THE ORIGIN, EVOLUTION AND CONSERVATION OF THE ARRAN *SORBUS* MICROSPECIES.

ASHLEY ROBERTSON

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

Institute of Ecology and Resource management, University of Edinburgh

July 2000

I would like to dedicate this thesis to my mother, Yvonne Patricia Robertson.

Acknowledgements

I would like to thank my supervisors Dr. Adrian Newton, Dr Richard Ennos, Dr Chris Sydes and Dr Juliet Vickery for their time, advice and encouragement. I would especially like to thank Dr Richard Ennos for his advice on the genetic and evolutionary aspects of the project and for his time and patience during the writing of this thesis.

Xin- Sheng Hu, John Morman, Amanda Gillies, Anthony Langdon, Nichola, Jenny, Joan, Aileen, Graham French, Sarah Beckesey, Tonny, Sylvia Heredia, Sarah Rendell, Cecile, Rhizana and Graeme Pigott thank you all for your friendship and support.

A special thanks to Mick Carr for his friendship, advice on cloning and the odd beer, to Sarah Beckesey and Adèle for proof reading, to John Morman for his help and advice throughout, to Nichola Preston for her help and advice with the DNA sequencing, to Tony Church and Philip Lusby for their help, enthusiasm and advice on the Arran whitebeams and to Vicky and Tam for making my trips to Arran very enjoyable and comfortable.

To Adèle and my family and friends in Somerset a very big thank you.

Declaration

I declare that this thesis has been composed solely by myself. The work and results reported within are my own, unless otherwise acknowledged.

ASHLEY ROBERTSON

July 2000, Edinburgh

Abstract

In theory, populations of agamospermous plants are genetically invariant and consequently hold little evolutionary potential. To test this hypothesis, molecular markers were used to examine evolutionary processes in the *Sorbus* microspecies *S. pseudofennica* and *S. arranensis*, which are endemic to the Isle of Arran, Scotland.

The proposed hybrid origin of these microspecies was investigated using a variety of molecular markers. Two isozyme systems (AAT and 6PGD), a nuclear DNA marker and a chloroplast DNA marker all provided markers that were specific to the putative parental species, *S. aucuparia* and *S. rupicola*. These markers could not separate *S. rupicola* from *S. aria* (a potential parent) or fully distinguish between *S. arranensis* and *S. pseudofennica*. Individuals of both *S. arranensis* and *S. pseudofennica* were screened and results were compatible with these microspecies being derived from a cross between *S. aucuparia* and *S. rupicola/S. aria*. DNA sequences of the chloroplast trnL intron and the trnL-trnF intergenic spacer for the Arran *Sorbus* microspecies were identical to those of *S. aucuparia*. In contrast, nucleotide substitutions were evident between the *Sorbus* microspecies and *S. rupicola* at three bases.

Nuclear DNA markers (five microsatellites and a Rubisco intron primer pair) revealed variation among individuals within *S. arranensis* and *S. pseudofennica* microspecies. Only three *S. arranensis* clones were detected from a total of 179 individuals. In contrast eight clones were detected from a total of 140 *S. pseudofennica* individuals. Six families of *S. arranensis* and *S. pseudofennica* were screened with three nuclear DNA markers. No variation in banding pattern was found among any of the *S. arranensis* families. In contrast, segregation of molecular makers was found in three out of the six *S. pseudofennica* families.

During the field study, three individual *Sorbus* trees with a phenotype unlike any previously recorded on Arran were observed at two locations. The gross leaf morphology of these plants appeared to be intermediate between *S. aucuparia* and *S.*

pseudofennica and resembled that of *S. meinichii sensu lato*. It was not possible to determine the precise origin of the novel Arran *Sorbus* from the results of the molecular marker assays. However, their genomes had unique banding patterns and contained markers specific to both *S. aucuparia* and *S. aria sensu lato*. The presence of this novel Arran *Sorbus* hybrid represents previously unrecorded genetic exchanges within the Arran *Sorbus* complex and these genetic exchanges could possibly lead to the formation of new apomictic taxa.

The implications of these results for the understanding of evolutionary processes within agamospermous species, and for the conservation of these threatened taxa are discussed.

Contents

List of Tables

List of Figures

Chapter One

Introduction: Evolutionary potential of agamospermous taxa

1.1 Introduction

Agamospermous taxa, asexual taxa which produce seeds without sexual fusion of male and female gametes, are capable of reproducing favourable heterozygous gene combinations which have arisen by hybridisation. Faithful reproduction of these advantageous gene combinations occurs by avoiding mechanisms that will lead to their break up, such as meiosis and genetic recombination.

Formally it was believed that agamospermous populations were genetically invariant a view that contributed to the hypothesis that agamospermous taxa held little evolutionary potential. Even though asexual reproduction of permanent hybrids could be advantageous in certain environments, and lead to expansion of the hybrid group, it was considered inevitable that the hybrid group would enter a period of decline leading to possible extinction (Stebbins 1950). The cause for this decline was proposed to be that of a slow response to changing environments, resulting from the failure of asexual organisms to exchange and compile favourable mutations that can lead to adaptive evolution. The theory was well supported, as it was compatible with the then prevalent view that sexual reproduction did have the required mechanisms which could exchange and compile favourable mutations and greatly accelerate the rate of adaptive evolution (Fisher 1930; Muller 1932).

More recently, analyses using starch-gel electrophoresis have documented levels of genetic diversity within asexual populations that are in fact comparable to those found in sexually reproducing populations (Ellstrand and Roose 1987; Hamrick and Godt 1990). These discoveries have prompted a reassessment of the evolutionary role of asexual reproduction (Mogie 1992) and inspired general interest in the evolution and ecology of asexual plant populations.

1.2 Evolution

Definitions of biological or organic evolution vary depending from which discipline researchers approach the subject. From a genetic perspective evolution consists of

alterations in gene frequencies within and between populations; from an ecological perspective the essence of evolution is the production of adaptive diversity. Futuyma (1986) proposed a more balanced view: 'Biological evolution is change in the properties of populations of organisms that transcends the lifetime of a single individual'.

Two fundamental aspects of biological inheritance lie at the heart of evolutionary processes. First, the ability of organisms to replicate themselves in a constant fashion and second, the fact that this replication process is not always exact. The evolutionary significance of these processes is that the information required to survive and reproduce (ranging from biochemical to behavioural) which has been amassed and stored over time by previous generations, will be passed on to new generations. The small errors (mutations) made in these replication processes will also be passed on to new generations, producing inherited biological changes.

The majority of these replication errors are either neutral in their effect or harmful only when they occur in a homozygous state. These neutral and deleterious mutations can therefore accumulate within evolutionary lineages without any immediate unfavourable effects. Any variation created by these mutations, among individuals, will in effect create a pool of genetic variability within a population. The significance of genetic variability is that it fuels evolution by allowing organisms to differ from each other and from their ancestors and provides raw material upon which natural selection can act.

During the first part of the 20th century Hardy (1908) and Weinberg (1908) emphasised that populations rather than individuals were an important evolutionary focus. Attention had to be paid to population gene frequencies (the proportion of a particular allele among all alleles at a gene locus) rather than only to whether a gene was absent or present. It became clear that the collection of gametes that a population contributes to the next generation could be considered as a giant gene pool from which offspring draw their various genotypic combinations at random. In the absence

of factors that could change gene frequencies, these frequencies tend to be conserved.

1.2.1 Natural selection

The wide variety of mechanisms that affect the survival and reproductive success of a genotype are known collectively as selection, and the extent to which a genotype contributes to the offspring of the next generation is commonly known as its fitness or adaptive value. It is possible to describe selection as a composite of the forces that limit the reproductive success of a genotype and fitness as the comparative ability of a genotype to withstand selection.

If evolution is defined as hereditary changes over time, selection although important is not the only process that can cause such changes. Other evolutionary mechanisms such as mutation, migration and random genetic drift can also affect gene frequencies. In general however, populations do tend to change genetically in directions that improve fitness for their environment. Endler (1986) lists more than 160 cases in natural populations where selection has been demonstrated. Given the persistent and recurrent genetic variability found in factors such as differential mortality, differential fecundity and differential mating success, gene frequency change caused by natural selection must be a constant feature of most if not all populations. This seems especially so since populations cannot remain immune to the repeated environmental changes that affect these components of fitness.

How then do populations change genetically in directions that improve fitness for their environment? As mentioned above natural selection acts on fitness differences that arise because different genotypes have different probabilities of survival and/or reproduction. If there is an optimum value of an adaptive trait that improves the population's fitness for its environment and if there is additive genetic variance for that adaptive trait then the mean value of the trait will move towards the optimum until the two coincide (Fisher 1930). The increase in mean fitness for a population because of natural selection is proportional to the additive genetic variance in fitness (Fisher 1930).

One consequence of Fisher's theorem is that it would be expected that populations subjected to selection for long periods of time would have little remaining variability for genes that affect fitness (Ennos 1983). Selection would tend to diminish such variability.

How then do populations maintain a balance between moving towards an optimum for an adaptive trait and maintain sufficient additive genetic variation to move towards future optimum for adaptive traits ? It seems intuitive that as sexual reproduction promotes genetic variability (particularly between parent and offspring) it will facilitate natural selection and the potential for a population to evolve and adapt to new and changing environments. Conversely, as asexual reproduction tends to conserve genotypes between parent and offspring, the potential of an asexually reproducing population to adapt will not be as great. It may be possible that sexually reproducing populations will be able to adapt to changing environments at a higher rate than asexual populations, and possibly of greater importance is the fact that whilst doing so they can maintain a sufficiently large pool of additive genetic variance required for future environmental changes. The following section considers how sexual reproduction may facilitate natural selection.

1.2.2 Sexual reproduction

The essential feature of the sexual process is that genetic material from different ancestors is brought together in a single descendant and there is subsequent genetic recombination between the two ancestral genomes. In most eukaryotes there is an alternation between haploid and diploid states. Two haploid cells (gametes) fuse to form a diploid zygote in the process of fertilisation and the diploid zygote produces gametes by meiosis. It is during the production of gametes that recombination takes place and these recombined gametes will consist of genes from both ancestors.

Fisher (1930) and Muller (1932) pointed out that favourable mutations that arise in different individuals could only be brought together by recombination. An asexual population must fix favourable mutations one by one, whereas a sexual population can establish them more rapidly by bringing them together. However, the argument

does not apply to small populations in which each favourable mutation that occurs will be fixed before the next one occurs (Maynard Smith 1989).

Williams (1975) and Maynard Smith (1978) proposed that if genotypes vary in their use of limiting resources, segregation and recombination can increase the number of surviving offspring by reducing competition among family members or by increasing the families probability of producing a successful offspring. However this cannot be a universal effect of sex as competition between siblings is confined to certain organisms (Maynard Smith 1989).

In a population without recombination, there is a tendency for slightly deleterious mutations to accumulate (Muller 1964; Felenstein 1974). If the frequency of the mutational class containing the lowest number of mutations is sufficiently small, it will be lost from the population after a finite number of generations. If there is no opportunity for genetic recombination, this class cannot be reconstituted and will be replaced by the class with one more mutation. This class is now vulnerable to stochastic loss in the same way and there is a repetitive process of loss of successive classes with minimum numbers of mutations (termed Muller's rachet) leading to a decline in the fitness of a non-recombining or asexual population (Muller 1964). Muller's argument applies not only to asexual organisms, but also to any region of DNA that never recombines, for example, the Y chromosome. It also applies to populations with 100% self-fertilisation (Maynard Smith 1989).

In theory, negative associations among alleles at different loci, resulting in favourable genes being found in different individuals more often than is expected by chance, will increase additive genetic variance and facilitate natural selection (Feldman et al. 1980; Kondrashov 1982). With this form of interaction the effect on fitness of a deleterious allele increases, the larger the number of deleterious alleles at other loci; conversely, as more loci acquire favourable alleles, fitness increases by a diminishing factor (Feldman et al. 1980; Kondrashov 1982). The immediate effect of recombination is to reduce mean fitness because it breaks up the negative gene combinations favoured by selection. This is counterbalanced by the increased

variance in fitness, which in the longer term increases the population mean fitness by speeding up the response to selection. Barton (1995) proposes that it is the balance between these opposing forces that determines whether recombination will increase the efficiency of natural selection.

It may be possible that these opposing forces of recombination (reducing mean fitness and increasing variance in fitness) could provide a mechanism allowing populations to maintain a balance between moving towards an optimum for an adaptive trait and still maintain sufficient additive genetic variation to move towards a future optimum.

Without recombination, asexual reproduction is likely therefore to be at a selective disadvantage when compared to sexual reproduction, which may help to explain the limited distribution of asexually reproducing taxa. However, there is thought to be an advantage of asexual reproduction that may offset this selective disadvantage. It has been termed the cost of meiosis (Williams 1975). In the absence of recombination and segregation maternal energy is not wasted on the aftercare of unfit zygotes for all offspring are as fit as the mother. Also the genetic contribution of an asexual mother to her offspring is 100 % as opposed to 50 % for a sexual outbreeder.

1.2.3 Asexual reproduction

The occurrence of parthenogenic species among taxa of vertebrates, invertebrates and flowering plants indicate that asexuality is mechanistically possible and has evolved independently many times. However, it is unusual to find a family or higher taxon consisting wholly of parthenogens. The bdelloid rotiferes, an order in which males have never been reported, are an exception (Maynard Smith 1989).

The restricted distribution of asexuality suggests that asexual taxa may be successful in the short term, but that (with a few exceptions such as the bdelloids) they have been unable to survive for a long time, or to undergo adaptive radiation (Maynard Smith 1978). This is consistent with the view that a successful modification of female germ cell production to allow parthenogenesis is difficult to accomplish and

that selective disadvantages to asexual populations may cause their more rapid extinction (Maynard Smith 1978).

There are three major asexual processes: parthenogenesis, the development of a new individual from an unfertilised egg; adventitious embryony, the development of a new individual from a single somatic cell and vegetative reproduction, the development of a new individual from group of somatic cells.

There are many forms and variations of these processes however, the focus of the remainder of this review will be asexual reproduction in seed plants, in particular agamospermy, the formation of seeds without sexual reproduction.

Asexual reproduction in seed plants, apomixis, can be divided into two main classes, vegetative reproduction and agamospermy. Agamospermy can be defined as the production of seeds without the sexual fusion of male and female gametes. In agamospermy, a full reductional meiosis is usually absent and thus the chromosomes do not segregate. Agamospermous embryos can originate from sporophytic tissue by adventitious embryony (sporophytic apomixis) in which somatic embryos are formed without a gametophytic stage (because the sexual megagametophytes are also functional, both sexual and asexual adventitous embryos can co-occur in a single seed) or from unreduced megagametophytes (gametophytic apomixis). Gametophytic apomixis results in the megagametophyte or embryo-sac producing a parthenogenetic embryo. Gametophytic apomixis is achieved either by Apospory, where the reduced megaspore degenerates and is replaced by a somatic cell (usually from the nucellus), or by Diplospory, where the reductional meiosis is replaced by a restitutional meiosis or by mitosis. A comprehensive review of these mechanisms can be found in Richards (1997).

In all cases the agamospermous embryo bypasses both meiosis and egg cell fertilisation producing offspring with a genome derived entirely from its mother. However, true obligate agamospermy, in which all possibility of sexuality has been lost, is a rare phenomenon and some sexuality continues to be a feature of most

agamospermous plants (Nogler 1984a; Asker and Jerling 1992). Before addressing the evolutionary potential of agamospermous populations, the following sections will consider their origins, distribution and genetic variation.

1.3 Distribution and origins of agamospermous taxa

1.3.1 Distribution of Agamospermy

Agamospermy is a common mode of reproduction in at least 300 species belonging to 35 plant families (Asker and Jerling 1992). The wide taxonomic distribution and the multiplicity of mechanisms underlying asexual seed formation suggests that this mode of reproduction has arisen independently on many different occasions. Although agamospory is widespread in Pteridophytes (Manton 1950) agamospermy is limited to angiosperms among seed plants, as it does not occur in the gymnosperms. Even though agamospermy is recorded in 35 families (approximately 15%) at least 75% of these agamospermous taxa belong to only three families, *Compositae*, *Gramineae*, and *Rosaceae* (Richards 1997).

Stebbins (1950) proposed that the limited distribution and rarity of agamospermous taxa was a measure of its long-term failure. However, there is little doubt that agamospermy is currently very successful in genera such as *Taraxacum*, *Hieracium* and *Rubus* in which some agamospecies may be of a great age dating back 10,000 years (Wendelbo 1959) or even to the early Pleistocene (Richards 1973).

1.3.2 Origins of Agamospermy

Agamospermy is a complex developmental trait consisting of different elements that may be encoded by multiple genes (Mogie 1992). Genes that control successful agamospermy when they occur in combination can be found singly as aberrations in sexual populations (Asker 1979, 1980; Asker and Jerling 1992; Mogie 1992), for example non-disjunction at meiosis I or II, in *Zea* L. and *Datura* L. (Rhoades 1956), total inhibition of meiosis in *Zea* (Palmer 1971) and haploid parthenogenesis in *Zea* (Chase 1969).

Similarly, agamospermous individuals that carry only one agamospermous gene are likely to be unsuccessful. For example, an individual that has a gene to halt meiosis but lacks a gene for the autonomous development of an egg (or somatic cell) into an embryo, without fertilisation, will produce unbalanced high number polyploids. Thus, the ploidy level increases with the fertilisation of each new generation (Stebbins 1950). Conversely, an individual that has a gene that permits the production of embryos without fertilisation but lacks a gene to halt meiosis will produce weak haploid offspring. However, when two of these particular mutants come together in a single individual, agamospermy may be successfully achieved.

Simultaneous mutations at two or more loci could be expected to be rare in practice. Individual mutants are unlikely to accumulate in populations since each would be at selective disadvantage without the other complementary member genes controlling the asexual process. On this basis a lack of suitable mutants would appear to be the major factor hampering the establishment of agamospermy in plant populations (Marshall and Brown 1981). The close association between agamospermy, polyploidy and interspecific hybridisation, observed in practice (Stebbins 1950; Grant 1971), can be explained under this hypothesis. As mentioned above, any apomictic cycle must involve at least two genetic changes from the normal sexual cycle, one that provides a substitute for meiosis, and one that provides a substitute for fertilisation. Hybrids are often partially or completely sterile (Reiseberg and Carney 1998) removing the need for one of the simultaneous mutations required for the development of agamospermy (Marshall and Brown 1981).

Results from Grimanelli (1998) support this hypothesis. Molecular markers linked to diplospory were used to analyse various generations of maize x *Tripsacum* L. hybrids. The results suggest that the gene or genes controlling apomixis in *Tripsacum,* were linked to a segregation disorder-type system promoting the elimination of the apomixis alleles when transmitted through haploid gametes. However, in a polyploid system deleterious recessives are masked at the gametophyte stage and both they and any linked apomixis genes are preserved.

Other more empirical evidence to support the close association of agamospermy with polyploidy and hybridisation is that gametophytic agamosperms are widely credited with possessing elevated levels of heterozygosity when compared to related diploids. In the case of allopolyploids this is due to fixed heterozygosity and in autopolyploids to polysomic inheritance (Soltis and Soltis 1993). Examples of allopolyploid agamosperms are *Erigeron compositus* (Noyes and Soltis 1996), *Antennaria rosea* Greene (Bayer 1989), and *Taraxacum albidum* Dahlst. (Hughes and Richards 1988). Examples of autopolyploid agamosperms are *Antennaria friesiana* ssp. *friesiana* (Trautv.) Ekman (Bayer 1991), *Eupatorium altissimum* L. (Yahra et al 1991) and *Panicum maximum* Jacq. (Assienan and Noiroy 1995).

1.3.3 Genetic control of agamospermy

Various models have been proposed to explain the regulation of apomixis. Most involve the existence of an apomictic allele or alleles that would be present and expressed in apomictic plants and absent or silent in sexual plants. Under this assumption apomixis would have arisen as the result of mutations at one or several loci playing a role in the reproductive process. The number of loci involved is still being strongly debated. Data obtained in, *Ranunculus* (Nogler 1984b), and *Brachiaria* (do Valle and Savidan 1996) demonstrate that a single dominant gene controls the dysfunction or absence of meiosis (apomeiosis).

Mogie (1992) and Leblanc and Savidan (1994) have hypothesised that no specific genes are required for the parthenogenic development of the unreduced gametes. The initiation of parthenogenesis occurs as a consequence of the shortening of megasporogenesis, resulting from dysfunctional meiosis. As a result the process of apomictic reproduction relies on the control of apomeiosis.

Van Dijk et al. (1999) concluded that the elements of apomixis, diplospory and parthenogenesis could be uncoupled. In the case of the facultative apomict, *Taraxacum* section *Palustria* G.H. Weber ex Wiggers, the breakdown of apomixis into its elements occurred at the ovule level within a single flower head. In the case of obligate apomict *Taraxacum* section *Ruderalia* G.H. Weber ex Wiggers the

breakdown occurred at the plant level between hybrid progeny within crosses. The former can be explained by recombination during restitutional megasporogenesis and the latter by recombination during reductional microsporogenesis. Both cases suggest a multigenic control of apomixis with several loci being involved in the genetic control of apomixis in *Taraxacum*.

A second category of hypotheses concerning the genetic control of apomixis postulates that no specific mutations are required to induce apomixis. Carman (1997) proposed that apomixis arose spontaneously as a result of asynchronous signals in the female reproductive pathway disturbing sexual reproduction following interspecific hybridisation. In this asynchrony hypothesis, allopolyploidy is therefore invoked as a causal factor.

1.4 Variation within agamospermous taxa

1.4.1 Variation within and between populations

Empirical studies, using starch-gel electrophoresis, showed that genetic diversity within some asexual populations was comparable to that found in sexually reproducing populations (Ellstrand and Roose 1987; Hamrick and Godt 1990). Ellstrand and Roose (1987) compiled and analysed data on the population genetic structure of 27 clonal plant taxa, covering a broad range of reproductive mechanisms. Of the 27 surveys of genotypic variation all but two revealed multiple clones both within and among populations.

Noyes and Soltis (1996) used allozyme assays to assess the genetic structure of the *Erigeron compositus* Pursh. agamospermous species complex. Fifteen enzyme loci were used to identify 24 unique multilocus genotypes in seven populations. The average number of genotypes detected per population (G) and the proportion of distinct genotypes (G/N) (where G is the number of distinct multilocus genotypes and N is the number of individuals sampled per population) for *E. compostius* was 3.4 and 0.1 respectively. These values are lower than corresponding mean values of

16.1 and 0.17 reported by Ellstrand and Roose (1987) for their study on 27 clonal taxa, but compare favourably to measures of 5.1 and 0.06 for the six taxa known to reproduce via agamospermy. Multiple genotypes were detected in five out of the seven populations of presumed agamic *E. compositus*.

Gornall (1999) compiled and analysed data from four major agamospermous taxa *Rubus*, *Hieracium*, *Taraxacum* and *Sorbus*. The data appear to show a similar pattern. In *Rubus* and *Sorbus* each microspecies contains an average of 2.1 and 1.4 genotypes respectively; in *Hieracium* the average is 1.8 and in *Taraxacum* 5.0. The high figure for *Taraxacum* is inflated by the extreme value for *T. vindobonense* Van Soest in which the unusually high number of 64 genotypes was detected.

1.4.2 Genetic structure in asexual versus sexual populations

Four studies that contrasted genetic variation in closely related sexual and agamospermous plant populations found no apparent trends (Hughes and Richards 1988; Bayer 1989; Yahara 1990; Noyes and Soltis 1996). Yahara (1990) obtained allozyme data for five triploid agamospermous and eight outcrossing diploid populations of *Eupatorium altissimum*, in which no significant differences in H or % P were observed (where H = observed heterozygosity, % P = percentage polymorphic loci). Bayer (1989) obtained significantly greater values of %P, H and A in agamospermous populations of tetraploid *Antennaria rosea* (where A = the average number of alleles per locus) in contrast with its sexual diploid progenitor species. Hughes and Richards (1988) Noyes and Soltis (1996) reported double the value of H for agamospermous *Taraxacum* and *Erigeron compositus* populations respectively and similar levels of % P and A when compared to their sexual relatives.

Agamospermous *Eupatorium altissimum* is thought to be derived from non-divergent sexual populations, characterised by very low levels of heterozygosity (Yahara 1990). Consequently, agamic *Eupatorium altissimum* populations do not differ genetically from related sexual populations. *Erigeron compositus* and *Antennaria rosea* are thought to be derived from divergent sexual progenitors. For these taxa the polyploid condition allows for additivity of the divergent ancestral genomes. Thus,

agamic populations of *Eigeron compositus* and *Antennaria rosea* exhibit higher levels of heterozygosity than sexual progenitors.

Values of H (observed heterozygosity), % P (percentage polymorphic loci) and A (the average number of alleles per locus) are informative measures of genetic diversity, and agamospermous populations may have higher or lower values of H, % P and A, when compared with sexual relatives. However, due to sex and recombination there is likely to be large differences in the way genetic variation is arranged into multilocus genotypes. The number of genotypes is likely to be small in apomictic taxa and large in sexual taxa. For example Hughes and Richards (1988) reported double the value of H for agamospermous *Taraxacum* populations and similar levels of % P and A when compared to their sexual relatives. In contrast and possibly of greater significance was that genotypic diversity was very high in the sexual populations and very low in the agamospermous population. As natural selection acts on additive genetic variance between individuals, genotype diversity (as opposed to gene diversity) within populations is, in evolutionary terms, an important measurement.

1.4.3 Origins of clonal diversity

Mutation and inheritance from sexual progenitors are considered to be the principal sources of genetic variation in agamic taxa (Grant 1981). Allozyme patterns for the agamospermous *Erigeron compositus* are very similar to those of the sexual diploid *E. compositus*, suggesting that most of the genetic variation found in *E. compositus* agamospermous populations is ultimately derived from the diploid progenitor populations rather than by mutation (Noyes and Soltis 1996).

This hypothesis is supported by allozyme evidence showing that all but nine of the 42 alleles observed in agamic populations, were also found in sexual diploid populations of *E. compositus*, and that all alleles common in agamospermous populations were also common in sexual populations. The nine unique alleles detected in the agamospermous *E. compositus* populations could have arisen by mutation, or occasional sexual reproduction either among facultative agamic clones

or between agamic clones and sexual plants (e.g. Gustaffson 1946/7; Grant 1981; Asker and Jerling 1992), or that corresponding alleles may not have been detected in the diploid populations due to inadequate sampling.

1.5 Evolutionary potential of agamospermy and agamospermous populations

One of the earliest theories, first proposed by Darlington (1939) and promoted by (Stebbins 1950), is that agamospermy offers no real selective advantage over normal sexual reproduction. It was argued that it restricts recombination and hence, the evolutionary potential in species in which it has become established. From this viewpoint the only, or at least principle, advantage of apomixis is that it restores the fertility of individuals that would otherwise be sexually sterile as a consequence of polyploidy and hybridisation. This hypothesis is likely to be true, particularly with respect to obligate agamosperms. However, to judge its significance it is important to first consider what the alternative consequences of interspecific hybridisation are.

1.5.1 Interspecific hybridisation

In addition to agamospermy, interspecific hybridisation can lead to: the introgressive origin of new intrapecific taxa, in the form of varieties, ecotypes and subspecies (Abbott et al. 1992) the origin of new fertile homoploid derivative species (Gallez and Gottlieb 1982; Reiseberg et al.1990) and hybrid breakdown which manifests itself in the form of hybrid inviability or sterility (Dobzhansky 1937; Wu and Palopoli 1994; Li et al.1997).

1.5.2 Introgression

In his monograph on introgressive hybridisation, Anderson (1949) reasoned that hybrids might serve as a bridge for the exchange of genetic adaptations between otherwise isolated species. Given the large number of plant and animal species that at least occasionally hybridise (Elstrand 1996; Dowling and Secor 1997), this process could under Anderson's model, potentially contribute to the adaptive evolution in many plant and animal groups.

However, there are theoretical difficulties associated with the transfer of adaptations across species barriers in nature. To move a gene across a reproductive barrier, the allele must recombine into a new genetic background before it is eliminated by selection against the alleles with which it is initially associated (Barton and Hewitt 1985). As a result successful introgression will depend in part on the genetic architecture of the reproductive barriers. If many genes contribute to hybrid unfitness, then much of the genome may be resistant to introgression because of linkage (Whitmore and Schaal 1991; Rieseberg and Wendell 1993) particularly if recombination rates are low. This problem may be exaggerated if the genes interact in an epistatic nature or are coadpated (Harlan 1936; Carson 1975). However, if reproductive barriers are under simple genetic control, then most of the genome should be permeable to introgression. Only those traits tightly linked to sterility or inviability genes will be difficult to introgress.

Studies on North American sunflowers (*Helianthus* L. spp.) (Rieseberg et al. 1990; Rieseberg et al. 1999b; Rieseberg et al. 2000) and British ragworts and groundsels (*Senecio* L. spp.) (Abbott et al.1992; Ashton and Abbott 1992) provide evidence for the introgressive nature of plant hybridisation.

1.5.3 Homopoloid hybrid derivatives

The most widely accepted model for homoploid hybrid speciation is the recombinational model (Grant 1958). The sorting of parental chromosomal and genic sterility factors in hybrid populations can, under appropriate conditions, lead to the formation of hybrid species that are homozygous for some combination of parental sterility factors. The new hybrid lineage would be fertile, but at least partially inter-sterile with both parental species.

Factors that appear to play a critical role in recombinational speciation include strong natural selection for the most fertile or viable hybrid segregates (Templeton 1981; McCarthy et al. 1995); rapid chromosome evolution (Rieseberg et al. 1995) and the availability of habitat suitable for the establishment of hybrid neospecies (Templeton 1981; Arnold 1997).

1.5.4 Hybrid breakdown

Natural selection acts to retain favourably interacting gene combinations (Wright 1931) and evolution may lead to the production of what Dobzhansky (1970) termed coadapted gene complexes. In contrast, Fisher (1930) argued that natural selection acts primarily on single genes rather than on gene complexes. In Fisher's model, natural selection favours alleles that elevate fitness, on average, across all possible genetic backgrounds within a lineage. For both Wright's and Fisher's models adaptive evolution is assumed to proceed independently within distinct evolutionary lineages. Therefore, crosses between divergent lineages may result in unfavourable gene combinations and as a consequence hybrid offspring are generally expected to be less viable and/or less fertile, relative to their parental taxa (Mayr 1963).

Hybrid breakdown, which manifests itself in the form of hybrid inviability or sterility, was first recognised by Dobzhansky (1937). In his model a gene from one species interacts negatively with a gene from another species, causing some degree of inviability or sterility. Wu and Palopoli (1994) argued that the most plausible interpretation of this model was that the hybrid sterility/inviability gene acts like a mutation whose deleterious effects are suppressed by another gene in the source species. However, when placed in the genetic background of another species the deleterious effects of the sterility/inviability gene are expressed. Data to support this model come from many studies of plant hybrids in which one or two genes appear to have major effects on hybrid sterility or inviability (e.g. Sano and Kita 1978; Wan et al. 1996).

A model, proposed by Wu and Palopoli (1994), for the evolution of hybrid inviability/sterility was that a much larger number of diverging loci interact negatively in a hybrid background and that these weak interactions act cumulatively to cause inviability or sterility. Li et al. (1997) provided evidence that hybrid breakdown in rice fits this polygenic model and resulted from the break-up of coadapted gene complexes by recombination.

1.5.5 Evolutionary potential of agamospermy

A precondition for introgressive hybridisation and the formation of new homoploid species is that the hybrid must exhibit some fertility, either as an ability to backcross to a parent or a capacity to reproduce via sib-crossing or inbreeding. As the majority of interspecific hybrids are highly or completely sterile e.g. British hybrids (Stace 1991) it would appear that introgression and recombinational speciation are rare phenomena and the most likely out come of interspecific hybridisation is hybrid sterility. If agamospermy can rescue a hybrid from sterility, then the evolutionary potential of the hybrid is significantly increased. The function of agamospermy in evolutionary terms can therefore only be advantageous, for at the very least the hybrid is reproducing.

1.5.6 Evolutionary potential of agamospermous populations

Agamosperms have the potential to fix a relatively fit genotype as a result of interspecific hybridisation. First generation hybrids, particularly between geographic races or closely related species, tend to exceed their parents in vegetative vigour or robustness (Grant 1975). This phenomenon of hybrid vigour (heterosis) is often used to maximise yields in crop plants and partly explains the success of allopolyploid species and many clonal hybrid lineages (Huskins 1931; Grootjans et al. 1987).

Although heterosis is a likely contributor to the evolutionary success of hybrids, the genetic basis is still poorly understood. Possible models for the evolution of heterosis are listed in Mitchell-Olds (1995) and include dominance (the masking of deleterious alleles), overdominance (single locus heterosis), and epistasis (enhanced performance of traits derived from different lineages).

A common assumption is that hybrids are intermediate, phenotypically, with respect to their parental species and thus have little to contribute to adaptive evolution. Even when extreme characters are reported, they are considered to be rare or as a consequence of developmental instability. However, studies of quantitative traits in first generation and segregating hybrid populations have reported the presence of phenotypes that are extreme (10% in first generation hybrids and 30% in later generation hybrids) relative to those of either parental type (deVincente and Tanksley

1993; Reiseberg and Ellstrand 1993). The generation of extreme phenotypes is referred to as transgressive segregation and is a major mechanism by which extreme or novel adaptations in new hybrid ecotypes or species are thought to have arisen (Reiseberg et al. 1999). Even though transgressive segregation will have greatest effect in segregating hybrid populations, facultative agamospermous taxa could easily fit into this category.

Therefore agamosperms have the potential to fix a relatively fit genotype through heterosis or the generation of novel or extreme phenotypes. This relatively fit and advantageous genotype can be replicated faithfully generation after generation without fear of being disrupted by recombination. Such a genotype has the potential to survive as long as its environment does not change beyond its tolerance range. This fit agamospermous genotype will survive as long as its environment does not change. If its environment changes it will either survive or become extinct. It is unlikely that without genetic variation and recombination it will adapt to any environmental change that does occur.

Therefore, a population without genetic variation, or more importantly without genetic variation at loci that affect fitness, will not evolve. The previous section has shown that agamospermous populations can be made up of more than one genotype, which could provide a basis of diversity on which selection could act. However, the definition of a population in an agamic complex cannot be the same as in a sexual population. For example, if the genotypes of an agamic population reproduce via obligate agamospermy they are not in genetic contact with each other. Thus, although an agamic complex can show considerable diversity within a location, the diversity and subsequent evolutionary potential within a given population gene pool will be low (Hughes and Richards 1989).

However, true obligate agamospermy, in which all possibility of sexuality has been lost, is a rare phenomenon and some sexuality continues to be a feature of most agamospermous plants (Nogler 1984a; Asker and Jerling 1992). Clausen (1954) was among the first to recognise that facultative apomixis does not necessarily lead to a

loss of variation and evolutionary potential. He argued that a combination of sexuality, which allows for the progressive production of new genotypes, and apomixis, which permits the unlimited and faithful reproduction of the fittest genotypes, would enhance rather than diminish a species capacity for adaptive change.

1.6 Conclusion

The most likely outcome of interspecific hybridisation is hybrid sterility. Agamospermy can restore the fertility of these otherwise sterile individuals. Obligate agamospermous populations have the potential to persist for long periods but appear to have little evolutionary potential. In contrast, facultative agamospermous populations can reproduce both sexually and asexually they therefore have the ability to recombine genotypes and hence, have the potential to evolve.

Even though obligate agamospermous populations have little evolutionary potential, it may be possible that an obligate agamospermous individual can play an import role in the evolution of closely related taxa. For example if a triploid derivative of an interspecific cross, between tetraploid and diploid progenitors, was predominantly an obligate agamosperm it may be possible, on occasion, that an unreduced triploid embryo-sac has the potential to be fertilised by haploid pollen (supplied by a sexual diploid relative). The outcome of such an event would be to produce a tetraploid agamospermous hybrid.

An example of this possible scenario can be found on the isle of Arran, Scotland. It is thought that a sexual diploid (*Sorbus aucuparia* L.) crossed with an apomict tetraploid (*Sorbus rupicola* (Syme) Hedl.) to produce a triploid agamosperm (*Sorbus arranensis* Hedl.), that subsequently backcrossed with the sexual diploid (*Sorbus aucuparia*) to produce a tetraploid agamosperm (*Sorbus pseudofennica* E.F. Warburg).

1.7 The endemic *Sorbus* species on Arran

1.7.1 Introduction

The Arran whitebeam (*Sorbus arranensis*) and the Bastard Mountain Ash (*Sorbus pseudofennica*) were first described by Landsbourgh (1897) who noted the occurrence and distribution of two kinds of cut-leafed whitebeam growing in the gorges of the north part of Arran. For many years *S. arranensis* and *S. pseudofennica* were not separated taxonomically and there has been a certain amount of taxonomic confusion (Bignal 1980). The fourth edition of Bryce's "Geology of Arran" (1872) lists *Pyrus pinnatifida* Sm. as occurring in Gleann Easan Biorach, while Landsborough (1897) mentioned a number of locations in north Arran for *Pyrus aria* (L.) Ehrh. and its varieties. In 1901, Hedlund described *Sorbus arranensis*, but it was not until Clapham et al. "Flora of the British Isles" was published in 1952 that *Sorbus pseudofennica* received authoritative recognition as a separate taxon.

Due to the scarcity of pollen produced by insect-pollinated plants, and difficulties in the identification of *Sorbus* pollen to species level, it cannot be clearly seen when Arran's endemic whitebeams were first present in the woodlands of North Arran. It is certain however, that *Sorbus* has been on Arran since at least 6000 bp and it was more common and widespread in the vicinity of loch a'Mhuilinn 1000 years ago than it is today (Boyd & Dickson 1987). Macroscopic remains (including seeds, wood and charcoal) and pollen records confirm the presence of *S. aucuparia* in Scotland dating from the late Weichselian glacial period (10,000-15,000 bp) (Godwin 1975). *Sorbus rupicola* pollen has been identified in Scotland (Isle of Skye) dating from the early interglacial Flandrian period (8,500-10,000 bp) (Godwin 1975).

1.7.2 Phylogeny

Sorbus species belong to the *Rosaceae* subfamily *Maloideae* that also includes the apples (*Malus* P. Mill. spp.), pears (*Pyrus* L. spp.), hawthorns (*Crataegus* L. spp.), cotoneasters (*Cotoneasters* spp.), and firethorns (*Pyracantha* M. Roemer spp.). The British species of *Sorbus* comprise three sexually reproducing diploids, rowan (*Sorbus aucuparia*), whitebeam (*S. aria* (L.) Crantz), and wild service tree (*S. torminalis* (L.) Crantz) and many interspecific hybrids (Richards 1975; Rich 1998). It

has been suggested, partly on the basis of the morphological resemblance between these species, and partly on the basis of cytological studies made on similar groups of plants in Scandinavia (Liljefors 1953, 1955), that *S. arranensis* is a result of a hybridisation event between rowan (*Sorbus aucuparia*) and rock whitebeam (*S. rupicola*) whilst *S. pseudofennica* may have resulted from a back cross of *S. arranensis* with *S. aucuparia* (Hull and Smart 1984) (Fig.1).

1.7.3 Genetic formulae

The genomic formulae of the hybrid groups in Britain are assumed to be similar to those found in Scandinavian material. It has however, been established that at least some of the individuals growing on Arran and classified on morphological grounds as *S. arranensis* have a chromosome number of 2n = 3x = 51 (Bignal 1980).

Liljefors (1955) suggests that the polyploid derivative of *S. aria* (RR, 2n =2x = 34) in the British case is *S. rupicola* (RRRR, 2n = 4x = 68). Although individuals of this species have not been found in the immediate neighbourhood of the hybrid plants on Arran, individuals do occur elsewhere on Arran and on the neighbouring mainland. This derivative could have hybridised with *S. aucuparia* (AA, 2n = 2x = 34) on different occasions and in different areas to give triploid hybrids (ARR, 2n = 3x = 51). These are known as *S. leyana* Wilmott and *S. minima* (A. Ley) Hedl. in Wales (Wilmott 1934) and *S. arranensis* on Arran and in Norway (Liljefors 1953). Apomictic triploid hybrids may have then backcrossed with *S. aucuparia* to give a tetraploid group (AARR, 2n = 4x = 68) known as *S. hybrida* L. in Scandinavia and *S. pseudofennica* in Arran. This may have occurred by the union of an unreduced egg in an aposporic embryo-sac from *S. arranensis* (ARR) with pollen from *S. aucuparia* (A).

1.7.4 Breeding systems

The supposed triploid plants in the *S. arranensis* group are assumed to reproduce exclusively by apomixis (Liljefors 1953) producing little if any viable pollen, thus depending on pollination by another species for seed set (as they are pseudogamous, requiring pollination and fertilisation for the production of the endosperm before the

apomictic embryo can develop). Seed set is therefore often rather sparse, though some clones are parthenocarpic, producing fruits without pollination and which do not contain viable seeds (McAllister 1986).

If the suggested origin of *S. pseudofennica* is correct then these trees have a diploid set of chromosomes from the *S. rupicola* parent and a full diploid set from the *S. aucuparia* parent. As a result each chromosome has a partner which can pair at meiosis and the trees can therefore produce fertile pollen. This can be seen in the analogous Scandinavian species *S. hybrida* where apospory is partial (Liljefors 1953) and 17-20 per cent of the pollen produced could be germinated (Liljefors 1955). Therefore there is a possibility that some sexual reproduction and genetic exchange can occur within the species (Hull and Smart 1984).

1.7.5 Variation between species

Hull and Smart (1984) used morphometric analysis to study the variation within and between the two Arran *Sorbus* species. It was found that there was a considerable overlap of about 35 % between the two hybrid groups where individuals could not be readily assigned to either species. These phenotypically intermediate individuals may be morphologically intermediate between the two groups because they are genetically so. This would suggest that there is some degree of genetic exchange between them. This might be considered unlikely if the chromosome complement of these groups is as suggested by the Scandinavian material (i.e. *S. arranensis* 2n = 3x = 51 and *S. pseudofennica* 2n = 4x = 68, on the other hand hybridisation would not be impossible even then (Hull and Smart 1984).

1.8 Objectives

The major question of interest is whether these Arran hybrid taxa represent an evolutionary dead end (as a consequence of hybridisation and apomictic reproduction leading to limited genotypic diversity), or whether they are they actively evolving? Do they produce (on occasion) new genotypes via sexual reproduction within and

between the taxa? Genetic markers will be used to address these questions. The answer to the major question is of fundamental importance for the development of conservation policy for this species. If no active evolution is occurring, the conservation policy will be aimed at preserving the unique products of hybridisation that are not to be repeated. On the other hand if evolution can be shown to be progressing then conservation policy should aim to ensure that management allows this evolutionary process to continue. With these overall objectives in mind, a number of questions have been formulated and investigated in this thesis, and are dealt with in chapters 2-5. These questions are:

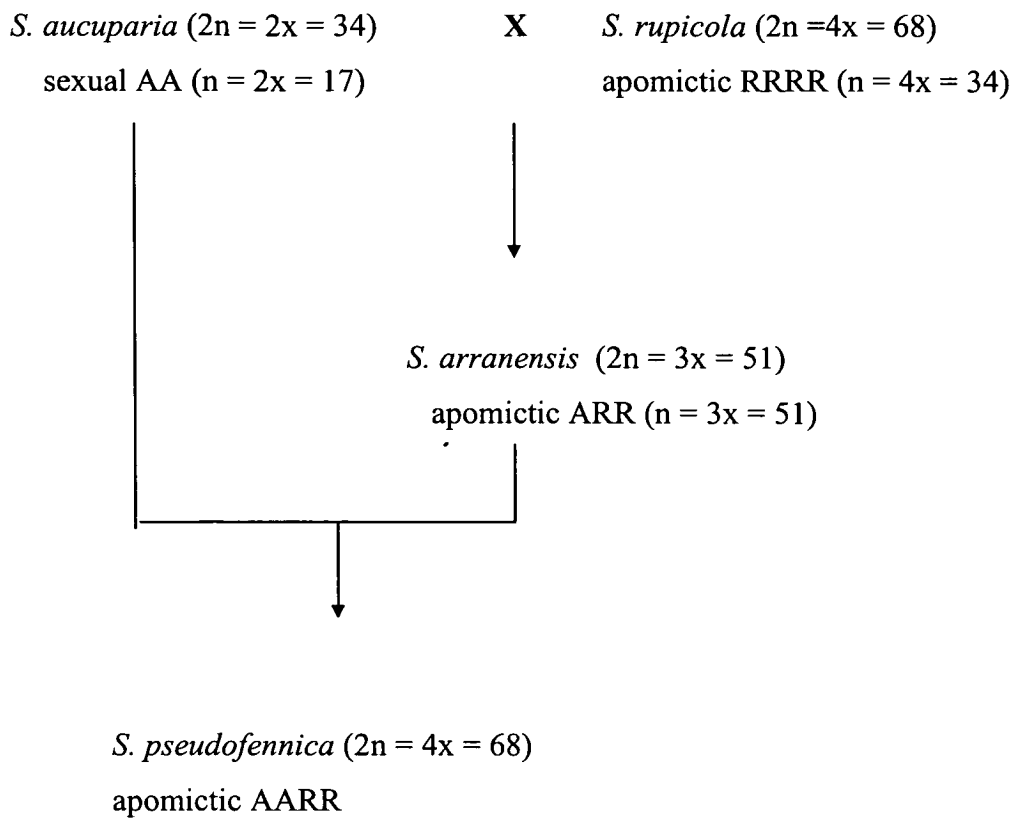
1. What is the origin of the Arran *Sorbus* taxa? Is it as hypothesised in figure1?

2. How many clones (representing genotypic diversity) are there within each of the Arran *Sorbus* taxa?

3. Are the taxa completely apomictic or are they capable of producing offspring that are different from the parents?

4. Is there any evidence that new genotypes are being established and that there is active evolution in this group?

5. What is the best strategy for conservation of the Arran whitebeam taxa?



Key to basic 17 chromosome set genotype.

A = *S. aucuparia*

R = *S. rupicola* (and *S. aria*)

2n = number of chromosomes

n = number of chromosomes in gametes (pollen grains and ovules).

Figure 1. Probable origins of Arran *Sorbus* endemics (adapted from McAllister 1986).

Chapter Two

Testing the proposed hybrid origin of the Arran *Sorbus* Microspecies

2.1 INTRODUCTION

2.1.1 Definition of hybridisation

Harrison (1993) discussed various definitions of the term hybridisation and each relates to levels of divergence between individuals that undergo reproduction. Extremes of these definitions are crosses between genetically distinct individuals and crosses between individuals from different species. Harrison (1990) defined hybrids as the offspring between individuals from two populations that are distinguishable on the basis of one or more heritable characters. Arnold (1997) emphasised that heritable characteristics did not include crosses between individuals from the same gene pool that happen to possess alternate states of polymorphic characters.

2.1.2 Distribution of hybrid taxa

Ellstrand et al. (1996) found natural hybridisation to be non-randomly distributed among taxa, and concluded that the taxonomic distribution of natural hybrids suggests hybridisation is not as widespread as was commonly believed, but concentrated in a small fraction of families and an even smaller fraction of genera. These genera were primarily outcrossing perennials with reproductive mechanisms that are able to stabilise hybridity such as agamospermy and vegetative reproduction.

Grant (1981) suggested, based on the frequency of allopolyploidy, that the majority of plant species have been derived from past hybridisation events. However, Arnold (1997) emphasised that although greater than 50 % of angiosperms may be of hybrid origin, it does not necessarily follow that the number of hybridisation events leading to speciation is equal to at least half the number of angiosperm species. For example, even though each of the five species of *Gossypium* are of allopolyploid origin it is hypothesised that a single allopolyploid ancestor gave rise to all of these species (Dejoode and Wendel 1912).

2.1.3 Study of hybridisation

Arnold (1992) noted that investigations of natural hybridisation have been used to study different aspects of evolutionary biology. Taxonomic biologists have studied the occurrence and extent of natural hybridisation to infer evolutionary relationships

between taxa of interest (e.g. Clausen et al. 1939; Heiser et al. 1969). Mechanisms that permit or limit gene flow between taxa have been of great interest to biologists studying speciation (e.g. Baker and Baker 1990; Howard and Waring 1991). The role of hybridisation in generating novel genotypes that may lead to adaptive evolution and or the formation of new evolutionary lineages is a study of the direct or potential consequences of hybridisation (e.g. Anderson 1949; Rieseberg and Wendel 1993; Arnold and Hodges 1995a).

2.1.4 The use of molecular markers to study plant hybridisation

Most hybrids have been identified on the basis of their morphology (Knobloch 1971) whereby a species that appears intermediate in many respects to two related species may often be assumed to be the outcome of hybridisation. However, there are problems with relying solely on morphological markers as hybrids may express intermediate, parental or even novel characters, making it difficult to determine parentage (Rieseberg and Ellstrand 1993).

In addition, the phenotype of an organism is a function of multiple genetic and environmental effects. Thus, the source of any observed morphological variation among individuals is difficult to define. In contrast, molecular markers provide differences among individuals and/ or populations that generally have an unambiguous genetic basis, and they also tend to be selectively neutral (Kimura 1983), whereas morphological characters often converge when exposed to similar selective pressures. When used alone, molecular markers can be a useful tool in determining hybrid status. However, when used in conjunction with morphological and cytological markers, the significance of the data is increased.

2.1.5 Types of molecular marker

Molecular markers can be directed at the nuclear genome and/or organelle genomes (chloroplast and mitochondria). Organelle markers used alone do not detect hybridisation events as they are usually uniparentally inherited and subsequently only the contribution of one parent is analysed (Soltis and Soltis 1993). However, organelle markers can determine the maternal and paternal genomic contributions to

the hybrid. The direction of gene flow can therefore be established. In separate hybridisation events, either both parental species act as the maternal parent, or there is a bias towards one parental type (e.g. Brysting et. al 2000). Song et al. (1988) predicted that the paternal genome would change more than the maternal genome in a derived allopolyploid hybrid, because the maternal genome donates both the DNA and the cytoplasm. The maternal cytoplasm may provide selection pressure on portions of the foreign nuclear genome, and such cytoplasmic–nuclear interactions may be important in the genetic stabilisation of newly formed hybrids and polyploids (Gill 1991; Jiang and Gill 1994).

Nuclear markers are normally inherited in simple Mendelian fashion and subsequently will be additively combined in the hybrid genotype, allowing the identification of both parental contributors. There are two groups of nuclear marker suitable for the study of hybridisation: isozyme markers, which examine neutral protein sequences, and DNA markers, which generally examine non-coding regions within and between genes or anonymous regions of the genome. Both classes of DNA marker are assumed to be neutral. Isozymes tend to be codominant single copy markers, which can be genetically interpreted. However, their relatively slow rate of mutation can limit the number of species-specific markers found in closely related parental species (Rieseberg et al. 1988).

Nuclear DNA markers that target non-coding regions, in theory, should highlight more genetic variation between the parental species than isozymes. Restriction fragment length polymorphic markers (RFLPs) (Botstien et al. 1980) and microsatellite markers (Morgante and Olivieri 1993) are potentially codominant and therefore genetic interpretations can be made. DNA markers such as RAPDS (Williams et al. 1993) and AFLPS (Vos et al. 1993) are not single copy, and are inherited in a dominant fashion, making genetic interpretations difficult. Nuclear DNA markers that have high mutation rates such as microsatellites may provide polymorphic markers between parent species, but in turn these would also likely to be polymorphic within the parental species.

2.1.6 Hybridisation within the genus Sorbus

Sorbus species belong to the Rosaceae subfamily Maloideae. In Britain the genus *Sorbus* consists of three widespread sexual diploid species *Sorbus aucuparia* (rowan) *S. aria* (common whitebeam) and *S. torminalis* (wild service tree) and one rare sexual diploid species (*S. domestica* L.). There are also a number of morphologically distinct taxa, which are known or presumed to be apomictic polyploids (Richards 1975; McAllister 1986).

Warburg (1938) noted that seven of these apomictic taxa closely resembled *S. aria* (*S. rupicola, S. leptophylla* E.F. Warburg, *S. eminens* E.F. Warburg, *S. hiberinca* E.F. Warburg, *S. lancastriensis* E.G. Warburg, S. *porrigentiformis* E.F. Warburg and *S. wilmottiana* E.F. Warburg). The morphological evidence combined with cytological analysis on analogous Scandinavian forms (*S. salicifolia* (Myrin) Hedl. (= *S. rupicola.*) and *S. obtusifolia* (DC.) Hedl.) by Liljefors (1953, 1955) suggested that the seven taxa are autotetraploid apomicts and likely to have been derived from intraspecific *S. aria* crosses. Warburg (1938) also noted that four taxa were intermediate between *S. aria sensu lato* and *S. aucuparia* (*S. arranensis, S. leyana, S. minima* and *S. anglica* Hedl.). As for the previous group, the morphological data combined with cytological analysis on analogous Scandinavian forms (*S. arranensis* (Norwegian type) *S. neglecta* Hedl., *S. subpinnata* Hedl. and *S. lancifolia* Hedl.) by Liljefors (1953, 1955) suggested that the four taxa are triploid allopolyploids derived from interspecific crosses between *S. aucuparia* and *S. rupicola.* Warburg (1938) finally noted three taxa that were intermediate between *S. aria* and *S. torminalis* (*Sorbus latifolia* (Lam.) Pers. *sensu lato*).

Warburg (1938) in addition proposed that *S. pseudofennica* had arisen as a result of hybridisation between *S. aucuparia* and *S. arranensis.* A proposal that was again supported by Liljefors (1953, 1955) who showed that the analogous Scandinavian *S. hybrida* to be an allotetraploid likely to have been formed from a cross between *S. aucuparia* and *S. arranensis.*

All but one of the apomict taxa native to Britain are considered to be endemic (even though some are morphologically very similar to overseas taxa) and have restricted distributions (McAllister 1986). The single non-endemic member of this group, the tetraploid *S. rupicola*, has a wide but scattered distribution throughout Britain, Ireland, Norway, southern Sweden and Estonia.

2.2 Objectives

The proposed hybrid origin of the Arran *Sorbus* microspecies, where *S. aucuparia* (2n = 2x = 34, n = 2x = 17) crossed with *S. rupicola* (2n = 4x = 68, n = 4x = 34) to produce the triploid apomict hybrid *S. arranensis* (2n = 3x = 51) and subsequently *S. arranensis* (with an unreduced gamete (n = 3x = 51)) crossed with *S. aucuparia* to produce the tetraploid apomict *S. pseudofennica* (2n = 4x = 68), will be tested using three different molecular marker systems: isozyme systems AAT, 6PGD and LAP, chloroplast DNA sequencing and restriction digest analysis, and a nuclear DNA marker.

Under certain conditions it is possible that *S. aria* could have been a potential *S. aria sensu lato* type parent (as opposed to *S. rupicola*) in the primary cross with *S. aucuparia* to produce *S. arranensis*. However, for this event to have been possible it would require *S. aria* to have produced a diploid unreduced gamete (n = 2x = 34) that could have crossed with a haploid *S. aucuparia* gamete (n = 2x = 17). A cross between *S. aria* and *S. aucuparia* under normal conditions (both with reduced haploid gametes, n = 2x = 17) produces a diploid hybrid *S. aucuparia* x *S. aria* = *S. thuringiaca* (Ilse) Fritsch (2n = 2x = 34), with a phenotype and genotype that is significantly different from *S. arranensis* (Richards 1975). Even though it is unlikely that *S. aria* is the *S. aria sensu lato* type parent it is included in this study as a potential progenitor species.

The nuclear markers will be used to identify species-specific markers that taxonomically separate the proposed parental species, and subsequently these species-specific markers will be used to screen the proposed hybrids to determine if

they inherit any of these species-specific markers from the putative parents. If only one set of parental species-specific markers are revealed in the putative hybrids it will be difficult to support the proposed hypothesis. However, if both parental sets of species-specific markers are found in the proposed hybrid taxa, then the hybrid hypothesis based on morphological and cytological data can be supported.

Liljefors (1955) proposed that an autotetraploid such as *S. rupicola* (RRRR) was likely to be the pollen donor in an interspecific cross with *S. aucuparia* (AA) that produced the triploid *S. arranensis* group (RRA). Evidence to support this hypothesis came from embryological studies and pollen analysis, showing that *S. rupicola* was an aposporous obligate apomict (producing only unreduced embryo-sacs) and 20% of the pollen was viable (Liljefors1953, 1955). For *S. aucuparia* to have been the pollen donor in a cross with *S. rupicola*, that created *S. arranensis* (RRA), it would have required *S. rupicola* to have produced a meiotically reduced embryo sac (RR) (in a facultative manner) that could be subsequently fertilised by a male *S. aucuparia* gamete (A). An unreduced *S. rupicola* embryo-sac (RRRR) fertilised by a haploid *S. aucuparia* gamete (A) would produce a pentaploid hybrid (RRRRA). There are no records of such a hybrid.

This hypothesis can be tested using maternally inherited chloroplast markers. As with the nuclear marker assays, species-specific markers will first be determined which taxonomically separate the putative parent species. Subsequently these species-specific markers will be used to screen the proposed hybrid taxa in order to determine which parental chloroplast type has been inherited. By screening a large number of putative hybrid individuals it should be possible to ascertain if *S. rupicola* was the pollen donor in all such hybridisation events (if more than one occurred), or reciprocal crosses had occurred whereby *S. aucuparia* was the pollen donor.

2.3 MATERIALS AND METHODS

2.3.1 Populations of *Sorbus arranensis* and *S. pseudofennica*

The present range of Arran *Sorbus* microspecies is limited to two glens in the north part of Arran, Glen Catacol and Gleann Easan Biorach (Fig. 2.1). These glens are marked in the west by a ridge of high ground overlooking the Kilbrannan sound. This ridge includes the peaks of Beinn Bharrain (715 m and 722 m), Beinn Bhreac (712 m), Meall Biorch (549 m) and Meall nan Dahm (570 m). Smaller peaks occur through the area ranging in height from 309 m (Clachan) to 573 m (Bienn Breach). The main rivers draining the area are the Catacol Burn and Allt Easan Biorach. Notable tributaries are the Diomhan Burn and Allt nan Calman, which flow into Catacol and Allt Dubh, which flows into Alt Easan Boirach.

Granite crags occur sporadically throughout the area and many burns dissect the landscape flowing into steep-sided rocky valleys and gorges. These gorges are the only situations where the *Sorbus* microspecies are found. Previous botanical surveys have identified 14 sites where these microspecies could be located. The present survey was confined to nine of these sites, denoted A, B, C, D, E, I, J, L, and M (Fig. 2.1). These sites range from 10 m (site M) to 350 m (site L) in altitude and vary in aspect, exposure and slope. The climate is wet, mild and windy with a mean annual rainfall ranging from 1778 mm near the coast to 2159 mm inland (Ordnance Survey 1967). The soils at all sites are humus iron podsols and soils of alluvial origin (Macaulay Institute for Soil Research 1982).

The current vegetation of this area is predominantly that of upland moorland, largely dominated by *Molinia caerulea* (L.) Moench and *Calluna vulgaris* (L.) Hull (NCC 1986a). Birch (*Betula* spp.) and rowan (*Sorbus aucuparia*) dominate the fragmented semi-natural woodland, with sporadic occurrences of holly (*Ilex aquifolium* L.), aspen (*Populus tremula* L.), juniper (*Juniperus* spp.) and willow (*Salix* spp.). The woodland is generally sparse and of low stature (less than 8 m tall). At sites where *Sorbus arranensis* and *S. pseudofennica* occur they contribute significantly to the species composition of the woodland.

2.3.2 Plant material

2.3.2.1 Sorbus microspecies

Samples were collected from 179 trees at eight sites and from 142 trees at five sites for *Sorbus arranensis* and *S. pseudofennica* respectively (Table 2.1). Samples were collected in a non-random method (selectively sampled from each patch) in order to maximise the representation of potential genotypes (Bayer 1990).

2.3.2.2 Putative parents

The sexual *Sorbus aucuparia* (rowan) is native and widespread in Britain and has a widespread distribution throughout the island of Arran. Populations in Glen Catacol and Gleann Easan Biorach are sympatric with *S. arranensis* and *S. pseudofennica*. In total 72 samples were collected from trees at seven of these sympatric sites (Table 2.1). Samples from two trees were also collected from the Royal Botanical Garden Edinburgh (origin unknown).

Sorbus aria is native to Britain, common within its range in the south but rare in the north. It is not native to Arran but is likely to be present in private gardens. Samples were collected from eight trees at two sites; seven from Bawsinch nature reserve in Edinburgh, and one from the Royal Botanical Garden Dawyck, Peeblesshire (Table 2.2). It was originally thought that *S. aria* was the only whitebeam species planted at Bawsinch however, on further investigation it became apparent that *S. rupicola* had also been planted. There is a problem with identification, as the sexual *S. aria* is very similar and closely related to the rare native apomict *S. rupicola*. Individual twigs, leaves and herbarium specimens of these species can be very difficult or impossible to distinguish from each other (McAllister 1986).

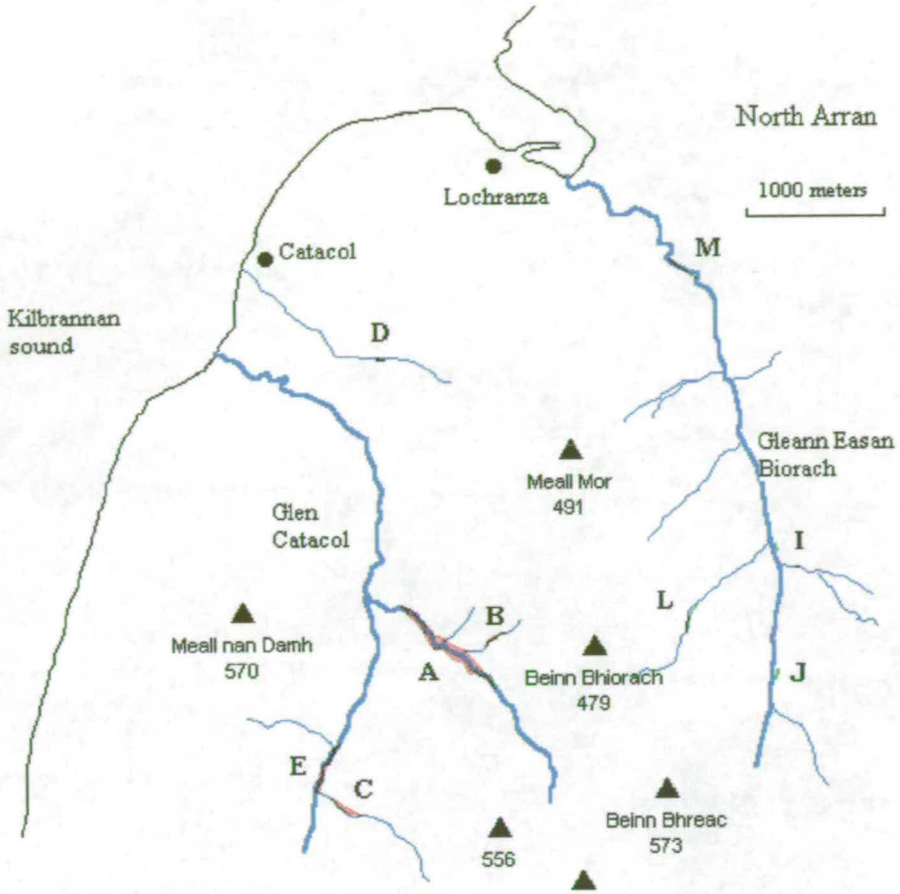


Figure 2.1. Geographic location of sampled sites. The area is covered by Ordnance Survey map, Pathfinder 441, Arran (North). Site A Diomhan Burn grid reference NR 925467. Site B tributary flowing into Diomhan Burn NR 930467. Site C Allt nan Calman NR 918454. Site D Abhainn Bheag NR 924489. Site E Catacol Burn NR 917457. Site I Alt Easan Boirach NR 953474. Site J Alt Easan Boirach 952467. Site L Allt Dubh NR 944470. Site M Alt Easan Boirach NR 945496. Grid references relate to the approximate centre of each site.

Table 2.1. Population and sample sizes for *Sorbus arranensis* and *S. pseudofennica*, and sample size for *S. aucuparia*, collected from different populations on Arran.

	Sorbus arranensis		*S. pseudofennica*		*S. aucuparia*
Site	Population size	Sample size	Population size	Sample size	Sample size
A	220	90	260	79	18
B	30	16	12	7	6
C	20	14	23	23	15
D	0	0	40	14	0
E	24	21	45	19	10
I	13	8	0	0	12
J	3	3	0	0	12
L	28	23	0	0	11
M	4	4	0	0	0
Total	342	179	380	142	72

Population sizes were determined by mapping each individual tree and subsequent transferral of data to Map Maker 2.0 GIS software (Dudley and Map Maker Ltd 1996, 1997)

Table 2.2. Location and origin of *S rupicola* and *S. aria* plant material.

Taxa	Project ID	Collection site	Origin	Acc/GR/locality
*S. rupic*ola	Rua	Herbarium, Edinburgh	Gleann Dubh Arran	GR No. 985335
S. rupicola	Ruj	Herbarium, Edinburgh	North West Jura	GR No. 9967
S. rupicola	R1	Garden, Dawyck	Inverness	Acc No. 1988001
S. rupicola	R2	Garden, Dawyck	Visby, Sweden	Acc No. 19931962
S. rupicola	Ra	Holyrood Park, Edinburgh	Scotland	Dunsapie Ridge
S. rupicola	Rb	Holyrood Park, Edinburgh	Scotland	Dunsapie Ridge
S. rupicola	Rc	Holyrood Park, Edinburgh	Scotland	Samson Ribs
S. rupicola	Rd	Holyrood Park, Edinburgh	Scotland	Powderhouse Gully
S. rupicola	Re	Holyrood Park, Edinburgh	Scotland	Redrock (old tree)
S. rupicola	Rf	Holyrood Park, Edinburgh	Scotland	Redrock (young tree)
S. aria	A1	Bawsinch, Edinburgh	South England	
S. aria	A2	Bawsinch, Edinburgh	South England	
S. aria	A3	Bawsinch, Edinburgh	South England	
S. aria	A4	Bawsinch, Edinburgh	South England	
S. aria	A6	Bawsinch, Edinburgh	South England	
S. aria	A7	Bawsinch, Edinburgh	South England	
S. aria	A9	Bawsinch, Edinburgh	South England	
S. aria	AL	Garden, Dywck	Scotland	Acc. No 19699204

Acc = accession number. Gr = grid reference number

As mentioned above, *Sorbus rupicola* is similar in appearance to the common whitebeam *S. aria.* It is also very similar in appearance to other tetraploid apomicts, presumably derived from *S. aria*: *S. lancastriensis*, *S. porregentiformis*, *S. leptophylla* and *S. vexans* E.F. Warburg (McAlister 1986). *S. rupicola* is the only one of this group that is native to Scotland. Only one tree remains on Arran, at Glen Dubh, approximately 10 miles south of the *Sorbus* microspecies. It is found to the north of Arran on the mainland and on Holy Island.

Leaf samples from *S. rupicola* at Glen Dubh, Arran and the North West of Jura were collected from the herbarium, Royal Botanical Garden Edinburgh. Six samples were collected from Holyrood Park, Edinburgh and two from the Royal Botanical Garden Dawyck. The samples collected from the botanical gardens are of known provenance (Table 2.2). The trees from Holyrood Park are thought to be native *S. rupicola.* However, many cultivated *S. aria* types, sexual and apomict, have been planted in close proximity, leading to the possibility of gene flow between the native and planted groups, and subsequent establishment of hybrid types. There are differences in leaf type among the *S. rupicola* individuals from Holyrood Park, which may be the result of geneflow and subsequent hybrid establishment, or the variation may solely reflect different clones of *S. rupicola.*

Plant material was collected from Arran on two occasions, the last week of June 1997 and the first week of May 1998. Samples from the Royal Botanical Garden Edinburgh were collected in the first week of May 1997. Samples from the Royal Botanical Garden Dawyck were collected in the first week of May 1999. Samples from Bawsinch nature reserve (which were cultivated) and Holyrood Park were collected in the third week of June 1999.

Material was sampled from young freshly clipped leaves and placed into individual bags. While in the field, leaves were kept on blue ice in a cool box. Upon return to the laboratory, samples were stored at 4 °C. For isozyme electrophoresis, proteins were extracted and stored at – 80 °C, and for DNA analysis whole leaf samples were stored at – 20 °C. Material from dried herbarium leaf samples were placed into individual bags and stored at 4 °C.

2.3.3 Electrophoresis procedures

A small section of leaf sample (10 mm^2) was placed in a cooled pestle and crushed in a few drops of chilled extraction buffer (Appendix 1.1). The homogenate was then absorbed on 5 x 3 mm wicks, made from Whatman chromatography paper. The wicks were stored at - 80 °C until further use.

Electrophoresis was performed using a horizontal gel tank system. Gel moulds and running trays were constructed, to the specifications of Cheliak and Pitel (1984). Sample wicks were inserted into gels made up of polymerised starch solution in a buffer of the required pH and salt concentration (gel buffer). Eight enzyme systems were tested in five buffer systems (Table 2.3).

2.3.3.1 Method

32.52 g of starch (potato starch, electrophoresis grade) and 270 ml of an appropriate gel buffer (Appendix 1.2) were mixed in a 500 ml side-arm flask. The mixture was heated slowly over a Bunsen burner, shaking constantly, until the starch had dissolved and the solution began to boil. A tap vacuum-line was applied to the side arm and a vacuum applied for approximately 30 seconds to de-gas the solution. The gel solution was poured smoothly and quickly into the gel mould. Gels were covered with cling film (after cooling) and left to set overnight. The next morning sample wicks and markers were loaded. Gels were run in a refrigerator at 4 °C with 50 mA of constant current for eight hours, for system 6PGD, and with 55 mA of constant current for five hours for systems AAT and LAP (Raspe et al. 1998)

After electrophoresis gels were sliced using a fixed cheese wire and stained (as described by Cheliak and Pitel 1984). The gels were incubated at 37 °C until bands appeared (approximately an hour for AAT and up to 2 hours for GPGD and LAP). Scoring was achieved over a light box where necessary.

2.3.4 DNA extraction and quantification

DNA was extracted from frozen leaf material, collected from Arran during the last week of June 1997, using the Scotlab phytopure plant DNA extraction kit (Scotlab 1996). This method proved successful for all PCR applications. The kit was again used to extract DNA from leaf samples collected from Arran during the first week of May 1998. However, on this occasion only 50% of the DNA samples were successful in producing an amplified PCR product. The DNA samples appeared to be more viscous, a result common in DNA extract having a high carbohydrate content (which can act as a PCR inhibitor). This higher carbohydrate content may be due to the 1998 samples being collected seven weeks earlier in the year than the 1997 samples.

A DNA mini-prep method adapted from Doyle and Doyle (1990) and Gillies (1999) was successful in overcoming this problem. The major differences with this protocol compared to the Scotlab kit was that dichloromethane was used in place of chloroform to extract proteins, a proteinase K digest step was introduced, and an additional protein extraction using phenol:chloroform was used. To achieve consistency, DNA for all samples (with the exception of the herbarium samples) was re-extracted using this method. DNA was extracted from the herbarium samples, using the Scotlab phytopure plant DNA extraction kit.

2.3.4.1 DNA extraction method

0.2% v/v B-mercaptoethanol was added to the extraction buffer (1 ml per sample) prior to use (Appendix 1.3). 0.1 g of leaf material was transferred to a 1.5 ml Eppendorf tube (with 500 µl of extraction buffer and a small amount of sand previously added). An electric drill with a plastic bit attached (designed to fit the Eppendorf tube) was used to grind down the leaf sample. A further 500 µl of extraction buffer was added, and the mix was ground again to achieve a homogenous mixture. Samples were kept on ice until incubation, which was for 30 min at 65 °C. Samples were shaken every 5 min. Tubes were then taken out and left to cool on the bench.

500 µl of dichloromethane was added and mixed gently to achieve an emulsion. Samples were centrifuged for 10 min at 1300 rpm. The upper phase of the supernatant was carefully removed and placed into a fresh Eppendorf. This sequence was repeated. 500 µl of ice-cold isopropanol was added and mixed gently. Samples were placed in the freezer for 1 hour. Tubes were centrifuged for 10 min at 1300 rpm. The supernatant was removed and the tubes were left upside down on a Kim Wipe tissue to dry for 5min. 1 ml of 70% ethanol was added to wash the pellet. Tubes were centrifuged at 1300 rpm for 10 min. The supernatant was carefully removed and tubes left to dry on the bench for 15 min. 300 µl of 1 x TE buffer was added and samples were left in the fridge overnight to allow the pellet to resuspend. 10% v/v RNaseA and 8 µl of protienase K was then added and samples were incubated for 2-3 hours at 37 °C or 1 hour at 55 °C. 200 µl of TE and 500 µl of phenol:chloroform were added and mixed gently. The upper phase of the supernatant was carefully transferred to a fresh tube. 10% v/v of 3 M Sodium acetate pH 5.2 was added and mixed. Two volumes of 100% ethanol were added and gently mixed. Samples were left at – 20 °C overnight or – 80°C for 1 hour.

Tubes were centrifuged for 10 min at 1300 rpm. The supernatant was removed and the tubes were left upside down on Kim Wipe tissue to dry for 5 min. 1 ml of 70% ethanol was added to wash the pellet. Tubes were centrifuged at 1300 rpm for 10 min. The supernatant was carefully removed and tubes left to dry on the bench for 15 min. Pellets were resuspended in 200 ul of 1 x TE buffer.

DNA was quantified by electrophoresis through an agarose gel and ethidium bromide staining, which causes DNA to fluoresce with an intensity dependant on its concentration. DNA standards of known concentration were run alongside experimental samples for comparison and judgements of DNA quantity were made by eye.

2.3.4.2 Method for agarose gel electrophoresis

Horizontal agarose gels were prepared, 1% w/v, in 0.5 xTBE using a microwave to heat and dissolve the agarose. 2 µl of sample DNA, mixed with 11 µl of 0.5 x TBE

and 1 µl of loading buffer, was loaded in the gel. Samples were run at a constant voltage 90 v for 1 hour (70 v for mini gel). Staining was achieved by placing gels in a tray containing 0.5 µg/ml ethidium bromide for 20 min. Gels were destained in distilled water for 20 min before visualisation by UV transillumination. Gels were photographed with the Bio Image system to retain a permanent record.

2.3.5 DNA amplification

2.3.5.1 Rubisco introns

PCR primers were designed to amplify an intron region within the Rubisco gene complex (Langdon 1999). The enzyme ribulose-1, 5-bisphosphate carboxylase (Rubisco) is responsible for fixation of carbon dioxide in the Calvin cycle. The holoenzyme is formed by a 16-mer structure that includes eight identical chloroplast-encoded large subunit polypeptides (*rbcL*) and eight small subunit polypeptides (*rbcS*). In green algae and in land plants, the genetic information for the small subunit is encoded in the nuclear genome typically in a small multigene family (Meagher et al. 1989; Palmer 1991).

Six genes of the Rubisco small subunit multigene family were reported in *Mesembryanthemum crystallinum* L. (common ice plant) (DeRocher et al. 1993). All six genes share similar structures, with two introns interrupting the coding sites at regions that are conserved not only among these genes, but also among *rbcS* genes from other species (Dean et al. 1989). The first introns are diverse in sequence and length ranging from 122 bp to 1092 bp. Five of the six second introns are highly conserved. The six *rbcS* genes of *M. crystallinum* are present at either one or two loci. The *rbcS* genes of other species are distributed among genetic loci in a variety of ways. The eight *rbcS* genes of petunia are located to at least three loci (Dean et al. 1985). In tomato the four *rbcS* genes have been mapped to three sites on two chromosomes and in pea all five genes are at a single locus.

The nuclear coding region for the small subunit was the target for the primers.

Nucleotide Sequence of the primers:

Forward ATGCATGCAGGTGTGGC
Reverse GTTGTCGAATCCGATGAT

PCR reaction mixtures (20 µl total volume) contained the following components / concentrations: one unit of Taq DNA polymerase (Bioline), 10x reaction buffer (Bioline) 1.5 mM $MgCl_2$ (Bioline), 0.2 mM of each dNTP (Sigma), with 0.2 µM of each forward and reverse primer and 20-50 ng of genomic DNA.

PCR reactions were carried out in a Perkin-Elmer 480 thermal cycler, in thin-walled 0.6 ml reaction tubes. The following programme was used: 1 step at 95 °C for 3 min, followed by 30 cycles of 1 min at 94 °C, 1 min at 55 °C, 1 min at 72 °C, and a final 5-min extension step of 72 °C.

PCR amplification products were separated by horizontal gel electrophoresis. Agarose gels were prepared, 1.7% w/v, in 1 xTBE using a microwave to heat and dissolve the agarose. 12 µl of PCR product, mixed with 1 µl of loading buffer, was loaded in the gel. Samples were run at a constant voltage 80 v for 2.5 hours. Staining was achieved by placing gels in tray containing 0.5 µg/ml EtBr for 20 min. Gels were destained in distilled water for 20 min before visualisation by UV transillumination. Gels were photographed with the Bio Image system to retain a permanent record.

2.3.5.2 Chloroplast non-coding regions

Universal primers were used to amplify three non-coding regions of chloroplast DNA via PCR (Taberlet et al. 1991). Primers anneal to highly conserved regions, within single copy chloroplast tRNA genes, which flank the non-coding regions.

Regions amplified by primers: primers a-b amplify an intergenic spacer between *trn*T (UGU) and the *trn*L (UAA) 5' exon. Primers c-d amplify the *trn*L (UAA) intron, and primers e-f amplify the intergenic spacer between the *trn*L (UAA) 3' exon and *trn*F (GAA).

Nucleotide sequence of the Primers:

a. CATTACAAATGCGATGCTCT
b. TCTACCGATTTCGCCATATC
c. CGAAATCGGTAGACGCTACG
d. GGGGATAGAGGGACTTGAAC
e. GGTTCAAGTCCCTCTATCCC
f. ATTTGAACTGGTGACACGAG

These primers are commonly used in surveys of organelle variation, for example the expected size of PCR products for tobacco are 773 bp with primers a and b, 577 bp with primers c and d and 438 bp with primers e and f (Taberlet et al. 1991).

PCR reaction mixtures (20 µl total volume) contained the following components / concentrations: one unit of Taq DNA polymerase (Promega), 10x reaction buffer (Promega), 1.5 mM $MgCl_2$ (Promega), 0.2 mM of each dNTP (Sigma), with 0.2 µM of each forward and reverse primer and 20-50 ng of genomic DNA. PCR reactions were carried out in a Perkin-Elmer 480 thermal cycler, in thin-walled 0.6 ml reaction tubes. The following programme was used: 1 step at 94 °C for 5 min, followed by 35 cycles of 1 min at 94 °C, 1 min at 50 °C, 2 min at 72 °C, and a final 5-min extension step of 72°C. PCR products were purified using QIAquick spin columns (Qiagen) to prepare for sequencing.

2.3.6 Alu restriction enzyme digests

Restriction digest Alu reaction mixtures contained the following components/ concentrations: 2 µl of restriction enzyme10x reaction buffer (Promega), 0.2 µl of BSA (10 mg/ml), 0.5 µl of restriction enzyme (10 µ/µl), 4 µl of DNA (PCR sample), made up to a final volume 20 µl.

Samples were mixed gently by pipetting, centrifuged briefly and incubated overnight at 37 °C. Restriction digest products were separated by horizontal gel electrophoresis

using the Rubisco intron protocol, with the exception of 18 µl of digest product being loaded in the gel.

2.4 RESULTS

2.4 1 Isozymes

Of the eight enzyme systems tested, three (AAT, 6PGD and LAP) showed at least one clear and sufficiently stained zone that could be used in subsequent analysis (Table 2.3). These three enzyme systems were then used to screen 74 *Sorbus aucuparia*, eight *S. rupicola,* three *S. aria,* 145 *S. arranensis* and 105 *S. pseudofennica* individuals.

Owing to the complex, multibanded, non-segregating and often unbalanced, electrophoretic banding patterns produced by the *Sorbus* hybrids, direct genetic interpretation of these patterns was not possible. Instead banding patterns were compared to those produced by *Sorbus aucuparia* and genetic interpretations were inferred from Raspé *et al.* (1998).

2.4.1.1 Leucine aminopeptidase (LAP)

A single LAP locus (*Lap-1*) was observed for each of the five *Sorbus* taxa. All *Sorbus aucuparia, S. rupicola, S. aria, S. arranensis* and *S. pseudofennica* individuals had the same homozygous *Lap-1* genotype.

2.4.1.2 Aspartate amino transferase (AAT)

Three phenotypes, two single-banded and one three-banded, were observed for *Sorbus aucuparia* (Table 2.4). The single-banded phenotypes correspond to homozygotes for either of the two putative alleles $Aat\text{-}1^1$ and $Aat\text{-}1^2$. The three-banded phenotype represents a heterozygote with alleles $Aat\text{-}1^1$and $Aat\text{-}1^2$ and a single heterodimeric band $Aat\text{-}1^1 1^2$ (Fig. 2.2). *Sorbus rupicola* and *S. aria* shared the same, fixed, heterozygous genotype, with alleles $Aat\text{-}1^3$ and $Aat\text{-}1^4$ and heterodimeric band $Aat\text{-}1^3 1^4$ (Fig. 2.2). *Sorbus arranensis* and *S. pseudofennica* had an identical,

fixed, heterozygous genotype, with alleles *Aat-1*^{2} and *Aat-1*^{3} and heterodimeric band *Aat-1*^{2}*1*^{3} (Fig. 2.2).

When comparing the genotypes of the putative parents, *Sorbus aucuparia* and *S. rupicola*, alleles *Aat-1*^{1} *Aat-1*^{2} and the heterodimer *Aat-1*^{1}*1*^{2} were specific to *S. aucuparia,* and alleles *Aat-1*^{3} and *Aat-1*^{4} and the heterodimeric *Aat-1*^{3}*1*^{4} were specific to *S. rupicola*. Alleles specific to *S. aucuparia* (*Aat-1*^{2}) and to *S. rupicola* and *S. aria* (*Aat-1*^{3}) were combined in the *Sorbus* hybrids as a heterozygous genotype (Figure 2.2, Table 2.4).

2.4.1.3 6- Phosphogluconate dehydrogenase (6PGD)

A fixed four-banded phenotype was observed for *Sorbus aucuparia*. The single bands *6Pgd-1*, *6Pgd-2*^{1} and *6Pgd-3* represent homozygous loci, with the fourth band *6Pgd-2*^{1}*3* being the product of an intergenic heterodimer between *6Pgd 2*^{1} and *6Pgd-3* (Fig. 2.3).

Sorbus aria and *S. rupicola* shared the same, fixed, five-banded phenotype with the single bands *6pgd-1* and *6pgd-4* representing homozygous loci and the three bands *6pgd-2*^{1} and *6Pgd-2*^{2} and their heterodimer *6pgd*-2^{1}2^2 representing a heterozygous locus (Fig. 2.3).

Sorbus arranensis and *S. pseudofennica* shared the same, fixed, seven-banded phenotype, loci *6Pgd-1*, *6Pgd-3* and *6Pgd-4* being homozygous and loci *6Pgd-2* heterozygous with allele *6pgd-2*^{1} and the heterodimer *6pgd-2*^{1}*2*^{2} (it was possible that the allele *6pgd-2*^{2} was also present but as a result of a lower copy number was too weak to be resolved). Two intergenic heterodimers *6Pgd-2*^{1}*3* and *6Pgd-2*^{2}*3* (or *6Pgd-2*^{1}*2*^{2}*3*) were also resolved (Fig. 2.3).

When comparing the genotypes of the putative parents *Sorbus aucuparia* and *S. rupicola*, loci *6Pgd-3* and the intergenic heterodimer *6Pgd-2*^{1}*3* were specific to *S. aucuparia,* loci *6Pgd-4*, allele *6Pgd-2*^{2} and the heterodimer *6pgd-2*^{1}*2*^{2} were specific

to *S. rupicola*. The putative hybrids *S. arranensis* and *S. pseudofennica* showed an additive banding pattern representative of the parental species-specific markers.

Table 2.3. Results of initial isozyme assays.

Enzyme	Buffer System				
	LiBo pH 8.3	LiBo pH 7.6/8.0	MoCi 6.1	B	Soltis pH 7.0
AAT	A	B	B	B	-
ADH	-	-	C	-	C
IDH	-	-	D	-	C
LAP	B	C	-	A	-
6PGD	-	-	A	-	B
PGI	B	C	-	B	C
PGM	C	C	B	C	B
SDH	-	-	B	C	B

A = good resolution and staining; B = weak resolution; C = no signal
D = inconsistent resolution; - = not tried; AAT = aspartate amino transferase; ADH = alcohol dehydorgenase; IDH = isocitrate dehydrogenase; Lap = leucine aminopeptidase; 6PGD = 6-phosphogluconate dehydrogenase; PGI = phosphoglucoisomerase; PGM = phosphoglucomutase; SDH = sorbitol dehydrogenase.

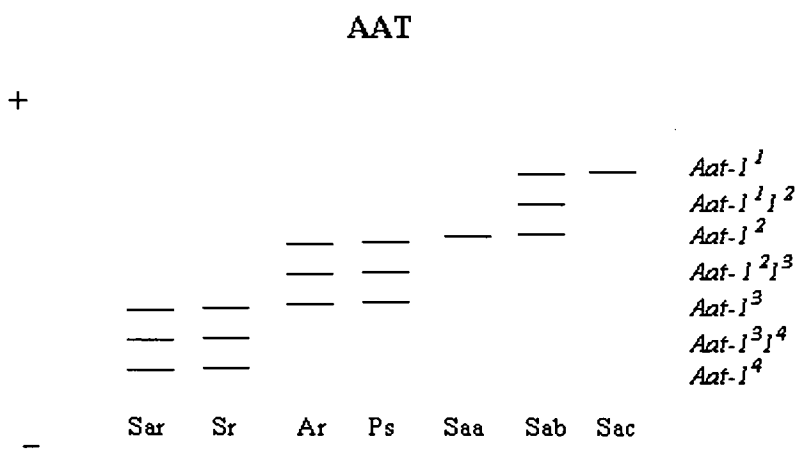


Figure 2.2. AAT isozyme banding patterns. Sar = *S. aria*, Sr = *S. rupicola*, Ar = *S. arranensis*, Ps = *S. pseudofennica*. Saa, Sab, Sac represent the three genotypes found in *S. aucuparia*.

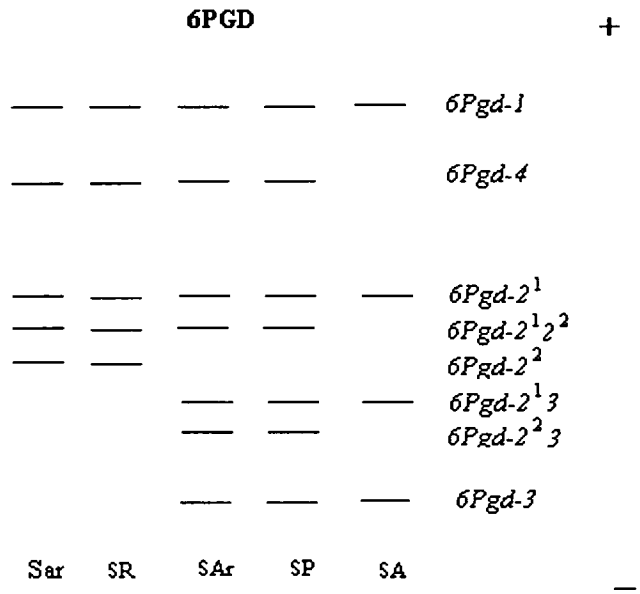


Figure 2.3. 6PGD isozyme banding patterns. Sar = *S. aria*, Sr = *S. rupicola,* Ar = *S. arranensis,* SP = *S. pseudofennica.* SA = *aucuparia.*

Table 2.4. Frequencies of *Sorbus* putative phenotypes for AAT and 6PGD.

Enzyme phenotype	Sa (74)	Sr (8)	Sar (3)	Ar (145)	Ps (105)
Aat-1^1	0.013	0.0	0.0	0.0	0.0
Aat-1^2	0.973	0.0	0.0	0.0	0.0
Aat-1^1, *Aat-1*^{2}, *Aat-1*^{1}*1*^{2}	0.013	0.0	0.0	0.0	0.0
Aat-1^3, *Aat-1*^{4}, *Aat-1*^{3}*1*^{4}	0.0	1.0	1.0	0.0	0.0
Aat-1^2, *Aat-1*^{3}, *Aat-1*^{2}*1*^{3}	0.0	0.0	0.0	1.0	1.0

Enzyme phenotype	Sa (74)	Sr (8)	Sar (3)	Ar (145)	Ps (105)
6Pgd-1, *6Pgd-2*^{1}, *6Pgd-3* *6Pgd-2*^{1}*3*	1.0	0.0	0.0	0.0	0.0
6pgd-1, *6pgd-4*, *6pgd-2*^{1} *6Pgd-2*^{2}, *6pgd-2*^{1}*2*^{2}	0.0	1.0	1.0	0.0	0.0
6Pgd-1, *6Pgd-3*, *6Pgd-4*, *6pgd-2*^{1} *6pgd-2*^{1}*2*^{2} *6Pgd-2*^{1}*3*, *6Pgd-2*^{2}*3*	0.0	0.0	0.0	1.0	1.0

Sa = *S. aucuparia*, Sr = *Sorbus rupicola*, Sar = *S. aria*
Ar = *S. arranensis*, Ps = *S. pseudofennica*
Sample number in brackets

2.4.2 Rubisco intron

The rbcS primers were used to screen 16 *Sorbus aucuparia,* ten *S. rupicola*, eight *S. aria,* 179 *S. arranensis* and 142 *S. pseudofennica* individuals. As might be predicted with a multigene family more than a single band was amplified. However, only three loci were evident and reproducible enough to be scored.

The genetic interpretations of the banding patterns are based on the following assumptions. The rbcS bands are amplification products of different genes, within the gene family, and not different alleles, and for each species the gene family is located in a tightly linked group which will in effect segregate as a single locus.

A fixed single-banded phenotype was observed for *S. aucuparia* (SA), with this single band, *rbcS-1*^{2}, representing a homozygous locus (Fig. 2.4). Two phenotypes were observed for *S. rupicola* (Table 2.5): type RA, a three-banded phenotype, with bands *rbcS-1*^{1}, *rbcS-3*^{1} and *rbcS-4* representing homozygous loci; and type RB, a five–banded phenotype with additional bands, *rbcS-2 and rbcS-5,* also representing homozygous loci (Fig.2.5). *S. aria* shared the same banding pattern as *S. rupicola* type RA (Fig. 2.5).

S. arranensis had a fixed, four-banded phenotype (AA). Locus *rbcS-1* was heterozygous (with alleles *rbcS-1*^{1} and *rbcS-1*^{2}). Loci *rbcS-3* (with allele *rbcS-3*^{1}) and *rbcS-4* were both homozygous (Fig. 2.4).

Three phenotypes were observed for *S. pseudofennica*: type PA was a four-banded phenotype with the same banding pattern as *S. arranensis* (Fig. 2.4); type PB was a five-banded phenotype with two heterzygous loci *rbcS-1* (with alleles *rbcS*-1^{1}and *rbcS-1*^{2}) and *rbcS-3* (with alleles *rbcS-3*^{1} and *rbcS-3*^{2}) and a single homozygous locus *rbcS*-4; and type PC was a four-banded phenotype heterozygous at locus *rbcS-1* (with alleles *rbcS*-1^{1}and *rbcS-1*^{2}) and homozygous at two loci *rbcS-3* (with allele *rbcS-3*^{2}) and *rbcS*-4 (Fig. 2.6).

When comparing the genotypes of the proposed parents *Sorbus aucuparia* and *S. rupicola*, allele *rbcS-1*^{2} was specific to *S. aucuparia* while alleles *rbcS-1*^{1}, *rbcS-3*^{1} and locus *rbcs-4* were specific to *S. rupicola*. *S. arranensis* and *S. pseudofennica* showed an additive banding pattern for rbcS that could have been produced either by crosses of *S. aucuparia* with *S. rupicola* or by crosses of *S. aucuparia* with *S. aria*. It is possible that the source of the additional allele, in *S. pseudofennica* type PB and PC, was the result of a new hybridisation event with a genetically different individual (not sampled) or the result of a mutation.

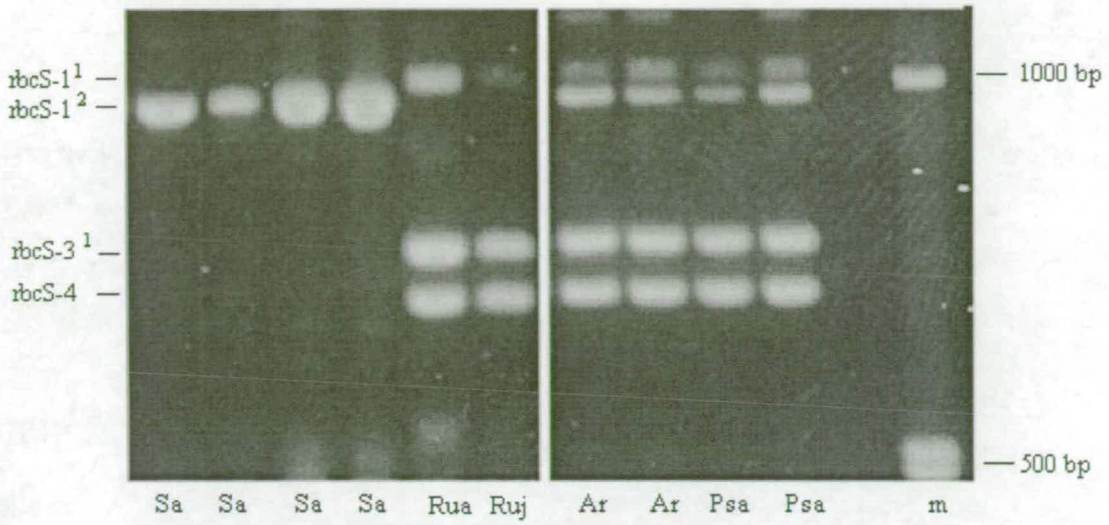


Figure 2.4. Rubisco intron banding patterns. Sa = *S. aucuparia* (SA), Rua, Ruj = *Sorbus rupicola* (Type RA), Ar = *S. arranensis* (AA), Psa = *S. pseudofennica* (Type PA), m = 1kb marker. *Sorbus rupicola* identification (Table 2).

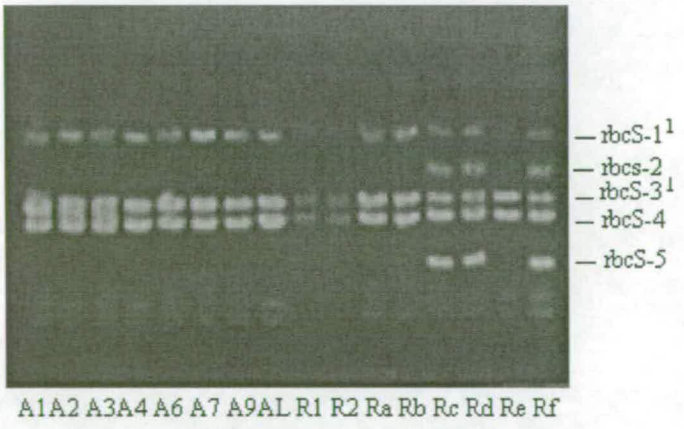


Figure 2.5. Rubisco intron banding patterns for *S. aria* and S. *rupicola*. A1, A2 A3 A4 A6 A7 A9 AL = *Sorbus aria*. R1, R2, Ra, Rb, Re = *S. rupicola* (Type RA). Rc, Rd, Rf = *S. rupicola* (Type RB). *Sorbus rupicola* identification, table 2.

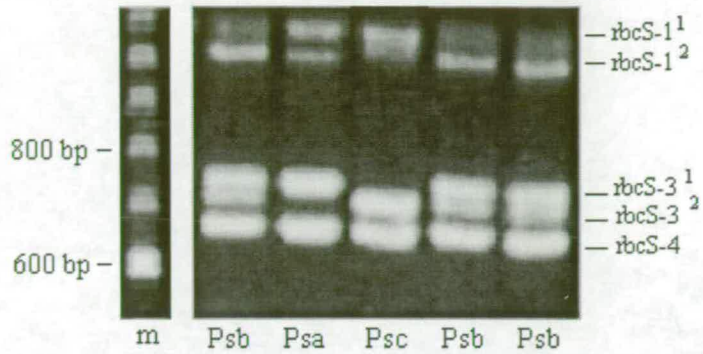


Figure 2.6. Rubisco intron banding patterns for *Sorbus pseudofennica*. Psa = *S. pseudofennica* (Type PA), Psb = *S. pseudofennica* (Type PB), Psc = *S. pseudofennica* (Type PC), m =100 bp marker.

Table 2.5. Frequencies of putative rubisco phenotypes in the *Sorbus* taxa phenotypes for the Rubisco intron.

Putative rubisco phenotypes	Sa (16)	Sr (10)	Sar (8)	Ar (179)	Ps (142)
SA	1.0	0.0	0.0	0.0	0.0
RA	0.0	0.70	1.0	0.0	0.0
RB	0.0	0.3	0.0	0.0	0.0
AA, PA	0.0	0.0	0.0	1.0	0.62
PB	0.0	0.0	0.0	0.0	0.37
PC	0.0	0.0	0.0	0.0	0.07

2.4.3 Chloroplast DNA markers

2.4.3.1 Interspecific chloroplast variation

Three individuals of both *Sorbus aucuparia* (two from the Royal Botanic Garden Edinburgh and one from site D Arran) and *S. rupicola* (Rua and Ruj and R1, Table 2.2) were screened with Taberlet primer pairs a-b, c-d, and e-f. Single PCR products were produced for primer pairs c-d (length 530bp) and e-f (length 483 bp); primer pair a-b did not amplify. Nucleotide sequences were determined by direct sequencing of the amplification products. Nucleotide substitutions were evident between *S. aucuparia* and *S. rupicola* at three bases, 154 e-f (T -G polymorphism, *S. aucuparia* and *S. rupicola* respectively), 297 e-f (A-T polymorphism, *S. aucuparia* and *S. rupicola* respectively) and 222 c-d (C-A polymorphism, *S. aucuparia* and *S. rupicola* respectively). Chloroplast sequences did not vary within taxa.

2.4.3.2 Maternal parent

Two individuals of both *S. arranensis* (A 24 and C 12) and *S. pseudofennica* (A9 and E33) were screened (and sequences determined) with the same primer pairs, c-d and e-f, used to identify interspecific chloroplast variation between *S. aucuparia* and *S. rupicola*. The nucleotide sequences *S. arranensis* and *S. pseudofennica* were compared to *S. aucuparia* and *S. rupicola*. The sequences of *Sorbus arranensis* and *S. pseudofennica* matched exactly (100%) the sequence of *S. aucuparia*. Chloroplast sequences did not vary within taxa.

2.4.3.3 Direction of interspecifc geneflow

The nucleotide polymorphism at base 154 (e-f) is situated within an Alu restriction digest site (AGCT) (Fig.7). Five *Sorbus aucuparia*, six *S. rupicola,* three *S. aria*, three *S. arranensis* and two *S. pseudofennica* individuals were again screened with primer pair e-f. On this occasion PCR products were digested with restriction enzyme Alu. As predicted, and confirming the sequence data, *S. aucuparia*, *S. arranensis* and *S. pseudofennica* individuals had two restriction sites, with a digest pattern of three bands 50, 154 and 286 bp and *S. rupicola* had one restriction site,

with a digest pattern of two bands 50 and 440 bp (Fig. 2.8). All three individuals of *S. aria* had the same restriction pattern as *S. rupicola* (Fig. 2.9).

This diagnostic PCR-RFLP marker provided a convenient and efficient method that identified the chloroplast haplotypes of 38 *S. pseudofennica* individuals (26 from site A, six from site C and six from site E) and 54 *S. arranensis* individuals (ten from site A, nine from site C, 11 from site E, four from site I, 18 from site L and two from site M). All *S. pseudofennica* and *S. arranensis* individuals were found to possess the same restriction pattern, and at this position the same chloroplast type as *S. aucuparia* (Fig. 2.9). This implies that *S. aucuparia* was the original maternal parent of the hybrids and there is no evidence for a reciprocal cross with *S. rupicola or S. aria* as the seed parent and *S. aucuparia* as the pollen donor leading to the formation of hybrids.

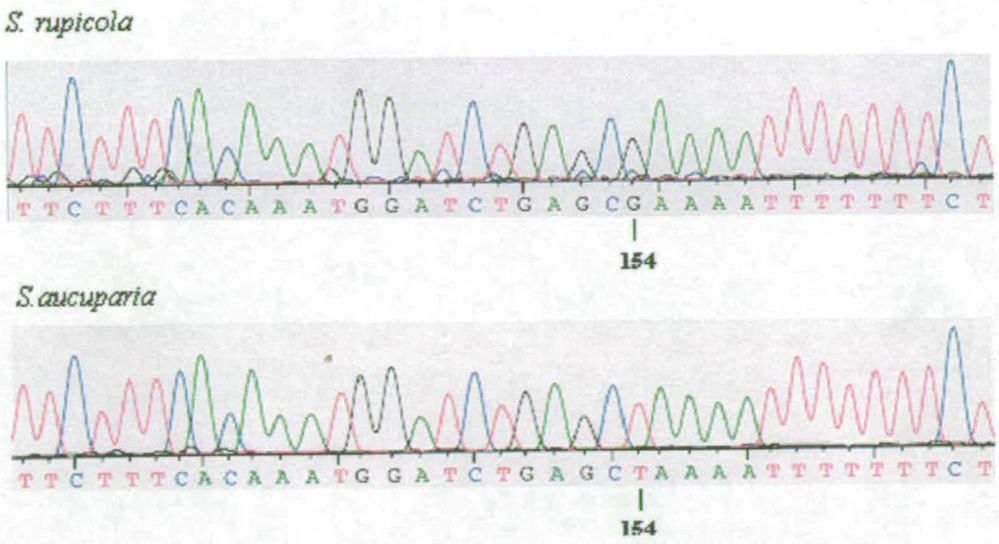


Figure 2.7. Nucleotide polymorphism at base 154 (e-f), between *Sorbus aucuparia* and *S. rupicola,* situated within an Alu restriction digest site (AGCT).

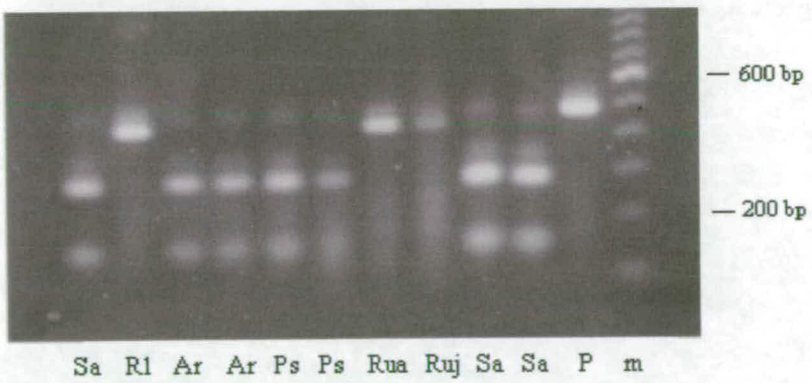


Figure 2.8. Chloroplast PCR products (e-f) digested with restriction enzyme Alu. *S. aucuparia*, *S. arranensis* and *S. pseudofennica* individuals have two restriction sites, with a digest pattern of three bands 50 (not illustrated), 154 and 286 bp and *S. rupicola* has one restriction site, with a digest pattern of two bands 50 and 440 bp. Sa = *S. aucuparia*, R1, Rua, Ruj = *Sorbus rupicola* (Table 2), Ar = *S. arranensis*, Ps = *S. pseudofennica,* P = undigested PCR product, m = 100bp ladder.

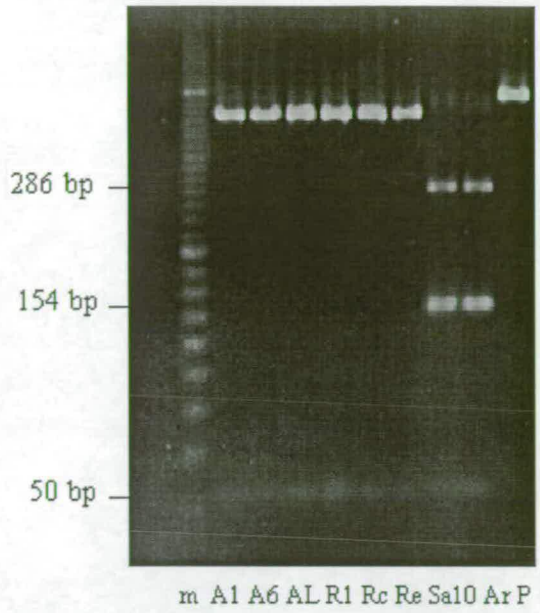


Figure 2.9. Chloroplast PCR products (e-f) digested with restriction enzyme Alu. *S. aucuparia* and *S. arranensis* have two restriction sites, with a digest pattern of three bands 50, 154 and 286 bp and *S. rupicola* and *S. aria* individuals have one restriction site, with a digest pattern of two bands 50 and 440 bp. A1, A6, AL = *S. aria.* R1, Rc, Re = *Sorbus rupicola*. Sa = *S. aucuparia,* Ar = *S. arranensis*, P = undigested PCR product, m = 20 bp ladder. *Sorbus rupicola* and *S. aria* identification (see Table 2.2).

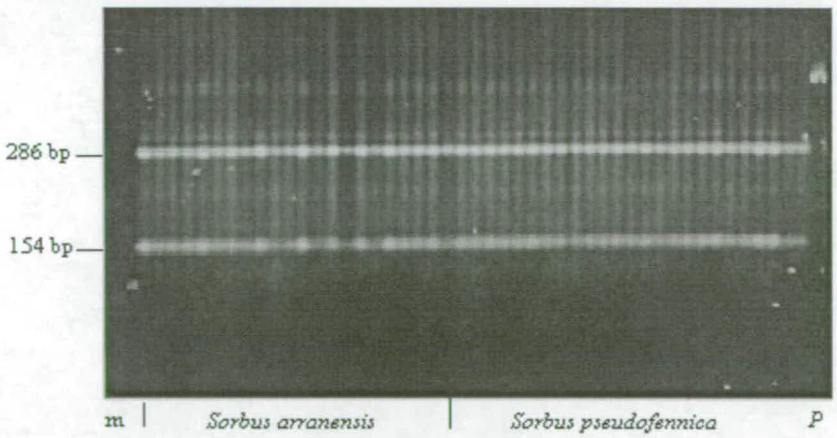


Figure 2.10. Diagnostic PCR-RFLP marker showing the e-f chloroplast haplotypes of the *Sorbus* hybrids to be the same type as *S. aucuparia.* P = undigested PCR product, m = 100bp ladder.

2.5 Discussion

2.5.1 Hybrid origins

Molecular markers were used to test the proposed hybrid origin of the Arran *Sorbus* microspecies, where *S. aucuparia* (2n = 2x = 34, n = 2x = 17) crossed with *S. rupicola* (2n = 4x = 68, n = 4x = 34) to produce the triploid apomict hybrid *S. arranensis* (2n = 3x = 51) and subsequently a second hybridisation event where *S. arranensis* (with an unreduced gamete (n = 3x = 51)) crossed with *S. aucuparia* to produce the tetraploid apomict *S. pseudofennica* (2n = 4x = 68).

Two out of the three isozyme systems AAT and 6PGD, the nuclear DNA Rubisco intron marker and the Chloroplast DNA marker, all proved successful in identifying markers that were specific to *S. aucuparia* and specific to *S. rupicola*. However, the markers could not separate *S. rupicola* from *S. aria* (2n = 2x =34) which could have been the parent in the primary cross with *S. aucuparia* to produce *S. arranensis* (see section 2.2). Similarly, as the markers were unable to fully distinguish between the *S. arranensis* and *S. pseudofennica* taxa, it was not possible to test the proposed second hybridisation event where *S. arranensis* crossed with *S. aucuparia* to create *S. pseudofennica*. It was however possible to test whether the genome of *S. pseudofennica* had ultimately been derived from both *S. aucuparia* and *S. rupicola/S. aria.*

Nuclear markers are particularly suitable markers to study hybridisation events as most nuclear genes are inherited in a simple Mendelian fashion and will be additively combined in the hybrids (Roose and Gottlieb 1976; Haufler 1985b; Lin et al. 1996; Sun 1996).

Aspartate amino transferase (AAT)

For isozyme system AAT the Arran *Sorbus* hybrid genotypes were compatible with a cross between *S. aucuparia* and *S. rupicola/S. aria*. Alleles specific to *S. aucuparia* (*Aat-1*^{2}) and to *S. rupicola/S. aria* (*Aat-1*^{3}) were combined in the *Sorbus* hybrids as a heterozygous genotype.

It is possible that the heterozygous triploid genotype of *S. arranensis* was derived from the fusion of a haploid *S. aucuparia* gamete, *Aat-1*^{2}, with a diploid *S. rupicola/S. aria* gamete homozygous for allele *Aat-1*^{3}. In turn it is possible that the heterozygous tetraploid genotype of *S. pseudofennica* was derived from the fusion of an unreduced triploid *S. arranensis* gamete with the haploid *S aucuparia* gamete *Aat-1*^{2}.

6- Phosphogluconate dehydrogenase (6PGD)

The Arran *Sorbus* hybrids showed an additive banding pattern for isozyme system 6PGD that could have been produced by *S. aucuparia* crossed with *S. rupicola/S. aria.* Locus *6Pgd-3* and the intergenic heterodimer *6Pgd-2*^{1}*3* were specific to *S. aucuparia* and loci *6Pgd-4*, allele *6Pgd-2*^{2} and the heterodimer *6pgd-2*^{1}*2*^{2} were specific to *S. rupicola/S. aria*. Both sets of these parental species-specific markers were expressed in the hybrids.

It is possible that the genotype of *S. arranensis* was derived from the fusion of a haploid *S aucuparia* gamete, with alleles *6Pgd-1*, *6Pgd 2*^{1} and *6Pgd-3,* with a diploid *S. rupicola/S.aria* gamete, homozygous for loci *6pgd-1, 6pgd-4* and heterozygous at *6Pgd-2*. In turn it is possible that the genotype of *S. pseudofennica* was derived from the fusion of an unreduced triploid *S. arranensis* gamete with a haploid *S. aucuparia* gamete, homozygous for loci *6Pgd-1*, *6Pgd 2* (with allele *6Pgd 2*^{1}*)* and *6Pgd-3.*

Rubisco intron

The Arran *Sorbus* hybrids showed an additive banding pattern for the nuclear DNA marker (rubisco intron) that could have been produced by *S. aucuparia* crossed with *S. rupicola/S. aria*. Allele *rbcS-1*^{2} was specific to *S. aucuparia* and alleles *rbcS-1*^{1} *rbcS-3*^{1} and locus *rbcS-4* were specific to *S. rupicola/S. aria*. Both sets of these parental species-specific markers were expressed in the hybrids.

It is possible that the banding pattern of *S. arranensis* was the result of allele *rbcS-1*^{2} being derived from a haploid *S. aucuparia* gamete and alleles *rbcS-1*^{1} and *rbcS-3*^{1} and locus *rbcS-4* were derived from a diploid *S. rupicola* (Type RA) gamete. In turn

it is possible that the banding pattern of *S. pseudofennica* (Type PA) was the result of allele *rbcS-1*^{2} being derived from a haploid *S. aucuparia* gamete and alleles *rbcS-1*^{1} and *rbcS-3*^{1} and locus *rbcS-4* derived from an unreduced triploid *S. arranensis* gamete.

The banding patterns of *S. pseudofennica* type PB and PC showed an additional allele, *rbcS-3*^{2}, which was not present in either *S. aucuparia* or *S. arranensis*. It is possible the source of this additional allele was the result of a second hybridisation event between *S. aucuparia* and *S. arranensis*, where a genetically different *S. arranensis* individual (not sampled) was involved in a cross with *S. aucuparia* to produce a new and genetically distinct *S. pseudofennica* clonal lineage. Or alternatively the source of the new allele could have been the result of a mutation event in a *S. pseudofennica* individual that was subsequently propagated via agamospermy. It is unlikely the additional allele had been derived from a genetically different *S. aucuparia* individual as the additional *S. pseudofennica* allele was derived from a locus that was specific to the *S. rupicola/S. aria* genome complex.

The above results show that the species-specific markers, found in both *S. aucuparia* and *S.rupicola/S.aria*, are combined in both *S. arranensis* and *S. pseudofennica* in an additive fashion that supports the hypothesis of their hybrid origin, with *S. aucuparia* and *S.rupicola/S.aria* as the progenitor species. These results are consistent with morphometric analysis of *S. aucuparia, S. rupicola*, *S. arranensis* and *S. pseudofennica*, carried out by Hull and Smart (1984) who found, using five leaf characteristics, that *S. arranensis* and *S. pseudofennica* had a hybrid index that was intermediate between that of *S. aucuparia* and *S. rupicola.*

The results are also consistent with Proctor *et al.* (1989) who, using a single isozyme system, studied primarily the taxonomy of *Sorbus* species in south-west England, but also included samples of *S. arranensis* and *S. pseudofennica* in their analysis (presumably only a single individual of each type). The banding patterns of *S. aucuparia* and *S. rupicola* were represented as an additive banding pattern in both *S. arranensis* and *S. pseudofennica*. Two bands specific to *S. rupicola* were found in both *S. arranensis* and *S. pseudofennica* and a third band also specific to *S. rupicola*

was additionally found in *S. pseudofennica*. A single band specific to *S. aucuparia* was found in both *S. arranensis* and *S. pseudofennica* and a second *S. aucuparia* specific band was additionally found in *S. pseudofennica*.

The additional S. *rupicola* species-specific band found in *S. pseudofennica* but absent in *S. arranensis* may reflect the result found in the current study, where the banding patterns of *S. pseudofennica* types PB and PC showed an additional allele, *rbcS-3*^{2}, which was not present in either *S. aucuparia* or *S. arranensis*. The additional *S. aucuparia* species-specific band found in *S. pseudofennica* may provide evidence for the proposed backcross hybridisation event between *S. aucuparia* and *S. arranensis* that produced *S. pseudofennica*. This would be possible if the *S. aucuparia* population, or a single individual, was heterozygous at a particular locus, with one allele type being present in the primary hybridisation event, *S. aucuparia* (reduced haploid gamete, A^1, n = 2x = 17) x *S. rupicola/S. aria* (reduced diploid gamete, RR, n = 4x = 34), producing triploid *S. arranensis*, RRA^1, and subsequently the alternative allele being present in the backcross hybridisation event *S. aucuparia* (haploid gamete, A^2, n = 2x = 17) x *S. arranensis* (unreduced triploid gamete, RRA^1, 3n = 51) producing tetraploid *S. pseudofennica*, RRA^1A^2. The additional *S. aucuparia* species-specific band found in *S. pseudofennica* could represent allele A^2.

Many other studies have successfully employed molecular markers to detect evidence of hybridisation (e.g. Roose and Gottlieb 1976; Aas et al. 1994; Kraft et al. 1994; Lin et al.1996; Sun 1996). Aas et al. (1994) found, using isozymes, that it was possible to support a hypothesis (based on morphological analysis) of a hybrid origin for a population of *Sorbus latifolia sensu lato*, with *Sorbus aria* (L.) and *Sorbus torminalis* as the progenitor species. Similarly, Gustafsson (1939) proposed that *Rubus vestericensis* had arisen as a hybrid between *R. pedemontanus* Pinkwart (= *R. bellardi* Weihe) and *R. grabowskii* Wiehe ex Gunther and al. (= *R. thyursanthus* Focke), Kraft et al. (1994) supported this hypothesis with the use of minisatellite nuclear DNA markers.

Lin et al. (1996) used isozymes to test the hypothesis, based on morphological and cytological data, that a hybrid between *Sphenomeris chinensis* (L.) Maxon var.

chinenisis and *S. biflora* (Kaulf.) occurred in Japan. Alleles were detected that were unique to each of the diploid parents and were recognised as species-specific marker alleles. It was found that these species-specific marker alleles were combined in the putative hybrids in a manner consistent with the postulated origin of the taxa, species-specific markers from each parent were present in an additive fashion in the hybrids.

Roose and Gottlieb (1976) found that the primary consequence of combining parental genomes of *Tragopogon* to form tetraploid hybrids *T. mirus* Ownbey and *T. miscellus* Ownbey was that the putative hybrids have a fixed heterozygous multi-enzyme phenotype (specified by 43% and 33 %, respectively, of the 21 duplicated genes examined). Similarly, Sun (1996) found the isozyme phenotypes of the putative hybrid *Spiranthes hongkongensis* confirmed the occurrence of natural hybridisation between and *S. sinensis* (Pers.) Ames and *S. spiralis* (L.) Chevall. Fixed heterozygosity was also a feature of the hybrids (at ten out of 28 loci).

Fixed heterozygosity

Fixed heterozygosity in putative hybrids is a significant indicator of an allopolyploid interspecific hybridisation event (Soltis and Rieseberg 1986). In contrast autopolyploids (not generally thought to have arisen via interspecific hybridisation) show polysomic inheritance, not fixed heterozygosity (Soltis and Rieseberg 1986). However, with significant inbreeding polysomic inheritance could also lead to fixed heterozygosity.

The Arran hybrid loci (*Aat-1* and *rbcS*-1) that were formed by the parental genomes supplying species-specific divergent alleles (as opposed to the parental genomes supplying species-specific loci (*6pgd-3*, *6pgd-4*, *rbcS-2* and *rbcS-3*)) were found to have fixed heterozygosity. Thus providing additional support for an interspecific hybrid origin of the endemic Arran *Sorbus* taxa.

2.5.2 Chloroplast DNA analysis

Organelle markers used alone will not detect hybridisation events as they are usually uniparentally inherited and subsequently only the contribution of one parent is analysed (Soltis et al. 1992; Soltis and Soltis 1993). However, organelle markers can determine the maternal and paternal genomic contributions to the hybrid. The direction of gene flow can be established where in separate hybridisation events, either both parental species act as the maternal parent or there is a bias towards one parental type (Brysting et al. 2000).

Assuming that chloroplast DNA is maternally inherited in *Sorbus* taxa, as in the majority of angiosperms (Palmer et al. 1988; Clegg et al. 991), the sequence data unambiguously demonstrates that *S. aucuparia* is the maternal parent of the Arran *Sorbus* hybrids. The sequences of the trnL intron and the trnL-trnF intergenic spacer for the Arran *Sorbus* hybrids (two samples of each type) were 100 % identical to those of *S. aucuparia*. In contrast nucleotide substitutions were evident between the *Sorbus* hybrids and *S. rupicola* at three bases. These results support the hypothesis that *S. rupicola* (RRRR) was likely to have be the male parent in an intespecific cross, with *S. aucuparia* (AA) acting as the female parent, that produced the triploid hybrid *S. arranensis* (RRA).

To test if the remaining *S. arranensis* and *S. pseudofennica* individuals each had the same maternally inherited *S. aucuparia* chloroplast type, and were not derived from a reciprocal cross, a random sample of 38 *S. pseudofennica* and 54 *S. arranensis* individuals were screened with a diagnostic PCR-RFLP marker. Each of the *S. arranensis* and *S. pseudofennica* individuals were found to possess the same chloroplast digest restriction pattern identical to that of *S. aucuparia*. It may be possible that this represents an example of biased geneflow among interspecific hybridisation events, where for each interspecific cross the species providing the maternal seed parent and the species providing the paternal pollen parent are constant or have a high frequency bias where the frequency of the reciprocal cross is very low. These results are comparable with Campbell et al. (1991) who found, using the Chloroplast PCR-RFLP method, biased geneflow from an apomict tetraploid (*Amelanchier laevis* Wieg.) to a sexual diploid (*A. bartramiana* (Tausch) M.

Roemer), where 13 of 14 hybrid individuals sampled shared the chloroplast restriction site of *A. bartraminana*. It was suggested that the self-incompatible sexual *A. bartramiana* was more likely to respond to foreign pollen than the self-compatible *A. laevis*.

The self-compatibility status of *S. rupicola* has not been tested, however, *S. aucuparia* is self-incompatible (Raspé et al.1998) which may part explain the direction of geneflow between *S. rupicola* and *S. aucuparia*. However, unlike other studies where controlled crosses to were used to detect the direction of interspecific gene flow (e.g. Hollingsworth et. al 1999; Brysting et al.2000), in the present study the chloroplasts of adult Arran *Sorbus* hybrids growing in situ, were analysed. The possibility that all *S. arranensis* individuals (and subsequently all *S. pseudofennica* individuals) were ultimately derived from a single hybridisation event, between *S. aucuparia* (female) and *S. rupicola* (male), must be taken into consideration. The chloroplast type of *S. pseudofennica* is likely to be derived from *S. arranensis*. It is unlikely that *S. arranensis* produces any viable pollen (Liljefors 1955) so when it hybridises with *S. aucuparia*, to create *S. pseudofennica*, it will always be the maternal parent.

Chapter Three

Genetic structure of *Sorbus* microspecies, *S. arranensis* and *S. pseudofennica*.

3.1 INTRODUCTION

It was formerly believed that agamospermous populations were genetically invariant and consequently held little evolutionary potential (Darlington 1939). It was also thought that asexual organisms responded slowly to changing environments, unable to exchange and compile favourable mutations that can lead to adaptive evolution. In contrast, it was believed that sexually reproducing populations were able to adapt to changing environments at a higher rate than asexual populations. Of greater importance perhaps, was the fact that whilst doing so they could maintain a sufficiently large pool of additive genetic variance required for future environmental changes. This is important because a population without genetic variation, or more strictly without genetic variation at loci that affect fitness, will not evolve.

The restricted distribution of asexuality suggests that asexual taxa may be successful in the short term, but that (with a few exceptions, such as the bdelloids) they have been unable to survive for a long time, or to undergo adaptive radiation (Maynard Smith 1978). This is consistent with the view that a successful modification of female germ cell production to allow parthenogenesis is difficult to accomplish and that selective disadvantages to asexual populations may cause their more rapid extinction (Maynard Smith 1978).

In Chapter One it was established that genotypic variation has been recorded within agamospermous populations of a variety of species and the source of this variation is likely to be via sexual events, either at the parental level or within the apomictic group due to facultative behaviour of agamospermous individuals.

The measurement of genotypic diversity within and between populations of an agamospermous microspecies is a prerequisite for understanding the dynamics and evolutionary potential of the microspecies. The evolutionary potential of agamospermous populations will be significantly increased if any new genotypes that do arise, via mutation or sexual recombination, have the ability to disperse to other populations and in so doing increase the genotypic diversity within populations.

Many researchers have argued that an understanding of the genetic composition of a species or population is important for the construction of any comprehensive conservation plan (Hamrick 1983; Falk and Holsinger 1991; Ellstrand and Elam 1993), whereas others have expressed doubt that genetic diversity plays a role in the survival of populations or species (Lande 1988; Schemske et al 1994). These latter authors argue that populations go extinct for ecological reasons (e.g. habitat loss or environmental changes) rather than because they lack genetic variation. Most biologists would agree that ecological factors and demographic characteristics of populations play a dominant role in the short-term survival of populations and species. However, it must be stressed that this applies to short-term survival only. The evolutionary time scale is a difficult concept to appreciate and a lack of concern for the preservation of natural levels of genetic diversity could lead to problems of species survival in the mid- to long-term.

The genotypic variation found within agamospermous taxa tends to be partitioned between populations with very few widespread genotypes (Ellstrand and Roose1987; Bayer 1990; Noyes and Soltis 1996; Gornall 1999). In contrast Proctor et al. (1989; reviewed in Gornall 1999) found that for three *Sorbus* agamospermous microspecies *S. anglica*, *S. devoniensis* E.F. Warburg and *S. porrigentiformis* genotypes were found distributed among more than 75 % of populations.

Hull and Smart (1984), using morphological markers, found a highly significant hybrid index difference between *S. arranensis* individuals from different locations, suggesting that each population is made up of a single unique genotype. However, in contrast, *S. pseudofennica* plants from different locations showed a very similar mean hybrid index, suggesting that each population contains the same set of genotypes (Hull and Smart 1984). In addition Hull and Smart (1984) found that although there was a significant difference in the mean hybrid index between S. *arranensis* and *S. pseudofennica* there was still a 35 % overlap of characters between the two hybrid taxa. Hull and Smart (1984) argued that it could have been possible, with the use of additional characters, to fully discriminate between *S. arranensis* and *S. pseudofennica*. However, it was also argued that the overlap of characters could

also reflect genetic exchange and further hybridisation events between *S. arranensis* and *S. pseudofennica*.

3.2 OBJECTIVES

The following hypotheses, based on morphometric analysis (Hull and Smart1984) were tested using molecular markers:

1. The overlap of morphological characters between *S. arranensis* and *S. pseudofennica* could indicate that there had been genetic exchange between the two taxa, promoting the genesis of new hybrid genotypes.
2. Each *S. arranensis* population consists of a single unique genotype that is genetically isolated and reproduces solely by apomixis.
3. *S. pseudofennica* populations each contain a similar set of genotypes, with the possibility of occasional genetic recombination among individuals, via sexual reproduction.

Nuclear DNA markers (microsatellite and Rubisco intron loci) will be used to determine the number of genotypes present within each population and within each microspecies. The geographical mapping of each individual genotype will permit spatial distribution analysis of these genotypes.

Microsatellites will be used in this part of the study as the three isozyme systems used to screen samples of 179 *S. arranensis* individuals and 142 *S. pseudofennica* individuals failed to detect any genotypic variation, and possibly of greater significance only three *S. aucuparia* genotypes were detected from a total of 72 individual plants sampled (Chapter Two, section 2.4.1).

3.3 MATERIALS AND METHODS

3.3.1 Sample material and DNA extraction

Populations of *S. arranensis* and *S. pseudofennica* were sampled as described in Chapter Two (section 2.3.1). DNA was extracted and quantified as described in Chapter Two (section 2.3.2).

3.3.2 Development of nuclear DNA based markers

3.3.2.1 PCR-RFLPs

Traditional approaches to identify restriction fragment length polymorphisms (RFLPs) have employed cloned DNA, as probes, in Southern blot analysis (Southern 1975). However, there is a considerable effort and cost required to produce and isolate clones that reveal informative single copy RFLPs. Large amounts of high quality DNA, large amounts of radio active material, and long exposure times required for the development of autoradiographs, are some of the drawbacks that have restricted the practical use of Southern blot analysis in large-scale population genetic studies.

Karl and Avise (1993) utilised the PCR process to overcome some of these technical problems. Similar to Southern blot analysis, a nuclear DNA library was constructed and clones were isolated. Instead of using these clones as probes in a Southern blot, their nucleotide sequences were determined. From these nucleotide sequences, unique oligonucleotide primers were designed to flank and hence define the cloned locus. These unique primers would then amplify sample genomic DNA homologous to the cloned locus. Digesting the subsequent PCR amplification product, with a selection of restriction enzymes, would reveal variation in the nucleotide sequence between individuals. Restriction digest products obtained from different individuals were then resolved on agarose gels with ethidium bromide staining.

3.3.2.2 Cloning and sequencing of anonymous Sorbus DNA

For the current study, anonymous nuclear DNA markers were selected using a similar approach as Karl and Avise (1993), except that Inter-SSR amplified fragments were used, instead of a genomic library, as the source of anonymous DNA to be cloned and sequenced.

A *Sorbus arranensis* individual, sample number 60, and a *S. pseudofennica* individual, sample number D59, were screened with Inter-SSR primers, 857 and 1204 (Table 3.1) (following the Inter-SSR procedure outlined in Charters et al. 1996). Six amplified fragments, one from sample 60 and four from D59 screened with primer 1204, and one from D59 screened with primer 857 were excised from the gels (Fig 3.1). Choice of fragment was based primarily on length as the automated sequencing technology has some difficulty in sequencing PCR products over 1000 bp. Gel slices containing the PCR fragments were weighed, and purified using a Qiagen gel extraction kit.

Table 3.1 Nucleotide sequence of Inter-SSR primers

Primer	Sequence
857	5' ACACACACACACACACYG 3'
1204	5' BDBGATGATGATGATGAT 3'

Y, B and D represent degenerate bases, where B = C, G or T (i.e. not A). Similarly, D = not C and Y = not A or G.

The six purified Inter-SSR PCR fragments were ligated into pGEM^{R}-T easy vectors (Promega) (Appendix 3.1). These vectors are convenient systems for cloning PCR products as they are cut with Eco RV and 3' terminal thymidines (T) are added to both ends. These 3'-T overhangs, at the product insertion site, are complementary to the single deoxyadenosine (A) bases that are added, as an artefact of PCR, to the 3' ends of the amplified products (Fig. 3.2).

Plasmids, with the Inter-SSR PCR fragments inserted, were used to transform competent *Escherichia coli* cells. Transformations of *E. coli* JM 109 Competent Cells were carried out as described in the Promega technical manual (1997) (Appendix 3.2).

Both strands of the cloned Inter-SSR fragment were sequenced using primers that anneal to the plasmid's T7 and SP6 promoter regions (which flank the insert site). The sequencing of the cloned Inter- SSR fragments was carried out in a manner as described in Chapter Two. As expected, the nucleotide sequences for each of the six cloned Inter SSR fragments revealed a terminal microsatellite at each end of the fragment (Figure3a - 3f). Each of the 12 terminal microsatellites was complementary to the microsatelite sequence of the Inter-SSR primer used to generate the original PCR fragment.

As a method for detecting nucleotide variation between individuals the PCR-RFLP method has many practical applications. However, it remains far from ideal because of the low level of polymorphism frequently detected in the PCR derived fragments, and the need for a post-amplification processing step to detect polymorphisms.

Three of the cloned Inter-SSR fragments (1204-2, 1204-5 and 857-1) had one or more additional microsatellites in the internal sequence (Figure 3b, 3e and 3f). Fragment 1204-2 had two internal microsatellites, repeat types $(CAA)_6$ and $(GT)_{12}$. Fragment 1204-5 had a single internal microsatellite, repeat type $(GCA)_7$. Fragment 857-1 had three internal microsatellites repeat types $(CA)_8$, $(CG)_6$ and $(GC)_6$.

Microsatellites are usually very polymorphic due to the high level of variation in the number of repeats between individuals (Levinson and Gutman 1987). Amplifying genomic DNA with unique oligonucleotide primers that flank, and hence define the microsatellite locus reveal this variation in the length of the repeat motif. The PCR products are run out on high resolution gels or sequenced directly.

Forward and reverse oligonucleotide primers were designed to amplify four of the above internal microsatellites, $(GCA)_7$, $(GT)_{12}$, $(CAA)_6$ and $(CA)_8$, instead of as originally planned the whole of the cloned Inter-SSR fragment (Table 3.2). The size of the PCR products were kept to within 100 bp-200 bp by judicious choice of the primer site, to allow accurate size differentiation.

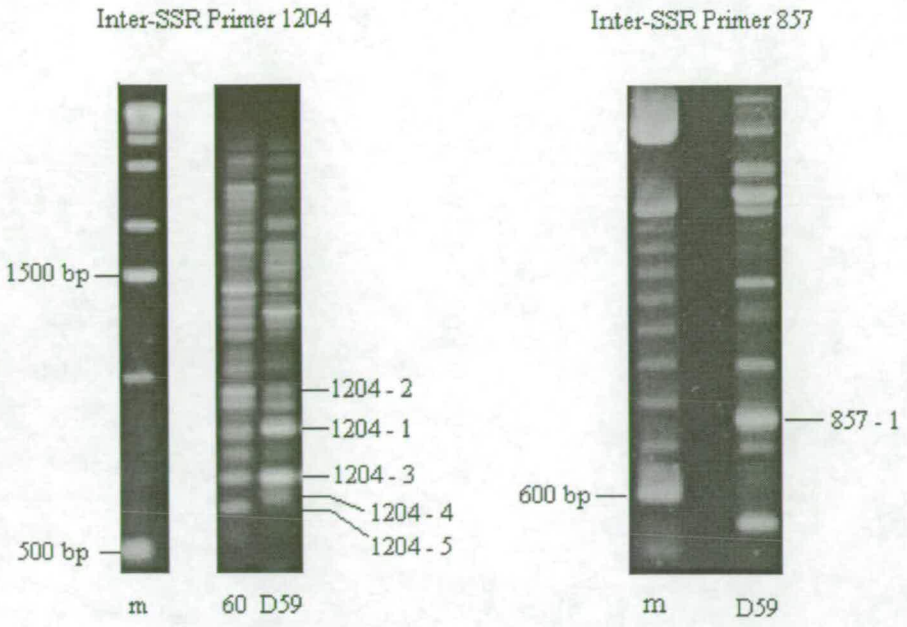


Figure 3.1. Inter-SSR bands used as a source of anonymous *Sorbus* DNA to be cloned and sequenced.

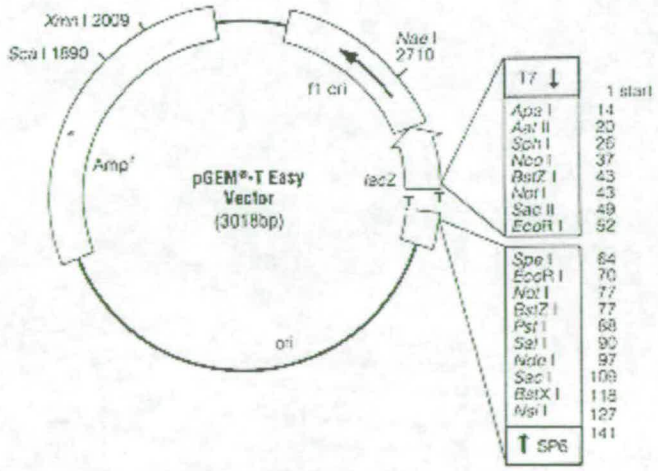


Figure 3.2. pGEM^{R}-T easy vector circle map (Promega 1997).

GGGGATGATGATGATGATGAGGAGCTTGGTAAGCCTTTGACATAGAGTT
AGAAATAGCTCCCACTAGACATGGCCACAGGAAATCCAATCCCGGTAGG
CCTATTAGTCGCAGCCCTCTCCAGCGTTTGGACCTGAGTCTTCAAAAACT
TGACATAGTGAATGGCCTCATCTAGCATGGAAGCGGTGTCCATTTTTGTC
CCGCCCGGGACAAGTCTCTGCAGGATTCGAATCCTCTCACTAATCCTCTC
CCTCCGATGCCGGGCCGCCACGCTCTGAGGGTCCTTAGAAATCTTCACGT
TCCTTCTCTTTGGCGGCTTAACAGATTCCGGGTCGATGTTGATTGGTTGCA
TGGCTGCAATCCGGAATATCATCTCCCTCATGGCTGCCATCGAAGCATGC
TTTTGCGACAGTGTTGAAAACGACCTGCCGTTTGTTTGGTCAGTAGGGTT
TGTGTTGGTGAATGAAATGGTGGATGGTATTGTTGGGAGGTTTAGGAAGG
ATTGTGAGTTTGGTGAATGAATATTATTGTTATGTGTGTTTTGGTCAGTGT
AATGTGACATGGTGGCCCTGAGACTATTAGGACTAGTTGTGGTGGATGAG
AAATTGTGATCGGACGGCGGCAGGTCGGAAACCACATCGTTGTAGGCAC
CGCAGAGCTCAGAGATTTTTTCCATTTGCATCATCATCATCATCCCG

Figure 3.3a. Full sequence for cloned Inter-SSR fragment 1204-1. Shaded areas indicate terminal microsatellites complementary to the microsatellite sequence of the Inter-SSR primer used to generate the original PCR product. Length of fragment 701 bases.

GGGGATGATGATGATGATGGGCAAGTGTGGAAAGAGTTGCTAACCAGCT
CGCTAGCACAGCTGACAGGTGGCGCCCCGCGGCA**CAACAACAACAACA**
ACAAATGACCCAATCACATTGAGTGATCTAAAAGCGATATGCCCGGCTC
GGTTCGGCCCGCTAGCTAGGCGATTAGATGACGGCAGCACTTGATGGATT
CGACATTCGAAGTGCGCGTCAAGCCTATCCGTTCCATTGCAAACGAGCCG
CCGGCCCGATGTGAAGCTGAAACAGGGCAAAAATTCACCCGGCTTGTGA
GATTGTGAGCCGAGCTCATCCGTTTGCGCTGGCTCGAGGCGTCACCCGAT
CGGCTTTTGCCGTCTC**GTGTGTGTGTGTGTGTATGTGTGT**GCGCGCGAG
CCGCATCGCTCAACCTTAATGGAAGTGATCCTTCATCTTGGTCTCGCCTCA
AGTGGCTGCCACCGCCGCGATCGTCTGACATCCGATCAATTATCTTCCAT
TCATCCTTGTTGGCGCGCAAGCTGACGGTGCCGCCGAACAGCTGCCATAT
TGATCGCCCGTATTGCGCGCATCAGCCGGTTGGCTGCAACTCACGCTAGC
CATTCCTCCCCAAAGGTTCGGATAATTCCCTCGACGACGTCATTGAAAGA
GCGCGCGAGTGAACTGGCATCATTATGCAACCCTCGACCACAACAATGG
TTAACCGACACGCATCACTTCCTAACTGCCGCCTTATGGTGGTGGCTATG
GTGGTGGCGACGCCTCGCGTGACGGATTTCGGTGCCGCTGCGTGCTGATG
AAGATAATAATAATAATCATTATCATCATCATCATCCCC

Figure 3.3b. Full sequence for cloned Inter-SSR fragment 1204-2. Shaded areas indicate terminal microsatellites complementary to the microsatellite sequence of the Inter-SSR primer used to generate the original PCR product. Bold letters underlined represent internal microsatellites. Length of fragment 839 bases.

GGGATGATGATGATGATGCAAATGGAAAAAATCTCTGAGCTCTGCGGTG
CCTACAACGATGTGGTTTCCGACCTGCCGCCGTCCGATCACAATTTCTCA
TCCACCACAACTAGTCCTAATAGTCTCAGGGCCACCATGTCACATTACAC
TGACCAAAACACACATAACAATAATATTCATTCACCAAACTCACAATCCT
TCCTAAACCTCCCAACAATACCATCCACCATTTCATTCACCAACACAAAC
CCTACTGACCAAACAAACGGCAGGTCGTTTTCAACACTGTCGCAAAAGC
ATGCTTCGATGGCAGCCATGAGGGAGATGATATTCCGGATTGCAGCCATG
CAACCAATCAACATCGACCCGGAATCTGTTAAGCCGCCAAAGAGAAGGA
ACGTGAAGATTTCTAAGGACCCTCAGAGCGTGGCGGCCCGGCATCGGAG
GGAGAGGATTAGTGAGAGGATTCGAATCCTGCAGAGACTTGTCCCGGGC
GGGACAAAAATGGACACCGCTTCCATGCTAGATGAGGCCATTCACTATGT
CAAGTTTTTGAAGACTCAGGTCCAAACGCTGGAGAGGGCTGCGACTAAT
AGGCCTACCGGGATTGGATTTCCTGTGGCCATGTCTAGTGGGAGCTATTT
CTAACTCTATGTCAAAGGCTTACCAAGCTCCTCATCATCATCATCATCCCC

Figure 3.3c. Full sequence for cloned Inter-SSR fragment 1204-3. Shaded areas indicate terminal microsatellites complementary to the microsatellite sequence of the Inter-SSR primer used to generate the original PCR product. Length of fragment 728 bases

TGGGATGATGATGATGATGACAATAATACTAAGCAAATCATTTAGATTTT
CTCTCACTAAATTTTCTAATGGGCCGTAACTAACAGTCTGGACTACACCT
TCTGACCGTCTGGGTACGAATGTCAATTACCAGATGCCAGATAAACAGG
ACACCCCGTGCATTAGTTCGTGACACTGCACCAATGTATCCAAGAGCACA
AGCTTAGTCTTGGCAGTCAGTTCATACGGAGTCAACTAGGCAAGGCAGGT
GCTGAGCTGACTGACTGGTAGACGAACGGGCATGCTAGCCAGTAAGTCA
ACCGCCGAGGAACAATAACCAAAGAGGTAAACTCTCGATAAGGCCGGGC
GGAACGCGTTCGCTCATTCGCTTCGTCCCCTGTACACGAAGCGACGAACT
TTCTATTCTATCGAAGAAAGCAATCGAAAGCACTCTCTTTTGTGTGATTG
CACCGTCGCTGCAGCAGCAGTGGTAAGCGCCGGCTACAAACACAAACAC
ACACGACCGCGCACGCACTCACTCGCTCACACGACCAGTGAAAGTGTAT
CTCGTCGGCGCCGACAAGAATGTGATCGTTGCCATGGCCACATCTTCTTG
ATCACGATGATCACATGCCAAATCATCATCATCATCCCC

Figure 3.3d. Full sequence for cloned Inter-SSR fragment 1204-4. Shaded areas indicate terminal microsatellites complementary to the microsatellite sequence of the Inter-SSR primer used to generate the original PCR product. Length of fragment 698 bases.

TTGGATGATGATGATGATGTTGCTCCTCTGCTTCGAGTGCTACCACCGCA
ACTATTGCAGCCGCAGCTGGTTGATTGAATTGGCGTCAGTGTCGTTTGTT
GGTGGCGCGCACATCGAGAGCAGCAGTAGCAGGAGCCAGACAAGAATC
ATGCAATTGCAAGTTGCAAAGACGACAACAACTCGTTCCCGAGAAACGC
GCTGCGGCTCTCTTCTCATCTTGGGCGCTCACGCGCGCGTGCTCTCTCACC
CCAAGCTCGTGCAGTGGTGCCACACAAACCCGAGGGGGATGCT**GCAGCA**
GCAGCAGCAGCAGCAACAAGATATGGCTCTCACCCCTCCGCCCCTGTTT
GCTTGGGCTGACATACTAACAGGAGGAGGAGTCGGAGCTGCGCAATCAT
TGGCGACAAATGTTTCTTTGATTCTCTCGCTCACTTGCAAGTGCAGATAA
ACCTCTCGACTTCAGAGAGGAATCATTCGAAGAAGAAGAATATAAGAAG
CGCACTAAACACTTCTACAACAATCGCACACAAGCGTTAATCAAGATATC
AGATTGAAACCAGATCCCAAATTAATTACCACCTGCCGAGCGAACACCA
ACAAAAAACCATCATCATCATCATCCCC

Figure 3.3e. Full sequence for cloned Inter-SSR fragment 1204-5. Shaded areas indicate terminal microsatellites complementary to the microsatellite sequence of the Inter-SSR primer used to generate the original PCR product. Bold letters underlined represent internal microsatellites. Length of fragment 659 bases

ATTACACACACACACACCGCAAATCGGCGCCATGTATTAGATAAAAG
CGGACGCGAGCAAGTTCAACTGGATGCGGCTAAACTTGGCTCACTAGAG
AACTAAAGAACTCGCGAACTCCCCCCGGACGCGACGAGAGGGTGGCTCG
CTCGTTGTTTTCTTGGCCTGGCGAA**CACACACACACACACA**GCTACAAG
CGACAAGGCAAACAATCGATACGAATCACACATCGCGGTTCCGGCCCCA
GCAGGAAAAAAATTCGTCCTCGTTGCTTTCGTGCCGCTTCCGTGCGGTTG
GTTAAGCGGCAGACGATCATCGTCGAGTGGGGGCCAATCCCGACGATGT
CACTGGGGCACGTCGCGCGATTGGGCCACTCGCAGACACGCAAACGAGC
GCATTTATAAT**CGCGCGCGCGCG**AGAGCCTAATGCGCGCGTATTTATAT
GCGCCGCGTGCGGGTCCTTAACCGGCGCACGCGCCACAACAGATAAAAG
TGTCGGCACACTTACATAAGAGCAAACTGTCACTAACTAACAAGTGGCG
AGGAGCCGCCGGGGCCCATTGGGAGAT**GCGCGCGCGCGC**TCTAAGCGG
CCGTTTTGCGCTCCCGGCCCGGCCGTCCCACGACGACACGGGCGCACAAA
AGCATCAAAGATCTCTTCTCTCCAGTGGCACTTTTAGAGAGGGCGCTAAC
AGTTGCGCGCACAGTGTGTGTGTGTGTGTAA

Figure 3.3f. Full sequence for cloned Inter-SSR fragment 857-1. Shaded areas indicate terminal microsatellites complimentary to the microsatellite sequence of the Inter-SSR primer used to generate the original PCR product. Bold letters underlined represent internal microsatellites Length of fragment 722 bases.

3.3.3 DNA amplification

3.3.3.1 Rubisco introns

Rubisco intron amplifications were carried out as described in Chapter Two (section 2.3.5)

3.3.3.2 Microsatellites

In addition to the development of the four *Sorbus* microsatellite markers, a subset of eight primers was chosen from a set of 16 primer pair sequences designed to amplify microsatellite loci within genomic DNA from the genus *Malus* (Gianfranceschi et al. 1998) (Table 3.3). *Malus* species (apple) belong to the *Rosaceae* subfamily *Maloideae* that also includes the *Sorbus* species. The ability to use these *Malus* microsatellite primers in *Sorbus* will depend on the extent to which primer sites flanking microsatellite repeat motifs are conserved between the related taxa and the stability of the microsatellite over time. Some studies have considered cross species amplification in rice (Wu and Tanksley 1993), grapevine (Thomas and Scott 1993) and citrus species (Kijas et al. 1995).

PCR reaction mixtures (20 µl total volume) contained the following components / concentrations: one unit of Taq DNA polymerase (Bioline) 10x reaction buffer (Bioline), 1.5 mM $MgCl_2$ (Bioline), 0.2 mM of each dNTP (Sigma), with 0.2 µM of each forward and reverse primer and 20-50 ng of genomic DNA.

PCR reactions were carried out in a Perkin-Elmer 480 thermal cycler, in thin-walled 0.6 ml reaction tubes. The following programme was used: an initial denaturation at 94 °C for 3 min, followed by four cycles of 94 °C for 1min, 65 °C for 1 min, 72 °C 1 min where the annealing temperature was reduced by 1 °C per cycle. The initial cycles were followed by 30 cycles of 94 °C for 1 min 60 °C for 1 min and 72 °C for 1 min. A final 5 min extension was included.

Primer pairs were preliminarily tested by running PCR products on a 1.5% agarose gel in 0.5 x TBE buffer, stained with Ethidium Bromide and visualised by UV illumination.

Microsatellite alleles were separated by vertical gel electrophoresis, using a Hoefer 600 midi gel rig. High resolution, Metaphor agarose gels were prepared, 4.0% w/v, in 1 x TBE buffer. Using this vertical gel system and 4% Metaphor agarose, it is possible to resolve a 1% size difference with DNA between 100 bp and 600 bp in length.

Glass plates were cleaned with soap and water, left to dry, cleaned again with 100 % ethanol, rinsed with distilled water and left to dry. Unlike polyacrylamide gels, agarose gels do not adhere to glass plates and will slide out during electrophoresis. To prevent this from happening a strip of Whatman chromatography paper, 1 mm thick and long enough to fit between the two spacers, was wetted with running buffer and placed at the bottom of the back plate in contact with the spacers on either side. The front plate was put on and both plates were clamped together to form the gel cassette. The assembled cassette was pre warmed to 55 °C.

Two 1 mm thick standard gels, each capable of holding 25 samples, were prepared, and run together in the same gel rig. Room temperature buffer, 70 ml of 1 x TBE, and a stir bar were added to a 250 ml conical beaker. 2.8 g of high resolution Metaphor agarose was added slowly while the solution was stirring rapidly to prevent the formation of clumps. The agarose was allowed to hydrate in the buffer for 15 min before heating reducing the tendency of the agarose solution to foam during heating. The solution was heated for 1 min on medium power taken out swirled gently and heated for a further 1min or until solution began to boil. The solution was left to cool to 60 °C.

The Metaphor agarose was drawn up into a preheated (55 °C) 60 ml syringe fitted with an adapted needle (designed to fit between the plates). The needle tip was wedged between the plates in the upper corner of the gel cassette. The agarose was

injected at a steady rate to prevent air bubble formation. Combs were inserted into the agarose to the minimum depth necessary to accommodate the samples. During gel casting, a fan heater was placed in front of the gel cassette (at 30 cm) to keep it a constant 55 °C. Thus preventing the agarose from cooling before casting is complete and allowing any air bubbles to rise to the surface. The gels were cooled at room temperature for 15 min, then placed at 4 °C for 20 min.
The screws at the top of the gel cassette were loosened. All excess agarose was removed from top of the gel and from the protruding gel comb with a scalpel to facilitate comb removal. Gel buffer was squirted in the spaces between the comb and the gel. The comb was slowly and gently lifted straight up allowing air or buffer to enter the well area to release the vacuum. The wells were cleaned by flushing with running buffer.

10 µl of PCR product, mixed with 1 µl of loading buffer, was loaded in the wells. Gels were run at a constant 300v. Running times depended on the length of PCR product in use. It was desirable to allow the product to migrate ¾ of the way down the gel. The loading buffer provided a guide, in a 4% gel the xylene cynanol FF dye migrates with DNA at 80 bp and the bromophenol blue with DNA at <20 bp. In general a 90 bp PCR product took 2.5 hours to migrate ¾ of the way down the gel.

Gels were stained with Ethidium Bromide for 15 min and destained for 15 min before visualisation by UV transillumination.

Table 3.2. Nucleotide sequences of microsatellite primers derived from cloned Inter-SSR PCR fragments

Locus	Primers	Repeat type
1204-2A	Forward CACTTCCATTAAGGTTGAGCG	$(GT)_{12}$
	Reverse TTGTGAGATTGTGAGCCGAG	
1204-2B	Forward GTCATCTAATCGCCTAGCTAGC	$(CAA)_6$
	Reverse TGGAAAGAGTTGCTAACCAGC	
1204-5	Forward CGGCTCTCTTCTCATCTTGG	$(GCA)_7$
	Reverse TGTTAGTATGTCAGCCCAAGC	
857-1	Forward ACTAAAGAACTCGCGAACTCC	$(CA)_8$
	Reverse GCGATGTGTGATTCGTATCG	

Table 3.3. Nucleotide sequences of microsatellite primers derived from *Malus x domestica* (Gianfranceschi et al. 1998)

Locus	Primers	Repeat type
CH02D12	Forward AACCAGATTTGCTTGCCATC	(GA)19
	Reverse GCTGGTGGTAAACGTGGTG	
CH02F06	Forward CCCTCTTCAGACCTGCATATG	(TG)10, (AG)20
	Reverse ACTGTTTCCAAGCGATCAGG	
CH02D11	Forward AGCGTCCAGAGCAACAGC	(AG)21
	Reverse AACAAAAGCAGATCCGTTGC	
CH02B03b	Forward ATAAGGATACAAAAACCCTACACAG	(GA)22
	Reverse GACATGTTTGGTTGAAAACTTG	
CH02B12	Forward CAAGGAAATCATCAAAGATTCAAG	(GA)19
	Reverse CAAGTGGCTTCGGATAGTTG	
CH01HI0	Forward TGCAAAGATAGGTAGATATATGCCA	(AG)21
	Reverse AGGAGGGATTGTTTGTGCAC	
CH01E12	Forward AAACTGAAGCCATGAGGGC	(AG)32
	Reverse TTCCAATTCACATGAGGCTG	

3.3.4 Data analysis

The majority of models used to detect the genetic structure of a species commonly assume the populations contain an infinite number of randomly mating diploid individuals. Agamosperms fail to meet these criteria because they are asexual and most are polyploid. Using models, out of context, such as heterozygosity expected under Hardy-Weinberg-Equilibrium, can produce misleading results (Gornall 1999).

An asexually reproducing genome effectively forms one linkage group, enabling the entire genotype to be used as the unit of measurement. This provides an alternative approach to that of dealing in terms of alleles and individual loci. Variability can then be calculated in terms of geneotypic or clonal diversity by using statistics often used in ecology (Peet 1974; Ellstrand and Roose 1987; Menken et al. 1995). The statistics used in the current study were those used by Gornall (1999) to review studies of agamospermous taxa.

Microsatellite and Rubisco intron loci were scored for each individual tree. Scores, 1 for the presence and 0 for the absence of an allele, were then compiled to represent individual genotypes. These individual genotypes were then used as the unit of measurement to identify the structure and distribution of *Sorbus* hybrid clones.

The proportion of clones detected estimated the genotype diversity within microspecies, Ng/Ni, where Ng is the number of clones detected among a total of Ni individuals (pooling data over all populations). Ng/Ni represents the proportion of unique clones in the total sample.

Clonal diversity within microspecies was estimated by Gini's (complement of Simpson's) Index (Pielou, 1969; Peet, 1974), Hg = 1-sum g_i^2, where g is the frequency of the i^{th} genotype, (again pooling data over all populations). Clonal diversity estimates take into account the frequency of the clones in relation to sample size giving a more balanced measure of genetic diversity within the microspecies.

Diversity within populations was expressed in terms of: a) the mean number of genotypes per population, Ng/pop; b) the proportion of populations that are multiclonal and c) mean clonal diversity per population, Hg/pop.

Diversity between populations was expressed as: a) the mean number of populations per genotype; b) the proportion of genotypes that are widespread (i.e. those occurring in at least 75% of the populations) or private (i. e. those restricted to one population).

The above statistical applications are relative measures and can be used to compare samples of different sizes. Parker (1979) discussed their use as indicators of clonal variation in detail. Since it is unlikely that any method or set of methods would uncover all genotype diversity the data described here represents a conservative estimate of diversity with the Arran microspecies.

3.4 RESULTS

3.4.1 Microsatellite amplification

The four primer pairs developed on *Sorbus* DNA and the eight pairs developed on *Malus* DNA were pre-screened, in 2% agarose gels, on 10 *S. pseudofennica* and 10 *S. arranensis* individuals. It was not the intention to search for polymorphism at this stage but to observe the nature of the PCR amplification product. Ideally, for all individuals a single locus would be amplified in a clear and reproducible manner.

Three of the microsatellite loci, amplified using the primer pairs developed on *Sorbus* (Table 3.1), proved to be problematic. For two loci, 1204-2a and 857-1, only 50 % of the samples gave a clear and reproducible PCR amplification product. For locus 1204-2 there was the additional problem of multiple banding patterns. Locus 1204-2b did not amplify at all in any of the samples.

Primer pair 1204-5, and each of the eight primer pairs developed on *Malus* (Table 3.2), produced clear and reproducible banding patterns for all samples. The banding patterns of three of the *Malus* primer pairs CH02D11, CH01H10, and CH02B03b appeared to be variable between a number of individuals.

Each *Malus* primer pair appeared to be suitable for use in screening *Sorbus* samples. However, only the potentially polymorphic primer-pairs CH02D11, CH01H10, CH02B03b, and one chosen at random CH02D12 were selected. Primer pair 1204-5 was also selected (on the basis of being the only the suitable primer pair developed from *Sorbus* DNA).

The five microsatellite primer-pairs and the Rubisco intron primer pair, were therefore chosen as nuclear DNA markers to assess the clonal structure of the Arran *Sorbus* microspecies. These primer pairs were then used to screen 179 *S. arranensis* and 142 *S. pseudofennica* individuals.

When investigating the genetic structure of a species, it is usual to consider the way in which any variability is partitioned within and among populations. Models that provide the conceptual framework for these studies generally use markers that deal in terms of alleles and individual loci. When making genetic interpretations of gel banding patterns, it is important to distinguish between bands that represent different alleles of a single locus, and bands that represent different loci. An understanding of the copy number of the target loci, and how the alleles segregate in controlled breeding experiments, would facilitate the task.

Multiple-banded gel patterns (commonly found in inter-specific polyploid species such as the *Sorbus* hybrids) complicate the task of understanding the genetic nature of the bands. However, as an asexually reproducing genome effectively forms one linkage group, the entire genotype can be used as the unit of measurement. This provides an alternative approach to that of dealing in terms of alleles and individual loci. The banding patterns can then be treated, and scored, as a set of characters and variability can be calculated in terms of genotypic or clonal diversity.

Accordingly the phenotypic variability in banding patterns found at each of the loci scored in this study is described below. The phenotypic variation was then used to delimit genotypes (clones) of the Arran *Sorbus* hybrids.

3.4.2 Nuclear DNA primer pair phenotypes

Microsatellite locus CH02D12

Sorbus arranensis and *S. pseudofennica* shared the same fixed five-banded phenotype, D12A (Fig. 3.4). Of the five bands amplified only a and b were clear and reliable enough to be scored.

Microsatellite locus CH02D11

S. arranensis had a fixed six-banded phenotype D11A (Fig 3.5). Two phenotypes were observed for *S. pseudofennica* (Fig 3.5); a six-banded phenotype, with the same pattern as *S. arranensis*, D11A (frequency 0.07, Table 3.3); and a seven-banded

phenotype, with the additional band b, D11B (frequency 0.93, Table 3.3). Of the seven bands amplified only a and b were clear and reliable enough to be scored.

Microsatellite locus CH02B3b

Sorbus arranensis had a fixed three-banded phenotype, B3A (Fig 3.6). Two phenotypes were observed for *S. pseudofennica* (Fig 3.6); a three-banded phenotype, with the same pattern as *S. arranensis*, B3A (frequency 0.99, Table 3.3); and a single-banded phenotype, with band a only, B3B (frequency 0.01, Table 3.3). Of the 3 bands amplified only a and b were clear and reliable enough to be scored.

Microsatellite locus CH01H01

Sorbus arranensis had a fixed single-banded phenotype H01A (Fig. 3.7). Three phenotypes were observed for *S. pseudofennica* (Fig. 3.7); a three-banded phenotype, with additional bands (in respect to *S. arranensis*) b and c, H01B (frequency 0.88, Table 3.3); a three banded-phenotype, showing length polymorphism in bands b and c, H01C (frequency 0.03, Table 3.3); and a three-banded phenotype, with a different class of length polymorphism for bands b and c, H01D (frequency 0.08, Table 3.3).

Microsatellite locus 1204-5

Three phenotypes were observed for *Sorbus arranensis* (Fig. 3.8); a single-banded phenotype, with band a, 1204A (frequency 0.95, Table 3.3); a single-banded phenotype, with band b, 1204B (frequency 0.01, Table 3.3); and a two-banded phenotype, with bands b and c, 1204C (0.04, Table 3.3). *S. pseudofennica* had two phenotypes (Fig 3.8); a single-banded phenotype, with band a, 1204A (frequency 0.99, Table 3.3); and a two banded phenotype, 1204B (frequency 0.01, Table 3.3).

Rubisco intron locus

S. arranensis had a fixed four-banded phenotype rbcSA (Fig.3.9). Three phenotypes were observed for *S. pseudofennica* (Fig. 3.9); a four-banded phenotype, with the same banding pattern as *S. arranensis*, rbcSA (frequency 0.62, Table 3.3); a five-banded phenotype, with the additional allele *rbcS-3*^{2}, rbcSB (frequency 0.37, Table 3.3); a four-banded phenotype with allele *rbcS-3*^{2}, rbcSC (frequency 0.01, Table 3.3).

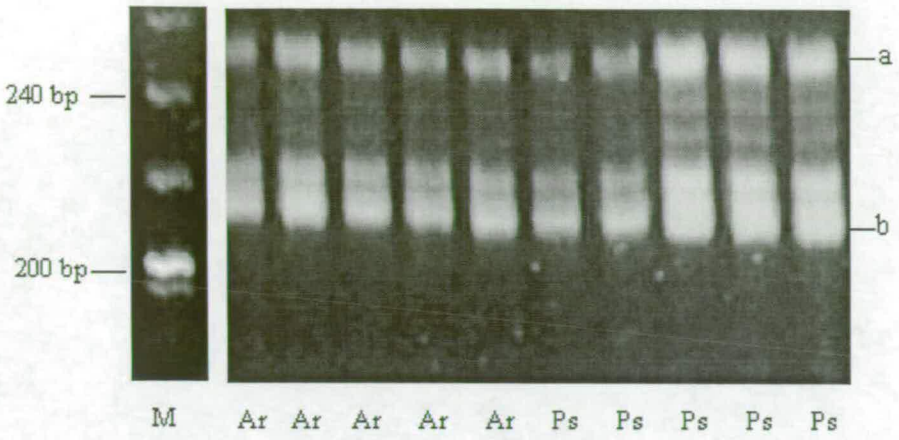


Figure 3.4. Microsatellite locus CH02D12 primer pair phenotypes; Ar = *Sorbus arranensis*, phenotype D12A; Ps = *S. pseudofennica*, phenotype D12A. m = 20 bp marker.

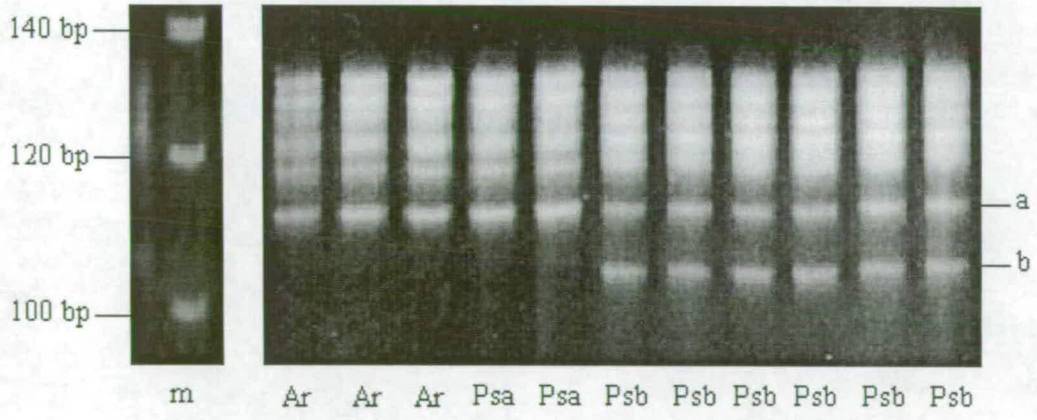


Figure 3.5. Microsatellite locus CH02D11 primer pair phenotypes; Ar = *Sorbus arranensis,* phenotype D11A; Psa = *S. pseudofennica*, phenotype D11A; Psb = *S. pseudofennica*, phenotype D11B. m = 20 bp marker.

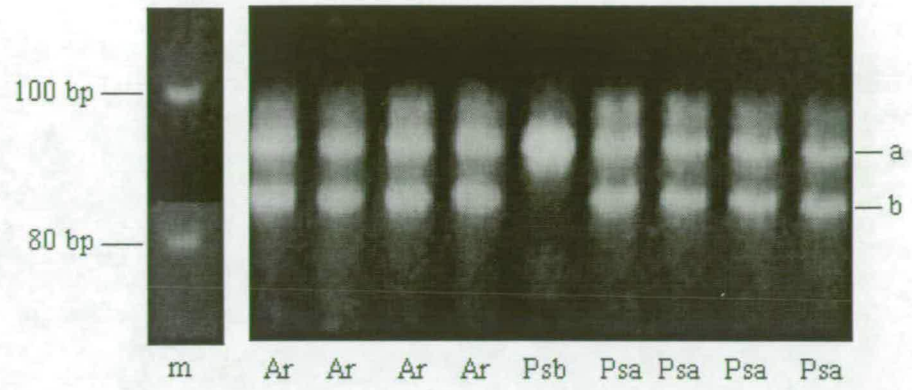


Figure 3.6. Microsatellite locus CH02B3b primer pair phenotypes; Ar = *Sorbus arranensis*, phenotype B3A; Psa = *S. pseudofennica*, phenotype B3A; Psb = *S. pseudofennica*, phenotype B3B. m = 20 bp marker.

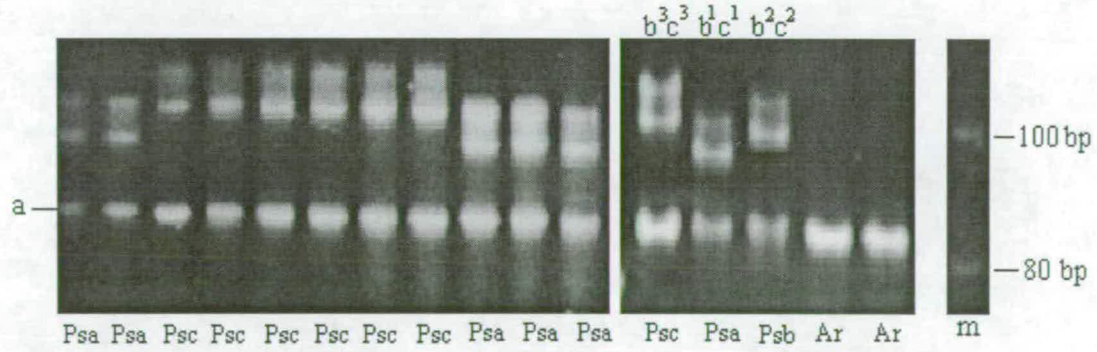


Figure 3.7. Microsatellite locus CH01H01 primer pair phenotypes; Ar = *Sorbus arranensis*, phenotype H01A; Psa = *S. pseudofennica*, phenotype H01B; Psb = *S. pseudofennica*, phenotype H01C; Psc = *S. pseudofennica*, phenotype H01D. m = 20 bp marker.

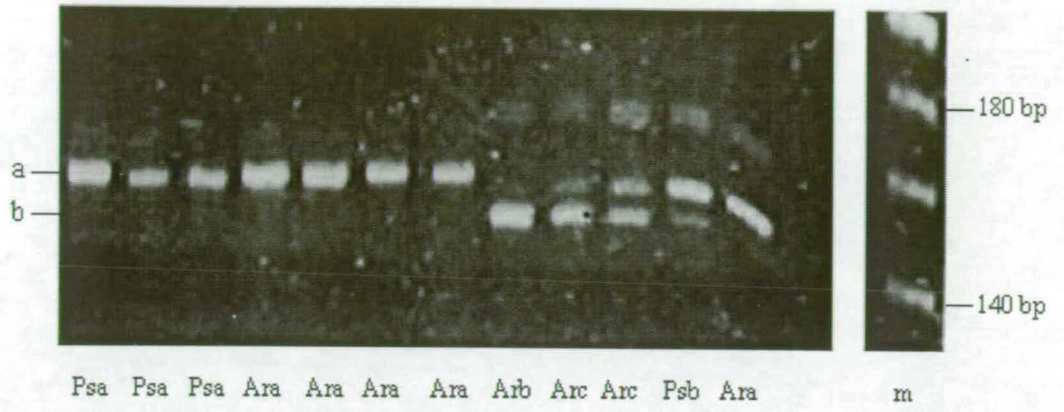


Figure 3.8. Microsatellite locus 1204-5 primer pair phenotypes; Ara = *Sorbus arranensis*, phenotype 1204A; Arb = *Sorbus arranensis*, phenotype1204B; Arc = *Sorbus arranensis*, phenotype 1204C; Psa = *S. pseudofennica*, phenotype1204A; Psb = *S. pseudofennica*, phenotype 1204C. m = 20 bp marker.

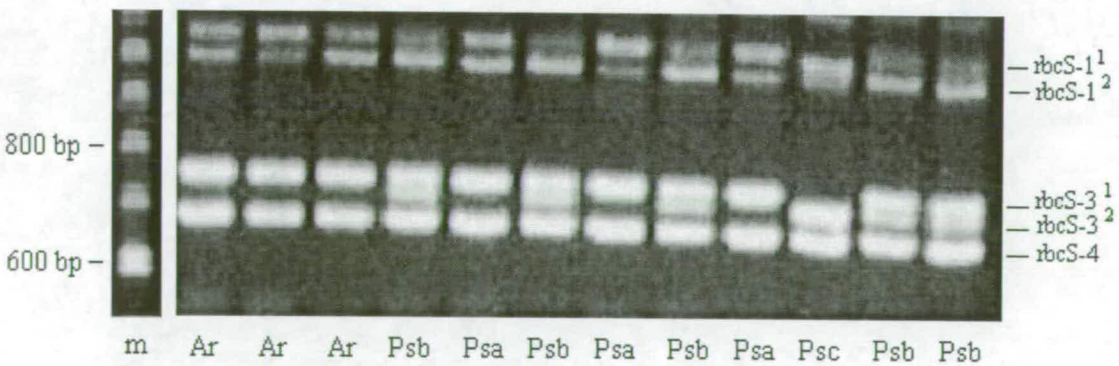


Figure 3.9. Rubisco intron locus primer pair phenotypes; Ar = *Sorbus arranensis*, phenotype rbcSA; Psa = *S. pseudofennica,* phenotype rbcSA; Psb = *S. pseudofennica* phenotype rbcSB; Psc = *S. pseudofennica*, phenotype rbcSC. m = 100 bp marker

Table 3.3. Frequency of *S. arranensis* and *S. pseudofennica* nuclear DNA primer pair phenotypes.

Primer pair	Phenotype	*S. arranensis* (n = 179)	*S. pseudofennica* (n = 140)
CH02D12	D12A	1.00	1.00
CH02D11	D11A	1.00	0.07
CH02D11	D11B	0.00	0.93
CH02B3b	B3A	1.00	0.99
CH02B3b	B3B	0.00	0.01
CH01H01	H01A	1.00	0.00
CH01H01	H01B	0.00	0.88
CH01H01	H01C	0.00	0.03
CH01H01	H01D	0.00	0.08
1204-5	1204A	0.95	0.99
1204-5	1204B	0.01	0.00
1204-5	1204C	0.04	0.01
rbcS	rbcSA	1.00	0.62
rbcS	rbcSB	0.00	0.37
rbcS-	rbcSC	0.00	0.01

3.4.3 *S. arranensis* and *S. pseudofennica* genotypic variation

3.4.3.1 Genotypic variation within S. arranensis and S. pseudofennica

The primer pair phenotypes were scored for each of the 179 *S. arranensis* and 140 *S. pseudofennica* individuals. The results were compiled to represent individual multilocus phenotypes, individuals with the same multilocus phenotype were classed as having the same genotype (or alternatively belonged to the same clone) (Table 3.4).

Three statistics were used to give an overall estimate of genotypic diversity within *S. arranensis* and *S. pseudofennica*; the proportion of distinct genotypes (Ng/Ni); overall genotype diversity (Hg); and overall genotype evenness (Eg).

From a total of 179 *S. arranensis* individuals screened only three genotypes were found, Ng/Ni = 0.016. The three genotypes each varied from the other at only a single primer-pair locus (Table 3.4). Overall clonal diversity (Hg) was 0.09, and clonal evenness (Eg) 0.36 (Table 3.5).

From a total of 140 *S. pseudofennica* individuals eight genotypes were found, Ng/Ni = 0.057. The degree of variation among the genotypes ranged from one to four loci (Table 3.4 and 3.6). Overall genotype diversity (Hg), was 0.63 for *S. pseudofennica* and overall genotype evenness (Eg) was 0.34 (Table 3.5).

S. arranensis genotype AA was distributed among each of the eight populations sampled, genotype AB was found in three populations (B, E and J), and genotype AC, consisting of only a single individual, was found in population J.

S. pseudofennica genotype PA was distributed among five populations (A,B,C,D, and E), genotype PB was restricted to population A, genotype PC was found in two populations (A and C), the remaining five *S. pseudofennica* genotypes were each restricted to a single population, PD in population B, PE in population A, PF in population C, PG in population B, and PH in population A.

3.4.3.2 Genotype diversity within populations

Genotype diversity was expressed in terms of: the number of clones per population; the proportion of populations that are multiclonal; population clonal diversity; and population clonal evenness.

The number of genotypes per population ranged between 1 and 3 with a mean of 1.6 for *S. arranensis* and between 1 and 5 with a mean of 2.6 for *S. pseudofennica* (Table 3.5 and 3.7). The proportion of multiclonal populations, was 0.5 for *S. arranensis* and 0.8 for *S. pseudofennica* (Table 3.5 and 3.7).

Clonal diversity estimates within each population (Hg/pop) ranged between 0 and 0.66 with a mean 0.15 for *S. arranensis* and between 0 and 0.58 with a mean of 0.31 for *S. pseudofennica* (Table 3.5 and 3.7). Hg/pop values of zero indicate a population composed of a single clone and a value of 1.0 would indicate that each individual sampled belongs to a different clone.

Clonal evenness (Eg/pop) within each population ranged between 0 and 1 with a mean of 0.3 for *S. arranensis* and between 0 and 0.47 with a mean of 0.25 for *S. pseudofennica* (Table 3.5 and 3.7). A value of zero would indicate a monoclonal population and a value of 1.0 would represent a multiclonal population (of infinite size) within which each clone is equally represented.

Table 3.4. Compilation of nuclear DNA primer pair phenotypes, representing *S. arranensis* and *S. pseudofennica* genotypes (clones).

	S. arranensis Genotypes			*S. pseudofennica* Genotypes							
Primer-pair	AA	AB	AC	PA	PB	PC	PD	PE	PF	PG	PH
	n=171	n=7	n=1	n=70	n=52	n=7	n=4	n=4	n=1	n=1	n=1
CH02D12 (D12)	A	A	A	A	A	A	A	A	A	A	A
CH02D11 (D11)	A	A	A	B	B	A	B	A	B	B	B
CH02B3b (B3)	A	A	A	A	A	A	A	A	B	A	B
CH01H01 (H01)	A	A	A	B	B	D	D	C	B	D	B
1204-5 (1204)	A	C	B	A	A	A	A	A	A	C	A
rbcS (rbcS)	A	A	A	A	B	A	A	A	A	A	C

Abbreviated primer pair name in brackets, corresponding to primer-pair phenotypes identified in Table 3.3.

Table 3.5. Genotypic diversity statistics for *S. arranensis* and *S. pseudofennica.*

Microspecies	Ng	Ng/ Ni	Hg	Eg	Ng/ pop	Multi-g pop	Hg/ pop	Eg/ pop	Pop/ g	Wide g	Priv g
S. arranensis	3	0.016	0.09	0.36	1.7	0.50	0.15	0.30	4.0	0.33	0.0
S. pseudofennica	8	0.057	0.63	0.34	2.6	0.80	0.31	0.25	1.5	0.12	0.37

Ni = total no. of individuals. Ng = total no of genotypes detected. Hg = overall clonal diversity. Eg = overall clonal evenness, Ng/pop = mean no of genotypes per population. Multi-g pop = proportion of populations that are multiclonal. Hg/pop = mean clonal diversity per population (where Ng > 1). Eg/pop = mean clonal evenness per population. Pop/g = mean no of populations per genotype. Wide g = proportion of genotypes that are widespread. Priv g = proportion of genotypes that are private.

Table 3.6. Number of primer pair loci that are variable between *S. pseudofennica* genotypes.

	PA	PB	PC	PD	PE	PF	PG	PH
PA	0	1	2	1	2	1	2	2
PB		0	3	3	3	2	3	2
PC			0	1	1	3	3	4
PD				0	2	2	1	3
PE					0	3	3	4
PF						0	3	1
PG							0	4
PH								0

Table 3.7. Genotypic diversity within *S. arranensis* and *S. pseudofennica* populations.

Population	*S. arranensis*			*S. pseudofennica*		
	Genotypes present	Hg	Eg	Genotypes present	Hg	Eg
A	1	0.00	0.00	5.00	0.50	0.40
B	2	0.45	0.90	3.00	0.58	0.47
C	1	0.00	0.00	3.00	0.48	0.38
D	-	-	-	1.00	0.00	0.00
E	2	0.09	0.50	1.00	0.00	0.00
I	1	0.00	0.00	-	-	-
J	3	0.66	1.00	-	-	-
L	1	0.00	0.00	-	-	-
M	1	0.00	0.00	-	-	-

3.4.3.3 Genotypic diversity between populations

Diversity between populations was expressed as: the number of populations in which a given clone occurred; the proportion of clones that were private; the proportion of clones that were widespread.

The number of populations in which a given clone occurred ranged between 1 and 8 with a mean of 4.0 for *S. arranensis* and between 1 and 5 with a mean of 1.5 for *S. pseudofennica* (Table 3.5 and 3.8). The number of private clones (those restricted to a single population, not including unique clones) was zero for *S. arranensis* and three for *S. pseudofennica* (PB, PD, PE). The number of widespread clones (defined as those occurring in more than 75% of the populations sampled) was one for *S. arranensis* (AA) and one for S. *pseudofennica* (PA).

Table 3.8. The number of populations in which a given clone occurred

Microspecies	Genotype	Number of populations
S. arranensis		
	AA	8
	AB	3
	AC	1
S. pseudofennica		
	PA	5
	PB	1
	PC	2
	PD	1
	PE	1
	PF	1
	PG	1
	PH	1

3.5 DISCUSSION

The nuclear DNA markers (six microsatellite primer pairs and the Rubisco intron primer pair) proved successful in detecting variation among individuals within *Sorbus arranensis* and *S. pseudofennica* microspecies. The DNA markers also detected species-specific marker variation between *S. arranensis* and *S. pseudofennica*.

3.5.1 Variation between S. arranensis and S. pseudofennica microspecies

The microsatellite marker CH01H01 proved to be a diagnostic marker that could unequivocally distinguish between *S. arranensis* and *S. pseudofennica* individuals (Fig. 3.7) and provided evidence that the differences observed between *S. arranensis* and *S. pseudofennica*, had a genetic basis which had been derived from different evolutionary events. Microsatellite marker CH02D11 provided supporting evidence with a band that was present in 92% of *S. pseudofennica* individuals but absent in 100% of *S. arranensis* individuals (Fig. 3.5). There was no evidence to support that genetic exchange between the two groups had occurred.

Propagation experiments on apomictic *Sorbus* taxa by Liljefors (1953) have shown that genetic exchange between aposporous obligate triploids (e.g. *S. arranensis* 2n =3x = 51) and facultative aposporous tetraploids (e.g. *S. pseudofennica* 2n = 4x = 68) could be possible. Liljefors (1953) found that from 24 progeny of the obligate triploid *Sorbus meinichii* (Lindeb.) Hedl. (2n = 3x = 51) 21 were triploid (2n = 3x = 51) and maternal in phenotype, two plants however, were non-maternal with one being tetraploid (2n = 4x = 68) and the other being pentaploid (2n = 5x = 85). Liljefors (1953) suggested that the two plants had resulted from the fertilisation of unreduced (2n =3x =51) maternal embryo sacs by male gametes, n =2x = 17 and n = 4x = 34 respectively. The male pollen donors were possibly the diploid *S. aucuparia* (2n = 2x = 34, n = 2x =17) and the tetraploid *S. hybrida* (2n = 4x = 68, n = 4x = 34) (cf. *S. pseudofennica*), both of which were growing near the maternal trees. The mother trees had very poor pollen quality (0%) so self-fertilisation was not possible.

3.5.2 Genotypic variation within S. arranensis

A total of 179 *S. arranensis* individuals, from eight populations, were screened with six nuclear DNA markers, only a single microsatellite primer pair (1204-5, Fig. 3.8) was successful in detecting variation among the *S. arranensis* individuals. Two bands in total were amplified (from primer-pair 1204-5) both of which were polymorphic, and these two bands were arranged into three combinations giving a total of three *S. arranensis* genotypes ((clones) AA, AB, and AC). 171 (95.5%) of the *S. arranensis* individuals were found to be genotype AA, seven individuals (3.9%) were genotype AB, and a single individual (0.5%) was genotype AC. Genotype AA was present in each of the eight populations, with genotype AB present in three, and genotype AC in one. These results do not support the hypothesis that each *S. arranensis* population consists of a single unique genotype, as population B (Diomhan Burn tributary) has two *S. arranensis* genotypes (AA, n = 11, AB, n = 5), population E (Catacol Burn) also has two genotypes (AA n = 20, AB n = 1) and population J (Alt Easan Biorach) has three genotypes (AA, n = 1, AB, n = 1 and AC, n = 1).

Ellstrand and Roose (1987) found the number of clones detected was generally more a function of the number of marker loci used, rather than sample size. It is a possibility therefore that with more marker loci used, the widespread *S. arranensis* genotype (AA) could be divided into genotypes that were each unique to a particular population. However, it would not alter the findings that three populations have more than one *S. arranensis* genotype present. Based on the present genotype distribution, it is more likely that if more *S. arranensis* genotypes were resolved the populations would have greater genotypic diversity within them and each population would be more similar to each other. This is in contrast to Hull and Smart's (1984) proposal, that each *S. arranensis* population consists of a single unique genotype.

The frequency distribution of these *S. arranensis* genotypes is comparable to that reported by Gornall (1999) who, in a review of the genetic structure of agamospermous microspecies (from six studies on the genus *Taraxacum* (Mogie 1985; Van Oostrum et. al. 1985; Ford and Richards 1985; Hughes and Richards 1989; Menken and Morita 1989; Battjes 1992), four studies on the genus *Rubus*

(Nybom and Schaal 1990; Kraft and Nybom 1995; Kraft et al. 1995; Kraft et al 1996), one study on the genus *Sorbus* (Proctor et al. 1989), previously unpublished data on the genus *Hieracium* (Gornall et al.) and 15 studies on agamospermous aggregates, where sampling was based on an agamic aggregate rather than individual agamospermous microspecies), found that where a microspecies was multiclonal there tends to be a dominant widespread genotype, with additional genotypes (1-4) at relatively low frequencies. Gornall (1999) also found that these low frequency genotypes usually differed from the dominant genotype by one allele at only one or two loci. These results are also comparable to the current study where *S. arranensis* genotypes AB and AC each differed from the dominant genotype (AA) at only one locus.

3.5.3 Origin of S. arranensis genotypic diversity

The observation that *S. arranensis* genotypes AB and AC differ from the dominant genotype (AA) at only one primer pair locus, implies that the differences between the genotypes have arisen as a result of mutation within the taxa, rather than having been derived from additional primary hybridisation events between the progenitors species *S. aucuparia* and *S. rupicola/S. aria*. If the genotypic diversity observed within *S. arranensis* was derived from additional primary hybridisation events, then the progenitor species *S. aucuparia* and *S. rupicola/S. aria* could only be polymorphic at the single locus that distinguish the three *S. arranensis* genotypes, and would have to be monomorphic at the other five loci scored. In theory this could of course be possible, especially within *S. rupicola*, which itself is a rare apomict. However, this is unlikely within *S. aucuparia,* which is a widespread sexual outcrosser. To support this hypothesis, ten *S. aucuparia* genotypes (Fig. 3.10) from a total 13 individuals (sampled from just two populations) were detected using just a single microsatellite prime-pair (CH01H01) and it could be predicted that if the 13 *S. aucuparia* individuals were screened with the additional primer pairs (used to screen *S. arranensis* and *S. pseudofennica* populations) then each individual would have a unique genotype. It is therefore likely that any additional primary hybridisation events between *S. aucuparia* and *S. rupicola* would be reflected in the additional *S.*

arranensis clones having significantly different genotypes, not just varying at a single locus.

The conclusion that the source of variation found between the *S. arranensis* genotypes was derived from within clone mutation events, as opposed to additional primary hybridisation events or selfing (as it is likely that the pollen viability is zero (Liljefors 1955)), is supported by studies on *Taraxacum* (Menken et al. 1995) and *Hieracium* (Shi et al. 1996). Both of these studies concluded that rare genotypes found within agamospermous microspecies, that differed from the principal genotype by only one or two alleles, were likely to have been derived from mutation events from within the microspecies rather that sexual events with progenitor species.

3.5.4 Genotypic variation within S. pseudofennica

A total of 140 *S. pseudofennica* individuals from five populations were screened with the same set of six nuclear DNA markers that were used to screen the *S. arranensis* samples. Five of these markers were successful in detecting variation among the *S. pseudofennica* individuals. From these five primer pairs eight bands were polymorphic, and were arranged into eight combinations giving a total of eight *S. pseudofennica* genotypes (PA, PB, PC, PD, PE, PF, PG, PH). 70 (50%) of the *S. pseudofennica* individuals were found to be genotype PA, 52 individuals (37.1%) were genotype PB, seven (5%) were genotype PC, genotypes PD and PE each had four (2.8%) individuals and the remaining three genotypes (PF, PG, PH) each had a single individual. Genotype PA was present in each of the five *S. pseudofennica* populations, genotype PC was present in two, and each of the remaining genotypes were restricted to a single population. These results do not support the hypothesis that each *S. pseudofennica* population consists of a similar set of genotypes with each genotype being freely distributed among the populations. Only a single genotype (PA) is distributed among each of the five populations.

If more marker loci were used and more genotypes detected as a result, the pattern of distribution would still be unlikely to reflect that proposed by Hull and Smart (1984). Based on the present distribution any additional *S. pseudofennica* genotypes found would tend to be restricted to single populations. Each population would therefore

consist of its own set of genotypes (with the exception of one or two individuals being dispersed), rather than populations that each consist of a similar set of genotypes.

The frequency distribution of *S. pseudofennica* genotypes differs from that of *S. arranensis*. *S. pseudofennica* has two dominant genotypes PA and PB with 50% and 37% of the total population respectively. The similarity with *S. arranensis* lies with there being a number of low frequency genotypes present; these six remaining genotypes represent only 13% of the total *S. pseudofennica* population.

3.5.5 Origin of S. pseudofennica genotypic diversity

The number of variable loci that separate the *S. pseudofennica* genotypes ranged from one to four (Table 3.6). Where two genotypes are separated from each other by only one locus the origin of this genotypic variation was likely to have been derived from a single mutation event within a clone. The genotypic variation at only one locus was unlikely to have been derived from an additional hybridisation event between genetically different *S. arranensis* and *S. aucuparia* individuals, as it would require the progenitor species *S. arranensis* and *S. aucuparia* to be polymorphic only at the locus that distinguishes the genotypes and would have to be monomorphic at the other loci scored. This is likely to be the case within *S. arranensis*, however it is unlikely for *S. aucuparia*, as it is likely that each individual represents a unique genotype (section 3.5.3). It is likely therefore that any additional primary hybridisation events between *S. arranensis* and *S. aucuparia* would be reflected in the additional *S. pseudofennica* clones having significantly different genotypes, not just varying at one locus.

Additional hybridisation events between genetically distinct *S. aucuparia* individuals and genetically invariant individuals of *S. arranensis* could explain the origin of variation found between *S. pseudofennica* genotypes that vary at four loci. An alternative hypothesis to explain the origin of relatively high genotypic variation among *S. pseudofennica* genotypes could be the proposed facultative apomictic nature of *S. pseudofennica* and a limited pollen viability of 17-20% (cf. *S. hybrida*,

Liljefors 1953). Aposporous facultative apomicts produce two types of embryo sac; unreduced embryo sacs which subsequently form parthenogenic embryos without fertilisation (any observed variation between the maternal and offspring genotypes will most likely have been derived via mutation); and reduced embryo sacs that require fertilisation before embryo formation (any observed variation between the maternal and offspring genotypes could be derived either through mutation or through genetic recombination).

The hypothesis that the origin of variation among *S. pseudofennica* clones is likely to be derived from a mixture of additional hybridisation events between genetically distinct *S. aucuparia* individuals and genetically invariant *S. arranensis* individuals, and from facultative intrapsecific sexual events among *S. pseudofennica* genotypes, is supported by Bayer (1990), Yhara et al. (1991) and Noyes and Soltis (1996). Bayer (1990), found a high number of genotypes among obligate and facultative agamic populations of *Antennaria rosea*, it was proposed that the high genetic diversity found within the related sexual taxa was the most likely origin of this variation. On a similar theme Yahara et.al (1991) explained that the low diversity within apomictic populations of *Eupatorium altissimum* could be directly related to the low genetic diversity with the sexual populations of *Eupatorium altissimum*. Noyes and Soltis (1996) identified 24 unique agamospermous *Erigeron compositus* genotypes distributed among seven populations. The geographical and ecological separation of the sexual and agamospermous populations made it unlikely that the origin of observed genotypic variation had come from geneflow via sexual populations. It was proposed that crossing between genetically distinct facultative agamosperms was the most likely source of the observed variation.

3.5.6 Comparisons of genotypic diversity

In general the genotypic diversity values observed for *S. pseudofennica* were greater than those found for *S. arranensis* (Table 3.5). The origins of genotypic diversity and the different reproductive mechanisms employed by *S. arranensis* and *S. pseudofennica* can provide an explanation for the observed variance in genotypic diversity between the two taxa. The origin of *S. arranensis* genotypes was likely to

be monophyletic and subsequent obligate agamospermous reproduction, contributing only a minimal amount of genotypic variation. In contrast the origin of *S. pseudofennica* genotypes was potentially polyphyletic and subsequent facultative agamospermous reproduction, recombining this genotypic diversity into new clonal lineages.

The two taxa have comparable results for overall clonal evenness (EG) *S. pseudofennica* 0.34 and *S. arranensis* 0.36 and for mean clonal evenness per population (EG/pop) *S. pseudofennica* 0.25 and *S. arranensis* 0.30. The values reflect that both taxa have one or two dominant genotypes that represent the majority of individuals and a number of genotypes each containing only a small number of individuals. This pattern is reflected at the species level and within the multiclonal populations.

Gornall (1999) found that approximately 50% of all microspecies studied consisted of only a single genotype (12 microspecies of *Taraxacum* (41%), 10 microspecies of *Rubus* (50%), 18 microspecies of *Hieracium* (56%) and five microspecies of *Sorbus* 62%). The mean number of genotypes (Ng), for *Rubus* was 2.1 (range 1-7), *Taraxacum* 5.0 (1-64), *Hieracium* 1.8 (1-7), and *Sorbus* 1.4 (1-2). The proportion of clones detected within each population (Ng/pop) for *Rubus* was 1.2 (range 1-3.5), *Taraxacum* 2.9 (1-16), *Hieracium* 1.2 (1-2) and *Sorbus* 1.0. The mean proportion of clones detected (Ng/Ni) for *Rubus* was 0.227 (range 0.07-0.33), *Taraxacum* 0.269 (0.017-1.0), *Hieracium* 0. 241 (0.006-1.0) and *Sorbus* 0.145 (0.037-0.5).

As noted by Gornall (1999) care must be taken when comparing and interpreting the results of studies where a variety of different marker systems, sample sizes and geographical scales were used to collect the data. However, Gornall (1999) found the results from the four genera reviewed were not dissimilar. The results from the current study do not vary greatly from Gornall (1999), the genotypic diversity values of *S. arranensis* and *S. pseudofennica* are placed within (or near) the range of values for each of the four genera reviewed by Gornall (1999) (Table 3.9)

Genotypic diversity comparisons between agamospermous and sexual taxa

As discussed in Chapter One natural selection acts on additive genetic variance between individuals. Genotype diversity (as opposed to gene diversity) within populations is in evolutionary terms perhaps the most important statistic to observe. As a consequence of sex and recombination there is likely to be large differences in the way genetic variation is arranged into multilocus genotypes between sexual taxa and agamospermous taxa, with the number of genotypes likely to be small in apomictic taxa and large in sexual taxa. For example, Proctor et al. (1989) detected single genotypes for *S. intermedia* (Ehrh.) Pers., *S. rupicola*, *S. subcuneata* Wilmott and *S. vexans* and two genotypes for *S. anglica* (Ng/Ni = 0.143), *S. devoniensis* (Ng/Ni = 0.08), and *S. porrigentiformis* (Ng/Ni = 0.154). In contrast, 12 genotypes from a total of 20 individual plants sampled (Ng/Ni = 0.6) were detected for the sexual species *S. aria* and six genotypes from a total of 16 individuals (Ng/Ni = 0.375) for the sexual species *S. aucuparia.*

Antonius and Nybom (1994) found (using DNA hybridisation probes, M13 and (AC)/(TG)) only a single genotype within the apomictic *Rubus nessensis* Hall from a total of nine sampled plants (Ng/Ni = 0.11) distributed among four populations. In comparison (using the same set of markers) 24 genotypes of the sexual *Rubus idaeus* L.were found from a total of 24 individual plants (Ng/Ni = 1.0) sampled from one of the four populations used in the *Rubus nessensis* assay.

Noyes and Soltis (1996), using eight enzyme systems, found 24 *Erigeron compositus* genotypes from a total of 231 individual plants sampled (Ng/Ni = 0.10). In contrast Ng/Ni = 0.61 for the sexual *E. compositus* populations studied (Noyes and Soltis 1995).

Hughes and Richards (1988) used ten isozyme systems to analyse the genetic structure of three apomictic *Taraxacum* microspecies, each from a single population and three sexual outbreeding *Taraxacum* populations. *T. unguilobum* Dahlst. had only a single genotype (from a total of 30 plants sampled Ng/Ni = 0.033). *T. brachyglossum* Dahlst. had only a single genotype (from a total of 38 plants sampled

Ng/Ni = 0.026) *T. pseudohamatum* Dahlst. had three genotypes (from a total of 45 plants sampled Ng/Ni = 0.06). In contrast, the sexual *Taraxacum* taxa had 102 genotypes sampled from 146 plants distributed among three populations (Ng/Ni =0.698 and Ng/pop = 34).

Menken et al. (1995), using two enzyme systems, found 57 *Tarxacum* section *Ruderalia* genotypes from a total 374 individual plants (Ng/Ni = 0.15) distributed among 11 populations. 47 genotypes were found in the Sexual *Taraxacum* section *Ruderalia* from a total of 158 individual plants (Ng/Ni = 0.29) distributed among four populations. The smaller difference in Ng/Ni values between the agamospermous taxa and sexual taxa for this study compared with the previous ones above may be explained by the aggregate nature of the agamospermous group studied.

The results from the current study are consistent with those described above; the proportion of genotypes (Ng/Ni) detected for *S. arranensis* and *S. pseudofennica* was 0.016 and 0.057 respectively, compared with 0.769 for the sexual outcrossing relative *S. aucupaia.*

The genotypic variation observed in the above agamospermous taxa when compared with related sexual taxa, in relative terms is very low and possibly from an evolutionary perspective ultimately restrictive.

Table 3.9. Genotype diversity statistics for *S. arranensis S. pseudofennica* and agamospermous genera *Rubus Taraxacum* and *Hieracium* (adapted from Gornall 1999).

Microspecies	*Ng*	*Ng/ Ni*	*Hg*	*Eg*	*Ng/ pop*	*Multi-g pop*	Hg/ pop	*Eg/ pop*	*Pop/ g*	*Wide g*	*Priv g*
S. arranensis	3	0.016	0.09	0.36	1.7	0.50	0.15	0.30	4.0	0.33	0.00
S. pseudofennica	8	0.057	0.63	0.34	2.6	0.80	0.31	0.25	1.5	0.12	0.37
Rubus	2.1	0.227	0.43	0.62	1.2	0.10	0.74	0.89	3.8	0.22	0.36
Taraxacum	5.0	0.269	0.55	0.69	2.9	0.51	0.50	0.73	2.9	0.14	0.69
Hieracium	1.8	0.241	0.39	0.66	1.2	0.10	0.38	0.73	2.2	0.30	0.55
Sorbus	1.4	0.145	0.31	0.76	1.0	0.00	na	na	3.6	0.50	0.33

Ni = total no. of individuals. Ng = total no of genotypes detected. Hg = overall clonal diversity. Eg = overall clonal evenness, Ng/pop = average no of genotypes per population. Multi-g pop = proportion of populations that are multiclonal. Hg/pop = mean clonal diversity per population (where Ng > 1). Eg/pop = mean clonal evenness per population. Pop/g = mean no of populations per genotype. Wide g = proportion of genotypes that are widespread. Priv g = proportion of genotypes that are private.

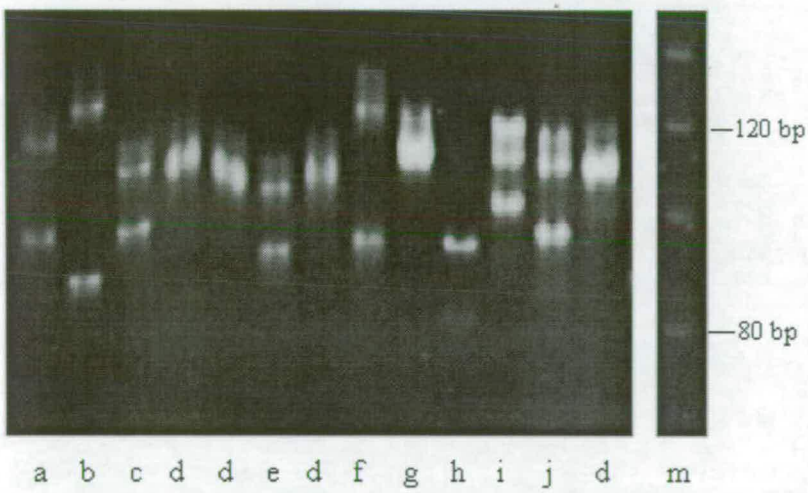


Figure 3.10. Banding patterns for microsatellite locus CH01H01. Ten *S. aucuparia* genotypes represented by the letters a–j. Genotype d is the only replicated genotype among the 13 individuals.

Chapter Four

Breeding System

4.1 INTRODUCTION

As discussed in Chapter One agamosperms have the potential to fix a relatively fit genotype (through heterosis or the generation of novel or extreme phenotypes) which can be replicated faithfully, generation after generation, without fear of being disrupted by recombination. Such a genotype has the potential to survive as long as its environment does not change beyond its tolerance range. It is unlikely that without genetic variation and recombination it will adapt to any environmental change that does occur.

Results in Chapter Three (section 3.4) have shown that populations of *S. pseudofennica,* and to a lesser extent *S. arranensis*, are made up of more than one clone and there is potential for genotype flow between the populations. In theory at least, natural selection has genotypic diversity on which to act and the problems associated with genetic drift could potentially be addressed (with the movement of these genotypes among the populations). Therefore, it should be possible for the populations to evolve. However, the definition of a population as an agamic complex may not be the same as in a sexual population, if the genotypes of an agamic population reproduce via obligate agamospermy there will be no genetic recombination between them. Thus, although an agamic complex can show considerable diversity within a location, the diversity within the population gene pool (the diversity on which natural selection acts) will be very low (Hughes and Richards 1989).

Recombination is thought to provide a mechanism allowing populations to maintain a balance between moving towards an optimum for an adaptive trait while still maintaining sufficient additive genetic variation to move towards a future optimum. Feldman et al. (1980) and Kondrashov (1982) suggested that even though the immediate effect of recombination is to reduce mean fitness, as it breaks up the gene combinations favoured by selection, this is counterbalanced by the increased variance in fitness, which in the longer term increases the population mean fitness by accelerating the response to selection.

To gain an understanding of the evolutionary potential of an agamospermous population it will be necessary to determine the reproductive nature of the agamospermous taxa. Obligate apomicts are not capable of producing reduced embryo-sacs whereas facultative apomicts are capable of producing both reduced and unreduced embryo-sacs. For example, within an obligate aposporous apomict (e.g. *Sorbus rupicola*, Liljefors (1953)) each (without exception) of the potentially meiotic megaspores within the ovule degenerate and are replaced with unreduced somatic cells derived from nucellar tissue. These unreduced somatic cells form the embryo sacs that subsequently give rise (without fertilisation) to a parthenogenic embryo (Liljefors 1953). Within a facultative aposporous apomict (*Sorbus hybrida*, Liljefors (1953)) a substantial proportion of the megaspores will not degenerate and will subsequently undergo meiosis to produce embryo-sacs that will require fertilisation before an embryo is formed. Within ovules where megaspores do degenerate they are replaced with unreduced somatic cells from which parthenogenic embryos arise in the same manner as those produced by obligate taxa.

Little is known about the reproductive mechanisms employed by the Arran *Sorbus* microspecies. However cytological studies carried out by Liljefors (1953, 1955) on analogous Scandinavian agamospermous taxa, have shown that the triploid *S. arranensis* group, comprising five apomictic taxa (*S. arranensis* Hedl., *S. neglecta S. subpinnata.*, *S. lancifolia.* and *S. lancifolia* f. *sognenis* Hedl.) are aposporous obligate agamosperms and the tetraploid *S. hybrida* is an aposporous facultative agamosperm.

4.2 Objectives

The observation that *S. arranensis* genotypes AB and AC differ from the dominant genotype (AA) at only one primer pair locus (Chapter Three), implies that the differences between the genotypes have arisen as a result of mutation within the taxa, rather than having been derived from additional primary hybridisation events between the progenitors species *S. aucuparia* and *S. rupicola/S. aria.*

The number of variable loci that separated the eight *S. pseudofennica* genotypes ranged from one to four (Chapter Three). Where two genotypes are separated from

each other by only one locus, the origin of this genotypic variation was likely to have been derived from a single mutation event within a clone. However, the origin of the variation found between *S. pseudofennica* genotypes that vary at four loci may reflect additional primary hybridisation events between genetically distinct *S. aucuparia* individuals and genetically invariant individuals of *S. arranensis*. An alternative hypothesis to explain the origin of relatively high genotypic variation among *S. pseudofennica* genotypes could be explained by the proposed facultative nature of *S. pseudofennica* and a limited pollen viability of 17-20% (cf. *S. hybrida*, Liljefors 1953).

The origins of genotypic diversity and the potentially different reproductive mechanisms employed by *S. arranensis* and *S. pseudofennica* could provide an explanation for the greater genotypic diversity values observed for *S. pseudofennica* (Chapter Three). The origin of *S. arranensis* genotypes was likely to be monophyletic and subsequent obligately agamospermous reproduction contributing only a minimal amount of genotypic variation. In contrast, the origin of *S. pseudofennica* genotypes was potentially polyphyletic and subsequent facultative agamospermous reproduction recombining this genotypic diversity into new clonal lineages.

The objective is to examine, using molecular markers, whether the observed genotypic variation in the Arran *Sorbus* microspecies is consistent with the putative reproductive mechanisms employed by *S. arranensis* (obligate agamospermy) and *S. pseudofennica* (facultative agamospermy).

4.3 MATERIALS AND METHODS

4.3.1 Plant material

Mature berries were collected from a total of nine *Sorbus arranensis* trees from three sites A, M and I, and from eight *S. pseudofennica* trees from one site, site A (Table 4.1). As the clonal composition of the *Sorbus* microspecies had not been established, at this stage, sample trees were chosen at random. Berries were collected during the first week of October 1996. Seeds were removed from the fruit, placed in brown envelopes and stored at 4 °C awaiting DNA extraction.

4.3.2 DNA extraction and quantification

Seeds were left to soak overnight, in sterile distilled water, allowing them to re-hydrate before DNA extraction. After soaking the seed embryo was carefully separated (using a scalpel and dissecting microscope) from the endosperm. At this stage it was possible to distinguish between developed and undeveloped embryos. A count was made of each type for both *S. arranensis* and *S. pseudofennica*. DNA was extracted from the embryos using the Scotlab Phytopure Plant DNA extraction kit (Scotlab 1996). DNA quantification was carried out as described in Chapter Two (section 2.3.4).

4.3.3 DNA amplification

Three nuclear DNA markers were used to asses the extent of genetic variation within the family groups.

Two microsatellite primer pairs (CH02D11, CH01HO1) and the Rubisco intron primer pair were selected from the set of six nuclear markers that were used to assess the clonal composition of the *Sorbus* microspecies (Chapter Three, section 3.4). These primer pairs were chosen as they were the three most polymorphic of the set. PCR reactions and resolution of PCR products were carried out as described in Chapter Three (section 3.3.3).

Table 4.1. Location of *S. arranensis* and *S. pseudofennica* seed material.

Tree	Site location	Tree	Site location
S. arranensis		*S. pseudofennica*	
D4	A	D1	A
D10	A	D2	A
D14	A	D3	A
D18	A	D5	A
D20	M	D7	A
22	M	D8	A
24	M	D9	A
28	I	D12	A
30	I		

Site locations referenced in Chapter Two

Table 4.2. Number of developed and undeveloped embryos found within *S. arranensis* and *S. pseudofennica* families.

	Undeveloped	Developed	Total
S. arranensis	49	59	108
S. pseudofennica	23	73	96
Total	72	132	204

$\chi^2_{(1)} = 10.19$ $P < 0.01$

4.4 RESULTS

There was a significantly lower number of undeveloped embryos found within the *S. pseudofennica* families (24%) compared with the number found in S. *arranensis* families (45%) ($P<0.01$) (Table 4.2).

Six *S. arranensis* and six *S. pseudofennica* families (a family consisting of the mother tree and between eight and 12 offspring) were screened with two microsatellite primer pairs (CH02D11, CH01H01) and the Rubisco intron primer pair. The interpretations of the gel banding patterns for each primer pair were as described in Chapter Three (section 3.4) and the same primer pair phenotype notations, as used to assess the clonal composition of the *Sorbus* microspecies, were adopted (Chapter Three, Table 3.3).

For the six *S. arranensis* families screened with the above primer pairs, no segregation was found (Table 4.3). These results were consistent with the results from the screening of the 179 *S. arranensis* adults, where no variation in banding patterns, between the individuals, was observed (Chapter Three, Table 3).

Segregation was found in three of the six *S. pseudofennica* families (D3, D8 and D9) (Table 4.4). Variation in banding patterns between the 140 *S. pseudofennica* adults, screened with the same three primer pairs, was also found (Chapter Three, Table 3.3).

Within each segregating *S. pseudofennica* family, 3 out of 11 (D3 and D9) and five out of ten (D12) individuals shared differences in banding pattern from the maternal tree. For ten of these 11 individuals, differences were represented by a loss of a band(s) (in comparison to the maternal phenotype) (Figs 4.1-4.5). The loss of a band(s) within four of these individuals created a novel phenotype not seen in any *S. arranensis* or *S. pseudofennica* adult and for the remaining six, the loss of a band created a phenotype consistent with that of the common *S. arranensis* type. For the odd individual the difference, between the offspring and maternal phenotype, was

represented by a variation in band length at two positions (Fig. 4.2). This length polymorphism (primer pair phenotype H01C) was also found in five *S. pseudofennica* adults (Chapter Three, Table 3.3).

4.4.1 Nuclear DNA primer pair phenotypes

Microsatellite locus CH02D11

For *S. arranensis* only a single phenotype was observed and no variation was found within or between any of the families. Each of the offspring, and each mother, shared the same fixed six-banded phenotype, D11A. This microsatellite phenotype was also present in 179 *S. arranensis* and 10 *S. pseudofennica* adults (Chapter Three, Fig. 3.5).

Five phenotypes were observed for *S. pseudofennica.* D11A (as described above), D11B (a seven-banded phenotype also found in 130 *S. pseudofennica* adults (Fig. 4.1)), D11C (a novel five-banded phenotype not found in any of the adults (Fig. 4.1)) and types D11D and D11E, (two novel four-banded phenotypes not found in any of the adults (Fig 4.1)).

For families D1, D2 and D8 each of the offspring, and each mother, shared the same, seven-banded, type D11B, phenotype. Family D3 had four phenotypes, 8 individuals of type D11B, one individual of type D11C, and one individual of types D11D and D11E. The phenotype of the D3 mother was D11B (Fig. 4.1). Family D9 had two phenotypes, 11 individuals of type D11B and one of type D11C. The phenotype of the D9 mother was D11B. Family D12 had two phenotypes, five individuals of type D11A and five individuals of type D11B. The phenotype of the D12 mother was type D11B.

Microsatellite locus CH01H01

For *S. arranensis* only a single phenotype was observed, no variation was found within or between any of the families. Each of the offspring, and the mother, shared

the same fixed single-banded phenotype, H01A. This microsatellite phenotype was also present in 179 *S. arranensis* adults (Chapter Three, Fig. 3.7).

Four phenotypes were observed for *S. pseudofennica.* Type H01A (as described above), H01 B (a three-banded phenotype found in 123 *S. pseudofennica* adults) (Fig. 4.2), type H01C (a three-banded phenotype found only in five *S. pseudofennica* adults) (Fig. 4.2) and type H01E (a novel single-banded phenotype not found in any of the adults) (Fig.4.2).

For families D1, D2 and D8 each of the offspring, and each mother, shared the same, three-banded, type H01B, phenotype. Family D3 had three phenotypes, 9 individuals of type H01B, one individual of type H01C, and two individuals of type H01E. The phenotype of the D3 mother was type H01B (Fig. 4.2). Family D9 had three phenotypes, one individual of type H01A, ten individuals of type H01B and one of type H01E. The phenotype of the D9 mother was type H01B (Fig 4.3). Family D12 had two phenotypes, five individuals of type H01A and five individuals of type H01B. The phenotype of the D12 mother was type H01B (Fig 4.4).

Rubisco intron locus

For *S. arranensis* only a single phenotype was observed, no variation was found within or between any of the families. Each of the offspring, and the mother, shared the same fixed four-banded phenotype, rbcSA. This Rubiso intron phenotype was also present in 179 *S. arranensis* and 87 *S. pseudofennica* adults (Chapter Three, Fig. 3.9). Three phenotypes were observed for *S. pseudofennica,* rbcSA (as described above), rbcSB (a five-banded phenotype found in 52 *S. pseudofennica* adults) (Fig 4.5), and type rbcSC a three-banded phenotype (found in only one *S. pseudofennica* adult) (Fig. 4.5).

For families D1, D2, D3, D8 and D12 each of the offspring, and each mother, shared the same, four-banded, type rbcSA, phenotype. Family D9 had two phenotypes, 11 individuals of type rbcSB, and one individual of type rbcSC. The phenotype of the D12 mother was type rbcSB (Fig. 4.5).

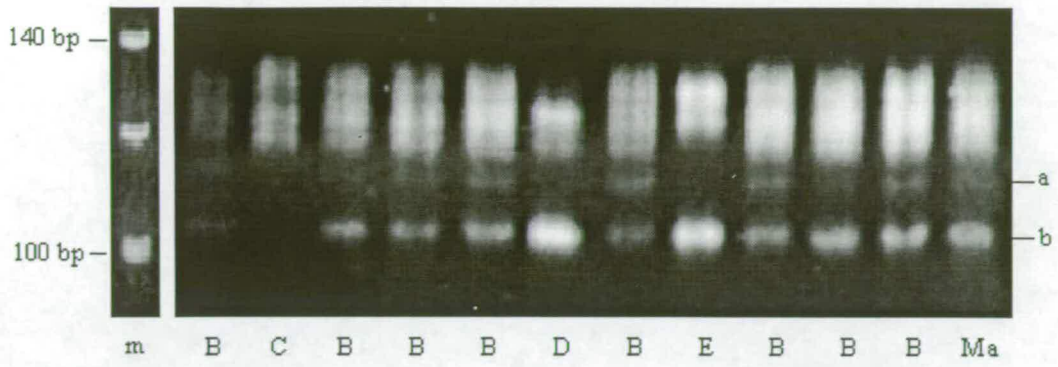


Figure 4.1. *S. pseudofennica* family D3. Maternal and offspring banding patterns for microsatellite locus CH02D11. D11B, D11C, D11D and D11E represent the four different phenotypes found within this family. Ma = maternal phenotype, m = 20bp ladder.

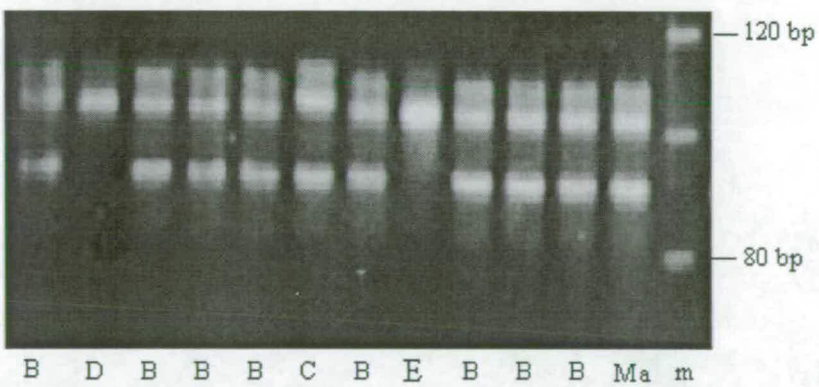


Figure 4.2. *S. pseudofennica* family D3. Maternal and offspring banding patterns for microsatellite locus CH01H01. H01B, H01C, and H01E represent the three different phenotypes found within this family. Phenotype C shows a band length polymorphism between the offspring and maternal phenotype. Ma = maternal phenotype, m = 20bp ladder.

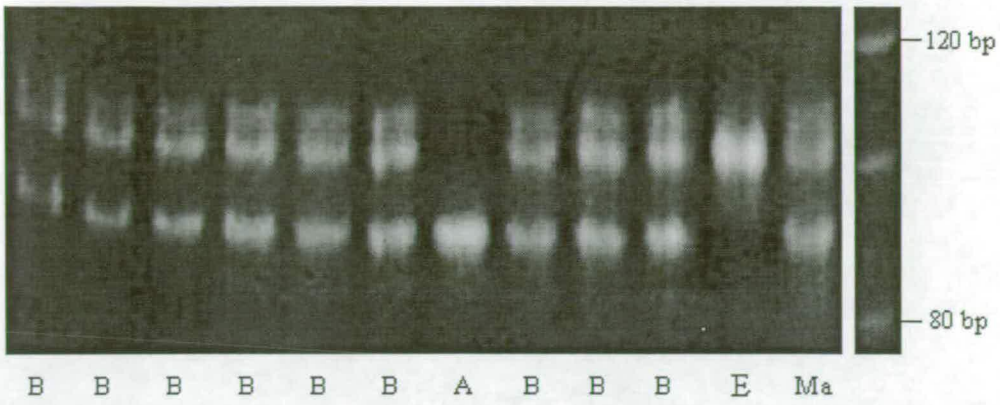


Figure 4.3. *S. pseudofennica* family D9. Maternal and offspring banding patterns for microsatellite locus CH01H01. H01A, H01B, and H01E represent the three different phenotypes found within this family. Ma = maternal phenotype, m = 20bp ladder.

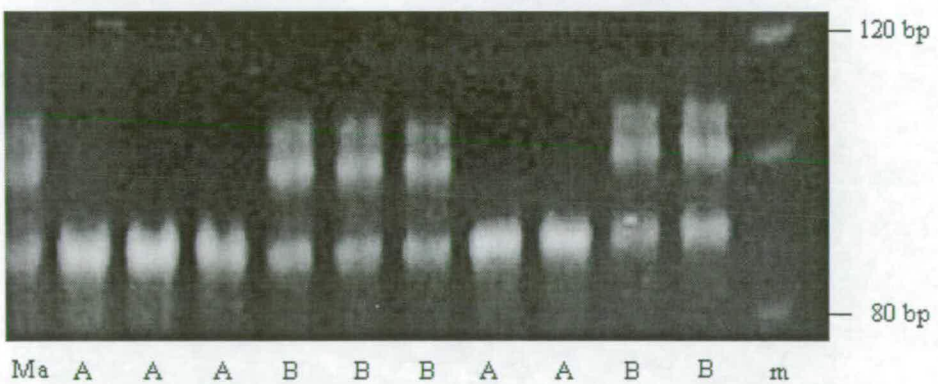


Figure 4.4. *S. pseudofennica* family D12. Maternal and offspring banding patterns for microsatellite locus CH01H01. H01A, and H01B = the two different phenotypes found within this family. Ma = maternal phenotype, m = 20bp ladder.

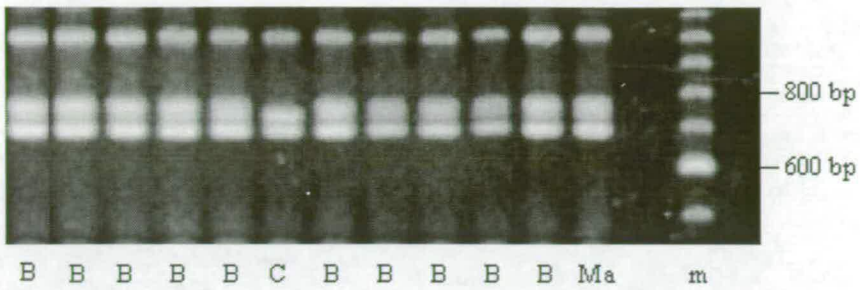


Figure 4.5. *S. pseudofennica* family D12. Maternal and offspring banding patterns for the Rubisco locus. rbcSB and rbcSC represent the two different phenotypes found within this family. Ma = maternal phenotype, m = 100 bp ladder.

Table 4.3 *Sorbus arranensis* maternal and offspring phenotypes for the microsatellite loci CH02D11, CH01H01 and the Rubisco intron locus.

Family	Locus Phenotype			Family	locus Phenotype		
	D11	H01	rbcS		D11	H01	rbcS
D10				**D14**			
Ma	A	A	A	Ma	A	A	A
1	A	A	A	1	A	A	A
2[u]	A	A	A	2[u]	A	A	A
3[u]	A	A	A	3	A	A	A
4[u]	A	A	A	4	A	A	A
5[u]	A	A	A	5	A	A	A
6	A	A	A	6[u]	A	A	A
7	A	A	A	7	A	A	A
8[u]	A	A	A	8	A	A	A
9	A	A	A				
D18				**22**			
Ma	A	A	A	Ma	A	A	A
1[u]	A	A	A	1	A	A	A
2[u]	A	A	A	2	A	A	A
3[u]	A	A	A	3u	A	A	A
4	A	A	A	4	A	A	A
5[u]	A	A	A	5	A	A	A
6[u]	A	A	A	6	A	A	A
7[u]	A	A	A	7	A	A	A
8	A	A	A	8	A	A	A
9u	A	A	A	9	A	A	A
10[u]	A	A	A	10	A	A	A
24				**30**			
Ma	A	A	A	Ma	A	A	A
1[u]	A	A	A	1[u]	A	A	A
2[u]	A	A	A	2[u]	A	A	A
3	A	A	A	3[u]	A	A	A
4	A	A	A	4[u]	A	A	A
5[u]	A	A	A	5[u]	A	A	A
6	A	A	A	6	A	A	A
7	A	A	A	7	A	A	A
8	A	A	A	8	A	A	A
9	A	A	A	9[u]	A	A	A
10[u]	A	A	A				

Ma = maternal phenotype
1,2,3 etc =offspring number.
D11= microsatellite locus CH02D11
H01= microsatellite locus CH01H01
rbcS = Rubisco intron locus
[u] = Undeveloped embryo
1,2,3 etc =offspring number.

Table 4.4. *Sorbus pseudofennica* maternal and offspring phenotypes for the microsatellite loci CH02D11, CH01H01 and the Rubisco intron locus.

Family	Locus Phenotype			Family	locus Phenotype		
	D11	H01	rbcS		D11	H01	rbcS
D1				**D2**			
Ma	B	B	A	Ma	B	B	A
1[u]	B	B	A	1[u]	B	B	A
2[u]	B	B	A	2	B	B	A
3[u]	B	B	A	3	B	B	A
4[u]	B	B	A	4	B	B	A
5[u]	B	B	A	5[u]	B	B	A
6	B	B	A	6[u]	B	B	A
7	B	B	A	7	B	B	A
8[u]	B	B	A	8	B	B	A
9	B	B	A	9	B	B	A
10	B	B	A	10[u]	B	B	A
11	B	B	A				
D3				**D8**			
Ma	B	B	A	Ma	B	B	A
1	B	B	A	1	B	B	A
2[u]	C*	E*	A	2	B	B	A
3	B	B	A	3	B	B	A
4[u]	B	B	A	4	B	B	A
5	B	B	A	5	B	B	A
6	D*	C	A	6	B	B	A
7	B	B	A	7	B	B	A
8	E*	E*	A	8	B	B	A
9	B	B	A	9	B	B	A
10	B	B	A	10	B	B	A
11	B	B	A				
D9				**D12**			
Ma	B	B	B	Ma	B	B	A
1	B	B	B	1[u]	A	A	A
2[u]	B	B	B	2[u]	A	A	A
3[u]	B	B	B	3	A	A	A
4	B	B	B	4	B	B	A
5[u]	B	B	B	5	B	B	A
6	C*	B	C	6	B	B	A
7	B	A	B	7[u]	A	A	A
8	B	B	B	8[u]	A	A	A
9	B	B	B	9	B	B	A
10	B	B	B	10	B	B	A
11	B	E*	B				

Ma = maternal phenotype
D11= microsatellite locus CH02D11
H01= microsatellite locus CH01H01
rbcS = Rubisco intron locus
[u] = Undeveloped embryo
1,2,3 etc =offspring number
* = novel phenotype

4.5 DISCUSSION

For the six *S. arranensis* families screened with the three nuclear DNA markers, no segregation of markers was found among the offspring and no variation in banding pattern was observed between any of the offspring and the maternal tree. The absence of diversity within the *S. arranensis* families was consistent with the absence of diversity among the 179 *S. arranensis* adults screened with the same DNA markers. In contrast, segregation of molecular makers was found in three out of the six *S. pseudofennica* families screened with the same set of markers. The presence of diversity within the *S. pseudofennica* families is consistent with there being diversity among 142 *S. pseudofennica* adults screened with the same three primer pairs. There was no association found among the number of undeveloped embryos and the occurrence of a non-maternal phenotype ($\chi_1^2 = 2.49$ n.s.).

The monomorphic banding patterns found within the families of *S. arranensis* are consistent with the hypothesis that *S. arranensis* is an obligate apomict not capable of producing reduced embryo sacs. The variable banding patterns, found within 50 % of the *S. pseudofennica* families, are consistent with the hypothesis that *S. pseudofennica* is a facultative apomict capable of producing both reduced and unreduced embryo sacs. These hypotheses were derived from cytological experiments performed by Liljefors (1953) on analogous Scandinavian *Sorbus* taxa, the triploid *S. arranensis* group (cf. *S. arranensis* on Arran) and the tetraploid *S. hybrida* (cf. *S. pseudofennica*). For the triploid *S. arranensis* group each (without exception) of the potentially meiotic megaspores, within the ovule, degenerated and were replaced with unreduced somatic cells derived from nucellar tissue. These unreduced somatic cells form the embryo-sacs, which could subsequently give rise (without fertilisation) to a parthenogenic embryo (Liljefors 1953). *S. hybrida* could produce meiotically reduced embryo-sacs (which would require fertilisation before an embryo was formed) from a substantial proportion of the megaspores that did not degenerate. Where the megaspores did degenerate, they were replaced with unreduced somatic cells from which parthenogenic embryos could arise in the same manner as those produced by the *S. arranensis* group (Liljefors 1953).

4.5.1 Obligate agamospermy

The absence of any variability within the *S. arranensis* families could possibly reflect obligate agamospermous reproduction. If the pollen viability of *S. arranensis* was zero (as was found to be the case for the analogous Scandinavian *S. arranensis* (Liljefors 1955)) then these results are unlikely to be the consequence of inbreeding. In turn, any variation found within the families would not necessarily represent a departure from obligate agamospermy; variable banding patterns could result from the fertilisation of unreduced *S. arranensis* embryo-sacs by *S. aucuparia* or *S. pseudofennica* pollen. For example Liljefors (1953) conducted progeny tests with the obligate apomict *Sorbus meinichii* (2n = 51) and found that from 24 progeny (from two mothers) 21 were triploid (2n = 51) and maternal. However, two plants were non-maternal with one being tetraploid (2n = 68) and the other pentaploid (2n = 85). Liljefors (1953) suggested that the two plants had resulted from fertilisation of unreduced (2n =51) eggs by male haploid gametes, n = 17- and n = 34 respectively. The male pollen donors were possibly the diploid *S. aucuparia* and the tetraploid *S. hybrida,* both of which were growing near the maternal trees. The mother trees had very poor pollen quality (0%) therefore self-fertilisation was not possible.

The results of current study vary with those from other studies on obligate agamosperms. For example the patterns of molecular marker variation vary among studies of obligate *Taraxacum* families (Mogie1985; Ford and Richards 1985; Hughes 1987; Battjes et al. 1992; reviewed in Richards 1996). Mogie (1985) using a single enzyme system (esterase) examined 232 seedlings from a single family of *Taraxacum unguilobum* and *T. pseudohamatum* Each of the offspring of *T. pseudohamatum* and all but one seedling of *T. unguilobum* had the same isozyme phenotype as the maternal genome. Hughes (1987) using 15 enzyme loci screened 56 seedlings from a single family of *T. pseudohamatum* and found no variation. Hughes (1987) did find among 31 siblings a single individual with a different esterase phenotype (differing at a single band) although it was shown not to be heritable as the next generation screened reverted back to the grandparental phenotype. Battjes (1992), using four isozymes systems to screen 11 families of the obligate apomict *Taraxacum hollandicum* Soest., found no variation within or among 11 families.

In contrast Battjes (1992) found that eight siblings had a non-parental phenotype among a family of 32 *T. vinobonense* Soest. offspring, and Ford and Richards (1985) found that out of 57 offspring grown from populations of *T. brachyglossum* and *T. lacistophyllum* Dahlst. 24 had non-parental phenotypes.

4.5.2 Facultative agamospermy

The presence of molecular marker variability within and among the *S. pseudofennica* families may reflect facultative agamospermous reproduction. If the pollen viability of *S. pseudofennica* was 20% (as was found in the analogous Scandinavian *Sorbus* hybrid *S. hybrida* (Liljefors 1955)), these results are possible via the fertilisation of a number of reduced embryo-sacs with either *S. pseudofennica* or *S. aucuparia* pollen. However, mutation, or possibly somatic recombination could generate the variation observed in the banding patterns. Automatic segregation is unlikely to explain the marker variation found in any unreduced embryo-sacs as the megaspores degenerate before being replaced by somatic cells. Automatic segregation may explain marker variation observed in meiotic diplosporous agamosperms, where the megaspores do not degenerate but undergo a restitutional meiosis before producing unreduced embryo-sacs.

The majority of non-maternal offspring banding patterns were represented by a loss of a band with respect to the maternal genotype. These non-maternal offspring could be derived from a sexual event with a pollen donor, not carrying any novel markers, or could also be derived from selfing, mutation or somatic segregation. For the remaining individual (with a non-maternal banding pattern) the variation was not in the number of bands but in the pattern of the bands. A novel marker was present, with respect to the maternal genotype that could have been derived via an outcrossing event with pollen from a *S. pseudofennica* tree carrying the novel marker. This novel marker was found in five *S. pseudofennica* genotypes, which could be the possible pollen donor.

When compared with other studies the results of the current are similar in that molecular marker variation found within facultative agamospermous families is not

uncommon (e.g. Mazzucato et al. 1995; Oritz et. al 1997; Nagib et al. 1998). However, the degree of variation can vary markedly. Nagib et al. (1998) used RAPD markers to study two putative facultative apomictic *Manihot esculenta* Crantz clones. 67 and 30 offspring from each clone were screened with 24 RAPD primers. Only a single individual from each progeny set had the maternal phenotype. Mazzucato et al. (1995) used 3 isozyme systems, six RAPD primers and morphological markers to screen (using controlled crosses) four families of the facultative *Poa pratensis* L. Each of the four families produced offspring with non-maternal phenotypes, The number of non-maternal offspring within each family varied with 15 %, 42 %, 58 %, and 96 % respectively. Oritz et al. (1997) used RFLP and RADP markers to identify non-maternal offspring from controlled crosses of the tetraploid apomictic *Paspalum notatum* Fluegge. 32 individuals from nine families were analysed. The number of non-maternal offspring within each family ranged from 3 % to 100%.

The results of the current study are consistent with the hypothesis that the putative reproductive mechanisms employed by *S. arranensis* and *S. pseudofennica*, have contributed to the greater genotypic diversity observed among *S. pseudofennica* genotypes.

Chapter Five

Novel Arran *Sorbus* Taxon

5.1 INTRODUCTION

During the field survey (described in Chapter Two) three individual *Sorbus* trees were encountered with a phenotype unlike any that has been previously recorded on Arran: the gross leaf morphology of these plants appears to be intermediate between *S. aucuparia* and *S. pseudofennica*. The leaves of *S. aucuparia* are pinnate with five to nine (normally six to seven) pairs of lateral leaflets with a single terminal leaflet equalling (but never larger than) the lateral leaflets (Stace 1991). The leaves of *S. pseudofennica* are oblong or ovate-oblong in outline, mostly with one to two pairs of free leaflets at the base and lobed above (Stace 1991). The free leaflets of *S. pseudofennica* may or may not be decurrent on the rachis (the lateral leaflets of *S. aucuparia* are never decurrent (Bolstad and Salvesen 1999)). The leaves of the novel Arran *Sorbus* were pinnate with four to five pairs of lateral leaflets. The terminal leaflet was lobed and much wider than the lateral leaflets. The leaves of the novel Arran *Sorbus* resemble *S. aucuparia* in leaf size, number of free leaflets and leaf form, but are more similar to *S. pseudofennica* in the relatively broader outline and terminal leaf (Figure 5.1). The leaflets tended not to be decurrent, especially the first four, the most terminal pair or pairs were in some cases and in others not.

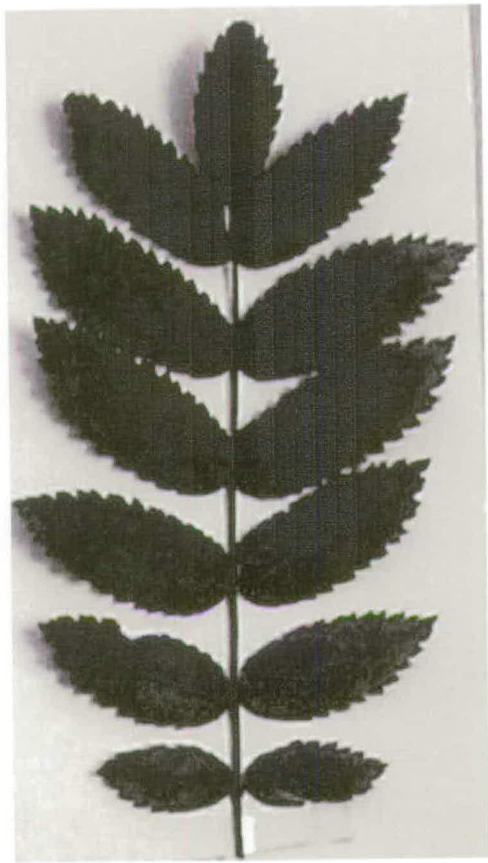
Lindenberg (in Hartman 1879, p. 271) was among the first to recognise a taxon intermediate between *S. aucuparia* and *S. hybrida* (cf. *S. pseudofennica*). Lindenberg (1879) considered it to be a form of *S. aucuparia* and gave it the name "*Sorbus Aucuparia* L. *Meinichii*"; the taxon was later ranked as a species by Hedlund (1901) *Sorbus meinichii* (Lindeb.) Hedl. The distribution of *Sorbus meinichii* is restricted to scattered localities in Norway (Lid and Lid 1994), Sweden and Finland (Hultén 1971; Hultén and Fries 1986).

Liljefors (1953), on the basis of morphology and cytological experiments, proposed that *S. meinichii* had been derived from a hybridisation event between *S. aucuparia* and *S. hybrida*. Liljefors's (1953) experiments also revealed that there were three forms of type *S. meinichii*, a tetraploid apomictic (to which the name was originally applied, a triploid apomict and a triploid sexual hybrid. Liljefors (1953) proposed that the name *S. meincihii* should be retained for the tetraploid and a new name

Sorbus teodori Liljef. given to the triploid form. Bolstad and Salvesen (1999) using morphometric analysis and cytology studied the relationships between the *S. meinichii* group, *S. hybrida* and *S. aucuparia.* They agreed with Liljefors (1953) that it was likely the *S. meinichii* forms had been derived from crosses between *S. aucuparia* and *S. hybrida,* but disputed the separation of the different ploidy levels into separate taxa.

5.2 Objectives

The novel phenotype suggests that previously unrecorded genetic exchanges have occurred within the Arran *Sorbus* complex and these genetic exchanges could have lead to the formation of new apomictic taxa. Without controlled crossing experiments and cytological analysis it will not be possible to establish the precise origin of the novel Arran phenotype. However, using the marker systems employed in Chapters Two and Three, it will be possible to test the hypothesis that the novel Arran *Sorbus* taxon was of hybrid origin, ultimately derived from the genomes of both *S. aucuparia* and *S. aria sensu lato*. It will be also possible to test the hypothesis that the novel phenotype is an Arran *Sorbus* hybrid with significantly different origins with respect to *S. arranensis* and *S. pseudofennica.*



Sorbus aucuparia

Novel Arran *Sorbus*

Sorbus pseudofennica

Figure 5.1. Leaf morphology for *Sorbus aucuparia*, novel Arran *Sorbus* and *S. pseudofennica* (scale: life size)

5.3 MATERIALS AND METHODS

Two of the novel Arran *Sorbus* plants were found at site E (H2 and H3) and the third was found at site C (H1) (Site references given in Fig. 2.1, Chapter Two). In total there were 23 *S. pseudofennica* and 20 *S. arranensis* trees at Site C and 45 *S. pseudofennica* and 24 *S. arranensis* trees at Site E. *S. aucuapria* trees were present at both sites.

Trees H1 and H2 are both old and damaged. H3 however, appeared to be a young tree, recently established. A small number of flowers and fruits were recorded on H2 but not on H1 or H3.

The complete set of marker systems (isozyme and DNA markers) employed in Chapter Two (Origin of Arran *Sorbus* microspecies) and Chapter Three (Clonal composition of Arran *Sorbus* microspecies) were used to screen the three novel phenotypes. Starch gel electrophoresis, DNA extraction and DNA amplification procedures were carried out exactly as described in Chapters Two and Three.

5.4 RESULTS

The three individuals of the putative novel Arran *Sorbus* taxon, H1 (site C), H2 and H3 (site E), were screened with three isozyme systems (LAP, AAT and 6PGD), five microsatellite primer pairs (CH02D12, CH02D11, CH02B3b, CH01H01 and 1204-5), the Rubisco intron primer pair and the chloroplast intergenic spacer primer pair e-f.

The interpretations of the gel banding patterns for each isozyme and DNA marker system are as described in Chapter Two (isozyme and Rubisco intron markers) and Chapter Three (microsatellite markers). The same genotype and primer pair phenotype notations were adopted (Chapter Two, Tables 2.4 and 2.5, Chapter Three Table 3.3).

5.4.1 Isozyme markers

Leucine aminopeptidase (LAP)

A single LAP locus (*Lap-1*) was observed for H1, H2 and H3. *S. arranensis* and *S. pseudofennica* individuals also had this homozygous *Lap-1* genotype.

Aspartate amino transferase (AAT)

H1, H2 and H3 each had the same heterozygous genotype, with alleles $Aat\text{-}1^2$ and $Aat\text{-}1^3$, and the heterodimeric band $Aat\text{-}1^2 1^3$. *S. arranensis* and *S. pseudofennica* also had this heterozygous genotype (Fig. 2, Chapter Two).

6- Phosphogluconate dehydrogenase (6PGD)

H1, H2 and H3 each had the same six-banded phenotype, with loci *6Pgd-1* and *6Pgd-3* being homozygous and loci *6Pgd-2* heterozygous (with allele $6pgd\text{-}2^1$ and the heterodimer $6pgd\text{-}2^1 2^2$) (Fig 5.2). This six-banded phenotype was unique to H1, H2 and H3 (with respect to *Sorbus aucuparia, S. rupicola, S. arranensis* and *S. pseudofennica*). *S. arranensis* and *S. pseudofennica* individuals had a seven-banded phenotype, with the six alleles of the novel taxon present. The additional *S. arranensis* and *S. pseudofennica* allele, *6Pgd-4,* not present in the novel hybrids, was derived from the parental type *S. rupicola* (Fig 3.3, Chapter Two).

Rubisco intron

H1, H2 and H3 each had the same three-banded phenotype, with homozygous loci *rbcS*-1 (with allele $rbcS\text{-}1^2$), *rbcS-3* (with allele $rbcS\text{-}3^1$), and *rbcS-4* (Fig. 3). This three-banded phenotype was unique to H1, H2 and H3. These three alleles were present in the four-banded phenotype of *S. arranensis,* and the phenotypes of *S. pseudofennica*. The additional *S. arranensis* and *S. pseudofennica* allele, $rbcS\text{-}1^2$, not present in the novel phenotype, was derived from the parental type *S. rupicola* (Fig. 2.4 Chapter Two).

5.4.2 Chloroplast DNA markers

Individual H2 was screened with chloroplast primer pairs c-d and e-f. Single PCR products were produced for both primer pairs, c-d (length 530 bp) and e-f (length 483 bp). Nucleotide sequences were determined by direct sequencing of the amplification products. The sequences of H2 matched exactly (100%) the sequences of *Sorbus arranensis*, *S. pseudofennica* ♂ *S. aucuparia* (full sequences Appendix 1).

5.4.3 Nuclear DNA markers

Microsatellite locus CH02D12

H1, H2 and H3 each had the same four-banded phenotype. The five-banded phenotype of *S. arranensis* and *S. pseudofennica* included these four bands and had an additional band, band a (Fig. 5.4).

Microsatellite locus CH02D11

H1 and H3 had the same six-banded phenotype. This class of six-banded phenotype was unique to these two individuals. However, it was similar to the six-banded phenotype found in *S. arranensis* with the exception of an apparent length mutation in band a (Fig. 5.5). This length mutation was not evident in H2. However, as band c was notably absent, H2 still had a unique phenotype (Fig. 5.5).

Microsatellite locus CH02B3b

H1, H2 and H3 each had the same single-banded phenotype. This single-banded phenotype was also found in two *S. pseudofennica* individuals (Fig. 3.6, Chapter Three).

Microsatellite locus CH01H01

H1 and H3 had the same three-banded phenotype. This class of three-banded phenotype was unique to these two individuals. H1 had a unique two-banded phenotype (Fig 5.6). Band a, which was present in all adults of both *S. arranensis*

and *S. pseudofennica* was notably absent in H1, H2 and H3. This band was also absent in three *S. pseudofennica* offspring (Figs 4.2 and 4.3 Chapter 4).

Microsatellite locus 12045

H1, H2 and H3 each had the same single-banded phenotype. This phenotype was the same single-banded phenotype that was found in *S. arranensis* (1204A) *and S. pseudofennica* (1204A) (Fig 3.8, Chapter Three).

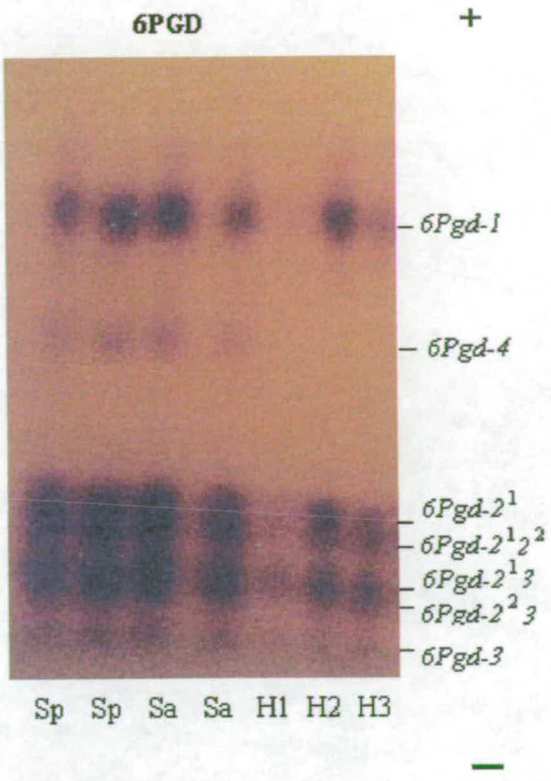


Figure 5.2. 6PGD isozyme banding pattern for the novel *Sorbus* taxon. H1, H2, H3 = novel *Sorbus* taxon individuals, Sa = *S. arranensis*, Sp = *S. pseudofennica.*

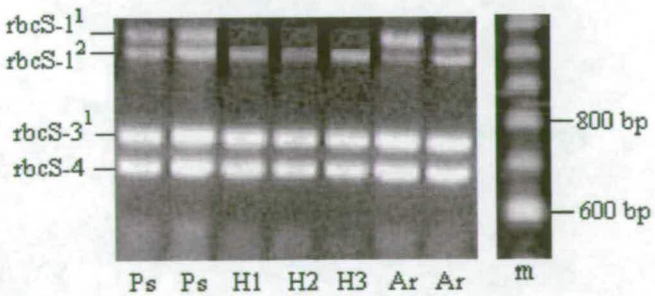


Figure 5.3. Rubisco banding pattern for the novel *Sorbus* taxon. H1, H2, H3 = novel *Sorbus* taxon individuals, Ar = *S. arranensis*, Ps = *S. pseudofennica,* m = 100 bp ladder.

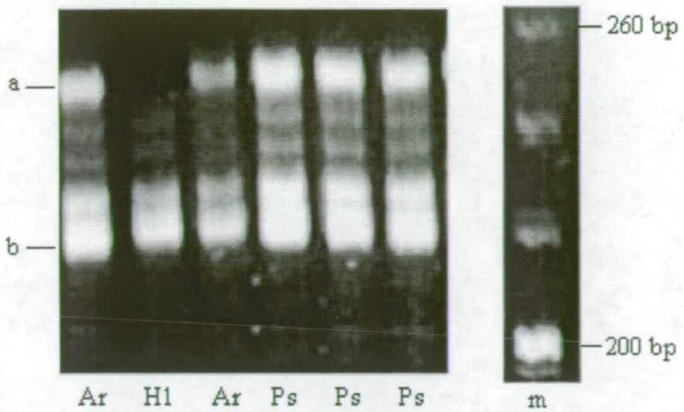


Figure 5.4. Microsatellite locus CH0D12. Banding pattern for the novel *Sorbus* taxon. H1 = novel *Sorbus* taxon, Ar = *S. arranensis*, Ps = *S. pseudofennica,* m = 20 bp ladder.

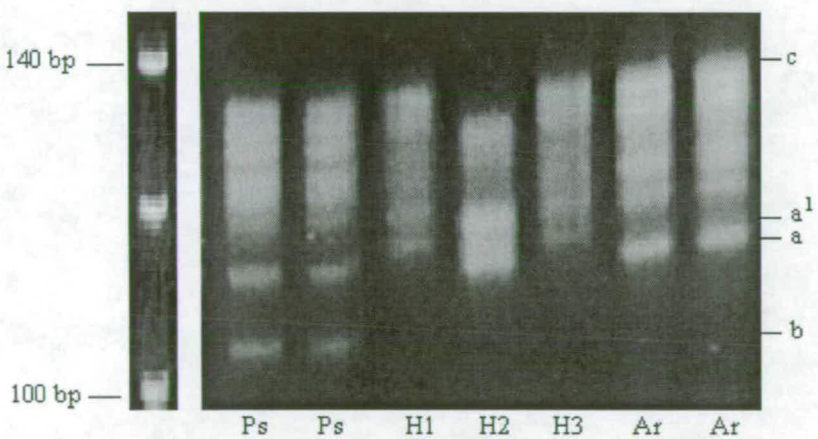


Figure 5.5. Microsatelite locus CH02D11. Banding pattern for the novel *Sorbus* taxon. H1, H2, H3 = novel *Sorbus* taxon, Ar = *S. arranensis*, Ps = *S. pseudofennica,* a1 represents a different length class of band a. m = 20 bp ladder.

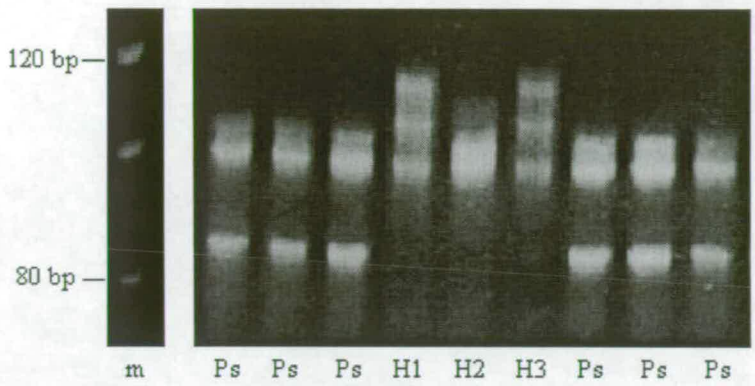


Figure 5.6. Microsatellite locus CH01H01. Banding pattern for the novel *Sorbus* taxon. H1, H2, H3 = novel *Sorbus* taxon, Ps = *S. pseudofennica*. m = 20 bp ladder.

5.5 DISCUSSION

5.5.1 Isozyme and Rubisco intron analysis

The three novel Arran *Sorbus* individuals each had the same banding pattern when screened with the three isozyme systems (AAT LAP and 6PGD) and the Rubisco intron marker. For isozyme system AAT and LAP the novel Arran *Sorbus* banding patterns were identical to those of both *S. arranensis* and *S. pseudofennica*. The 6PGD banding pattern was non-variable among the three novel Arran trees and was unique to this leaf type. Locus *6Pgd-4*, which was fixed in *S. arranensis*, *S. pseudofennica* and *S. rupicola/S. aria* was missing in the novel Arran phenotype. With the exception of missing locus *6Pgd-4* the remainder of the banding pattern was the same as *S. arranensis* and *S. pseudofennica*. The Rubisco banding pattern was also non-variable among the three novel Arran trees and was unique to this leaf type. Allele *rbcS-1*^{1}, which was fixed for *S. arranensis S. pseudofennica* and *S. rupicola/S. aria,* was missing in the novel leaf type individuals. With the exception of missing allele *rbcS-1*^{1} the remainder of the banding pattern was the same as that found in *S. arranensis* and *S. pseudofennica.*

These results support the hypothesis that the novel Arran *Sorbus* taxon was of hybrid origin, ultimately derived from the genomes of both *S. aucuparia* and *S. rupicola/S. aria*. Although it has not been possible to establish the precise origin of the novel Arran phenotype, the banding patterns are consistent with Liljefors (1953) hypothesis that *S. meinichii* (cf. novel Arran *Sorbus* phenotype) is triploid with a genomic formula AAR, with two copies ultimately inherited from *S. aucuparia* (AA) and a single copy ultimately inherited from *S. aria sensu lato* (R).

The novel Arran *Sorbus* phenotype has the same *S. aucuparia* derived banding pattern as was found in *S. pseudofennica.* However, two markers allele *rbcS-1* and loci *6Pgd-4*, that were ultimately derived from the *S. rupicola/ aria* genome, were absent in the novel phenotype but present in *S. pseudofennica*. This loss of *S. rupicola/S. aria* markers (with respect to *S. pseudofennica*) could be explained by the loss of a *S. rupicola/S. aria* chromosome compliment, which would be the case if the novel Arran phenotype was triploid and analogous to *S. meinichii*.

This hypothesis is possible if the novel Arran *Sorbus* was triploid type RAA, with R representing a single *S. rupicola/S. aria* genome and AA representing two *S. aucuparia* genomes. The results in Chapter Four (section 4.5.2) support the hypothesis that *S. pseudofennica* is a facultative apomict with the potential to produce reduced embryo sacs. If these reduced embryo-sacs are type AR and they are fertilised with *S. aucuparia* pollen type A, it would be possible for the resulting offspring to have a triploid genome type RAA.

5.5.2 Microsatellite analysis

The microsatellite banding patterns for the three novel Arran *Sorbus* phenotypes revealed variation among the three individuals. In addition the banding patterns of the novel Arran *Sorbus* individuals were unlike any patterns found among the *S. arranensis* and *S. pseudofennica* individuals. Without having established the equivalent *S. aucuparia* and *S. rupicola* microsatellite banding patterns it was not possible to determine the origins of the markers. However, the novel Arran *Sorbus* phenotypes had unique banding patterns, with respect to *S. arranensis* and *S. pseudofennica*, at three the of microsatellite loci scored.

The results of the above molecular marker assays support the hypothesis that the novel Arran *Sorbus* taxon was of hybrid origin, ultimately derived from the genomes of both *S. aucuparia* and *S. rupicola/S. aria* and support the hypothesis that previously unrecorded genetic exchanges have occurred within the Arran *Sorbus* complex and that these genetic exchanges could have lead to the formation of new apomictic taxa.

Chapter Six

Summary and general discussion

6.1 Introduction

Analysis of pollen taken from lake sediments at loch a'Mhuilinn near Lochranza, north Arran, suggests that the landscape's few scattered woods are remnants of an extensive mixed forest that formerly covered most of the area. Birch and rowan dominate the fragmented semi-natural woodland, with holly, aspen, juniper and willow also present. The woodland is generally sparse and of low stature.

Granite crags occur sporadically throughout the area and many burns dissect the landscape flowing into steep-sided rocky valleys and gorges where landslides and rock disturbances are common. These gorges are the predominant sites where the *Sorbus* microspecies are found. Previous botanical surveys have identified 14 of these sites where the microspecies can be located. The present survey was confined to nine of these sites from which a total 342 *S. arranensis* and 380 *S. pseudofennica* trees were identified and mapped (Chapter Two, section 2.3.2). The present population density for the nine sites was approximately the same as that recorded by the last survey (Burlinson 1986).

Natural hybrids tend to inhabit the fringes of the natural range of their progenitor species, often forming hybrid zones, but rarely ever outcompeting the parents. As ecological disturbances are common precursors for interspecific hybridisation events, it is possible for natural hybrids to become established in areas when the natural habitat of the established parent species has been disturbed.

The range of *S. aucuparia* includes that of the Arran *Sorbus* hybrids but also extends into many other habitats in addition to the steep gorges and banks where the hybrids are found. The Arran hybrids are present on some of the more gentle slopes, but in these locations they tend to be outnumbered by the common *S. aucuparia*. However, there is one site in particular, namely Glenn Diomhan, where the hybrids greatly outnumber *S. aucuparia*. This site has the largest population of *S. arranensis* and *S. pseudofennica* trees, 220 and 260 respectively, more than 50% of the total population for each microspecies. It is perhaps no coincidence that this site is the most inhospitable of them all; very few trees other than the Arran hybrids are present on

the harsh rock slopes. It is possible that the hybrid origin of *S. arranensis* and *S. pseudofennica* has led to their being better adapted to these rock faces than *S. aucuparia*, possibly occupying an ecological niche, allowing the hybrids to survive competition from *S. aucuparia*. *S. aucuparia* appears to be the dominant species in more hospitable environments where its much greater reproductive potential provides a competitive edge over the hybrid species.

Even though the range of *Sorbus* hybrids will always be restricted, it is apparent that many sites within the present range are sparsely populated and could potentially be suitable habitats for the hybrids to colonise. Many adult hybrid trees present in these sites rarely reached any substantial height, over a third of trees surveyed were less than a meter high, and over a three year period these stunted trees were not seen to flower or set fruit. The trees seem to be constantly browsed and growth is thereby restricted.

The principal aim of this study was to gain an understanding of some of the evolutionary processes that determined the fragmented and restricted distribution of these *Sorbus* hybrids. Two approaches were adopted address this issue; a direct approach to detect the number of genotypes present for each *Sorbus* microspecies and record their distribution within and among populations sampled and an indirect approach to investigate three evolutionary processes, namely the hybrid origin of the microspecies, the origin of genotypic diversity and the breeding system of each microspecies. It was assumed at the outset that gaining an understanding of these evolutionary processes would assist in the explanation of the observed pattern of genotypic distribution.

6.2 Hybrid origins

6.2.1 Nuclear markers

Molecular markers were used to test the proposed hybrid origin of the Arran *Sorbus* microspecies. It was hypothesized, that the sexual diploid *S. aucuparia* crossed with the apomict tetraploid *S. rupicola* to produce the triploid apomict *S. arranensis*, and

subsequently a second hybridisation event where *S. arranensis* (with an unreduced gamete) crossed with *S. aucuparia* to produce the tetraploid apomict *S. pseudofennica*.

Nuclear markers are particularly suitable markers to study hybridisation events, as most nuclear genes are inherited in a simple Mendelian fashion and will be additively combined in the hybrids. Two out of the three isozyme systems (AAT and 6PGD), the nuclear DNA Rubisco intron marker and the Chloroplast DNA marker, all proved successful in identifying species-specific markers for *S. aucuparia* and *S. rupicola*. These markers were then used to screen individuals of both *S. arranensis* and *S. pseudofennica* to test the above hypothesis. However, the markers could not separate *S. rupicola* from *S. aria*, which could have been the parent in the primary cross with *S. aucuparia* to produce *S. arranensis*. Similarly as the markers were unable to fully distinguish between *S. arranensis* and *S. pseudofennica* taxa, it was not possible to test the proposed second hybridisation event where *S. arranensis* crossed with *S. aucuparia* to create *S. pseudofennica*. However, it was possible to test whether the genome of *S. pseudofennica* had ultimately been derived from both *S. aucuparia* and *S. rupicola/S. aria*.

For isozyme system AAT the Arran *Sorbus* genotypes were compatible with being derived from a cross between *S. aucuparia* and *S. rupicola/S. aria*, with alleles specific to *S. aucuparia* and to *S. rupicola/S. aria* being combined in the *Sorbus* hybrids as a heterozygous genotype. Similarly, the Arran *Sorbus* 6PGD and Rubisco intron banding patterns were also compatible with being formed by a cross between *S. aucuparia* and *S. rupicola/S. aria*, with both sets of putative parental species-specific markers being combined in the Arran *Sorbus* as an additive banding pattern. The results of the molecular marker assays therefore support the hypothesis of the Arran *Sorbus* being of hybrid origin, with *S. aucuparia* and *S. rupicola/S. aria* as the progenitor species.

An additional allele present in the banding patterns of 37.1% of *S. pseudofennica* individuals was not present in any of the *S. aucuparia* or *S. arranensis* individuals

screened. It is possible that the source of this additional allele was the result of a second hybridisation event between *S. aucuparia* and *S. arranensis*, where a genetically distinct *S. arranensis* individual (not sampled) was involved in a cross with *S. aucuparia* to produce a new and genetically distinct *S. pseudofennica* clonal lineage. Alternatively, the source of the new allele could have been the result of a mutation event in a *S. pseudofennica* individual that was subsequently propagated via agamospermy. It is unlikely that the additional allele was derived from a genetically distinct *S. aucuparia* individual, as the additional *S. pseudofennica* allele was derived from a locus that was specific to the *S. rupicola/S. aria* genome complex.

6.2.2 Chloroplast DNA markers

Liljefors (1955) proposed that an autotetraploid such as *S. rupicola* is likely to have been the pollen donor in an interspecific cross with *S. aucuparia*, which created the triploid *S. arranensis* group. Evidence to support Liljefors (1955) hypothesis came from embryological studies and pollen analysis, showing that *S. rupicola* was an aposporous obligate apomict (producing only unreduced embryo sacs) and pollen was viable (20%) which could in theory fertilise a reduced female *S. aucuparia* gamete (Liljefors 1953, 1955).

Presuming that chloroplast DNA is maternally inherited in *Sorbus* taxa, as in the majority of angiosperms, the sequence data demonstrated that *S. aucuparia* is the maternal parent of the Arran *Sorbus* hybrids. The sequences of the trnL intron and the trnL-trnF intergenic spacer for the Arran *Sorbus* hybrids were 100 % identical to those of *S. aucuparia*. In contrast, nucleotide substitutions were evident between the *Sorbus* hybrids and *S. rupicola* at three bases. These results support the hypothesis that an autotetraploid such as *S. rupicola* (RRRR) was likely to be the pollen donor in an intespecific cross with *S. aucuparia* (AA).

To test if the remaining *S. arranensis* and *S. pseudofennica* individuals each had the same maternally inherited *S. aucuparia* chloroplast type, and were not derived from a reciprocal cross, a random sample of 38 *S. pseudofennica* and 54 *S. arranensis* individuals were screened with a diagnostic PCR-RFLP marker. Each of the *S.*

arranensis and *S. pseudofennica* individuals were found to possess the same chloroplast digest restriction pattern identical to that of *S. aucuparia*.

6.2.3 Limitations of molecular markers used in hybrid studies

The results of the current study highlight potential limitations when using molecular markers to study hybridisation events. It is clearly possible to determine the hybrid origins of taxa derived from interspecific events such as *S. aucuparia* and *S. aria sensu lato* (the progenitor species have been reproductively isolated from each other long enough for the genomes to have evolved separately), where molecular markers will readily be species-specific and the subsequent hybrids will clearly show a parentally derived additive banding pattern. However, there are problems in determining hybrid origins of taxa where the proposed progenitor taxa are closely related and subsequently the identification of species-specific markers will prove more elusive.

This is illustrated by the fact that the molecular markers were not able to determine the precise origin of *S. pseudofennica,* although it has been demonstrated that it was ultimately derived from the genomes of both *S. aucuparia* and *S rupicola/S. aria.* However, what is not clear is the realisation of the proposed second hybridisation event between *S. arranensis* and *S. aucuparia* that supposedly created *S. pseudofennica*.

To determine the hybrid origins of *S. pseudofennica*, markers are required that will detect nuclear DNA polymorphisms within *S. aucuparia*. The isozyme systems and the Rubisco intron markers did not detect nuclear DNA variation within the *S. aucuparia* population (with the exception of two individuals (2.7 %) with a different AAT genotype).

In theory, an additional *S. aucuparia* species-specific band found in *S. pseudofennica* but absent from *S. arranensis* could provide evidence for the origin of *S. pseudofennica* being the result of an hybridisation event between *S. aucuparia* and *S. arranensis*. This would be possible if the *S. aucuparia* population, or a single

individual, was heterozygous at a particular locus, with one allele type being present in the primary hybridisation event, *S. aucuparia* (reduced haploid gamete, A^1, n = 2x = 17) x *S. rupicola/S. aria* (reduced diploid gamete, RR, n = 4x = 34), producing triploid *S. arranensis*, RRA^1, and subsequently the alternative allele being present in the backcross hybridisation event *S. aucuparia* (haploid gamete, A^2, n = 2x = 17) x *S. arranensis* (unreduced triploid gamete, RRA^1, 3n = 51) producing tetraploid *S. pseudofennica*, RRA^1A^2.

Nuclear DNA markers that have high mutation rates, such as microsatellites, may provide such polymorphic markers. However, there are potential problems when using molecular markers with high mutation rates. For example a preliminary study using microsatellite markers revealed potential allelic variation among 13 *S. aucuparia* individuals that may provide a marker of a type described above. There were two potential problems with interpreting this allele as evidence that *S. pseudofennica* was derived from a cross between *S. arranensis* and *S. aucuparia*. In the first instance only six *S. rupicola* individuals were screened with the microsatellite marker (even though the allele was absent, a sample size of only six may not have revealed the complete set of *S. rupicola* alleles). Secondly there is a possibility that the additional allele detected in *S. pseudofennica* and *S. aucuparia* does not share a common descent but has acquired the same properties by the process of homoplasy (parallel evolution).

6.3 Genetic structure and origins of genotypic diversity

The nuclear DNA markers (five microsatellite primer pairs and the Rubisco intron primer pair) proved successful in detecting variation among individuals within *Sorbus arranensis* and *S. pseudofennica* microspecies. The DNA markers also detected species-specific marker variation between *S. arranensis* and *S. pseudofennica*, providing evidence that the differences observed between *S. arranensis* and *S. pseudofennica*, had a genetic basis which was derived from different evolutionary events.

6.3.1 Genotypic variation

A total of 179 *S. arranensis* individuals from eight populations were screened with six nuclear DNA markers. Only three genotypes ((clones) AA, AB, and AC) were detected. Genotype AA was present in each of the eight populations, with genotype AB present in three, and genotype AC in one.

These results do not support the hypothesis that each *S. arranensis* population consists of a single unique genotype, as population B (Diomhan Burn tributary) has two *S. arranensis* genotypes (AA, n = 11, AB, n = 5), population E (Catacol Burn) also has two genotypes (AA, n = 20, AB, n = 1) and population J (Alt Easan Biorach) has three genotypes (AA, n = 1, AB, n = 1 and AC, n = 1). It is a possibility that with more marker loci the widespread *S. arranensis* genotype (AA) could be divided into genotypes that were each unique to a particular population. However, it would not alter the findings that three populations have more than one *S. arranensis* genotype present. Based on the present genotype distribution it is more likely that if more *S. arranensis* genotypes were resolved the populations would have greater genotypic diversity within them and populations would be appear more similar. This is in contrast with Hull and Smart (1984), who proposed that each *S. arranensis* population consists of a single unique genotype.

A total of 140 *S. pseudofennica* individuals from five populations were screened with the same set of six nuclear DNA markers that were used to screen the *S. arranensis* samples. Eight genotypes were detected PA, PB, PC, PD, PE, PF, PG, and PH. Genotype PA was present in each of the five *S. pseudofennica* populations, genotype PC was present in two, and each of the remaining genotypes were restricted to a single population. These results do not support the hypothesis that each *S. pseudofennica* population consists of a similar set of genotypes with each genotype being freely distributed among the populations. Only a single genotype (PA) is distributed among each of the five populations.

If more marker loci were used and more *S. pseudofennica* genotypes were detected as a result, the distribution pattern would still be unlikely to reflect that proposed by

Hull and Smart (1984). Based on the present distribution it is predicted that additional *S. pseudofennica* genotypes found would tend to be restricted to single populations. Each population would therefore consist of its own set of genotypes, with the exception of one or two individuals being inter-dispersed, rather than all populations consisting of a similar set of genotypes.

6.3.2 Origin of genotypic diversity

The observation that *S. arranensis* genotypes AB and AC differ from the dominant genotype (AA) at only one primer pair locus, implies that the differences between the genotypes have arisen as a result of mutation within the taxa, rather than having been derived from additional primary hybridisation events between the progenitors species *S. aucuparia* and *S. rupicola/S. aria*. If the genotypic diversity observed within *S. arranensis* was derived from additional primary hybridisation events, then the progenitor species *S. aucuparia* and *S. rupicola/S. aria* could only be polymorphic at the single locus that distinguish the three *S. arranensis* genotypes, and would have to be monomorphic at the other five loci scored. In support of this hypothesis, ten *S. aucuparia* genotypes were detected from a total of 13 individuals screened. It is likely therefore that any additional primary hybridisation events between genetically distinct *S. aucuparia* and *S. rupicola* would be reflected in the additional *S. arranensis* clones having significantly different genotypes, not just varying at a single locus.

The number of variable loci that separated the *S. pseudofennica* genotypes ranged from one to four. Where two genotypes are separated from each other by only one locus, the origin of this genotypic variation is likely to have been derived from a single mutation event within a clone, as explained above for *S. arranensis*. However, the origin of the variation found between *S. pseudofennica* genotypes that vary at four loci may reflect additional primary hybridisation events between genetically distinct *S. aucuparia* individuals and genetically invariant individuals of *S. arranensis*.

The origins of genotypic diversity and the different reproductive mechanisms employed by *S. arranensis* and *S. pseudofennica* could provide explanations for the greater genotypic diversity values observed for *S. pseudofennica*. The origin of *S. arranensis* genotypes is likely to be monophyletic with subsequent obligate agamospermous reproduction contributing only a minimal amount of genotypic variation. In contrast the origin of *S. pseudofennica* genotypes is potentially polyphyletic with subsequent facultative agamospermous reproduction recombining this genotypic diversity into new clonal lineages.

6.3.3 Genotypic diversity comparisons between agamospermous and sexual taxa

As discussed in Chapter One natural selection acts on additive genetic variance among individuals and genotype diversity (as opposed to gene diversity) within populations is in evolutionary terms an important statistic to observe. As a consequence of sex and recombination there are likely to be large differences in the way genetic variation is arranged into multilocus genotypes between sexual and agamospermous taxa, with the number of genotypes likely to be small in apomictic taxa and large in sexual taxa.

Genotypic variation has been observed in many agamospermous taxa, including those in the present study. However, when compared with the genotypic diversity observed in related sexual taxa, the genotypic variation is relatively very low and possibly from an evolutionary perspective ultimately restrictive. For example, the proportion of genotypes (Ng/Ni) detected for *S. arranensis* and *S. pseudofennica* was 0.016 and 0.057 respectively, compared with 0.769 for the sexual outcrossing relative *S. aucuparia*. The results from six other studies (Chapter Three, section 3.5) were consistent with the current study with a mean Ng/Ni value of 0.09 for the agamospermous taxa and 0.61 for sexual taxa.

6.3.4 Limitations of molecular markers in studying genotypic diversity

Identifying a method to detect genetic variation among individuals of *S. arranensis* and *S. pseudofennica* proved to be problematic. Genetic markers, unlike morphological markers, only survey a limited amount of the target genome so it was

a question of trial and error before a suitable marker system was successful in detecting genetic variation between any two individuals. The first assay tested was isozymes, which proved successful in detecting variation between the putative parental species *S. aucuparia* and *S. rupicola* and was successful, to an extent, in testing the proposed hybrid origin of the Arran microspecies. However, they failed to detect any variation within or between the two agamospermous taxa. It may have been possible that if more isozyme systems were optimised some genetic variation may have been found. However, as there was only limited variation found among the sexual outcrosser, *S. aucuparia*, it was unlikely there was ever going to be a great deal of variation found among the agamospermous clonal lineages.

The second assay tested was internal simple sequence repeat PCR primers (Inter-SSR). Inter-SSR primers are complementary to genomic simple sequence repeats (1-4 nucleotides occurring in tandem repeats) and contain short oligonucleotide 'anchor' sequences that ensure the primers anneal to either the 5'or 3'end of the genomic repeat. If two sites are in the correct orientation with respect to each other and within range the region between them will be amplified. As the SSR primers target highly variable and numerous loci it was hoped that this assay would be capable of detecting and quantifying genetic variation among *S. arranensis* and *S. pseudofennica* individuals. The method proved successful in detecting variation among individuals. However, quantifying this variation was more problematic. The gel band positions where the variation was evident were faint and difficult to detect with any certainty, between individuals, or between repeat PCR amplifications. It was not possible to rule out that the variation was derived from PCR artefacts hence the bands were therefore not scored. The band positions that were clear, reliably amplified and scoreable were invariable. This pattern was repeated for 54 *S. arranensis* and 38 *S. pseudofennica* individuals screened with five Inter-SSR primers.

The third assay tested was a set of PCR primers designed to amplify the intron regions of isozyme genes (Strand et al. 1997). As introns are regions of DNA that are not expressed there is a theoretical possibility that they could accumulate mutations

at a faster rate than the expressed products of isozyme genes. Primer pairs designed to amplify introns within genes of Calmodulin (*Cam*), Glyceraldehyde 3-phosphate dehydrogenase (G3*pdH*) and the ribulose-1, 5-bisphosphate carboxylase small subunit (*rbcS*) were used to screen ten *S. arranensis* and ten *S. pseudofennica* individuals. Restriction digests (using a selection of five restriction enzymes) from the PCR amplification products of three of these primer pairs did not detect any variation between individuals. However, the *rbcS* intron revealed a length variation between *S. pseudofennica* individuals that was clear and prove to be repeatable and as described in Chapters Two and Three was suitable for use in the hybrid origin and genetic structure assays.

The final molecular assay to be tested was microsatellite analysis, which after considerable time, effort and cost (as described in Chapter Three) proved to be successful in detecting genomic variation among individuals within *Sorbus arranensis* and *S. pseudofennica* microspecies that could be quantified and used to delimit clones.

6.4 Population distribution of *S. pseudofennica* and *S. arranensis* genotypes

The results in the current study show that a single *S. arranensis* genotype AA, representing 95.5% of the *S. arranensis* population was distributed among each of the eight populations, genotype AB, representing 3.9% distributed among three populations, and genotype AC was represented by only a single individual. This pattern could be explained by genotype AA being the original genotype derived from the hypothesised single interspecific hybridisation event between *S. aucuparia* and *S. rupicola/S. aria* (Chapter Three, section 3.5.3) and that the genotypes AB and AC were subsequently formed via mutation at a significantly later date. Genotype AA, being the first formed, could have migrated via bird dispersal from its place of origin to all eight populations. If, as hypothesised the genotypes AB and AC were formed as a direct result of mutation within genotype AA at a much later date, it may be possible that these genotypes are only just beginning to expand their range and population numbers.

The above hypothesis could also explain the distribution of *S. pseudofennica* genotypes. Genotype PA represents 50% of the *S. pseudofennica* population and was distributed among each of the five populations. This widespread genotype could represent the *S. pseudofennica* genotype formed from the first hybridisation event between *S. arranensis* and *S. pseudofennica*. This genotype being the first formed could have migrated via bird dispersal from its place of origin to the five present populations. The formation of *S. pseudofennica* requiring a backcross hybridisation event between *S. arranensis* and *S. aucuparia* may explain why *S. arranensis*, being formed first, has a much greater distribution.

The *S. pseudofennica* genotypes restricted to single populations may reflect their formation at a later date. It possible that both *S. arranensis* and *S. pseudofennica* are just beginning to expand their range and population numbers. This expansion has been, and will continue to be possible via seed dispersal by frugivorous birds. It was observed during this study that birds do actually feed from both *S. arranensis* (Appendix 2.1) and *S. pseudofennica* (personal observation). However, the birds had a preference for the heavier fruit crops of *S. aucuparia* (Appendix 2.1) and only appeared to start feeding from the *S. pseudofennica* and *S. arranensis* trees when the berry crop of *S. aucuparia* was approximately equal to that of *S. arranensis* and *S. pseudofennica*. This preference for the heavier fruiting *S. aucuparia* trees may affect the seed dispersal of *S. arranensis* and *S. pseudofennica*. For example, if for any one year there was a low number of migrating frugivorous birds and the fruit crop of *S. aucuparia* was sufficiently high to feed them, it may be possible for birds to ignore the *S. arranensis* and *S. pseudofennica* trees.

6.5 Conservation of *S. arranensis* and *S. pseudofennica* microspecies

Investigations of the genetic structure and distribution of a rare species can provide important data on which conservation and management plans can be based (Holsinger and Gottlieb 1991). It is important to establish the extent and range of genotypes as seed collections, or transplants of rare genotypes to protected areas, may be necessary. The principle aim of the current study was to determine the genetic structure of the *Sorbus* microspecies. By studying their hybrid origins, breeding systems and origins of clonal diversity, it was hoped to gain an understanding of how this pattern may have been formed.

Looking at the pattern of genotypic diversity of *S. arranensis* and *S. pseudofennica* without an understanding of the processes involved could lead an observer to reasonably conclude that the number of clones for both microspecies were in decline. However, as described above and in Chapters Three and Four, the pattern may actually reflect that both microspecies are in the process of starting to expand their clonal composition, which could in turn lead to an increase in the species range and population size.

Any conservation plan should focus on the evolutionary potential of the group, in particular the importance of the role played by the sexual parent *S. aucuparia*. It is likely that *S. aucuparia*, by hybridising with *S. arranensis* on more than one occasion, provided the source of the observed genotypic diversity found among the *S. pseudofennica* genotypes. The presence of the novel Arran *Sorbus* phenotype suggests previously unrecorded hybridisation events within the group, possibly between *S. aucuparia* and *S. pseudofennica*. This would give even greater importance to the role played by *S. aucuparia*.

The continuing input of new *S. pseudofennica* genotypes through hybridisation of progenitor taxa and the ability to recombine this genetic variability through facultative agamospermy may provide sufficient genotypic diversity and the mechanism to maintain this diversity to enable *S. pseudofennica* populations to adapt

to changing environments. The obligately *S. arranensis* microspecies may not have the potential to evolve, but as discussed previously has an important evolutionary role in the formation of new *S. pseudofennica* genotypes.

From an evolutionary perspective *S. aucuparia*, *S. pseudofennica* and *S. arranensis* (and possibly the novel Arran *Sorbus* phenotype) have an integrated and dynamic role in creating new *Sorbus* genotypes. From a conservation perspective, it is important to maintain and facilitate this dynamic process. Possibly the first and most important steps are to ensure the *S. aucuparia* populations (in particular those sympatric with the Arran *Sorbus* microspecies) are maintained. This will not only preserve an important source of genetic diversity, but may also relax the browsing pressure on the *Sorbus* microspecies and subsequently improve the evolutionary potential of the group. To achieve this active planting of *S. aucuparia*, along with other native woodland species is required.

6.6 Further work

The results of the current study highlighted limitations when using molecular markers to study evolutionary mechanisms. The isozyme systems and the Rubisco intron markers were unable to distinguish between *S. rupicola* and *S. aria*, or *S. arranensis* and *S. pseudofennica*. Microsatellites are a system capable of achieving this aim but as discussed above (section 6.5.3) their rapid rate of evolution can limit their use in hybridisation studies. Many studies have employed nuclear ribosomal internal transcribed spacers (ITS) to test hybrid origins. However, care must be taken when using such a system as gene conversion appears to be able to replace one parental ITS type with the other, with the age of the hybrid complex determining the stage of gene conversion. It may be possible that by screening more isozyme systems suitable markers will be found.

The molecular marker studies on the novel Arran *Sorbus* phenotype suggest that previously unrecorded genetic exchanges have occurred within the Arran *Sorbus*

complex. Continued analysis of this *Sorbus* taxa using additional molecular markers and cytological analysis, will enable a more precise determination of its origin.

References

Aas G, Maier J, Baltisberger M, Metzger S (1994) Morphology, isozyme variation, cytology, and reproduction of hybrids between *Sorbus aria* (L.) Crantz and *S. torminalis* (L.) Crantz. *Botanica Helvetica*, **104**,195-214.

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

Abbott RJ, Ashton PA, Forbes DG (1992) Introgressive origin of the radiate groundsel, *Senecio vulgaris* L. var *hibernicus* Syme: Aat-3 evidence. *Heredity*, **68**, 425–435.

Anderson E (1949) *Introgressive hybridisation*. John Wiley, New York.

Antonius K and Nybom H (1994) DNA-fingerprinting reveals significant amounts of genetic-variation in a wild raspberry *Rubus-idaeus* population. *Molecular Ecology*, **3**, 177-180.

Arnold ML (1992) Natural hybridisation as an evolutionary process. *Annual Review of Ecology and Systematics*, **23**, 237-261.

Arnold ML (1997) *Natural hybridisation and evolution*. Oxford University Press, Oxford, UK.

Arnold ML, Hodges SA (1995) Are natural hybrids fit or unfit relative to their parents. *Trends in Ecology and Evolution*, **10**, 67-71.

Ashton PA, Abbot RJ (1992) Isozyme evidence and the origin of *Senecio vulgaris* (Compositae). *Plant Systematics and Evolution*, **179**, 167-74.

Asker SE (1979) Progress in apomixis research. *Hereditas*, **91**, 213-240.

Asker SE (1980) Gametophytic apomixis: elements and genetic regulation. *Hereditas*, **93**, 277-293.

Asker SE, Jerling L (1992) *Apomixis in plants*. CRC-Press, Boca Raton.

Assienan B, Noirot M (1995) Isozyme polymorphism and organisation of the agamic complex of the *Maximae* (*Panicum maximum* Jacq., *P. infestum* Anders., and *P. trichocladum* K. Schum.) in Tanzania. *Theoretical and Applied Genetics*, **91**, 672-680.

Baker MC, Baker AEM (1990) Reproductive behaviour of female buntings: isolating mechanisms in a hybridising pair of species. *Evolution*, **44**, 1272-1282.

Barton NH (1995) A general model for the evolution of recombination. *Genetical Research,* **65,** 123-144.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113-148.

Battjes J, Menken SBJ, Den Nijs HCM (1992) Clonal diversity in some microspecies of *Taraxacum* sect. *Palustria* (Lindb. Fil.) Dahlst. from Czechoslovakia. *Botanischer Jahrbucher Systematische*, **114**, 315-328.

Bayer RJ (1989) Patterns of isozyme variation in *Antennaria rosea* (Asteraceae: Inuleaea) polyploid agamic complex. *Systematic Botany*, **14**, 389-397.

Bayer, RJ (1990) Patterns of clonal diversity in the *Antennaria-rosea* (Asteraceae) Polyploid agamic complex. *American Journal Of Botany*, **77**, 1313-1319.

Bayer RJ (1991) Allozymic and morphological variation in *Antennaria* (Asteraceae: Inuleaea) from the low arctic of northwestern North America. *Systematic Botany*, **16**, 492-506.

Bignal E (1980) The endemic whitebeams of north Arran. *Glasgow Naturalist*, **20**, 59-64.

Bolstad AM and Slavesen PH (1999) Biosystematic studies of *Sorbus meinichii* (Rosaceae) at Moster, S. Norway. *Nordic Journal of Botany,* **19**, 547-559.

Botstein D, White RL, Skolnick M, Davis W (1980) Construction of a genetic map in man using restriction fragment length polymorphism. *American Journal of Human Genetics*, **32**, 314-331.

Bryce J (1872) *Geology of Arran and other Clyde islands*. Collins Glasgow.

Brysting AK, Holst-Jensen A, Leitch I (2000) Genomic origin and organisation of the hybrid *Poa jemtlandica* (Poaceae) verified by genomic *in situ* hybridisation and chloroplast DNA sequences. *Annals of Botany*, **85**, 439-445.

Burlison JP (1986) Arran *Sorbus* census. Unpublished file note, NCC 144/R

Campbell CS, Weeden NF, Dickson EE, Wright WA (1991) Biased gene flow from an apomictic to a sexual species of *Amelanchier* (Rosaceae: Maloideae) based on a chloroplast DNA restriction site difference. *American Journal of Botany, abstract supplement*, **78,** 438.

Carman JG (1977) Asynchronous expression of duplicated genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biological Journal of the Linnean Society*, **61**, 51-94.

Carson HL (1975) The genetics of speciation at the diploid level. *American Naturalist*, **109**, 73-92.

Charters YM, Robertson A, Wilkinson MJ, Ramsay G (1996) PCR analysis of oilseed rape cultivars (*Brassica napus* L. ssp. *oleifera*) using 5′-anchored simple sequence repeat (SSR) primers. *Theoretical and Applied Genetics*, **92**, 442-447.

Chase S (1969) Monoploids, and monoploid-derivatives of maize (*Zea mays* L.). *Biological Review*, **35**, 117-167.

Cheliack WM, Pitel JA (1984) Techniques for starch electrophoresis of enzymes from forest tree species. Information report PI-X-42, Petawawa National Forest Institute.

Clapham AR, Tutin TG, Warburg EF (1952) *Flora of the British Isles*. Cambridge University Press, Cambridge.

Clausen J (1954) Partial apomixis as an equilibrium system in evolution. *Caryologia*, **6**, 469-479.

Clausen J, Keck DD, Heisey WM (1939) The concept of species based on experiment. *American Journal of Botany*, **26**, 103-106.

Clegg MT, Learn GH, Golenberg EM (1991) Molecular evolution of chloroplast DNA. In: *Evolution at the molecular level* (eds Selander RK, Clark AG, Whittam TS), pp. 135-149. Sunderland: Sinauer Associates.

Darlington CD (1939) *The evolution of genetic systems*. Oliver and Boyd, Edinburgh.

De Vincente MC, Tanskley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics*, **134**, 585-596.

Dean C, Pichersky E, Dunsmuir P (1989) Structure, evolution and regulation of *rbcS* genes in higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, **40**, 415-439.

Dean C, van den Elzen P, Tamaki S, Dunsmuir P, Bedbrook J (1985) Linkage and homology analysis divides the eight genes for the small subunit of petunia ribulose 1,5-bis-phosphate carbyoxylase into three gene families. *Proceedings of the National Academy of Sciences USA,* **82**, 4964-4968.

Dejoode DR, Wendel JF (1992) Genetic diversity and origin of the Hawaiian Islands cotton, *Gossypium tomentosum. American Journal of Botany*, **79**, 1311-1319.

DeRocher EJ, Quigley F, Mache R, Bohnert HJ (1993) The six genes of the Rubisco small subunit multigene family from *Mesembryanthemum crystallinum*, a facultative CAM plant. *Molecular and General Genetics*, **239**, 450-462.

do Valle CB, Savidan YH, (1996) Genetics, cytogenetics and reproductive biology of *Brachiaria*. In: *Brachiaria: Biology, agronomy, and improvement*, (eds Miles JW, Maass BL, do Valle CB), pp.147-163. CIAT-EMBRAPA, Cali, Columbia.

Dobzhansky TH (1937) *Genetics and the origin of species*. Columbia University Press, New York.

Dobzhansky TH (1970) *Genetics of the evolutionary process*. Columbia University Press, New York.

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13-15.

Dudley and Map Maker (1996, 1997) Map Maker Ltd., 64 Tenison Rd., Cambridge

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

Ellstrand NC, Roose ML (1987) Patterns of genotypic diversity in clonal plant species. *American Journal of Botany*, **74**, 123-131.

Ellstrand NC, Whitkus R, Rieseberg LH (1996) Distribution of spontaneous plant hybrids. *Proceedings of the National Academy of Sciences USA*, **93**, 5090-5093.

Endler JA (1986) *Natural selection in the wild.* Princeton University Press, Princeton, N.J.

Ennos RA (1983) Maintenance of genetic-variation in plant populations. *Evolutionary biology*, **16**, 129-155.

Falk DA and Holsinger K (1991) *Genetics and conservation of rare plants*. Oxford University Press, New York.

Feldman MW, Christiansen FB, Brooks LD (1980) Evolution of recombination in a constant environment. *Proceedings of the National Academy of Sciences of the USA*, **77**, 4838-41.

Felenstein J (1974) The evolutionary advantage of recombination. *Genetics*, **78**, 737-756.

Fisher RA (1930) *The genetical theory of natural selection*. Clarendon Press, Oxford, England. Second edition 1958, Dover, New York.

Ford H, Richards AJ (1985) Isozyme variation within and between *Taraxacum* agamospecies in a single locality. *Heredity*, **55**, 289-291.

Futuyma DJ (1986) *Evolutionary biology*. Sinauer Associates, Sunderland, MA.

Gallez GP, Gottlieb LD (1982) Genetic evidence for the hybrid origin of the diploid plant *Stephanomeria diegensis*. *Evolution*, **36**, 1158-1167.

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998) Simple sequence repeats for genetic analysis. *Theoretical and Applied Genetics*, **96**, 1069-1076.

Gill BS (1991) Nucleo-cytoplasmic interaction (NCI) hypothesis of genome evolution and speciation in polyploid plants. In: *Proceedings of Dr. H. Kihara Memorial International Symposium on Cytoplasmic Engineering in Wheat* (eds Sasakuma T, Kinoshita T), pp. 48-53.

Godwin H (1975) The history of the British flora a factual basis for phytogeography. 2nd edition. Cambridge University Press, Cambridge.

Gornall RJ (1999) Population genetic structure in agamospermous plants. In: *Molecular systematics and plant evolution* (eds Hollingsworth PM, Bateman RM, Gornall RJ), pp. 118-138. Taylor and Francis, London.

Grant V (1958) The regulation of recombination in plants. *Cold Spring Harbour Symposium in Quantitative Biology*, **23**, 337-363.

Grant V (1971) *Plant speciation*. Columbia University Press, New York.

Grant V (1975) *Genetics of flowering plants*. Columbia University Press, New York.

Grant V (1981) *Plant speciation.* Second edition. Columbia University Press, New York.

Grimanelli D, Leblanc O, Espinosa E, Perotti E, González De León D, Savidan Y (1998) Non-Mendelian transmission of apomixis in maize-Tripsacum hybrids caused by a transmission ratio distortion. *Heredity*, **80**, 40-47.

Grootjans AP, Allersma RJ, Kirk C (1987) Hybridisation of the habitat in disturbed hay meadows. In: *Disturbance in grasslands* (ed Van Andel J), pp. 67-77. Dr W. Junk Publishers, Dordrecht, The Netherlands.

Gustaffson A (1946) *Apomixis in higher plants*. Gleerup, Lund.

Hamrick JL (1983) The distribution of genetic variation within and among natural plant populations. In: *Genetics and conservation* (eds Schoewald-Cox CM, Chamber SM, MacBryde B, Thomas WL), pp. 501-508. Benjamin/Cummings, Menlo Park, CA.

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

Hardy GH (1908) Mendelian proportions in a mixed population. *Science*, **28**, 49-59.

Harlan SC (1936) The genetical conception of the species. *Biological Review*, **11**, 83-112.

Harrison RG (1990) Hybrid zones: windows on evolutionary process. *Oxford Surveys in Evolutionary Biology*, **7**, 69-128.

Harrison RG (1993) Hybrids and hybrid zones: historical perspective. In: *Hybrid zones and the evolutionary process* (ed Harrison RG), pp. 3-12. Oxford university Press, Oxford.

Haufler CH (1985) Pteridophyte evolutionary biology: the electrophoretic approach. *Proceedings of the Royal Society of Edinburgh*, **86B**, 315-323.

Hedland T (1901) Monographie der Gattung *Sorbus*. *Kongliga Svenska Vetenskaps Academiens Handlingar*, **35**, 1-47.

Heisser CB Jr., Smith DM, Clevenger S, Martin WC Jr. (1969) The North American sunflowers (*Helianthus*). *The Memoirs of the Torrey Botanical Club*, **22**, 1-213.

Hollingsworth ML, Bailey JP, Hollingsworth PM, Ferris C (1999) Chloroplast DNA variation and hybridization between invasive populations of Japanese knotweed and giant knotweed (*Fallopia*, Polygonaceae). *Botanical Journal of the Linnean Society*, **129**, 139-154.

Holsinger KE, Gottlieb LD (1991) Conservation of rare and endangered plants: principles and prospects. In: *Genetics and conservation of rare plants* (eds. Falk DA, Holsinger KE), pp. 195-223. Oxford University Press, Oxford.

Howard DJ, Waring GL (1991) Topographic diversity, zone width, and the strength of reproductive isolation in a zone of overlap and hybridisation. *Evolution*, **47**, 789-800.

Hughes J (1987) Variability in sexual and apomictic *Taraxacum*. PhD Thesis, University of Newcastle upon Tyne, Newcastle upon Tyne, UK.

Hughes J, Richards AJ (1988) The genetic structure of populations of sexual and asexual *Taraxacum* (dandelions). *Heredity*, **60**, 161-171.

Hughes J, Richards AJ (1989) Isozymes and the status of Taraxacum agamospecies. *Botanical Journal of the Linnean Society*, **99**, 365-376.

Hull P, Smart GJB (1984) Variation in two Sorbus species endemic to the Isle of Arran, Scotland. *Annals of Botany*, **53**, 641-648.

Hultén E (1971) *Atlas över växternas utbredning i Norden.* Stockholm.

Hultén and Fries (1986*) Atlas of North European vascular plants of the Tropic of Cancer.* Koeltz Scientific Books, Königstein.

Huskins CL (1931) The origin of *Spartina townsendii. Nature*, **127**, 781.

Jiang J, Gill BS (1994) Different species-specific chromosome translocations in *Triticum timopheevii* and *T. turgidum* support the diphyletic origin of polyploid wheats. *Chromosome Research*, **2**, 59-64.

Karl SA, Avise JC (1993) PCR-based assays of Mendelian polymorphisms from anonymous single copy nuclear DNA- techniques and applications for population genetics. *Molecular Biology and Evolution*, **2**, 342-361.

Kijas JMH, Fowler JCS, Thomas MR (1995) An evaluation of sequence tagged microsatellite site markers for genetic analysis within *Citrus* and related species. *Genome,* **38**, 349-355.

Kimura M (1983) *The neutral theory of molecular evolution.* Cambridge University Press, Cambridge.

Knobloch IW (1972) Intergenic hybridisation in flowering in plants. *Taxon*, **21**, 97-103.

Kondrashov AS (1982) Selection against harmful mutations in large sexual and asexual populations. *Genetical Research, Cambridge*, **40**, 325-32.

Kraft T, Nubom H (1995) DNA fingerprinting and biometry can solve some taxonomic problems in apomictic blackberries (*Rubus* subgen. *Rubus*). *Watsonia*, **20**, 329-343.

Kraft T, Nubom H, Werlemark G (1994) *Rubus vestervicensis* (Rosaceae)- its hybrid origin revealed by DNA fingerprinting. *Nordic Journal of Botany*, **15**, 237-242.

Kraft T, Nybom H, Werlemark G (1996) DNA fingerprint variation in some blackberry species (*Rubus* subg *Rubus*, Rosaceae). *Plant Systematics and Evolution*, **199**, 93-108.

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

Landsbourgh D (1897) *Pyrus aria* and its varieties in Arran. *Transactions and Proceedings of the Botanical society Edinburgh*, **21**, 56-62.

Leblanc O, Savidan Y (1994) Timing of megasporegenesis in *Tripsacum* species (Poaceae) as related to the control of apomixis and sexuality. *Polish Botanical Studies*, **8**, 75-81.

Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*, **4**, 203-221.

Li Z, Pinson SRM, Paterson AH, Park WD, Stansel JW (1997) Genetics of hybrid sterility and hybrid breakdown in an intersubspecific rice (*Oryza sativa* L.) population. *Genetics*, **145**, 1139-1148.

Lid J, Lid DT (1994) *Norsk flora*, ed. 6 of R. Elvin. Det norske samaget, Oslo.

Liljefors A (1953) Studies on propagation, embryology and pollination in *Sorbus*. *Acta Horti Bergini*, **16**, 277-239.

Liljefors A (1955) Cytological studies in *Sorbus*. *Acta Horti Bergini*, **17**, 47-113.

Lin S-J, Kato M, Iwatsuki K (1996) Electrophoretic evidence for the origins of tetraploids and hybrids of *Sphenomeris chinensis sensu lato* (Lindsaeaceae) in Japan. *Journal of Plant Research*, **109**, 201-209.

Macauly Institute for Soil Research Aberdeen (1982) Soil survey of Scotland. 1:2500,000 sheet **6**, south west Scotland.

Manton I (1950) *Problems of cytology and evolution in the Pteridophyta*. Cambridge University Press, Cambridge.

Marshall DR, Brown HD (1981) The evolution of apomixis. *Heredity*, **47**, 1-15.

Maynard Smith J (1989) *Evolutionary genetics*. Oxford University Press, NewYork.

Maynard Smith J (1978) *The evolution of sex*. Cambridge University Press.

Mayr E (1963) *Animal species and evolution*. Belknap press, Cambridge, MA.

Mazzucato A, Barcaccia G, Pezzotti M, Falcinelli M (1995) Biochemical and molecular markers for investigating the mode of reproduction in the facultative apomict *Poa pratensis* L. *Sexual Plant Reproduction*, **8**, 133-138.

McAllister HA (1986) *The Rowan and its relatives (Sorbus spp.)*. Ness series 1. Ness Gardens, Neston.

McCarthy EM, Assumes MA, Anderson WW (1995) A theoretical assessment of recombinational speciation. *Heredity*, **74**, 502-509.

Meagher RB, Berry-Lowe S, Rice K (1989) Molecular evolution of the small subunit of ribulose bisphosphate-carboxylase from tobacco. *Nucleic Acids Research*, **13**, 2373-2386.

Menken SBJ, Morita T (1989) Uniclonal population structure in the pentaploid obligate agamosperm *Taraxacum albidum* Dahlst. *Plant Species Biology*, **4**, 29-36.

Menken SBJ, Smit E, and Dennijs HCM (1995) Genetical population structure in plants: Gene flow between diploid sexual and triploid asexual dandelions (*Taraxacum* section Ruderalia). *Evolution*, **49**, 1108-1118.

Mitchell-Olds T (1995) Interval mapping of viability loci causing heterosis in Arabidopsis. *Genetics*, **140**, 1105-1109.

Mogie M (1985) Morphological, developmental and electrophoretic variation within and between obligately apomictic *Taraxacum* species. *Biological Journal of the Linnean Society*, **24**, 207-216.

Mogie M (1992) *The evolution of asexual reproduction in plants*. Chapman and Hall, London, UK.

Morgante M, Olivieri A (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant Journal*, **3**, 175-182.

Muller HJ (1932) Some genetic aspects of sex. *American Naturalist*, **66**, 118-138.

Muller HJ (1964) The relation of recombination to mutational advance. *Mutation Research*, **1**, 2-9.

Nassar NMA, Vieira MA, Vieira C, Grattapaglia D (1998) Molecular and embryonic evidence of apomixis in cassava interspecific hybrids (*Manihot* spp.). *Canadian Journal of Plant Science,* **78**, 349-352.

NCC (1986) Glen Diomhan National Nature Reserve third management plan, 1986-1993. NNC 144/MP.

Nogler GA (1984a) Gametophytic apomixis. In: *Embryology of agamosperms*, (ed Johri BM) pp. 475-518. Springer-Verlag, Berlin.

Nogler GA (1984b) Genetics of apospory in apomictic *Ranunculus auricomus* V. conclusions. *Botanica Helvetica*, **94**, 411-422.

Noyes RD, Soltis DE (1996) Genotypic variation in agamospermous *Erigeron compositus* (Asteraceae). *American Journal of Botany*, **83**, 129-1303.

Nybom H, Schaal BA (1990) DNA 'fingerprints' reveal genotypic distributions in natural populations of blackberries and raspberries (*Rubus*, Rosaceae). *American Journal of Botany*, **77**, 883-888.

Ordnance Survey (1967) Ten mile map of Great Britain, north. 1:625,000. Rainfall. Annual average, 1916-1950.

Oritz JPA, Pessino SC, Leblanc O, Hayward MD, Quarin CL (1995) Genetic fingerprinting for determining the mode of reproduction in *Paspalum notatum*, a subtropical apomictic forage grass. *Theoretical and Applied Genetics*, **95**, 850-856.

Palmer JD (1991) Plastid chromosomes: structure and evolution. In: *Cell culture and somatic cell genetics of plants* (eds Bogorad L, Vasil K), pp. 5-53. Volume 7A, Academic, New York.

Palmer JD, Jansen RK, Michaels HJ, Chase MW, Manhart JR (1988) Chloroplast DNA variation and plant phylogeny. *Annals of Missouri Botanical Garden*, **75**, 1180-1206.

Palmer RG (1971) Cytological studies of ameiotic and normal maize with references to premeiotic pairing. *Chromosoma*, **35**, 233-246.

Parker ED Jr (1979) Ecological implications of clonal diversity in parthenogenetic morphospecies. *American Zoologist* **19**, 753-762.

Peet K (1974) The measurement of species diversity. *Annual Review of Ecological Systematics*, **5**, 285-307.

Pielou EC (1969) An introduction to mathematical ecology. Wiley-Interscience, New York.

Proctor MCF, Proctor ME and Groenhof AC (1989) Evidence from peroxidase polymorphism on the taxonomy and reproduction of some *Sorbus* populations in south west England. *New Phytologist*, **112**, 569-575.

Promega (1997) *Technical Manual*. Promega Corporation, WI, USA.

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

Rhoades MM (1956) Genetic control of Chromosome behaviour. *Maize Genetics Co-operative News Letter*, **30**, 38-42.

Rich TCG, Rich MDB (1998) The plant Crib. *Botanical Society of the British Isles*. London.

Richards AJ (1973) The origin of *Taraxacum* agamospecies. *Botanical Journal of the Linnean Society*, **66**, 189-211.

Richards AJ (1975) Sorbus. In: *Hybridisation and the flora of the British Isles* (ed Stace CA). Academic Press, London.

Richards AJ (1996) Genetic variability in obligate apomicts of the genus *Taraxacum*. *Folia Geobotanica and Phytotaxonomica* **31**, 405-414.

Richards AJ (1997) *Plant breeding systems*. Chapman and Hall, London.

Riesberg LH, Carney SE (1998) Plant hybridization. *New Phytologist*, **140**, 599-624.

Rieseberg LH, Ellstrand NC (1993) What can morphological and molecular markers tell us about plant hybridisation? *Critical Reviews in Plant Sciences*, **12**, 213-241.

Rieseberg LH, Wendel J (1993). Introgression and its consequences in plants. In: *Hybrid zones and the evolutionary process* (ed Harris R), pp. 70-109. Oxford University Press, New York.

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

Rieseberg LH, Baird SJE, Gardner KA (2000) Hybridisation, introgression and evolution. *Plant Molecular Biology*, **42**, 205-224.

Rieseberg LH, Carter R, Zona S (1990) Molecular tests of the hypothesised origin of two diploid *Helianthus* species (Asteraceae). *Evolution*, **44**, 1498-1511.

Rieseberg LH, Kim MJ, Seiler GJ (1999) Introgression between the cultivated sunflower and a sympatric wild relative, *Helianthus petiolaris* (Asteraceae). *International Journal of Plant Sciences*, **160**, 102-109.

Rieseberg LH, Soltis DE, Palmer JD (1988) A molecular re-examination of introgression between *Helianthus annuus* and *H. bolanderi*. *Evolution*, **42**, 227-238.

Rieseberg LH, Van Fossen C, Desrochers A (1995) Hybrid speciation accompanied by genomic reorganisation in wild sunflowers. *Nature*, **375**, 313-316.

Roose ML, Gottlieb LD (1976) Genetic and chemical consequences of polyploidy in *Tragopogon*. *Evolution*, **30**, 818-830.

Sano Y, Kita F (1978) Reproductive barriers distributed in *Melilotus* species and their genetic basis. *Canadian Journal of Cytology and Genetics*, **20**, 275-289.

Schemske DW, Husband BC, Ruckelshaus MH, Goodwillie C, Parker IM, Bishop JG (1994) Evaluating approaches to the conservation of rare and endangered plants. *Ecology* **75**, 584-606.

Scotlab (1996) *Extraction and purification protocols*. Nucleonbiosciences, Tepnel Life Sciences PLC, Manchester.

Shi Y, Gornall RJ, Draper J, and Stace CA (1996) Intraspecific molecular variation in *Hieracium* sect *Alpina* (Asteraceae), an apomictic group. *Folia Geobotanica and Phytotaxonomica*, **31**, 305-313.

Soltis DE, Rieseberg LH (1986) Autopolyploidy in *Tolmiea menziesii* (Saxifragaceae): genetic insights from enzyme electrophoresis. *American Journal of Botany*, **73**, 310.

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

Soltis PS, Doyle JJ, Soltis DS (1992) Molecular and polyploid evolution in plants. In: *Molecular systematics of plants* (eds Soltis PS, Soltis DS, Doyle JJ), pp. 177. Chapman and Hall, New York.

Song KM, Osborn TC, Williams PH (1988) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). *Theoretical and Applied Genetics*, **86**, 811- 821.

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, **98**, 503- 517.

Stace CA (1991) *New flora of the British Isles*. Cambridge University Press.

Stace CA, Gornall RJ, Shi Y (1997) Cytological and molecular variation in apomictic *Hieracium* sect. *Alpina*. *Opera Botanica*, **132**, 39-51.

Stebbins GL (1950) *Variation and evolution in plants*. Columbia University Press, New York.

Stebbins GL (1971) *Chromosomal evolution in higher plants*. Edward Arnold, London.

Sun M (1996) The allopolyploid origin of *Spiranthes honkongensis* (Orchidaceae). *American Journal of Botany*, **83**, 252-260.

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

Templeton AR (1981) Mechanisms of speciation – a population genetic approach. *Annual Review of Ecology and Systematics*, **12**, 218-217.

Thomas MR and Scott NS (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). *Theoretical and Applied Genetics* **86**, 985-990.

van Dijk PJ, Tas ICQ, Falque M, Bakx-Schotman T (1999) Crosses between sexual and apomictic dandelions (*Taraxacum*). II. The breakdown of apomixis. *Heredity*, **63**, 715-721.

Van Oostrum H, Sterk AA, Wijsman HJW (1985) Genetic variation in agamospermus microspecies of *Taraxacum* sect. *Erythrosperma* and sect. *Obliqua*. *Heredity*, **55**, 223-228.

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

Wan J, Yamaguchi Y, Kato H, Ikehashi H (1996) The new loci for hybrid sterility in cultivated rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*, **91**, 1153-1161.

Warburg EF (1938) The origin and distribution of the British forms of *Sorbus*. Report of the Annual Meeting, pp. 507-508. British Association for the Advancement of Science.

Weinberg W (1908) Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für Vaterländische Naturkunde in Württermburg*, **64,** 368-382.

Wendelbo P (1959) *Taraxacum gotlandicum*, a pre-boreal relic in the Norwegian Flora? *Nytt. Mag. Bot.*, **7**, 161-167.

Whittemore AT, Schaal BA (1991) Interspecific gene flow in oaks. *Proceedings of the National Academy of Sciences USA*, **88**, 2540-2544.

Williams GC (1975) *Sex and evolution*. Princeton University Press, Princeton, N.J.

Williams JGK, Hanafey MK, Rafalski JA, Tingey SV, (1993) Genetic analysis using random amplified polymorphic DNA markers. *Methods in Enzymology*, **218**, 704-740.

Wilmott AJ (1934) Some interesting British *Sorbi*. *Proceedings of the Linnean Society*, **146**, 73-79.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **6**, 97-159.

Wu CI, Palopoli M (1994) Genetics of post mating reproductive isolation in animals. *Annual Review of Genetics*, **27**, 283-308.

Wu KS and Tanskley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. *Molecular and General Genetics*, **241**, 225-235.

Yhara T (1990) Evolution of agamospermous races in *Boehmeria* and *Eupatorium*. *Plant Species Biology*, **5**, 183-196.

Yhara T, Ito M, Wantanabe K, Crawford DJ (1991) Very low genetic heterozygosities in sexual and agamospermous populations of *Eupatorium altissimum* (Asteraceae). *American Journal of Botany*, **78**, 706-710.

Appendix One Buffer recipes

1.1. Isozyme extraction buffer

8 % (w/v) PVP, 0.3 M sucrose, 0.5 mM EDTA, 1 mM dithiothreitol, 1 mM Ascorbic acid, 0.1 % (w/v) BSA, 0.2 mM Pyridoxal 5-phosphate, 0.4 mM NAD, 0.3 mM NADP

1.2 Isozyme gel and electrode buffers

Soltis, Histidine (pH 7.0) 3 x gel buffer, 4.19 g/l L- Histidine, pH to 7.0 with NaOH. Soltis, Histidine (pH 7.0) electrode buffer, 117.64 g/l Citric acid, pH to 7.0 with NaOH.

Ashton, Lithium borate (pH 8.3) gel buffer, 6.2 g/l Trisma base, 1.46 g/l Citric acid (add 27 ml of electrode buffer to 243 ml of gel buffer for gel preparation). Ashton, Lithium borate (pH 8.3) electrode buffer, 1.2 g/l Lithium hydroxide, 11.9 g/l Boric acid.

B gel buffer, 1.2 g/l Trisma base, pH to 8.5 with citric acid. B electrode buffer, 2 g/l NaOH, 18.5g/l Boric acid.

Morpholine citrate (pH 6.1) gel buffer (dilute electrode buffer x 20). Morpholine citrate (pH 6.1) electrode buffer, 7.69 g/l citric acid (anhydrate), pH to 6.1 with N-(3-aminopropyl)-Morpholine (approximately 10ml).

Lithium borate (pH 8.2) gel buffer, 1.05 g/l Citric acid, 3.8 g/l Trisma base (add 10ml l^{-1} of electrode buffer to gel buffer for gel preparation, pH 8.2 with Trisma base). Lithium borate (pH 8.1) electrode buffer, 2.51 g/l Lithium hydroxide, 19.16 g/l Boric acid.

1.3. DNA extraction buffer

ATMAB (Alkyltrimethylammonium bromide)	20 g/l
PVP 4000 (soluble)	10 g/l
EDTA 0.5 M (pH 8)	7.44 g/l
Tris HCl 1 M (pH 8)	12.11 g/l
NaCl 5 M	31.81 g/l

Appendix Two Bird survey data

2.1 Bird feeding times from timed watches on two *S. aucuparia* and two *S. arranensis* trees

	S. aucuparia	*S. arranensis*
Observation time (min)	349	400
Time with at least one bird feeding	60%	22%
Number of bird Visits	242	64
Average number of berries eaten per visit	9	3

Appendix Three DNA cloning methods

3.1 Ligation of Inter-SSR fragments into plasmids

Reagents

pGEM^{R}-T Easy Vector System II:

pGEM^{R}-T Easy Vector (50 ng/ul)	Control Insert DNA (4ng/ul)
T4 DNA Ligase	T4 Ligase 10 X Buffer
JM 109 Competent Cells, high efficiency	

Methods

Ligase reactions:

	Standard Reaction	Positive Control	Background Control
T4 DNA Ligase 10 x buffer	1 ul	1 ul	1 ul
pGEM^{R}Easy Vector (50ng)	1 ul	1 ul	1 ul
Isolated PCR fragment	7 ul	-	-
Control Insert DNA	-	2ul	-
T4 DNA ligase	1 ul	1 ul	1ul
Deionised H_20	0 ul	5 ul	7 ul
Final volume	10 ul	10 ul	10 ul

Mix reactions by pipteting gently. Incubate reactions overnight at 4 °C.

3.2. Transformations of *Escherichia coli* JM 109 Competent Cells

Reagents

A composition of LB medium, SOC medium, IPTG, and X-Gal Buffers and solutions used were exactly as described in Appendix B, Promega Technical Manual for pGEM^{R}-T and pGEM^{R}-T Easy Vector systems Part #TM042 (1997).

Method

1. Prepare 2 LB/ampicillin/IPTG/X-GAL plates for each ligation reaction.
2. Centrifuge the tubes containing the ligation reactions to collect contents at the bottom of the tube. Add 2 ul of each ligation reaction to a sterile 1.5 ml microcentrifuge tube on ice.
3. Remove tube of frozen JM 109 High Efficiency Competent Cells from – 70 °C storage and place them in an ice bath until just thawed (about 5 mins). Mix cells by very gently flicking the tube.
4. Carefully transfer 50 ul of cells into each tube prepared in step 2. Avoid excessive pipetting as the cells are very fragile.
5. Gently flick the tubes to mix and place them on ice for 20 min.
6. Heat-shock the cells for 40-45 seconds in a water bath at exactly 42 °C (do not shake).
7. Immediately return the tubes to ice for 2 minutes.
8. Add 950 ul of room temperature SOC to the tubes (LB broth may be substituted, but colony number may be lower).
9. Incubate for 1.5 hours at 37 °C with shaking (~ 150 rpm)
10. Plate 300 ul of each transformation culture onto antibiotic plates.
11. Incubate the plates overnight (16-24 hours) at 37 °C. Longer incubations or storage of plates at 4 °C (after 37 °C overnight incubation) may be used to facilitate blue/white screening. White colonies generally contain inserts however, inserts may also be present in blue colonies (Promega 1997).

Population sizes of transformed *E-coli* cells were increased to allow sufficient recombinant plasmid DNA to be extracted:

Reagents:

Falcon polypropylene tubes (50ml)
Ampicilin 50 mg/ml
LB medium
Sterile tooth picks
Qiagen Plasmid Miniprep Kit

Method

Recombinant clones were identified by colour screening on indicator plates as described in step 11 above. A single white colony, from each plate, was isolated using a sterile wooden toothpick and used to inoculate 5 mls of LB (in a 50-ml falcon tube with Ampicilin added). Solutions were incubated overnight, with shaking (~ 150 rpm), at 37 °C.

Tubes were spun at 2000 rpm for 10 min and the supernatant decanted. Extraction buffer (supplied with the Qiagen Plasmid Miniprep kit) was added to the pellet. Pellets were resuspended by flicking the tube. The Plasmid DNA containing the isolated PCR inserts was extracted according to manufactures instructions (Qiagen)

To confirm the insertion and cloning process of the Inter-SSR PCR fragments had been successful, 5 ul of plasmid DNA was digested with restriction enzyme EcoR 1. EcoR1 restriction sites flank the plasmid insert site (Fig 2). The cutting of the plasmid at these sites will release the cloned insert. The restriction digests products were run out on agarose gels and stained with ethidium bromide to test for the presence of the insert.