# "Evolutionary consequences of hybridization between *Geum urbanum* and *G. rivale*"

**Markus Ruhsam**

PhD Thesis
University of Edinburgh
2009

## Declaration

I, Markus Ruhsam, hereby declare that this thesis represents entirely my own work unless otherwise clearly acknowledged. This thesis has not been previously presented for any other degree or professional qualification and was conducted between February 2006 and February 2009.

Markus Ruhsam

Edinburgh, March 2009

# Acknowledgements

I am very much indebted to Richard Ennos and Pete Hollingsworth for their excellent supervision and support throughout the whole project. When I write 'excellent' then I mean excellent because I can not think of a way they could have done it better. This is not to be taken for granted as I have seen other PhD students struggling through their projects without much support from their supervisors, wondering how they will ever be able to finish. Richard was very encouraging and generous with his time to explain yet another thing I should have known about already and his constructive comments expanded my horizon on population genetics and evolution considerably. Pete has the admirable ability to distil ideas and difficult concepts into clear and well structured sentences and cut out a lot of 'blethering' in this thesis (no doubt my examiners will think 'not enough' after reading through it). He never hesitated to fit in meetings in an already very busy time schedule to give me feedback and answer my questions.

I am also indebted to many other people without whom this PhD would have been a lot more traumatic than it actually was. In particular the help of Will Goodall-Copestake, Michelle Hollingsworth and Sabrina Reignoux was vital at the start of my PhD when I was getting to grips with molecular techniques and lab related stuff. Special recognition goes to Alan Forrest for his help in my first year. Jane Squirrell was a star throughout and helped with many things from offering me her shoulder to cry on when AFLPs were not working to formatting data so that the annoying software program would finally accept my input file. Additionally, she managed to isolate microsatellites for *G. rivale* which will save a lot of time and effort for a follow up project. A big thank you goes to Dario Beraldi for WORD related questions, playing chess, drinking endless cups of green tea and for playing classical guitar together. Lel Eory was indispensable for keeping me physically fit with squash and Hungarian cake specialities.

Special recognition is given to Billy Adams and staff for looking after my plants at the University as well as Pete Brownless and staff who looked after them at the nursery in the Royal Botanic Garden Edinburgh. Many other people gave valuable advice, practical assistance and support amongst them Peter Schönswetter and Andi Tribsch (for sending me *Geum* samples from throughout Europe), Deborah Charlesworth, Silvia Perez Espona, Roberta Bergero, Jill Hamilton, David Marsden, Mike Church and Caroline Robertson. Finally, a big thank you to Emma for the time we shared.

Für meine Eltern die mich in all meinen Unternehmungen immer unterstützt haben.

---

Geheimnißvoll am lichten Tag,
Läßt sich Natur des Schleiers nicht berauben,
Und was sie deinem Geist nicht offenbaren mag,
Das zwingst du ihr nicht ab mit Hebeln und mit Schrauben.

Mysterious, even in broad daylight
Nature won't let her veil be raised,
And secrets not revealed freely to your spirit,
Won't be extorted from her by using screws or levers [or 454 sequencing].

Faust I, Johann Wolfgang von Goethe

The inherent limitation of all predictions (Heisenberg uncertainty principle)

$$\Delta x \, \Delta p \geq \frac{\hbar}{2}$$

## ABSTRACT

Theoretical considerations predict that hybridization between plants with different breeding systems could potentially be a powerful evolutionary force but very few studies have been carried out to investigate hybridization among outcrossing and selfing species in natural populations. The objective of this thesis was to use *G. urbanum* and *G. rivale*, two freely hybridizing species with putatively different breeding systems, and study the processes that take place when the two different mating systems meet. To establish the extent of morphological and genetic marker differences, the breeding system and levels of inbreeding depression in these two species, I examined reference populations from various locations in Great Britain using morphological (nine phenotypic traits) and genetic tools (AFLP and microsatellite markers) and also undertook crossing experiments. These investigations provided the framework for the morphological and genetic analyses of a long standing hybrid swarm between *G. urbanum* and *G. rivale*. The results revealed that the two parental species differed significantly in all nine traits measured and were clearly separated by 616 AFLP markers, of which 8% were species-specific. The breeding system analysis confirmed the highly selfing nature of *G. urbanum* and the outcrossing nature of *G. rivale*. No inbreeding depression was discovered in *G. urbanum* and only very low levels in *G. rivale*. The AFLP analysis (203 loci) of the hybrid swarm showed that four distinct genetic groups can be distinguished (1) a parental *G. urbanum* like group (2) an F1 like group (3) a Backcross Rivale group and (4) a parental *G. rivale* like group. No Backcross Urbanum individuals were detected. Morphological data were highly correlated with genetic data (Spearman Rank Order Correlation = 0.80, $p<0.0001$) but less informative as morphological data could not separate the Backcross Rivale samples from *G. rivale*. Analysis of seed progeny from the hybrid swarm showed that there were clear differences in the genotypic composition of the seed progeny and the plants from which they were collected, revealing additional genetic classes such as Backcross Urbanum individuals which were not represented in the adult plant stage. The Backcross Rivale parental group was the only one in which the average offspring position did not differ significantly from the mean parental position in a PCO plot using 203 AFLP markers, indicating that this group might be reproductively isolated from other groups in the hybrid swarm. Taken together the results suggest that hybridization between *G. urbanum* and *G. rivale* results in the production of a range of genotypic and phenotypic classes, some of which are selected against. There is evidence for introgression between these two species, and tentative evidence for partial reproductive isolation of some hybrids suggesting that hybridisation between these outcrossing and selfing *Geum* species may represent a creative evolutionary force.

**1. INTRODUCTION ..... 1**

**1.1 General introduction ..... 1**
**1.2 Hybridization ..... 3**
1.2.1 Extinction ..... 3
1.2.2 Introgression ..... 4
1.2.3 Speciation ..... 5
1.2.4 Stalemate ..... 8
**1.3 Hybridization between two different mating systems ..... 8**
1.3.1 Transgressive segregation ..... 9
1.3.2 Inbreeding depression ..... 10
1.3.3 Previous studies ..... 11
**1.4 Motivation for this research ..... 13**
**1.5 Study species ..... 14**
1.5.1 Geum urbanum ..... 14
1.5.2 Geum rivale ..... 15
1.5.3 Geum urbanum x G. rivale ..... 18
**1.6 Aims and objectives ..... 21**

**2. ANALYSIS OF PURE *GEUM URBANUM* AND *G. RIVALE* POPULATIONS ..... 23**

**Abstract ..... 23**
**2.1 Introduction ..... 24**
**2.2 Materials and Methods ..... 28**
2.2.1 Sampling material for morphological analysis ..... 28
2.2.2 Sampling material for genetic marker analysis ..... 28
2.2.3 Quantifying morphological variation ..... 28
2.2.4 Analysis of morphological data ..... 29
2.2.5 Genetic marker analysis ..... 31
2.2.6 Analysis of genetic differentiation ..... 34
**2.3 Results ..... 35**
2.3.1 Morphological variation ..... 35
2.3.2 Genetic marker differentiation ..... 38
**2.4 Discussion ..... 42**
2.4.1 Morphology ..... 42
2.4.2 Genetic markers ..... 43
**2.5 Summary ..... 46**

***3.* BREEDING SYSTEM OF *GEUM URBANUM* AND *G. RIVALE* ..... 47**

**Abstract ..... 47**
**3.1 Introduction ..... 48**
**3.2 Materials and Methods ..... 55**
3.2.1 Sampling material ..... 55
3.2.2 Auto-pollination and SI ..... 56
3.2.3 Microsatellites ..... 57
3.2.4 Data analysis ..... 58
**3.3 Results ..... 60**
3.3.1 Auto-pollination and SI ..... 60
3.3.2 Microsatellites ..... 61
**3.4 Discussion ..... 65**
3.4.1 Auto-pollination and SI ..... 65
3.4.2 Outcrossing rates ..... 66

3.4.3 GENETIC DIFFERENTIATION ..... 68
3.4.4 SUMMARY ..... 69

**4. INBREEDING DEPRESSION IN *G. URBANUM* AND *G. RIVALE* .....70**

**ABSTRACT ..... 70**
**4.1 INTRODUCTION ..... 71**
4.1.1 GENETIC HYPOTHESES OF INBREEDING DEPRESSION ..... 71
4.1.2 MEASURING INBREEDING DEPRESSION ..... 74
4.1.3 RATIONALE FOR THIS CHAPTER ..... 78
**4.2 MATERIALS AND METHODS ..... 79**
4.2.1 SAMPLING MATERIAL ..... 79
4.2.2 CONTROLLED CROSSES ..... 79
4.2.3 INBREEDING COEFFICIENT $F_{IS}$ ..... 80
4.2.4 DATA ANALYSIS ..... 81
**4.3 RESULTS ..... 82**
4.3.1 *GEUM URBANUM* ..... 82
4.3.2 *GEUM RIVALE* ..... 84
**4.4 DISCUSSION ..... 87**
4.4.1 *GEUM URBANUM* ..... 87
4.4.2 *GEUM RIVALE* ..... 88
4.4.3 SUMMARY ..... 90

**5. CONTROLLED F1S BETWEEN *GEUM URBANUM* AND *G. RIVALE* .....91**

**ABSTRACT ..... 91**
**5.1 INTRODUCTION ..... 92**
**5.2 MATERIALS AND METHODS ..... 94**
5.2.1 SAMPLE MATERIAL ..... 94
5.2.2 NON MOLECULAR DATA ..... 94
5.2.3 MOLECULAR DATA ..... 97
**5.3 RESULTS ..... 99**
5.3.1 NON MOLECULAR DATA ..... 99
5.3.2 MOLECULAR DATA ..... 102
**5.4 DISCUSSION ..... 105**
5.4.1 SUCCESS AND GERMINATION RATES OF CONTROLLED F1 CROSSES ..... 105
5.4.2 MORPHOLOGY ..... 106
5.4.3 POLLEN FERTILITY ..... 106
5.4.4 CHLOROPLAST INHERITANCE ..... 107
5.4.5 AFLP MARKER INHERITANCE PATTERNS ..... 107
5.4.6 SPECIES-SPECIFIC MARKERS AND NEWHYBRIDS ..... 109
5.4.7 SUMMARY ..... 110

**6. GENETIC COMPOSITION OF A HYBRID SWARM BETWEEN *GEUM URBANUM* AND *G. RIVALE* .....111**

**ABSTRACT ..... 111**
**6.1 INTRODUCTION ..... 112**
6.1.1 EXTINCTION ..... 112
6.1.2 INTROGRESSION ..... 113
6.1.3 SPECIATION ..... 114
6.1.4 STALEMATE ..... 115
6.1.5 HYBRIDIZATION BETWEEN SELFING AND OUTCROSSING TAXA ..... 116
**6.2 MATERIALS AND METHODS ..... 120**

6.2.1 Sampling material ..... 120
6.2.2 Molecular variation ..... 123
6.2.3 Non molecular variation ..... 124
6.2.4 Analysis ..... 126
**6.3 RESULTS ..... 128**
6.3.1 Molecular data ..... 128
6.3.2 Attributes of genetic groups ..... 132
**6.4 Discussion ..... 141**
6.4.1 AFLP error rates ..... 141
6.4.2 Genetic groups present in the hybrid swarm ..... 141
6.4.3 Chloroplast haplotypes ..... 142
6.4.4 Morphology ..... 142
6.4.5 Transgressive segregation ..... 143
6.4.6 Pollen fertility and auto-pollination ..... 145
6.4.7 Patterns of hybridization ..... 146
6.4.8 Summary ..... 148

**7. DYNAMICS OF THE HYBRID SWARM ..... 149**

**Abstract ..... 149**
**7.1 Introduction ..... 150**
**7.2 Materials and Methods ..... 153**
7.2.1 Sample material ..... 153
7.2.2 Germination rates ..... 153
7.2.3 Diversity of offspring generation ..... 154
7.2.4 Offspring performance ..... 155
**7.3 Results ..... 156**
7.3.1 Germination rates ..... 156
7.3.2 Hybrid offspring diversity ..... 156
7.3.3 Offspring performance ..... 167
7.3.4 Summary of the main results ..... 170
**7.4 Discussion ..... 171**
7.4.1 Separation along PCO2 (PCO2 top group) ..... 171
7.4.2 Hybrid offspring diversity ..... 173
7.4.3 Correlation between morphology and genetics ..... 178
7.4.4 Fitness of offspring ..... 180
7.4.5 Summary ..... 181

**8. SYNTHESIS ..... 182**

**8.1 Main findings of this study ..... 182**
**8.2 What are the possible evolutionary consequences? ..... 183**
8.2.1 Extinction ..... 183
8.2.2 Introgression ..... 183
8.2.3 Speciation ..... 186
8.2.4 Stalemate ..... 188
**8.3 Further research ..... 188**
**8.4 Conclusions ..... 191**

**REFERENCES ..... 193**

# 1. INTRODUCTION

## 1.1 General introduction

Our planet sustains an extraordinarily rich variety of life. Understanding this diversity and explaining its origin continues to be a central issue in evolutionary biology. The question 'Why are there so many species' runs like a red thread through many fields of the biological sciences (Burger, 1981; May, 1992; Simmons & Cowling, 1996; Galis & Metz, 1998; Pennings & van Opijnen, 2002; Morrell, 2005; Novotny *et al.*, 2006; Mayhew, 2007; Ruggiero & Hawkins, 2008; Crepet & Niklas, 2009). In particular the high number of flowering plant (angiosperm) species has puzzled biologists since Darwin and sparked many studies investigating the factors which might have caused such biodiversity to evolve (Darwin, 1859; Darwin, 1903; Burger, 1981; Hill & Hill, 2001; Leigh *et al.*, 2004; Aarssen *et al.*, 2006; Linder, 2008; Crepet & Niklas, 2009). The origin of the angiosperms is still controversial and many theories have been put forward as to their ancestry (reviewed in Stuessy, 2004). However, there is consensus that angiosperms are clearly united by a suite of synapomorphies (i.e., shared, derived features), including double fertilization and endosperm formation, the carpel, stamens with two pairs of pollen sacs, features of gametophyte structure and development, and phloem tissue composed of sieve tubes and companion cells. This evidence strongly suggests that all extant angiosperms share a common ancestor and do not have polyphyletic origins (Soltis & Soltis, 2004). The fossil record of the angiosperms dates back to the early Cretaceous, indicating that flowering plants are at least 130 million years old (Crane *et al.*, 2004; Stuessy, 2004).

The debate as to the exact figure of how many angiosperm species there are is still ongoing (Ungricht, 2004) and estimates range from about 250,000 species (Mabberley, 1997) to 300,000-320,000 (Prance *et al.*, 2000) to more than 400,000 (Govaerts, 2001). It might come as a surprise that the species estimates for angiosperms differ by more than 60%. Lack of communication between taxonomists (mainly before the invention of computers) and differing and changing views on the distinctness of a plant taxon to be described ('splitters and lumpers') has led to more than 1,000,000 published scientific names for seed plants, many of which are synonyms, i. e. multiple names for a single biological entity (Scotland & Wortley, 2003). How many of the binomials are synonyms, however, is uncertain because many plant taxa have not been subjected to taxonomic revisions leading to the discrepancy in species number estimates. However, it is undisputed that angiosperms have become the largest and most successful group of embryophytes since their origin, outcompeting and outnumbering spore-bearing plants and gymnosperms. They have evolved to occupy and dominate nearly

every habitat on Earth and have developed an unparalleled diversity in chemistry, reproductive morphology, genome size and organization in the plant kingdom (Soltis & Soltis, 2004). Numerous hypotheses have been put forward to explain why there are so many angiosperm species compared to other terrestrial or aquatic plants and why the flowering plants are so widespread ecologically (see references in Crepet & Niklas, 2009). Crepet & Niklas' (2009) survey of the literature indicates that no single vegetative, reproductive, or ecological feature taken in isolation can account for the evolutionary success of the angiosperms. They suggest that a key strength of the angiosperms is their capacity to "reinvent themselves" by mixing old vegetative and reproductive traits into different combinations in addition to their ability to innovate new traits throughout their history (Friis *et al.*, 2006; Crepet & Niklas, 2009).

Although biological diversity is measured in numbers of species (Diamond, 1985), "it is almost universally conceded that no satisfactory definition of what constitutes a species has ever been proposed" (Dobzhansky, 1935). After a multitude of different species concepts (Coyne & Orr (2004) discuss nine concepts) and more than 70 years later the definition of what exactly is a species still remains a contentious issue (Abbott *et al.*, 2008). Mayr's (1942) 'biological species concept' which defines species as reproductively isolated groups of populations is perhaps the most widely accepted among evolutionary biologists (Coyne & Orr, 2004). Complete reproductive isolation is at the centre of the biological species concept, indicating that diverging populations must be protected from gene flow throughout their genomes to be considered 'good species'. However, the fact that many morphologically and genetically distinct species continue to exchange genetic material long after they have embarked on different evolutionary trajectories has been known for many decades resulting in the widespread rejection of the biological species concept by botanists (Mishler, 1999; Rieseberg *et al.*, 2003a). It is now clear that divergence and speciation can occur in the presence of gene flow as long as the strength of selection exceeds the migration rate (Maynard Smith, 1966; Rieseberg *et al.*, 2003a).

This can be explained by the genic view of speciation which is based on the notion that species isolation may be achieved by a modest number of genes (Wu, 2001). Such 'speciation' genes do not move freely between species because they are selected against in a 'foreign' genetic background (Wu & Ting, 2004). In the genic model differential adaptations to different natural environments at the gene level are the driving force behind speciation. Wu (2001) therefore defines species as groups that are differentially adapted and, upon contact, are not able to share genes controlling these adaptive characters, by direct exchanges or through intermediate hybrid populations. Complete reproductive isolation is therefore not a requirement for species status anymore but remains important in that

adaptations and evolutionary divergence among species are preserved (Futuyma, 1989; Rieseberg *et al.*, 2003a).

Species are normally prevented from crossing via reproductive barriers which can be either pre- or postzygotic. Prezygotic barriers are gene-flow barriers or incompatibilities that act to prevent successful fertilization such as ecological (Emms & Arnold, 1997), temporal (Stace, 1975) and ethological/mechanical barriers (Grant, 1994), as well as gametic competition or incompatibility (Carney *et al.*, 1994; Rieseberg *et al.*, 1995). In contrast, postzygotic barriers are gene flow barriers that act after the successful fertilization and formation of a zygote, in either the first or subsequent hybrid generations. Common postmating barriers include hybrid weakness or inviability, hybrid sterility and hybrid breakdown in which first generation (F1) hybrids are robust and fertile, but later generation hybrids are weak or inviable (see Rieseberg, 1997 and examples therein). The breakdown of prezygotic reproductive barriers due to changes in selection or the environment which may lead to hybridization and/or the escape from postzygotic barriers might serve as one mechanism by which angiosperms are able to 'reinvent' themselves.

## 1.2 Hybridization

There are various definitions of the term 'hybridization' in the literature depending on the context in which it is used (Rieseberg & Ellstrand, 1993). For this study 'hybridization' refers to crosses, without the immediate manipulation by humans, between two distinct taxonomic entities which occur naturally in the wild. It has long been recognised that hybridization is an important force in angiosperm evolution and numerous studies have investigated its consequences. There can be four major outcomes: extinction, introgression, speciation or stalemate (Lotsy, 1916; Anderson & Stebbins, 1954; Lewontin & Birch, 1966; Grant, 1971; Rieseberg & Ellstrand, 1993; Arnold, 1997; Arnold *et al.*, 2001; Rieseberg *et al.*, 2003b; Sweigart & Willis, 2003; Karrenberg *et al.*, 2007; Minder *et al.*, 2007).

### 1.2.1 Extinction

Hybridization can lead to extinction by demographic swamping and genetic assimilation of one species by another (Levin *et al.*, 1996; Rhymer & Simberloff, 1996; Wolf *et al.*, 2001). This was first discussed by Harper *et al.* (1961). The growth rate of a less common species is slowed down by the production of hybrid seeds at the expense of conspecific seeds. This will also lead to a relatively higher percentage of hybrid seeds produced by the rarer species compared to the more common congener ('minority disadvantage'). The minority

disadvantage does not depend on the viability of hybrid seeds as the production of viable and not viable seeds will reduce the possible number of conspecific seeds (Levin, 2000). If the crossing barriers between two species are weak then this may result in a significant reduction of conspecific seeds of the rarer species followed by a decline in population size (Heiser, 1979; Warwick et al., 1989). Successfully established hybrids will further decrease the relative population size of the less common species and contribute to its genetic assimilation (Rieseberg *et al.*, 1989; Rieseberg & Gerber, 1995). This process of extinction by hybridization is a particular threat to island floras as previously geographically separated but interfertile species often have been brought into contact by man's activities (Levin et al., 1996).

## 1.2.2 Introgression

Introgression was first defined by Anderson and Hubricht (1938) as the exchange of genetic material facilitated by repeated backcrossing of fertile hybrids to the parental species. Since then the definition of 'introgression' has been subjected to slight alterations and is now commonly used for the permanent incorporation of genes from one set of differentiated populations (or species) into another (Rieseberg & Wendel, 1993). Rieseberg and Wendel (1993) list 165 proposed cases of introgression but since then many more (molecular) studies of hybridization have highlighted its widespread occurrence in plants (Howard *et al.*, 1997; Wolfe *et al.*, 1998a; Ellstrand *et al.*, 1999; Milne & Abbott, 2000; Petit *et al.*, 2003).

Introgression can help hybrids to colonise new habitats by transferring adaptive traits from one species to the other (Lewontin & Birch, 1966; Abbott, 1992). Klier *et al.* (1991) studied introgression in two orchid species, *Cypripedium pubescens* which mainly grows in woodlands and *C. candidum* which survives only on a small number of prairie remnants in the United States. Their data demonstrated the existence of bidirectional gene flow wherever the two species are sympatric and hybridize. Interestingly, *C. pubescens* has been previously noted to occasionally occur on prairies in Iowa and allozyme and morphological data as well as ecological considerations suggested that the prairie ecotype of *C. pubescens* may have arisen through introgression of genetic material from *C. candidum* (Klier *et al.*, 1991). This indicates that *C. pubescens* plants growing in prairie habitat are actually later-generation backcrosses to the prairie-dwelling *C. candidum* and may represent an example of the transfer of an adaptation (or suite of adaptations) which allowed these plants to survive in this new habitat.

In a similar study using evidence from rDNA and cpDNA data, Milne & Abbott (2000) detected that naturalized British hybrids between *Rhododendron ponticum* and *R. catawbiense* were significantly more abundant in Britain's coldest region, eastern Scotland, than elsewhere. This

could indicate that introgression from the cold tolerant North American *R. catawbiense* conferred improved cold tolerance to *R. ponticum*, a species from the warmer Iberian Peninsula, thus enabling the hybrids to survive in colder climates (Milne & Abbott, 2000).
Introgression may also increase genetic diversity (Percy & Wendel, 1990), reduce the strength of reproductive barriers (Stebbins & Daly, 1961; Meyn & Emboden, 1987) and lead to genetic assimilation or extinction in the long term (Rieseberg *et al.*, 1989; Rieseberg & Gerber, 1995).

## 1.2.3 Speciation

*Polyploid speciation*

Crosses between two species often produce sterile offspring due to different parental chromosome numbers and/or large genomic differences. The main reason for hybrid sterility is the problematic alignment of homologous chromosomes during hybrid meiosis. However, hybrid fertility can be restored by allopolyploidy. Allopolyploidy is the spontaneous doubling of chromosome numbers in organisms such as hybrids, which consist of two or more different genomes (in contrast to autoplolyploidy where the doubling occurs within populations of the same species). The hybrid is then at least partially reproductively isolated from the parents but can also reproduce sexually as aligning problems have been removed by the doubling of chromosomes. If there are also distinct morphological differences between hybrid and parents, a new species might be the outcome of the hybridization process.

The relevance of polyploidy in plant speciation has long been recognized (Müntzing 1936; Grant, 1971) and recent studies (Lowe & Abbott, 1996; Ainouche *et al.*, 2003; Guo *et al.*, 2006) and reviews (Ramsey & Schemske, 1998; Soltis *et al.*, 2003) have further highlighted its evolutionary importance (Soltis, 2005).

*Homoploid speciation*

Homoploid or recombinational speciation is the process of forming and establishing new recombinational hybrid lineages which have the same ploidy level as their parents. To avoid the effects of homogenising gene flow, the new hybrid lineage needs to be reproductively isolated from the parental species. One way this can be achieved is via the formation of an intermediate flower type through hybridization. This could lead to ethological and/or mechanical isolation where the new hybrid floral phenotype is adapted for a pollinator with different body size and shape, or pollen is picked up by different parts of the pollinator's body

compared to pollen from the parental taxa (Grant, 1949, 1994). It is also possible that the hybrid flower morphology has been altered so much that the parental pollinators avoid it altogether. Alternatively, reproductive isolation from the progenitors can be achieved via ecological separation where the new hybrid lineage is able to colonise new habitats which are unsuitable for the growth of either parent (Buerkle *et al.*, 2000; Lexer *et al.*, 2003).

If the two parental taxa differ by two or more chromosomal rearrangements then another mechanism to protect the new hybrid lineage from parental gene flow is possible (Stebbins, 1957c; Grant, 1958). The main process underlying this model can be explained by the following example (Figure 1.1). The two parents P1 (AA BB CC DD) and P2 ($A_bA_b$ $B_aB_a$ $C_dC_d$ $D_cD_c$) differ in terms of two reciprocal translocations on four chromosomes, which means that on four chromosomes of P2 segments of non-homologous chromosomes can be found, i. e. chromosome A has parts of chromosome B incorporated and vice versa. These reciprocal translocations lower the fertility of an individual plant if they are not homozygous due to pairing difficulties at meiosis. The F1 hybrid ($AA_b$ $BB_a$ $CC_d$ $DD_c$) resulting from a hybridization event between P1 and P2 is heterozygous for two reciprocal translocations and therefore partially sterile. The expected production rate of fertile and balanced gametes (gametic fertility) is 25% whereas 75% of the F1 gametes will be infertile and unbalanced (Figure 1.1). The balanced and fertile gametes include the two parental types which can only recreate the parental type in the F2 and hence are not capable of generating a novel hybrid lineage. The other two fertile gametes, however, can give rise to new homozygous types in the F2 (AA BB $C_dC_d$ $D_cD_c$ and $A_bA_b$ $B_aB_a$ CC DD) which are fully fertile but are partially sterile with one another and the parents.

It is noteworthy that the parents have to differ by two or more translocations as a difference by only one translocation can not produce a novel and fertile F2 generation. The fertile F1 gametes will only be able to reconstitute the parental genotype. Apart from reciprocal translocations there is a number of chromosomal rearrangements which can act as sterility barriers like fusion or fission of chromosomes, the duplication or deletion of a chromosomal segment and the re-inversion of a segment where the original gene order on a chromosome is changed, for example from ABCDE to AEBCD (Rieseberg, 2001). The novel hybrid derivative is fertile, homozygous for the chromosomal rearrangement and on the same ploidy level as the parents but isolated from them by a more or less strong chromosomal sterility barrier.

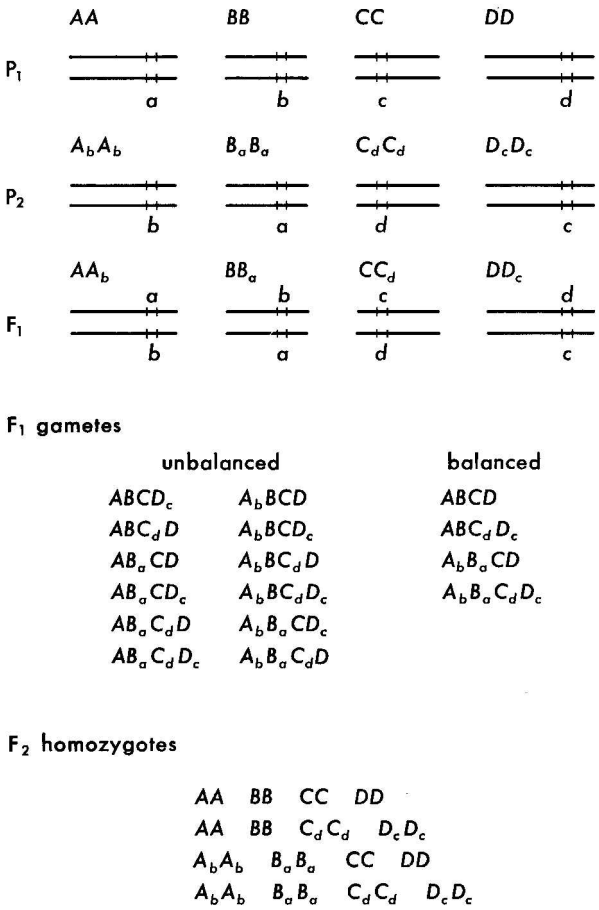



**Figure 1.1.** Model of homoploid speciation for two independent translocations. The first translocation involves segments a and b in chromosomes A and B, and the second translocation involves c and d in chromosomes C and D. Parent 1 (P1) has the standard arrangement in all four chromosomes, while parent 2 (P2) carries both translocations. The F1 hybrid produces 16 classes of gametophytes as shown. In 12 cases there are deficiencies and duplications for the a-b segments, the c-d segments, or both. Four other cases of functioning gametes produce various types of translocation heterozygotes in F2 (not shown) along with the four types of structural homozygotes (shown). Further explanation see text. From Grant (1971).

In a study undertaken by Rieseberg (1997) more than 50 putative examples of homoploid speciation were identified in the literature. However, in the opinion of the author only eight studies have convincingly documented homoploid hybrid speciation (Rieseberg, 1997). In a more recent study by Gross and Rieseberg (2005) this number was doubled to about 20 good examples.

### 1.2.4 Stalemate

Hybridization may be just an ephemeral event without any further evolutionary consequences. This outcome can be explained by the Tension Zone Model (Barton & Hewitt, 1985). In this model hybrid zones are viewed as clines which are maintained by a balance between dispersal and selection against hybrids. Clines are regions in which a variety of characters change gradually resulting in a cluster of parallel zones with different gene frequencies, grading from a maximum to a minimum between these zones. The Tension Zone model assumes that hybrids are less fit than their parental genotypes independent of the habitat they grow in because of genomic incompatibilities (Barton & Hewitt, 1985). This leads to a 'tension zone' between the conflicting effects of dispersal of parental forms on the one hand and the creation and selection against less fit hybrids on the other, resulting in an equilibrium between selection against hybrids and their constant production from inter-parental crosses. The outcome could be compared to a 'stalemate' situation with little evolutionary consequences as there is a very restricted potential for introgression of neutral genes between the parental species and because no viable new hybrid lineages are produced.

## 1.3 Hybridization between two different mating systems

It has recently been suggested that hybridization between selfing and outcrossing taxa may have been a keyfactor in the production of biodiversity and taxonomic complexity in the widely hybridizing genera of *Euphrasia* (French *et al.*, 2008) and *Epipactis* (Hollingsworth *et al.*, 2007). The chances of survival and perseverance of novel hybrid lineages have often been connected to the availability of suitable ecological habitats (Anderson, 1948; Rieseberg, 1991; Buerkle *et al.*, 2000) and especially in homoploid speciation, where hybrids are at the same ploidy level as their parents, to a selfing mating system (Grant, 1958; McCarthy *et al.*, 1995; Buerkle *et al.*, 2000). The rationale behind these assumptions is that the likelihood for the persistence of a new hybrid lineage will depend on how well it is insulated from homogenizing gene flow. A different habitat and a selfing mating system reduce gene flow with the parental population ensuring that the genetic and phenotypic integrity of the new lineage is maintained. Selfing hybrid lineages are most likely to occur where hybridization takes place between closely related inbreeding and outcrossing plant taxa. The outcrossing parent maximizes the chances of initial hybrid formation while genes for selfing contributed by the inbreeding parent ensure that a proportion of later generation hybrids will be selfing. This hypothesis is particularly plausible if seen in the light of transgressive segregation.

## 1.3.1 Transgressive segregation

Transgressive segregation is a phenomenon where certain hybrids display extreme phenotypes which fall outside the natural range of variation observed in the parents. These individuals exceed the parental phenotypic range in a positive or negative direction, i. e. some characters in individual hybrids are much larger or smaller than in the parents. There are several different genetic explanations like elevated mutation rates in hybrids, overdominance, reduced developmental stability in hybrids, epistasis, the unmasking of rare recessive alleles which are usually heterozygous in the parents and chromosome number variations. However, all these theories can only explain a small part of the observed cases or are mechanistically unrealistic (Rieseberg *et al.*, 1999). Marker based QTL studies show that the most likely mechanism seems to be additive complementary action of alleles which are dispersed in parental lines (Rieseberg *et al.*, 1999). Different species or parental lines have loci which are often fixed for sets of alleles with opposing effects on the phenotype, an important prerequisite for complementary action (Rieseberg *et al.*, 1999). These alleles have either a positive (increasing) or negative (decreasing) effect on the value of a certain phenotypic trait.

Figure 1.2 shows a hypothetical example of how complementation among loci might work. The four unlinked loci A, B, C and D of the parental populations I and II are fixed for different alleles (+1 or -1; + or – denotes an increasing or decreasing effect on the phenotype exerted by that allele, respectively). The overall effect (net score) on a certain phenotypic trait in population I and II is zero as the additive values of loci A to D cancel each other out, creating the 'normal' parental phenotype. The same is true for the F1 hybrid (between population I and II) even though it is heterozygous at all loci as all alleles have opposing effects. Hybrids of the F2 generation, however, can have net scores between +8 and -8 because of different additive combinations facilitated by independent segregation of alleles by a heterozygous F1 hybrid (three different allele combinations of the four loci shown in Figure 1.2 with net scores of +8, 0 and -8, respectively; possible net score values for all combinations below). These hybrids exceed the parental phenotypic range and are the 'hopeful monsters' of evolution.

| Locus | Population I | Population II | F1 Hybrid | F2 Hybrid |
|---|---|---|---|---|
| A | +1, +1 | −1, −1 | +1, −1 | +1, +1; +1, −1; −1, −1 |
| B | +1, +1 | −1, −1 | +1, −1 | +1, +1; +1, −1; −1, −1 |
| C | −1, −1 | +1, +1 | +1, −1 | +1, +1; +1, −1; −1, −1 |
| D | −1, −1 | +1, +1 | +1, −1 | +1, +1; +1, −1; −1, −1 |
| Net score | 0 | 0 | 0 | +8, +6, +4, +2, 0, −2, −4, −6, −8 |

**Figure 1.2.** Hypothetical example of transgressive segregation involving four locus, two allele genotypes in two parental populations and their hybrid descendants. For explanation see text. (from Bell & Travis, 2005).

What is the evidence for transgressive segregation in nature? A recent review examined 171 studies (113 dealing with plants and 58 with animals) which reported phenotypic variation in segregating hybrid populations (Rieseberg *et al.*, 1999). Rieseberg *et al.*'s (1999) review shows that 44% of the 1229 examined characters showed transgressive segregation and that 91% of the 171 studies report at least one transgressive trait. Astonishingly, 97% of the papers studying segregating plant hybrids reported at least one transgressive character. Even though transgressive segregation is much more common in hybrids from inbred, domesticated lines (92%), it still is a common feature in wild populations (38%). More than half (65%) of the transgressive characters were of morphological nature whereas the rest was more or less evenly divided between physiology, biochemistry, life history and ecological factors (Rieseberg *et al.*, 1999). A later study revealed that the genetic architecture necessary for transgressive segregation appears to be a common feature in plants (Rieseberg *et al.*, 2003c).

If transgressive segregation alters characters which are under strong selection, like flower shape or colour, then transgressive hybrids can diverge sufficiently from their progenitors and establish lineages which are able to coexist sympatrically with them.

### 1.3.2 Inbreeding depression

There is one serious difficulty to the theory that selfing hybrids which show novel and transgressive traits could lead to the production of new lineages and increase taxonomic diversity in the long run. This arises because deleterious mutations hidden in heterozygous form will be rendered homozygous through selfing, which results in their expression and consequently will lead to lower fitness of these individuals (inbreeding depression δ). Inbreeding depression can also be caused by heterozygote advantage, i. e. if heterozygous individuals in a population are fitter than either homozygote at a viability locus. Selfing will increase the frequency of homozygotes every generation thus producing offspring with lower fitness compared to heterozygotes which is then observed as inbreeding depression. These two competing models for the genetic basis of inbreeding depression are referred to as partial dominance and overdominance hypothesis, respectively (Charlesworth & Charlesworth, 1987). In both models inbreeding depression will prevent the spread of alleles promoting selfing if $\delta \geq 50\%$. Although the two hypotheses have the same qualitative outcome, i. e. low fitness is observed in the next generation, they differ greatly in the prediction of the evolutionary consequences for populations.

According to the partial dominance hypothesis where inbreeding depression is due to a mutational load of deleterious genes, theory predicts that autogamous (selfing) plants will

exhibit low inbreeding depression. Habitual selfing should purge deleterious recessive genes from the population over time as individuals with lower fitness due to disadvantageous mutations will be selected against (Schemske & Lande, 1985; Charlesworth & Charlesworth, 1987). In contrast, when inbreeding depression is due to a segregational load of less fit homozygotes (overdominance hypothesis), theory predicts that inbreeding depression will increase over time because of the loss of heterozygote advantage at many loci (Charlesworth & Charlesworth, 1990).

A review comparing levels of inbreeding depression in 79 populations of 54 plant species found that on average inbreeding depression in mainly selfing species was significantly less than in mainly outcrossing species (Husband & Schemske, 1996), suggesting that purging does occur in selfing taxa. However, there are ambiguous results from other studies which did not detect lower levels of inbreeding depression in selfing populations (Dole & Ritland, 1993; Rathcke & Real, 1993; Johnston & Schoen, 1996; Michalski & Durka, 2007; Ishida, 2008) and there is also evidence of inbreeding depression in highly selfing plant taxa such as *Leavenworthia uniflora* (Charlesworth *et al.*, 1994), *Mimulus nasutus* (Latta & Ritland, 1994) and *Begonia semiovata* and *B. hirsuta* (Ågren & Schemske, 1993). A review by Byers & Waller (1999) on inbreeding depression and purging comparing 52 studies found that there was no conclusive evidence for purging. The results of their study showed that purging did not seem to act as a consistent and major force in natural populations, thus questioning the predicted powerful force of purging by theoretical models (Byers & Waller, 1999). The authors come to the conclusion that under some circumstances inbreeding will tend to purge populations of enough deleterious recessive mutations to reduce inbreeding depression but that it is impossible to make universal predictions (Byers & Waller, 1999).

There is still no conclusive evidence as yet to which hypothesis of the genetic basis of inbreeding depression might be the correct one. However, there seems to be growing consensus in favour for the dominance model in the most recent reviews on this topic (Charlesworth & Charlesworth, 1999; Carr & Dudash, 2003)

### 1.3.3 Previous studies

Despite the potentially important evolutionary consequences of hybridization between selfing and outcrossing plant taxa, we know very little about it. In a study carried out in naturally occurring sympatric populations of outcrossing *Mimulus guttatus* and selfing *M. nasutus*, Martin & Willis (2007) found that prezygotic barriers acted asymmetrically to completely prevent the formation of F1 hybrids among seeds produced by outcrossing *M. guttatus*. F1 hybrid production among seeds produced by selfing *M. nasutus* was reduced to about 1%.

Postzygotic isolation contributed very little to the total isolation between the two species. Despite the strong total reproductive isolation (0.99 to 1.0) between *M. guttatus* and *M. nasutus*, substantial opportunity for interspecific gene flow still exists. More than 1 in 1000 seeds are highly fit *M. nasutus* F1 individuals which are likely to backcross to *M. guttatus*. Martin & Willis (2007) think it plausible that the introgression rate from *M. nasutus* to *M. guttatus* is on the order of 0.001 for neutral regions not tightly linked to selected sites in the genome. That level of introgressive gene flow is high in population genetic terms (Slatkin, 1987) and is expected to cause neutral *M. nasutus* alleles to increase in frequency in *M. guttatus*. This was supported by a recent study which reported introgression at the molecular level in sympatric populations of *M. guttatus* and *M. nasutus* (Sweigart & Willis, 2003). Martin & Willis (2007) therefore conclude that there is greater potential for introgression from selfing *M. nasutus* to outcrossing *M. guttatus* than for the reverse direction.

Studies on hybridization in the genus *Senecio* have elegantly demonstrated that the 'meeting of mating systems' can be a creative force in plant evolution (Abbott *et al.*, 1992; Lowe & Abbott, 1996; Lowe & Abbott, 2000; Abbott & Lowe, 2004; Lowe & Abbott, 2004; Abbott *et al.*, 2005). Hybridization between the introduced diploid *S. squalidus* (outcrossing) and the native tetraploid *S. vulgaris* (inbreeding) has resulted in the generation of two new inbreeding species in the British Isles, namely tetraploid *S. eboracensis* and allohexaploid *S. cambrensis*. *S. eboracensis* was first recorded in the British Isles in 1979, formally described in 2003 and possesses an intermediate morphological phenotype for capitulum and leaf characters (Lowe & Abbott, 2003). In a later study Lowe & Abbott (2004) concluded that strong breeding barriers exist between the new hybrid species and its two parents. Prezygotic isolation between *S. eboracensis* and its parents is likely to be largely due to the predominant selfing mating system but differences in germination rates, seedling survival, time of flowering and characters associated with pollinator attraction are also contributing to reproductive isolation and ecological differentiation (Lowe & Abbott, 2004). It is noteworthy that *S. eboracensis* has not been found at its presumed place of origin in York or anywhere else in Britain since 2000 and is believed to be extinct in the wild (Abbott *et al.*, 2005).

The other hybrid species which originated in the British Isles, *S. cambrensis*, was described as a new species in 1955 after it had been recorded in Wales in 1948 (Abbott *et al.*, 2005). *S. cambrensis* exhibits intermediate morphology and also an additive isozyme and RAPD profile of its parents (Lowe & Abbott, 1996). Like *S. eboracensis*, *S. cambrensis* is mainly self-fertilising and occurs in populations which frequently become extinct but also are established by colonisation of new sites (Abbott *et al.*, 2005). There is also evidence that hybridization between non-radiate *Senecio vulgaris* and radiate *S. squalidus* has led to the introgression of the radiate trait into some populations of *S. vulgaris* which promotes outcrossing. This

suggests that unlike in *Mimulus*, gene flow occurs from the outcrossing into the selfing species (Kim *et al.*, 2008).

## 1.4 Motivation for this research

Novel phenotypes can arise through hybridization via the combination of two different genomes and/or transgressive segregation. The likelihood that a new phenotype will persist and ultimately develop into a new species partly depends on how well it is insulated from homogenizing gene flow. Various authors have predicted that the breeding system of hybridizing plants is an important factor in hybrid speciation (Grant, 1958; 1971: 203; Templeton, 1981) as selfing reduces gene flow considerably and can therefore maintain the genetic and phenotypic novelty of the hybrid. As mentioned previously, hybrids with high selfing rates are most likely to occur where hybridization happens between closely related inbreeding and outbreeding species. The outcrossing species maximises the chances of the initial hybrid formation, while genes for selfing donated by the selfing parent ensure that a proportion of later generation hybrids will be inbreeding. Hybridization between selfing and outcrossing plant taxa is therefore a potentially important evolutionary force.

This could come at a considerable fitness cost, however, as selfing will lead to the expression of deleterious recessive mutations resulting in inbreeding depression if either of the parental species carries a high genetic load. Thus, on the one hand hybridization between selfing and outcrossing plant taxa might be of importance in the generation of taxonomic novelty. On the other hand the genetic load carried by the parental species might be a major barrier for the establishment of new hybrid lineages. In order to resolve this dilemma, empirical data are needed. However, the vast majority of literature on hybridization is on outcrossing plant species. To better understand the processes which take place when inbreeding meets outcrossing, investigations using an appropriate study system are clearly needed. The objective of this thesis is to use the putatively selfing *G. urbanum* and putatively outcrossing *G. rivale*, two freely hybridizing species which are native to Great Britain, and study the processes that take place when the two different mating systems meet.

# 1.5 Study species

## 1.5.1 Geum urbanum

**Figure 1.3**. *Geum urbanum*. **A:** Habit **B**: Flower detail

*Geum urbanum* (Rosaceae) or Wood Avens (Figure 1.3 and Figure 1.6) is a common perennial herb and occurs throughout Europe and parts of Asia (Figure 1.9). In Great Britain it is a very abundant species and is only absent from northern and central Scotland (Figure 1.10). *G. urbanum* is a self-compatible species with a 'weedy' tendency and grows naturally in shaded places like deciduous woodlands, scrub and hedgerows but is also found in open, disturbed habitats (Taylor, 1997b). The Wood Avens prefers moist soil and is absent from places that dry out or are continuously wet. The rhizome system is generally poorly developed and usually produces only one single plant above ground (Figure 1.4). The small yellow, open and erect flowers do not have any pollinator preferences, are visited by various insects, mostly Muscid and Syrphid flies and appear to be mostly self-pollinated (Briggs & Walters, 1997; Taylor, 1997b). They are hermaphrodite and weakly protogynous (stigmas are receptive before pollen is shed) but male-only flowers (andromonoecious) and rarely male-only plants (androdioecious) have been recorded (Hegi, 1975). Nectar is secreted by a green fleshy disc situated internal to the filaments (Taylor, 1997b).

**Figure 1.4.** Rhizome system of *G. urbanum* and *G. rivale*.

*G. urbanum* flowers from end of May to October and sets fruit from July to September (October). Seeds are either animal dispersed or just fall to the ground. In the fruiting stage the stigma breaks off at a predetermined joint along the style modifying it into a hooked awn which then readily adheres to any rough surface like animal fur or clothing (Taylor, 1997b). *G. urbanum* is hexaploid with 2n = 42 (Gajewski, 1957).

## 1.5.2 Geum rivale

**Figure 1.5.** *Geum rivale*. **A:** Habit **B:** Flower detail

*G. rivale* or Water Avens (Figure 1.5 and Figure 1.7) is a perennial herb and occurs naturally in Europe, parts of Asia and North America (Figure 1.9). In Great Britain it is an abundant species but is absent from some areas in central and southern England (Figure 1.10). *G. rivale* is a self-compatible species and grows on slow-draining and wet or sometimes water-logged soils, usually along stream sides, wet hay meadows and wet flushes in oak- or alderwoods (Taylor, 1997a). It has a well developed branching rhizome system (Figure 1.4) which provides a means of vegetative spread and can produce numerous flowering stems (Taylor, 1997a).

**Figure 1.6.** *G. urbanum*. A, plant; B leaf from rosette; C, upper part of flowering stem; D, longitudinal section of flower, petals removed; E, carpel; F, upper part of fruiting stem; G, achene; H, seed. From Ross-Craig (1955).

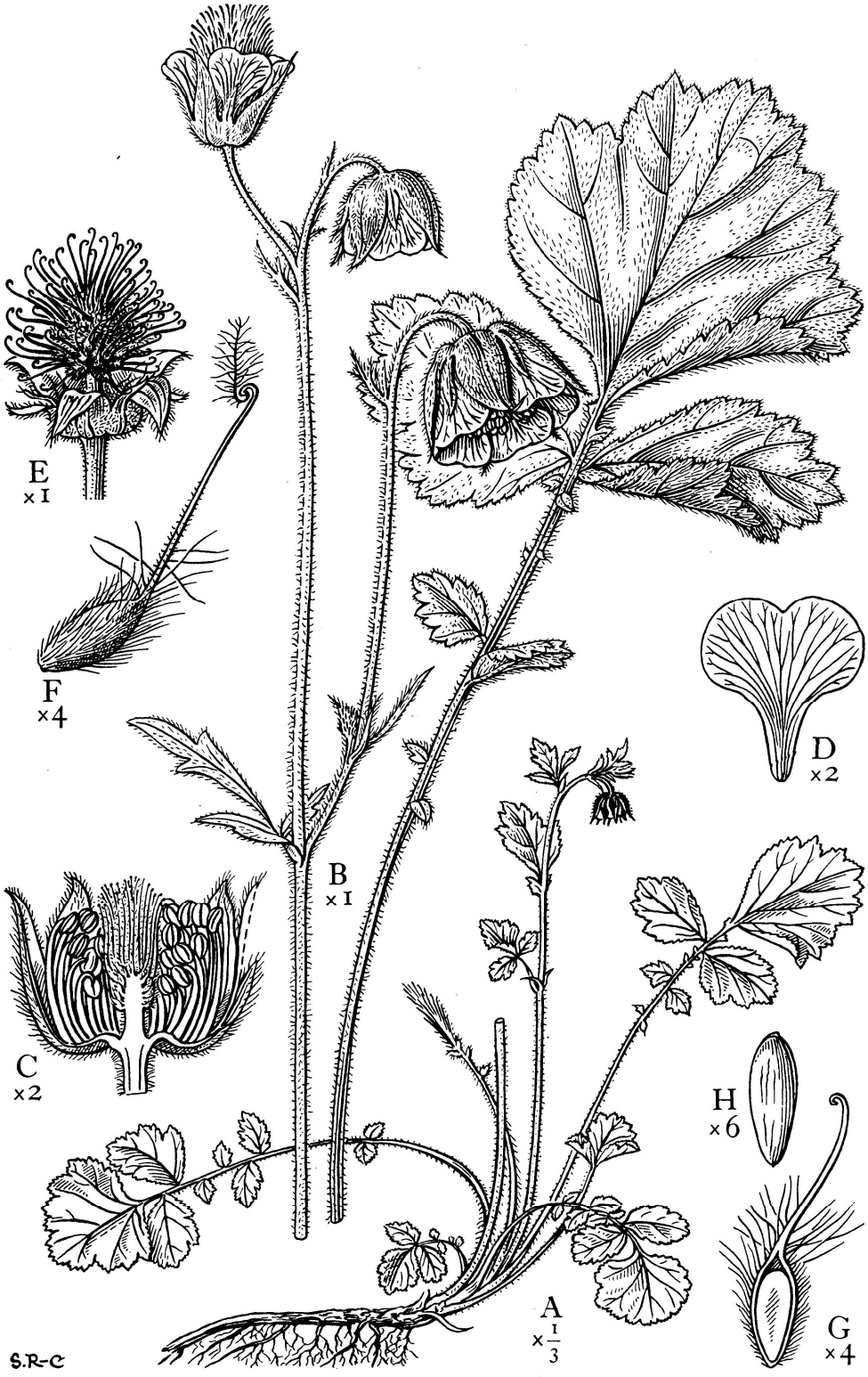


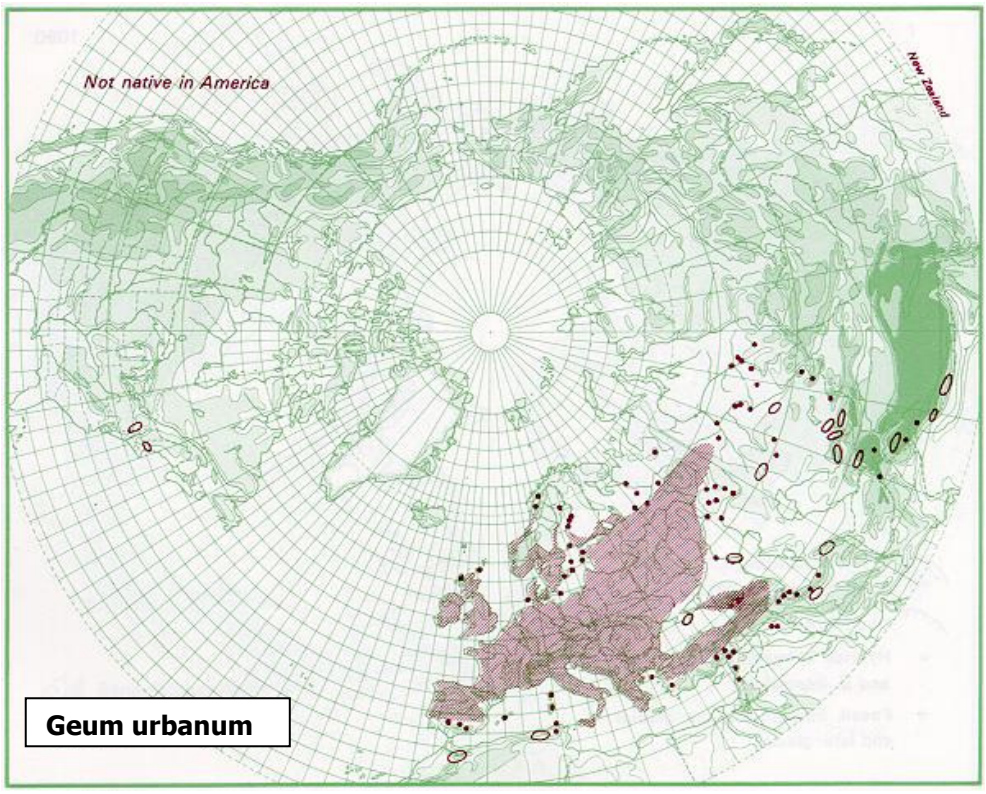

**Figure 1.7.** *G. rivale*. A, plant; B leaf from rosette and flowering stem; C, longitudinal section of flower, petals removed; D, petal; E, fruit; F, young achene; G, longitudinal section of achene after the upper style has fallen; H, seed. From Ross-Craig (1955).

The purplish, campanulate flowers are mostly protogynous and produce nectar in numerous droplets on the receptacle (Briggs & Walters, 1997). Langlotz (2007: 228) reported that *G. rivale* flowers stay fertile for about 4 days with the female phase lasting 2.1 days, the male phase 2.6 days and the two phases overlapping by 1.1 days. Species of *Bombus* are recorded as the most common visitors but flowers are also visited by other insects such as Syrphid flies and beetles (Taylor, 1997a). In the later stages of floral development self-pollination can be readily effected but cross pollination seems to be the major mode of fertilisation (Taylor, 1997a). As in *G. urbanum* seeds are either animal dispersed (hooked awn) or just fall to the ground. In a study conducted on the dispersal distances of adhesive seeds, *G. rivale* seeds were dispersed on average 27 ± 7 m by fallow deer (Kiviniemi, 1996). Based on the maximum time seeds were attached to the animals (c. 3 hours) and the observed average speed of deer (4 m/min) a maximum dispersal distance of 700 ± 165 m was reported by Kiviniemi (1996). Flowering starts about 3-4 weeks before *G. urbanum* at the beginning of May and lasts till July and fruits ripen from July to September. *G. rivale* is hexaploid with 2n = 42 and differs from its sister species *G. urbanum* (Smedmark *et al.*, 2003) by two to three chromosomal rearrangements (Gajewski, 1957).

### 1.5.3 Geum urbanum x G. rivale

**Figure 1.8.** Hybrid phenotypes (F1) of *Geum urbanum* x *G. rivale.*

*G. urbanum* and *G. rivale* occupy the same broad geographic area throughout most of Europe and Asia (Figure 1.9). Even though they are separated by their habitat requirements and partially by pollinator preferences and flowering time, hybrids are often formed wherever the two species meet. The hybrids are reported to be highly fertile (Stace, 1975), often form hybrid swarms (Waldren et al., 1989), show a wide range of variation between the parental phenotypes (Marsden-Jones, 1930; Prywer, 1932; Gajewski, 1957) and occur widely throughout Great Britain (Figure 1.10). They are on the same ploidy level as their parents (hexaploid, 2n = 42).



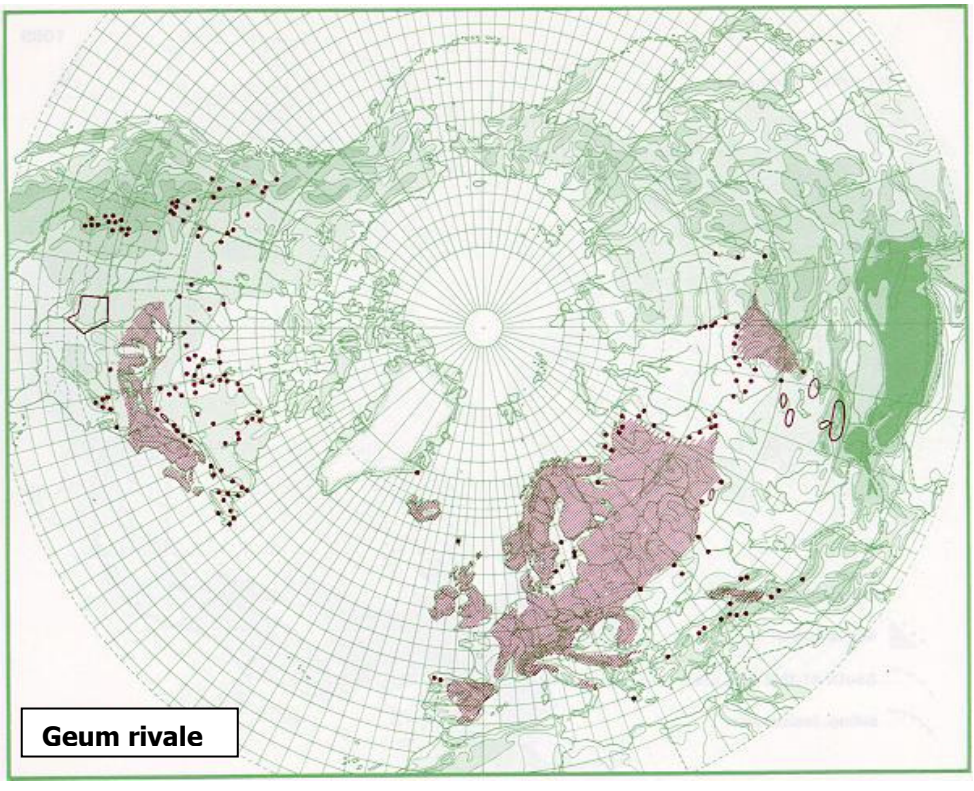




**Figure 1.9.** Natural distribution of *Geum urbanum* and *Geum rivale.* From Hulten & Fries (1986).

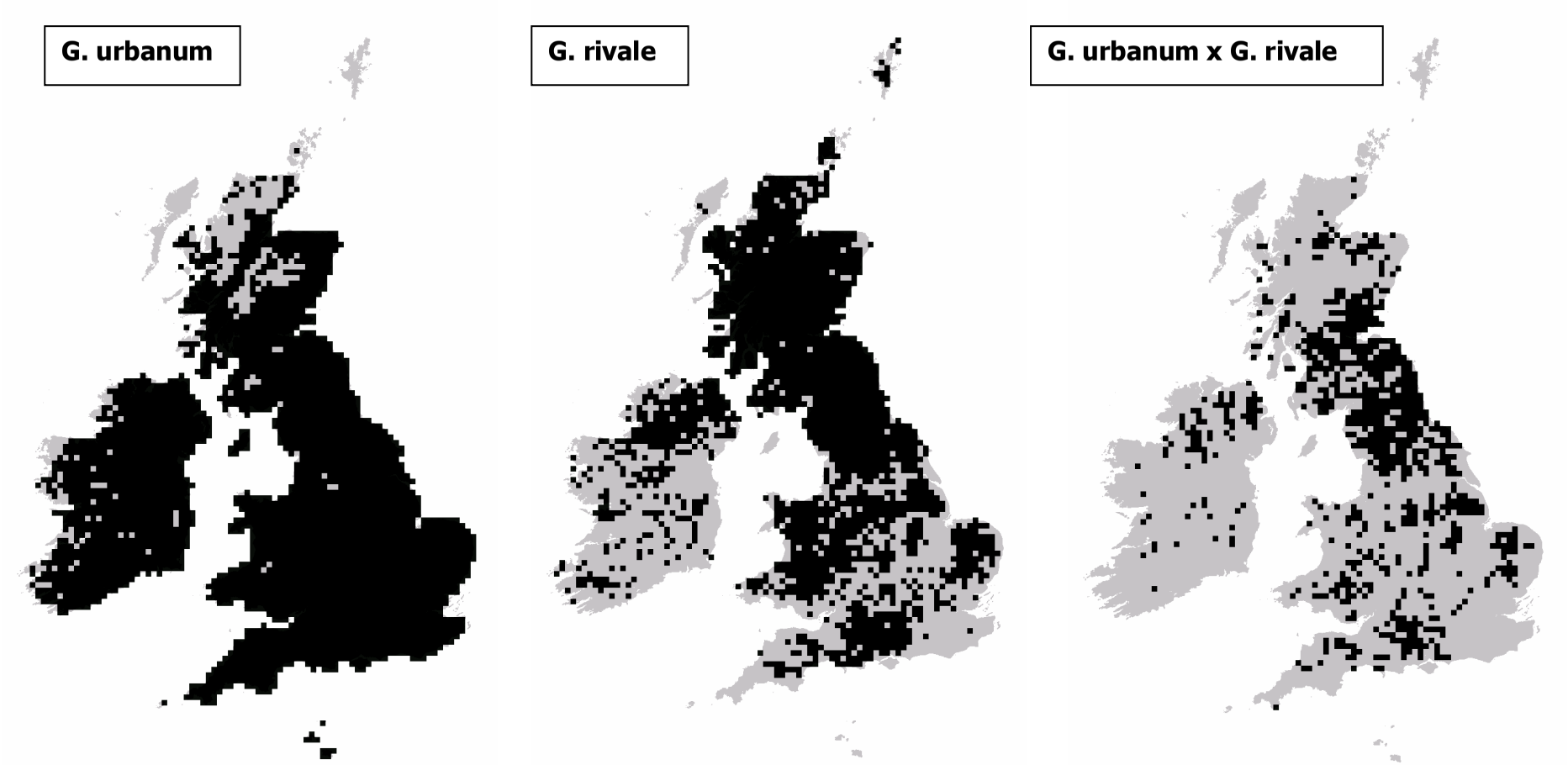



**Figure 1.10.** Distribution of *G. urbanum*, *G. rivale* and their hybrid *G. urbanum* x *G. rivale* in Great Britain. All native and alien records (from pre-1970 to 1999) are included. From Preston *et al.* (2002).

## 1.6 Aims and objectives

Hybridization between taxonomically distinct plant taxa is an important force in plant evolution and it has recently been suggested that hybridization between inbreeding and outcrossing plant taxa may have been an important keyfactor in generating taxonomic complexity in genera such as *Epipactis* and *Euphrasia*. However, the vast majority of studies which have investigated hybridization concerned hybridization between outbreeding plant taxa. The overall aim of this thesis is to study the processes which take place when species with two different mating systems (selfing and outbreeding) hybridize.

The six specific questions of this research are outlined below and represent the following chapters of the thesis. In the final chapter, the results from this study will be synthesized and discussed in the light of possible evolutionary consequences.

1. **Are *G. urbanum* and *G. rivale* genetically and morphologically distinct species despite ongoing hybridization between the two species?**

To address this question, samples from reference ('pure') populations of both species were collected from various locations throughout Great Britain, and AFLP markers were used to identify species-specific markers, assessing the genetic integrity of the two species. Nine morphological measurements were taken on Scottish reference populations grown in a common environment for a year to establish species delimiting phenotypic characters.

2. **Is the mating system of *G. urbanum* mainly selfing and of *G. rivale* mainly outcrossing, i.e. is it justified to say that two different mating systems meet when the two species hybridize?**

To address this question, direct breeding system estimates using microsatellites were carried out in reference populations of *G. urbanum* and *G. rivale*. Indirect estimates of the mating system were obtained by calculating $F_{is}$ in the adult generation. Additionally, controlled crosses were performed in *G. urbanum* and *G. rivale* to establish whether there is evidence for a self-incompatibility system in either species.

3. **What levels of inbreeding depression are present in pure *G. urbanum* and *G. rivale* populations?**

To address this question, reference populations of either species were used to perform controlled crosses within each population to establish whether fitness differences between outcrossed and selfed progeny can be detected.

4. **What are the genetic and morphological characteristics of F1 hybrids compared to pure *G. urbanum* and *G. rivale*?**

To address this question, reciprocal controlled crosses were performed between pure *G. urbanum* and *G. rivale* individuals. F1 hybrids were assayed with genetic and morphological tools to establish their genetic and phenotypic properties which were compared to pure parental populations.

5. **What is the genetic and morphological structure of a hybrid swarm between *G. urbanum* x *G. rivale*?**

To address this question, one hybrid population was examined using AFLP markers to establish which genetic groups are present in the hybrid zone. Morphological measurements of pure and hybrid individuals taken in a common environment were used to test for transgressive segregation in the hybrid zone. Phenological observations were carried out in the hybrid swarm to estimate the ease of hybrid production by the parental taxa.

6. **What is the dynamic of a hybrid swarm? Do we find the same genotypic classes produced in the offspring generation also in the adult generation?**

To study the dynamics of a hybrid zone between *G. urbanum* and *G. rivale*, open-pollinated seeds were collected from adults covering the whole morphological spectrum in the hybrid swarm. A random choice of seedlings per maternal parent was planted into a common garden environment and their genotype, phenotype and fitness was established. This allowed me to compare the phenotypic and genetic composition of adult and offspring generations and ascertain whether the diversity present in the progeny is different to the one in the adult generation.

Collectively I intend to use the answers to these questions to predict the possible evolutionary long term consequences of hybridization between *G. urbanum* and *G. rivale*.

# 2. ANALYSIS OF PURE *GEUM URBANUM* AND *G. RIVALE* POPULATIONS

## Abstract

Theoretical considerations predict that hybridization between plants with different breeding systems could potentially be a powerful evolutionary force but very few studies have been carried out to investigate the processes involved in natural hybridizing populations. *Geum urbanum* (mainly selfing) and *G. rivale* (mostly outcrossing) form an example of hybridization between two species with different breeding systems. The main aim of this chapter is to establish the extent of morphological and genetic marker differences between pure populations of these two species to provide a baseline for studying hybridization between them. 39 *G. urbanum* and 32 *G. rivale* individuals from three Scottish populations were measured for nine morphological characters taken in a common environment. 30 individuals from ten British reference populations per species were screened for 616 AFLP markers and 154 samples for two chloroplast regions using RFLPs (trnL-F/*Eco*R1, trnC-D/*Mse*1). The morphological analysis showed that the two species differed significantly in all nine traits measured. The AFLP analysis revealed that 8% of the markers were species-specific and that *G. urbanum* and *G. rivale* were clearly and unambiguously separated in a PCO plot, indicating their genetic distinctiveness. One chloroplast haplotype (A=100%) was detected in *G. urbanum* whereas four haplotypes (A=25.3%, B=1.3%, C=70.9%, and D=2.5%) were observed in *G. rivale*. No species-specific chloroplast markers were detected. These results suggest that despite widespread hybridization, the morphological and genetic boundaries which clearly separate *G. urbanum* and *G. rivale* have not been eroded and that the two species remain distinct biological entities.

## 2.1 Introduction

Hybridization between plants has attracted considerable interest among biologists as it can result in various evolutionary outcomes ranging from extinction to speciation (see Chapter 1). However, the vast majority of studies have focused on hybridization between outcrossing plants despite the fact that hybridization between plants with different mating systems (i. e. selfing and outcrossing) is not rare and has been reported to occur in various genera such as *Geum* (Marsden-Jones, 1930; Prywer, 1932), *Mimulus* (Vickery, 1956), *Phlox* (Levin, 1967), *Carex* (Waterway, 1990) and *Lolium* (Charmet et al., 1996). Theoretical considerations predict that hybridization between plants with different breeding systems could potentially be a powerful evolutionary force but very few studies have been carried out to investigate the processes involved in natural hybridizing populations.

It has long been recognized that the transition from outcrossing to selfing is a frequent evolutionary event in flowering plants (Stebbins, 1974: 51) and can lead to rapid reproductive isolation of a selfing neospecies from its progenitor. If reproductive isolation is coupled with morphological variation due to the expression of recessive alleles and the accumulation of mutations resulting in further phenotypic variation then this can lead to an increase in taxonomic diversity over time. Every generation of complete selfing will decrease heterozygosity by 50% and recessive alleles previously hidden in heterozygous form will be expressed. Theoretical models and empirical studies have shown that even moderate levels of selfing increase the chance of fixation of a favourable recessive allele in a large population (Charlesworth, 1992).

Hybrid lineages with novel character combinations embedded in a selfing background are at least partially reproductively isolated from gene flow with congeners and therefore represent the 'hopeful monsters' of evolution with the potential to develop into something new. Selfing hybrids are most likely to occur where hybridisation takes place between closely related outcrossing and selfing species. The outcrossing parent maximises the chance of initial hybrid formation while genes for selfing donated by the inbreeding parent ensure that a proportion of the offspring will be selfing. Recent studies in the taxonomically complex and widely hybridizing genera of *Euphrasia* and *Epipactis* have led to suggestions that breeding system transitions from outcrossing to selfing have contributed to the recurrent evolution of autogamous (selfing) lines and that subsequent hybridization between outcrossing and inbreeding taxa could have been the cause for the high intrageneric diversity in these genera (Squirrell *et al.*, 2002; Hollingsworth *et al.*, 2007; French *et al.*, 2008).

The main aim of this thesis is to gain a better understanding of the hybridization process between inbreeding and outcrossing plant species using the naturally hybridizing *Geum urbanum* (mainly selfing) and *G. rivale* (mostly outcrossing) as study species. To be able to investigate the patterns and the extent of genetic exchange, levels of phenotypic and genetic differences between *G. urbanum* and *G. rivale* need to be established first. This can be done by looking at (1) the genetically determined morphological differences separating the two species using a quantitative approach and by (2) genetic marker differences using molecular fingerprinting techniques. Some key phenotypic traits are believed to be controlled by a small number of important genes of large effect (Tanksley, 1993; Bradshaw *et al.*, 1998) which are dominant and under strong selection. In contrast, markers detected by molecular fingerprinting techniques may not be under selection and are usually believed to be neutral. This approach to studying hybridization is only likely to be informative if the two hybridizing species differ significantly for both genetically determined morphological variation and genetic markers. The aim of this chapter is to establish those differences by answering the following questions:

1. How different are the two parental species in terms of their genetically determined morphological variation? Can the two species be separated morphologically using leaf, floral and seed characters or do they overlap in some of these characters?

2. How different are the two parental species in terms of their overall genetic makeup?

To address whether *G. urbanum* and *G. rivale* are distinct in their genetically determined morphology nine characters were chosen. These characters were selected because they capture the maximum quantifiable morphological differentiation between the two species, have relatively low phenotypic plasticity and have also been used in previous studies (Prywer, 1932; Waldren *et al.*, 1989). To exclude morphological variation due to different environmental conditions samples included in this analysis were dug up and grown in a common environment for at least one season before measurements were taken. To my knowledge no statistical analysis of genetically determined morphological variation between the two species has been carried out in any previous study.

To address the second question of genetic distinctiveness, Amplified Fragment Length Polymorphism (AFLP) and chloroplast DNA Restriction Fragment Length Polymorphism (cpDNA PCR-RFLP) have been used. AFLPs are genetic markers based on the Polymerase Chain Reaction (PCR) generated by digestion of DNA using frequently and rarely cutting restriction enzymes (Vos *et al.*, 1995). The main advantage of AFLPs is that they allow for rapid screening of genetic diversity without prior sequence knowledge by generating

hundreds of replicable genetic markers with low error levels. AFLPs have a high multiplex ratio which is the number of different genetic loci that can be simultaneously analysed per experiment (Rafalski *et al.*, 1996). In the case of AFLPs bands are assumed to stem from different areas of the genome (Robinson & Harris, 1999). Unfortunately there are a number of problems associated with AFLPs. The main disadvantage is that they are dominant markers (polymorphisms are scored as presence or absence of bands) and can not reveal whether the screened loci are heterozygous or homozygous but this is compensated by the large numbers of independent markers from throughout the genome.

AFLPs are particularly useful in analyses of genetic variation between species (Carolan *et al.*, 2002; Scotti-Saintagne *et al.*, 2004) and below the species level for assessing population differentiation and variation within populations (Tribsch *et al.*, 2002; Schönswetter *et al.*, 2003) but have also been successfully employed to detect hybridization (Gobert *et al.*, 2002; Salmon *et al.*, 2005; Guo *et al.*, 2006).

The second kind of marker chosen for this study is chloroplast DNA PCR-RFLPs. These markers provide information about the genetic differentiation of organelle DNA and are used to detect inter- and intraspecific DNA variation employing restriction enzymes when prior sequence knowledge is available. They have the advantage that populations can be quickly and cheaply screened for polymorphisms once suitable variable sites have been identified during sequence analysis (Rafalski *et al.*, 1996). Chloroplast DNA is generally maternally inherited in angiosperms (Reboud & Zeyl, 1994), and can therefore also be used to determine the direction of hybridization. However, this is only possible if species-specific cpDNA markers are identified in the parental species. Species-specific markers are markers which occur in all individuals of one species and in none of the other.

To answer the questions above pure populations need to be located to establish reference points for each species. Both parent species have large geographic ranges (Chapter 1.5) so many different populations could be chosen. However, the chosen populations should be as close as possible genetically to the hybrid population under investigation for this study in order to be representative in terms of genetic markers. Because the location of the study hybrid zone used for this thesis is located just outside the city of Edinburgh a number of pure populations where chosen in the surroundings of Edinburgh.

This was fairly easy for pure *G. urbanum* populations but proved to be more difficult for *G. rivale* populations as in most cases *G. urbanum* individuals could be found in the vicinity of *G. rivale* populations. It was therefore decided to collect samples from *G. rivale* populations further afield, i. e. Stirling (Ben Lawers, 115 km) and the Scottish Borders (Moorfoot Hills, 30

km). In addition to the higher density sampling (about 10 individuals or more) from these populations low density sampling (one individual per population) was also carried out throughout Great Britain (Table 2.1) in order to ascertain that species-specific markers are indeed species-specific and not only population specific.

A byproduct of this approach using duplicate pure populations with moderate sample sizes is that some information on phenotypic and genetic marker differentiation between populations can be gleaned from the data set.

Summarising, the aim of this chapter is to establish to which extent *G. urbanum* and *G. rivale* differ in their genetically determined morphology and their genetic markers. Individuals from pure populations of both species close to a hybrid zone in Edinburgh and further afield were grown under common conditions and scored for nine morphological characters. Some of these samples and individuals from various other pure populations in Great Britain were scored for nuclear AFLP and maternally inherited cpDNA markers to establish the extent of genetic marker variation between the two species. This information will ultimately be used to interpret patterns of exchange of genes with phenotypic effect and molecular markers by analysing plants from a natural hybrid zone between selfing *G. urbanum* and outcrossing *G. rivale.*

# 2.2 Materials and Methods

## 2.2.1 Sampling material for morphological analysis

Three pure *Geum urbanum* and three pure *G. rivale* populations were identified from different sites around Edinburgh and Stirling, Scotland (Table 2.1). British populations were considered pure if the other parental species was absent within an area of at least 2 km x 2 km surrounding the collection site according to Smith et al. (2002) for populations around Edinburgh and 10 km x 10 km according to Preston et al. (2002) for populations elsewhere. Where these conditions were not satisfied but personal field observations confirmed the absence of the other parental species in the immediate surrounding and only individuals resembling the parental phenotype were present in the sampled population, then these populations were also considered pure. A total of 39 individuals from three pure *G. urbanum* populations (n=11, 18, 10) and 32 individuals from three pure *G. rivale* populations (n=13, 16, 3) were dug up for morphological measurements in spring and summer 2006, respectively (Table 2.1). Plants were transferred to an unheated greenhouse and grown in a common environment in adequately sized pots (diameter c. 15 cm) filled with peat free compost, fertilised once a week during the active growing period and watered when necessary.

## 2.2.2 Sampling material for genetic marker analysis

Ten pure *G. urbanum* and ten pure *G. rivale* populations were identified from different sites throughout Great Britain including the populations U1, U2, U3, R1 and R2 which were used for the morphological analyses, resulting in a sample size of 30 individuals per species (Table 2.1). The same criteria for pureness of populations were applied as detailed above.

## 2.2.3 Quantifying morphological variation

A total of nine floral, leaf and fruit characters were recorded to identify non overlapping species-specific characters and to document intraspecific morphological variation. All measurements were taken after plants had been growing in a randomised array in a common environment (unheated greenhouse) for at least one season to exclude morphological variation due to different growing environments. The characters measured were the length and width of stipule, petal and sepal in centimetres, and the length of the petal claw, gynophore and upper style joint in millimetres (Figure 2.1). Between three and 18 individuals

from each of three pure Scottish *G. urbanum* and *G. rivale* populations were measured, totalling 39 *G. urbanum* and 32 *G. rivale* individuals (Table 2.1).

### 2.2.4 Analysis of morphological data

To identify non overlapping species-specific morphological characters for *G. urbanum* and *G. rivale* a Mann-Whitney test was used. To establish if populations within one species differed significantly from one another a Kruskal-Wallis test was applied. This test is a generalization of the Mann-Whitney test (same assumptions apply such as samples are randomly and independently drawn and that the dependent variable is intrinsically continuous) but is not restricted to only two values of the nominal variable (i. e. more than two populations are possible). Non parametric tests had to be used as the measurements were significantly different from being normally distributed thus violating the assumptions of more powerful parametric tests such as analysis of variance (ANOVA). A principal component analysis (PCA) was performed using all nine characters. PCA is a multivariate method that reduces data dimensionality helping to understand the underlying structure in a data set with many variables.

The program MINITAB 14.1 (Minitab Inc., 2003) was used to conduct all statistical tests and analyses.

| Species | Country | Location | Altitude | Coordinates | No. morph. | No. genetic AFLP/cpDNA |
|---|---|---|---|---|---|---|
| **Geum urbanum** | **Scotland** | Edinburgh (U1) Braid Hills | 170 m | $55^0$ 56′ N $03^0$ 12' W | 11 | 10/31 |
| | | Edinburgh (U2) Colinton Ruderal | 120 m | $55^0$ 55' N $03^0$ 14' W | 18 | 10/24 |
| | | Edinburgh (U3) Craigmillar | 70 m | $55^0$ 56′ N $03^0$ 09′ W | 10 | 3/9 |
| | **England** | Yorkshire Swainby | 200 m | $54^0$ 24 N' $01^0$ 16' W | n/a | 1/1 |
| | | Gloucestershire Cotswolds | 200 m | $52^0$ 03' N $01^0$ 48' W | n/a | 1/1 |
| | | Gloucestershire Cotswolds | 240 m | $51^0$ 46' N $02^0$ 08' W | n/a | 1/1 |
| | | London Kew | 10 m | $51^0$ 28' N $00^0$ 16' W | n/a | 1/1 |
| | | Surrey Westcott | 130 m | $51^0$ 13' N $00^0$ 22' W | n/a | 1/1 |
| | | Devon Dawlish | 20 m | $50^0$ 34′ N $03^0$ 28' W | n/a | 1/1 |
| | | Cornwall Truro | 50 m | $50^0$ 12′ N $05^0$ 03' W | n/a | 1/1 |
| | **Austria** | Burgenland Forchtenstein | 480 m | $47^0$ 42' N $16^0$ 19' E | n/a | 0/1 |
| | | Salzburg Werfen | 560 m | $47^0$ 28' N $13^0$ 11' E | n/a | 0/1 |
| **Geum rivale** | **Scotland** | Stirling (R1) Ben Lawers | 600 m | $56^0$ 32′ N $04^0$ 17' W | 13 | 11/30 |
| | | Edinburgh (R2) Moorfoot Hills | 300 m | $55^0$ 44′ N $03^0$ 01' W | 16 | 11/32 |
| | | Edinburgh Pentland Hills | 300 m | $55^0$ 48′ N $03^0$ 21' W | 3 | 0/3 |
| | | Scottish Borders Moffat Dale | 300 m | $55^0$ 23′ N $03^0$ 19' W | n/a | 2/5 |
| | | Orkney Berridale Hoy | 40 m | $58^0$ 53′ N $03^0$ 23' W | n/a | 1/1 |
| | | Ullapool Inchnadamph | 180 m | $58^0$ 09′ N $04^0$ 56' W | n/a | 1/1 |
| | | Ullapool Knockan Crag | 250 m | $58^0$ 02′ N $05^0$ 03' W | n/a | 1/1 |
| | | Thurso Betty Hill | 20 m | $58^0$ 31′ N $04^0$ 14' W | n/a | 1/1 |
| | **England** | Yorkshire Cross Fell | 800 m | $54^0$ 42′ N $02^0$ 29' W | n/a | 1/1 |
| | | Yorkshire Teesdale | 260 m | $54^0$ 39′ N $02^0$ 09' W | n/a | 1/1 |
| | **Austria** | Lower Austria Altenberg | 800 m | $47^0$ 41' N $15^0$ 38' E | n/a | 0/1 |
| | | Lower Austria Klostertal | 616 m | $47^0$ 49' N $15^0$ 50' E | n/a | 0/1 |

**Table 2.1.** Location of individuals collected from pure *Geum urbanum* and *G. rivale* populations in Great Britain used for the analysis of morphological (in common environment) and genetic differentiation between species. U and R: *G. urbanum* and *G. rivale* populations sampled more intensively, respectively. No. morph. and No. genetic: number of individuals sampled for morphological and genetic marker analysis (AFLP, cpDNA), respectively. Numbers in genetic column refer to AFLP and cpDNA survey, respectively.

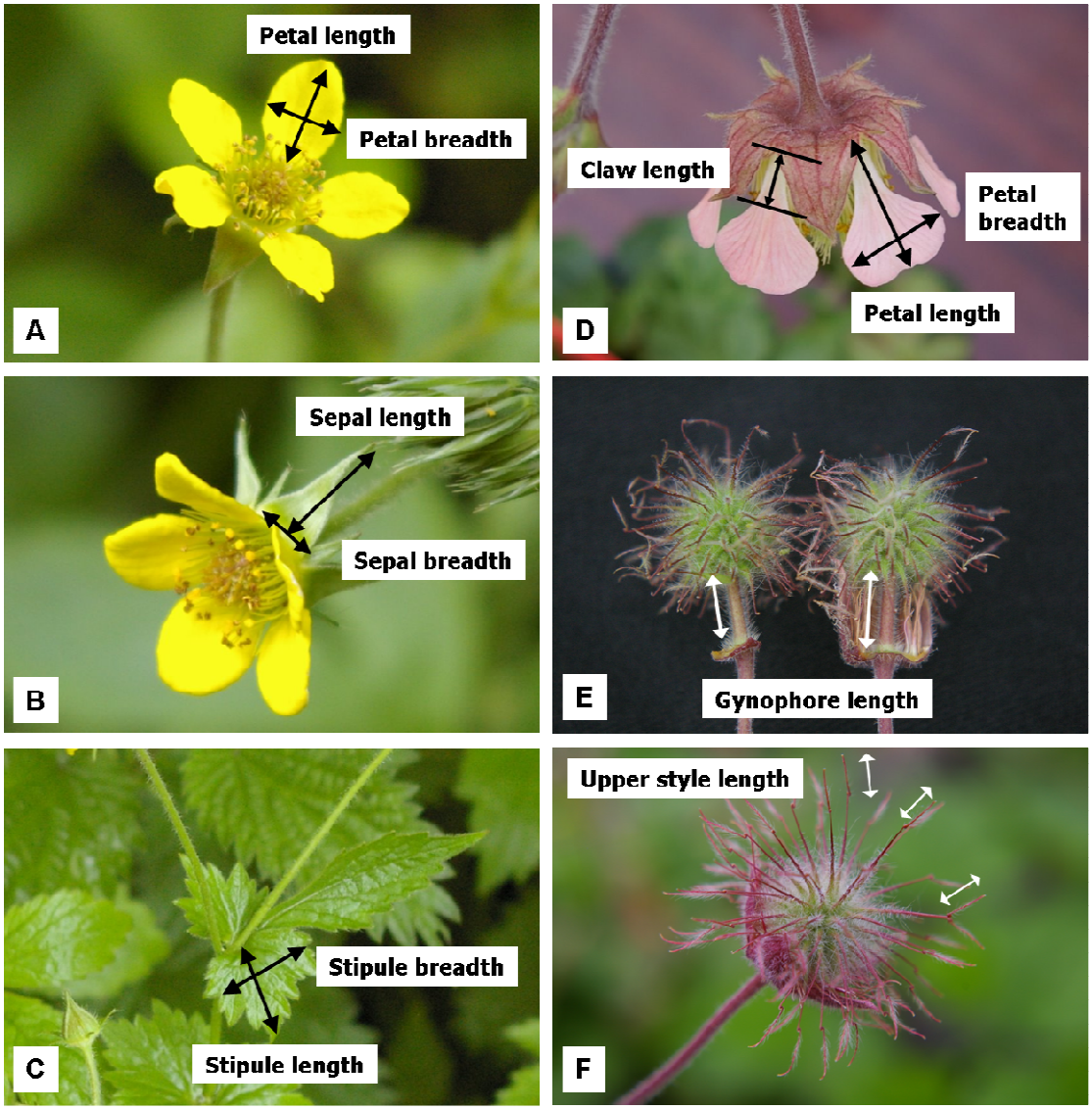



**Figure 2.1.** Morphological measurements taken on *G. urbanum* and *G. rivale* individuals. A to C - *G. urbanum*; D to F - *G. rivale*. All measurements were taken on both species apart from claw length (D) and gynophore length (E) which are *G. rivale* characters only.

## 2.2.5 Genetic marker analysis

*DNA extraction*

DNA was extracted from either silica dried leaf material or fresh leaves (live plants in greenhouse) and was based on the protocol of Doyle and Doyle (1990) using CTAB (cetyltrimethyl ammonium bromide) with the inclusion of 0.1% 2-mercaptoethanol and 0.1% insoluble PVPP to the 2×CTAB buffer. Two chloroform/isoamyl alcohol extraction steps were performed; after precipitation with freezer-cold isopropanol, the resultant DNA pellet was re-suspended in 150 µL TE. DNA quantity was checked by running 3 µl of resuspended DNA and 2 µl loading buffer on a 1% agarose gel.

*Chloroplast DNA (cpDNA)*

To search for species-specific chloroplast markers, 13 cpDNA regions were surveyed using universal chloroplast primers (Table 2.2). Polymerase chain reactions (PCR) were performed in 10µl volumes using the following protocol: 5.2 µl purified water, 1 µl 10x buffer (Bioline), 3 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 µl bovine serum albumin (BSA, 10 mg/ml), 0.25 µM forward primer, 0.25 µM reverse primer, 0.025 U Bio*Taq* (Bioline) and 1.5 µl DNA (c. 2ng/µl). The thermocycle was 4 min at 94°C, 30 cycles of 45 s denaturing at 94°C, 45 s annealing at 50°C to 62°C (according to primer annealing temperature, Table 2.2), and 1 to 4 min extension at 72°C (according to base pair length of region, Table 2.2), ending with 10 min at 72°C to complete extension and subsequent storage at 4°C.

Regions that could be successfully amplified were then sequenced for four *G. urbanum* and four *G. rivale* individuals from populations in Britain and Austria (Table 2.1). To purify the PCR product prior to sequencing 2 µl of 1:10 diluted ExoSAP-IT (USB) was added to 5 µl of PCR product and incubated at 37°C for 2 h and stopped at 80°C for 15 m. The sequencing PCR was performed in 10 µl volumes for forward and reverse primer separately using the following protocol: 1.75 µl purified water, 1 µl BigDye Terminator v1.1 mix (Applied Biosystems), 2 µl BigDye buffer (Applied Biosystems) and 0.05 µM of primer. Sequencing of PCR products was carried out on an ABI 3730 automated sequencer following the manufacturer's instructions. Data were then analysed using the program Sequencher 4.7 (Gene Codes Corporation) to identify sequence variation which could potentially be exploited for species-specific markers. Four regions which showed the desired variation (Table 2.2) were detected yielding four possible species-specific markers. PCR restriction fragment length polymorphisms (PCR-RFLP) were designed for these regions using the following enzymes: *Eco*R1 (trnL-F), *Mse*1 (trnC-D), *Spe*1 (atpH-I) and *Hinf*1 (trnS-Q). For trnC-D, an internal primer C1 (5'-ACT TCC GTG AAA CAT GTG AAA-3') was designed in order to amplify only the region of interest, which was a c. 440 bp fragment at the 3' end of the region (trnC-D is in total about 3000 bp). This reduced the extension time from 4 min to 1 min. The annealing temperature was lowered from 58°C to 55°C to be able to amplify the trnC-D region in the same PCR cycle as the trnL-F region.

PCR-RFLPs were carried out on all pure populations collected in Great Britain to test the species-specific nature of these markers (Table 2.1). The following protocol was used for 10 µl PCR-RFLP reaction volumes: 3.9 µl purified water, 1 µl 10x buffer (Bioline), 0.1 µl bovine serum albumin (BSA, 10 mg/ml), 2.5 U enzyme (*Mse*1, *Eco*R1, *Spe*1 or *Hinf*1, depending on region) and 5 µl of PCR product. A total of 151 individuals were used from pure populations comprising 73 pure *G. urbanum* and 78 *G. rivale* samples. PCR-RFLPs were run out on a 1.5% agarose gel and banding patterns for all four regions were scored by eye for each sample.

| Primer | Annealing temp.$T_m$ [°C] | Extension time [m] | Approx. basepair length | Reference |
|---|---|---|---|---|
| psbA-trnH | 53 | 1 | 350 | (Hamilton, 1999) |
| atpB-rbcl | 50 | 2 | 900 | (Chiang *et al.*, 1998) |
| **trnC-D** | 58 | 4 | 3000 | (Demesure *et al.*, 1995) |
| trnD-T | 55 | 2 | 1600 | (Demesure *et al.*, 1995) |
| trnH-K | 62 | 2 | 1700 | (Demesure *et al.*, 1995) |
| **trnL-F** | 55 | 2 | 1300 | (Taberlet *et al.*, 1991) |
| **trnS-Q** | 47 | 3 | 1900 | (Grivet *et al.*, 2001) |
| **atpH-I** | 45 | 2 | 900 | (Grivet *et al.*, 2001) |
| psbC-trnS | 57 | 2 | 2000 | (Demesure *et al.*, 1995) |
| trnK1-K2 | 53.5 | 3 | 2600 | (Demesure *et al.*, 1995) |
| trnK-Q | 47.5 | 3 | 3000 | (Grivet *et al.*, 2001) |
| trnS-fM | 62 | 2 | 1700 | (Demesure *et al.*, 1995) |
| trnT-psbC | 52.5 | 4 | 3200 | (Grivet *et al.*, 2001) |

**Table 2.2.** cpDNA regions surveyed for species-specific markers with respective annealing temperatures, extension times and approximate base pair lengths. Primer pairs in bold were used for the survey of all pure *Geum urbanum* and *G. rivale* populations.

*AFLP fingerprinting*

To establish whether there are genetic differences between pure populations of *G. urbanum* and *G. rivale*, 30 individuals from nine British populations per species were screened for AFLPs (Table 2.1). The AFLP protocol was based on Vos *et al.* (1995) using the same adapters in the ligation process and the same primers in the pre-amplification stage. Restriction digestion included the following ingredients in a 15 µl reaction per sample: 2.15 µl purified water, 2 µl NEB2 buffer (New England Biolabs), 0.3 U *Eco*R1 (20.000 U/ml, New England Biolabs), 0.3 U *Mse*1 (50.000 U/ml, New England Biolabs), 0.50 µl bovine serum albumin (200 µg/ml, New England Biolabs) and 10 µl extracted DNA (about 50 ng of DNA on average). The mixture was then incubated at 37°C for two hours. For the ligation process 5.33 µl of following ingredients were added to the restriction digestion and left for three hours at room temperature: 2 µl T4 10 x ligase buffer (New England Biolabs), 0.5 µM *Eco*R1 adapter, 2.7 µM *Mse*1 adapter and 0.5 U T4 ligase (400 U/µl, New England Biolabs).

After diluting the digestion-ligation product ten times pre-amplification PCR was performed in 13 µl reaction volumes with following ingredients: 6 µl purified water, 1.3 µl 10x buffer (Bioline), 1.9 mM $MgCl_2$, 0.2 mM dNTPs, 0.3 µM *Eco*R1 primer + 1 (A), 0.3 µM *Mse*1 primer + 1 (C), 0.04 U Bio*Taq* (Bioline) and 3 µl diluted digestion-ligation product. PCR reactions were performed using the following profile: 2 min at 72°C, 25 cycles of 30 s denaturing at 94°C, 30 s annealing at 54°C, and 2 min extension at 72°C, ending with 30 min at 60°C to complete extension and subsequent storage at 4°C. After checking for the presence of a smear of fragments on a 1% agarose gel the amplification product was diluted ten times in purified

water. A selective PCR reaction with 2 µl diluted pre-amplification product was then performed in 10 µl reaction volume consisting of: 4.66 µl purified water, 1 µl 10x buffer (Bioline), 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.5 µM *Eco*R1 primer + 3 (AXX), 0.5 µM *Mse*1 primer + 3 (CXX) and 0.02 U Bio*Taq* (Bioline). PCR reactions were performed using the following profile: 2 min at 94°C, 30 cycles of 20 s denaturing at 94°C, 30 s annealing and 2 min elongation at 72°C, ending with 30 min at 60°C to complete extension. Annealing was initiated at a temperature of 66°C, which was then reduced by 1°C for the next 10 cycles and maintained at 56°C for the next 20 cycles.

All PCR reactions were performed on an MJ Research Dyad thermal cycler. Selective PCR products were run on an ABI 3730 automated sequencer following the manufacturer's instructions. LIZ-500 labelled internal size standard (PE Applied Biosystems) was added to each sample to size fragments. The data generated were analysed by the software GENEMAPPER 3.7 (PE Applied Biosystems), producing a data matrix for presence (1) or absence (0) of bands. This data matrix was exported into Quattro Pro 12 (Corel Corporation, 2004) and checked by eye against the electropherograms for each sample making manual adjustments to the data matrix wherever necessary. Markers which occurred only in one species but not in the other were considered to be species-specific. 67 primer combinations were tested on eight samples from each species. Eight primer combinations yielding 616 markers were chosen based on the number of species-specific markers produced and their ability to generate clear peaks: 6FAM-*Eco*R1 AAC/ *Mse*1 CTA, PET-*Eco*R1 AAT/ *Mse*1 CAGA, VIC-*Eco*R1 AGC/ *Mse*1 CTT, NED-*Eco*R1 ATC/ *Mse*1 CGA and 6FAM-*Eco*R1 AAC/ *Mse*1 CGA, PET-*Eco*R1 AAT/ *Mse*1 CTC, VIC-*Eco*R1 AGC/ *Mse*1 CGA, NED-*Eco*R1 ATC/ *Mse*1 CGA The selected eight *Eco*R1 primers were labelled with four different dyes to allow multiplexing.

## 2.2.6 Analysis of genetic differentiation

*AFLP markers*

To estimate species-specific genetic differences basic population genetic analysis was carried out using the program AFLP-Surv v1.0 (Vekemans *et al.*, 2002). To identify genetic differences between populations of the same species a subset of the populations screened was analysed for this purpose. This subset included two *G. urbanum* populations (U1, U2) and two *G. rivale* populations (R1, R2) with sample sizes of ten and eleven individuals per population, respectively (Table 2.1). To visualise the genetic structure of pure populations a Principal Coordinates Analysis (PCO) of AFLP data was performed with the program PAST (Hammer *et al.*, 2006). PCO is a multivariate technique which operates on a distance matrix of similarities or associations between a set of objects. The result is a two-dimensional image that approximates the distances in the genetic distance matrix. A Jaccard index was applied.

# 2.3 Results

## 2.3.1 Morphological variation

All nine floral, leaf and fruit measurements of individuals collected from three pure *G. urbanum* and three pure *G. rivale* populations taken in a common environment differed significantly between species and did not overlap for any of the recorded traits (Table 2.3). *G. urbanum* has larger stipules but smaller sepals and petals and a shorter upper style compared to *G. rivale*. The presence of a gynophore and a petal claw are unique to *G. rivale*.

| | **Stipule length** | **Stipule breadth** | **Sepal length** | **Sepal breadth** | **Petal length** | **Petal breadth** | **Upper style length** | **Claw length** | **Gyno-phore** |
|---|---|---|---|---|---|---|---|---|---|
| **Species means** | **(cm)** | **(cm)** | **(cm)** | **(cm)** | **(cm)** | **(cm)** | **(mm)** | **(mm)** | **(mm)** |
| *Geum urbanum* | 1.89 ±0.27 | 2.09 ±0.34 | 0.63 ±0.07 | 0.33 ±0.06 | 0.72 ±0.09 | 0.62 ±0.10 | 1.56 ±0.17 | absent | absent |
| *Geum rivale* | 0.76 ±0.17 | 0.45 ±0.22 | 0.93 ±0.17 | 0.53 ±0.10 | 1.14 ±0.14 | 0.95 ±0.17 | 4.82 ±0.94 | 5.13 ±0.83 | 6.17 ±1.67 |
| Sp. diff. | sig*** | sig*** | sig*** | sig*** | sig*** | sig*** | sig*** | sig*** | sig*** |
| **Population means** | | | | | | | | | |
| U1 (Braid Hills) | 1.94 ±0.34 | 1.96 ±0.36 | 0.64 ±0.08 | 0.34 ±0.05 | 0.73 ±0.10 | 0.59 ±0.10 | 1.55 ±0.15 | n/a | n/a |
| U2 (Colinton) | 1.94 ±0.25 | 2.21 ±0.34 | 0.61 ±0.08 | 0.33 ±0.08 | 0.70 ±0.06 | 0.61 ±0.08 | 1.56 ±0.16 | n/a | n/a |
| U3 (Craigmillar) | 1.76 ±0.23 | 2.04 ±0.28 | 0.64 ±0.05 | 0.36 ±0.05 | 0.75 ±0.11 | 0.67 ±0.13 | 1.60 ±0.21 | n/a | n/a |
| Pop. urb. | ns | ns | ns | ns | ns | ns | ns | n/a | n/a |
| R1 (Ben Lawers) | 0.87 ±0.10 | 0.58 ±0.30 | 0.96 ±0.16 | 0.59 ±0.06 | 1.19 ±0.14 | 0.99 ±0.13 | 4.77 ±0.56 | 5.19 ±1.03 | 5.35 ±1.77 |
| R2 (Moorfoot Hills) | 0.64 ±0.15 | 0.34 ±0.05 | 0.88 ±0.17 | 0.47 ±0.09 | 1.08 ±0.12 | 0.88 ±0.16 | 4.75 ±1.20 | 5.09 ±0.66 | 6.94 ±1.29 |
| Pop. riv. | sig*** | sig*** | ns | sig*** | sig* | sig* | ns | ns | sig* |

**Table 2.3.** Mean and standard deviation of morphological measurements taken in the greenhouse on 39 individuals from three pure *G. urbanum* and 32 individuals from three *G. rivale* populations. Measurements of all traits differed highly significantly between species (Sp. diff., Mann-Whitney test, $p \leq 0.001$). Pop. urb. and Pop. riv.: significant (sig*: $0.01 < p \leq 0.05$, sig**: $0.001 < p \leq 0.01$, sig***: $p \leq 0.001$) or not significant (ns) differences of traits between three *G. urbanum* and two *G. rivale* populations (third population excluded because of small sample size), respectively.

The three pure *G. urbanum* populations did not differ significantly from each other in these characters whereas the two larger *G. rivale* populations (because of small sample size R3 was excluded from this analysis) were more diverse and showed significant differences among populations in 67% of the traits (apart from sepal length, upper style length and claw length). Apart from the gynophore length the Ben Lawers population (R1) was larger than the Moorfoot Hills population (R2) in all characters measured (Table 2.3).

A PCA plot revealed that the two species *G. urbanum* and *G. rivale* were clearly and unambiguously separated by their morphology along the first component which accounted for more than 80% of the variation (Figure 2.2). All nine characters were more or less equally weighted for the first component thus contributing equally to the differentiation of the two species. The second component accounted for about 8% of the variation with stipule length, sepal and petal breadth being the most informative traits. A PCA plot of three pure Scottish *G. urbanum* populations using seven morphological traits (gynophore and claw length are *G. rivale* characters only) showed that all populations intermingled along the first and second component (37% and 17% of the variation explained, respectively) and did not show any appreciable grouping (Figure 2.3). A PCA plot for two pure Scottish *G. rivale* populations R1 and R2 using nine morphological traits revealed that although there was some grouping of the two populations no unambiguous clustering was apparent along the first and second component, which accounted for 35% and 19% of the variation, respectively (Figure 2.4).

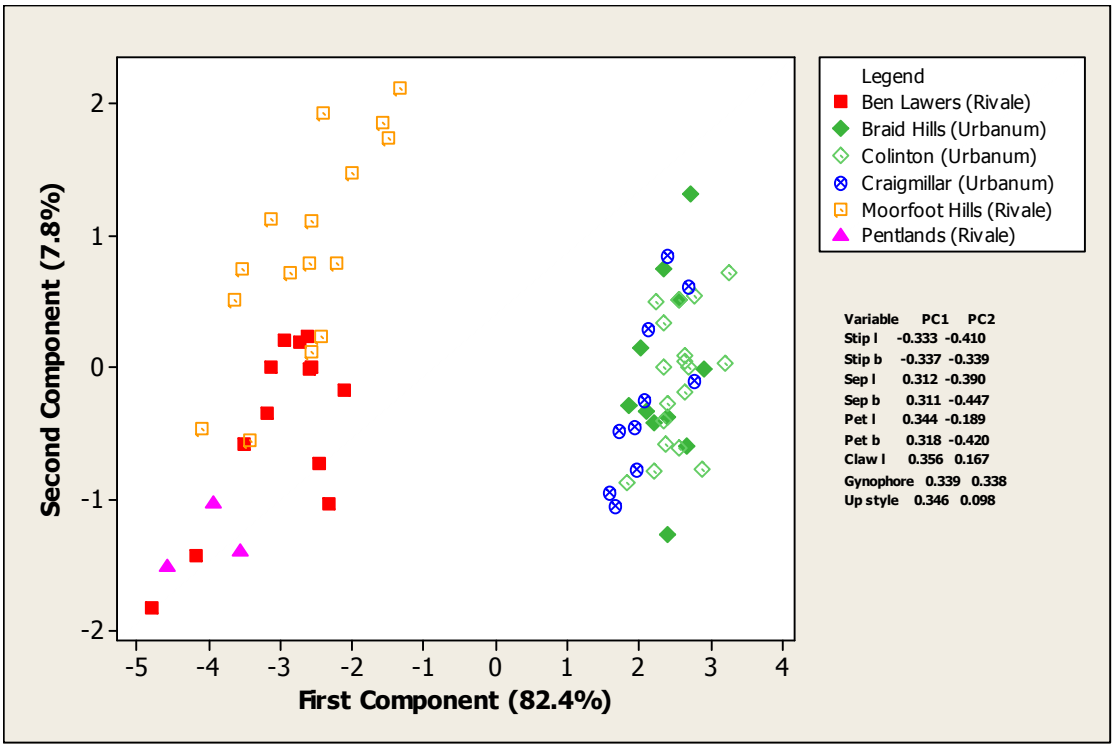



**Figure 2.2.** PCA plot using nine morphological characters taken in a common environment. Individuals were collected from three pure *G. urbanum* (39 samples) and three *G. rivale* (32 samples) populations. Inset specifies weighting of characters for the first and second component.

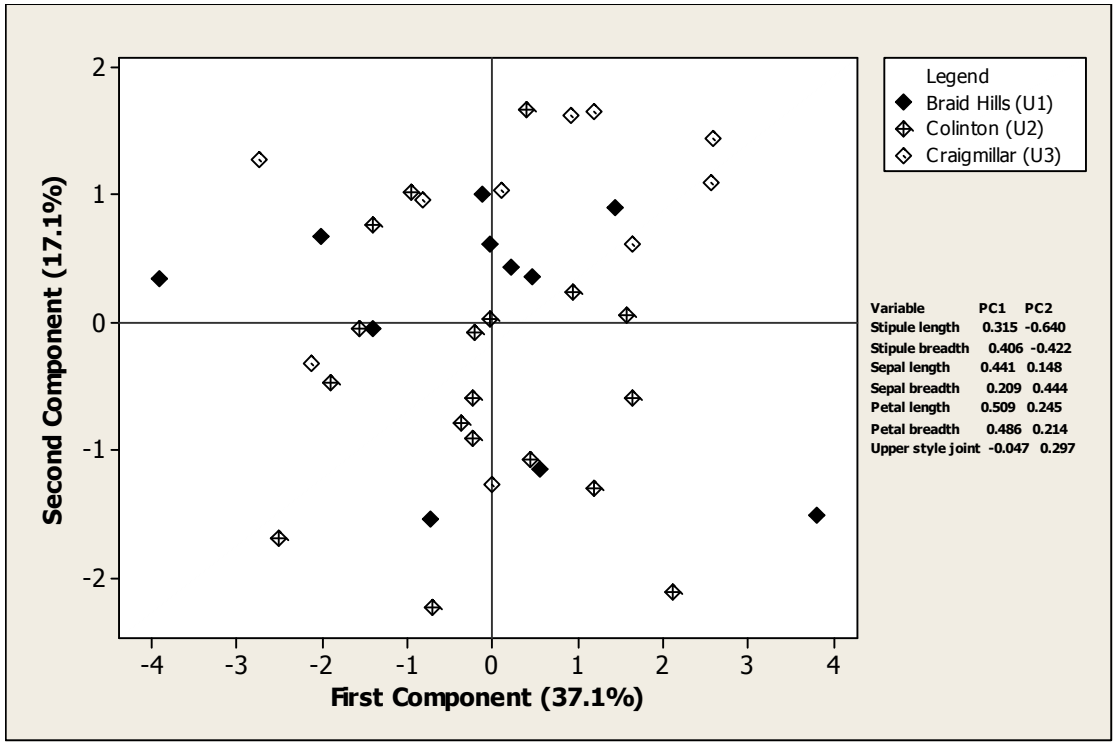



**Figure 2.3.** PCA of three pure *G. urbanum* Scottish populations U1 (11 samples), U2 (18 samples) and U3 (10 samples) using seven morphological characters. Inset specifies weighting of characters for the first and second component.

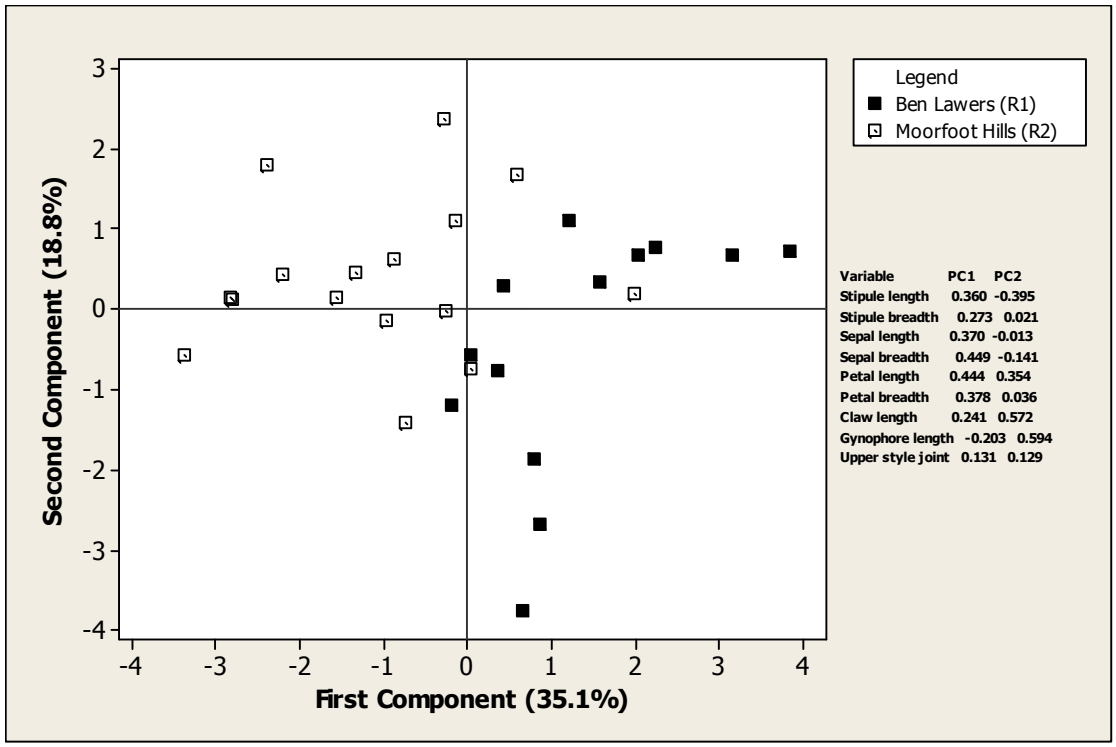



**Figure 2.4.** PCA of two pure *G. rivale* Scottish populations R1 (13 samples) and R2 (15 samples) using nine morphological characters. Inset specifies weighting of characters for the first and second component.

## 2.3.2 Genetic marker differentiation

*Chloroplast markers*

Out of 13 regions surveyed nine could be amplified successfully and four regions revealed four exploitable markers (Table 2.2). The PCR-RFLP analysis of all pure populations showed that none of the four markers were species-specific. It also showed that the four markers were not independent as the banding patterns of *Spe*1 and *Hinf*1 were correlated with the banding patterns of *Eco*R1 and could therefore be considered redundant. A total of four different haplotypes (A, B, C and D; Table 2.4) were present in the 151 individuals sampled from pure populations comprising 73 pure G. urbanum and 78 *G. rivale* samples (Table 2.5).

| | *Eco*R1(trnL-F) | | *Mse*1(trnC-D) | |
|---|---|---|---|---|
| **Haplotype** | **cut (390/630 bp)** | **no cut (1020 bp)** | **cut (100/340 bp)** | **no cut (440 bp)** |
| **A** | ✓ | | | ✓ |
| **B** | ✓ | | ✓ | |
| **C** | | ✓ | ✓ | |
| **D** | | ✓ | | ✓ |

**Table 2.4.** Haplotypes observed in pure *G. urbanum* and *G. rivale* populations performing cpDNA RFLPs using *Eco*R1 (trnL-F) and *Mse*1 (trnC-D).

| **Species** | **A** | **B** | **C** | **D** |
|---|---|---|---|---|
| *Geum urbanum* | 100.0% | 0.0% | 0.0% | 0.0% |
| *Geum rivale* | 25.3% | 1.3% | 70.9% | 2.5% |
| ***G. rivale* population** | | | | |
| R1 (Ben Lawers) | 0.0% | 0.0% | 96.7% | 3.3% |
| R2 (Moorfoot Hills) | 37.5% | 3.1% | 56.3% | 3.1% |

**Table 2.5.** Chloroplast haplotype (A, B, C, D) frequencies in pure *G. urbanum* and pure *G. rivale* populations in Great Britain.

In pure *G. urbanum* only haplotype A occurred whereas all haplotypes (A to D) occurred in pure *G. rivale* populations with haplotypes A and C being the most common (25% and 71%, respectively) and haplotypes B and D being very rare (1% and 3%, respectively; Table 2.5). There was a marked frequency difference of haplotypes between the two *G. rivale* populations R1 and R2. Haplotype C (97%) occurred in R1 in all but one individual which had haplotype D (3%) instead (haplotypes A and B were absent). Although haplotype C was also dominant in R2 it occurred at a lower frequency (56%) than in R1. More than a third of the samples in R2 had haplotype A (37.5%) with haplotypes B and D being very rare and B being unique to this population (3.1%). The $F_{st}$ value based on cpDNA data for the two *G. rivale* populations R1 and R2 was $F_{st}$=0.34.

*AFLP markers – species comparison*

616 markers using eight primer combinations were recovered, analyzing 30 individuals each from ten pure British *G. urbanum* and *G. rivale* populations (Table 2.1). 52 markers (8%) were species-specific out of which 40 (6%) were species-specific *G. urbanum* and 12 (2%) were species-specific *G. rivale* markers. There were 160 AFLP bands (26%) with a frequency difference of more than 70% and 124 bands (20%) with a frequency difference of more than 85% between the two species. The number of species limited markers (i. e. markers which occur only in one species but not in all individuals) in *G. rivale* (209) was much higher than in *G. urbanum* (79). The two species shared 272 loci (44%). The number of polymorphic loci was P=51% in *G. urbanum* and P=79% in *G. rivale* and the corresponding expected heterozygosity was He=0.18 and He=0.26, respectively (Table 2.6).

| **Species** | **# loci** | **#sp-sp** | **# sp-lim** | **# loci P** | **%P** | **He** |
|---|---|---|---|---|---|---|
| *Geum urbanum* | 390 | 31 | 79 | 199 | 51 | 0.18 |
| *Geum rivale* | 493 | 12 | 209 | 390 | 79 | 0.26 |

**Table 2.6.** Population genetic analysis based on AFLP markers of ten pure *G. urbanum* and ten pure *G. rivale* populations comprising 30 individuals per species. **# loci**: number of loci with at least one individual showing a band, **#sp-sp**: number of species-specific markers, i. e. markers which occur in all individuals of one species and in none of the other, **# sp-lim**: number of species limited markers, i. e. markers which occur in one species only but not in all individuals, **# loci P**: number of polymorphic loci, **%P**: percentage of polymorphic loci at the 5% level, **He**: expected heterozygosity.

| **Population** | **# samples** | **# loci** | **# loci P** | **%P** | **He** | **$F_{st}$** |
|---|---|---|---|---|---|---|
| *Geum urbanum* | | | | | | |
| U1 (Braid Hills) | 10 | 355 | 190 | 54% | 0.22 | 0.14 ± 0.14 |
| U2 (Colinton) | 10 | 355 | 124 | 35% | 0.16 | |
| *Geum rivale* | | | | | | |
| R1 (Ben Lawers) | 11 | 465 | 400 | 86% | 0.29 | 0.04 ± 0.02 |
| R2 (Moorfoot Hills) | 11 | 465 | 394 | 85% | 0.28 | |

**Table 2.7.** Population genetic analysis based on AFLP markers of two pure *G. urbanum* and two pure *G. rivale* populations. **# loci**: number of loci scored, **# loci P**: number of polymorphic loci, **%P**: percentage of polymorphic loci at the 5% level, **He**: expected heterozygosity, **$F_{st}$**: Fixation index with standard deviation.

*AFLP markers – population comparison*

The two *G. urbanum* populations U1 and U2 with 10 individuals each were scored for 355 markers. U1 showed a higher number of polymorphic loci (54%) and higher gene diversity He (0.22) compared to U2 (35% and 0.16, respectively; Table 2.7). The fixation index $F_{st}$ was 0.14 and not significantly different from zero due to a high standard error of 0.14 (calculated with the assumption of $F_{is}$ = 0.9, see Chapter 3). In contrast the two *G. rivale* populations R1

and R2 with 11 individuals each and scored for 465 markers (Table 2.7) displayed a very similar number of polymorphic loci (86% and 85%, respectively) and very similar gene diversity (0.29 and 0.28, respectively). $F_{st}$ equalled 0.04 ± 0.024 and was not significantly different from zero.

*AFLP markers – PCO analysis*

A PCO analysis showed that the two species are clearly and unambiguously separated along PCO1, which accounted for nearly 60% of the variation (Figure 2.5). Individuals from pure *G. rivale* populations formed a very loose group and spread considerably along PCO2, which accounted for 7 % of the variation. *G. urbanum* samples clustered much more closely forming a tight group along PCO1 and PCO2. A PCO plot for two pure populations of *G. urbanum* (U1 and U2) using 355 AFLP markers showed that U1 and U2 formed non overlapping groups along PCO1 and PCO2 which accounted for 27% and 20% of the variation, respectively (Figure 2.6). In a PCO plot for *G. rivale* populations R1 and R2 using 465 AFLP markers a similar picture emerges (Figure 2.7). The two populations were separated along PCO1 which accounted for 14% of the variation.

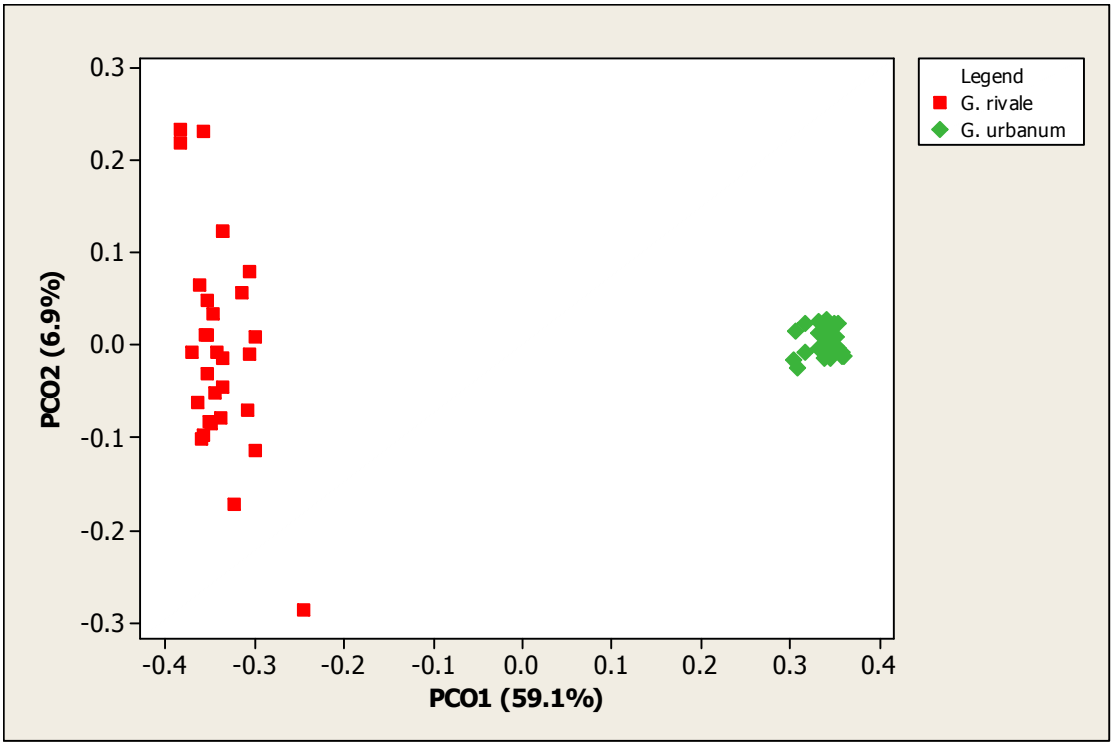



**Figure 2.5.** PCO plot of ten pure *Geum urbanum* (30 samples) and nine pure *G. rivale* (30 samples) populations collected in Great Britain based on 616 AFLP markers.

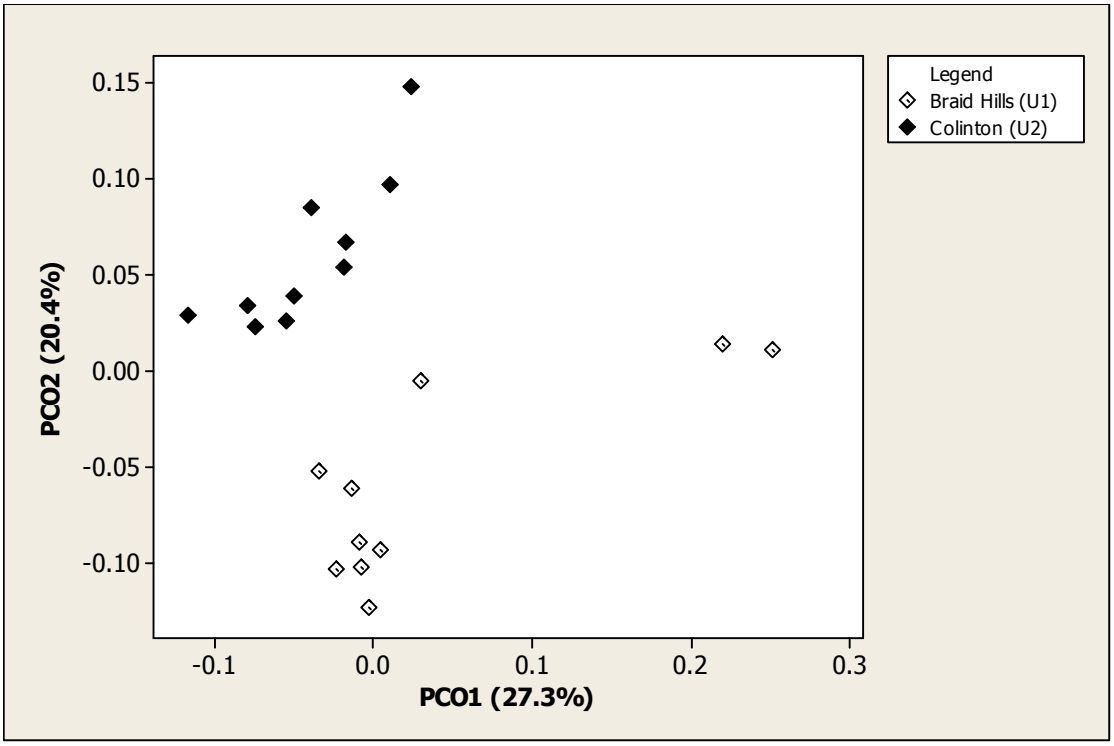



**Figure 2.6.** PCO plot of two pure *Geum urbanum* populations U1 and U2 (10 samples/population) collected in Scotland based on 355 AFLP markers. $F_{st}$=0.14 (±0.14)

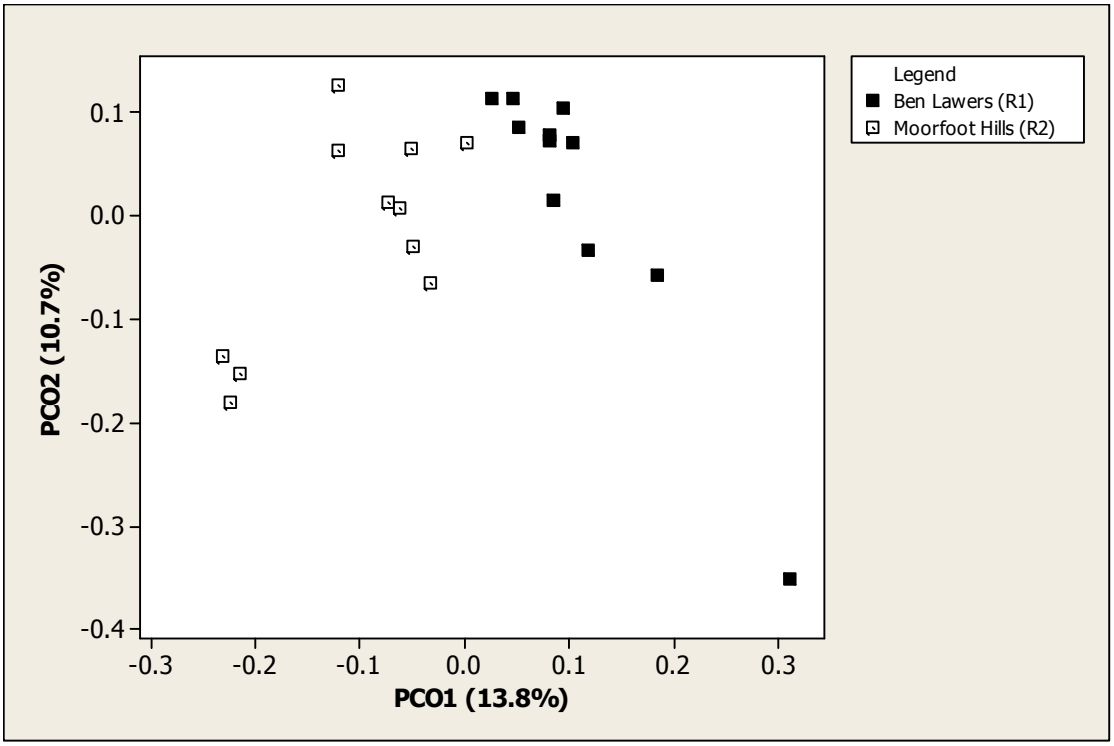



**Figure 2.7.** PCO plot of two pure *Geum rivale* populations R1 and R2 (11 samples/population) collected in Scotland based on 465 AFLP markers. $F_{st}$=0.04 (±0.02)

# 2.4 Discussion

## 2.4.1 Morphology

Results from the morphological analysis clearly show that individuals from pure *G. urbanum* and *G. rivale* populations are distinct. They do not overlap in any of the measured leaf, floral and seed characters which corresponds to the findings of earlier authors (Marsden-Jones, 1930; Prywer, 1932; Waldren *et al.*, 1989). A PCA plot revealed that the two species were unambiguously separated along the first component which accounted for more than 80% of the variation (Figure 2.2). All nine characters were more or less equally weighted for the first component thus contributing equally to the differentiation of the two species. This suggests that all nine traits will be useful in determining the identity of plants in hybrid zones. The second component accounted for 8% of the variation with stipule length, sepal and petal breadth being the most informative traits.

There was no intraspecific variation in morphology among the three *G. urbanum* populations. In contrast, there was significant morphological variation among the two *G. rivale* populations (Ben Lawers and Moorfoot Hills) with larger sample sizes. It has long been recognized that *G. rivale* is a very variable species showing considerable genecological variation and phenotypic plasticity in terms of leaf size, height, flower shape and length of gynophore. These differences are of a hereditary nature and have persisted in plants grown in cultures for years (Turesson, 1925, 1930; Gajewski, 1957). This has led to the recognition of various subspecies by some authors but Gajewski (1957) was convinced that no subspecies had developed within this species as morphological variation could not be correlated with geographical distribution.

In transplant experiments using Scandinavian *G. rivale* plants Turesson (1925) reported three ecotypes (alpine, subalpine and lowland). He demonstrated that the alpine type is expressed in lower stature, smaller basal leaves and earlier flowering compared to the lowland type. Turreson (1925) also noted that the subalpine plants are taller and have larger flowers and leaves than lowland plants which seems to be the case for most European subalpine populations (Taylor, 1997a). This is in line with the on average larger leaf and floral characters in the subalpine Ben Lawers population (600 m above sea level) compared to the lower lying (300 m above sea level) Moorfoot Hills population. Although there is some intermingling of individuals between the two populations in the PCA plot along the first component, the tendency to form two groups is also apparent (Figure 2.4).

In contrast, the three *G. urbanum* populations (Colinton, Braid Hills and Craigmillar) do not show any significant differences in their morphological characters. This can also be seen in the PCA plot where all individuals form an intermingling cloud without any appreciable grouping (Figure 2.3). One explanation for the lack of morphological differentiation could be gene flow via seeds as the sites are much closer together (between 2 and 5 km apart) than the two *G. rivale* populations (110 km apart). However, although *G. urbanum* generally shows marked variability in hairiness and leaf shape (Gajewski, 1957), characters which have not been measured in this study, no ecotypes or variability in other morphological traits have been reported in Great Britain.

## 2.4.2 Genetic markers

*Chloroplast DNA markers*

The chloroplast DNA analysis did not detect any species-specific markers. However, it showed that only one haplotype (A) occurs in *G. urbanum* whereas all four detected haplotypes occur in *G. rivale* at marked frequency differences. The two main haplotypes in *G. rivale* samples are C (71%) and A (25%) with B (1.3%) and D (2.5%) being very rare.

A possible explanation for the observed pattern could be that haplotype C was originally present in both species but subsequently was lost by genetic drift in British *G. urbanum* populations. This is supported by the analysis (not shown) of five *G. urbanum* samples from five countries in south and south-eastern Europe (Italy, Bulgaria, Macedonia, Montenegro and Serbia) which display haplotype C. Another explanation could be that haplotype A is a species-specific *G. urbanum* marker and *G. rivale* samples with this haplotype are the products of a past hybridization event with repeated backcrossing to *G. rivale*. This is supported by the absence of haplotype A in the subalpine *G. rivale* population R2 (Ben Lawers), a habitat in which no *G. urbanum* individual has ever been recorded and the moderate frequency of haplotype A (38%) in the lower lying *G. rivale* population R2 (Moorfoot Hills). Although no *G. urbanum* populations have been recorded in an area of 6x6 $km^2$ around R2, which is also dominated by moorland, a habitat unsuitable for *G. urbanum*, it is possible that the hybridization event in question could have predated the colonization of the Moorfoot Hills by *G. rivale*. More extensive sampling with larger sample sizes and additional data sources are therefore needed to gain more robust evidence for either of these speculations.

*AFLP markers*

A PCO plot using 616 AFLP markers (Figure 2.5) separates the two species clearly and unambiguously along PCO1 which accounted for 59% of the variation. *G. urbanum* and *G. rivale* also showed genetic differences in respect to the percentage of polymorphic loci (51% and 79%, respectively) and gene diversity (0.18 and 0.26, respectively). The reason for the generally lower number of polymorphic loci and gene diversity of *G. urbanum* can be understood by looking at the mating system. *G. urbanum* is mainly selfing whereas *G. rivale* is mainly outcrossing (Chapter 3).

Selfing plant species generally tend to exhibit lower genetic diversity (Hamrick & Godt, 1996) but it is noteworthy that selfing as such does not reduce genetic diversity. Selfing rearranges genetic diversity in terms of genotypes but does not change allele frequencies. However, genetic drift is a more effective force in selfing populations because under complete selfing the effective population size is halfed compared to complete outcrossing populations. Additionally gene flow between selfing populations is restricted and intrapopulational genetic diversity can not be shared as effectively as in outcrossing populations. Selfing plant species also have a much higher chance to establish new populations from only one or a few individuals as no cross fertilization for the production of offspring is needed (Baker, 1955; Stebbins, 1957b). However, if the founding of such a new population is successful a lot of the genetic variation present in the original source population will have been lost leading to a genetic bottleneck situation. The combination of these factors generates the observed pattern of generally lower genetic diversity in selfing species.

How does the genetic differentiation between *G. urbanum* and *G. rivale* compare to other hybridizing species? In a recent study looking at the hybridization process between *Silene dioica* and *S. latifolia* 3.3% of the AFLP markers were species-specific (Minder *et al.*, 2007) and 10% showed a frequency difference of more than 70% between the two species (Karrenberg & Favre, 2008). These figures are much lower than those for this study (8% and 26%, respectively) despite the more inclusive definition of 'species-specific' by Minder *et al.* (2007) as markers were included as species-specific when the genotype frequency difference between the two species was at least 0.85. Bodénès *et al*. (1997) investigated the geographic variation of the species differentiation throughout their natural range in *Quercus petraea* and *Q. robur*. After screening 2800 PCR amplification products using random primers, they found only 2% of the amplified fragments that exhibited significant frequency differences between the two species and none of them was species specific. A similar study carried out on two hybridizing Japanese oak species (*Q. dentata* and *Q. crispula*) did not detect any diagnostic

AFLP markers revealing a low degree of genetic differentiation between the two species (Ishida *et al.*, 2003).

The hybridizing species *Salix fragilis* and *S. alba* displayed a very similar percentage of discriminating AFLP bands (Beismann *et al.*, 1997). 11% of the markers were species-specific with 8% only present in *S. fragilis* and 3% in *S. alba* (in comparison 6% are species-specific *G. urbanum* and 2% *G. rivale* markers). Much higher numbers of species-specific AFLP markers were found in hybridizing *Spartina alterniflora* (accidentally introduced to Europe in the 19$^{th}$ century from eastern North American) and *Spartina maritima* (native to Europe) in Great Britain: 54% of the markers were diagnostic fragments (Salmon *et al.*, 2005). However, this has to be seen in the context of the extreme allopatry of the two parental species and their sterile hybrid, which offers much less opportunity for gene exchange. Three species of *Vasconcellea* also showed a larger percentage of species-specific AFLP bands in comparison to this study ranging from 19% to 30% (Van Droogenbroeck *et al.*, 2006).

This comparison shows that a wide spectrum of interspecific genetic differences between hybridizing species exist, ranging from hardly any differentiation at all in *Quercus* to high levels in *Spartina*. However, the key point that emerges is that hybridization between *G. urbanum* and *G. rivale* has not eroded the genetic integrity of the two species as it appears to have happened in *Quercus*. The selected reference populations of *G. urbanum* and *G. rivale* examined here were clearly genetically distinct.

There was no significant genetic differentiation at the population level (i. e. $F_{st}$ was not significantly different from zero) within *G. urbanum* and *G. rivale*. This is in contrast with PCO plots where the two respective populations of each species did not overlap and formed more or less clearly separated groups. Hollingsworth & Ennos (2004) illustrated in simulation studies that there was no straightforward relationship between partitioning of genetic variation among populations on the one hand, and clustering of individuals on a neighbour joining tree on the other. If enough markers were employed (e.g. 250 AFLP loci) then mutually exclusive grouping of individuals into their respective populations can occur despite low differentiation between populations (e.g. $F_{st}$=0.14 in their simulations). This helps explain the apparent contradiction between non-significant $F_{st}$ values and population groupings on the ordination plots detected here.

In terms of morphology *G. urbanum* populations showed no significant differences in the nine characters measured whereas *G. rivale* populations differed significantly in six out of nine traits. However, these interpopulational differences were of the order of one magnitude less

than interspecific differences and did not put the morphological distinctiveness of the two species in question.

## 2.5 Summary

*G. urbanum* and *G. rivale* are morphologically and genetically distinct species, although *G. rivale* is a more diverse species in terms of morphology and genetic diversity than *G. urbanum*. The two species do not overlap in any of the nine measured leaf, floral and seed characters and form unambiguously separated genetic clusters in a PCO plot using 616 AFLP markers. The distinct morphology of the parental species and the intermediate level of genetic differentiation based on species-specific AFLP markers (8% of all markers scored were species specific) suggest that the detection of first generation and early backcross hybrids on a phenetic and molecular level is possible.

# *3.* BREEDING SYSTEM OF *GEUM URBANUM* AND *G. RIVALE*

## Abstract

To be able to study the evolutionary consequences of hybridization between selfing and outbreeding plant taxa the mating system of the study species *G. urbanum* and *G. rivale* have to be established beyond circumstantial evidence. Direct breeding system estimates based on four microsatellite loci in *G. urbanum* and one locus in *G. rivale* were carried out in two pure Scottish populations per species using mother-progeny arrays consisting of 20 maternal parents per population and five of their open-pollinated offspring. Indirect estimates of the mating system were obtained by calculating $F_{is}$ in the adult generation. Additionally, controlled crosses were performed in two *G. urbanum* and two *G. rivale* populations to establish whether there is evidence for a self-incompatability (SI) system in either species. Direct and indirect estimates of the breeding system yielded similar results and despite variation in outcrossing rates between populations of *G. urbanum* and *G. rivale* former species is highly selfing with only 6% to 23% of its progeny outcrossed in contrast to latter where 66% to 100% of the progeny are outcrossed. Auto-pollination as well as controlled selfing and outcrossing experiments suggested that *G. urbanum* is a completely self-compatible species (SCI=1) whereas *G. rivale* is a moderately self-incompatible species (SCI=0.30) with a high variability in individual genotypic expression of SI.

# 3.1 Introduction

Breeding systems have long been a topic of interest among evolutionary biologists (Darwin, 1876; Fisher, 1941). The evolution of mixed mating systems in hermaphrodites where plants have the ability to reproduce by both self-fertilization (selfing) and mating with other individuals (outcrossing), has drawn a lot of attention (Jain, 1976; Lande & Schemske, 1985; Charlesworth *et al.*, 1990b; Charlesworth, 1992; Goodwillie *et al.*, 2005). Selfing offers reproductive assurance, allows locally adapted genotypes to exist and also has a 3:2 transmission advantage compared to outcrossing (Charlesworth & Silvertown, 2001). This cost of outcrossing or the selective advantage of selfing was first pointed out by Fisher (1941) and arises because a maternal parent which reproduces by selfing can also sire offspring by pollinating a different plant. Selfing transmits both maternal alleles in the selfed progeny and another maternal allele in the outcrossed offspring thus contributing three alleles to the next generation. In contrast, outcrossing individuals can only provide two alleles (one as pollen donor and one as ovule donor) and therefore selfers have a 50% transmission advantage. Consequently alleles enabling for self fertilization are expected to spread unless selfed individuals display a compensating disadvantage in survival or reproduction.

This compensating disadvantage of selfing is inbreeding depression. A major consequence of selfing is that it leads to increased homozygosity. Deleterious or lethal mutations previously hidden as heterozygotes are rendered homozygous and are therefore expressed which results in low fitness of selfed compared to outcrossed progeny. This suggests that outcrossing will only be maintained if selfed offspring have half or less the fitness of outcrossed offspring (i. e. if inbreeding depression $\delta \geq 0.5$ where $\delta$ is defined as 1 – [fitness of selfed progeny/fitness of outcrossed progeny]). This argument assumes that self-fertilization does not diminish the selfer's ability to serve as an outcross pollen parent. However, selfing plants often have reduced flower size (Lloyd, 1965; Dole, 1992) and less pollen output (Cruden, 1977) than outcrossing taxa. Pollen discounting, the extent to which selfing reduces an individual's contribution to the outcross pollen pool will therefore diminish the 50% transmission advantage so that lower levels of inbreeding depression can block the evolution of selfing (Holsinger, 2000; Charlesworth & Silvertown, 2001).

Lande & Schemske (1985) showed that inbreeding depression should evolve jointly with the level of selfing so that inbreeding populations purge their genetic load over time leading to two stable outcomes of mating system evolution (i. e. predominantly selfing and predominantly outcrossing). However, later studies analyzing the distribution of mating systems in seed plants showed that despite a weak bimodality in the distribution (Figure 3.1), mixed mating systems are common with about 40 to 50% of the studied species showing

outcrossing rates of 20 to 80% (Vogler & Kalisz, 2001; Goodwillie *et al.*, 2005). The high frequency of intermediate selfing and outcrossing rates indicates that mixed mating systems can be evolutionary stable although some phylogenetic studies show that transitions between mostly selfing and mostly outcrossing are common in some plant taxa (Kohn *et al.*, 1996; Goodwillie, 1999). This suggests that the model of Lande & Schemske (1985) is based on oversimplified theories and in reality various factors act together differently in different plants creating the more complex picture we see in terms of plant mating systems.

Factors which can lead to evolutionary stable mixed mating systems include local adaptations (Schoen & Lloyd, 1984; Holsinger, 1986), biparental inbreeding (Uyenoyama, 1986), declining fitness of selfed progeny over successive generations (Jarne & Charlesworth, 1993), frequency- and density dependant reproductive success of genotypes (mass action model, Holsinger, 1991) and differences in male and female fertility, for instance if individuals with high male function have lower levels of femal function (Charlesworth & Silvertown, 2001). Apart from the mass-action model put forward by Holsinger (1991) most other plant mating evolution models include inbreeding depression as a major force opposing selfing (reviewed in Goodwillie *et al.*, 2005).

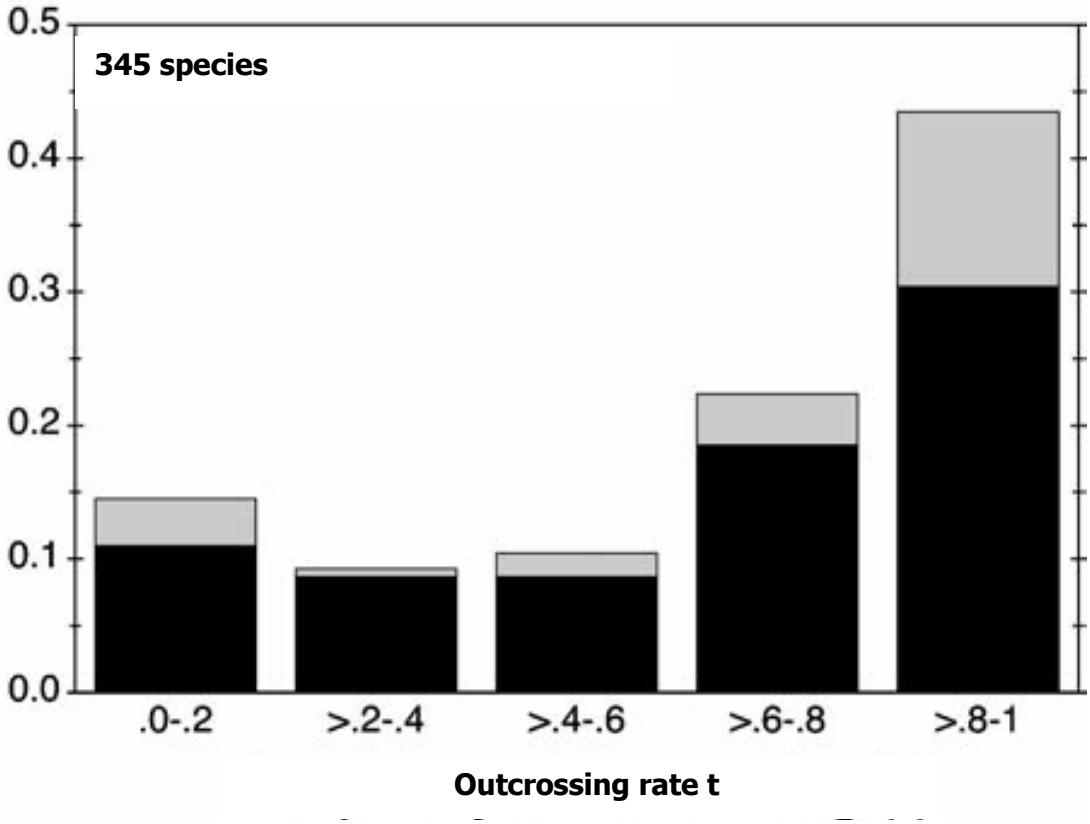



**Figure 3.1.** Distribution of estimated outcrossing rates in 345 seed plant species. Histograms are divided in biotically pollinated (insects, birds or bats, black) and abiotically pollinated (wind or water, grey). Taken from Goodwillie *et al.* (2005).

It is well known that outcrossing rates can vary considerably among closely related species and even among populations within species (Jain, 1976; Barrett & Harder, 1996; Cheptou *et al.*, 2002). The outcrossing rate t among populations of the annual *Gilia achilleifolia* ranged from 0.15 to 0.96 (Schoen, 1982) and varied even more from 0.002 to 0.96 in populations of the perennial *Eichhornia paniculata* (Barrett & Husband, 1990) and span the whole range from 0 to 1 in *Lupinus nanus* (Harding *et al.*, 1974). Outcrossing rates can also vary significantly between years as a study on the mating system of *Mimulus guttatus* revealed where the two analysed populations experienced a drop in t from 0.91 to 0.76 and 1 to 0.66 (Willis, 1993).

An important determining factor on how selfing a plant can be is its degree of self-incompatibility (SI). Self-incompatible plants reject their own pollen (i. e. self pollen does not germinate on the stigma or pollen tube growth of self pollen is slowed down compared to outcross pollen which will result in a failure of ovule fertilization). Self-sterility is controlled by self-incompatibility loci (S-locus) which may be comprised of one or several linked genes usually with many different alleles (Nasrallah & Nasrallah, 1986; Stone & Goring, 2001). Pollen is rejected when it comes from a plant with the same allele as the one that is being pollinated. Two different self-incompatibility systems are widely recognized although recent molecular data suggest that other molecular mechanisms of SI exist as well (Hiscock & McInnis, 2003). In the more inclusive gametophytic (GSI) system, the recognition mechanism only depends on the haploid genotype of the pollen grain. A pollen grain produced by a plant heterozygous (e. g. S1/S2) for the S locus will, for example produce pollen grains with the incompatibility alleles S1 and S2. S2 pollen grains are rejected by diploid S1/S2 stigmas but not by S1/S3 stigmas (Figure 3.2 A). In the more exclusive sporophytic system (SSI), stigmas can recognize the diploid origin of the pollen grain via the pollen exine, the outer pollen wall which was formed by the diploid tissue. S1 pollen grains of S1/S2 origin are therefore not only rejected by S1/S2 stigmas but also by S2/S3 stigmas (Figure 3.2 B).

Gametophytic SI systems have been observed in families such as Rosaceae, Solanaceae, Papaveraceae and Scrophulariaceae whereas sporophytic SI is present in Brassicaceae, Convolvulaceae and Asteraceae (Igic & Kohn, 2001; Hiscock & McInnis, 2003). It is estimated that about 60% of angiosperms have a SI system in place (Hiscock *et al.*, 2000). SI is usually not a qualitative trait (i. e. if a species is self-incompatible it is automatically 100% outcrossing). It is more often a quantitative trait with considerable phenotypic plasticity and can vary with the composition of the pollen load (pure self pollen versus a mixture of self and outcross pollen), time of the year, the absence or presence of developing fruits, flower age and environmental conditions such as temperature (Lloyd & Schoen, 1992; Vogler *et al.*, 1998; Stephenson *et al.*, 2000; Goodwillie *et al.*, 2004; Travers *et al.*, 2004; Good-Avila *et al.*,

2008). Sometimes an increase of self-fertility on plants with no developing fruits is observed which might be due to an increase in floral longevity. Because no resources are invested in developing fruits, slower growing self pollen tubes have enough time to fertilize the ovary before the flower wilts (Good-Avila *et al.*, 2008).

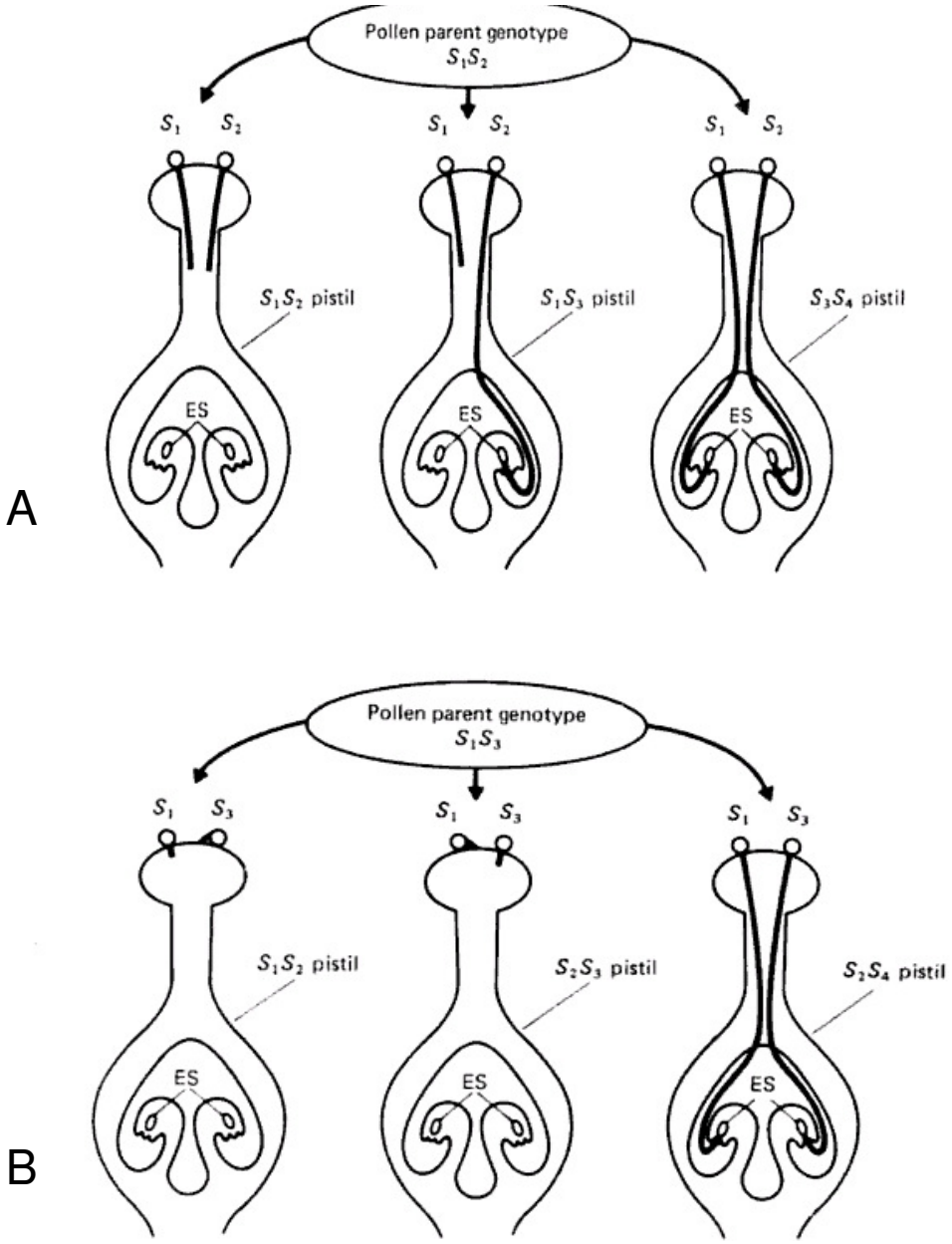



**Figure 3.2.** Gametophytic (A) and sporophytic (B) self-incompatibility systems. From Briggs & Walters (1997).

Degrees of SI can also vary between populations (Tsukamoto *et al.*, 1999; Stone *et al.*, 2006; Good-Avila *et al.*, 2008) and some studies have documented examples of SI species with large population sizes located near the centre of their distribution range and derivative self-compatible (SC) species or populations characterized by colonizing lifestyles in peripheral habitats (Stebbins, 1957b; Rick *et al.*, 1979; Barrett *et al.*, 1989; Lipow *et al.*, 1999). However, little is still known about the genetic and molecular mechanisms that cause the variation in the strength of SI or the consequences of such variation for the breeding systems (Stephenson *et al.*, 2000). Self-incompatibility systems have a major influence on outcrossing rates but there are many other factors which can significantly determine levels of selfing and outcrossing such as population density (Karron *et al.*, 1995; Lu, 2000), physical separation of anthers and stigmas in individuals (Karron *et al.*, 1997), environmentally induced variation in floral traits (Elle & Hare, 2002) and influence of pollinator type and behaviour (Brunet & Sweet, 2006). In animal-pollinated species the reliability of pollinators can also play a central ecological role and determine if plant populations are primarily outcrossed or self-pollinated (Levin, 1972; Holsinger, 1996; Schoen *et al.*, 1996; Takebayashi & Morrell, 2001).

The study species *G. urbanum* and *G. rivale* are reported to have a mixed mating system with *G. urbanum* being mainly selfing and *G. rivale* mainly outcrossing (Knuth, 1898: 363). In the seminal work on the floral biology of plants, Knuth (1898) states that although outcrossing is possible in the early and later stages of the weakly protogynous flowers of *G. urbanum*, spontaneous selfing is common due to the close proximity of the stamens to the anthers and a lack of pollinator visits (which are attracted to larger flowers). In contrast, *G. rivale* is reported to be mainly outcrossing as the flowers are weakly protogynous and stigmas are well above the anthers in the early stages of flowering. However, selfing is possible in the later stages of anthesis because elongating stamens reach the same height as the stigmas and pollen can then be deposited on nearby stigmas when flowers are closing (Knuth, 1898). Langlotz (2007) established the pollen-ovule (P/O) ratio in two Alpine *G. rivale* populations which, despite considerable variation between the two populations (P/O=999±77 and 3,052±344) are indicative of facultative outcrossing (Cruden, 1977).

In a recent detailed account of *G. urbanum* (Taylor, 1997b) and *G. rivale* (Taylor, 1997a) information on the breeding system of the two species is taken from Knuth (1898) and no new evidence is presented. Arens *et al.* (2004) developed 10 microsatellite loci for *G. urbanum* and surveyed 688 individuals in 45 European populations. The $F_{is}$ value of 0.93 (averaged over all loci and populations) is the first non circumstantial evidence for the selfing nature of this species. Another study using six microsatellite loci on 18 Belgian populations reports an equally high $F_{is}$ value of 0.89 (Vandepitte *et al.*, 2007). However, no direct estimates of the breeding system (i. e. mother – progeny arrays) have been conducted so far

for *G. urbanum* and no data at all are available for the breeding system of *G. rivale*. Molecular techniques provide a very efficient tool to estimate gene flow and infer breeding systems directly or indirectly. The direct approach to establish breeding systems is to screen mother – progeny arrays for highly variable codominant molecular markers such as microsatellites applying the mixed mating model. This model assumes that a certain percentage of offspring is derived from self pollination (s) whereas the remainder is derived from outcrossing at random (t) so that s+ t = 1. The assumptions made are (1) that excess of self-fertilization is the only factor which may cause the actual mating system to deviate from random mating, (2) that t is uniform for maternal genotypes and pollen alleles (i.e., there is no incompatibility and mating events are independent) and (3) that pollen frequency is uniform over maternal genotypes (Brown & Allard, 1970).

The indirect approach to infer the breeding system uses the inbreeding coefficient $F_{is.}$ The observed genetic structure in the adult population is compared to the genetic structure expected under Hardy-Weinberg assumptions and the difference between the two (expressed by $F_{is}$) is an indication of the mating system. For sexually reproducing populations using neutral markers, $F_{is}$ values lie between 0 and 1 ranging from complete outbreeding to complete selfing, respectively. The advantage of this method is that seeds from adult plants do not need to be collected and germinated providing a quick albeit less powerful approach to infer the breeding system.

To study the breeding system of *G. urbanum* and *G. rivale* pure populations of either species were sampled from Scotland. The same considerations as detailed in Chapter 2.1 were applied to choosing populations using pure populations of either species close to the study hybrid zone outside Edinburgh. A byproduct of using multiple populations for the breeding system analysis is that information regarding the genetic differentiation of populations can be gathered. This can in turn be used to test the findings of the mating system analysis as populations of selfing plant taxa generally show higher genetic differentiation than outcrossing taxa (Hamrick & Godt, 1996). High levels of selfing can be a very effective isolation mechanism as gene flow is very restricted and can only take place via seed dispersal. Beneficial mutations and alleles are kept within the physical boundaries of populations and can not be exchanged between populations. This leads to high genetic differentiation over time in selfing plant taxa. In contrast, outcrossing promotes gene flow and alleles are easily shared between neighbouring populations leading to little genetic differentiation between outcrossing populations.

To be able to study the evolutionary consequences of hybridization between selfing and outbreeding plant taxa the mating system of the study species *G. urbanum* and *G. rivale* have

to be established beyond circumstantial evidence. Are there truly two different mating systems involved when the two species hybridize? The aim of this chapter is to gain a better understanding of the mating system of *G. urbanum* and *G. rivale* by answering the following questions:

1. Do *G. urbanum* and *G. rivale* differ in their ability to auto-pollinate and if so could this be due to a self-incompatibility system?
2. Do *G. urbanum* and *G. rivale* differ significantly in their mating system?
3. Is there evidence for interpopulational variability in the breeding system?
4. Is $F_{st}$ larger in presumed selfing *G. urbanum* than in presumed outcrossing *G. rivale*?

# 3.2 Materials and Methods

## 3.2.1 Sampling material

Leaf samples of 20 to 36 individuals per population from four pure Scottish *G. urbanum* populations and two pure Scottish *G. rivale* populations were collected in silica gel in late summer 2005 and 2006. Four of the Scottish populations used for the analysis of morphological and genetic variation in pure populations of *G. urbanum* and *G. rivale* (Chapter 2) were included for the breeding system analysis (U1, U2, R1 and R2). Two additional *G. urbanum* populations were collected in the surroundings of Edinburgh (U2a and U4). Population U2a (Colinton Woods) was collected from a semi natural woodland area in Colinton about 50 m from population U2 (Colinton Ruderal), which is located along a public footpath in a disturbed, ruderal area. Samples of population U4 (Gorebridge) were taken from a natural woodland outside Edinburgh.

| Species | Location | Altitude | Coordinates | Sampling | Used in |
|---|---|---|---|---|---|
| **Geum urbanum** | Edinburgh (U1)<br>Braid Hills | 170 m | $55^0$ 56′ N<br>$03^0$ 12' W | 31 adults<br>98 offspring<br>129 total | Autopoll.<br>SI<br>$F_{is}$<br>MP-array |
| | Edinburgh (U2)<br>Colinton Ruderal | 120 m | $55^0$ 55' N<br>$03^0$ 14' W | 20 adults<br>95 offspring<br>115 total | $F_{is}$<br>MP-array |
| | Edinburgh (U2a)<br>Colinton Woods | 120 m | $55^0$ 51' N<br>$03^0$ 14' W | 29 adults | Autopoll.<br>SI<br>$F_{is}$ |
| | Edinburgh (U3)<br>Craigmillar | 70 m | $55^0$ 56′ N<br>$03^0$ 09′ W | 9 adults | Autopoll. |
| | Edinburgh (U4)<br>Gorebridge | 90 m | $55^0$ 55' N<br>$03^0$ 04' W | 36 adults | $F_{is}$ |
| **Geum rivale** | Stirling (R1)<br>Ben Lawers | 600 m | $56^0$ 32′ N<br>$04^0$ 17' W | 30 adults<br>95 offspring<br>125 total | Autopoll.<br>SI<br>$F_{is}$<br>MP-array |
| | Edinburgh (R2)<br>Moorfoot Hills | 300 m | $55^0$ 44′ N<br>$03^0$ 01' W | 36 adults<br>86 offspring<br>122 total | Autopoll.<br>SI<br>$F_{is}$<br>MP-array |

**Table 3.1.** Location of individuals from pure Scottish *Geum urbanum* and *G. rivale* populations used for the breeding system analysis. Autopoll.: auto-pollination; SI: self-incompatibility analysis, MP-array: mother-progeny array.

The same criteria for pureness of populations were applied as in Chapter 2. Open pollinated seeds were also collected from 20 individuals from two pure *G. urbanum* (U1 and U2) and two pure *G. rivale* populations (R1 and R2, Table 3.1). Seeds were sown in pots filled with peat free compost and overwintered outside. In spring 2007 five seedlings per parent were randomly chosen and potted on, resulting in 20 families per population for a detailed breeding system analysis (Table 3.1).

## 3.2.2 Auto-pollination and SI

To establish whether spontaneous selfing is a possible means of seed production in *G. urbanum* and *G. rivale* one to two flowers per maternal individual were bagged with bridal veil (mesh size 0.5 x 0.5 mm) and the absence or presence of seeds was recorded at the end of the growing season. The experiment took place in an unheated greenhouse and was carried out using populations U1, U2a and U3 (*G. urbanum*) and R1 and R2 (*G. rivale*).

Controlled crosses were performed within *G. urbanum* populations U1 and U2a as well as *G. rivale* populations R1 and R2 to establish whether there is evidence for a SI system in either species. Each maternal parent was selfed using pollen from a flower on the same plant and outcrossed with two other maternal parents from the same population.

To establish the right pollination time the receptivity of stigmas was tested using a hydrogen peroxide test (Zeisler, 1938). Receptive stigmas are characterized by high enzymatic activity and the presence of peroxidase coincides with this developmental stage (Dafni *et al.*, 2005). Hydrogen peroxide ($H_2O_2$) produces a bubble as a result of peroxidase presence indicating that the stigma is receptive. Stigmas from flowers in various stages of opening were immersed in a 6% $H_2O_2$ solution and the presence or absence of bubbles was noted. A disadvantage of this test is that bubbles are also produced by pollen grains on the stigmatic surface or by injured stigmatic tissue and thus lead to unreliable results. To avoid this from happening, stigmas were checked for deposited pollen or tissue rupture under a stereo microscope (40x) prior to testing. The hydrogen peroxide test clearly showed that *G. urbanum* and *G. rivale* are protogynous as stigmas were already receptive in closed buds when all stamens were still closed. Stigmas remained receptive until some time after the first pollen had been shed.

The pollination procedure involved two steps. In the first step the sepals and petals of pollen donating flowers (male parents) were removed and the flowers of ovule donating female parents were carefully emasculated with a pair of tweezers. Only flowers of maternal parents where all stamens were still firmly closed were emasculated. In the second step all the pollen

shedding anthers of the male parent were brushed against the stigmas of the receptive maternal parent repeatedly to transfer a sufficient pollen load on all stigmas. Both steps were carried out under a stereo microscope. Pollinated flowers were then labelled and plants were kept in a pollinator free cage. Seeds were harvested when ripe and seed numbers were established for selfed and outcrossed seed heads.

## 3.2.3 Microsatellites

DNA was extracted from the collected samples as detailed in Chapter 2.2.5 (p 31). Microsatellite markers developed for *G. urbanum* by Arens *et al*. (2004) were used for this study and tested on a subsample of *G. urbanum* individuals from different populations. A total of six primer pairs yielding consistent polymorphic patterns with a maximum number of two alleles per individual were chosen to genotype further samples (Table 3.2).

| Locus | Repeat motif | Primer sequence | $T_m$ |
|---|---|---|---|
| WGU1-33 | $CT_{14}$ | F:TTTTATGGCTGAAGCAGTTTATTTT<br>R:ATGGAGTAGTTGGGTCTCTTCTTTT | 53°C |
| WGU2-10 | $AAC_{20}$ | F:TGGGCATTGCCATTGTACATACTTT<br>R:GAACAGTCTGCTATAGTTGGAGAAA | 53°C |
| WGU2-48 | $GTT_{12}$, $GTT_7$ | F:TAAAAATTAGGGCATACGGAGAAAT<br>R:TTAGACAAAACAATAACCCACAGT | 53°C |
| WGU6-1 | $GAA_{15}$ | F:CCGCCGATATAAAAGGCACACCAA<br>R:GAGAGCCGCCGTGAGGGAATGA | 60°C |
| WGU6-5 | $AGA_{28}$ | F:AACCCCTGACTAAAGAGAAAAA<br>R:AGAAGAGGGTATCAAAAAGATGACT | 53°C |
| WGU8-1 | $GAA_4$, $GAA_6$ | F:AAGACCCTCCAGCCGCTTCATC<br>R:CAGATCCGTCGTCGTCTTCGTCATC | 60°C |

**Table 3.2.** Microsatellite primers used for the breeding system analysis. F: forward primer, R: reverse primer, $T_m$: annealing temperature. Nomenclature follows Arens *et al*. (2004).

| Population/Locus | 1-33 | 2-10 | 2-48 | 6-1 | 6-5 | 8-1 |
|---|---|---|---|---|---|---|
| Braid Hills (U1) | n/a | n/a | 3 | 2 | 5 | 3 |
| Colinton Ruderal (U2) | n/a | n/a | 2 | monom. | 2 | 2 |
| Colinton Woods (U2a) | 2 | 2 | 2 | monom. | 4 | monom. |
| Gorebridge (U4) | 2 | 3 | n/a | 6 | n/a | n/a |
| Ben Lawers (R1) | >2 peaks | >2 peaks | 6 | >2 peaks | >2 peaks | >2 peaks |
| Moorfoot Hills (R2) | >2 peaks | >2 peaks | 5 | >2 peaks | >2 peaks | >2 peaks |

**Table 3.3.** Number of alleles per locus in sampled populations. n/a: population not screened for this locus, monom.: monomorphic loci.

Despite the reported transferability of the selected primers to *G. rivale* (Arens *et al.*, 2004) only locus 2-48 could be used as all other loci produced ambiguous multiple peaks in the tested *G. rivale* samples (Table 3.3).

Microsatellite PCRs were performed in 10µl volumes using the following protocol: 5.4 µl purified water, 1 µl 10x buffer (Bioline), 2 mM $MgCl_2$, 0.2 mM dNTPs, 0.175 mg/ml bovine serum albumin (BSA), 0.25 µM forward primer, 0.25 µM reverse primer, 0.025 U Bio*Taq* (Bioline) and 1.5 µl DNA (c. 2ng/µl). The mixture was then cycled through the following profile: 2 min at 95°C, 30 cycles of 30 s denaturing at 94°C, 1.5 min annealing at 53°C or 60°C (according to primer annealing temperature, see Table 3.2), 1.5 min extension at 72°C, ending with 10 min at 72°C to complete extension and subsequent storage at 4°C. All PCR reactions were performed on an MJ Research Dyad thermal cycler and run on an ABI 3730 automated sequencer following the manufacturer's instructions. LIZ-500 labelled internal size standard (PE Applied Biosystems) was added to each sample to size microsatellites. The data generated were analysed by the software GENEMAPPER 3.7 (PE Applied Biosystems) and electropherograms were scored by eye.

## 3.2.4 Data analysis

*Auto-pollination and SI*

The self-incompatibility index (SCI) was calculated for either species as the average number of selfed seeds per attempted cross divided by the average number of outcrossed seeds per attempted cross following Zapata & Arroyo (1978). SCI was determined for every maternal individual and then averaged over all individuals. SCI can have values from 0 to 1 with completely self-compatible species scoring 1 and completely self-incompatible species 0. Stephenson *et al*. (2000) define plants with an SCI of less than 0.15 as strong SI plants while plants with an SCI of more than 0.40 are regarded as weak SI plants.

*Genetic structure of populations*

The average number of alleles per locus, the inbreeding coefficient $F_{is}$ and the expected heterozygosity $H_e$ of maternal parents and offspring were calculated using the program GDA (Lewis & Zaykin, 1999). 95% confidence intervals for $F_{is}$ are based on bootstrapping over all loci (10,000 replicates).

*Indirect estimates of breeding system*

A method to estimate levels of selfing (s) and outcrossing (t=1-s) in a population without prior knowledge of offspring genotypes is via the inbreeding coefficient $F_{is}$. Under Hardy-Weinberg conditions (complete random mating) the expected frequency of heterozygotes $H_e$

= 2pq. If selfing rates are not zero (mixed mating) then a lower proportion of heterozygotes due to selfing is observed occurring at the frequency $H_o = 2pq(1-F_{is})$. The differences in the number of heterozygotes observed and expected within a population can therefore be used to calculate levels of inbreeding such that $F_{is} = 1 - (H_o/H_e)$. If a state of equilibrium is reached in a mixed mating population (i. e. if the frequencies of selfing and outcrossing do not change over time and if there is no differential selection against homozygotes and heterozygotes) then $F_{is} = s/(2-s)$ (Spiess, 1977: 247).

### *Direct estimates of breeding system*

To estimate single ($t_s$) and multilocus ($t_m$) outcrossing rates, populations were analysed separately, under the mixed mating model of Ritland & Jain (1981) and Ritland (1989), implemented by the software MLTR (Ritland, 2002). This program estimates single and multilocus outcrossing rates using maximum likelihood methods by fitting the observed proportions of genotypes descended from a known maternal genotype to the proportions expected under the mixed mating model. The mixed mating model assumes that (1) each mating represents a random event of outcrossing or selfing, with probabilities equal to t and (1 - t), respectively; (2) no mutation and selection following fertilization occurs; (3) there is no assortative (non random) mating (i. e. the probability of outcrossing is independent of the maternal or paternal genotypes) and (4) no variability in pollen pool frequencies occurs (Ritland & Jain, 1981).

To determine the degree of biparental inbreeding (mating with close relatives) $t_m$ and $t_s$ values were compared. In the absence of biparental inbreeding $t_m = t_s$, whereas in the presence of biparental inbreeding $t_s < t_m$ because outcrossing events that are not detected at a single locus have a higher probability of being detected as more loci are examined (Ritland, 1990). The standard errors for these parameters were calculated from 1000 bootstraps without resampling among families. To determine whether the values calculated were significantly different from zero or one, the mean ± 1.96 x SE (standard error) was used. MLTR was also used to estimate the pollen and ovule allele frequencies for each locus.

### *Genetic differentiation*

Genetic differentiation between populations ($F_{st}$) was established using the software program Fstat 2.9.3 (Goudet, 2002).

# 3.3 Results

## 3.3.1 Auto-pollination and SI

*Auto-pollination*

There was a highly significant difference ($\chi^2$=40.8, df=1, p<0.001) in seed set of bagged flowers between *G. urbanum* and *G. rivale*: 3% (2 out of 66) of the bagged *G. rivale* flowers in populations R1 and R2 and 100% (50 out of 50) of the bagged *G. urbanum* flowers using populations U1, U2a and U3 produced spontaneously selfed seeds. Successful seed set in *G. rivale* was only observed in one individual from R2 (Moorfoot Hills, two flowers bagged). Although another individual produced a small number of seeds by auto-pollination, it was not counted as successful since very few seeds ripened. All other bagged *G. rivale* flowers were aborted (Figure 3.3).

**Figure 3.3.** Aborted *G. rivale* flower

*SI analysis*

Selfed flowers produced full seed heads in 100% of the crosses in both populations whereas outcrossing produced seed heads in 98.8% and 98.4% of the crosses in U1 and U2a, respectively but the difference in success rate was not significant (U1: $\chi^2$=0.247, df=1,

p=0.62 and U2a: $\chi^2$=0.024, df=1, p=0.88). The self-incompatibility index for *G. urbanum* was SCI=1, based on 22 individuals averaging over 43 selfed (22 different maternal genotypes) and 79 outcrossed events from U1 and U2a.

R1 and R2 showed a highly significant difference in success of seed set between controlled selfed and outcrossed flowers. In R1 only 10.4% of the selfed but 52.2% of the outcrossed flowers developed a full set head ($\chi^2$=13.3, df=1, p<0.001) whereas in R2 19.2% and 51.2% of attempted crosses were successful, respectively ($\chi^2$=11.2, df=1, p<0.001). The self-incompatibility index for *G. rivale* was SCI=0.33, based on 18 individuals averaging over 18 selfed (18 different maternal genotypes) and 35 outcrossed events from R1 and R2. The average population SCI was 0.24 for R1 and 0.44 for R2 but the values were not significantly different from each other (Mann-Whitney test, p=0.27). Out of 18 maternal parents 39% never set any seeds despite repeated attempts of selfing (SCI=0), 11% had a SCI=1 and 50% had intermediate SCI, i. e. 0<SCI<1.

## 3.3.2 Microsatellites

*Geum urbanum*

Genetic structure and indirect estimates of breeding system

Four pure Scottish *G. urbanum* adult populations (20 to 36 adult samples per population) were screened for three to four microsatellite loci (Table 3.4). Average number of alleles per locus equalled 2 to 3.7 and average expected heterozygosity ranged from 0.32 to 0.56. $F_{is}$ values for U1, U2a and U4 were similar with 0.72, 0.67 and 0.63, respectively (significantly different from zero and 1, p<0.05) in contrast to the markedly higher value of 0.95 (significantly different from zero but not from 1, p<0.05) in U2 (Table 3.4). Under equilibrium conditions these $F_{is}$ values correspond to selfing rates of 84%, 80% and 77% for U1, U2a and U4, respectively and 97% for U2.

Progeny arrays from 20 families were collected additionally to the adult samples in U1 and U2 (98 and 95 progeny, respectively) and screened for the same loci as the adult population (Table 3.4). Progeny in U2 had the same number of alleles per locus as the adults (2) and very similar expected heterozygosity (0.31) whereas in U1 progeny showed a higher number of alleles per locus than the adults (3.3 and 3, respectively) and lower expected heterozygosity (0.30 and 0.35, respectively). The offspring $F_{is}$ values of U1 and U2 were 0.67 (s=80%) and 0.93 (s=96%), respectively and both significantly different from zero and one (p<0.05). In both populations adult and offspring $F_{is}$ were not significantly different (p<0.05) although offspring $F_{is}$ was slightly lower than adult $F_{is}$ (Table 3.4).

| Population | | N | #loc | A | $H_e$ | $F_{is}$ | 95%($F_{is}$) | inferred s |
|---|---|---|---|---|---|---|---|---|
| **U1 (Braid Hills)** | | | | | | | | |
| | Adult | 31 | 4 | 3 | 0.35 | 0.72* | 0.67≤ $F_{is}$ ≤0.83 | 84% |
| | Offspring | 96.3 | 4 | 3.3 | 0.30 | 0.68* | 0.44≤ $F_{is}$ ≤0.77 | 81% |
| **U2 (Colinton Ruderal)** | Adult | 20 | 3 | 2 | 0.32 | 0.95† | 0.86≤ $F_{is}$ ≤1.00 | 97% |
| | Offspring | 94 | 3 | 2 | 0.31 | 0.93* | 0.88≤ $F_{is}$ ≤0.97 | 96% |
| **U2a (Colinton Woods)** | | 29 | 4 | 2.5 | 0.45 | 0.67* | 0.41≤ $F_{is}$ ≤0.88 | 80% |
| **U4 (Gorebridge)** | | 32.3 | 3 | 3.7 | 0.56 | 0.63* | 0.44≤ $F_{is}$ ≤0.84 | 77% |

**Table 3.4.** Population genetic parameters for pure Scottish *G. urbanum* populations **N**: average number of individuals sampled over all loci, **#loc**: number of microsatellite loci screened, **A**: average number of alleles per locus, $\mathbf{H_e}$: average expected heterozygosity over all loci, $\mathbf{F_{is}}$: inbreeding coefficient, **95%($\mathbf{F_{is}}$)**: 95% confidence interval for inbreeding coefficient, **s**: inferred selfing rate under equilibrium condition. *Significantly > 0% ($p<0.05$), † significantly different from 0% but not from 100% ($p<0.05$)

## Direct outcrossing estimates

*G. urbanum* populations U1 (20 families, 98 progeny) and U2 (20 families, 94 progeny) were screened for four and three (locus 6-1 monomorphic in U2) microsatellite loci, respectively (Table 3.3). Two U1 offspring individuals from one maternal mother showed impossible genotypes (no maternal allele present) and were excluded from the analysis, all other progenies displayed at least one of the maternal alleles, confirming Mendelian inheritance.

Outcrossing rates varied considerably among the two populations (Table 3.5). U1 showed a multilocus outcrossing rate of 17.7% (significantly different from zero, $p<0.05$) whereas the multilocus outcrossing rate for U2 was much lower (5.8%) and not significantly different from 0% ($p<0.05$). There was no evidence for biparental inbreeding (Table 3.5).

| Population | #fam | # pro | # loci | $t_m$ (SE) | $t_s$ (SE) | $t_m$-$t_s$ (SE) |
|---|---|---|---|---|---|---|
| U1 | 20 | 98 | 4 | 17.7* (±4.6) | 16.8* (±4.6) | 0.8 (±1.9) |
| U2 | 20 | 95 | 3 | 5.8 (±4.8) | 4.6 (±3.7) | 1.2 (±1.7) |

**Table 3.5.** Mating system parameters for two pure Scottish populations of *G. urbanum* (U1: Braid Hills and U2: Colinton) **#fam**: number of families per population, **# pro**: number of progeny per population, **# loci**: number of loci used, $\mathbf{t_m}$: multilocus outcrossing rate in %, $\mathbf{t_s}$: single locus outcrossing rate in %, $\mathbf{t_m}$-$\mathbf{t_s}$: biparental inbreeding rate in %. *Significantly > 0% ($p<0.05$).

## Genetic differentiation

Only U1, U2 and U2a were used to establish genetic differentiation between populations as there was only one common locus available for U4 (Gorebridge). Pairwise $F_{st}$ values for U1, U2 and U2a ranged from 0.21 to 0.57 (Table 3.6). Differentiation was highest between populations U2 and U2a (0.57) and lowest between U1 and U2 (0.21).

| **Population** | Colinton Ruderal U2 | Colinton Woods U2a |
|---|---|---|
| Braid Hills U1 | 0.21* | 0.50* |
| Colinton Ruderal U2 | | 0.57* |

**Table 3.6.** Pairwise $F_{st}$ values for three *G. urbanum* populations based on four microsatellite loci.*Significantly > 0% ($p<0.05$).

## *Geum rivale*

### Genetic structure and indirect estimates of breeding system

Two pure Scottish *G. rivale* populations R1 (Ben Lawers, 30 samples) and R2 (Moorfoot Hills, 36 samples) were screened for one locus only (2-48, see Chapter 3.2.3). In R1 and R2 the number of alleles equalled 6 and 5, respectively and expected heterozygosity was 0.72 and 0.55, respectively (Table 3.7). Both populations had the same $F_{is}$ value (-0.07) corresponding to a selfing rate of 0% under equilibrium conditions.

Progeny arrays from 20 families were collected additionally to the adult samples in R1 and R2 (95 and 86 progeny, respectively) and screened for the same locus as the adult population. Progeny in R1 had the same number of alleles per locus as the adults (6) and very similar expected heterozygosity (0.71) whereas in R2 progeny showed a lower number of alleles per locus than the adults (4) and had the same expected heterozygosity (0.55). The R1 offspring $F_{is}$ value of -0.02 (s=0%) was very similar to the parental $F_{is}$ of -0.07 (s=0%) whereas in R2 the offspring $F_{is}$ of 0.13 (s=77%) was higher than the adult $F_{is}$ (-0.07, s=0%).

| **Population** | | **N** | **#loc** | **A** | **$H_e$** | **$F_{is}$** | **inferred s** |
|---|---|---|---|---|---|---|---|
| **R1 (Ben Lawers)** | Adult | 30 | 1 | 6 | 0.72 | -0.07 | 0% |
| | Offspring | 95 | 1 | 6 | 0.71 | -0.02 | 0% |
| **R2 (Moorfoot Hills)** | Adult | 36 | 1 | 5 | 0.55 | -0.07 | 0% |
| | Offspring | 86 | 1 | 4 | 0.55 | 0.13 | 23% |

**Table 3.7.** Population genetic parameters for pure Scottish *G. rivale* populations **N**: average number of individuals sampled over all loci, **#loc**: number of microsatellite loci screened, **A**: number of alleles per locus, **$H_e$**: expected heterozygosity, **$F_{is}$**: inbreeding coefficient, **s**: inferred selfing rate under equilibrium condition.

### Direct outcrossing estimates

*G. rivale* populations R1 (20 families, 95 samples) and R2 (20 families, 86 samples) were screened for one locus (2-48, see Chapter 3.2.3). The single locus outcrossing rate was 90.5% for R1 and 65.6% for R2. Both were significantly different from 0% but not from 100% ($p<0.05$, Table 3.8).

| Population | #fam | # pro | # loci | $t_m$ (SE) | $t_s$ (SE) | $t_m$-$t_s$ (SE) |
|---|---|---|---|---|---|---|
| R1 | 20 | 95 | 1 | n/a | 90.5† (±12.1) | n/a |
| R2 | 20 | 86 | 1 | n/a | 65.6† (±21.1) | n/a |

**Table 3.8.** Mating system parameters for two pure Scottish populations of *G. rivale* (R1: Ben Lawers and R2: Moorfoot Hills). **#fam**: number of families per population, **# pro**: number of progeny per population, **# loci**: number of loci used, $t_m$: multilocus outcrossing rate in %, $t_s$: single locus outcrossing rate in %, $t_m$-$t_s$: biparental inbreeding rate in %. † significantly different from 0% but not from 100% ($p<0.05$).

## Genetic differentiation

Genetic differentiation between the two *G. rivale* populations R1 and R2 based on a single locus was $F_{st} = 0.04$.

# 3.4 Discussion

## 3.4.1 Auto-pollination and SI

*Geum urbanum*

Results showed that *G. urbanum* is a fully self compatible species (SCI=1) and is 100% successful in producing seeds by auto-pollination. Out of 50 bagged flowers 50 produced full seed heads. This confirms earlier morphological observations reported by Knuth (1898) that despite weak protogyny *G. urbanum* is likely to be mainly selfing due to the close proximity of anthers and stigmas and the lack of frequent visits by pollinators.

*Geum rivale*

Only two out of 66 bagged flowers (3%) produced seeds in *G. rivale* both of which were from the same individual. There are several possible explanations for this low seed set. Firstly, in order for self pollination to occur *G. rivale* might need the assistance of a pollinator to transfer self pollen onto the stigma. Morphological observations by Knuth (1898) suggest that only in the later stages of anthesis male and female reproductive organs are close enough for self pollination to occur spontaneously. Anther stigma distance might vary between individuals which could make self-pollination more difficult without the help of a pollinator. There is evidence for insect mediated selfing from other species such as *Phyla incisa* (Estes & Brown, 1973) and *Calyptridium monospermum* (Hinton, 1976) where insect visits are necessary for successful self-pollination. It is also possible that wind might partly contribute to self-pollination as movement of the drooping flowers could release pollen from ripe anthers depositing self-pollen on receptive stigmas below the anthers (Langlotz, 2007). However, the SI analysis in R1 and R2 using controlled selfing and outcrossing experiments revealed that full seed set of selfed flowers was generally very low (10% and 20%, respectively) despite manual saturation of stigmas with pollen. This suggests that a lack of insect mediated pollen is unlikely to be the main reason.

Secondly, the low seed set of bagged flowers could be due to the abortion of selfed seeds after fertilization by very early acting inbreeding depression. However, this seems unlikely as some of the same maternal parents which aborted the bagged flowers produced seed heads when selfed manually. Other pollination experiments found that after manual self-pollination only very few pollen tubes reached the ovary (fluorescence microscopy) and fertilized the ovules (Erbar, pers. comm., 2008).

Thirdly, a self-incompatibility system could be in place which prevents self pollen from germinating or slows down pollen tube growth of self pollen. It was generally more difficult to pollinate *G. rivale* than *G. urbanum* flowers independent from crosstype (selfing or outcrossing) as even outcrossing had on average only a 50% success rate compared to more than 98% in *G. urbanum*. However, there was a significant difference in the success rate between successful selfing and outcrossing in *G. rivale*. In R1 about 10% and in R2 about 20% of selfing attempts were successful compared to about 50% for outcrossing in either population. Results from the SI experiment (SCI=0.30) suggest that according to Stephenson *et al.* (2000) *G. rivale* can be classified as intermediate or moderate self-incompatible. Large variation in the expression of SI was detected between different genotypes which has also been observed in many other plant taxa (Levin, 1996; Vogler *et al.*, 1998; Tsukamoto *et al.*, 1999; Stephenson *et al.*, 2000; Stone *et al.*, 2006; Mena-Ali & Stephenson, 2007) . Nearly 40% of the selfed individuals did not produce any seeds despite repeated attempts of selfing, only 6% were fully compatible yielding a full seed head for every selfing attempt and 54% of the genotypes had an intermediate SCI.

This suggests that *G. rivale* has a leaky self-incompatibility system which depending on genotype and environmental factors does not necessarily prevent selfed flowers from producing seeds. If flowers are bagged then the combination of a self-incompatibility system and the lack of a pollen vector depositing larger amounts of self pollen on the stigma might prevent seed set in all but a few cases. Individual differences in stamen-anther distance and/or protogyny could also be a contributing factor in terms of successful auto-pollination (Schoen, 1982; Karron *et al.*, 1997). However, with the data at hand it is not possible to decide what the exact causes for the observed pattern are but it is clear that *G. rivale* has very limited abilities to auto-pollinate.

## 3.4.2 Outcrossing rates

Outcrossing rates varied between 6% and 18% in two *G. urbanum* and between 66% and 91% in two *G. rivale* populations. The two approaches used to calculate the outcrossing rates (software MLTR and $F_{is}$) yielded similar results which did not differ significantly from each other ($p<0.05$). This suggests that the analysed populations are in equilibrium and have more or less constant outcrossing and selfing rates every generation.

Out of the four *G. urbanum* populations analysed U1, U2a and U4 have similar outcrossing rates of c. 20% whereas U2 shows a markedly lower rate of 6%. Despite the varying outcrossing rates among populations the figures are a clear indication of the highly selfing nature of *G. urbanum* where at least 80% of the progeny are the product of self fertilization.

It is noteworthy that the three populations with outcrossing rates of c. 20% were collected in woodlands whereas U2 (c. 6%) was collected from a ruderal area along a paved path, although very close (c. 50 m) to the semi-natural forested area of population U2a. *G. urbanum* is essentially a plant of deciduous woodlands (Taylor, 1997b) and it could be that the higher selfing rate of population U2 is related to a lack of suitable pollinators outside the preferred habitat of this species.

Inferring outcrossing rates from average $F_{is}$ values of *G. urbanum* populations reported in previous studies yielded similar figures to that found in U2. Arens *et al.* (2004) developed ten microsatellite loci for *G. urbanum* and surveyed 688 individuals from 45 European populations (15 samples on average per population) reporting an overall $F_{is}$ value of 0.93 (t=4%). Vandepitte *et al.* (2007) analysed 18 *G. urbanum* populations from woodlands in Belgium (20 to 23 samples per population) using six microsatellite loci yielding an overall $F_{is}$ value of 0.89 (t=6%). No $F_{is}$ figures were given for individual populations by Arens *et al.* (2004) but Vandepitte *et al.* (2007) reported a range from $F_{is}$ = 0.79 (t=12%) to 1 (t=100%) for individual populations with the exception of one population ($F_{is}$ = 0.45, t=38%).

This suggests that the outcrossing rates of c. 20% in the woodland populations U1, U2a and U4 are on average higher compared to other European populations. As detailed in the introduction to this chapter outcrossing rates can be influenced by many factors and vary considerably among populations and between years. Whether the higher outcrossing rates of the analysed Scottish woodland populations are habitat related or are due to a combination of chance events promoting outcrossing in a particular year can not be established without further investigations.

Offspring $F_{is}$ values in U1 and U2 (0.67 and 0.93, respectively) are slightly lower than adult $F_{is}$ values (0.72 and 0.95, respectively) but are not significantly different at the 5% level, providing no circumstantial evidence for inbreeding depression (see Chapter 4 for a full discussion). However, the differing $F_{is}$ values indicate that there is a slight variation in selfing rates between generations suggesting less selfing in the year the offspring generation was produced and more selfing in the year the adult generation originated from. This lends support to the hypothesis that the on average higher outcrossing rates in woodland populations might have been at least partly due to beneficial circumstances promoting outcrossing in the year of seed collection. Outcrossing rates based on MLTR estimates in *G. rivale* populations R1 and R2 were 91% and 66%, respectively but did not differ significantly from 100% ($p<0.05$). Despite the fact that only one locus was available for the estimates these figures are a clear indication for the highly outcrossing nature of *G. rivale*. Both adult populations had slightly negative $F_{is}$ values (-0.07, t=100%). Theoretically $F_{is}$ lies between 0

(100% outcrossing) and 1 (0% outcrossing) but negative $F_{is}$ values (more heterozygotes than expected) can result from selection favouring heterozygotes or negative assortive mating, that is, where individuals tend to mate with phenotypically different individuals to produce heterozygous genotypes (Hartl, 1987). However, since $F_{is}$ is very close to zero it is more likely that the slightly negative value is an artefact with no biological meaning. In R2 the offspring $F_{is}$ (0.13, t=77%) was larger than the adult $F_{is}$ (-0.07, t=100%) providing some circumstantial evidence for reduced fitness of selfed offspring (see Chapter 4).

## 3.4.3 Genetic differentiation

Genetic differentiation (based on four loci) between *G. urbanum* populations U1, U2 and U2a was high with $F_{st}$ values ranging from 0.21 to 0.57. U2 (Colinton Ruderal) and U2a (Colinton Woods) are only c. 50 m apart from each other but were the most differentiated populations ($F_{st}$ = 0.57). With the assumption of Wright's island model (Wright, 1931) values of $F_{st}$ can be directly related to the number of migrants per generation and thus gene flow. $F_{st}$ = 0.2 corresponds to one migrant per generation which indicates that the effect of gene flow is greater than the effect of drift. $F_{st}$ values less than 0.2 are taken to indicate high gene flow whereas values greater than 0.2 indicate low gene flow (Lowe *et al.*, 2004). $F_{st}$ = 0.57 corresponds to 0.2 migrants every generation or one migrant per five generations increasing population differentiation over time as the number of migrants is too low to offset divergence due to genetic drift.

However, the use of $F_{st}$ to infer gene flow has been heavily criticized in terms of the assumptions of the model and the inability to differentiate between contemporary and historical gene flow (Whitlock & McCauley, 1999; Lowe *et al.*, 2004). It is unrealistic to assume a non-geographical structure which implies that migration is equally likely between all populations and that all populations are of equal size. Further assumptions are that no mutations occur and that there is equilibrium between migration and genetic drift.

Despite all the limitations of the Island model it is clear that the high $F_{st}$ value indicates very low to no gene flow at all between U2 and U2a. Nearly 50% of the alleles present in U2a are missing from U2 and the allele of the locus for which U2 is monomorphic is absent in U2a. There is also high genetic differentiation between U2a and U1 (3 km apart, $F_{st}$ = 0.5). Genetic differentiation between U1 and U2 is less pronounced (3 km apart, $F_{st}$ = 0.21) but still large enough to suggest only low levels of gene flow. This contrasts somewhat with the $F_{st}$ value of 0.14 based on AFLP markers (Chapter 2.3.2) which was not significantly different from zero because of a large standard error. Taking various lines of evidence into account such as the higher microsatellite $F_{st}$ value (based on four loci and samples sizes of over 100 individuals

per population) and the clear genetic differentiation of U1 and U2 in a PCO plot (Chapter 2.3.2) it seems likely that the lower AFLP $F_{st}$ value is an artefact due to the smaller sample size of 10 individuals per population.

The results of the *G. urbanum* mating system analysis and the levels of genetic differentiation between the *G. urbanum* populations based on the fixation index $F_{st}$ corresponds well with the highly selfing nature of this species as genetic differentiation is generally high between populations in selfing plant taxa (Hamrick & Godt, 1996). Selfing effectively restricts gene flow and exchange of mutations between populations thus leading to high genetic differentiation over time.

In contrast, the two *G. rivale* populations R1 and R2 although more than 100 km apart have a much lower $F_{st}$ value (0.04, the same value was obtained using AFLP data) suggesting very little genetic differentiation. This is in agreement with the highly outcrossing mating system of this species as outcrossing taxa generally display low genetic variation between populations (Hamrick & Godt, 1996). Alleles and mutations can be passed on between even remote populations through an interconnecting network of 'stepping stone' populations.

## 3.4.4 Summary

Despite variation in outcrossing rates between populations of *G. urbanum* and *G. rivale,* the former species is highly selfing with only 6% to 23% of its progeny outcrossed in contrast to latter where 66% to 100% of the progeny are outcrossed. Auto-pollination as well as controlled selfing and outcrossing experiments suggest that *G. urbanum* is a completely self-compatible species (SCI=1) whereas *G. rivale* is a moderately self-incompatible species (SCI=0.33) with a high variability in individual genotypic expression of SI. High genetic differentiation based on $F_{st}$ values between *G. urbanum* populations and low genetic differentiation between *G. rivale* populations are in agreement with the predominant breeding system of either species.

# 4. INBREEDING DEPRESSION IN *G. URBANUM* AND *G. RIVALE*

## Abstract

Inbreeding depression is a potentially important factor in establishing the long term evolutionary consequences of hybridization as the genetic load carried by the parental species could be a major barrier for the establishment of new hybrid lineages. The main aim of this chapter is to establish levels of inbreeding depression in pure reference populations of two hybridizing species, the selfing *G. urbanum* and the outcrossing *G. rivale.* Two pure Scottish populations per species were identified and ten individuals per population were transferred to an unheated greenhouse. Controlled crosses were performed within each population to establish whether fitness differences (seed set and mass, germination and survival rate and growth vigour) between outcrossed and selfed progeny could be detected. No inbreeding depression ($\delta_t$=0) could be detected in the two sampled selfing *G. urbanum* populations and in one of the two outcrossing *G. rivale* populations. Low levels of inbreeding depression were observed in the other *G. rivale* population ($\delta_t$=0.10, survival rate). Results for the selfing *G. urbanum* populations are in line with theoretical considerations about purging the genetic load and breeding system evolution. Factors which might have led to the unexpectedly low levels of inbreeding depression in *G. rivale* populations are discussed.

# 4.1 Introduction

## 4.1.1 Genetic hypotheses of inbreeding depression

Deleterious mutations, i. e. the generation of mutations which are harmful or lethal to health and their dominance level are very important factors in determining the fitness of organisms. It has been estimated that there is one deleterious mutation on average per diploid genome and generation in plants (Johnston & Schoen, 1995; Charlesworth & Silvertown, 2001). Most of these mutations are recessive and do not affect an individual's fitness adversely in heterozygous form because the wild-type allele, the original non mutated allele, is dominant and suppresses the effects of the mutant allele (Charlesworth & Silvertown, 2001).

A major consequence of selfing, the most extreme form of inbreeding, is that it leads to increased homozygosity. If the heterozygous loci are unlinked, an individual heterozygous for 100 loci will be 98% homozygous for these loci if 12 generations of complete selfing occur (Richards, 1997). If loci heterozygous for deleterious alleles in the parental generation are rendered homozygous through inbreeding in the offspring generation then the deleterious allele will be expressed resulting in lower fitness of these individuals. This decline in fitness with inbreeding is known as inbreeding depression δ and has attracted the attention of evolutionary biologists since Darwin.

Inbreeding depression can also be caused by heterozygote advantage, i. e. if heterozygous individuals in a population are fitter than either homozygote at a viability locus. Selfing will increase the frequency of homozygotes every generation thus producing offspring with lower fitness compared to heterozygotes which is then observed as inbreeding depression. These two competing models for the genetic basis of inbreeding depression are referred to as partial dominance and overdominance hypothesis, respectively (Charlesworth & Charlesworth, 1987). In both models inbreeding depression will prevent the spread of alleles promoting selfing if $\delta \geq 50\%$ (see also Chapter 3.1, p 48). Although the two models result in low fitness in the next generation, they differ greatly in the prediction of the evolutionary consequences for populations.

According to the partial dominance hypothesis where inbreeding depression is due to a mutational load of deleterious genes, theory predicts that autogamous (selfing) plants will exhibit low inbreeding depression. Habitual selfing should purge deleterious recessive genes from the population over time as individuals with lower fitness due to disadvantageous mutations will be selected against (Schemske & Lande, 1985; Charlesworth & Charlesworth, 1987). In contrast, when inbreeding depression is due to a segregational load of less fit

homozygotes (overdominance hypothesis), theory predicts that inbreeding depression will increase over time because of the loss of heterozygote advantage at many loci (Charlesworth & Charlesworth, 1990).

A major difficulty in distinguishing between the two mechanisms is that so-called 'pseudo-overdominance' can produce an effect which appears to be overdominance at a single locus (Mangelsdorf, 1952; Stuber *et al.*, 1992). Pseudo-overdominance is caused by two closely linked loci 1 and 2 which segregate their deleterious alleles 'a' and 'b', respectively in repulsion phase meaning that 'a' and 'b' are borne on opposite homologous chromosomes. The dominant allele (A or B) on one homologue always segregates with the deleterious allele (b or a, respectively) on the other (Figure 4.1). Individuals heterozygous for this linkage group (Ab/aB) will appear to have higher fitness because dominant alleles A and B occur at both loci whereas homozygous individuals (Ab/Ab or aB/aB) will always express the deleterious recessive alleles at locus 2 or locus 1, respectively and be less fit. This creates a situation where heterozygotes appear to be fitter than homozygotes because of heterozygote advantage (overdominance hypothesis) whereas in reality the lower fitness of homozygotes is caused by the expression of two recessive alleles (dominance) and not because of an intrinsic heterozygote advantage.

| | Chromosome C | Chromosome C' |
|---|---|---|
| Locus 1 | A | a |
| Locus 2 | b | B |

**Figure 4.1.** Pseudo-overdominance. Locus 1 and 2 represent one linkage group segregating deleterious alleles a and b in repulsion, i. e. they are borne on opposite homologous chromosomes C and C' so that the dominant allele of locus 1 is always segregating with the deleterious allele of locus 2 and vice versa (Ab and aB, respectively). Further explanations see text.

There are also two additional factors which can affect the evolution of inbreeding depression and mating systems in a population (Carr & Dudash, 2003). Firstly, epistasis, the interaction among loci, can influence how fitness levels change with increasing homozygosity. In the case of no epistasis, fitness declines log-linearly as a population becomes more inbred. In the case of reinforcing epistasis the individual deleterious effect of homozygosity at a given locus becomes greater with increasing homozygosity of the genome resulting in lower fitness at the same inbreeding coefficient compared to the 'no epistasis' scenario. Alternatively, with diminishing epistasis individual effects become weaker with increasing genomic homozygosity leading to higher fitness compared to no epistasis. If either type of epistasis exists, the effect of any mutation is not independent of its genetic background (Carr & Dudash, 2003).

Secondly, the number of loci and the magnitude of their effects will have an impact on the evolution of inbreeding depression in a population. Mutations of small effect are not easily removed by selection and might even become fixed by drift in selfing populations (Barrett & Charlesworth, 1991; Willis, 1999). If this occurs, crosses between inbred lines fixed for deleterious alleles will show greater fitness than the parental inbred lines.

The dominance theory has received support from a review comparing levels of inbreeding depression found in 79 populations of 54 plant species with different mating systems (Husband & Schemske, 1996). The average inbreeding depression in mainly selfing species ($\delta$=0.23) was significantly less (43%) than in mainly outcrossing species ($\delta$=0.53), suggesting that purging occurs in selfing populations. However, there are also ambiguous results from studies which did not detect lower levels of inbreeding depression in selfing populations (Ågren & Schemske, 1993; Dole & Ritland, 1993; Rathcke & Real, 1993; Charlesworth *et al.*, 1994; Latta & Ritland, 1994; Johnston & Schoen, 1996; Michalski & Durka, 2007; Ishida, 2008).

There is also evidence of inbreeding depression in highly selfing plant taxa such as *Leavenworthia uniflora* (Charlesworth *et al.*, 1994), *Mimulus nasutus* (Latta & Ritland, 1994) and *Begonia semiovata* and *B. hirsuta* (Ågren & Schemske, 1993). In both *Begonia* species, the seed production of selfed flowers was on average 12% lower than that of outcrossed flowers. There was no significant effect of crosstype on germination rate or survival, but selfed offspring had a lower dry mass than outcrossed offspring. Both species showed significant inbreeding depression for total fitness (estimated as the product of seed production, germination rate, survival and dry mass at 18 weeks) equalling 22% in *B. hirsuta*, and 43% in *B. semiovata* (Ågren & Schemske, 1993). These findings are consistent with the hypothesis (Charlesworth *et al.*, 1990a; Charlesworth *et al.*, 1991) and findings of other studies (Willis, 1999; Carr & Dudash, 2003) that a high mutation rate to mildly deleterious alleles with intermediate degree of dominance can contribute to the maintenance of inbreeding depression in selfing populations.

A review by Byers & Waller (1999) on inbreeding depression and purging comparing 52 studies found that there was no conclusive evidence for purging. The results of their study showed that purging did not seem to act as a consistent and major force in natural plant populations, thus questioning the predicted powerful force of purging by theoretical models (Byers & Waller, 1999). The authors come to the conclusion that under some circumstances inbreeding will tend to purge populations of enough deleterious recessive mutations to reduce inbreeding depression but that it is impossible to make universal predictions (Byers & Waller, 1999). Selfing populations can harbour considerable inbreeding depression and Goodwillie et al. (2005) highlight that the disparity between the studies conducted by Husband &

Schemske (1996) and Byers & Waller (1999) could be a reflection of different time scales. Purging might be easier to detect in species that have been selfing for a long time than in recently diverged and closely related species or populations.

A later review by Crnokrak and Barrett (2002) which only included studies where populations had been serially inbred for at least two generations, found that 72% of the traits showed a decrease in inbreeding depression in successive generations with a moderate trait rebound in inbred lines. 53% of the reported traits revealed evidence of complete purging with inbreeding and 47% revealed no purging. Although the study by Crnokrak and Barrett (2002) provided good support for purging the authors pointed out that the frequency and magnitude of purging with inbreeding appeared to depend upon the manner in which trait fitness components were experimentally assessed and that little correspondence between measures of purging were found. They also suggest that adaptations to laboratory conditions may have accounted for rebounds in fitness and caution against general conclusions about purging in plant and animal populations (Crnokrak & Barrett, 2002).

Studies using molecular markers have frequently found support for the overdominance hypothesis but none could exclude pseudo-overdominance conclusively (Stuber *et al.*, 1992; Mitchell-Olds, 1995; Yu *et al.*, 1997; Li *et al.*, 2001). Carr and Dudash (2003) pointed out that all of these studies involved crosses between highly divergent lines suggesting that they might not be relevant in answering the question which type of genetic load (mutational or segregational) is maintained within populations. In contrast, molecular studies on pines (Remington & O'Malley, 2000b, a) and *Arabis petrea* (Kärkkäinen *et al.*, 1999) seem to support the dominance hypothesis but some loci which acted in an overdominant manner were also detected although pseudo-overdominance could not be ruled out.

There is still no conclusive evidence as to which hypothesis accounting for the genetic basis of inbreeding depression might be the correct one. This is partly due to the difficulty of distinguishing overdominance from pseudo-overdominance. However, there seems to be growing consensus in favour for the dominance model in the most recent reviews on this topic (Charlesworth & Charlesworth, 1999; Carr & Dudash, 2003)

## 4.1.2 Measuring inbreeding depression

To detect inbreeding depression in plant populations hand pollinations are usually performed to generate offspring by selfing and outcrossing. Fitness parameters at early and later life stages are then collected and compared between crosstypes and inbreeding depression δ is calculated as $1 - (w_s/w_x)$ where $w_s$ is the fitness of inbred progeny and $w_x$ the fitness of

outcrossed progeny. Early life stage parameters include number of seeds produced, seed mass and germination rate. Later life stage parameters are survival rate of seedlings, above ground biomass and number of flowers and seeds produced. Assuming that the traits expressed at different life stages are independent of one another, the overall effect of inbreeding on fitness can be calculated in a multiplicative fashion so that δ is 1 minus the product of all individual fitness parameters (Charlesworth & Silvertown, 2001).

Circumstantial evidence for inbreeding depression can also be gained by comparing inbreeding coefficient ($F_{is}$) values using mother-progeny arrays and assuming that populations are at inbreeding equilibrium (same amount of inbreeding every generation). If the $F_{is}$ value of the adult population is lower than the $F_{is}$ value of the offspring generation then there is evidence for inbreeding depression. The adult population consists only of fit individuals which were able to survive to adulthood whereas less fit individuals from the offspring generation will be selected against and die before reaching maturity. A lower adult $F_{is}$ compared to offspring $F_{is}$ is therefore a sign that inbred progeny have a fitness disadvantage compared to outcrossed progeny and have not survived to contribute to the adult generation.

The level of inbreeding depression can vary over the different life stages and may be expressed at any stage (Dudash, 1990; Goodwillie & Knight, 2006). In their review of 54 plant species Husband & Schemske (1996) reported that most selfers expressed the majority of inbreeding depression late in their life cycles at the growth and reproduction stage, whereas many outcrossing species exhibited substantial inbreeding depression either early or late in their life cycles (Figure 4.2). There was also an increase of inbreeding depression in selfers from early to late life stages compared to outcrossers. These results suggest that most early-acting inbreeding depression is associated with relatively few recessive lethals that can be purged easily through inbreeding. In contrast, inbreeding depression later in the life cycle seems to result from many weakly deleterious mutations, which are more difficult to purge (Husband & Schemske, 1996). However, the relative performance of selfed and outcrossed offspring in early life-history traits may be uncorrelated to the performance in later life-history traits suggesting that a number of loci are likely to make independent contributions to the genetic load affecting these traits (Carr & Dudash 2003).

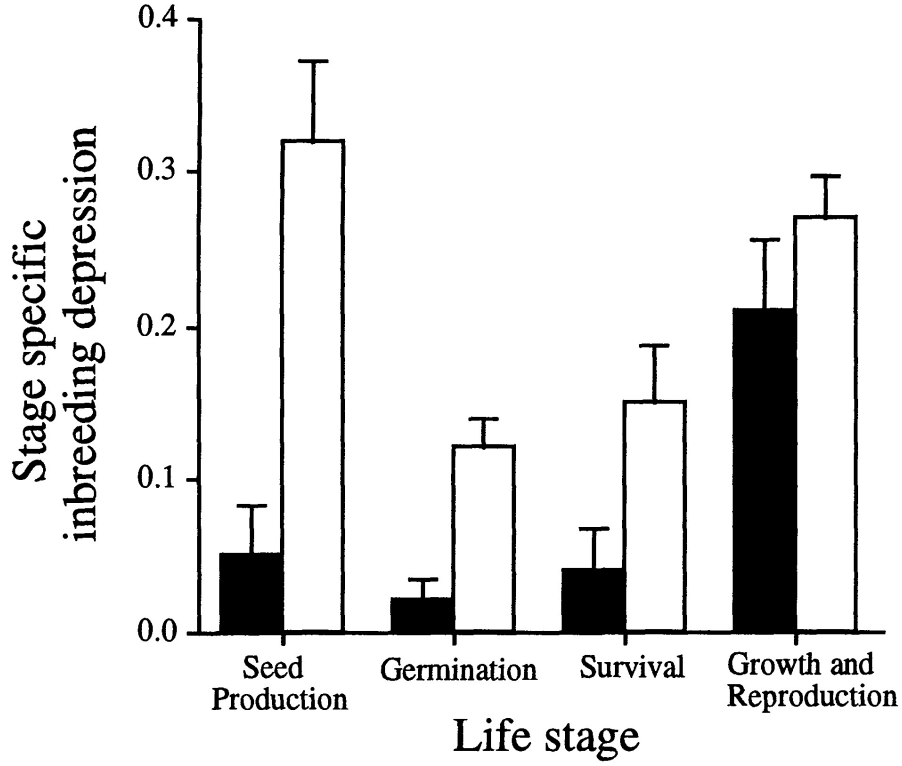



**Figure 4.2.** Mean life stage specific expression of inbreeding depression for selfing (shaded bars) and outcrossing (open bars) populations of angiosperm and gymnosperm species. From Husband and Schemske (1996).

Levels of inbreeding depression among populations can vary as some studies have demonstrated. Jain (1978) reported varying degrees of inbreeding depression using number of flowers per plant, plant height and seed set per flower in seven *Limnanthes alba* populations with differing outcrossing rates (t=0.43-0.97), heterozygosity ($H_e$=0.12-0.27) and percentage of polymorphic loci (P=0.29-0.57). Three populations showed inbreeding depression for all three characters, three populations for one or two traits only and one population did not show any inbreeding depression at all. There was no association between observed patterns of inbreeding response and predictions from the genetic variation parameters (Jain, 1978). A similar result was reported by Latta & Ritland (1994) where two populations of *Mimulus* with intermediate inbreeding coefficients differed greatly in levels of inbreeding depression for all characters measured apart from germination rate and two other populations with $F_{is}$ values close to 0 and 1 did not differ at all in inbreeding depression.

The level of inbreeding depression can also vary according to the environmental conditions of the study population and it is often assumed that inbreeding depression is generally greater in more stressful environments (Frankham *et al.*, 2002). A recent review on this topic using the results of 34 published studies on plants and animals found that in 69% of the cases inbreeding depression was more pronounced in stressful compared to more benign

environments (Armbruster & Reed, 2005). However, in 24% of the studies inbreeding depression actually decreased in stressful environments. One clear pattern which emerged from Armbruster and Reed's (2006) review was that different populations of the same species and different inbred lineages from the same population, often exhibited highly variable responses to inbreeding under stressful conditions. This suggests that in different populations different alleles and loci determine the stress response and that it might be difficult to predict how populations of the same species will respond to inbreeding, particularly under environmentally variable conditions (Hedrick & Kalinowski, 2000; Armbruster & Reed, 2005).

Holsinger (1988) has also pointed out that the fitness of inbred progeny is not independent of the genotype as frequently assumed and many studies have found variation of inbreeding depression at the family level (Ågren & Schemske, 1993; Pico *et al.*, 2004; Kelly, 2005; Oakley & Winn, 2008). Another factor influencing the level of inbreeding depression is polyploidy. Theoretical models predict less inbreeding depression in polyploids than in their diploid relatives (Lande & Schemske, 1985; Ronfort, 1999). A few studies have compared levels of inbreeding depression in natural populations of polyploids and closely related diploids, and these have produced ambiguous results. Lower levels of inbreeding depression were found in autotetraploid *Chamerion angustifolium* (Husband & Schemske, 1997) and allotetraploid *Anthericum liliago* populations (Rosquist, 2001) when compared to diploid populations of the same species or closely related species, respectively. In contrast, Johnston and Schoen (1996) found higher levels of inbreeding depression in allotetraploid than in diploid populations of *Amsinckia* (Boraginaceae), as did Barringer and Geber (2008) using allotetraploid and diploid species of *Clarkia.*

In a recent study investigating the consequences of polyploid evolution in flowering plants, Husband *et al.* (2008) pointed out that it is likely that inbreeding depression in newly formed polyploids (neopolyploids) will differ from established polyploids. The results of their study using synthetic neopolyploids of *Chamerion angustifolium* provided some evidence for this theory as inbreeding depression, based on three fitness components, rose from effectively 0 to 0.80 in neopolyploids after one generation of selfing (Husband *et al.*, 2008). Low inbreeding depression in neopolyploids is consistent with the partial dominance model. Despite possessing the same frequency of deleterious mutations compared to diploids, neopolyploids have less inbreeding depression because polyploidization generates high levels of heterozygosity, thus preventing the expression of deleterious mutations. Neopolyploids differ from established polyploids because there has been insufficient opportunity for the frequency of deleterious mutations to reach equilibrium. However as selfing events increase over time, heterozygotes are increasingly of the AAaa or Aaaa form, which are most likely to yield full homozygotes after selfing, thus increasing inbreeding depression (Husband *et al.*, 2008). Summarising, it is evident from many empirical studies that the amount of inbreeding

depression found in a population depends on many interacting genetic and life history factors and that theoretical models can only point to likely outcomes but are not as yet powerful enough to reliably predict the evolution of inbreeding depression in relation to a combination of these factors.

### 4.1.3 Rationale for this chapter

Inbreeding depression is a key factor in establishing the long term evolutionary consequences of hybridization between inbreeding and outcrossing plant taxa. Either parental species will pass on their phenotypic traits which may potentially create a novel phenotype in a later generation hybrid via transgressive segregation (see Chapter 1). The parental species will also pass on their genetic load to their hybrid offspring and thus influence its overall fitness. The likelihood that a new phenotype will persist and ultimately develop into a new species partly depends on how well it is insulated from homogenizing gene flow. Various authors have predicted that the breeding system of the hybridizing plants is an important factor in hybrid speciation (Grant, 1958; 1971: 203; Templeton, 1981) as selfing reduces gene flow considerably and can therefore maintain the genetic and phenotypic novelty of the hybrid. Selfing hybrids are most likely encountered where inbreeding and outcrossing plants hybridize. However, this could come at a considerable fitness cost as selfing will lead to the expression of deleterious mutations resulting in inbreeding depression if either of the parental species carries a high genetic load. Thus, on the one hand hybridization between selfing and outcrossing plant taxa might be of critical importance in the generation of taxonomic novelty. On the other hand the genetic load carried by the parental species might be a major barrier for the establishment of new hybrid lineages.

To be able to predict possible evolutionary outcomes for novel hybrid lineages it is essential to know about levels of inbreeding depression and the genetic load in the parental species. *G. urbanum* is a highly selfing species and as many studies on inbreeding depression in selfing taxa suggest is likely to have purged most of its genetic load, expressing only low levels of inbreeding depression (Charlesworth & Charlesworth, 1987; Husband & Schemske, 1996; Crnokrak & Barrett, 2002). In contrast, *G. rivale* is a highly outcrossing species with a leaky self-incompatibility system possibly harbouring a large genetic load hidden in heterozygous form. The aim of this chapter is to answer the following questions:

1. Do *G. urbanum* and *G. rivale* populations display inbreeding depression, i. e. can we detect fitness differences between selfed and outcrossed progeny?
2. Is there appreciable variation in inbreeding depression among populations of either species?

# 4.2 Materials and Methods

## 4.2.1 Sampling material

Two pure *G. urbanum* (U1, U2a) and two pure *G. rivale* (R1, R2) populations were identified from different sites around Edinburgh and Stirling/Ben Lawers (Table 4.1). Populations U1, R1 and R2 were previously used for the genetic and morphological analysis of pure *G. urbanum* and *G. rivale* populations (Chapter 2) and populations U1, U2a, R1 and R2 for the breeding system analysis (Chapter 3). Between ten and 20 individuals per population were dug up in spring (U1) and summer (U2a, R1, R2) 2006. Plants were transferred to an unheated greenhouse and grown in a common environment in adequately sized pots (c. 15 cm diameter) filled with peat free compost, fertilised once a week during the active growing period and watered when necessary.

| Species | Location | Altitude | Coordinates | Crosstype |
|---|---|---|---|---|
| **Geum urbanum** | Edinburgh (U1) | 170 m | $55^0$ 56′ N | 10 selfed |
| | Braid Hills | | $03^0$ 12' W | 20 outcrossed |
| | Edinburgh (U2a) | 120 m | $55^0$ 55' N | 12 selfed |
| | Colinton Woods | | $03^0$ 14' W | 24 outcrossed |
| **Geum rivale** | Stirling (R1) | 600 m | $56^0$ 32′ N | 14 selfed |
| | Ben Lawers | | $04^0$ 17' W | 46 outcrossed |
| | Edinburgh (R2) | 300 m | $55^0$ 44′ N | 20 selfed |
| | Moorfoot Hills | | $03^0$ 01' W | 45 outcrossed |

**Table 4.1.** Location of individuals from pure Scottish *Geum urbanum* and *G. rivale* populations used for the inbreeding depression analysis. Crosstype: number of individuals which were selfed and outcrossed.

## 4.2.2 Controlled crosses

Controlled crosses were performed within populations to establish whether fitness differences between outcrossed and selfed progeny could be detected. Each maternal parent was selfed using pollen from a flower on the same plant and outcrossed with two other plants from the same population. Pollination procedures were the same as described in Chapter 3.2.2 (p 56). Controlled crosses were carried out between April and July 2007 and replicated in 80% of the crosses in *G. urbanum* populations. Due to the limited number of available flowers in *G. rivale* populations, crosses could only be carried out once. Seed heads were harvested when mature and stored in paper envelopes at $4^oC$ for at least three months to allow for drying. Completely dried seeds were weighed (Satorius electronic scales) and the total number of seeds per seed head and the average individual seed mass (total seed mass/total seed number) was established. During this process it became apparent that nearly all seed heads of both species

contained some seeds which were significantly smaller and lighter with only c. 20% or less of the average seed weight (Figure 4.3). These stunted seeds which seemed to be a mixture of aborted and maybe unfertilized ovules were counted and weighed separately. To test whether the number of stunted seeds per seed head produced is related to the efficiency of manually pollinating flowers, two flowers per maternal parent in population U1 (Braid Hills) were bagged and not interfered with (auto-pollinated). The number of normal and stunted seeds produced by auto- and manual pollination were then compared.

**Figure 4.3.** Different seed sizes found in a *G. urbanum* seed head.

Seeds of each cross were sown according to category (normal and stunted) in pots with peat free compost and left outside over the winter period. If duplicate crosses were available only the cross with the higher seed number was used for sowing. After germination, pots were moved to an unheated greenhouse and the number of germinated seeds was recorded. Stunted seeds were never observed to germinate. Five offspring per controlled cross were grown on and planted into a common garden environment in June 2008 to establish survival rates and the fitness parameter ´volume occupied' (in $cm^3$) which is the average area occupied by the rosette leaves in $cm^2$ multiplied by their average height above ground in cm. Measurements were taken four months after planting out in October 2008.

## 4.2.3 Inbreeding coefficient $F_{is}$

For *G. urbanum* population U1 and *G. rivale* populations R1 and R2, $F_{is}$ values for parental and naturally occurring progeny generations were derived from the analyses in Chapter 3. Parental and offspring $F_{is}$ values were compared for each population to establish whether they provided indirect evidence for inbreeding depression. No data of this kind were available for *G. urbanum* population U2a.

## 4.2.4 Data analysis

Analysis of variance (ANOVA) was used to test crosstype effects (selfed or outcrossed) within populations of *G. urbanum* (U1, U2a) and *G. rivale* (R1, R2) for the following explanatory variables: number of normal seeds produced per cross, percentage of normal seeds per seed head produced, mean individual seed mass, germination rate of normal seeds and volume occupied by five seedlings per cross in a common garden environment. All replicate values for each maternal genotype were averaged within crosstype resulting in one mean genotypic value for each of the fitness traits for selfed and outcrossed events. Percentage data were arcsine transformed and growth volume had to be square root transformed for U2a to comply with ANOVA normality assumptions. A Kruskal Wallis test was performed on number of stunted seeds per cross as the data were not normally distributed despite various transformations.

Variation in germination rates between maternal genotypes within populations were tested for significance using a $\chi^2$ test. Due to the low number of perished selfed and outcrossed seedlings, survival rates were tested for significance using Fisher's Exact Test. Inbreeding depression $\delta_t$ over all life stages was calculated by $1 - \Pi(w_s/w_x)$, where $w_s$ and $w_x$ are fitness estimates of selfed and outcrossed progeny, respectively. Only significant ($p<0.05$) fitness terms were used for this calculation. Statistical tests were performed with the software Minitab 14.1 (Minitab Inc., 2003) apart from Fisher's Exact Test which was conducted using a web based program (http://faculty.vassar.edu/lowry/webtext.html).

# 4.3 RESULTS

## 4.3.1 *Geum urbanum*

*Seed set*

Selfed flowers produced full seed heads in 100% of the crosses in both populations whereas outcrossing produced seed heads in 98.8% and 98.4% of the crosses in U1 and U2a, respectively (no significant difference in success rate in U1: $\chi^2$=0.247, df=1, p=0.62 and U2a: $\chi^2$=0.024, df=1, p=0.88). No significant difference in the percentage of stunted seeds produced by manual- (14.6%) and auto-pollination (10.7%) was detected in the test population U1 ($\chi^2$=3.36, df=1, p=0.07).

| | U1 (Braid Hills) | | | U2a (Colinton Woods) | | |
|---|---|---|---|---|---|---|
| | **Self** | **Outcross** | **δ** | **Self** | **Outcross** | **δ** |
| Number of normal seeds per seedhead | 64.2 (±3.2) | 63.9 (±2.3) | 0.00 | 84.5 (±3.2) | 84.1 (±3.1) | 0.00 |
| Number of stunted seeds per seedhead | 9.4 (±2.0) | 6.5 (±1.1) | 0.45 | 9.4 (±1.5) | 10.6 (±1.9) | -0.11 |
| Percentage of normal seeds per seedhead (%) | 87.1% (±2.7) | 90.3% (±1.6) | 0.04 | 88.1% (±2.4) | 90.4% (±1.0) | 0.03 |
| Mean individual seed mass of normal seeds (mg) | 2.16 mg (±0.09) | 2.06 mg (±0.08) | -0.05 | 2.28 mg (±0.10) | 2.35 mg (±0.08) | 0.03 |
| Germination rate of normal seeds sown (%) | 26.4% (±3.5) | 35.8% (±3.2) | 0.26 | 55.0% (±4.1) | 52.9% (±4.4) | -0.04 |
| Survival rate of seedlings (%) | 96% | 100% | 0.04 | 91.7% | 88.4% | -0.04 |
| Volume occupied by progeny after 4 months ($cm^3$) | 1262 $cm^3$ (±132) | 1253 $cm^3$ (±112) | -0.01 | 1074 $cm^3$ (±155) | 1251 $cm^3$ (±134) | 0.14 |

**Table 4.2.** Mean (standard error) performance and inbreeding depression δ for various life stages of selfed and outcrossed progeny in Scottish *G. urbanum* populations U1 ($n_s$=10 and $n_x$=10) and U2a ($n_s$=12 and $n_x$=12). n: number of genotypes selfed ($n_s$) and outcrossed ($n_x$). No significant differences between selfed and outcrossed progeny were detected in either population.

There was no significant difference between outcrossed and selfed seeds in U1 and U2a for the number (p=0.98 and p=0.83, respectively) and the percentage (p=0.38 and p=0.64, respectively) of normal seeds per seedhead (Table 4.2 and Table 4.3) and in number of stunted seeds per cross (Kruskal-Wallis test; U1, p=0.23 and U2a, p=0.63).

*Seed mass*

An ANOVA revealed that there was no significant mass difference between outcrossed and selfed seeds in U1 (p=0.48) and U2a (p=0.74) (Table 4.2 and Table 4.3).

| Fitness trait | Source | U1 df | U1 MS | U1 p | U2a df | U2a MS | U2a p |
|---|---|---|---|---|---|---|---|
| **Number of normal seeds per seedhead** | Crosstype | 1 | 0.00 | 0.98 | 1 | 10 | 0.83 |
| | Error | 18 | 77.5 | | 22 | 199 | |
| **Percentage of normal seeds per seedhead** | Crosstype | 1 | 0.017 | 0.38 | 1 | 0.005 | 0.64 |
| | Error | 18 | 0.022 | | 22 | 0.021 | |
| **Mean individual seed mass of normal seeds** | Crosstype | 1 | 0.047 | 0.48 | 1 | 0.011 | 0.74 |
| | Error | 18 | 0.090 | | 22 | 0.101 | |
| **Germination rate of normal seeds** | Crosstype | 1 | 0.050 | 0.08 | 1 | 0.003 | 0.76 |
| | Error | 18 | 0.015 | | 22 | 0.035 | |
| **Volume occupied by progeny** | Crosstype | 1 | 360 | 0.96 | 1 | 48.8 | 0.33 |
| | Error | 18 | 150,129 | | 22 | 49.5 | |

**Table 4.3.** ANOVA results for the effect of crosstype (selfed and outcrossed) on various fitness traits in two pure Scottish *G. urbanum* populations U1 and U2a. Volume occupied was square root transformed in U2a prior to analysis. No significant differences between selfed and outcrossed progeny were detected.

## *Germination rate*

No significant difference in germination rates between outcrossed and selfed seeds was observed in U1 ($p=0.08$) and U2a ($p=0.76$) (Table 4.2 and Table 4.3). There was a highly significant difference in terms of germination rates of selfed and outcrossed seeds between maternal genotypes in U1 ($\chi^2=41.5$, $df=9$, $p<0.001$ and $\chi^2=108$, $df=19$, $p<0.001$, respectively) and U2a ($\chi^2=479.6$, $df=11$, $p<0.001$ and $\chi^2=373.6$, $df=23$, $p<0.001$, respectively). Germination rates of selfed seeds varied from 9.2% to 45.7% between maternal genotypes in U1 (n=10) and from 30% to 83.3% in U2a (n=12) whereas outcrossed seeds varied from 13.8% to 63.3% in U1 (n=20) and from 9.9% to 92.6% in U2a (n=24). Stunted seeds were never observed to germinate.

## *Survival rate*

In U1 two out of 50 selfed seedlings (4%) died before planting out into a common garden plot whereas all of the 100 outcrossed individuals survived ($\delta=0.04$). It was not possible to perform a $\chi^2$ test or a Fisher's exact test to calculate the significance level due to the violation of test assumptions but it is unlikely that the 4% survival difference is significant. In U2a no significant difference in survival rate before planting out was detected between selfed (91.7%) and outcrossed seedlings (88.4%, $\chi^2= 0.447$, $df=1$, $p=0.50$, $\delta=-0.04$). All seedlings from populations U1 and U2a which were planted into a common garden plot survived until volume measurements were taken four months later.

*Volume occupied*

No significant difference for volume occupied was detected between selfed and outcrossed seeds in U1 (p=0.96) and U2a (p=0.33) (Table 4.2 and Table 4.3).

*Inbreeding depression $\delta_t$ over all life stages*

Inbreeding depression over all life stages was $\delta_t = 0$ in U1 and U2a.

*$F_{is}$ values for mother-progeny array*

Adult and offspring $F_{is}$ were not significantly different in U1 ($F_{is}$=0.72 and $F_{is}$=0.68, respectively). No data were available for U2a.

## 4.3.2 *Geum rivale*

*Seed set*

R1 and R2 showed a highly significant difference in success of seed set between controlled selfed and outcrossed flowers. In R1 only 10.4% of the selfed but 52.2% of the outcrossed flowers developed a full set head ($\chi^2$=13.3, df=1, p<0.001) whereas in R2 19.2% and 51.2% of attempted crosses were successful, respectively ($\chi^2$=11.2, df=1, p<0.001).

| | R1 (Ben Lawers) | | | R2 (Moorfoot Hills) | | |
|---|---|---|---|---|---|---|
| | **Self** | **Outcross** | **δ** | **Self** | **Outcross** | **δ** |
| Number of normal seeds per seedhead | 92.2 (±11.4) | 70.8 (±10.4) | -0.30 | 81.3 (±7.4) | 64.6 (±6.5) | -0.26 |
| Number of stunted seeds per seedhead | 29.5 (±6.5) | 20.9 (±4.5) | 0.41 | 20.1 (±7.1) | 22.1 (±4.6) | -0.09 |
| Percentage of normal seeds per seedhead (%) | 75.8% (±4.8) | 72.2% (±8.0) | -0.05 | 80.4% (±6.2) | 75.5% (±5.1) | -0.06 |
| Mean individual seed mass of normal seeds (mg) | 1.40 mg (±0.14) | 1.19 mg (±0.11) | -0.18 | 1.00 mg (±0.13) | 0.98 mg (±0.10) | -0.02 |
| Germination rate of normal seeds sown (%) | 60.9% (±10.9) | 39.9% (±8.2) | -0.53 | 33.4% (±10.3) | 36.1% (±11.0) | 0.07 |
| Survival rate of seedlings (%) | 100% | 97% | -0.03 | 90% | 100% | **0.10*** |
| Volume occupied by progeny after 4 months ($cm^3$) | 1932 $cm^3$ (±257) | 2658 $cm^3$ (±256) | 0.27 | 2047 $cm^3$ (±525) | 2452 $cm^3$ (±261) | 0.17 |

**Table 4.4.** Mean (standard error) performance and inbreeding depression δ of selfed and outcrossed progeny in Scottish *G. rivale* populations R1 ($n_s$=6 and $n_x$=9) and R2 ($n_s$=7 and $n_x$=8). Maternal parents which never produced seeds on selfing (SCI=0) could not be included in the analysis. n: number of genotypes selfed ($n_s$) and outcrossed ($n_x$). *0.01<p<0.05.

For both number and percentage of normal seeds per seedhead there was no significant effect of crosstype within R1 (p=0.20 and p=0.97, respectively) or R2 (p=0.11 and p=0.58, respectively) (Table 4.4 and Table 4.5). Similarly, stunted seeds per cross did not differ significantly between selfed and outcrossed progeny (Kruskal-Wallis test; R1, p=0.18 and R2, p=0.60).

*Seed mass*

There were no significant differences in seed mass between selfed and outcrossed seeds within R1 (p=0.24) and R2 (p=0.91) (Table 4.4 and Table 4.5).

| | | R1 | | | R2 | | |
|---|---|---|---|---|---|---|---|
| **Fitness trait** | **Source** | **df** | **MS** | **p** | **df** | **MS** | **p** |
| **Number of normal seeds per seedhead** | Crosstype | 1 | 1637 | 0.20 | 1 | 1041 | 0.11 |
| | Error | 13 | 900 | | 13 | 360 | |
| **Percentage of normal seeds per seedhead** | Crosstype | 1 | 0.000 | 0.97 | 1 | 0.032 | 0.58 |
| | Error | 13 | 0.106 | | 13 | 0.102 | |
| **Mean individual seed mass of normal seeds** | Crosstype | 1 | 0.163 | 0.24 | 1 | 0.001 | 0.91 |
| | Error | 13 | 0.110 | | 13 | 0.092 | |
| **Germination rate of normal seeds** | Crosstype | 1 | 0.293 | 0.11 | 1 | 0.005 | 0.84 |
| | Error | 14 | 0.101 | | 14 | 0.115 | |
| **Volume occupied by progeny** | Crosstype | 1 | 1,965,525 | 0.07 | 1 | 530,600 | 0.48 |
| | Error | 13 | 496,034 | | 11 | 1,011,295 | |

**Table 4.5.** ANOVA results for the effect of crosstype (selfed and outcrossed) on various fitness traits in two pure Scottish *G. rivale* populations R1 and R2. No significant differences between selfed and outcrossed progeny were detected.

*Germination rate*

No significant differences in germination rate between selfed and outcrossed seeds were detected in R1 (p=0.11) and R2 (p=0.84) (Table 4.4 and Table 4.5). There was a highly significant effect of maternal genotype on germination rate in both selfed and outcrossed seeds within population R1 ($\chi^2$=155.9, df=5, p<0.001 and $\chi^2$=316.1, df=15, p<0.001, respectively) and R2 ($\chi^2$=207.8, df=7, p<0.001 and $\chi^2$=337, df=12, p<0.001, respectively). Germination rates of selfed seeds varied considerably between genotypes from 4.9% to 92.9% in R1 (n=6) and from 0% to 76.6% in R2 (n=8) whereas outcrossed seeds had germination rates from 0% to 90.2% in R1 (n=13) and from 0% to 94.2% in R2 (n=16). Stunted seeds were never observed to germinate.

### *Survival rate*

In R1 all of the 33 selfed seedlings survived whereas two out of 68 outcrossed seedlings (3%) died before planting out into a common garden plot (Fisher's Exact Test, p=0.48). In R2 the difference in survival rate until planting out between selfed (26 out of 29, 90%) and outcrossed seedlings (50 out of 50, 100%) was just significant (Fisher's Exact Test, p=0.046, δ=0.10). All seedlings (R1 and R2) planted into a common garden plot survived until volume measurements were taken four months later.

### *Volume occupied*

No significant difference could be detected in R1 (p=0.07) and R2 (p=0.48) for the volume occupied by selfed and outcrossed plants (Table 4.4 and Table 4.5).

### *Inbreeding depression over all life stages*

Inbreeding depression over all life stages was $\delta_t = 0$ in R1 and $\delta_t = 0.10$ in R2 (survival rate).

### *$F_{is}$ values for mother-progeny array*

Adult and offspring $F_{is}$ were -0.07 and -0.02, respectively in R1 and -0.07 and 0.13, respectively in R2.

# 4.4 Discussion

## 4.4.1 *Geum urbanum*

In line with the theory on purging in selfing populations (Schemske & Lande, 1985) and the findings of many other studies (reviewed in Husband & Schemske, 1996; Crnokrak & Barrett, 2002), U1 and U2a did not show significant levels of inbreeding depression for any of the measured life stages. Germination rates showed highly significant variation between maternal genotypes within selfed and outcrossed groups which suggests that genotype is an important factor in determining germination success in the two populations.

There was no significant difference in number of stunted (infertile) seeds between selfed and outcrossed flowers in both populations. The production of stunted and infertile seeds is caused by selective seed abortion which has long been recognized to be a common phenomenon in many plant taxa (Stephenson, 1981). Various causes have been identified such as pollen limitation (Burd, 1994), climate conditions (Lee & Bazzaz, 1982), the position of ovules in the developing fruit (Reed & Singletary, 1989; Mena-Ali & Rocha, 2005) and characteristics related to the reproductive system or genetic makeup of the plant like self-incompatibility (De Nettancourt, 1997), inbreeding or outbreeding depression (Charlesworth & Charlesworth, 1987) or differential genotype selection (Melser & Klinkhamer, 2001). However, in many cases seed abortion is simply caused by resource limitation (Stephenson, 1981; Herrera, 1991; Gimenez-Benavides *et al.*, 2007; Hirayama *et al.*, 2008). Seeds are competing for a limited amount of maternal resources and those which were initiated first and have the faster growth rate can capture more resources and are thus less likely to be aborted than those that are initiated later.

The percentage of unfertilized ovules due to pollen limitation in U1 and U2a is probably very low as stigmas were manually saturated with pollen and checked for pollen coverage under a stereomicroscope. There was no significant difference in the percentage of stunted seeds per seed head between manually selfed and auto-pollinated flowers suggesting that the manual pollination procedure was carried out efficiently. *G. urbanum* is a fully self-compatible species (see Chapter 3) and it seems therefore likely that the production of stunted seeds is due to resource limitation.

Although it is possible that the two populations harbour some mildly deleterious mutations which might be expressed later in the reproductive life stage, the results from this study suggest that most of the genetic load has been purged in U1 and U2a. For population U1 this is also supported by results from mother-progeny arrays (Chapter 3.3.2, p 61; no data were

available for U2a). Parental and offspring $F_{is}$ values ($F_{is}$=0.72 and $F_{is}$=0.68, respectively) did not differ significantly in this population, indicating that inbred progeny do not have a fitness disadvantage compared to outcrossed progeny.

### 4.4.2 *Geum rivale*

Although there was a highly significant difference in crossing success between controlled selfed and outcrossed flowers in both populations, this is more likely to be due to a leaky self-incompatibility system than inbreeding depression as some of the same maternal parents which aborted bagged flowers produced seed heads when selfed manually (Chapter 3.3.1, p 60). No inbreeding depression was detected in R1 at any life stage in contrast to R2 which showed significant levels of inbreeding depression for survival rate (δ=0.10). However, the 37% larger volume of outcrossed compared to selfed progeny in R1 was approaching significance (p=0.07), indicating that significant inbreeding depression (δ=0.27) for growth might have been detected if a larger sample size had been available.

Various studies have reported variability in the expression of inbreeding depression for different variables and between populations of the same species. Johnston (1992) reported that selfed and outcrossed flowers in three *Lobelia* populations produced an indistinguishable number of seeds and that there was no significant difference between crosstypes in the average seed mass per fruit. However, significant inbreeding depression in germination rates was detected in two populations but not in the third one whereas reduced flower numbers of selfed progeny was observed in all three populations (Johnston, 1992).

Levels of inbreeding depression found in R1 ($\delta_t$=0) and R2 ($\delta_t$=0.10) contrasted with the findings of other studies as the amount of inbreeding depression over all life stages in the two *G. rivale* populations was much lower than expected for outcrossing populations. In a review comparing 54 plant species Husband & Schemske (1996) reported an average of $\delta_t$=0.53 for outcrossing plant taxa with the lowest recorded value being $\delta_t$=0.27 in *Sidalcea oregano* subsp. *spicata*. One possible explanation for this discrepancy could be that inbreeding depression is expressed at a later life stage in *G.* rivale, which was not captured by the experiments of this study and/or that the low levels of inbreeding depression are an artefact due to the relatively low numbers of individuals used in the controlled crosses (R1: $n_{self}$=6 and $n_{outx}$=9; R2: $n_{self}$=7 and $n_{outx}$=8). The latter explanation might especially apply to R1 where the difference in growth vigour between selfed and outcrossed progeny was close to being significant.

Alternatively, it is possible that the low levels of inbreeding depression in hexaploid *G. rivale* are connected to the polyploid level of this plant as theoretical models predict less inbreeding depression in polyploids than in their diploid relatives (Lande & Schemske, 1985; Ronfort, 1999). Empirical studies have so far provided only equivocal support for these predictions (Johnston & Schoen, 1996; Husband & Schemske, 1997; Rosquist, 2001; Barringer & Geber, 2008), although investigations using synthetic neopolyploids provided some evidence for low inbreeding depression in early generations (Husband *et al.*, 2008). However, there is only limited applicability of these results to *G. rivale*, as it is considered to be an ancient polyploid (Smedmark *et al.*, 2003) and, as AFLP marker inheritance patterns suggest (Chapter 5), about 80% of the genome behave as a functional diploid.

In a study carried out in the self-incompatible alpine plant *Campanula thyrsoides*, Ægisdottir *et al.* (2007) could not detect inbreeding depression for seed set, seed weight, germination percentage, seedling survival and size between outcrossed and sister-crossed offspring. The authors argue that the fragmented landscape of the Alps and the spatial isolation from other populations could have caused frequent bottlenecks during the colonization of isolated habitats. There is evidence that bottlenecks can lead to a reduction in inbreeding depression, although the decrease of inbreeding depression immediately after a sudden reduction of population size is expected to be only several percent at the most (Kirkpatrick & Jarne, 2000). However, it could also be about 25% with a very high lethal mutation rate and a very slow population growth (Kirkpatrick & Jarne, 2000). Repeated bottlenecks are therefore necessary to reduce the genetic load considerably and if *C. thyrsoides* populations have been exposed to such a scenario frequently then it could explain the low inbreeding depression in this outbreeding species (Ægisdottir *et al.*, 2007).

It is possible that this might have also happened to the *G. rivale* populations used for this study. R1 was collected from an exclosure around a subalpine cliff face at Ben Lawers in the Scottish Highlands. Most areas in the Scottish Highlands have been grazed by a combination of red deer and sheep for much of the past two hundred years and the environmental impact of deer and sheep management on the ecosystem has been much discussed. During this period, heather cover has declined and the prevention of tree regeneration by browsing has led to the gradual erosion of many of the remaining native woodlands (Clutton-Brock *et al.*, 2004). Grazing pressure has put a strain on wild plant populations and their reproductive abilities. It is therefore possible that *G. rivale* populations have gone through repeated bottlenecks in some areas of the Scottish Highlands which could have led to a selective advantage of selfing over outcrossing. Removal of flowers by grazing reduces the effective population size and lowers the chances for successful outcrossing. Selfing is potentially the only option left for some plants to reproduce in an otherwise poorly flowering *G. rivale* population which could lead to subsequent purging of the genetic load in selfed offspring over

time. R1 might be a population which was founded by genotypes purged of most of their genetic load. This scenario could also apply to R2, a population which was sampled along a stream in the upland area of the Moorfoot Hills in the Scottish Borders. Like the Scottish Highlands, the Scottish Borders have been traditionally used as grazing pastures for sheep, potentially creating isolated *G. rivale* populations. To gain more evidence for this hypothesis, inbreeding depression in *G. rivale* populations from areas not subjected to grazing need to be compared to populations from grazed habitats.

There is circumstantial evidence for inbreeding depression in R2 as $F_{is}$ values of parental and offspring generations were different ($F_{is}$=-0.07 and $F_{is}$=0.13, respectively). However, the level of significance for this difference could not be established because only one microsatellite locus was available. No difference between parental and offspring $F_{is}$ values was observed in population R1 ($F_{is}$=-0.07 and $F_{is}$=-0.02, respectively). This is not surprising as the single locus outcrossing rate in R1 was t=90.5% (Chapter 3.2.3, p 63) which was not significantly different from 100%, suggesting that the production of selfed progeny is likely to be very low.

It also has to be born in mind that all measurements were taken in the more benign environments of an unheated greenhouse and a common garden plot. In a recent review on this topic it has been shown that measurements taken under harsher field conditions have resulted in higher levels of inbreeding depression in about 70% of the cases (Armbruster & Reed, 2005). It is therefore possible that inbreeding depression has been underestimated in the studies of *G. urbanum* and *G. rivale* populations.

## 4.4.3 Summary

In line with theoretical predictions no inbreeding depression was observed in selfing *G. urbanum* populations U1 and U2a. Inbreeding depression was absent (R1) or low (R2) in outcrossing *G. rivale* populations, which is in contrast to the findings of other studies investigating levels of inbreeding depression in outcrossing plant taxa.

# 5. CONTROLLED F1s BETWEEN *GEUM URBANUM* AND *G. RIVALE*

## Abstract

The genetic and morphological structure of controlled F1 crosses between *G. urbanum* and *G. rivale* was analysed and compared to pure parental populations. A total of 14 different reciprocal crosses were performed using individuals from two *G. urbanum* and three *G. rivale* populations resulting in nine *G. urbanum* x *G. rivale* and five *G. rivale* x *G. urbanum* crosses. 41 controlled F1 samples and 39 individuals from pure *G. urbanum* and 32 individuals from pure *G. rivale* populations were measured for nine morphological characters. 15 controlled F1s as well as 12 samples per species from various pure populations throughout Great Britain were screened for 203 AFLP markers. Additionally, genetic marker inheritance patterns were analysed using 16 offspring from one controlled ♀*G. urbanum* x ♂*G. rivale* and 18 offspring from one controlled ♀*G. rivale* x ♂*G. urbanum* cross for 118 and 136 AFLP markers, respectively. Controlled crosses were successful in both directions but germination rates differed significantly ($\chi^2$=198.22, df=1, p<0.0001) between *G. urbanum* x *G. rivale* (83%) and *G. rivale* x *G. urbanum* (31%) crosses. Mean pollen fertility was significantly lower ($\chi^2$=443.5, df=2, p<0.001) in controlled F1s (70%) compared to parental species (>90%). A Principal Components Analysis (PCA) and a Principal Coordinate Analysis (PCO) showed that the controlled F1 crosses are intermediate between the parental species and can be clearly identified as hybrids morphologically and genetically. One transgressive character (brown stamens with indumentum) which is a trait not observed in pure parental populations, was recorded in 7% (3 out of 41) of the controlled crosses. AFLP analysis of offspring from controlled F1 crosses revealed that about 80% of polymorphic markers were inherited in a Mendelian fashion (p<0.05) and that 13% and 25% of species-specific *G. urbanum* and *G. rivale* markers segregated, respectively, indicating cryptic heterozygosity. Chloroplast inheritance was established to be predominantly maternal.

# 5.1 Introduction

The production of first generation hybrids (F1s) is the first step in hybridization and the successful survival and productivity of F1s is a significant determinant of the evolutionary consequences of hybridization. For example F1s usually have lower pollen fertility than their parents which reduces the ability of hybrids to sire offspring. In *Helianthus annuus* x *H. petiolaris* hybrids pollen fertility is reported to be about 6% (Heiser, 1947), in *Gilia malior* x *G. modocensis* hybrids about 2% (Grant, 1966), in *Mimulus lewisii* x *M. cardinalis* about 30% (Ramsey *et al.*, 2003) and in *Armeria pseudoarmeria* x *A. welwitschii* most F1 hybrids have pollen fertility between 40 to 60% (Tauleigne-Gomes & Lefebvre, 2008). However, F1s are not necessarily less fit in all fitness parameters than their parents as a study in *Iris* has shown. Burke *et al.* (1998) investigated components of fitness such as seed germination, vegetative and clonal growth as well as sexual reproduction in *Iris fulva*, *I. hexagona* and their reciprocal F1 hybrids and reported that F1s performed as well as, or significantly better than, both of their parents. Similarly, other studies have found that F1s can display fitness that is comparable to their parents (Arnold & Hodges, 1995; Campbell *et al.*, 2005).

In previous studies of *Geum* it has been reported that the generation of F1s was not independent of the direction of the cross (i. e. which species served as the maternal parent). Prywer (1932) was not able to obtain any seeds if *G. rivale* was the maternal parent and Marsden-Jones (1930) reported that "for two years in succession I tried to make the hybrid with *G. rivale* as the seed parent, making many crosses each year, but failed to obtain a single viable seed, indeed frequently the achenes never developed at all." However, Marsden-Jones (1930) also cited that "it is of considerable interest that Dr Bell Slater obtained a generation by crossing this way". Asymmetric hybrid formation (Rieseberg *et al.*, 1991; Broyles, 2002; Van Der Velde & Bijlsma, 2004; Mahelka *et al.*, 2007; Miglia *et al.*, 2007; Turelli & Moyle, 2007) or asymmetric survival chances (Campbell & Waser, 2001; Tiffin *et al.*, 2001; Kirk *et al.*, 2005) are very common in plants. A recent review of asymmetrical crossing barriers in angiosperms reported that within 14 diverse angiosperm genera significant asymmetry in 35–45% of all species pairings, evaluated at three different postmating stages of reproductive isolation, was detected (Tiffin *et al.*, 2001). These asymmetries can be caused by a variety of prezygotic mechanisms such as differences in style length, mating system (self compatible versus self-incompatible) and differential fruit abortion (Tiffin *et al.*, 2001), complex genotype-environment interactions (Miglia *et al.*, 2007; Ruane & Donohue, 2007), cytoplasmic incompatibilities, i. e. incompatibilities between the cytoplasm of the maternal parent and the nuclear genome of the paternal parent (Tiffin *et al.*, 2001) and cytoplasmic-environment interactions (Campbell & Waser, 2007). Nuclear–cytoplasmic interactions occur in all organisms where the function of haploid cytoplasmic organelles (mitochondria,

chloroplasts and plastids) depends on the coordinated expression with the diploid nuclear genome (Turelli & Moyle, 2007). These incompatibilities are often the result of divergent evolution of nuclear and cytoplasmic genes coding for proteins which interact in various physiological processes such as photosynthesis or respiration (Michaelis, 1954; Levin, 1978) but can also be caused by transposable elements (Tiffin *et al.*, 2001).

In order to study the evolutionary consequences of hybridization in a natural hybrid zone between *G. urbanum* and *G. rivale*, it is important to establish the genetic and morphological properties of controlled F1s first. These data in conjunction with data on pure populations (Chapter 2) will then provide the framework against which samples from the hybrid zone can be compared and assessed.

To establish such a framework reciprocal controlled F1 crosses between pure *G. urbanum* and *G. rivale* individuals were performed and the following questions were addressed:

1. What is the phenotypic and genetic structure of controlled F1 individuals?
2. What is their pollen fertility and their survival rate?
3. Is there evidence for asymmetric hybrid formation?

It is also important to test whether certain assumptions about the methods used to gather these data are correct for the study species. For example, in AFLP assays it is generally assumed but rarely tested that markers are inherited in a predictable Mendelian way and that species-specific markers are non-segregating (i. e. homozygous). Statistical software packages for identifying species hybrids using data on multiple, unlinked markers such as NEWHYBRIDS (Anderson & Thompson, 2002) assume Mendelian inheritance and need species-specific markers to classify hybrid individuals. Deviations from both assumptions could therefore potentially lead to the incorrect classification of samples. To test whether these assumptions are fulfilled in the study species the following additional questions were addressed:

4. Does the AFLP marker inheritance pattern observed in controlled F1s correspond to Mendelian expectations?
5. Are parental species-specific markers homozygous?
6. Can the software NEWHYBRIDS allocate controlled F1 crosses to the correct hybrid category even if deviations from Mendelian inheritance and heterozygous species-specific markers exist?

# 5.2 Materials and Methods

## 5.2.1 Sample material

Controlled F1 crosses were performed in an unheated greenhouse between individuals collected from five pure Scottish populations (Table 5.1). A total of 14 different reciprocal crosses were performed at the beginning of June 2006 (nine *G. urbanum* x *G. rivale* crosses and five *G. rivale* x *G. urbanum* crosses) using individuals from two *G. urbanum* and three *G. rivale* populations.

Pollination procedures were the same as outlined in Chapter 3.2.2 (p 56). Seeds of successful crosses were collected, sown in pots with peat free compost and left outside over the winter period. After germination, pots were moved into an unheated greenhouse and leaf material for DNA extraction was collected from 59 seedlings when about 2 cm tall. 35 offspring were derived from six crosses that used *G. urbanum* as the maternal parent and 24 offspring were derived from five crosses using *G. rivale* as the female parent. Subsequently, all seedlings were planted out into a common nursery environment.

To compare the morphological and genetic structure of the controlled F1 samples with the variation found in pure populations, eight pure populations additional to the parental populations were sampled for reference purposes (Table 5.1).

## 5.2.2 Non molecular data

*Morphology*

To establish the extent of the phenotypic variation found in controlled F1s a total of nine floral, leaf and fruit characters were recorded. All measurements were taken after seedlings had been growing in a randomised array in a common nursery environment until flowering and fruiting. The nine characters used were the same traits measured in pure populations from Chapter 2.2.3 (p 28) and included the length and width of stipule, petal and sepal in centimetres, and the length of the petal claw, gynophore and upper style joint in millimetres (Figure 2.1, p 31). The means for the above mentioned characters were compared between controlled F1 crosses and pure species to establish inheritance patterns for each trait. Differences were tested for significance with a Kruskal Wallis test. A principal components analysis (PCA) using all morphological measurements was performed for pure populations and controlled F1s.

| Species | Country | Location | Altitude | Coordinates | Controlled F1 | Morphology | AFLP | cpDNA | Pollen |
|---|---|---|---|---|---|---|---|---|---|
| **Geum urbanum** | **Scotland** | Edinburgh, Braid Hills | 170 m | 55$^0$ 56′ N/ 03$^0$ 12' W | 4 | 11 | 2 | 1 | 11 |
| | | Edinburgh, Colinton Ruderal | 120 m | 55$^0$ 55' N/ 03$^0$ 14' W | | 5 | | | 2 |
| | | Edinburgh, Colinton Wood | 120 m | 55$^0$ 55' N/ 03$^0$ 14' W | | 13 | 1 | | 4 |
| | | Edinburgh, Craigmillar | 70 m | 55$^0$ 56′ N/ 03$^0$ 09′ W | 6 | 10 | 2 | | 5 |
| | | Edinburgh, Gorebridge | 90 m | 55$^0$ 51′ N/ 03$^0$ 04′ W | | | 1 | | |
| | **England** | Yorkshire, Swainby | 200 m | 54$^0$ 24 N'/ 01$^0$ 16' W | | | 1 | | |
| | | Gloucestershire, Cotswolds | 200 m | 52$^0$ 03' N/ 01$^0$ 48' W | | | 1 | | |
| | | Gloucestershire, Cotswolds | 240 m | 51$^0$ 46' N/ 02$^0$ 08' W | | | 1 | | |
| | | London, Kew | 10 m | 51$^0$ 28' N/ 00$^0$ 16' W | | | 1 | | |
| | | Devon, Dawlish | 20 m | 50$^0$ 34′ N/ 03$^0$ 28' W | | | 1 | | |
| | | Cornwall, Truro | 50 m | 50$^0$ 12′ N/ 05$^0$ 03' W | | | 1 | | |
| **Geum rivale** | **Scotland** | Orkney, Berridale Hoy | 40 m | 58$^0$ 53′ N/ 03$^0$ 23' W | | | 1 | | |
| | | Ullapool, Inchnadamph | 180 m | 58$^0$ 09′ N/ 04$^0$ 56' W | | | 1 | | |
| | | Ullapool, Knockan Craig | 250 m | 58$^0$ 02′ N/ 05$^0$ 03' W | | | 1 | | |
| | | Thurso, Betty Hill | 20 m | 58$^0$ 31′ N/ 04$^0$ 14' W | | | 1 | | |
| | | Stirling, Ben Lawers | 600 m | 56$^0$ 32′ N/ 04$^0$ 17' W | | 13 | 1 | | 13 |
| | | Edinburgh, Pentland Hills, Carlops | 300 m | 55$^0$ 48′ N/ 03$^0$ 21' W | 1 | 2 | | | |
| | | Edinburgh, Pentland Hills, Bonaly | 210 m | 55$^0$ 53′ N/ 03$^0$ 16' W | 1 | 1 | 1 | | 1 |
| | | Scottish Borders, Moorfoot Hills | 300 m | 55$^0$ 44′ N/ 03$^0$ 01' W | | 16 | 1 | | 8 |
| | | Scottish Borders, Moffat Dale | 300 m | 55$^0$ 23′ N/ 03$^0$ 19' W | 5 | | 3 | 3 | 1 |
| | **England** | Yorkshire, Cross Fell | 800 m | 54$^0$ 42′ N/ 02$^0$ 29' W | | | 1 | | 1 |
| | | Yorkshire, Teesdale | 260 m | 54$^0$ 39′ N/ 02$^0$ 09' W | | | 1 | | |

**Table 5.1.** Location of individuals collected from pure *Geum urbanum* and *G. rivale* populations used for controlled F1 crosses (Controlled F1), chloroplast inheritance (cpDNA) and for reference purposes (Morphology, AFLP, Pollen fertility) with number of individuals used.

## *Germination and survival rates of controlled F1 crosses*

To test whether there are significant differences in germination rates between *G. urbanum* x *G. rivale* and *G. rivale* x *G. urbanum* crosses, ANOVA was performed on arcsine transformed germination data. Survival rates until the first fruit set were established in August 2008.

## *Pollen fertility*

To establish whether pollen fertility is affected by hybrid origin, pollen fertility was measured in pure *G. urbanum* (n=22, Table 5.1), pure *G. rivale* (n=24, Table 5.1) and controlled F1s (n=14). The mean group fertility was established by averaging over all individual fertility counts in that group. Pollen analysis was conducted by collecting pollen from dehiscing anthers of a flower onto a microscope slide. Pollen grains were stained with 2% acetocarmine for about 10 minutes and examined under a light microscope to determine the percentage of stained pollen grains as an indicator for pollen fertility (Figure 5.1). The pollen nucleus is rich in chromatin and viable pollen stains pink to red with acetocarmine, in contrast to sterile pollen which does not take any stain (McKellar & Quesenberry, 1992; Marutani *et al.*, 1993). Stained and unstained pollen grains were counted in two to five different areas of the slide and the respective numbers were then summed up to a total figure.

**Figure 5.1.** Fertile and infertile pollen grains in a *G. urbanum* x *G. rivale* hybrid.

## 5.2.3 Molecular data

DNA was extracted from either silica dried leaf material or fresh leaves as detailed in Chapter 2.2.5 (p 31).

*Chloroplast DNA (cpDNA)*

To establish whether chloroplast is predominantly maternally or paternally inherited in the two *Geum* species one *G. urbanum* x *G. rivale* and two *G. rivale* x *G. urbanum* crosses and six of their offspring each (18 offspring in total) were surveyed for their chloroplast haplotypes following the method outlined in Chapter 2.2.5 (p 32) using *Eco*R1 restriction digest of trnL-F. The parents for these crosses were chosen for their different haplotypes ('1 band' for the *G. urbanum* and '2 bands' for the *G. rivale* crossing parent, see Chapter 2.3.2, p 38).

*AFLP fingerprinting*

To establish the genetic structure of controlled F1 crosses (n=15) and to be able to put it into context with the genetic structure found in pure British populations (n=12 per species, Table 5.1) AFLP fingerprinting was carried out as detailed in Chapter 2.2.5 (p 31). Four primer combinations were chosen based on the number of species-specific markers produced and their ability to generate clear peaks: 6FAM-*Eco*R1 AAC/ *Mse*1 CTA, PET-*Eco*R1 AAT/ *Mse*1 CAGA, VIC-*Eco*R1 AGC/ *Mse*1 CTT, NED-*Eco*R1 ATC/ *Mse*1 CGA.

The selected four *Eco*R1 primers were labelled with four different dyes to allow multiplexing. The species-specific nature of markers for the above mentioned primer pairs was established in Chapter 2.2.5 (p 31) using 30 pure *G. urbanum* and 30 pure *G. rivale* samples collected from various British populations (Table 2.1, p 30).

To visualise the genetic structure of pure populations and controlled F1s a Principal Coordinates Analysis (PCO) of AFLP data was performed with the program PAST using a Jaccard similarity index (Hammer *et al.*, 2006).

*Testing NEWHYBRIDS*

NEWHYBRIDS is a software program implementing statistical methods for identifying species hybrids using data on multiple, unlinked markers (Anderson & Thompson, 2002). The program employs a Bayesian approach using Markov Chain Monte Carlo simulations to calculate the posterior probability that each individual belongs to the following six classes:

pure *G. urbanum*, pure *G. rivale*, F1, F2, backcross to *G. urbanum* and backcross to *G. rivale*. It does not require prior information on genetic composition of the parental populations, but assigns individuals to putative populations based on minimization of departures from Hardy–Weinberg and linkage equilibrium. However, the program requires species-specific markers for distinguishing between the genotype frequency classes of hybrid individuals, e. g. F1 vs F2 (Anderson & Thompson, 2002).

To test the predictive power of the program 59 controlled F1 offspring from ten different controlled hybrid crosses (six *G. urbanum* x *G. rivale* and four *G. rivale* x *G. urbanum*) scored for 185 AFLP markers were analysed using NEWHYBRIDS. The program was run using the default parameters for the six genotype class frequencies and Jeffreys priors, no individual or allele frequency prior information was specified, and a burn-in phase of 100 000 steps and 100 000 sweeps was used.

### *Marker inheritance patterns*

To establish marker inheritance patterns and numbers of segregating and non-segregating AFLP markers (species-specific and non-species-specific), DNA of 16 offspring from one controlled *G. urbanum* x *G. rivale* and 18 offspring from one *G. rivale* x *G. urbanum* cross was extracted. AFLP inheritance patterns were established using only loci which were heterozygous in one parent (i. e. polymorphic in the offspring) and homozygous (no band present) in the other parent. If markers segregate in a Mendelian fashion one would expect to find a 1:1 ratio of band presence/absence in the offspring. However, this is only true if a large number of offspring is available. To ascertain whether markers are likely to show Mendelian segregation in finite sample sizes a $\chi^2$ test was used to calculate the probability of the observed deviation from a strict 1:1 ratio being due to chance.

# 5.3 Results

## 5.3.1 Non molecular data

*Morphology*

A PCA plot of controlled F1 crosses (n=23 from eight *G. urbanum* x *G. rivale* crosses and n=18 from five *G. rivale* x *G. urbanum* crosses) and pure populations (n=39 for *G. urbanum* and n=32 for *G. rivale*) using nine morphological characters clearly showed that controlled F1s were morphologically intermediate and did not overlap with either parent species (Figure 5.2). All traits contributed more or less equally to the differentiation along the first component (76.3% of the variation explained). Stipule length and petal breadth were the most distinguishing traits along the second component (11.5% of the variation explained).

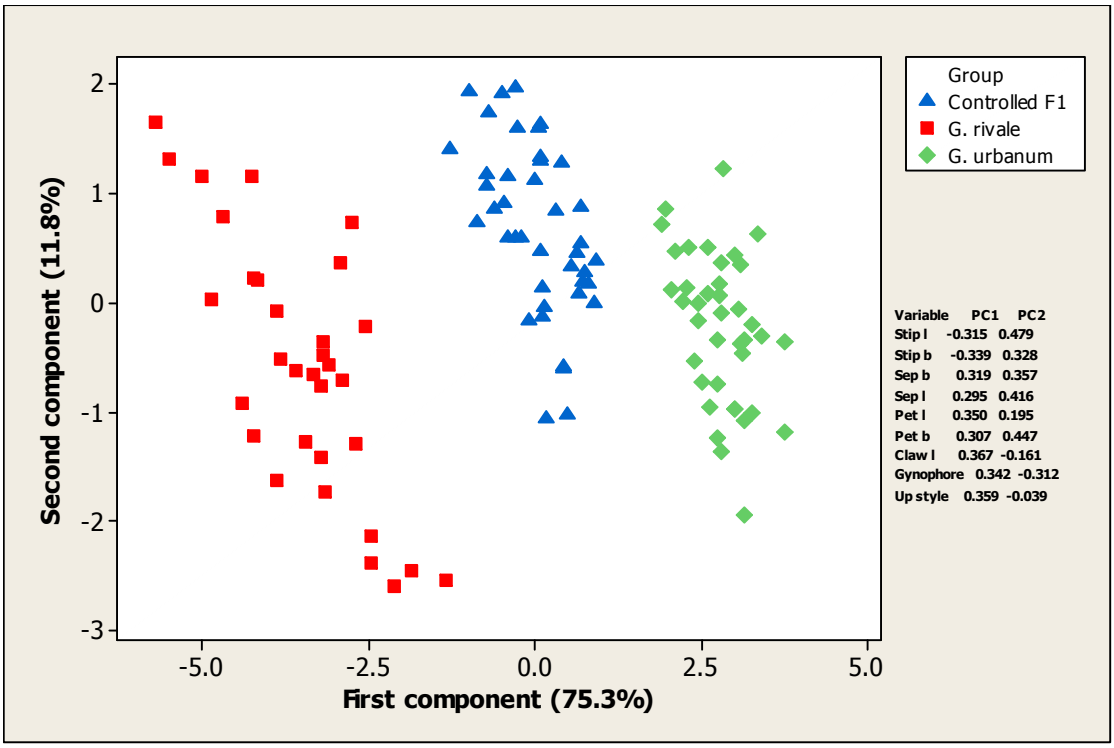



**Figure 5.2.** PCA plot using nine morphological characters taken from individuals in pure *G. urbanum* (n=39), pure *G. rivale* (n=32) and controlled F1 crosses (n=41). Measurements were taken in an unheated greenhouse (pure species) and in a common nursery environment (controlled F1 crosses).

If a character is inherited dominantly (from one species only), then it is expected that the trait average in the hybrid offspring does not differ significantly from the trait average of the contributing species. In contrast, if additive inheritance (from both species) is present then the trait mean of the hybrid is expected to differ significantly from each species mean but not

from the average of the two species. The majority of trait means (67%) in controlled F1 individuals was not significantly different from the overall average of *G. urbanum* and *G. rivale* characters indicating that they are inherited additively (Table 5.2). One character (11%) showed dominant *G. urbanum* inheritance (stipule length) and two characters (22%) dominant *G. rivale* inheritance (sepal length and petal breadth, Table 5.2).

No transgressive segregation (i. e. trait values beyond the parental ranges in either species) for the nine analysed morphological characters was observed in controlled F1s. However, brown stamens with indumentum, a character which has not been observed in either pure *G. urbanum* or pure *G. rivale* populations, was detected in 7% (two *G. urbanum* x *G. rivale* and one *G. rivale* x *G. urbanum* cross) of all (n=41) controlled F1s (Figure 5.3). Stamens in pure populations are yellow and glabrous.

| Trait | Average (±SE) Urbanum | Average (±SE) controlled F1 | Average (±SE) Rivale | Average (±SE) Urb_Riv |
|---|---|---|---|---|
| Stipule length (cm) | 1.89 (±0.04) | 1.78 (±0.05) (U) | 0.76*** (±0.03) | 1.38*** (±0.07) |
| Stipule breadth (cm) | 2.09*** (±0.05) | 1.51 (±0.05) (F1) | 0.45*** (±0.04) | 1.35 (±0.10) |
| Sepal length (cm) | 0.63*** (±0.01) | 0.87 (±0.02) (R) | 0.93 (±0.03) | 0.76** (±0.02) |
| Sepal breadth (cm) | 0.34*** (±0.01) | 0.47 (±0.01) (F1) | 0.54*** (±0.02) | 0.43 (±0.02) |
| Petal length (cm) | 0.72*** (±0.01) | 0.95 (±0.02) (F1) | 1.14*** (±0.02) | 0.91 (±0.03) |
| Petal breadth (cm) | 0.62*** (±0.02) | 0.87 (±0.02) (R) | 0.95 (±0.03) | 0.77** (±0.03) |
| Claw length (mm) | 0.00*** (±0.00) | 1.97 (±0.10) (F1) | 5.13*** (±0.15) | 2.31 (±0.31) |
| Gynophore (mm) | 0.00*** (±0.00) | 1.25 (±0.11) (F1) | 6.17*** (±0.30) | 2.78 (±0.39) |
| Upper style joint (mm) | 1.56*** (±0.03) | 3.09 (±0.05) (F1) | 4.82*** (±0.17) | 3.04 (±0.21) |

**Table 5.2.** Average values (standard error) for nine morphological characters taken in a common environment in pure *G. urbanum* (39), controlled F1 crosses (41) and pure *G. rivale* (32) populations. Average Urb_Riv: overall average of pure *G. urbanum* and *G. rivale* measurements. (U): trait inherited dominantly from *G. urbanum* (significantly different from *G. rivale* and Average Urb/Riv but not from *G. urbanum* means); (R): trait inherited dominantly from *G. rivale* (significantly different from *G. urbanum* and Average Urb/Riv but not from *G. rivale* means); (F1): trait inherited additively (significantly different from *G. rivale* and *G. urbanum* but not from Average Urb_Riv values); Kruskal Wallis test for significant differences between F1 means and means of other groups : **0.001<p<0.01, ***p≤0.001.

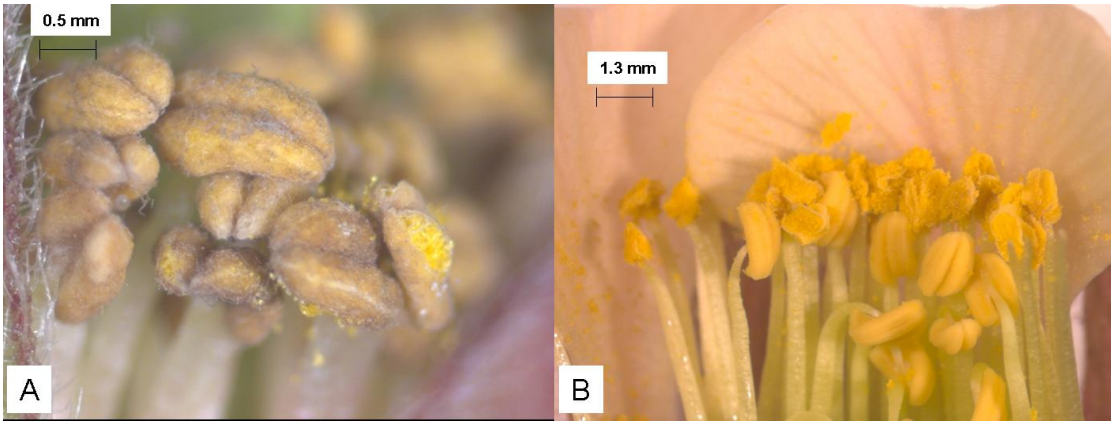



**Figure 5.3.** Stamens of **(A)** a controlled F1 showing brown surfaces with indumentum and **(B)** a pure *G. rivale* sample showing yellow and glabrous surfaces.

## *Germination and survival rates of controlled F1 crosses*

All nine *G. urbanum* x *G. rivale* and all five *G. rivale* x *G. urbanum* crosses were successful. ANOVA revealed that the overall germination rates for the two cross directions differed significantly ($F_{1,12}$=9.52, p=0.009) and was 83% for *G. urbanum* x *G. rivale* and 31% for *G. rivale* x *G. urbanum* crosses. Apart from one *G. urbanum* x *G. rivale* cross where no seeds germinated at all, individual germination rates with *G. urbanum* as the female parent lay above 84%. In contrast, none of the *G. rivale* x *G. urbanum* crosses showed germination rates higher than 69%. There was no significant difference in number of seeds produced per seed head between the two cross directions ($F_{1,12}$=0.0, p=0.99). Out of the 59 planted F1s only one individual (1.7%) died before the first fruit set.

## *Pollen fertility*

Mean pollen fertility was above 90% in pure populations and 70% in controlled F1s (Table 5.3). ANOVA revealed that there was a significant difference in pollen fertility between pure populations and controlled F1s (*G. urbanum* vs F1s: $F_{1,32}$=46.92, p<0.001 and *G. rivale* vs F1s: $F_{1,36}$=24.04, p<0.001). The difference in fertility between *G. urbanum* (94%) and *G. rivale* (90%) populations was not significant ($F_{1,42}$=2.07, p=0.157).

| Group | Samples used | Mean fertility [%(±SE)] | Min. fertility [%] | Max. fertility [%] |
|---|---|---|---|---|
| **Pure populations** | | | | |
| *G. urbanum* | 20 | 94.3(±2.0) | 54.7 | 100 |
| *G. rivale* | 24 | 90.3(±3.0) | 20.0 | 99.0 |
| **Controlled F1s** | 15 | 70.0(±4.0) | 28.9 | 86.2 |

**Table 5.3.** Pollen fertility in pure *G. urbanum* and *G. rivale* populations and controlled F1 crosses.

## 5.3.2 Molecular data

*cpDNA inheritance*

A survey of 18 controlled F1 offspring from three different crosses and their parents using cpDNA-RFLPs (*Eco*R1 on trnL-F) revealed that the maternal banding pattern was present in all offspring (Figure 5.4).

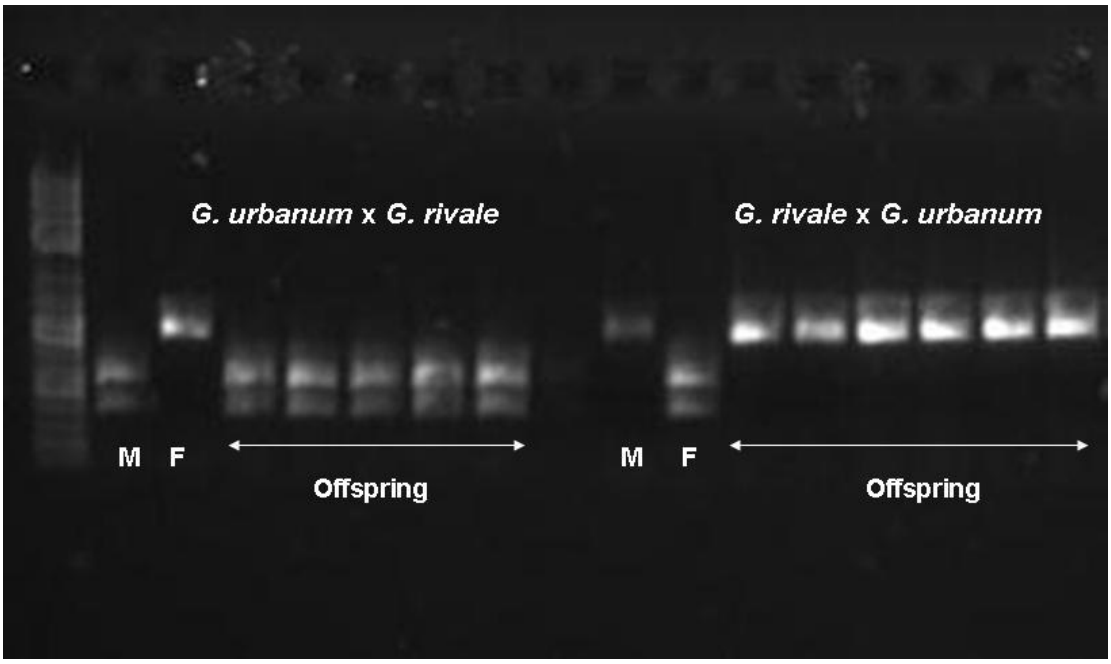



**Figure 5.4.** cpDNA inheritance (*Eco*R1 on trnL-F) in two controlled hybrid (F1) crosses. M: Mother, F: Father.

*AFLP marker inheritance patterns in controlled F1 crosses*

A total of 118 and 136 markers (four primer combinations) were scored for the parents and offspring of one controlled F1 cross in each direction. Loci where the null allele showed a frequency of more than 6%, i. e. loci where the band absence could be found in at least two samples, were scored as polymorphic. There is strong evidence from the AFLP data that accidental maternal self-pollination can be excluded and that all analysed offspring were outcrossed, as all offspring contained a substantial number of bands that were absent in the female parent but present in the male parent. The percentage of polymorphic markers in the Urb x Riv cross was half (25.4%) of that found in the Riv x Urb cross (50.0%). More than 75% of the polymorphic markers in both crosses displayed segregation ratios expected under Mendelian inheritance ($p<0.05$, Table 5.4).

| Female x Male | Markers scored | No. polym. mark. (f>6%) | % polym. mark. (f>6%) | No. polym. Mendel (f>6%) | % polym. Mendel (f>6%) |
|---|---|---|---|---|---|
| **Urb x Riv** | 118 | 30 | 25.4 % | 24 | 80% |
| **Riv x Urb** | 136 | 68 | 50.0% | 52 | 76.5% |

**Table 5.4.** Total number of polymorphic markers and number of polymorphic markers showing Mendelian inheritance found in controlled hybrid crosses *G. urbanum* x *G. rivale* (Urb x Riv, 16 offspring) and *G. rivale* x *G. urbanum* (Riv x Urb, 18 offspring). **No./% polym. mark. (f>6%)**: number/percentage of polymorphic markers where the null allele has a frequency of more than 6%; **No./% polym. Mendel (f>6%)**: number/percentage of polymorphic markers which segregate in a Mendelian fashion (p>0.05).

| UrbxRiv and RivxUrb | Species specific *G. urbanum* markers | Species specific *G. rivale* markers |
|---|---|---|
| **Markers scored** | 15 | 8 |
| **Number non-segregating** | 13 | 6 |
| **% non-segregating** | 86.7% | 75.0% |
| **Number segregating** | 2 | 2 |
| **% segregating** | 13.3% | 25.0% |

**Table 5.5.** Number of segregating and non-segregating species-specific markers in two controlled hybrid crosses *G. urbanum* x *G. rivale* (Urb x Riv) and *G. rivale* x *G. urbanum* (Riv x Urb). Segregating markers are markers where the frequency of the null allele > 6%.

23 species-specific markers were used comprising 15 species-specific *G. urbanum* and eight species-specific *G. rivale* markers to analyse segregation patterns: 13.3% of *G. urbanum* and 25.0% of *G. rivale* species-specific markers showed segregation, suggesting heterozygosity (Table 5.5). Only 0.8% of the markers occurred in the male parent but in none of the offspring which may indicate an organelle genome origin of these markers.

## *Genetic structure of pure and controlled F1 samples*

The PCO analysis using 203 AFLP markers revealed that the controlled F1 samples (n=15) were clearly intermediate between the parental species (n=12 per species) and did not overlap with either cluster.

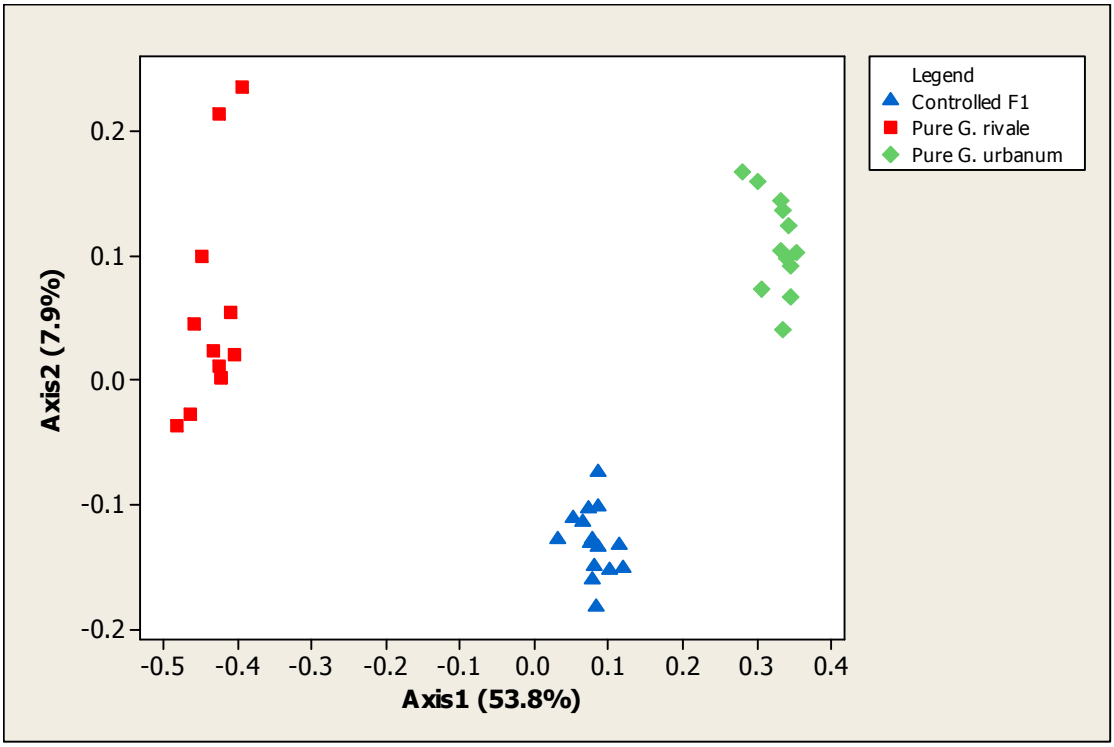



**Figure 5.5.** PCO plot using 203 AFLP markers of 15 controlled F1 crosses and 12 pure individuals per species collected throughout Great Britain.

## *Testing NEWHYBRIDS*

59 controlled F1 crosses and their parents were scored for 185 markers and analysed with NEWHYBRIDS to test the predictive power of this software. 96.6% of the controlled F1 samples were correctly identified as 100% F1s and only two individuals (3.4%) were classified as F2s, highlighting the overall predictive power of the NEWHYBRIDS analysis.

# 5.4 Discussion

## 5.4.1 Success and germination rates of controlled F1 crosses

Despite the difficulty reported in the literature to obtain seeds if *G. rivale* is the maternal parent (Marsden-Jones, 1930; Prywer, 1932) all five *G. rivale* x *G. urbanum* crosses were successful. Although crosses in both directions were successful in terms of seed set in this study, there was a significant difference in germination rates between *G. urbanum* x *G. rivale* (83%) and *G. rivale* x *G. urbanum* crosses (31%).

Asymmetric hybrid formation (Rieseberg *et al.*, 1991; Broyles, 2002; Van Der Velde & Bijlsma, 2004; Mahelka *et al.*, 2007; Miglia *et al.*, 2007; Turelli & Moyle, 2007) or asymmetric survival chances (Campbell & Waser, 2001; Tiffin *et al.*, 2001; Kirk *et al.*, 2005) are very common in plants. Campbell and Waser (2001) crossed naturally hybridizing *Ipomopsis aggregata* and *I. tenuituba* and reciprocally planted the offspring into a range of natural environments across a hybrid zone. All types of crosses produced similar numbers and weights of seeds but survival of the offspring after 5 years differed markedly among crosstypes. On average, the F1 hybrids had survival and growth rates as high as the average for their parents. However, hybrid survival depended strongly on the direction of a cross, i. e. which species served as the maternal parent. This fitness difference between reciprocal hybrids appeared only in the parental environments, suggesting cytonuclear gene interactions that are environment specific (Campbell & Waser, 2001). Similar results were obtained in a different study where the cross direction influenced the fitness of F1 hybrids between *Senecio jacobaea* and *S. aquaticus* (Kirk *et al.*, 2005). Offspring from *S. jacobaea* mothers exhibited higher fitness than those from *S. aquaticus* mothers when compared with parental species and natural hybrids.

The difficulty reported in the literature to produce hybrids if *G. rivale* is the maternal parent was not verified by the results of this study. However, the differential germination success between the two cross directions suggests that some kind of incompatibility is in place if *G. rivale* is the maternal parent. This is possibly due to cytoplasmic-nuclear incompatibilities although complete cytoplasmic-nuclear incompatibility can be excluded in the present case as reciprocal crosses were possible.

## 5.4.2 Morphology

Low phenotypic variability has long been recognized to be a general feature of F1s (Briggs & Walters, 1997) and was also observed in this study. The morphological analysis showed that the controlled F1s are uniform in their appearance, which has also been reported by previous studies in *Geum* (Marsden-Jones, 1930; Prywer, 1932). F1s were clearly identifiable as hybrids in a PCA plot as they clustered inbetween their parents and did not overlap with either parental species.

Phenotypic comparisons of controlled F1 samples with pure samples from either species revealed that two thirds of the nine analysed characters are inherited additively, i. e. are expressed as the average of the two parental mean trait values. One character (stipule length) was dominated by *G. urbanum* and two (sepal length and petal breadth) by *G. rivale*. The expression of a gynophore seems to be a variable trait in F1s. Although all of the 41 controlled F1s measured for this study displayed a gynophore, Prywer (1932) did not observe this character in his F1 crosses whereas Marsden-Jones (1930) reported a gynophore in some crosses. It is possible, however, that Prywer (1932) did not consider the, on average, very short length of the gynophore in F1s (1.25 mm) as distinct enough to deserve reporting.

There was no evidence for transgressive segregation for any of the nine measured traits in the F1 generation. However, one transgressive character was observed. Pure populations of either parental species only display yellow and glabrous anthers whereas 7% of the controlled F1s had brown stamens with indumentum, indicating a transgressive origin. Although the presence of this transgressive character in some F1 individuals did not alter the overall phenotype it could be of some importance for pollinators and potentially influence pollination behaviour (see Chapter 6.1, p 112 for a discussion of the role of transgressive segregation in hybridization).

## 5.4.3 Pollen fertility

Controlled F1s showed significantly lower pollen fertility (70.0%) than individuals from pure *G. urbanum* (94.3%) and *G. rivale* (90.3%) populations. However, F1s between *G. urbanum* x *G. rivale* have on average much higher pollen fertility compared to F1s in other species such as *Helianthus* (6%) (Heiser, 1947), *Gilia* (2%) (Grant, 1966), *Mimulus* (30%) (Ramsey *et al.*, 2003) and *Armeria* (40 to 60%) (Tauleigne-Gomes & Lefebvre, 2008). The lowest fertility observed in controlled F1s from this study was one individual with just under 30%, while the rest showed fertility rates between 55% and c. 80%. This suggests that the contribution of fertile pollen from F1 hybrids to the pollen pool in a hybrid zone could be relatively high and be a significant source for the production of later generation hybrid individuals.

## 5.4.4 Chloroplast inheritance

A survey of 18 controlled F1 offspring and their parents which differed in their chloroplast haplotype revealed that the maternal banding pattern was present in all offspring samples. Chloroplast inheritance is a complex phenomenon but is believed to be strictly maternal in most angiosperm species (Reboud & Zeyl, 1994), however numerous exceptions have been discovered (Milligan, 1992; Yang *et al.*, 2000; McCauley *et al.*, 2005; Ellis *et al.*, 2008). Milligan (1992) pointed out that many studies rarely use big enough sample sizes to be able to distinguish between alternative modes of organelle inheritance such as strict maternal inheritance, in which all progeny contain organelles derived from the same parent and paternal or biparental inheritance, in which at least some progeny contain organelles derived from different parents.

The probability to falsely accept the strictly maternal inheritance hypothesis is given by the equation $Pr = (1 - P)^N$, where P is the probability of paternal transmission and N the number of offspring used (Milligan, 1992). The sample size of N=18 individuals used for this analysis is only likely to detect paternal inheritance at the Pr=5% level if the probability of paternal inheritance P is larger than 15%. The results from this study therefore suggest that chloroplast inheritance is predominantly maternal in the two *Geum* species but can not exclude occasional paternal inheritance if P≤15%.

## 5.4.5 AFLP marker inheritance patterns

Marker inheritance patterns were investigated using 16 offspring from one controlled F1 *G. urbanum* x *G. rivale* and 18 offspring from one *G. rivale* x *G. urbanum* cross. More than 75% of the polymorphic markers in both crosses showed segregation ratios expected under Mendelian inheritance ($p<0.05$). *G. rivale* and *G. urbanum* are hexaploid (genome consisting of three sub-genomes, e. g. A, B and C) and it is therefore likely that some polymorphic markers will not segregate in a Mendelian fashion because not all loci on the three sub-genomes will have diverged enough to allow for Mendelian segregation. The ratio of segregation for each locus will depend on how many of the three sub-genomes have the same specific locus. For Mendelian segregation to occur the locus must be present on only one sub-genome (Figure 5.6) whereas different ratios occur if the locus is present on two or three sub-genomes.

| Sub-Genome A | Sub-Genome B | Sub-Genome C |
|---|---|---|
| AAAA**LOCUSX1**AAAA | BBBB**BBBB**BBBB | CCCC**CCCC**CCCC |
| AAAA**LOCUSX0**AAAA | BBBB**BBBB**BBBB | CCCC**CCCC**CCCC |

**Figure 5.6.** Hypothetical scenario for Mendelian segregation in a hexaploid plant. The whole plant genome consists of three sub-genomes A, B and C. Sub-genome A has a locus X which is heterozygous (X1 and X0). This locus is absent in sub-genomes B and C (through divergence) and segregation is therefore expected to be Mendelian (1:1 ratio of X1:X0) as the whole genome of the hexaploid plant is comparable to a diploid genome in respect to locus X.

Figure 5.7 highlights a particular scenario where all sub-genomes have the same heterozygous locus X, equally contributing gametes. The eight possible gamete combinations for locus X are X1X1X1, X1X1X0, X1X0X0, X0X1X1, X0X0X1, X0X0X0, X0X1X0, X1X0X1. If this plant crosses with another plant that is homozygous for the recessive allele for locus X (X0X0X0) then only one cross out of the eight possible (12.5%) will be homozygous for the null allele X0 and therefore be absent in an AFLP electropherogram.

This can explain why not all polymorphic markers show a 1:1 segregation, although there are many more possible scenarios which can lead to Non-Mendelian segregation ratios. However, the high proportion of markers segregating in a Mendelian way suggests that hexaploid *G. urbanum* and *G. rivale* behave largely as functional diploids.

| Sub-Genome A | Sub-Genome B | Sub-Genome C |
|---|---|---|
| AAAA**LOCUSX1**AAAA | BBBB**LOCUSX1**BBBB | CCCC**LOCUSX1**CCCC |
| AAAA**LOCUSX0**AAAA | BBBB**LOCUSX0**BBBB | CCCC**LOCUSX0**CCCC |

**Figure 5.7.** One of the possible hypothetical scenarios for Non-Mendelian segregation in a hexaploid plant. The whole plant genome consists of three sub-genomes A, B and C. Sub-genome A, B and C have the locus X (no divergence) which is heterozygous (X1 and X0). All three sub-genomes contribute gametes in respect to this locus.

A large proportion of the markers which only occurred in *G. urbanum* parents (79.1%) did not segregate compared to 42.6% of non segregating markers which only occurred in *G. rivale* parents. Similarly, the percentage of non segregating species-specific *G. urbanum* markers was higher (86.7%) than the percentage of non segregating species-specific *G. rivale* markers (75.0%). This is in line with the predominant breeding system of the two species. *G. urbanum* is mostly selfing which can lead to a rapid loss of heterozygosity. If the

heterozygous loci are not linked then for example an individual heterozygous for 100 loci will be 98% homozygous for these loci if 12 generations of complete selfing occur (Richards, 1997). In contrast, *G. rivale* is mainly outcrossing and is therefore less likely to have large numbers of homozygous markers.

## 5.4.6 Species-specific markers and NEWHYBRIDS

A PCO plot of pure and F1 samples using 203 AFLP markers showed that controlled F1s clearly clustered between the parental species and did not overlap with either species. A total of 24 species-specific markers were detected comprising 15 diagnostic *G. urbanum* and 9 diagnostic *G. rivale* markers. Simulation studies carried out by Boecklen & Howard (1997) suggest that as few as four or five species-specific markers should provide enough power to broadly classify individuals into parental, F1 and simple backcross categories. However, 50–70 species-specific markers are needed for unequivocal identification of advanced generation hybrid genotypes (Boecklen & Howard, 1997). This suggests that 24 diagnostic markers are sufficient to reliably classify individuals from a hybrid zone into parental, F1 and backcross categories.

It was somewhat unexpected to find a relatively large percentage of species-specific markers segregating in *G. urbanum* (13.3%) and *G. rivale* (25.0%). However, segregating species-specific markers have also been reported by other studies (Caraway *et al.*, 2001; Van Droogenbroeck *et al.*, 2006). Theoretically species-specific markers are homozygous (non-segregating) as they must occur in all individuals of one species (and in none of the other species). The results imply that some populations can be heterozygous for at least some of the species-specific loci. The frequency of the non wild type allele is likely to be low because 30 individuals per species collected from various populations throughout Great Britain (Table 2.1, p 30) were used to establish the species-specific nature of the markers. It is important to take the possible heterozygous nature of apparent species-specific markers into account when using these markers to establish the hybrid nature of samples.

F1s are expected to inherit 100% of the species-specific markers of each parental species. Although only one out of 15 controlled F1 crosses was missing one species-specific *G. urbanum* marker, more than half the samples were missing two to three species-specific *G. rivale* markers (data not shown). The heterozygous nature of some species-specific markers and the fact that c. 20% of the markers deviate from strict Mendelian inheritance could potentially distort the group allocation of hybrid samples. However, the software NEWHYBRIDS has proven robust in this respect. All F1 samples with missing species-specific markers were allocated with a very high probability (> 99%) to their correct group and only

two (3.3%) out of 59 individuals were misclassified as F2s. These results suggest that the overall number of 203 AFLP markers used contained enough information to compensate for the deviations from theoretically expected species-specific marker frequencies and strict Mendelian inheritance.

## 5.4.7 Summary

Controlled F1 crosses between *G. urbanum* and *G. rivale* were successful in both directions but germination rate was significantly lower (31%) in *G. rivale* x *G. urbanum* compared to *G. urbanum* x *G. rivale* crosses (83%). Analyzing the crosses with molecular and morphological tools revealed that F1s are intermediate between their parents with two thirds of the nine measured phenotypic traits being additively inherited. Pollen fertility was significantly lower (70%) in F1s compared to parental values (>90%) but survival rate of F1s was very high (c. 98%). One transgressive trait, brown stamens with indumentum, was discovered in 7% of the F1s. AFLP analysis revealed that about 80% of polymorphic markers were inherited in a Mendelian fashion and that 13% and 25% of species-specific *G. urbanum* and *G. rivale* markers were segregating, respectively. However, this did not compromise the predictive power of the software NEWHYBRIDS as about 97% of the controlled F1 crosses were allocated with a very high probability (>99%) to their correct category.

# 6. GENETIC COMPOSITION OF A HYBRID SWARM BETWEEN *GEUM URBANUM* AND *G. RIVALE*

## Abstract

The morphological and genetic structure of a hybrid swarm between *G. urbanum* and *G. rivale* located in Edinburgh/Scotland was analysed. 34 samples from the hybrid swarm covering the total morphological spectrum present, 39 and 32 samples from pure *G. urbanum* and *G. rivale* populations, respectively, and 32 controlled F1 crosses were measured for nine morphological traits in a common environment. To study the genetic structure of the hybrid zone 34 individuals from the hybrid zone were screened for 203 AFLP markers. 12 samples per species collected from various pure populations throughout Britain and 15 controlled F1s were included in the analysis for reference purposes. AFLP analysis of the hybrid zone showed that four distinct genetic groups can be distinguished (1) a parental *G. urbanum* like group (2) an F1 group (3) a Backcross Rivale group and (4) a parental *G. rivale* like group. No F2 or Backcross Urbanum individuals were detected. Morphological data were highly correlated with genetic data (Spearman Rank Order Correlation = 0.915, $p<0.0001$) but less informative as the Backcross Rivale samples could not be separated from *G. rivale* individuals in a Principal Component Analysis (PCA) using nine morphological traits. There was evidence for transgressive segregation (broader sepals than parental species) in the Backcross Rivale group. A transgressive character, brown stamens with indumentum, was found in the hybrid zone. This trait was not restricted to hybrid individuals but occurred in all four groups, although it was very rare in *G. urbanum* like samples. This suggests that pure looking *G. rivale* individuals with this character may represent advanced backcrosses.

# 6.1 Introduction

Hybridization (apart from polyploid hybridization) has traditionally been regarded as either an evolutionary dead end or the result of a breakdown in reproductive isolation caused by disturbance, often related to human activities (Mayr, 1963). These views, largely based on animal studies, have recently been revised considerably, and it is now recognized that hybridization is an important force in angiosperm evolution (Arnold, 1992, 2006).

The frequency of hybridization in plants has been estimated by Ellstrand *et al.* (1996) using data from five biosystematic floras and ranges from about 22% for the British flora to 5.8% for the intermountain flora of North America, with an average of 11% over all five floras. This suggests that there are a total of more than 30,000 hybrid combinations among the 300,000 plant species described worldwide. In a review on hybridization in plants, Rieseberg (1997) indicates that this figure could represent a substantial underestimate as many hybrids might have gone undetected due to inadequate systematic attention to certain groups in various floras. Natural hybrids are found to be unevenly distributed taxonomically and only 16% to 34% of plant families and 6% to 16% of genera have one or more reported hybrids (Ellstrand *et al.*, 1996; Rieseberg, 1997). It seems that hybridization is concentrated in a small fraction of families and genera which usually show life-history traits such as perennial habit, outcrossing breeding systems and asexual reproductive modes that allow stabilization of hybrids.

Hybridization often results in the formation of hybrid swarms which are spatially limited areas where phenotypically pure looking parental species grow together with a mixture of more or less intermediate hybrids. Hybrid swarms are either the result of secondary contact of two formerly allopatric taxa or of an ongoing speciation process in sympatry or parapatry (Hewitt, 1988). They can be of considerable evolutionary importance as the long term consequences of hybridization may have four major outcomes.

## 6.1.1 Extinction

Hybridization can lead to extinction by demographic swamping and genetic assimilation of one species by another (Levin *et al.*, 1996; Rhymer & Simberloff, 1996; Wolf *et al.*, 2001). The growth rate of a less common species is slowed down by the production of hybrid seeds at the expense of conspecific seeds. This will also lead to a relatively higher percentage of hybrid seeds produced by the rarer species compared to the more common congener ('minority disadvantage'). The minority disadvantage does not depend on the viability of

hybrid seeds as the production of viable and not viable seeds will reduce the possible number of conspecific seeds (Levin, 2000). If the crossing barriers between two species are weak then this may result in a significant reduction of conspecific seeds of the rarer species followed by a decline in population size (Heiser, 1979; Warwick et al., 1989). Successfully established hybrids will further decrease the relative population size of the less common species and contribute to its genetic assimilation (Rieseberg *et al.*, 1989; Rieseberg & Gerber, 1995). This process of extinction by hybridization is a particular threat to island floras as previously geographically separated but interfertile species often have been brought into contact by human activities (Levin et al., 1996).

## 6.1.2 Introgression

Hybridization may lead to introgression (Anderson, 1949; Howard *et al.*, 1997; Wolfe *et al.*, 1998a; Ellstrand *et al.*, 1999; Milne & Abbott, 2000; Petit *et al.*, 2003). Introgression was first defined by Anderson and Hubricht (1938) as the exchange of genetic material facilitated by repeated backcrossing of fertile hybrids to the parental species and is now commonly used for the permanent incorporation of genes from one set of differentiated populations (or species) into another (Rieseberg & Wendel, 1993). Rieseberg and Wendel (1993) list 165 proposed cases of introgression but since then many more (molecular) studies of hybridization have highlighted its widespread occurrence in plants (Howard *et al.*, 1997; Wolfe *et al.*, 1998a; Ellstrand *et al.*, 1999; Milne & Abbott, 2000; Petit *et al.*, 2003).

Introgression can help hybrids to colonise new habitats (Abbott, 1992) by transferring adaptive traits from one species to the other (Lewontin & Birch, 1966). Klier *et al.* (1991) studied introgression in two orchid species, *Cypripedium pubescens* which mainly grows in woodlands and *C. candidum* which survives only on a small number of prairie remnants in the United States. Their data demonstrated the existence of bidirectional gene flow wherever the two species are sympatric and hybridize. Interestingly, *C. pubescens* has been previously noted to occasionally occur on prairies in Iowa and allozyme and morphological data as well as ecological considerations suggested that the prairie ecotype of *C. pubescens* may have arisen through introgression of genetic material from *C. candidum* (Klier *et al.*, 1991). This indicates that *C. pubescens* plants growing in prairie habitat are actually later-generation backcrosses to the prairie-dwelling *C. candidum* and may represent an example of the transfer of an adaptation (or suite of adaptations) which allowed these plants to survive in this new habitat. In a similar study using evidence from rDNA and cpDNA data, Milne & Abbott (2000) detected that naturalized British hybrids between *Rhododendron. ponticum* and *R. catawbiense* were significantly more abundant in Britain's coldest region, eastern Scotland, than elsewhere. This could indicate that introgression from the cold tolerant North

American *R. catawbiense* conferred improved cold tolerance to *R. ponticum*, a species from the warmer Iberian Peninsula, thus enabling the hybrids to survive in colder climates (Milne & Abbott, 2000).

Introgression may also increase genetic diversity (Percy & Wendel, 1990), reduce the strength of reproductive barriers (Stebbins & Daly, 1961; Meyn & Emboden, 1987) and lead to genetic assimilation or extinction in the long term (Rieseberg *et al.*, 1989; Rieseberg & Gerber, 1995).

### 6.1.3 Speciation

Hybridization may lead to an increase in biodiversity due to the production of novel hybrid lineages which can develop into a new species (Stebbins, 1957c; Grant, 1971; Arnold, 1993, 1994; Lowe & Abbott, 1996; Rieseberg, 1997; Ainouche *et al.*, 2003; Guo *et al.*, 2006). This can happen either via allopolyploid or homoploid speciation. Allopolyploidy is the spontaneous doubling of chromosome numbers in organisms such as hybrids, which consist of two or more different genomes (in contrast to autopolyploidy where the doubling occurs within populations of the same species). The hybrid is then at least partially reproductively isolated from the parents but can also reproduce sexually as meiotic alignment problems have been removed by the doubling of chromosomes. The relevance of polyploidy (auto- and allopolyploidy) in plant speciation has long been recognized (Müntzing 1936; Grant, 1971) and recent studies (Lowe & Abbott, 1996; Ainouche *et al.*, 2003; Guo *et al.*, 2006) and reviews (Ramsey & Schemske, 1998; Soltis *et al.*, 2003) have further highlighted its evolutionary importance (Soltis, 2005).

In contrast, homoploid or recombinational speciation is the process of forming and establishing new recombinational hybrid lineages which have the same ploidy level as their parents. Reproductive isolation of the hybrids from their parents can be achieved if the parental species differ by chromosomal rearrangements such as reciprocal translocations, fusion or fission of chromosomes, the duplication or deletion of a chromosomal segment and inversions (Stebbins, 1957c; Grant, 1958). F1 and later hybrid individuals which are heterozygous for these rearrangements are partially sterile and/or inviable. However, selfing or interbreeding among F1 individuals may give rise to some novel, chromosomally balanced genotypes that have restored fertility. Because crosses with the parental species will produce heterozygous progeny, these novel genotypes are at least partially reproductively isolated from the parental species, and if they can persist as a lineage, may constitute a novel hybrid species (Buerkle *et al.*, 2000).

Alternatively, isolation of the novel hybrid lineage from homogenizing gene flow can also be achieved through ecological isolation (occupation of habitats which are different to parental ones) or through a shift to a selfing mating system, thus reducing gene flow between individuals. This is also supported by computer simulations which suggest that a chromosomal sterility barrier is unlikely to lead to hybrid speciation without additional isolation mechanisms (McCarthy *et al.*, 1995; Buerkle *et al.*, 2000). Buerkle *et al.* (2000) highlighted that strong ecological and geographical barriers to gene flow between parent species and novel hybrid lineages were necessary in their model for hybrid speciation to occur in about 20% of the simulation runs. Additionally, both simulation models confirmed that homoploid speciation was more likely to occur under conditions of inbreeding (McCarthy *et al.*, 1995; Buerkle *et al.*, 2000). In a recent study Gross and Rieseberg (2005) reviewed the literature to evaluate the importance of ecological selection in the circa 20 well documented cases of homoploid speciation. Essentially many, if not all, homoploid hybrid species are ecologically divergent from their parent's ecological requirements and in some cases have colonised extreme habitats (Gross & Rieseberg, 2005).

### 6.1.4 Stalemate

Hybridization can lead to a stalemate situation with no (evolutionary) consequences. This outcome can be explained by the Tension Zone Model (Barton & Hewitt, 1985). In this model hybrid zones are viewed as clines which are maintained by a balance between dispersal and selection against hybrids. Clines are regions in which a variety of characters change gradually resulting in a cluster of parallel zones with different gene frequencies, grading from a maximum to a minimum between these zones. The Tension Zone model assumes that hybrids are less fit than their parental genotypes independent of the habitat they grow in because of genomic incompatibilities (Barton & Hewitt, 1985). This leads to a 'tension zone' between the conflicting effects of dispersal of parental forms on the one hand and the creation and selection against less fit hybrids on the other, resulting in an equilibrium between selection against hybrids and their constant production from inter-parental crosses. The outcome could be compared to a 'stalemate' situation with little evolutionary consequences as there is a very restricted potential for introgression of neutral genes between the parental species and because no viable new hybrid lineages are produced.

Summarising, the long term consequences of hybridization can decrease (extinction of one species) or increase (speciation via a novel and viable hybrid lineage) biodiversity, broaden the genetic base of the parental taxa (introgression) but also be an ephemeral event with little evolutionary impact. However, the outcome can not be predicted.

### 6.1.5 Hybridization between selfing and outcrossing taxa

It has recently been suggested that hybridization between selfing and outcrossing taxa may have been a keyfactor in the production of biodiversity and taxonomic complexity in the widely hybridizing genera of *Euphrasia* (French *et al*., 2008) and *Epipactis* (Hollingsworth *et al*., 2007). The study by Hollingsworth *et al*. (2007) revealed that mixed populations of the autogamous taxa *E. phyllanthes* and *E. dunensis* showed no evidence of gene exchange. However, if *E. dunensis* occurs sympatrically with the allogamous taxa *E. helleborine* and *E. youngiana* then *E. dunensis* is genetically variable and heterozygous for the same alleles found in local populations of *E. helleborine* and *E. youngiana* (Hollingsworth *et al*., 2007). Plants with intermediate morphology were also found at these sites. In contrast, all other populations of *E. dunensis* studied were homozygous and uniform. In the case of *E. phyllanthes* plants of a site growing sympatrically with *E. helleborine* and *E. youngiana* possessed an allele that was also present in the other two species but absent from the other 26 studied *E. phyllanthes* populations (Hollingsworth *et al*., 2007).

In a study on British *Euphrasia* species using AFLP markers French *et al.* (2008) found that widespread inbreeding taxa like *E. confusa*, *E. tetraquetra*, *E. foulaensis*, *E. micrantha*, *E. scottica* and *E. ostenfeldii* were clearly separated from outbreeding taxa such as *E. arctica* and *E. nemorosa*. However, individuals from the local endemic inbreeding taxa *E. marshallii*, *E. campbelliae* and *'E. fharaidensis'* scattered widely across the PCO plot and overlapped with individuals from the outcrossing *E. arctica* and *E. nemorosa* complex. This pattern could be explained by recurrent hybridization between the widespread outcrossing *E. arctica/nemorosa* complex and a widespread inbreeding species with the local endemic taxa representing a collection of more or less inbreeding segregants of this process (French *et al*., 2008). On the basis of morphological observations it has also been suggested that the diploid endemic taxa *E. vigursii* and *E. rivularis* have originated due to introgression of genes from the inbreeding tetraploid *E. micrantha* into the outcrossing diploid *E. officinalis* (Yeo, 1968). A single triploid individual was discovered in a mixed population of *E. micrantha* and *E. officinalis* indicating that hybridization between these taxa does occur in the wild (Yeo, 1954). French *et al.* (2008) found that occasional gene flow between the ploidy levels occurs which could lead to the origin of novel taxa.

Breeding system transitions from outcrossing to selfing might have therefore contributed to the recurrent evolution of autogamous lines and subsequent hybridization between outcrossing and inbreeding taxa could have been the cause for the high intrageneric diversity in genera such as *Epipactis* and *Euphrasia* (Squirrell *et al*., 2002; Hollingsworth *et al*., 2007; French *et al*., 2008). A selfing mating system leads to increased homozygosity and morphological characters encoded by recessive alleles are then expressed and can be

selected for. Selfing can therefore lead to the fixation of advantageous recessive mutations which might produce morphological differences between selfers and their outbreeding progenitors (Charlesworth, 1992; Levin, 2000; Hollingsworth *et al.*, 2007).

The hypothesis that hybridization between inbreeding and outcrossing plant taxa might generate taxonomic complexity is particularly plausible if seen in the light of transgressive segregation. Transgressive segregation is a phenomenon where certain hybrids display extreme phenotypes which fall outside the natural range of variation observed in the parents. These individuals exceed the parental phenotypic range in a positive or negative direction meaning that some characters in individual hybrids are much larger or smaller than in the parents. A recent literature survey on transgressive traits in hybrid populations by Rieseberg *et al.* (1999) showed that transgressive segregation is a very common phenomenon. 91% of the surveyed studies reported at least one transgressive trait and 44% of the traits reviewed were transgressive. A later study revealed that the genetic architecture necessary for transgressive segregation appears to be a common feature in plants (Rieseberg *et al.*, 2003c). It therefore does not come as a surprise that transgressive segregation can largely explain the important role hybridization plays in plant evolution (Rieseberg *et al.*, 1999; Bell & Travis, 2005). If transgressive segregation alters characters which are under strong selection, like flower shape or colour, then transgressive hybrids can diverge sufficiently from their progenitors and establish lineages which are able to coexist sympatrically with them. This process might be supported if functionally related traits do not segregate independently and are conserved together (Bell & Travis, 2005).

The chances of survival and perseverance of novel and transgressive hybrid lineages have often been associated with the availability of suitable ecological habitats (Anderson, 1948; Rieseberg, 1991; Buerkle *et al.*, 2000). Additionally, especially in homoploid speciation, where hybrids are at the same ploidy level as their parents, a selfing mating system decreases the chances of gene exchange with the parents (Grant, 1958; McCarthy *et al.*, 1995; Buerkle *et al.*, 2000). The rationale behind these assumptions is that the likelihood for the persistence of a new hybrid lineage will depend on how well it is insulated from homogenizing gene flow. Hybrid species are able to form in sympatry with parental species but then occupy divergent niches giving rise to ecological isolation. Dorken & Pannell (2007) suggested that ecological gradients in habitat disturbance could be particularly important in contact zones in which the interacting lineages differ in their sexual system (selfing versus outcrossing) because selfing promotes the colonization of open patches as only one individual is required to found a new population (Baker, 1955; Pannell & Barrett, 1998). Additionally, a selfing mating system reduces gene flow with the parental population ensuring that the genetic and phenotypic integrity of the new lineage is maintained. Selfing hybrid lineages are most likely to occur where hybridization takes place between closely related inbreeding and outcrossing plant

taxa. The outcrossing parent maximizes the chances of initial hybrid formation while genes for selfing contributed by the inbreeding parent ensure that a proportion of later generation hybrids will be selfing.

Despite the potentially important evolutionary consequences of hybridization between selfing and outcrossing plant taxa, we know very little about it. In a study carried out in naturally occurring sympatric populations of outcrossing *Mimulus guttatus* and selfing *M. nasutus*, Martin & Willis (2007) found that prezygotic barriers acted asymmetrically to completely prevent the formation of F1 hybrids among seeds produced by outcrossing *M. guttatus*. F1 hybrid production among seeds produced by selfing *M. nasutus* was reduced to about 1%. Postzygotic isolation contributed very little to the total isolation between the two species. Despite the strong total reproductive isolation (0.99 to 1.0) between *M. guttatus* and *M. nasutus*, substantial opportunity for interspecific gene flow still exists. More than 1 in 1000 seeds are highly fit *M. nasutus* F1 individuals which are likely to backcross to *M. guttatus*. Martin & Willis (2007) think it plausible that the introgression rate from *M. nasutus* to *M. guttatus* is on the order of 0.001 for neutral regions not tightly linked to selected sites in the genome. That level of introgressive gene flow is high in population genetic terms (Slatkin, 1987) and is expected to cause neutral *M. nasutus* alleles to increase in frequency in *M. guttatus*. This was supported by a recent study which reported introgression at the molecular level in sympatric populations of *M. guttatus* and *M. nasutus* (Sweigart & Willis, 2003). Martin & Willis (2007) therefore conclude that there is greater potential for introgression from selfing *M. nasutus* to outcrossing *M. guttatus* than for the reverse direction.

To better understand the processes which take place when inbreeding meets outcrossing, further investigations using different study systems are clearly needed. In previous chapters it has been shown that *G. urbanum* and *G. rivale* are mostly selfing and outcrossing, respectively (Chapter 3, p 47) and that they are morphologically and genetically clearly distinct species (Chapter 2, p 23). Controlled F1 crosses between *G. urbanum* and *G. rivale* revealed that survival rates were very high and that pollen fertility, although lower compared to the parental species, was substantial compared to other species (Chapter 5, p 91). The two species occupy the same broad geographic area throughout most of Europe and Asia (Figure 1.9, p 19). They are partially separated by their habitat requirements, pollinator preferences and flowering time but form hybrids wherever the two species meet. *G. urbanum* and *G. rivale* therefore are an ideal study system for the purpose discussed above. The objective of this chapter is to investigate the current patterns of hybridization between selfing *G. urbanum* and outcrossing *G. rivale* in a long lived hybrid swarm in Scotland. My aim is to document the immediate consequences of this hybridization event and to predict the likely future long term evolutionary consequences of this process.

The hybrid swarm was analysed using morphological and genetic tools in order to give a comprehensive description of its composition. Phenotypic variation of pure and hybrid individuals was studied under uniform greenhouse and field conditions and the genetic structure was established using AFLP markers. The flowering times of the different morphs in the hybrid zone were recorded to obtain insight into the ease with which they are able to interbreed.

Controlled F1 crosses and pure populations were included in the analyses of the hybrid swarm to address the following questions:

1. What is the phenotypic and genetic structure of the hybrid swarm? In particular, do hybrids form a morphologic and genetic continuum between the parents or do they fall into distinct classes?
2. What is the pollen fertility in the hybrid swarm?
3. How much do parental species and hybrids overlap in their flowering time?
4. Are hybrids able to auto-pollinate (i. e. produce seeds without being pollinated by other individuals)?
5. Is there evidence for transgressive segregation in the hybrid swarm?

# 6.2 Materials and Methods

## 6.2.1 Sampling material

### *Study site*

The study site of the hybrid population between *G. urbanum* and *G. rivale* is located 11 km SW of the city centre of Edinburgh (Scotland) near the village of Balerno (Figure 6.1). It stretches about 400 m along the Water of Leith walkway (extending about 4 m on average to either side of the path) which runs alongside a river of the same name at an elevation of 150 m ($55^0$ 53' 28" N/$03^0$ 19' 26" W to $55^0$ 53' 24" N/$03^0$ 19' 49" W). The walkway passes through semi-natural vegetation on disturbed ground which forms a more or less dense canopy on either side of the path apart from a partially shaded to sunny 100 m stretch of open ground at the western end of the study site. The population comprises more than 1000 individuals presenting a wide range of parental and intermediate phenotypes. There are no records of the exact age of the studied hybrid swarm but it is clear that *G. urbanum* and *G. rivale* have been hybridizing for a long time in this area. Balfour (1863) mentioned the occurrence of hybrids between the two taxa in Currie, a place about 1 km away from the hybrid swarm used for this study. In the database for plant records of the Royal Botanic Garden Edinburgh, one record refers to hybrids between *G. urbanum* and *G. rivale* in Currie in June 1955. Observations from the warden responsible for the stretch of the Water of Leith walkway which passes through the hybrid swarm suggest a minimum age of 20 years (Church, pers. comm.).

### *Hybrid samples*

54 individuals, representing the total morphological spectrum present in the hybrid population, were dug up in July 2006, transferred to an unheated greenhouse and treated as described in Chapter 2.2.1 (p 28).

### *Pure populations*

To put the morphological and genetic variation found in the hybrid swarm into context, individuals from ten pure populations per species were also included as reference points in the morphological (39 *G. urbanum* and 32 *G. rivale* samples, Table 6.1) and genetic (12 individuals per species, Table 6.1) analyses of the hybrid swarm.

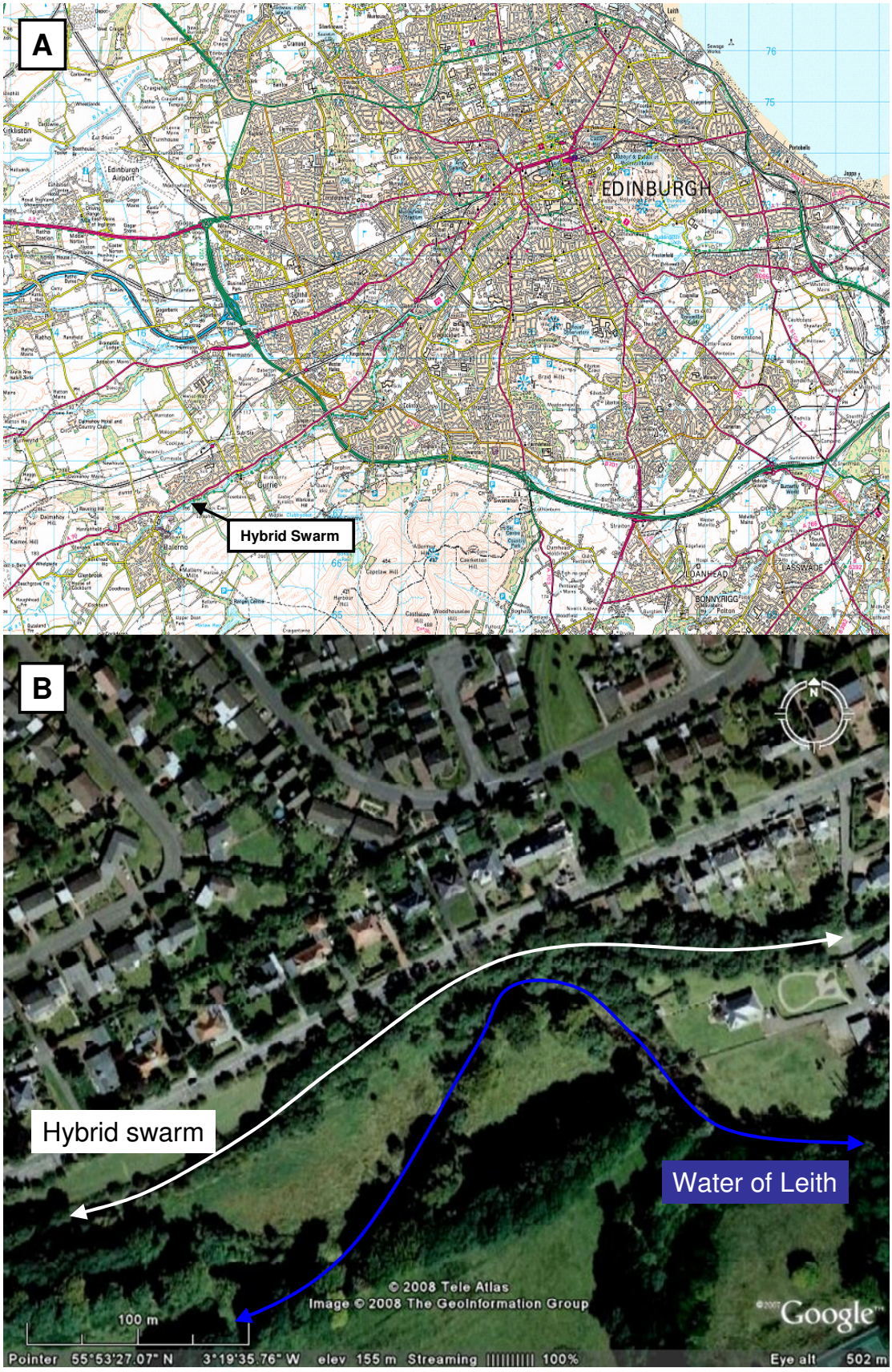



**Figure 6.1. A.** Location of the studied hybrid swarm near Balerno just outside Edinburgh (Scotland) **B.** Aerial photo of the hybrid swarm along the Water of Leith. The two white arrow heads mark the end points of the hybrid swarm used for this project.

| Species | Country | Location | Altitude | Coordinates | Morphology | | AFLP | Pollen |
|---|---|---|---|---|---|---|---|---|
| | | | | | Field | Greenhouse | | |
| **Geum urbanum** | **Scotland** | Edinburgh, Braid Hills | 170 m | 55$^0$ 56′ N/ 03$^0$ 12' W | 8 | 11 | 2 | 11 |
| | | Edinburgh, Colinton Ruderal | 120 m | 55$^0$ 55' N/ 03$^0$ 14' W | | 5 | | 2 |
| | | Edinburgh, Colinton Wood | 120 m | 55$^0$ 55' N/ 03$^0$ 14' W | | 13 | 1 | 4 |
| | | Edinburgh, University | 60 m | 55$^0$ 55' N/ 03$^0$ 10' W | 4 | | | |
| | | Edinburgh, Craigmillar | 70 m | 55$^0$ 56′ N/ 03$^0$ 09′ W | 3 | 10 | 2 | 5 |
| | | Edinburgh, Gorebridge | 90 m | 55$^0$ 51′ N/ 03$^0$ 04′ W | | | 1 | |
| | **England** | Yorkshire, Swainby | 200 m | 54$^0$ 24 N'/ 01$^0$ 16' W | | | 1 | |
| | | Gloucestershire, Cotswolds | 200 m | 52$^0$ 03' N/ 01$^0$ 48' W | | | 1 | |
| | | Gloucestershire, Cotswolds | 240 m | 51$^0$ 46' N/ 02$^0$ 08' W | | | 1 | |
| | | London, Kew | 10 m | 51$^0$ 28' N/ 00$^0$ 16' W | | | 1 | |
| | | Devon, Dawlish | 20 m | 50$^0$ 34′ N/ 03$^0$ 28' W | | | 1 | |
| | | Cornwall, Truro | 50 m | 50$^0$ 12′ N/ 05$^0$ 03' W | | | 1 | |
| **Geum rivale** | **Scotland** | Orkney, Berridale Hoy | 40 m | 58$^0$ 53′ N/ 03$^0$ 23' W | | | 1 | |
| | | Ullapool, Inchnadamph | 180 m | 58$^0$ 09′ N/ 04$^0$ 56' W | | | 1 | |
| | | Ullapool, Knockan Craig | 250 m | 58$^0$ 02′ N/ 05$^0$ 03' W | | | 1 | |
| | | Thurso, Betty Hill | 20 m | 58$^0$ 31′ N/ 04$^0$ 14' W | | | 1 | |
| | | Stirling, Ben Lawers | 600 m | 56$^0$ 32′ N/ 04$^0$ 17' W | | 13 | 1 | 13 |
| | | Edinburgh, Pentland Hills, Harlaw | 230 m | 55$^0$ 53′ N/ 03$^0$ 18' W | 4 | | | |
| | | Edinburgh, Pentland Hills, Carlops | 300 m | 55$^0$ 48′ N/ 03$^0$ 21' W | | 2 | | |
| | | Edinburgh, Pentland Hills, Bonaly | 210 m | 55$^0$ 53′ N/ 03$^0$ 16' W | 4 | 1 | 1 | 1 |
| | | Scottish Borders, Moorfoot Hills | 300 m | 55$^0$ 44′ N/ 03$^0$ 01' W | 2 | 16 | 1 | 8 |
| | | Scottish Borders, Moffat Dale | 300 m | 55$^0$ 23′ N/ 03$^0$ 19' W | | | 3 | 1 |
| | **England** | Yorkshire, Cross Fell | 800 m | 54$^0$ 42′ N/ 02$^0$ 29' W | | | 1 | 1 |
| | | Yorkshire, Teesdale | 260 m | 54$^0$ 39′ N/ 02$^0$ 09' W | | | 1 | |
| **Hybrid swarm** | **Scotland** | Edinburgh, Balerno | 150 m | 55$^0$ 53′ N/ 03$^0$ 19′ W | 26 | 34 | 54 | 34 |

**Table 6.1.** Location of the hybrid zone and individuals collected from pure *Geum urbanum* and *G. rivale* populations used as reference samples in various analyses (Morphology, AFLP, Pollen fertility) with number of individuals used. Field: morphological measurements were taken in the field, Greenhouse: morphological measurements were taken in a common environment.

*Controlled F1 crosses*

To provide a reference point to which hybrids found in the hybrid swarm could be compared, synthetic F1s from controlled crosses conducted in a greenhouse between pure *G. urbanum* and *G. rivale* populations (Chapter 5, p 94) were included in the morphological (n=41) and genetic (n=15) analyses of the hybrid zone.

## 6.2.2 Molecular variation

DNA was extracted from either silica dried leaf material or fresh leaves as detailed in Chapter 2.2.5 (p 31).

*AFLP fingerprinting*

To establish the genetic structure of the hybrid population and to test whether individuals fall into genetically distinct categories, 34 samples from the hybrid zone were used for AFLP fingerprinting as detailed in Chapter 2.2.5 (p 31). To be able to put the genetic variation found into context, individuals from pure populations (n=12 per species, Table 6.1) and from controlled F1 crosses (n=15) were included in the assay.

After an initial survey established that eight primer combinations yielding 616 markers did not reveal more information than four primer combinations yielding 203 markers, the following combinations were chosen based on the number of species-specific markers produced and their ability to generate clear peaks: 6FAM-*Eco*R1 AAC/ *Mse*1 CTA, PET-*Eco*R1 AAT/ *Mse*1 CAGA, VIC-*Eco*R1 AGC/ *Mse*1 CTT, NED-*Eco*R1 ATC/ *Mse*1 CGA.

The four selected *Eco*R1 primers were labelled with four different dyes to allow multiplexing. Species-specific markers for the above mentioned primer pairs were identified from data presented in Chapter 2.2.5 (p 31) using 30 pure *G. urbanum* and 30 pure *G. rivale* samples collected from various British populations (Table 2.1, p 30).

*Chloroplast DNA (cpDNA)*

To compare the haplotype composition in the hybrid swarm with that in pure populations 34 hybrid individuals were surveyed for the two informative polymorphic cpDNA regions trnL-F (*Eco*R1) and trnC-D (*Mse*1) identified in Chapter 2.3.2 (p 38) using cpDNA-RFLPs. Haplotype frequencies for pure populations (based on 154 individuals from pure populations comprising 75 pure *G. urbanum* and 79 *G. rivale* samples) were taken from Table 2.5 (p 38). CpDNA-RFLPs were executed as described in Chapter 2.2.5 (p 31).

## 6.2.3 Non molecular variation

*Morphology*

To document the morphological variation found in the field in the hybrid swarm and in pure parental populations, five floral and two vegetative characters were measured on plants. The characters used were stipule, petal and sepal length and width (in cm) and claw of petal length (in mm). 26 individuals from the hybrid swarm, 15 *G. urbanum* individuals from three pure populations (Table 6.1) and ten *G. rivale* individuals from four pure populations were measured (Table 6.1).

To minimize the influence of environmentally induced variation on morphological measurements, measurements were also taken in a common environment using 34 individuals from the hybrid zone, 39 individuals from three pure *G. urbanum* and 32 individuals from four pure *G. rivale* populations (Table 6.1). All samples were dug up in 2006, grown in an unheated greenhouse following the procedure outlined in Chapter 2.2.1 (p 28) and measured the following year. Measurements from 32 controlled F1 crosses grown in a common garden plot were included in the analysis for reference purposes. In addition to the seven characters measured in the field, gynophore length and length of upper style (in mm) were also recorded using a stereo microscope equipped with a graticule.

*Pollen fertility*

Pollen fertility was measured on a total of 40 individuals from the genetic groups identified via AFLP fingerprinting in the hybrid swarm. All samples were grown in a common environment (unheated greenhouse) before anthers were removed for analysis in summer 2008. Pollen fertility tests were conducted following the procedure outlined in Chapter 5.2.2 (p 96).

*Auto-pollination*

To establish whether spontaneous self pollination is a possible means of seed production in the absence of pollinators, auto-pollination experiments were conducted as described in Chapter 3.2.2 (p 56) in a common environment (unheated greenhouse) on 39 individuals covering the whole phenotypic spectrum present in the hybrid swarm.

*Phenology*

To establish the flowering times of parental and hybrid individuals, 11 plots of variable sizes ranging from 20 to 52 $m^2$ were established within the hybrid swarm (Figure 6.2). The plots covered about 10% of the total hybrid swarm area (c. 330 $m^2$ of plots out of an estimated area of 3200 $m^2$ hybrid swarm). The hybrid swarm was visited weekly during the flowering period between May and July 2008. At each of the 12 recording dates the number of open flowers within each plot was estimated for three main phenotypic classes: *G. urbanum*, *G. rivale* and hybrid. To establish possible phenological differences between different classes of hybrids, the latter category was subdivided on the basis of morphology into F1, Backcross Rivale and Backcross Urbanum classes. Plants which could not be readily allocated to a class were classified as uncertain.

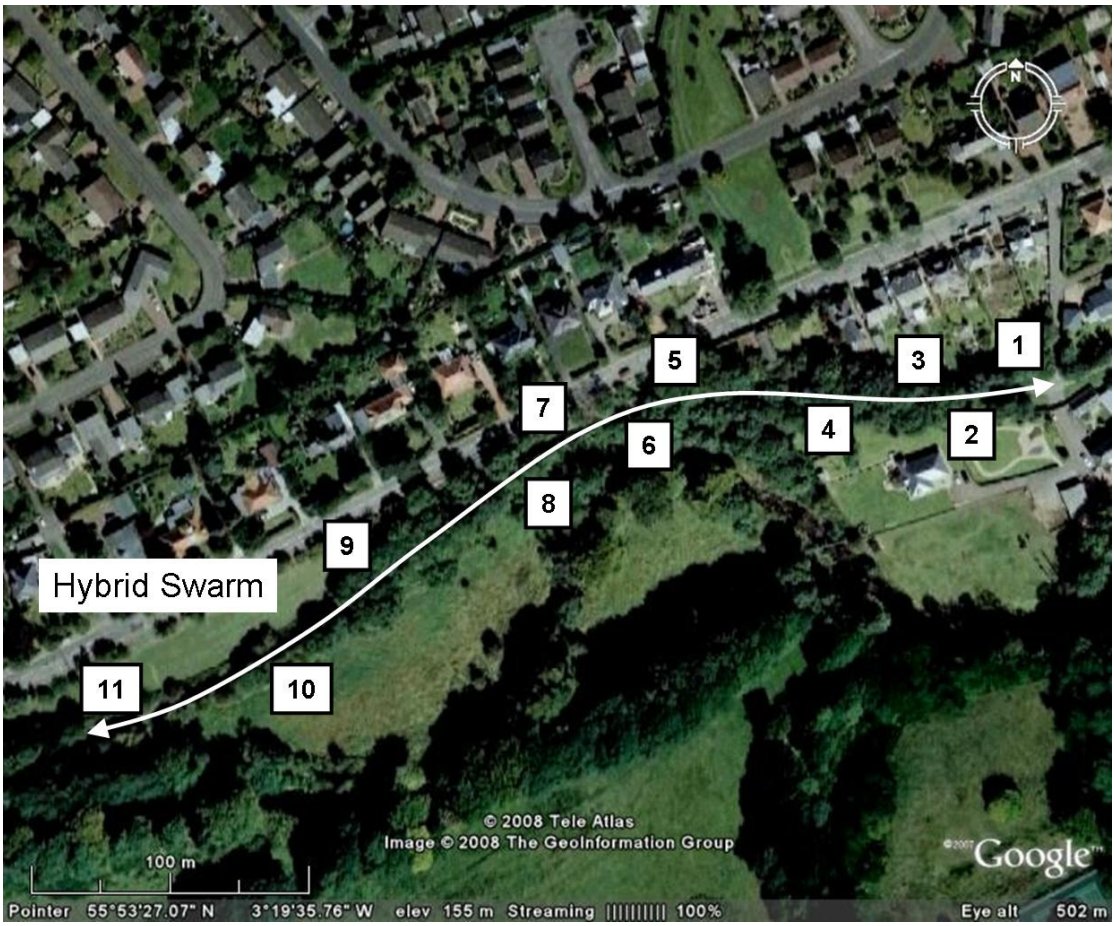



**Figure 6.2.** Location of 11 plots used for monitoring flowering times in the hybrid zone.

## 6.2.4 Analysis

*AFLP error rates*

AFLP runs were repeated three times for 24 samples from the hybrid zone to estimate error rates. Two of the runs were performed after using the same pre-amplification product for the selective PCR. The third run was performed using a pre-amplification product for the selective PCR from a completely new restriction-ligation process. However, the same source of DNA for the initial stage of the restriction process was used for all three runs. Error rates were estimated following Bonin *et al.* (2004) as the ratio between observed number of phenotypic differences and total number of phenotypic comparisons. One sample with no DNA (negative control) was incorporated in each run.

*Genetic difference*

To visualise the genetic structure of the study hybrid population a Principal Coordinates Analysis (PCO) of AFLP data from 54 samples collected in the hybrid zone was performed. To be able to put the genetic structure in the hybrid zone in context, 12 *G. urbanum* and 12 *G. rivale* samples from pure populations (Table 6.1) as well as 15 individuals from controlled F1 crosses were included in the analysis. The program PAST was used for the PCO analysis applying a Jaccard similarity index (Hammer *et al.*, 2006).

To establish whether individuals in the hybrid swarm fall into distinct genetic classes the program NEWHYBRIDS version 1.1 beta (Anderson & Thompson, 2002) was used with the default parameters for the six genotype class frequencies (pure parental, F1, F2 and backcross categories) and Jeffreys priors, no individual or allele frequency prior information was specified, and a burn-in phase of 100 000 steps and 100 000 sweeps. Individuals from pure populations were marked as pure according to the requirements of the program and run with individuals from the hybrid population in the same simulation.

*Morphology*

A principal component analysis (PCA) was performed for pure and hybrid individuals on field and greenhouse measurements using seven and nine morphological characters, respectively, all of which were identified as species-specific (Chapter 2.3.1, p 35). PCA is a multivariate method that reduces data dimensionality helping to understand the underlying structure in a data set with many variables. The program MINITAB 14.1 (Minitab Inc., 2003) was used to

conduct the PCA analysis. Differences in means of morphological traits between pure and hybrid groups were tested for significance with a Kruskal-Wallis test.

### *Pollen fertility and auto-pollination*

The mean group pollen fertility was established by averaging over all individual fertility counts in that group. Pollen fertility (percentage of fertile pollen produced per individual) and the ability to auto-pollinate (percentage of bagged flowers which auto-pollinated per individual) were arcsine transformed and differences between pure and hybrid groups were tested for significance using Analysis of Variance (ANOVA) plus a Tukey pairwise comparison test. Within group variability of pollen fertility was tested for significance with a $\chi^2$ test (pollen fertile/pollen infertile).

### *Correlation of morphological and genetic data set*

To establish the extent to which the morphological and genetic data set are correlated the Spearman Rank Order Correlation was calculated between the first PCA component and the first PCO axis.

# 6.3 RESULTS

## 6.3.1 Molecular data

*AFLP error rates*

AFLP runs were repeated three times for 24 individuals from the hybrid swarm to establish error rates. All samples were scored for the same 154 polymorphic loci per run resulting in 3696 phenotypic comparisons between two runs. The negative control never showed any amplification.

Error rates were lowest (2.6%) for runs using the same pre-amplification product and highest (5.1%) for runs using pre-amplification products from a different restriction-ligation process. The maximum individual error rates were about twice the overall error rate (Table 6.2).

Despite the range of error rates between runs and samples the effect on the outcome of the PCO analysis was minimal. A PCO plot using 24 individuals from the hybrid swarm revealed that all the samples from each of the three different repeat runs formed the same groupings (Figure 6.3).

| Different preamp | Mean error rate | No. of phenotypic differences | Total no. of comparisons | Minimum individual error rate | Maximum individual error rate |
|---|---|---|---|---|---|
| Run1 vs Run 3 | 4.1% | 151 | 3696 | 1.9% | 8.4% |
| Run2 vs Run3 | 5.1% | 187 | 3696 | 1.3% | 7.8% |
| **Same preamp** | | | | | |
| Run1 vs Run 2 | 2.6% | 95 | 3696 | 0% | 5.2% |

**Table 6.2.** Error rates of three replicate AFLP runs using 154 polymorphic loci and 24 samples. Run 1 and 2 were performed using the same pre-amplification product for the selective PCR. Run 3 was performed using a pre-amplification product for the selective PCR from a completely new restriction-ligation process. However, the same source of DNA for the initial stage of the restriction process was used for all three runs.

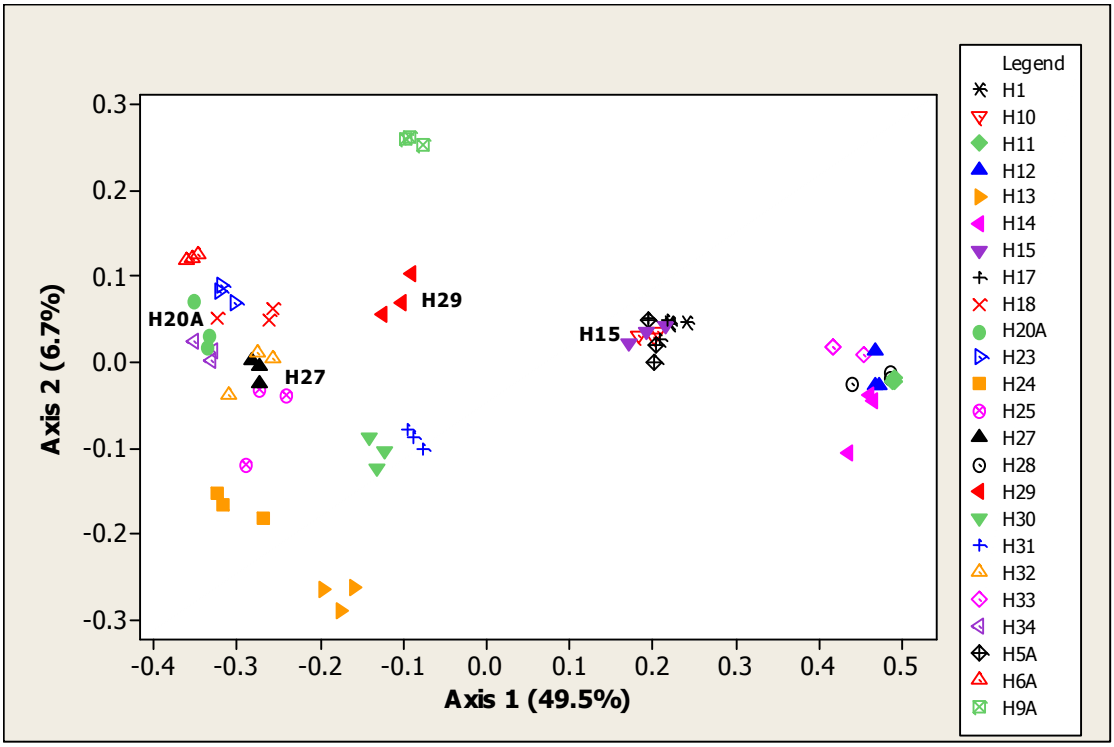



**Figure 6.3.** PCO plot comparing three different AFLP runs using 24 samples collected in a hybrid swarm (Edinburgh). Two runs were performed using the same pre-amplification product for the selective PCR. A third run was performed using a pre-amplification product for the selective PCR from a completely new restriction-ligation process. However, the same source of DNA for the initial stage of the restriction process was used for all three runs. Individuals with the highest error rates are labelled (H15: 8.4%, H27 and H29: 7.8%, H20A: 7.1%).

## *Genetic structure of the hybrid zone*

The AFLP run included 54 samples from the hybrid zone as well as 12 pure *G. urbanum* and 12 pure *G. rivale* individuals and 15 controlled F1 crosses for reference purposes. Four primer combinations yielded a total of 203 markers including 15 species-specific *G. urbanum* and nine species-specific *G. rivale* markers. The PCO analysis showed that the parental species were clearly separated along the first axis which accounted for 51% of the total variation (Figure 6.4).

Individuals from the hybrid zone did not form a continuum in the genetic space between the parents but clustered in four more or less well defined groups. The analysis of 203 AFLP markers for pure and hybrid individuals with the program NEWHYBRIDS showed that the four groups can be classified as pure *G. urbanum* (48.1%), F1s (14.8%), backcross to *G. rivale* (14.8%) and pure *G. rivale* (22.2%). None of the sampled individuals had the genetic characters of a 'backcross *G. urbanum*' genotype. The controlled F1 crosses grouped together with the individuals from the hybrid swarm which were identified as F1 genotypes by the program NEWHYBRIDS.

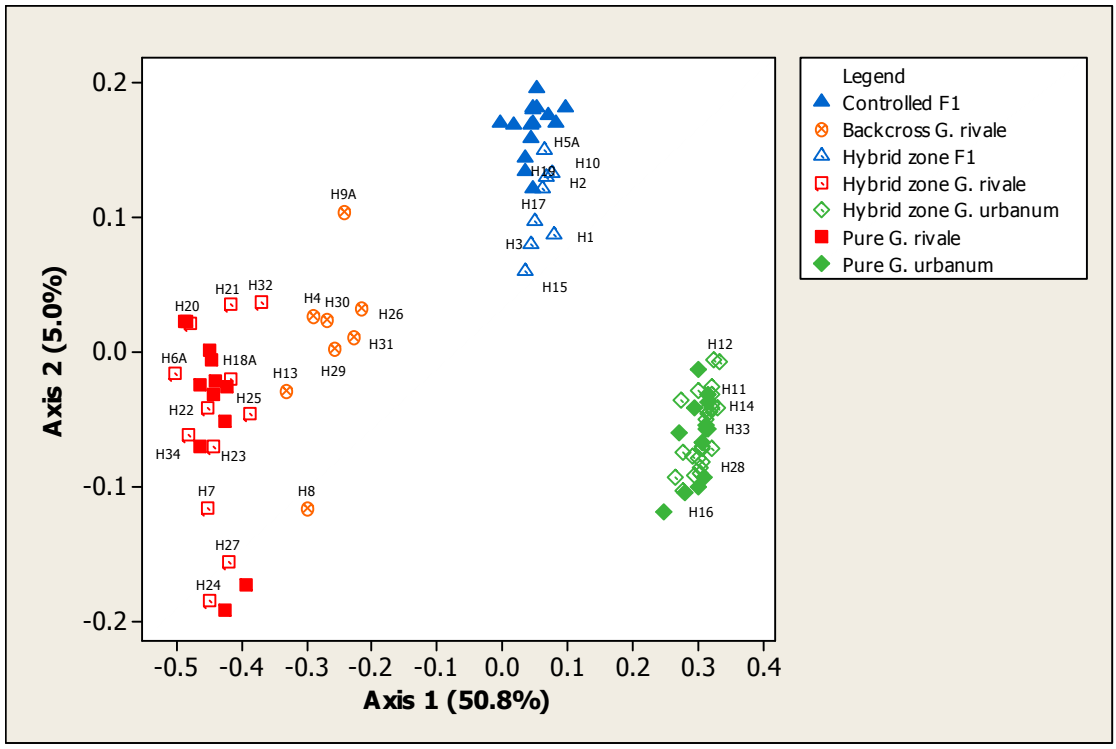



**Figure 6.4.** PCO plot of a hybrid population in Edinburgh (Scotland) (54 plants) using 203 AFLP markers (not all *G. urbanum* like individuals collected in the hybrid zone are labelled individually (H) for clarity reasons). Solid shapes are reference samples from pure *G. urbanum* (n=12) and pure *G. rivale* populations (n=12) as well as from controlled F1 crosses (n=15). Groups are according to the output of the program NEWHYBRIDS.

## *Species-specific marker frequencies in the hybrid swarm*

Analysis of 15 species-specific *G. urbanum* and nine species-specific *G. rivale* markers in the *G. urbanum* like group revealed that apart from one individual, which was missing one *G. urbanum* marker, all samples possessed 100% of the *G. urbanum* and none of the *G. rivale* markers (Table 6.3). In the *G. rivale* like group none of the samples had *G. urbanum* markers and two missed one *G. rivale* marker.

Controlled F1 crosses and F1 like samples from the hybrid swarm differed in their number of *G. rivale* markers (91.1% and 83.3%, respectively) but the difference was not significant (Kruskall Wallis test, p=0.146). Only one individual from the hybrid swarm F1 group had 100% of the species-specific *G. rivale* markers with the remainder of the samples displaying between 66.7% and 88.9% of the markers. Figures for the controlled F1 group were higher with 46.7% of the individuals showing 100% and 53.3% from 77.8% to 88.9% of the *G. rivale* markers. In contrast, all samples in the hybrid F1 group and 99.6% of the controlled F1s displayed 100% of the species-specific *G. urbanum* markers, highlighting their mostly homozygous nature.

| Group | sp-sp *G. urbanum* markers | sp-sp *G. rivale* markers |
|---|---|---|
| *G. urbanum* like samples (6) | 98.9% | 0% |
| *G. rivale* like samples (9) | 0.0% | 97.5% |
| F1 (8) | 100% | 83.3% |
| Controlled F1 (15) | 99.6% | 91.1% |
| Backcross Rivale (7) | 34.3% | 92.1% |

**Table 6.3.** Group frequencies of species-specific markers in four hybrid groups from a hybrid swarm in Edinburgh (Scotland) and controlled F1 crosses (number of individuals).

In the Backcross Rivale group 34.3% of the *G. urbanum* and 92.1% of the *G. rivale* markers were present on average. *G. urbanum* markers were present in variable frequencies ranging from 13.3% (two markers) to 60.0% (nine markers) compared to *G. rivale* specific markers with frequencies of 88.9% (one marker missing) or 100%.

### *CpDNA haplotypes in pure and hybrid populations*

Out of the four different haplotypes A, B, C and D which were detected in pure *G. urbanum* and *G. rivale* populations (screening 151 pure individuals using chloroplast regions trnL-F and trnC-D; Chapter 2.3.2, p 38), only two (A and C) could be found in the hybrid population (Table 6.4). In pure *G. urbanum* only haplotype A occurred whereas all haplotypes (A to D) occurred in pure *G. rivale* populations with haplotypes A and C being the most common (25% and 71%, respectively) and haplotypes B and D being very rare (1% and 3%, respectively).

In the hybrid population haplotype A occurred in 82.4% of the individuals scored. Only 17.6% of the sampled individuals possessed haplotype C, all of which had G*. rivale* morphology (Table 6.4). None of the individuals with hybrid morphology (F1, backcross to *G. rivale*) possessed the private *G. rivale* haplotype C.

| Frequency | A | B | C | D |
|---|---|---|---|---|
| **Pure *G. urbanum* (73)** | 100.0% | 0.0% | 0.0% | 0.0% |
| **Pure *G. rivale* (78)** | 25.3% | 1.3% | 70.9% | 2.5% |
| **Hybrid population (34)** | 82.4% | 0.0% | 17.6% | 0.0% |
| *G. urbanum* like samples (6) | 100.0% | 0.0% | 0.0% | 0.0% |
| *G. rivale* like samples (12) | 50.0% | 0.0% | 50.0% | 0.0% |
| F1 (8) | 100.0% | 0.0% | 0.0% | 0.0% |
| Backcross Rivale (8) | 100.0% | 0.0% | 0.0% | 0.0% |

**Table 6.4.** Chloroplast haplotype (A, B, C, D) frequencies in pure *G. urbanum*, pure *G. rivale* and one hybrid population in Edinburgh (number of individuals used in brackets).

## 6.3.2 Attributes of genetic groups

*Morphology*

A PCA plot of seven morphological traits measured in the field on 26 hybrid, 15 pure *G. urbanum* and ten pure *G. rivale* individuals showed that the parental species are clearly separated along the first component (77.5% of the variation explained, all characters equally weighted). The samples from the hybrid swarm filled the whole phenetic space between the parents and did not show any clear grouping (Figure 6.5).

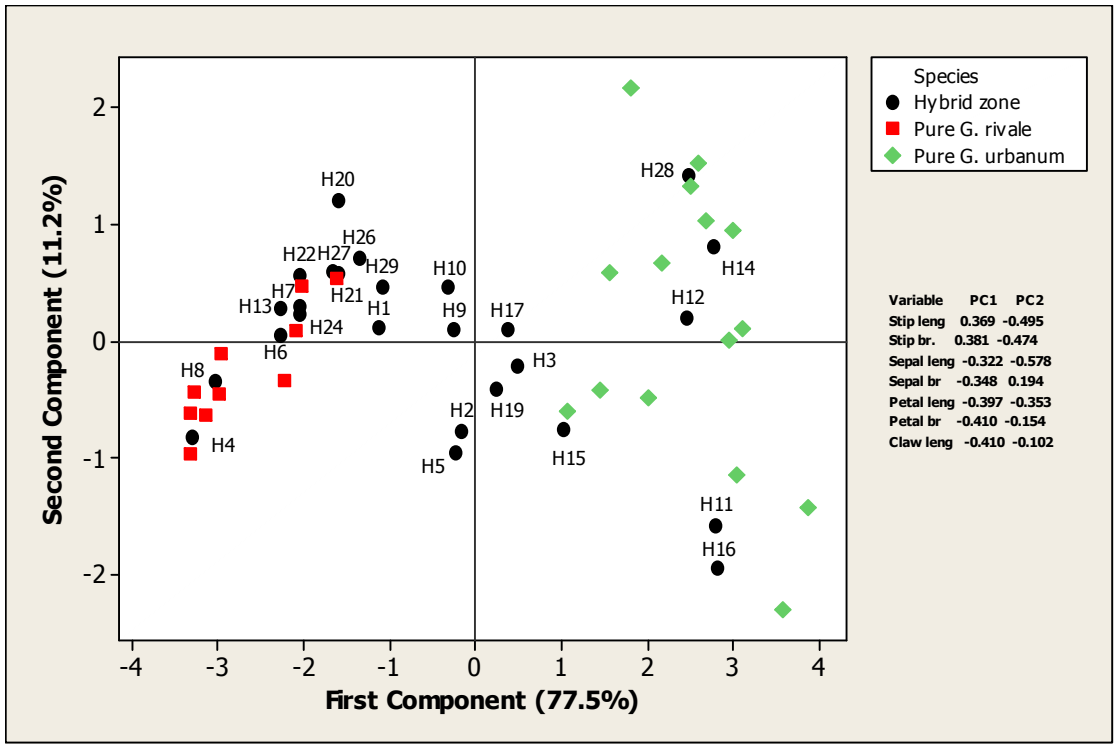



**Figure 6.5.** PCA plot using field measurements of seven morphological characters of individuals from pure *G. urbanum* and pure *G. rivale* populations from Scotland and a hybrid population in Edinburgh (Scotland).

A PCA plot of hybrid, pure and controlled F1 individuals using nine morphological characters taken in a common environment clearly separated the parental species. However, individuals collected in the hybrid swarm did not fill the phenetic space between the two parent species but clustered together in three groups: pure *G. urbanum*, pure *G. rivale* and an intermediate group which coincided with the controlled F1 samples (Figure 6.6). The nine characters were more or less equally weighted for the first component which accounted for nearly 80% of the variation whereas stipule length and petal breadth were the most important contribution for the second component which accounted for about 11% of the variation.

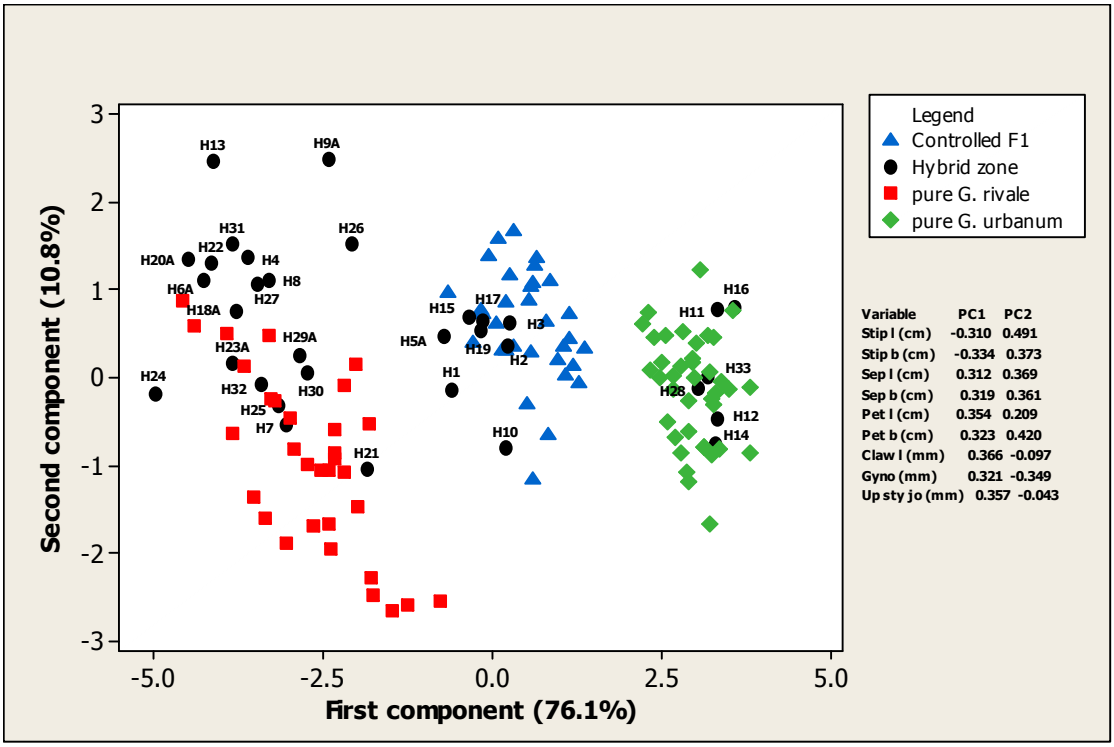



**Figure 6.6.** PCA plot using nine morphological characters measured in a common environment taken from individuals collected in a hybrid population in Edinburgh (Scotland) (34 plants) as well as reference samples from pure *G. urbanum* (39 plants), pure *G. rivale* populations (32 plants) and controlled F1 crosses (32).

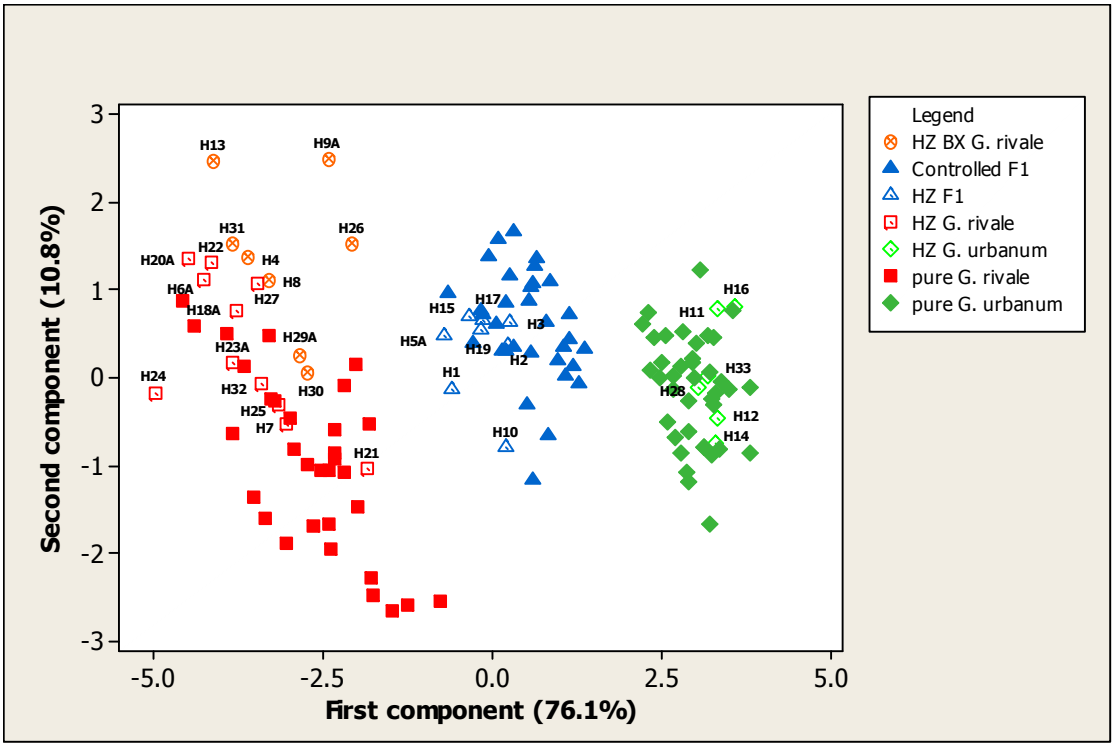



**Figure 6.7.** PCA plot using nine morphological characters taken from individuals collected from a hybrid population in Edinburgh (Scotland) with superimposed genetic grouping as given by NEWHYBRIDS. For reference purposes pure *G. urbanum* and pure *G. rivale* populations as well as controlled F1 crosses were included in the graph. Filled symbols represent individuals from reference populations, empty symbols represent samples from the hybrid swarm (HZ). BX: Backcross Rivale.

*Correlation between morphological and genetic classification*

There was a highly significant correlation between the first axes of the PCA plot for morphology and the PCO plot based on AFLP data (Spearman Rank Order Correlation = 0.915; $p < 0.0001$). Groupings according to the output of NEWHYBRIDS using AFLP data were superimposed on the PCA plot allowing a comparison between the morphological and the genetic data set (Figure 6.7). *G. urbanum* and *G. rivale* like plants from the hybrid swarm clustered clearly with the reference individuals collected from pure populations. All individuals which were classed as F1s by NEWHYBRIDS formed the intermediate group between the parental species in the PCA plot. The Backcross Rivale group, more or less genetically distinct from the pure *G. rivale* reference individuals in the PCO plot, lacked morphological distinction and graded into the pure reference samples in the morphology based PCA plot (Figure 6.7).

*Group comparison*

1. Pure *G. urbanum* and *G. urbanum* like samples from the hybrid swarm

No significant differences in any of the nine characters measured were detected between pure *G. urbanum* and *G. urbanum* like samples from the hybrid swarm (Table 6.5).

| Trait | Average (±SE) pure *G. urbanum* | Average (±SE) *G. urbanum* like |
|---|---|---|
| Stipule length (cm) | 1.89 (±0.04) | 2.08 (±0.07) |
| Stipule breadth (cm) | 2.09 (±0.05) | 2.33 (±0.23) |
| Sepal length (cm) | 0.63 (±0.01) | 0.63 (±0.02) |
| Sepal breadth (cm) | 0.34 (±0.01) | 0.33 (±0.02) |
| Petal length (cm) | 0.72 (±0.01) | 0.70 (±0.03) |
| Petal breadth (cm) | 0.62 (±0.02) | 0.57 (±0.02) |
| Claw length (mm) | 0.00 (±0.00) | 0.00 (±0.00) |
| Gynophore (mm) | 0.00 (±0.00) | 0.00 (±0.00) |
| Upper style joint (mm) | 1.56 (±0.03) | 1.50 (±0.00) |

**Table 6.5.** Average values (± standard error) for nine morphological characters taken in a common environment on pure *G. urbanum* (39) and *G. urbanum* like (6) individuals from the hybrid swarm. No significant differences were detected (Kruskal Wallis test).

2. Controlled and hybrid swarm F1 samples

The mean values of nine morphological characters measured in a common environment were compared between controlled F1s (n=32) and the eight individuals from the hybrid swarm grouping morphologically with the controlled F1 crosses (Figure 6.6). This showed that the hybrid swarm F1 group had significantly shorter and narrower stipules and significantly broader sepals than the controlled F1. However, none of the measurements taken from F1 like samples in the hybrid swarm were outside the parental range.

| Trait | Average (±SE) controlled F1 | Average (±SE) Hybrid swarm F1 |
|---|---|---|
| Stipule length (cm) | 1.78 (±0.05) | **1.34 ***** (±0.09) |
| Stipule breadth (cm) | 1.51 (±0.05) | **1.19*** (±0.10) |
| Sepal length (cm) | 0.87 (±0.02) | 0.86 (±0.04) |
| Sepal breadth (cm) | 0.47 (±0.01) | **0.53*** (±0.02) |
| Petal length (cm) | 0.95 (±0.02) | 0.99 (±0.03) |
| Petal breadth (cm) | 0.87 (±0.02) | 0.89 (±0.03) |
| Claw length (mm) | 1.97 (±0.10) | 2.00 (±0.21) |
| Gynophore (mm) | 1.25 (±0.11) | 0.95 (±0.38) |
| Upper style joint (mm) | 3.09 (±0.05) | 3.31 (±0.16) |

**Table 6.6.** Average values (± standard error) for nine morphological characters taken in a common environment on controlled F1 crosses (37) and eight individuals from the hybrid swarm which grouped with the controlled F1s morphologically in a PCA analysis (Figure 6.6). Measurements which differed significantly between the two groups are marked *0.05≤p≤0.01 and ***p≤0.001 (Kruskal Wallis test).

3. Backcross Rivale samples

The means of three traits (sepal breadth, petal length and breadth) in the Backcross Rivale group were significantly higher than the means of pure *G. urbanum* and *G. rivale* individuals, suggesting transgressive segregation (Table 6.7). Additionally, sepal length and claw length exceeded parental values but were not significantly different from the respective parental measurements. However, only the mean of sepal breadth was significantly larger in Backcross Rivale compared to the *G. rivale* like group in the hybrid swarm (Table 6.7).

| Trait | Average (±SE) pure *G. urbanum* | Average (±SE) Backcross Rivale | Average (±SE) pure *G. rivale* | Average (±SE) *G. rivale* like |
|---|---|---|---|---|
| Stipule length (cm) | 1.89$^{a}$ (±0.04) | 1.29$^{b}$ (±0.11) | 0.76$^{c}$ (±0.03) | 0.99$^{d}$ (±0.05) |
| Stipule breadth (cm) | 2.09$^{a}$ (±0.05) | 0.98$^{b}$ (±0.08) | 0.45$^{c}$ (±0.04) | 0.69$^{d}$ (±0.01) |
| Sepal length (cm) | 0.63$^{a}$ (±0.01) | 1.06$^{bc}$ (±0.05) | 0.93$^{b}$ (±0.03) | 1.07$^{c}$ (±0.05) |
| Sepal breadth (cm) | 0.34$^{a}$ (±0.01) | 0.69$^{b}$ (±0.04) | 0.54$^{c}$ (±0.02) | 0.58$^{c}$ (±0.02) |
| Petal length (cm) | 0.72$^{a}$ (±0.01) | 1.31$^{b}$ (±0.02) | 1.14$^{c}$ (±0.02) | 1.32$^{b}$ (±0.06) |
| Petal breadth (cm) | 0.62$^{a}$ (±0.02) | 1.26$^{b}$ (±0.06) | 0.95$^{c}$ (±0.03) | 1.20$^{b}$ (±0.05) |
| Claw length (mm) | 0.00$^{a}$ (±0.00) | 5.63$^{bc}$ (±0.42) | 5.13$^{b}$ (±0.15) | 5.88$^{c}$ (±0.25) |
| Gynophore (mm) | 0.00$^{a}$ (±0.00) | 4.13 $^{b}$ (±0.48) | 6.17$^{c}$ (±0.30) | 6.17$^{c}$ (±0.97) |
| Upper style joint (mm) | 1.56$^{a}$ (±0.03) | 4.31$^{b}$ (±0.16) | 4.82$^{bc}$ (±0.17) | 5.21$^{c}$ (±0.29) |

**Table 6.7.** Average values (± standard error) for nine morphological characters taken in a common environment in pure *G. urbanum* (39) and pure *G. rivale* (32) populations as well as for *G. rivale* like (12) and Backcross Rivale individuals (8) collected in a hybrid swarm in Edinburgh/Scotland. Measurements which differed significantly ($p<0.05$) between the groups are marked with different letters (Kruskal Wallis test).

4. Pure *G. rivale* and *G. rivale* like samples from the hybrid swarm

The means of six traits (stipule and petal length and breadth, sepal and claw length) were significantly larger in the *G. rivale* like group of the hybrid swarm compared to pure *G. rivale* individuals (Table 6.7).

## *Pollen fertility*

Pollen fertility was measured in pure populations, controlled F1s and the four genetic categories detected in the hybrid swarm (number of individuals measured in parenthesis): pure *G. urbanum* (20), pure *G. rivale* (24), controlled F1s (14) and 40 individuals from the hybrid swarm belonging to *G. urbanum* like (9), *G. rivale* like (15), F1 (8) and backcross *G. rivale* (8) groups. Pollen fertility was highest in pure populations and parental like groups in the hybrid swarm, intermediate in the hybrid swarm backcross *G. rivale* group and lowest in the controlled F1 and hybrid swarm F1 group (Table 6.8). There was a highly significant difference in pollen fertility between pure, controlled F1 and hybrid groups ($F_{6,91}=12.53$, $p<0.001$).

| Groups | Samples used | Group fertility [%(±SE)] | Minimum fertility [%] | Maximum fertility [%] |
|---|---|---|---|---|
| **Pure populations** | | | | |
| *G. urbanum* | 20 | 94.3(±2.4) | 54.7 | 100 |
| *G. rivale* | 24 | 90.3(±3.5) | 20.0 | 99.0 |
| **Controlled F1 crosses** | 15 | 70.0(±3.9) | 28.9 | 86.2 |
| **Hybrid population** | | | | |
| *G. urbanum* like samples | 9 | 94.7(±1.4) | 89.6 | 98.8 |
| *G. rivale* like samples | 15 | 88.0(±3.6) | 62.5 | 98.8 |
| F1 | 8 | 62.9(±5.2) | 43.2 | 78.9 |
| Backcross to *G. rivale* | 8 | 83.4(±4.0) | 63.1 | 95.6 |

**Table 6.8.** Pollen fertility in pure *G. urbanum* and *G. rivale* populations, controlled F1 crosses and one hybrid population (Edinburgh).

A Tukey pairwise comparison between groups revealed that pollen fertility in controlled F1s as well as F1 like samples from the hybrid swarm was significantly different (i. e. lower) compared to pure populations and parent like groups in the hybrid swarm (Table 6.9). No significant differences in pollen fertility between controlled F1s and F1 like samples from the hybrid swarm were detected.

| **Groups** | Pure Urb | Pure Riv | ContF1 | HZ Urb | HZ Riv | HZ F1 |
|---|---|---|---|---|---|---|
| Pure Riv | ns | | | | | |
| ContF1 | *** | *** | | | | |
| HZ Urb | ns | ns | *** | | | |
| HZ Riv | ns | ns | ** | ns | | |
| HZ F1 | *** | *** | ns | *** | *** | |
| HZ BXRIV | ns | ns | ns | ns | ns | ns |

**Table 6.9.** Tukey pairwise group comparison of pollen fertility between pure *G. urbanum* (Pure Urb) and pure *G. rivale* (Pure Riv) populations, controlled F1 (ContF1) and hybrid groups *G. urbanum* like (HZ Urb), *G. rivale* like (HZ Riv), F1 (HZ F1) and Backcross Rivale (HZ BXRIV). **$0.01 \leq p < 0.001$ and ***$p \leq 0.001$.

A $\chi^2$ test showed that there was significant heterogeneity in pollen fertility within all groups (results not shown). The number of individuals with intermediate pollen fertility (between 40% and 80%) was largest in controlled F1s and F1 like samples from the hybrid swarm and smallest in pure populations and parent like samples from the hybrid swarm (Figure 6.8). Brown stamens with indumentum, a character which has not been observed in pure populations (where stamens are yellow and glabrous), was detected in the hybrid swarm. Stamens with indumentum occurred in 50% of F1s (4 out of 8), in 82% of the Backcross Rivale samples (9 out of 11) and 25% of the *G. rivale* like group (4 out of 16). None of the sampled individuals (6) from the *G. urbanum* like group possessed this character, however,

brown stamens with indumentum were observed very rarely in this group in the field (pers. observ.).

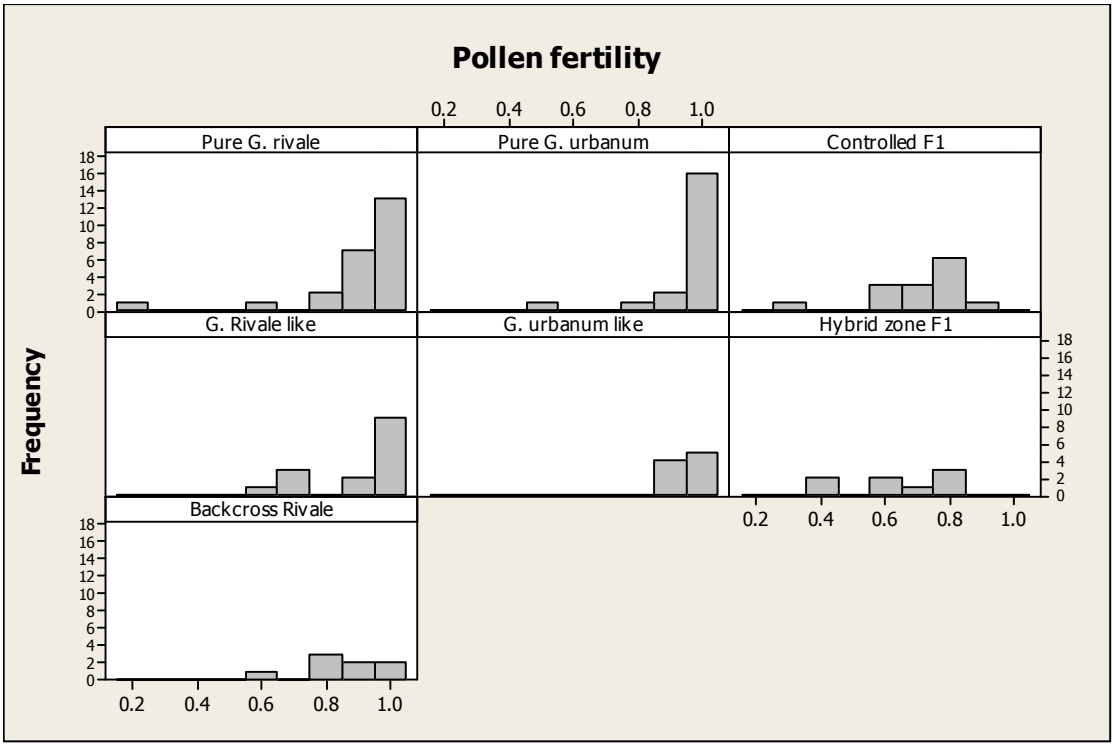



**Figure 6.8.** Histogram of pollen fertility in pure populations, controlled F1s and hybrid groups *G. urbanum* like, *G. rivale* like, hybrid swarm F1 and Backcross Rivale.

## *Auto-pollination*

Auto-pollination success was measured in two ways. Firstly, as the average success rate to produce seeds by auto-pollination within each group and secondly, as the percentage of individuals (genotypes) within each group, which possessed at least one bagged flower that gave rise to seeds by auto-pollination (mean % autopoll. flowers and % individuals autopoll., respectively, in Table 6.10). There was a highly significant difference in auto-pollination success rates (overall ability/inability of bagged flowers to auto-pollinate) between the four different groups collected in the hybrid swarm ($F_{3,35}$=16.04, p<0.001). In *G. urbanum* like samples 100% of the bagged flowers produced seed heads. In contrast, the mean success rate in the *G. rivale* like group was 10.2% (Table 6.10). Only three out of 15 individuals were able to set seed by auto-pollination. Backcross Rivale samples had a mean auto-pollination rate of 17.1%. Six out of 11 individuals were able to auto-pollinate. All individuals in the F1 group were able to produce seeds by auto-pollination. However, the mean frequency of auto-pollination success was only 46.9%. A Tukey pairwise comparison of all four groups revealed that there was a significant difference in the ability to auto-pollinate between the *G. urbanum* like group and all other groups (Table 6.11).

| Hybrid group | Samples used | Average number bagged flowers | Mean % autopoll. flowers | % of individuals autopoll. |
|---|---|---|---|---|
| *G. urbanum* like samples | 5 | 3.6 | 100 | 100 |
| *G. rivale* like samples | 15 | 3.3 | 10.6 | 20 |
| F1 | 8 | 7 | 46.9 | 100 |
| Backcross to *G. rivale* | 11 | 3.7 | 20.3 | 54.5 |

**Table 6.10.** Auto-pollination success rates of 39 bagged individuals collected in the hybrid swarm. Mean % autopoll. flowers: group average of the percentage of bagged flowers which auto-pollinated per individual. % individuals autopoll: percentage of individuals with the ability to auto-pollinate (i. e. at least one bagged flower auto-pollinated successfully).

| Groups | *G. urbanum* like | *G. rivale* like | F1 |
|---|---|---|---|
| *G. rivale* like | *** | | |
| F1 | *** | ns | |
| Backcross Rivale | *** | ns | ns |

**Table 6.11.** Tukey pairwise group comparison of the ability to auto-pollinate between the four genetic groups detected in the hybrid swarm. ***p≤0.001.

*Phenology*

*G. rivale* started flowering at the beginning of May and peaked towards the end of the month (435 open flowers recorded) when the first *G. urbanum* plants started to open their flowers (Figure 6.9). At flowering peak time of *G. urbanum* in the middle of June, 709 open *G. urbanum* flowers were recorded. The two species overlapped in flowering times for about four weeks during the month of June. Hybrid individuals flowered throughout the recording period from the beginning of May till the middle of July and peaked at the end of May/beginning of June (340 open hybrid flowers, Figure 6.9).

Looking at different hybrid classes separately revealed that F1 individuals have a peak flowering time (250 open F1 flowers) of about two weeks from the end of May until the middle of June whereas Backcross Rivale individuals flower at a low but more or less constant rate (30 to 50 open flowers) from the beginning of May until the end of June (Figure 6.10). Individuals which could not be allocated unambiguously to either the Backcross Rivale or *G. rivale* class made up 90% of the 'Uncertain' category whereas the rest (10%) could not be allocated clearly to the Backcross Urbanum or *G. urbanum* group. No Backcross Urbanum individuals were detected.

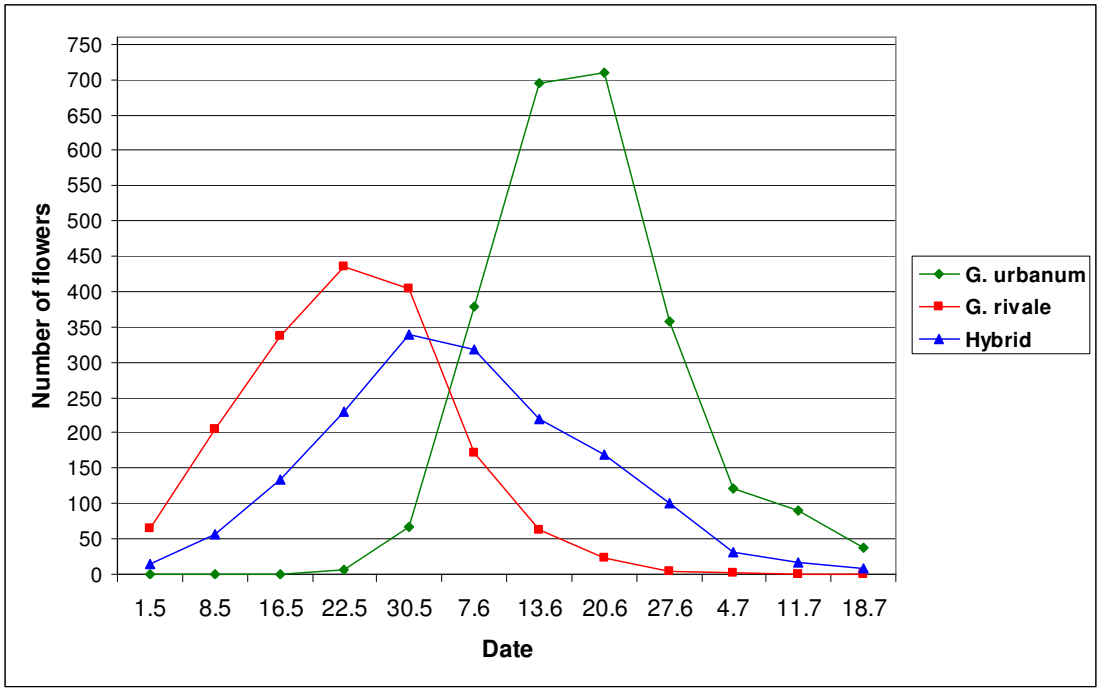



**Figure 6.9.** Phenology of a hybrid swarm in Edinburgh/Scotland using two parental classes (*G. urbanum*, *G. rivale*) and one hybrid class.

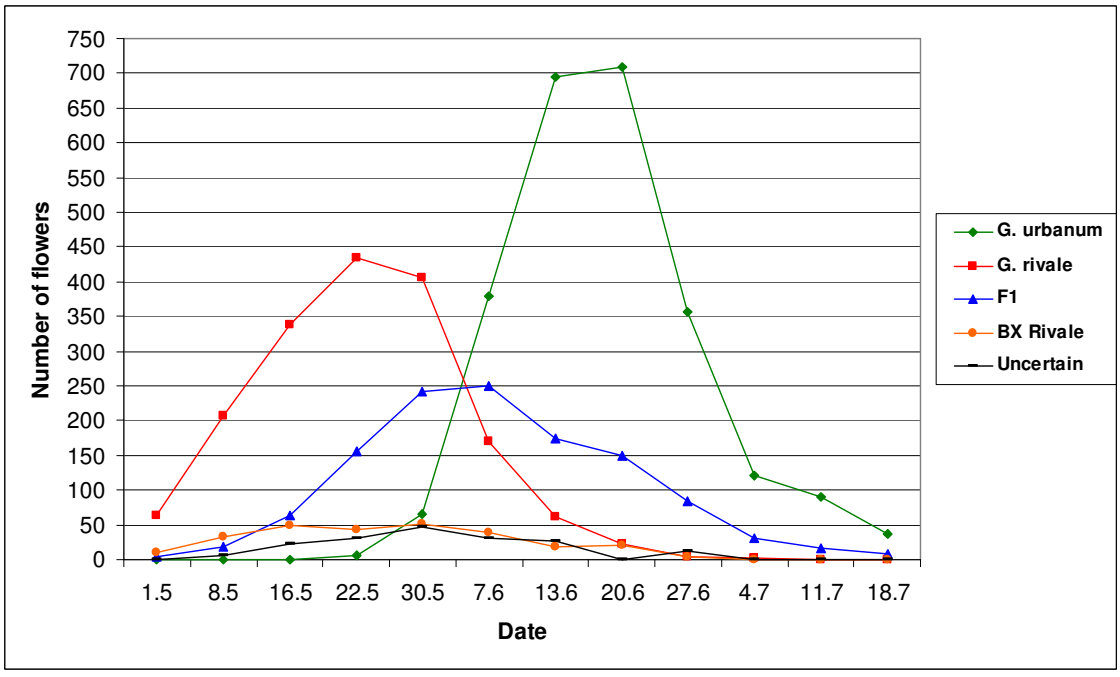



**Figure 6.10.** Phenology of a hybrid swarm in Edinburgh/Scotland using five classes: two parental (*G. urbanum*, *G. rivale*) and two hybrid classes (F1 and Backcross Rivale) as well as the category Uncertain for individuals which could not positively allocated to either a Backcross or a parental class.

# 6.4 Discussion

## 6.4.1 AFLP error rates

AFLP runs were repeated three times for 24 samples. Two runs were performed after using the same pre-amplification product for the selective PCR whereas the third run was performed using a pre-amplification product from a new restriction-ligation process. Average error rates fell in the 2 - 5% range typically cited in the literature for AFLPs (Bonin *et al.*, 2007). However, individual error rates could be about twice as high as the average error rate but never exceeded the maximum acceptable error of 10% recommended by Bonin *et al.* (2007). Average (4.1% and 5.1%) and individual error rates (1.9% to 8.4% and 1.3% to 7.8%, respectively) were higher between runs using different pre-amplification products compared to same pre-amplification runs (average error 2.6% and individual error rates from 0% to 5.2%). This suggests that variation in the digestion-ligation process can influence the reproducibility of AFLPs.

Despite the relatively high individual error rates between various runs the effect on the group allocation of the samples was minimal. Individuals from different runs showed minor shifts within groups. However, the slightly different positions of the same individual did not affect their group allocation, even for samples with the highest error rates. The robustness of the inferred biological message in terms of error rates was probably due to the redundancy of the information contained in the large number of AFLP markers (more than 200 polymorphic loci were screened in total).

## 6.4.2 Genetic groups present in the hybrid swarm

A PCO plot using 203 AFLP markers and 54 samples (including the 34 samples used for the morphological analysis) showed that four groups can be distinguished in the hybrid swarm (1) a *G. urbanum* like group clustering with pure *G. urbanum* reference populations (2) an F1 like group clustering with controlled F1s (3) a Backcross Rivale group and (4) a *G. rivale* like group clustering with pure *G. rivale* reference populations. These groups were fully supported by the analysis of AFLP data with the software NEWHYBRIDS (i. e. all individuals in each genetic cluster of the PCO plot were assigned to the same group by NEWHYBRIDS).

## 6.4.3 Chloroplast haplotypes

Although it was not possible to establish the direction of hybridization directly as no species-specific chloroplast markers for *G. urbanum* and *G. rivale* were detected, it seems likely that most initial hybrid crosses have *G. urbanum* as maternal parent. This hypothesis is supported by haplotype frequencies in pure and hybrid populations. A marked frequency difference between the two most common haplotypes (A and C) of pure *G. urbanum* and *G. rivale* populations was observed in the hybrid swarm. In pure *G. urbanum* populations only haplotype A can be found whereas in pure *G. rivale* populations haplotype C is much more frequent (70.9%) than haplotype A (25.3%), with haplotypes B and D being very rare (1.3% and 2.5%, respectively).

The only haplotypes detected in the hybrid swarm were A (82.4%) and C (17.6%). All six individuals with the private *G. rivale* haplotype C are genetically and morphologically indistinguishable from pure *G. rivale* individuals. None of the hybrid pheno- and genotypes (F1s and Backcross Rivale) possessed haplotype C which would indicate a *G. rivale* maternal parent. Given that A is the only haplotype in pure *G. urbanum* populations and that its frequency is low in pure *G. rivale* populations, it seems plausible that most A haplotypes found in F1s and Backcross Rivale individuals are of *G. urbanum* origin. For F1s, this is also supported by the difficulty reported in the literature to produce hybrids if *G. rivale* is the maternal parent (Marsden-Jones, 1930; Prywer, 1932) and by the significantly lower germination success of *G. rivale* x *G. urbanum* crosses in this study.

## 6.4.4 Morphology

Reference populations of pure *G. urbanum*, pure *G. rivale* and controlled F1s were clearly separated in a PCA plot using nine morphological characters taken in a common environment. None of the three groups overlapped and the F1 group was intermediate between the two parental populations. 34 individuals collected in a hybrid swarm in Edinburgh did not fill the phenetic space between the parents but grouped with the three reference populations, i. e. pure *G. urbanum*, pure *G. rivale* and controlled F1s suggesting that only parental and F1 phenotypes are present in the hybrid swarm. This was in contrast with a PCA plot using seven morphological measurements taken in the field where hybrid samples covered the whole phenetic space between the parents.

Morphological characters with environmentally induced phenotypic plasticity are commonly assumed to exhibit the greatest degree of plastic response in variable environments, whereas less variation is expected in constant environments (Mitchell-Olds & Rutledge, 1986). In a common environment the genetically determined phenotype is less variable as confounding

plastic responses to environmental conditions are minimized. The above results show that in order to maximize the interpretability of the biological information present in a morphological data set, it is important to take measurements in a common environment.

In contrast to the genetic analysis using 203 AFLP markers, morphological data did not reveal the presence of Backcross Rivale individuals, which clustered with the pure *G. rivale* group in the PCA plot and could not be distinguished as a separate cluster. However, all individuals in the hybrid swarm which were classed as pure *G. urbanum,* F1s and pure *G. rivale* genetically were in agreement with their respective phenotypes, suggesting that morphology is generally a reliable indicator for the identification of at least two of the groups, i. e. pure *G. urbanum* and F1s.

It is noteworthy that most (but not all) of the Backcross Rivale samples could be identified as such in the field using the 'Gestalt' of the individual. Identification was based on an overall impression including a combination of qualitative characters such as calyx and petal colour as well as other traits which are difficult to quantify exactly, such as hairiness of style joints, calyx and petal angle. This suggests that a morphological analysis using a restricted number of easily quantifiable traits is likely to underestimate the diversity in a hybrid swarm and that molecular analysis is a more informative and reliable tool. Admittedly, a more sophisticated recording of shape and form might improve the morphological discrimination.

### 6.4.5 Transgressive segregation

Comparisons of means of nine morphological traits revealed that individuals from the Backcross Rivale group had on average significantly larger petals and broader sepals than pure parental individuals, indicating transgressive segregation. However, compared to the means of the *G. rivale* like group in the hybrid swarm, only sepal breadth was significantly larger in the Backcross Rivale group.

It is possible that there is more phenotypic variation present in pure *G. rivale* populations than has been captured by using a restricted number of reference populations for this study. *G. rivale* individuals in the hybrid swarm could have been derived from populations with larger phenotypic traits. This seems somewhat unlikely, however, as one of the pure *G. rivale* reference populations is a subalpine population collected at Ben Lawers/Scotland. Subalpine populations have long been known to possess on average larger flowers and leaves than lowland plants (Turesson, 1925). Comparing the means of 13 individuals from the Ben Lawers population with the means of 15 *G. rivale* like individuals from the hybrid swarm, showed that

the hybrid swarm samples exceeded the Ben Lawers individuals in all measurements apart from sepal breadth, which was equal (data not shown). Additionally, Prywer (1932) observed similar, albeit slightly smaller, average petal length (1.0 cm) and breadth (0.9 cm) in the *G. rivale* samples of his study compared to the average of pure *G. rivale* samples in this study (1.14 cm and 0.97 cm, respectively). Similarly, Waldren *et al.* (1989) reported an average petal length and breadth of 1.16 cm and 1.06 cm (field measurements), respectively in a group of *G. rivale* like individuals in a hybrid swarm in North England.

It is therefore possible that the six larger phenotypic traits in *G. rivale* like individuals collected in the hybrid swarm could be larger due to transgressive segregation which has been carried over by Backcross Rivale individuals. The means of all the six characters in question are also larger in the Backcross Rivale group compared to the means of pure *G. rivale* populations, although only four significantly (stipule and petal length and breadth). This would require that many *G. rivale* like individuals in the hybrid swarm are actually advanced backcrosses which can not be distinguished from pure *G. rivale* samples anymore but have retained some of the transgressive characters found in early backcross generations. That speculation is also supported by the transgressive character 'brown stamens with indumentum'. This trait was very rare in *G. urbanum* like individuals but occurred frequently in F1s, Backcross Rivale and *G. rivale* like samples. It is very likely that all individuals exhibiting brown stamens with indumentum are of hybrid origin, suggesting that the *G. rivale* like group in the hybrid swarm is an assemblage of advanced backcrosses and 'pure' individuals. About 25% of the individuals in the *G. rivale* like group had this character possibly indicating that these individuals are advanced backcrosses to *G. rivale*. It is therefore feasible that the characters discussed above have found their way into pure looking *G. rivale* individuals in the hybrid swarm via the Backcross Rivale group, where more than 80% of the samples display that trait.

Marsden-Jones (1930) reported a transgressive character regarding a novel "short plumose" hair type on the style in nearly 50% of the controlled Backcross Urbanum individuals, which otherwise where very like *G. urbanum* and "indeed in the field many of them would have passed for *G. urbanum* ". However, no Backcross Urbanum group was detected in the analysed hybrid swarm.

The means of three traits in the hybrid swarm F1 group differed significantly from the means observed in controlled F1 samples although these differences did not alter their F1 group allocation in a PCA plot using nine morphological characters. Hybrid swarm F1 individuals had smaller stipules and broader sepals. Comparing these results to the measurements of controlled hybrid crosses in the literature (Marsden-Jones, 1930; Prywer, 1932), it is evident

that despite the overall uniformity of the F1 generation measured here, some variation exists between F1 hybrids.

Prywer (1932) did not observe a gynophore in his F1 crosses whereas Marsden-Jones (1930) reported a gynophore in some crosses. The average petal size in F1 hybrids obtained by Prywer (1932) was much smaller (petal length and breadth: 0.70 cm and 0.60 cm, respectively) than the petal sizes in either of the controlled F1 (0.95 cm and 0.87 cm, respectively) or hybrid swarm F1 individuals (0.99 cm and 0.90 cm, respectively) in this study. In contrast, Prywer (1932) reported longer sepals in F1 samples (1.0 cm) compared to smaller sepals in this study (0.87 cm and 0.88 cm, respectively). It could be that the differences between controlled F1 crosses and hybrid swarm F1 samples are due to the morphological variability expressed in F1 hybrids. Alternatively, it could also be due to the limited number of populations used in the controlled F1 crosses (two populations per species) and/or because of the differences in sample sizes between the two groups (32 samples from the controlled F1 crosses compared to 8 samples from the hybrid swarm F1 group). The genetic analysis conducted here found no evidence for F2s. However, it is also possible that some of the F1 like samples are not first but second or later generation hybrids which were not identifiable with the methods employed.

## 6.4.6 Pollen fertility and auto-pollination

Pollen fertility in the *G. urbanum* and *G. rivale* like samples from the hybrid swarm did not differ significantly from the fertility found in the respective pure populations. Controlled F1 (70.0%) and hybrid swarm F1 phenotypes (62.9%) displayed the lowest pollen fertility of all groups but fertility was somewhat restored in the backcross *G. rivale* group (83.4%).

These results are in line with other studies where F1 hybrids displayed low pollen fertility like in *Helianthus* (6%) (Heiser, 1947), *Gilia* (2%) (Grant, 1966), *Mimulus* (30%) (Ramsey *et al.*, 2003) and *Armeria* (40 to 60%) (Tauleigne-Gomes & Lefebvre, 2008). Higher pollen fertility in later backcross generations is a common phenomenon (Arnold *et al.*, 1999) and has been observed in various plant taxa (Grant, 1966; Ungerer *et al.*, 1998; Tauleigne-Gomes & Lefebvre, 2005). Rieseberg *et al.* (1996) conducted pollen fertility experiments with synthetic greenhouse-grown *Helianthus annuus* x *H. petiolaris* hybrids and reported that although F1s were mostly sterile (<10%), pollen fertility was rapidly re-established after only four generations of sib-mating or backcrossing to more than 90%. Pollen fertility showed significant heterogeneity in all groups indicating that fertility varied amongst individuals considerably. This variability was largest in F1s where the majority of samples (80% in controlled F1s and 100% in F1 phenotypes from the hybrid swarm) displayed low to

intermediate pollen fertility (21%-80%). In contrast, only about 15% had less than 90% fertility in pure *G. rivale* and *G. urbanum* populations, indicating that most pure samples are fully fertile in either species. The overall high number of flowering F1 phenotypes in the hybrid swarm suggests that despite the significantly lower pollen fertility compared to parental phenotypes in the hybrid swarm, F1 hybrid pollen is likely to be a significant source for the production of later generation hybrids. However, individual contributions will largely depend on individual fertility rates.

The results from the auto-pollination experiment using samples from the hybrid swarm showed that all five *G. urbanum* like samples had a 100% success rate in producing seed heads by auto-pollination whereas only three out of 15 *G. rivale* like individuals where able to auto-pollinate. This is in line with the auto-pollination results from pure populations in Chapter 3.3.1 (p 60) where the ability to auto-pollinate was severely restricted in *G. rivale* and completely unrestricted in *G. urbanum*. All eight F1 individuals from the hybrid swarm produced seed heads via auto-pollination, in contrast to only 54% of Backcross Rivale individuals. This suggests that the ability to auto-pollinate depends on the percentage of *G. urbanum* genes incorporated in a particular hybrid and is high in F1s and declines in backcrosses to *G. rivale*. F1 individuals are therefore likely to be able to reproduce successfully even if suitable pollinators are rare or avoid certain intermediate phenotypes.

## 6.4.7 Patterns of hybridization

The genetic analysis did not detect a Backcross Urbanum or an F2 group in the hybrid swarm. In a study of a hybrid swarm between *Populus fremontii* and *P. angustfolia* a similar situation was observed (Keim *et al.*, 1989). Within the hybrid swarm individual trees were either F1 hybrids or backcrosses to *P. angustfolia* but no progeny were found that could be attributed to crossings between F1s or backcrosses between F1s and *P. fremontii.* Keim *et al.* (1989) could exclude phenology as the reason for the lack of backcrosses to *P. fremontii* and concluded that genetic incompatibilities were the most likely cause for the observed asymmetric pattern of hybridization as greenhouse crosses showed a much lower survival rate of backcross 'Fremontii' individuals.

In a more recent study of a hybrid swarm between *Rhododendron hirsutum* and *R. ferrugineum* in the Austrian Alps, a similar situation to the one in the *Geum* hybrid swarm was encountered (Milne & Abbott, 2008). Among all hybrid derivatives examined about half were designated as F1s or a similar class whereas the rest of the remaining hybrid classes were identified as backcrosses towards *R. hirsutum*. Backcrosses to *R. ferrugineum* were not detected although viable seeds of this class could be generated by hand, which showed a

similar seedling survivorship (to 6 months) to backcrosses towards *R. hirsutum* (Milne & Abbott, 2008). Milne & Abbott (2008) pointed out that the highly asymmetric backcrossing pattern could be explained in terms of phenology as *R. ferrugineum* plants nearly finish flowering at a time when hybrids are in full flower and *R. hirsutum* flowers are just starting to open. All three taxa are protandrous, so hybrid stigmas would mainly receive pollen from *R. hirsutum* and other hybrids reducing the likelihood of backcrossing to *R. ferrugineum*. The absence of F2 individuals is probably due to a habitat mediated fitness disadvantage of F1 offspring as hand pollination experiments showed that F1s were fertile (Milne & Abbott, 2008).

The lack of a Backcross Urbanum group in the hybrid swarm is somewhat puzzling. Studies have shown that if one parent is adapted to disturbed sites (Bleeker & Hurka, 2001; Lexer *et al.*, 2005) or early succession stages (Bacilieri *et al.*, 1996; Caraway *et al.*, 2001), introgression tends to be towards this parent as many hybrid swarms occur on disturbed sites (Anderson, 1948; Rieseberg & Carney, 1998; Milne & Abbott, 2008). *G. urbanum* occurs naturally in woodlands but has a 'weedy' tendency and can often be found in ruderal habitats whereas *G. rivale* is less tolerant of disturbed sites and usually grows in its natural habitat in wet meadows, subalpine communities and along streams. This suggests that *G. urbanum* is the more likely parent for backcrossing. However, earlier studies based on morphological observations have also observed a lack of backcrosses to *G. urbanum* in hybrid swarms in England (Waldren *et al.*, 1989) and Finland (Ravanko, 1979).

In contrast to the *Rhododendron* study (Milne & Abbott, 2008), phenology can not explain the lack of a Backcross Urbanum group in the hybrid swarm. *G. urbanum* and F1 individuals overlap in flowering time for more than one month and flowers of both groups are abundant in the hybrid swarm, suggesting that there is ample opportunity to interbreed. It is possible that pollinator preferences keep the two groups apart as F1 phenotypes overall resemble *G. rivale* more than *G. urbanum*, which could lead to the observed pattern of backcrossing to *G. rivale.* Alternatively, habitat mediated selection against Backcross Urbanum individuals could explain their absence in the hybrid swarm. One explanation for the lack of F2 individuals in the hybrid swarm could be fertility problems in the F1 generation. This seems to be unlikely as pollen fertility, although significantly lower than in the parental species, is generally high (average 62.9% fertility) and many F1s in the hybrid swarm displayed full seed heads at the end of the season. However, it is possible that F1 seeds have a low germination rate and are inviable or are mainly derived by backcrossing to *G. rivale*. There are thus multiple potential reasons for the absence of the F2 and Backross Urbanum group. As alluded to earlier, at least for F2 detection, method sensitivity may also be an issue. Screening seeds from plants in the hybrid swarm may shed further light on this and will enable me to establish whether all

genotypic classes are formed, but only some become established, and also whether F1 plants produce offspring which are erroneously classed as F1s.

## 6.4.8 Summary

Analyzing a hybrid swarm between *G. urbanum* x *G. rivale* with molecular and morphological tools showed that there is a high correlation between the morphology and the genetic make-up of studied individuals. Molecular markers provided a better resolution than a limited set of morphological data, especially when taken in the field.

The genetic analysis revealed that the hybrid swarm is comprised of four distinct genetic groups, i. e. (1) a parental *G. urbanum* like group (2) an F1 like group (3) a Backcross Rivale group and (4) a parental *G. rivale* like group. No F2 or Backcross Urbanum individuals were detected.

Pollen fertility was high in parental species (c. 90%), lowest (62.9%) in F1 individuals and somewhat restored in Backcross Rivale samples (83.4%). There was evidence for transgressive segregation for at least one phenotypic trait (sepal breadth) in the Backcross Rivale group. A transgressive character, brown stamens with indumentum, was found in the hybrid swarm. This trait was not restricted to obvious hybrid individuals but occurred in all four groups, although it was very rare in *G. urbanum* like samples. This suggests that pure looking *G. rivale* individuals with this character are advanced backcrosses.

Results from the chloroplast DNA analysis showed that the private *G. rivale* haplotype C is less common in the hybrid swarm than haplotype A, which occurs at a much higher frequency in *G. urbanum* than in *G. rivale* populations. The haplotype frequency differences in the hybrid swarm and reports from the literature about difficulties obtaining seeds if *G. rivale* is the maternal parent suggest that *G. urbanum* serves as the main maternal parent in hybridization events.

Understanding the absence of Backcroos Urbanums and F2s requires a more detailed account of the dynamics of the hybrid swarm.

# 7. DYNAMICS OF THE HYBRID SWARM

## Abstract

Understanding the processes which generate and maintain the genetic and phenotypic structure of a hybrid swarm can provide insight into the evolutionary potential of hybridization. The morphological and genetic structure of 30 adults collected in a hybrid swarm between *Geum urbanum* and *G. rivale* and five of their open-pollinated seeds were compared. The seedlings were planted into a common garden environment and their phenotype (nine morphological traits), fitness (seed head production, growth volume) and genotype (203 AFLP markers) was established. The standing adult population showed a strong correlation between AFLP genotypes and morphological phenotypes (Spearman Rank Order 0.915, $p<0.001$). In contrast the progeny of open-pollinated mothers showed a reduced correlation between genotype and phenotype suggesting segregation in advanced generation hybrids (Spearman Rank Order 0.642, $p<0.001$). In the adult population, the genotypic diversity could be classed into four genotypic classes (*G. urbanum* like, F1 like, Backcross Rivale and *G. rivale* like), whereas their offspring showed a much greater diversity of genotypes indicating that selection acts against some genotypic classes to prevent their establishment in the field. There was evidence for the generation of Backcross Urbanum and F2 individuals in the offspring generation. However, common garden fitness experiments measuring reproductive output and plant growth did not detect a selective disadvantage in any genotypic class, suggesting that any fitness differences are unlikely to be due to intrinsic genetic problems, but instead are more likely to be attributable to extrinsic genotype-environment interactions. Finally, the genotypic composition of *G. urbanum* like, F1 like and *G. rivale* like progeny was significantly more varied than the genotypic composition of their parents. Curiously, adult plants classed as Backcross Rivale only gave rise to Backcross Rivale like offspring suggesting some form of assortative mating for this group. Potential mechanisms for this and possible causes for the differences between the genotypic structure of adult and offspring generations are discussed.

# 7.1 Introduction

The heterogeneity of hybrid swarms in terms of their pheno- and genotypic structure has been well documented in the literature. Many hybrid swarms are dominated by hybrids from later generations. Examples include *Populus* (Lexer et al., 2005), *Arctostaphylos* (Nason *et al.*, 1992), *Helianthus* (Rieseberg & Carney, 1998) and *Pinus* (Watano *et al.*, 2004) with only a few F1s, and *Iris* (Cruzan & Arnold, 1993; Hodges *et al.*, 1996) and *Liparis* (Chung et al., 2005) where F1s are absent. In contrast, a range of hybrid swarms comprise mainly F1 like individuals such as in *Rhododendron* (Milne *et al.*, 2003), *Phyllodoce* (Kameyama *et al.*, 2008) and *Quercus* (Nason *et al.*, 1992). From an evolutionary point of view it is important to know the frequency of these different hybrid classes as they provide information about the potential for gene flow and can help us to understand the processes that generate and maintain these differences (Minder *et al.*, 2007). If the hybrid zone is dominated by F1 hybrids then there is little potential for gene exchange between the parental species, whereas an abundance of backcross individuals is likely to lead to substantial interspecific gene flow and introgression.

Despite the potential evolutionary importance of the presence or absence of certain genotypic classes, very few studies have gone beyond the analysis of the adult generation or controlled crosses. However, to establish whether there are genotypic differences between the standing adult population and the potential future adult population is essential to gain a deeper insight into mating events and possible selection against certain genotypic groups in the hybrid swarm. There are various causes which can explain the absence of certain groups such as intrinsic or endogenous selection (i. e. lower fitness and viability of certain classes) and extrinsic or exogenous selection (i. e. environmentally mediated discrimination due to the advantageous or disadvantageous interaction between genes and the environment). To be able to distinguish between endogenous and exogenous forces it is necessary to measure the fitness of individuals in multiple environments with reciprocal transplant experiments. Exogenous selection is indicated by a difference in relative fitness of hybrids depending on the environment they grow in whereas endogenous isolation would be indicated by low hybrid fitness in all environments (Kimball *et al.*, 2008).

Another possible reason for the absence of genotypic classes is assortative mating, which is the non random mating of one group with another. This minimizes the chances that certain groups are generated in the first place and can be either caused by pollinator behaviour or by selection (Cruzan & Arnold, 1994). It has long been known that reduced survival or fitness of hybrids beyond the F1 stage can result from a breakdown in co-adapted gene complexes, which is mainly attributable to the epistatic effects described by Dobzhansky–Muller

incompatibilities (Dobzhansky, 1936; Orr, 1995; Fenster *et al.*, 1997; Demuth & Wade, 2005). Recombination causes the break-up of co-adapted gene complexes or creates new maladapted combinations, thus producing hybrids with lower average fitness compared to their progenitors, leading to endogenous selection against these individuals (Demuth & Wade, 2005; Johansen-Morris & Latta, 2006).

Cruzan and Arnold (1994) examined the dynamics of a hybrid zone between *Iris fulva* and *I. brevicaulis* by genotyping open-pollinated seeds from adults collected in the hybrid zone. The *Iris* hybrid zone was dominated by parental and advanced hybrid genotypes but lacked early generation genotypes. However, the genotypic distribution of offspring collected as open pollinated seeds contained a high proportion of early generation genotypes which were not present in the adult generation (Cruzan & Arnold, 1994). *I. brevicaulis* and hybrid genotypes produced a mixture of offspring randomly mating with different genotypic classes, whereas *I. fulva* types only mated with genetically similar individuals and did not produce offspring that were the result of crossings with the most genetically dissimilar classes. Cruzan and Arnold (1994) argued that the observed asymmetric pattern is most likely to be due to intrinsic selection such as pollen-style interactions (reduction of fertilization ability of genetically dissimilar pollen) or preferential abortion of genetically intermediate zygotes by *I. fulva* like genotypes.

Exogenous selection can play an important role in terms of relative fitness of parents and hybrids as recent studies have suggested (Emms & Arnold, 1997; Johansen-Morris & Latta, 2006; Campbell & Waser, 2007). An example for extrinsic selection against hybrids was documented in two hybridizing subspecies of *Artemisia tridentata* by Wang *et al.* (1997). Reciprocal transplant experiments in three natural gardens (two parental and one hybrid environment) involving seeds collected from parental and natural hybrid sites were used to establish differences in various fitness components such as germination rate, survivorship, growth and reproduction (Wang *et al.*, 1997). The analyses revealed that the fittest type in each planting, including the hybrid site, was the type that naturally grows there, suggesting that selection was environmentally dependent (Wang *et al.*, 1997). Similarly, reciprocal transplant experiments of hybrids between *Ipomopsis aggregata* and *I. tenuituba* combined with the examination of pollination and other fitness components indicated that the survival of hybrids depended on both genotype and environment (Campbell, 2003).

Results from the previous chapter revealed that four genotypic classes were present in the studied *Geum* hybrid zone, that is 'pure' *G. urbanum* and *G. rivale,* F1s and Backcross Rivale but that other classes such as F2 and Backcross Urbanum appeared to be missing. What causes this obvious punctuation of the range of potential genotypes and phenotypes – are the missing classes due to lack of production, or lack of establishment? To address this over-arching question the following questions were addressed:

1. Can all four genotypic classes present in the hybrid swarm produce offspring and give rise to fertile seeds? If so, do germination rates differ between classes?
2. What is the diversity of the offspring generation? When the offspring from these seeds are grown up, is there only evidence for the four genotypic classes found in the adult population, or is there a greater continuum of genotypes and phenotypes?
3. What is the fitness of the offspring? If there is greater diversity in the offspring generation than in the established adults in the field, is there evidence for intrinsic fitness problems in some offspring genotypic classes that might explain their absence in the established adults?

The results of these investigations are then used to interpret the current genetic composition of the *Geum* hybrid swarm and to predict possible long term evolutionary outcomes (Chapter 8).

# 7.2 Materials and Methods

## 7.2.1 Sample material

The study investigated the same hybrid population from Edinburgh described in Chapter 6.2.1 (p 120). To establish whether the limited number of observed hybrid categories in the hybrid zone can be accounted for by restrictive mating behaviour, open pollinated seeds were collected from 30 individuals covering the whole morphological range in the hybrid zone in August 2006 (Table 7.1). The maternal plants were those used for the morphological and molecular analyses in the previous chapter and belonged to the following four categories: *G. urbanum* like, F1, Backcross Rivale and *G. rivale* like.

| Class | n | Total number sown | Range of seed numbers sown per parent |
|---|---|---|---|
| ***G. urbanum* like** | 6 | 558 | 65 to 113 |
| **F1** | 8 | 320 | 17 to 58 |
| **Backcross Rivale** | 7 | 333 | 38 to 93 |
| ***G. rivale* like** | 9 | 586 | 24 to 115 |

**Table 7.1.** Total number of seeds and range of seeds sown per parent in each genotypic class using n parents.

Seeds were sown in pots and overwintered outside. Once germinated five seedlings per maternal plant were randomly selected and grown on in an unheated greenhouse for one year. In autumn 2007 all five seedlings per maternal parent were planted out in a fully randomised design into a common nursery environment. Seedlings were planted 15 to 20 cm apart in a bed 12.4 m x 1.30 m x 0.2 m filled with a mixture of silty loam, grit and compost (Figure 7.1).

## 7.2.2 Germination rates

To assess whether all four genotypic classes gave rise to fertile seeds, germination rates were established in spring 2007 as the number of germinated seeds over the number of seeds sown. Differences in germination rates between the four maternal genotypic groups (*G. urbanum* like, F1, Backcross Rivale and *G. rivale* like) were tested for significance with ANOVA after the arcsine transformation of the data set.

**Figure 7.1.** Common nursery environment for five offspring per maternal parent grown from open pollinated seeds collected in the hybrid swarm (Edinburgh) from 30 mothers.

## 7.2.3 Diversity of offspring generation

*Morphological data*

To analyse and compare the phenotypic variation of parental and offspring groups produced in the hybrid zone, all plants in the parent and offspring groups were scored for the same nine phenotypic characters used for the morphological analysis in pure (Chapter 2.2.1, p 28) and hybrid (Chapter 6.2.3, p 124) populations: length and width of stipule, petal and sepal in centimetres, and the length of the petal claw, gynophore and upper style joint in millimetres (Figure 2.1, p 31). The data were then subjected to a principal component analysis (PCA) including reference samples from pure *G. urbanum* (39 plants), pure *G. rivale* populations (37 plants) and controlled F1 crosses (32) used in previous chapters (Chapter 5.2.1, p 94). To establish the extent to which progeny differ from their parents within the four genotypic classes, means and variances of the first principal component value (PCA1) for parents and offspring were calculated.

*Molecular data*

To establish the genetic structure of the progeny produced in the hybrid swarm DNA was extracted from five offspring per maternal plant and subjected to AFLP fingerprinting as described in Chapter 2.2.5 (p 31). The maternal parents were the same plants used in the

morphological analysis (see above). Mothers and progeny were scored for the same 203 AFLP markers used in Chapter 6.2.2 (p 123). To minimize within family error rates, maternal parents and their offspring were screened in the same AFLP run. To establish the extent to which progeny differ from their parents within the four genotypic classes, means and variances of the first principal coordinate value (PCO1) for parents and offspring were calculated (excluding pure and F1 reference samples). Differences between means and equality of variances were tested for significance using a Kruskal Wallis and an F test, respectively.

*Morphological versus genetic data*

To establish whether the observed high correlation between the morphological and genetic data sets in the adult generation is also present in the progeny (Chapter 6.3.2, p 134), the Spearman Rank Order Correlation was calculated between the first PCA component and the first PCO axis.

## 7.2.4 Offspring performance

To test for intrinsic fitness problems of offspring in the genotypic classes the performance parameter 'number of seed heads produced' was established by counting the number of full seed heads per offspring at the end of the fruiting season. To quantify the variation in plant size between offspring individuals and classes the volume occupied in $cm^3$, that is the area covered by the rosette leaves times the average leaf height above ground, was recorded at the end of the growing season in November 2008. Variation in 'number of seed heads produced' and 'volume occupied' between groups and between families within the same group was tested for significance with a Kruskal Wallis test. The correlation between the two parameters was estimated with a Spearman rank test. PCA and statistical tests were performed using the software MINITAB 14 (Minitab Inc., 2003).

# 7.3 Results

## 7.3.1 Germination rates

All groups produced fertile seeds with mean group germination rates ranging from 36.4% in the *G. urbanum* like class to 59.7% in *G. rivale* like class (Table 7.2). Germination rates did not differ significantly between the four genotypic classes ($F_{3,25}$=2.30, p=0.102).

| Class | n | Mean germination rate (%) | Standard error (%) |
|---|---|---|---|
| ***G. urbanum* like** | 6 | 36.4 | ±6.1 |
| **F1** | 7 | 50.7 | ±9.6 |
| **Backcross Rivale** | 7 | 41.9 | ±7.6 |
| ***G. rivale* like** | 9 | 59.7 | ±3.8 |

**Table 7.2.** Mean germination rates and standard error of seeds collected from four genotypic classes in the hybrid swarm.

## 7.3.2 Hybrid offspring diversity

*PCA analysis*

A PCA analysis of nine morphological characters measured in a common nursery environment was performed on all parents and five of their offspring. To highlight the morphological variation present in the offspring and their parents of the four genotypic groups (*G. urbanum* like, F1, Backcross Rivale and *G. rivale* like), each group was then subjected to a PCA analysis separately but included pure populations for reference purposes (pure *G. urbanum* and *G. rivale* populations as well as controlled F1s).

1. Parents and their offspring

All nine characters contributed more or less equally to the differentiation of the samples along the first component (68.1% of the variation explained). Stipule length was the single most important trait to distinguish samples along the second component (9.3% of the variation explained). The phenotypic spectrum of the offspring covered a wider range than the parental one (Figure 7.2).

Offspring of all genotypic classes apart from the F1 class had significantly different mean positions along PCA1 compared to their parents (Table 7.3). *G. rivale* like and Backcross Rivale offspring were shifted towards the right of the average parental position (i. e. towards the F1 group). The mean position of *G. urbanum* offspring was located to the right of the

parental position, which means away from the F1 group and towards the more extreme end of pure *G. urbanum*. Offspring of the F1 and *G. urbanum* like group possessed significantly higher variability than their parents. In contrast, no significant difference in phenotypic variability was observed between Backcross Rivale and *G. rivale* like progeny and their parents (Table 7.3).

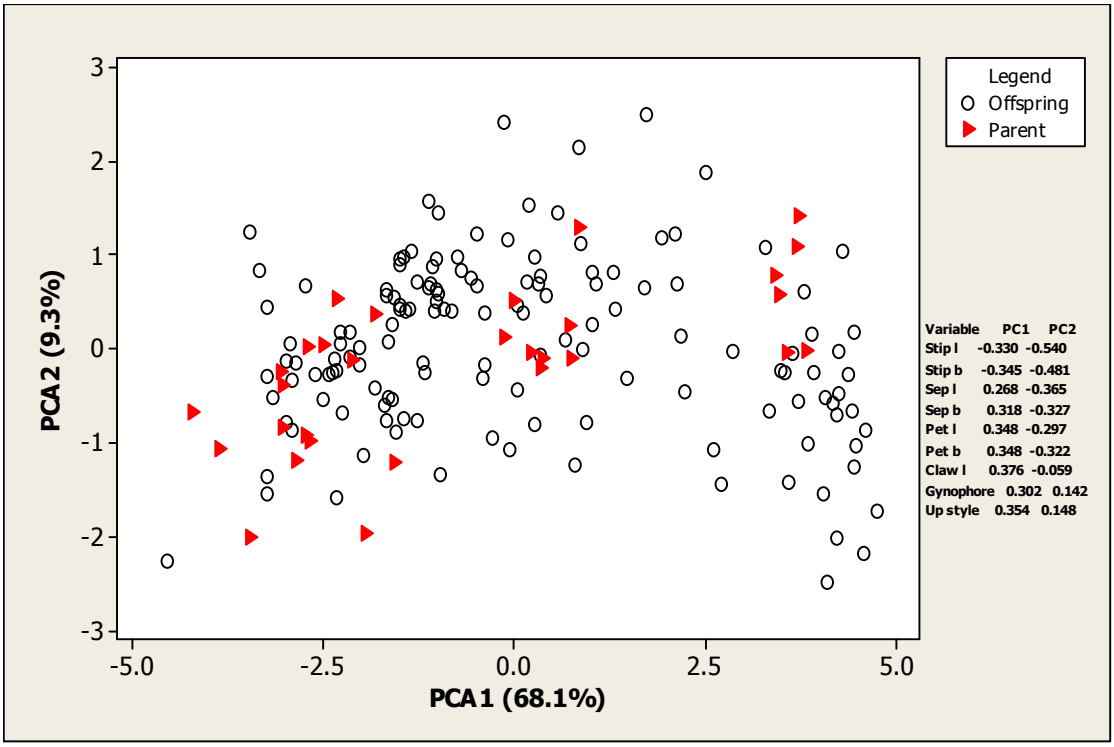



**Figure 7.2.** PCA using nine morphological traits (common environment) of parents from four genotypic classes plus five of their open pollinated offspring collected in the hybrid zone. Inset: weighting of characters.

| | PCA1 | | | |
|---|---|---|---|---|
| **Class** | **Mean** | **p** | **Variance** | **p** |
| ***G. urbanum* like** | 3.59 | **0.04** | 0.0178 | **0.01** |
| *G. urbanum* like offspring | 3.93 | | 0.2616 | |
| **F1** | 0.41 | 0.70 | 0.1340 | **0.00** |
| F1 offspring | 0.45 | | 1.6180 | |
| **Backcross Rivale** | -2.35 | **0.01** | 0.4470 | 0.18 |
| Backcross Rivale offspring | -1.17 | | 1.3280 | |
| ***G. rivale* like** | -3.00 | **0.00** | 0.5320 | 0.45 |
| *G. rivale* like offspring | -1.92 | | 0.8920 | |

**Table 7.3.** Principal component axis 1 (PCA1) means and variances of parents from four different hybrid classes and their offspring. Differences between parental and offspring values were tested for significance (p) with a Kruskal Wallis test (means) and an F test (equality of variances).

## 2. Offspring from *G. urbanum* like parents

All offspring produced by the six maternal parents grouped with the pure *G. urbanum* reference samples along the first component, which explained 74.7% of the variation (Figure 7.3). About half of the progeny were scattered above the reference group along the second component, which accounted for 11.3% of the variation.

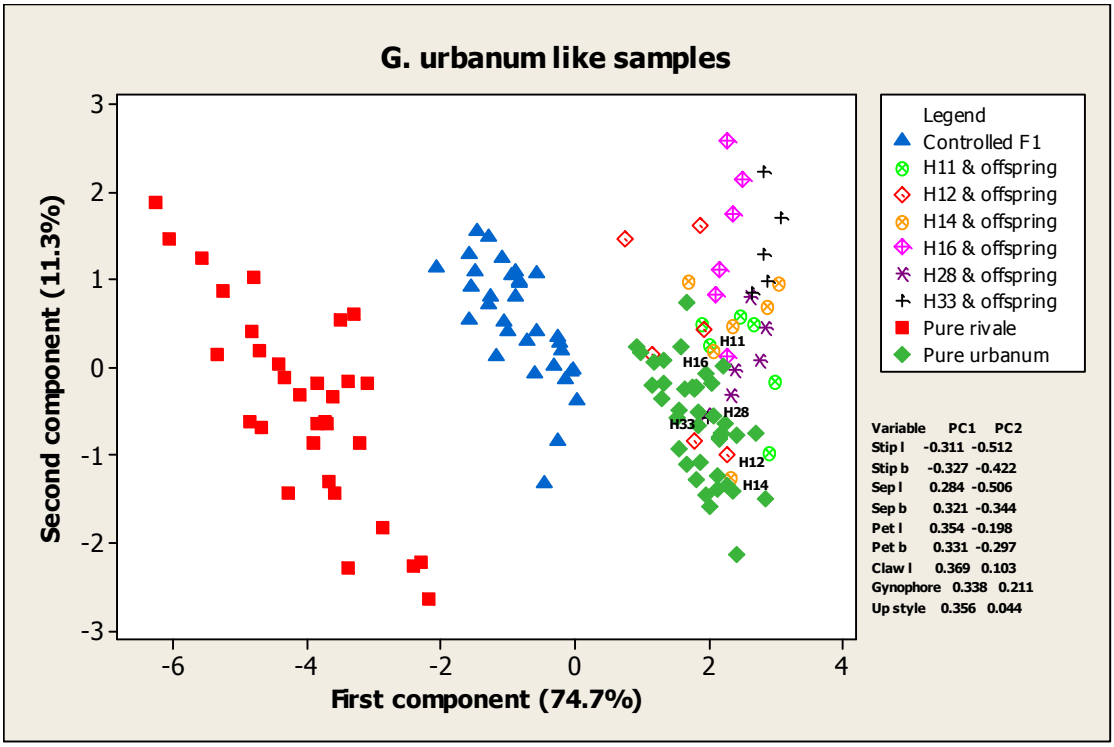



**Figure 7.3.** PCA using nine morphological traits (common environment) of six *G. urbanum* like individuals (labelled) collected in the hybrid zone plus five of their open pollinated offspring. Solid shapes represent individuals from reference populations, empty symbols represent samples from the hybrid zone. Inset: weighting of characters.

## 3. Offspring from F1 parents

Offspring of F1 maternal parents filled the phenetic space between the two parental species along the first component, which accounted for 67.5% of the variation (Figure 7.4). Most individuals were scattered in and around the controlled F1 group with some samples approaching either pure *G. rivale* or pure *G. urbanum* reference populations.

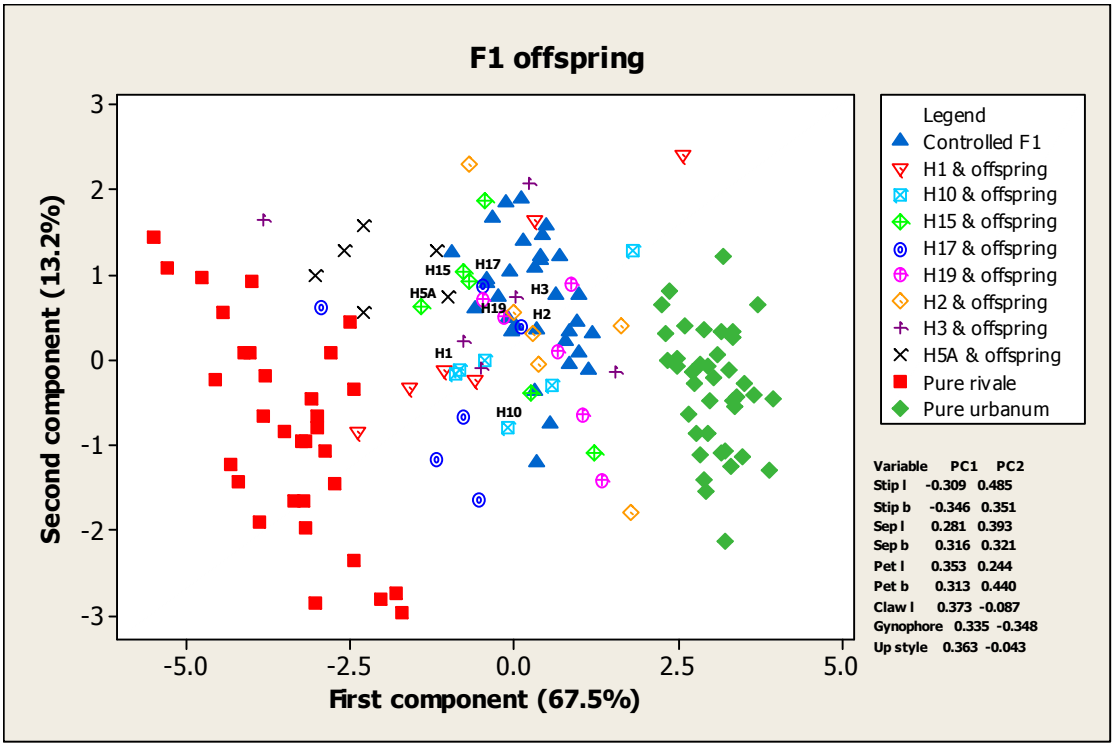



**Figure 7.4.** PCA using nine morphological traits (common environment) of eight F1 individuals (labelled) collected in the hybrid zone plus five of their open pollinated offspring. Further details see Figure 7.3.

## 4. Offspring from Backcross Rivale parents

The majority of the Backcross Rivale progeny grouped in and around the pure *G. rivale* reference populations with a few samples scattered close to the controlled F1 group (Figure 7.5). The first component accounted for 72.7% and the second for 12.5% of the variation.

## 5. Offspring from *G. rivale* like parents

All of the *G. rivale* like offspring grouped in and around the pure *G. rivale* reference populations along the first component which accounted for 73.8% of the variation (Figure 7.6).

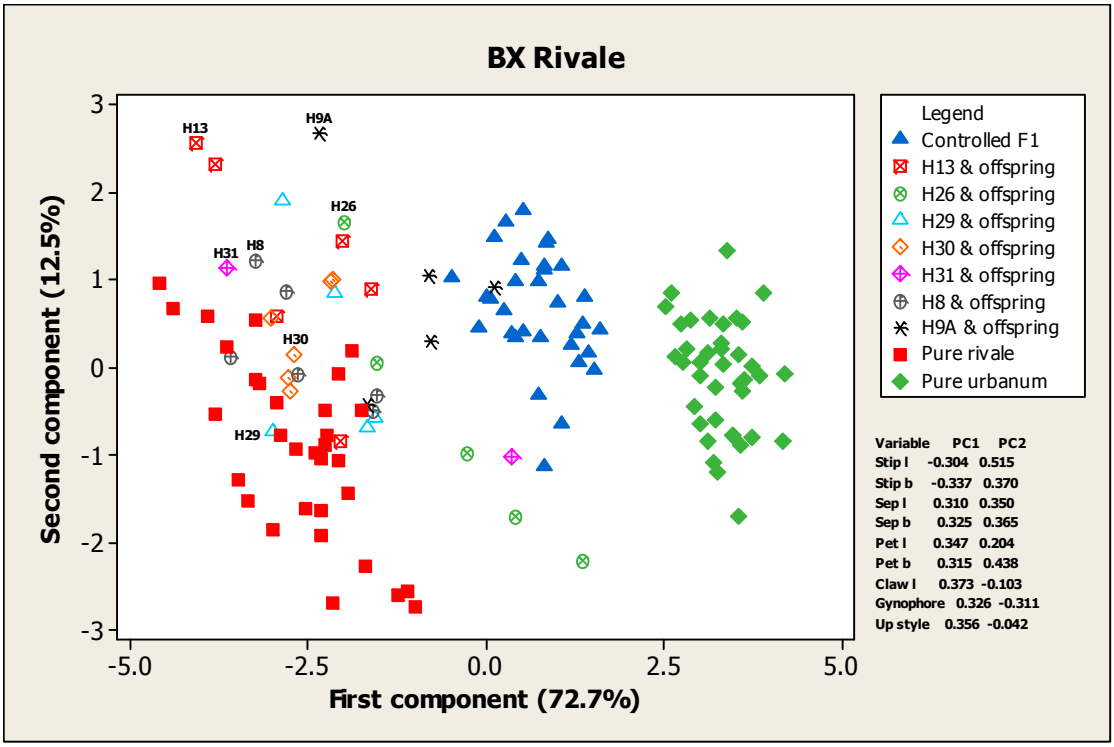



**Figure 7.5.** PCA using nine morphological traits (common environment) of seven Backcross Rivale individuals (labelled) collected in the hybrid zone plus five of their open pollinated offspring. Further details see Figure 7.3.

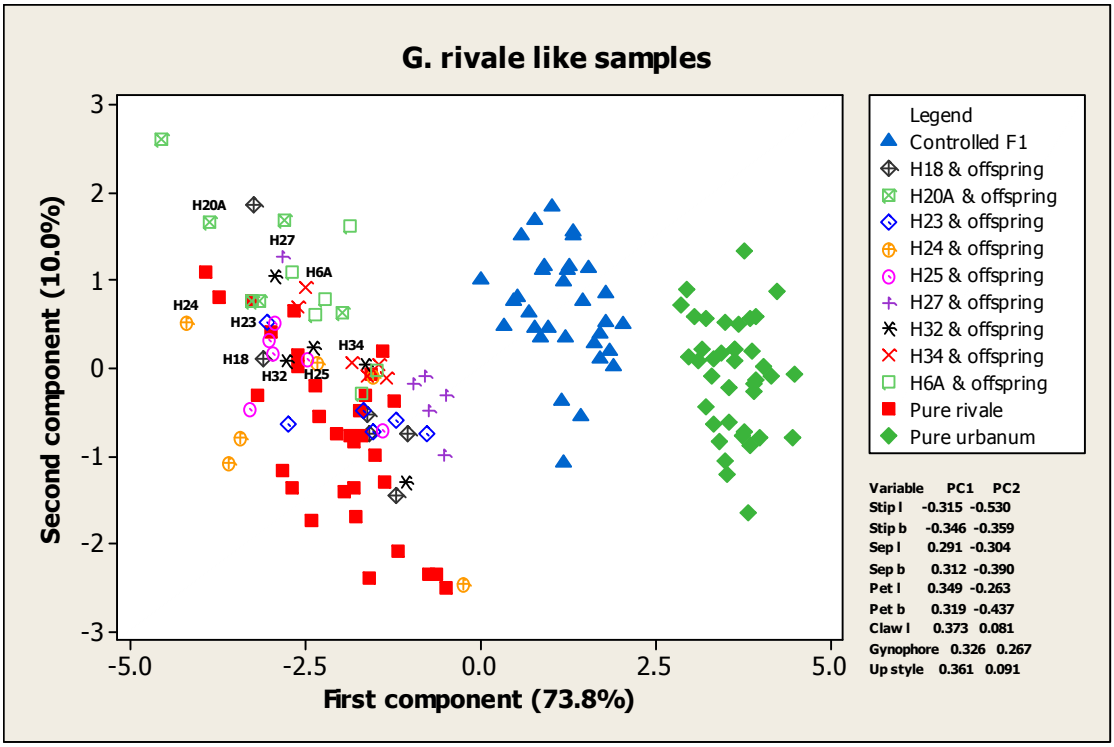



**Figure 7.6.** PCA using nine morphological traits (common environment) of nine *G. rivale* like individuals (labelled) collected in the hybrid zone plus five of their open pollinated offspring. Further details see Figure 7.3.

## *PCO analysis of AFLP data*

A PCO analysis using 203 AFLP markers was performed on 30 maternal parents used in the morphological analysis plus five of their open-pollinated offspring each. To highlight the genetic variation present in the offspring and their parents of the four genotypic groups (*G. urbanum* like, F1, Backcross Rivale and *G. rivale* like), each group was then subjected to a PCO analysis separately but included pure populations for reference purposes (pure *G. urbanum* and *G. rivale* populations as well as controlled F1s).

### 1. Parents and offspring

The PCO plot revealed that offspring filled the genetic space between the parents along axis 1 (PCO1), which accounted for 21.4% of the variation (Figure 7.7). Additionally, individuals clustered together in a group along axis 2 (PCO2, 11.4% of the variation explained) centred around PCO2=0.35. Individuals were separated along PCO1 by their species-specific markers, that is pure *G. urbanum* samples with 100% species-specific *G. urbanum* and 0% species-specific *G. rivale* markers were at one end of the plot, whereas pure *G. rivale* samples with 0% species-specific *G. urbanum* and 100% species-specific *G. rivale* markers could be found at the opposite end. F1s displaying about 100% of species-specific markers from either species were found inbetween the pure species although shifted slightly to the *G. urbanum* end as more species-specific *G. urbanum* (15) than *G. rivale* (9) markers were discovered.

Samples were separated along PCO2 according to the total number of bands with samples showing the highest number (F1s, average band number n=117) at the bottom and the lowest number (average band number n=40) at the top. None of the parents had PCO2 values > 0.1.

Individuals from the top group along PCO2 (with PCO2 values between 0.2<PCO2<0.5) displayed on average only 48.1% of the total number of maternal bands although individual percentages varied between 25% and 71% with the exception of H24/5 which possessed 97% of the bands. Not all of the four genetic classes contributed equally to the PCO2 top group (Table 7.4). The *G. rivale* like parental group contributed the highest percentage of offspring (40%) to the PCO2 top group whereas *G. urbanum* like and F1 classes contributed about 20% each. Apart from one offspring individual (3%) Backcross Rivale parents did not produce progeny in the PCO2 top group (Table 7.4).

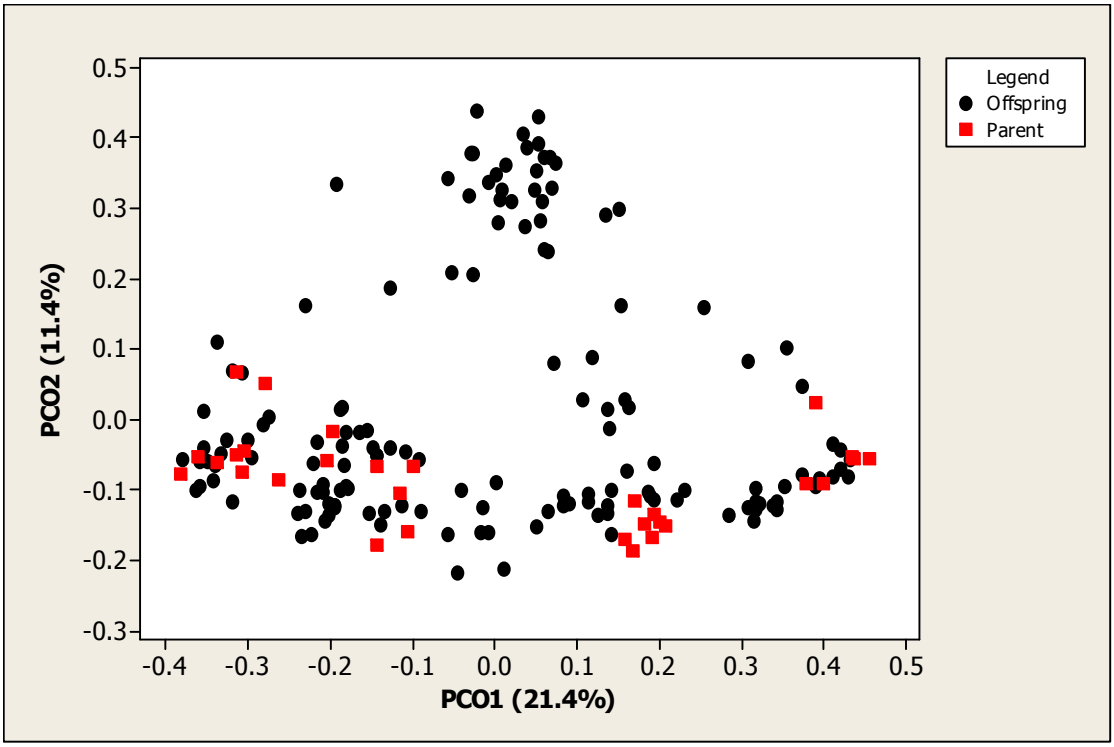



**Figure 7.7.** PCO plot of 30 maternal parents from the hybrid zone plus five of their open-pollinated offspring each using 203 AFLP markers

| | % PCO2 mothers | Total offspring in class | # progeny in PCO2 | % of PCO2 offspring |
|---|---|---|---|---|
| ***G. urbanum* like** | 33.3% | 30 | 6 | 20.0% |
| ***G. rivale* like** | 66.7% | 45 | 18 | 40.0% |
| **F1** | 37.5% | 40 | 7 | 17.5% |
| **Backcross Rivale** | 14.2% | 34 | 1 | 2.9% |
| **Total** | 40% | 149 | 32 | 21.5% |

**Table 7.4.** PCO2 top group statistics. % PCO2 mothers: percentage of maternal parents contributing offspring to the PCO2 group within each class, Total offspring in class: total number of offspring analysed, # progeny in PCO2: number of offspring in PCO2 top group within each class, % of PCO2 offspring: percentage of offspring in PCO2 group within in each class.

Offspring in the PCO2 top group were not randomly distributed among maternal genotypes (Table 7.4). If maternal parents gave rise to PCO2 top group progeny, then in most cases more than 50% of their offspring clustered with that group.

Numbers of species-specific markers varied among some offspring from the same family. This was especially apparent in progeny from F1 parents but also occurred in *G. urbanum* like and *G. rivale* like groups. Parents which appeared to be genetically pure *G. urbanum* and *G. rivale* produced offspring with fewer corresponding species-specific markers. For example, H11 (*G. urbanum* like group) had 100% of the species-specific *G. urbanum* markers and gave rise to offspring (H11/4 and H11/5) with only 66.7% of these markers. Similarly, some *G. rivale* like

parents with 100% of species-specific *G. rivale* markers produced progeny with lower species-specific marker frequencies (66.7% to 88.9%).

| | PCO1 | | | |
|---|---|---|---|---|
| **Class** | **Mean** | **p** | **Variance** | **p** |
| ***G. urbanum* like** | 0.42 | **0.00** | 0.0009 | **0.00** |
| *G. urbanum* like offspring | 0.29 | | 0.0191 | |
| **F1** | 0.18 | **0.01** | 0.0003 | **0.00** |
| F1 offspring | 0.07 | | 0.0185 | |
| **Backcross Rivale** | -0.17 | 0.14 | 0.0023 | 0.83 |
| Backcross Rivale offspring | -0.14 | | 0.0018 | |
| ***G. rivale* like** | -0.33 | **0.04** | 0.0011 | **0.00** |
| *G. rivale* like offspring | -0.15 | | 0.0376 | |

**Table 7.5.** Principal coordinate axis 1 (PCO1) means and variances of parents from four different hybrid classes and their offspring. Differences between parental and offspring values were tested for significance (p) with a Kruskal Wallis test (means) and an F test (equality of variances).

Comparison of the mean position and its variance along PCO axis 1 (PCO1) of parents and offspring from four different hybrid classes revealed that apart from the Backcross Rivale group all classes produced offspring which occupied a significantly different mean position to that from the parents along PCO1 and also showed significantly higher positional variability in this respect (Table 7.5).

## 2. Offspring from *G. urbanum* like parents

The majority of offspring produced from genotypically *G. urbanum* maternal parents clustered with the pure *G. urbanum* reference populations (Figure 7.8). However, two parents (H11 and H12) gave rise to progeny which fell between the pure *G. urbanum* and the F1 group, suggesting a Backcross Urbanum identity. H14 and H16 had offspring which fell in the top group along PCO2 with equally low numbers of species-specific markers from either parent (Figure 7.8).

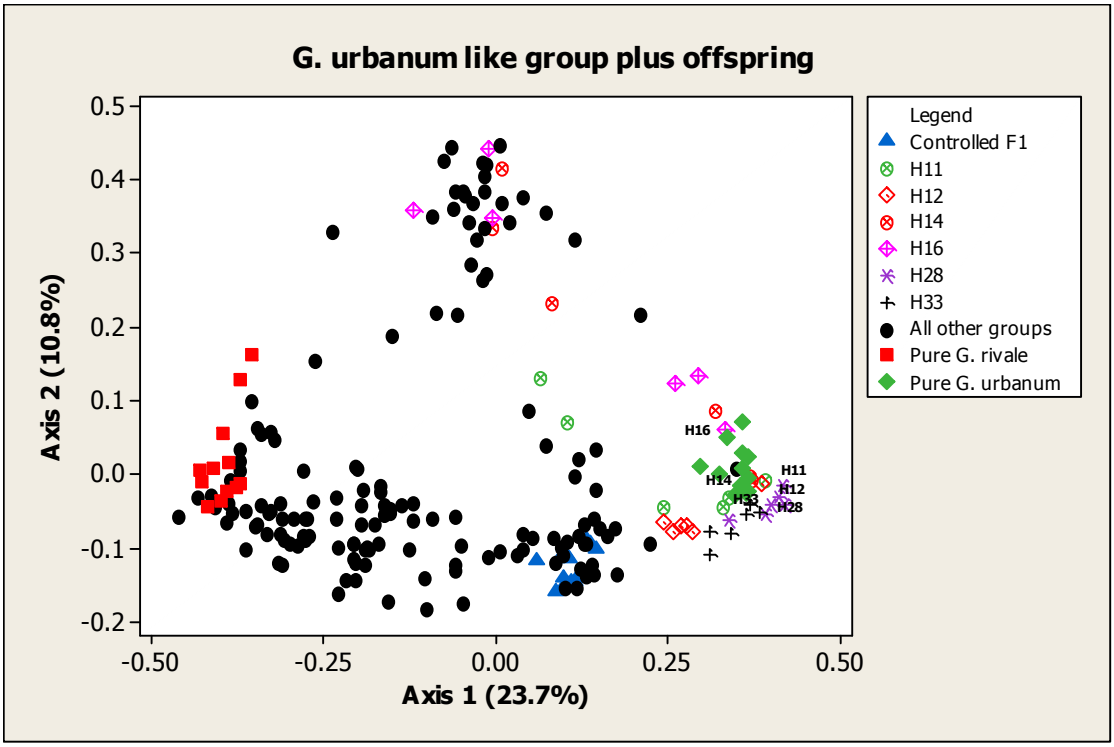



**Figure 7.8.** PCO plot of six *G. urbanum* like parents (labelled) plus five offspring each using 203 AFLP markers. For reference purposes pure *G. urbanum* and pure *G. rivale* populations, controlled F1s and the offspring of all other groups are included in the graph.

## 3. Offspring from F1 parents

Offspring from F1 maternal parents showed a large diversity of genetic backgrounds (Figure 7.9). Most individuals clustered with F1 controlled crosses and spread towards the Backcross Rivale group but three samples also exhibited a 'pure' parental genetic identity (two resembled *G. rivale* and one *G. urbanum*). Apart from one sample, there was little evidence for the production of Backcross Urbanum genotypes. Several progeny grouped with the top group along PCO2.

## 4. Offspring from Backcross Rivale parents

Nearly all offspring fell between the pure *G. rivale* and controlled F1 reference populations and clustered with the parental Backcross Rivale group. Only one sample was found outside this cluster close to the top group along PCO2 (Figure 7.10).

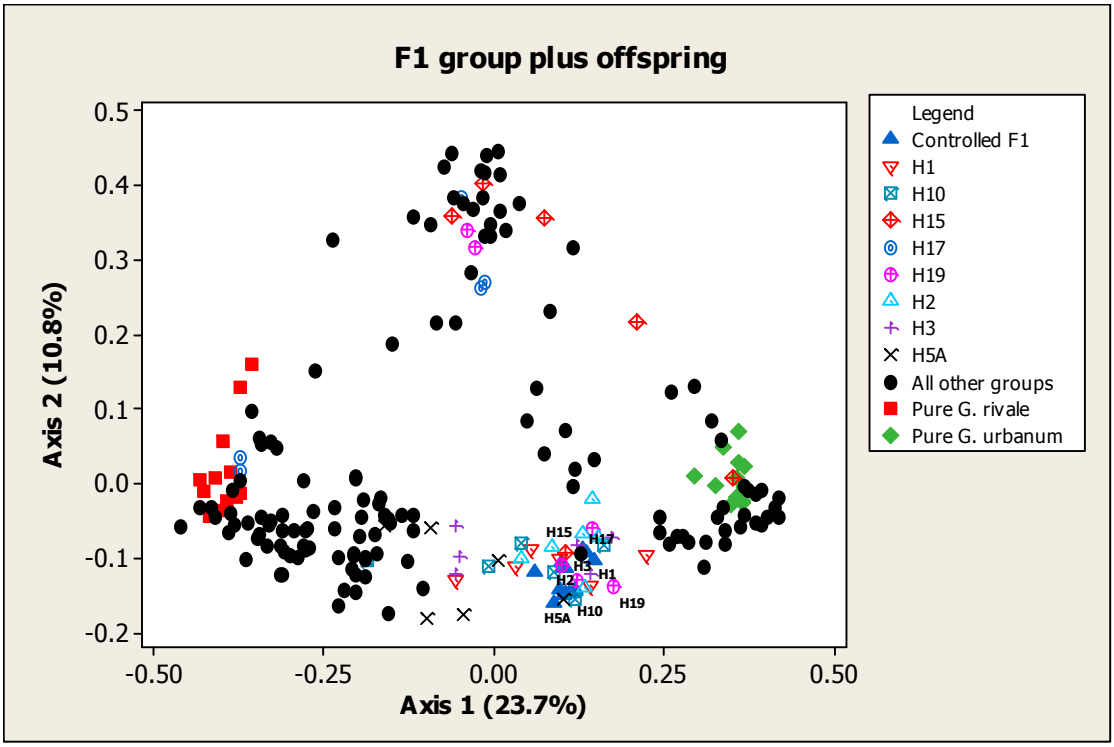



**Figure 7.9.** PCO plot of eight F1 parents (labelled) plus five offspring each using 203 AFLP markers. For reference purposes pure *G. urbanum* and pure *G. rivale* populations, controlled F1s and the offspring of all other groups are included in the graph.

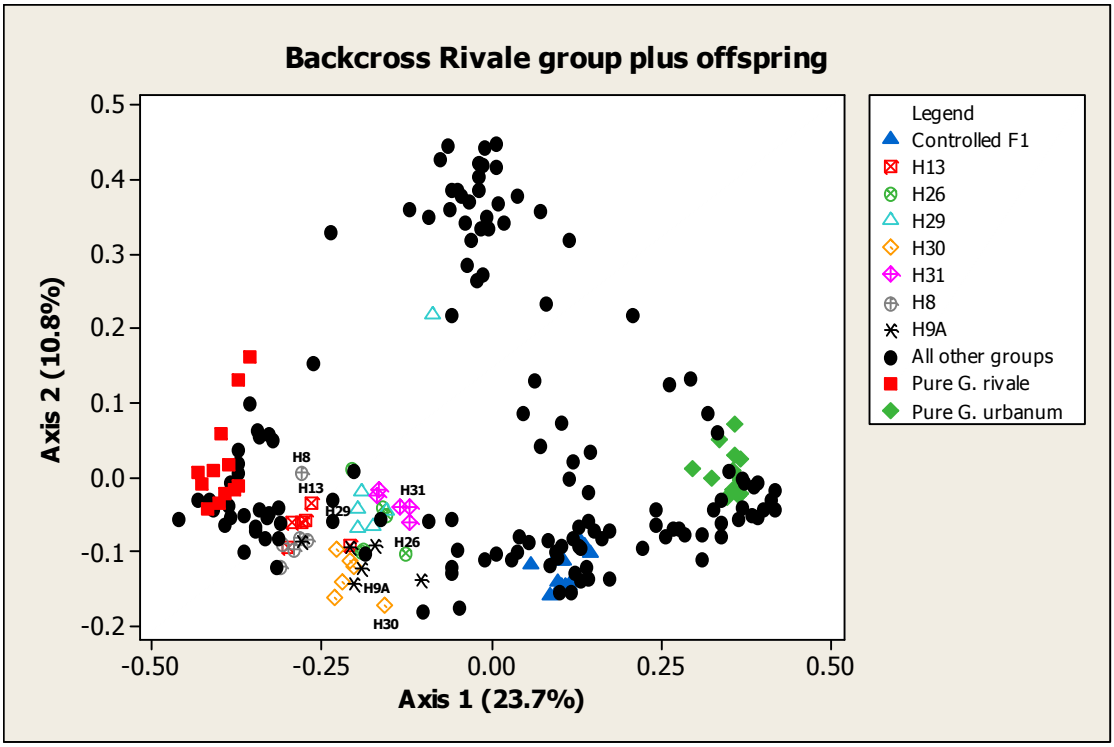



**Figure 7.10.** PCO plot of seven Backcross Rivale parents (labelled) plus five offspring each using 203 AFLP markers. For reference purposes pure *G. urbanum* and pure *G. rivale* populations, controlled F1s and the offspring of all other groups are included in the graph.

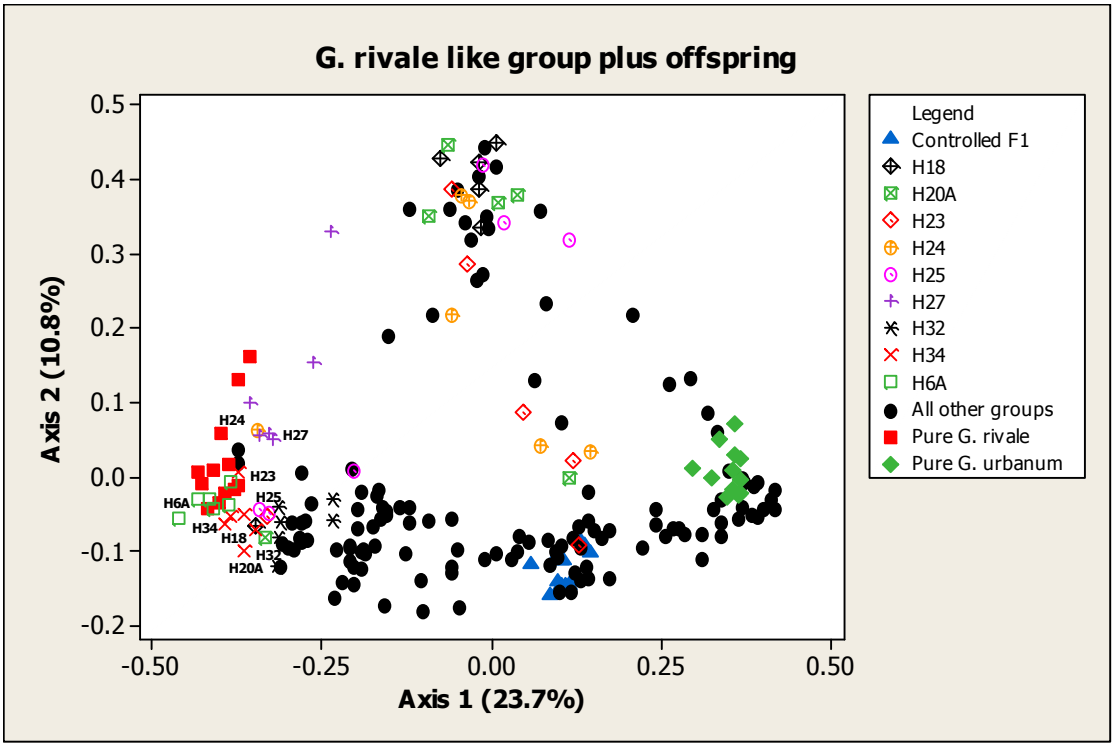



**Figure 7.11.** PCO plot of nine *G. rivale* like parents (labelled) plus five offspring each using 203 AFLP markers. For reference purposes pure *G. urbanum* and pure *G. rivale* populations, controlled F1s and the offspring of all other groups are included in the graph.

## 5. Offspring from *G. rivale* like parents

*G. rivale* like maternal parents produced offspring clustering in three different groups. Most samples grouped with the pure *G. rivale* and the PCO2 top group whereas six individuals (13%) grouped with (n=1) or were scattered just above (n=5) the controlled F1s (Figure 7.11).

### *Morphological versus genetic data*

There was a fairly low but significant correlation between PCO1 and PCA1 in the offspring generation (Spearman Rank Order Correlation = 0.642; $p < 0.0001$; 41% of the variation explained) which contrasted with the high correlation between PCO1 and PCA1 in the adult generation (Spearman Rank Order Correlation = 0.915; $p < 0.0001$; 84% of the variation explained). Some major discrepancies between morphological and genetic data were observed in some samples (Figure 7.12). *G. rivale* offspring which scattered above the controlled F1 group genetically, clustered phenotypically with the pure *G. rivale* group. One individual (H15/1) appeared to be a pure *G. urbanum* in the PCO plot but morphology suggested an F1 phenotype whereas H9A/3 grouped with the Backcross Rivale cluster but looked like an F1 (Figure 7.12).

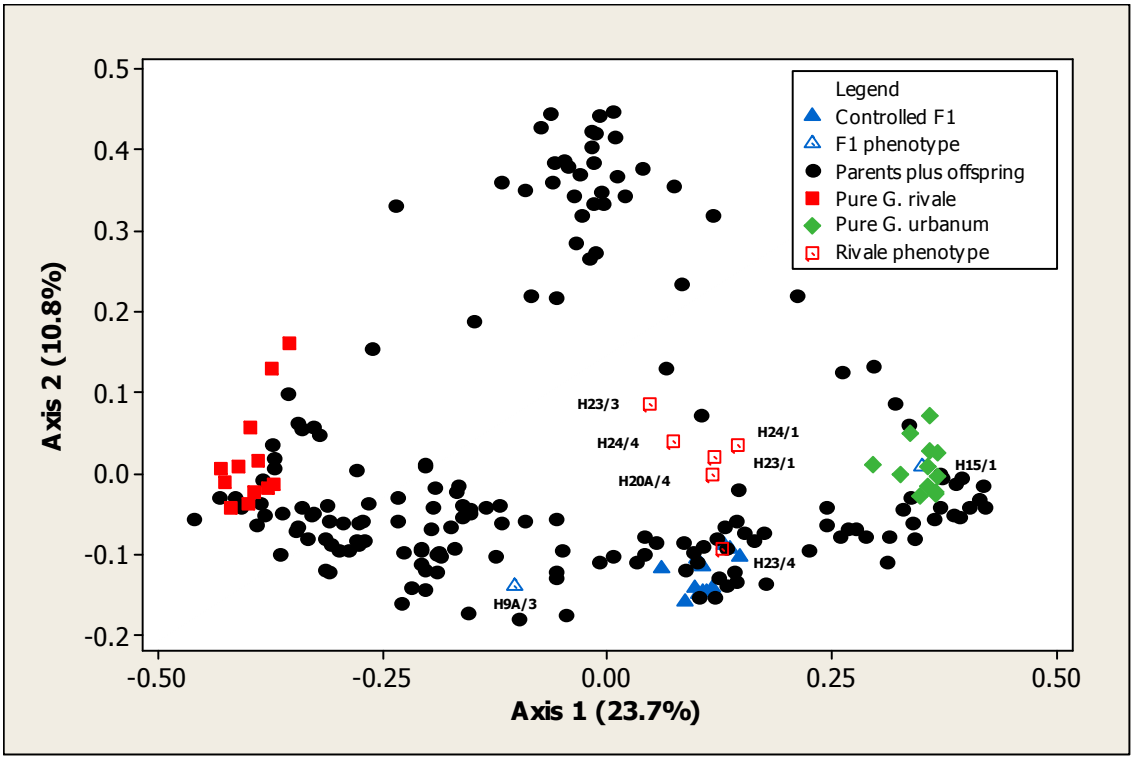



**Figure 7.12.** PCO plot of 30 maternal parents from the hybrid zone plus five of their open-pollinated offspring each using 203 AFLP markers. Samples where major differences between morphological and genetic data were observed are labelled.

## 7.3.3 Offspring performance

All progeny planted into a common environment survived until measurements were taken. Offspring from *G. urbanum* like maternal parents produced about 20 times more seed heads and occupied only 1/13 of the volume of progeny from *G. rivale* like parents. Offspring from F1 and Backcross Rivale parents lay between those extremes. A Kruskal Wallis test detected a significant difference in the mean seed head number ($H=100.66$, $df=3$, $p<0.0001$) and the mean volume occupied ($H=64.17$, $df=3$, $p<0.0001$) between progeny from the four different genetic classes in the hybrid zone (Table 7.6). Across maternal genotype groups there was a negative but not significant correlation (Pearson correlation factor = -0.873, $p = 0.127$) between the values for the two mean performance measurements (Figure 7.13). The non significant correlation is likely to be due to the small data set of four values. A scatter plot for 143 offspring collected from 30 maternal parents in the hybrid zone (Figure 7.14) revealed that progeny from *G. urbanum* like parents varied comparatively little in terms of volume occupied but greatly in number of seed heads produced (25 to 230). In contrast, offspring from *G. rivale* like parents showed large variation in volume occupied but little variability in seed heads produced (0 to 21).

| Family | Seed heads | | Volume occupied ($cm^3$) | |
|---|---|---|---|---|
| | Mean | SE | Mean | SE |
| ***G. urbanum* like** | **126.3** | **±10.0** | **887** | **±138** |
| **H11** | 72.4 | ±20.0 | 582 | ±125 |
| **H12** | 149.4 | ±28.8 | 480 | ±99.1 |
| **H14** | 138.8 | ±13.2 | 758 | ±129 |
| **H16** | 109 | ±14.7 | 1025 | ±140 |
| **H28** | 199.2 | ±5.2 | 1854 | ±660 |
| **H33** | 89 | ±9.2 | 623 | ±96 |
| **F1** | **53.3** | **±7.4** | **5105** | **±743** |
| **H1** | 68.8 | ±41.2 | 7840 | ±2773 |
| **H2** | 60 | ±9.1 | 3345 | ±865 |
| **H3** | 75.4 | ±21.2 | 7023 | ±2673 |
| **H5A** | 72.8 | ±19.5 | 11406 | ±1293 |
| **H10** | 30.6 | ±17.7 | 4529 | ±1169 |
| **H15** | 23.3 | ±4.8 | 1116 | ±491 |
| **H17** | 32.4 | ±5.4 | 2516 | ±765 |
| **H19** | 57.4 | ±19.1 | 2262 | ±1035 |
| **Backcross Rivale** | **12.8** | **±1.9** | **5929** | **±1005** |
| **H8** | 10.8 | ±1.6 | 4811 | ±748 |
| **H9A** | 11.8 | ±4.2 | 2602 | ±1398 |
| **H13** | 16.2 | ±4.9 | 12404 | ±2942 |
| **H29** | 12.3 | ±3.0 | 4482 | ±1302 |
| **H30** | 23.6 | ±5.3 | 10664 | ±1190 |
| **H31** | 0.2 | ±0.2 | 325 | ±56.8 |
| ***G. rivale* like** | **±6.0** | **±0.6** | **11517** | **±933** |
| **H6A** | 7.4 | ±1.5 | 12115 | ±4651 |
| **H18** | 5.2 | ±1.2 | 13022 | ±1969 |
| **H20A** | 7.4 | ±1.7 | 18014 | ±1735 |
| **H23** | 10.6 | ±3.2 | 14609 | ±1575 |
| **H24** | 6.8 | ±2.1 | 11863 | ±3241 |
| **H25** | 4 | ±0.5 | 7920 | ±1408 |
| **H27** | 5 | ±1.1 | 4449 | ±748 |
| **H32** | 3.2 | ±0.9 | 12057 | ±3104 |
| **H34** | 4.8 | ±1.7 | 9606 | ±1459 |

**Table 7.6.** Basic statistics of the performance parameter 'number of seed heads' and 'volume occupied' for all families collected in the hybrid swarm within each of the four classes.

| Class | df | H (Seed heads) | H (Volume occupied) | Spearman correlation |
|---|---|---|---|---|
| ***G. urbanum* like** | 5 | **16.58**** | **12.47*** | **0.451*** |
| **F1** | 7 | 11.95 | **20.12**** | 0.305 |
| **Backcross Rivale** | 5 | **14.23*** | **14.23*** | **0.729***** |
| ***G. rivale* like** | 8 | 12.36 | **18.33*** | **0.404**** |

**Table 7.7.** Kruskal-Wallis test results for offspring variation between families within four genetic groups for number of seed heads produced and volume occupied and the Spearman rank correlation. *$0.01 \leq p < 0.05$, **$0.001 \leq p < 0.01$, ***$p<0.001$.

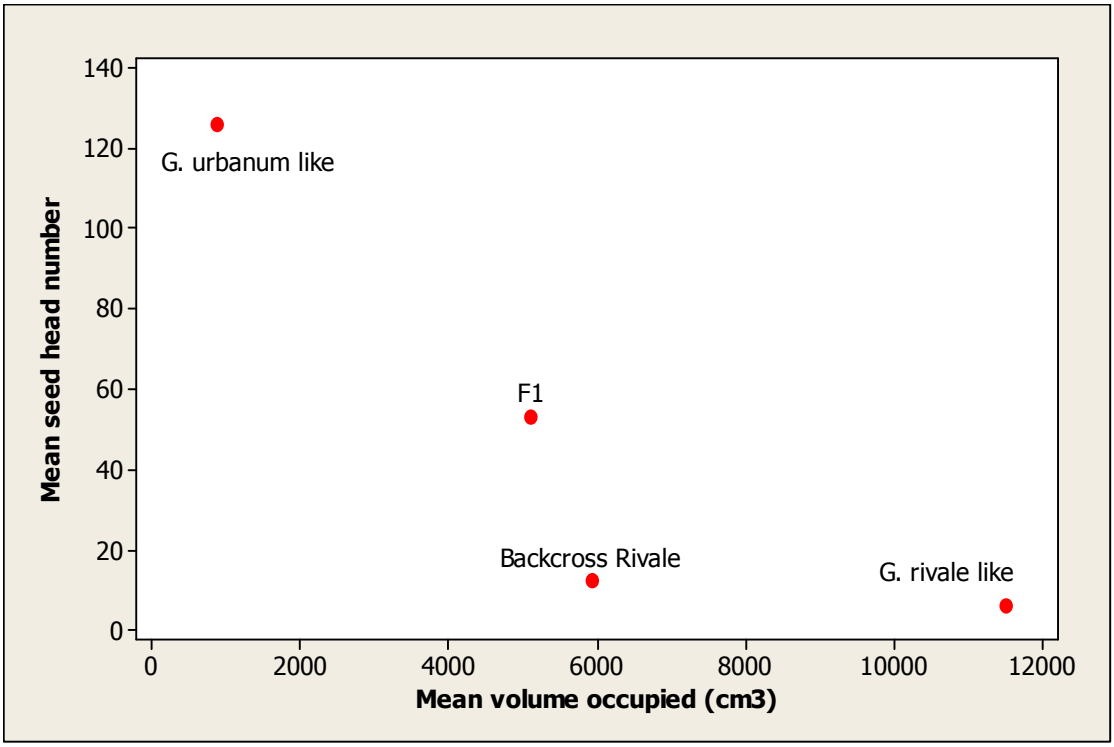



**Figure 7.13.** Scatterplot of the mean seed head number versus the mean volume occupied of progeny from the four genetic classes detected in the hybrid zone.

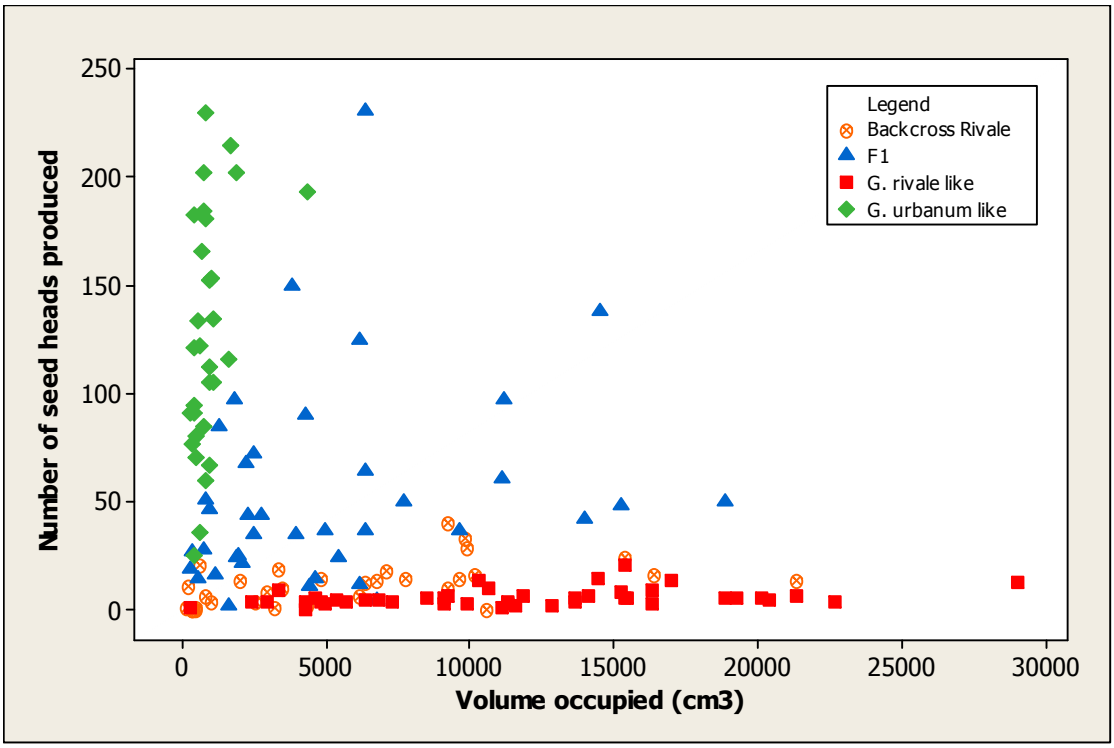



**Figure 7.14.** Scatter plot of volume occupied versus number of seed heads produced for about five open-pollinated offspring (n=143) per maternal parent (n=30) collected in the hybrid zone. Progeny is coded according to the genetic class of the maternal parent.

Progeny from F1 maternal parents scattered between those two extremes and showed a large number of novel combinations, which were not observed in the parental groups. Progeny from the Backcross Rivale parents displayed variation similar to the *G. rivale* like group but also showed a range of combinations which did not occur in that group (Figure 7.14). Significant differences in the number of seed heads produced were found between families of progeny from both *G. urbanum* like and Backcross Rivale maternal parents (Table 7.7). No such difference was discovered among families from F1 or *G. rivale* like maternal parents. In contrast, there was significant variation for volume occupied among the families derived from all four classes of maternal genotypes (Table 7.7).

Although no significant differences in seed production were discovered among families from F1 maternal parents, considerable within family variation was observed. For example, offspring of maternal parent H10 differed in their seed head output nearly 50-fold (H10/3 produced 2 seed heads, H10/1 97) (Table 7.6). No F1 family produced offspring with no or consistently low seed head numbers. In contrast, four individuals in the Backcross Rivale family H31 did not produce any seed heads and only one gave rise to one seed head. Additionally, all of the offspring from maternal parent H31 had the lowest values for 'volume occupied' within the Backcross Rivale group. There was a positive correlation between seed head numbers and volume occupied within all maternal genotype groups although this correlation was not significant in the F1 group and did not exceed a value of 0.729 (Backcross Rivale offspring) in any of the four groups (Table 7.7). Overall no evidence for intrinsic selection was found between the groups.

## 7.3.4 Summary of the main results

All four genotypic classes found in the hybrid swarm produced viable seeds that did not differ significantly in their germination success. The morphological and genotypic variation found in the offspring generation was much greater than that observed in the parental generation growing in the hybrid swarm. Although there was a significant correlation between the genetic (PCO1) and phenotypic (PCA1) measures of variation in the progeny generation, only half the variation was explained by this relationship compared to the equivalent parental data set. Significant differences were discovered for the number of seed heads produced and the volume occupied by offspring between the four genotypic groups in the hybrid swarm. Offspring from *G. urbanum* like maternal parents had 20 times the number of seed heads produced by *G. rivale* maternal parents, but occupied only $1/13^{th}$ of the volume. Against this background of different life history strategies between the two parents no consistent evidence for intrinsic fitness problems in any of the offspring genotypic classes was detected.

# 7.4 Discussion

Prior to discussing the implications of the results in the light of the chapter aim, I first look at the unexpected separation of some offspring samples along the second axis of the PCO plot.

## 7.4.1 Separation along PCO2 (PCO2 top group)

The genotypic composition of some progeny revealed a curious phenomenon. About 20% of offspring showed a circa 50% reduction in AFLP bands of parental genotypes from which they were derived. This led to their separation along the top of the PCO2 axis. Offspring in the PCO2 top group resembled the parental phenotypic class (e. g. individuals with a *G. urbanum* like maternal parent looked like *G. urbanum*). Three (*G. urbanum*, *G. rivale* and F1) out of the four genetic classes encountered in the hybrid zone contributed offspring to the PCO2 top group which resulted in great phenotypic heterogeneity despite the genetic similarity of the samples in this group.

It is difficult to explain how offspring can loose on average 50% (and in individual cases up to 75%) of the bands present in the maternal parent. PCR artefacts are a possible cause. However, this does not seem to be a likely explanation. AFLP runs were carried out on two 96 well plates with families (maternal parent plus five offspring) included in the same assay. If the loss of bands was due to PCR artefacts then we would expect to see fewer bands in individuals randomly distributed throughout the four genotypic classes. This was not the case as the Backcross Rivale parents contributed only one individual to the PCO2 top group whereas the three other classes contributed between 17.5% and 40% of their offspring. The generation of PCO2 offspring was also restricted to certain maternal parents and was not randomly distributed among parental genotypes (Table 7.4) so that on average more than 50% of the progeny from mothers with PCO2 top group offspring clustered in this group. Additionally, none of the maternal parents showed a substantial loss of bands as their AFLP phenotypes matched the phenotypes of previous runs within the expected error rate (Chapter 6.3.1, p 128).

Assuming that PCR artefacts are not a likely explanation for the elimination of bands, what else could account for it? It is possible that the loss of bands was caused by the elimination of chromosomes during hybridization. Jørgensen (1928) reported the phenomenon of pseudogamy in the genus *Solanum* where pollen tubes grew normally, but sperm failed to fuse with the egg cells and degenerated in the cytoplasm of the eggs which resulted in the production of a haploid embryo. A similar case of abnormal fertilization is called semigamy. The male nucleus penetrates the egg cell but does not fuse with the egg nucleus, leading to

the formation of an embryo with separate tissue of maternal or paternal origin which can result in haploid offspring (Coe, 1953; Turcotte & Feaster, 1967).

The complete or partial elimination of the chromosomes from one parent during the mitotic divisions of the zygotes after normal fertilization is a well documented albeit poorly understood phenomenon (Ho & Kasha, 1975). In the genus *Hordeum* the loss of *H. bulbosum* chromosomes from the interspecific hybrid (2n=14) between tetraploid (2n=28) *H. vulgare* and tetraploid *H. bulbosum* has long been known (Kasha & Kao, 1970; Barclay, 1975; Ho & Kasha, 1975). Chromosome counts carried out one to six days after pollination indicated that fertilization was normal. However, at a later stage the presence of micronuclei in both embryo and endosperm tissues as well as lagging chromosomes in endosperm divisions showed the subsequent elimination of *H. bulbosum* chromosomes (Barclay, 1975). Hybrids with the haploid complement from one parent and one to several chromosomes from the other parent were observed in crosses between oat and maize (Riera-Lizarazu *et al.*, 1996) as well as between cultivated *Brassica* species and another crucifer *Orychophragmus violaceus* (Li & Ge, 2007). It is possible that a similar process might have caused the loss of chromosomes (and thus AFLP bands) in some offspring from the *Geum* hybrid zone.

An alternative explanation for the substantial loss of bands in hybrid offspring could be DNA methylation, the addition of a methyl group to the C5 position of a cytosine residue, which is the most common DNA modification in plants and animals (Xu *et al.*, 2000; Cervera *et al.*, 2003). It has been reported that changes in DNA methylation frequently occur in the genes of newly formed allopolyploids (Comai *et al.*, 2000; Shaked *et al.*, 2001). Restriction enzyme isoschizomers like *Hpa*II and *Msp*I differ in their sensitivity to methylation at their recognition sequences (5'-CCGG-3') and are frequently used to detect cytosine methylation as they cannot cleave if one or both cytosines are methylated (Cervera *et al.*, 2003). Studies in *Cucumis* revealed that methylation changed at 2.0% to 6.4% of total sites in both the F1 and the allotetraploid hybrids compared with their corresponding parents (Chen & Chen, 2008). Similarly, F1 or allopolyploid hybrids in the genus *Aegilops* showed alterations in cytosine methylation in 13% of the loci (Shaked *et al.*, 2001).

Although *Eco*RI, the restriction enzyme employed in the AFLP assay of this study, is commonly assumed to be methylation insensitive (Kale & Abbott, 2008), its activity has been shown to be impaired or blocked by 5-methyl-C residue (http://rebase.neb.com/rebase/rebase.html). It is therefore possible that extensive methylation in some of the offspring individuals has influenced the ability of *Eco*RI to cut at the recognition sequence and thus led to a significant loss of AFLP bands. Another mechanism which could result in the loss of AFLP bands is sequence elimination. Shaked *et al.* (2001) reported that sequence elimination is one of the major responses of the wheat

genome to hybridization or allopolyploidy, which can affect a large fraction of the genome. In a cross between *Aegilops sharonensis* and *A. umbellulata* 14% of the loci from *A. sharonensis* and 0.5% from *A. umbellulata* were eliminated with most changes occurring in the F1 hybrids.

It is possible that a combination of DNA methylation and sequence elimination could result in a significant loss of AFLP bands in hybrid offspring. However, considering the magnitude of the two processes reported in the literature (see above), it seems unlikely that the combined effect of DNA methylation and sequence elimination could account for the loss of 50% or more of parental bands in hybrid offspring.

With the data at hand it is not possible to explain the observed loss of AFLP bands in individuals from the PCO2 top group. Further work such as chromosome counts, methylation sensitive AFLP assays and sequencing are required to explore the cause for the loss of AFLP bands in offspring individuals from the *Geum* hybrid zone in more detail.

## 7.4.2 Hybrid offspring diversity

All genotypic classes gave rise to viable seeds and 100% of the offspring planted into a common nursery environment survived until performance measurements were taken. The major finding of this chapter is that the range of both phenotypes and genotypes present within the offspring was greater than that observed in the adult generation. The results from the various analyses are discussed for each of the four genotypic classes separately.

*Offspring from* G. rivale *like parents*

The majority of offspring from *G. rivale* like maternal parents clustered genetically in more or less equal numbers with the *G. rivale* reference populations and the PCO2 top group, whereas 13% of the progeny were scattered just above the controlled F1 group. Interestingly, none of the progeny which were genetically similar to F1s had the phenotype of F1s, suggesting that the production of individuals with an F1 genotype and phenotype by *G. rivale* like parents is rare.

Certain *G. rivale* like parents gave rise to offspring with fewer (two to three) species-specific *G. rivale* markers (excluding offspring from the PCO2 top group where many bands were missing). Some of the missing species-specific markers were different markers compared to the ones which were found to be heterozygous species-specific *G. rivale* markers in controlled F1 crosses (Chapter 6.3.1, p 128). One possible explanation for this observation could be that

a higher percentage of species-specific *G. rivale* markers is heterozygous. However, an alternative and probably more likely explanation is that the seemingly 'pure' *G. rivale* parents with a higher number of missing species-specific markers are actually advanced backcrosses to *G. rivale* or offspring with a *G. rivale* like genotype and morphology derived from F1 parents. This is supported by the observation in the hybrid zone that an F1 like parent (e. g. H17) can give rise to offspring with pure phenotypic and genotypic attributes within one generation (H17/3 and H17/4). Although 'pure' within the scope of the morphological and genetic methods applied in this study, these progeny are likely to possess a considerable number of species-specific markers in heterozygous form. Offspring from these seemingly pure parents could thus exhibit the observed segregating pattern of species-specific markers.

### *Offspring from* G. urbanum *like parents*

Individuals genetically intermediate between the pure *G. urbanum* reference populations and the controlled F1s were produced, indicating the generation of a Backcross Urbanum class in the hybrid zone. As the PCA analysis showed, these samples could not be distinguished morphologically from pure *G. urbanum* plants. Marsden-Jones (1930) carried out a total of 94 crosses between F1 ("*G. intermedium*") and *G. urbanum* plants in both directions, which gave identical results independent from the cross direction. He commented that "with the exception of four plants […], the whole generation was superficially very like *urbanum*, indeed in the field many of them would have been passed for *urbanum*." (Marsden-Jones, 1930).

This suggests that the detection of Backcross Urbanum individuals using phenotypic traits may be very difficult. The studies which have investigated natural hybrid zones between *G. urbanum* and *G. rivale* so far have relied on morphology and reported the absence of Backcross Urbanum phenotypes (Ravanko, 1979; Waldren *et al.*, 1989). Results from Chapter 6.3.1 (p 128) using AFLP markers suggest that Backcross Urbanum individuals were not 'overlooked' in the hybrid zone due to the difficulty of identifying individuals of that class phenotypically but are either absent or very rare in the hybrid zone. Results from this chapter showed that Backcross Urbanums are actually produced indicating that their absence in the hybrid zone is not pollinator mediated. It is possible that Backcross Urbanums have lower fitness compared to the four classes observed in the hybrid zone. However, no significant difference between the Backcross Urbanum samples and offspring which clustered with the pure *G. urbanum* reference population in terms of number of seed heads produced or volume occupied could be detected in a common garden experiment (data not shown). Even if hybrids are completely fertile and viable in common garden experiments, it is possible that under natural conditions hybrids may have low fitness because they are poorly adapted to the physical habitat of the parents (Schluter, 1998). The absence of significant differences in

fitness between Backcross Urbanums and *G. urbanum* like offspring in a common garden experiment seems to suggest that endogenous selection is unlikely to play a major part.

Results from the offspring analysis revealed that at least three *G. urbanum* like parents (H11, H14 and H16) were heterozygous for a number of species-specific *G. urbanum* markers as some of their offspring possessed up to a third fewer species-specific *G. urbanum* markers. Although results from Chapter 6.3.1 (p 128) suggest that at least 13% of species-specific *G. urbanum* markers are heterozygous in pure *G. urbanum* reference populations, it is unlikely that up to 33% of the markers would be heterozygous in a pure *G. urbanum* individual. This is similar to the findings in the *G. rivale* like group discussed above and suggests that samples from the hybrid zone grouping genetically and morphologically with *G. urbanum* reference populations are not necessarily pure. The F1 genotype H15 gave rise to one offspring (H15/1) which grouped with the pure *G. urbanum* reference populations which could explain how seemingly genetically pure samples produce progeny with more than expected segregating species-specific markers. It is noteworthy, however, that H15/1 was phenotypically similar to an F1 and would not have been mistaken for a pure *G. urbanum* individual in the field although it could be possible that genetically and morphologically pure looking *G. urbanum* individuals are generated sometimes by F1s as it has been observed to occur for a pure *G. rivale* like morphology (see above). No progeny with F1 genotype and phenotype were observed suggesting that the generation of F1s by *G. urbanum* like parents is rare.

### *Offspring from F1 parents*

F1 like parents generated morphologically and genetically diverse offspring due to the large number of new different gene combinations possible between segregating *G. urbanum* and *G. rivale* genes. If the interpretation is correct that parental individuals from the hybrid zone which grouped with controlled F1 crosses in the PCO plot are truly F1s and that AFLP inheritance in backcross individuals is straightforward, then the expectation would be that F2s will spread considerably from the parental point in a PCO plot whereas backcross progeny will display greater intermediacy between parental and F1 clusters. Making these assumptions, an obvious explanation for the scattering of offspring around the F1 parental group is likely to be attributable to F2s.

So why are F2s missing in the adult generation? PCO and PCA plots revealed that some F1 offspring grouped morphologically and genetically with the controlled F1 group and would therefore be classified as F1s. However, it is clear that progeny from F1 parents can not be F1s. This suggests that samples which were classified genetically and morphologically as F1 individuals in the adult generation are not necessarily F1s and could be F2s or later

generation hybrids with the phenotype and genotype of F1s. The majority of F1 individuals in the adult generation did not possess 100% of species-specific *G. rivale* markers which could be due to the heterozygous nature of some of these markers. It is also possible that species-specific *G. rivale* markers are missing because some of the adult F1 parents are actually F2 or later generation hybrids. This is also supported by the presence of numerous individuals with F1 like morphology in the hybrid zone (Figure 6.10, p 140) suggesting that the production of F1s by *G. rivale* and/or *G. urbanum* like individuals occurs at an appreciable frequency. However, neither the *G. urbanum* nor the *G. rivale* parents generated offspring which could be identified as F1s morphologically, indicating that this is a relatively rare event. It is therefore likely that the absence of F2s in the adult generation is an artefact and that the adult group which was classified as F1s are actually a mixture of F1s, F2s and possibly later generation hybrids.

Progeny from F1 like parents did not scatter symmetrically around the parental position but showed a significant shift towards the *G. rivale* side along PCO1 with only a few samples towards *G. urbanum*. If the scattering was mainly attributable to F2s then we would expect to find more or less equal numbers of offspring towards the *G. urbanum* and *G. rivale* side of the PCO plot. The asymmetrical form of the scatter cloud therefore suggests that a considerable part of the scattering towards *G. rivale* might also be due to mating events with *G. rivale* like individuals. Waldren *et al.* (1989) reported that in an experimental garden, bees appeared to show flower colour specificity and visited predominantly purple campanulate flowers of *G. rivale* and related hybrids or the yellow *G. urbanum* flowers and flowers of similar looking hybrids. If the hybrids with yellow petals had a campanulate purple calyx, which largely obscured the yellow petals (like in F1s) then 'purple' bees also visited this flower. With their campanulate, drooping flowers and purple calyx F1s are morphologically more similar to *G. rivale* and it is therefore likely that F1s will receive a mixture of pollen from *G. rivale* like and F1 individuals but rarely from *G. urbanum* in outcrossing events. It is noteworthy that two F1 like parents gave rise to three offspring which appeared to be genetically pure *G. urbanum* (H15/1) and *G. rivale* (H17/3 and H17/4). This suggests that one generation of backcrossing (if one dismisses the possibility of the statistically unlikely event of extreme segregation which could produce pollen grains or ovules with an exclusively high *G. urbanum* or *G. rivale* gene content in the F1 parent) could generate progeny which appears to be genetically pure. However, in the case of H15/1 genetic data did not correspond with morphological data as phenotypically this individual resembled an F1 rather than a pure *G. urbanum*.

*Offspring from Backcrosses Rivale parents*

The key point which emerged from analyzing the Backcross Rivale group was the high genetic similarity of offspring and parents. There was no significant difference between the mean position and the variance of parents and offspring along PCO1 as observed in all other groups. If F1 parents can give rise to seemingly genetically pure offspring in one generation then it is improbable that 'pure' *G. rivale* individuals constitute a major source of pollen for Backcross Rivale parents as many more progeny, which are genetically indistinguishable from reference *G. rivale* populations would be expected. This seems to indicate that most Backcross Rivale individuals are likely to mate with other Backcross Rivale individuals.

Phenological observations in the hybrid zone (Chapter 6.3.2, p 139) revealed that Backcross Rivale samples flowered at a more or less constant rate (30 to 50 flowers) throughout most of the flowering season until the end of June and would theoretically have ample opportunity to receive pollen from any of the other groups in the hybrid zone (F1, *G. urbanum* and *G. rivale* like). The genetic similarity of Backcross Rivale parents and offspring therefore suggests that this group is either mostly selfing and/or mainly pollinated by an agent which is different to the one pollinating the other three groups. Data on auto-pollination indicated (Chapter 6.3.2, p 136) that only about 20% of flowers on 50% of Backcross Rivale individuals were able to produce seeds by spontaneous selfing, making it unlikely that a large part of the Backcross Rivale seed output is selfed. The speculation of a different pollinating agent is however supported by the analysis of nine morphological characters taken in a common environment on pure and hybrid individuals (Chapter 6.3.2, p 136). The analysis revealed that the Backcross Rivale group had significantly larger petals and broader sepals as well as (not significantly) longer sepals and claws than pure *G. rivale* reference individuals. Additionally, petal colour and the degree of flower openness were sometimes different between those two groups. Backcross Rivale individuals often displayed more open flowers with spreading sepals and salmon like to orange as well as creamlike to pinkish petals compared to the mostly pink campanulate *G. rivale* flowers (Figure 7.15). Backcross Rivale morphology is also considerably different from F1 morphology. It is possible that the differences in morphology between Backcross Rivale and individuals from other groups in the hybrid swarm attract different pollinators and therefore partly contribute to the observed lack of Backcross Rivale offspring which is genetically very similar to *G. rivale*.

**Figure 7.15.** Phenotypes of **(A)** Backcross Rivale and **(B)** *G. rivale*.

If the interpretation that Backcross Rivale individuals are mainly mating within their group is correct, then this could have important evolutionary implications as a reproductively isolated and morphologically distinct group of Backcross Rivale individuals could potentially develop into a new species over time. However, the data at hand only allow speculations regarding the evolutionary potential of the Backcross Rivale group. A detailed mating system analysis with larger sample sizes is needed to test whether this group is reproductively isolated from the other groups within the hybrid zone. Additionally, more Backcross Rivale genotypes from the studied and other hybrid zones need to be measured morphologically to be sure that the transgressive nature of above characters is not due to a sampling artefact.

## 7.4.3 Correlation between morphology and genetics

Another major finding was that the high correlation between morphological and genetic data in the adult population (see Chapter 6.3.1, p 128) was reduced to half of its original value in the offspring generation (c. 80% versus c. 40% of the variation explained, respectively). Although the correlation in the offspring generation was high enough to produce a good match between genotypic and morphological data for most progeny, for some samples contradicting results were obtained. Samples H15/1 and H9A/3 had the phenotype of an F1 whereas genetically they grouped with pure *G. urbanum* and Backcross Rivale samples, respectively. Similarly, some offspring from *G. rivale* like parents clustered phenotypically with *G. rivale* reference populations but where very close to the F1 group genetically. The partial breakdown of the correlation between adult and offspring generation was also demonstrated by attempts to use the program NEWHYBRIDS to classify the genotypic variation within the offspring generation (data not shown). Allocating offspring to six classes (pure *G. urbanum* and *G. rivale*, backcrosses to either parental species, F1 and F2) yielded sometimes impossible results (e. g. *G. urbanum* like parents produced offspring which were classified as 100% Backcross Rivale). Additionally, there were many individuals which had more or less similar probabilities of belonging to two or three classes. This is in stark contrast to the results obtained by NEWHYBRIDS analyzing 34 individuals from the adult hybrid zone (Chapter 6.3.1, p 128) where class allocation of samples was unambiguous and consistent

with morphology. Analyzing the same offspring data set with NEWHYBRIDS using up to 45 different possible hybrid classes (Milne & Abbott, 2008) did not alter the results significantly in terms of their interpretability, suggesting that the lack of correlation between phenotype and genotype is not due to the assignment of individuals to a limited number of genotypic classes.

Incongruence between genotype and phenotype has also been reported in a hybrid zone between the polyploid taxa *Cardamine pratensis* and *C. raphanifolia* (Lihova *et al.*, 2007). Superimposing morphology-based classification on AFLP and cpDNA data revealed a partial discordance between the data sets. Although morphology attributable to *C. raphanifolia* appeared to be generally a rather good indicator of AFLP patterns, there was no correlation between phenotype and genotype in other cases. Plants, which morphologically were identified as *C. pratensis*, were in fact hybrids or introgressed individuals genetically whereas some intermediate looking specimens corresponded genetically to *C. raphanifolia*. This suggests that parental phenotypes were more or less affected by hybridization and introgression, while intermediate genotypes could be genetically closer to one of the parents, indicating that morphology does not allow explanation or reliable prediction of AFLP patterns and cpDNA haplotypes, and vice versa (Lihova *et al.*, 2007). Considering the reticulate and polyploid history of the *C. raphanifolia* complex Lihova *et al.* (2007) suggested that the genome of *C. pratensis* might have been included in *C. raphanifolia*. This could be the reason why individuals which appear to be hybrids genetically might resemble *C. pratensis* morphologically rather than having intermediate morphology as the genome of *C. pratensis* prevailed in the genetic make up of the hybrids. An alternative explanation could be the overall pattern of introgression, where backcrossing may occur asymmetrically and hybrids hybridize more frequently with *C. pratensis* than with *C. raphanifolia* individuals (Lihova *et al.*, 2007).

In contrast to the *Cardamine* hybrid zone, the partial incongruence between genotype and phenotype in *Geum* is only observed in the offspring but not in the adult generation. Assuming that this is not due to a sampling artefact, this seems to indicate that although the mating patterns in this longstanding hybrid zone give rise to a multitude of different genotypes with variable phenotypic expression, only individuals with matching genotypes and phenotypes survive to adulthood. One possible explanation for this phenomenon is that the break down of a tight correlation between genotype and phenotype in some hybrids is linked to a break up of favourable gene combinations through recombination and/or to unfavourable parental gene interactions in the different genotype combinations of the hybrids resulting in low fitness and the inability to become established. More extensive sampling and transplant experiments in the hybrid zone are needed to establish whether the absence of mismatch individuals is a sampling artefact or is due to selection.

## 7.4.4 Fitness of offspring

Number of seeds produced and vegetative growth is often used to investigate fitness differences between families or populations within the same species. However, this is not possible if offspring from parents with fundamental differences in life history characters are involved. *G. urbanum* produces on average 20 times more seed heads than *G. rivale* but its rosette leaves only occupy about $1/13^{th}$ of the leaf volume occupied by *G. rivale*. This indicates that *G. urbanum* invests heavily in early reproduction with little investment in vegetative structure (r selected), while *G. rivale* invests heavily into vegetative structure and rather little into early seed production (K selected). In these circumstances, where groups of plants differ in their life history strategies, neither seed heads produced nor volume of vegetative structure can be used to compare fitness of *G. urbanum* and *G. rivale* and their hybrids with intermediate life history traits. However, the two parameters can still be used to assess variation in fitness of families within the same genotypic groups. When this was done the key finding was that significant differences were found among families for vegetative volume for all maternal parent groups (*G. urbanum* like, F1, Backcross Rivale and *G. rivale* like) and for seed head number for families from *G. urbanum* like and Backcross Rivale maternal parents. Correlation between seed head number and volume were also positive within groups. Intrafamilial variation was large in all four groups and many maternal parents produced offspring with a wide range of seed productivity. The largest range was found in offspring from F1 maternal parents where progeny from the same parents varied up to 50-fold in their seed output. In the Backcross Rivale group one maternal genotype was observed (H31) which produced progeny with consistently very low seed head numbers as only one out of five offspring set seed. These results imply that there is variation in fitness between and within offspring families and therefore the opportunity for selection of the fittest families or individuals within any one maternal genotype group.

Fitness variation among hybrids is a well documented phenomenon and depends on favourable or unfavourable parental gene interactions in the different genotype combinations of the hybrids as well as on environment x genotype interactions (Arnold, 1997; Fritz *et al.*, 2006; Tauleigne-Gomes & Lefebvre, 2008). Fritz *et al.* (2006) investigated the performance of the two willow species *Salix sericea* and *S. eriocephala* and their controlled F1, F2, and backcross hybrids using a common garden experiment with six replicated gardens that differed in soil moisture. Their findings suggested that fitness of hybrids was not uniformly low and that exogenous factors influenced plant performance and components of fitness but that endogenous factors were also relevant. Kimball *et al.* (2008) evaluated *Penstemon* hybrid performance in the greenhouse and in a reciprocal transplant experiment in the field measuring fruit set, seed set, seed weight, time required for fruits to mature and seedling growth for potted plants. To test for exogenous isolation, survival was also measured for pure

parents, reciprocal F1 hybrids and later generation hybrids at three different elevations. The results revealed that interspecific crosses produced as many seeds as conspecific crosses and that hybrid performance was equal to or greater than parental performance in all environments, including the greenhouse and all field gardens, indicating a lack of endogenous isolation (Kimball *et al.*, 2008). In each field garden, the hybrid with the native cytoplasm had a higher survival rate, suggesting local adaptation to different elevations. Kimball *et al.* (2008) concluded that exogenous factors which differ along elevational gradients can be more important than intrinsic genetic incompatibilities in determining the fitness of plant hybrids.

The genotypic constitution of the offspring generation in the *Geum* hybrid zone is far more complex than the one encountered in the adult generation which clearly shows that selection against certain hybrids must occur, although with the data at hand it is not possible to distinguish between exogenous and endogenous selection. However, as results from the common garden environment suggest, it is likely that environmental selection could be more important than endogenous selection. No significant differences could be detected in number of seed heads produced between Urbanum Backcross and *G. urbanum* genotypes. Similarly, offspring with significantly fewer parental bands (PCO2 top group) and offspring with more or less equal parental band numbers did not differ significantly in their seed output. This indicates that the difference between the genotypic structure of adult and offspring generation could be due to selection against certain genotypes and the interaction between genotype and environment.

## 7.4.5 Summary

Progeny collected from 30 maternal parents in the hybrid zone filled the gaps between the four genotypic classes (*G. urbanum* like, F1, Backcross Rivale and *G. rivale* like) observed in the adult generation. Four key points emerged from the analyses of open-pollinated seeds. Firstly, the high correlation between morphological and genetic data in the adult generation is partially corroded in the offspring generation. Secondly, more hybrid classes (such as Backcross Urbanums and F2s) are produced in the offspring generation than appear in the adult generation, suggesting that some progeny classes do not contribute to the adult plant stage. Thirdly, what appears to be a genetically pure parental sample is not necessarily pure and fourthly, significant genotypic differences exist within hybrid classes in terms of number of seed heads produced and volume occupied. The Backcross Rivale group was the only group out of the four genotypic classes in the hybrid zone where the average offspring position did not differ significantly from the mean parental position in a PCO plot, indicating the close genotypic relatedness of parents and progeny. In contrast, parents in all other classes gave rise to genetically more diverse offspring.

# 8. SYNTHESIS

This thesis aims to contribute to our knowledge about the possible evolutionary consequences of hybridization between inbreeding and outcrossing plants using *G. urbanum* and *G. rivale* as a study system. Morphological and genetic tools were applied to study the structure of pure parental populations and controlled F1 crosses. The results from these analyses provided the framework against which the morphological and genetic structure of a hybrid swarm in Scotland could be compared. This final chapter summarises the key findings of this thesis and explores the possible evolutionary consequences of hybridization between selfing *G. urbanum* and outcrossing *G. rivale*.

## 8.1 Main findings of this study

*G. urbanum* and *G. rivale* are two morphologically and genetically distinct taxa with different breeding systems, namely highly selfing and highly outcrossing, respectively. Controlled crosses (selfing and outcrossing) within two pure parental populations per species could not detect inbreeding depression in the two *G. urbanum* populations. Inbreeding depression was absent in one *G. rivale* population and present only in low levels in the other *G. rivale* population.

The analyses of a hybrid swarm between *G. urbanum* and *G. rivale* with morphological and genetic tools revealed that four distinct genetic groups can be distinguished, namely a (1) *G. urbanum* like, (2) *G. rivale* like, (3) F1 like and (4) Backcross Rivale group. The F1 like group was likely to comprise F1, F2 and maybe later generation hybrids possessing F1 morphology and an equal contribution of genomes from both parents.

Two transgressive characters were detected in the hybrid swarm. The character 'brown stamens with indumentum' occurred in all the four hybrid groups, although it was very rare in the *G. urbanum* like class. This suggests that pure looking individuals could actually be advanced backcrosses. The second transgressive character was observed for 'sepal breadth'. Individuals from the Backcross Rivale group possessed sepals which on average exceeded the parental mean.

Sampling genotypes between seed progeny and the plants from which they were collected, showed that some progeny classes did not contribute to the adult plant stage. In particular, Backcross Urbanum individuals were produced as seeds but were not observed in the adult generation.

## 8.2 What are the possible evolutionary consequences?

As outlined in the Introduction (Chapter 1, p 3) hybridization can lead to extinction, introgression, speciation and stalemate. Based on the findings of this thesis and previous studies I will discuss the possible evolutionary consequences of hybridization between *G. urbanum* and *G. rivale* in the light of the four major outcomes.

### 8.2.1 Extinction

The likelihood that the hybridization process between *G. urbanum* and *G. rivale* will lead to the extinction of one species by demographic swamping and/or genetic assimilation in the long run seems to be very small. The morphological and genetic analyses of pure *G. urbanum* and *G. rivale* populations revealed that both taxa are clearly separated by numerous morphological characters and species-specific AFLP markers. This also applied to the analysis within the long-lived hybrid swarm where both parental species are morphologically and genetically clearly distinguishable. The two forms persist even in the presence of gene flow. Results from earlier studies showed that the challenge lay in separating *G. rivale* like hybrid phenotypes from pure looking *G. rivale* individuals rather than identifying the two parental taxa (Marsden-Jones, 1930; Ravanko, 1979; Waldren *et al.*, 1989). This suggests that the long term evolutionary consequence of hybridization between *G. urbanum* and *G. rivale* is unlikely to lead to extinction.

### 8.2.2 Introgression

Introgression can have wide reaching evolutionary consequences if the introgressed loci are adaptive. For example, it has recently been shown by Kim *et al.* (2008) that hybridization between *Senecio vulgaris*, an allotetraploid non-radiate species in the family Asteraceae and *S. squalidus*, a diploid radiate species originating from Sicily, has led to the introgression of the radiate trait into some populations of *S. vulgaris* which promotes outcrossing (Figure 8.1). Self fertilization may impose long-term limitations responding to changing environmental conditions and Kim *et al.* (2008) argue that the reintroduction of genes that promote outcrossing could therefore allow *S. vulgaris* to prevent extinction in the long term. Another example of introgression of adaptive loci from one species into the another is the transfer of increased shade tolerance in *Helianthus* (Heiser, 1979). Heiser (1979) examined three hybrid populations between *Helianthus microcephalus* (a species of shaded areas) and *H. divaricatus* (a species of open habitats) over a period of 22 years. At the end of this period, although the habitat had become more closed at two sites, *H. divaricatus* was dominant at one site and the sole species at the other one. This suggests that introgression of loci which increase shade

tolerance were transferred from *H. microcephalus* into *H. divaricatus* (Heiser, 1979). Similarly, evidence from the Louisiana irises indicate that there is a significant association between shade tolerance in the hybrids and the number of markers from the shade loving parent found in that hybrid (Arnold, 2004).

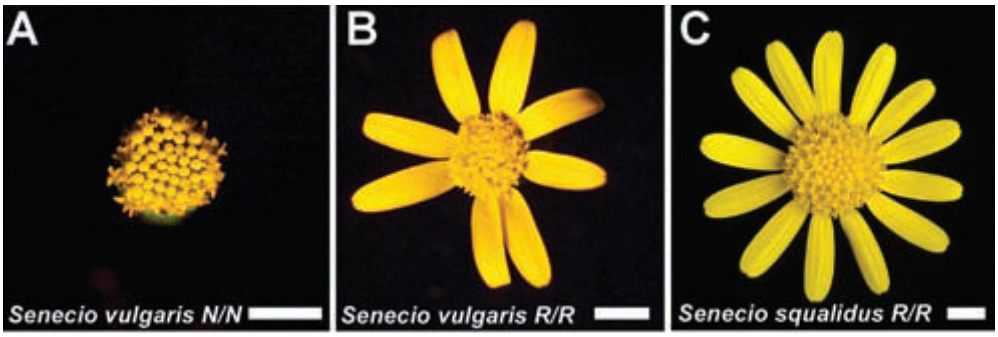



**Figure 8.1.** Flower head of nonradiate *S. vulgaris* (A), radiate *S. vulgaris* (B), and *S. squalidus* (C). Scale bars, 3 mm. From Kim *et al.* (2008).

## *Evidence for introgression from previous* Geum *studies*

*G. urbanum* and *G. rivale* differ in a number of physiological traits such as in their ability to grow in shade as well as in their tolerance to waterlogging and levels of manganese in the soil (Waldren *et al.*, 1987b; Waldren *et al.*, 1987a; Waldren *et al.*, 1987c). A study on the responses to flooding by *G. urbanum* and *G. rivale* found that although waterlogging restricted the growth of both species, the responses of the two species differed consistently (Waldren *et al.*, 1987a). The growth of *G. urbanum* was more severely reduced by flooding than that of *G. rivale*. Similarly, a study on the manganese tolerance in *G. urbanum* and *G. rivale* found that *G. rivale* was significantly more tolerant of higher manganese concentrations than *G. urbanum* (Waldren *et al.*, 1987b).

To establish the responses to waterlogging Waldren *et al.* (1988) and manganese tolerance (Waldren *et al.*, 1987b) in hybrid individuals between *G. urbanum* and *G. rivale*, the authors used cuttings from hybrids collected in a natural hybrid swarm. The cuttings were propagated under greenhouse conditions and allowed to flower freely. Fruits were then collected and seedling families raised. The results of these studies revealed that the responses of hybrids to waterlogging or manganese were either *G. urbanum* or *G. rivale* like (no intermediate tolerance levels for either trait were detected) and generally related to their predominant morphology (whether they looked more like *G. urbanum* or *G. rivale*). However, some hybrid individuals did not follow this pattern and showed levels of waterlogging and manganese tolerance which corresponded to one parental species despite having a morphology more akin to the other (Waldren *et al.*, 1987b; Waldren *et al.*, 1988). No simple relationship between

manganese and waterlogging tolerance could be found in hybrid families, indicating that the loci involved in the expression of either trait are not linked.

These findings suggest that adaptive traits can be transferred between *G. urbanum* and *G. rivale* and that hybrids are unlikely to have ecological preferences which are intermediate between the parents (Waldren *et al.*, 1988). However, this does not necessarily mean that hybrids are restricted to parental habitats only as combinations of different physiological traits inherited from either parent may allow hybrids to thrive in habitats other than parental ones.

### *Evidence for introgression from this study*

The molecular analysis of the hybrid zone indicates that there is limited opportunity for gene exchange and introgression from *G. rivale* into *G. urbanum*. Although Backcross Urbanum individuals are produced by *G. urbanum* like parents and to a lesser extent by F1 like plants, there is no evidence that these individuals survive to adulthood. However, observations in the hybrid swarm in relation to the transgressive trait 'brown stamens with indumentum' suggest that introgression of genes from *G. rivale* into *G. urbanum* is probably not completely prevented and may happen occasionally.

Gene exchange and introgression from *G. urbanum* into *G. rivale* is a much likelier event than vice versa. A Backcross Rivale group was detected in the hybrid swarm which could act as a stepping stone for repeated backcrossing to *G. rivale*, possibly resulting in the increase of genetic diversity in *G. rivale* (allelic polymorphism and/or allelic diversity). Phenological observations in the hybrid zone revealed that *G. rivale* phenotypes flower for about three weeks before the appearance of the first *G. urbanum* flowers, suggesting that offspring sired during this period are likely to be 'pure'. However, hybrids were also observed to flower at the same time as the first *G. rivale* individuals thus reducing the chances of 'pure' *G. rivale* offspring. Similarly, the high frequency (25%) of the putative transgressive character 'brown stamens with indumentum' in pure looking *G. rivale* individuals indicates that many of the parental phenotypes in the hybrid zone are possibly advanced backcrosses. It seems therefore likely that the concept of 'pure' *G. rivale* is somewhat tenuous in this hybrid zone and that numerous genetically and morphologically pure looking *G. rivale* individuals might contain some *G. urbanum* genes.

This could also apply to some British *G. rivale* populations as hybridization between *G. urbanum* and *G. rivale* is widespread in Great Britain. Briggs & Walters (1997) pointed out that the much higher frequency of hybrids in Great Britain compared to other European countries might be due to a complex history of land use. They argue that the landscape of

Great Britain has been modified extensively through human activities in a way which has led to the widespread fragmentation of the original vegetation and has created large areas of disturbed marginal habitats suitable for *G. urbanum* (Briggs & Walters, 1997). That in turn has resulted in the partial breakdown of the ecological isolation between *G. urbanum* and *G. rivale* as their habitats have been brought into close proximity in large parts of the country. This 'hybridization of the habitat' has provided the opportunity for the two species to hybridize more freely in Great Britain than for example in Poland, where large areas have been arguably less affected by human activity (Briggs & Walters, 1997). In Austria, where human land use might not have changed the original vegetation as extensively as in Great Britain due to topography, hybrids between *G. urbanum* and *G. rivale* are also rare (Fischer & Niklfeld, pers. comm., 2008). Apart from larger areas in northern Scotland and southern England hybrids between *G. urbanum* and *G. rivale* have been reported throughout Great Britain (Figure 1.10). It is therefore possible that many British *G. rivale* populations could have been exposed to hybridization at some point in their history and might have been introgressed by *G. urbanum* genes. Thus populations from areas of highest elevations in Scotland, England and Wales from which *G. urbanum* is absent, might have retained their 'purity'. This could mean that some British *G. rivale* populations which occur in areas of hybridization have a broader genetic base than *G. rivale* populations untouched by any form of hybridization. The acquisition of *G. urbanum* genes by more or less unidirectional introgression may have also allowed such populations to become better adapted to different environmental conditions like a shaded habitat for example.

Summarising, one of the consequences of hybridization between *G. urbanum* and *G. rivale* could be the introgression of neutral and/or adaptive loci from *G. urbanum* into *G. rivale* increasing the genetic diversity and possibly the colonising potential of *G. rivale* populations. Introgression from *G. rivale* into *G. urbanum* is not completely prevented but is less likely to happen.

## 8.2.3 Speciation

*G. urbanum* and *G. rivale* are hexaploid (2n=42) as are the hybrids between them. The two species differ by two to three chromosomal rearrangements (Gajewski, 1957) which according to the verbal model of homoploid speciation (Stebbins, 1957a; Grant, 1958) are necessary to establish a reproductive barrier between parental taxa and novel hybrid lineages. *G. urbanum* and *G. rivale* therefore have the genetic architecture which could lead to the generation of novel taxonomic diversity despite the same ploidy level of hybrids and parents. Another way to gain reproductive isolation from the parents is via the formation of an intermediate flower type through hybridization and subsequent pollination by flower-

constant pollinators (Grant, 1949, 1994; Rieseberg & Willis, 2007). This could lead to speciation even if the parents and hybrids are on the some ploidy level. One example where this mode of speciation has been postulated is *Penstemon clevelandii*, a homoploid hybrid species derived from *P. spectabilis* and *P. centranthifolius* (Straw, 1955; Wolfe *et al.*, 1998a, b). The hybrid species' flowers are of much the same shape as those of *P. spectabilis* but are significantly smaller, of a different colour and are mainly pollinated by solitary bees and sometimes hummingbirds. In contrast, the parental taxa are pollinated by wasps and hummingbirds, respectively (Straw, 1955).

Molecular analysis of seeds collected from hybrids in the hybrid swarm between *G. urbanum* and *G. rivale* revealed that the Backcross Rivale group only gave rise to offspring which were more or less of a Backcross Rivale genotype. This seems to indicate that most Backcross Rivale individuals are reproductively isolated from any other group in the hybrid zone by either mainly outcrossing with other Backcross Rivale individuals, or by selfing, or by a combination of the two mechanisms. Most Backcross Rivale individuals are distinguished from 'pure' *G. rivale* plants by a number of morphologically characters such as the way flowers are held, openness of petals and calyx and often petal colour (Figure 7.15, p 178). Additionally, two transgressive floral characters (sepal breadth and 'brown stamens with indumentum') were discovered in this group, although brown stamens were not exclusive to the Backcross Rivale group. As highlighted in Chapter 6.4.4 (p 142) the overall phenotype of many (but not all) Backcross Rivale individuals is noticeably different from pure looking *G. rivale* plants, despite the difficulty of distinguishing the two groups on the basis of a limited number (nine) of morphological characters.

It is possible that the morphological differences between *G. rivale* and the Backcross Rivale class are large enough to attract different pollinators most of the time, which could explain why Backcross Rivale offspring are genetically very similar to parental Backcross Rivale genotypes. The reproductive isolation between the two groups could therefore be pollinator mediated. This, however, conflicts with the hypothesis that Backcross Rivale individuals act as a bridge for introgression of *G. urbanum* genes into *G. rivale.* A possible explanation for this phenomenon is that the Backcross Rivale group is a heterogeneous assemblage of backcross individuals which behave in different ways. Certain combinations of *G. urbanum* and *G. rivale* genes might lead to the expression of a 'typical' Backcross Rivale phenotype which is easily recognizable by the characters described above. Other combinations might result in less obvious phenotypes which correspond much more to a 'pure' *G. rivale* morphology and go therefore unnoticed. Support for this speculation comes from the analyses of F1 offspring in the hybrid swarm. Some progeny from a putative F1 parent possessed a more or less *G. rivale* like phenotype and could not be distinguished genetically from pure *G. rivale* reference individuals (with the genetic tools used). In contrast, other F1 like parents produced offspring

which were morphologically and genetically close to a 'typical' Backcross Rivale phenotype. This seems to suggest that Backcross Rivale individuals can be roughly divided into two phenotypic subgroups. The 'typical' Backcross Rivale subgroup is morphologically distinct from other groups in the hybrid swarm and might attract different pollinators. The other subgroup which is more akin to plants with a pure *G. rivale* morphology is likely to facilitate gene exchange and introgression of *G. urbanum* genes into *G. rivale*.

### 8.2.4 Stalemate

As outlined above there is evidence for introgression and some support for the reproductive isolation of a morphologically more or less distinct hybrid group. This suggests that the likelihood of a stalemate situation where hybridization is just an ephemeral event, which gives rise to transient hybrids without any further evolutionary consequences in the long run, is small.

## 8.3 Further research

This thesis has combined molecular and morphological approaches to generate a baseline survey of the hybridization process between selfing *G. urbanum* and outcrossing *G. rivale*, and assessed the results in terms of its evolutionary potential. However, the findings presented here suggest that further research would deepen our knowledge of the evolutionary consequences of hybridization between different mating systems in general and the two *Geum* taxa in particular.

*Is asymmetric introgression a common pattern in hybridizing plant taxa with different mating systems?*

This and other studies (Ravanko, 1979; Waldren *et al.*, 1989; Sweigart & Willis, 2003) revealed that asymmetric introgression from the selfing into the outcrossing taxon is likely to occur when the two mating systems meet. To establish whether this is a general pattern more hybridizing species with different breeding systems need to be investigated.

*Are Backcross Urbanums generally missing in hybrid swarms?*

Only one hybrid swarm was analysed in this thesis. To be able to generalise the findings of this study, it is necessary to investigate the genetic and morphological patterns of hybridization between *G. urbanum* and *G. rivale* in more hybrid swarms. This is essential to

establish whether the same four genetic groups are generally present in other hybrid swarms or whether there is evidence for a different genotypic composition and/or additional groups such as Backcross Urbanums.

*Can a reference frame work of controlled crosses help us to interpret the patterns found in hybrid swarms?*

Controlled F1s were generated in the greenhouse to provide a reference point for the analyses of the hybrid swarm. Although this was a very useful first reference point, it became clear that controlled backcrosses in both directions (F1 x *G. urbanum* and F1 x *G. rivale*) and controlled F2s are needed to provide a more detailed framework for the genetic and morphological analyses. Of particular interest are the morphology and the genetic composition of these controlled crosses. To be able to identify Backcross Urbanum individuals in the field it is essential to ascertain whether they differ significantly in any measurable trait from a pure *G. urbanum* morphology. Controlled Backcross Rivale individuals would help to establish the morphological variation within this group and to test for the generation of two phenotypic subgroups, i. e. one similar to *G. rivale* and one which is more or less morphological differentiated from *G. rivale* displaying transgressive characters. Results from analyses of the hybrid swarm suggested that some putative F1 individuals might actually be F2s or later generation hybrids. Controlled crosses between individuals of known F1 origin are therefore necessary to establish the genetic and morphological composition of F2s.

*Can we detect hybrid lineages with predominantly selfing mating systems?*

It was not possible to establish the mating system of hybrids in the hybrid swarm as the only microsatellite markers available were developed for *G. urbanum* and did not give unambiguous results for hybrid individuals. A selfing breeding system could be of critical importance for the establishment of a novel hybrid lineage as homogenising gene flow is minimised. Knowing the mating system of hybrids belonging to various genotypic groups in a hybrid swarm is therefore essential to predict their evolutionary potential. A breeding system analysis of hybrid individuals in the hybrid swarm would provide information about which genotypic groups are mainly selfing, mainly outcrossing or have a mixed mating system. Establishing the allele frequencies of the pollen pool could also elucidate mating patterns between the different genotypic groups. In particular, investigating the mating system of Backcross Rivale individuals would help to clarify whether the overall genetic similarity between Backcross Rivale maternal parents and their offspring is mainly due to selfing or likely to be pollinator mediated.

*Is there evidence for pollinator constancy among genotypic groups?*

Pollinator behaviour plays a crucial part in facilitating hybridization, especially when the parental taxa differ significantly in their floral morphologies. Experimental blocks stocked with a varying composition of individuals from the four genotypic groups set up in the field or in a garden environment would provide data on the influence of different floral morphs on pollinator behaviour. Currently we do not know to what extent pollinators differentiate between F1, Backcross Rivale, *G. urbanum* and *G. rivale* morphs. Such an experiment would provide information on whether the Backcross Rivale group is likely to be pollinated by a different pollinator compared to the other genotypic groups. Additionally, one could test whether mating events between parental taxa and/or different hybrid groups are influenced by a higher density of different floral morphs. It is possible that hybridization and outcrossing rates are increased compared to a hybrid swarm with a limited range of morphologies as pollinators might find it more difficult to identify their preferred floral phenotype reliably.

*What is the frequency of novel combinations of adaptive traits in hybrids?*

*G. urbanum* and *G. rivale* differ in their shade tolerance and their ability to cope with adverse soil conditions such as waterlogging and high manganese concentrations, physiological characters which can be measured readily (Waldren *et al.*, 1987b; Waldren *et al.*, 1987a; Waldren *et al.*, 1987c; Waldren *et al.*, 1988). To be able to assess the transfer of these species-specific traits into a hybrid background would provide information about their potential for evolution, especially if they display a novel suite of adaptations. It is therefore particularly interesting to evaluate the physiological properties of individuals in the Backcross Rivale group as the results from this thesis indicate that they are likely to be reproductively isolated from other groups in the hybrid swarm. The ability of a lineage, which is partly morphologically differentiated from any other group, to colonise new habitats due to a novel set of physiological adaptations would greatly enhance their evolutionary potential.

*Can we detect fitness differences between different hybrid offspring classes in reciprocal transplant experiments?*

Comparing the genetic structure of adult plants and their progeny from the four genotypic groups in the hybrid swarm showed that a wider range of genotypic classes was produced in the offspring generation than could be observed in the adult generation. This suggests that some genotypic groups which are not present in the hybrid swarm although they are produced (e. g. Backcross Urbanum group) are at a fitness disadvantage and are selected against. However, a preliminary common garden experiment did not reveal any significant

fitness differences between these groups. It is therefore necessary to investigate the possible reason for the absence of certain genotypic groups by transplanting individuals from controlled crosses into their natural environment where they are exposed to the unmitigated effects of selection and competition.

*Is there evidence for introgression in 'pure' populations?*

The degree of introgression is likely to depend on the sympatry of *G. urbanum* and *G. rivale* populations and the frequency of hybridization events. A future research project could investigate whether *G. rivale* populations from areas that have been affected by hybridization are more introgressed than populations from areas, where hybridization between the two taxa is unlikely to have happened (such as upland areas in Northern Scotland). This can be done by comparing allele frequencies in populations from pure and mixed areas. It has also been speculated that hybridization between the two taxa is more frequent in Great Britain due to the higher fragmentation of the natural landscape by human activities compared to other countries in Europe (Briggs & Walters, 1997). This hypothesis could be tested in the same way as described above using populations from Great Britain and countries such as Poland and Austria where hybridization between *G. urbanum* and *G. rivale* is allegedly a rarer event.

## 8.4 Conclusions

This study has provided a valuable first insight into the genetic structure and the dynamics of a hybrid swarm between *G. urbanum* and *G. rivale*. The interpretation of the morphological and genetic data set suggests that the Backcross Rivale group is likely to be the most important group in terms of its evolutionary impact in the hybrid swarm. On the one hand introgression and the possible transfer of *G. urbanum* specific adaptations into a *G. rivale* background might be facilitated by a subgroup of Backcross Rivale individuals which are morphologically close to pure *G. rivale* phenotypes. This could potentially enable introgressed *G. rivale* individuals to colonise new habitats such as waterlogged and shaded places and/or widen the adaptive potential and genetic base of *G. rivale*. On the other hand individuals from a different Backcross Rivale subgroup with large enough morphological differentiation and some new phenotypic characters which reproductively isolate these individuals (despite their putatively predominantly outcrossing breeding system) from pure *G. rivale* and other groups in the hybrid zone via different pollinators, could potentially develop into a new species over time. This is a somewhat unexpected finding as theoretical predictions suggest (see Introduction, p 8) that the obvious route to the generation of taxonomic diversity is via novel hybrid lineages with a selfing mating system. However, a Backcross Urbanum group (i. e. the

group with the highest chances to produce such selfing novel hybrid lineages) was not detected in the hybrid swarm, although observations indicate that they must survive sometimes to reproductive age. If a novel selfing Backcross Urbanum individual should survive to adulthood then the evolution of an inbreeding line containing *G. rivale* genes is unlikely to be strongly hindered by the expression of deleterious recessive genes as data on the levels of inbreeding depression in *G. urbanum* and *G. rivale* indicate. Theoretically there are very few barriers for the development of inbreeding hybrid lineages but practically it seems to be an unlikely event as the group with the highest chances to produce such selfing novel hybrid lineages (e. g. Backcross Urbanums) is missing from the hybrid swarm. Given the rarity of Backcross Urbanum individuals it seems more probable that the potential evolutionary impact of these individuals is likely to lead to rare introgression of *G. rivale* genes into a *G. urbanum* background rather than to the generation of taxonomic diversity in the long run.

This thesis has shown that even after considerable ecological and morphological differentiation, genetic exchange between diverging lineages can occur and might potentially lead to the generation of taxonomic diversity. More and more studies find evidence for past hybridization events highlighting that evolution is in many cases a reticulate process and can not be depicted as a tree for some taxa (Cook *et al.*, 2008; Fortune *et al.*, 2008; Kim & Donoghue, 2008; Lohne *et al.*, 2008; Peng & Wang, 2008; Stefanovic & Costea, 2008; Yi *et al.*, 2008; Garcia *et al.*, 2009). Phylogenetic systematics generally presumes the dichotomous splitting of species from a putative ancestor. However, the assumption that species diverge and form independent evolutionary lineages which do not 'communicate' with each other anymore is too simplistic. This thesis contributes to the growing body of literature that hybridization is a creative force in plant evolution and highlights that the evolutionary role of selfing plant taxa, which have been hypothesised to be evolutionary dead ends (Stebbins, 1957b; Uyenoyama, 1986), might have been underestimated.

# REFERENCES

**Aarssen LW, Schamp BS, Pither J** (2006) Why are there so many small plants? Implications for species coexistence. *Journal of Ecology* **94**, 569-580.

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

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

**Abbott RJ, Ireland HE, Joseph L, Davies MS, Rogers HJ** (2005) Recent plant speciation in Britain and Ireland: Origins, establishment and evolution of four new hybrid species. *Biology and Environment* **105B**, 173-183.

**Abbott RJ, Lowe AJ** (2004) Origins, establishment and evolution of new polyploid species: *Senecio cambrensis* and *S. eboracensis* in the British Isles. *Biological Journal of the Linnean Society* **82**, 467-474.

**Abbott RJ, Ritchie MG, Hollingsworth PM** (2008) Introduction. Speciation in plants and animals: pattern and process. *Philosophical Transactions of the Royal Society London B: Biological Sciences* **363**, 2965-2969.

**Ægisdottir HH, Jespersen D, Kuss P, Stöcklin J** (2007) No inbreeding depression in an outcrossing alpine species: the breeding system of *Campanula thyrsoides*. *Flora* **202**, 218-225.

**Ågren J, Schemske DW** (1993) Outcrossing rate and inbreeding depression in two annual monoecious herbs, *Begonia hirsuta* and *B. semiovata*. *Evolution* **47**, 125-135.

**Ainouche ML, Baumel A, Salmon A, Yannic G** (2003) Hybridization, polyploidy and speciation in *Spartina* (Poaceae). *New Phytologist* **161**, 165-172.

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

**Anderson E** (1949) *Introgressive hybridization.* John Wiley & Sons, New York.

**Anderson E, Hubricht L** (1938) Hybridization in *Tradescantia*. III. The evidence for introgressive hybridization. *American Journal of Botany* **25**, 396-402.

**Anderson E, Stebbins GL** (1954) Hybridization as an evolutionary stimulus. *Evolution* **8**, 378-388.

**Anderson EC, Thompson EA** (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**, 1217-1229.

**Arens P, Durka W, Wernke-Lenting JH, Smulders MJM** (2004) Isolation and characterization of microsatellite loci in *Geum urbanum* (Rosaceae) and their transferability within the genus *Geum*. *Molecular Ecology Notes* **4**, 209-212.

**Armbruster P, Reed DH** (2005) Inbreeding depression in benign and stressful environments. *Heredity* **95**, 235-242.

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

**Arnold ML** (1993) *Iris nelsonii* (Iridaceae) - origin and genetic composition of a homoploid hybrid species. *American Journal of Botany* **80**, 577-583.

**Arnold ML** (1994) Natural hybridization and Louisiana irises. *Bioscience* **44**, 141-147.

**Arnold ML** (1997) *Natural Hybridization and Evolution.* Oxford University Press, New York, Oxford.

**Arnold ML** (2004) Transfer and origin of adaptations through natural hybridization: Were Anderson and Stebbins right? *Plant Cell* **16**, 562-570.

**Arnold ML** (2006) *Evolution through Genetic Exchange.* Oxford University Press, New York, Oxford.

**Arnold ML, Bulger MR, Burke JM, Hempel AL, Williams JH** (1999) Natural hybridization: how low can you go and still be important? *Ecology* **80**, 371-381.

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

**Arnold ML, Kentner EK, Johnston JA, Cornman S, Bouck AC** (2001) Natural hybridization and fitness. *Taxon* **50**, 93-104.

**Bacilieri R, Ducousso A, Petit RJ, Kremer A** (1996) Mating system and asymmetric hybridization in a mixed stand of European oaks. *Evolution* **50**, 900-908.

**Baker HG** (1955) Self compatibility and establishment after long distance dispersal. *Evolution* **9**, 347-349.

**Balfour DH** (1863) *Flora of Edinburgh - Being a List of Plants Found in the Vicinity of Edinburgh.* Adam & Charles Black, Edinburgh.

**Barclay IR** (1975) High-frequencies of haploid production in wheat (*Triticum aestivum*) by chromosome elimination. *Nature* **256**, 410-411.

**Barrett SCH, Charlesworth D** (1991) Effects of a change in the level of inbreeding on the genetic load. *Nature* **352**, 522-524.

**Barrett SCH, Harder LD** (1996) Ecology and evolution of plant mating. *Trends in Ecology & Evolution* **11**, A73-A79.

**Barrett SCH, Husband BC** (1990) Variation in outcrossing rates in *Eichhornia paniculata* the role of demographic and reproductive factors. *Plant Species Biology* **5**, 41-56.

**Barrett SCH, Morgan MT, Husband BC** (1989) The dissolution of a complex genetic polymorphism - the evolution of self-fertilization in tristylous *Eichhornia paniculata* (Pontederiaceae). *Evolution* **43**, 1398-1416.

**Barringer BC, Geber MA** (2008) Mating system and ploidy influence levels of inbreeding depression in *Clarkia* (Onagraceae). *Evolution* **62**, 1040-1051.

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

**Beismann H, Barker JHA, Karp A, Speck T** (1997) AFLP analysis sheds light on distribution of two *Salix* species and their hybrid along a natural gradient. *Molecular Ecology* **6**, 989-993.

**Bell MA, Travis MP** (2005) Hybridization, transgressive segregation, genetic covariation, and adaptive radiation. *Trends in Ecology & Evolution* **20**, 358-361.

**Bleeker W, Hurka H** (2001) Introgressive hybridization in *Rorippa* (Brassicaceae): gene flow and its consequences in natural and anthropogenic habitats. *Molecular Ecology* **10**, 2013-2022.

**Bodénès C, Joandet S, Laigret F, Kremer A** (1997) Detection of genomic regions differentiating two closely related oak species *Quercus petraea* (Matt) Liebl and *Quercus robur* L. *Heredity* **78**, 433-444.

**Boecklen WJ, Howard DJ** (1997) Genetic analysis of hybrid zones: numbers of markers and power of resolution. *Ecology* **78**, 2611-2616.

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

**Bonin A, Ehrich D, Manel S** (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology* **16**, 3737-3758.

**Bradshaw HD, Jr., Otto KG, Frewen BE, McKay JK, Schemske DW** (1998) Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (Mimulus). *Genetics* **149**, 367-382.

**Briggs D, Walters SM** (1997) *Plant Variation and Evolution.* 3rd edn. Cambridge University Press, Cambridge.

**Brown AHD, Allard RW** (1970) Estimation of mating system in open-pollinated maize populations using isozyme polymorphisms. *Genetics* **66**, 133-&.

**Broyles SB** (2002) Hybrid bridges to gene flow: a case study in milkweeds (*Asclepias*). *Evolution* **56**, 1943-1953.

**Brunet J, Sweet HR** (2006) Impact of insect pollinator group and floral display size on outcrossing rate. *Evolution* **60**, 234-246.

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

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

**Burger WC** (1981) Why are there so many kinds of flowering plants? *BioScience* **31**, 572-581.

**Burke JM, Carney SE, Arnold ML** (1998) Hybrid fitness in the Louisiana irises: analysis of parental and F1 performance. *Evolution* **52**, 37-43.

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

**Campbell D, Waser N** (2007) Evolutionary dynamics of an *Ipomopsis* hybrid zone: confronting models with lifetime fitness data. *American Naturalist* **169**, 298-310.

**Campbell DR** (2003) Natural selection in *Ipomopsis* hybrid zones: implications for ecological speciation. *New Phytologist* **161**, 83–90.

**Campbell DR, Galen C, Wu CA** (2005) Ecophysiology of first and second generation hybrids in a natural plant hybrid zone. *Oecologia* **144**, 214-225.

**Campbell DR, Waser NM** (2001) Genotype-by-environment interaction and the fitness of plant hybrids in the wild. *Evolution* **55**, 669-676.

**Caraway V, Carr GD, Morden CW** (2001) Assessment of hybridization and introgression in lava-colonizing Hawaiian *Dubautia* (Asteraceae: Madiinae) using RAPD markers. *American Journal of Botany* **88**, 1688-1694.

**Carney SE, Cruzan MB, Arnold ML** (1994) Reproductive interactions between hybridizing irises - analyses of pollen tube growth and fertilization success. *American Journal of Botany* **81**, 1169-1175.

**Carolan JC, Hook ILI, Walsh JJ, Hodkinson TR** (2002) Using AFLP markers for species differentiation and assessment of genetic variability of in vitro-cultured *Papaver bracteatum* (section Oxytona). *In Vitro Cellular & Developmental Biology-Plant* **38**, 300-307.

**Carr DE, Dudash MR** (2003) Recent approaches into the genetic basis of inbreeding depression in plants. *Philosophical Transactions from the Royal Society London B: Biological Sciences* **358**, 1071-1084.

**Cervera MT, Ruiz-Garcia L, Martinez-Zapater JM** (2003) Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Molecular Genetics and Genomics* **268**, 832-833.

**Charlesworth B** (1992) Evolutionary rates in partially self-fertilizing species. *The American Naturalist* **140**, 126-148.

**Charlesworth B, Charlesworth D** (1999) The genetic basis of inbreeding depression. *Genetical Research* **74**, 329-340.

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

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

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

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

**Charlesworth D, Lyons EE, Litchfield LB** (1994) Inbreeding depression in two highly inbreeding populations of *Leavenworthia*. *Proceedings of the Royal Society of London B: Biological Sciences* **258**, 209-214.

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

**Charlesworth D, Silvertown J** (2001) *Plant Population Biology.* 4th edn. Blackwell Science, Oxford.

**Charmet G, Balfourier F, Chatard V** (1996) Taxonomic relationships and interspecific hybridization in the genus *Lolium* (grasses). *Genetic Resources and Crop Evolution* **43**, 319-327.

**Chen LZ, Chen JF** (2008) Changes of cytosine methylation induced by wide hybridization and allopolyploidy in *Cucumis*. *Genome* **51**, 789-799.

**Cheptou PO, Lepart J, Escarre J** (2002) Mating system variation along a successional gradient in the allogamous and colonizing plant *Crepis sancta* (Asteraceae). *Journal of Evolutionary Biology* **15**, 753-762.

**Chiang TY, Schaal BA, Peng CI** (1998) Universal primers for amplification and sequencing a noncoding spacer between the atpB and rbcL genes of chloroplast DNA. *Botanical Bulletin of Academia Sinica* **39**, 245-250.

**Chung MY, Nason JD, Chung MG** (2005) Patterns of hybridization and population genetic structure in the terrestrial orchids *Liparis kumokiri* and *Liparis makinoana* (Orchidaceae) in sympatric populations. *Molecular Ecology* **14**, 4389-4402.

**Clutton-Brock TH, Coulson T, Milner JM** (2004) Red deer stocks in the Highlands of Scotland. *Nature* **429**, 261-262.

**Coe GE** (1953) Cytology of reproduction in *Cooperia pedunculata*. *American Journal of Botany* **40**, 335-343.

**Comai L, Tyagi AP, Winter K*, et al.*** (2000) Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* **12**, 1551-1567.

**Cook LG, Morris DC, Edwards RD, Crisp MD** (2008) Reticulate evolution in the natural range of the invasive wetland tree species *Melaleuca quinquenervia*. *Molecular Phylogenetics and Evolution* **47**, 506-522.

**Coyne JA, Orr AH** (2004) *Speciation.* Sinauer Associates Inc., Sunderland, MA.

**Crane PR, Herendeen P, Friis EM** (2004) Fossils and plant phylogeny. *American Journal of Botany* **91**, 1683-1699.

**Crepet WL, Niklas KJ** (2009) Darwin's second "abominable mystery": Why are there so many angiosperm species? *American Journal of Botany* **96**, 366-381.

**Crnokrak P, Barrett SCH** (2002) Perspective: Purging the genetic load: a review of the experimental evidence. *Evolution* **56**, 2347-2358.

**Cruden RW** (1977) Pollen-ovule ratios - conservative indicator of breeding systems in flowering plants. *Evolution* **31**, 32-46.

**Cruzan MB, Arnold ML** (1993) Ecological and genetic associations in an *Iris* hybrid zone. *Evolution* **47**, 1432-1445.

**Cruzan MB, Arnold ML** (1994) Assortative mating and natural selection in an *Iris* hybrid zone. *Evolution* **48**, 1946-1958.

**Dafni A, Kevan PG, Husband BC** (2005) *Practical Pollination Biology.* Enviroquest Ltd., Cambridge, Canada.

**Darwin C** (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.*, 1st ed. edn. John Murray, London.

**Darwin C** (1876) *The Effects of Cross and Self-fertilization in the Vegetable Kingdom.* Murray, London.

**Darwin F** (1903) *More letters of Charles Darwin, a Record of His Work in Hitherto Unpublished Letters.* John Murray, London.

**De Nettancourt D** (1997) Incompatibility in angiosperms. *Sexual Plant Reproduction* **10**, 185-199.

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

**Demuth JP, Wade MJ** (2005) On the theoretical and empirical framework for studying genetic interactions within and among species. *American Naturalist* **165**, 524-536.

**Diamond JM** (1985) How many unknown species are yet to be discovered? *Nature* **315**, 538-539.

**Dobzhansky T** (1935) A critique of the species concept in biology. *Philosophy of Science* **2**, 344-355.

**Dobzhansky T** (1936) Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* **21**, 113-135.

**Dole J, Ritland K** (1993) Inbreeding depression in two *Mimulus* taxa measured by multigenerational changes in the inbreeding coefficient. *Evolution* **47**, 361-373.

**Dole JA** (1992) Reproductive assurance mechanisms in three taxa of the *Mimulus guttatus* complex (Scrophulariaceae). *American Journal of Botany* **79**, 650-659.

**Dorken ME, Pannell JR** (2007) The maintenance of hybrid zones across a disturbance gradient. *Heredity* **99**, 89-101.

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

**Dudash MR** (1990) Relative fitness of selfed and outcrossed progeny in a self-compatible, protandrous species, *Sabatia angularis* L. (Gentianaceae): a comparison in three environments. *Evolution* **44**, 1129-1139.

**Elle E, Hare JD** (2002) Environmentally induced variation in floral traits affects the mating system in *Datura wrightii*. *Functional Ecology* **16**, 79-88.

**Ellis JR, Bentley KE, McCauley DE** (2008) Detection of rare paternal chloroplast inheritance in controlled crosses of the endangered sunflower *Helianthus verticillatus*. *Heredity* **100**, 574-580.

**Ellstrand NC, Prentice HC, Hancock JF** (1999) Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics* **30**, 539-563.

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

**Emms SK, Arnold ML** (1997) The effect of habitat on parental and hybrid fitness: Transplant experiments with Louisiana irises. *Evolution* **51**, 1112-1119.

**Estes JR, Brown LS** (1973) Entomophilous, intrafloral pollination in *Phyla incisa*. *American Journal of Botany* **60**, 228-230.

**Fenster CB, Galloway LF, Chao L** (1997) Epistasis and its consequences for the evolution of natural populations. *Trends in Ecology & Evolution* **12**, 282-286.

**Fisher RA** (1941) Average excess and average effect of a gene substitution. *Annals of Eugenics* **11**, 53-63.

**Fortune PM, Pourtau N, Viron N, Ainouche ML** (2008) Molecular phylogeny and reticulate origins of the polyploid *Bromus* species from section Genea (Poaceae). *American Journal of Botany* **95**, 454-464.

**Frankham R, Ballou JD, Briscoe DA** (2002) *Introduction to Conservation Genetics.* Cambridge University Press, Cambridge.

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

**Friis EM, Pedersen KR, Crane PR** (2006) Cretaceous angiosperm flowers: innovation and evolution in plant reproduction. *Palaeogeography Palaeoclimatology Palaeoecology* **232**, 251-293.

**Fritz RS, Hochwender CG, Albrectsen BR, Czesak ME** (2006) Fitness and genetic architecture of parent and hybrid willows in common gardens. *Evolution* **60**, 1215-1227.

**Futuyma DJ** (1989) Macroevolutionary consequences of speciation: inferences from phytophagous insects. In: *Speciation and its Consequences* (eds. Otte D, Endler JA), pp. 557-578. Sinauer Associates, Sunderland, USA.

**Gajewski W** (1957) A cytogenetic study on the genus *Geum* L. *Monographiae Botanicae* **4**, 1-416.

**Galis F, Metz JAJ** (1998) Why are there so many cichlid species? *Trends in Ecology & Evolution* **13**, 1-2.

**Garcia PE, Schönswetter P, Aguilar JF, Feliner GN, Schneeweiss GM** (2009) Five molecular markers reveal extensive morphological homoplasy and reticulate evolution in the *Malva alliance* (Malvaceae). *Molecular Phylogenetics and Evolution* **50**, 226-239.

**Gimenez-Benavides L, Escudero A, Iriondo JM** (2007) Reproductive limits of a late-flowering high-mountain Mediterranean plant along an elevational climate gradient. *New Phytologist* **173**, 367-382.

**Gobert V, Moja S, Colson M, Taberlet P** (2002) Hybridization in the section *Mentha* (Lamiaceae) inferred from AFLP markers. *Am. J. Bot.* **89**, 2017-2023.

**Good-Avila SV, Mena-Ali JI, Stephenson AG** (2008) Genetic and environmental causes and evolutionary consequences of variation in self-incompatible species. In: *Self-Incompatibility in Flowering Plants. Evolution, Diversity, Mechanisms.* (ed. Franklin-Tong VE), pp. 33-51. Springer, Berlin, Heidelberg.

**Goodwillie C** (1999) Multiple origins of self-compatibility in Linanthus section *Leptosiphon* (Polemoniaceae): phylogenetic evidence from internal-transcribed-spacer sequence data. *Evolution* **53**, 1387-1395.

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

**Goodwillie C, Knight MC** (2006) Inbreeding depression and mixed mating in *Leptosiphon jepsonii*: a comparison of three populations. *Annals of Botany* **98**, 351-360.

**Goodwillie C, Partis KL, West JW** (2004) Transient self-incompatibility confers delayed selfing in *Leptosiphon jepsonii* (Polemoniaceae). *International Journal of Plant Sciences* **165**, 387-394.

**Goudet J** (2002) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

**Govaerts R** (2001) How many species of seed plants are there? *Taxon* **50**, 1085-1090.

**Grant V** (1949) Pollination systems as isolating mechanisms in angiosperms. *Evolution* **3**, 82-97.

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

**Grant V** (1966) Selection for vigor and fertility in progeny of a highly sterile species hybrid in *Gilia*. *Genetics* **53**, 757-&.

**Grant V** (1971) *Plant Speciation.* Columbia University Press, New York and London.

**Grant V** (1994) Modes and origins of mechanical and ethological isolation in angiosperms. *Proceedings of the National Academy of Sciences USA* **91**, 3-10.

**Grivet D, Heinze B, Vendramin GG, Petit RJ** (2001) Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Molecular Ecology Notes* **1**, 345-349.

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

**Guo Y-P, Vogl C, Van Loo M, Ehrendorfer F** (2006) Hybrid origin and differentiation of two tetraploid *Achillea* species in East Asia: molecular, morphological and ecogeographical evidence. *Molecular Ecology* **15**, 133-144.

**Hamilton MB** (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* **8**, 521-523.

**Hammer Ø, Harper DAT, Ryan PD** (2006) PAST - Paleontological Statistics ver. 1.42. Published at: http://folk.uio.no/ohammer/past.

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

**Harding J, Mankinen CB, Elliott MH** (1974) Genetics of *Lupinus*. Part 7. Outcrossing auto fertility and variability in natural populations of the Nanus group. *Taxon* **23**, 729-738.

**Harper JL, McNaughton IH, Sagar GR, Clatworthy JN** (1961) Evolution and ecology of closely related species living in same area. *Evolution* **15**, 209-227.

**Hartl DL** (1987) *A Primer of Population Genetics.* 2nd edn. Sinauer Associates, Sunderland.

**Hedrick PW, Kalinowski ST** (2000) Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics* **31**, 139-162.

**Hegi G** (1975) *Illustrierte Flora von Mitteleuropa.* Paul Parey, Berlin, Hamburg.

**Heiser CB** (1979) Hybrid populations of *Helianthus divaricatus* and *Helianthus microcephalus* after 22 Years. *Taxon* **28**, 71-75.

**Heiser CB, Jr.** (1947) Hybridization between the sunflower species *Helianthus annuus* and *H. petiolaris*. *Evolution* **1**, 249-262.

**Herrera CM** (1991) Dissecting factors responsible for individual variation in plant fecundity. *Ecology* **72**, 1436-1448.

**Hewitt GM** (1988) Hybrid zones - natural laboratories for evolutionary studies. *Trends in Ecology & Evolution* **3**, 158-167.

**Hill JL, Hill RA** (2001) Why are tropical rain forests so species rich? Classifying, reviewing and evaluating theories. *Progress in Physical Geography* **25**, 326-354.

**Hinton WF** (1976) The evolution of insect-mediated self-pollination from an outcrossing system in *Calyptridium* (Portulacaceae). *American Journal of Botany* **63**, 979-986.

**Hirayama D, Nanami S, Itoh A, Yamakura T** (2008) Individual resource allocation to vegetative growth and reproduction in subgenus Cyclobalanopsis (*Quercus*, Fagaceae) trees. *Ecological Research* **23**, 451-458.

**Hiscock SJ, Kues U, Jeon KW** (2000) Cellular and molecular mechanisms of sexual incompatibility in plants and fungi. In: *International Review of Cytology*, pp. 165-295. Academic Press Limited; Academic Press Limited.

**Hiscock SJ, McInnis SM** (2003) The diversity of self-incompatibility systems in flowering plants. *Plant Biology* **5**, 23-32.

**Ho KM, Kasha KJ** (1975) Genetic control of chromosome elimination during haploid formation in barley. *Genetics* **81**, 263-275.

**Hodges SA, Burke JM, Arnold ML** (1996) Natural formation of *Iris* hybrids: experimental evidence on the establishment of hybrid zones. *Evolution* **50**, 2504-2509.

**Hollingsworth PM, Ennos RA** (2004) Neighbour joining trees, dominant markers and population genetic structure. *Heredity* **92**, 490-498.

**Hollingsworth PM, Squirrell J, Hollingsworth ML** (2007) Taxonomic complexity, conservation and recurrent origins of self-pollination in *Epipactis* (Orchidaceae). In: *Current taxonomic research on the British and European flora* (eds. Bailey J, Ellis RG), pp. 27-44. BSBI, London.

**Holsinger KE** (1986) Dispersal and plant mating systems: the evolution of self-fertilization in subdivided populations. *Evolution* **40**, 405-413.

**Holsinger KE** (1988) Inbreeding depression doesn't matter - the genetic basis of mating system evolution. *Evolution* **42**, 1235-1244.

**Holsinger KE** (1991) Mass-action models of plant mating systems - the evolutionary stability of mixed mating systems. *American Naturalist* **138**, 606-622.

**Holsinger KE** (1996) Pollination biology and the evolution of mating systems in flowering plants. In: *Evolutionary Biology* (eds. Hecht MK, Steere WC, Wallace B), pp. 107-149. Plenum Press, New York.

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

**Howard DJ, Preszler RW, Williams J, Fenchel S, Boecklen WJ** (1997) How discrete are oak species? Insights from a hybrid zone between *Quercus grisea* and *Quercus gambelii*. *Evolution* **51**, 747-755.

**Hulten E, Fries M** (1986) *Atlas of North European Vascular Plants North of the Tropic of Cancer.* Koeltz Scientific Books, Königstein.

**Husband B, Ozimec B, Martin S, Pollock L** (2008) Mating consequences of polyploid evolution in flowering plants: current trends and insights from synthetic polyploids. *International Journal of Plant Sciences* **169**, 195-206.

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

**Husband BC, Schemske DW** (1997) The effect of inbreeding in diploid and tetraploid populations of *Epilobium angustifolium* (Onagraceae): Implications for the genetic basis of inbreeding depression. *Evolution* **51**, 737-746.

**Igic B, Kohn JR** (2001) Evolutionary relationships among self-incompatibility RNases. *Proceedings of the National Academy of Sciences USA* **98**, 13167-13171.

**Ishida K** (2008) Effects of inbreeding on the magnitude of inbreeding depression in a highly self-fertilizing tree, *Magnolia obovata*. *Ecological Research* **23**, 995-1003.

**Ishida TA, Hattori K, Sato H, Kimura MT** (2003) Differentiation and hybridization between *Quercus crispula* and *Q. dentata* (Fagaceae): insights from morphological traits, amplified fragment length polymorphism markers, and leafminer composition. *American Journal of Botany* **90**, 769-776.

**Jain SK** (1976) Evolution of inbreeding in plants. *Annual Review of Ecology and Systematics* **7**, 469-495.

**Jain SK** (1978) Breeding system in *Limnanthes alba*: several alternative measures. *American Journal of Botany* **65**, 272-275.

**Jarne P, Charlesworth D** (1993) The evolution of the selfing rate in functionally hermaphrodite plants and animals. *Annual Review of Ecology and Systematics* **24**, 441-466.

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

**Johnston MO** (1992) Effects of cross and self-fertilization on progeny fitness in *Lobelia cardinalis* and *L. siphilitica*. *Evolution* **46**, 688-702.

**Johnston MO, Schoen DJ** (1995) Mutation rates and dominance levels of genes affecting total fitness in two angiosperm species. *Science* **267**, 226-229.

**Johnston MO, Schoen DJ** (1996) Correlated evolution of self-fertilization and inbreeding depression: an experimental study of nine populations of *Amsinckia* (Boraginaceae). *Evolution* **50**, 1478-1491.

**Jørgensen CA** (1928) The experimental formation of heteroploid plants in the genus *Solanum*. *Journal of Genetics* **19**, 133-211.

**Kale C, Abbott AG** (2008) *Principles and Practices of Plant Genomes. Volume 1: Gene Mapping.* Science Publishers, Enfield.

**Kameyama Y, Kasagi T, Kudo G** (2008) A hybrid zone dominated by fertile F1s of two alpine shrub species, *Phyllodoce caerulea* and *Phyllodoce aleutica*, along a snowmelt gradient. *Journal of Evolutionary Biology* **21**, 588-597.

**Kärkkäinen K, Kuittinen H, Treuren Rv*, et al.*** (1999) Genetic basis of inbreeding depression in *Arabis petraea*. *Evolution* **53**, 1354-1365.

**Karrenberg S, Favre A** (2008) Genetic and ecological differentiation in the hybridizing campions *Silene dioica* and *S. latifolia*. *Evolution* **62**, 763-773.

**Karrenberg S, Lexer C, Rieseberg LH** (2007) Reconstructing the history of selection during homoploid hybrid speciation. *American Naturalist* **169**, 725-737.

**Karron JD, Jackson RT, Thumser NN, Schlicht SL** (1997) Outcrossing rates of individual *Mimulus ringens* genets are correlated with anther-stigma separation. *Heredity* **79**, 365-370.

**Karron JD, Thumser NN, Tucker R, Hessenauer AJ** (1995) The influence of population density on outcrossing rates in *Mimulus ringens*. **75**, 175-180.

**Kasha KJ, Kao KN** (1970) High frequency haploid production in barley (*Hordeum vulgare* L.). *Nature* **225**, 874-876.

**Keim P, Paige KN, Whitham TG, Lark KG** (1989) Genetic analysis of an interspecific hybrid swarm of populus: occurrence of unidirectional introgression. *Genetics* **123**, 557-565.

**Kelly JK** (2005) Family level inbreeding depression and the evolution of plant mating systems. *New Phytologist* **165**, 55-62.

**Kim M, Cui ML, Cubas P*, et al.*** (2008) Regulatory genes control a key morphological and ecological trait transferred between species. *Science* **322**, 1116-1119.

**Kim ST, Donoghue MJ** (2008) Incongruence between cpDNA and nrITS trees indicates extensive hybridization within *Eupersicaria* (Polygonaceae). *American Journal of Botany* **95**, 1122-1135.

**Kimball S, Campbell DR, Lessin C** (2008) Differential performance of reciprocal hybrids in multiple environments. *Journal of Ecology* **96**, 1306-1318.

**Kirk H, Vrieling K, Klinkhamer PGL** (2005) Maternal effects and heterosis influence the fitness of plant hybrids. *New Phytologist* **166**, 685-694.

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

**Kiviniemi K** (1996) A study of adhesive seed dispersal of three species under natural conditions. *Acta Botanica Neerlandica* **45**, 73-83.

**Klier K, Leoschke MJ, Wendel JF** (1991) Hybridization and introgression in White and Yellow Ladyslipper Orchids (*Cypripedium candidum* and *C. pubescens*). *Journal of Heredity* **82**, 305-318.

**Knuth P** (1898) *Handbuch der Blütenbiologie.* Vol. 2, part 1. Verlag Wilhelm Engelmann, Leipzig.

**Kohn JR, Graham SW, Morton B, Doyle JJ, Barrett SCH** (1996) Reconstruction of the evolution of reproductive characters in *Pontederiaceae* using phylogenetic evidence from chloroplast DNA restriction-site variation. *Evolution* **50**, 1454-1469.

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

**Langlotz M** (2007) *Aspekte der Reproduktionsbiologie chorikarper Blütenpflanzen.* E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart.

**Latta R, Ritland K** (1994) The relationship between inbreeding depression and prior inbreeding among populations of four *Mimulus* taxa. *Evolution* **48**, 806-817.

**Lee TD, Bazzaz FA** (1982) Regulation of fruit and seed production in an annual legume, *Cassia fasciculata*. *Ecology* **63**, 1363-1373.

**Leigh EG, Davidar P, Dick CW*, et al.*** (2004) Why do some tropical forests have so many species of trees? *Biotropica* **36**, 447-473.

**Levin D** (2000) *The Origin, Expansion, and Demise of Plant Species.* Oxford University Press, New York.

**Levin DA** (1967) Hybridization between annual species of *Phlox* - population structure. *American Journal of Botany* **54**, 1122-&.

**Levin DA** (1972) Competition for pollinator service - stimulus for evolution of autogamy. *Evolution* **26**, 668-669.

**Levin DA** (1978) The origin of isolating mechanisms in flowering plants. *Evolutionary Biology* **11**, 185-317.

**Levin DA** (1996) The evolutionary significance of pseudo-self-fertility. *American Naturalist* **148**, 321-332.

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

**Lewis PO, Zaykin D** (1999) Genetic Data Analysis: Computer program for the analysis of allelic data, version 1.0. Free program distributed by the authors over the internet from the GDA home page at http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php.

**Lewontin RC, Birch LC** (1966) Hybridization as a source of variation for adaptation to new environments. *Evolution* **20**, 315-336.

**Lexer C, Fay MF, Joseph JA, Nica M-S, Heinze B** (2005) Barrier to gene flow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen): the role of ecology and life history in gene introgression. *Molecular Ecology* **14**, 1045–1057.

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

**Li ZK, Luo LJ, Mei HW*, et al.*** (2001) Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* **158**, 1737-1753.

**Li ZY, Ge XH** (2007) Unique chromosome behavior and genetic control in *Brassica* x *Orychophragmus* wide hybrids: a review. *Plant Cell Reports* **26**, 701-710.

**Lihova J, Kucera J, Perny M, Marhold K** (2007) Hybridization between two polyploid *Cardamine* (Brassicaceae) species in north-western Spain: discordance between morphological and genetic variation patterns. *Annals of Botany* **99**, 1083-1096.

**Linder HP** (2008) Plant species radiations: where, when, why? *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 3097-3105.

**Lipow SR, Broyles SB, Wyatt R** (1999) Population differences in self-fertility in the "self-incompatible" milkweed *Asclepias exaltata* (Asclepiadaceae). *American Journal of Botany* **86**, 1114-1120.

**Lloyd DG** (1965) Evolution of self compatability and racial differentiation in *Leavenworthia* (Cruciferae). *Contributions to the Gray Herbarium* **195**, 3-134.

**Lloyd DG, Schoen DJ** (1992) Self-fertilization and cross-fertilization in plants. 1. Functional dimensions. *International Journal of Plant Sciences* **153**, 358-369.

**Lohne C, Borsch T, Jacobs SWL, Hellquist CB, Wiersema JH** (2008) Nuclear and plastid DNA sequences reveal complex reticulate patterns in Australian water-lilies (*Nymphaea* subgenus Anecphya, Nymphaeaceae). *Australian Systematic Botany* **21**, 229-250.

**Lotsy JP** (1916) *Evolution by Means of Hybridization.* Nijhoff, The Hague, The Netherlands.

**Lowe A, Abbott RJ** (2003) A new British species, *Senecio eboracensis* (Asteraceae), another hybrid derivative of *S. vulgaris* L. and *S. squalidus* L. *Watsonia* **24**, 375-388.

**Lowe A, Harris SA, Ashton P** (2004) *Ecological Genetics: Design, Analysis and Application.* Blackwell Publishing, Oxford.

**Lowe AJ, Abbott RJ** (1996) Origins of the new allopolyploid species *Senecio cambrensis* (Asteraceae) and its relationship to the Canary Islands endemic *Senecio teneriffae*. *American Journal of Botany* **83**, 1365-1372.

**Lowe AJ, Abbott RJ** (2000) Routes of origin of two recently evolved hybrid taxa: *Senecio vulgaris* var. *hibernicus* and York radiate groundsel (Asteraceae). *American Journal of Botany* **87**, 1159-1167.

**Lowe AJ, Abbott RJ** (2004) Reproductive isolation of a new hybrid species, *Senecio eboracensis* Abbott & Lowe (Asteraceae). *Heredity* **92**, 386-395.

**Lu YQ** (2000) Effects of density on mixed mating systems and reproduction in natural populations of *Impatiens capensis*. *International Journal of Plant Sciences* **161**, 671-681.

**Mabberley DJ** (1997) *The Plant Book: A Portable Dictionary of the Vascular Plants.* 2nd ed. edn. Cambridge University Press, Cambridge.

**Mahelka V, Fehrer J, Krahulec F, Jarolimova V** (2007) Recent natural hybridization between two allopolyploid wheatgrasses (*Elytrigia*, Poaceae): ecological and evolutionary implications. *Annals of Botany* **100**, 249-260.

**Mangelsdorf AC** (1952) Gene interaction in heterosis. In: *Heterosis*, pp. 320-329. Iowa State College Press, Ames, Iowa.

**Marsden-Jones EM** (1930) The genetics of *Geum intermedium* Willd. haud Ehrh., and its backcrosses. *Journal of Genetics* **23**, 377-395.

**Martin NH, Willis JH** (2007) Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution* **61**, 68-82.

**Marutani M, Sheffer RD, Kamemoto H** (1993) Cytological analysis of *Anthurium andraeanum* (Araceae), its related taxa and their hybrids. *American Journal of Botany* **80**, 93-103.

**May RM** (1992) How many species inhabit the earth? *Scientific American* **267**, 42-48.

**Mayhew PJ** (2007) Why are there so many insect species? Perspectives from fossils and phylogenies. *Biological Reviews* **82**, 425-454.

**Maynard Smith J** (1966) Sympatric speciation. *American Naturalist* **100**, 637-&.

**Mayr E** (1942) *Systematics and Origin of Species.* Columbia University Press., New York.

**Mayr E** (1963) *Animal Species and Evolution.* Belknap Press, Cambridge, MA.

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

**McCauley DE, Bailey MF, Sherman NA, Darnell MZ** (2005) Evidence for paternal transmission and heteroplasmy in the mitochondrial genome of *Silene vulgaris*, a gynodioecious plant. *Heredity* **95**, 50-58.

**McKellar MA, Quesenberry KH** (1992) Chromosome pairing and pollen viability in *Desmodium ovalifolium* Wall X *Desmodium heterocarpon* (L) Dc hybrids. *Australian Journal of Botany* **40**, 243-247.

**Melser C, Klinkhamer PGL** (2001) Selective seed abortion increases offspring survival in *Cynoglossum officinale* (Boraginaceae). *American Journal of Botany* **88**, 1033-1040.

**Mena-Ali JI, Rocha OJ** (2005) Effect of ovule position within the pod on the probability of seed production in *Bauhinia ungulata* (Fabaceae). *Annals of Botany* **95**, 449-455.

**Mena-Ali JI, Stephenson AG** (2007) Segregation analyses of partial self-incompatibility in self and cross progeny of *Solanum carolinense* reveal a leaky S-allele. *Genetics* **177**, 501-510.

**Meyn O, Emboden WA** (1987) Parameters and consequences of introgression in *Salvia apiana* X *S. mellifera* (Lamiaceae). *Systematic Botany* **12**, 390-399.

**Michaelis P** (1954) Cytoplasmic inheritance in Epilobium and its theoretical significance. *Advances in Genetics Incorporating Molecular Genetic Medicine* **6**, 287-401.

**Michalski SG, Durka W** (2007) High selfing and high inbreeding depression in peripheral populations of *Juncus atratus*. *Molecular Ecology* **16**, 4715-4727.

**Miglia KJ, McArthur ED, Redman RS*, et al.*** (2007) Genotype, soil type, and locale effects on reciprocal transplant vigor, endophyte growth, and microbial functional diversity of a narrow sagebrush hybrid zone in Salt Creek, Canyon, Utah. *American Journal of Botany* **94**, 425-436.

**Milligan BG** (1992) Is organelle DNA strictly maternally inherited - power analysis of a binomial-distribution. *American Journal of Botany* **79**, 1325-1328.

**Milne RI, Abbott RJ** (2000) Origin and evolution of invasive naturalized material of *Rhododendron ponticum* L. in the British Isles. *Molecular Ecology* **9**, 541-556.

**Milne RI, Abbott RJ** (2008) Reproductive isolation among two interfertile *Rhododendron* species: low frequency of post-F1 hybrid genotypes in alpine hybrid zones. *Molecular Ecology* **17**, 1108-1121.

**Milne RI, Terzioglu S, Abbott RJ** (2003) A hybrid zone dominated by fertile F1s: maintenance of species barriers in *Rhododendron*. *Molecular Ecology* **12**, 2719-2729.

**Minder AM, Rothenbuehler C, Widmer A** (2007) Genetic structure of hybrid zones between *Silene latifolia* and *Silene dioica* (Caryophyllaceae): evidence for introgressive hybridization. *Molecular Ecology* **16**, 2504-2516.

**Mishler BD** (1999) Getting rid of species. In: *Species: New Interdisciplinary Essays.* (ed. Wilson RA), pp. 307-316. Harvard University Press, Cambridge, USA.

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

**Mitchell-Olds T, Rutledge JJ** (1986) Quantitative genetics in natural plant-populations - a review of the theory. *American Naturalist* **127**, 379-402.

**Morrell JJ** (2005) Molds and stain fungi: so many species, but what do they all mean? *Abstracts of Papers American Chemical Society* **229**, U298.

**Müntzing A** (1936) The evolutionary significance of autopolyploidy. *Hereditas* **21**, 263-378.

**Nason JD, Ellstrand NC, Arnold ML** (1992) Patterns of hybridization and introgression in populations of oaks, manzanitas, and irises. *American Journal of Botany* **79**, 101-111.

**Nasrallah ME, Nasrallah JB** (1986) Molecular biology of self-incompatibility in plants. *Trends in Genetics* **2**, 239-244.

**Novotny V, Drozd P, Miller SE*, et al.*** (2006) Why are there so many species of herbivorous insects in tropical rainforests? *Science* **313**, 1115-1118.

**Oakley CG, Winn AA** (2008) Population-level and family-level inbreeding depression in a cleistogamous perennial. *International Journal of Plant Sciences* **169**, 523-530.

**Orr AH** (1995) The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* **139**, 1805-1813.

**Pannell JR, Barrett SCH** (1998) Baker's law revisited: reproductive assurance in a metapopulation. *Evolution* **52**, 657-668.

**Peng D, Wang XQ** (2008) Reticulate evolution in Thuja inferred from multiple gene sequences: implications for the study of biogeographical disjunction between eastern Asia and North America. *Molecular Phylogenetics and Evolution* **47**, 1190-1202.

**Pennings P, van Opijnen T** (2002) Why are there so many species of arthropods? *Entomologische Berichten (Amsterdam)* **62**, 20-23.

**Percy RG, Wendel JF** (1990) Allozyme evidence for the origin and diversification of *Gossypium barbadense* L. *Theoretical and Applied Genetics* **79**, 529-542.

**Petit RJ, Bodénès C, Ducousso A, Roussel G, Kremer A** (2003) Hybridization as a mechanism of invasion in oaks. *New Phytologist* **161**, 151-164.

**Pico FX, Ouborg NJ, Van Groenendael JM** (2004) Evaluation of the extent of among-family variation in inbreeding depression in the perennial herb *Scabiosa columbaria* (Dipsacaceae). *American Journal of Botany* **91**, 1183-1189.

**Prance GT, Beentje H, Dransfield J, Johns R** (2000) The tropical flora remains undercollected. *Annals of the Missouri Botanical Garden* **87**, 67-71.

**Preston CD, Pearman DA, Dines TD** (2002) *New Atlas of the British and Irish Flora: An Atlas of the Vascular Plants of Britain, Ireland, The Isle of Man and the Channel Islands.* Oxford University Press, London, Oxford.

**Prywer C** (1932) Genetische Studien über die Bastarde zwischen *Geum urbanum* L. und *Geum rivale* L. *Acta Societatis Botanicorum Poloniae* **9**, 87-114.

**Rafalski JA, Vogel JM, Morgante M*, et al.*** (1996) *Generating and Using DNA Markers in Plants.* Academic Press, London.

**Ramsey J, Bradshaw HD, Schemske DW** (2003) Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution* **57**, 1520-1534.

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

**Rathcke B, Real L** (1993) Autogamy and inbreeding depression in mountain laurel, *Kalmia latifolia* (Ericaceae). *American Journal of Botany* **80**, 143-146.

**Ravanko O** (1979) Observations from uniform experimental plot on introgression between *Geum urbanum* and *Geum rivale* in SW Finland. *Annales Botanica Fennici* **16**, 43–49.

**Reboud X, Zeyl C** (1994) Organelle inheritance in plants. *Heredity* **72**, 132-140.

**Reed AJ, Singletary GW** (1989) Roles of carbohydrate supply and phytohormones in maize kernel abortion. *Plant Physiology* **91**, 986-992.

**Remington DL, O'Malley DM** (2000a) Evaluation of major genetic loci contributing to inbreeding depression for survival and early growth in a selfed family of *Pinus taeda*. *Evolution* **54**, 1580-1589.

**Remington DL, O'Malley DM** (2000b) Whole-genome characterization of embryonic stage inbreeding depression in a selfed loblolly pine family. *Genetics* **155**, 337-348.

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

**Richards AJ** (1997) *Plant Breeding Systems.* 2. ed. edn. Chapman & Hall, London, New York, Tokyo.

**Rick CM, Fobes JF, Tanksley SD** (1979) Evolution of mating systems in *Lycopersicon hirsutum* as deduced from genetic variation in electrophoretic and morphological characters. *Plant Systematics and Evolution* **132**, 279-298.

**Riera-Lizarazu O, Rines HW, Phillips RL** (1996) Cytological and molecular characterization of oat x maize partial hybrids. *Theoretical and Applied Genetics* **93**, 123-135.

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

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

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

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

**Rieseberg LH, Carney SE** (1998) Plant hybridization. *New Phytologist* **114**, 599-624.

**Rieseberg LH, Choi HC, Ham D** (1991) Differential cytoplasmic versus nuclear introgression in *Helianthus*. *Journal of Heredity* **82**, 489-493.

**Rieseberg LH, Church SA, Morjan CL** (2003a) Integration of populations and differentiation of species. *New Phytologist* **161**, 59–69.

**Rieseberg LH, Desrochers AM, Youn SJ** (1995) Interspecific pollen competition as a reproductive barrier between sympatric species of *Helianthus* (Asteraceae). *American Journal of Botany* **82**, 515-519.

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

**Rieseberg LH, Gerber D** (1995) Hybridization in the Catalina island Mountain Mahogany (*Cercocarpus traskiae*) - RAPD evidence. *Conservation Biology* **9**, 199-203.

**Rieseberg LH, Raymond O, Rosenthal DM*, et al.*** (2003b) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**, 1211–1216.

**Rieseberg LH, Sinervo B, Linder CR, Ungerer MC, Arias DM** (1996) Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. *Science* **272**, 741-745.

**Rieseberg LH, Wendel JF** (1993) Introgression and its consequences in plants. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp. pp. 70-109. Oxford University Press, Oxford.

**Rieseberg LH, Widmer A, Arntz MA, Burke JM** (2003c) The genetic architecture necessary for transgressive segregation is common in both natural and domesticated

populations. *Philosophical Transactions of the Royal Society London B: Biological Sciences* **358**, 1141-1147.

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

**Rieseberg LH, Zona S, Aberbom L, Martin TD** (1989) Hybridization in the island endemic, *Catalina mahogany*. *Conservation Biology* **3**, 52-58.

**Ritland K** (1989) Correlated matings in the partial selfer *Mimulus guttatus*. *Evolution* **43**, 848-859.

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

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

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

**Robinson JP, Harris SA** (1999) *Amplified Fragment Length Polymorphisms and Microsatellites: A phylogenetic perspective.* Final Compendium of the Research Project Development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity., http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm.

**Ronfort J** (1999) The mutation load under tetrasomic inheritance and its consequences for the evolution of the selfing rate in autotetraploid species. *Genetical Research* **74**, 31-42.

**Rosquist G** (2001) Reproductive biology in diploid *Anthericum ramosum* and tetraploid *A. liliago* (Anthericaceae). *Oikos* **92**, 143-152.

**Ross-Craig S** (1955) *Drawings of British Plants.* G. Bell & Sons, London.

**Ruane LG, Donohue K** (2007) Environmental effects on pollen-pistil compatibility between *Phlox cuspidata* and *P. drummondii* (Polemoniaceae): implications for hybridization dynamics. *American Journal of Botany* **94**, 219-227.

**Ruggiero A, Hawkins BA** (2008) Why do mountains support so many species of birds? *Ecography* **31**, 306-315.

**Salmon A, Ainouche ML, Wendel JF** (2005) Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). *Molecular Ecology* **14**, 1163-1175.

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

**Schluter D** (1998) Ecological causes of speciation. In: *Endless Forms: Species and Speciation.* (eds. Howard DJ, Berlocher SH), pp. 114-129. Oxford University Press, Oxford.

**Schoen DJ** (1982) The breeding system of *Gilia achilleifolia* - variation in floral characteristics and outcrossing rate. *Evolution* **36**, 352-360.

**Schoen DJ, Lloyd DG** (1984) The selection of cleistogamy and heteromorphic diaspores. *Biological Journal of the Linnean Society* **23**, 303-322.

**Schoen DJ, Morgan MT, Bataillon T** (1996) How does self-pollination evolve? Inferences from floral ecology and molecular genetic variation. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **351**, 1281-1290.

**Schönswetter P, Tribsch A, Niklfeld H** (2003) Phylogeography of the high alpine cushion plant *Androsace alpina* (Primulaceae) in the European Alps. *Plant Biology* **5**, 623-630.

**Scotland RW, Wortley AH** (2003) How many species of seed plants are there? *Taxon* **52**, 101-104.

**Scotti-Saintagne C, Mariette S, Porth I*, et al.*** (2004) Genome scanning for interspecific differentiation between two closely related oak species (*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.). *Genetics* **168**, 1615-1626.

**Shaked H, Kashkush K, Ozkan H, Feldman M, Levy AA** (2001) Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* **13**, 1749-1759.

**Simmons MT, Cowling RM** (1996) Why is the Cape Peninsula so rich in plant species? An analysis of the independent diversity components. *Biodiversity and Conservation* **5**, 551-573.

**Slatkin M** (1987) Gene flow and the geographic structure of natural populations. *Science* **236**, 787-792.

**Smedmark JEE, Eriksson T, Evans RC, Campbell CS** (2003) Ancient allopolyploid speciation in Geinae (*Rosaceae*): evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. *Systematic Biology* **52**, 374-385.

**Smith PM, Dixon ROD, Cochrane MP** (2002) *Plant Life of Edinburgh and the Lothians.* Edinburgh University Press, Edinburgh.

**Soltis DE, Soltis PS, Tate JA** (2003) Advances in the study of polyploidy since Plant Speciation. *New Phytologist* **161**, 173-191.

**Soltis PS** (2005) Ancient and recent polyploidy in angiosperms. *New Phytologist* **166**, 5-8.

**Soltis PS, Soltis DE** (2004) The origin and diversification of angiosperms. *American Journal of Botany* **91**, 1614-1626.

**Spiess E** (1977) *Genes in Populations.* John Wiley & Sons, New York.

**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.* Academic Press, London, New York, San Francisco.

**Stebbins GL** (1957a) The hybrid origin of microspecies in the *Elymus glaucus* complex. *Cytologia (Supplement)* **36**, 336-340.

**Stebbins GL** (1957b) Self fertilization and population variability in the higher plants. *American Naturalist* **91**, 337-354.

**Stebbins GL** (1957c) *Variation and Evolution in Plants.* Columbia University Press, New York.

**Stebbins GL** (1974) *Flowering Plants: Evolution Above the Species Level.* Belknap Press, Cambridge, Mass.

**Stebbins GL, Daly K** (1961) Changes in variation pattern of a hybrid population of *Helianthus* over an 8-year period. *Evolution* **15**, 60-71.

**Stefanovic S, Costea M** (2008) Reticulate evolution in the parasitic genus *Cuscuta* (Convolvulaceae): over and over again. *Botany-Botanique* **86**, 791-808.

**Stephenson AG** (1981) Flower and fruit abortion: proximate causes and ultimate functions. *Annual Review of Ecology and Systematics* **12**, 253-279.

**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.

**Stone JL, Sasuclark MA, Blomberg CP** (2006) Variation in the self-incompatibility response within and among populations of the tropical shrub *Witheringia solanacea* (Solanaceae). *American Journal of Botany* **93**, 592-598.

**Stone SL, Goring DR** (2001) The molecular biology of self-incompatibility systems in flowering plants. *Plant Cell Tissue and Organ Culture* **67**, 93-114.

**Straw RM** (1955) Hybridization, homogamy, and sympatric speciation. *Evolution* **9**, 441-444.

**Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES** (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* **132**, 823-839.

**Stuessy TF** (2004) A transitional-combinational theory for the origin of angiosperms. *Taxon* **53**, 3-16.

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

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

**Takebayashi N, Morrell PL** (2001) Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and macroevolutionary approach. *American Journal of Botany* **88**, 1143–1150.

**Tanksley SD** (1993) Mapping Polygenes. *Annual Review of Genetics* **27**, 205-233.

**Tauleigne-Gomes C, Lefebvre C** (2005) Natural hybridisation between two coastal endemic species of *Armeria* (Plumbaginaceae) from Portugal. 1. Populational *in situ* investigations. *Plant Systematics and Evolution* **250**, 215-230.

**Tauleigne-Gomes C, Lefebvre C** (2008) Natural hybridisation between two coastal endemic species of *Armeria* (Plumbaginaceae) from Portugal. 2. Ecological investigations on a hybrid zone. *Plant Systematics and Evolution* **273**, 225-236.

**Taylor K** (1997a) Biological Flora of the British Isles: *Geum rivale*. *Journal of Ecology* **85**, 721–731.

**Taylor K** (1997b) Biological Flora of the British Isles: *Geum urbanum*. *Journal of Ecology* **85**, 705–720.

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

**Tiffin P, Olson MS, Moyle LC** (2001) Asymmetrical crossing barriers in angiosperms. *Proceedings of the Royal Society of London Series B: Biological Sciences* **268**, 861-867.

**Travers SE, Mena-Ali J, Stephenson AG** (2004) Plasticity in the self-incompatibility system of *Solanum carolinense*. *Plant Species Biology* **19**, 127-135.

**Tribsch A, Schönswetter P, Stuessy TF** (2002) *Saponaria pumila* (Caryophyllaceae) and the Ice Age in the European Alps. *American Journal of Botany* **89**, 2024-2033.

**Tsukamoto T, Ando T, Kokubun H*, et al.*** (1999) Breakdown of self incompatibility in a natural population of *Petunia axillaris* (Solanaceae) in Uruguay containing both self-incompatible and self compatible plants. *Sexual Plant Reproduction* **12**, 6-13.

**Turcotte EL, Feaster CV** (1967) Semigamy in Pima cotton. *Journal of Heredity* **58**, 55-57.

**Turelli M, Moyle LC** (2007) Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics* **176**, 1059-1088.

**Turesson G** (1925) The plant species in relation to habitat and climate. Contributions to the knowledge of genecological units. *Hereditas* **6**, 147-236.

**Turesson G** (1930) The selective effect of climate upon the plant species. *Hereditas* **14**, 99-152.

**Ungerer MC, Baird SJ, Pan J, Rieseberg LH** (1998) Rapid hybrid speciation in wild sunflowers. *Proceedings of the National Academy of Science USA* **95**, 11757-11762.

**Ungricht S** (2004) How many plant species are there? And how many are threatened with extinction? Endemic species in global biodiversity and conservation assessments. *Taxon* **53**, 481-484.

**Uyenoyama MK** (1986) Inbreeding and the cost of meiosis: the evolution of selfing in populations practicing biparental inbreeding. *Evolution* **40**, 388-404.

**Van Der Velde M, Bijlsma R** (2004) Hybridization and asymmetric reproductive isolation between the closely related bryophyte taxa *Polytrichum commune* and *P. uliginosum*. *Molecular Ecology* **13**, 1447-1454.

**Van Droogenbroeck B, Kyndt T, Romeijn-Peeters E*, et al.*** (2006) Evidence of natural hybridization and introgression between *Vasconcellea* species (Caricaceae) from southern Ecuador revealed by chloroplast, mitochondrial and nuclear DNA markers. *Annals of Botany* **97**, 793-805.

**Vandepitte K, Jacquemyn H, Roland-Ruiz I, Honnay O** (2007) Landscape genetics of the self-compatible forest herb *Geum urbanum*: effects of habitat age, fragmentation and local environment. *Molecular Ecology* **16**, 4171-4179.

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

**Vickery RK** (1956) Data on interracial and interspecific hybridizations in the sections Simiolus of the genus *Mimulus* (Scrophulariaceae). *Utah Academy of Sciences, Arts, and Letters* **33**, 45-64.

**Vogler DW, Das C, Stephenson AG** (1998) Phenotypic plasticity in the expression of self-incompatibility in *Campanula rapunculoides*. *Heredity* **81**, 546-555.

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

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

**Waldren S, Davies MS, Etherington JR** (1987a) Comparative studies of plant growth and distribution in relation to waterlogging .11. Growth of *Geum rivale* L and *Geum urbanum* L and soil chemical condition in experimentally flooded soil. *New Phytologist* **105**, 551–562.

**Waldren S, Davies MS, Etherington JR** (1987b) The effect of manganese on root extension of *Geum rivale* L., *G. urbanum* L. and their hybrids. *New Phytologist* **106**, 679-688.

**Waldren S, Etherington JR, Davies MS** (1987c) Comparative studies of plant growth and distribution in relation to waterlogging .14. Iron, Manganese, Calcium and Phosphorus concentrations in leaves and roots of *Geum rivale* L and *Geum urbanum* L grown in waterlogged soil. *New Phytologist* **106**, 689–696.

**Waldren S, Etherington JR, Davies MS** (1988) Comparative studies of plant growth and distribution in relation to waterlogging. XV. The effect of waterlogging on growth of various populations of and hybrids between *Geum rivale* L. and *Geum urbanum* L. *New Phytologist* **109**, 97–106.

**Waldren S, Etherington JR, Davies MS** (1989) Hybridization and ecology of *Geum rivale* L and *Geum urbanum* L (Rosaceae) at Malham Tarn, Yorkshire. *Botanical Journal of the Linnean Society* **101**, 199-211.

**Wang H, McArthur ED, Sanderson SC, Graham JH, Freeman DC** (1997) Narrow hybrid zone between two subspecies of Big Sagebrush (*Artemisia tridentata*: Asteraceae). IV. Reciprocal transplant experiments. *Evolution* **51**, 95-102.

**Warwick SI, Bain JF, Wheatcroft R, Thompson BK** (1989) Hybridization and introgression in *Carduus nutans* and *Carduus acanthoides* re-examined. *Systematic Botany* **14**, 476-494.

**Watano Y, Kanai I, Tani N** (2004) Genetic structure of hybrid zones between *Pinus pumila* and *P. parviflora* var. *pentaphylla* (Pinaceae) revealed by molecular hybrid index analysis. *American Journal of Botany* **91**, 65-72.

**Waterway MJ** (1990) Genetic differentiation and hybridization between *Carex gynodynama* and *Carex mendocinensis* (Cyperaceae) in California. *American Journal of Botany* **77**, 826-838.

**Whitlock MC, McCauley DE** (1999) Indirect measures of gene flow and migration: $F_{st}$ not equal 1/(4Nm+1). *Heredity* **82**, 117-125.

**Willis JH** (1993) Partial self-fertilization and inbreeding depression in two populations of *Mimulus guttatus*. *Heredity* **71**, 145-154.

**Willis JH** (1999) The role of genes of large effect on inbreeding depression in *Mimulus guttatus*. *Evolution* **53**, 1678-1691.

**Wolf DE, Takebayashi N, Rieseberg LH** (2001) Predicting the risk of extinction through hybridization. *Conservation Biology* **15**, 1039-1053.

**Wolfe AD, Xiang QY, Kephart SR** (1998a) Assessing hybridization in natural populations of *Penstemon* (Scrophulariaceae) using hypervariable intersimple sequence repeat (ISSR) bands. *Molecular Ecology* **7**, 1107-1125.

**Wolfe AD, Xiang QY, Kephart SR** (1998b) Diploid hybrid speciation in *Penstemon* (Scrophulariaceae). *Proceedings of the National Academy of Sciences USA* **95**, 5112-5115.

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

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

**Wu CI, Ting CT** (2004) Genes and speciation. *Nature Reviews Genetics* **5**, 114-122.

**Xu ML, Li XQ, Korban SS** (2000) AFLP-based detection of DNA methylation. *Plant Molecular Biology Reporter* **18**, 361-368.

**Yang TW, Yang YA, Xiong ZG** (2000) Paternal inheritance of chloroplast DNA in interspecific hybrids in the genus *Larrea* (Zygophyllaceae). *American Journal of Botany* **87**, 1452-1458.

**Yeo PF** (1954) The cytology of British species of *Euphrasia*. *Watsonia* **3**, 101-108.

**Yeo PF** (1968) Evolutionary significance of speciation of *Euphrasia* in Europe. *Evolution* **22**, 736-747.

**Yi T, Wen J, Golan-Goldhirsh A, Parfitt DE** (2008) Phylogenetics and reticulate evolution in *Pistacia* (Anacardiaceae). *American Journal of Botany* **95**, 241-251.

**Yu SB, Li JX, Tan YF*, et al.*** (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences USA* **94**, 9226-9231.

**Zapata TR, Arroyo MTK** (1978) Plant reproductive ecology of a secondary deciduous tropical forest in Venezuela. *Biotropica* **10**, 221-230.

**Zeisler M** (1938) Über die Abgrenzung der eigentlichen Narbenfläche mit Hilfe von Reaktionen. *Beihefte zum Botanischen Zentralblatt* **A58**, 308-318.