Evolutionary relationships and reproductive ecology of endemic *Sorbus* species in south west UK: Implications for conservation.

Submitted by Tracey Jane Hamston to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Biological Sciences
In September 2016

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: …………………………………………………..

Abstract

The genus *Sorbus* is an example of a taxonomically complex group (TCG) with diversity derived from hybridisation, polyploidy and apomixis. The focus of this study was to elucidate the evolutionary relationships among nine *Sorbus* species including endemics of the Devon and north Somerset region of the south west UK, determine main routes of polyploid formation and investigate reproductive sustainability in order to make recommendations for *Sorbus* conservation.

Molecular analysis showed that genetic structure patterns and genotypic diversity support the hypothesis that the study polyploids are a product of rare interspecific hybridisation, of single origins and are maintained through apomictic reproduction. PCoA, Neighbour Joining analysis and parental simulations reveal a reticulated relationship, with diversification the result of hybridisations between sexual diploid *Sorbus torminalis* and both tetraploid and triploid species. Hybridisation between *S. torminalis* and tetraploid *Sorbus margaretae* (subgenus *Aria*) have likely given rise to the study members of subgenus *Tormaria* through production of a triploid which has subsequently backcrossed to *Sorbus torminalis* to form further tetraploids. The discovery of a cryptic hybrid in subgenus *Aria* also suggests occasional hybridisation events among tetraploids are a possible route for further tetraploid formation These events illustrate key routes of polyploid formation, both illustrating the role of triploids in tetraploid formation via the triploid bridge and the key role in sexual diploids in diversification in *Sorbus.* Hand pollination experiments showed that self-incompatibility in the triploid species (*Sorbus subcuneata*) means reliance on congeneric pollen from sympatric tetraploid species for seed production.
2

Reproductive sustainability in this species is severely compromised through spatial isolation from compatible congeners. Our findings are strong support for the development of conservation strategies that aim to safeguard current diversity through actions that increase reproductive sustainability and recruitment opportunities, and promote opportunities for on-going hybridisation for future diversification of *Sorbus* in this region.

3

Acknowledgments

4

5

CONTENTS

Chapter 2:   Apomixis and hybridisation drives reticulate evolution

and phyletic differentiation in *Sorbus* L.: Implications for

conservation **80**

Chapter 3: Inheritance patterns and hybrid origins of apomictic polyploids: *Sorbus* in the southwest UK.............................. 143

8

9

LIST OF FIGURES

Chapter 1 **General Introduction**

Chapter 2   Pseudogamous apomixis and hybridisation drives
reticulate evolution and phyletic differentiation in
***Sorbus* L.: Implications for conservation.**

16

LIST OF TABLES

17

Chapter 1: General Introduction

1.1   Plant conservation

The Global Strategy for Plant Conservation 2011-2020 (GSPC) grew out of the Convention on Biological Diversity (CBD) with the aim of supporting global plant conservation and halting the continuing loss of plant diversity. It is estimated that 25% of plant species globally are threatened with extinction (https://www.bgci.org/plant-conservation/threats/) and in order to meet set targets within the strategy, objective one requires assessment of the conservation status of all known plant species to guide conservation action. Where status has been evaluated, as in the IUCN red lists of threatened species, this often forms the basis for prioritising national conservation efforts and resources (IUCN, 2001). However, the determination of factors such as population size, demography and distribution of species, depends on the ability to recognise delineated species, primarily in the field. For complex plant groups which are undergoing active diversification, taxa are often not well differentiated phenotypically and genetically. Where such complexes contain species or variants of conservation concern, problems arise for both conservationists and taxonomists. There are frequently problems with species delineation due to a lack of clear morphological differences between taxa, causing much debate over what constitutes a species with a knock-on effect on conservation policy (Hollingsworth, 2003, Zink, 2004, Frankham *et al*., 2012) and representation within conservation strategies (Hollingsworth et al., 2006, Rich et al., 2008). Whilst we cannot expect to effectively conserve species if we cannot recognise and describe them (Mace, 2004), for species complexes, conserving individual entities ignores the fundamental evolutionary processes underlying such diversification. Taxonomic complexity has led to difficulties in implementing

appropriate conservation plans for groups such as *Epipactis, Euphrasia* (Hollingsworth *et al.*, 2006) and *Hieracium* (Rich *et al.*, 2008).Therefore a more process-based approach to the conservation of taxonomically complex groups (TCG's) has been proposed (Ennos *et al.*, 2012), whereby the focus is the evolutionary process of diversification and speciation rather than selected rare or threatened components of a larger complex. This approach appears to be a more holistic and pragmatic solution, particularly where resource allocation for many individual priority species may be unreasonable. However, in order to apply this approach, we firstly need to understand the process that gives rise to biodiversity and elucidate factors that impact upon the process before threats can be evaluated and appropriate conservation strategies put in place. Therefore the study of diversifying TCG's represents opportunities to both investigate evolutionary divergence and contribute to a more informed approach to their conservation.

1.2 Taxonomic complexity and diversification

Taxonomically complex groups are often characterised by departure from random sexual mating and have mechanisms allowing rapid diversification such as hybridisation and polyploidy (Squirrell *et al.*, 2002). These factors often produce complex reticulated evolutionary relationships confounding their placement in hierarchical classification (Hörandl *et al.*, 2009). In Britain, such mechanisms have allowed for the development of endemic taxa since recent ice ages. Indeed, a large proportion of the British endemic higher plant species are contained within TCG's (Hollingsworth *et al.*, 2006), many associated with hybridisation, uniparental lineages and polyploidy (Stace, 1975).

21

1.3   Hybridisation and polyploidy

Sexual hybridisation between divergent lineages is a significant route to plant speciation but is often given as a reason why species definitions are hotly debated (Grant, 1981, Rieseberg & Willis, 2007) since hybrids are often poorly defined morphologically and genetically. Because of hybrid speciation, the history of plant evolution forms a reticulated network for many groups, rather than a 'tree of life' (Linder & Rieseberg, 2004). Hybridisation may occur between species but also divergent populations of the same species (Soltis & Soltis, 2009, Abbott *et al.*, 2013). There may be no resulting change in chromosomal number (homoploid speciation) or on occasion hybridisation may involve duplication of the genome (polyploid speciation) or indeed occur between species of different ploidy where sterility is usually the result and further genome duplication is required to restore fertility (Soltis & Soltis, 2009). In order for speciation to occur, novel hybrids must be able to establish and their genomic integrity protected through isolation mechanisms. Sympatric homoploid hybridisation will often be followed by backcrossing, creating hybrid swarms, if sexual reproduction is maintained (Mallet, 2007). For example, offspring of hybridisation between *Geum rivale* (Rosaceae) and *Geum urbanum* show an entire spectrum of genetic and morphological variation between the parent species (Ruhsam *et al.*, 2013). Barriers to gene exchange are required for speciation to occur. This may be brought about by gene expression changes induced by hybridisation that allow rapid ecological and spatial divergence from the parent species through natural selection (Abbott *et al.*, 2013) as seen in the recently formed diploid hybrid species *Senecio squalidus* (Hegarty *et al.*, 2009, Abbott *et al.*, 2010).

22

Polyploid hybridisation results in offspring with three or more sets of homologous chromosomes that are strongly, although often incompletely reproductively isolated from the parental taxa due to differing ploidy (Abbott *et al.*, 2013). As such, polyploidy has long been recognised as playing an important role in speciation and plant evolution (Hegarty & Hiscock, 2007, Soltis & Soltis, 2009). In fact, it is estimated that 70-80% of plant species have polyploid origins (Soltis *et al.*, 2004) as determined through stomatal size of fossilised angiosperms (Masterson, 1994), which tends to be significantly larger in polyploids, and the frequency of even numbers compared to odd, of haploid chromosomes present in current species (Otto & Whitton, 2000). This high frequency of polyploidy suggests an evolutionary advantage in possessing more than one genome. It is generally considered that increased fitness is likely due to fixed heterozygosity (Brochmann *et al.*, 2004) together with a greater pool of genes and alleles for natural selection (Hegarty & Hiscock, 2007). Indeed, the frequency of polyploids that are found in different environments to that of their diploid progenitors and that also possess novel life history characteristics (Ramsey & Schemske, 1998) suggests they often have the adaptive capacity to colonise new environmental niches (McIntyre, 2012, Laport *et al.*, 2013, Theodoridis *et al.*, 2013). Despite the ubiquity of polyploidy there are still gaps in knowledge regarding formation and establishment of naturally occurring polyploids (Ramsey & Schemske, 1998, Soltis *et al.*, 2016).

1.4 Polyploid formation

Polyploids may arise within populations of individual species (autopolyploidy) or as a result of interspecific hybridisation (allopolyploidy) (Ramsey & Schemske, 1998). There is debate about reliable indicators by which each of these types of polyploids are recognised and this classification is somewhat dependant on the

23

criteria used to define taxa as species. Ramsey & Schemske (2002) base recognition on the degree of pre and / or post-zygotic reproductive isolation according to the biological species concept. Since this is also open to interpretation and varies among taxa, these two terms (allopolyploidy and autopolyploidy) represent extremes of a spectrum (Obbard *et al.*, 2006) with many polyploids occupying points in between.

Two distinct routes of polyploid formation are recognised; somatic doubling and the production of unreduced ($2n$) gametes. Polyploid cells are thought to exist in the non-meristematic tissues of most angiosperms (Bennett, 2004) however chromosome doubling in meristem tissue may give rise to tetraploid shoots (Ramsey & Schemske, 1998). Gametes with the somatic chromosome number ($2n$ gametes) is considered to be the most common route to polyploid formation (Bretagnolle & Thompson, 1995). Whilst it is known that the production of unreduced gametes is frequent in plants, this frequency is highly variable even within species (Bretagnolle & Thompson, 1995, Burton & Husband, 2001, Carputo *et al.*, 2003) with the tendency of sexual diploids to produce such gametes being genetically inherited and environmentally induced (Brownfield & Köhler, 2010). A review by Ramsey & Schemske (1998) suggested that the natural rates of non-reduction are similar in microsporogenesis and megasporogenesis.

Potentially, unreduced gametes could fuse with other unreduced or reduced gametes, providing pathways to higher ploidy (Ramsey & Schemske, 1998). The formation of autotetraploids from diploids in one step, via the fusion of *2n* gametes, is thought to be rare in natural populations since it involves the combination of unlikely events (Husband, 2004). Allotetraploids have been

produced directly from crosses between distinct diploid subspecies of *Dactylis glomerata* (Sato *et al.*, 1993). However, increasingly, the study of natural, mixed ploidy systems is revealing the role that triploids play in the production of tetraploids (Husband, 2004, Robertson *et al.*, 2004b, Sabara *et al.*, 2013). These often occur in zones where diploids and tetraploid ranges overlap and provide an intermediate step to ongoing tetraploid formation through backcrosses with diploid progenitors, a phenomenon known as the 'triploid bridge' (Ramsey & Schemske, 1998). However, reproductive isolation from progenitors is required for polyploid hybridisation to result in new species.

1.5   Polyploid speciation

Traditionally, polyploid speciation has been described as instantaneous, as incompatibilities between ploidy levels leading to hybrid offspring sterility are expected to cause immediate reproductive isolation of the newly formed polyploid (Soltis & Soltis, 2009). Difficulties arising from the pairing of odd-numbered chromosomes at meiosis and the formation of unbalanced endosperm tissue can both restrict backcrossing with diploid progenitors. The term 'triploid block' has been used to describe this scenario when diploid × tetraploid hybridisations produce sterile triploid offspring (Ramsey & Schemske, 1998). In this case it is likely that rare tetraploids would suffer minority cytotype exclusion in a larger diploid population, as most pollination events would lead to sterile or no offspring, leading to eventual extinction (Levin, 1975). However, triploids are often found in mixed ploidy populations and are rarely completely sterile, as discussed earlier and Felber & Bever (1997) theorised how the evolution of tetraploids in diploid populations would depend on the relative frequency and fitness of both the tetraploid and the diploid × tetraploid hybrid. Empirical evidence for this was provided by Burton & Husband (2000) who

compared seed production of various ploidy hybrids resulting from crosses between naturally occurring diploid and tetraploid *Chamerion angustifolium.*

1.5.1 Reproductive isolation

Polyploid speciation requires reproductive isolation of novel polyploids from their progenitors. Where this is the case, it is usually unclear whether this is due to polyploidy *per se* or whether barriers have formed after the hybrid event (Sobel *et al.*, 2010). Barriers may be pre-zygotic, preventing zygotes from forming. Some may occur pre-pollination such as geographic isolation, flowering asynchrony, pollinator fidelity (Husband & Schemske, 2000) and climatic niche separation (Thompson *et al.*, 2014). Post pollination barriers such as pollen incompatibility and apomixis (Ludwig *et al.*, 2013) may also prevent hybridisation. Post zygotic failure is often due to endosperm imbalance whereby there is deviation of the maternal to paternal genome ratio (from 2 maternal: 1 paternal) in the endosperm tissue, causing failure of the endosperm in interploidy crosses (Köhler *et al.*, 2010). Reproductive isolation mechanisms that lead to speciation are usually multiple and the primary isolating mechanism is often difficult to determine, particularly for species that have been long established (Sobel *et al.*, 2010). Where isolating mechanisms are partial it allows for occasional hybridisations resulting in a reticulate network of taxa. For this reason we may view the production of polyploid complexes as a balance of isolating factors. Polyploid speciation depends on sufficient reproductive isolation to establish novel lineages, but for ongoing hybridisation and diversification to occur strong isolating barriers may inhibit aspects of this evolutionary process.

1.5.2 Evolution of new breeding systems: Polyploidy and apomixis

The transition from sexual to an apomictic breeding system is often associated with polyploidization (Otto & Whitton, 2000). Apomixis (here synonymous with agamospermy; asexual seed production) overcomes some reproductive problems associated with polyploidy described above, since it bypasses meiosis during female gamete formation. It also enables novel polyploids to establish reproductively isolated populations in sympatry with progenitors.

Breeding systems shape the nature of evolution with flowering plants displaying a huge variety of reproductive systems, which give rise to countless patterns of variation even among populations of the same species (Briggs & Walters, 1997, Richards, 1986). A combination of asexual and sexual systems, largely within perennial plant populations, allows for adaptation to different ecological settings due to the benefits imbued by each breeding system. This may be illustrated by the fact that many asexual or apomictic plants co-occur with their sexual counterparts with complete or obligate apomixis a rare phenomenon (Richards, 2003, Vallejo-Marin *et al.*, 2010). The short term advantages of apomixis such as reproductive assurance allowing for rapid colonisation, results in largely clonal populations of low diversity. However, even where apomixis is thought to be obligate, clonal lineages have been shown to display divergent mutational variation (Paun *et al.*, 2006, Majesky *et al.*, 2012) adding to their genetic and ecological variability. Theory predicts that obligate apomixis is only of short term advantage over sexuality due to the lack of adaptive potential and the fact that it does not predominate over sexuality in major systematic groupings (Richards, 2003). Nevertheless, obligate apomicts may still produce functional pollen, retaining some sexuality in their male function.

27

1.6 Methodological challenges for genetic studies of apomictic polyploids.

1.6.1 Molecular markers

Apomictic, polyploid organisms present particular challenges for selection of appropriate markers for use in molecular ecology. They have complex genomes potentially with multiple copies of alleles, indeed one of the main problems when working with polyploids is the uncertainty of allele dosage which presents difficulties for calculating frequency based statistics. Polyploids also have mixed inheritance patterns (disomic and polysomic) sometimes at different loci (Otto & Whitton, 2000), for reasons explained in section 3.1. When they also exhibit mixed breeding systems (apomixis, obligate and facultative, and sexual) they are unable to exhibit random mating. Both these factors mean they violate the assumptions for Hardy-Weinberg equilibrium (Freeland *et al.*, 2011) another prerequisite for most population genetic analysis. The majority of analyses for population genetics were developed for sexual diploids, and present difficulties for use in polyploids. A recent review of molecular and statistical tools available for use with polyploids reveals that the newer sequencing technologies are still hampered by these problems (Dufresne *et al.*, 2014).

Microsatellite markers are tandemly repeated sequences typically shorter than 100bp (Schlötterer, 2004). They are still often a marker of choice for population genetic studies as they are highly polymorphic due to high mutation rates, which makes them suitable for use in closely related species, such as in this study. The advantages and disadvantages of microsatellites for use in molecular ecology compared to other markers have been thoroughly reviewed in many texts e.g. Freeland *et al.* (2011) and Lowe *et al.* (2009). Microsatellites remain a

28

popular option despite the huge advances in sequencing technology over recent years. Indeed, next generation sequencing technology now enables detection of large numbers of microsatellites, with high throughput, reducing the cost and time traditionally associated with the development of these markers in addition to reducing genotyping errors (Zhan *et al.*, 2016).

The problems of uncertain allele dosage common to polyploids exacerbate some of the disadvantages of microsatellites such as genotyping error due to stutter bands and increased chances of size homoplasy due to increased numbers of potentially similar sized alleles. Null alleles may remain undetected since the software for identification of null alleles, scoring of stutter peaks and allele dropout uses allele frequencies based on diploids; for example MICROCHECKER (Van Oosterhout *et al.*, 2004). Despite these problems microsatellites are still powerful tools which have been widely used for population studies on polyploid organisms with sexual and/or apomictic breeding systems, for example; (Andreakis *et al.*, 2009, Cunha *et al.*, 2011, Garcia-Verdugo *et al.*, 2013). The problem of unknown allele dosage means they are often scored as dominant markers with alleles scored as present or absent which reduces their usefulness as co-dominant markers. This problem has been addressed in some studies by calculating the likely frequencies of alleles using the MAC-PR (microsatellite DNA allele counting-peak ratios) method devised by Esselink *et al.* (2004) and outlined in Chapter 3. This enabled exploration of microsatellite inheritance patterns in allopolyploid *Bordera spp.*(Catalán et al., 2006), determination of genotypic configurations and comparison of genetic diversity of diploid, triploid and tetraploid *Crataegus spp.*(Lo *et al.*, 2009). This method becomes less reliable at higher ploidy levels beyond tetraploid and depends on each sample producing clear repeatable

29

electrophoretic peaks during fragment analysis as well as sufficient comparable allele pairs of which to compare peak area ratios. More recently developed software specifically designed to analyse polyploid microsatellite data e.g.POLYSAT (Clark & Jasieniuk, 2011) adopt various ways of estimating allele frequencies, see Dufresne *et al.* (2014) for a review of these approaches.

1.6.2 Development of molecular methods for this study

I decided to use a selection of previously developed nuclear microsatellite markers already shown to be useful in determining evolutionary relationships of mixed ploidy groups of *Sorbus;* see Robertson *et al.* (2010) and González-González *et al.* (2010). In order to minimise scoring errors due to stutter peaks and unknown allele numbers we set up a sequence of checks. Initially, a test panel of the nine species were used test the optimal specificity of each primer pair based on the literature. A 4 μl aliquot of each PCR product was subjected to electrophoresis on a 1% agarose gel containing ethidium bromide and the resulting bands were visualised by UV illumination (G:BOX Sygene, Cambridge UK). Negative controls were always run to test for contamination. For loci that produced clean consistent bands I selected 14 forward primers for labelling with a florescent tag (WellRED D2, D3 or D4, which are indicated as black, green and blue, respectively). Capillary electrophoresis of the PCR products was carried out on an eight capillary Beckman Coulter sequencer (Beckman Coulter, Fullerton, USA) and fragment analysis, to determine allele size, was performed using CEQ 8000 Genetic Analysis system (Beckman Coulter) followed by a manual verification of each call to ensure proper peak designation. Primer pairs were grouped into three multiplex groups according to size and tag colour and re-optimised using a touchdown PCR cycle; see Tables S2.2 and S2.3 for details. Amplification and fragment analysis of the test panel was repeated

several times to check for consistency of allele scoring and genotyping. Automated binning of the raw allele lengths into size classes defined by the repeat unit of each locus was set up based on the size ranges of the diploid samples since these were most variable.

The electropherograms of the diploids were used to characterise the allele peak shapes at each locus and these were used as references when deciding which peaks represented true alleles rather than stutter peaks for the polyploid samples. For those polyploid species with mixed genomes from both diploid species it was occasionally visible which peaks were inherited from which genome (for an example see Fig. 3.4). We predicted the maximum allele number through the use of flow cytometry (Chapter 2) to determine ploidy level for the majority of samples. Unfortunately, it was not possible to confirm allele dosage for all samples at all loci using the MAC-PR method due to variation in the peak quality of the samples, however, the mean values for the polyploid species allowed for resolution of allele dosage at some loci at a species level for use in the parentage simulations of Chapter 3.

1.6.3 Molecular analysis of mixed ploidy groups containing sexual and apomictic individuals.

The violation of the basic assumptions of random mating and gene frequencies remaining constant from one generation to the next makes many traditional measures of population genetic diversity inappropriate for polyploid apomicts where individual loci are not free to recombine. The genetic studies of many fungal and bacterial pathogens commonly encounter these issues and the literature concerning these groups can be fruitful for investigating suitable approaches to genetic analysis in this situation. For the investigation of

31

population structure the multiple occurrences of the same multilocus genotypes (MLG's) can complicate analysis, therefore 'clone mates' are censored (Milgroom, 1996) and all duplicates removed from the analysis which is then performed on single MLG's. This accounts for the lack of non-random mating in populations caused by asexual (or apomictic in this case) reproduction of many clones and more closely approximates panmictic populations (Kamvar *et al.*, 2014).

Analysis of diversity cannot be based on allele frequencies since these are not generally known in polyploids (allele dosage uncertainty) and therefore involves calculation of the numbers of genotypes observed (richness), evenness and diversity. Thus, typically studies use diversity measures from ecology such as Shannon-Wiener, or Stoddart and Taylor (Arnaud-Haond *et al.*, 2005, Kamvar *et al.*, 2014), although there are problems with these when comparing genotypic diversity between different sized samples and when diversity is low (Grünwald *et al.*, 2003). A review of the commonly used richness, evenness and diversity measures is given by Arnaud-Haond *et al.* (2007). The most widely used index measuring clonal richness and evenness is the Simpson index **(**λ) which can be modified to account for sample size and vary positively with heterogeneity by using the complement (1- λ), see Chapter 2.

The remaining methodological approaches are detailed in the relevant following chapters.

1.7   Pollination and population sustainability

When insect pollination is required for both seed production in the first instance polyploid formation via hybridisation, it is both ecologically and evolutionarily important. This is discussed in detail in Chapter 3; I therefore only provide a short overview of a number of pertinent factors here. For apomictic plants that rely on pollination for seed set (pseudogamy; see 1.7.2), pollen transfer between anther and stigma may be assured if there are sufficient pollinators and the plants are self-compatible. However, when seed production is dependent on pollen movement between plants, i.e. for self-incompatible species, sufficient pollen movement becomes a requirement for population sustainability. It is also a requirement for hybridisation to occur between different lineages (sometimes species) thus sustaining an evolutionary process dependent on hybridisation.

Pollen flow in the landscape is affected by a number of factors. It is well known that pollinator mediated gene flow is heavily influenced by distance, with the majority of pollen moving relatively short distances. Dispersal patterns typically follow leptokurtic decay from a point source (Fig. 1.1) which also means a small amount of long distance pollen dispersal still occurs (fat-tailed curve) (Cresswell, 2006). The implications are that both reproductive sustainability and frequency of hybridisation events will be greatly affected by fragmentation and isolation of progenitors, although the fat-tail to the dispersal curve will mean that for large, long lived perennial plants a small amount of long distance pollen flow may sustain isolated individuals (Lander *et al*., 2010). The relative numbers of flowers produced by progenitors may also affect the likelihood of pollinator mediated gene flow since increased floral display size results in increased

33

geitonogamous (within plant) pollination (Richards *et al.*, 1999). For self-incompatible plants this results in high pollen wastage and failure of seed production. Differing relative fitness and pollen viability of sympatric progenitors may play a significant role in determining the frequency and nature of hybridisations via the impact of 'pollen pressure' i.e. the relative abundance of pollen from different hybridising species in the pollen cloud. For example the frequency of hybridisation between diploid and tetraploid cytotypes was high when the diploids were in a minority (Hajrudinović *et al.*, 2015b), with an increased probability of tetraploid pollen successfully fertilising a diploid egg cell despite the likely poor outcome of this interploidy cross (Ludwig *et al.*, 2013).
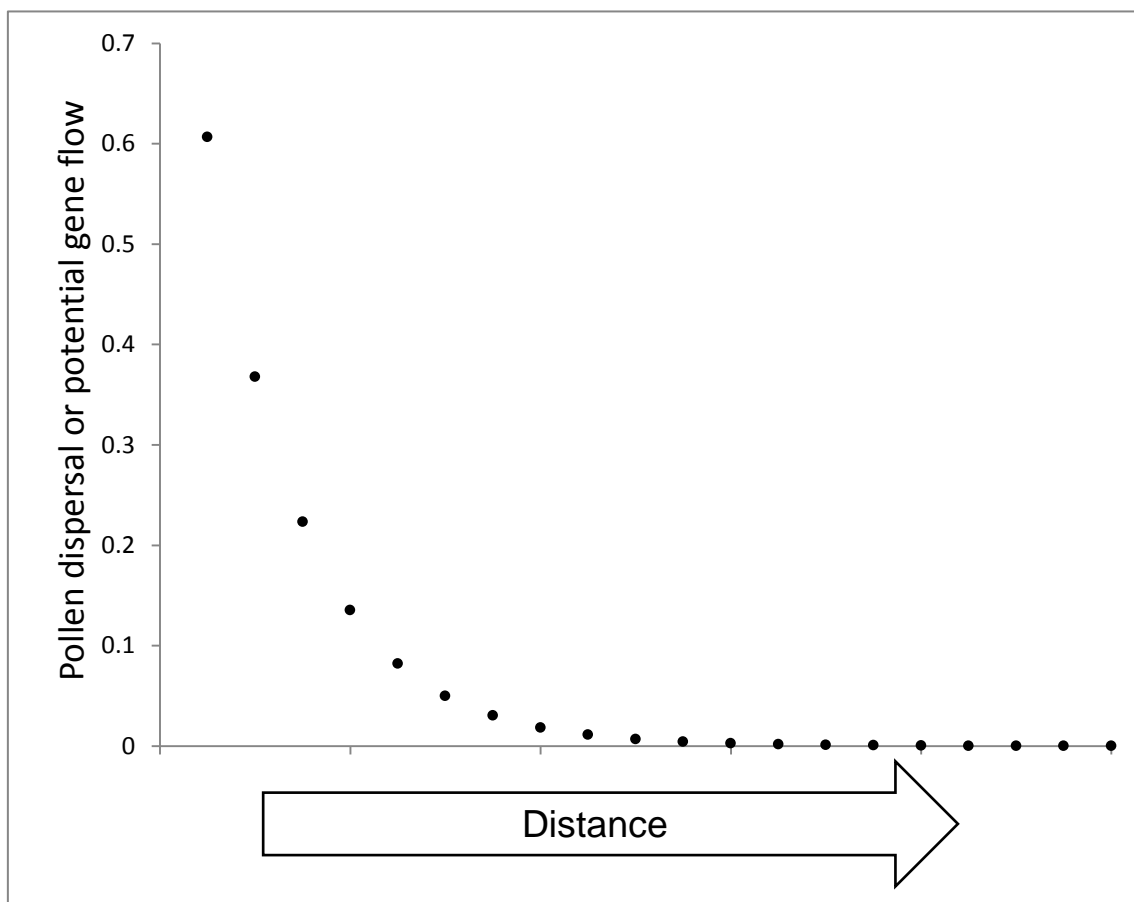
Figure 1.1 Typical pollinator mediated gene flow model using an exponential power function with leptokurtic decay.

For TCG's that rely on pollen movement for both reproductive sustainability and diversification through hybridisation, factors that affect pollen flow in the landscape are of great importance and form a vital component of the study of these groups.

1.8 *Sorbus* as a model system of a Taxonomically Complex Group

Summary

The genus *Sorbus* L. (Rosaceae) is contained within the subtribe *Pyrinae* (formerly the *Maloideae*) and is an example of a taxonomically complex group (TCG), due to a combination of interspecific hybridisations, associated polyploidy and mixed breeding systems which drive diversification and reticulated evolution (Liljefors, 1955, Richards, 1975, Nelson-Jones *et al.*, 2002, Rich *et al.*, 2010, Robertson *et al.*, 2010). It is a good example of a genus with recently diverging species complexes; novel genotypes have been recorded within the last 100 years e.g. *Sorbus × motleyi* (Rich & Proctor, 2009) thus it makes a useful model for research into the evolutionary processes of TCG's.

Gene flow within this genus is dependent on pollen transfer between individuals by pollinating insects and seed dispersal. Therefore both pollen flow and seed production are vital components in both sustaining current diversity of *Sorbus* populations and in the generation of future diversity for long term evolution and adaptation in a changing environment. Successful hybridisation events are rare with reproductive isolation occurring at the individual (for the apomictic species) and species level. The components of reproductive isolation may be pre and post-zygotic and are likely a combination of ecological and non-ecological. However, the contributions of different reproductive barriers in *Sorbus* are poorly understood.

35

The interest in gaining a better understanding of evolutionary processes in *Sorbus* also stems from a desire for the development of more informed conservation strategies which are currently mainly targeted at individual species of conservation concern. Therefore, we need to both unravel the process of diversification and try to elucidate environmental factors that may affect the survival of the current diversity and future evolutionary potential if we are to conserve *Sorbus* diversity in perpetuity.

This research project aims to inform a process-based conservation strategy for *Sorbus* in Britain by investigating further how diversification occurs in this genus and what barriers may exist to ongoing hybridisation events.

1.8.1 Life history

Sorbus is a genus of small, mostly deciduous trees with a mainly temperate distribution across the northern hemisphere (Aldasoro *et al.*, 1998). Flowering occurs in the spring with the timing of flowering varying with species, year and situation (Rich *et al.*, 2010). The flowers are pollinated by insects, mainly bees (Ludwig, 2013) and borne on inflorescences arranged as slightly convex panicles to corymbs (Fig. 1.2). *Sorbus* flowers are hermaphrodite with most having five petals and sepals and 10 - 20 stamens. They appear to be slightly protandrous, with anthesis occurring some hours before maturation of the stigma (pers. obs.). Figure 1.3 shows an inflorescence of *S. admonitor* with yellow anthers at anthesis and the immature stigma.

Figure 1.2 *Sorbus vexans* inflorescence with close up of single flower with creamy pink anthers. Oxen Tor, north Devon.

Figure 1.3 *Sorbus admonitor* in flower.

The two styles (more in some species) may be fused together, at least at the base and carpels number two to five, each forming a locule that contains two ovules (Bednorz, 2007) (Fig. 1.4; *S. torminalis* has two carpels). Pollen size and morphology differs between subgenera with pollen from *S. aucuparia* and associated species being the smallest (Boyd & Dickson, 1987, Bednorz *et al.*, 2005).
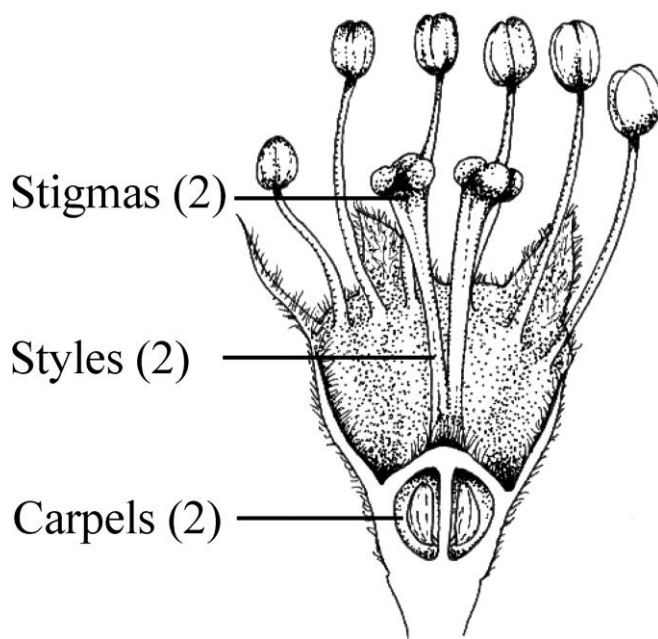
Figure 1.4 Illustration of hermaphroditic flower of *Sorbus torminalis*. Illustrator: Gaëtan Oddou (Oddou-Muratorio *et al.*, 2006).

Fruits (pomes) vary in colour, shape and size and are a diagnostic feature used for identification. There is variability in the production of fruit without developed seed. Mostly, these are aborted, but some species e.g. *S. minima, S. bristoliensis* (Rich *et al.*, 2010) and *S. subcuneata* (pers. obs.) produce fruits of two sizes (notably these species are all triploid). The smaller fruits are either without seed or contain undeveloped seed and the larger fruit are as other species which contain 1 to 4 seed, rarely 5 (Hajrudinović *et al.*, 2015b). Seed is

38

dispersed primarily by birds and small mammals which feed off the fruits in autumn (Snow & Snow, 1988, Bednorz, 2007).

Figure 1.5 *Sorbus devoniensis* in fruit. Little Haldon (lat.50.56476, long.-3.52589) south Devon.

The habitat type for *Sorbus* varies widely. Rich *et al.* (2010) provides an overview of the ecology of British and Irish *Sorbus*. The key points to note are that the three native British sexual diploid species, *S. aria, S. aucuparia* and *S. torminalis* are associated with different habitat types and this is probably reflected in their hybrid derivatives although the author is not aware of any detailed studies to investigate niche differentiation of these intermediate species. The majority of British *Sorbus* appear to prefer open sunlit conditions

where flowering and fruiting is more prolific. Many of the rarer polyploid species are found on cliffs and open slopes where they also escape the grazing and browsing of sheep, deer and goats. *S. torminalis* is an exception, being a forest tree and is more shade-tolerant than the other sexual species, a feature that appears to have been conferred to at least some members of subgenus *Tormaria* (see section 1.7.3). *S. aucuparia* occupies a wide range of habitats and the saplings show shade tolerance. *S. aria* is a characteristic species of open woodland and scrub on chalk and limestone in southern England, although it can be found on a wide range of soil types and has now been widely planted and naturalised; as such, its native U.K. distribution is difficult to ascertain.

1.8.2   Reproduction in *Sorbus*

There are close links between polyploidy, self-compatibility, gametophytic apomixis and hybridisation in the subtribe *Pyrinae* which contains the genus *Sorbus* (Dickinson *et al.*, 2007). The type of breeding system shapes the patterns of evolution in *Sorbus.* Sexual diploid *Sorbus* are typically outcrossing and self-incompatible (Oddou-Muratorio *et al.*, 2005; Pías & Guitián, 2006) relying on pollen flow among individuals for seed production. Polyploid *Sorbus* exhibit gametophytic apomixis in which embryo-sacs develop apomeiotically from somatic cells containing a female gametophyte with the somatic chromosome number. This unfertilized egg cell then goes on to form an embryo (Talent, 2009) with an identical genotype to the maternal tree. The unreduced central cell still requires fertilisation by pollen for normal endosperm development (pseudogamy) although there is no contribution of male genetic material to the embryo (Liljefors, 1953, Ludwig *et al.*, 2013). The requirement for pollination to produce seed could render polyploids succeptible to pollen limited

40

seed production if self pollen is incompatible, as would be the case for the sexual species. However, tetraploid *Sorbus* species show a break down of self-incompatibility enabling self pollination to assure seed production, although some triploid species are self-incompatible (Ludwig *et al.*, 2013), potentially exposing them to constraints on seed production. Self-incompatibility results from the pollen either being rejected on the stigma or pollen tubes failing to grow down the style. In some triploid *Sorbus*, incompatible pollen tubes fail to reach the bottom of the style due to gametophytic self-incompatibility (GSI). GSI was considered the most likely reason for this failure since male sterility had been discounted (Ludwig *et al.*, 2013). GSI resuls from an *S* allele product in the pollen interacting with RNase of the style (Horandl, 2010).

Apomixis in polyploid *Sorbus* may be obligate or facultative and the requirement for pollination (pseudogamy) provides a pathway for sexual reproduction on occasion. Robertson *et al.* (2004a) showed that although tetraploid *S. pseudofennica* reproduced primarily by apomixis, 17.5% of its seed was of sexual origin and that it was involved in ongoing hybridisation with diploid *S. aucuparia*. Apomixis has allowed the establishment and persistence of isolated populations that consist of only a few trees. For example the global population of *S. leyana* is approximately 17 trees (Rich *et al.,* 2010). The genes conferring apomixis appear to be carried on the *Aria* genome (Rich *et al.*, 2010) although it is unknown which are the main factors that affect the frequency of sexual reproduction in apomictic *Sorbus*. Indeed, some apomictic triploid *Sorbus* have shown high rates of hybridisation when brought into cultivation, for example *S. leyana*, a triploid species, was observed to produce offsping with a wide range of leaf morphology when grown in proximity to a diversity of other *Sorbus* species (both native and non-native to the UK) (N. deVere, pers. comms.). This

41

suggests environmental factors on the wild sites contribute greatly to the likelihood of sexual reproduction, perhaps availability of compatible pollen.

Pollen viability as estimated via stainability using Alexander's stain varies among the different cytotypes. Pollen viability is highest in the diploid species although many tetraploids have similar viability. Triploids show the lowest stainability; however, this is highly variable between species (2.4 - 95% across studies). The hybrids within subgenus *Aria* have higher stainability than those in other subgenera (Bednorz *et al.*, 2005, Rich, 2009, Hajrudinović *et al.*, 2015a, Hajrudinović *et al.*, 2015b). Seed set per fruit is again highest in the diploid species with *S. aria* having almost twice the number as triploid and tetraploid cytotypes (Hajrudinović *et al.*, 2015b). *S. torminalis* produces slightly fewer seed per fruit than *S. aria* (Price & Rich, 2007). The relative high fitness of *S. aria* as measured by pollen viability and seed production may affect the rates of successful hybridisation at certain sites; although where it outnumbers polyploid species it would not be unreasonable to imagine it would be a competitor for resources, unless there is some degree of ecological niche separation whereby the polyploids have certain adaptive advantage in particular microhabitats.

1.8.3 Taxonomic complexity of *Sorbus*

Research has shown the main centres of *Sorbus* diversity in Europe to be Scandinavia, Great Britain and south east Europe (Aldasoro *et al.*, 1998, Rich *et al.*, 2010, Hajrudinović *et al.*, 2015b). The genus contains obligate outcrossing diploids and hybridogenous apomictic, generally self-compatible polyploids. Current taxonomy of *Sorbus* is based firstly on morphology, primarily leaf shape, veining patterns and abaxial tomentum, together with fruit size, shape, lenticels and colour. The parental diploid taxa have distinct leaf and fruit

42

features and the intermediate characters of hybrids largely reflect their broad origins. Cytological and molecular studies have further informed *Sorbus* classification and biology, enabling detailed exploration of the evolutionary relationships within some *Sorbus* groups resulting in ongoing taxonomic revision and the discovery of newly formed hybrids and species (Lepší *et al.*, 2008, Rich *et al.*, 2009, Rich *et al.*, 2014, Hajrudinović *et al.*, 2015a, Lepší *et al.*, 2015). Thus evidence shows diversification within this genus is active and how use of molecular tools is able to reveal the complexity of polyploid speciation.

In Europe, *Sorbus* is represented by five diploid, outcrossing sexual species *Sorbus aria* L., *S. torminalis* (L.) Crantz ,*S. aucuparia* L., *S. chamaemespilus* (L.) Crantz and *S. domestica* L. This group of species (except *S. domestica*) is responsible for the formation of numerous hybridogenous polyploids that are largely apomictic (Liljefors, 1953, Aldasoro *et al.*, 1998). In Britain, *S. aria, S. torminalis* and *S. aucuparia* have given rise to upwards of 35 described apomictic polyploid species of hybridogenous origin. The British polyploid taxa are classified into subgenera according to their broad origins and here I follow that described by Rich *et al.* (2010). The subgenus *Aria* contains the sexual *S. aria s.s.* but also a number of polyploid species of *S. aria* origin. This group is often referred to as the *S. aria* aggregate or *S. aria s.l.* The remaining polyploid taxa originate from hybridisations involving members of subgenus *Aria* and either *S. torminalis* or *S. aucuparia* and are classified into subgenera *Tormaria* and *Soraria* respectively. *Sorbus domestica* is genetically distinct from the other diploid species and does not hybridise with them (Rich *et al.*, 2010).

43

1.8.4 What constitutes a species in *Sorbus*?

The question of what constitutes a species in *Sorbus* has implications for both taxonomy and conservation for reasons discussed earlier (1.1). There is ongoing debate on the definition of the species category, particularly in hybridogenous apomictic groups (Lepší *et al.*, 2008). The most popular and often quoted concept is the Biological Species Concept (BSC) as presented by Mayr (1942) "Species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups." This definition suits the discontinuous patterns of relationships found in many animals and some higher plants. This concept has been constantly reviewed with various interpretations over time (De Queiroz, 2005) and one reason for this is that the central themes of interbreeding and reproductive isolation makes this an inadequate definition for organisms exhibiting uniparental reproduction or those that occasionally hybridise (Grant, 1981), features particularly associated with recently diverging species complexes and common to the genus *Sorbus.* A solution to this is to accept any polyploid hybrid as species if it forms a sufficiently morphologically discrete group (although this can be somewhat challenging) with distinct geographic distribution and is biologically sustaining (Rich *et al.*, 2010). There are some areas of uncertainty such as whether to classify products from repeated hybridisation involving the same parents as individual species with single origins or one species with multiple origins. It is not known how common this recurrent evolution is within *Sorbus* and the outcomes of multiple origins are dealt with differently depending on the morphological distinction of each hybrid product. For example, on Arran, Robertson *et al.* (2004b) demonstrated that repeated hybridisation events between *S. rupicola* and *S. aucuparia*, have given rise to *S. arranensis* at least

44

three times, each of which is genetically identifiable but morphologically indistinct. This may be compared with the ongoing hybridisations within Avon Gorge. Here, a number of direct or indirect hybridisations between diploid *S. aria* and tetraploid *S. porrigentiformis* have resulted in *S. porrigentiformis* like clones e.g. *S. leighensis* and *S. × avonensis,* which are recognised as distinct morphological and genetic entities (Robertson *et al.*, 2010).

1.8.5   Diversification in *Sorbus*

The diversity seen in extant *Sorbus* across Europe is a product of hybridisation, allopolyploidy, autopolyploidy and mixed sexual and apomictic (both obligate and facultative) mating systems (Liljefors, 1953, Nelson-Jones *et al.*, 2002, Robertson *et al.*, 2004a, Chester *et al.*, 2007, Robertson *et al.*, 2010, Ludwig *et al.*, 2013, Hajrudinović *et al.*, 2015b). Hybridisation is the key factor in creating *Sorbus* diversity, leading to large numbers of intermediate species across Europe (Aldasoro *et al.*, 1998) via hybridisations and backcrosses among the four widely distributed diploid, sexual species; *Sorbus aria, S. torminalis, S. aucuparia* and *S. chamaemespilus* (Liljefors, 1955, Richards, 1975, Rich *et al.*, 2010). In the UK, all the hybridogenous polyploid species are derived from hybrid events involving a member of subgenus *Aria* as one of the parents. Subgenera *Tormaria* and *Soraria* have maternally inherited chloroplast DNA haplotypes corresponding to *S. torminalis* and *S. aucuparia* respectively (Nelson-Jones *et al.*, 2002, Chester *et al.*, 2007) which may suggest that there is a tendency for hybridisations to occur predominantly in one direction. However, where the diploids *S. torminalis* and *S. aria* produce the primary diploid hybrid *S. × tomentella, S. torminalis* is the female parent in the UK (Fay *et al.*, 2002, Chester *et al.*, 2007) but in France hybridisations in both directions appear to occur but predominately (75%) with *S. aria* as the female parent

(Oddou-Muratorio *et al.*, 2001). In addition, the endemic polyploids of Arran, Scotland were produced with sexual diploid *S. aucuparia* as female and male parent for triploid *S. arranensis* and tetraploid *S. pseudofennica* respectively (Robertson *et al.*, 2004b) illustrating that hybridisation can occur with the sexual diploid as the male or female parent. The initial hybrid events between sexual *S. aucuparia* and tetraploid *S. rupicola* (subgenus *Aria*) were with *S. rupicola* (subgenus *Aria*) as the male parent. Tetraploid *S. rupicola* is the most widespread polyploid (UK, north-west Europe and Scandinavia) and thought to be a parental species for other hybridogenous polyploids throughout its range (Liljefors, 1955, Robertson *et al.*, 2004b, Rich *et al.*, 2010). Liljefors (1955), proposed an autopolyploid origin for *S. rupicola* from *S. aria*, which may have been the case, perhaps via a spontaneous *S. aria* triploid backcrossing to the diploid form or fusion of two 2*n* gametes (Ramsey & Schemske, 1998). It is unknown how often spontaneous triploid forms of the sexual diploids are produced and whether they are fertile. Interploidy hybridization in *Sorbus* can lead to speciation as apomixis prevents introgression. However, the facultative nature of apomixis in *Sorbus*, allowing occasional ongoing hybridisation has produced the current array of reticulate relationships. The activity of this process i.e. frequency of successful hybridisation, will be partly due to inherent characteristics of the species involved; such as mating system, mate compatibility and pollen and seed viability; however, these processes take place in environments where many other factors e.g. pollinator abundance, may affect stages of the hybrid process including survival of the component individuals and species.

46

An aim of this research project is to investigate how some of these factors may affect short term species survival and discuss the implications for long term evolution of the complex.

1.8.6   Conservation status

Due to the endemic nature and small population sizes of many apomictic *Sorbus* species [all features of vulnerable species (IUCN, 2001)], many are considered of conservation concern with threats primarily from changing land management, browsing and invasive non-native species. Indeed, approximately 54% of native British *Sorbus* fall into the IUCN threatened categories (Rich *et al.*, 2010) although their status is currently being re-assessed (T. Rich pers. comms.). The allocation of currently scarce resources available for conservation is impractical if each species is managed separately, indeed, *Sorbus* has been highlighted as a model TCG for a process-based approach to conservation on Arran (Ennos *et al.*, 2012) and the Avon Gorge (Robertson *et al.*, 2010, Ludwig *et al.*, 2013), particularly as the commonly occurring diploids appear to perform a key role in polyploid formation (Liljefors, 1953, Robertson *et al.*, 2010, Hajrudinović *et al.*, 2015b) and would therefore need to be included within such plans.

Whilst a process-based approach may put the emphasis on the evolutionary process rather than necessarily the products of that process, there is worth to the recognition and naming of regional variants such as the Devon whitebeam (*S. devoniensis*) as it gives cultural significance to some of these species. This is of great value when securing funding for conservation management and engaging public support for their protection. It also illustrates the detail of

variation in this genus which would otherwise be lost if they were taxonomically combined together into broader groups (Rich *et al.*, 2010).

1.9   Sorbus in Devon and north Somerset

The presence of a number of naturally occurring polyploid *Sorbus* species in this region makes this an ideal area to study the evolutionary dynamics of *Sorbus* diversification over an extensive area. Here, I review their ecology, systematics and conservation status followed by an outline of identification and sampling methodology.

1.9.1   Distribution

 The stretch of coastline along north Devon and into the western parts of north Somerset is known for its endemic *Sorbus* species. The underlying geology is largely slates, shales and grits of varying pH (Rich *et al.*, 2010). Coastal cliffs and woodlands are the main habitat types for the rarer species and there are several key sites in this region where multiple polyploid species occur together. There are also a number of sites in south Devon where multiple polyploid *Sorbus* species are found, again in woodland and on cliffs primarily on Devonian limestone. A summary of sample sites is given in Table 1.1 from information compiled by Rich *et al.* (2010).

There are eight polyploid species native to Devon and north Somerset; *Sorbus subcuneata* Wilmott*, S. admonitor* M.C.F. Proctor*, S. vexans* E.F. Warb and *S. margaretae* M.C.F. Proctor are endemic to the north coastal stretch of Devon and Somerset. *S. devoniensis* E.F.Warb*, S. porrigentiformis* E.F.Warb and *S. rupicola* (Syme) Hedlund have wider distributions, although *S. devoniensis* is largely restricted to Devon and assumed to have arisen there. *S. porrigentiformis* has a south-west England and South Wales distribution whilst

48

S. rupicola is the second most widespread *Sorbus* in Britain. *S. anglica* Hedl. is widely scattered in south-west England, Wales and Killarney in Ireland (Rich *et al.*, 2010). Of the three native sexual diploid species, *S. aucuparia* L. is common throughout this region, occurring alongside all the polyploid species at many of their sites but curiously there are no apparent associated syngameons in this region with *S. anglica* the only local hybrid derivative (see pg 48). *S. torminalis* (L.) Crantz., is also widespread in this region but rather infrequent and often associated with hedgerows, as is *S. devoniensis,* with which it is occasionally found. *S. torminalis* also occurs with *S. porrigentiformis* at one site in Torbay. *S. aria* L. is not thought to be native this far southwest (Rich *et al.*, 2010). The western limit of its native distribution range is further to the east near Cheddar Gorge where it grows with other polyploid taxa (Houston *et al.*, 2009). Local distribution has influenced hypothesised relationships in this group to an extent so clarification is required.

Table 1.1 Study species with a summary of sample sites. For full details see sample map Fig 2.2, page 88 and Table S2.1 pages 116-124. Site designations are also given. SSSI = Sites of Special Scientific Interest.

Study species	Common Name	Reported ploidy	Geographic distribution	Sample collection sites with location (Site designation (SSSI) and NGR only given at first mention)
S. aria	Common whitebeam	diploid	Central & southern Europe; southern England in UK	Leigh Woods & Avon Gorge SSSI (ST5573-5574);Cheddar Gorge SSSI (ST4754).
S. torminalis	Wild service tree	diploid	Central Europe, rare in north Africa and near East. Central and southern England and Wales in UK.	Numerous sites across S Wales, Avon, Somerset, Devon and Cornwall (see Table S2.1)
S. rupicola	Rock whitebeam	tetraploid	NW Europe; Scattered in Wales, SW England, Scotland & northern and western Ireland in UK.	Neck wood SSSI (SS6348); Churston, Torbay (SX9165-9257); Babbacombe, Torbay SSSI (SX9265); Creagh Dhubh SSSI (NN6795); Darren Fach, Brecon Beacons SSSI (SO0110); Penmoelallt, Brecon Beacons SSSI (SO0109).
S. porrigentiformis	Grey-leaved whitebeam	tetraploid	SW England & South Wales	Fishermans car park, Watersmeet SSSI (SS7348);Woody Bay SSSI (SS6649); Leigh Woods & Stokeleigh Camp , Avon Gorge; Babbacombe, Torbay; Walls Hill SSSI, Torbay (SX9365); Redgate SSSI, Torbay (SX9364); Cheddar Gorge; Broadridge Woods SSSI, Newton Abbot (SX8371); Darren Fach, Brecon Beacons.

Study species	Common Name	Reported ploidy	Geographic distribution	Sample collection sites with location
S. vexans	Bloody whitebeam	tetraploid	Coast of N Devon and S Somerset	Fishermans car park, Watersmeet SSSI; Oxen Tor SSSI (SS7249); Neck wood;Dogsworthy nr Desolation SSSI (SS7749); Culbone (SS8348); Woody bay SSSI (SS6748).
S. margaretae	Margaret's whitebeam	tetraploid	Coast of N Devon and S Somerset	Watersmeet SSSI (SS 7448); Neck wood; Desolation (SS7749-7849); Culbone; Embelle woods, nr Culbone (SS8149)
S. subcuneata	Slender whitebeam	triploid	Coast of N Devon and S Somerset	Watersmeet; Neck wood; Woody bay SSSI (SS6749); Culbone (SS8448); Greencliff (SS9647) & Greenaleigh (SS9746) nr Minehead
S. devoniensis	Devon whitebeam	tetraploid	Devon, Cornwall & Somerset. Also SE Ireland	Eleven sites across Devon see Table S2.1 pgs 117-118
S. admonitor	No Parking whitebeam	tetraploid	N Devon	Watersmeet

Figure 1.6 The steep river valleys of Watersmeet (lat. 51.22675, long.-3.79965) near to the north Devon coast feature oak woodland, rock outcrops and scree slopes.

Figure 1.7 Desolation (lat. 51.2337, long. -3.7434) is a coastal cliff site in north Devon with open scrub and rock, home to many *S. margaretae* individuals.

1.9.2   History and origins of the Devon and north Somerset Sorbus.

Seven of the eight polyploid species of this region along with diploids *S. torminalis* and *S. aria* are thought to be closely related. *S. anglica*, a tetraploid, occurs in Devon (as part of a widely scattered SW England and Wales distribution) but is a member of subgenus *Soraria* derived from *S. aucuparia* × *S. porrigentiformis* (Robertson *et al.*, 2010). The remaining seven polyploids are from subgenera *Tormaria* and *Aria* and these two groups form the basis of this study.

Subgenus *Tormaria,* formerly known as the *S. latifolia* aggregate, has fruit and leaf morphology intermediate to *S. aria* and *S. torminalis* (Sell, 1989). Wilmott (1934) recognised the forms of *S. latifolia* found in Devon and Cornwall as separate from the form in the Avon Gorge at Bristol (*S. bristoliensis* A. J. Wilmott). These Devon and Cornish forms showed some variation at Watersmeet, Devon, and a form with particularly narrow leaves was recorded near Minehead, Somerset. The narrow leaves with whiter tomentum beneath than other forms suggested it arose from *S. rupicola* × *S. torminalis* (or the Devon form of *S. latifolia*) (Sell, 1989). These forms were separated by their uniform appearance and noted that they came 'true' from seed. Wilmott described the Minehead variety as *S. subcuneata* and stated that it should be distinguished as a species. He also made a note that if they were apomictic (he refers to parthenogenesis) and long established, it would be expected that they would occur in much greater numbers. The Devon form was later described by Warburg as *S. devoniensis* (Warburg, 1962) and the Watersmeet variety, *S. admonitor*, was finally separated from *S. devoniensis* and described by Rich & Proctor (2009). *S. admonitor* was long maintained as a variant of *S. devoniensis,* or even *S. subcuneata,* having a similar leaf form to S. *devoniensis*

53

but more sharply lobed (Fig's. 1.12 to 1.14). The broad leaf shape has also prompted suggestions that *S. aria* may be a parental species along with *S. torminalis* (Sell, 1989).

S. rupicola, S. porrigentiformis, S. vexans and *S. margaretae* all belong to subgenus *Aria* as they are all thought to be direct or indirect derivatives of *S. aria s.s.,* which is also contained in this subgenus (Rich *et al.*, 2010). If this is the case their genome is entirely *S. aria* derived. *S. rupicola* and *S. porrigentiformis* are undoubtedly the oldest species of this group with their larger distribution ranges and have long been recognised as species with various synonyms (Wilmott, 1934, Keble & Fraser, 1939, Liljefors, 1955). *S. rupicola* has a more northern limit to its distribution than *S. aria* and is thought to have given rise to a number of polyploid species through hybridisation with *S. aucuparia* both in Britain (Robertson *et al.*, 2004b) and Scandinavia (Liljefors, 1955). *S. porrigentiformis* has also been found to be a parental species for other polyploids via hybridisation with *S. aria* and *S. aucuparia,* most likely as the pollen donor (Robertson *et al.*, 2010). *S. vexans* was described as the Devon form of *S. rupicola* in the Devon Flora of 1939 (Keble & Fraser, 1939) due to their similar leaf morphology and *S. margaretae* was only fully differentiated and described recently by Rich & Proctor (2009).

This group of species was studied by Proctor *et al.* (1989) using peroxidase isoenzymes. S*. rupicola, S. porrigentiformis* from Devon and *S. vexans* all consisted of individuals with the same phenotype with minor variations, sharing all bands with *S. aria. S. vexans* had very similar banding patterns to *S. rupicola* suggesting a close relationship. A new form of *S. vexans*, 'Taxon D', now known as *S. margaretae* was confirmed, which also showed some variation with

a 'western form'. *S. subcuneata* and *S. devoniensis* both shared bands with *S. aria* and *S. torminalis* confirming their taxonomic position as intermediate to the two diploid species. *S. admonitor* was also found to have unique banding patterns, although at this time it was still considered a variant of *S. devoniensis* and in fact only differed by the lack of one band. Whilst the isoenzymes had proved useful as a taxonomic tool the precise hybrid origins of these species were still unclear. RFLP analysis confirmed the above close relations (Nelson-Jones *et al.*, 2002) and cpDNA patterns revealed the female parent of subgenus *Tormaria* to be *S. torminalis* (Nelson-Jones *et al.*, 2002, Chester *et al.*, 2007). An extensive study of the *Sorbus* species found in the Avon Gorge (Robertson *et al.*, 2010) using molecular analyses and parentage mating simulations showed that the majority of polyploid species arose as a result of hybridisations involving the diploid species. Ploidy screening of seed endosperm and embryos from sexual diploid and putatively apomictic polyploid *Sorbus* supported the prominent role of sexual diploids in hybrid events (Hajrudinović *et al.*, 2015b). Pellicer *et al.* (2012) established ploidy levels for the study polyploids as tetraploid except for *S. subcuneata* which is triploid. However, for some species (*S. admonitor, S. margaretae, S. devoniensis* and *S. vexans*) sample sizes were small (4-6 individuals) and unexpected ploidy levels were found for some more intensively sampled species, e.g. a pentaploid (5x) *S. porrigentiformis* in Wales, potentially formed from union between an unreduced gamete from a triploid ($3n$) with a reduced tetraploid gamete ($2n$).

1.9.3 Conservation and threats

Of the nine study species (seven polyploids and two diploids), *S. vexans, S. margaretae* and *S. admonitor* are considered Endangered (*circa* 70, 100 & 110 plants respectively); *S. subcuneata* is Vulnerable (*c.* 300 plants) and the

remainder are of Least Concern (IUCN, 2001). Whilst species based

conservation may be appropriate for some *Sorbus,* it does not necessarily

consider the interactions among species which may result in further

diversification. As mentioned previously, process-based action plans have been

suggested to be more appropriate for *Sorbus* and these may also include

common species of least concern, such as the sexual diploids (Ennos *et al.*,

2012). However, it is unknown to what extent these study species are of

common ancestry and whether they arose from single hybrid events or from

several events in different areas. Their hybrid origins are largely unknown

despite the several theories put forward, as summarised above. These

problems have implications for the development of conservation strategies

aiming to promote gene flow and *Sorbus* diversification in this region. If

evolutionary processes are ongoing in this region, it is vital to know which

species predominantly take part. These should form the focus of any

conservation actions in order to preserve the potential for further gene flow.

However, if these species are remnants of a once active process, then we

should seek to conserve these remaining species.

It is unknown whether the populations of these study species are declining and

comprehensive surveys are required for *S. margaretae* and *S. vexans* to fully

ascertain population sizes (Rich *et al.*, 2010). Since these species occur over

an extensive area, the threats to their short term survival and long term

evolutionary potential are likely to be different at the various sites. The main

threats come from the invasive *Rhododendron ponticum* (Fig. 1.8) which has

formed extensive stands along the north Devon / Somerset coastline, effectively

fragmenting *Sorbus* habitat by shading out other tree species and preventing

regeneration. This is especially pertinent for the light demanding *Sorbus*

species of this area. This has also happened with *Quercus ilex*, another invasive evergreen species with a dense canopy which has had a negative impact on the calcareous grassland and scrub habitats of the Torbay coast where *S. rupicola* and *S. porrigentiformis* are found. *Sorbus* is highly palatable and browsing is a problem on many *Sorbus* sites (Rich *et al.*, 2010) and this is almost certainly the case on many of the study sites in Exmoor where deer numbers can be high and likely prevent regeneration.

Figure 1.8 Invasive *Rhododendron ponticum* in flower, seen as a purple swathe along the Exmoor coast of north Devon.

Changing land management allowing development of dense woodland canopies, or secondary woodland on once open scrub and grassland sites both create shaded conditions which has resulted in reduction in numbers of shade intolerant *Sorbus* species (Hamston, T. unpubl. MSc.). Land slips and rock falls may also eradicate whole groups of trees (pers. obs) but create new, open colonising opportunities.

These threats all potentially impact on the short term survival of these species by destroying individual trees and preventing regeneration. Fragmentation of populations also reduces potential gene flow via pollen and seed among the species and may affect colonisation of new areas. Indeed, the current pollinator decline may impact on pollinator services to wild plant communities (Potts *et al.*, 2010) including those containing *Sorbus*.

1.9.4 Tree identification and sampling

Identification of taxa is problematic due to the cryptic nature of this genus, containing species of very similar morphology. For this study, identification training was given by expert botanist, Tim Rich and Botanical Society of the British Isles (BSBI) vascular plant recorder, Roger Smith. *Sorbus* identification is based on leaf and fruit characteristics, with leaf shape and vein number as key components. Fully expanded leaves from a non-flowering lateral side shoot (Fig. 1.9) in a sunny position are required and young seedlings or saplings may be impossible to identify with confidence due to their juvenile leaf forms.
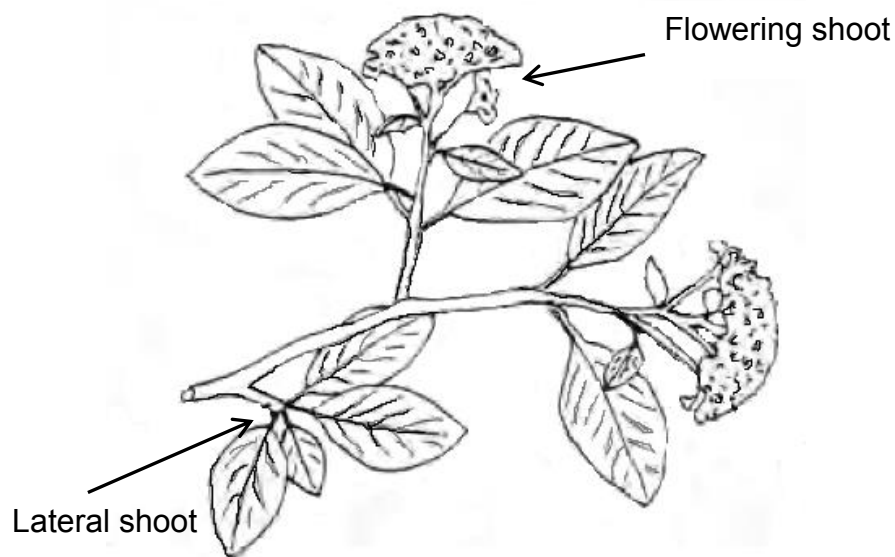
Figure 1.9 *Sorbus* branch showing short lateral shoot. Re-drawn from Rich & Jermy (1998).

These side shoots were collected as voucher specimens for all trees sampled

for DNA. These provide a reference for later examination of morphological

features. The vouchers were labelled accordingly, pressed and dried and sent

to the Welsh National Herbarium in Cardiff (NMW). Due to difficult access to

many trees, particularly those growing on cliffs, long-handled loppers were used

to collect vouchers. See Appendix 1 for a selection of herbarium images of all

study species. The identity of all sampled trees were determined via their

vouchers and subsequently confirmed by their genetic profile. The shape, size

and colour of fruits also provide important identification characters but often

samples were collected at other times of the year, so identification was made on

leaf characters only. The identification keys in 'Whitebeams, Rowans and

Service Trees of Britain and Ireland. A monograph of British and Irish Sorbus L.

B.S.B.I. Handbook No. 14.'(Rich *et al.*, 2010) and 'Plant Crib' (Rich & Jermy,

1998) were used. *S. aucuparia* was easily distinguished from the nine study

species by its compound leaf (Fig. 1.10).
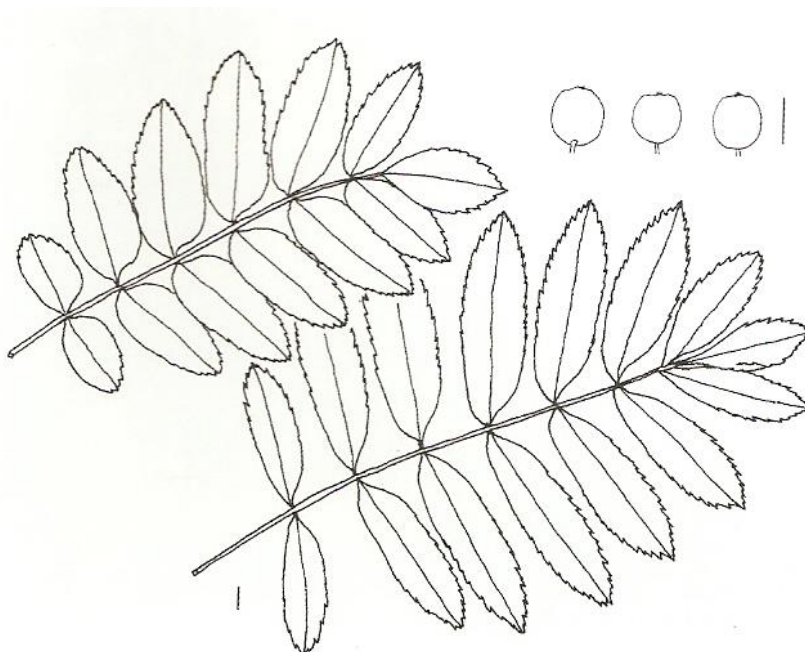
Figure 1.10 *Sorbus aucuparia* from Rich *et al.* (2010). The bar represents 1cm.

Figures 1.11 to 1.19 show the leaf shapes for all nine study species and were
provided by Tim Rich. The bar represents 1cm.
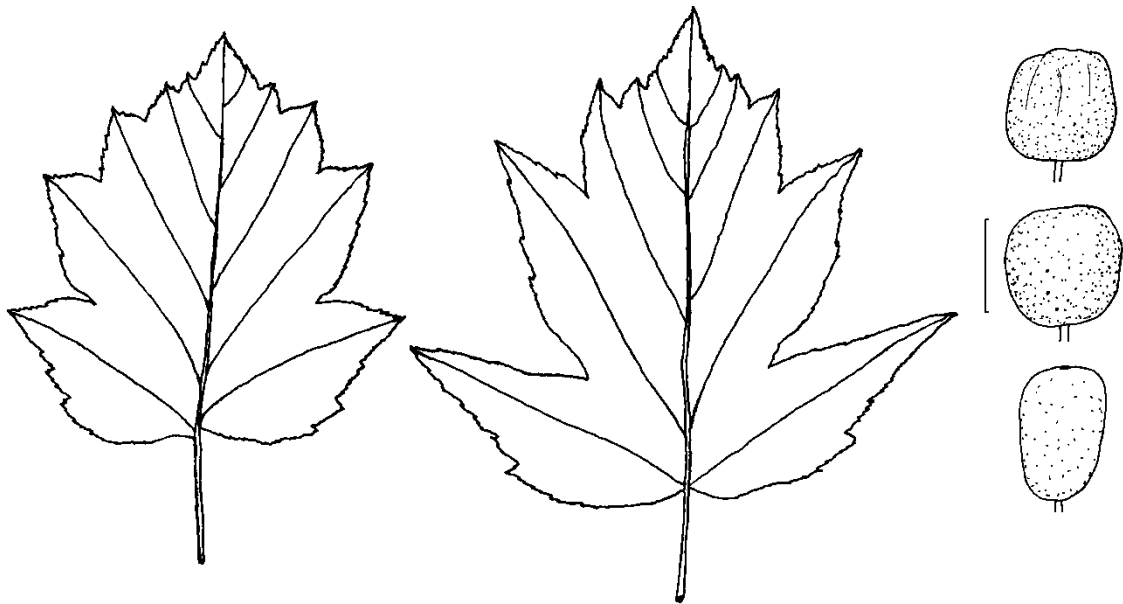
Figure 1.11 Wild service tree *Sorbus torminalis* with lobed leaves
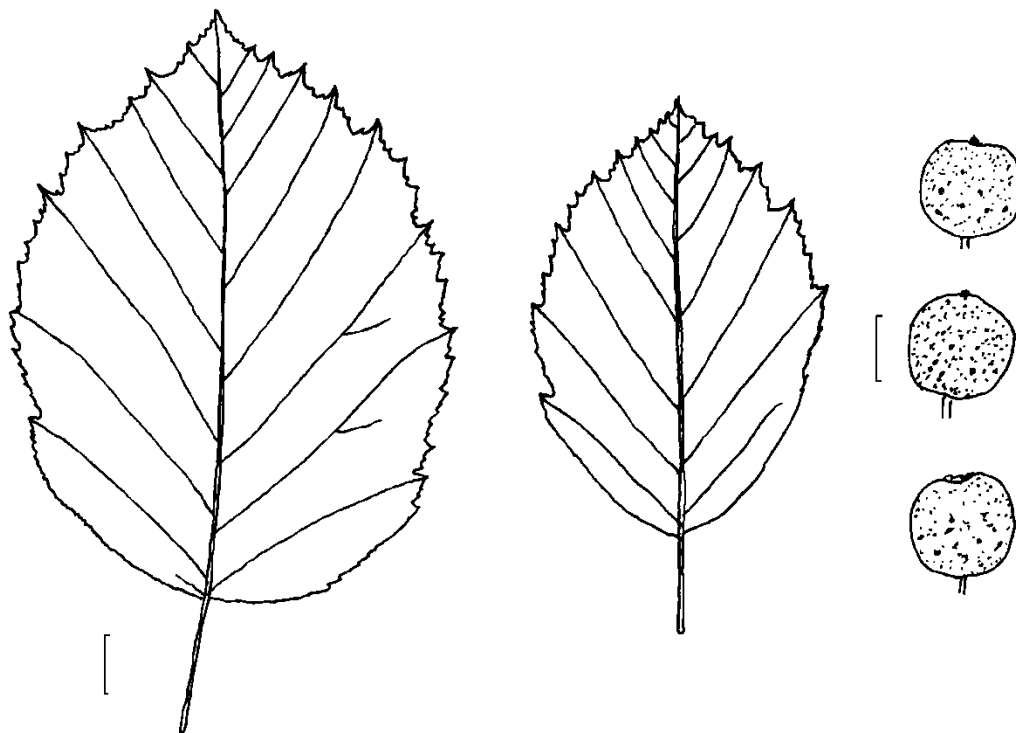
Figure 1.12 Devon whitebeam *Sorbus devoniensis*
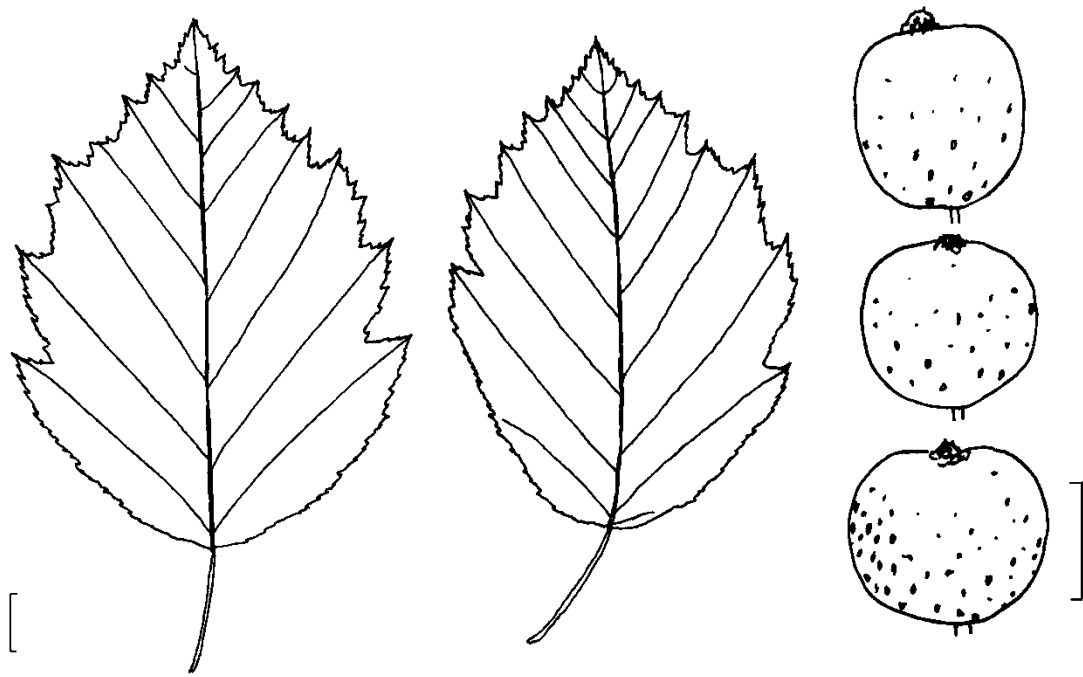
Figure 1.13 No Parking whitebeam *Sorbus admonitor*
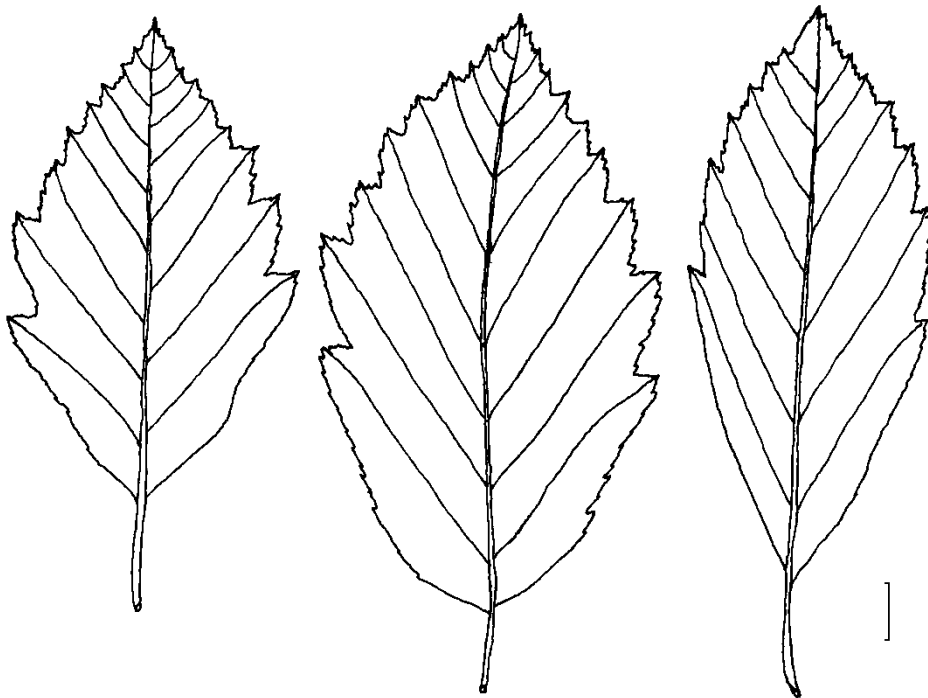
Figure 1.14 Slender whitebeam *Sorbus subcuneata*
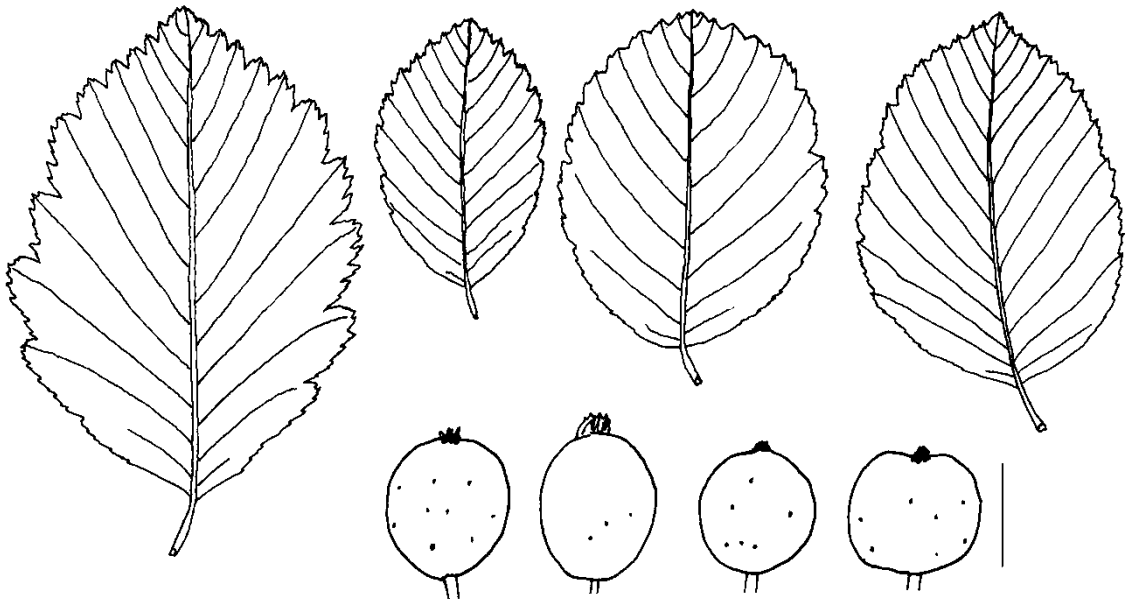
Figure 1.15 Common whitebeam *Sorbus aria* with very variable simple leaves
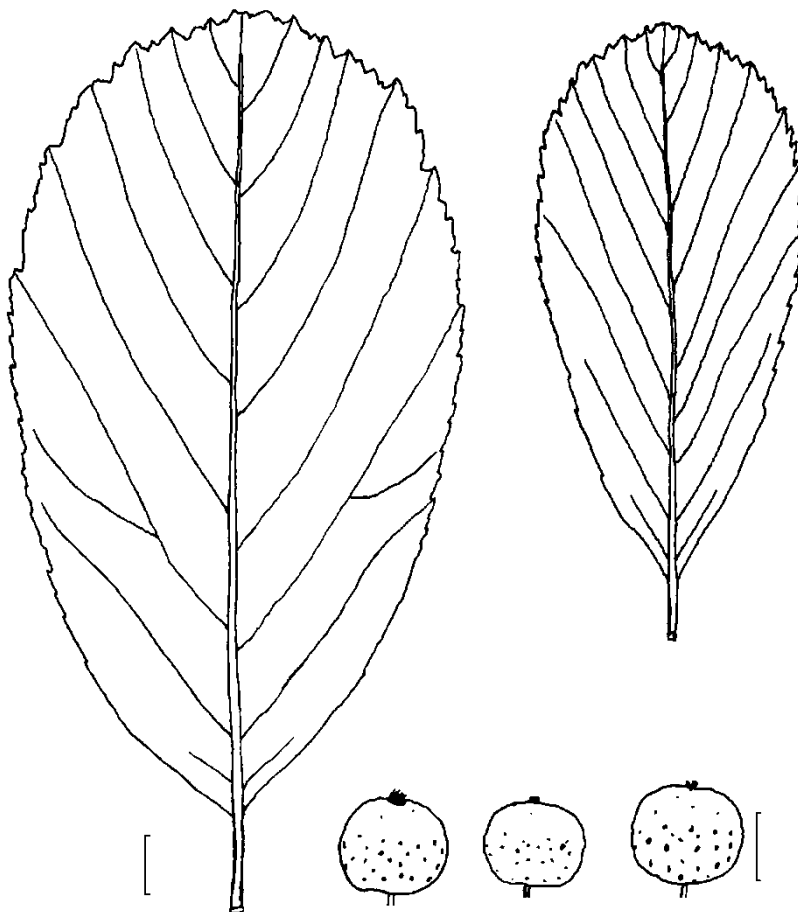
Figure 1.16 Rock whitebeam *Sorbus rupicola*
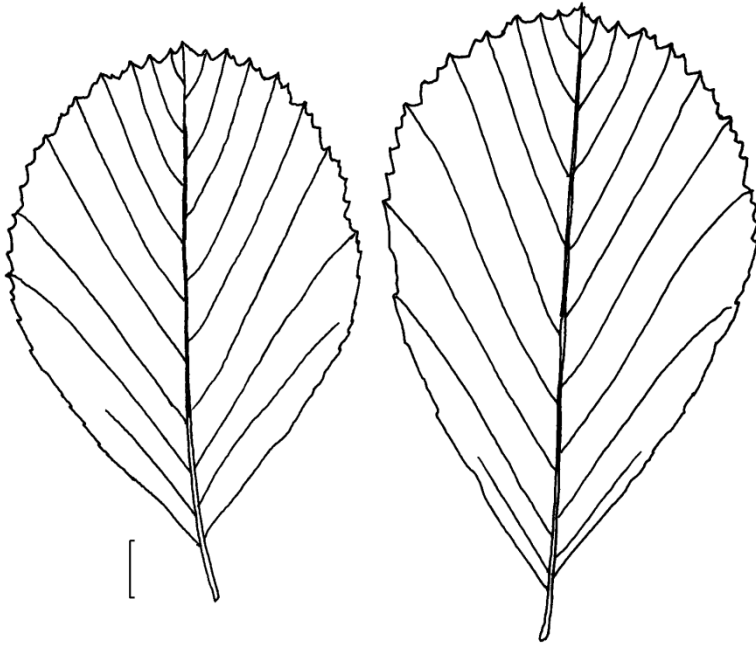
Figure 1.17 Grey-leaved whitebeam *Sorbus porrigentiformis*
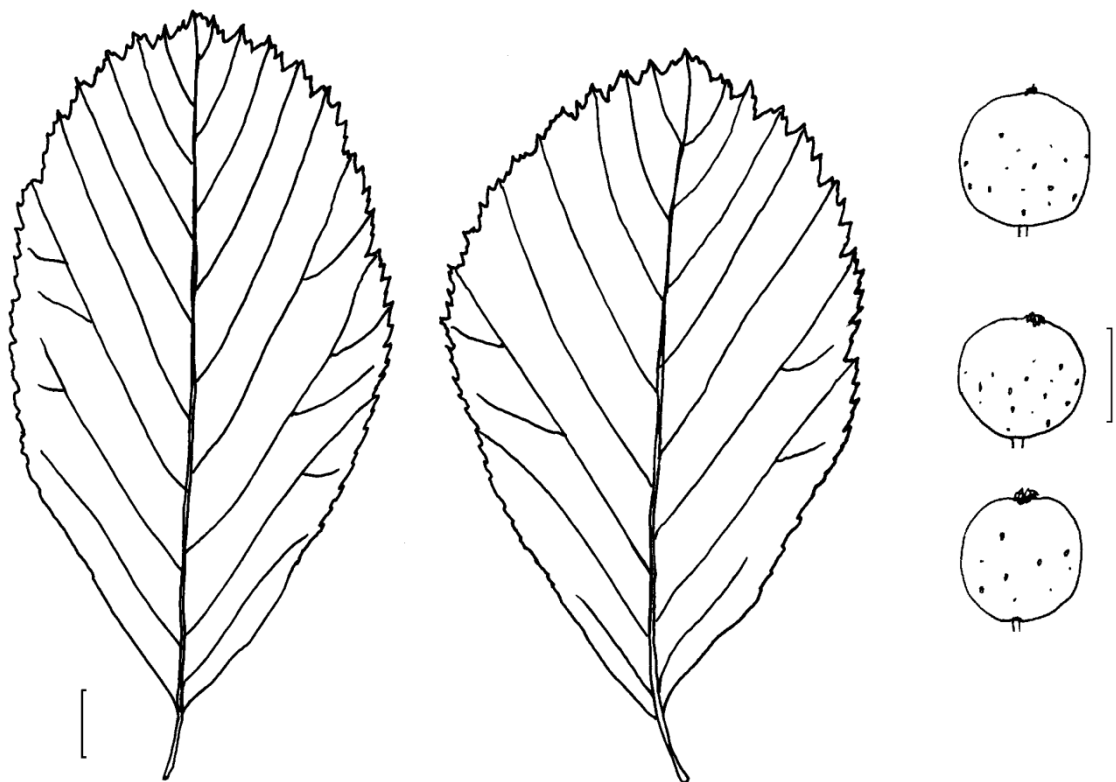
Figure 1.18 Bloody whitebeam *Sorbus vexans*
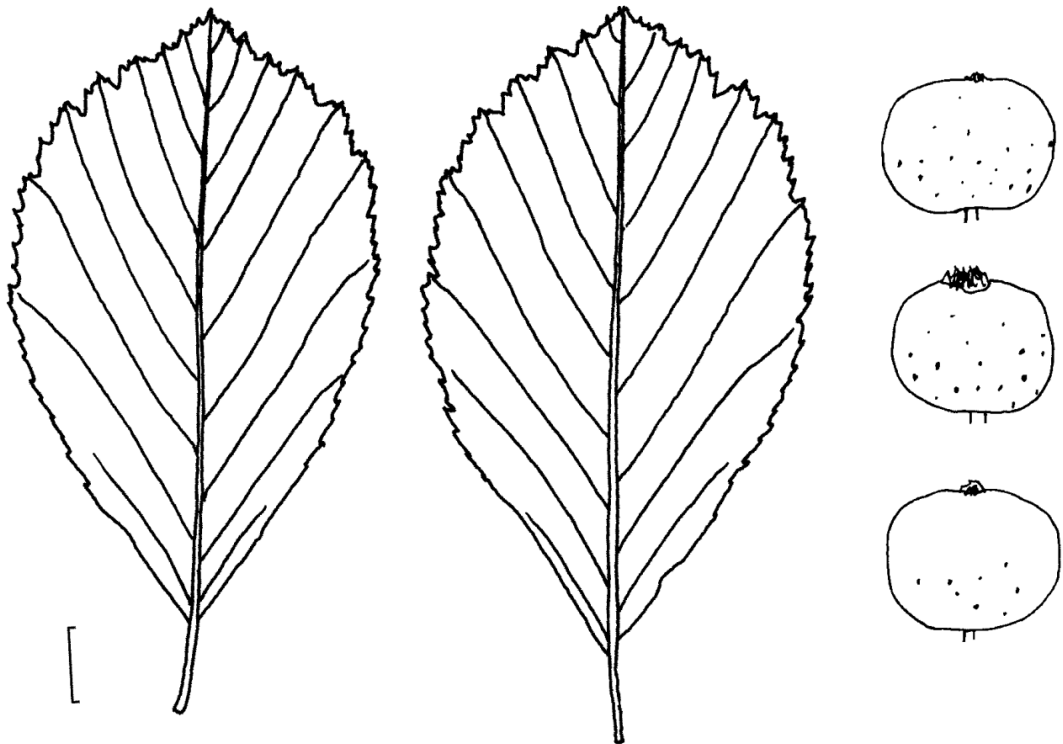
Figure 1.19 Margaret's whitebeam *Sorbus margaretae* with broader fruits than *S. vexans.*

Sampling sites varied in their habitat composition; hedgerow, deciduous woodland, scrub and woodland edge, cliffs and rocky outcrops. The sample sites were chosen based on prior information from the (BSBI) vice-county recorders and the extensive personal knowledge of Tim Rich who also contributed his own database of records, hand-drawn maps and location descriptions which proved invaluable for locations where only one or two trees existed. Previous surveys of the study species on both north and south coasts, carried out by M. E. Proctor in1984 under contract for the Nature Conservancy Council, also provided detail on population sizes, identification and location. Tim Rich, Libby Houston, Martin Lepší, Natasha deVere and Jack Hamston-Goodfellow assisted with the sampling.

Site owners were approached where necessary (i.e. SSSI sites) and necessary permissions granted with Natural England via the landowners if requested. None of the actions on site were in breach of SSSI restrictions so this was out of courtesy. The main study site of Watersmeet (lat. 51.22675, long. -3.79965) is a National Trust property and site personnel were kept informed of experimental plans and progress.

In order to capture any variation within and among species I sampled across the range of the south west endemic species at as many sites as possible where species were known to occur. Where species occupied many sites (i.e. *S. devoniensis*) I sampled a sub set of sites to represent the full geographic range. Sampling at sites where populations exceeded 50 individuals was performed following a non-random sampling method, where all parts of the site were sampled so as to encompass any within site variation and avoid re-sampling of clonal groups (Bayer, 1990). In order to try and encapsulate the widest range of alleles available to this taxon group the non-endemic taxa, *S. torminalis*, S*. aria, S. rupicola* and *S. porrigentiformis* were sampled across the wider south west area. Although *S. aria* was included, its current range does not naturally extend westwards of Cheddar (Lat. 51.2879, long. -2.74631). Maps and location details are provided in Chapter 2. Access and variable population sizes resulted in fewer samples for some species. Samples were also collected for all trees participating in all studies; hence more samples of these species were collected e.g. *S. torminalis* (chapter 4) and *S. subcuneata* (chapter 5).

1.10 Conclusion

The presence of naturally occurring polyploid *Sorbus* species in this region, four of which are endemic to north Devon /Somerset, makes this an ideal area to study the evolutionary dynamics of *Sorbus* diversification over an extensive area.

The prevalence of apomixis in polyploid *Sorbus* maintains the integrity of the hybrid genome with the offspring of each successful hybridisation maintained in the environment. This is in contrast to sexual hybrids which may be quickly subsumed within a hybrid swarm or systems where genomic downsizing quickly follows polyploidisation. These features give us opportunities to study the processes involved in generating polyploid diversity and unravelling the steps of polypoid formation.

The pseudogamous nature of apomictic *Sorbus* enables the investigation of pollen flow among species occurring at relatively low densities. This is vital if we are to develop appropriate conservation strategies aimed to maximise opportunities for further polyploid formation. We need to be able to predict the effects of known threats such as habitat degradation and fragmentation on pollen flow, reproductive sustainability and potentially gene flow among species.

1.11 Thesis aims and structure

This study concerns evolutionary diversification and reproductive sustainability in a TCG using *Sorbus* as a model system. Using a combination of molecular techniques and field experiments I investigate the reticulate patterns of evolution among a regional group of *Sorbus* species and determine the reproductive sustainability of threatened *Sorbus* species.

66

This study has three main aims.

- Firstly, to elucidate the evolutionary relationships of a TCG of *Sorbus* species in the Devon and north Somerset region of England, determine hybrid origins and likely route of formation.
- The second aim is to investigate the reproductive sustainability of a key triploid population occurring at low density, with implications for both viability and long term evolutionary potential.
- Thirdly, in light of my findings, I aim to determine whether the current approach to the conservation of the individual threatened species is appropriate or if the process-based approach advocated for TCG's will better conserve the diversity of *Sorbus* in this region.

The first aim is addressed in chapters 2, 3 and 4. In chapter 2 I use nuclear DNA microsatellite markers and flow cytometry to investigate the reticulate relationships among seven polyploid and two diploid *Sorbus* species. Analysis of patterns of genetic diversity within and among these species across their range is used to determine the number of origins for the endemic polyploid species and the dominant breeding system of the study species. The investigation into the relationships among the study species is expounded in Chapter 3 using genome-specific markers and parentage simulations to examine the hybrid origins and routes of polyploid formation in *Sorbus*. I used this method to determine whether certain species are responsible for producing hybrids and are possibly important in driving diversification in this region and identify if particular sites are hot spots for the generation of novel diversity. In Chapter 4 I use flow cytometry to determine the frequency of triploid forms of

the normally diploid progenitor *S. torminalis* across the region, to assess whether this is a potential route for polyploid formation.

The second aim is addressed in chapter 5. Having noticed that seed production was exceedingly low in the only triploid species of this group and identified that this vulnerable endemic has performed a key role in the production of tetraploids, I decided to conduct a series of pollination experiments and molecular analysis of embryo and endosperm to identify how breeding system and factors affecting pollination may cause reproductive failure.

In chapter 6, I synthesize the findings of the experimental chapters and use them to address the third more philosophical aim of conservation approach and ethics. Should we be concerned less about the rarity of some of these species, accepting them as transient stepping stones in a diversifying process and focussing more on promoting the interaction among key species? Or is there a role for a more traditional species based approach with perhaps *ex situ* components performed by botanic gardens and living collections.

The research chapters are intended to be self-contained and independent documents are formatted for submission to relevant journals; however some cross referencing has been allowed to avoid repetition of common methods.

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R, Butlin RK, Dieckmann U, Eroukhmanoff F, Grill A, Cahan SH, Hermansen JS, Hewitt G, Hudson AG, Jiggins C, Jones J, Keller B, Marczewski T, Mallet J, Martinez-Rodriguez P, Möst M, Mullen S, Nichols R, Nolte AW, Parisod C, Pfennig K, Rice AM, Ritchie MG, Seifert B, Smadja CM, Stelkens R, Szymura JM, Väinölä R, Wolf JBW, Zinner D. 2013. Hybridization and speciation. *Journal of Evolutionary Biology,* **26**: 229-246.

Abbott RJ, Hegarty MJ, Hiscock SJ, Brennan AC. 2010. Homoploid hybrid speciation in action. *Taxon,* **59**: 1375-1386.

Aldasoro JJ, Aedo C, Navarro C, Garmendia FM. 1998. The Genus *Sorbus* (Maloideae, Rosaceae) in Europe and in North Africa: Morphological analysis and systematics. *Systematic Botany,* **23**: 189-212.

Andreakis N, Kooistra WH, Procaccini G. 2009. High genetic diversity and connectivity in the polyploid invasive seaweed *Asparagopsis taxiformis* (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. *Mol Ecol,* **18**: 212-26.

Arnaud-Haond S, Alberto F, Teixeira S, Procaccini G, Serrão EA, Duarte CM. 2005. Assessing genetic diversity in clonal organisms: low diversity or low resolution? Combining power and cost efficiency in selecting markers. *Journal of Heredity,* **96**: 434-440.

Arnaud-Haond S, Duarte CM, Alberto F, Serrão EA. 2007. Standardizing methods to address clonality in population studies. *Molecular Ecology,* **16**: 5115-5139.

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

Bednorz L. 2007. Morphological variability of fruits and seeds of *Sorbus torminalis* in Poland. *Dendrobiology,* **57**: 3-14.

Bednorz L, Maciejewska-Rutkowska I, Wronska-Pilarek D, Fujiki T. 2005. Pollen morphology of the Polish species of the genus *Sorbus* L. *Acta Societatis Botanicorum Poloniae,* **74**.

Bennett MD. 2004. Perspectives on polyploidy in plants – ancient and neo. *Biological Journal of the Linnean Society,* **82**: 411-423.

Boyd WE, Dickson JH. 1987. A post-glacial pollen sequence from Loch a'Mhuilinn, north Arran: A record of vegetation history with special reference to the history of endemic *Sorbus* species. *The New Phytologist,* **107**: 221-244.

Bretagnolle F, Thompson JD. 1995. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologist,* **129**: 1-22.

Briggs D, Walters SM. 1997. *Plant variation and evolution.* Cambridge: Cambridge University Press.

Brochmann C, Brysting AK, Alsos IG, Borgen L, Grundt HH, Scheen AC, Elven R. 2004. Polyploidy in arctic plants. *Biological Journal of the Linnean Society,* **82**: 521-536.

Brownfield L, Köhler C. 2010. Unreduced gamete formation in plants: mechanisms and prospects. *Journal of Experimental Botany*: erq371.

Burton TL, Husband BC. 2000. Fitness differences among diploids, tetraploids, and their triploid progeny in *Chamerion angustifolium:* mechanisms of inviability and implications for polyploid evolution. *Evolution,* **54**: 1182-1191.

Burton TL, Husband BC. 2001. Fecundity and offspring ploidy in matings among diploid, triploid and tetraploid *Chamerion angustifolium* (Onagraceae): consequences for tetraploid establishment. *Heredity,* **87**: 573-582.

Carputo D, Frusciante L, Peloquin SJ. 2003. The Role of 2n Gametes and Endosperm Balance Number in the Origin and Evolution of Polyploids in the Tuber-Bearing Solanums. *Genetics,* **163**: 287-294.

Catalán P, Segarra-Moragues JG, Palop-Esteban M, Moreno C, González-Candelas F. 2006. A Bayesian approach for discriminating among alternative inheritance hypotheses in plant polyploids: the allotetraploid origin of genus *Borderea* (Dioscoreaceae). *Genetics,* **172**: 1939-1953.

Chester M, Cowan RS, Fay MF, Rich TCG. 2007. Parentage of endemic *Sorbus* L. (Rosaceae) species in the British Isles: evidence from plastid DNA. *Botanical Journal of the Linnean Society,* **154**: 291-304.

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

Cresswell JE. 2006. Models of pollinator-mediated gene dispersal in plants. In: Harder LD, Barrett SCH, eds. *Ecology and Evolution of Flowers.* Oxford: Oxford University Press.

Cunha C, Doadrio I, Abrantes J, Coelho MM. 2011. The evolutionary history of the allopolyploid *Squalius alburnoides* (Cyprinidae) complex in the northern Iberian Peninsula. *Heredity,* **106**: 100-112.

De Queiroz K. 2005. Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences,* **102**: 6600-6607.

Dickinson TA, Lo E, Talent N. 2007. Polyploidy, reproductive biology, and Rosaceae: understanding evolution and making classifications. *Plant Systematics and Evolution,* **266**: 59-78.

Dufresne F, Stift M, Vergilino R, Mable BK. 2014. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Molecular Ecology,* **23**: 40-69.

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

Esselink GD, Nybom H, Vosman B. 2004. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting—peak ratios) method. *Theoretical and Applied Genetics,* **109**: 402-408.

Fay MF, Gernandt DS, Cowan RS, Kitchen MAR, Kitchen C, Rich TCG. 2002. Parentage of an unknown member of the *Sorbus latifolia* (Lam.) Pers. group (Rosaceae). *Watsonia,* **24**: 91-100.

Felber F, Bever JD. 1997. Effect of triploid fitness on the coexistence of diploids and tetraploids. *Biological Journal of the Linnean Society,* **60**: 95-106.

Frankham R, Ballou JD, Dudash MR, Eldridge MDB, Fenster CB, Lacy RC, Mendelson JR, Porton IJ, Ralls K, Ryder OA. 2012. Implications of different species concepts for conserving biodiversity. *Biological Conservation,* **153**: 25-31.

Freeland JR, Kirk H, Petersen S. 2011. *Molecular Markers in Ecology.* Oxford: Wiley-Blackwell.

Garcia-Verdugo C, Calleja JA, Vargas P, Silva L, Moreira O, Pulido F. 2013. Polyploidy and microsatellite variation in the relict tree *Prunus lusitanica* L.: how effective are refugia in preserving genotypic diversity of clonal taxa? *Molecular Ecology,* **22**: 1546-57.

González-González EA, González-Pérez MA, Rivero E, Sosa PA. 2010. Isolation and characterization of microsatellite loci in *Sorbus aria* (Rosaceae). *Conservation Genetics Resources,* **2**: 341-343.

Grant V. 1981. *Plant speciation.* New York: Columbia University Press

Grünwald NJ, Goodwin SB, Milgroom MG, Fry WE. 2003. Analysis of genotypic diversity data for populations of microorganisms. *Phytopathology,* **93**: 738-746.

Hamston T. 2009. An investigation into the population size, local distribution and reproductive potential of *Sorbus rupicola* (Syme) Hedl. and *Sorbus porrigentiformis* E.F. Warb. in Torbay. (Unpublished MSc thesis), University of Birmingham, Birmingham

Hajrudinović A, Frajman B, Schönswetter P, Silajdžić E, Siljak-Yakovlev S, Bogunić F. 2015a. Towards a better understanding of polyploid *Sorbus* (Rosaceae) from Bosnia and Herzegovina (Balkan Peninsula), including description of a novel, tetraploid apomictic species. *Botanical Journal of the Linnean Society*.

Hajrudinović A, Siljak-Yakovlev S, Brown SC, Pustahija F, Bourge M, Ballian D, Bogunić F. 2015b. When sexual meets apomict: genome size, ploidy level and reproductive mode variation of *Sorbus aria s.l.* and *S. austriaca* (Rosaceae) in Bosnia and Herzegovina. *Annals of Botany,* **116**: 301-312.

Hegarty M, Hiscock S. 2007. Polyploidy: doubling up for evolutionary success. *Current Biology,* **17**: R927-R929.

Hegarty MJ, Barker GL, Brennan AC, Edwards KJ, Abbott RJ, Hiscock SJ. 2009. Extreme changes to gene expression associated with homoploid hybrid speciation. *Molecular Ecology,* **18**: 877-889.

Hollingsworth P, Squirrell J, Hollingsworth M, Richards A, Bateman R. 2006. Taxonomic complexity, conservation and recurrent origins of self-pollination in *Epipactis* (Orchidaceae). *Current taxonomic research on the British and European flora. London: Botanical Society of the British Isles*: 27-44.

Hollingsworth PM. 2003. Taxonomic complexity, population genetics, and plant conservation in Scotland. *Transactions and Proceedings of the Botanical Society of Edinburgh and Botanical Society of Edinburgh Transactions,* **55**: 55-63.

Hörandl E, Greilhuber J, Klímová K, Paun O, Temsch E, Emadzade K. 2009. Reticulate evolution and taxonomic concepts in the *Ranunculus auricomus* complex (Ranunculaceae): insights from analysis of morphological, karyological and molecular data. *Taxon,* **58**: 1194-1216.

Horandl E. 2010. The evolution of self-fertility in apomictic plants. *Sexual Plant Reproduction,* **23**: 73-86.

Houston L, Robertson A, Jones K, Smith SCC, Hiscock S, Rich TCG. 2009. An account of the Whitebeams (*Sorbus* L., Rosaceae) of Cheddar Gorge, England, with description of three new species. *Watsonia,* **27**: 283.

Husband B, C. , Schemske D, W. . 2000. Ecological mechanisms of reproductive isolation between diploid and tetraploid *Chamerion angustifolium*. *Journal of Ecology,* **88**: 689-701.

Husband BC. 2004. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. *Biological Journal of the Linnean Society,* **82**: 537-546.

IUCN. 2001. *IUCN Red List categories and criteria version 3.1.* Gland, Switzerland: IUCN.

Kamvar ZN, Tabima JF, Grunwald NJ. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ,* **2**: e281.

Keble MW, Fraser GT. 1939. *Flora of Devon.* Arbroath.

Köhler C, Mittelsten Scheid O, Erilova A. 2010. The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics,* **26**: 142-148.

Lander TA, Boshier DH, Harris SA. 2010. Fragmented but not isolated: Contribution of single trees, small patches and long-distance pollen flow to genetic connectivity for *Gomortega keule,* an endangered Chilean tree. *Biological Conservation,* **143**: 2583-2590.

Laport RG, Hatem L, Minckley RL, Ramsey J. 2013. Ecological niche modeling implicates climatic adaptation, competitive exclusion, and niche conservatism among Larrea tridentata cytotypes in North American deserts. *The Journal of the Torrey Botanical Society,* **140**: 349-363.

Lepší M, Lepší P, Koutecky P, Bila J, Vit P. 2015. Taxonomic revision of *Sorbus* subgenus *Aria* occurring in the Czech Republic. *Preslia,* **87**: 109-162.

Lepší M, Ví t P, Lepší P, Boublík K, Suda J. 2008. *Sorbus milensis*, a new hybridogenous species from northwestern Bohemia. *Preslia,* **80**: 229-244.

Levin DA. 1975. Minority cytotype exclusion in local plant populations. *Taxon*: 35-43.

Liljefors A. 1953. Studies on propagation, embryology, and pollination in *Sorbus. Acta Horti Bergiani,* **16**: 277-329.

Liljefors A. 1955. Cytological studies in *Sorbus. Acta Horti Bergiani,* **17**: 47-113.

Linder CR, Rieseberg LH. 2004. Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany,* **91**: 1700-1708.

Lo EYY, Stefanovic S, Dickinson TA. 2009. Population genetic structure of diploid sexual and polyploid apomictic hawthorns (*Crataegus*; Rosaceae) in the Pacific Northwest. *Molecular Ecology,* **18**: 1145-1160.

Lowe A, Harris S, Ashton P. 2009. *Ecological genetics: design, analysis, and application*: John Wiley & Sons.

Ludwig S, Robertson A, Rich TCG, Djordjevic M, Cerovic R, Houston L, Harris SA, Hiscock SJ. 2013. Breeding systems, hybridization and continuing evolution in Avon Gorge *Sorbus. Annals of Botany*.

Ludwig SC. 2013. Breeding systems, pollen flow and continuing evolution in Avon Gorge *Sorbus* (whitebeams, rowans and service trees)*,* PhD Thesis, University of Bristol, Bristol.

Mace GM. 2004. The role of taxonomy in species conservation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* **359**: 711-719.

Majesky L, Vasut RJ, Kitner M, Travnicek B. 2012. The pattern of genetic variability in apomictic clones of *Taraxacum officinale* indicates the alternation of asexual and sexual histories of apomicts. *PLoS One,* **7**: e41868.

Mallet J. 2007. Hybrid speciation. *Nature,* **446**: 279-283.

Masterson J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science-AAAS-Weekly Paper Edition-including Guide to Scientific Information,* **264**: 421-423.

Mayr E. 1942. *Systematics and the origin of species, from the viewpoint of a zoologist*: Harvard University Press.

McIntyre PJ. 2012. Polyploidy associated with altered and broader ecological niches in the *Claytonia perfoliata* (Portulacaceae) species complex. *American Journal of Botany,* **99**: 655-662.

Milgroom MG. 1996. Recombination and the multilocus structure of fungal populations. *Annual review of phytopathology,* **34**: 457-477.

Nelson-Jones E, Briggs D, Smith A. 2002. The origin of intermediate species of the genus *Sorbus*. *Theoretical and Applied Genetics,* **105**: 953-963.

Obbard DJ, Harris SA, Pannell JR. 2006. Simple allelic-phenotype diversity and differentiation statistics for allopolyploids. *Heredity,* **97**: 296-303.

Oddou-Muratorio S, Klein EK, Demesure-Musch B, Austerlitz F. 2006. Real-time patterns of pollen flow in the wild-service tree, *Sorbus torminalis* (Rosaceae). III. Mating patterns and the ecological maternal neighborhood. *American Journal of Botany,* **93**: 1650-1659.

Oddou-Muratorio S, Petit RJ, Le Guerroue B, Guesnet D, Demesure B. 2001. Pollen- versus seed-mediated gene flow in a scattered forest tree species. *Evolution,* **55**: 1123-1135.

Otto SP, Whitton J. 2000. Polyploid incidence and evolution. *Annual Review of Genetics,* **34**: 401-437.

Paun O, Greilhuber J, Temsch EM, Hörandl E. 2006. Patterns, sources and ecological implications of clonal diversity in apomictic *Ranunculus carpaticola* (*Ranunculus auricomus* complex, Ranunculaceae). *Molecular Ecology,* **15**: 897-910.

Pellicer J, Clermont S, Houston L, Rich TCG, Fay MF. 2012. Cytotype diversity in the *Sorbus* complex (Rosaceae) in Britain: sorting out the puzzle. *Annals of Botany,* **110**: 1185-1193.

Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, Kunin WE. 2010. Global pollinator declines: trends, impacts and drivers. *Trends in ecology & evolution,* **25**: 345-353.

Price D, Rich T. 2007. One-way introgressive hybridisation between *Sorbus aria* and *S. torminalis* (Rosaceae) in southern Britain. *Watsonia,* **26**: 419-432.

Proctor MCF, Proctor ME, 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.

Ramsey J, Schemske DW. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics,* **29**: 467-501.

Ramsey J, Schemske DW. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics,* **33**: 589-639.

Rich TCG. 2009. Pollen stainability in British *Sorbus* L. (Rosaceae). *Plant Ecology & Diversity,* **2**: 85-88.

Rich TCG, Green D, Houston L, Lepší M, Ludwig S, Pellicer J. 2014. British *Sorbus* (Rosaceae): six new species, two hybrids and a new subgenus. *New Journal of Botany,* **4**: 2-12.

Rich TCG, Harris SA, Hiscock SJ. 2009. Five new *Sorbus* (Rosaceae) taxa from the Avon Gorge, England. *Watsonia,* **27**: 217.

Rich TCG, Houston L, Robertson A, Proctor MCF. 2010. *Whitebeams, rowans and service trees of Britain and Ireland. A monograph of British and Irish Sorbus L. B.S.B.I. Handbook No. 14.* London: Botanical Society of the British Isles.

Rich TCG, Jermy AC. 1998. *Plant crib 1998*: Botanical Society of the British Isles (BSBI).

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

Rich TCG, Proctor MCF. 2009. Some new British and Irish *Sorbus* L. taxa (Rosaceae). *Watsonia,* **27**: 207-216.

Richards A. 1986. *Plant breeding systems.* London: Chapman and Hall.

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

Richards AJ. 2003. Apomixis in flowering plants: an overview. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **358**: 1085-1093.

Richards CM, Church S, McCauley DE. 1999. The influence of population size and isolation on gene flow by pollen in *Silene alba. Evolution*: 63-73.

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

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

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

Robertson A, Rich TCG, Allen AM, Houston L, Roberts CAT, Bridle JR, Harris SA, Hiscock SJ. 2010. Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus. Molecular Ecology,* **19**: 1675-1690.

Ruhsam M, Hollingsworth PM, Ennos RA. 2013. Patterns of mating, generation of diversity, and fitness of offspring in a *Geum* hybrid swarm. *Evolution,* **67**: 2728-2740.

Sabara HA, Kron P, Husband BC. 2013. Cytotype coexistence leads to triploid hybrid production in a diploid—tetraploid contact zone of *Chamerion angustifolium* (Onagraceae). *American Journal of Botany,* **100**: 962-970.

Sato T, Maceira N, Lumaret R, Jacquard P. 1993. Flowering characteristics and fertility of interploidy progeny from normal and 2n gametes in *Dactylis glomerata* L. *New phytologist,* **124**: 309-319.

Schlötterer C. 2004. The evolution of molecular markers—just a matter of fashion? *Nature Reviews Genetics,* **5**: 63-69.

Sell PD. 1989. The *Sorbus latifolia* (Lam.) Pers. aggregate in the British Isles. *Watsonia,* **17**: 385-399.

Snow B, Snow D. 1988. *Birds and berries–A study of ecological interactions.* London: Poyser

Sobel JM, Chen GF, Watt LR, Schemske DW. 2010. The biology of speciation. *Evolution,* **64**: 295-315.

Soltis DE, Soltis PS, Tate JA. 2004. Advances in the study of polyploidy since Plant speciation. *New Phytologist,* **161**: 173-191.

Soltis DE, Visger CJ, Marchant DB, Soltis PS. 2016. Polyploidy: Pitfalls and paths to a paradigm. *American Journal of Botany,* **103**: 1146-1166.

Soltis PS, Soltis DE. 2009. The role of hybridization in plant speciation. *Annual Review of Plant Biology,* **60**: 561-588.

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

Stace CA. 1975. *Hybridization and the Flora of the British Isles.* London: Academic Press.

Talent N. 2009. Evolution of gametophytic apomixis in flowering plants: an alternative model from Maloid Rosaceae. *Theory in Biosciences,* **128**: 121-138.

Theodoridis S, Randin C, Broennimann O, Patsiou T, Conti E. 2013. Divergent and narrower climatic niches characterize polyploid species of European primroses in *Primula* sect. *Aleuritia. Journal of Biogeography,* **40**: 1278-1289.

Thompson KA, Husband BC, Maherali H. 2014. Climatic niche differences between diploid and tetraploid cytotypes of *Chamerion angustifolium* (Onagraceae). *American Journal of Botany,* **101**: 1868-1875.

Vallejo-Marin M, Dorken ME, Barrett SCH. 2010. The ecological and evolutionary consequences of clonality for plant mating. *Annual Review of Ecology, Evolution, and Systematics,* **41**: 193-213.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes,* **4**: 535-538.

Warburg EF. 1962. *Flora of the British Isles.* Cambridge: Cambridge University Press.

Wilmott AJ. 1934. Some interesting British Sorbi. *Proceedings of the Linnean Society of London* **146**: 73–79.

Zhan L, Paterson IG, Fraser BA, Watson B, Bradbury IR, Nadukkalam Ravindran P, Reznick D, Beiko RG, Bentzen P. 2016. MEGASAT: automated inference of microsatellite genotypes from sequence data. *Molecular Ecology Resources.*

Zink RM. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proceedings of the Royal Society of London B: Biological Sciences,* **271**: 561-564.

Chapter 2: Apomixis and hybridisation drives reticulate evolution and phyletic differentiation in *Sorbus* L.: Implications for conservation

*Tracey J. Hamston[1,2], Natasha de Vere[3], R. A. King[1], J. Pellicer[4], M. F. Fay[4,5], James E. Cresswell[1] and Jamie R. Stevens[1]

[1] *Biosciences, College of Life & Environmental Sciences, University of Exeter, Geoffrey Pope Building, Stocker Road, Exeter, EX4 4QD, UK*

[2] *Whitley Wildlife Conservation Trust, Paignton Zoo, TQ4 7EU, UK*

[3] *National Botanic Garden of Wales & Institute of Biological Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, UK*

[4] *Jodrell Laboratory, Royal Botanic Gardens, Kew, TW9 3AB, UK*

[5] *School of Plant Biology, University of Western Australia, Crawley, WA 6009, Australia*

For submission to Heredity

Abstract

Hybridisation and polyploidy are major forces in the evolution of plant diversity and the study of these processes is of particular interest to understand how novel taxa are formed and maintain genetic integrity. *Sorbus* is an example of a genus where active diversification and speciation are ongoing and as such, represents an ideal model to investigate the roles of hybridisation, polyploidy and apomixis in a reticulate evolutionary process. To elucidate breeding systems and evolutionary origins of a complex of closely related *Sorbus* taxa we assessed genotypic diversity and population structure within and among taxa combining data from nuclear DNA microsatellite markers and flow cytometry. Clonal analysis and low genotypic diversity within the polyploid taxa suggest apomixis is obligate. However, microsatellite profiles and site demographics suggest hybridisation events among apomictic polyploid *Sorbus* may have contributed to the extant diversity of recognised taxa in this region. The patterns of mutational variation previously undocumented in polyploid *Sorbus* represent a source of genetic diversity within apomictic lineages. Clonal variation has led to groups of 'clone-mates' within apomictic taxa that strongly suggests mutation is responsible for the genotypic diversity of these clonal lineages. This research demonstrates that both macro- and micro-evolutionary processes are active within this reticulate *Sorbus* complex. Conservation measures should be aimed at maintaining this process and should therefore be prioritised for those areas of *Sorbus* species richness where potential for interspecific gene flow is greatest.

Key words

Hybridisation, polyploidy, apomixis, *Sorbus,* evolution, diversification

81

2.1 Introduction

Hybridisation between species resulting in the formation of new polyploid populations that are distinct and reproductively isolated from the parental taxa is the most common mechanism for sympatric speciation (Grant, 1981; Mallet, 2007; Hendry, 2009). However, the frequency and the main formation routes of polyploid taxa remain unclear (Soltis *et al*, 2010) and studies of hybridisation processes in polyploid species complexes may help to understand this form of speciation.

Apomixis (synonymous with agamospermy; asexual seed production) is often associated with polyploidy (Whitton *et al*, 2008) and effectively causes instant reproductive isolation of novel polyploids from sexual progenitors, enabling sympatric establishment while maintaining the heterozygosity associated with hybridisation. Where apomixis is partial or facultative it allows for occasional exchange of genetic material where such apomicts co-occur with sexual counterparts (Richards, 2003). Apomictic groups develop an intricate variety of morphologically uniform clonal lineages which may be designated as species or microspecies (Grant, 1981), hence leading to much debate over species delineation; examples include: *Rubus* L. (Newton, 1980)*, Taraxacum* (Hughes and Richards, 1988), *Crataegus* L. (Dickinson *et al*, 2008) and *Sorbus* L. (Liljefors, 1955; Rich *et al*, 2010)*.* Hybridisation, polyploidy and apomixis are all features of these and other complex genera and those groups that contain evolutionary young species represent good models to investigate the roles of these processes in plant speciation.

Sorbus (Rosaceae) is a good study group to test the extent of hybridisation among species of varying ploidy and elucidate the role of breeding system in

creation of novel polyploids and establishment of polyploid populations, as the ongoing speciation in *Sorbus* is well described, particularly in Britain (Rich *et al*, 2010; Robertson *et al*, 2010; Ludwig *et al,* 2013). Sexual diploid taxa are pivotal in the creation of novel *Sorbus* polyploids (Liljefors, 1953; Robertson *et al*, 2004b; Lepší *et al*, 2008; Lepší, 2009; Robertson *et al*, 2010; Hajrudinović *et al*, 2015b). Contact zones between sexual diploids and partially apomictic polyploids has produced a reticulation of allopolyploids (polyploids produced from interspecific hybridisation) with varying fertility and ploidy (Ludwig *et al*, 2013; Hajrudinović *et al*, 2015a), In *Sorbus,* where polyploids are geographically isolated from diploids, the role of hybridisation among allotetraploids and divergent mutation of polyploids, both of which may have contributed to the genetic diversity of the *Sorbus* complex has not been fully investigated.

The interest in this genus stems from its evolutionary biology and the conservation status of many *Sorbus* species. Many of the apomictic polyploid taxa are narrow endemics and only exist as small populations; 12 UK species are threatened according to IUCN (2015), making them a priority for conservation. Since the production of these endemic taxa relies on hybridisation there is a growing awareness that process-based conservation is most appropriate, focusing on the evolutionary mechanisms that generate taxonomic complexity rather than a collection of possibly ill-defined individual taxa (Ennos *et al,* 2005). Indeed, such a plan has been proposed for the endemic *Sorbus* of Arran, Scotland (Ennos *et al*, 2012). However, the development of appropriate conservation strategies depends on a detailed knowledge of the processes concerned.

Several important UK sites for *Sorbus* diversity occur in Devon and along the north Somerset coast (Rich *et al*, 2010). This study focuses on a group of seven polyploid taxa, four of which are narrowly endemic and of conservation concern. Their sexual diploid progenitors, *S. torminalis* (L.) Crantz and *S. aria* L. exist at low densities or are not currently native to the region, respectively.

Sorbus rupicola (Syme) Hedlund and *S. porrigentiformis* E.F.Warb. are thought to be the oldest polyploids in our study group, based on their wide distribution (Rich *et al*, 2010) and are possible progenitors for other polyploid *Sorbus* in this region as they have been elsewhere (Liljefors, 1955; Robertson *et al*, 2004b; Robertson *et al*, 2010). *Sorbus vexans* E.F.Warb., *S. margaretae* M.C.F.Proctor*, S. admonitor* M.C.F.Proctor and *S.subcuneata* Wilmott are restricted to areas along the north coast of Devon and Somerset.

Sorbus devoniensis E.F.Warb. is largely found in Devon, however a number are found on sites in southeast Ireland (Rich *et al*, 2010). A schema of proposed relationships between the study species is presented in Fig.2.1a and b. Evidence for these relationships comes from both morphological (Sell, 1989; Warburg, 1962) and molecular studies; Nelson-Jones *et al* (2002) used restriction fragment length polymorphisms (RFLP) to assign hybridogenous polyploid *Sorbus* taxa to various subgenera and plastid DNA identified the ancestral maternal parent (Chester *et al*, 2007). These taxonomic groupings were also supported by previous peroxidase isozyme studies (Proctor *et al*, 1989) which also suggested there may be variation within some polyploid taxa. However, the hybrid origins of our study taxa and in particular the pollen donors were not identified.
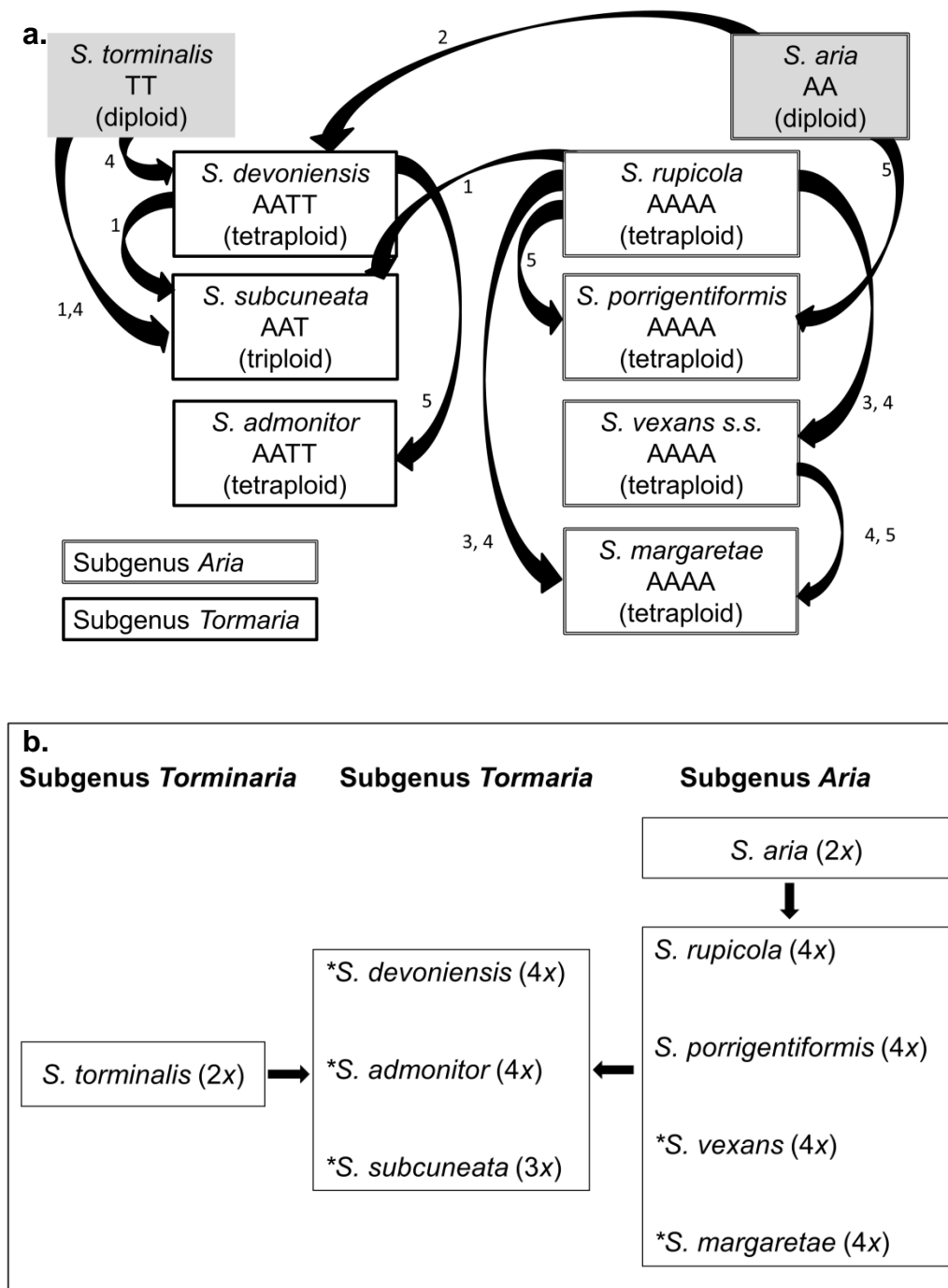
Figure 2.1a. Previously hypothesised relationships among the study *Sorbus* species based on the following literature. 1. Wilmott (1934), 2. Sell (1989), 3. Proctor (1989), 4. Chester (2007), 5. Rich et al. (2010)

b. Summarised relationships among study polyploid and diploid members of three subgenera from Nelson-Jones *et al* (2002). *Sorbus torminalis* is the ancestral maternal parent for subgenus *Tormaria* (Chester *et al*, 2007) and members of subgenus *Aria* are derived from *S. aria*. Assumed ploidy is in parenthesis, from Pellicer *et al* (2012). * Taxa largely restricted to study region.

85

To differentiate among closely allied species with possible common ancestry we chose nuclear DNA microsatellite markers as they are codominant and highly polymorphic due to high mutation rates which allow identification of hybrid parentage (Freeland *et al*, 2011). We used flow cytometry to determine relative nuclear DNA contents and infer ploidy for our species. Flow cytometry is a useful tool for rapidly screening of samples and has been used increasingly to explore hybrid speciation (Siljak-Yakovlev, 2010; Pellicer *et al*, 2012; Hajrudinović *et al*, 2015b). Our sampling strategy sought to encompass the geographical ranges of *S. admonitor, S. subcuneata, S. devoniensis, S. vexans* and *S. margaretae* whilst the remaining potential parental species were sampled more widely to obtain a representative selection of alleles for these taxa (Fig. 2.2).

The principal aims of this study were to elucidate evolutionary relationships among the study taxa and to determine breeding systems within this species complex; in addressing these aims, we explored patterns of genetic structure and diversity. Specifically we addressed the following questions. (1) What are the most likely hybrid origins of the polyploid taxa? (2) Are single or multiple origins evident for the polyploid taxa? (3) Is the apomictic breeding system of polyploid taxa obligate or facultative? (4) What is the source of genetic diversity within and among the polyploid taxa? Finally, we draw on our genetic findings to make robust recommendations for conservation and management of these often rare and complex taxa.

2.2 Materials and Methods

2.2.1 Plant material

Molecular analysis was carried out on 207 individuals of nine *Sorbus* species from 35 sites (Fig. 2.2). Fully expanded disease-free leaf material was collected and small pieces, approximately 1cm^2, were torn into small zip lock plastic bags containing self-indicating silica gel, which dries the sample within 12 hours (Chase & Hills, 1991). They were subsequently stored at room temperature until required. In addition, 145 trees were re-sampled to provide fresh leaf material for use in flow cytometry which was carried out by Jaume Pellicer and Tracey Hamston at the Jodrell Laboratory, Royal Botanic Gardens, Kew. The fresh samples were stored in moist tissue, in polythene bags at 4°C for up to 7 days before use. Voucher specimens were placed in the Welsh National Herbarium, Cardiff (NMW). Each tree had its location described and recorded with a GPS unit. Full details of sample locations, site codes and herbarium accession numbers can be found in Table S2.1, supplementary information.

2.2.2 DNA extraction and molecular markers

DNA was extracted from dried leaf samples with the Qiagen DNeasy plant mini kit (Qiagen, Hilden, Germany) following the manufacturer's protocol with lysis buffer added to samples before being processed at room temperature using a Qiagen TissueLyser bead mill (Qiagen, Hilden, Germany) set at 30Hz for two two-minute cycles with the tube racks rotated between cycles to ensure even levels of tissue disruption. The pure DNA was eluted into 200 μl AE buffer (10 mM Tris·Cl, 0.5 mM EDTA, pH 9.0) and then frozen at -20°C until use.

87

Fourteen previously published microsatellite loci were used; CH01F02, CH01F09 and CH02D11 were developed for use in *Malus × domestica* (Gianfranceschi *et al*, 1998); MSS5, MSS16 and MSS13 for *S. torminalis* (Oddou-Muratorio *et al*, 2001); SA01, SA19.1, SA03, SA06, SA02, SA08, SA09 and SA14 for *S. aria* (González-González *et al*, 2010). Primers for CH01F02, CH01F09, CH02D11 and MSS16 were redesigned by Robertson *et al* (2010) for use in a wide range of *Sorbus* taxa.
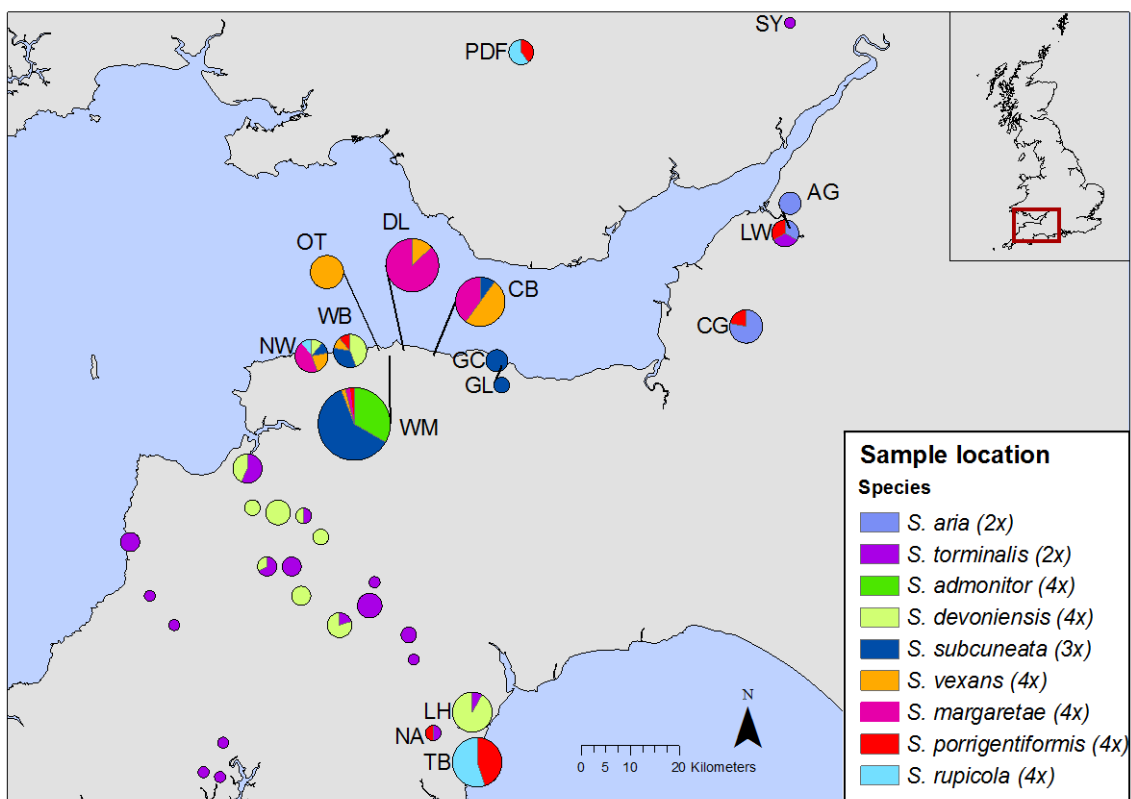
Figure 2.2. Geographic distribution of samples included in our study. Each pie chart represents a site, with pie size relative to site sample size. The inset shows the area covered by the map. Site codes match those in supplementary Table S2.1. (The map was created using ArcGIS Desktop version 10.2.2, ESRI, California, USA, URL: http://www.esri.com/).

Thirteen of the loci were combined in three multiplex groups according to fragment size ranges and dye colour at the amplification stage. SA14 was amplified separately. Microsatellite primer details and multiplex design are given

88

in Tables S2.2 and S2.3. PCR conditions followed a touchdown cycle modified from Hamilton *et al* (2014) and consisted of 95°C for 5 min followed by 30 cycles of 95°C for 30 s, annealing temperature for 1 min 30s and 72°C for 3 min. Annealing temperatures were 62°C (4 cycles), 58°C (4) , 55°C (7), 53°C (12), 51°C (5), 49°C (5), 47°C (5) with a final extension of 72°C for 10 min. Capillary electrophoresis of the PCR products was carried out on a Beckman Coulter sequencer (Beckman Coulter, Fullerton, USA) and fragment analysis was performed using CEQ 8000 Genetic Analysis system (Beckman Coulter) followed by a manual verification of each call to ensure proper peak designation. To determine a standard genotype for each polyploid taxon, a reference selection of all study species was subjected to three PCR repeats and used as an internal standard against which we checked for mutations and scoring error. Any subsequent inconsistent samples were repeated to ensure observed allele sizes were not artefacts of PCR amplification or scoring error.

2.2.3 Estimation of ploidy

Since variation in ploidy within species indicates different routes of formation and origins we investigated cytotype diversity within the study species, extending that of Pellicer *et al*. (2012). Confident ploidy allocation for our study species also avoided potential problems arising from unknown or inconsistent ploidy when interpreting microsatellite amplification patterns. We used flow cytometry to estimate nuclear DNA content and infer ploidy for our species samples. Propidium iodide flow cytometry (FMC) analysis was performed as described by (Pellicer *et al*, 2012) at the Jodrell Laboratory (Royal Botanic Gardens, Kew, UK). Ploidy was inferred by means of the ratio between the target sample peak and that of a known internal standard (*Oryza sativa* 'IR36',

2C = 1 pg, (Bennett and Smith, 1991). Previous FCM sample: standard ratios of all the diploid *Sorbus* species (Pellicer *et al*, 2012) provided an additional baseline against which to compare our samples.

2.2.4 Data analyses

The 14 primer pairs successfully amplified across all polyploid taxa, but primers for four loci (SA19.1, SA02, SA09 and CH01F09) failed to amplify alleles in many *S. aria* and *S. torminalis* individuals (the two putative ancestral diploid taxa). Therefore, analyses which included both diploid and polyploid taxa were based on only ten loci. The additional four loci were included for analysis of polyploids only.

Population genetic structure within Sorbus

A prevalence of apomictic reproduction within the polyploid taxa results in clonal groups of genetically identical individuals. To investigate relationships among these groups the samples were assembled into 82 multi-locus genotypes (MLG's) and the following analyses were performed on these genotypes. STRUCTURE was used to carry out Bayesian clustering analyses to assign a probability to each genotype of belonging to each of K genetic clusters. It allowed identification of putative hybrids and admixed individuals. This was carried out in two steps; firstly with microsatellite data using ten loci for all taxa and secondly using all sample data from fourteen loci for the polyploid taxa. This approach allowed us to describe the patterns of genetic variation among the polyploid groups after allowing for the genetic divergence between diploid and polyploid taxa.

An admixture model was used with a burn-in period of 50,000 with 150,000 iterations on a range of K values (1 to 15) and maximum ploidy was set to 4. This analysis was repeated seven times to enable determination of the optimum number of population clusters (K). The numbers of clusters (K) that best fits the data is inferred by simulating a range of K values. STRUCTURE HARVESTER (Earl and vonHoldt, 2012) executes the 'Evanno' method (Evanno *et al*, 2005) where the mean posterior probability of the data for each given K, L(K) is calculated and the point of maximum rate of change for this value (ΔK) is returned showing a peak corresponding to an optimal K value. For visual representation of the aligned cluster assignments we used the programme STRUCTURE PLOT (Ramasamy *et al*, 2014).

Genetic distance

Where analyses were carried out on matrixes of pair-wise genetic distances between genotypes they were constructed using the Bruvo distance (Bruvo *et al*, 2004) in the POLYSAT package in R (Clark and Jasieniuk, 2011). Bruvo genetic distance is calculated to take into account step-wise mutation processes without the requirement for allele copy number and for individuals to be the same ploidy, thus making it appropriate for use with mixed ploidy samples (Dufresne *et al*, 2014). Different sample sets were used to investigate the various aspects of genetic structure and diversity of our study species. Ploidy information was added according to the results of flow cytometry (FCM).

To infer evolutionary relationships among the study taxa and determine likely hybrid origins, patterns of genetic structure were examined using a principal coordinate analysis (PCoA) of genetic distances using the single MLG microsatellite data at ten loci. The results of the PCoA were visualised in 3D

using the R package 'pca3d' (Weiner, 2015). To summarise the relationships among the groups, the distance matrix was also used to construct a neighbour-joining (NJ) tree using SplitsTree 4 (Huson and Bryant, 2006).

Genotypic diversity

To determine whether the polyploid taxa have simple or multiple origins, identify the breeding systems prevalent within the group of study species and identify sources of any genetic diversity within the polyploid taxa, we analysed the genotypic diversity within and among our study taxon group.

The following diversity calculations were carried out using the microsatellite data at 10 loci from all samples and implemented in POLYSAT. To ensure sample sizes were comparable for allelic diversity statistics, which are affected by sample size, rarefaction was applied to the diploid species *S. torminalis* with the 33 samples randomly sub-sampled to match the sample size of *S. aria* (13 individuals) (Pruett and Winker, 2008). Genotypic diversity within and among study taxa was determined by calculating allelic richness or total number of unique alleles for each species summed across ten loci (A), total number of MLG's (Ng) and genotypic diversity for each of the polyploid species, in terms of the complement of Simpson's index λ (1- λ) (Arnaud-Haond *et al*, 2007). The actual number of MLG's present (Ng) was determined using the 'assign clone function' in POLYSAT with zero as threshold, which considers all pairs of individuals with a non-zero genetic distance as separate MLG's. The threshold value of zero was used to assign individuals to genotypes before calculating Simpson's diversity index (λ) as follows;

92

$$\lambda = \sum \frac{p_i(p_i - 1)}{N(N-1)}$$

This equation gives an unbiased estimator of λ for a sample size of N where p is a vector of genotypes. This calculated a value for genotypic diversity based on the number of MLG's which varies positively with clonal heterogeneity. The complement of λ (1- λ) was used to compare genotypic diversity among all species and described the probability of encountering distinct MLG's when taking two units at random from the sample. The Simpson's complement was recalculated using the polyploid sample data at 14 loci for comparison.

To determine whether any of the diversity seen within the polyploid taxa was best explained by genetic recombination either as a result of an interspecific or intraspecific hybridisation or may be attributed to somatic mutation, we assigned each polyploid sample to a clonal lineage, within which any diversity was considered due to mutation using the method of Douhovnikoff and Dodd (2003). In this case we used data from the 160 polyploid samples at 14 microsatellite loci. To establish a threshold of genetic distance, above which a recombination event would be indicated, we plotted a frequency histogram of all pairwise genetic distances between samples. Such histograms are often multi-modal due to highly uneven relative abundance of clones in the dataset. The position of the valley between the first peak which is close to zero and represents nearly identical genotypes perhaps due to the presence of somatic mutations or scoring errors in the data set, and the second, which represents distinct but closely related clones each deriving from a single reproductive event, is considered an appropriate threshold (Meirmans and Van Tienderen, 2004;

Arnaud-Haond *et al*, 2005). The resulting threshold was then employed to assign all samples to clonal groups using POLYSAT.

Flow cytometry data

Differences among species were tested with one-way ANOVA's with *post-hoc* Tukey comparisons of means to determine where differences lay. The normality of the data distributions was tested using the Shapiro-Wilk test and the homogeneity of variances by Levene's test. All statistical analysis was performed using R (R Development Core Team, 2015).

2.3 Results

2.3.1 Microsatellite markers

Fourteen microsatellite loci were successfully amplified across the seven polyploid species with ten of these successful across all nine study species. Ten loci yielded a total of 154 unique alleles from 186 *Sorbus* samples, ranging between eight (MSS13 and SA03) and 24 (SA14) per locus. The alleles observed at all loci for each polyploid species are given in Table S2.4 (supplementary information). The maximum number of alleles for any individual sample at all loci corresponded with expected ploidy with the exception of one sample of the normally diploid *S. torminalis* which had three alleles at two loci (SA19.1 and SA1; see ploidy analysis).

2.3.2 Population genetic structure and evolutionary relationships.

Our investigation of genetic structure revealed that each of the study species is genetically differentiated although clustering patterns varied between the two subgenera.

94

STRUCTURE analysis indicated that five clusters best described our data when both diploid and polyploid species were analysed together (Fig S2.1). Of these five clusters, the two sexual, diploid species *S. aria* and *S. tormin*alis are distinct but the seven polyploid species are grouped into three clusters irrespective of ploidy. *S. porrigentiformis* clustered separately and contains some microsatellite alleles not found in the other polyploids. It also occupied a position furthest away from all other polyploid taxa in the PCoA. When the polyploid data was analysed separately using an additional four loci, STRUCTURE assigned all individuals to two clusters (K=2), corresponding to the two subgenera. However, one species, *S. margaretae*, clearly shows an intermediate position, along with two samples within *S. vexans*, which we have referred to hereafter as vex2. The sexual diploid taxa *S. aria* and *S. torminalis* were both differentiated from each other and all the polyploid individuals in the PCoA and NJ trees (Figs. 2.3 and 2.4). Samples from the polyploid taxa fall into two groups in the NJ tree which correspond to the two subgenera. However, the three members of subgenus *Tormaria* [one triploid *(S. subcuneata)* and two tetraploids (*S. admonitor* and *S. devoniensis*)], are closely grouped in the PCoA, particularly the latter two tetraploids. Their separation is also weak in the NJ tree where, although distinct clusters are observed they are positioned at the tips of short branches. Individuals of *S. vexans* (4*x)* and *S. rupicola* (4*x)* are also tightly grouped in the PCoA with the exception of vex2 that clearly occupies an intermediate position between *S. vexans* and *S. margaretae* (4*x). *These intermediate positions mirror those seen in the PCoA and NJ tree analyses (Figs. 2.3 and 2.4). *Sorbus margaretae* and *S. porrigentiformis* (4*x*) samples all conform to single, highly differentiated clusters in both analyses.
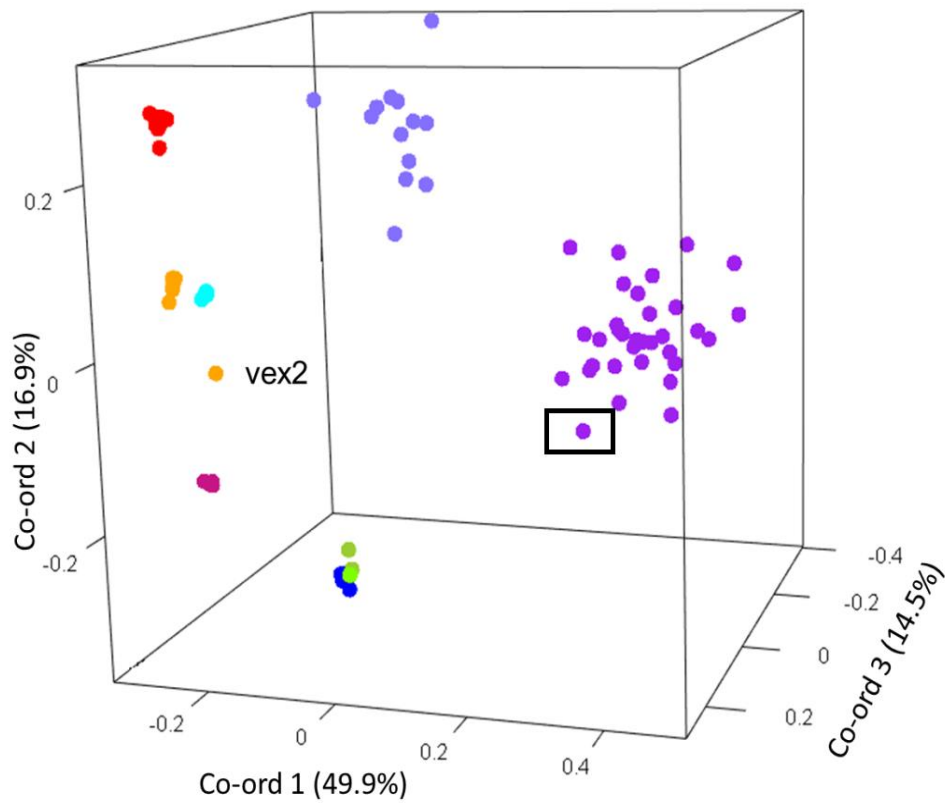
Figure 2.3. Principal Coordinate Analysis of the Bruvo distance matrix of 82 MLG's from nine species based on ten microsatellite loci. Percentages of total variance explained by the co-ordinates are given in parentheses explaining a total of 81.3% of the variation in the matrix. The triploid *S. torminalis* individual is indicated by the black box. ● *S. aria,* ● *S. torminalis,* ● *S. subcuneata,* ● *S. devoniensis,* ● *S. admonitor,* ● *S. margaretae,* ● *S. vexans,* ● vex2, ● *S. rupicola,* ● *S. porrigentiformis.*
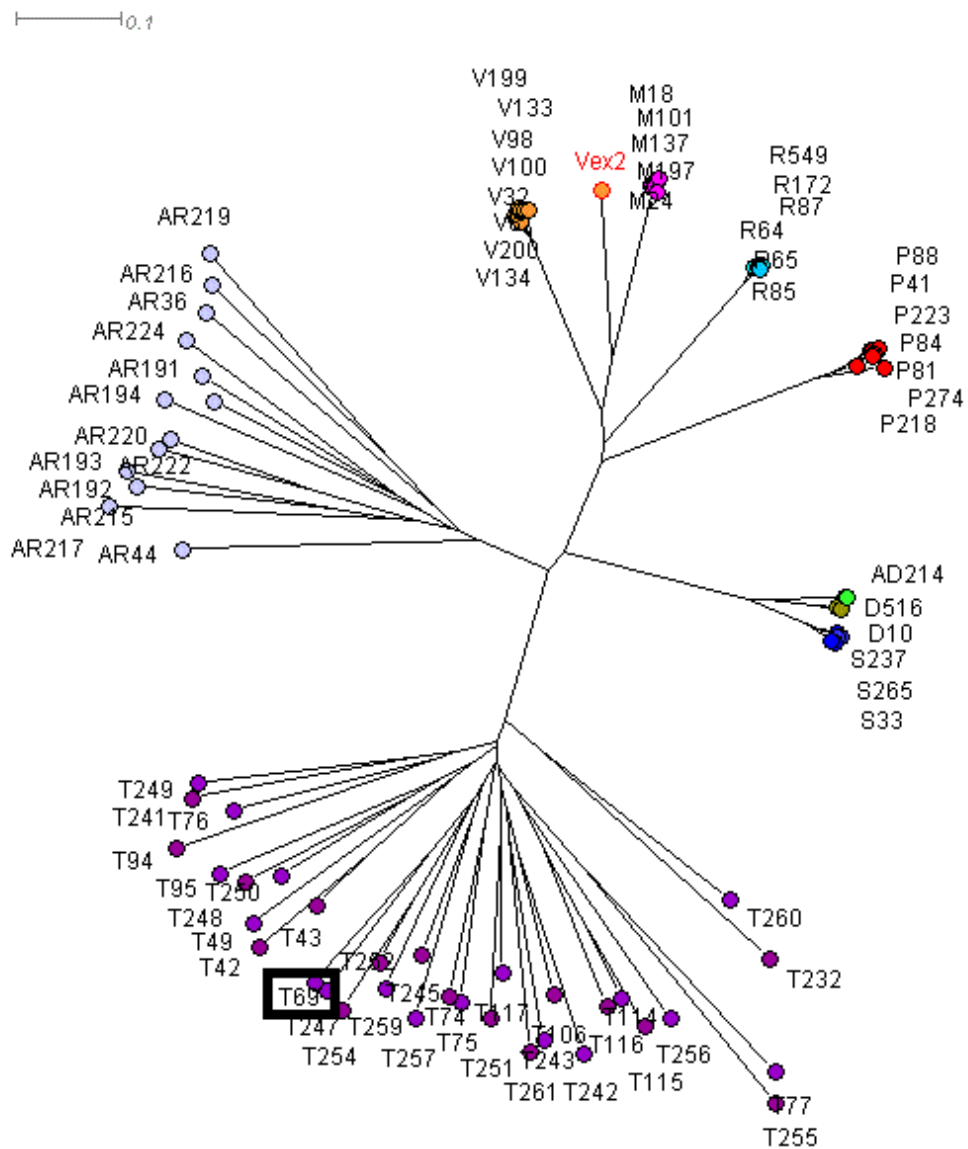
Figure 2.4. Neighbour-joining (NJ) tree of 82 *Sorbus* MLG's constructed using a Bruvo distance matrix in SplitsTree 4.0. Colours are the same as for Fig. 2.3. Vex2 indicates the second *S. vexans* clone. The triploid *S. torminalis* individual is indicated by the black box.

2.3.3 Genotypic diversity within and among taxa

The outcrossing diploid species show high levels of genotypic variation (1- λ) in comparison to the polyploid species which are characterised by few MLG's.

Allelic richness (A), measured across ten loci, was highest for the diploid, sexual species *S. aria* and *S. torminalis.* In contrast, the polyploid taxa had approximately half the number of alleles, despite the larger genome size. There were no private alleles (present in no other taxa) in any of the seven polyploid taxa across the ten loci and there were high levels of allele sharing across all species. Only four of the 95 alleles present in the polyploid taxa were not sampled within the diploid species. The polyploids together contained 62% of the total alleles observed. When we used a zero threshold to identify unique MLG's (N_g), the number of MLG's observed in the diploids equalled the sample number as would be expected for sexual, outcrossing taxa. In contrast, low numbers of MLG's were detected within the apomictic polyploid species, although all showed more than one MLG except tetraploid *S. admonitor.* The complement of Simpson's diversity ($1 - \lambda$) at ten loci also reflects this pattern with members of subgenus *Tormaria* ranging from 0 (*S. admonitor*) to 0.27 for triploid *S. subcuneata*. *Sorbus margaretae* shows the lowest genotypic diversity in subgenus *Aria*, and *S. vexans s.l.* the highest. At 14 loci, the values of $1 - \lambda$ are higher; reflecting the greater number of loci but the relative diversity of the polyploid taxa follows a similar pattern. These diversity statistics are summarised in Table 2.1.

The frequency histogram of all pairwise distances is multi-modal with a clear peak at zero indicating the abundance of replicate genotypes due to an apomictic mode of reproduction (Fig. 2.5). The threshold distance between the first and second peak is 0.09, and this was used to assign all polyploid genotypes to a clonal lineage. Each polyploid species corresponded to a single clonal lineage with the exception of *S. vexans* which had two clones which conformed to two distinct genotypes that differed at all 14 loci. This result is

98

inconsistent with only mutational variation and reveals a separate sexual origin

for each *S. vexans* clone. Vex2 forms the second clone, consisting of two

identical samples. The diversity statistics for *S. vexans* were calculated with

and without vex2; *S. vexans s.l.* and *S. vexans s.s.*, respectively.
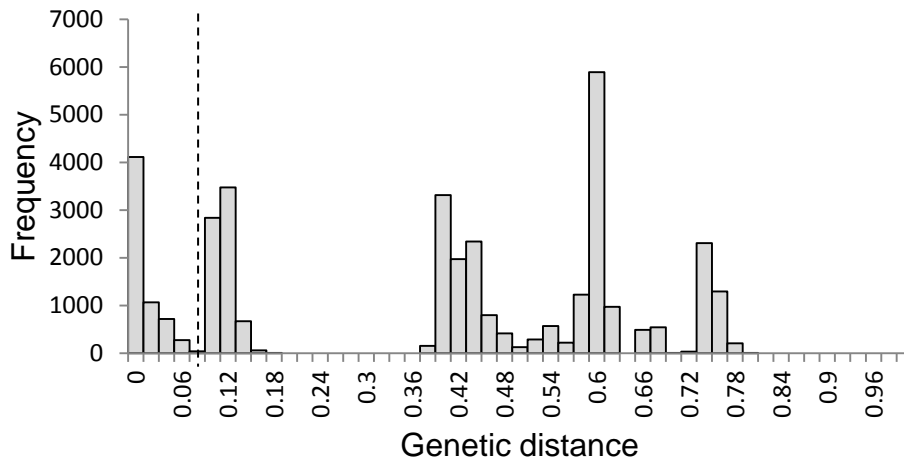
Figure 2.5. Distribution of pair-wise genetic distances between all polyploid

individuals. The dashed line represents the threshold distance applied to

separate asexually and sexually related individuals = 0.09.

2.3.4   Ploidy analysis

We allocated ploidies for 145 samples from all nine species. The flow cytometric

analysis revealed three cytotypes; diploid ($2x$), triploid ($3x$) and tetraploid ($4x$).

The *Sorbus* samples and the internal size (*Oryza sativa*) produced clear flow

histogram peaks with low coefficients of variation (Table 2.3. CV%: 1.91-3.74;

mean = 2.54 ± 0.39).

Determination of ploidy levels. The ratios between the peaks and internal size

(S) were used to infer the ploidy of each sample based on previous results in

Sorbus (Pellicer *et al*, 2012). Comparison of the mean peak ratios confirmed

that each of the three ploidies is of significantly different size [Fig. S2.2; One-

99

way analysis of means (not assuming equal variances); $F = 15132$, DF = 2, $n = 145$, $p = < 2.2\text{e-}16$].

Table 2.1. Allelic and genotypic diversity found in the studied *Sorbus* species. N = approximate population size in study region across sampled sites (Rich *et al*, 2010); N_i = number sampled; X = ploidy from flow cytometry and total number of alleles at any locus; A = total number of alleles observed at ten loci; Ng = Total number of multi-locus genotypes; 1- λ = Simpson's complement for all species at both ten and 14 loci with values in parenthesis for *S. vexans s.s.*; Nc = number of clonal linages when a threshold of 0.09 is applied to delineate between sexual and asexual relations between samples (apomictic polyploid species only). Diversity indices for *S. torminalis* were calculated on a sub sample of 13 of the 33 sampled individuals.

Taxon	N	N_i	X	A	Ng	1- λ (10 loci)	1- λ (14 loci)	Nc
S. aria		13	2	65	13	1	n/a	-
S. torminalis		13	2, 3	74	13	1	n/a	-
S. admonitor	c.110	19	4	29	1	0	0	1
S. devoniensis	>450	31	4	30	2	0.06	0.17	1
S. subcuneata	c.300	27	3	26	3	0.27	0.34	1
S. margaretae	c.100	29	4	33	5	0.43	0.48	1
S. porrigentiformis	>100	17	4	36	7	0.74	0.78	1
S. rupicola	c.40	13	4	35	6	0.79	0.79	1
S. vexans s.l.	c.70	24(22)	4	46 (35)	9 (8)	0.86(0.84)	0.86(0.84)	2

Apomictic taxa were all polyploid and the *S. torminalis* sample analysed, which showed three alleles at two loci had a 2C = 2.381 pg confirming its triploid status. We also confirmed that *S. subcuneata* is a triploid as all samples had nuclear DNA contents of 2.286 - 2.386 pg, which is consistent with a triploid cytotype. Previously, *S. subcuneata* had been thought to have both triploid and tetraploid cytotypes. The remaining six polyploid species were all confirmed as tetraploid. However, the tetraploids had a larger variance than the other groups, with the subgenus *Tormaria* species (*S. admonitor* and *S. devoniensis*) displaying significantly larger genome sizes than the four tetraploids of subgenus *Aria* (Fig. S2.3; ANOVA; F = 16.738, n = 87, p = <0.001).

Table 2.2 Nuclear DNA content (pg) of each of the clusters with inferred ploidy and chromosome number (2*n*). Triploid *S. torminalis* is shown separately. N = sample size.

| Ploidy | 2*n* | N | Relative 2C DNA (pg) | | CV |
			Min-max (pg)	Mean ± s.d.	(%)
2*x (S. torminalis)*	34	26	1.600 - 1.678	**1.63** ± 0.018	3.25
3*x (S. torminalis)*	51	1	2.38	**2.38**	2.98
3*x (S. subcuneata)*	51	31	2.286 - 2.386	**2.33** ± 0.022	2.52
4*x*	68	87	2.857 - 3.231	**3.07** ± 0.084	2.53

2.4 Discussion

Our investigation of relationships among often sympatric populations of polyploid *Sorbus* taxa revealed that each taxon is genetically differentiated and characterised by few multi-locus genotypes. Each MLG is composed largely of alleles common to other taxa resulting in high levels of allele sharing among all study taxa. These results support the broad hypothesis based on previous studies (Robertson *et al*, 2004b; Robertson *et al*, 2010; Hajrudinović *et al*, 2015b) that the route of polyploid formation in *Sorbus* taxa is via interspecific hybridisation and the genetic integrity of each polyploid taxon is maintained via apomixis.

2.4.1 Relationships among polyploid and diploid taxa

The use of nuclear microsatellite markers has enabled us to refine the hypothesised evolutionary relationships shown in Fig. 2.1a and b. Using the data from this study we now propose the evolutionary relationships among our study taxa shown in Fig. 2.6. This is discussed below and routes of formation are explored further in Chapter 3.

The currently classified species correspond to discrete genetic clusters shown in the PCoA and NJ trees (Figs. 2.3 & 2.4). The close relationships among polyploid taxa indicate probable linkage through ancestral hybrid events or ongoing gene flow. Shared hybrid origins are the most likely reason for the many shared alleles observed with frequent gene flow an unlikely explanation due to the predominance of apomixis among the study polyploid taxa.

Sorbus rupicola is the most likely parental polyploid species for endemics *S. vexans* and *S. margaretae,* either directly or indirectly, rather than *S. porrigentiformis.* Although they both occur across the range of the local endemic

103

taxa within our group, the PCoA shows *S. rupicola* is more closely related to other members of subgenus *Aria* in our study group. This is in contrast to other sites within its range where *S. porrigentiformis* is thought to be one of the primary parental taxa, hybridising with diploid *S. aria s.s.* (Houston *et al*, 2009; Robertson *et al*, 2010)*.
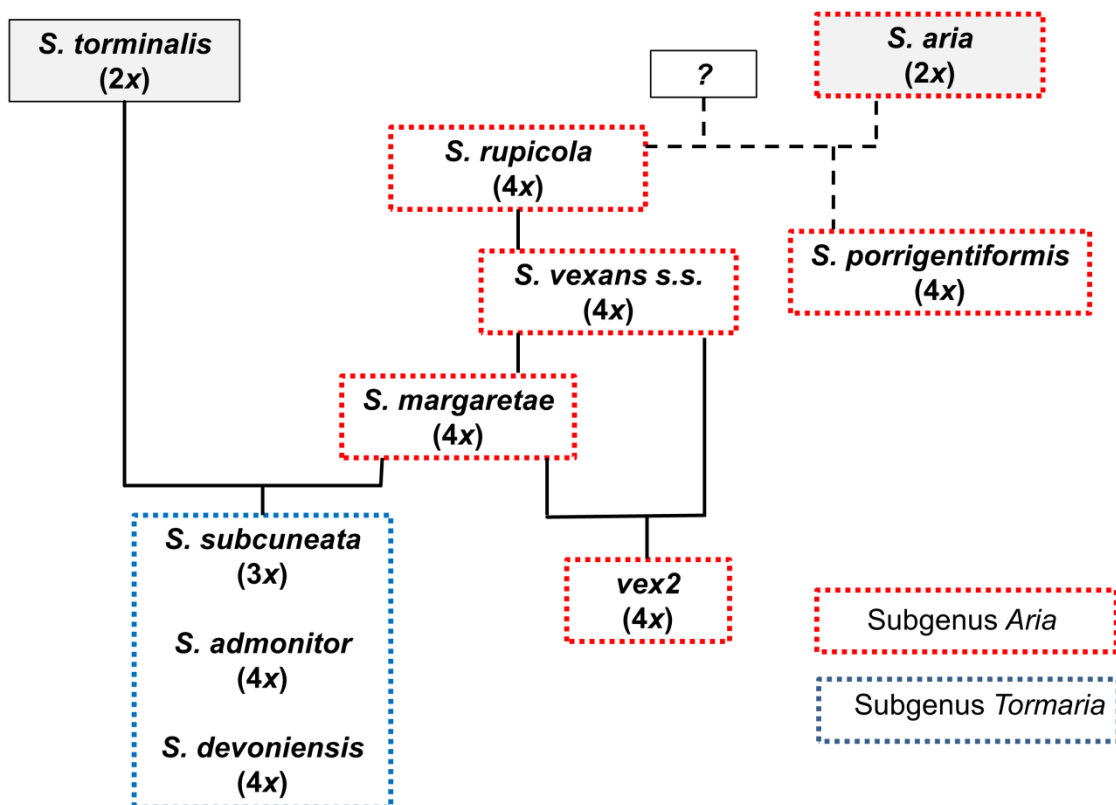
Figure 2.6 Proposed relationships among the south-west *Sorbus* taxa. Ploidy levels as determined by FCM are given in parentheses. Dashed lines indicate speculative relationships and ? indicates possible missing intermediate species.

Sorbus porrigentiformis shares alleles with *S. rupicola* at 11 of the 14 loci, so it seems likely that *S. porrigentiformis* is also derived from *S. rupicola*, maybe indirectly, in line with proposed theories (Rich *et al*, 2010). However, as with previous studies (Robertson *et al*, 2010), their relationship remains unclear. Our results also suggest that *S. margaretae* may have a more recent origin than *S.*

104

vexans s.s. Their relative genotypic variation due to mutation (1- λ values: *S. margaretae,* 0.43 vs *S. vexans s.s.,* 0.84) indicated a more recent origin, since mutations accrue over time since divergence (Ellegren, 2000b). We postulate that if this is so, *S. margaretae* is most likely derived from a hybridisation involving *S. vexans s.s* and unknown taxa. They are closely related, sharing alleles at every locus except SA01 (Table S2.4) and have very similar leaf morphologies, being hard to distinguish in the field. However, they are genetically distinct and our clonal analysis attributes their genetic differences to sexual reproduction rather than genetic mutation. *Sorbus vexans s.s.* showed greater affinity to *S. rupicola* in the PCoA analysis which would be explained if it is directly derived from *S. rupicola.* Both *S. vexans* and *S. margaretae* are endemic to this region and in the absence of *S. aria* their origin could be via allotetraploid hybridisation rather than a diploid × polyploid cross.

The second *S. vexans* clone (vex2) represents a separate genotype resulting from interspecific hybridisation rather than sexual reproduction within the taxon, since alleles from more than one diploid taxon are present. The intermediate position of the vex2 between *S. vexans* and *S. margaretae* in the PCoA, suggests it may be a hybrid involving these two tetraploids, especially since the ancestral diploid progenitor for subgenus *Aria*, *S. aria,* is not present in the locality. Vex2 occurs on a small (<3 ha) coastal site (Neck Wood, north Devon, Fig. 2.2), which has a high diversity of polyploid *Sorbus* species: *S. subcuneata, S. devoniensis, S. margaretae,* S*. rupicola* and *S. vexans*, and specimens of *S. intermedia* (Ehrh) Pers*.,* a non-native that has become naturalised. The vex2 variant is genetically unique and its possible derivation from polyploid taxa in the absence of parental diploid forms suggests a possible route for polyploid

Sorbus formation and provides strong evidence of ongoing diversification in the region.

Our analyses confirm the accepted view that members of subgenus *Tormaria* (*S. subcuneata, S. admonitor* and *S. devoniensis*) are distinct from, but intermediate to *S. torminalis* and subgenus *Aria,* in line with the hybrid origins proposed by (Nelson-Jones *et al*, 2002). The larger genome size of *S. admonitor* and *S. devoniensis*, when compared to the tetraploid members of subgenus *Aria,* adds weight to their hybrid origin with *S. torminalis* as an ancestral parent (Chester *et al*, 2007) since *S. torminalis* has the largest genome size of the three diploid *Sorbus* species tested by Pellicer *et al* (2012). The intermediate position of these members of subgenus *Tormaria* between *S. torminalis* and tetraploid *S. margaretae* (subgenus *Aria*) suggests *S. margaretae* may be the male ancestral parent of these representatives of subgenus *Tormaria* rather than *S. aria* or *S. rupicola* as proposed by Sell (1989). This relationship would explain the many shared alleles among *S. margaretae, S. devoniensis, S. admonitor* and *S. subcuneata.* Subgenus *Tormaria* forms a tight group of very closely related taxa in all our cluster analyses. *Sorbus subcuneata*, a triploid, shares all its alleles with both tetraploids *S. devoniensis* and *S. admonitor,* across all 14 loci, implying a common recent origin for the group, possibly with *S. subcuneata* as an ancestral species for *S. admonitor* and *S. devoniensis*, which poses the taxonomic question as to whether these should be considered variants of the same species with multiple origins as suggested by Proctor *et al* (1989) and Sell (1989). The geographical distribution of triploid *S. subcuneata* overlaps with *S. admonitor* and *S. devoniensis*, but the spatial separation of the latter two suggest they may have arisen in different locations.

106

These relationships suggest the triploid *S. subcuneata* may be involved in tetraploid formation. Closer analyses of allelic patterns to test whether this is via the 'triploid bridge' are performed in Chapter 3. However, this route would be consistent with the formation of tetraploid *Sorbus* elsewhere (Hajrudinović *et al*, 2015b; Robertson *et al*, 2004b; Robertson *et al*, 2010).

Based on our evidence, *S. torminalis* must have historically occurred along the north coastal areas of our study region in sympatry with *S. subcuneata* and *S. admonitor*, although there are no records of it having done so.It currently co-occurs with *S. devoniensis* at a number of sites, which may indicate a similar ecology; *S. devoniensis* is found on a wider range of geologies and soil types than the other study polyploid species (Rich *et al*, 2010).

The sexually reproducing *S. aria* and *S. torminalis* are clearly differentiated from each other and from the polyploid taxa in all our cluster analyses. The triploid form of *S. torminalis* clusters with its diploid forms (Figs. 3), suggesting its origin is due to intraspecific rather than interspecific hybridisation. It is thought that such cryptic autopolyploids, often formed via the fusion of unreduced gametes (Ramsey and Schemske, 1998) are a more common and important component of plant diversity than historic views suggest (Soltis *et al*, 2007; Barker *et al*, 2015). Indeed, Pellicer *et al* (2012) identified a number of polyploid *S. aria* samples. However, the triploid *S. torminalis* was found close to tetraploid *S. devoniensis* (Hamston et al, 2015)*,* so the fusion of gametes from diploid and polyploid *Sorbus* cannot be ruled out. Indeed, wide-scale screening of *Sorbus* seed ploidy showed this to be the most likely origin of polyploid seed embryos from *S. aria* occurring in the Balkan peninsula (Hajrudinović *et al*, 2015b).

2.4.2 Breeding system & genotypic variation in *Sorbus*

The patterns of genetic diversity within and among our study taxa are a consequence of breeding system and mutational load. Our results demonstrate that the polyploid *Sorbus* populations in our study are predominantly apomictic, in accordance with the findings from isoenzyme studies (Proctor *et al*, 1989). This is evident in the low levels of genetic variability within each polyploid species in contrast to the sexual diploids *S. aria* and *S. torminalis* which were highly variable, having a unique genotype for each sample (Table 2.1) as would be expected for self-incompatible, outcrossing species. Sexual reproduction is likely to be a rare event within the polyploid taxa as we were unable to find evidence of it within the individuals we sampled. However, due to the few allelic combinations of gametes produced among plants of the same clone, recombination will only be revealed if certain alleles disappear from the offspring which will, therefore, only represent a portion of those produced via sexual reproduction. We also sampled established trees, representatives of viable seeds and successful seedlings. It may be that the apomictic clones sampled are those best adapted to their environment with other genetic combinations less viable.

Each clonal lineage has arisen from a single hybridisation event rather than multiple origins as seen elsewhere (e.g. Arran, Scotland; Robertson *et al* (2004b) and, with the exception of vex2, represents a delineated polyploid species. If we accept that the principal route of polyploid formation in *Sorbus* is hybridisation involving a diploid parental species and a facultative apomict (Robertson *et al*, 2004a; Robertson *et al*, 2010; Hajrudinović *et al*, 2015b), the rate of novel polyploid formation will depend on the abundance and relative

distributions of the parental taxa and to what degree apomixis is facultative. The sexual diploids, *S. torminalis* and *S. aria,* currently rarely co-occur with any of our endemic study polyploid taxa so opportunities for hybridisation between diploid and polyploid taxa would be rare, although this may not always have been the case.

2.4.3 Source of genetic variability within apomicts

Our study has revealed genetic variability in each apomictic polyploid. Although the levels of variation were low, this study nonetheless revealed groups of 'clone mates' associated with particular sites. In the absence of recombination events, mutation plays a key role in the generation of genetic variation in apomictic lineages (Paun *et al*, 2006; Majesky *et al*, 2012). Polymorphisms at a number of loci were sufficient to identify divergent 'clone mates' within some apomictic species. *Sorbus subcuneata, S. rupicola, S. porrigentiformis, S. margaretae* and *S. vexans* all showed small numbers of site-associated mutations. One of these mutational variants (*S. subcuneata* from Greenaleigh, near Minehead, Somerset; Fig.2.2) had been identified previously as having some variation in leaf morphology compared to those at other sites (T.C.G. Rich, pers. comm.). If so, this could suggest a greater level of phenotypic variation than that detected with our microsatellite loci. The small site of Neck Wood was associated with specific clonal variants for *S. margaretae* and *S. rupicola* (at loci SA06 and SA09)*.* Wider interpretation of spatial patterns evident using the SA06 locus should be cautioned against since there is high likelihood of allele size homoplasy due to combinations of expansion and contraction in different lineages, a feature linked to high mutation rates particularly of dinucleotide repeats (Schlötterer *et al*, 1998; Ellegren, 2000a; Vigouroux *et al*,

2002). However, the use of more variable markers may reveal spatial patterns that relate to possible colonisation routes, a potentially interesting line of investigation. The variation seen previously in isozyme banding patterns in *S. margaretae* at the western end of its distribution (Proctor *et al*, 1989) could correspond with some of the site-specific mutational variation associated with *S. margaretae* at Neck Wood (the most western location for this species) or indeed be the vex2 variant, also of Neck Wood.

Members of subgenus *Tormaria* show little mutational variation compared to *Aria.* They may be of more recent origin than the members of subgenus *Aria*, particularly *S. admonitor* which has the most restricted distribution of all our study species.

2.4.4 Conservation

Strategies developed for the conservation of polyploid complexes that contain threatened species need to encompass any local adaptation of particular groups, together with the long-term ability of the complex to evolve through natural selection in a changing environment. Our results show discrete species with close evolutionary relationships derived from hybridisation and mixed mating systems which should be accounted for when devising conservation plans to optimise future diversification. There should be some assessment of the status of progenitor species, however common, to ensure they are protected from detrimental human activities and conservation measures should also be targeted at high diversity sites containing many constituents of species complexes.

Acknowledgments

References

Arnaud-Haond S, Alberto F, Teixeira S, Procaccini G, Serrão EA, Duarte CM (2005). Assessing genetic diversity in clonal organisms: low diversity or low resolution? Combining power and cost efficiency in selecting markers. *Journal of Heredity* **96**(4)**:** 434-440.

Arnaud-Haond S, Duarte CM, Alberto F, Serrão EA (2007). Standardizing methods to address clonality in population studies. *Molecular Ecology* **16**(24)**:** 5115-5139.

Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA (2015). On the relative abundance of autopolyploids and allopolyploids. *New Phytologist*.

Bennett MD, Smith JB (1991). Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society B: Biological Sciences* **334**(1271)**:** 309-345.

Bruvo R, Michiels NK, D'Souza TG, Schulenburg H (2004). A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* **13**(7)**:** 2101-2106.

Chase MW, Hills HH (1991). Silica Gel: An Ideal Material for Field Preservation of Leaf Samples for DNA Studies. Taxon 40(2): 215-220.

Chester M, Cowan RS, Fay MF, Rich TCG (2007). Parentage of endemic *Sorbus* L. (Rosaceae) species in the British Isles: evidence from plastid DNA. *Botanical Journal of the Linnean Society* **154**(3)**:** 291-304.

Clark LV, Jasieniuk M (2011). POLYSAT: an R package for polyploid microsatellite analysis. *Molecular Ecology Resources* **11**(3)**:** 562-566.

Dickinson TA, Lo EYY, Talent N, Love RM (2008). Black-fruited hawthorns of western North America - one or more agamic complexes? *Botany-Botanique* **86**(8)**:** 846-865.

Douhovnikoff V, Dodd RS (2003). Intra-clonal variation and a similarity threshold for identification of clones: application to *Salix exigua* using AFLP molecular markers. *Theoretical and Applied Genetics* **106**(7)**:** 1307-1315.

Dufresne F, Stift M, Vergilino R, Mable BK (2014). Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Molecular Ecology* **23**(1)**:** 40-69.

Earl DA, vonHoldt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**(2)**:** 359-361.

Ellegren H (2000a). Heterogeneous mutation processes in human microsatellite DNA sequences. *Nature genetics* **24**(4)**:** 400-402.

Ellegren H (2000b). Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics* **16**(12)**:** 551-558.

Ennos R, French G, Hollingsworth P (2005). Conserving taxonomic complexity. *Trends in Ecology & Evolution* **20**(4)**:** 164-168.

Ennos RA, Whitlock RAJ, Fay MF, Jones B, Neaves LE, Payne R *et al* (2012). Process-Based Species Action Plans: an approach to conserve contemporary evolutionary processes that sustain diversity in taxonomically complex groups. *Botanical Journal of the Linnean Society* **168**(2)**:** 194-203.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**(8)**:** 2611-2620.

Freeland JR, Kirk H, Petersen S (2011). *Molecular Markers in Ecology*, 2nd edn. Wiley-Blackwell: Oxford.

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998). Simple sequence repeats for the genetic analysis of apple. *Theoretical and Applied Genetics* **96**(8)**:** 1069-1076.

González-González EA, González-Pérez MA, Rivero E, Sosa PA (2010). Isolation and characterization of microsatellite loci in *Sorbus aria* (Rosaceae). *Conservation Genetics Resources* **2**(S1)**:** 341-343.

Grant V (1981). *Plant speciation*, 2nd edn. Columbia University Press New York.

Hajrudinović A, Frajman B, Schönswetter P, Silajdžić E, Siljak-Yakovlev S, Bogunić F (2015a). Towards a better understanding of polyploid *Sorbus* (Rosaceae) from Bosnia and Herzegovina (Balkan Peninsula), including description of a novel, tetraploid apomictic species. *Botanical Journal of the Linnean Society*.

Hajrudinović A, Siljak-Yakovlev S, Brown SC, Pustahija F, Bourge M, Ballian D *et al* (2015b). When sexual meets apomict: genome size, ploidy level and reproductive mode variation of *Sorbus aria s.l.* and *S. austriaca* (Rosaceae) in Bosnia and Herzegovina. *Annals of Botany* **116**(2)**:** 301-312.

Hamilton PB, Nicol E, De-Bastos ESR, Williams RJ, Sumpter JP, Jobling S *et al* (2014). Populations of a cyprinid fish are self-sustaining despite widespread feminization of males. *BMC Biology* **12**(1)**:** 1-13.

Hamston TJ, Pellicer J, Fay MF (2015). Polyploid wild service tree: first record of a triploid *Sorbus torminalis* (Rosaceae) in Britain. *New Journal of Botany* **5**(1)**:** 34-36.

Hendry AP (2009). Evolutionary biology: Speciation. *Nature* **458**(7235)**:** 162-164.

Houston L, Robertson A, Jones K, Smith SCC, Hiscock S, Rich TCG (2009). An account of the Whitebeams (*Sorbus* L., Rosaceae) of Cheddar Gorge, England, with description of three new species. *Watsonia* **27**(4)**:** 283.

Hughes J, Richards AJ (1988). The genetic structure of populations of sexual and asexual *Taraxacum* (dandelions). *Heredity* **60**(2)**:** 161-171.

Huson DH, Bryant D (2006). Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution* **23**(2)**:** 254-267.

IUCN. 2015. The IUCN Red List of Threatened Species. [ONLINE] Available at: http://www.iucnredlist.org/. [Accessed 12 February 2016].

Lepší M (2009). *Sorbus portae-bohemicae* and *Sorbus albensis*, two new endemic apomictic species recognized based on a revision of *Sorbus bohemica*. *Preslia* **81:** 63-89.

Lepší M, Ví t P, Lepší P, Boublík K, Suda J (2008). *Sorbus milensis*, a new hybridogenous species from northwestern Bohemia. *Preslia* **80**(2)**:** 229-244.

Liljefors A (1953). Studies on propagation, embryology, and pollination in *Sorbus. Acta Horti Bergiani* **16:** 277-329.

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

Ludwig S, Robertson A, Rich TCG, Djordjevic M, Cerovic R, Houston L *et al* (2013). Breeding systems, hybridization and continuing evolution in Avon Gorge *Sorbus. Annals of Botany*.

Majesky L, Vasut RJ, Kitner M, Travnicek B (2012). The pattern of genetic variability in apomictic clones of *Taraxacum officinale* indicates the alternation of asexual and sexual histories of apomicts. *PLoS One* **7**(8)**:** e41868.

Mallet J (2007). Hybrid speciation. *Nature* **446**(7133)**:** 279-283.

Meirmans PG, Van Tienderen PH (2004). GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* **4**(4)**:** 792-794.

Nelson-Jones E, Briggs D, Smith A (2002). The origin of intermediate species of the genus *Sorbus*. *Theoretical and Applied Genetics* **105**(6-7)**:** 953-963.

Newton A (1980). Progress in British *Rubus* studies. *Watsonia* **13**(1)**:** 35-40.

Oddou-Muratorio S, Aligon C, Decroocq S, Plomion C, Lamant T, Mush-Demesure B (2001). Microsatellite primers for *Sorbus torminalis* and related species. *Molecular Ecology Notes* **1**(4)**:** 297-299.

Paun O, Greilhuber J, Temsch EM, Hörandl E (2006). Patterns, sources and ecological implications of clonal diversity in apomictic *Ranunculus carpaticola* (*Ranunculus auricomus* complex, Ranunculaceae). *Molecular Ecology* **15**(4)**:** 897-910.

Pellicer J, Clermont S, Houston L, Rich TCG, Fay MF (2012). Cytotype diversity in the *Sorbus* complex (Rosaceae) in Britain: sorting out the puzzle. *Annals of Botany* **110**(6)**:** 1185-1193.

Proctor MCF, Proctor ME, Groenhof AC (1989). Evidence from peroxidase polymorphism on the taxonomy and reproduction of some *Sorbus* populations in south-west England. *New Phytologist* **112**(4)**:** 569-575.

Pruett CL, Winker K (2008). The effects of sample size on population genetic diversity estimates in song sparrows *Melospiza melodia*. *Journal of Avian Biology* **39**(2)**:** 252-256.

R Development Core Team. (2015). R Foundation for Statistical Computing: Vienna, Austria.

Ramasamy RK, Ramasamy S, Bindroo BB, Naik VG (2014). STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus* **3**(1)**:** 431.

Ramsey J, Schemske DW (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* **29:** 467-501.

Rich TCG, Houston L, Robertson A, Proctor MCF (2010). *Whitebeams, rowans and service trees of Britain and Ireland. A monograph of British and Irish Sorbus L. B.S.B.I. Handbook No. 14.* Botanical Society of the British Isles: London.

Richards AJ (2003). Apomixis in flowering plants: an overview. *Philosophical Transactions of the Royal Society B: Biological Sciences* **358**(1434)**:** 1085-1093.

Robertson A, Newton AC, Ennos RA (2004a). Breeding systems and continuing evolution in the endemic *Sorbus* taxa on Arran. *Heredity* **93**(5)**:** 487-495.

Robertson A, Newton AC, Ennos RA (2004b). Multiple hybrid origins, genetic diversity and population genetic structure of two endemic *Sorbus* taxa on the Isle of Arran, Scotland. *Molecular Ecology* **13**(1)**:** 123-134.

Robertson A, Rich TCG, Allen AM, Houston L, Roberts CAT, Bridle JR *et al* (2010). Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus. Molecular Ecology* **19**(8)**:** 1675-1690.

Schlötterer C, Ritter R, Harr B, Brem G (1998). High mutation rate of a long microsatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates. *Molecular Biology and Evolution* **15**(10)**:** 1269-1274.

Sell PD (1989). The *Sorbus latifolia* (Lam.) Pers. aggregate in the British Isles. *Watsonia* **17**(4)**:** 385-399.

Siljak-Yakovlev S, Pustahija, F., Šolić, E. M., Bogunić, F., Muratović, E., Bašić, N., Catrice, O. & Brown, S. C. (2010). Towards a Genome Size and Chromosome Number Database of Balkan Flora: C-Values in 343 Taxa with Novel Values for 242. *Advanced Science Letters* **3**(2)**:** 190-213.

Soltis DE, Buggs RJA, Doyle JJ, Soltis PS (2010). What we still don't know about polyploidy. *Taxon* **59**(5)**:** 1387-1403.

Soltis DE, Soltis PS, Schemske DW, Hancock JF, Thompson JN, Husband BC *et al* (2007). Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon***:** 13-30.

Vigouroux Y, Jaqueth JS, Matsuoka Y, Smith OS, Beavis WD, Smith JSC *et al* (2002). Rate and Pattern of Mutation at Microsatellite Loci in Maize. *Molecular Biology and Evolution* **19**(8)**:** 1251-1260.

Warburg EF (1962). *Flora of the British Isles*, 2nd edn. Cambridge University Press: Cambridge.

Weiner J. (2015). Three dimensional PCA plots. R package version 0.8. Accessed April 26, 2015. https://CRAN.R-project.org/package=pca3d.

Whitton J, Sears CJ, Baack EJ, Otto SP (2008). The dynamic nature of apomixis in the angiosperms. *International Journal of Plant Sciences* **169**(1)**:** 169-182.

Chapter 2: Supplementary Information

Table S2.1 Site location and herbarium voucher accession numbers for DNA samples.

Sample no:	Accession no:	Taxon	Lat.	Long.	Site	Site code
5	V.2014.003.143	*S. admonitor*	51.22675	-3.79965	Watersmeet	WM
6	V.2014.003.142	*S. admonitor*	51.22501	-3.79907	Watersmeet	WM
56	V.2014.003.140	*S. admonitor*	51.22467	-3.80003	Watersmeet	WM
57	V.2014.003.141	*S. admonitor*	51.22494	-3.80024	Watersmeet	WM
161	V.2014.003.171	*S. admonitor*	51.22509	-3.80029	Watersmeet	WM
162	V.2014.003.181	*S. admonitor*	51.22474	-3.80033	Watersmeet	WM
163	V.2014.003.168	*S. admonitor*	51.22469	-3.80019	Watersmeet	WM
164	V.2014.003.182	*S. admonitor*	51.22443	-3.80045	Watersmeet	WM
168	V.2014.003.166	*S. admonitor*	51.22477	-3.80015	Watersmeet	WM
169	V.2014.003.167	*S. admonitor*	51.22459	-3.80014	Watersmeet	WM
170	V.2014.003.169	*S. admonitor*	51.22661	-3.79892	Watersmeet	WM
171	V.2014.003.170	*S. admonitor*	51.22456	-3.79831	Watersmeet	WM
174		*S. admonitor*	51.22451	-3.79974	Watersmeet	WM
178	V.2014.003.162	*S. admonitor*	51.22301	-3.79608	Watersmeet	WM
180	V.2014.003.164	*S. admonitor*	51.22436	-3.79841	Watersmeet	WM
183	V.2014.003.165	*S. admonitor*	51.22504	-3.80014	Watersmeet	WM
214	V.2014.003.180	*S. admonitor*	51.22669	-3.79968	Watersmeet	WM
234		*S. admonitor*	51.22681	-3.79968	Watersmeet	WM
553	V.2014.003.163	*S. admonitor*	51.22451	-3.80147	Watersmeet	WM
191		*S. aria*	51.46904	-2.62975	Avon Gorge	AG
192		*S. aria*	51.46867	-2.63051	Avon Gorge	AG

Sample no:	Accession no:	Taxon	Lat.	Long.	Site	Site code
193		*S. aria*	51.46867	-2.63051	Avon Gorge	AG
194		*S. aria*	51.46955	-2.63484	Avon Gorge	AG
215	V.2014.003.175	*S. aria*	51.2879	-2.74631	Cheddar gorge	CG
216	V.2014.003.176	*S. aria*	51.2879	-2.75157	Cheddar gorge	CG
217	V.2014.003.173	*S. aria*	51.28788	-2.75276	Cheddar gorge	CG
219	V.2014.003.179	*S. aria*	51.28812	-2.75229	Cheddar gorge	CG
220	V.2014.003.174	*S. aria*	51.28674	-2.75364	Cheddar gorge	CG
222	V.2014.003.177	*S. aria*	51.2867	-2.75541	Cheddar gorge	CG
224	V.2014.003.172	*S. aria*	51.28748	-2.75555	Cheddar gorge	CG
36	V.2014.003.178	*S. aria*	51.45552	-2.63963	Leigh woods	LW
44		*S. aria*	51.4661	-2.63564	Leigh woods	LW
105	V.2014.003.096	*S. devoniensis*	50.9119	-4.07305	Beaford	
246	V.2014.003.016	*S. devoniensis*	50.8323	-4.13558	Highampton	
103	V.2014.003.093	*S. devoniensis*	50.88709	-3.98226	Hollocombe	
78	V.2014.003.014	*S. devoniensis*	50.77751	-4.02974	Inwardleigh	
79	V.2014.003.011	*S. devoniensis*	50.77758	-4.02977	Inwardleigh	
80	V.2014.003.008	*S. devoniensis*	50.77732	-4.02965	Inwardleigh	
51	V.2014.003.017	*S. devoniensis*	50.56476	-3.52589	Little Haldon	LH
52	V.2014.003.018	*S. devoniensis*	50.56479	-3.52554	Little Haldon	LH
89	V.2014.003.006	*S. devoniensis*	50.56573	-3.5364	Little Haldon	LH
90	V.2014.003.003	*S. devoniensis*	50.56619	-3.53573	Little Haldon	LH
91	V.2014.003.007	*S. devoniensis*	50.56491	-3.53645	Little Haldon	LH
92	V.2014.003.004	*S. devoniensis*	50.56546	-3.534	Little Haldon	LH
93	V.2014.003.005	*S. devoniensis*	50.56544	-3.53406	Little Haldon	LH
184	V.2014.003.084	*S. devoniensis*	50.11504	-3.52031	Little Haldon	LH
228	V.2014.003.083	*S. devoniensis*			Little Haldon	LH

Sample no:	Accession no:	Taxon	Lat.	Long.	Site	Site code
229	V.2014.003.086	*S. devoniensis*	50.56541	-3.5341	Little Haldon	LH
230		*S. devoniensis*			Little Haldon	LH
231		*S. devoniensis*			Little Haldon	LH
110	V.2014.003.085	*S. devoniensis*	50.92121	-4.1573	Little Torrington	
516		*S. devoniensis*			Luckbarrow ENHS collection	
517		*S. devoniensis*			Luckbarrow ENHS collection	
19	V.2014.003.073	*S. devoniensis*	51.21864	-3.95849	Neck wood	NW
70	V.2014.003.009	*S. devoniensis*	50.73087	-3.8987	South Tawton	
71	V.2014.003.012	*S. devoniensis*	50.72905	-3.90449	South Tawton	
72	V.2014.003.010	*S. devoniensis*	50.73076	-3.89058	South Tawton	
73	V.2014.003.015	*S. devoniensis*	50.73155	-3.89127	South Tawton	
118	V.2014.003.094	*S. devoniensis*	51.00979	-4.20639	Upcott	
113	V.2014.003.089	*S. devoniensis*	50.93744	-4.18177	Watergate Bridge	
10	V.2014.003.070	*S. devoniensis*	51.22129	-3.89988	Woody Bay	WB
11	V.2014.003.071	*S. devoniensis*	51.22084	-3.89829	Woody Bay	WB
15	V.2014.003.072	*S. devoniensis*	51.22832	-3.90799	Woody Bay	WB
101		*S. margaretae*	51.22463	-3.66655	Culbone	CB
102		*S. margaretae*	51.22416	-3.66752	Culbone	CB
197		*S. margaretae*	51.22473	-3.66602	Culbone	CB
198		*S. margaretae*	51.22471	-3.66685	Culbone	CB
204		*S. margaretae*	51.22501	-3.66679	Culbone	CB
221		*S. margaretae*	51.22450	-3.667753	Culbone	CB
26	V.2014.003.066	*S. margaretae*	51.2337	-3.7434	Desolation	DL
27	V.2014.003.065	*S. margaretae*	51.23388	-3.74425	Desolation	DL

Sample no:	Accession no:	Taxon	Lat.	Long.	Site	Site code
29	V.2014.003.082	*S. margaretae*	51.23509	-3.75604	Desolation	DL
30		*S. margaretae*	51.23509	-3.75604	Desolation	DL
31	V.2014.003.126	*S. margaretae*	51.23509	-3.75604	Desolation	DL
35	V.2014.003.067	*S. margaretae*	51.22761	-3.73556	Desolation	DL
54		*S. margaretae*	51.23509	-3.75604	Desolation	DL
122		*S. margaretae*	51.23516	-3.75698	Desolation	DL
123		*S. margaretae*	51.23511	-3.7567	Desolation	DL
124		*S. margaretae*	51.23547	-3.75818	Desolation	DL
125		*S. margaretae*	51.23522	-3.75512	Desolation	DL
126		*S. margaretae*	51.2353	-3.75522	Desolation	DL
127		*S. margaretae*	51.2352	-3.7553	Desolation	DL
128		*S. margaretae*	51.23521	-3.75532	Desolation	DL
129		*S. margaretae*	51.23544	-3.75565	Desolation	DL
130		*S. margaretae*	51.2352	-3.75586	Desolation	DL
131		*S. margaretae*	51.23509	-3.75604	Desolation	DL
132		*S. margaretae*	51.23509	-3.75604	Desolation	DL
137		*S. margaretae*	51.22913	-3.6964	Embelle woods	CB
18	V.2014.003.146	*S. margaretae*	51.21869	-3.95807	Neck wood	NW
24	V.2014.003.068	*S. margaretae*	51.2193	-3.95883	Neck wood	NW
25	V.2014.003.069	*S. margaretae*	51.21927	-3.95887	Neck wood	NW
3	V.2014.003.062	*S. margaretae*	51.22528	-3.79756	Watersmeet	WM
53		*S. porrigentiformis*	50.48037	-3.51315	Babbacombe slopes	TB
86	V.2014.003.152	*S. porrigentiformis*	50.47946	-3.51364	Babbacombe slopes, Torbay	TB
88	V.2014.003.150	*S. porrigentiformis*	50.47946	-3.51364	Babbacombe slopes, Torbay	TB

Sample no:	Accession no:	Taxon	Lat.	Long.	Site	Site code
218	V.2014.003.151	*S. porrigentiformis*	51.28799	-2.7524	Cheddar gorge	CG
223	V.2014.003.158	*S. porrigentiformis*	51.28732	-2.75596	Cheddar gorge	CG
7	V.2014.003.159	*S. porrigentiformis*	51.22465	-3.81605	Fishermans car park	WM
47	V.2014.003.161	*S. porrigentiformis*	51.46611	-2.63517	Leigh woods	LW
533		*S. porrigentiformis*			Luckbarrow ENHS (Torbay)	
67	V.2014.003.153	*S. porrigentiformis*	50.4742	-3.50361	Redgate, Torbay	TB
188		*S. porrigentiformis*	50.47497	-3.50271	Redgate, Torbay	TB
81	V.2014.003.157	*S. porrigentiformis*	50.47382	-3.50242	Redgate, Torbay	TB
82	V.2014.003.155	*S. porrigentiformis*	50.47371	-3.50248	Redgate, Torbay	TB
83	V.2014.003.156	*S. porrigentiformis*	50.47336	-3.50251	Redgate, Torbay	TB
84	V.2014.003.183	*S. porrigentiformis*	50.47152	-3.50211	Redgate, Torbay	TB
41	V.2014.003.160	*S. porrigentiformis*	51.45721	-2.63513	Stokeleigh Camp	AG
68	V.2014.003.154	*S. porrigentiformis*	50.47691	-3.50228	Walls Hill, Torbay	TB
14		*S. porrigentiformis*	51.22802	-3.90778	Woody Bay	WB
910		*S. porrigentiformis*	51.78454	-3.42401	Darren Fach	PDF
B9-F10		*S. rupicola*	51.78454	-3.42401	Darren Fach	PDF
B9-910		*S. rupicola*	51.77683	-3.42755	Penmoelallt	PDF
85	V.2014.003.050	*S. rupicola*	50.47946	-3.51364	Babbacombe slopes, Torbay	TB
87	V.2014.003.055	*S. rupicola*	50.48036	-3.51367	Babbacombe slopes, Torbay	TB
62	V.2014.003.019	*S. rupicola*	50.40265	-3.52442	Churston, Torbay	TB
63	V.2014.003.144	*S. rupicola*	50.4013	-3.52397	Churston, Torbay	TB
64	V.2014.003.149	*S. rupicola*	50.40164	-3.52092	Churston, Torbay	TB
65	V.2014.003.051	*S. rupicola*	50.40312	-3.52474	Churston, Torbay	TB

Sample no:	Accession no:	Taxon	Lat.	Long.	Site	Site code
66	V.2014.003.147	*S. rupicola*	50.40228	-3.52467	Churston, Torbay	TB
172	V.2014.003.052	*S. rupicola*	50.40248	-3.52423	Churston, Torbay	TB
173	V.2014.003.053	*S. rupicola*	50.40349	-3.52598	Churston, Torbay	TB
267		*S. rupicola*	57.03369	-4.19312	Creagh Dhubh, Scotland	
121		*S. rupicola*			Luckbarrow ENHS collection (Neck Wood)	
22	V.2014.003.145	*S. rupicola*	51.21917	-3.95852	Neck wood	NW
265		*S. subcuneata*	51.2205	-3.49021	Greencliff, Minehead	GC
266		*S. subcuneata*	51.2205	-3.49021	Greencliff, Minehead	GC
33	V.2014.003.022	*S. subcuneata*	51.21227	-3.47574	Greenleigh Wood, Minehead	GL
34	V.2014.003.122	*S. subcuneata*	51.21227	-3.47574	Greenleigh Wood, Minehead	GL
23	V.2014.003.123	*S. subcuneata*	51.21917	-3.95874	Neck wood	NW
2	V.2014.003.121	*S. subcuneata*	51.22534	-3.79637	Watersmeet	WM
4	V.2014.003.077	*S. subcuneata*	51.22642	-3.80044	Watersmeet	WM
58	V.2014.003.078	*S. subcuneata*	51.22539	-3.80008	Watersmeet	WM
59	V.2014.003.120	*S. subcuneata*	51.22575	-3.80083	Watersmeet	WM
144	V.2014.003.125	*S. subcuneata*	51.22343	-3.79845	Watersmeet	WM
145	V.2014.003.103	*S. subcuneata*	51.22436	-3.79837	Watersmeet	WM
146	V.2014.003.107	*S. subcuneata*	51.2242	-3.79833	Watersmeet	WM
147	V.2014.003.102	*S. subcuneata*	51.22439	-3.79816	Watersmeet	WM
157	V.2014.003.114	*S. subcuneata*	51.22379	-3.79796	Watersmeet	WM
165	V.2014.003.111	*S. subcuneata*	51.22469	-3.79219	Watersmeet	WM
166	V.2014.003.109	*S. subcuneata*	51.22556	-3.80064	Watersmeet	WM

Sample no:	Accession no:	Taxon	Lat.	Long.	Site	Site code
167	V.2014.003.112	S. subcuneata	51.22566	-3.80075	Watersmeet	WM
181	V.2014.003.075	S. subcuneata	51.22468	-3.79828	Watersmeet	WM
182	V.2014.003.056	S. subcuneata	51.22459	-3.79833	Watersmeet	WM
186	V.2014.003.117	S. subcuneata	51.22644	-3.80035	Watersmeet	WM
187	V.2014.003.116	S. subcuneata	51.22533	-3.79862	Watersmeet	WM
212		S. subcuneata	51.22394	-3.79374	Watersmeet	WM
226		S. subcuneata	51.22453	-3.79459	Watersmeet	WM
269		S. subcuneata	51.22365	-3.79811	Watersmeet	WM
280		S. subcuneata	51.22493	-3.79729	Watersmeet	WM
235	V.2014.003.115	S. subcuneata	51.22144	-3.8989	Woody Bay	WB
237	V.2014.003.113	S. subcuneata	51.22839	-3.90739	Woody Bay	WB
106	V.2014.003.037	S. torminalis	50.91148	-4.0721	Beaford	
251		S. torminalis	50.716897	-4.408505	Beardon	
261		S. torminalis	50.707611	-3.716723	Berryhead plantation	
273		S. torminalis	50.52844	-3.63954	Broadridge wood, Newton Abbot	NA
256		S. torminalis	50.876113	-4.5402991	Coombe Valley, nr stibb	
250		S. torminalis	50.503475	-4.2524213	Halton Barton	
243	V.2014.003.046	S. torminalis	50.83352	-4.10751	Hatherleigh	
242		S. torminalis	51.64024	-4.1388192	Hatherleigh	
245		S. torminalis	50.830912	-4.1418027	Highampton	
247		S. torminalis	50.832512	-4.1349456	Highampton	
255		S. torminalis	50.869228	-4.5317944	Houndapitt, Nr Stibb	
42	V.2014.003.135	S. torminalis	51.46256	-2.63963	Leigh woods	LW
232		S. torminalis	50.562964	-3.533977	Little Haldon	LH

Sample no:	Accession no:	Taxon	Lat.	Long.	Site	Site code
241		*S. torminalis*	50.787998	-3.817363	Little Langford	
95		*S. torminalis*			Luckbarrow ENHS collection	
94		*S. torminalis*	51.203371	-3.5844902	Luckbarrow, ENHS	
252	V.2014.003.048	*S. torminalis*	50.76732	-4.48058	Odd Mill	
254	V.2014.003.049	*S. torminalis*	50.76711	-4.47862	Odd Mill	
249	V.2014.003.127	*S. torminalis*	50.44776	-4.30243	Pillaton Mill	
260		*S. torminalis*	50.666518	-3.7051913	Plaston Green	
69	V.2014.003.039	*S. torminalis* (3n)	50.7294	-3.9037	South Tawton	
74	V.2014.003.042	*S. torminalis*	50.7547	-3.83881	Spreyton	
75		*S. torminalis*	50.758138	-3.8243868	Spreyton	
76		*S. torminalis*	50.761542	-3.8136719	Spreyton	
77		*S. torminalis*	50.769495	-3.8169577	Spreyton	
259		*S. torminalis*	50.764303	-3.8400693	Spreyton	
49		*S. torminalis*	51.841978	-2.6377618	Symonds Yat	SY
257	V.2014.003.138	*S. torminalis*	50.85228	-4.52043	Tiscott- nr Stibb	
117		*S. torminalis*	51.008554	-4.2065583	Upcott	
116		*S. torminalis*	51.008409	-4.2066227	Upcott	
114	V.2014.003.034	*S. torminalis*	51.00523	-4.20671	Upcott Wood	
115	V.2014.003.036	*S. torminalis*	51.0052	-4.20681	Upcott Wood	
96		*S. vexans*	51.22402	-3.66306	Culbone	CB
97		*S. vexans*	51.22459	-3.66623	Culbone	CB
98		*S. vexans*	51.22471	-3.66642	Culbone	CB
99		*S. vexans*	51.2247	-3.6665	Culbone	CB
100		*S. vexans*	51.22471	-3.66642	Culbone	CB
199		*S. vexans*	51.22471	-3.66685	Culbone	CB

Sample no:	Accession no:	Taxon	Lat.	Long.	Site	Site code
200		*S. vexans*	51.22434	-3.66681	Culbone	CB
201		*S. vexans*	51.22442	-3.6667	Culbone	CB
202		*S. vexans*	51.22442	-3.6667	Culbone	CB
205		*S. vexans*	51.22501	-3.66645	Culbone	CB
32	V.2014.003.148	*S. vexans*	51.23509	-3.75604	Desolation	DL
28	V.2014.003.057	*S. vexans*	51.23385	-3.75227	Dogsworthy Combe, nr Desolation	DL
8	V.2014.003.064	*S. vexans*	51.22456	-3.81578	Fishermans car park	WM
20	V.2014.003.063	*S. vexans*	51.21893	-3.95847	Neck wood	NW
21	V.2014.003.059	*S. vexans*	51.21893	-3.95847	Neck wood	NW
60	V.2014.003.001	*S. vexans*	51.22595	-3.82097	Oxen tor	OT
61	V.2014.003.025	*S. vexans*	51.226	-3.82156	Oxen tor	OT
133	V.2014.003.023	*S. vexans*	51.22615	-3.82165	Oxen tor	OT
134	V.2014.003.028	*S. vexans*	51.2291	-3.83256	Oxen tor	OT
135	V.2014.003.024	*S. vexans*	51.22613	-3.82128	Oxen tor	OT
136	V.2014.003.026	*S. vexans*	51.22609	-3.82127	Oxen tor	OT
138	V.2014.003.029	*S. vexans*	51.22611	-3.82171	Oxen tor	OT
139	V.2014.003.027	*S. vexans*	51.22611	-3.82164	Oxen tor	OT
9	V.2014.003.060	*S. vexans*	51.22148	-3.89927	Woody Bay	WB

Table S2.2. Nucleotide sequences of nuclear microsatellite primers used in this study. [1]Microsatellite primers were redesigned by Robertson *et al* (2010); [2]Microsatellite primers from *S. aria* (González-González *et al*, 2010); [3]Microsatellite primers from *S. torminalis* (Oddou-Muratorio *et al*, 2001); [4]Microsatellite primers derived from *Malus domestica* (Gianfranceschi *et al*, 1998); + SA19.1, CH01F09, SA02 and SA09 were excluded from analyses involving the diploid species. MS14 was only used in Chapter 3 for determination of inheritance patterns using primers developed for *Malus domestica* (Nelson-Jones *et al*, 2002).

Multiplex	Locus	Primer sequence (5' to 3')	Dye	Repeat
MPLX 1	CH01F02[1]	F - CCACATTAGAGCAGTTGAGGATGA	D4-PA	$(AG)_{22}$
		R - ATAGGGTAGCAGCAGATGGTTGT		
	SA01[2]	F - ATGGAGTTGAGCTCCACATC	D2-PA	$(GA)_{13}$
		R - GGTGGAGGGACAATTGTGTC		
	SA19.1+[2]	F - AAGTTTACAAGAGTGTGTTCAG	D3-PA	$(GA)_{24}$
		R - GAATTCATGAAAGCAGCTAATG		
	MSS5[3]	F - CCCCAACAACATTTTTCTCC	D2-PA	$(GA)_{19}$
		R - CCTCTCGCTCTTTGCCTCT		
	MSS16[1]	F - ATGTCACATCTCTCCCCTTGTGT	D3-PA	$(GA)_{28}$
		R - TTTTGCCCTCAAAGAATGCCTTA		
MPLX 2	CH01F09 +[1]	F - ATGTACATCAAAGTGTGGATTG	D3-PA	$(AG)_{22}$
		R - GGCGCTTTCCAACACATC		
	CH02D11[1]	F - AAATAAGCGTCCAGAGCAACAG	D4-PA	$(AG)_{21}$
		R - GGGACAAAATCTCACAAACAGA		
	SA03[2]	F - CACTTCTTCCTGCTGTTTGG	D2-PA	$(GA)_{12}$
		R - ACTACTGCTACTTCTGTGGG		
	SA06[2]	F - ATTTGATCCATGTGCGACTGCA	D4-PA	$(GA)_{32}$
		R - TGCAGCGGTTGCAGATTGCA		
	MSS13[4]	F - GAAAATTCCTTCCCGAACTTCAT	D3-PA	$(GA)_{12}$
		R - AACTCACTCGGATTTTGGAACCT		

Multiplex	Locus	Primer sequence (5' to 3')	Dye	Repeat
MPLX 3	SA02+[2]	F - CTAGGTATCATCTCCGACCA R - ACGTAGCACTGAATGGTATAG	D2-PA	$(GA)_{16}$
	SA08[2]	F - CAGAGAGAGTGCACTGCCT R - GAATTCTTGGCAGTTTGCCT	D3-PA	$(CT)_{16}$
	SA09+[2]	F - CTTGTTGGACGGATTTCTTC R - CCAATACTTGAGTAGCATAC	D3-PA	$(AG)_{17}$
	MS14	F - CGCTCACCATCGTAGACGT R - ATGCAATGGCTAAGCATA3	D4-PA	
Single	SA14[2]	F - ATGGATTTAGGTTAACAGTTGTC R - GAGGTAAAACCTACCAGTATAC	D4-PA	$(TC)_{30}$

Table S2.3. Multiplex design for PCR reaction of *Sorbus.* MS14 was only used for analysis of inheritance patterns in chapter 3.

Multiplex	Marker	Final concentration in PCR (µM)	Dilution for capillary electrophoresis
MPLX 1	CH01F02	0.06	20%
	SA01	1.25	
	SA19.1	1.25	
	MSS5	0.375	
	MSS16	0.125	
MPLX 2	CH01F09	0.375	0%
	CH02D11	0.125	
	SA03	0.25	
	SA06	0.125	
	MSS13	0.25	
MPLX 3	SA02	0.125	0%
	SA08	0.075	
	SA09	0.125	
	MS14	0.025	
SINGLE	SA14	0.25	80%
PCR 10 sample reaction mix: 50 µl HotStarTaq Master mix, 10 µl MPLX primer mix, 30µl water. 1 µl DNA sample template + 9 µl PCR reaction mix			**Capillary electrophoresis mix:** 25µl SLS + internal size standard, 5 µl PCR product at specified dilution.

Table S2.4 Multilocus genotypes for the polyploid study taxa at 14 loci. X = ploidy level, Fq. = number of individuals sampled, ▢ = rare allele sizes. Loci in **bold** are the ten loci used in the combined diploid and polyploid analysis.

Taxon	X	Fq.	**CH01F02**				SA19.1				**MSS16**				**SA01**			
											Alleles at each microsatellite loci							
S. admonitor	4	19	187	195	199		224	232			158	160	188	204	224	232	242	
S. devoniensis	4	30	187	195	199		224	232			158	160	198	204	224	232	234	242
S. devoniensis	4	1	187	195	199		224	232			158	160	198	204	224	232	234	242
S. subcuneata	3	21	187	195	199		224	232			158	160	204		224	232	242	
S. subcuneata	3	3	187	195	199		224	232			158	160	204		224	232	242	
S. subcuneata	3	2	187	195	199		224	232			158	160	202		224	232	242	
S. margaretae	4	23	191	195	199	221	216	224	232	250	158	160	162		224	232		
S. margaretae	4	1	191	195	199	221	216	224	232	250	158	160	162		224	232		
S. margaretae	4	2	191	195	199	221	216	224	232	250	158	160	162		224	232		
S. margaretae	4	1	191	195	199	221	216	224	232	250	158	160	162		224	232		
S. margaretae	4	2	191	195	199	221	216	224	232	250	158	160	162		224	232		
S. porrigentiformis	4	5	191	197	201	203	216	228	238		158	162	170		214	236	244	
S. porrigentiformis	4	9	191	197	201	203	216	228	238		158	162	170		214	236	244	
S. porrigentiformis	4	1	191	197	201	203	216	228	238		158	162	170		214	236	244	
S. porrigentiformis	4	1	191	197	201	203	216	228	238		158	162	170		214	236	244	
S. porrigentiformis	4	1	191	197	201	203	216	228	238		158	162	170		214	236	246	
S. porrigentiformis	4	1	191	197	201	203	216	228	238		158	162	170		214	236	244	
S. porrigentiformis	4	1	191	197	201	203	216	228	238		158	162	170		214	236	244	

			Alleles at each microsatellite loci														
Taxon	X	Fq.	**CH01F02**				SA19.1				**MSS16**			**SA01**			
S. rupicola	4	4	191	199	201	209	216	250			158	162		224	230	234	236
S. rupicola	4	1	191	199	201	209	216	250			158	162		224	230	234	236
S. rupicola	4	1	191	199	201	209	216	250			158	162		224	230	234	236
S. rupicola	4	5	191	199	201	209	216	250			158	162		224	230	234	236
S. rupicola	4	1	191	199	201	209	216	250			158	162		224	230	234	236
S. rupicola	4	1	191	199	201	209	216	250			158	162		224	230	234	236
S. vexans	4	1	191	195	201	203	216	224	236	250	158	160	162	236			
S. vexans	4	1	191	195	201	203	216	224	236	250	158	160	162	236			
S. vexans	4	1	191	195	201	203	216	224	236	250	158	160	162	236			
S. vexans	4	7	191	195	201	203	216	224	236	250	158	160	162	236			
S. vexans	4	2	191	195	201	203	216	224	236	250	158	160	162	236			
S. vexans	4	7	191	195	201	203	216	224	236	250	158	160	162	236			
S. vexans	4	1	191	195	201	203	216	224	236	250	158	160	162	236			
S. vexans	4	2	191	195	201	203	216	224	236	250	158	160	162	236			
vex2	4	2	191	195	203	221	224	234	236	256	158	160		232	236	238	

Table S2.4

			Alleles at each microsatellite loci														
Taxon	X	Fq.	**MSS5**				CH01F09				**CH02D11**			**MSS13**			
S. admonitor	4	19	119	121	123	127	113	123			152	162	182	187	193	195	
S. devoniensis	4	30	119	121	123	127	113	123			152	162	182	189	193	195	
S. devoniensis	4	1	119	121	123	127	113	123			152	162	182	189	193	195	
S. subcuneata	3	21	119	121	123		113	123			152	162	182	193	195		
S. subcuneata	3	3	119	121	123		113	123			152	162	182	193	195		
S. subcuneata	3	2	119	121	123		113	123			152	162	182	193	195		
S. margaretae	4	23	119	121	135		113	115	123		152	154	182	193	195	197	
S. margaretae	4	1	119	121	135		113	115	121	123	152	154	182	193	195	197	
S. margaretae	4	2	119	121	135		113	115	123		152	154	182	193	195	197	
S. margaretae	4	1	119	121	135		113	115	123		152	154	182	193	195	197	
S. margaretae	4	2	119	121	135		113	115	123		152	154	182	193	195	197	
S. porrigentiformis	4	5	115	127	131	137	115	123	125	129	152	198		193	195	197	203
S. porrigentiformis	4	9	115	127	131	137	115	123	125	129	152	198		193	195	197	203
S. porrigentiformis	4	1	115	127	131	137	115	123	125	129	152	198		193	195	197	203
S. porrigentiformis	4	1	115	125	127	131	115	123	125	129	152	198		193	195	197	203
S. porrigentiformis	4	1	115	127	131	137	115	123	125	129	152	198		193	195	197	203
S. porrigentiformis	4	1	115	127	131	137	115	123	125	129	152	198		193	195	197	203
S. porrigentiformis	4	1	115	127	131	137	115	123	125	129	152	198		193	195	197	203

Taxon	X	Fq.	MSS5				CH01F09			CH02D11			MSS13			
													Alleles at each microsatellite loci			
S. rupicola	4	4	119	127	131	173	113	115	121	152	162	168	193	195	197	
S. rupicola	4	1	119	127	131	173	113	115	121	152	162	168	193	195	197	
S. rupicola	4	1	119	127	131	173	113	115	121	152	162	170	193	195	197	
S. rupicola	4	5	119	127	131	173	113	115	121	152	162	168	193	195	197	
S. rupicola	4	1	119	127	131	173	113	115	121	152	162	168	193	195	197	
S. rupicola	4	1	119	127	131	173	113	115	121	152	162	168	193	195	197	
S. vexans	4	1	119	121	127		115	121		150	154		193	195	197	199
S. vexans	4	1	119	121	127		115	121		150	154		193	195	197	199
S. vexans	4	1	119	121	127		115	121		150	154		193	195	197	199
S. vexans	4	7	119	121	127		115	121		150	154		193	195	197	199
S. vexans	4	2	119	121	127		115	121		150	154		193	195	197	199
S. vexans	4	7	119	121	127		115	121		150	154		193	195	197	199
S. vexans	4	1	119	121	127		115	121		150	154		193	195	197	199
S. vexans	4	2	119	121	127		115	121		150	154		193	195	197	199
vex2	4	2	121	127	129	135	115	123		152	154	182	193	195	199	

Table S2.4

Alleles at each microsatellite loci

Taxon	X	Fq.	SA03			SA06				SA09				SA14			
S. admonitor	4	19	224			258	268			162	194			170	178	208	226
S. devoniensis	4	30	224			258	268			162	194			170	204	208	226
S. devoniensis	4	1	224			258	268			162	186			170	204	208	226
S. subcuneata	3	21	224			258	268			162	194			170	208	226	
S. subcuneata	3	3	224			258	268			162	194			170	212	226	
S. subcuneata	3	2	224			258	268			162	194			170	208	226	
S. margaretae	4	23	224			258	268	280	312	162	182	194		194	208	226	
S. margaretae	4	1	224			258	268	280	312	162	182	194		194	208	226	
S. margaretae	4	2	224			258	268	280	314	162	182	194		194	208	226	
S. margaretae	4	1	224			258	268	282	312	162	182	194		194	208	226	
S. margaretae	4	2	224			258	268	284	314	162	182	194		194	208	226	
S. porrigentiformis	4	5	228	240		256	264	270		174	176	184	186	196	222	224	226
S. porrigentiformis	4	9	228	240		256	264	270		174	176	184	186	196	222	224	
S. porrigentiformis	4	1	228	240		256	264	270		174	176		186	196	222	224	
S. porrigentiformis	4	1	228	240		256	264	270		174	176	184	186	196	222	224	
S. porrigentiformis	4	1	228	240		256	264	270		174	176	184	186	196	222	224	226
S. porrigentiformis	4	1	228	240	242	256	264	270		174	176	184	186	196	222	224	
S. porrigentiformis	4	1	228	240		256	264	270		174	176	184	186	196	222	224	226

			Alleles at each microsatellite loci												
Taxon	X	Fq.	**SA03**		**SA06**				SA09				**SA14**			
S. rupicola	4	4	224	240	258	264	280	312	162	164	176		194	206	208	
S. rupicola	4	1	224	240	258	264	280	300	162	164	176		194	206	208	
S. rupicola	4	1	224	240	258	264	280	300	162	164	176		194	206	208	
S. rupicola	4	5	224	240	258	264	282	300	162	164	178		194	206	208	
S. rupicola	4	1	224	240	258	264	284	300	162	164	178		194	206	208	
S. rupicola	4	1	224	240	258	264	282	302	162	164	178		194	206	208	
S. vexans	4	1	224	240	258	268	280	312	162	176	178	182		206	224	226
S. vexans	4	1	224	240	258	280	290	302	162	176	178	182		206	224	226
S. vexans	4	1	224	240	258	280	294		162	176	178	182		206	224	226
S. vexans	4	7	224	240	258	280	302		162	176	178	182		206	224	226
S. vexans	4	2	224	240	258	280	302		162	176	178	182		206	224	226
S. vexans	4	7	224	240	258	280	312		162	176	178	182		206	224	226
S. vexans	4	1	224	240	258	280	312		162	176	178	182		206	224	226
S. vexans	4	2	224	240	258	282	312		162	176	178	182		206	222	226
vex2	4	2	224		258	264	268	278	160	178	182	194	204	212	224	226

Table S2.4

Taxon	X	Fq.	Alleles at each microsatellite loci						
			SA08			SA02			
S. admonitor	4	19	261	285		292	294		
S. devoniensis	4	30	261			292	294		
S. devoniensis	4	1	261			292	294		
S. subcuneata	3	21	261			292	294		
S. subcuneata	3	3	261			292	294		
S. subcuneata	3	2	261			292	294		
S. margaretae	4	23	247	263		278	282	292	294
S. margaretae	4	1	247	263		278	282	292	294
S. margaretae	4	2	247	263		278	282	292	294
S. margaretae	4	1	247	263		278	282	292	294
S. margaretae	4	2	247	263		278	282	292	294
S. porrigentiformis	4	5	249	257	277	294	300	324	
S. porrigentiformis	4	9	249	257	277	294	300	324	
S. porrigentiformis	4	1	249	257	277	294	300	324	
S. porrigentiformis	4	1	249	257	277	294	300	324	
S. porrigentiformis	4	1	249	257	277	294	300	324	
S. porrigentiformis	4	1	249	257	277	294	296	300	324
S. porrigentiformis	4	1	249	257	277	294	300	324	

Taxon	X	Fq.	Alleles at each microsatellite loci SA08			SA02			
S. rupicola	4	4	263	277		282	292		
S. rupicola	4	1	263	277		282	292		
S. rupicola	4	1	263	277		282	292		
S. rupicola	4	5	263	277		282	292		
S. rupicola	4	1	263	277		282	292		
S. rupicola	4	1	263	277		282	292		
S. vexans	4	1	247	263	275	286	292	294	
S. vexans	4	1	247	263	275	286	292	294	
S. vexans	4	1	247	263	275	286	292	294	
S. vexans	4	7	247	263	275	286	292	294	
S. vexans	4	2	247	263	279	286	292	294	
S. vexans	4	7	247	263	275	286	292	294	
S. vexans	4	1	247	263	279	286	292	294	
S. vexans	4	2	247	263	275	286	292	294	
vex2	4	2	247	275		278	286	294	300

Table S2.5 Microsatellite allele sizes for the sexual diploid species

Microsatellite loci			
Taxon	CH01F02	MSS16	SA01	MSS5
S. torminalis	157 167 175 187 189 209	154 166 170 178 182 184 186 188 190 194 196 198 200 202 204 206 208 210 216 222 230	190 192 212 216 226 230 234 236 238 240 242 244 246 256	105 113 117 119 123 125 127 129 135 137 139 141
S. aria	191 195 197 201	156 158 160 164	212 220 230 232 234 240 242 246	115 121 127 129 135 139 141
Taxon	CH02D11	SA03	SA06	MSS13
S. torminalis	148 150 152 154 156 162 164 170 172 176 178 194 196	214 224 234	258 260 268 270 278 308 310	181 183 187 189 191 193 195 197
S. aria	154 156 164 172 186	224 240 242 250 252 253 254	256 258 260 268 280 282 288	189 195 197 199 203

	Microsatellite loci											
Taxon	SA14				SA08				CH01F09			
S. torminalis	170	176	178	180	229	232	259	260				
	182	184	186	188	261	265	267	269				
	190	198	200	202	271	273	277	281				
	204	206	208	210	283	285						
	212	214	224	226								
S. aria	194	202	210	212	249	251	253	259	111	115	117	121
	216	230	240	242	273	275	277		125	133		
	258											
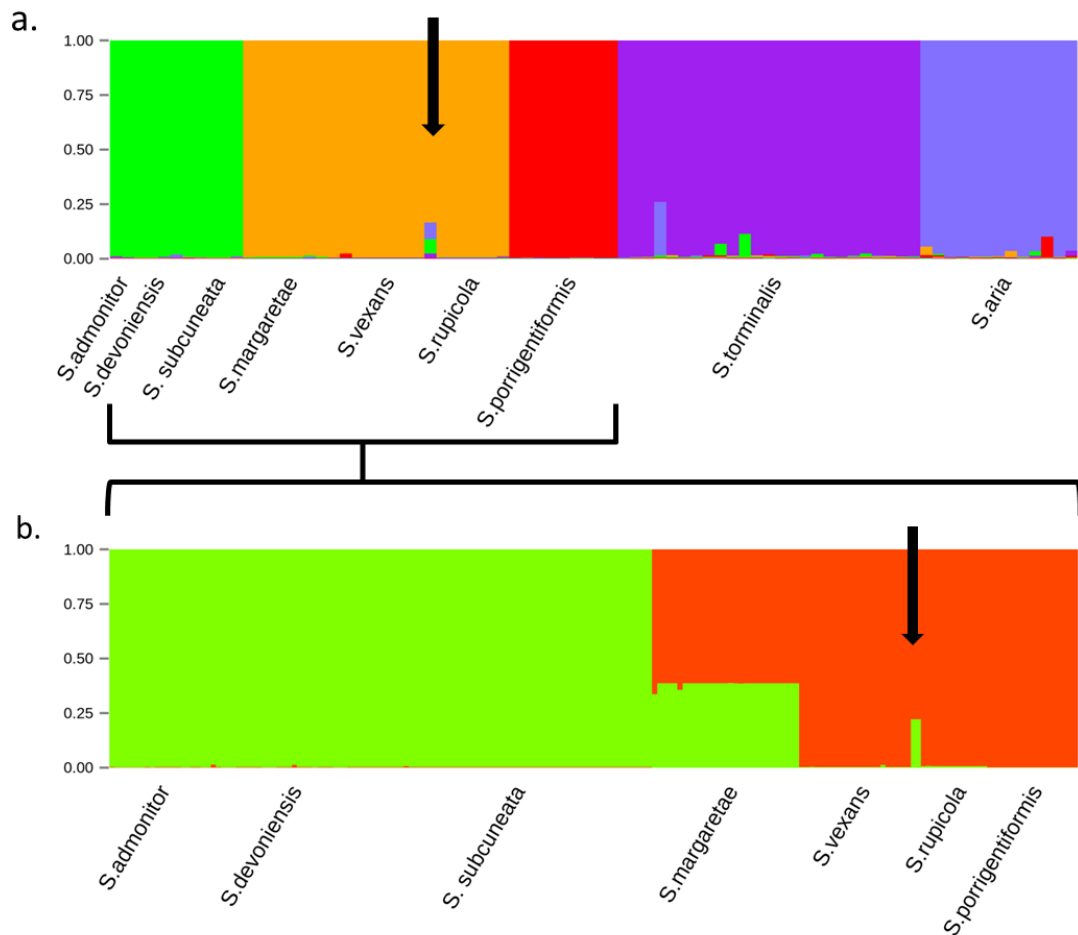
Figure S2.1 Results of Bayesian clustering for all *Sorbus* genotypes and assignment to genetic clusters as determined by STRUCTURE. The colours represent the probability of assignment to different clusters. Arrows represent the second *S. vexans* clone (vex2).

a). Diploid and polyploid taxa. K=5 (mean L(K) = -3274.64) each vertical bar represents a single multilocus genotype. 82 samples at 10 microsatellite loci.

b). Polyploid taxa only. K=2(mean L(K) = -12481.6) each vertical bar represents a polyploid individual. 160 samples at 14 microsatellite loci.

141
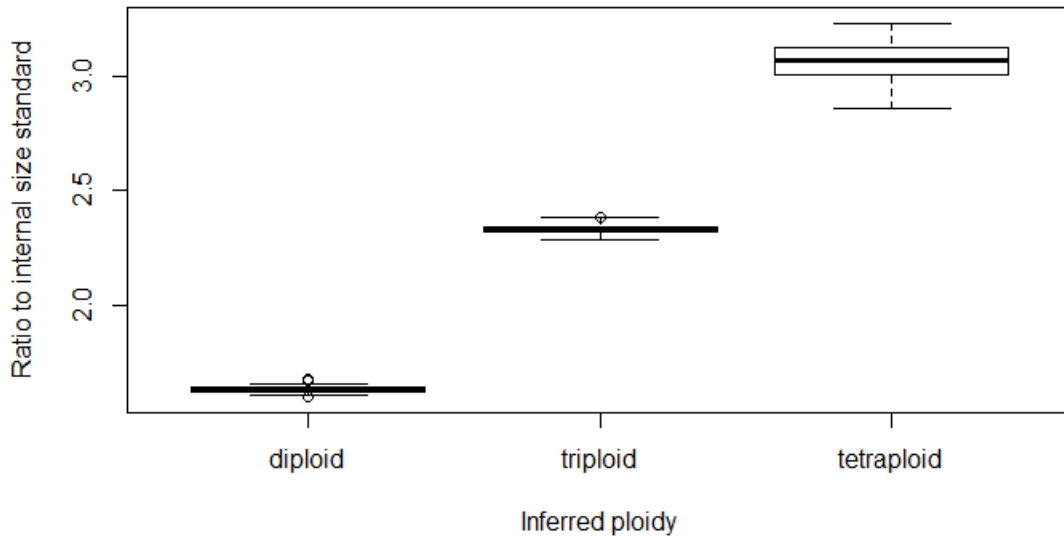
Figure S2.2 Relative nuclear DNA for each of the distinct clusters corresponding with chromosome number (ploidy level). ANOVA p= < 2.2e-16
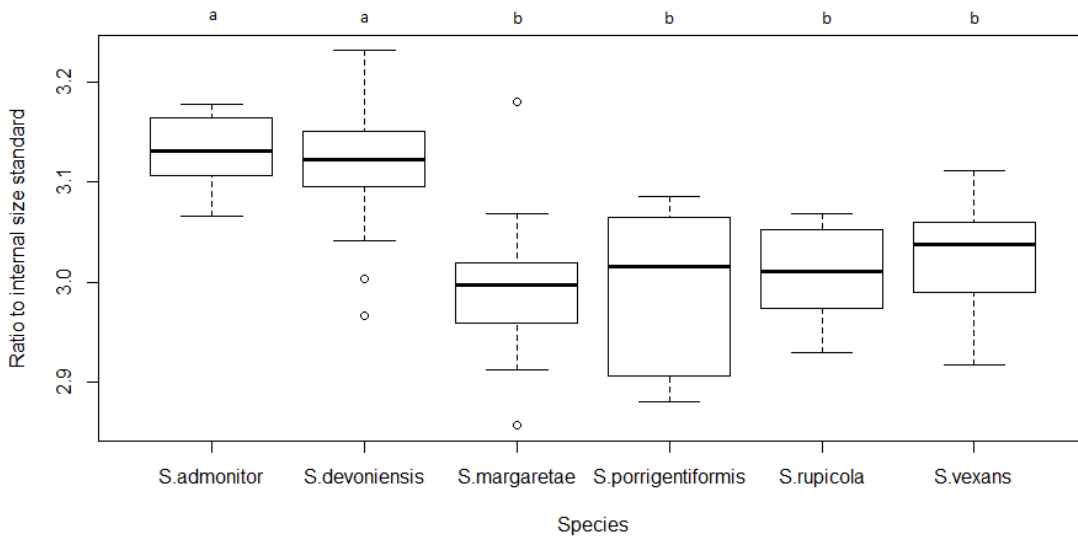
Figure S2.3 Relative nuclear DNA content of the tetraploid species. a and b indicate which comparisons were statistically significant in Tukey post hoc pairwise analysis with p< 0.01. Sample sizes = 12, 31, 13, 7, 11, 13, left to right.

Chapter 3: Inheritance patterns and hybrid origins of apomictic polyploids: *Sorbus* in the southwest UK

*Tracey J. Hamston[1,2], Natasha de Vere[3], R. A. King[1], James E. Cresswell[1] and Jamie R. Stevens[1]

[1] *Biosciences, College of Life & Environmental Sciences, University of Exeter, Geoffrey Pope Building, Stocker Road, Exeter, EX4 4QD, UK*

[2] *Whitley Wildlife Conservation Trust, Paignton Zoo, TQ4 7EU, UK*

[3] *National Botanic Garden of Wales & Institute of Biological Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, UK*

*E-mail: thamston@gmail.com

Formatted for submission to Botanical Journal of the Linnean Society.

143

Abstract

The hybrid formation of polyploids is accepted as a key mode of speciation in plants. Unravelling the processes that give rise to polyploid taxa is essential to our understanding of plant evolution. In this study we use microsatellites to reconstruct parentage and elucidate the hybrid origins for seven polyploid *Sorbus* species endemic to the south west of England. Exact pairwise parentage matches for tetraploids *S. devoniensis* and *S. admonitor* reveal their formation is most likely via the 'triploid bridge' with triploid *S. subcuneata* and diploid *S. torminalis* as parental species. Distribution patterns of this group indicate a wider historic range for *S. torminalis* in southwest England. Allele composition of the highly endemic study members of subgenus *Aria* hints at possible links with other polyploid species approximately 100km east. This may suggest possible colonisation routes. Our results demonstrate a dynamic system of diversification in *Sorbus*, but that the production of persistent novel species is a rare event.

3.1 Introduction

It has long been recognised that polyploidy, in the presence of more than two sets of chromosomes, has been of major importance in the diversification of angiosperm lineages (Grant, 1981, Soltis *et al.*, 2009). Indeed, estimates suggest 70% of flowering plants are descended from polyploid ancestors with different ploidy levels seen within and among closely related species, often arising on multiple occasions (Ramsey & Schemske, 1998, Soltis, 2005). This apparent profusion of polyploidization and its significance for plant evolution has led to increasing interest in processes involved in the formation and establishment of polyploids. However, a recent review highlights the huge gaps in coverage of well-known polyploids and the need for more studies of natural evolutionary models, as most of what is known about polyploidy comes from a few crop and genetic model systems (Soltis *et al.*, 2016).

Polyploids are generally classified into autopolyploids and allopolyploids. Following the definition of Ramsey & Schemske (1998), those that have arisen within populations of single species are termed autopolyploids, and allopolyploids are those arising from interspecific hybridisation events. The type of polyploid also affects the mode of genome inheritance. Autopolyploids, will generally have chromosomes with similar structure derived from a common parental species, so duplicated chromosomes can pair at random (Fig. 3.1a). This gives rise to polysomic inheritance, for example alleles at a given locus in an autotetraploid ABCD, may segregate A-B, A-C, A-D, B-C, B-D or C-D. Allopolyploids have homeologous (partially homologous) chromosomes derived from different lineages (Fig 3.1b). The homologous pairs tend to remain differentiated and inheritance is similar to a diploid organism, termed disomic

145

(Ramsey & Schemske, 2002). Some groups may exhibit a mixture of disomic and polysomic inheritance patterns, sometimes at different loci (Lerceteau-Köhler et al, 2003).

a)

Autotetraploid

Possible gametes. Polysomic inheritance

b)

Allotetraploid

Gametes showing disomic inheritance
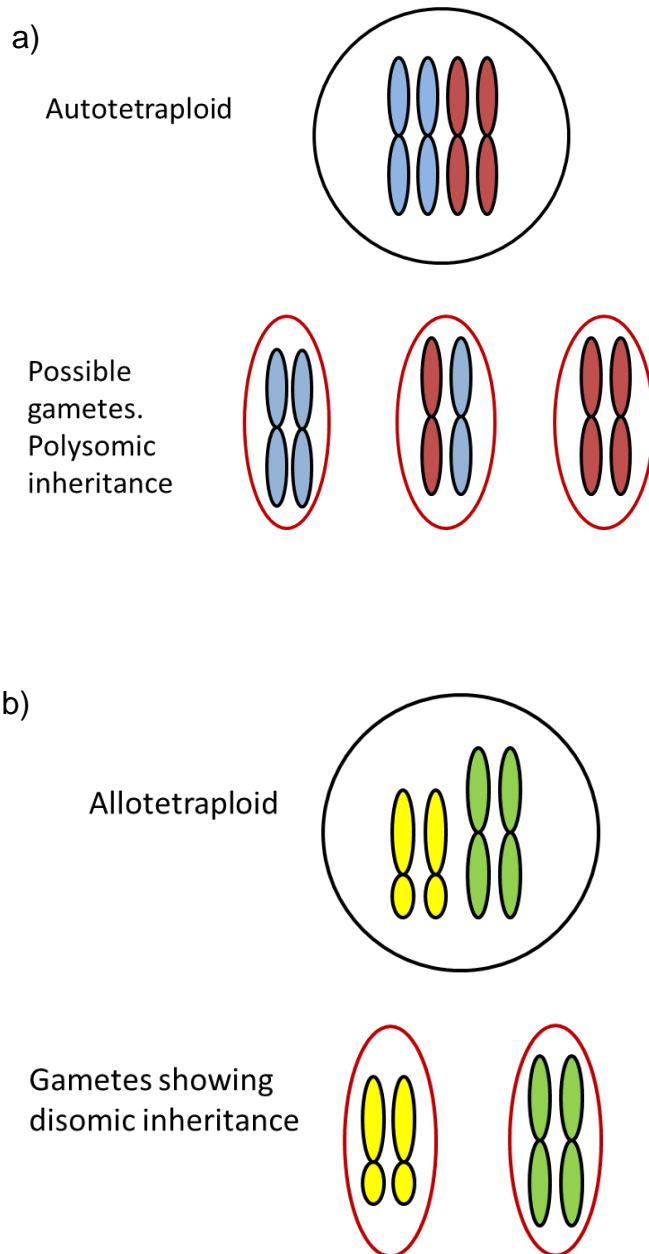
Figure 3.1. Polysomic a) and disomic b) inheritance of auto and allotetraploids showing possible allele combinations in gametes.

Polyploid hybridization results in novel combinations of genotypes combined with high levels of heterozygosity which may increase adaptive potential of

novel polyploids compared to diploid progenitors (Sobel *et al.*, 2010) and there are a growing number of examples of ecological divergence associated with polyploidy (Ramsey & Ramsey, 2014, Thompson *et al.*, 2014, Segraves & Anneberg, 2016). Despite its evolutionary and ecological importance, it may be difficult to determine the steps involved in the polyploid speciation process due to the complex and often reticulate patterns of hybridisations and the cryptic nature of many hybridogenous taxa. Indeed, the process of polyploidization often leads to extensive chromosomal rearrangements and loss of duplicated genes within a few generations (Paun *et al.*, 2007, Hegarty & Hiscock, 2009, Soltis *et al.*, 2016). Therefore, in order to study polyploid evolution, it is necessary to choose a polyploid complex with a high level of genome stability so patterns of inheritance and hybrid origins may be determined unambiguously. The study of largely asexual or apomictic populations of closely related polyploids offers this potential as asexual reproduction will maintain the original genomic composition, with less subsequent reshaping of novel hybrids seen in fully sexual polyploids (Wendel, 2000, Mandáková *et al.*, 2016).

Apomixis, asexual seed production, synonymous with agamospermy is often associated with polyploidy (Whitton *et al.*, 2008) in combination with a loss of self-incompatibility (Ramsey & Schemske, 1998). The asexual formation of seed from the maternal ovule tissue bypasses meiosis and thus potentially enables establishment of viable novel polyploids in sympatry with their progenitors re-enforcing reproductive isolation and generating clonal lineages. This eliminates rapid introgression and allows investigation into formation routes by studying allele segregation among the clonal species. Since apomixis is generally considered facultative at some level (Nogler, 1984), occasional sexual reproduction offers opportunities for further polyploid production and even

147

where apomixis is obligate, apomicts may still pollinate sexual plants (Richards, 2003). The development of apomictic lineages of some polyploid hybrids means that combinations of genotypes may be maintained within the polyploid genome and provide a snapshot of the evolutionary process at the moment of inception. Such polyploids could be seen as reservoirs of alleles from ancestral sexual progenitors.

A primary route of polyploid formation is via unreduced *2n* gametes that retain the somatic chromosome number (Bretagnolle & Thompson, 1995, Rieseberg & Willis, 2007). The tendency of sexual diploids to produce *2n* gametes can be environmentally induced and genetically inherited (Ramsey & Schemske, 1998, Köhler *et al.*, 2010). Successful fusion of such gametes with reduced gametes within the same population will produce triploids. However, the presence of post zygotic reproductive barriers such as hybrid sterility, often brought about by the inability for chromosomes to pair during meiosis due to the presence of odd numbered ploidy, a phenomenon known as 'triploid block', has been proposed to prevent establishment of novel polyploids in a sexual system (Marks, 1966). Triploid block has also been proposed as a mechanism which may prevent establishment of tetraploids in a diploid population as they can suffer minority cytotype exclusion brought about via a frequency-dependant mating disadvantage as the majority of pollinations will produce sterile triploids (Levin, 1975, Husband & Schemske, 2000). However, triploids contribute significantly to the ongoing production of polyploids both in sexual and asexual or apomictic systems (Ramsey & Schemske, 1998). Although their fitness is generally lower than either diploids or tetraploids they may still enhance the production of tetraploids when present in sufficient numbers (Husband, 2004), often via the

148

'triploid bridge' where unreduced triploid gametes combine with reduced diploid gametes (Fig. 3.2).
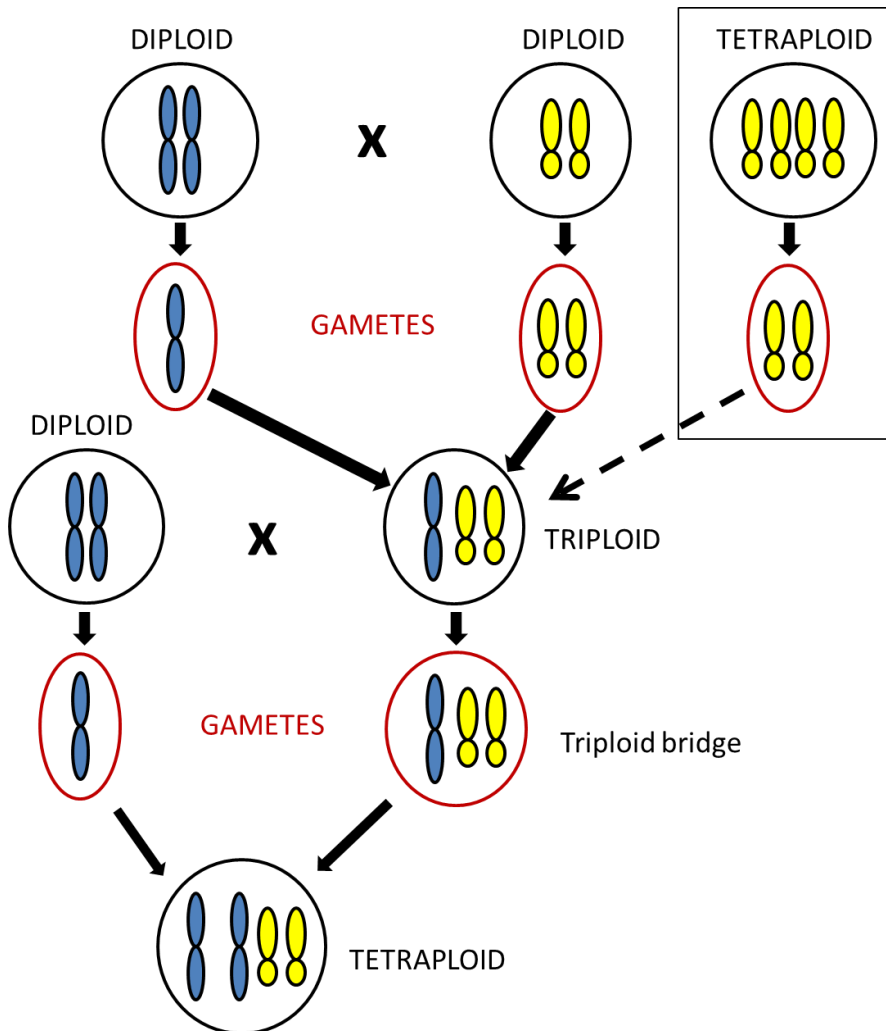
Figure 3.2. One possible route of triploid and tetraploid formation illustrating the 'triploid bridge'.

The genus *Sorbus* contains both sexual diploids and apomictic polyploids of hybrid origins. Although the polyploids are largely apomictic they still require pollination for the initiation of seed (pseudogamy). *Sorbus* are long lived tree species, thus novel hybrids can persist for tens or hundreds of years, combined with potentially long generation times they offer a window into hybrid events that have occurred over long time periods. Genome sizes in *Sorbus* are very

consistent with little evidence of downsizing (Pellicer *et al.*, 2012, Hajrudinović *et al.*, 2015) which suggests a degree of stability making the genus suitable for the investigation of hybrid origins in a polyploid complex.

The primary route for polyploid formation in *Sorbus* is via backcrossing of triploid and tetraploid with diploid parental taxa with the polyploid as either the maternal or paternal parent (Nelson-Jones *et al.*, 2002, Robertson *et al.*, 2004b, Robertson *et al.*, 2010, Ludwig *et al.*, 2013). Chloroplast and mitochondrial DNA have identified the ancestral maternal lineages of many UK polyploid species (Nelson-Jones *et al.*, 2002, Chester *et al.*, 2007) but the pollen donors are largely unknown as is the nature of any further hybridisation events.

Devon and north Somerset in south-west England contain four apomictic polyploid *Sorbus* species endemic to the region (*Sorbus subcuneata* Wilmott*, S. admonitor* M.C.F. Proctor*, S. vexans* E.F. Warb and *S. margaretae* M.C.F. Proctor) along with three other closely related polyploid taxa with wider distributions (*S. devoniensis* E.F. Warb*, S. porrigentiformis* E.F. Warb and *S. rupicola* (Syme) Hedlund) and one putative parental sexual diploid *S. torminalis* (L.) Crantz*. S. aria* L. which is also closely related, although it is not thought to be native to this southwest peninsular. It occurs naturally further to the east in Somerset, where it grows with other polyploid taxa (Fig. 3.3). It is, however, widely planted across the region so its native distribution is difficult to ascertain (Rich *et al.*, 2010).
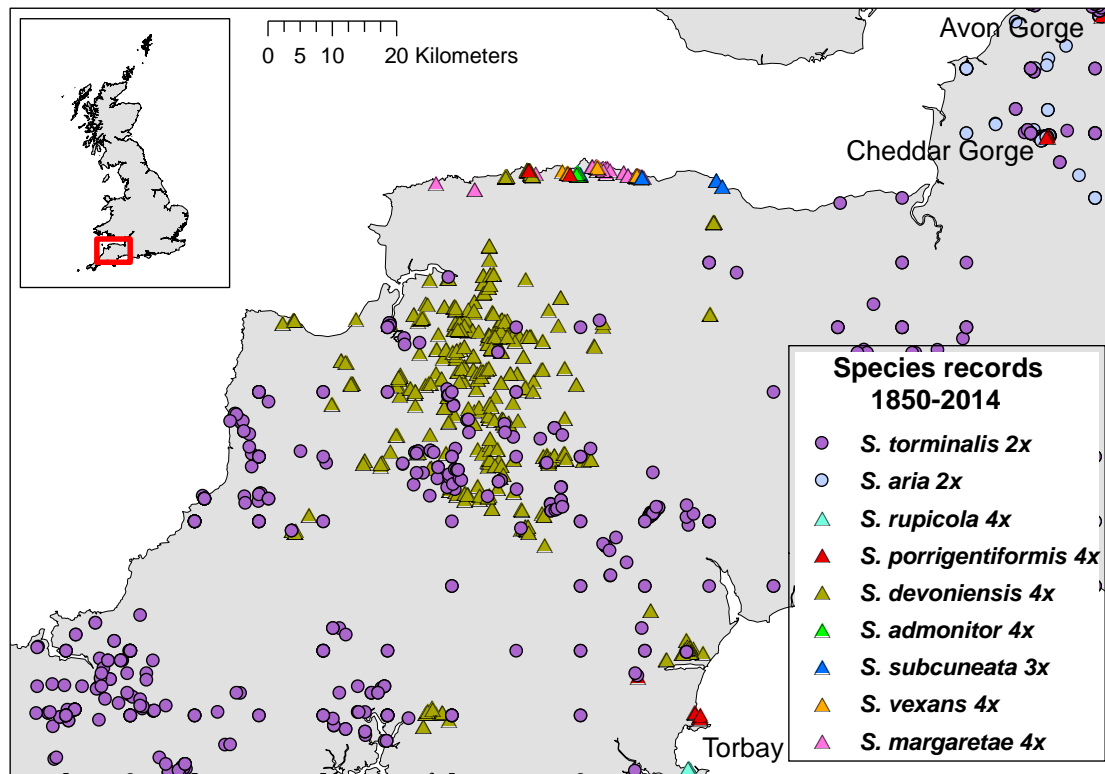
Figure 3.3. Species distribution records for study species supplied by T. C. G. Rich. Note that records for *S. aria* are restricted to eastern areas of this region. (The map was created using ArcGIS Desktop version 10.2.2, ESRI, California, USA, URL: http://www.esri.com/).

Following the classification described by Rich *et al.* (2010), the polyploid species in this study are taxonomically divided into two subgenera; subgenus *Aria* is composed of diploid *S. aria* and its polyploid derivatives (*S. rupicola, S. porrigentiformis, S. margaretae* and *S. vexans*); subgenus *Tormaria* is represented by three species (*S. devoniensis, S. admonitor and S. subcuneata*) resulting from hybridisation between diploid *S. torminalis* as the ancestral maternal parent (Chester *et al.*, 2007) and members of sub genus *Aria*. All polyploids are tetraploid with the exception of *S. subcuneata* which is a triploid (see section 2.3). The exact hybrid origins of all the polyploid species are unknown. Morphology and AFLP analysis place diploid *S. aria* and / or

151

tetraploid *S. rupicola,* which has a northern European distribution and likely originated outside the UK, as parental species' for the remaining study polyploid species of subgenus *Aria* (Sell, 1989, Lemche, 1999)*.*

By determining the hybrid origins of recently formed hybridogenous polyploids it may be possible to determine the genome contributions of each parental species which would give insight into the mode of inheritance and possible routes of polyploid formation. Due to their high variability, microsatellites are often used for pedigree or parentage analyses where often only a few informative loci are required for identification of parents through the comparison of genotypes (Gerber *et al.*, 2000). This chapter extends the microsatellite analysis of the 207 samples of nine *Sorbus* study species described in Chapter 2.

The main aims of this study were to determine the hybrid origins of this group of polyploid species, to ascertain genome configurations of subgenus *Tormaria* and to identify likely routes of polyploid formation in *Sorbus*.

3.2 Materials and methods

3.2.1 Molecular methods

In order to reconstruct the hybrid origins of the seven polyploid taxa we used 12 nuclear DNA microsatellite loci previously used in the genus *Sorbus* to determine genotypes for 206 individuals from the nine study species as described in Chapter 2 (2.2). Loci used for this study were CH01F02, MSS16, SA01, MSS5, CH02D11, SA03, SA06, MSS13, SA14 and SA08, and to elucidate the inheritance patterns in subgenus *Tormaria,* we used an additional two loci which had been found to only amplify in either *S. aria* or *S. torminalis*, CH01F09 (Robertson *et al.*, 2010) and MS14 (Nelson-Jones *et al.*, 2002) respectively. DNA extraction and PCR conditions are described in Section 2.2.2. For details of the microsatellite primers and multiplex design see Tables S2.2 and S2.3 in supplementary information for Chapter 2.

3.2.2 Allele dosage

To overcome the problem of defining which alleles occur in more than one copy, allele dosage for the majority genotype of each of the polyploid species was determined using the MAC-PR (microsatellite DNA allele counting-peak ratios) method (Esselink *et al.*, 2004). This method utilises the quantitative values for microsatellite allele peak areas provided by the fragment analysis software (CEQ 8000 Genetic Analysis system, Beckman) to determine allele copy number, therefore only loci which produced unambiguous, consistent peaks may be used. As polyploid *Sorbus* are primarily apomictic, each species has a multi-locus genotype profile and allele copy was calculated at species level. Using the mean peak area for each species allows for some variation in peak quality which precludes use of this method for individual samples. The alleles

from all species were inspected at each locus and ratios between the mean peak areas for pairs of alleles occurring in more than one species were analysed in pair-wise combinations. Species where the maximum number of alleles equalled ploidy, the allele ratio thus equalling 1:1, were used as a baseline. The mean ratios for each species of unknown copy number were divided by the mean baseline taxon ratio to give relative proportions. Thus, if differences in peak amplification produce a peak ratio of 1.2 from a known 1:1 allele ratio used as a baseline, all unknown ratios were then divided by 1.2. This allowed for variation in amplification between different sized alleles. ANOVA was used to test for variation among species for each allele pair and if this was significant the relative peak ratios were plotted for further inspection to compare size ranges and only those peak ratios that were clearly separated without overlap were used to infer allele copy.

This method can only be used where the same pairs are present in more than one group and where a robust baseline can be established. The MAC-PR method is also useful for the detection of null alleles as, unlike diploids, there is currently no software available to test for the presence of null alleles in a polyploid population.

3.2.3 Inheritance patterns of subgenus *Tormaria*

To identify the relative contributions of *S. aria* and *S. torminalis* genomes to study members of subgenus *Tormaria,* we used the genome specific loci CH01F09 and MS14 in addition to any loci where allele copy number could be determined for members of this subgenus. All alleles were matched to either *S. torminalis* or *S. aria* wherever possible. The proposed genome contributions are summarised in Table 3.1.

154

Table 3.1. A summary of ploidy level, breeding systems and proposed genome composition of the study species (commonly suggested genome compositions are indicated by a capital letter A = *aria* genome and T = *torminalis* genome). X = ploidy based on flow cytometry data (see Section 2.3.4), ♀ = maternal chloroplast type from Chester *et al.* (2007) and Nelson-Jones *et al.* (2002).

Taxon	X	Breeding system	♀	Proposed genome composition
S. torminalis	2*x*	Out-crossing	n/a	TT
S. subcuneata	3*x*	Apomictic	*S.torminalis*	AAT
S. devoniensis	4*x*	Apomictic	*S.torminalis*	AATT
S. admonitor	4*x*	Apomictic	*S.torminalis*	AATT
S. aria	2*x*	Out-crossing	n/a	AA
S. porrigentiformis	4*x*	Apomictic	*S. aria*	AAAA
S. rupicola	4*x*	Apomictic	*S. aria*	AAAA
S. vexans	4*x*	Apomictic	*S. aria*	AAAA
S. margaretae	4*x*	Apomictic	*S. aria*	AAAA

3.2.4 Parentage analysis

The most likely parental species for each polyploid taxon were assessed by compiling and comparing multi-locus genotypes. The accumulation of mutations is likely to be an issue for determining hybrid origins as it may be many generations since the occurrence of the hybrid event. Therefore, the majority genotype was used for each apomictic polyploid species where >80% of individuals follow type, since small allele variations, ascribed to mutations were observed in the minority (chapter 2). The exception to this was the highly variable locus SA06 where several allele size variants were included for *S.*

rupicola and *S. vexans.* For the sexual parental species all alleles from all samples were considered together as a pool of potential alleles, since the original genotype was presumed extinct. The diploid species sample sites included Cheddar Gorge, Somerset (lat. 51.287249, long. -2.7470425) and Avon Gorge, Bristol (lat. 51.468669, long. -2.6305114) and we included alleles reported in previous studies of *Sorbus* at these sites for microsatellite markers common to our study; CH01F02, CH02D11, MSS5 and MSS16 (Houston *et al.*, 2009) plus MSS13 (Robertson *et al.*, 2010). Since reported allele sizes depends on the calibration of the fragment analysis process, we verified allele sizes from these studies using *S. porrigentiformis* as baseline genotypes for the Avon Gorge study and *S. anglica* for the Cheddar study and adjusted the reported allele sizes accordingly. *S. anglica* was not part of our study group but had also been sampled and genotyped by the author. This also enabled us to make comparisons at these common loci between our study species and other highly endemic polyploid taxa at these other sites. To investigate the most likely parental species for each polyploid species, each species was considered a putative parent in turn and the genotype was compared to a target offspring species and the number of loci where there was a contribution of at least one allele from the putative parental species was recorded.

We then compiled the various possible parental pairwise combinations of alleles and scored each combination based on the number of mismatches of alleles missing from the pairing required to generate the putative offspring genotype. This follows the method of Robertson *et al.* (2010) where an exact match with no missing alleles scores zero, two missing alleles scores two etc. This was achieved most simply by converting all genotypes to a binary presence / absence matrix (1, 0) where presence of an allele = 1 and absence = 0, which

156

was carried out in the POLYSAT package in R (Clark & Jasieniuk, 2011). Where an allele was missing from both putative parents (0, 0) the sum equalled zero and was scored as a mismatch. The number of mismatches or missing alleles was compiled for each putative parent pair.

All analysis was done at the species level except for *S. vexans* which has two distinct genotypes *S. vexans s.s.* and vex2, both of which were analysed separately.

The results from all these analyses, relative levels of genetic variation (See Chapter 2) plus morphological features are combined to suggest the sequence of hybrid events and probable parentage for each species.

3.3 Results

3.3.1 Microsatellites

All ten loci amplified alleles for all samples. The additional two loci, CH01F09 and MS14 proved to be genome specific, but both loci yielded alleles for all three members of subgenus *Tormaria* confirming their status as hybrids between subgenus *Aria* and *S. torminalis*. The tetraploid members of subgenus *Aria* all had four alleles at the *Aria* genome specific locus CH01F09 and no alleles at locus MS14 (*Torminalis* genome specific). Thus, *S. torminalis* and members of subgenus *Tormaria* are eliminated as potential parents of any member of subgenus *Aria*. The relative allele sizes for *S. porrigentiformis* corresponded with those from the Avon gorge study with matching size increments between alleles at all the common loci. This was also the case for the *S. anglica* samples from the Cheddar Gorge study. For the two diploid species, seven alleles were added from Cheddar and Avon for *S. aria* and seven alleles from Avon for *S. torminalis* from previous studies (see method section 3.2.4). These two diploid species had overlapping allele size ranges with common alleles at all loci. The alleles used in this study are shown in Table 3.2. Microsatellite electropherograms for the polyploid species showed maximum peaks which corresponded to ploidy (e.g.Fig. 3.4).

3.3.2 Allele dosage

Where allele copy number was determined, it enabled the inference of inheritance patterns for subgenus *Tormaria,* with alleles allocated to either *S. torminalis* or *S. aria* where possible. Mean peak area ratio (MAC-PR) analysis successfully elucidated all allele copy numbers for all the polyploid species at four loci, CH01F02, MSS16, MSS5 and SA14 and subgenus *Aria* at locus

158

CH01F09 (Table 3.2). Allele copy number was not elucidated where there were ambiguous relative peak area ratios or insufficient comparable allele pairs. Table 3.3 summarises the relative peak ratios and ANOVA results. Typical box plots used to inspect the relative peak ratios among species are shown in Fig. 3.5.

3.3.3 Inheritance patterns for subgenus *Tormaria*

The most likely genome composition of members of subgenus *Tormaria* was determined where alleles were matched exclusively to either *S. torminalis* or subgenus *Aria.* This was achieved at the genome specific loci CH01F09 and MS14, i.e. triploid *S. subcuneata* = AAT; tetraploids *S. admonitor* and *S. devoniensis* = AATT. These combinations were also supported by the allele configurations at locus MSS16. This pattern was not contradicted at any other loci, although due to overlap in size ranges and common alleles to *S. aria* and *S. torminalis,* genome composition was ambiguous at the remaining loci. Tetraploids *S. admonitor* and *S. devoniensis* shared the three alleles of triploid *S. subcuneata* at all loci plus one additional allele, suggesting a common origin for all three species. This additional allele could be detected in seven loci and attributed to *S. torminalis* at these loci; however, it was also shared with *S. aria* at three loci for *S. devoniensis* and one locus for *S. admonitor.*

159
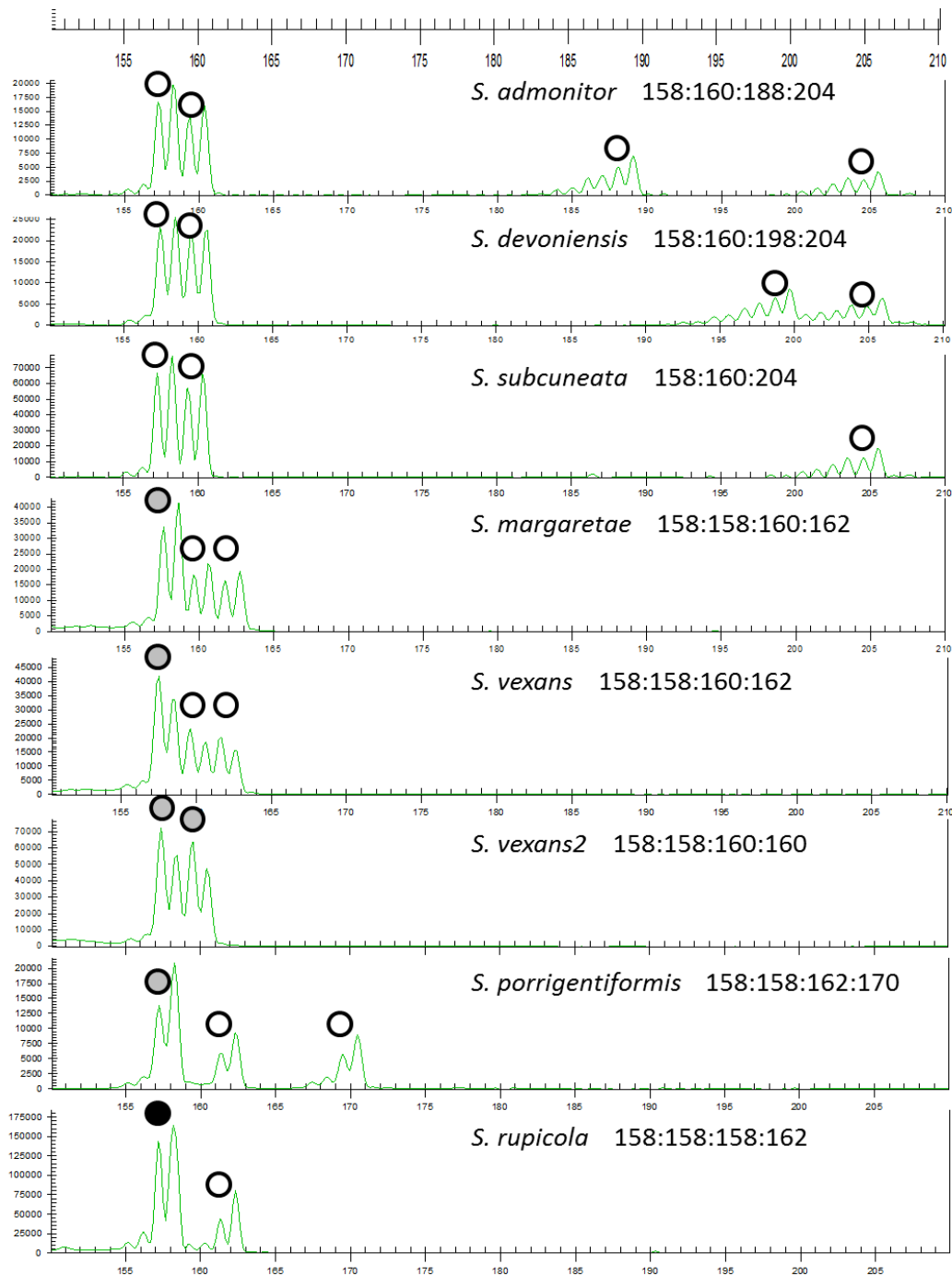
Figure 3.4 Sample electropherogram of one locus (MSS16) showing the polyploid genotype patterns for the study group. Allele sizes are in base pairs (bps). Note, the larger alleles (188-204) derived from diploid *S. torminalis* in the members of subgenus *Tormaria*. The copy number for each allele was inferred from a comparison of peak area ratios as described by Esselink *et al.* (2004)

○ = one copy; ◒ = two copies; ● = three copies.
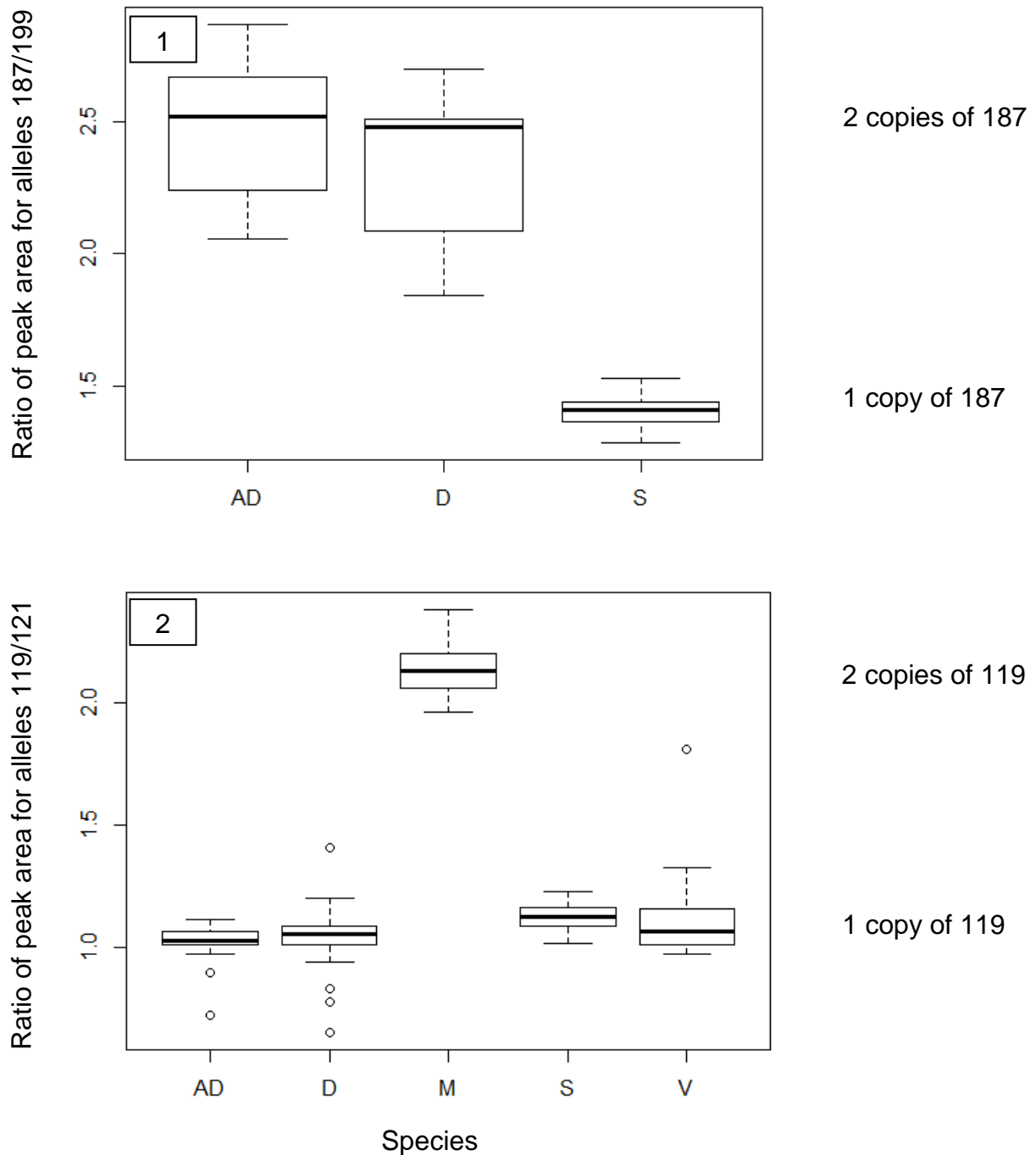
Figure 3.5 Exemplar box plots for visual comparison of the peak area ratios of allele pairs from two microsatellite loci. The allele pair is labelled on the y axis. Species codes are: AD = *S. admonitor,* D = *S. devoniensis,* M = *S. margaretae,* S = *S. subcuneata* (the baseline 1:1 for both plots)*,* V = *S. vexans s.s.*

1. Locus CH01F02; ANOVA, F = 21.13, p = 0.000117, df = 2, n = 5, 5, 5. *S. devoniensis* and *S. admonitor* both have two copies of allele187.

2. Locus MSS5; ANOVA, F = 322.0, p <2e-16, df = 4, n = 13, 27, 22, 31, 19. *S. margaretae* has two copies of allele 119.

Table 3.2. Genome composition of SW endemic taxa. X = ploidy level. Allele sizes are from the majority (80%) genotype for polyploid taxa. Red text indicates where allele copy number has been determined using the MAC-PR method (Esselink *et al.*, 2004), see Table 3.3. Underlined alleles are from Avon Gorge and Cheddar Gorge (Houston *et al.*, 2009, Robertson *et al.*, 2010). Coloured cells indicate possible sources of alleles for the six endemic taxa. ▧ Alleles associated with *S. torminalis* samples, ▧ alleles associated with *S. aria* samples, ▧ alleles associated with non-endemic polyploids *S. rupicola* and *S. porrigentiformis*, no colour indicates alleles common to more than one of the above groups and ▧ alleles unique to the six local endemic polyploid taxa.

Taxon	X	CH01F02				MSS16				SA01				MSS5			
									Microsatellite loci								
S. admonitor	4	187	187	195	199	158	160	188	204	224	232		242	119	121	123	127
S. devoniensis	4	187	187	195	199	158	160	198	204	224	232	234	242	119	121	123	127
S. subcuneata	3	187		195	199	158	160		204	224	232		242	119	121	123	
S. margaretae	4	191	195	199	221	158	160	162	158	224	232	224	232	119	121	119	135
S. vexans	4	191	195	201	203	158	160	162	158	236	236	236	236	119	121	127	127
S. vexans (vex2)	4	191	195	203	221	158	160	160	158	232	236	238		121	127	129	135
S. porrigentiformis	4	191	197	201	203	158	162	170	158	214	236	244		115	127	131	137
S. rupicola	4	191	199	201	209	158	162	158	158	224	230	234	236	119	127	131	
S. torminalis	2	157	167	175	187	154	166	170	178	190	192	212	216	105	113	117	119
		189	209	191	195	182	184	186	188	226	230	234	236	123	125	127	129
						190	194	196	198	238	240	242	244	135	137	139	141
						200	202	204	206	246	256						
						208	210	216	222								
						230	192	214									
S. aria	2	191	195	197	201	156	158	160	164	212	220	230	232	115	121	127	129
		193	207							234	240	242	246	135	139	141	

		Microsatellite loci			
Taxon	X	CH02D11	SA03	SA06	MSS13
S. admonitor	4	152 162 182	224	258 268	187 193 195 193
S. devoniensis	4	152 162 182	224	258 268	189 193 195 193
S. subcuneata	3	152 162 182	224	258 268	193 195 193
S. margaretae	4	152 154 182 152	224	258 268 280 312	193 195 197 195
S. vexans	4	150 154	224 240	258 280 302 312 / 294 282	193 195 197 199
S. vexans (vex2)	4	152 154 182 152	224	258 264 268 278	193 195 199 195
S. porrigentiformis	4	152 198	228 240	256 264 270	193 195 197 203
S. rupicola	4	152 162 168 162	224 240	258 264 280 282 / 300 312 302	193 195 197
S. torminalis	2	148 150 152 154 / 156 162 164 170 / 172 176 178 194 / 196 174 193 200	214 224 234	258 260 268 270 / 278 308 310	181 183 187 189 / 191 193 195 197
S. aria	2	154 156 164 172 / 186 150 174 176 / 180	224 240 242 250 / 252 253 254	256 258 260 268 / 280 282 288	189 195 197 199 / 203

164

Taxon	X	SA14				SA08				CH01F09				MS14			
																	Microsatellite loci
S. admonitor	4	170	178	208	226	261	285			113	123			123	131		
S. devoniensis	4	170	204	208	226	261				113	123			123	133		
S. subcuneata	3	170		208	226	261				113	123			123			
S. margaretae	4	194	208	208	226	247	263			113	115	123	115				
S. vexans	4	206	206	224	226	247	263	275		115	115	121	121				
S. vexans (vex2)	4	204	212	224	226	247	275			115	115	123	123				
S. porrigentiformis	4	196	222	224	226	249	257	277		115	123	125	129				
S. rupicola	4	194	206	208	206	263	277			113	115	121	121				
S. torminalis	2	170 182 190 204 212	176 184 198 206 214	178 186 200 208 224	180 188 202 210 226	229 261 271 283	232 265 273 285	259 267 277	260 269 281					122 129	123 131	125 133	127 135
S. aria	2	194 216 258	202 230	210 240	212 242	249 273	251 275	253 277	259	111 125	115 133	117	121				

Table 3.3. Mean peak ratio values for each allele pair tested. Shaded boxes indicate the baseline ratio and those marked with $^\$$ are derived from the pairwise comparisons of other allele pairs. *F* values from one way ANOVA's are shown with significance level. *$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.The upper values for each taxon are the mean peak area ratios and the lower values are the relative values to the known ratio. Sample sizes are in parenthesis. All taxa are tetraploids except S = *S. subcuneata* which is triploid. AD = *S. admonitor*, D = *S. devoniensis,* M = *S. margaretae,* V = *S. vexans s.s.,* V2 = vex2, P = *S. porrigentiformis,* R = *S. rupicola.*

Locus	Allele pair	*F*	Mean peak ratio for each taxon							
			S	AD	D	M	V	V2	P	R
CH01F02	187/199	21.13 ***	1.41 (5)	2.47 (5)	2.32 (5)					
			1:1	1.94	1.86					
	195/199	1.03 n.s.	1.23 (29)	1.24 (14)	1.21 (27)	1.27 (21)				
			1:1	1	0.98	1:1				
	187/195	125.7***	1.05 (29)	1.89 (14)	2.01 (27)					
			1:1	1.78	1.86					
MSS16	158/160	260.3***	1.1 (31)	1.08 (12)	1.1 (27)	1.88 (23)	1.85 (17)	1.06 (2)		
			1:1	1:1	1:1	1.71	1.68	0.96		
	158/162	90.1***				2.11 (23)	2.1 (17)		2.23 (9)	3.1 (11)
						2:1 $^\$$	2:1 $^\$$		1.1	1.5

Locus	Allele pair	F	Mean peak ratio for each taxon							
			S	AD	D	M	V	V2	P	R
CH02D11	162/152	90.81***	0.24 (19)	0.17 (10)	0.17 (18)					0.78 (9)
			1:1	0.69	0.72					3.27
	152/182	31.39***	1.88 (19)	2.37 (10)	2.43 (18)	2.89 (24)		2.95 (2)		
			1:1	1.26	1.29	1.54		1.57		
MSS5	119/127	71.99***		1.82 (13)	1.89 (27)		1 (18)			2.23 (11)
				1:1	1:1		0.5			1.22
	119/121	322.9***	1.1 (31)	1 (13)	1.03 (27)	2.1 (22)	1.1 (19)			
			1:1	1:1	1:1	2.1	1			
SA14	208/226	50.1***	1.45 (10)	1.5 (2)	1.45 (16)	0.7 (11)				
			1:1	1:1	1:1	0.5				
	194/208	173.9***				1.7 (11)				0.7 (7)
						2:1				0.41
CH01F09	115/123	364***				2.69 (24)		1.28 (2)	0.95	
						2.82		1.34	1:1	
	121/115	581***					0.79 (20)			1.6 (11)
							0.49			1:1, 2:1 or 1:2
	115/113	214.9***				1.57 (24)				0.98 (11)
						1.6				1:1 [$]

Locus	Allele pair	F	Mean peak ratio for each taxon							
			S	AD	D	M	V	V2	P	R
MSS13	193/195	88.08***	1.98 (13)	1.85 (3)	2.04 (6)	0.5 (18)	1.1 (16)	0.7 (2)	0.93 (4)	0.9 (8)
			1.8	1.7	1.9	0.5	1:1	0.6	0.8	0.8
	195/199	39.07***					1.64 (16)	3.13 (2)		
							1:1$^{\$}$	1.9		
SA01	224/232	0.19 n.s.	1.78 (9)	0.55 (12)	0.48 (22)	1.18 (19)				
			1:1	1:1		1.05				

3.3.4 Parentage analysis

The three study members of subgenus *Tormaria*, *S. subcuneata, S. devoniensis* and *S. admonitor* had matching alleles at all 10 loci. They also had common alleles with *S. torminalis* at every locus, as mentioned in section 3.3.3. For these three species, the pairwise comparisons generated multiple genotype matches including a match of S*. margaretae* ✕ *S. torminalis* (Table 3.4)*.* The most likely of these potential matches are discussed below and are highlighted in Table 3.4.

Of the study members of subgenus *Aria*, only *S. margaretae* and the second *S. vexans* clone, vex2, had matching alleles with any other study taxa at all ten loci. The pairwise comparisons generated no exact matches for potential parentage of members of subgenus *Aria* among the study group*.* All missing alleles for each parentage match were subsequently checked against the minority genotypes for each potential progenitor without resolution. The next best match was recorded with a score for the number of missing allele matches against each pairwise cross. Tables for the pair wise comparison of each species can be found in supplementary information Tables S3.1 to S3.8. Table 3.4 summarises the best parentage matches for each polyploid species and vex2.

Table 3.4 Summary of the parentage simulation. For each study polyploid taxon the best matched pairs of potential parent genotypes are shown. The best matches have the lowest number of mismatched or missing alleles as shown in the right column. Based on the results from all analyses the most likely pairings are highlighted and discussed in 3.4.

Taxon	Parent pair X		Number of mismatched alleles
S. subcuneata	S. admonitor	S. devoniensis	0
	S. admonitor	S. margaretae	0
	S. admonitor	S. vexans	0
	S. admonitor	S. porrigentiformis	0
	S. admonitor	S. rupicola	0
	S. admonitor	S. torminalis	0
	S. admonitor	S. aria	0
	S. torminalis	S. margaretae	0
S. admonitor	S. torminalis	S. devoniensis	0
	S. torminalis	S. subcuneata	0
	S. torminalis	S. margaretae	0
S. devoniensis	S. torminalis	S. devoniensis	0
	S. torminalis	S. subcuneata	0
	S. torminalis	S. margaretae	0
S. margaretae	S. vexans	S. admonitor	3
	S. vexans	S. devoniensis	3
	S. vexans	S. subcuneata	3
	S. aria	S. rupicola	4
	S. vexans	S. rupicola	5
S. vexans	S. margaretae	S. porrigentiformis	5
	S. aria	S. rupicola	6
	S. margaretae	S. rupicola	6
	S. margaretae	S. aria	6
	S. margaretae	S. torminalis	6
Vex2	S. vexans	S. torminalis	4
	S. margaretae	S. torminalis	5
	S. vexans	S. margaretae	6
S. porrigentiformis	S. aria	S. torminalis	12
	S. aria	S. rupicola	13
	S. aria	S. vexans	13
	S. torminalis	S. vexans	13

3.4   Discussion

The parentage approach in this study has provided strong support for some of the relationships proposed in Chapter 2. The allelic configurations of the polyploid study species confirm their close associations. The many shared alleles among the south west endemics suggest they have common ancestry or are derived from each other. The parentage analysis (Table 3.4) upholds the close relationship of *S. rupicola* with the other members of the study group compared to *S. porrigentiformis* proposed in Chapter 2. It shares more alleles (Table 3.2) and is therefore a possible parent for more species (Table 3.4).

3.4.1   Origins and genome configuration of *Tormaria* members in the southwest UK

Our results confirm the findings of previous studies that subgenus *Tormaria* is derived from ancestral hybridisations between diploid *S. torminalis* and members of subgenus *Aria* (Nelson-Jones *et al.*, 2002, Chester *et al.*, 2007).

That *S. subcuneata* shares all its alleles with both *S. devoniensis* and *S. admonitor* suggests it is the common link between the two and that they are both derived from *S. subcuneata* rather than the reverse. The fact that the additional alleles found in *S. devoniensis* and *S. admonitor* match those seen in *S. torminalis* suggest that *S. torminalis* is the source of these alleles via hybridisation events. This formation of *S. admonitor* and *S. devoniensis* via a common triploid (*S. subcuneata*) hybridising with diploid *S. torminalis* places strong support for the formation of *Sorbus* tetraploids via the triploid bridge.

A less parsimonious explanation for the pairwise matches for *S. subcuneata* is from a fusion of diploid gametes from *S. devoniensis* or *S. admonitor* with

171

reduced n gametes from *S. torminalis, S. aria or* any member of subgenus *Aria.* These possibilities have less validity primarily because if either *S. devoniensis* or *S. admonitor* gave rise to triploid *S. subcuneata* they would contribute a pair of alleles at each locus along with a third allele from a contributing diploid or tetraploid. This would fail to explain how both tetraploid species share the complete set of alleles common to *S. subcuneata.* A hybrid event among any pair of tetraploids would most likely result in a further tetraploid ($2x + 2x$), rather than a triploid which would require the production of unbalanced gametes. *S. devoniensis* has been proposed as a progenitor for *S. subcuneata* (Wilmott, 1934) on morphological grounds and possibly due to its far larger distribution range (Fig. 3.3), but we dismiss this as the least likely scenario due to the above reasons. Its relatively large distribution is most likely due to fitness and a possible wider ecological niche. Seeds readily germinate and are produced in abundance most years (pers. obs.).

The combination of two gametes with double sets of chromosomes to produce a novel tetraploid AATT cannot be ruled out and the discovery of a triploid form of *S. torminalis* (Hamston *et al.*, 2015) suggests unreduced gamete formation in diploid *Sorbus* has occurred in this region, however the allelic configuration of the study members of subgenus *Tormaria* suggests this was not their direct route of origin. The production of unreduced *3x* gametes from triploids leading to the formation of allotetraploids by backcrossing is seen in other systems (Ramsey & Schemske, 1998). The maternally inherited chloroplast haplotype found in *S. admonitor* and *S. devoniensis* (Chester *et al.*, 2007) could have been inherited from either *S. torminalis* or *S. subcuneata* so it is unclear in which direction this hybridisation occurred. The prevalence of apomixis in *S. subcuneata* would suggest it was most likely the pollen donor rather than the

172

maternal progenitor species since it does produce stainable pollen, a proxy for viability (Rich, 2009).

If we accept that *S. subcuneata* gave rise to *S. devoniensis* and *S. admonitor*, *S. admonitor* can be eliminated as a progenitor of *S. subcuneata.* Therefore, our parentage analysis suggest that *S. margaretae* is the most likely parental species to have given rise to *S. subcuneata,* via an original single hybrid event with *S. torminalis*, since they all share a common pair of *Aria* derived alleles at ten of the eleven loci that amplified in *S. margaretae* (Table 3.2). No common alleles were amplified at SA08 although the presence of null alleles at this locus cannot be excluded since we were unable to resolve allele dosage at this locus. This deviates from previous hypotheses based on morphology which place *S. rupicola* as progenitor of *S. subcuneata,* and *S. aria* s.s. as progenitor for *S. devoniensis* with a doubling of chromosomes (Sell, 1989). Lemche (1999) also suggests *S. rupicola* and *S. vexans* as probable progenitors for this group.

The most likely scenario is that a reduced $2x$ pollen grain from *S. margaretae* fused with a normal reduced female gamete from *S. torminalis* (Chester *et al.*, 2007) to produce a triploid AAT, e.g. *S. subcuneata*, which is supported by the common pair of AA alleles across the group. Since *Sorbus* polyploids are pseudogamous they generally retain male fertility, thus tetraploids produce viable haploid pollen through regular microsporogenesis (Gornall, 1999). The prevalence of apomixis among these polyploids (see Section 2.4.2) would imply that pollen from an apomictic is more likely to fertilise an ovum of a sexual species than vice versa. These results are consistent with the findings of Proctor *et al.* (1989) who found that *S. admonitor* and *S. devoniensis* shared a common peroxidase phenotype, similar to *S. subcuneata* and *S. margaretae*.

173

Hajrudinović *et al.* (2015) showed that this type of interploidy cross was frequent (31% of seed from diploid *S. aria* at one site) where diploids were in the minority and pollen pressure from tetraploids was high. Therefore is seems that the successful hybridisation between diploid *S. torminalis* and *S. margaretae* was most likely where *S. torminalis* was rare, relative to *S. margaretae*. However, cross pollination experiments performed by Ludwig (2013) where pollen from tetraploids was used to pollinate sexual diploid *S. aria* resulted in pollen tube growth but only a few deformed seeds which showed a triploid profile. It may be that successful hybrid seed production may be a low proportion of overall hybridisation events.

Our findings reiterate the roles of tetraploids and triploids in the repeated cycle of polyploid formation in the Rosaceae (Robertson *et al.*, 2004b, Dickinson *et al.*, 2007, Ludwig *et al.*, 2013). The repeated hybridisation events involving sexual diploid *S. torminalis* create opportunities for new alleles to 'refresh' the polyploid gene pool. The likely sequence of hybrid events and origins for this group is shown in figure 3.6.
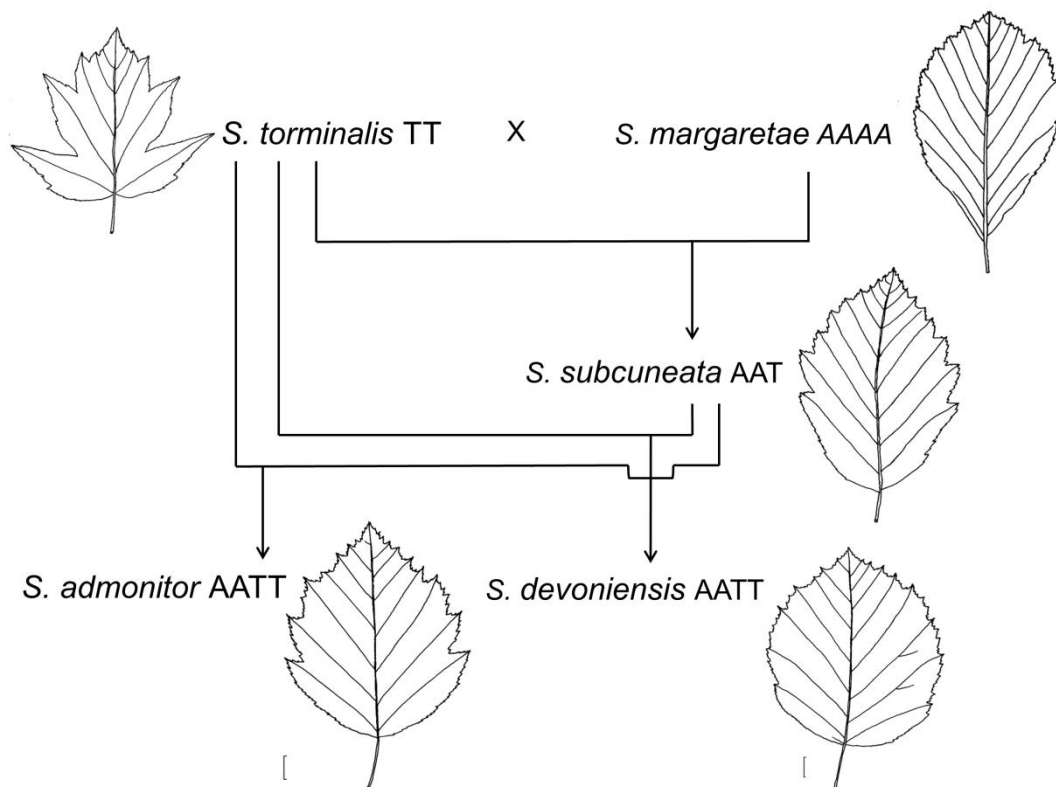
Figure 3.6 Hypothesis for the hybrid origins of the study members of subgenus *Tormaria* in the southwest UK based on parentage simulation data and inheritance patterns

3.4.2 Origins of members of subgenus *Aria* in southwest UK

We confirm the placement of *S. rupicola, S. porrigentiformis, S. vexans* and *S. margaretae* within subgenus *Aria.* Resolution of allele dosage at locus CH01F09, which only amplifies for the *Aria* genome, revealed four alleles for all four species and vex2. Therefore, we can rule out *S. torminalis* as a potential progenitor from Table 3.4. The reasons for a lack of exact matches for all members of this group are discussed below but the nearest matches can still indicate possible hybrid origins.

If we accept that *S. margaretae* gave rise to *S. subcuneata, S. devoniensis* and *S. admonitor* we must look to other members of subgenus *Aria* for the origins of *S. margaretae. S. vexans* has a comparable distribution to *S. margaretae* and they occur together on the majority of their sites. This fact and their similar morphology suggest that one species is derived from the other. The probability that *S. aria* is a progenitor for these two species cannot be ruled out as it may once have occurred in this region but the sympatric distributions of the polyploid species make them more likely contenders; however the matching of *S. aria* alleles not found in other polyploids could suggest involvement of unsampled genotypes.

S. margaretae has the lowest genetic diversity of this group (Table 2.1) and is highly endemic to this region. These features point towards a more recent origin than *S. porrigentiformis* and *S. rupicola* and possibly *S. vexans.* The closest match for *S. margaretae* is *S. rupicola* × *S. aria* (4 mismatches, Table 3.4) followed by *S. rupicola* × *S. vexans* (6 mismatches). If *S. margaretae* is more recent than *S. vexans* and possibly derived from *S. vexans* (rather than *S. aria*), the other most closely matched parent is then *S. rupicola*. This then makes *S. rupicola* × *S. aria* (6 mismatches) the most likely origin for *S. vexans.*

These relationships are more speculative than for members of subgenus *Tormaria* due to the lack of complete matches.

The second clone, vex2, was identified in Section 2.5 as a product of hybridisation rather than mutations or sexual recombination within one of the apomictic species. As its origins are within subgenus *Aria* the most likely parent pair is *S. vexans* × *S. margaretae* (6 mismatches). There are only two known trees with this genotype that occur on a small (<3ha) highly diverse site with six

176

other polyploid *Sorbus* species so it may be assumed to be of recent origin so it is surprising that no parental match could be ascertained for vex2. However, all alleles are present in the wider study group, which may suggest intermediate unsampled genotypes.

The hybridisation of two tetraploid *Sorbus* species is thought to be a rare event (Rich *et al.*, 2010). Where hybrid parentage has been determined it has involved diploid species (Robertson *et al.*, 2010), although sexual reproduction within tetraploid *Sorbus* species (Robertson *et al.*, 2004a) suggests that it may be possible especially between closely related tetraploids. Therefore, where heterospecific pollen pressure is high, which is the case for the vex2 site, cross pollination becomes more likely.

It is not known where *S. porrigentiformis* arose but it is present on the south coast of Devon (*circa* 20 trees; T. Hamston unpublished data, MSc thesis) and in smaller numbers along the north coast (Fig. 3.3), often occurring with *S. rupicola.* Within the study members of subgenus *Aria* its two closest parentage pairings of *S. rupicola* × *S. aria* and *S. vexans* × *S. aria,* however the large mis-match (Table 3.4; 13 mismatches of a total of 32 alleles) suggest it arose elsewhere.

That exact matches for the members of subgenus *Aria* could not be achieved could be due to a number or reasons such as mutations in the offspring since inception; non-sampling of progenitor variants within the sampled species; progenitors are extinct or occur outside the sampling region; the presence of null alleles or genotyping errors.

177

Mutations in study members of subgenus *Aria* that have occurred since origins most likely account for the variation in allele sizes at some loci e.g. SA06. In Chapter 2 we showed that low levels of genetic variation within *S. rupicola, S. porrigentiformis, S. margaretae* and *S. vexans* (Section 2.3.3) were attributed to mutational changes rather than recombination events. Even small mutations mean exact parental matches may not be possible even if all parental genotypes were sampled.

The likelihood of exact parental matches decreases with increasing numbers of markers; with mutations, null alleles and genotyping errors possibly leading to the false exclusion of true parents (Jones & Ardren, 2003). It is also possible that other offspring of hybridisations will be more closely related than detectable parents (Marshall *et al.*, 1998), leading to false conclusions unless the sequence of speciation events can be determined. For this reason error rates are generally used to determine the number of mis-matches allowed before excluding a single parent (Gerber *et al.*, 2000). The two phase model and generalised stepwise model are considered to be the most realistic models for microsatellites describing the occurrence of mutation steps by an absolute number which may be more than a single tandem repeat (Estoup *et al.*, 2002). All microsatellite loci used in this study are dinucleotide repeats and many of the alleles present at any one locus differ in size by two base pairs (Table 3.2). In fact, fifty eight percent of the reported 'missing alleles' reported above are two base pairs different from the nearest match (24% are 4bp removed) and only two alleles present in the endemic taxa are not found in any other sampled species. This means that mutations occurring in two or four base pair steps are likely to lead to size homoplasy between species, obscuring relationships among very similar genotypes.

178

It is unknown how old these species are; pollen analysis on Arran found non diploid *Sorbus* pollen present from *c.* 5400 B.P. (Boyd & Dickson, 1987) so it is possible that some of these south west endemics may have been in existence that length of time. Therefore, the likelihood of being able to sample the exact parental genotypes will decrease with the age of the polyploid. There is a probability that there may be missing genotype 'links' either now extinct or existing as cryptic hybrids within the current *Sorbus* distributions. The discovery of vex2, which has always been identified as *S. vexans* indicates the likelihood for such a scenario. These may include once native *S. aria* or extinct triploids acting as the bridge between diploids and tetraploid formation. Triploid *Sorbus* in common with other triploids in Rosaceae are less fertile than diploids and tetraploids when seed production is measured (Talent & Dickinson, 2005) and they may occupy a more transient position in polyploid formation particularly as some have been shown to be self-incompatible (Ludwig *et al.*, 2013).

Long distance gene flow via seed dispersal may account for the presence of alleles found in endemic polyploid species but which appear to be absent in the diploid species of the locality. The two alleles unique to the endemic polyploid species were compared to alleles of other polyploid species at Cheddar which is approximately 60 -100km east of the north Somerset coast sites. We were able to match allele 182 at locus CH02D11 (see Table 3.2) with *S. rupicoloides* (Houston et al., 2009), which also had this as a unique allele (scored as 189) among the Cheddar polyploid taxa. It also had three of the seven alleles within our study species that were not present in either *S. porrigentiformis* or *S. rupicola,* which we had matched to *S. aria* (some of which were sampled at Cheddar). Whilst this does not confirm a link it does illustrate the close relationships among *Sorbus* populations in the broader southwest region of the
179

UK. Bird mediated seed dispersal could link these south west sites particularly since *Sorbus* fruit are a known food for migrating thrush species *Turdus spp.*(Snow & Snow, 1988, Huttunen, 2004; pers.obs) as they move across the country from Scandinavia during autumn when *Sorbus* fruits are ripe.

Null alleles were not detected in those loci where all allele copy numbers were determined so this problem can be discounted in those loci. However, it is possible at other loci and the missing alleles at SA08 causing the mis-match of *S. margareate* as a parent of *S. subcuneata* seems most likely due to this as discussed above. The probability of genotyping errors were eliminated by repeating PCR amplification of inconsistent samples to ensure observed allele sizes were not artefacts of PCR amplification or scoring error.

The lack of obvious pairing of inherited alleles in subgenus *Aria* (see Table 3.2) suggest a deviation from disomic inheritance at some loci (refer to Fig. 3.1). The common pairs of *Aria* derived alleles in *S. margaretae* and *S. subcuneata* are not necessarily associated elsewhere. Whilst this observation cannot substantiate inheritance patterns due to the lack of clarity in parent/offspring pairs, progeny analysis of sexual *S. aria* × tetraploid *S. porrigentiformis* showed polysomic inheritance patterns at the only locus with parent specific alleles, MSS5 (Ludwig, 2013); both species had common alleles at the remaining loci. Polysomic inheritance might be expected in this subgenus of very closely related species where structurally similar chromosomes may exhibit the multivalent pairing characteristic of autopolyploids.

Conclusion

Evidence was obtained for the formation of hybridogenous polyploids both via the triploid bridge route and via reduced gametes from tetraploids combining

180

with those of sexual diploids to form triploid cytotypes. Therefore the role of outcrossing diploids in refreshing the genetic 'pool' of apomictic polyploid complexes is apparent. The closely related genomes of the study members of subgenus *Aria* suggest tetraploid hybridisations have occurred and similarities with those of more geographically distant members could suggest colonisation routes but also that many of the *Sorbus* polyploids in southwest England may have common polyploid ancestors.

ACKNOWLEDGEMENTS

3.5 References

Boyd WE, Dickson JH. 1987. A post-glacial pollen sequence from Loch a'Mhuilinn, north Arran: A record of vegetation history with special reference to the history of endemic *Sorbus* species. *The New Phytologist,* **107**: 221-244.

Bretagnolle F, Thompson JD. 1995. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *The New Phytologist,* **129**: 1-22.

Chester M, Cowan RS, Fay MF, Rich TCG. 2007. Parentage of endemic *Sorbus* L. (Rosaceae) species in the British Isles: evidence from plastid DNA. *Botanical Journal of the Linnean Society,* **154**: 291-304.

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

Dickinson TA, Lo E, Talent N. 2007. Polyploidy, reproductive biology, and Rosaceae: understanding evolution and making classifications. *Plant Systematics and Evolution,* **266**: 59-78.

Esselink GD, Nybom H, Vosman B. 2004. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting—peak ratios) method. *Theoretical and Applied Genetics,* **109**: 402-408.

Estoup A, Jarne P, Cornuet JM. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular ecology,* **11**: 1591-1604.

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

Gornall RJ. 1999. Population genetic structure in agamospermous plants. In: Hollingsworth PM, Bateman RM, Gornall RJ, eds. *Molecular systematics and plant evolution.* London: Taylor & Francis.

Grant V. 1981. *Plant speciation.* New York: Columbia University Press

Hajrudinović A, Siljak-Yakovlev S, Brown SC, Pustahija F, Bourge M, Ballian D, Bogunić F. 2015. When sexual meets apomict: genome size, ploidy level and reproductive mode variation of *Sorbus aria s.l.* and *S. austriaca* (Rosaceae) in Bosnia and Herzegovina. *Annals of Botany,* **116**: 301-312.

Hamston TJ, Pellicer J, Fay MF. 2015. Polyploid wild service tree: first record of a triploid *Sorbus torminalis* (Rosaceae) in Britain. *New Journal of Botany,* **5**: 34-36.

Hegarty MJ, Hiscock SJ. 2009. The complex nature of allopolyploid plant genomes. *Heredity,* **103**: 100-101.

Houston L, Robertson A, Jones K, Smith SCC, Hiscock S, Rich TCG. 2009. An account of the Whitebeams (*Sorbus* L., Rosaceae) of Cheddar Gorge, England, with description of three new species. *Watsonia,* **27**: 283.

Husband BC. 2004. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. *Biological Journal of the Linnean Society,* **82**: 537-546.

Husband BC, Schemske DW. 2000. Ecological mechanisms of reproductive isolation between diploid and tetraploid *Chamerion angustifolium. Journal of Ecology,* **88**: 689-701.

Huttunen MJ. 2004. Autumn migration of thrushes over eastern Finland: a comparison of visible migration and ringing recovery patterns. *Ringing and Migration,* **22**: 13-23.

Jones AG, Ardren WR. 2003. Methods of parentage analysis in natural populations. *Molecular Ecology,* **12**: 2511-2523.

Köhler C, Mittelsten Scheid O, Erilova A. 2010. The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics,* **26**: 142-148.

Lemche EB. 1999. The origins and interactions of British *Sorbus* species. Unpublished PhD thesis, University of Cambridge, Cambridge.

Lerceteau-Köhler E, Guérin G, Laigret F, Denoyes-Rothan B. 2003. Characterization of mixed disomic and polysomic inheritance in the octoploid strawberry (*Fragaria x ananassa*) using AFLP mapping. *Theoretical and Applied Genetics,* **107**: 619-628.

Levin DA. 1975. Minority cytotype exclusion in local plant populations. *Taxon*: 35-43.

Ludwig S, Robertson A, Rich TCG, Djordjevic M, Cerovic R, Houston L, Harris SA, Hiscock SJ. 2013. Breeding systems, hybridization and continuing evolution in Avon Gorge *Sorbus. Annals of Botany*.

Ludwig SC. 2013. Breeding systems, pollen flow and continuing evolution in Avon Gorge *Sorbus* (whitebeams, rowans and service trees)*,* PhD Thesis, University of Bristol, Bristol.

Mandáková T, Gloss AD, Whiteman NK, Lysak MA. 2016. How diploidization turned a tetraploid into a pseudotriploid. *American Journal of Botany,* **103**: 1187-1196.

Marks GE. 1966. The origin and significance of intraspecific polyploidy: experimental evidence from *Solanum chacoense. Evolution*: 552-557.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular ecology,* **7**: 639-655.

Nelson-Jones E, Briggs D, Smith A. 2002. The origin of intermediate species of the genus *Sorbus. Theoretical and Applied Genetics,* **105**: 953-963.

Nogler GA. 1984. *Gametophytic apomixis.* Berlin: Springer-Verlag.

Paun O, Fay MF, Soltis DE, Chase MW. 2007. Genetic and epigenetic alterations after hybridization and genome doubling. *Taxon,* **56**: 649-656.

Pellicer J, Clermont S, Houston L, Rich TCG, Fay MF. 2012. Cytotype diversity in the *Sorbus* complex (Rosaceae) in Britain: sorting out the puzzle. *Annals of Botany,* **110**: 1185-1193.

Proctor MCF, Proctor ME, 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.

Ramsey J, Ramsey TS. 2014. Ecological studies of polyploidy in the 100 years following its discovery. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences,* **369**.

Ramsey J, Schemske DW. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics,* **29**: 467-501.

Ramsey J, Schemske DW. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics,* **33**: 589-639.

Rich TCG. 2009. Pollen stainability in British *Sorbus* L. (Rosaceae). *Plant Ecology and Diversity,* **2**: 85-88.

Rich TCG, Houston L, Robertson A, Proctor MCF. 2010. *Whitebeams, rowans and service trees of Britain and Ireland. A monograph of British*

and Irish Sorbus L. B.S.B.I. Handbook No. 14. London: Botanical Society of the British Isles.

Richards AJ. 2003. Apomixis in flowering plants: An overview. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **358**: 1085-1093.

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

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

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

Robertson A, Rich TCG, Allen AM, Houston L, Roberts CAT, Bridle JR, Harris SA, Hiscock SJ. 2010. Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus. Molecular Ecology,* **19**: 1675-1690.

Segraves KA, Anneberg TJ. 2016. Species interactions and plant polyploidy. *American Journal of Botany,* **103**: 1326-1335.

Sell PD. 1989. The *Sorbus latifolia* (Lam.) Pers. aggregate in the British Isles. *Watsonia,* **17**: 385-399.

Snow B, Snow D. 1988. Birds and berries–A study of ecological interactions. *Poyser, London.*

Sobel JM, Chen GF, Watt LR, Schemske DW. 2010. The biology of speciation. *Evolution,* **64**: 295-315.

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, Depamphilis CW, Wall PK, Soltis PS. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany,* **96**: 336-48.

Soltis DE, Visger CJ, Marchant DB, Soltis PS. 2016. Polyploidy: Pitfalls and paths to a paradigm. *American Journal of Botany,* **103**: 1146-1166.

Soltis P. 2005. Ancient and recent polyploidy in angiosperms. *New Phytologist,* **166**: 5 - 8.

Talent N, Dickinson TA. 2005. Polyploidy in *Crataegus* and *Mespilus* (Rosaceae, Maloideae): evolutionary inferences from flow cytometry of

nuclear DNA amounts. *Canadian Journal of Botany-Revue Canadienne De Botanique,* **83**: 1268-1304.

Thompson KA, Husband BC, Maherali H. 2014. Climatic niche differences between diploid and tetraploid cytotypes of *Chamerion angustifolium* (Onagraceae). *American Journal of Botany,* **101**: 1868-1875.

Wendel JF. 2000. Genome evolution in polyploids. *Plant Molecular Biology,* **42:** 225-249.

Whitton J, Sears CJ, Baack EJ, Otto SP. 2008. The dynamic nature of apomixis in the angiosperms. *International Journal of Plant Sciences,* **169**: 169-182.

Wilmott AJ. 1934. Some interesting British Sorbi. *Proceedings of the Linnean Society of London* **146**: 73–79.

Chapter 3: Supplementary Information

The following tables show results for the pairwise match to simulate possible parentage. Species codes: sub = *S. subcuneata*, adm = *S. admonitor,* dev = *S. devoniensis,* mar = *S. margaretae,* vex = *S. vexans,* por = *S. porrigentiformis,* rup = *S. rupicola,* torm = *S. torminalis,* aria = *S. aria.*   Shaded cells show the best matches.

Table S3.1a Parentage match for *S. subcuneata* at 10 loci (24 alleles)

Most likely parent	Missing loci in parent / loci with >0% contribution
S. admonitor	0/10
S. devoniensis	0/10
S. torminalis	0/10
S. margaretae	1/10
S. rupicola	1/10
S. vexans s.s.	3/10
S. aria	3/10
S. porrigentiformis	4/10

Table S3.1b Parent pair allele matches at 10 loci for *S. subcuneata* (24 alleles).

	adm	dev	mar	vex	por	rup	torm	aria
S. admonitor	0							
S. devoniensis	0	0						
S. margaretae	0	0	7					
S. vexans s.s.	0	0	7	14				
S. porrigentiformis	0	0	7	13	19			
S. rupicola	0	0	6	9	12	13		
S. torminalis	0	0	0	4	7	6	9	
S. aria	0	0	6	11	11	7	4	14

Table S3.2a. Parentage match for *S. admonitor* at 10 loci (29 alleles)

Most likely parent	Missing loci in parent / loci with >0% contribution
S. subcuneata	0/10
S. devoniensis	0/10
S. torminalis	0/10
S. margaretae	1/10
S. rupicola	1/10
S. vexans s.s	3/10
S. aria	3/10
S. porrigentiformis	4/10

Table S3.2b Parent pair allele matches at 10 loci for *S. admonitor* (29 alleles).

	dev	sub	mar	vex	por	rup	torm	aria
S. devoniensis	4							
S. subcuneata	4	5						
S. margaretae	4	4	12					
S. vexans s.s.	4	4	11	18				
S. porrigentiformis	4	4	11	17	23			
S. rupicola	4	4	8	13	16	17		
S. torminalis	0	0	0	6	9	6	10	
S. aria	4	4	9	16	15	11	4	19

Table S3.3a Parentage match for *S. devoniensis* at 10 loci (29 alleles)

Most likely parent	Missing loci in parent / loci with >0% contribution
S. admonitor	0/10
S. subcuneata	0/10
S. torminalis	0/10
S. rupicola	1/10
S. margaretae	1/10
S. aria	3/10
S. vexans s.s.	3/10
S. porrigentiformis	5/10

Table S3.3b. Parent pair allele matches at 10 loci for *S. devoniensis* at 10 loci

(29 alleles)

	adm	sub	mar	vex	por	rup	torm	aria
S. admonitor	4							
S. subcuneata	4	5						
S. margaretae	4	5	12					
S. vexan s.s.	4	4	11	18				
S. porrigentiformis	4	4	11	14	23			
S. rupicola	3	3	9	8	15	16		
S. torminalis	0	0	0	2	7	6	9	
S. aria	2	2	8	11	13	9	4	16

Table S3.4a Parentage match for *S. margaretae* at 10 loci (24 alleles)

Most likely parent	Missing loci in parent / loci with >0% contribution
S. rupicola	0/10
S. admonitor	1/10
S. devoniensis	1/10
S. subcuneata	1/10
S. vexans s.s.	1/10
S. aria	2/10
S. torminalis	4/10
S. porrigentiformis	5/10

Table S3.4b Parent pair allele matches at 10 loci for *S. margaretae*.

	adm	dev	sub	vex	por	rup	torm	aria
S. admonitor	11							
S. devoniensis	11	11						
S. subcuneata	11	11	11					
S. vexans s.s.	3	3	3	10				
S. porrigentiformis	8	8	8	9	20			
S. rupicola	4	4	4	5	9	13		
S. torminalis	8	8	8	6	13	8	17	
S. aria	6	6	5	6	10	4	8	14

Table S3.5a Parentage match for *S. vexan s.s.* at 10 loci (29 alleles)

Most likely parent	Missing loci in parent / loci with >0% contribution
S. rupicola	1/10
S. margaretae	1/10
S. torminalis	2/10
S. admonitor	3/10
S. devoniensis	3/10
S. subcuneata	3/10
S. porrigentiformis	3/10
S. aria	3/10

Table S3.5b. Parent pair allele matches at 10 loci for *S. vexans s.s.*

	adm	dev	sub	mar	por	rup	torm	aria
S. admonitor	18							
S. devoniensis	18	18						
S. subcuneata	18	18	19					
S. margaretae	10	10	11	11				
S. porrigentiformis	10	10	10	5	16			
S. rupicola	8	8	8	6	9	12		
S. torminalis	12	12	12	6	10	10	17	
S. aria	11	11	11	6	8	6	8	14

Table S3.6a Parentage match for vex2 at 10 loci (29 alleles)

Most likely parent	Missing loci in parent / loci with >0% contribution
S. vexans s.s.	0/10
S. margaretae	0/10
S. admonitor	1/10
S. devoniensis	1/10
S. subcuneata	1/10
S. aria	1/10
S. rupicola	2/10
S. porrigentiformis	2/10
S. torminalis	3/10

Table S3.6b. Parent pair allele matches at 10 loci for vex2

	adm	dev	sub	mar	vex s.s.	por	rup	torm	aria
S. admonitor	16								
S. devoniensis	14	15							
S. subcuneata	15	14	17						
S. margaretae	11	10	12	12					
S. vexans s.s.	8	7	8	6	12				
S. porrigentiformis	11	10	11	7	10	19			
S. rupicola	13	12	13	9	10	17	20		
S. torminalis	7	7	7	5	4	9	12	15	
S. aria	10	9	10	7	7	7	10	7	14

Table S3.7a Parentage match for *S. porrigentiformis* (32 alleles)

Most likely parent	Missing loci in parent / loci with >0% contribution
S. rupicola	1/10
S. torminalis	2/10
S. aria	3/10
S. vexans s.s.	3/10
S. margaretae	5/10
S. devoniensis	5/10
S. admonitor	5/10
S. subcuneata	6/10

Table S3.7b. Parent pair allele matches at 10 loci for *S. porrigentiformis.*

	adm	dev	sub	mar	vex s.s.	rup	torm	aria
S. admonitor	26							
S. devoniensis	26	26						
S. subcuneata	26	26	27					
S. margaretae	23	23	24	24				
S. vexans s.s.	18	18	18	18	19			
S. rupicola	17	17	17	17	15	18		
S. torminalis	18	18	18	16	13	14	20	
S. aria	16	16	16	15	13	13	12	19

Chapter 4: Polyploid wild service tree: first record for a triploid *Sorbus torminalis* in the UK

T. J. Hamston*[1], J. Pellicer[2] and M. F. Fay[2,3]

[1]*Biosciences, College of Life and Environmental Sciences, University of Exeter, EX4 4QD, UK,*

[2]*Jodrell Laboratory, Royal Botanic Gardens, Kew, TW9 0ED, UK*

[3]*School of Plant Biology, University of Western Australia, Crawley, WA 6009, Australia*

An earlier version of this paper was published in the New Journal of Botany

TH collected the samples, JP and TH carried out the flow cytometry, TJH wrote the paper with comments from JP & MFF

194

Abstract

The genus *Sorbus* is known for its complex taxonomy involving many polyploid species. However, *Sorbus torminalis* has long been assumed to be uniformly diploid. An analysis of DNA content using flow cytometry revealed a triploid individual within 0.5 km of tetraploid *S. devoniensis* at South Tawton, Devon. A hypothesis based on the leaf morphology and genetic genotypes described in Chapter 2, suggests the novel *Sorbus* to be a spontaneous *S. torminalis* triploid rather than the result of interspecific hybridisation.

Key words: evolution, flow cytometry, polyploid, *Sorbus*, triploid.

4.1 Introduction

Polyploidy is widely accepted to be an important factor in the evolution of angiosperms and the spontaneous formation of novel species (Grant, 1981; Soltis *et al.*, 2009). Polyploidy can arise in several ways but occurs most commonly via the fertilisation of unreduced gametes containing the somatic chromosome number (Ramsey & Schemske; 1998; Köhler *et al.*, 2010). This process can result in triploid progeny that may, in turn, give rise to further polyploids through back-crossing or via hybridisation events involving other polyploid taxa (Husband, 2004).

In the British Isles the genus *Sorbus* is represented by four sexually reproducing diploid species – *Sorbus aria* (L.) Crantz*, S. torminalis* (L.) Crantz*, S. aucuparia* L. and *S. domestica* L. plus, at least 30 recognised polyploid species that primarily reproduce asexually via apomixis (Robertson *et al.*, 2010; Ludwig *et al.*, 2013). Within this group there are high levels of endemism, often represented as small populations or even a few individuals restricted to one or few sites e.g. *Sorbus leyana* Wilmott (Rich *et al.*, 2010). This complex array of morphologically diverse taxa arose from hybridisation events involving crossings between the diploid species and other polyploid apomicts (Rich *et al.*, 2010), resulting in a taxonomically intricate polyploid network (Pellicer *et al.*, 2012).

Chromosome numbers can provide useful information about the origin of certain species and the mechanism of speciation. In this case, chromosome counts have been carried out on some of the *Sorbus* taxa with diploid ($2n=34$), triploid ($2n=51$) and tetraploid ($2n=68$) individuals found among the genus (Bailey *et al.*,

2008). The use of flow cytometry to estimate the relative DNA content, and hence to infer DNA ploidies, has recently been used to analyse the cytotype diversity in a large-scale survey of UK *Sorbus* species (Pellicer *et al.*, 2012). This study revealed the high incidence of polyploidy in the genus ($3x$, $4x$ and $5x$), while also confirming the constancy of $2x$ cytotypes in parental taxa in the British territory, including all samples of *S. torminalis* evaluated. Although there have been previous reports of polyploid *S. torminalis* from the Balkans and Spain (Aldasoro *et al.*, 1998; Siljak-Yakovlev *et al.*, 2010) (tetraploid and triploid, respectively), no such plants have ever been reported from the UK. It has been suggested that these individuals may be of hybrid origin but this hypothesis has yet to be confirmed (Pellicer *et al.*, 2012).

4.2 Methods and study species

In Devon, *S. torminalis* occurs largely as a relatively widespread but infrequent hedgerow tree, occasionally growing alongside *Sorbus devoniensis* E.F.Warb., a tetraploid species endemic to the region. Indeed, genetic analyses using nuclear microsatellites (see Chapters 2 & 3) and chloroplast microsatellites (Chester *et al.*, 2007) have suggested that *S. torminalis* is a maternal parental taxon of *S. devoniensis*.

During a study of the polyploid *Sorbus* species endemic to Devon and Somerset, we assessed ploidy levels of 106 samples from six polyploid species using flow cytometric analyses at the Jodrell Laboratory (RBG, Kew). We initially included two samples of *S. torminalis* for comparison, but a significant deviation in the relative DNA content of one of these samples prompted a wider collection of *S. torminalis* samples from Cornwall and Devon. Fresh leaf material was collected from 27 individuals during September 2012, herbarium

vouchers being deposited in the Welsh National Herbarium in Cardiff (NMW).

Locations and voucher accession numbers can be found in supplementary

information, Table S2.1. The absolute nuclear DNA content was measured

using propidium iodide flow cytometry, following the method described by

Pellicer *et al*. (2012). DNA ploidy levels of all samples were determined by

comparing their fluorescence profile with that of *Oryza sativa*, which was used

as an internal standard of predetermined genome size (Fig. 4.1).

4.3   Results

The fluorescence peak ratios between the *O. sativa* and diploid *S. torminalis*

individuals were fairly constant (R = *ca.*1.58). However, one sample from a

suckering individual in a tightly flailed hedgerow near South Tawton, Devon

(SX6693) presented a ratio of R = 2.29, approximately 1.5 × the relative DNA

content of the diploid samples, thus indicating that it was a triploid cytotype

($2n = 3x = 51$). To illustrate the results, a flow cytometric histogram from a

combined run that included a diploid *S. torminalis*, tetraploid *S. devoniensis* and

the suspected triploid *S. torminalis* is shown in Figure 4.1. Also see flow

cytometry results in Chapter 2; Table 2.2.
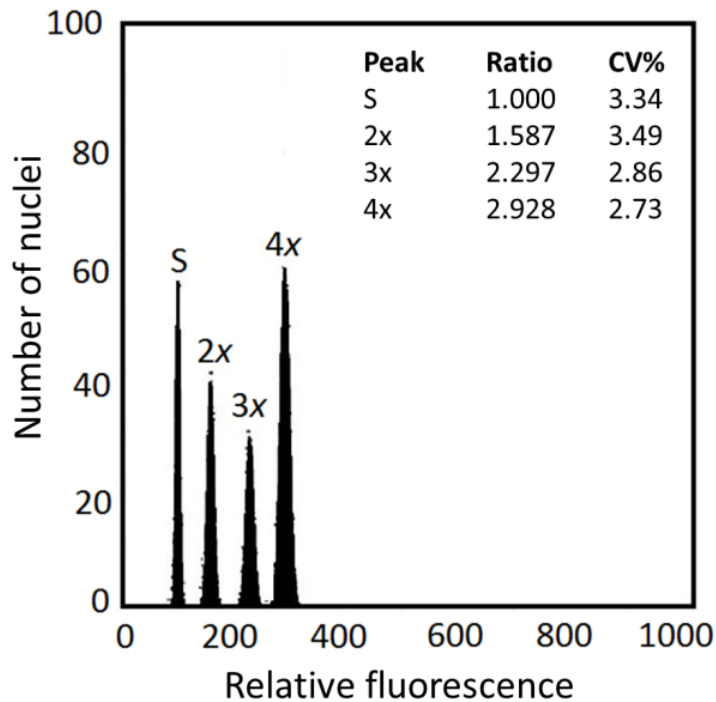
Figure 4.1. Flow cytometric histogram of a combined sample of diploid *S. torminalis* (2*x*), triploid *S. torminalis* (3*x*), tetraploid *S. devoniensis* (4*x*) and the internal standard *Oryza sativa* (S).

A detailed search of the hedgerows within 1 km radius of the triploid individual failed to find further specimens of *S. torminalis.* The nearest confirmed record, approximately 5 km away, was included in the sampling, with a diploid result. However, several tetraploid *S. devoniensis* trees occur within 400 m, so one plausible hypothesis is that this triploid individual was the result of a back-cross between *S. devoniensis* and *S. torminalis.* However, Figure 2 shows the leaf with characteristic deep triangular lobes, consistent with *S. torminalis* (Rich *et al.*, 2010). Table 4.1 shows data compiled from the study in Chapter 2 at the 9 loci where alleles in the triploid form amplified. The triploid form has common alleles with both *S. devoniensis* and other diploid *S. torminalis* at 6 loci, but 3 loci show no alleles common with *S. devoniensis.*

199

4.4   Discussion

The occurrence of a triploid *S. torminalis* individual has implications for the production of further polyploid *Sorbus* species in this geographic area, assuming that it is able to produce viable gametes. It is within pollination distance of *S. devoniensis* and possibly unrecorded specimens of *S. torminalis.* Polyploids could result from a union of reduced (n) and unreduced ($2n$ and $3n$) gametes of these species if sexual outcrossing were to occur. However, given the predominance of apomictic reproduction in *Sorbus* polyploids, hybridisation is not a common occurrence (Rich *et al.*, 2010). The genetic data suggests a hybrid origin of *S. torminalis x S. devoniensis* is unlikely since the triploid has no common alleles with *S. devoniensis* at 3 of the 9 loci and all alleles in common with *S. devoniensis* are also found in the wider populations of *S. torminalis*. Grafted material grown on at Paignton Zoo will eventually enable us to investigate the pollen viability and breeding system of this plant and thus further evaluate this likelihood.
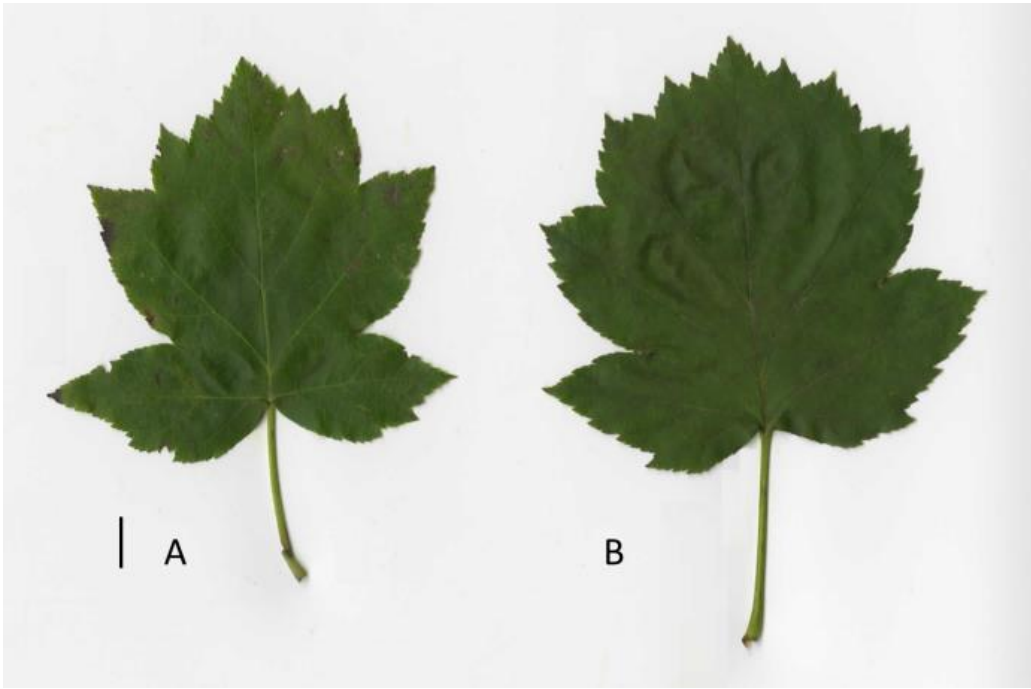
Figure 4.2. View of adaxial leaf surface of (a) diploid *S. torminalis* (sampled from Spreyton, Devon SX7297) and (b) triploid *S. torminalis*, showing broadly similar leaf morphology. Scale bar 1 cm.

Table 4.1 Genome composition of *S. devoniensis* and *S. torminalis* populations at 9 loci showing the alleles for the triploid form. X = ploidy level. See Chapter 2 for methods. [green] = Common alleles

Microsatellite loci

Taxon	X	CH02D11				MSS13				CH01F02				MSS16			
S. devoniensis	4	**152**	**162**	182		**189**	193	195	**193**	**187**	187	195	199	158	160	198	204
S. torminalis	2	148	150	**152**	154	181	183	187	**189**	157	167	175	**187**	154	166	170	178
		156	**162**	164	170	191	**193**	195	197	189	209			182	**184**	186	188
		172	176	178	194									190	194	196	198
		196												200	202	204	206
														208	210	216	222
S. torminalis	3	**152**	**162**			**189**	**193**			**187**				**184**	**190**		

Taxon	X	SA01				MSS5				SA14				SA08				MS14			
S. devoniensis	4	224	232	234	242	119	121	**123**	127	170	**204**	208	226	261				**123**	133		
S. torminalis	2	190	192	212	216	105	113	117	119	170	176	178	180	229	232	259	260	122	**123**	125	127
		226	**230**	234	**236**	**123**	125	127	129	182	184	186	188	261	265	267	269	129	**131**	**133**	135
		238	240	242	**244**	135	**137**	139	141	190	198	**200**	202	271	273	277	281				
		246	256							**204**	206	208	210	**282**	**284**						
										212	214	224	226								
S. torminalis	3	**230**	**236**	**244**		**123**	**137**			**200**	**204**			**282**	**284**			**123**	**131**		

References

Aldasoro, J. J., Aedo, C., Navarro, C. & Garmendia, F. M. 1998. The genus *Sorbus* (Maloideae, Rosaceae) in Europe and in North Africa: morphological analysis and systematics. *Systematic Botany,* 23**:** 189–212.

Bailey, J., Kay, Q., McAllister, H. & Rich, T. 2008. Chromosome numbers in *Sorbus* L. (Rosaceae) in the British Isles. *Watsonia,* 27**:** 69–72.

Chester, M., Cowan, R. S., Fay, M. F. & Rich, T. C. G. 2007. Parentage of endemic *Sorbus* L. (Rosaceae) species in the British Isles: evidence from plastid DNA. *Botanical Journal of the Linnean Society,* 154**:** 291–304.

Grant, V. 1981. *Plant speciation,* 2nd ed. New York: Columbia University Press.

Husband, B. C. 2004. The role of triploid hybrids in the evolutionary dynamics of mixed- ploidy populations*. Biological Journal of the Linnean Society*, 82(4): 537–546.

Köhler, C., Scheid, O. M. & Erilova, A. 2010. The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics,* 26**:** 142–148.

Ludwig, S., Robertson, A., Rich, T. C. G., Djordjevic, M., Cerovic, R., Houston, L., Harris, S. A. & Hiscock, S. J. 2013. Breeding systems, hybridization and continuing evolution in Avon Gorge *Sorbus. Annals of Botany*, 111: 563–575.

Pellicer, J., Clermont, S., Houston, L., Rich, T. C. G. & Fay, M. F. 2012. Cytotype diversity in the *Sorbus* complex (Rosaceae) in Britain: sorting out the puzzle. *Annals of Botany,* 110**:** 1185–1193.

Ramsey, J. & Schemske, D. W. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics,* 29**:** 467–501.

Rich, T. C. G., Houston, L., Robertson, A. & Proctor, M. C. F. 2010. *Whitebeams, rowans and Ssrvice trees of Britain and Ireland. A monograph of British and Irish* Sorbus *L.,* 1st ed. London: Botanical Society of the British Isles.

Robertson, A., Rich, T. C. G., Allen, A. M., Houston, L., Roberts, C. A. T., Bridle, J. R., Harris, S. A. & Hiscock, S. J. 2010. Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus*. *Molecular Ecology, 19*(8): 1675–1690.

Siljak-Yakovlev, S., Pustahija, F., Šolić, E. M., Bogunić, F., Muratović, E., Bašić, N., Catrice, O. & Brown, S. C. 2010. Towards a genome size and chromosome number database of Balkan flora: C-Values in 343 taxa with novel values for 242. *Advanced Science Letters, 3***:** 190–213.

Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., Sankoff, D., DePamphilis, C. W., Wall, P. K. & Soltis, P. S. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany,* 96(1): 336–348.

Chapter 5: Breeding system and spatial isolation from congeners strongly constrain seed set in an insect-pollinated apomictic tree: *Sorbus subcuneata* (Rosaceae)

*Tracey J. Hamston[1, 2], Natasha de Vere[3, 4], Tim C. G. Rich[5], Jamie R. Stevens[1] and James E. Cresswell[1]

[1]Biosciences, University of Exeter, Exeter, Devon, U.K .

[2] Whitley Wildlife Conservation Trust, Paignton Zoo, Paignton, U.K.

[3] National Botanic Garden of Wales, Llanarthne, U.K.

[4] Institute of Biological Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, U.K.

[5] 57, Aberdulais Road, Cardiff, U.K.

A later version of this paper has been reformatted and accepted with revisions in Scientific Reports

205

Abstract

1. In plants, apomixis results in the production of clonal offspring via seed and can provide reproductive assurance for isolated individuals in fragmented populations. However, many apomicts require pollination to fertilise their endosperms for successful seed set (pseudogamy) and therefore risk pollination-limitation, particularly in self-incompatible species that require heterospecific pollen.

2. We investigated pollen-limitation in *Sorbus subcuneata* (slender whitebeam)*,* a threatened endemic tree of conservation concern that co-occurs with its congener, *S. admonitor,* in the southwestern United Kingdom. We used microsatellite analysis of paternity and hand pollinations to investigate its breeding system and pollination ecology.

3. We confirmed that *S. subcuneata* is an obligate apomict, because all embryos studied had identical genotypes comprising maternal alleles. In woodland, open-pollinated flowers of *S. subcuneata* rarely produced seed (flower-to-seed conversion < 1%) even though they rapidly accumulated pollen on their stigmas.

5. Manual self-pollination rarely produced seed (<3% flower-to-seed conversion), but manual heterospecific pollination by *S. admonitor* resulted in a high flower-to-seed conversion rate (65%). However, paternity from *S. subcuneata* and *S. admonitor* was almost equally represented among the endosperms of seeds from open-pollinated flowers, so we estimate that the ratio of self: congeneric pollination in open-pollinated flowers was at least 22:1.

6. Despite the efficacy of heterospecific pollination, the contribution of *S. admonitor* trees to paternity in seed from open-pollinated flowers of *S.*
206

subcuneata decreased rapidly with the spatial separation between paternal and maternal trees.

Synthesis. Our study indicates that seed set in this principally self-incompatible pseudogamous apomict was limited by the spatial proximity of its congener. Conservation efforts aimed at maintaining species with this breeding system must therefore manage the distributions of congeners in tandem. Management of this kind will also maintain the potential for rare heterospecific fertilisation, thereby also preserving the evolutionary processes that typically cause rapid diversification in these lineages.

Key-words

Apomixis, connectivity, conservation, paternity analysis, pollen limitation, pollination, polyploidy, pseudogamy, reproductive ecology, self-incompatibility.

5.1   Introduction

Asexual reproduction through clonal seed (apomixis) offers reproductive assurance to isolated individuals (Richards, 2003) and is often associated with colonising species that have wider geographic distributions than their sexual counterparts (van Dijk, 2003). Apomixis is widespread among angiosperm families (Campbell & Dickinson, 1990), but occurs more frequently in the Asteraceae, Poaceae and Rosaceae (Bicknell & Koltunow, 2004; van Dijk & Vijverberg, 2005), where it is associated with polyploidy (Whitton *et al.*, 2008). Many apomicts are also pseudogamous, which means that they require pollen to develop functional endosperm for the maturation of their otherwise clonal seed. For species with a pseudogamous apomictic (PA) breeding system, the requirement for pollination may limit seed set just as it often does in sexual species (Burd, 1994). The degree of limitation can be measured by the increase in seed produced when pollen is added to the stigma by hand (Knight *et al.*, 2005), but the ecology of pollen-limited seed set in PA species has rarely been investigated.

Pollen-limited seed set can have various causes that include both low quantity and low quality of pollen available to females (Wilcock & Neiland, 2002). Factors that limit the quantity of pollen delivered to stigmas include flowering asynchrony between males and females and inadequate service from pollen vectors (insects or wind). Pollen quality limits seed set when too much of the pollen that reaches a female's stigmas is incompatible. In the Rosaceae, many polyploid PA's are self-compatible (Dickinson *et al.*, 2007), which enables seed set through autogamous (within flower) or geitonogamous (within individual)

pollen transfer. Curiously, however, some PA species in the Rosaceae are self-incompatible (Ludwig *et al.*, 2013), which exposes them to the pressures of pollen-limitation without conferring any obvious adaptive benefits. We investigated the performance of this perplexing breeding system in a woodland community of rare, putatively self-incompatible triploid *Sorbus* species.

The genus *Sorbus* L. comprises small and medium-sized trees that produce corymbs of showy, hermaphrodite flowers during late spring and early summer. In common with many other members of the Rosaceae (Guitián *et al.*, 1993; Pías & Guitián, 2006), the floral architecture in *Sorbus* is entomophilous and its flowers attract generalist flower-visiting insects, mainly bees (Ludwig, 2013). Diploid *Sorbus* species are typically self-incompatible out-crossers (Oddou-Muratorio *et al.*, 2005; Pías & Guitián, 2006), but *Sorbus* also contains apomictic polyploids derived from hybridisation (Campbell & Dickinson, 1990; Rich *et al.*, 2010). Recent speciation is evident in *Sorbus* (Robertson *et al.*, 2004; Rich & Proctor, 2009; Robertson *et al.*, 2010) and was likely favoured by breeding systems where apomixis is facultative and possibly coupled with triploid self-incompatibility (Ludwig *et al.*, 2013). This infrequent sexual route for gene exchange among otherwise clonal species provides the raw variation for adaptation (Nogler, 1984), whilst apomixis maintains the new gene combinations and enables sympatric speciation (van Dijk & Vijverberg, 2005). Where clonal self-incompatible PA breeding systems exist, females therefore require heterospecific cross-pollination for seed set (the males are congeners). Consequently, the proximity of suitable mates in both space and time may constrain seed production.

In *Sorbus*, many polyploid species have small population sizes and a high degree of endemism, which are both features of conservation priority species (IUCN, 2001). Threats to the persistence of rare *Sorbus* species arise principally from changing land use, browsing by herbivores (which prevents recruitment) and competition from invasive non-native plant species (Rich *et al.*, 2010). Conserving this evolutionarily dynamic group relies on understanding the factors that affect population viability, which include the influence of pollination on seed production. In order to measure the extent to which pollen limitation constrains seed production in a rare PA species in *Sorbus*, we investigated the pollination and breeding system of the slender whitebeam, *Sorbus subcuneata* Wilmott, which is putatively self-incompatible.

Sorbus subcuneata is a rare triploid species endemic to nine sites in Devon and Somerset (southwestern United Kingdom), where it co-occurs with six closely related tetraploid congeners and the common diploid *S. aucuparia* L. (Proctor *et al.*, 1989; Rich *et al.*, 2010). *Sorbus subcuneata* typically occurs at low relative density as an understorey tree, so spatial isolation may be a constraint on seed production. Here, we report an investigation of the breeding system and pollination ecology of *S. subcuneata* in which we determined: (1) the species' pollination requirements (i.e. compatibility with conspecific and heterospecific pollen); (2) the extent to which seed set is limited by pollination; and (3) the factors that imposed pollination-limited seed set. Specifically, we evaluated whether pollination limitation arose through either temporal isolation imposed by flowering asynchrony between compatible pollination partners or by spatial isolation of females from suitable male pollen donors.

5.2   Materials and methods

5.2.1   Study system and sites

We studied a natural population of the triploid *Sorbus subcuneata* system in an area of ancient woodland dominated by sessile oak (*Quercus petraea* (Matt.) Liebl.) at Watersmeet, Devon (-3.7975, 51.2243 WGS84). At this site, the *S. subcuneata* population of *c.* 300 trees occur together with *c.*100 trees of tetraploid *S. admonitor* M.C.F. Proctor (Rich & Cann, 2009) and individuals grow both as stunted specimens on thin soils and rocky outcrops and also in woodland as the understorey below an oak canopy. These two species occur as scattered individuals along with relatively abundant *S. aucuparia*, but they have somewhat disjunct distributions chiefly on opposite sides of a steep river valley (see Results, Fig. 5.1). *Sorbus subcuneata* is closely related to *S. admonitor* (Proctor *et al.*, 1989), which is likely to be a source of compatible pollen for *S. subcuneata*. *S. aucuparia* is a distantly related diploid species, but also may provide compatible pollen because it pollinates other polyploids within the *Sorbus* genus (Robertson *et al.*, 2004; Rich *et al.*, 2010). *Sorbus admonitor* and *S. aucuparia* appeared to be the most likely sources of cross-pollination for *S. subcuneata*, so we focussed our study on these three species, although very small numbers of three other tetraploid *Sorbus* species are also found on site: *S. porrigentiformis* E.F. Warb.; *S. margaretae* M.C.F. Proctor and *Sorbus vexans* E.F. Warb*.* All study species were morphologically distinguishable by leaf and fruit features. The trees in the study were tagged with a unique identifying number and location details were recorded using maps and a hand held GPS unit. Leaf samples were collected and herbarium vouchers placed at the National Museum of Wales (NMW), Cardiff. GPS readings of sample tree

locations are included in Table S5.1 in Supplementary Information. Leaf or seed samples used for genetic analysis were collected from the study sites, dried and stored in silica gel until use. For logistical convenience, we conducted hand-pollinations to confirm compatibility among the three study species on individual trees held in a collection belonging to the Exmoor Natural History Society (Luckbarrow, West Luccombe, Somerset, UK).

5.3.2   Investigation of the breeding system of triploid *Sorbus subcuneata*

In order to test whether *S. subcuneata* requires pollen for apomictic seed production (pseudogamy) and to compare its pollen compatibility with the two focal congeners, we used two individuals of *S. subcuneata* as maternal trees to conduct four pollination treatments as follows: (1) pollen added from the same species to test for conspecific compatibility; (2) pollen added from *S. admonitor* (4x); (3) pollen added from *S. aucuparia* (2x); (4) no pollen added to test for autonomous apomixis. Microsatellite analysis confirmed that these trees are representative of the type genotype, and are representative of the species (Table S5.2). All maternal *S. subcuneata* flowers were emasculated on opening (anthesis) and pollen supplementation was conducted 24 hours later to allow stigma maturation. Twenty inflorescences on each of the two maternal trees were randomly assigned among the four pollination treatments. To prevent pollination occurring before hand-supplementation, we excluded insect pollinators by placing a mesh bag over each inflorescence when flower buds were beginning to 'balloon' just prior to anthesis. At this stage, all but 10 buds were removed; those remaining were used for the pollination experiment. To reduce the potential effect of resource limitation on seed set, all other inflorescences were removed from the experimental branches. Pollen donor

212

inflorescences were bagged prior to pollen collection in order to ensure sufficient was available for the experiment and to avoid potential contamination from other species. Hand pollination was achieved by gently wiping fully dehisced donor anthers across the receiving stigma until pollen was clearly visible on the stigma surface. The pollen exclusion bags were then replaced to prevent natural pollination. The bags were removed after one week to allow fruits to develop. In order to assess the levels of fruit abortion in each treatment, fruit set was recorded at three intervals during July, September and finally in October when the fruit was ripe. Mesh bags were placed over the infructescences to prevent predation and fruit loss. Successful pollination of a flower was measured by the production of at least one seed and calculated as a flower-to-seed conversion rate. Where there was a failure to produce seed in any treatment, we used the binomial theorem to determine the maximum successful conversion rate that would be statistically consistent with this observation at this sample size. The binomial function can be used to calculate $P(x)$, the probability of x successes in n trials. If the probability of success in a single trial is denoted p, then $P(x)$ is given by:

$$P(x) = \frac{n!}{x!\,(n-x)!}\, p^x (1-p)^{n-x}$$

Eq. 1

Thus, given the condition that $x = 0$, we solved Eq. 1 for the maximum value of p such that $P(x) \geq 0.05$, which identified the greatest success rate at which n flowers producing zero seed is not statistically significant at the conventional level. In effect, we determined the upper 95% confidence interval on p. In

213

addition, we also calculated the overall flower-to-seed conversion rate of naturally pollinated flowers from three trees on the woodland site.

To verify that *S. subcuneata* is an obligate apomict and to determine whether interspecific pollen flow contributes to endosperm fertilisation, we also investigated breeding patterns in the natural population. To this end, we used nuclear DNA microsatellite markers to genotype the embryo and endosperm of wild-collected, naturally-pollinated seed. Seed embryos produced by apomixis will have a genotype identical to that of the maternal tree. Pollen donors were identified through the presence of their unique alleles in a seed's endosperm and the proportion of seed resulting from pollination by each congener was calculated.

Due to the large quantities of small, parthenocarpic fruit that remained on the trees into autumn, we restricted our study to fruits of $\geq$ 9 mm diameter that contained seed because the hand pollination experiment had demonstrated that smaller fruits were invariably seedless (Table 5.1). During a systematic search of the woodland, large fruits were collected from various trees in two fruiting seasons (2013 and 2014). Seed was extracted from fruits, air dried and stored in silica gel until use. Before dissection, seeds were soaked for 24 hours in deionised water on filter paper to soften them before the seed coat was removed, which enabled the embryo and endosperm to be cleanly separated under a dissecting microscope. All *Sorbus* taxa within 2 km of the study population were subjected to genetic analysis to provide reference samples for comparison genotypes. Since every polyploid *Sorbus* species on the study site is clonal, microsatellite alleles could be matched only to species, not individuals.

Microsatellite analyses were conducted as follows. In order to discriminate among all species sampled, we selected ten nuclear microsatellites previously used for *Sorbus* taxa (Table S5.2). DNA extraction of leaf and seed material followed QIAGEN DNeasy Plant Minikit protocol with lysis buffer added to samples prior to disruption using a QIAGEN TissueLyser bead mill set at 30 Hz for 4 minutes. The primer pairs were combined into three multiplex reactions and a touchdown polymerase-chain-reaction was carried out in a MyCycler thermal cycler (Bio-Rad, California, USA) according to the following cycling program: 95°C for 5 min; 4 cycles of 95°C for 30 sec, 62°C for 1 min 30 sec and 72°C for 3 min, followed by 4 cycles at decreased annealing temperature of 58°C; 7 cycles at 55°C annealing temperature; 12 cycles at 53°C annealing temperature followed by 3 further sets of 5 cycles at decreasing annealing temperatures in increments of 2°C, and final extension at 72°C for 10 min. The amplified products were analysed using CEQ 8000 Genetic Analysis system (Beckman Coulter, Fullerton, CA, USA). To identify the pollen donor, identity of each species was confirmed by comparing multilocus genotypes of each species with the alleles present in the endosperm tissue (Table S5.2). Only samples that successfully amplified across all loci were included in the endosperm paternity analysis with the exception of locus MS14, which only amplified in *S. admonitor, S. subcuneata* and *S. aucuparia.*

In order to estimate a conspecific: heterospecific pollination ratio we combined the proportion of seed resulting from pollination by each congener with the flower-to-seed conversion rates of the hand pollinated flowers. The estimated ratio of self: outcross pollination, denoted R, was calculated by:

215

$$R = \frac{S}{C_s} \bigg/ \frac{A}{C_a}$$

Eq. 2

Where S = total proportion of seed sired by *S. subcuneata,* C_s = the flower-to-seed conversion rate for *S. subcuneata* pollination, A = total proportion of seed sired by *S. admonitor*, and C_a = the flower-to-seed conversion rate for *S. admonitor* pollination. *S. admonitor* was the only species to pollinate

5.2.3 Testing pollen limitation of seed set in the natural population

In order to investigate whether pollen quantity limited seed production in the woodland population of *S. subcuneata,* we recorded pollen deposition on the stigmas of target flowers over their three-day blooming period. Pollen deposition was measured among flowers on two individuals of *S. subcuneata* over six days from the onset of the trees' flowering until the peak. The trees were visited each day between 12.00 and 16.00hrs and the calyces of buds about to open were uniquely colour-marked and their stigmas were subsequently collected at durations of up to 3 days post-anthesis. The collected flowers were carefully emasculated to prevent pollen transfer and each flower's stigma was placed in a 1.5ml collection tube containing damp cotton wool in the base to enable pollen germination. The stigmas were refrigerated at 4°C and frozen within 3 days of collection. To determine the amount of pollen on each stigma, a squash preparation was made by softening each sample in sodium hydroxide (8 M NaOH for 10 mins. at 60°C) before counting the number of pollen grains at $\times$ 10 magnification. The accumulation of pollen on the flowers of both trees was modelled using a negative binomial generalised linear model (GLM) with square-root link function, which allowed for over-dispersed data. We

216

tested the effect of tree and flower age on pollen counts using likelihood ratio tests. This analysis was implemented in the statistical software R (R Development Core Team, 2015).

To investigate whether seed production was limited by pollen quality, we tested whether the application of supplementary compatible pollen increased seed production in open-pollinated flowers of woodland *S. subcuneata*.   Based on the outcome of compatibility testing (see 5.3.1), we used tetraploid *S. admonitor* as the most favourable male donor. On two maternal trees of *S. subcuneata,* ten inflorescences were selected and each was randomly assigned to one of two treatment groups. Group 'O' was left to open pollinate naturally whereas group 'H' had heterospecific pollen applied from mature, dehisced anthers that were collected from *S. admonitor.* The anthers were wiped gently across the stigmas of five open flowers per inflorescence. Both groups of inflorescences were left to develop fruit, which was collected during October and checked for the presence of seeds. We assessed the extent of the pollen deficit by comparing the flower-to-seed conversion rate between the H and O treatments.

To investigate whether temporal isolation (asynchronous flowering) could limit pollen flow between maternal trees of *S. subcuneata* and their congeneric pollinator, *S. admonitor,* we quantified the flowering periods of six trees of each species, which were observed directly using binoculars at two day intervals throughout the 17-day flowering period during May to estimate the proportion of flowers in bloom. On each occasion, we estimated the proportion of flowers in bud, fully open or in senescence, which was evident because the anthers and petals turned visibly brown. To test for flowering asynchrony between *S. subcuneata* and *S. admonitor,* the relationship of the cumulative proportion of

217

opened buds on each tree over time was fitted by a sigmoid curve, which we used to estimate the time to 50% of flowers open, denoted F_{50} (Fig. S5.1). Variation in F_{50} between the two species was analysed using a Wilcoxon Rank-Sum test. For each species, we also determined the mean flowering day (MFD), Eq 3; (Barbour *et al.*, 2006), in which % Fl_i is the mean percent of the flowers at anthesis on observation day i and $Days_i$ is the number of days elapsed at day$_i$ from the start of the observation period:

$$MFD = \frac{\sum_{i=1}^{n}(Days_i \times \%Fl_i)}{\sum_{i=1}^{n}\%Fl_i}$$

Eq. 3

To investigate whether spatial isolation (i.e. spatial separation of maternal trees from compatible pollen donors) could limit seed set in *S. subcuneata*, we used the genetic paternity data (see microsatellite analysis above) and correlated the proportion of seed attributable to *S. admonitor* with the maternal tree's 'connectivity' to all potential *S. admonitor* males. The connectivity (S_i) for each maternal tree i to every potential paternal tree j was determined by:

$$S_i = A_i^c \sum \exp(-\alpha d_{i,j}) A_j^b$$

Eq. 4

where α modulates the effect of distance between trees i and j, denoted $d_{i,j}$, in a negative exponential decay kernel; A_i is the size of maternal tree I; A_j is the size of pollen donor j; and b and c are exponents that scale the impact of tree size on pollen export and import (Moilanen & Nieminen, 2002). Connectivity analysis was executed using the SI software (Moilanen, 2000). To implement the model, we set $\alpha = 0.01$ as extrapolated from pollen dispersal patterns
218

measured for *S. torminalis* (Oddou-Muratorio *et al.*, 2005) and we assumed that connectivity was unaffected by the size of the maternal tree because female function is typically satisfied by few pollinator visits (Bell, 1985) (i.e. $c = 0$) and that male success saturated with increasing tree size ($b = 0.5$) (Oddou-Muratorio *et al.*, 2005). Distances and tree size were based on raw data collected during a comprehensive survey that had measured the association between size and fruiting for all the polyploid *Sorbus* species on the study site (Rich & Cann, 2009). Tree size was estimated as crown area, which was derived from a measurement of diameter at breast height (DBH) using the standard forestry relationship for the architecturally similar congener, *S. aucuparia* (Hemery *et al.*, 2005). In the analysis, we included only trees of flowering size (trees >3 cm DBH for *S. admonitor* and >2 cm DBH for *S. subcuneata*).

Using values calculated from Eq. 4, we tested whether the proportion of seed derived from heterospecific pollen (here *S. admonitor*) produced by each maternal tree (% of total seed) was explained by the connectivity of the maternal seed trees to the most effective pollen donor. To test this hypothesis, we used a GLM with a binomial error and probit link function in which the contribution of each measurement of the response variable was weighted by the sample size (number of seeds per tree). This analysis was implemented in the statistical software R (R Development Core Team, 2015). All relevant model assumptions were checked by examining the model deviance residuals and AIC (Akaike information criterion) values were used for model comparisons.

219

5.3   Results

5.3.1   Breeding system

We confirmed that triploid *S. subcuneata* is an obligate pseudogamous apomict, because hand-pollinations demonstrated that seed production required pollination (Table 5.1) and microsatellite analysis revealed that all of the 95 embryos studied had microsatellite phenotypes identical to both their maternal tree and the *S. subcuneata* reference samples shown in Table S5.2.

We found that *S. subcuneata* nevertheless depends on outcross pollinations, because 93 hand self-pollinations failed to produce even a single seed. Instead, only fruits resulting from addition of tetraploid *S. admonitor* pollen contained seed, with a corresponding flower-to-seed conversion rate of 65% (Table 5.1). One seed was also produced by a flower in the control group (emasculated and no pollen added), which also contained alleles from *S. admonitor* in the endosperm, presumably due to pollen contamination.

Table 5.1. Fruit production at three time intervals after manual pollination with final seed production for four pollination treatments carried out on cultivated specimens of *S. subcuneata*. The two fruit sizes represent small parthenocarpic fruit and larger fruit containing seed at three time intervals after pollination. Pollen donors with ploidy level in parenthesis represent the four pollination treatments, N = flowers treated, S/F = percentage of flowers that produced seed

Pollen donor	**N**	Fruit set of *S. subcuneata*						**S/F (%)**
		July		August		October		
♂ ♀		<9mm	≥9mm	<9mm	≥ 9mm	Total fruit	Total seed	
S. admonitor (4*x*)	**98**	11	46	0	40	40	64	**65.3**
S. aucuparia (2*x*)	**92**	13	1	11	0	4	0	**0.0**
S. subcuneata (3*x*)	**93**	39	0	8	0	7	0	**0.0**
None	**85**	33	1	10	1	11	1	**1.2**

Microsatellite analyses of wild-collected seed showed that endosperm formation frequently resulted from heterospecific pollination by *S. admonitor* (Table 5.2 and Fig. 5.1). Very rarely (one seed among 95), endosperm resulted from pollination by *S. margaretae,* which is one of the other tetraploid species on the study site (<10 individuals). One endosperm contained paternal alleles from a genotype of *S. aucuparia* that we did not find on the study site. However, endosperm formation was initiated by conspecific pollination in about half of the seeds in open-pollinated fruits (Table 5.2). Our hand-pollinations showed that no seed was set in 93 separate self-pollinations, so using the binomial theorem (Eq. 2) we sought the largest rate of flower-to-seed conversion such that observing 93 seedless fruits is not statistically significant.   Using $x=0$ and $n=93$,

221

we solved Eq. 1 at this threshold inequality, which yielded $p = 0.032$. Thus, a small rate of seed production from self-pollination of approximately 3% is statistically consistent with the observation that a sample of 93 hand-pollinations failed to yield even a single fruit.

Table 5.2. Summary of paternity results of microsatellite analysis of seed endosperm from a total of ten maternal *S. subcuneata* trees over two years. The proportions (%) of seed resulting from the various pollen donor species are shown for each tree in both years. N = number of seed sampled. Proportions in parenthesis are when only seed from *S. subcuneata* and *S. admonitor* are considered, to generate values for Eq 1.

| Maternal tree ID No. | Year | N | Paternity (%) | | |
			S. subcuneata (3x)	*S. admonitor* (4x)	Other
S01	2013	8	13	75	12
S02	2013	3	0	100	0
S269	2013	12	0	100	0
S58	2013	1	0	100	0
S01	2014	14	50	43	7
S02	2014	10	90	10	0
S269	2014	5	20	80	0
S58	2014	1	0	100	0
S156	2014	5	40	60	0
S280	2014	16	63	37	0
S282	2014	1	100	0	0
S283	2014	5	80	20	0
S284	2014	8	88	12	0
S285	2014	6	83	17	0
	TOTAL	95	50(51)	48(49)	2
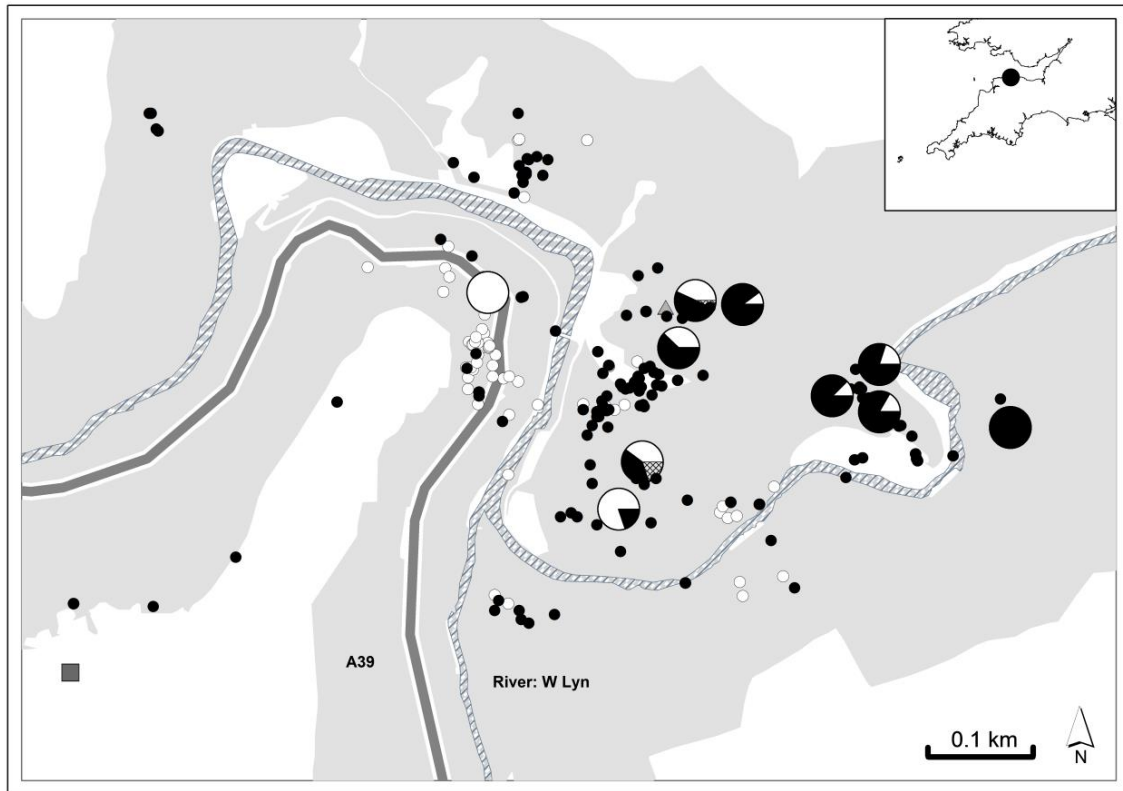
Figure 5.1. Distribution map of polyploid *Sorbus* species at the woodland site. The locations of *S. subcuneata* maternal seed trees are indicated by pie charts**.** The sections of the pie charts represent the proportion of seed resulting from pollination (2014 data) by; ●= *S. subcuneata* or ○ = *S. admonitor.* Hatching denotes other. ▲= *S. margaretae,* ■= *S. porrigentiformis.* Survey data supplied by Rich and Cann, (2009). (The map was created using ArcGIS Desktop version 10.2.2, ESRI, California, USA, URL: http://www.esri.com/).

5.3.2   Causes of pollen-limited seed set

Virtually no seed was produced by 1534 open-pollinated *S. subcuneata* flowers on three trees (mean flower-to-seed conversion rate = 0.5 %, SD = 0.46, n = 3), which probably was not due to a lack of pollen deposition because both test trees showed a significant accumulation of pollen on the stigmas with flower age over three days (likelihood ratio test for flower age χ^2 = 12.11, d.f. = 1, P < 0.001; Fig. 5.2). There was a large variation in pollen deposition between test

trees (χ^2 = 7.7, d.f. =1, P = 0.005), and between flowers, but nevertheless 88%
of two-day old flowers had more than 50 pollen grains on their stigma.

Table 5.3. Summary of fruit and seed production for two pollination
supplementation treatments carried out on two maternal *S. subcuneata*
individuals (S01 and S02). Each treatment was applied to 25 flowers. H =
heterospecific pollen from *S. admonitor (4x)*, O= naturally open pollinated, S/F =
percentage of flowers that produced seed.

S. subcuneata ♀	Pollen treatment ♂	No. fruit	No. seed	S/F (%)
S01	H	3	3	**12.0**
S02	H	2	4	**16.0**
S01	O	0	0	**0**
S02	O	0	0	**0**

This pollen accumulation probably was in part due to the activities of the insect
pollinators (*Bombus spp., Apis melifera*, Diptera and Lepidoptera) that we
observed visiting the flowers. However, when supplementary pollen from *S.
admonitor* was applied, flower-to-seed conversion was much higher than in
open-pollinated flowers, i.e. 12 % and 16 % for treated flowers on the two test
trees respectively, compared with zero conversion in open pollinated flowers
(Table 5.3). This suggests that a lack of compatible pollen from *S. admonitor*
limited seed production in open-pollinated flowers of *S. subcuneata*.
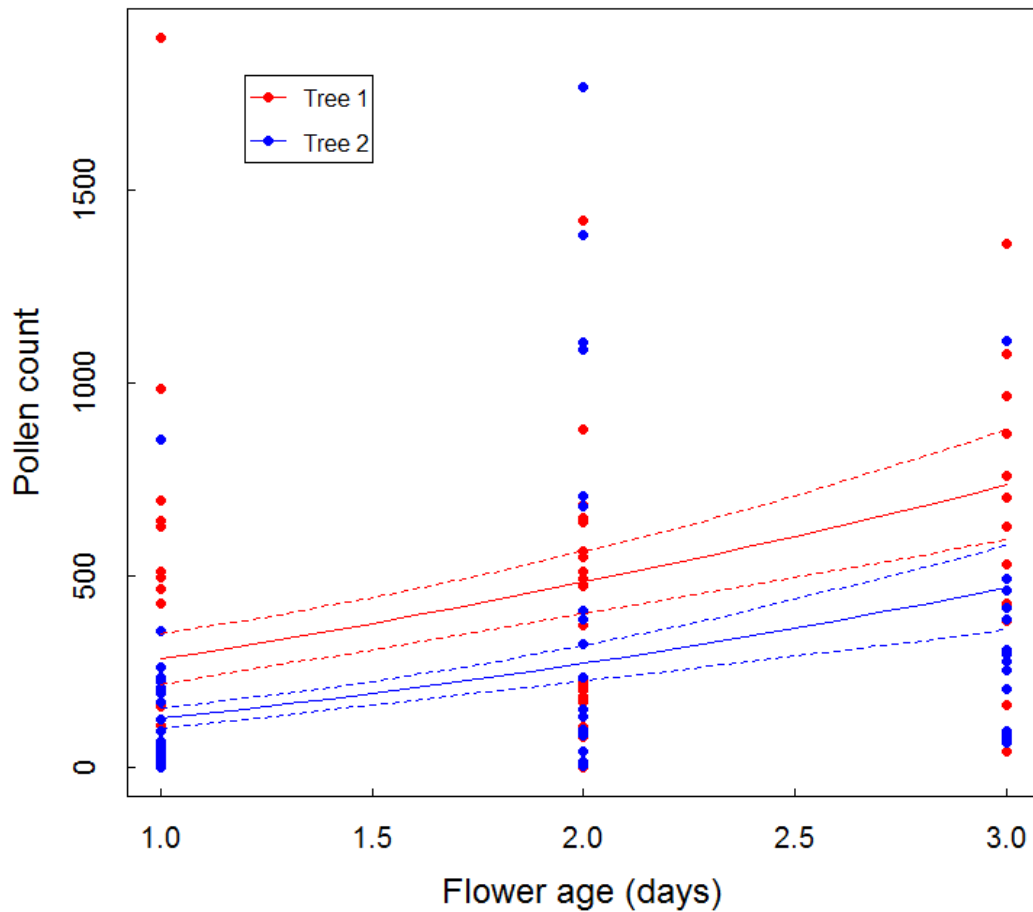
Figure 5.2. Pollen deposition on the stigmas over the three day life of *S. subcuneata* flowers. Lines show fitted values (±1 SE) from a negative binomial GLM (*square root pollen count* =11.64 [±2.84] + 5.15 [± 1.41] *age* - 5.47 [± 2.06] *tree*, where tree 1 is the reference category). Pseudo R^2 = 0.102. Tree 1 (upper line): *n* = 58 flowers; tree 2 (lower line): *n* = 68.

The lack of compatible pollen in flowers was not the result of asynchrony in blooming of *S. subcuneata* with its congeneric cross-pollinator, *S. admonitor*. Pollen flow between these species is certainly possible as the mean flowering day (MFD) was similar for both species (9.4 for *S. admonitor, n* = 6; 11.0 for *S.*

subcuneata, n = 7) and there was no significant difference in F_{50} (occasion of 50% of flowers open, Wilcoxon test, W = 12; *P* = 0.23; Figs. 5.3a and b).
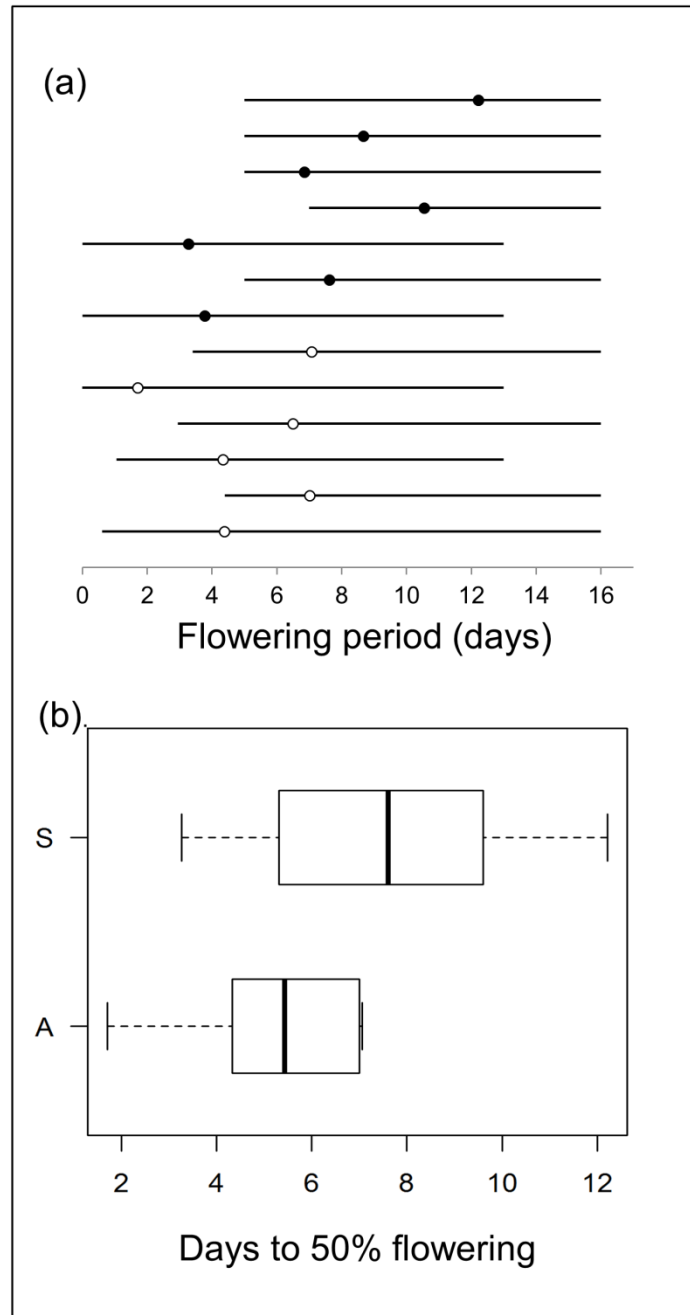
Figure 5.3. (a) Flowering phenology of individual trees of both species over a 16 day flowering period. Each line indicates the flowering period of an individual tree; dots represent the 50 percentile stage of the cumulative flowering curve; *S.*

subcuneata (●) *n* = 7 and *S. admonitor* (o) *n* = 6. (**b**) Flowering phenology of both species combined S = *S. subcuneata,* A = *S. admonitor.*

Instead, spatial isolation of the maternal trees from compatible pollen donors limits pollen flow between them because the proportion of seed initiated by heterospecific pollination decreased significantly with increasing spatial isolation (i.e. decreasing connectivity) of maternal *S. subcuneata* seed trees from *S. admonitor* individuals (Binomial GLM: connectivity likelihood ratio test χ^2 = 4.78, d.f. = 1, *P* = 0.029; pseudo-R^2 = 0.33; Fig. 5.4).
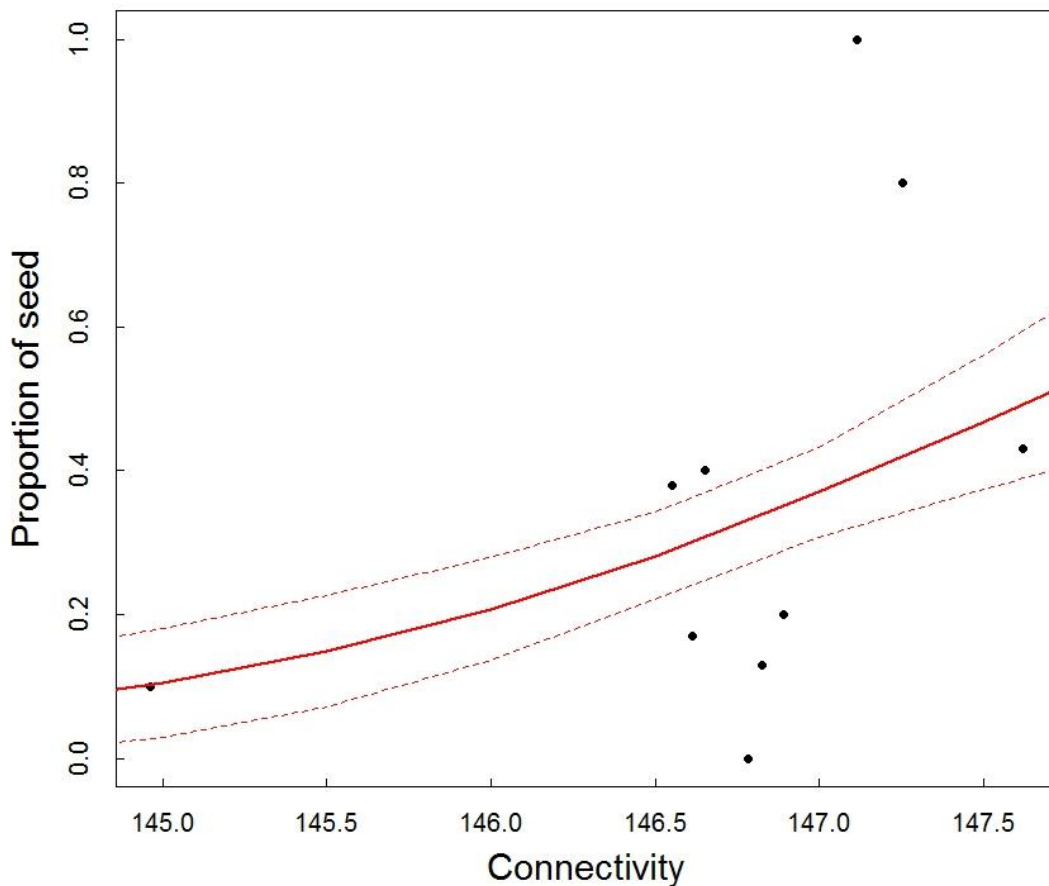
Figure 5.4. Relationship between the connectivity of the maternal *S. subcuneata* seed trees to *S. admonitor* and the proportion of seed sampled with

S. admonitor as pollinator (2014 data). The line shows the fitted values (± 1SE)

for a weighted binomial GLM of: logit (proportion of seed) = -119.1 (± 60.1) +

0.81 (± 0.41)* *connectivity*. N = 71 seeds from 10 trees.

The greater efficacy of heterospecific S*. admonitor* pollen in initiating seed

compared to that from *S. subcuneata* (Table 5.1) was not reflected in the

composition of the endosperm genotypes in open-pollinated flowers, where the

two species were represented fairly equally when total seed production from

both years was considered (Table 5.2). Solving Eq. 2 to estimate the ratio of

conspecific: heterospecific pollination with values derived from our study (see

Tables 5.1, 5.2 and equation 1; S = 49.5, C_s = 0.03, A = 48.5, and C_a = 0.65),

we find that R = 22:1, which suggests that the trees incurred very high levels of

conspecific pollination probably by within-plant selfing (geitonogamy). The

overwhelming amounts of conspecific pollination were sufficient to initiate some

seed production despite the very low efficacy of *S. subcuneata* pollen.

5.4   Discussion

Our study has characterised a system where seed production is severely limited by the availability of compatible pollination. Our findings provide understanding of the factors affecting seed production for partially self-incompatible pseudogamous apomicts. This is particularly relevant for *Sorbus* where, in the UK, thirteen of the forty polyploid species tested by Pellicer *et al.* (2012) were triploid. Despite their frequency, the majority of them exist as small populations sometimes numbering only a few trees e.g. c.17 individuals of *S. leyana* Wilmott *(*Rich *et al.*, 2005), so that effective strategies of conservation management are critical, which we discuss below.

5.4.1   Breeding system

We found that *S. subcuneata* is normally an obligate apomict that requires pollination for successful seed production in common with other pseudogamous *Sorbus* spp. (Liljefors, 1953; Ludwig *et al.*, 2013). Thus, successful pollination of *S. subcuneata* flowers will normally only affect the genotype of the endosperm except, very rarely when it will determine the genotype of the embryo of resulting seed. Our hand pollination experiments showed that the sympatric tetraploid *S. admonitor* is by far the most effective pollinator for *S. subcuneata* on our study site. Indeed, *S. subcuneata* is almost completely dependent on its congener because self-pollen is so poorly compatible.

The most likely explanation for the lack of seed produced by conspecific hand pollination is that *S. subcuneata* has a system of gametophytic self-incompatibility (GSI), which is common in other triploid *Sorbus* species (Ludwig *et al.*, 2013). GSI is caused by a disruption of the pollen tube before it reaches

229

the ovules and it is controlled by a multi-allelic *S*-locus. GSI occurs when the pollen tube has an *S*-allele in common with the pistil and it is the cause of GSI in diploid *Sorbus* but appears to break down in polyploids with tetraploids exhibiting self-compatibility (Horandl, 2010). High prevalence of inviable pollen is an alternative explanation for the widespread failure to set seed by self-pollination, but this is unlikely because the pollen grains from *S. subcuneata* observed during our study appeared rounded and well-formed and had a higher than average (of triploids tested) stainability (Rich, 2009), a proxy for pollen viability. We are unable to exclude the possibility that seed failure after selfing is also due to an imbalance between maternal and paternal genome components (m: p) in the endosperm tissue. The 2m:1p genome ratio of maternal to paternal contributions to the endosperm in sexual species is considered a prerequisite for successful seed development (Köhler *et al.*, 2010). In pseudogamous apomicts, the lack of meiotic division leads to unreduced central cell polar nuclei in the ovary and fertilisation of these results in m:p ratios that greatly exceed 2m:1p. However, triploid *Sorbus* such as *S. subcuneata* can tolerate an unbalanced endosperm, which is an attribute shared with other triploid apomicts in the Rosaceae (Talent & Dickinson, 2007; Ludwig *et al.*, 2013; Hajrudinović *et al.*, 2015). Thus it seems unlikely that the self-incompatibility that we observed in *S. subcuneata* is due to unbalanced endosperm and, instead, the low rate of seed production by self-pollination is likely the result of the operation of GSI.

Despite the greater efficacy of pollination from *S. admonitor* compared to self-pollination, about half of the seeds in open-pollinated flowers arose from conspecific pollination. This finding contrasts to other studies on triploid *Sorbus* (Ludwig *et al.*, 2013) and suggests that the GSI system in *S. subcuneata* is

230

partial, but with a very low rate of success. Self-incompatability can be a quantitive and plastic trait affected by environmental conditions such as temperature at flowering, the composition and density of the pollen load and the internal stylar conditions which change with flower age (Stephenson *et al.*, 2000). A mentor effect whereby a mix of heterospecific and self pollen on the stigma may reduce heterospecific pollination from plants of different ploidy and promote increased selfing in natural mixed ploidy populations (KouteckÝ *et al.*, 2011). Mentor pollination is widely used as a tool to overcome incompatibility in commercial plant breeding (Shivanna *et al.*, 2005). *S. subcuneata* stigmas pollinated towards the end of the flower life may show reduced self-incompatibility since pollen /pistil interactions during pollen tube growth change with flower age and can allow successful self pollination as opportunities for cross pollination reduce (Stephenson *et al.*, 2000). Whilst these mechanisms may be responsible for a weakening of the recognition of self-pollen (pseudo-self compatibility), we cannot rule out the possibility that potentially unsampled genotypes in the *S. subcuneata* population may carry alternative *S*-alleles, conferring pollen / pistil compatability and allowing succesfull pollination, although we were unable to find evidence of genetic variation among both seed embryos and adult trees sampled at the study site. The low rate of seed production despite ample pollen deposition in naturally pollinated flowers, however, suggests self-compatibility is a rare phenomenon in *S. subcuneata*. At our study site, *S. subcuneata* relied heavily for seed set on heterospecific pollination from a congener, *S. admonitor*.   The diploid congener, *S. aucuparia*, was more common than *S. admonitor* on our study site, but it was an ineffectual male parent. *S. aucuparia* can successfully pollinate other triploid *Sorbus* (Robertson *et al.*, 2004; Rich *et al.*, 2010), which shows that compatibility is not

231

determined solely by the comparative ploidy of the parents. The relative efficacy of pollination by tetraploid *S. admonitor* in enabling seed set in closely related triploid *S. subcuneata* probably reflects a broad gametophytic compatibility that permits the successful growth of pollen tubes. Despite the greater efficacy of pollination from *S. admonitor* compared to self-pollination, about half of the seeds in open-pollinated flowers probably arose from geitonogamous (within-individual) self-pollination. This observation has two implications. First, it implies that *S. subcuneata* exhibits a pseudo-self-compatibility or 'leakiness' of self-pollen recognition (discussed above), in the sense that self-pollination *can* result in production of seed, but only with a very low rate of success (flower-to-fruit conversion rate in open-pollinated flowers was <0.5% despite apparently abundant self-pollination). Second, it implies that levels of pollen transfer between the two species are low. This is usual when insect-pollinated plants have many flowers simultaneously in bloom, because the tendency of floral foragers to minimize travel time means that most inter-flower transitions are on the same plant (Zimmerman, 1988).

5.4.2 Causes of pollen-limited seed set

We found open-pollinated flowers of *S. subcuneata* frequently produced parthenocarpic fruits, suggesting that whereas resources were available for fructification, pollination had limited seed production. We found that seed production in our study population was not limited by the quantity of pollen accumulated and, instead, it appeared that an active system of insect-pollination was capable of delivering ample pollen to the stigmas of flowers at the study site. Despite the potential effectiveness of heterospecific pollination by *S. admonitor* that was demonstrated by our hand-supplementation

232

experiments, the low rate of flower-to-seed conversion among open-pollinated flowers suggests that the availability of this congeneric pollen constrained seed set. We have shown that pollination from *S. admonitor* was not restricted by asynchronous flowering with *S. subcuneata*, but instead spatial isolation of the maternal trees from congeners (i.e. connectivity) was the limiting factor.

The impact of spatial isolation appears largely due to the disjunct distribution patterns of *S. subcuneata* and *S. admonitor*. Despite the fact that insect-mediated pollen flow can occur over long distances (Kamm *et al.*, 2009; Lander *et al.*, 2010; Fuchs & Hamrick, 2011), Oddou-Muratorio *et al.* (2005),found that for forest populations of *S. torminalis* the majority of an individual's incoming pollen came from its nearest neighbours, although occasionally a male pollen donor could be up to 2.2km away from the maternal tree. This contrasts to open habitats where the maximum distance for mating pairs was only 11.8 m among various *Sorbus* spp. (Ludwig, 2013, p. 195). Our study site is less than 2 km across, so in theory pollen has the potential to be transported between any two (or more) trees, but given that both *S. admonitor* and *S. subcuneata* individuals can produce hundreds of flowers that bloom simultaneously, the majority of pollen flow probably occurs within individuals because of the area-restricted movements of flower visitors. Thus, geitonogamy and short distance dispersal probably dominated pollen transfer.

The proportion of seed produced from self-pollination was unrelated to the maternal tree's proximity to other *S. subcuneata* trees, which further suggests that the large amounts of *Sorbus* pollen that accumulated on the flowers of the maternal trees was principally geitonogamous in origin. Moreover, of *S.*

233

subcuneata and *S. admonitor* on the study site, the closest extrinsic sources of pollen for maternal *S. subcuneata* seed trees were normally other individuals of *S. subcuneata*, not congeners (Fig. 5.1). Since *S. subcuneata* is effectively clonal, due to its apomictic mode of reproduction, pollen from these conspecific individuals is likely to be as incompatible as that from the maternal tree itself. Together, these attributes could place additional barriers to heterospecific pollination since high levels of geitonogamous pollination will result in increased competition for space on the stigma from unsuitable self-pollen, or 'stigma clogging' (Groom, 2001; Murphy & Lovett-Doust, 2004). Whilst pollen carryover will, to some extent alleviate the effects of geitonogamous pollinator movements (Morris *et al.*, 1994), the normally rapid attenuation of carryover (Thomson, 1986; Morris *et al.*, 1995) and localised movements of insect vectors make it inevitable that insect-mediated pollen transfer among large floral displays remains highly restricted in space (Levin & Kerster, 1967).

5.4.3 Evolutionary implications

The extremely low reproductive output of *S. subcuneata* in our study population is seemingly inherent to its breeding system, which is maladaptive in a population comprised of a single-cytotype whose individuals have limited access to compatible heterospecific pollen donors. Indeed, although it has been previously predicted on theoretical grounds that SI pseudogamous plants should not persist (Noirot *et al.*, 1997), we have discovered at our study site an established population of c. 300 such trees. However, although triploids such as *S. subcuneata* may be transcient in the long term, where apomixis is facultative, they may form important evolutionary stages as part of a recurring process of polyploid formation that continually diversifies a lineage (Ramsey & Schemske,

234

1998) because the dependence on heterospecific pollination may increase opportunities for hybrid events, although the predominantly apomictic breeding system of *S. subcuneata* implies that such events will be rare. In the case of *S. subcuneata*, for example, the high level of allele sharing between *S. subcuneata* and *S. admonitor* suggests recently shared origins from an active process of reticulate evolution by hybridisation.

5.4.4 Conservation implications

This system provides valuable insights into the reproductive ecology of self-incompatible polyploid plants, which may elucidate the conservation ecology of similar species. Our study suggests that for a self-incompatible pseudogamous apomict that exists at low densities, seed production is likely to be limited. Furthermore, *Sorbus* is a palatable species in which recruitment from seeds is threatened by browsing from deer (Rich *et al.*, 2010) which occur in high densities at our study site.

Given the vulnerability of species that are, in effect, locally endemic clonal cytotypes, conservation strategies that aim to enhance the short term population viability of rare *Sorbus* species should also safeguard the evolutionary process in which extinctions are offset by diversifying hybridisation. This may be achieved through maintaining compatible congeneric species in close proximity to allow pollen exchange with the dual benefit of increasing the potential for hybridisation events and also resulting in increases in seed production for self-incompatible species. Given that seed production is likely to be chronically low because of the breeding system (self-incompatible) and architecture of the floral display (many flowers promoting geitonogamous

pollination), conservation strategies should also seek to increase recruitment opportunities to maximise the germination chances of the few seeds currently produced and control herbivore populations. Some introductions of compatible pollinating species should be considered at the dwindling isolated populations found in this area. *S. subcuneata* exists as a few trees at two sites further east of the study site where they have been seen to be slowly reducing in number within recent years (Rich *et al.*, 2010).

References

Barbour, R. C., Potts, B. M., Vaillancourt, R. E. & Tibbits, W. N. (2006) Gene flow between introduced and native *Eucalyptus* species: Flowering asynchrony as a barrier to F1 hybridisation between exotic *E. nitens* and native Tasmanian *Symphyomyrtus* species. *Forest Ecology and Management,* **226,** 9-21.

Bell, G. (1985) On the function of flowers. *Proceedings of the Royal Society of London B: Biological Sciences,* **224,** 223-265.

Bicknell, R. A. & Koltunow, A. M. (2004) Understanding apomixis: Recent advances and remaining conundrums. *Plant Cell,* **16,** S228-S245.

Burd, M. (1994) Bateman's principle and plant reproduction: the role of pollen limitation in fruit and seed set. *The Botanical Review,* **60,** 83-139.

Campbell, C. S. & Dickinson, T. A. (1990) Apomixis, patterns of morphological variation, and species concepts in subfam. Maloideae (Rosaceae). *Systematic Botany,* **15,** 124-135.

Dickinson, T. A., Lo, E. & Talent, N. (2007) Polyploidy, reproductive biology, and Rosaceae: understanding evolution and making classifications. *Plant Systematics and Evolution,* **266,** 59-78.

Fuchs, E. J. & Hamrick, J. L. (2011) Mating system and pollen flow between remnant populations of the endangered tropical tree, *Guaiacum sanctum* (Zygophyllaceae). *Conservation Genetics,* **12,** 175-185.

Groom, M. J. (2001) Consequences of subpopulation isolation for pollination, herbivory, and population growth in *Clarkia concinna concinna* (Onagraceae). *Biological Conservation,* **100,** 55-63.

Guitián, J., Guitián, P. & Sánchez, J. M. (1993) Reproductive biology of two *Prunus* species (Rosaceae) in the Northwest Iberian Peninsula. *Plant Systematics and Evolution,* **185,** 153-165.

Hajrudinović, A., Siljak-Yakovlev, S., Brown, S. C., Pustahija, F., Bourge, M., Ballian, D. & Bogunić, F. (2015) When sexual meets apomict: genome size, ploidy level and reproductive mode variation of *Sorbus aria s.l.* and *S. austriaca* (Rosaceae) in Bosnia and Herzegovina. *Annals of Botany,* **116,** 301-312.

Hemery, G. E., Savill, P. S. & Pryor, S. N. (2005) Applications of the crown diameter–stem diameter relationship for different species of broadleaved trees. *Forest Ecology and Management,* **215,** 285-294.

Horandl, E. (2010) The evolution of self-fertility in apomictic plants. *Sexual Plant Reproduction,* **23,** 73-86.

IUCN (2001) *IUCN Red List categories and criteria version 3.1.* IUCN, Gland, Switzerland.

Kamm, U., Rotach, P., Gugerli, F., Siroky, M., Edwards, P. & Holderegger, R. (2009) Frequent long-distance gene flow in a rare temperate forest tree (*Sorbus domestica*) at the landscape scale. *Heredity,* **103,** 476-482.

Knight, T. M., Steets, J. A., Vamosi, J. C., Mazer, S. J., Burd, M., Campbell, D. R., Dudash, M. R., Johnston, M. O., Mitchell, R. J. & Ashman, T.-L. (2005) Pollen limitation of plant reproduction: Pattern and Process. *Annual Review of Ecology, Evolution, and Systematics,* **36,** 467-497.

Köhler, C., Mittelsten Scheid, O. & Erilova, A. (2010) The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics,* **26,** 142-148.

KouteckÝ P, BaĎUrovÁ T, ŠTech M, KoŠNar JAN, KarÁSek J. (2011). Hybridization between diploid *Centaurea pseudophrygia* and tetraploid *C. jacea* (Asteraceae): the role of mixed pollination, unreduced gametes, and mentor effects. *Biological Journal of the Linnean Society,* **104**: 93-106.

Lander, T. A., Boshier, D. H. & Harris, S. A. (2010) Fragmented but not isolated: Contribution of single trees, small patches and long-distance pollen flow to genetic connectivity for *Gomortega keule,* an endangered Chilean tree. *Biological Conservation,* **143,** 2583-2590.

Levin, D. A. & Kerster, H. W. (1967) An analysis of interspecific pollen exchange in *Phlox. American Naturalist,* 387-399.

Liljefors, A. (1953) Studies on propagation, embryology, and pollination in *Sorbus. Acta Horti Bergiani,* **16,** 277-329.

Ludwig, S., Robertson, A., Rich, T. C. G., Djordjevic, M., Cerovic, R., Houston, L., Harris, S. A. & Hiscock, S. J. (2013) Breeding systems, hybridization and continuing evolution in Avon Gorge *Sorbus. Annals of Botany.*

Ludwig, S. C. (2013) *Breeding systems, pollen flow and continuing evolution in Avon Gorge Sorbus (whitebeams, rowans and service trees).* Unpublished PhD thesis Unpublished PhD thesis, University of Bristol, Bristol.

Moilanen, A. (2000) The Si program. *Metapopulation research group, Helsinki.*

Moilanen, A. & Nieminen, M. (2002) Simple connectivity measures in spatial ecology. *Ecology,* **83,** 1131-1145.

Morris, W. F., Mangel, M. & Adler, F. R. (1995) Mechanisms of pollen deposition by insect pollinators. *Evolutionary Ecology,* **9,** 304-317.

Morris, W. F., Price, M. V., Waser, N. M., Thomson, J. D., Thomson, B. & Stratton, D. A. (1994) Systematic increase in pollen carryover and its consequences for geitonogamy in plant populations. *Oikos,* 431-440.

Murphy, H. T. & Lovett-Doust, J. (2004) Context and connectivity in plant metapopulations and landscape mosaics: does the matrix matter? *Oikos,* **105,** 3-14.

Nogler, G. A. (1984) Gametophytic apomixis. *Embryology of Angiosperms* (ed B. Johri), pp. 475-518. Springer Berlin Heidelberg.

Noirot, M., Couvet, D. & Hamon, S. (1997) Main role of self-pollination rate on reproductive allocations in pseudogamous apomicts. *Theoretical and Applied Genetics,* **95,** 479-483.

Oddou-Muratorio, S., Klein, E. K. & Austerlitz, F. (2005) Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. II. Pollen dispersal and heterogeneity in mating success inferred from parent–offspring analysis. *Molecular Ecology,* **14,** 4441-4452.

Pellicer, J., Clermont, S., Houston, L., Rich, T. C. G. & Fay, M. F. (2012) Cytotype diversity in the *Sorbus* complex (Rosaceae) in Britain: sorting out the puzzle. *Annals of Botany,* **110,** 1185-1193.

Pías, B. & Guitián, P. (2006) Breeding system and pollen limitation in the masting tree *Sorbus aucuparia* L. (Rosaceae) in the NW Iberian Peninsula. *Acta Oecologica,* **29,** 97-103.

Proctor, M. C. F., Proctor, M. E. & Groenhof, A. C. (1989) Evidence from peroxidase polymorphism on the taxonomy and reproduction of some *Sorbus* populations in south-west England. *New Phytologist,* **112,** 569-575.

R Development Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramsey, J. & Schemske, D. W. (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual review of Ecology and Systematics,* **29,** 467-501.

Rich, T. C. G. (2009) Pollen stainability in British *Sorbus* L. (Rosaceae). *Plant Ecology & Diversity,* **2,** 85-88.

Rich, T. C. G. & Cann, D. C. G. (2009) A survey of *Sorbus* species at Watersmeet, north Devon, September 2007 *Transactions of the Devonshire Association for the Advancement of Science,* **140,** 185-198.

Rich, T. C. G., Houston, L., Robertson, A. & Proctor, M. C. F. (2010) *Whitebeams, rowans and service trees of Britain and Ireland. A monograph of British and Irish Sorbus L. B.S.B.I. Handbook No. 14.* Botanical Society of the British Isles, London.

Rich, T. C. G., Motley, G. S. & Kay, Q. O. N. (2005) Population sizes of three rare Welsh endemic *Sorbus* species (Rosaceae). *Watsonia,* **25,** 381-388.

Rich, T. C. G. & Proctor, M. C. F. (2009) Some new British and Irish *Sorbus* L. taxa (Rosaceae). *Watsonia,* **27,** 207-216.

Richards, A. J. (2003) Apomixis in flowering plants: an overview. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **358,** 1085-1093.

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

Robertson, A., Rich, T. C. G., Allen, A. M., Houston, L., Roberts, C. A. T., Bridle, J. R., Harris, S. A. & Hiscock, S. J. (2010) Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus*. *Molecular Ecology,* **19,** 1675-1690.

Shivanna KR, Sawhney VK, Knox RB. (2005). *Pollen biotechnology for crop production and improvement*: Cambridge University Press.

Stephenson AG, Good SV, Vogler DW. (2000). Interrelationships among inbreeding depression, plasticity in the self-incompatibility system, and

the breeding system of *Campanula rapunculoides* L.(Campanulaceae). *Annals of Botany,* **85:** 211-219

Talent, N. & Dickinson, T. A. (2007) Endosperm formation in aposporous *Crataegus* (Rosaceae, Spiraeoideae, tribe Pyreae): parallels to Ranunculaceae and Poaceae. *New Phytologist,* **173,** 231-249.

Thomson, J. D. (1986) Pollen transport and deposition by bumble bees in *Erythronium*: Influences of floral nectar and bee grooming. *The Journal of Ecology*, 329-341.

van Dijk, P. J. (2003) Ecological and evolutionary opportunities of apomixis: insights from *Taraxacum* and *Chondrilla*. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* **358,** 1113-1121.

van Dijk, P. J. & Vijverberg, K. (2005) The significance of apomixis in the evolution of the angiosperms: a reappraisal. *Plant Species-Level Systematics: New Perspectives on Pattern & Process* (eds F. T. Bakker, L. W. Chatrou, B. Gravendeel & P. B. Pelser), pp. 101-116. Gantner Verlag, Ruggell [Regnum Vegetabile 143], Liechtenstein.

Whitton, J., Sears, C. J., Baack, E. J. & Otto, S. P. (2008) The dynamic nature of apomixis in the angiosperms. *International Journal of Plant Sciences,* **169,** 169-182.

Wilcock, C. & Neiland, R. (2002) Pollination failure in plants: Why it happens and when it matters. *Trends in Plant Science,* **7,** 270-277.

Zimmerman, M. (1988) Nectar production, flowering phenology, and strategies for pollination. *Plant reproductive ecology: patterns and strategies* (eds J. Lovett-Doust & L. Lovett-Doust), pp. 157-178. Oxford University Press, Oxford, UK.

A206

S209

A214

S02

A174

S01

A6

S208
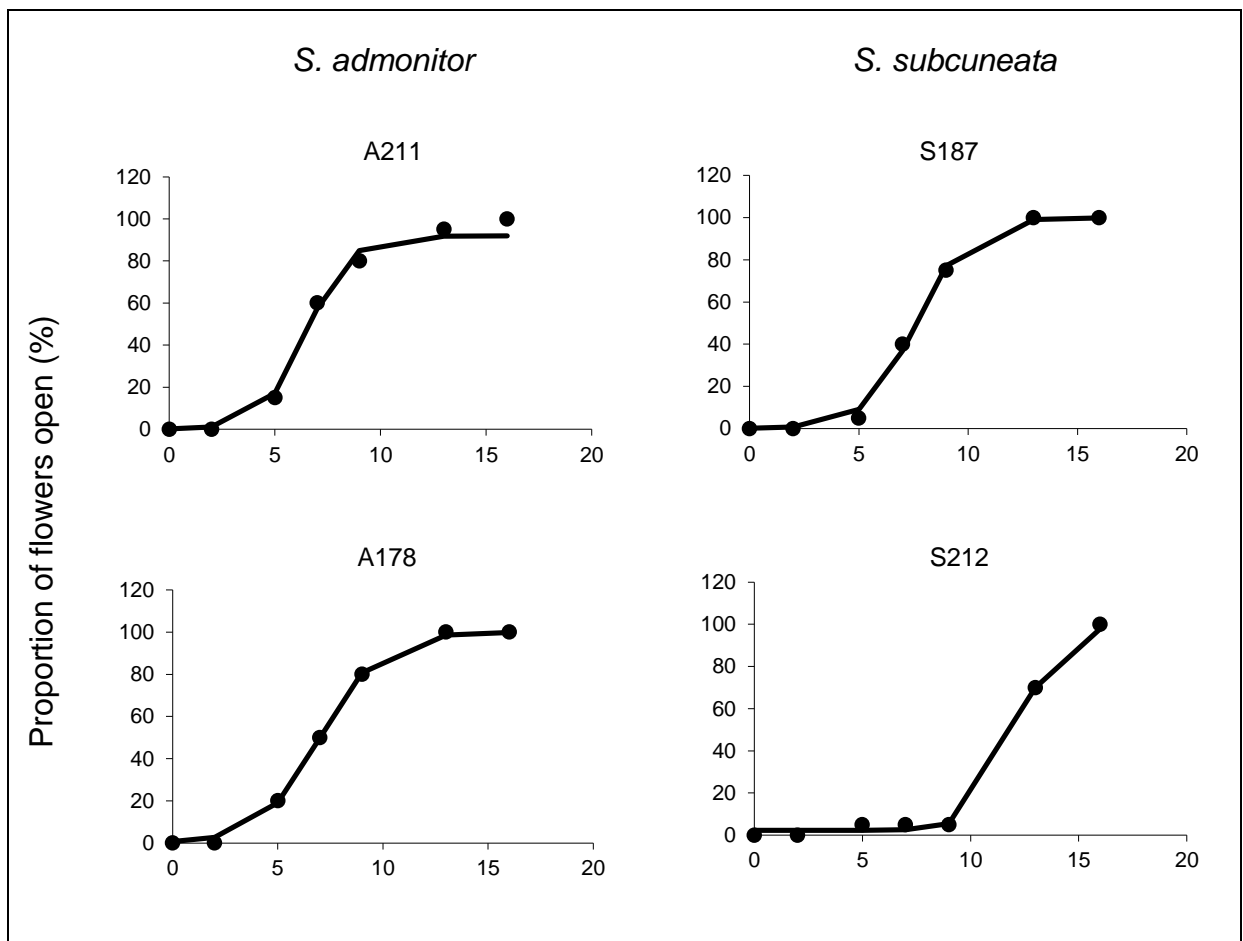
Proportion of flowers open (%)
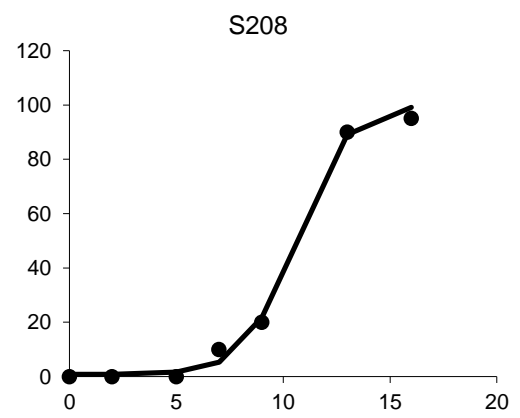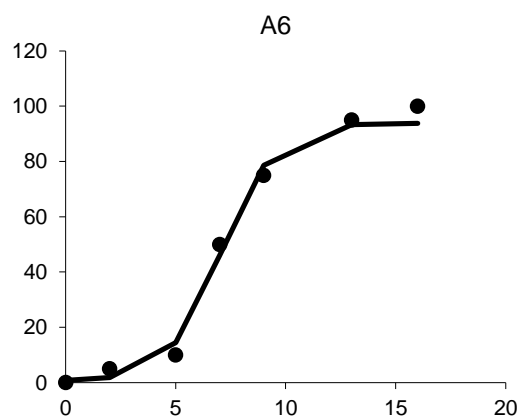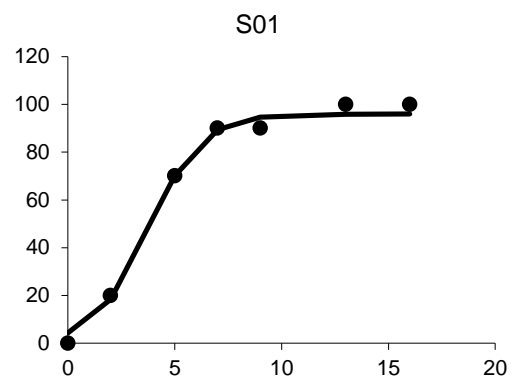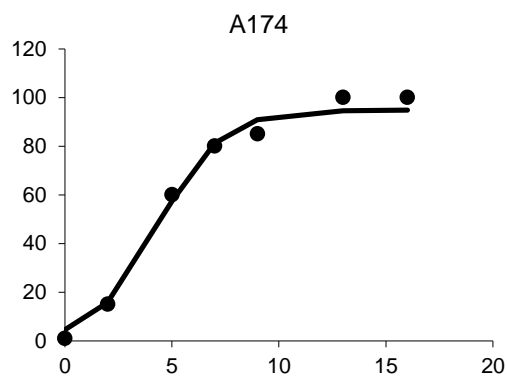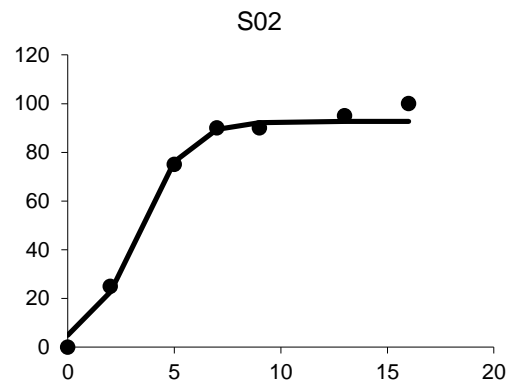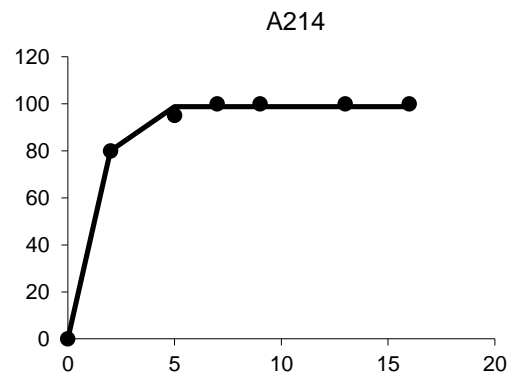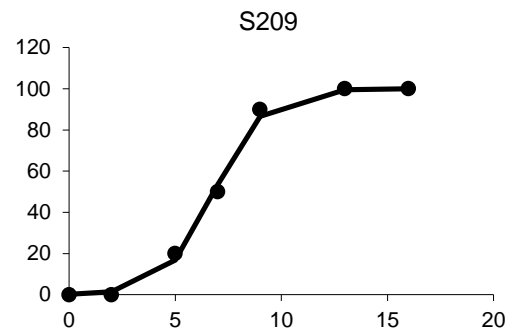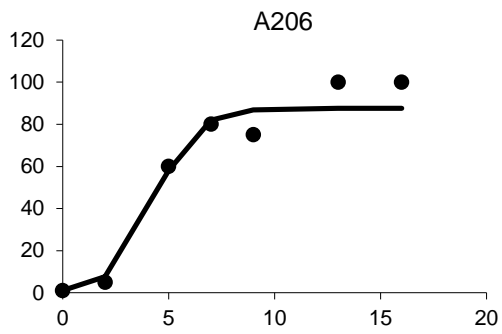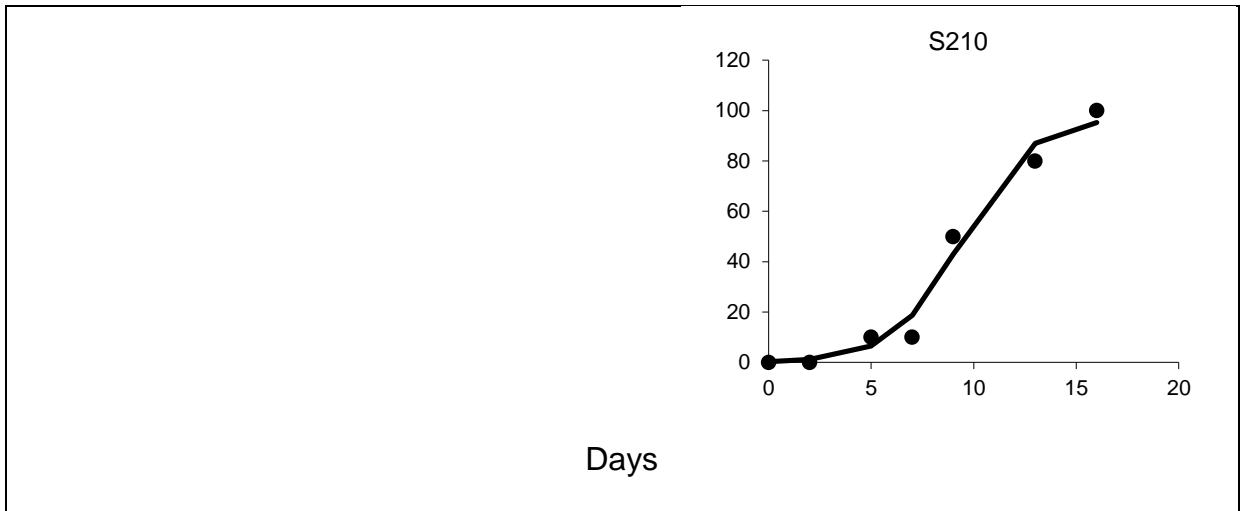
Figure S5.1. Cumulative frequency of flowers opening over the flowering period. Time to 50% of flowers open (F50) is calculated from a sigmoid curve expression:

y (% flowering) = a + [(b – a) / (1 + exp(c/d)]

Where a= y intercept, b = asymptote, c = F50 - time, d = Hill's slope of the curve (i.e. this is related to the steepness of the curve at the inflexion point c).

244

Table S5.1. Location details for study trees.

Tree ID	Species	x (OS GB)	y (OS GB)	Lat.	Long.	Site	Data		
A174	*S. admonitor*	274421	148797	51.224511	-3.7997354	Watersmeet	Phenology	Molecular	-
A178	*S. admonitor*	274672	148624	51.223011	-3.7960823	Watersmeet	Phenology	Molecular	-
A183	*S. admonitor*	274431	148995	51.226292	-3.7996617	Watersmeet	Phenology	Molecular	-
A211	*S. admonitor*	274648	148664	51.223365	-3.7964398	Watersmeet	Phenology	Molecular	-
A214	*S. admonitor*	274431	149039	51.226688	-3.7996772	Watersmeet	Phenology	Molecular	-
A6	*S. admonitor*	274469	148851	51.225007	-3.7990673	Watersmeet	Phenology	Molecular	-
Au262	*S. aucuparia*	289584	145973	51.202265	-3.5818184	Horner wood	-	Molecular	-
Au263	*S. aucuparia*	274762	149051	51.226868	-3.7949435	Watersmeet	-	Molecular	-
Au264	*S. aucuparia*	274681	148705	51.223741	-3.7959819	Watersmeet	-	Molecular	-
Au507	*S. aucuparia*	289400	146100	51.203371	-3.5844902	Luckbarrow	-	Molecular	-
M03	*S. margaretae*	274575	148879	51.225282	-3.7975599	Watersmeet	-	Molecular	-
S01	*S. subcuneata*	274651	148878	51.225289	-3.7964718	Watersmeet	Phenology	Molecular	Seed analysis
S02	*S. subcuneata*	274658	148883	51.225336	-3.7963734	Watersmeet	Phenology	Molecular	Seed analysis
S156	*S. subcuneata*	274556	148723	51.223875	-3.7977773	Watersmeet	-	Molecular	Seed analysis
S187	*S. subcuneata*	274501	148886	51.225328	-3.7986216	Watersmeet	Phenology	Molecular	-
S208	*S. subcuneata*	274441	148962	51.225998	-3.799507	Watersmeet	Phenology	Molecular	-
S209	*S. subcuneata*	274353	148926	51.225655	-3.800754	Watersmeet	Phenology	Molecular	-
S210	*S. subcuneata*	274648	148664	51.223365	-3.7964398	Watersmeet	Phenology	Molecular	-

Tree ID	Species	x (OS GB)	y (OS GB)	Lat.	Long.	Site	Data		
S212	*S. subcuneata*	274838	148723	51.223937	-3.7937411	Watersmeet	Phenology	Molecular	-
S269	*S. subcuneata*	274532	148698	51.223645	-3.798112	Watersmeet	-	Molecular	Seed analysis
S280	*S. subcuneata*	274593	148839	51.224926	-3.7972883	Watersmeet	-	Molecular	Seed analysis
S282	*S. subcuneata*	274930	148758	51.224272	-3.7924365	Watersmeet	-	Molecular	Seed analysis
S283	*S. subcuneata*	274780	148789	51.224518	-3.7945943	Watersmeet	-	Molecular	Seed analysis
S284	*S. subcuneata*	274779	148790	51.224526	-3.7946089	Watersmeet	-	Molecular	Seed analysis
S285	*S. subcuneata*	274797	148774	51.224387	-3.7943457	Watersmeet	-	Molecular	Seed analysis
S58	*S. subcuneata*	274399	148895	51.225387	-3.8000847	Watersmeet	-	Molecular	Seed analysis

Table S5.2. Allele sizes for each study *Sorbus* species at 10 loci. Ploidy levels in parenthesis. Shaded alleles are not present in *S. subcuneata.*

Locus	*S. subcuneata* (3x)			*S. admonitor* (4x)				*S. margaretae* (4x)				*S. porrigentiformis* (4x)				*S.aucuparia* (2x)			
MSS5	119	121	123	119	121	123	127	119	121	135		115	127	131	137	115			
MSS16	158	160	204	158	160	188	204	158	160	162		158	162	170		154	156		
CH01F09	113	123		113	123			119	121	135		115	123	125	129	-	-		
SA06	258	268		258	268			258	268	280	312	256	264	270		246	264	272	280
																326			
MSS13	193	195		187	193	195		193	195	197		193	195	197	203	183	187	193	195
																199	203		
SA02	292	294		292	294			278	282	292	294	294	300	324		-	-		
SA08	261			261	285			247	263			249	257	277		257	265	267	
SA09	162	194		162	194			162	182	194		174	176	184	186	168	180		
MS14	123			123	131			-	-	-	-	-	-	-	-	111	123	125	127
SA14	170	208	226	170	178	208	226	194	208	226		196	222	224	226	212	214	267	

Chapter 6: General discussion

This study concerns evolutionary diversification and reproductive sustainability in a taxonomically complex group (TCG) using *Sorbus* as a model system. We used a combination of molecular techniques and field experiments to elucidate the reticulate patterns of evolution among a regional group of *Sorbus* species and determine the reproductive sustainability of a threatened *Sorbus* species. Here I discuss the findings of this study under each of the aims presented in the thesis introduction Chapter.

6.1   AIM 1. Evolutionary relationships, hybrid origins and polyploid formation of *Sorbus* in Devon and north Somerset.

In order to elucidate evolutionary relationships among seven polyploid and two diploid *Sorbus* species native to the Devon and north Somerset region of Britain we characterized 230 samples into species and multilocus genotypes using nuclear DNA microsatellite markers (Chapter 2). These data also provided an assessment of genetic diversity within and among the study species which revealed an apomictic breeding system and single hybrid origins for each of the polyploid species. No evidence for sexuality within the polyploid species was found and whilst there is confirmation of sexuality in polyploid *Sorbus* seed (Robertson *et al.*, 2004a , Ludwig *et al.*, 2013, Hajrudinović *et al.*, 2015), this is not necessarily evident in the adult population (Robertson *et al.*, 2010) possibly due to low fitness either inherently due to post zygotic barriers, or being mal-adapted to the environment, but see Robertson *et al.*, (2004b).

Each species was genetically differentiated with an additional interspecific hybrid discovered (vex2), representing a novel polyploid lineage. The relative genetic distances among the species implied close evolutionary relationships

248

among the five endemic species (including *S. devoniensis*) and *S. rupicola;* in particular the parental role of *S. margaretae* for the three study members of subgenus *Tormaria* (*S. subcuneata, S. devoniensis* and *S. admonitor*). Diploid *S. torminalis* is the other parental species. Parental simulations confirmed these hybrid origins (Chapter 3) and also revealed the key role of triploid *S. subcuneata* in the production of both tetraploids *S. devoniensis* and *S. admonitor* via hybridisation with *S. torminalis.* The inheritance patterns in members of subgenus *Tormaria* combined with plastid haplotype information from previous studies (Chester *et al.*, 2007) reveal that triploid formation in this group was via fertilisation of a normally reduced *S. torminalis* gamete with reduced pollen from tetraploid *S. margaretae.* The subsequent formation of the tetraploids from *S. subcuneata* was via fusions of unreduced triploid gametes with reduced *S. torminalis* gametes. Due to the prevalence of apomixis in triploid *S. subcuneata* (also see 5.3.1) we suggest that sexual *S. torminalis* was likely the female progenitor. The lack of exact parental matched pairs for the endemics *S. margaretae, S. vexans* and the hybrid vex2 suggest mutational variation or potentially missing intermediary genotypes (perhaps triploid forms) due to insufficient sampling or extinction. The presence of a small number of alleles unique to the endemic species also suggests missing parental genotypes. It is likely that all three scenarios may account for the variation seen in the study members of subgenus *Aria.* Flow cytometry confirmed the ploidy of each species and revealed a spontaneous triploid form of the normally diploid *S. torminalis* (Chapter 4).

These findings suggest that the four endemic polyploid species and *S. devoniensis* all arose from single hybridisation events and subsequently colonised other sites. It is likely that the high unstable coastal cliffs provided a

249

colonisation route with many chances for establishment where disturbances due to landslips and quarrying provided open conditions for seedling establishment.

The key roles of sexual diploids in the diversification process in *Sorbus* is supported by our data, however *S. torminalis* currently only occurs with *S. devoniensis* and not *S. margaretae, S. subcuneata* or *S. admonitor.* Given the restricted distributions of the latter three species it seems that the distribution of *S. torminalis* has changed since the origin of these polyploids. This is not unlikely as it may have been an infrequent tree in this area, as indeed it still is across Devon. Its absence from the north Devon and Somerset sites implies that further triploid and tetraploid formations via the above routes are unlikely. The presence of cryptic hybrids and missing genotypes present a complex situation within subgenus *Aria* and is worthy of further investigation. This group is the most challenging of the subgenera with high diversity and ongoing evolution (Lepší *et al.*, 2015) and our findings reiterate its critical role in speciation within the genus *Sorbus.* The patterns of reticulate evolution we see in this study group are brought about by gene flow via cross pollination among diploid and polyploid *Sorbus.* This is unsurprising since our findings show there is reliance on heterospecific pollination for seed production in the triploid species that performs a key role in the diversification of this group.

6.2   AIM 2. Reproductive sustainability of a threatened *Sorbus* species.

Pollen flow among the study species provides the conduit for gene flow in the process of reticulate evolution we have described in this thesis. It also sustains seed production in the threatened triploid *S. subcuneata* which is shown to have played a key role in *Sorbus* diversification in this region. The reproductive potential of *Sorbus* is of importance since it directly impacts population viability.

The key role of triploid *S. subcuneata* in the production of endemic tetraploids also makes it of interest in an evolutionary context. Reasons for low seed production in *S. subcuneata* were investigated with a series of pollination experiments in addition to molecular analysis of embryo and endosperm to identify how breeding system and factors affecting pollination may cause reproductive failure. We found *S. subcuneata* to be largely self-incompatible as with other triploid *Sorbus* (Ludwig *et al.*, 2013) and therefore reliant on sufficient pollen from other sources to produce seed. Its main compatible partner was tetraploid *S. admonitor* which is the most prevalent other polyploid *Sorbus* on our study site (Chapter 5). Diploid *S. torminalis* does not occur there and the frequent diploid *S. aucuparia* was an ineffectual father. Pollination by *S. admonitor* resulted in a 65% flower-to-seed conversion rate in hand pollination trials compared to <0.5% of open pollinated flowers at the field site, so provided sufficient pollen from *S. admonitor* is available, sustainable seed production should be possible despite a self-incompatible breeding system. Whilst seed production may be limited by a number of factors such as resource availability, pollen supplementation experiments showed that pollination limitation was the reason for the critically low natural levels of seed set. Our results showed that spatial isolation from compatible congener *S. admonitor* was the reason for pollen limitation rather than insufficient pollinator activity or flowering asynchrony between the two species. The low density of *S. subcuneata* and the disjunct distributions of compatible pollination partners meant limited pollen flow from *S. admonitor* to *S. subcuneata.* Whilst this low level may be sufficient to promote adequate seed production to sustain the population during the lifetime of an individual *S. subcuneata* (if regeneration opportunities were not limiting), the potential for further hybridisation is low without *S. torminalis.* Whilst our

study site contains the largest group of *S. subcuneata* (*c*. 300 individuals), at other sites it exists as only a few trees and the risk of extinction on those sites is therefore high. It is unknown if other local tetraploids are effective males and this could be important since it only occurs with *S. admonitor* on the single site.

Reproductive failure has profound implications for the long term survival of this population of triploids. Triploids may be seen as transient entities, in an evolutionary sense, due to low relative fitness (as discussed in Chapter 5), performing a 'stepping stone' role to further polyploid formation and diversification (as seen in Chapter 3). However, S*. torminalis* and *S. margaretae* no longer co-occur, so ongoing production of *S. subcuneata* is unlikely in the current situation. This makes conservation of the current population of *S. subcuneata* a priority. It needs to be managed in tandem with other compatible pollen donors to ensure sufficient seed production will offset threats to seedling establishment and survival.

6.3   AIM 3. Determine if the current approach to the conservation of the individual threatened species is appropriate or if the process-based approach advocated for TCG's will better conserve the diversity of *Sorbus* in this region.

To address this aim I first discuss process-based approaches to the conservation of other taxonomically complex groups then how it might be applied to this group in the light of our findings.

6.3.1   A process-based approach to the conservation of taxonomically complex groups

Plant conservation aims to protect biodiversity both in its present state but also to ensure there is the genetic capacity for diversification and adaptation in a

changing environment. For taxonomically complex groups (TCG's) this can be far from straightforward. High levels of endemism, often in specific habitats combined with small population sizes are features of many TCG's largely as a consequence of rapid evolutionary divergence (Rich *et al.*, 2008, Ennos *et al.*, 2012). A process-based approach is advocated for some of these groups where the production of novel lineages is ongoing and diversification is associated with particular areas which allow for the development of a targeted conservation programme (Ennos *et al.*, 2012). Conservation should be a balance between conserving the sources of novel lineages to preserve the ongoing process of diversification and conserving local lineages with particular adaptations or features where appropriate. These concepts apply to actively diversifying groups such as *Euphrasia, Dactylorhiza* and *Sorbus*. In *Euphrasia,* rare endemics are produced through hybridisation of widespread largely selfing and outcrossing species. Backcrossing produces hybrid swarms in places and identification of all individual lineages is impractical. Endemics appear adapted to particular microhabitats. Currently advocated conservation is to target 'hot spot' sites where *Euphrasia* biodiversity but also environmental heterogeneity is high, providing opportunities for interactions between the widespread progenitors and allowing natural selection of lines for adaptation to the various microhabitats (Ennos *et al.*, 2005, Stone, 2012). Hybridisation between diploids and autotetraploids in European *Dactylorhiza* is frequent and has resulted in a diverse array of allotetraploids, some of which are also involved in subsequent hybridisations. The conservation advocated to preserve ongoing diversification within *Dactylorhiza* is to focus on the progenitors in particular geographic regions where genetic diversity is high (southern Greece, Alps and northern

Sweden) and habitats such as base-rich fens (Nordström & Hedrén, 2009, Ennos *et al.*, 2012) rather than conserving individual taxa.

Ennos *et al.* (2012) developed a detailed process-based action plan for the *Sorbus* of Arran, Scotland. This underlines some of the key steps in developing and undertaking such a plan which apply to other complex groups and a few points are outlined below for comparative purposes. The *Sorbus* system is well understood and relatively simple; hybridisations between two widespread species of least concern (Rich *et al.*, 2010), diploid *S. aucuparia* and *S. rupicola,* have produced triploid *S. arranensis.* Subsequent back crosses to *S. aucuparia* have produced tetraploid *S. pseudofennica* and triploid *S. pseudomeinichii.* These individual species have arisen on several occasions (Robertson *et al.*, 2004b). One of the key issues is that a main progenitor, S. *rupicola* only occurs as one tree approximately 14.5km south of the current populations of hybrids. Distribution, parentage and the recurrence of hybrid events suggest it has been much more widespread. Therefore one action is to re-introduce *S. rupicola* into closer proximity to the hybrid species to allow resumption of hybridisation to generate *S. arranensis.* Proximity of progenitors as well as site management to promote flowering, pollination and seedling recruitment are all advocated. Monitoring on Arran would allow for adaptive management and focus on outcomes such as fruiting, recruitment and the presence of individuals with a range of morphological features indicative of ongoing production of novel lineages rather than the taxing identification of each species and variant.

For all these cases there is an emphasis on conservation of the progenitors in areas of contact to promote interaction. Targeted management of these sites

may also be required. An underlying understanding of the processes that generate diversity and identification of threats to that process is required for successful conservation. There may be less emphasis on maintaining particular lineages that may be transient 'stepping stones' to further diversification.

6.3.2 Conservation of *Sorbus* diversity in Devon and north Somerset.

The current diversity of polyploid *Sorbus* species in Europe has arisen through hybridisation and is maintained via apomixis (Lemche, 1999, Nelson-Jones *et al.*, 2002, Robertson *et al.*, 2004a, Lepší *et al.*, 2008, Robertson *et al.*, 2010). The results we have presented in this study indicate that this is also the case for our study group. Therefore, rather than conserving the individual products of an evolutionary process, species in this case, we advocate conservation of the processes that have resulted in the current diversity of *Sorbus.* Therefore, we need to focus on maintaining (or restoring) the conditions that promote interaction among the various species to maximise opportunities for hybridisation and for the establishment of newly formed polyploids.

Whilst some of the species occur on many sites and have wider distributions than the north Devon/ Somerset coast, this is where the four endemic species are found and where *Sorbus* diversity is highest in the region. Site based conservation focussed on those with multiple species represents a pragmatic approach as many of the key sites are designated SSSI's with rare *Sorbus* as conservation priorities on those sites (Table 1.1). However, an overarching plan that would operate at the landscape scale would reflect the scale at which these processes have taken place and thought should be given to increasing connectivity between these sites. Critical aspects of a process-based action plan for *Sorbus* in this region are to ensure the current diversity is maintained

through management that promotes reproductive sustainability; ensure that there are opportunities for on-going hybridisation and speciation by maintaining or restoring a diversity of species in close spatial proximity to ensure compatible pollination. Monitoring targets should be focussed on these aspects rather than purely population size.

The findings from this study have highlighted a number of key considerations that apply to this specific group. We have shown that pollen flow occurs over relatively short distances, this is similar to the findings in the Avon Gorge (Ludwig, 2013) however, the Devon / Somerset species occur over a much more extensive area, often at low densities in a range of habitats. We also show that in the woodland situation the majority of pollination is geitonogamous. To increase interspecific pollen flow, critical for the survival of triploid *S. subcuneata* and essential for gene flow among species, management to increase flowering and fruiting should be targeted where multiple species occur. Such management could consist of thinning around individual trees to allow more light for flowering, however, re-instating woodland management such as coppicing (cutting tree species to a stump or stool at ground level) or targeted clear felling of competing tree species will open up larger areas also creating opportunities for seedling establishment. More open conditions may allow development of other flowering species to support pollinator populations. Many of the cliff sites are more open but large landslips have obliterated significant groups of endemic *Sorbus* (pers. obs) over recent years. Whilst this is a natural occurrence and probably provides regeneration opportunities, when combined with habitat fragmentation caused by *Rhododendron ponticum* it may present a serious threat. A combination of continuing eradication of *Rhododendron ponticum* with some re-planting of saplings from local stocks may offset this

256

threat but also create 'stepping stone' groups for occasional longer distance pollen flow, re-connecting some of the closer sites.

Our results show that the process underpinning the reticulate evolution of *Sorbus* in this region is the hybridisation of diploid *S. torminalis* and tetraploid *S. margaretae* to create triploid *S. subcuneata* followed by subsequent hybridisations also involving *S. torminalis*. *S. torminalis* is a widespread species of least conservation concern. A serious consideration to promote further diversification would be to re-introduce *S. torminalis* into some of the sites where *S. subcuneata* and *S. margaretae* are found. Based on our parentage data *S. torminalis* must have been present at some of these sites in the past. Some research may be required to select areas best suited to the ecological requirements of S*. torminalis* before any planting takes place. Indeed, we do not know why it disappeared and *S. torminalis* may have been at the limit of its ecological tolerance on the sites where the polyploids are found. Sites where polyploids are in greater abundance than diploids promote interploidy sexual hybridisation (Hajrudinović *et al.*, 2015) with the diploid as the maternal parent, especially since the diploid species are largely self-incompatible (Ludwig *et al.*, 2013) therefore low rather than high density S*. torminalis* may be more effectual at producing hybrid offspring.

An effective conservation policy would combine aspects of species-based conservation, including some re-introduction, with a wider process-based approach in landscape scale context.

6.4   Conclusion

The present study has demonstrated that this group of species has provided a good model for investigating how polyploid evolution and routes of formation

257

occur at a regional scale. The evidence for past evolutionary events has given insight into historic distribution ranges of progenitor species which has implications for the production of new polyploids. This thesis provides evidence for the 'triploid bridge' as a significant route for allotetraploid formation but that reproductive sustainability of triploids is at risk when compatible pollen is limited. Therefore triploid range expansion can only occur in tandem with that of compatible pollination partners and is the most likely reason for the dwindling numbers of isolated groups of triploids in our study. Findings from this study will be used to inform a conservation approach that can enable conservationists to protect and sustain beautiful endemic species' that possess such a complex and interesting evolutionary history

6.5   Future research directions

There remains much further work that could be done both to extend the work done in this thesis and using *Sorbus* as a model to answer broader questions about polyploidy and the ecological implications of polyploidy. Further work is needed to clarify the role of the diploid progenitors in polyploid formation. There appear to be regional patterns and *Sorbus* hotspots for polyploid diversity yet progenitors co-occur at other sites where hybridisation has not occurred. Further pollen compatibility experiments as in chapter 5, along with wider genetic screening of the diploids at more high diversity areas such as some of the Welsh sites, may allow greater understanding of why some sites appear to be particularly diverse. We have shown that spontaneous triploid forms of diploids occur, and a useful direction of study would be to ascertain their pollen fertility and compatibility with other species to investigate whether this is a possible route for further polyploid formation.

Subgenus *Aria* is of great evolutionary interest. Whilst our use of microsatellite markers was able to illuminate relationships among study members of this group, questions remain regarding the extent of tetraploid hybridisation for example. Further analysis of pollen flow on the high diversity sites coupled with increased sampling at these sites may reveal both potential for, and actual *Sorbus* diversification via this route. Investigations into spatial patterns of genetic variation could give clues to colonisation routes. However, for both these areas of research the use of molecular markers with greater resolving power would be required to elucidate variation among individuals of the same species. The rapid development of sequencing technology combined with reducing costs now makes the identification of large numbers of microsatellites more achievable; however there is increasing use of single nucleotide polymorphisms (SNP's) which would potentially give greater resolution, although there are still problems caused by the number of repeated elements in polyploids as reviewed by Dufresne *et al.* (2014).

The ecological implications of polyploidization are of great interest and this area has not been extensively studied in *Sorbus*. Many sites of high *Sorbus* diversity also have high environmental heterogeneity. It is unknown whether these factors are linked and closer examination of ecological niche diversity among the study species may help shed light on the ecological implication of *Sorbus* evolution. This would also be of importance in the conservation of *Sorbus* diversity especially where re-planting for mitigating loss of rare species is considered. Ecological niche modelling with increasingly fine scale models (Maclean *et al.*, 2016) could be used to investigate how environmental change may affect the distribution of current *Sorbus* ecological niches and how this

would affect potential for future survival but also hybridisation opportunities (Vallejo-Marín & Hiscock, 2016).

References

Chester M, Cowan RS, Fay MF, Rich TCG. 2007. Parentage of endemic *Sorbus* L. (Rosaceae) species in the British Isles: evidence from plastid DNA. *Botanical Journal of the Linnean Society,* **154**: 291-304.

Dufresne F, Stift M, Vergilino R, Mable BK. 2014. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Molecular Ecology,* **23**: 40-69.

Ennos R, French G, Hollingsworth P. 2005. Conserving taxonomic complexity. *Trends in Ecology & Evolution,* **20**: 164-168.

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

Hajrudinović A, Siljak-Yakovlev S, Brown SC, Pustahija F, Bourge M, Ballian D, Bogunić F. 2015. When sexual meets apomict: genome size, ploidy level and reproductive mode variation of *Sorbus aria s.l.* and *S. austriaca* (Rosaceae) in Bosnia and Herzegovina. *Annals of Botany,* **116**: 301-312.

Lemche EB. 1999. The origins and interactions of British *Sorbus* species*,* Unpublished PhD thesis, University of Cambridge, Cambridge.

Lepší M, Lepší P, Koutecky P, Bila J, Vit P. 2015. Taxonomic revision of *Sorbus* subgenus *Aria* occurring in the Czech Republic. *Preslia,* **87**: 109-162.

Lepší M, Ví t P, Lepší P, Boublík K, Suda J. 2008. *Sorbus milensis*, a new hybridogenous species from northwestern Bohemia. *Preslia,* **80**: 229-244.

Ludwig S, Robertson A, Rich TCG, Djordjevic M, Cerovic R, Houston L, Harris SA, Hiscock SJ. 2013. Breeding systems, hybridization and continuing evolution in Avon Gorge *Sorbus. Annals of Botany*.

Ludwig SC. 2013. Breeding systems, pollen flow and continuing evolution in Avon Gorge *Sorbus* (whitebeams, rowans and service trees)*,* PhD thesis, University of Bristol, Bristol.

261

Maclean IMD, Suggitt AJ, Wilson RJ, Duffy JP, Bennie JJ. 2016. Fine-scale climate change: modelling spatial variation in biologically meaningful rates of warming. *Global Change Biology*. Doi: 10.1111/gcb.13343

Nelson-Jones E, Briggs D, Smith A. 2002. The origin of intermediate species of the genus *Sorbus. Theoretical and Applied Genetics,* **105**: 953-963.

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

Rich TCG, Houston L, Robertson A, Proctor MCF. 2010. *Whitebeams, rowans and service trees of Britain and Ireland. A monograph of British and Irish* Sorbus *L. B.S.B.I. Handbook No. 14.* London: Botanical Society of the British Isles.

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

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

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

Robertson A, Rich TCG, Allen AM, Houston L, Roberts CAT, Bridle JR, Harris SA, Hiscock SJ. 2010. Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus. Molecular Ecology,* **19**: 1675-1690.

Stone H. 2012. The evolution and conservation of tetraploid *Euphrasia* L. in Britain*,* PhD PhD, University of Edinburgh.

Vallejo-Marín M, Hiscock SJ. 2016. Hybridization and hybrid speciation under global change. *New Phytologist,* **211**: 1170-1187.

Appendix: Herbarium voucher specimens.

Mounted and scanned by the Welsh National Herbarium, Amgueddfa Cymru – National Museum Wales, Cardiff (NMW).

S. torminalis, Wild Service-tree. Beaford, north Devon.

S. torminalis triploid form, South Tawton, north Devon. Note 'bulky' appearance of leaves which had a leathery texture.
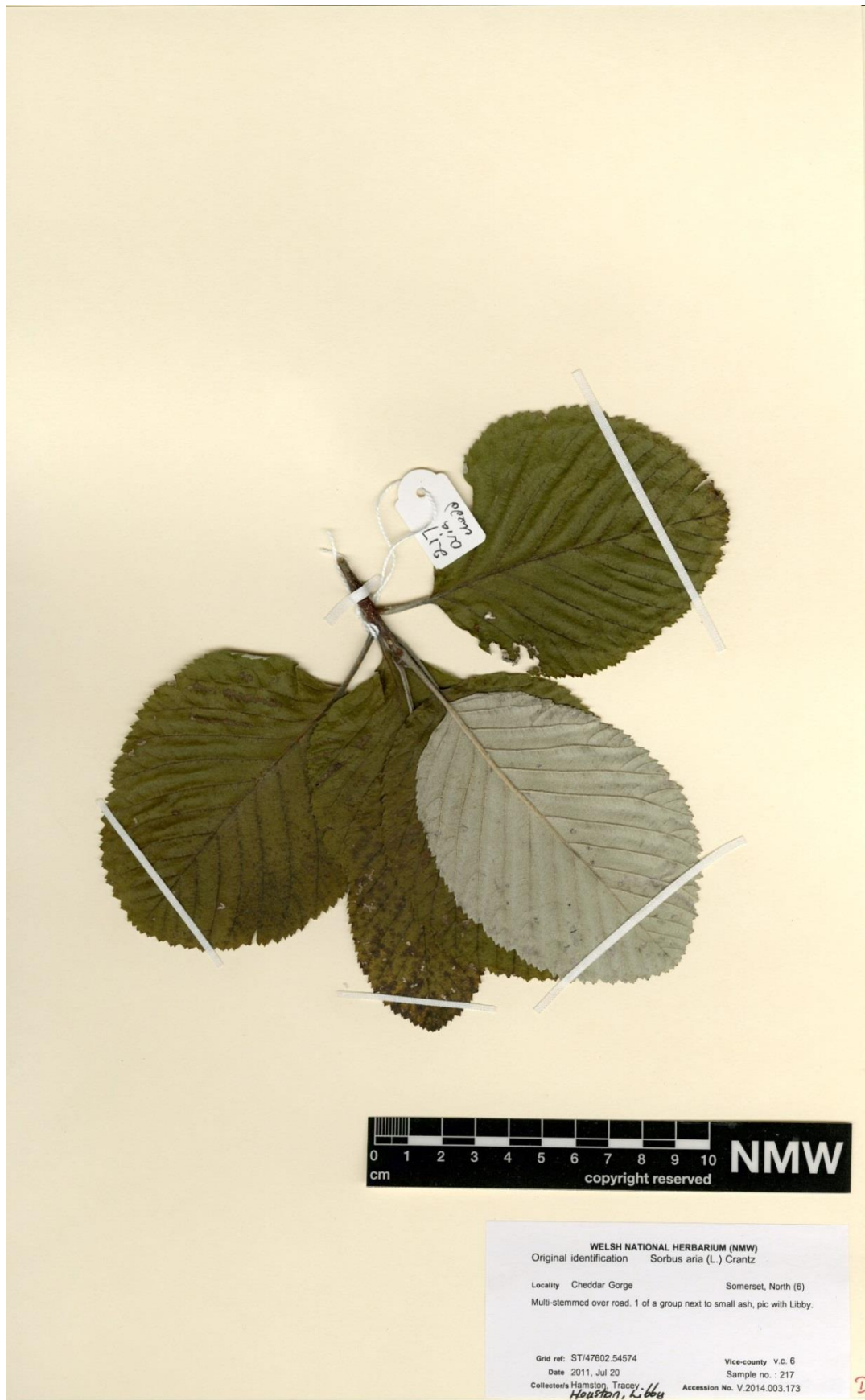
S. aria, Common Whitebeam. Cheddar Gorge, Somerset.

dev.
229
fr.

WELSH NATIONAL HERBARIUM (NMW)
Original identification   Sorbus devoniensis E.Warb.

Locality   Little Haldon   Devon, South (3)
North side of roadd nearr 92 and 93, see map NB.

Grid ref: SX/91446.75086
Date   2011, Sep 29   Vice-county V.C. 3
                                        Sample no. : 229
Collector/s Hamston, Tracey   Accession No. V.2014.003.086

S. devoniensis, Devon Whitebeam. Little Haldon, south Devon.

WELSH NATIONAL HERBARIUM (NMW)
Original identification   Sorbus admonitor M.C.F.Proctor

Locality   Watersmeet   Devon, North (4)
Next to 158, NGR taken from 158.

Grid ref: SS/743.488   Vice-county v.c. 4
Date   2012, Jul 23   Sample no. : 238
Collector/s Hamston, Tracey   Accession No. V.2014.003.163

S. admonitor, 'No Parking' Whitebeam. Watersmeet, north Devon.

267

S. subcuneata, Slender Whitebeam. Watersmeet, north Devon.

S. rupicola, Rock Whitebeam. Churston Cove, south Devon.

S. porrigentiformis, Grey-leaved Whitebeam*.* Redgate Beach, south Devon.

WELSH NATIONAL HERBARIUM (NMW)
Original identification   Sorbus vexans E.F. Warb.

Locality   Oxen Tor   Devon, North (4)

Single stemmed tree dividing into 2 @ 50cm. Approx 4m below138 - see
map in TH Nbook.

Grid ref: SS/72895.49017   Vice-county v.c. 4
Date 2010, Aug 23   Sample no. : 133
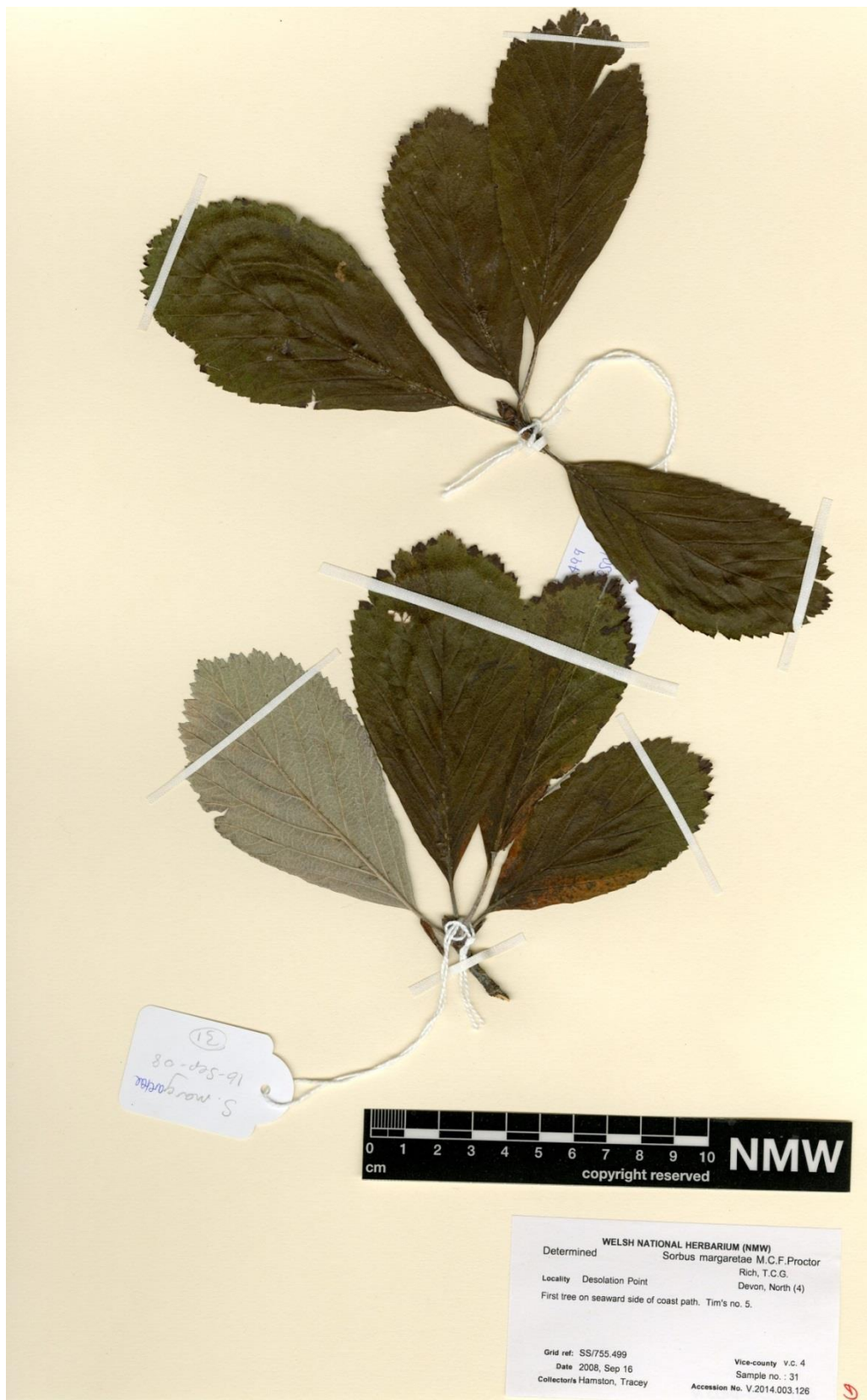Collector/s Hamston, Tracey   Accession No. V.2014.003.023

S. vexans, Bloody Whitebeam. Oxen Tor, nr Watersmeet, north Devon.

S. vexans (vex2) 2nd clone. Neck Wood, north Devon.

S. margaretae, Margaret's Whitebeam. Desolation, north Devon.