetic consequences of strong fertility selection due to pollinator loss in an isolated population of *Primula sieboldii*. *Conservation Biology* **10**, 59-64.

Weir, B.S. and Cockerham, C. (1984) Estimating *F*-Statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Wells, G.P. and Young, A.G. (2002) Effects of seed dispersal on spatial genetic structure in populations of *Rutidosis leptorrhychoides* with different levels of correlated paternity. *Genetical Research* **79**, 219-226.

Wendel, J. and Weeden, N. (1989) Visualization and interpretation of plant isozymes. In: *Isozymes in plant biology*, (Eds. Soltis, D. and Soltis, P.S.) Portland: Discirides Press

White, G.M., Boshier, D.H. and Powell, W. (1999) Genetic variation within a fragmented population of *Swietenia humilis* Zucc. *Molecular Ecology* **8**, 1899-1909.

White, G.M., Boshier, D.H. and Powell, W. (2002) Increased pollen flow conteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. *Proceedings of the National Academy of Science USA* **99**, 2038-2042.

Whitlock, M.C. and McCauley, D.E. (1999) Indirect measures of gene flow and migration: F_{st} different 1/(4Nm+1). *Heredity* **82**, 117-125.

Wilcove, D.S., McLellan, C.H. and Dobson, A.P. (1986) Habitat fragmentation in the temperate zone. In: Soule, M., (Ed.) *Conservation Biology*, pp. 237-256.

Wilcox, B.A. and Murphy, D.D. (1985) Conservation strategy : the effect of fragmentation on extinction. *American Naturalist* **125**, 875-887.

Wildwood Group of the Borders Forest Trust (2000) *Carrifran Wildwood Project. Native woodland restoration in the Southern Uplands of Scoland. Management Plan*. Newton, A.C. and Ashmole, P., (Eds.) Borders Forest Trust, Monteviot Nurseries, Ancrum, Jedburgh, Scotland, UK.

Worrell, R. (1992) A comparison between European continental and British provenances of some British native trees: Growth, survival and stem form. *Forestry* **65**, 253-283.

Worrell, R., Cundall, E.P., Malcolm, D.C. and Ennos, R.A. (2000) Variation among seed sources of silver birch in Scotland. *Forestry* 73, 419-435.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics* **16**, 97-159.

Wright, S. (1951) The genetic structure of populations. *Annals of Eugenics* **15**, 322-354

Wright, S. (1982) Character change, speciation, and the higher taxa. *Evolution* **36**, 427-443.

Young, A.G. and Clarke, G. Eds. (2000) *Genetics, demography and the viability of fragmented populations.* Cambridge. Cambridge University Press.

Young, A. and Mitchell, N. (1994) Microclimate and vegetation edge effects in a fragmented podocarp-broadleaf forest in New Zealand. *Biological conservation* **67**, 63-72.

Young, A., Boyle, T. and Brown, A.H.D. (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution* **11**, 413-418.

Young, A.G., Merriam, H.G. and Warwick, S.I. (1993) The effects of forest fragmentation on genetic variation in *Acer saccharum* Marsh. (sugar maple) populations. *Heredity* **71**, 277-289.

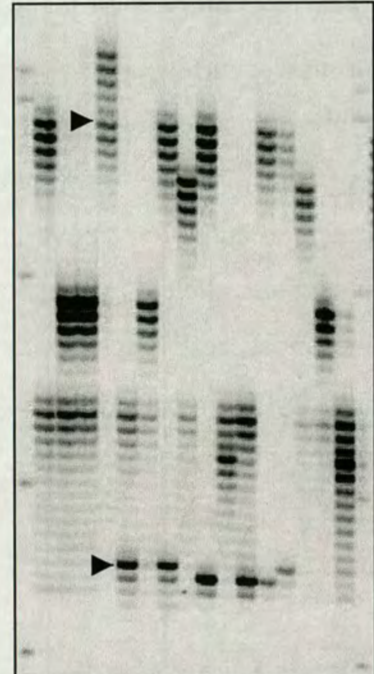
Young, A.G. and Brown, A.H.D. (1999) Paternal bottleneck in fragmented populations of the grassland daisy *Rutidosis leptorrhynchoides. Genetical Research* **73**, 111-117.

Young, A.G., Brown, A.H.D. and Zich, F.A. (1999) Genetic structure of fragmented populations of the endangered daisy *Rutidosis leptorrhynchoides. Conservation Biology* **13**, 256-265.

APPENDIX A: DESCRIPTION OF GENOTYPING PROCEDURE FOR EIGHT *FRAXINUS EXCELSIOR* MICROSATELLITE LOCI: ADDRESSING THE PROBLEM OF NULL ALLLELES AND GENOTYPING ERROR

I Genotyping strategy.

1 M2-30B



Original publication: Brachet *et al.* (1999)

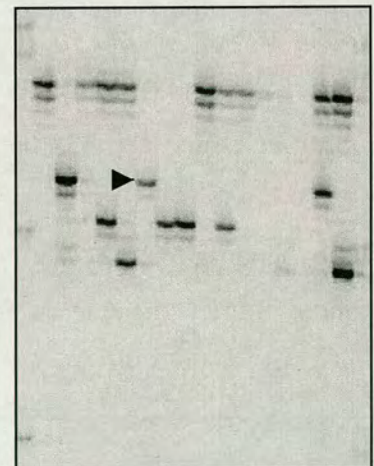
Repeat motif: $(TG)_{15}(AG)_{23}$

Expected range: 171-237 bp, 18 alleles

Observed range: 129-209 bp, 35 alleles

Nice banding pattern with 2 bp stutters of 3 to 4 weaker bands which allows good size calibration. Stutter reduced to one band for shorter alleles. Easy to score. Band of strongest intensity (2nd largest) chosen as true allele size[1].

2 3.1



Original publication: Brachet *et al.* (1999)

Repeat motif: $(AG)_{14}$

Expected range: 174-204 bp, 15 alleles

Observed range: 160-208 bp, 17 alleles

Nice banding pattern with very few stutters, almost always double band. Easy to calibrate and score. Band of strongest intensity (largest) chosen as true allele size.

[1] In each picture, arrows (▶) indicate examples of true alleles as scored for each marker.

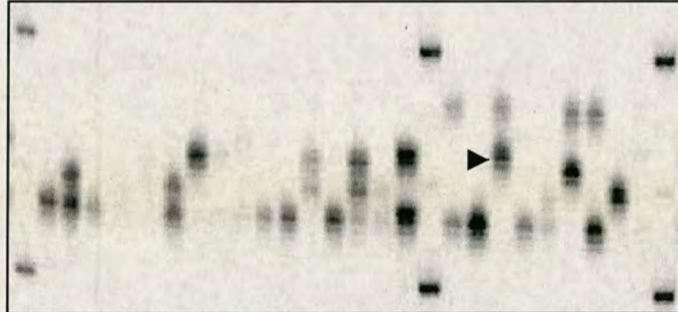
3 1.19

Original publication: Brachet *et al*. (1999)

Repeat motif: $(T)_{19}$

Expected range: 231-265 bp, 13 alleles

Observed range: 236-261 bp, 10 alleles



Very difficult to score because of the occurrence of a large number of alleles with only 1 bp size difference in a limited range. 1 bp stutter above and below true allele (strongest intensity band). Faced with difficulties in calibrating gels, the most easily scored gel was chosen as a reference, calibrating others by identifying repeated individuals between gels. However, two alleles were eventually systematically binned (246 was binned with 245 and 241 with 240) because of disagreement between gels.

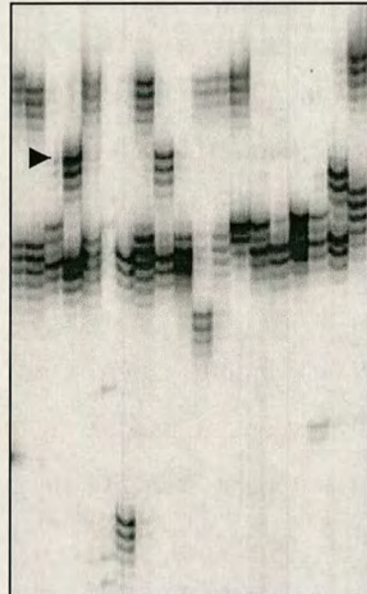
4 FEMSATL2

Original publication: Lefort *et al*. (1999)

Repeat motif: $(CT)_{28}$

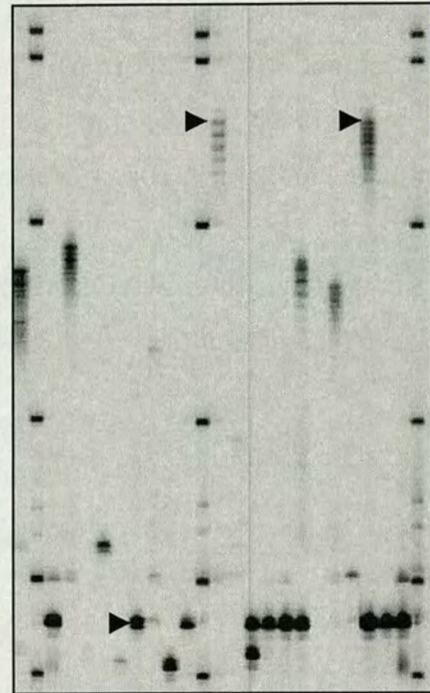
Expected range: 174-224 bp, 5 alleles

Observed range: 186-338 bp, 57 alleles



Nice banding pattern with 2 bp stutters of 3 to 4 bands. Stutter disappears for smaller sized alleles. Easy to score. Strongest intensity (largest) band chosen as true allele size.

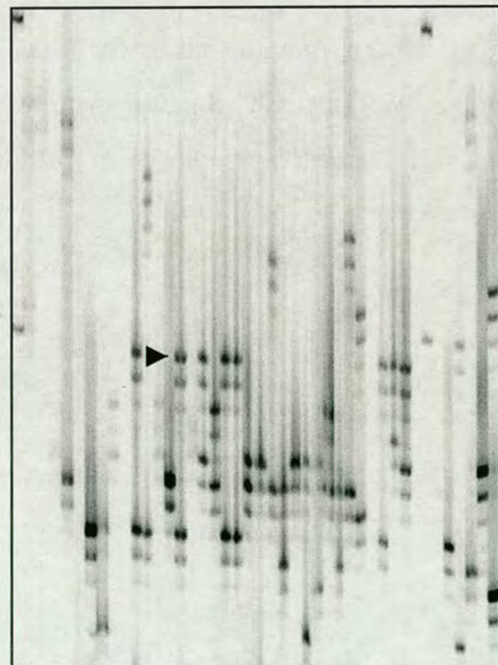
5 FEMSATL5



Original publication : Lefort *et al.* (1999)
Repeat motif : $(GA)_{41}$
Expected range : 107-183 bp, 12 alleles
Observed range: 105-231, 46 alleles

Moderately easy to score. Very different behaviour between shorter and larger alleles. Shorter alleles generally display two bands, the upper band being not clearly defined. Larger alleles display several stutter bands of 2 or 1 bp. When 1bp stutter occurs, every 2 bp, bands of stronger intensity. Largest band of strongest intensity chosen as true allele size.

6 FEMSATL8



Original publication: Lefort *et al.* (1999)
Repeat motif: $(GT)_{23}$
Expected range: 138-188 bp, 13 alleles
Observed range: 142-194 bp, 24 alleles

Nice banding pattern but size calibration of gel not so easy. Each allele displays two bands which do not correspond to a 2 bp stutter. Product at 138 bp was systematically ignored because it was observed frequently for individuals for which two other alleles could be identified.

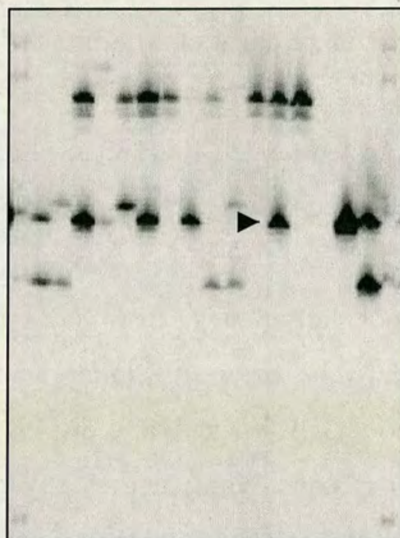
7 FEMSATL16

Original publication: Lefort *et al.* (1999)

Repeat motif: $(CA)_3(CG)(CA)_{10}(TA)_2(CA)_3$

Expected range: 180-200 bp, 4 alleles

Observed range: 174-204 bp, 7 alleles



Nice banding pattern with nearly no stutter. Easy to score and to calibrate.

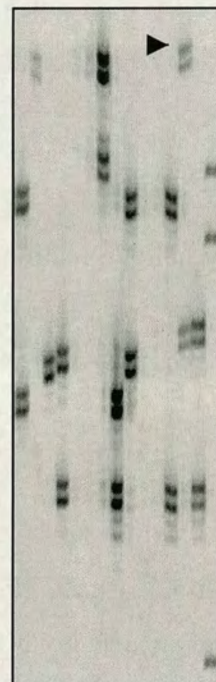
8 FEMSATL19

Original publication: Lefort *et al.* (1999)

Repeat motif: $(CA)_6cggc(CA)_{13}$

Expected range: 174-214 bp, 12 alleles

Observed range: 174-218 bp, 21 alleles



Nice banding pattern but as for FEMSATL 8, size calibration of gel was difficult. Banding pattern of two bands of similar intensity, largest band chosen as true allele size.

II Null allele frequencies at eight microsatellite loci.

Null allele frequencies were estimated under the assumption of Hardy-Weinberg equilibrium in CERVUS 2.0 (Marshall *et al.*, 1998), considering the 88 mature individuals sampled in Moffat Dale only, and are presented in Table A.1.

Table A.1: Number of alleles observed (*k*), number of genotyped individuals (*N*) and null allele frequencies at eight microsatellite loci

Locus	*k*	*N*	Null frequency[#]
3.1	**12**	**74**	**0.4179**
M2-30B	30	87	0.1255
1.19	10	77	0.1413
FEMSATL2	40	84	0.0481
FEMSATL5	**34**	**79**	**0.3321**
FEMSATL8	21	77	0.2623
FEMTSAL16	6	68	0.0833
FEMSATL19	19	82	0.2236

[#]Null allele frequencies as estimated in CERVUS 2.0 (Marshall *et al.*, 1998) under the assumption of Hardy-Weinberg equilibrium at each locus, considering a sample of 88 mature *F. excelsior* individuals.

These frequencies are indicative of the occurrence of null alleles at eight microsatellite used to screen the sampled adults, seedlings and seeds, but may not reflect true null allele frequencies as departure from Hardy-Weinberg equilibrium may be affected by other factors, such as, for instance, non-random mating patterns in the sampled populations or sub-structuring of populations leading to heterozygote deficit.

III Genotypes of progeny arrays at five microsatellite loci and estimation of error rates.

Genotypes at five microsatellite markers of seeds collected from 19 trees producing fruits in the year of collection in Moffat Dale are presented in Table A.2. In some arrays, a large number of mismatches that cannot be solely attributed to sampling or genotyping errors are observed at loci 3.1 and FEMSATL5 (Table A.2) Occurrence of one or several null alleles at these loci is a strong possibility. However, non-Mendelian segregation is not entirely resolved by taking into account such phenomenon, therefore loci 3.1 and FEMSATL5 were discarded from analyses.

Table A.2: Genotypes of 19 *F excelsior* progeny arrays at five microsatellite loci. [§]Maternal parent. Genotypes of progeny which mismatch the genotype of the maternal parent are highlighted in bold.

			Locus		
Individual	**M2-30B**	**1.19**	**FEMSATL2**	**3.1**	**FEMSATL5**
CDa102[§]	**197149**	**243236**	**250196**	**182182**	**207125**
1055	197171	243236	250250	182182	205125
1056	197171	243243	**260240**	196182	**203203**
1057	197197	243236	**222194**	196182	125125
1058	**153153**	**244244**	**252244**	182170	**203155**
1059	**171171**	243236	**252218**	182182	**205205**
1060	197131	243236	222196	182182	203125
1133	197133	236236	276250	182170	207109
1134	197131	243239	222196	182182	**205205**
1135	197131	236236	196196	182170	125125
1136	197171	243239	250222	196182	**203203**
1137	197197	242236	276250	182170	125125
1138	149131	243236	250222	182182	205125
1139	149149	242236	262250	182170	**203109**
1140	197197	243236	250222	182182	125109
1141	149131	243236	250222	182182	125109
1142	173149	243243	230196	194182	125125
1156	149131	243243	250222	182182	207125
1205	171149	243236	250244	182182	203125
1206	171149	243236	250244	196182	203125
1207	197197	243239	264196	194182	139125
1208	197133	236236	276196	182170	**203109**
1209	197149	236236	276196	182170	**203109**
1210	197197	243239	222196	182182	205125
1211	**131131**	243243	222196	196182	203125
1212	**171171**	239236	262250	182182	125125
1213	197171	243239	250250	182182	125109
1214	197191	243243	208196	182182	**203113**
1215	197171	239236	222196	182182	205125
1217	**171171**	243243	250244	182182	205125
CDa105[§]	**197189**	**236236**	**260250**	**196180**	**169169**
1246	000000	236236	280260	180176	**109109**
1247	197131	243236	260222	196182	**109109**
1248	189189	236236	260260	180180	**109109**
1249	189171	236236	250250	196182	169109
1250	207197	239236	260222	196176	**187109**
1251	197197	236236	260250	180180	169109
1252	197171	236236	260246	182180	169109
1253	197197	244236	258250	180180	**163163**
1254	189171	236236	260260	196180	169109
1255	189133	236236	276260	196176	**167109**
1256	197171	236236	250230	196196	169169
1257	**199165**	236236	258250	180180	169109
1258	197189	239236	250222	196176	169109
1301	189131	236236	260222	196182	**109109**
1302	197171	236236	260222	000000	169109
1061	197131	236236	250222	196182	169169
1062	197171	243236	260246	196180	**167167**
1063	197189	236236	264250	182180	169169

Individual	M2-30B	1.19	FEMSATL2	3.1	FEMSATL5
CDa105§	**197189**	**236236**	**260250**	**196180**	**169169**
1064	189189	236236	274260	194180	**109109**
1065	197171	236236	250222	196196	**167167**
1066	189189	239236	260222	196190	187169
1067	189171	236236	250196	182180	**109109**
1068	000000	243236	260244	196180	203169
1239	197189	236236	260222	196176	**187187**
1240	189131	236236	260222	196180	169169
1241	197131	243236	250222	196180	169109
1242	189189	243236	266250	196170	**117117**
1243	207197	239236	250216	190180	**187187**
1244	197171	243236	260232	182180	**167167**
1245	189171	243236	260232	196196	**167167**
CDa108§	**189145**	**236236**	**250236**	**180180**	**151151**
1036	189145	245236	**302280**	180180	**141109**
1037	205145	239236	236218	180176	**109109**
1038	189133	243236	262250	180176	151109
1039	163145	236236	258236	180180	**141141**
1040	145133	243236	250238	182180	**141141**
1041	145145	236236	276250	180170	000000
1042	189163	245236	**260252**	180180	**109109**
1043	207189	236236	**252218**	**196176**	**187109**
1098	207189	236236	250218	180176	151109
1218	189145	236236	260250	180180	151109
1219	189189	236236	258250	180180	**141141**
1220	205189	239236	254250	180180	**109109**
1221	189189	240236	**258240**	186180	195151
1222	189171	239236	**244240**	182180	203151
1223	189189	239236	**222222**	180176	187151
1224	145145	245236	260236	180180	**109109**
1225	145135	244236	264236	180180	151151
1226	189189	236236	258250	180180	**109109**
1227	145145	243236	276250	180176	**141141**
1228	145145	239236	250224	180180	**141109**
1229	189163	239236	**238238**	182180	151151
1230	171145	249236	244236	180176	**109109**
1231	145145	249236	250222	180170	**109109**
1232	189133	249236	250218	180180	**107107**
1233	177145	236236	266236	184180	**117117**
1234	199145	243236	236206	180180	185151
1235	189163	236236	258250	180180	151151
1236	145145	243236	276250	180170	**141109**
1237	163145	245236	260236	180180	**141109**
1238	189189	245236	260250	180180	**109109**
CDa111§	**189173**	**249236**	**250220**	**196196**	**169125**
1099	207189	249239	**280260**	**176176**	187169
1100	189173	239236	220220	**190190**	125109
1101	207189	249239	250216	**176176**	125109
1102	189149	242236	250218	**170170**	169169
1103	173131	249236	262250	196196	169109
1104	189133	236236	256220	**176176**	169109
1105	173173	249236	250220	196196	125109

Individual	**M2-30B**	**1.19**	**FEMSATL2**	**3.1**	**FEMSATL5**
CDa111[§]	**189173**	**249236**	**250220**	**196196**	**169125**
1106	207173	239236	220220	**176176**	125109
1107	173131	249243	220220	196182	201125
1108	189189	236236	220220	**176176**	187169
1277	189153	249243	250188	000000	169169
1278	173131	249236	222220	000000	125109
1279	189171	249236	250250	**182182**	167125
1280	189189	249243	260250	196196	125125
1281	189171	249243	250250	**182182**	167125
1282	173173	249243	220220	196182	169109
1283	173131	243236	222220	196182	125109
1284	173173	243236	290220	196192	169169
1285	173173	243236	246220	196196	167125
1286	189173	249239	222220	196190	125109
1287	189189	249239	268250	198198	125109
1288	189171	249236	250244	196196	169109
1289	173173	249249	268250	196196	125109
1290	000000	236236	250246	000000	169169
1291	207173	249243	250250	196190	169109
1292	199189	249239	250218	000000	125109
1293	189149	236236	250250	196170	177125
1294	189171	242236	276250	000000	169169
1295	173131	239236	250250	**182182**	169109
1296	189173	243236	246220	196196	167125
CDa112[§]	**197189**	**237236**	**252222**	**196176**	**167109**
1044	189131	242236	246222	196196	167109
1045	000000	240236	246222	182176	125109
1046	207189	236236	**250218**	196176	109109
1047	199197	236236	258222	196196	109109
1048	189149	236236	**262250**	176176	167167
1157	189131	243236	252232	196182	167109
1158	189189	236236	258222	180176	109109
1159	000000	240236	252222	196170	109109
1160	197171	243237	222222	182176	167109
1161	189189	244236	246222	176170	201109
1162	197131	239236	252222	196182	125109
1163	189189	243237	252246	196176	167109
1164	207197	239237	252218	176176	109109
1165	197189	244237	258222	180176	**163163**
1166	197189	239236	222222	196176	187167
1197	000000	236236	252230	196176	109109
1303	199189	245236	260222	176176	167109
1304	189189	249237	280252	176176	167109
1305	197183	245237	256252	176176	**163163**
1306	189165	245236	258252	180176	169109
1307	189173	245237	252220	196180	169169
1308	189189	245237	260252	180176	167167
1309	189189	249237	**260250**	180176	169109
1310	000000	245237	222222	176176	167167
1311	197189	245237	**260220**	196196	167167
1312	197131	236236	222222	196196	167109
1313	189171	236236	250222	196196	167167
1314	189133	236236	276222	196170	109109

Individual	M2-30B	1.19	FEMSATL2	3.1	FEMSATL5
CDa112[§]	**197189**	**237236**	**252222**	**196176**	**167109**
1315	**199169**	236236	276222	182176	109109
1316	**199199**	242237	222202	190176	187109
CDa118[§]	**197155**	**240236**	**250192**	**196196**	**149109**
1071	171155	244240	232192	196182	109109
1072	171155	239236	250250	196196	203109
1073	207197	236236	250216	196176	109109
1074	155133	243236	262250	196170	109109
1075	171155	244236	250232	**182182**	149109
1076	197133	244236	250232	**176176**	187109
1077	189155	240236	276192	**176176**	109109
1078	197131	244240	250250	196196	149125
1079	155149	243240	262192	196176	149109
1080	155131	244236	222192	196196	109109
1333	197131	236236	**222194**	196196	169149
1334	197133	243236	276192	196170	149149
1335	189155	240236	260192	000000	109109
1336	171155	243240	250250	196182	203149
1337	155149	243236	276192	196170	109109
1338	197189	244240	260192	000000	109109
1339	201155	240239	272192	000000	109109
1340	155149	240236	260192	196176	109109
1341	207155	240236	216192	**190190**	149109
1342	197165	244240	256192	196196	149149
1343	000000	240236	250222	000000	149109
1344	189155	236236	258192	196180	149149
1345	197131	243240	222192	182182	149109
1346	197171	243236	250232	196196	169149
1347	207197	236236	216192	196176	149109
1348	197197	243240	222192	196182	169109
1349	155131	243236	250250	196196	125109
1350	189155	240240	222192	196190	149109
1351	197131	243240	**222194**	196196	201149
1352	155149	236236	**262194**	**176176**	109109
CDa123[§]	**189155**	**243236**	**296250**	**176176**	**169117**
1069	000000	242236	296276	000000	117117
1070	189133	242236	262250	176170	169169
1081	155149	236236	296262	176176	169169
1082	155149	236236	296262	176170	117109
1083	155133	236236	296262	**170170**	117109
1084	189149	242236	260250	**180180**	183117
1085	155131	242236	296222	196176	117109
1086	155133	236236	296276	176170	169109
1372	155131	**239239**	296222	182170	125117
1373	189189	244236	264250	180170	169169
1374	000000	**242242**	262250	000000	117109
1375	000000	000000	222194	000000	000000
1376	189133	236236	296262	176170	169169
1377	199155	239236	296296	176176	183169
1378	189131	243242	250222	**182170**	169109
1379	155133	242236	296262	176170	169169
1380	155139	243242	260250	182176	**143115**

Individual	M2-30B	1.19	FEMSATL2	3.1	FEMSATL5
CDa123[§]	**189155**	**243236**	**296250**	**176176**	**169117**
1381	155149	242236	296262	**170170**	**109109**
1382	155131	**242239**	250222	**196170**	125117
1383	189171	243242	250250	196176	169109
1384	171155	242239	296222	**182170**	169125
1385	189189	245236	296260	**180170**	169109
1386	155149	242236	296262	176176	169169
1387	189131	243242	**298222**	**170170**	117117
1388	155155	**242239**	000000	**170170**	169169
1389	155149	**242242**	276250	176170	169169
1390	189131	242236	250222	182176	**165109**
1391	207155	237236	296216	**190170**	169109
1392	155149	243236	**298298**	180176	169169
1393	155149	236236	264250	176170	**109109**
CDa125[§]	**189149**	**243236**	**260248**	**174170**	**167159**
1049	189149	243240	260222	**176176**	167109
1050	207149	243236	248222	190174	**187187**
1051	207189	236236	260216	**190190**	187167
1052	189135	243243	260230	**180180**	167159
1053	207189	243240	260216	**176176**	187187
1054	189189	240236	260216	**176176**	187187
1204	173149	236236	252248	174170	169169
1394	149131	243240	248222	**196196**	125125
1395	207149	236236	248222	**190190**	167109
1396	149149	240236	248214	**176176**	173167
1415	189135	243240	260260	174170	183167
1416	165149	243240	260216	**180180**	**187187**
1417	207189	236236	248216	**176176**	167109
1418	189149	243240	260216	**190190**	**187187**
1419	189189	243236	260216	**190190**	**109109**
1420	207189	236236	260216	176170	**109109**
1421	209189	240236	260216	**190190**	**187187**
1422	207189	236236	260216	**190190**	**109109**
1423	207189	000000	248216	**176176**	187167
1424	000000	000000	260222	000000	187167
1425	189149	240236	248216	**176176**	187167
1426	197149	243243	264248	000000	187167
1427	209189	243243	248222	**176176**	**187187**
1428	189149	245236	258248	**180180**	167109
1429	207149	236236	260222	**176176**	167109
1430	189189	236236	260222	**190190**	187167
1431	207149	243240	248216	**190190**	167109
1432	189149	236236	260216	**190190**	**109109**
1433	189189	243240	260216	**176176**	187167
1434	000000	240236	260222	000000	**109109**
CDa126[§]	**189149**	**236236**	**274250**	**180170**	**165109**
1109	207149	236236	274216	190180	185165
1110	171149	236236	274244	182170	109109
1111	189189	236236	250222	190170	185165
1112	207189	236236	250216	176170	185109
1113	189189	239236	250222	180176	185165
1114	207149	236236	274222	190180	185109

Individual	**M2-30B**	**1.19**	**FEMSATL2**	**3.1**	**FEMSATL5**
CDa126[§]	**189149**	**236236**	**274250**	**180170**	**165109**
1115	149131	244236	250232	182170	109109
1116	189145	242236	250240	194180	109109
1117	207149	239236	250216	176170	185165
1118	207149	239236	250216	190170	165109
1259	189149	236236	250222	**182176**	185165
1260	189171	236236	250232	196170	165109
1261	189189	236236	274262	000000	165165
1262	189149	239236	274216	190180	109105
1263	207189	236236	250216	190180	185165
1264	149139	236236	290250	170170	113109
1265	149131	249236	274208	180170	165113
1266	189149	236236	274222	190180	165109
1267	189149	236236	250222	176170	109109
1268	207189	236236	274216	190170	185109
1269	189149	236236	274222	180176	165109
1270	189149	236236	274222	180176	185165
1271	207149	236236	274216	176170	185109
1272	149131	236236	274222	196180	169109
1273	207149	236236	274216	190180	165109
1274	207189	236236	250216	190170	165109
1275	189189	236236	274222	190170	185109
1276	207189	000000	274216	190170	109109
1299	189165	243236	258250	180180	165165
1300	207189	236236	250216	190170	185109
CDa129[§]	**197131**	**244237**	**280250**	**176176**	**163109**
1087	**155133**	261237	280280	**170170**	169169
1088	**189133**	243237	280222	176170	000000
1089	207131	237237	280280	190176	163109
1090	197133	244237	280280	176170	109109
1091	197133	244237	258250	**170170**	109109
1092	197171	244236	250232	196176	163109
1093	161131	244236	**276194**	176176	163163
1094	207131	244239	250216	190176	109109
1095	155131	261237	280280	176176	109109
1096	197131	244237	280222	182170	125109
1097	197197	237237	280280	**170170**	109109
1353	133131	244237	280280	176170	109109
1354	189131	244237	258250	**170170**	163163
1355	197171	244236	250232	182176	163109
1356	197171	244236	250250	176176	109109
1357	171131	237236	280250	**170170**	163163
1358	131131	249237	250250	176170	109109
1359	197149	244244	262250	176170	109109
1360	189131	244244	280264	180176	000000
1361	**189133**	244236	280262	176170	109109
1362	149131	244236	276250	176176	163109
1363	197189	244236	280216	176176	163109
1364	197197	245237	264250	180176	169163
1365	189131	240237	250222	190176	109109
1366	133131	243237	276250	176176	163109
1367	197131	244239	250222	**196170**	203163
1368	131131	244244	280222	**196170**	163163

Individual	M2-30B	1.19	FEMSATL2	3.1	FEMSATL5
CDa129§	**197131**	**244237**	**280250**	**176176**	**163109**
1369	197171	244244	250250	**196170**	203109
1370	197189	244244	280264	180176	109109
1371	197171	237237	250250	176170	163109
CDa130§	**149133**	**239236**	**276262**	**176170**	**125109**
1186	189133	244236	262262	176176	109109
1187	149149	242236	276262	176176	109109
1188	**171171**	239239	262248	192170	195109
1189	**189189**	242236	262262	180170	109109
1190	149131	243236	276276	182176	167109
1191	**189189**	243243	276260	180170	167000
1192	000000	242236	262262	000000	109109
1193	149149	242236	276262	170170	109109
1194	165133	236236	262262	170170	109109
1195	133131	239236	276276	196170	125125
1196	149131	236236	262222	182170	**167167**
1198	191133	239236	276222	176176	185109
1199	191133	239236	276222	176176	**201201**
1203	149149	236236	262262	170170	000000
1317	**171171**	239236	262244	196176	125109
1318	000000	243236	262216	196170	**167167**
1319	171149	242236	262232	000000	109109
1320	207133	236236	000000	176170	185109
1321	171149	239236	262222	196176	**201201**
1322	149131	239239	262222	182176	125109
1323	171149	239239	262244	182176	125109
1324	189149	236236	262216	190176	109109
1325	189149	239236	276220	176170	185109
1326	000000	244236	276276	000000	**201201**
1327	000000	239239	276220	000000	125125
1328	149133	**242242**	276262	176170	109109
1329	149149	242236	262216	176170	109109
1330	**207207**	242239	276222	190170	109109
1331	207133	**242242**	276232	176176	177109
1332	133133	242239	276222	182170	125109
CMa24§	**149131**	**249236**	**246242**	**194194**	**113113**
1173	000000	236236	246186	000000	**171121**
1174	147131	249244	242226	194194	000000
1175	195131	236236	256246	000000	**125125**
1176	165149	240236	246204	**176170**	**107107**
1177	185149	249243	246204	**190170**	113113
1178	000000	249236	242242	000000	000000
1179	**135135**	**240240**	246222	194194	**107107**
1180	149135	245236	246206	**170170**	113113
1181	201149	245236	242220	194170	**191191**
1183	149133	249243	242232	194174	113109
1184	149131	249236	252242	194194	113113
1185	171149	249243	252246	194182	113113
1435	163131	249249	246246	**170170**	169113
1436	171131	000000	246246	194170	000000
1437	145131	239236	242242	**190190**	**165165**
1438	161131	244236	242226	194194	**121121**

Individual	M2-30B	1.19	FEMSATL2	3.1	FEMSATL5
CMa24§	**149131**	**249236**	**246242**	**194194**	**113113**
1439	197131	249240	246204	**198198**	191113
1119	189149	243236	242222	194194	113105
1120	163149	249242	254242	194194	121113
1122	149149	249249	246236	000000	**109109**
1440	171131	243236	246222	**170170**	113113
CMa26§	**197131**	**249240**	**246222**	**194182**	**113113**
1182	159131	245240	246198	194194	000000
1121	197131	249242	256246	194194	113113
1123	139131	244240	222222	194190	113113
1124	189131	249244	296222	182182	113107
1125	149131	249244	222208	182182	**115111**
1126	169131	249249	302246	182182	165113
1127	149131	249236	246246	000000	**161161**
1128	163131	249236	246246	000000	000000
1200	131131	244240	296246	182180	**137121**
1201	169131	249239	302246	194190	**201201**
1202	145131	249240	302246	190182	113113
SCa33§	**137137**	**243243**	**274256**	**182182**	**189107**
1023	137137	243243	**272266**	182170	107107
1024	137137	243243	266256	**170170**	189189
1025	137137	243243	**272254**	182182	189189
1026	165137	243243	274244	182182	107107
1027	137137	243243	**272254**	**170170**	173107
1028	199137	000000	278256	182170	000000
1029	137137	243243	256256	182182	173107
1030	137137	243243	256256	182180	153107
1031	137137	243240	256256	000000	107107
1297	137131	243236	318256	182170	189125
1298	137137	243243	274202	**170170**	189161
1397	153137	243243	260256	182182	189125
1398	167137	243243	256256	**170170**	189189
1399	137137	000000	256256	**170170**	189151
1400	137137	243243	256256	182182	107107
1401	137137	243243	274254	182180	153107
1402	137131	243239	256250	182170	161107
1403	137137	243239	**272244**	**190190**	155107
1404	173137	243243	262256	000000	113107
1405	137137	243243	**272254**	**180180**	107107
1406	161137	243243	**272266**	**170170**	189173
1407	137137	243243	**272254**	182182	153107
1408	159137	243243	256256	182182	153107
1409	159137	000000	256256	182182	000000
1410	159137	000000	**276254**	182180	153107
1411	137137	243243	256256	182182	189173
1412	137137	243243	256256	**180180**	107107
1413	167137	243243	268256	**170170**	189173
1414	167137	243243	272256	182170	189173
1006	000000	243243	266256	000000	189189
SCa34§	**185185**	**243243**	**280250**	**170170**	**183183**
1032	185139	243243	280250	170170	**181181**
1033	185167	243243	266250	170170	00000

Individual	M2-30B	1.19	FEMSATL2	3.1	FEMSATL5
SCa34[§]	**185185**	**243243**	**280250**	**170170**	**183183**
1034	**139139**	243243	280252	170170	**181181**
1035	**155155**	243243	**256256**	000000	**175175**
1143	**133133**	243236	280206	**180180**	**181121**
SCa35[§]	**185147**	**261243**	**272250**	**170170**	**151151**
1013	185185	261243	274250	000000	151151
1014	185137	261261	272254	**180180**	**153153**
1015	147131	261239	272206	000000	151151
1016	185159	243243	254250	000000	151151
1017	159147	261243	272254	**180180**	**153153**
1018	147137	243243	272254	000000	**153153**
1019	185135	261243	**274274**	**188188**	**149149**
1020	159147	261261	272254	000000	**153153**
1021	185185	243240	272206	000000	151113
1022	185173	249243	254250	172170	**159159**
1441	185137	261243	272254	000000	**153153**
1442	185159	261243	272254	**180180**	151151
1443	185167	243243	272254	170170	000000
1444	185159	243243	254250	000000	**153153**
1445	147147	261243	264250	**184184**	**163163**
1446	145137	000000	250250	**180180**	**163163**
1447	145137	000000	272254	**180180**	**153153**
1448	185165	000000	272266	170170	**173153**
1449	000000	000000	272254	180170	**153153**
1450	185137	243243	272266	**180180**	173151
1451	185155	243240	272210	**196196**	151151
1452	145145	243240	300272	170170	163151
1453	201185	261243	296250	**194176**	205151
1454	145137	243243	272254	000000	**153153**
1475	185159	243243	**254254**	000000	000000
1476	185159	243243	254250	**180180**	**153153**
1477	185165	243243	250218	000000	117117
1478	193185	243243	272214	170170	**187187**
1007	185137	261243	272254	000000	151151
1008	185159	261243	254250	000000	153153
Sca36[§]	**185177**	**240240**	**276250**	**170170**	**153153**
1129	000000	000000	250244	000000	000000
1130	177137	000000	254250	000000	000000
1131	000000	000000	250196	**174174**	000000
1132	185137	243240	276254	170170	153153
1152	177137	243240	276254	180170	153153
1153	201185	243240	250222	180170	**165165**
1154	197185	240240	276250	**180180**	**167167**
1155	207185	240240	250250	170170	000000
1455	177159	**242242**	276254	180170	153153
1456	185153	**236236**	276220	000000	153153
1457	177137	000000	**278254**	180170	153153
1458	177137	242240	276254	180170	153153
1459	197177	240240	256250	180170	**171171**
1460	177177	245240	266250	194170	**125125**
1461	177137	242240	276254	**180180**	**151125**
1462	185135	243240	250210	170208	**165151**

Individual	**M2-30B**	**1.19**	**FEMSATL2**	**3.1**	**FEMSATL5**
Sca36[§]	**185177**	**240240**	**276250**	**170170**	**153153**
1463	000000	240240	276250	**180180**	**151151**
1464	185137	240240	266250	170170	173153
1465	185137	240240	254250	170170	153153
1466	185177	240240	276242	170170	153125
1467	177137	242240	276254	000000	153153
1468	177137	**243243**	276254	180170	153153
1469	185137	240240	254250	**180180**	153125
1470	185135	249240	296250	000000	**155155**
1471	185159	**242242**	254250	**180180**	153153
1472	185159	000000	250250	000000	**151151**
1473	185153	240240	260250	**180176**	153153
1474	185137	240240	276256	170170	153153
1009	177163	240240	**280280**	174170	**121121**
1010	177131	243240	276222	170170	153153
SCa38[§]	**195139**	**243243**	**262250**	**170170**	**000000**
1167	187139	243243	250222	000000	173173
1168	185139	243243	254250	**188188**	173173
1170	185139	243243	254250	**180180**	117117
1171	**187153**	243243	250222	**188182**	153153
1172	000000	243243	**242196**	000000	153109
SCa41[§]	**205185**	**243243**	**250196**	**188170**	**000000**
1144	205139	243243	254196	188180	151151
1145	185161	243243	254250	170170	153153
1146	205161	243243	254196	180170	151151
1147	185139	243243	250250	188188	000000
1148	205139	243243	254196	188188	121121
1149	000000	243243	254250	000000	151151
1150	185161	243243	254250	188180	151125
1151	205161	243243	208196	188180	121121
1479	205139	243243	294254	180170	151151
1480	205139	243243	254250	180170	151151
1481	205161	243243	254250	170170	151151
1482	185161	243243	254250	000000	151151
1483	205133	243243	244196	188188	173173
1484	205139	243243	254250	188180	151151
1485	185161	243243	254250	188180	151151
1486	205161	243243	254196	188170	151151
1487	185161	243243	250250	180170	151151
1488	185139	243243	254250	170170	121121
1489	205139	243243	252250	180170	151151
1490	185161	243243	254250	180170	151151
1491	185155	243240	254250	170170	175175
1492	205139	243243	254250	188180	000000
1493	185167	243243	222196	182170	175175
1494	205139	243243	254196	188188	151151
1495	185161	243243	254250	188180	151151
1496	205139	243243	254196	000000	121121
1497	185139	243243	254196	170170	151151
1498	205161	243243	254250	188180	151151
1011	205189	243243	222196	188188	113113
1012	185153	243243	232196	188188	125125

Excluding information at loci 3.1 and FEMSATL5, a total of 70 seeds were discarded from analyses because their genotype did not have any allele in common with the genotype of their maternal parent at one or more of the three loci, M2-30B, 1.19 and FEMSATL2 (genotypes highlighted in bold in Table A.2).

Based on mismatches between offspring and maternal parents' genotypes, error rates at each of five microsatellites were calculated in CERVUS 2.0, and are presented in Table A.3. The mean error rate (*e*) across loci M2-30B, 1.19 and FEMSATL2, and assuming that all parent-offspring pairs compared are equally independent, was estimated to be less than 5% (*e*=0.0420).

Table A.3: Global error rate estimated in CERVUS 2.0 (Marshall *et al.*, 1998) at each of five *F. excelsior* microsatellite loci from open-pollinated progeny arrays.

Locus	N seeds	N mismatches	Error rate
FEMSATL5	437	132	0.2226
3.1	437	133	0.2890
M2-30B	466	26	0.0373
1.19	474	15	0.0380
FEMSATL2	491	38	0.0508

Sources of error may be sampling, genotyping or actual mutation at a particular marker. To try untangling the importance of each of these factors, I attempted to provide an estimation of error rates for each of these factors by considering a few further assumptions and excluding information at loci 3.1 and FEMSATL5. It is reasonable to assume that, if a sampling error has been made, it is most likely that a progeny will not match the genotype of the presumed maternal parent at more than one locus; while genotyping error or mutations are more likely to result in mismatches at only one locus. Furthermore, while the proportion of seeds with mismatches at only one locus provide a combined estimate of genotyping error and mutation, it is possible to provide a crude estimate of stepwise mutation rate at each locus by considering the proportion of mismatched alleles that only differ from alleles in the maternal genotype by the length of a mono or a dinucleotide repeat depending on the microsatellite sequence (but considering a maximum of one mismatch for a particular genotype). Obviously such differences may be due to

genotyping errors, however, amplification and run for all seed presented mismatches were repeated, thus minimising the potential for genotyping error.

Only 3 seeds out of 493 tested mismatch the genotype of their maternal parent at more than one (out of three loci), thus providing a sampling error rate of less than 1%. (0.6%). 18 individuals mismatched the genotype of their maternal parent at M2-30B only, and six of these mismatches correspond to a 2 bp difference in length. This would estimate that the 'genotyping error' rate at this locus is about 1.8% and that the 'stepwise mutation rate' is about 0.6%. 13 individuals mismatched the genotype of their maternal parent at 1.19 only, and five of these mismatches correspond to a 1 bp difference in length. This would estimate that the 'genotyping error' rate at this locus is about 1.3% and that the 'stepwise mutation' rate is about 0.5%. Finally, 33 individuals mismatched the genotype of their maternal parent at FEMSATL2 only, and 23 of these mismatches correspond to a 2 bp difference in length. This would estimate that the 'genotyping error' rate at this locus is about 3.3% and that the 'stepwise mutation' rate is about 2.3%.

IV Procedure for allele binning

Since high genotyping error rates (whether due to mutation or actual scoring mistake) were estimated from progeny arrays at three of the five loci used in paternity and parentage analyses, it seems necessary to find a strategy to minimise the impact of these factors on estimates of gene flow for the studied populations.

Both genotyping errors and stepwise mutations are most likely to arise between alleles of similar length, the actual length of rarer alleles may also be more difficult to determine accurately. Therefore, I chose to account for error by binning rare alleles with more common alleles of nearest size for both paternity (Chapters 4 and 5) and parentage (Chapter 6) analyses, following the procedure detailed below. Such procedure results in reducing the exclusion power at a particular locus and consequently provides more conservative estimates of gene flow. Considering the focus of this thesis, it seemed most appropriate to provide conservative estimates of gene flow and dispersal.

Since a number of different alleles occurred in seed and seedling samples at loci M2-30B and FEMSATL2, the binning procedure at these loci varies slightly for data sets used in paternity and parentage analyses. In general, at all loci except 1.19, alleles found in frequency of less than 0.01 were binned with more common alleles of nearest size. At 1.19, a procedure reflecting difficulties in gel scoring was used. For paternity analyses, allele frequencies were estimated in FSTAT 2.9.3 (Goudet, 2001) considering adult and seed data sets pooled together. For parentage analyses, allele frequencies were estimated considering adult and seedling data sets pooled together.

1 M2-30B

Table A.4: Allele binning procedure and allele frequencies at locus M2-30B for paternity analysis.

allele	frequency before binning	binned with allele	frequency after binning
129	**0.002**	131	--
131	0.096		0.098
133	0.030		0.030
135	0.020		0.020
137	0.068		0.068
139	0.022		0.022
145	0.036		0.036
147	0.013		0.013
149	0.095		0.105
151	**0.005**	149	--
153	**0.005**	155	--
155	0.036		0.041
159	0.013		0.013
161	0.013		0.022
163	**0.009**	161	--
165	0.013		0.020
167	**0.007**	165	--
169	**0.009**	171	--
171	0.065		0.074
173	0.028		0.028
175	**0.002**	177	--
177	0.014		0.016
183	**0.001**	185	--
185	0.063		0.064
187	**0.003**	189	--
189	0.160		0.166
191	**0.003**	189	--
193	**0.002**	197	--
195	**0.002**	197	--
197	0.093		0.104
199	**0.007**	197	--
201	**0.004**	205	--
205	0.018		0.022
207	0.039		0.041
209	**0.002**	207	--

Table A.5: Allele binning procedure and allele frequencies at locus M2-30B for parentage analysis.

allele	frequency before binning§	binned with allele	frequency after binning§
129	**0.006**	131	--
131	0.087		0.093
133	**0.009**	135	--
135	0.069		0.078
137	0.105		0.105
139	0.021		0.021
145	0.027		0.027
147	0.018		0.018
149	0.060		0.060
151	0.021		0.021
153	**0.006**	155	--
155	0.021		0.027
159	0.039		0.039
161	**0.003**	163	--
163	0.015		0.018
165	0.027		0.027
167	**0.003**	169	--
169	0.021		0.024
171	0.060		0.060
173	0.012		0.012
175	**0.006**	177	--
177	0.030		0.042
179	**0.006**	177	--
183	**0.003**	185	--
185	0.066		0.069
187	0.018		0.018
189	0.099		0.099
193	0.012		0.012
195	**0.003**	197	--
197	0.075		0.078
199	0.012		0.012
201	0.018		0.018
205	0.012		0.018
207	**0.006**	205	--

2 1.19

Table A.6: Allele binning procedure and allele frequencies at locus 1.19 for paternity analysis.

allele	frequency before binning	binned with allele	frequency after binning
236	0.344		0.378
237	0.034	236	--
239	0.052		0.052
240	0.094		0.094
242	0.038	243	--
243	0.289		0.327
244	0.051	245	--
245	0.022		0.073
249	0.061		0.061
261	0.015		0.015

Table A.7: Allele binning procedure and allele frequencies at locus 1.19 for parentage analysis.

allele	frequency before binning	binned with allele	frequency after binning
236	0.137		251
237	0.114	236	--
239	0.016		0.016
240	0.133		0.133
242	0.056	243	--
243	0.330		0.386
244	0.082	245	--
245	0.020		0.102
249	0.088		0.088
261	0.013		0.013

3 FEMSATL2

Table A.8: Allele binning procedure and allele frequencies at locus FEMSATL2 for paternity analysis.

allele	frequency before binning	binned with allele	frequency after binning
186	**0.001**	192	--
188	**0.001**	192	--
192	0.019		0.021
196	0.027		0.028
198	**0.001**	196	--
202	**0.002**	206	--
204	**0.006**	206	--
206	0.008		0.022
208	**0.004**	206	--
210	**0.002**	206	--
212	**0.004**	216	--
214	**0.002**	216	--
216	0.043		0.049

allele	frequency before binning	binned with allele	frequency after binning
218	**0.008**	220	--
220	0.029		0.037
222	0.102		0.105
224	**0.001**	222	--
226	**0.002**	222	--
230	**0.006**	232	--
232	0.014		0.020
234	**0.001**	236	--
236	0.012		0.015
238	**0.002**	236	--
240	**0.003**	242	--
242	0.016		0.019
244	0.014		0.014
246	0.037		0.037
248	0.014		0.014
250	0.207		0.207
252	0.023		0.023
254	0.053		0.053
256	0.049		0.049
258	0.016		0.016
260	0.057		0.057
262	0.039		0.039
264	0.011		0.011
266	0.010		0.014
268	**0.004**	266	--
272	0.022		0.022
274	0.022		0.022
276	0.040		0.042
278	**0.002**	276	--
280	0.024		0.031
282	**0.006**	280	--
284	**0.001**	280	--
286	0.002		0.002
290	**0.003**	296	--
296	0.022		0.029
298	**0.004**	296	--
300	0.001		0.004
302	**0.003**	300	--
306	0.001		0.003
318	**0.001**	306	--
320	**0.001**	306	--

Table A.9: Allele binning procedure and allele frequencies at locus FEMSATL2 for parentage analysis.

allele	frequency before binning	binned with allele	frequency after binning
192	0.012		0.015
194	**0.003**	192	--
196	0.031		0.031
202	**0.003**	204	--
204	0.012		0.015
206	0.019		0.019
210	**0.003**	212	--
212	0.016		0.019
216	0.022		0.022
218	**0.009**	220	--
220	0.025		0.036
222	0.109		0.109
230	0.006		0.006
234	**0.003**	236	--
236	0.019		0.028
238	**0.006**	236	--
240	**0.006**	242	--
242	0.028		0.034
244	0.012		0.012
246	0.038		0.038
248	**0.009**	246	--
250	0.161		0.046
252	0.043		0.043
254	0.062		0.062
256	0.059		0.059
258	**0.009**	260	--
260	0.028		0.037
262	0.028		0.049
264	**0.006**	262	--
266	**0.006**	262	--
268	**0.009**	262	--
270	**0.006**	274	--
272	**0.006**	274	--
274	0.019		0.031
276	0.025		0.025
278	0.012		0.012
280	0.022		0.022
282	0.025		0.025
284	**0.006**	286	--
286	0.006		0.009
290	**0.003**	286	--
296	**0.006**	300	--
298	**0.012**	300	--
300	0.006		0.024
306	0.003		0.009
320	**0.003**	306	--
338	**0.003**	306	--

4 FEMSATL8

Table A.10: Allele binning procedure and allele frequencies at locus FEMSATL8 for parentage analysis.

allele	frequency before binning	binned with allele	frequency after binning
142	0.051		0.051
144	0.006		0.009
146	**0.003**	144	--
150	0.035		0.035
154	0.025		0.025
156	0.019		0.019
158	0.105		0.105
160	0.019		0.019
162	0.083		0.083
164	0.258		0.258
166	0.032		0.032
168	0.022		0.022
170	**0.006**	172	--
172	0.073		0.079
174	**0.003**	176	--
176	0.108		0.111
178	0.013		0.013
180	**0.003**	182	--
182	0.013		0.016
184	**0.003**	186	--
186	0.025		0.028
190	0.025		0.025
192	0.054		0.067
194	**0.013**	192	--

5 FEMSATL19

Table A.10: Allele binning procedure and frequencies at FEMSATL19 for parentage analysis.

allele	frequency before binning	binned with allele	frequency after binning
174	0.041		0.041
176	0.035		0.035
178	0.050		0.050
180	**0.013**	182	--
182	0.038		0.051
184	0.066		0.066
186	0.110		0.110
188	0.126		0.126
190	0.101		0.101
192	0.167		0.167
194	0.013		0.013
196	0.016		0.025
198	**0.009**	196	--
200	0.088		0.088
202	0.016		0.016
204	0.016		0.016
206	**0.006**	208	--
208	0.047		0.053
210	0.016		0.016
212	0.025		0.028
218	**0.003**	212	--

V References

Brachet S, Jubier MF, Richard M, Jung-Muller B, Frascaria-Lacoste, N. (1999) Rapid identification of microsatellite loci using 5'anchored PCR in the common ash *Fraxinus excelsior*. *Molecular Ecology,* **8**, 157-168.

Goudet J (2001) *FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3).* Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995)

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

Lefort F, Brachet S, Frascaria-Lacoste N, Edwards KJ, Douglas GC (1999) Identification and characterisation of microsatellite loci in ash (*Fraxinus excelsior* L.) and their conservation in the olive family (Oleaceae). *Molecular Ecology,* **8**, 1075-1092.

Marshall TC, Slate J, Kruuk LEB, Pemberton J (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology,* **7**, 639-655.

Molecular Ecology (2004) 13, 573–584 doi: 10.1046/j.1365-294X.2004.02093.x

Genetic effects of chronic habitat fragmentation on tree species: the case of *Sorbus aucuparia* in a deforested Scottish landscape

C. F. E. BACLES,*† A. J. LOWE†‡ and R. A. ENNOS*
*Institute of Cell, Animal and Population Biology, Ashworth Laboratories, School of Biological Sciences, University of Edinburgh, Edinburgh, EH9 3JT, UK, †Centre for Ecology and Hydrology-Edinburgh, Bush Estate, Penicuik, Midlothian, EH26 0QB, UK, ‡School of Life Sciences, University of Queensland, Brisbane, Qld., 4072, Australia

Abstract

Sustainable forest restoration and management practices require a thorough understanding of the influence that habitat fragmentation has on the processes shaping genetic variation and its distribution in tree populations. We quantified genetic variation at isozyme markers and chloroplast DNA (cpDNA), analysed by polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP) in severely fragmented populations of *Sorbus aucuparia* (Rosaceae) in a single catchment (Moffat) in southern Scotland. Remnants maintain surprisingly high levels of gene diversity (H_E) for isozymes ($H_E = 0.195$) and cpDNA markers ($H_E = 0.490$). Estimates are very similar to those from non-fragmented populations in continental Europe, even though the latter were sampled over a much larger spatial scale. Overall, no genetic bottleneck or departures from random mating were detected in the Moffat fragments. However, genetic differentiation among remnants was detected for both types of marker (isozymes $\Theta_n = 0.043$, cpDNA $\Theta_c = 0.131$; *G*-test, *P*-value < 0.001). In this self-incompatible, insect-pollinated, bird-dispersed tree species, the estimated ratio of pollen flow to seed flow between fragments is close to 1 ($r = 1.36$). Reduced pollen-mediated gene flow is a likely consequence of habitat fragmentation, but effective seed dispersal by birds is probably helping to maintain high levels of genetic diversity within remnants and reduce genetic differentiation between them.

Keywords: cpDNA, isozymes, *F*-statistics, habitat fragmentation, pollen vs. seed migration, *Sorbus aucuparia* L.

Received 31 July 2003; revision received 14 November 2003; accepted 14 November 2003

Introduction

Extensive anthropogenic use of natural resources to meet the needs of the ever-expanding human population has led to growing concern about the detrimental effects of deforestation and habitat fragmentation. A substantial body of literature has documented the influence of deforestation on community and ecosystem processes (Wilcove *et al.* 1986; Robinson *et al.* 1992; Benitez-Malvido & Martinez-Ramos 2003). However, few studies have measured empirically the genetic consequences of habitat fragmentation on population structure, and most of the available data are from tropical forest tree species (Aldrich *et al.* 1998; White *et al.* 1999; Aguirre-Plantier *et al.* 2000; Dick 2001). In addition, little effort has been made to integrate genetic factors into forest restoration and management programmes (Ennos *et al.* 1998).

Habitat fragmentation decreases the size and potentially increases the isolation of forest fragments (Andren 1994). It also reduces the availability of suitable colonization sites for the foundation of new populations (Wilcox & Murphy 1985). Theoretical predictions are that habitat fragmentation will decrease genetic variation within population and increase interpopulation genetic differentiation, affecting population viability in the short and long term (Young *et al.* 1996). These effects are due mainly to an increase in the level of inbreeding and genetic drift in remnants associated with small census size and reduced gene flow between fragments (Couvet 2002).

Correspondence: Cecile Bacles. Fax: 0131 4453943; E-mail: cfb@ceh.ac.uk

© 2004 Blackwell Publishing Ltd

Loss of genetic variation (Dayanandan *et al.* 1999; White *et al.* 1999) and increased inbreeding (Fuchs *et al.* 2003) have been observed in fragmented populations of several tropical tree species. However, there is growing empirical evidence for enhanced gene flow between isolated trees in forest fragments (Young *et al.* 1993; White *et al.* 2002; Dick *et al.* 2003). It appears that the effects of habitat fragmentation on the genetic behaviour of tree populations are more varied and complex than first expected (Aldrich & Hamrick 1998). Only by studying these issues in a wide range of species and ecological conditions will we gain the information needed to formulate sustainable management practices, especially for temperate species, for which few data are available.

The objectives of the present study are to assess the genetic diversity and genetic structure of highly fragmented populations of the temperate tree species, *Sorbus aucuparia*, in the southern Uplands of Scotland. In particular, we are interested in the amount of genetic variation remaining within remnant populations, the extent of differentiation between remnants, and the inferred levels of historical gene flow occurring among fragments.

S. aucuparia L. (Rosaceae, Maloideae), rowan or mountain ash, is a small tree, native to most of Europe. In southern Scotland it grows commonly along riverbeds on rocky mountain slopes and cliffs. At lower altitude, the species behaves as a hardy pioneer or post-pioneer and is usually succeeded by late successional species (Raspé *et al.* 2000a). At high altitude, it is one of the few species to maintain the tree habit where it forms part of the late successional vegetation. It can grow at altitudes up to 1000 m in Scotland, higher than any other tree (Mitchell 1989). Seed bearing begins at 15 years of age and the maximum lifespan of *S. aucuparia* is 150 years. Small, hermaphrodite flowers grouped into inflorescence are pollinated mainly by flies but also attract beetles and bees (Raspé *et al.* 2000b). Selfing is prevented by a gametophytic system of self-incompatibility (Raspé & Kohn 2002). The species produces fleshy fruits eaten by birds, such as thrushes, blackbirds and starlings, which are its main seed dispersers (Snow & Snow 1988).

In southern Scotland, forest cover is extremely limited with, in some instances, less than 1% of native woodland remaining (Newton & Ashmole 1998). Major human impact in this area has been estimated to have begun as early as 6000 years ago (Wildwood Group of the Borders Forest Trust 2000). Ancient land use for pasture and much more recent establishment of conifer plantations have fragmented the forested landscape greatly. Many native tree species, including *S. aucuparia*, are confined to steep riversides not accessible to grazing herds. Remnants are very small, comprising typically 10–30 mature individuals with no or very little natural regeneration in grazed areas. Remnants are separated typically by hundreds of metres and may be isolated from the nearest remnant by a distance of more than 1 km. As a result of their exposed situation in a windy climate and under continuous grazing, trees are shrub-like, undersized and display a convoluted shape. There is an immediate need to assess the population genetic structure of these forest fragments because they are being considered as potential seed sources for native woodland restoration in Scotland.

Previous isozymes studies have shown that European populations of *S. aucuparia* ranging from Finland to the Pyrenees maintain high gene diversity within population and exhibit low interpopulation differentiation compared to other temperate tree species (Raspé & Jacquemart 1998). CpDNA variation in populations sampled in several regions of France and Belgium also demonstrates low differentiation and high diversity (Raspé *et al.* 2000a). However, one study of Belgian populations separated by only a few kilometres found atypically high genetic differentiation at isozyme markers (Raspé & Jacquemart 1998) and this was explained by a founder effect in these recently colonized populations (Raspé & Jacquemart 1998).

In this study, we use isozymes and cpDNA markers to quantify genetic variation and to assess interpopulation genetic differentiation between *S. aucuparia* remnants. The contrasting mode of inheritance of isozymes, coded by biparentally inherited nuclear genes, and the chloroplast genome, inherited maternally in *S. aucuparia* (Raspé 2001), is especially helpful to infer the relative importance of pollen and seed migration in historical gene flow processes (Ennos 1994). Previous studies describing the genetic structure of non-fragmented European populations provide a useful comparison to assess the extent to which habitat fragmentation has altered patterns of gene diversity in *S. aucuparia* remnants in southern Scotland.

Materials and methods

Study site and populations

The study site is located 80 km south of Edinburgh in Dumfries and Galloway (N55°23′51″, W3°19′50″). The study focuses on a single catchment that has been sampled exhaustively for *S. aucuparia*. Eight populations were defined within the Carrifran valley and its immediate surroundings (Fig. 1). Remnants are typically small in size (10–30 individuals, Table 1), are separated by a distance of at least 300 m and are on average 1.5 km apart. Trees tend to grow on steep river sides, away from grazing herds, along with other native species such as *Fraxinus excelsior*, *Corylus avellana*, *Betula pubescens* and *Crataegus monogyna*. One fragment, in Swine Cleuchs, is a pure stand of 10 trees situated within a 30-year-old conifer plantation.

In five of the eight populations, all identified mature trees were sampled. In the two larger remnants, Spoon Burn and Whitewells (Fig. 1), samples of 19 and 20 randomly

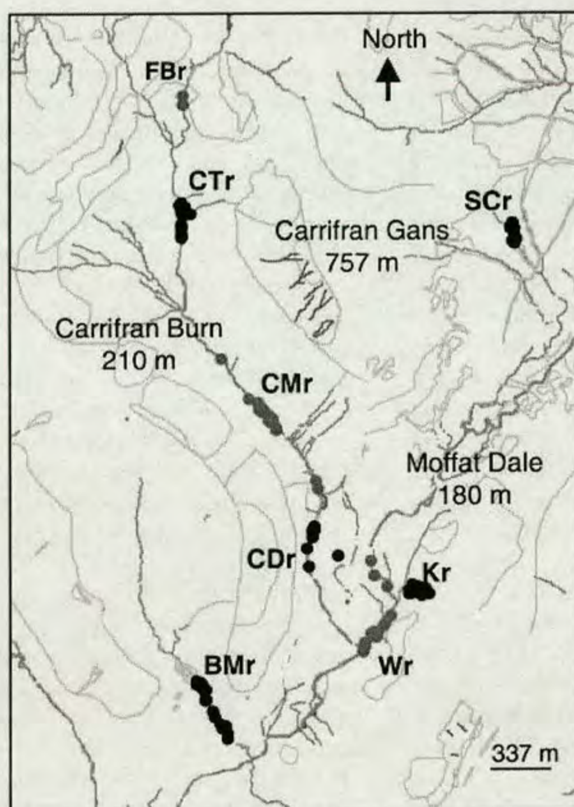


Fig. 1 Distribution of *S. aucuparia* individuals sampled in Moffat Dale remnants. Trees are grouped into populations as described in Table 1. Each dot represents a tree. The background map is a section of Ordnance Survey product Land-Line. Plus-nt11 © Crown copyright Ordnance Survey. An EDINA digimap/JISC supplied service.

Table 1 Location, abbreviated name, landscape characteristics, altitude and size of eight *S. aucuparia* remnants present in Moffat Dale. Altitude given in metres, size expressed as the approximate number of individuals censed. Census sampling in all fragments but *; 19, 2 and 20 individuals were sampled in BMr, FBr and Wr, respectively

Location	Abbreviation	Site characteristics	Altitude	Size
Spoon Burn	BMr	Grazed riverside	262	50–70*
Carrifran (down)	CDr	Grazed riverside	186	13
Carrifran (middle)	CMr	Grazed riverside	215	21
Carrifran (top)	CTr	Grazed riverside	332	27
Firthope Burn	FBr	Grazed riverside	464	< 10*
Bodesbeck	Kr	Grazed pasture	215	12
Whitewells	Wr	Protected riverbank	170	40–50*
Swine Cleuchs	SCr	Grazed riverside in conifer plantation	384	10

distributed individuals were taken, respectively. In Firthope Burn, tiny shrub-like individuals hang onto very steep cliffs in poor conditions and only two individuals (in a remnant of less than 10 trees) were accessible for sampling, bringing the total sample to 123 individuals. The sampling strategy reflects the patchy nature of the landscape (Table 1). The *S. aucuparia* remnants sampled vary in size, spatial isolation, altitude and accessibility to grazing and are ideal for documenting the genetic effects of long-standing habitat fragmentation.

Plant material

Plant material collection was carried out in winter 1999 and spring 2000. Dormant buds were sampled initially to allow optimization of molecular marker methods. In spring a resampling was performed. Twigs with breaking buds or freshly emerged leaves were collected from the 123 individuals. The latter were retained for isozyme screening and provided the optimum material for electrophoretic analysis.

Protein extraction and electrophoretic procedure

Protein extraction was performed on fresh material, immediately after collection. Small buds or young leaves were crushed on ice using extraction buffer (Cheliak & Pitel 1984) prepared freshly and stored at −20 °C. Immediately prior to extraction 3.2 μL/mL of 2-β mercaptoethanol and 0.03 g of polyvinyl-pyrrolidone 40 (PVP-40) were added to the extraction buffer. Extracted proteins were absorbed on filter paper wicks and stored at −80 °C until electrophoresis. Electrophoresis was carried out at +4 °C, on 13% starch (Sigma) gels using standard gel and electrode buffers (Soltis *et al.* 1983; Wendel & Weeden 1989) and staining recipes (Cheliak & Pitel 1984). Choice of isozyme systems and electrophoretic conditions was guided by the developmental work of Raspé *et al.* (1998), based on scorability and reproducibility of results.

Conditions were optimized successfully for eight enzyme systems: aspartate-aminotransferase (AAT, EC 2.6.1.1), leucine-aminopeptidase (LAP, EC 3.4.11.1), malate dehydrogenase (MDH, EC 1.1.1.37), menadione reductase (MNR, EC 1.6.99.2), peroxidase (PRX, EC 1.11.1.7), 6-phosphogluconate dehydrogenase (6-PGD, EC 1.1.1.44), phosphoglucomutase (PGM, EC 2.7.5.1) and shikimic acid dehydrogenase (SDH, EC 1.1.1.25). Fifteen putative loci were identified and phenotypes at AAT, MDH, PRX, 6-PGD and PGM were scored in accordance with zymograms and genetic analysis reported previously (Raspé *et al.* 1998). Eight loci encoding six enzymes found to be polymorphic either in this study (prx-1, pgm-1, pgm-2, pgm-3, sdh-1, aat-1) or previously (Raspé & Jacquemart 1998; 6-pgd-1, 6-pgd-2), were retained for detailed analysis.

DNA isolation and PCR-RFLP analysis of cpDNA

DNA was isolated from 100 mg of frozen leaf or bud material following a standard extraction protocol adapted from Doyle & Doyle (1987) and described by Rendell & Ennos (2002). Extracted DNA was dissolved in 100 μL of Tris-ethylene-diamine-tetraacetic acid (TE, pH 7.6). All individuals were characterized by restriction analysis of three cpDNA fragments AS, DT and HK amplified by PCR with the universal primers of Demesure *et al.* (1995). PCR reactions were carried out in a GeneAmp® PCR system 9700 (PE Applied Biosystems) following the conditions of Raspé *et al.* (2000a). PCR products of fragments AS, DT and HK were digested with *Hin*fI (Raspé *et al.* 2000a) and run on 1.8% agarose gels in 0.5× TBE buffer; 7.5 μL of each PCR product was loaded and 500 ng of 1 kb ladder (Gibco-BRL Life Technologies) was included as a size marker. Agarose gels were stained with ethidium bromide and photographed with a DOC-IT gel documentation system (UVP) under ultraviolet light.

Statistical analysis

Genetic diversity. Unbiased estimates of gene diversity (H_E) per population and for the total sample were calculated for each isozyme and cpDNA locus using FSTAT 2.9.3 (Goudet 2001). Mean H_E estimates over eight isozyme loci at the population and species level as well as other standard estimates of genetic diversity [percentage of loci polymorphic at the 95% level (*P*95%) and mean number of alleles observed per locus (A_0)] were obtained using GENETIX 4.01 (Belkhir *et al.* 1998). However, these measures are prone to bias due to variation in sample size (Leberg 2002). To correct for variation in sample size, we used the rarefaction method described by El Mousadik & Petit (1996). Total allelic richness (R_t) was estimated for each isozyme locus in FSTAT 2.9.3. Allelic richness, over all isozyme loci per population (R_s), was also computed excluding populations FBr and BMr (because of missing values at one locus) and including all populations for chloroplast markers.

Departure from random mating. The inbreeding coefficient (F_{IS}) was estimated over all isozyme loci for each population and departure from Hardy–Weinberg equilibrium was assessed using FSTAT 2.9.3. Tests for deficit in heterozygotes (*HD*) and excess in heterozygotes (*HE*) were performed by randomizing alleles among individuals. Luikart *et al.* (1998) demonstrated that populations that have gone through a severe and recent genetic bottleneck show a transient excess of heterozygotes. To test whether fragmentation in *S. aucuparia* populations has caused a genetic bottleneck, a sign test (Cornuet & Luikart 1996) and Wilcoxon test (Luikart *et al.* 1998) were performed using BOTTLENECK (Cornuet & Luikart 1996). In the analysis, all individuals were pooled together as a single sampled population, i.e. the bottlenecked population, and all enzyme loci are assumed to fit an infinite allele model of mutation (IAM).

Genetic differentiation. Global inbreeding (f) and genetic differentiation (Θ_n) coefficients were estimated following Weir & Cockerham (1984) for each isozyme locus and overall loci using FSTAT 2.9.3. The genetic differentiation coefficient was also estimated for the cpDNA data set (Θ_c). All estimations excluded population FBr (where only two individuals were sampled). Departure from Hardy–Weinberg equilibrium within samples was tested by comparing the distribution of the F_{IS} statistic (f) for the observed data set with its distribution for a randomized data set obtained after 10 000 permutations of alleles among individuals within samples. A *G*-test (Goudet *et al.* 1996), based on 10 000 permutations of genotypes among samples, was performed to test for population differentiation at isozyme loci and at cpDNA markers. Whenever possible, pairwise tests of differentiation were performed for each population

pair (excluding FBr) using the G-statistic (Goudet *et al.* 1996) by randomization of multilocus genotypes for isozymes, and for cpDNA markers using FSTAT 2.9.3.

Independence of loci. Calculation of mean estimates of F-statistics assumes independence of loci. Genotypic linkage equilibrium between all polymorphic isozyme loci and between isozyme and cpDNA loci was tested using GENEPOP 3.2a (Raymond & Rousset 1995). Furthermore, homogeneity of F-statistics estimates between isozyme loci was tested by jackknifing in GENETIX 4.01.

Pollen vs. seed migration ratio. Several authors have derived equations to infer the relative rate of pollen- and seed-mediated gene flow (Petit *et al.* 1993; Ennos 1994; Hamilton & Miller 2003). Ennos (1994) showed that under the assumption of an island model of migration (Wright 1951), at migration-drift equilibrium, a ratio (r) of the amount of migration by pollen (m_P) by the amount of migration by seed (m_s) can be inferred from F-statistics estimated for biparental markers (F_{STB}) and for uniparentally inherited maternal markers (F_{STM}) as:

$$r = \frac{m_P}{m_s} = \frac{\left(\frac{1}{F_{STB}} - 1\right)(1 + F_{IS}) - 2\left(\frac{1}{F_{STM}} - 1\right)}{\frac{1}{F_{STM}} - 1} \quad \text{(eqn 1)}$$

We estimated the r-ratio for our populations of *S. aucuparia* by substituting in eqn 1 F_{STB} with Θ_n and F_{IS} with f estimated over all loci, while Θ_c provided an estimate of F_{STM} as maternal inheritance of the chloroplast genome has been demonstrated in *S. aucuparia* (Raspé 2001).

Results

Genetic diversity at isozyme loci

Across the 123 individuals scored, six polymorphic isozyme loci were identified (prx-1, sdh-1, pgm-1, pgm-2, pgm-3, aat-1) and a total of 15 alleles were recorded (Table 2). Allele frequencies (p) per locus and population are detailed in Table 2. The number of alleles at polymorphic loci varies between two (sdh-1, pgm-3, aat-1) and three (prx-1, pgm-2, pgm-3). 6-pgd-1 and 6-pgd-2 were monomorphic in the sampled populations. Among polymorphic loci, H_E varies between 0.073 (aat-1) and 0.549 (pgm-2), whereas R_t varies between 1.145 (aat-1) and 2.409 (pgm-2). A_o varies between 1.200 for population FBr and 2.000 for Kr (Table 2).

No population is monomorphic for all loci but no population includes all 15 alleles. Population Kr is most polymorphic ($P95\%$ = 0.667) while FBr is least polymorphic ($P95\%$ = 0.200). H_E and R_s are relatively homogeneous across populations, although CTr is most diverse when diversity is measured by H_E and Kr is most diverse when R_s is considered (Table 2). Values of $P95\%$, A_o and H_E were 0.421, 1.392 and 0.163 when averaged over populations and 0.500, 2.000 and 0.195 overall. F_{IS} ranges between −0.500 (FBr) and 0.344 (CDr) and a significant deficit of heterozygotes was detected solely for the latter suggesting a departure from random mating in this population (Table 2).

No genetic bottleneck was detected for the sampled *S. aucuparia* fragmented populations either with a sign test (P-value = 0.222) or with a one-tailed Wilcoxon test (P-value (HE) = 0.055), although this was close to significance at the 5% level.

CpDNA haplotypes and diversity

Three indels were detected using three primer/restriction enzyme combinations. Single locus haplotype nomenclature was attributed to polymorphic fragments according to Raspé *et al.* (2000a). Combining information from those three indels allowed the identification of four haplotypes in the sampled populations (Table 3).

Haplotype A was most frequent overall [P(A) = 0.681, Table 4] whereas haplotype D was found in only two individuals, one in population CTr and one in CDr (Table 4). Haplotype B was more frequent in CMr than in any other population, and haplotype C was not found in populations CMr and Wr (Fig. 2). Estimation of allelic richness based on a minimum of two genes is highest for population CTr, with a mean over populations of 2.021 (Table 4). Gene diversity varies between 0 (population FBr) and 0.530 (population K) and is high overall (H_E = 0.490).

Population genetic structure in S. aucuparia remnants

Across isozyme loci, significant genotypic association was found only for one pair of loci across all populations (sdh-1 and pgm-2; P-value = 0.013), suggesting that all except these loci are independent. Homogeneity of inbreeding parameters among isozyme loci was verified because jackknifed estimates of f and Θ for isozyme loci were within the 95% confidence interval of the jackknife mean (Results not shown). Excluding one or other locus did not alter the trend observed in global estimates and multilocus estimates reported include all six polymorphic isozyme loci.

Estimates of f per locus vary between −0.146 (sdh-1) and 0.190 (pgm-1). Departure from Hardy–Weinberg equilibrium was detected at locus pgm-1 only (Table 5). f estimated over all loci and populations is 0.024 and not significantly different from zero. On average, there is no tendency towards heterozygote excess or deficiency in the sampled populations and therefore no evidence of inbreeding.

Estimates of Θ are consistently low among nuclear isozyme loci, but population differentiation was detected for loci prx-1, pgm-2 and aat-1 (Table 5), whether assuming

Table 2 Allele frequencies (p(i)), gene diversity (H_E) and total allelic richness (R_t) for six isozyme loci in eight *S. aucuparia* remnants of Moffat Dale and averaged over loci. Also shown are mean over loci gene diversity (H_E), percentage of polymorphic loci ($P95\%$) and number of alleles (A_o) per population and overall, allelic richness (R_s), inbreeding coefficient estimates (F_{IS_EST}), P-values for heterozygote deficit (HD) and excess (HD). NS: non-significant at 5% level. †Based on a minimum sample size of 89. ‡Based on a minimum sample size of 4. All analyses were conducted using FSTAT 2.9.3 (Goudet 2001) excluding estimation of H_E, $P95\%$ and A_o per population over all loci performed in GENETIX 4.01 (Belkhir *et al.* 1998). Mean values include loci monomorphic for the sampled populations

	Population								
Locus	CDr	CMr	SCr	CTr	FBr	Kr	Wr	BMr	Overall
prx-1	$N = 13$	20	10	27	2	10	20	19	121
$p(1)$	0.500	0.750	0.750	0.537	1.000	0.750	0.825	0.816	0.702
2	0.500	0.200	0.150	0.426	0.000	0.200	0.175	0.053	0.248
3	0.000	0.050	0.100	0.037	0.000	0.050	0.000	0.132	0.050
H_E	0.526	0.404	0.433	0.538	0.000	0.411	0.297	0.325	0.443
R_t†									1.860
sdh-1	13	18	5	11	2	7	20	19	95
1	0.731	0.639	0.500	0.682	0.500	0.714	0.500	0.711	0.637
2	0.269	0.361	0.500	0.318	0.500	0.286	0.500	0.289	0.363
H_E	0.404	0.467	0.550	0.455	0.500	0.452	0.511	0.424	0.463
R_t									1.822
pgm-1	9	14	4	26	0	8	19	12	92
1	0.833	0.786	0.750	0.846	NA	0.688	0.632	0.792	0.767
2	0.167	0.179	0.000	0.115		0.250	0.316	0.208	0.190
3	0.000	0.036	0.250	0.038		0.063	0.053	0.000	0.043
H_E	0.306	0.368	0.417	0.274		0.518	0.509	0.356	0.375
R_t									2.126
pgm-2	12	21	10	27	0	10	20	13	113
1	0.625	0.476	0.650	0.370	NA	0.750	0.550	0.654	0.540
2	0.375	0.524	0.350	0.574		0.150	0.200	0.346	0.394
3	0.000	0.000	0.000	0.056		0.100	0.250	0.000	0.066
H_E	0.508	0.512	0.478	0.538		0.422	0.611	0.468	0.549
R_t									2.409
pgm-3	13	17	8	22	2	8	19	0	89
1	0.962	0.971	1.000	0.886	0.750	0.938	0.974	NA	0.944
2	0.038	0.029	0.000	0.114	0.250	0.063	0.026		0.056
H_E	0.077	0.059	0.000	0.206	0.500	0.125	0.053		0.106
R_t									1.250
aat-1	13	21	7	26	2	10	20	19	118
1	1.000	1.000	1.000	1.000	1.000	0.900	0.825	1.000	0.962
2	0.000	0.000	0.000	0.000	0.000	0.100	0.175	0.000	0.038
H_E	0.000	0.000	0.000	0.000	0.000	0.189	0.297	0.000	0.073
R_t									1.145
Overall									
H_E	0.174	0.158	0.143	0.210	0.075	0.183	0.204	0.153	0.195
$P95\%$	0.333	0.333	0.333	0.500	0.200	0.667	0.500	0.400	0.500
A_o	1.500	1.667	1.500	1.833	1.200	2.000	1.833	1.600	2.000
R_s‡	1.513	1.571	1.556	1.651		1.795	1.710		
F_{IS_EST}	0.344	−0.062	−0.012	−0.081	−0.500	0.180	−0.047	0.203	
P-value HD	< 0.01	NS	NS	NS	NS	NS	NS	NS	
P-value HE	NS	NS	NS	NS	NS	NS	NS	NS	

Hardy–Weinberg equilibrium or not. Overall differentiation for nuclear markers is low ($\Theta_n = 0.043$) but significant (P-value < 0.001). For chloroplast markers, Θ_c is higher ($\Theta_c = 0.131$) and genetic differentiation was detected (P-value < 0.001) among the sampled populations. Tests for pairwise differentiation show that population SCr is the least differentiated for isozyme loci, while differentiation for the chloroplast genome is significant only for pairs

Table 3 Description of four cpDNA haplotypes found in eight *S. aucuparia* populations of Moffat Dale as a combination of polymorphism observed at three PCR-RFLP fragments. †For nomenclature and size of polymorphic fragments, see Raspé *et al.* (2000a)

CpDNA haplotype	Polymorphism observed for each primer/enzyme pair†		
	HK / *Hin*fI	DT / *Hin*fI	AS / *Hin*fI
A	1	1	1
B	1	1	2
C	1	2	2
D	2	1	2

including CMr (Table 6). These results suggest that overall genetic structure at isozyme loci is the result of differences among most of the populations sampled, whereas genetic structure for chloroplast markers is influenced mainly by population CMr.

Pollen vs. seed-mediated gene flow

Under the assumptions of the island model, levels of differentiation at biparentally inherited ($\Theta_n = 0.043$) and maternal ($\Theta_c = 0.131$) markers indicate that migration by pollen is of the same order of magnitude as migration by seed between the *S. aucuparia* remnants in Moffat Dale, $r = 1.36$. The difference increases slightly when taking into account the estimated value for the inbreeding coefficient ($f = 0.024$), $r = 1.44$.

Discussion

Habitat fragmentation is expected to have a detrimental impact on the genetic diversity of plant species, but empirical evidence suggests that population responses to fragmentation are idiosyncratic (Young *et al.* 1996). Remnants of *S. aucuparia* in the deforested landscape of the southern Uplands of Scotland have suffered severe long-standing fragmentation associated with anthropogenic land use. This provides an unique opportunity to assess empirically the extent to which fragmentation affects the amounts and patterns of gene diversity in a temperate tree species.

Our findings show that fragmented populations of *S. aucuparia* in Moffat Dale maintain high levels of isozyme diversity both within populations ($H_{EP} = 0.163$) and overall ($H_{ES} = 0.195$).

Table 4 CpDNA haplotype frequencies, haplotype diversity (H_E) and allelic richness (R_s) per population and overall for eight *S. aucuparia* remnants sampled in Moffat Dale. †CpDNA haplotype nomenclature refers to Table 3. ‡Based on a minimum sample of two haplotypes

Population	N	CpDNA haplotype frequency				H_E	R_s‡
		A†	B	C	D		
CDr	12	0.833	0.083	0.000	0.083	0.318	1.623
CMr	21	0.381	0.619	0.000	0.000	0.495	1.850
SCr	10	0.800	0.100	0.100	0.000	0.378	1.737
CTr	23	0.696	0.087	0.174	0.043	0.498	2.024
FBr	2	1.000	0.000	0.000	0.000	0.000	1.000
Kr	12	0.667	0.083	0.250	0.000	0.530	2.017
Wr	20	0.700	0.000	0.300	0.000	0.442	1.771
BMr	19	0.789	0.105	0.105	0.000	0.374	1.743
Overall	119	0.681	0.168	0.134	0.017	0.490	2.021

Table 5 Estimates of global inbreeding (f) and genetic differentiation (Θ) coefficients for each of six isozyme loci, for cpDNA markers and over all isozyme loci for *S. aucuparia* populations sampled in Moffat Dale. †HW: Hardy–Weinberg equilibrium. ‡PD: population differentiation, test not assuming random mating. NS: non-significant at 5% level. Estimates obtained and tests performed using FSTAT 2.9.3 (Goudet 2001)

Locus	Smallf (f)	Test HW†	Theta (Θ)	Test PD‡
prx-1	0.060	NS	0.077	P-value < 0.05
sdh-1	−0.146	NS	0.009	NS
pgm-1	0.190	P-value < 0.05	0.004	NS
pgm-2	0.028	NS	0.069	P-value < 0.001
pgm-3	−0.052	NS	0.004	NS
aat-1	0.107	NS	0.115	P-value < 0.01
overall	0.024	NS	0.043	P-value < 0.001
cpDNA	—	—	0.131	P-value < 0.001

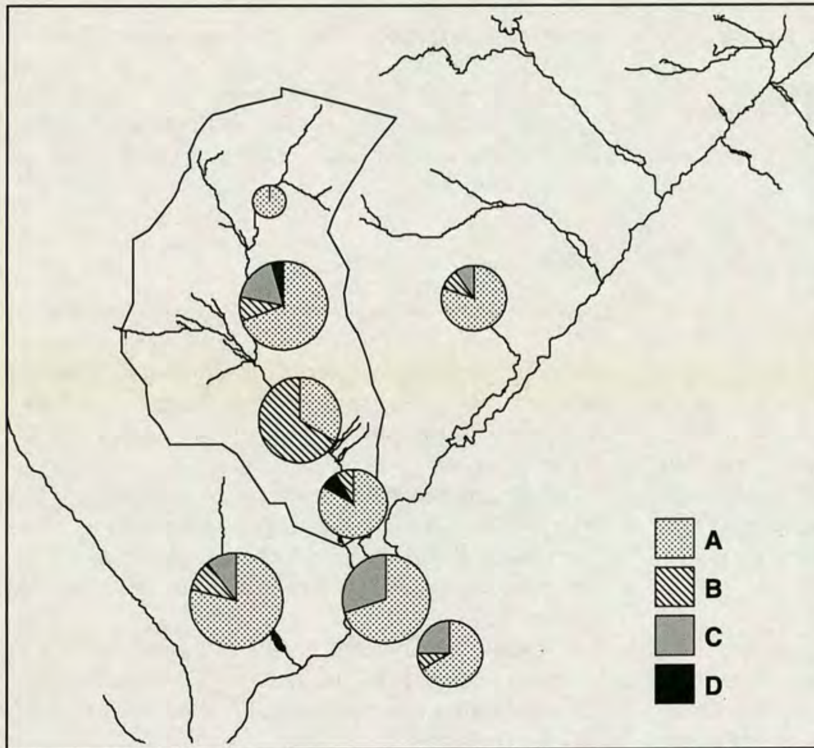


Fig. 2 Map of the distribution of four cpDNA haplotypes in sampled populations of *S. aucuparia* according to their location in Moffat Dale. For population names and characteristics, see Fig. 1 and Table 1. Haplotype nomenclature refers to Table 4. Pie chart diameter reflects relative sampling effort in each population.

Table 6 *P*-values for pairwise tests of differentiation between *S. aucuparia* remnants sampled in Moffat Dale. Above diagonal: test based on isozyme loci prx-1, pgm-2 and aat-1. Below diagonal: test based on cpDNA data. *P*-values obtained after 420 permutations in FSTAT 2.9.3 (Goudet 2001); *indicative adjusted nominal level (5%) for multiple comparisons is: 0.002381

Population	CDr	CMr	SCr	CTr	Kr	Wr	BMr
CDr		0.06905	0.17381	0.16905	0.02619	**0.00238***	**0.00238***
CMr	0.00476		0.51905	0.08095	0.00476	**0.00238***	0.18333
SCr	0.84048	0.00476		0.02857	0.10000	0.021430	0.70000
CTr	0.42143	**0.00238***	0.71905		**0.00238***	**0.00238***	**0.00238***
Kr	0.21905	0.00714	0.81667	0.88095		0.49048	0.01905
Wr	0.03333	**0.00238***	0.18810	0.20000	0.66667		**0.00238**
BMr	0.49524	**0.00238***	1.00000	0.81190	1.00000	0.20000	

Values are only marginally lower than those of Raspé & Jacquemart (1998) for non-fragmented European populations of *S. aucuparia*. They found that the species maintain high levels of isozyme diversity, $H_{EP} = 0.212$ within populations and $H_{ES} = 0.229$ at the species level, compared to species with similar ecological traits.

Estimates of diversity for the chloroplast genome can also be compared with those of non-fragmented populations in France and Belgium (Raspé *et al.* 2000a). To obtain comparable estimators of variation, we used data provided by Raspé *et al.* (2000a; Tables 4 and 5) to redefine cpDNA haplotype distribution in each population using only information provided by primer/enzyme combination in common with the present study (polymorphic fragments DT1, AS2 and HK3). At the wider scale of France and Belgium, *S. aucuparia* displays higher levels of gene diversity (H_E = 0.657) than in Moffat Dale (H_E = 0.490). This is also the case for Belgian populations separated by several kilometres (H_E = 0.671). Lower diversity is expected in Moffat Dale populations because of the smaller spatial scale and reduced size of individual populations. Nevertheless it is most remarkable that, on the basis of the same primer/enzyme combinations, just as many chloroplast haplotypes (i.e. four) are found within Moffat Dale as in much larger continental populations spread over a much greater area. We can conclude that despite severe habitat

fragmentation, *S. aucuparia* remnants of Moffat Dale maintain substantial levels of genetic diversity in both nuclear and chloroplast genomes.

S. aucuparia is a self-incompatible species (Raspé & Kohn 2002) which makes it an obligate outcrosser. Populations of Moffat Dale typically have a low inbreeding coefficient ($f = 0.024$) with no evidence for departure from Hardy–Weinberg equilibrium (P-value > 0.05). Autogamy can arise if a severe genetic bottleneck leads to a breakdown of self-incompatibility (Reinhartz & Les 1994). Furthermore, in a situation where few compatible mates are available, selection for self-compatibility may be strong (Stephenson *et al.* 2000). Self-compatibility has been reported previously in Swedish populations of *S. aucuparia* (Sperens 1996 cited in Raspé *et al.* 2000b). Based on the results of this study there is no evidence for biparental inbreeding or selfing in the fragmented populations in Moffat Dale. However, inbred progeny are very likely to suffer from inbreeding depression (Charlesworth & Charlesworth 1987) and may not have survived to maturity. If this were the case, inbred individuals and changes in the mating system of *S. aucuparia* would not be detected with the indirect approach taken here.

We found no strong support for the hypothesis that a genetic bottleneck has accompanied habitat fragmentation in the studied populations of *S. aucuparia* (one tailed-Wilcoxon sign rank test, P-value = 0.055). It is important to note that the test relies on the transient excess of heterozygotes that recently bottlenecked populations will exhibit at marker loci (Luikart *et al.* 1998). However, the tests for heterozygosity excess can detect bottlenecks only for a limited window of time after a bottleneck has occurred (0.2–2.5 times $2N_e$, discussed by Cornuet & Luikart 1996). Therefore, only recent population declines are detectable. It is possible, however, that populations have gone through a genetic bottleneck not recent enough to be detected. This interpretation would be supported by anecdotal evidence that habitat fragmentation in the area is ancient, because anthropogenic influence on the forest habitat can be dated back 6000 years (Wildwood Group of the Borders Forest Trust 2000).

Populations from Moffat Dale are slightly less differentiated ($\Theta_n = 0.043$) at isozyme markers than populations sampled within five different regions across Europe, ranging from Finland to the Pyrenees ($G_{STN} = 0.060$, Raspé & Jacquemart 1998). Furthermore, Raspé & Jacquemart (1998) found that, within regions, non-fragmented populations of *S. aucuparia* separated by at least 20 km displayed much less genetic differentiation ($G_{STN} \sim 0.010$) than populations in Moffat Dale. For the chloroplast genome, comparison with data reported by Raspé *et al.* (2000a; Tables 4 and 5) shows that genetic differentiation between populations sampled in France and Belgium ($\Theta_c = 0.258$) and between Belgian populations sampled within a single region ($\Theta_c = 0.218$) is higher than in Moffat Dale ($\Theta_c = 0.131$). *S. aucuparia* populations of Moffat Dale are differentiated more genetically at isozyme markers than would be expected in non-fragmented populations separated by less than a few kilometres and remarkably, within a single catchment, statistically significant differentiation is detected for both nuclear and cytoplasmic markers (G-test, P-value < 0.01). It is therefore likely that the genetic structure observed has been influenced by habitat fragmentation.

There is no evidence for predominance of pollen migration over seed migration for *S. aucuparia* in Moffat Dale, as is found normally in tree species. Inference of historical gene flow from F-statistics assumes an infinite island model of migration. Because natural populations deviate necessarily from the unrealistic assumptions of the model, estimates lack precision (Whitlock & McCauley 1999), especially for species with a long generation time, when disturbance has occurred and when pollen flow is not predominant (Ennos 1994) although they may be reasonable when the spatial scale is small (Whitlock & McCauley 1999). Furthermore, estimates should be accurate within a factor of two, even for population structure quite different from an island model (Slatkin & Barton 1989). Bearing these considerations in mind and that there is no evidence of a pattern of isolation by distance for *S. aucuparia* in Moffat Dale (results not shown), the reported estimate of pollen vs. seed-mediated migration is relevant for comparison with other species (Ennos 1994).

The ratio ($r = 1.36$) is one of the lowest ever recorded and comparable to that for the related species *S. torminalis* ($r = 2.21$, Oddou-Muratorio *et al.* 2001) and for a species with dust-like seeds *Epipactis helleborine* ($r = 1.43$, Squirrell *et al.* 2001). This ratio may be explained largely by *S. aucuparia* reproductive biology and ecology. The low cpDNA population structure observed suggests that seed dispersal between populations by birds is very effective. In Moffat Dale, it is conceivable that the interpopulation dispersal of seeds by birds is actually enhanced by habitat fragmentation. Birds feeding in small remnants are likely to move regularly between these isolated populations in the landscape, as seed resources at any one site become quickly exhausted.

However, the ratio does not provide a quantification of gene flow *per se*. The exceptionally low pollen flow to seed flow ratio may also be explained by reduced pollen flow among population fragments. Low pollen dispersal may arise if, as a result of increased isolation between remnants, pollinators tend to forage more locally and travel less between remnants. In the exposed situation of Moffat Dale long-distance insect flights between fragments may well be rare, as has been reported for tropical species (Powell & Powell 1987; Jennersten 1988) and this is the most probable cause of the observed genetic differentiation (however, see Dick *et al.* 2003).

S. aucuparia is often a pioneer that will be replaced later by late successional tree species (Raspé *et al.* 2000a). This extinction/recolonization dynamics is likely to be enhanced in the fragmented landscape of Moffat Dale, as small population size increases the risk of extinction. Pannel & Charlesworth (2000) have shown that as a result of such population turnover, genetic differentiation is increased if colonization occurs through a similar process to migration and particularly if colonization follows a propagule-pool model of dispersal more closely. Several empirical studies (reviewed in Giles & Goudet 1997) have confirmed these theoretical predictions and found higher F_{ST} values among younger populations than among older populations (Pannel & Charlesworth 2000). This is also likely to be the case for bird dispersed *S. aucuparia*.

Indeed, Raspé (unpublished data) described a metapopulation of *S. aucuparia* in Belgium where genetic diversity is lower in young populations and genetic differentiation is higher between recently founded populations for both nuclear and maternal markers. There are substantial peculiarities in the genetic structure of *S. aucuparia* remnants of Moffat Dale that suggest a similar dynamic. We could hypothesize that populations of Moffat Dale which are most diverse (Kr and CTr, Tables 2 and 4) and least differentiated (SCr, Table 5) are older than populations which are least diverse (Wr), depart from random mating (CDr) or are distinct in their cpDNA haplotype composition (CMr; Fig. 2 and Table 6). The latter three populations are situated at lower altitude where *S. aucuparia* is more likely to be outcompeted by late successional species and where population turnover may therefore be especially important. It is also conceivable that birds travel preferentially between remnants at higher altitude (CTr, SCr, BMr) or where the species occurs on its own (SCr, Kr) which would increase diversity and decrease differentiation between these populations. Further research is evidently needed to test these predictions.

The overall conclusions of this study are that despite being highly fragmented for very many generations, the populations of *S. aucuparia* in Moffat Dale have retained substantial genetic diversity for both isozyme and chloroplast DNA markers. In the extant population there is no evidence of inbred individuals. However, significant genetic differentiation is present among the population fragments implying that gene flow among them has been restricted. A very low value for the ratio of pollen to seed flow among fragments indicates that it is reduced pollen flow that has been principally responsible for the increased genetic differentiation of fragments.

From the point of view of forest restoration these results are encouraging. They indicate that within Moffat Dale as a whole there has been no dramatic loss of genetic variation. However, the data provide only information about the adult population of trees. In order to determine whether the seed would be suitable for including in a forest restoration scheme we need to measure contemporary mating patterns within and between fragments. Only with such information will it be possible to determine whether inbreeding is occurring within the population, and whether the genetic variation present in the adult population is effectively transmitted to the seed generation. These questions are the subject of a companion paper.

Acknowledgements

The study was carried out with financial support from The Scottish Forestry Trust (Registered Charity no. SC008465). CFEB was supported in part by an *Entente Cordiale* scholarship of the British Council and by a NERC PhD studentship. We would like to thank A. C. Newton and The Wildwood Group of the Borders Forest Trust. We also thank Tilhill Forestry and C. Moffat for field access in Moffat Dale, especially during the foot-and-mouth outbreak, J. D. Morman for assistance in the laboratory and M. Young for help with sample collection. S. Rendell and two anonymous referees provided useful comments on earlier versions of the manuscript.

References

Aguirre-Plantier E, Furnier GR, Eguiarte LE (2000) Low levels of genetic variation within and high levels of genetic differentiation among populations of species of *Abies* from southern Mexico and Guatemala. *American Journal of Botany*, **87**, 365–371.

Aldrich PR, Hamrick JL (1998) Reproductive dominance of pasture trees in a fragmented tropical forest mosaic. *Science*, **281**, 103–105.

Aldrich PR, Hamrick JL, Chavarriaga P, Kochert G (1998) Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*. *Molecular Ecology*, **7**, 933–944.

Andren H (1994) Effects of habitat fragmentation on birds and mammals in landscape with different proportion of suitable habitat — a review. *Oikos*, **71**, 355–366.

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (1998) GENETIX *logiciel sous Windows™ pour la génétique des populations*. Laboratoire Génome et Population, CNRS-UPR, 9060, Montpellier, France.

Benitez-Malvido J, Martinez-Ramos M (2003) Impact of forest fragmentation on understorey plant species richness in Amazonia. *Conservation Biology*, **17**, 389–400.

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

Cheliak WM, Pitel JA (1984) *Techniques for Starch Gel Electrophoresis of Enzymes from Forest Tree Species*. Petawa National Forest Institute, Canada.

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

Couvet D (2002) Deleterious effects of restricted gene flow in fragmented populations. *Conservation Biology*, **16**, 369–376.

Dayanandan S, Dole J, Bawa K, Kesseli R (1999) Population structure delineated with microsatellite markers in fragmented

populations of a tropical tree, *Carapa guianensis* (Meliaceae). *Molecular Ecology*, **8**, 1585–1592.

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

Dick CW (2001) Genetic rescue of remnant tropical trees by an alien pollinator. *Proceedings of the Royal Society London B*, **268**, 2391–2396.

Dick CW, Etchelecu G, Austerlitz F (2003) Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rain forest. *Molecular Ecology*, **12**, 753–764.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.

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

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

Ennos RA, Worrell R, Malcolm DC (1998) The genetic management of native species in Scotland. *Forestry*, **71**, 1–23.

Fuchs EJ, Lobo JA, Quesada M (2003) Effects of forest fragmentation and flowering phenology on the reproductive success and mating patterns of the tropical dry forest tree *Pachira quinata*. *Conservation Biology*, **17**, 149–157.

Giles BE, Goudet J (1997) A case study of genetic structure in a plant metapopulation. In: *Metapopulation Biology: Ecology, Genetics and Evolution* (eds Hanski I, Gilpin ME), pp. 429–454. Academic Press, San Diego, CA.

Goudet J (2001) FSTAT version 1.2: a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**, 485–486.

Goudet J (2001) *FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices, Version 2.9.3.* Available from: http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

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

Hamilton MB, Miller JR (2003) Comparing relative rates of pollen and seed gene flow in the island model using nuclear and organelle measures of population structure. *Genetics*, **162**, 1897–1909.

Jennersten O (1988) Pollination in *Dianthus deltoides* (Caryophyllaceae): effects of habitat fragmentation on visitation and seed set. *Conservation Biology*, **2**, 359–366.

Leberg PL (2002) Estimating allelic richness: Effects of sample size and bottlenecks. *Molecular Ecology*, **11**, 2445–2449.

Luikart G, Allendorf FW, Cornuet J-M, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238–247.

Mitchell A (1989) *Trees of Britain and Northern Europe*, 2nd edn. Collins, London.

Newton AC, Ashmole P (1998) How may native woodlands be restored to southern Scotland. *Scottish Forestry*, **52**, 168–171.

Oddou-Muratorio S, Petit RJ, Le Guerroué B, Guesnet D, Demesure B (2001) Pollen versus seed-mediated gene flow in a scattered forest tree species. *Evolution*, **55**, 1123–1135.

Pannell JR, Charlesworth B (2000) Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society of London B*, **355**, 1851–1864.

Petit RJ, Kremer A, Wagner DB (1993) Finite island model for organelle and nuclear genes in plants. *Heredity*, **71**, 630–641.

Powell AH, Powell GVN (1987) Population dynamics of male englossine bees in Amazonian forest fragments. *Biotropica*, **19**, 176–179.

Raspé O (2001) Inheritance of the chloroplast genome in *Sorbus aucuparia* L. (Rosaceae). *Journal of Heredity*, **92**, 507–509.

Raspé O, Findlay C, Jacquemart A-L (2000b) Biological flora of the British Isles, 214 *Sorbus aucuparia* L. *Journal of Ecology*, **88**, 910–930.

Raspé O, Jacquemart A-L (1998) Allozyme diversity and genetic structure of European populations of *Sorbus aucuparia* L. (Rosaceae: Maloideae). *Heredity*, **81**, 537–545.

Raspé O, Jacquemart A-L, de Sloover JR (1998) Isozymes in *Sorbus aucuparia* (Rosaceae: Maloideae): genetic analysis and evolutionary significance of zymograms. *International Journal of Plant Science*, **159**, 627–636.

Raspé O, Kohn JR (2002) S-allele diversity in *Sorbus aucuparia* and *Crataegus monogyna* (Rosaceae: Maloideae). *Heredity*, **88**, 458–465.

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

Raymond M, Rousset F (1995) GENEPOP: population genetics software for exact test and ecumenism. *Journal of Heredity*, **86**, 248–249.

Reinhartz JA, Les DH (1994) Bottleneck induced dissolution of self-incompatibility and breeding systems consequences in *Aster furcatus* (Asteraceae). *American Journal of Botany*, **81**, 446–455.

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

Robinson GR, Holt RD, Gaines MS *et al.* (1992) Diverse and contrasting effects of habitat fragmentation. *Science*, **257**, 524–526.

Slatkin M, Barton NH (1989) A comparison of three indirect methods for estimating average levels of gene flow. *Evolution*, **43**, 1349–1368.

Snow B, Snow D (1988) *Birds and Berries*. Poyser, London.

Soltis D, Haufler C, Darrow D, Gastony G (1983) Starch gel electrophoresis of ferns: a compilation of grinding buffers, gel and electrode buffers. *American Fern Journal*, **73**, 410–418.

Sperens U (1996) Is fruit and seed production in *Sorbus aucuparia* L. (Rosaceae) pollen limited? *Ecoscience*, **3**, 325–329.

Squirrell J, Hollingsworth PM, Bateman RM *et al.* (2001) Partitioning and diversity of nuclear and organelle markers in native and introduced populations of *Epipactis helleborine* (Orchidaceae). *American Journal of Botany*, **88**, 1409–1418.

Stephenson AG, Good SV, Vogler DW (2000) Interrelationships among inbreeding depression, plasticity in the self-incompatibility system, and the breeding system of *Campanula rapunculoides* L. (Campanulaceae). *Annals of Botany*, **85**, 211–219.

Weir BS, Cockerham C (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wendel J, Weeden N (1989) Visualization and interpretation of plant isozymes. In: *Isozymes in Plant Biology* (eds Soltis D, Soltis PS), pp. 5–45. Dioscorides Press, Portland.

White GM, Boshier DH, Powell W (1999) Genetic variation within a fragmented population of *Swietenia humilis* Zucc. *Molecular Ecology*, **8**, 1899–1909.

White GM, Boshier DH, Powell W (2002) Increased pollen flow counteracts fragmentation in a tropical dry forest: an example

from *Swietenia humilis* Zuccarini. *Proceedings of the National Academy of Science USA*, **99**, 2038–2042.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. *Heredity*, **82**, 117–125.

Wilcove DS, McLellan CH, Dobson AP (1986) Habitat fragmentation in the temperate zone. In: *Conservation Biology: the Science of Scarcity and Diversity* (ed. Soulé M), pp. 237–256. Sinauer, Sunderland, MA.

Wilcox BA, Murphy DD (1985) Conservation strategy: the effect of fragmentation on extinction. *American Naturalist*, **125**, 875–887.

Wildwood Group of the Borders Forest Trust (2000) Carrifran Wildwood Project. Native woodland restoration in the Southern Uplands of Scotland. *Management Plan* (eds Newton AC, Ashmole P). Borders Forest Trust, Jedburgh, Scotland.

Wright S (1951) The genetic structure of populations. *Annals of Eugenics*, **15**, 322–354.

Young AG, Boyle T, Brown AHD (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution*, **11**, 413–418.

Young AG, Merriam HG, Warwick SI (1993) The effects of forest fragmentation on genetic variation in *Acer saccharum* Marsh. (sugar maple) populations. *Heredity*, **71**, 277–289.

This study is part of Cecile Bacles' doctoral research on the consequences of habitat fragmentation on contemporary and historical gene flow for Scotland's native tree species and their implications for forest restoration. Andrew Lowe is particularly interested in the application of genetic markers to the management of genetic resources for a range of tropical and temperate plant species. Richard Ennos' interests are centred around population and ecological genetics, including gene flow in plant populations, host–pathogen interactions, hybridization and introgression.