ORIGINAL ARTICLE

Early evolution in a hybrid swarm between outcrossing and selfing lineages in *Geum*M Ruhsam^{1,2}, PM Hollingsworth² and RA Ennos¹¹Ashworth Laboratories, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK and ²Department of Genetics and Conservation, Royal Botanic Garden Edinburgh, Edinburgh, UK

Although often considered as evolutionary dead ends, selfing taxa may make an important contribution to plant evolution through hybridization with related outcrossing lineages. However, there is a shortage of studies examining the evolutionary dynamics of hybridization between outcrossing and selfing taxa. On the basis of differential pollinator attractiveness, production and competitive ability of pollen, as well as levels of inbreeding depression, we predict that the early products of hybridization between outcrossing and selfing lineages will be F1s and first-generation backcrosses sired mainly by the outcrossing lineage, together with selfed F2s containing a limited genetic contribution from the outcrosser. These predictions were tested using amplified fragment length polymorphism and chloroplast markers to analyze the composition of a recent hybrid swarm between predominantly outcrossing *Geum rivale* and predominantly selfing *Geum*

urbanum. In line with predictions, the hybrid swarm comprised both parental species together with F1s and first-generation backcrosses to *G. rivale* alone. Chloroplast data suggested that *G. rivale* was the pollen parent for both observed hybrid classes. However, there was no evidence for F2 individuals, despite the fact that the F1 was fully self-compatible and able to auto-pollinate. The pollen fertility of F1s was only 30% lower than that of the parental taxa, and was fully restored in backcross hybrids. Predicting future evolution in the hybrid swarm will require an understanding of the mating patterns within and among the mix of parental, F1 and backcross genotypes that are currently present. However, these results support the hypothesis that introgression is likely to be asymmetrical from selfing to outcrossing lineages.

Heredity (2011) **107**, 246–255; doi:10.1038/hdy.2011.9; published online 30 March 2011

Keywords: hybridization; breeding system; hybrid zone; AFLP

Introduction

The most common mating system transition that occurs in flowering plants is that from outcrossing to selfing (Goldberg *et al.*, 2010). This switch leads to the coexistence of closely related lineages, which evolve independently under contrasting breeding systems. Selfing lineages arising from mating system transitions are often considered evolutionary dead ends, making no further contribution to taxonomic diversification and condemned to early extinction as a consequence of their low effective population sizes and restricted potential for genetic exchange and recombination (Glemin *et al.*, 2006; Schoen and Busch, 2008). However, there are situations in which self-fertilizing lineages, after a period of independent evolution, hybridize with their outcrossing relatives. Such hybridization has the potential to generate novel biodiversity based on recombination of genetic information derived from the contrasting selfing and outcrossing taxa (Ennos *et al.*, 2005; Hollingsworth *et al.*, 2007; French *et al.*, 2008; Chapman and Abbott, 2010).

To predict the immediate evolutionary consequences of hybridization between outcrossing and selfing plant species, it is essential to understand the ways in which selfing and outcrossing lineages differ genetically from one another. In relation to their outcrossing relatives, selfing plants generally show a reduced investment in pollinator attraction, lower pollen-to-ovule ratio, a reduction in competitive fitness of pollen and a greater ability to auto-pollinate (Lloyd, 1987; Charlesworth and Morgan, 1991; Diaz and Macnair, 1999; Fishman *et al.*, 2008). In terms of their genomic attributes, selfing taxa are expected to possess a lower genetic load than their outcrossing relatives (Schemske and Lande, 1985; Charlesworth and Charlesworth, 1987; Husband and Schemske, 1996; Igic *et al.*, 2008; Wright *et al.*, 2008).

With knowledge of these systematic differences between selfing and outcrossing lineages, it is possible to make some predictions about the early course of hybridization between them. As a result of the greater pollinator attractiveness and pollen production of outcrossing lineages, we expect pollinators to carry more pollen originating from outcrossing species than from selfing species when these species co-occur. Moreover, pollen from the outcross lineage is likely to be more competitive than that from the selfing lineage. Other factors being equal, it follows that F1 embryos are more likely to be the product of fertilization by pollen from the outcrossing lineage than from the selfing lineage.

Making *a priori* predictions about the identity of second-generation hybrids is more problematic. F1

Correspondence: Dr M Ruhsam, Department of Genetics and Conservation, Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK.

E-mail: m.ruhsam@rbge.ac.uk

Received 14 September 2010; revised 21 January 2011; accepted 1 February 2011; published online 30 March 2011

individuals could potentially reproduce by selfing. Formation of selfed progeny would require that the F1 be self-compatible, can auto-pollinate and be fertile. However, selfed genotypes homozygous for large portions of the outcrosser's genome could be of low fitness if the genetic load in the outcross genome was substantial. Without knowledge of the reproductive attributes of the F1, and the genetic load present in the outcrossing lineage, it is difficult to predict whether selfing lines segregating from F1s will constitute a significant component of second-generation hybrids.

F1s could alternatively reproduce by outcrossing, either as pollen or seed parents. When acting as a pollen parent, the F1 will be in direct competition with the often numerically superior outcrossing parent. In addition, assuming intermediate characteristics of the F1 between the outcrosser and the selfer, the hybrid is likely to be less attractive to pollinators and to have lower pollen production, pollen fitness and fertility than the outcrossing parent. Therefore, success of the F1 as a pollen parent is likely to be low. Thus, the F1 is most likely to contribute offspring to the hybrid swarm as an outcrossing maternal parent. The superior male fitness of the outcrossing lineage relative to the F1 and the inbreeding lineage is likely to ensure that most of the seed progeny of the F1 will be sired by the outcrossing parent.

In summary, the composition of a recent hybrid swarm involving an outcrossing and an inbreeding plant lineage is predicted to be a mixture of the two parental genotypes together with F1s and backcrosses sired by the outcrossing species. In addition, there is the possibility that inbreeding F2s may be present, possessing genetic contributions from both parents. Their presence is contingent on the ability of F1s to set seed successfully by auto-pollination, and their frequency will depend on the extent of inbreeding depression expressed in the F2. The potential genetic consequences of continued evolution in the hybrid swarm are widespread introgression of both maternally inherited and nuclear genomes from the inbreeding lineage into the outcrossing lineage, and more limited formation of novel inbreeding lineages, which contain nuclear genes from the outcrosser.

To date, there have been very few empirical studies that have documented the process of hybridization between outcrossing and inbreeding plant taxa, and none have been explicitly designed to test the predictions outlined above. Two studies do provide some support for these predictions. Evidence from hybridizing taxa of *Cyclamen* suggested asymmetric introgression of nuclear markers from an inbreeding into an outcrossing taxon, in line with expectations (Thompson *et al.*, 2010). Similarly, molecular analysis of species-specific markers has produced evidence of effective asymmetric transfer of nuclear genes from self-fertilizing *Mimulus nasutus* to outcrossing *Mimulus guttatus* well outside their current areas of sympatry (Sweigart and Willis, 2003). Unfortunately, the rarity of F1 hybrid formation in the wild and low F1 fertility precluded a detailed study of the early generations that followed hybridization in the *Mimulus* system (Martin and Willis, 2007). Contrary evidence to the above predictions comes from a study of hybridization between selfing and outcrossing *Phlox* species, which indicated that F1 hybrids were produced by crosses in both directions. However, no analysis of later-generation hybrids was conducted (Ferguson *et al.*, 1999).

Similarly, in *Senecio*, adaptively important genes coding for ray florets have been introgressed from outcrossing *Senecio squalidus* to inbreeding *Senecio vulgaris*, rather than in the direction predicted above (Chapman and Abbott, 2010), but this system has the additional complicating factor of ploidy level difference between the two species. Given the general shortage of studies, and their conflicting findings, further studies are required to understand the evolutionary consequences of hybridization between outcrossing and selfing species.

The objective of the present research is to explore in detail the early genetic consequences of natural hybridization between predominantly outcrossing *Geum rivale* and predominantly inbreeding *Geum urbanum*, to test and refine the predictions made above (Waldren *et al.*, 1989; Taylor, 1997a,b; Ruhsam *et al.*, 2010). We use a combination of maternal (cpDNA) and nuclear (AFLP) markers to determine the genetic composition of a naturally occurring hybrid swarm. Controlled crosses between the parental taxa are used to verify the molecular identification of hybrid classes and to measure the viability of F1 seeds from reciprocal pollinations. Having identified the parental and hybrid classes in the swarm, we measure their pollen fertility, self-compatibility and their ability to auto-pollinate. These data are important for assessing the likelihood that novel selfing lines will evolve in the hybrid swarm. Finally, we characterize the genotypic classes in the hybrid swarm in terms of their floral and vegetative morphology. We use this information to identify genotypic classes in the field and to score their flowering phenology, so as to understand possible ecological constraints on the genetic exchange between them, and the likely path of future evolution in the hybrid swarm.

Materials and methods

Study species

G. urbanum L., wood avens and *G. rivale* L., water avens are closely related, perennial, herbaceous, insect-pollinated plant species widely distributed in the British Isles (Taylor, 1997a,b). The species are ancient allohexaploids ($2n=42$) and show disomic inheritance (Smedmark *et al.*, 2003; Vandepitte *et al.*, 2007; Ruhsam *et al.*, 2010). *G. urbanum* possesses a weakly protogynous inflorescence (Knuth, 1898), which is held upright and has yellow petals that are attractive to flies (Supplementary Figure 1a). *G. urbanum* is completely self-compatible and its outcrossing rate is low, that is, between 5 and 20% (Ruhsam *et al.*, 2010). In contrast, the inflorescence of *G. rivale* is strongly protogynous, pendulous, with light purple/cream petals and is visited predominantly by bumble bees (Supplementary Figure 1b). *G. rivale* possesses a 'leaky' self-incompatibility system and shows a high outcrossing rate (70–90%) (Ruhsam *et al.*, 2010).

G. urbanum occurs naturally in shaded, well-drained and often wooded habitats, whereas *G. rivale* is found in more open and poorly drained sites along stream sides. Where these two habitats occur in close proximity, morphological hybrids, named *Geum × intermedium* Ehrh. (Supplementary Figure 1c), are commonly found (Taylor, 1997a). Hybrid swarms are frequent and have been recorded throughout the United Kingdom since the nineteenth century (Balfour, 1863; Marsden-Jones, 1930; Waldren *et al.*, 1989).

Plant material

Pure populations: In all, 10 putatively pure populations of *G. urbanum* and of *G. rivale* were identified from sites located throughout Great Britain (Table 1). Populations were considered pure if the other parental species had not been recorded in the surrounding $2 \times 2 \text{ km}^2$ area (Edinburgh populations, Smith *et al.* (2002)) or $10 \times 10 \text{ km}^2$ (all other populations, Preston *et al.* (2002)). A leaf sample was taken from at least one individual in each population and dried in silica gel for later DNA extraction. In spring 2007, 1–18 plants were dug up from 3 pure *G. urbanum* and 5 pure *G. rivale* populations, potted in peat-free compost and grown in a randomized design in an unheated glasshouse. These plants were used for controlled crosses and studies of flower morphology, and pollen fertility (Table 1).

Controlled F1 crosses: Controlled interspecific crosses were made between individuals previously collected from pure populations, using procedures described in the study by Ruhsam *et al.* (2010), to establish the morphology and genetic marker composition of F1 genotypes (Table 1). Nine crosses had *G. urbanum* and five crosses had *G. rivale* as the maternal parent. Seeds were sown by family in pots with peat-free compost, overwintered outside and then moved to an unheated greenhouse where the percentage germination of each family was scored. A total of 35 offspring from 6 crosses with *G. urbanum* as maternal parent, and 24 offspring from 4 crosses with *G. rivale* as maternal parent, were grown to maturity in the greenhouse. DNA extractions were obtained from silica-dried leaves. When referring to these controlled crosses in the paper, we use the standard notation of listing the female parent first.

Hybrid swarm: The *Geum* hybrid swarm studied is located 11 km southwest of Edinburgh at an elevation of 150 m (Table 1). It extends for 400 m along the southern bank of the Water of Leith occupying semi-natural woodland vegetation beside a railway track that was abandoned in the 1960s. The population numbers >1000 individuals and comprises a roughly equal proportion of both parents together with a range of intermediate phenotypes. A total of 34 individuals, representing the total morphological spectrum present in the hybrid population, were dug up in July 2006, transferred to individual pots containing peat-free compost and grown in a randomized array in an unheated glasshouse.

Molecular markers

DNA extraction: DNA was extracted from silica-dried or fresh leaf material using the protocol of Doyle and Doyle (1990), with the inclusion of 0.1% 2-mercaptoethanol and 0.1% insoluble polyvinylpyrrolidone (PVPP) to the 2% cetyl trimethylammonium bromide (CTAB) buffer.

AFLP analysis: A total of 30 individuals of each taxon collected from 10 pure populations, F1 individuals from controlled crosses ($n = 15$) and the 34 individuals sampled from the hybrid swarm were scored for AFLP marker variation (Table 1). The AFLP protocol was based on the work by Vos *et al.* (1995) and is described in detail in the study by Ruhsam (2009). Four primer combinations (namely 6FAM-*Eco*R1 AAC/*Mse*1 CTA, PET-*Eco*R1

AAT/*Mse*1 CAGA, VIC-*Eco*R1 AGC/*Mse*1 CTT and NED-*Eco*R1 ATC/*Mse*1 CGA) yielded 202 markers. After the initial results were obtained, an additional 20 *G. urbanum*-like samples from the hybrid zone were assayed using the same methods to test for the presence of backcrosses to *G. urbanum*. AFLP runs were repeated 3 times for 24 randomly selected samples from the hybrid swarm to estimate error rates. Two repeat runs used the same pre-amplification product for selective PCR, whereas the third involved a completely new restriction-ligation process to generate the PCR template. Following the study by Bonin *et al.* (2004), error rates were estimated as proportion of phenotypic comparisons within individuals that were different.

Classification of genotypic classes in the hybrid swarm: A matrix of genetic distances among individuals based on Jaccard's similarity was calculated from AFLP genetic marker data and subjected to principal coordinates analysis (PCO) with the program PAST (Hammer *et al.*, 2006) to visualize the genetic groupings present in the hybrid swarm. Data obtained from the 34 hybrid swarm individuals plus a subset of reference individuals were included in the PCO analysis (12 pure *G. urbanum*, 12 pure *G. rivale* and 15 F1 individuals from controlled crosses). In addition, the Bayesian assignment program NEWHYBRIDS version 1.1 beta (Anderson and Thompson, 2002) was used to assign individuals within the hybrid swarm to most likely genotypic classes using default parameters (such as pure parental, F1, F2 and backcross categories) and Jeffreys priors. The hybrid swarm sample together with the same set of reference individuals (12 pure *G. urbanum*, 12 pure *G. rivale* and 15 F1 individuals from controlled crosses) were included in this analysis. No individual or allele frequency prior information was specified for individuals in the hybrid swarm or controlled cross-F1s, but reference individuals from pure populations were marked as pure according to the requirements of the program. A burn-in phase of 100 000 steps and 100 000 sweeps was used. To verify that the NEWHYBRIDS program can accurately classify hybrid classes when one of the parental species involved in hybridization is inbreeding, genotypic data for F1, F2, *G. rivale* backcross and *G. urbanum* backcross classes were simulated using the same number of markers and degree of interspecific differentiation, as found in the *Geum* system and assuming that one parent is 100% outcrossing and the other 95% selfing. The NEWHYBRIDS program classified simulated genotypes into their correct hybrid classes with high probability under a wide range of assumptions about hybrid swarm composition (Donnelly, 2010).

cpDNA variation: cpDNA variation in *Geum* was scored using a restriction fragment length polymorphism (PCR-RFLP) assay involving *Eco*R1 digestion of the amplified trnL-F region (Ruhsam, 2009). Individuals possessing a restriction site gave two fragments of 390 and 630 bp (haplotype A), whereas individuals lacking a restriction site gave a single 1020-bp fragment (haplotype B). To test the predominant mode of inheritance, parents and progeny from three interspecific crosses were scored. In each of the three crosses, the *G. urbanum* parent was of haplotype A and the *G. rivale* parent was of haplotype B. In one of the crosses, *G. urbanum* was the female parent,

Table 1 Details of sites from which pure populations of *Geum urbanum* and *Geum rivale*, together with a hybrid swarm, were sampled

Species	Location	Altitude (m)	Coordinates	Number of individuals scored for AFLP	Number of individuals scored for cpDNA	Number of individuals scored for morphology	Parents for controlled F1 crosses	Pollen fertility
Pure <i>Geum urbanum</i>	Edinburgh, Braid Hills	170	55° 56' N/03° 12' W	10	31	11	4	11
	Edinburgh, Colinton	120	55° 55' N/03° 14' W	10	20	18		4
	Edinburgh, Craigmillar	70	55° 56' N/03° 09' W	3	9	10	6	5
	Edinburgh, Gorebridge	90	55° 51' N/03° 04' W	1				
	Yorkshire, Swainby	200	54° 24' N/01° 16' W	1	1			
	Gloucestershire, Cotswolds	200	52° 03' N/01° 48' W	1	1			
	Gloucestershire, Cotswolds	240	51° 46' N/02° 08' W	1	1			
	London, Kew	10	51° 28' N/00° 16' W	1	1			
	Surrey, Westcott	130	51° 13' N/00° 22' W	1	1			
	Devon, Dawlish	20	50° 34' N/03° 28' W	1	1			
	Cornwall, Truro	50	50° 12' N/05° 03' W	1	1			
				30	67	39	10	20
Total <i>G. urbanum</i>								
Pure <i>Geum rivale</i>	Orkney, Berrisdale Hoy	40	58° 53' N/03° 23' W	1	1			
	Ullapool, Inchnadamph	180	58° 09' N/04° 56' W	1	1			
	Ullapool, Knockan Crag	250	58° 02' N/05° 03' W	1	1			
	Thurso, Betty Hill	20	58° 31' N/04° 14' W	1	1			
	Stirling, Ben Lawers	600	56° 32' N/04° 17' W	11	30	13		13
	Edinburgh, Pentland Hills, Carlops	300	55° 48' N/03° 21' W		2	2	1	
	Edinburgh, Pentland Hills, Bonaly	210	55° 53' N/03° 16' W		1	1	1	1
	Scottish Borders, Moorfoot Hills	300	55° 44' N/03° 01' W	11	32	16		8
	Scottish Borders, Moffat Dale	300	55° 23' N/03° 19' W	2	5		5	1
	Yorkshire, Cross Fell	800	54° 42' N/02° 29' W	1	1	1		1
	Yorkshire, Teesdale	260	54° 39' N/02° 09' W	1	1			
				30	76	32	7	24
Total <i>G. rivale</i>								
Hybrid swarm	Edinburgh, Balerno	150	55° 53' N/03° 20' W	34 (+20)	34	34		40

Abbreviations: N, north; W, west.

For each pure population, the numbers of individuals scored for AFLP and cpDNA markers, the number of individuals used for morphological measurements and as parents for F1 hybrids are shown (total number of individuals used per species in bold).

whereas in the remaining two crosses, *G. rivale* was the female parent. In all crosses, the offspring inherited the maternal plastid type. To test the species specificity of the cpDNA variants, the same PCR-RFLP assay was carried out on 67 *G. urbanum* and 76 *G. rivale* individuals collected from pure populations sampled throughout the United Kingdom (Table 1). The 34 individuals sampled from the hybrid swarm were also screened for this cpDNA variation.

Phenotypic analysis of genetic classes

Morphological variation: Nine morphological traits previously found to distinguish between *G. urbanum* and *G. rivale* (Marsden-Jones, 1930; Prywer, 1932; Waldren *et al.*, 1989) were measured on individuals grown in the glasshouse that had been sampled from pure populations (*G. urbanum* $n=39$, *G. rivale* $n=32$), from controlled F1s ($n=32$, including the same 15 individuals examined with AFLPs) and from the hybrid swarm ($n=34$, Table 1). The characters measured were length and width of stipules, petals and sepals, as well as length of the petal claw, gynophore and upper style joint.

To see whether the genetic classes identified by molecular markers in the hybrid swarm could be distinguished morphologically, the complete morphological data set obtained from hybrid swarm and reference individuals was subjected to principal component analysis (PCA). Morphological groupings were visualized by plotting the PCA1 score against the PCA2 score for each individual in the sample. To determine whether the genetic classes recognized by molecular marker analysis differed for individual morphological attributes in the hybrid zone, means and s.e. for each of the morphological traits were calculated for individuals classified as pure *G. urbanum*, pure *G. rivale*, F1 and backcrosses to *G. rivale*.

Pollen fertility: To establish whether pollen fertility is influenced by hybrid origin, it was measured in pure *G. urbanum* ($n=20$), pure *G. rivale* ($n=24$), controlled F1s ($n=14$) and 34 individuals from the hybrid swarm. The mean group fertility was established by averaging over all individual fertility counts in that group. All samples had been grown in a common environment (an unheated greenhouse) before anthers were removed for analysis in the summer of 2008. Pollen was collected from dehiscing anthers onto a microscope slide. Pollen grains were stained with 2% acetocarmine for 10 min and examined under a light microscope (McKellar and Quesenberry, 1992; Marutani *et al.*, 1993). The percentage of stained pollen grains was determined in two to five fields on the slide.

Auto-pollination: The ability to set seed in the absence of pollinators was measured on 33 plants from the hybrid swarm grown under standard conditions in an unheated greenhouse. Flowers were bagged using a bridal veil (mesh size 0.5×0.5 mm) before they had opened. Emasculated and bagged flowers were included as controls to ensure that self-pollen was in fact responsible for seed set. The presence or absence of seeds was recorded at the end of the growing season. Multiple flowers were bagged on each plant, and the percentage of individuals within each genetic group that successfully gave rise to at least one seed head was recorded (auto-pollination ability). In addition, each individual was scored for the proportion of its bagged

flowers that successfully set seed (% auto-pollination). The ability to auto-pollinate was measured on individuals from each of the genotypic classes recognized in the hybrid swarm on the basis of AFLP analysis.

Phenology: To establish the flowering periods of parental and hybrid phenotypes in the field, 11 representative plots of variable sizes ranging from 20 to 52 m² were established within the hybrid swarm. The plots covered ~10% of the area occupied by the hybrid swarm. Plots were 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 the 4 distinct phenotypic classes present: *G. urbanum*, *G. rivale*, F1 hybrid and backcross to *G. rivale*.

Results

Controlled F1 crosses

All nine *G. urbanum* $\times$ *G. rivale* and all five *G. rivale* $\times$ *G. urbanum* crosses were successful. There was no significant difference in the number of seeds produced per seed head between the two cross-directions ($F_{1,12}=0.0$, $P=0.99$). However, the overall germination rates for the two cross-directions differed significantly ($F_{1,12}=9.52$, $P=0.009$). The germination rate was 83% ($n=481$) where *G. urbanum* and 31% ($n=268$) where *G. rivale* was the female parent. Apart from one *G. urbanum* $\times$ *G. rivale* cross in which no seeds germinated at all, individual germination rates with *G. urbanum* as the female parent lay $>84\%$. In contrast, none of the *G. rivale* $\times$ *G. urbanum* crosses showed germination rates $>69\%$. Out of the 59 F1s that were raised in a greenhouse, only 1 individual died before the first fruit set.

Molecular markers

AFLP analysis: AFLP error rates: All 24 samples were scored for the same 154 polymorphic AFLP loci, resulting in 3696 phenotypic comparisons. Error rates were the lowest (2.6%) for runs using the same pre-amplification product, and the highest (5.1%) for runs using pre-amplification products from a different restriction-ligation process. The maximum individual error rate was 8%. There was little effect of this rate of error on placement of individuals on PCO plots (data not shown).

Classification of genotypic classes in the hybrid swarm: AFLP analysis of the total sample from pure populations using 4 primer-enzyme combinations yielded a total of 202 markers. These included 15 species-specific *G. urbanum* and 9 species-specific *G. rivale* markers. PCO analysis of the data from the hybrid swarm and reference individuals showed that the reference groups (pure *G. urbanum*, pure *G. rivale* and F1s from controlled crosses) were clearly separated along the first axis, which accounted for 51% of the total variation (Figure 1). Individuals from the hybrid swarm clustered in four more or less well-defined groups. Three of these groups overlapped with respectively the pure *G. urbanum* ($n=6$), the controlled F1 hybrids ($n=8$) and the pure *G. rivale* ($n=12$) samples. The fourth group ($n=8$) lay between the F1 hybrid and the pure *G. rivale* clusters and was tentatively recognized as backcrosses to *G. rivale*. No backcrosses to *G. urbanum* were detected, despite the assay of 20 additional *G. urbanum*-like individuals from the hybrid swarm (data not shown).

Figure 1 The PCO plot of individuals from a hybrid swarm in Edinburgh ($n = 34$, open symbols) based on variation at 202 AFLP markers. Reference samples (filled symbols) of *G. urbanum* ($n = 12$), *G. rivale* ($n = 12$) and F1 individuals from controlled crosses ($n = 15$) are also included. Hybrid classes are identified according to NEWHYBRIDS (Anderson and Thompson, 2002).

Table 2 shows the proportion of AFLP bands, classified as species specific based on 30 individuals per species collected from populations throughout the United Kingdom (Table 1), which are present in the 4 genetic groups tentatively recognized in the PCO plot within the hybrid swarm. Individuals classified as pure species contain only species-specific bands expected in their taxon. In the F1 group, some individuals lack a full complement of species-specific bands from *G. rivale*, suggesting that these markers may not be fixed in that species. The group recognized as backcrosses to *G. rivale* contain only 35% of the species-specific AFLP bands from *G. urbanum*, noticeably $<50\%$ expected in a first-generation backcross under random segregation of species-specific markers in the F1.

NEWHYBRIDS analysis of the same AFLP data set, comprising individuals from the hybrid swarm plus reference individuals, showed that four genetic groups can be recognized within the hybrid swarm. These are pure *G. urbanum* ($n = 6$, 17.7%), F1 hybrids ($n = 8$, 23.5%), backcross to *G. rivale* ($n = 8$, 23.5%) and pure *G. rivale* ($n = 12$, 35.3%). These correspond exactly to the four genetic groups postulated on the basis of the PCO analysis (Figure 1). The F1 individuals from controlled crosses grouped together with the eight individuals from the hybrid swarm, which had been identified as F1 genotypes by the program NEWHYBRIDS. None of the individuals in the hybrid swarm were classified either as an F2 or as a backcross to *G. urbanum*.

cpDNA variation: *cpDNA differentiation between taxa:* The survey of pure population samples showed that the cpDNA variation scored was species limited rather than species specific. *G. urbanum* was fixed for haplotype A, whereas *G. rivale* was polymorphic for haplotypes A and B. Among all samples, 28% of *G. rivale* individuals possessed haplotype A and 72% possessed haplotype B. Haplotype frequencies varied significantly between the

Table 2 Proportions of species-specific (sp-sp) AFLP bands, identified in the widespread population survey of pure *Geum urbanum* and *Geum rivale* populations, that were present in the different genetic classes recognized in the hybrid swarm ($n = 34$) by NEWHYBRIDS

Group	sp-sp <i>G. urbanum</i> bands (%)	sp-sp <i>G. rivale</i> bands (%)
<i>G. urbanum</i> -like samples (6)	98.9	0
<i>G. rivale</i> -like samples (12)	0.0	97.2
F1 (8)	100.0	72.2
Backcross Rivale (8)	35.0	93.1

Figure 2 The PCA plot based on 9 morphological characters measured in a common environment for 34 individuals collected from a hybrid swarm in Edinburgh (open symbols), as well as reference samples (filled symbols) from pure *G. urbanum* ($n = 39$), pure *G. rivale* populations ($n = 32$) and controlled F1 crosses ($n = 32$). The genetic group into which individuals are classified by the NEWHYBRIDS program is indicated on the PCA plot.

two populations of *G. rivale* that had been extensively sampled. Haplotype B was fixed in the population of *G. rivale* at Ben Lawers, but had a frequency of only 59% at the Moorfoot Hills site.

cpDNA variation in the hybrid swarm: In the hybrid swarm, all individuals classified on the basis of AFLP variation as *G. urbanum* ($n = 6$), as F1 hybrids ($n = 8$) or as backcrosses to *G. rivale* ($n = 8$) possessed haplotype A. In contrast, individuals classified as pure *G. rivale* ($n = 12$) were polymorphic for haplotypes A and B, each haplotype being present at a frequency of 0.5.

Phenotypic analysis of genotypic classes: *Morphological variation:* Figure 2 shows the results of the PCA analysis based on the morphological measurements of pure populations, F1s from controlled crosses and individuals from the hybrid swarm. The genetic group into which individuals are classified by the NEWHYBRIDS program is indicated on the PCA plot. The first two PCA axes account for 76.1 and 10.8% of the variation, respectively. Pure populations of *G. urbanum* and *G. rivale* are clearly separated at either end of the PCA1 axis, and F1 individuals from controlled crosses form a

Table 3 Mean ($\pm$ s.e.) for morphological characters measured under standard conditions for genetic classes recognized in the hybrid swarm by NEWHYBRIDS (Anderson and Thompson, 2002)

Trait	Hybrid swarm <i>Geum urbanum</i> n = 6	Hybrid swarm <i>Geum rivale</i> n = 12	Hybrid swarm F1 n = 8	Hybrid swarm backcrosses to <i>G. rivale</i> n = 8
Stipule length (cm)	2.08 ($\pm$ 0.07)	0.99 ($\pm$ 0.05)	1.34 ($\pm$ 0.09)	1.29 ($\pm$ 0.11)
Stipule breadth (cm)	2.33 ($\pm$ 0.23)	0.69 ($\pm$ 0.01)	1.19 ($\pm$ 0.10)	0.98 ($\pm$ 0.08)
Sepal length (cm)	0.63 ($\pm$ 0.02)	1.07 ($\pm$ 0.05)	0.86 ($\pm$ 0.04)	1.06 ($\pm$ 0.05)
Sepal breadth (cm)	0.33 ($\pm$ 0.02)	0.58 ($\pm$ 0.02)	0.53 ($\pm$ 0.02)	0.69 ($\pm$ 0.04)
Petal length (cm)	0.70 ($\pm$ 0.03)	1.32 ($\pm$ 0.06)	0.99 ($\pm$ 0.03)	1.31 ($\pm$ 0.02)
Petal breadth (cm)	0.57 ($\pm$ 0.02)	1.20 ($\pm$ 0.05)	0.89 ($\pm$ 0.03)	1.26 ($\pm$ 0.06)
Claw length (mm)	0.00 ($\pm$ 0.00)	5.88 ($\pm$ 0.25)	2.00 ($\pm$ 0.21)	5.63 ($\pm$ 0.42)
Gynophore (mm)	0.00 ($\pm$ 0.00)	6.17 ($\pm$ 0.97)	0.95 ($\pm$ 0.38)	4.13 ($\pm$ 0.48)
Upper style joint (mm)	1.50 ($\pm$ 0.00)	5.21 ($\pm$ 0.29)	3.31 ($\pm$ 0.16)	4.31 ($\pm$ 0.16)

Table 4 Mean ($\pm$ s.e.) for pollen fertility of plants grown under standard conditions and derived from pure populations of *Geum urbanum* and *Geum rivale*, for F1s produced by control crosses, and for the four genotypic classes recognized by NEWHYBRIDS analysis in a *Geum* hybrid swarm (Anderson and Thompson, 2002)

Groups	Sample size	Group fertility (% ($\pm$ s.e.))	Minimum fertility (%)	Maximum fertility (%)
Pure populations				
<i>G. urbanum</i>	20	94.3 ($\pm$ 2.4) ^a	54.7	100.0
<i>G. rivale</i>	24	90.3 ($\pm$ 3.5) ^a	20.0	99.0
Controlled F1 crosses	15	70.0 ($\pm$ 3.9) ^b	28.9	86.2
Hybrid swarm				
<i>G. urbanum</i>	6	94.0 ($\pm$ 1.4) ^a	89.6	98.6
<i>G. rivale</i>	12	88.8 ($\pm$ 3.6) ^a	62.5	98.8
F1	8	62.9 ($\pm$ 5.2) ^b	43.2	78.9
Backcrosses to <i>G. rivale</i>	8	83.4 ($\pm$ 4.0) ^a	63.1	95.6

Group fertility values with different superscripts (a, b) are significantly different ($P < 0.05$).

Table 5 Ability of individual flowers (auto-pollinating flowers %) and individual plants (auto-pollination capacity %) to set seed in the absence of pollinators for each of the four genetic groups found in the *Geum* hybrid swarm ($n = 33$)

Genetic group in the hybrid swarm	Number of individuals	Average number of bagged flowers per individual	Auto-pollinating flowers (%)	Auto-pollination capacity (%)
<i>Geum urbanum</i>	5	3.6	100.0	100.0
<i>Geum rivale</i>	12	3.3	13.2	25.0
F1	8	7.0	46.9	100.0
Backcross to <i>G. rivale</i>	8	4.0	23.8	62.5

third and distinct cluster between the pure parents. Individuals from the hybrid swarm classified as *G. urbanum*, *G. rivale* and F1s cluster with their respective reference groups. Those classified as backcrosses to *G. rivale* by AFLP analysis form a fourth cluster, which overlaps somewhat with the *G. rivale* cluster.

Within the hybrid swarm, individuals classified by AFLP markers as *G. urbanum* and *G. rivale* are highly significantly different for all nine morphological characters measured (Table 3). Individuals classified as F1 hybrids are morphologically intermediate between the parental taxa for every character. There are various degrees of dominance in the F1, with two traits showing dominance of *G. rivale* morphology (sepal length and petal breadth), one trait showing dominance of *G. urbanum* morphology (stipule length) and the remaining traits showing intermediate morphology. For seven of the nine characters, the *G. rivale* backcross group has mean values that are closer to the hybrid swarm *G. rivale* mean values than to the mean values of the hybrid swarm F1 group. For the remaining two characters, sepal breadth and petal breadth, the *G. rivale* backcross group

shows mean values that exceed those of the *G. rivale* parent, indicating transgressive segregation. However, this difference is significant only in the case of sepal breadth ($P < 0.05$).

Pollen fertility: Among pure populations of the parental species, for individuals classified as *G. urbanum* and *G. rivale* in the hybrid swarm, and individuals classified as backcrosses to *G. rivale* in the hybrid swarm, pollen fertility was uniformly high (close to 90%) (Table 4). However, controlled F1 hybrids and individuals classified as F1 hybrids in the hybrid swarm showed significantly lower pollen fertility (70.0 and 62.9%, respectively) than did the other groups tested ($P < 0.001$).

Auto-pollination: In the hybrid swarm sample, every individual in the *G. urbanum* group set seed when pollinators were excluded (Table 5). In contrast only 25% of individuals in the *G. rivale* group were capable of seed set when flowers were bagged, and the overall proportion of flowers producing seed after bagging was only 13.2%. In the F1 group, all plants were capable of setting seed after bagging. However, only 46.9% of flowers produced seed heads after auto-pollination. In the backcross

Figure 3 The flowering schedule measured in 2008 for five genetic groups within a hybrid *Geum* swarm in Edinburgh. Groups were identified on the basis of flower morphology. The 'Uncertain' group comprises a mixture of *G. rivale* and backcrosses to *G. rivale*. The overall contribution of flowers by the five genetic groups throughout the period was 42% for *G. urbanum*, 29% for *G. rivale*, 21% for F1s, 5% for backcrosses to *G. rivale*, 0% for backcrosses to *G. urbanum* and 3% for the uncertain group. A full color version of this figure is available at the *Heredity* journal online.

G. rivale group, only 62.5% of individuals were capable of setting seed when flowers were bagged, and the overall proportion of flowers successfully setting seed was only 23.8%. Differences in percentage auto-pollination among genetic groups were highly significant ($F_{3,35} = 16.04$, $P < 0.001$).

Phenology: Figure 3 shows flowering schedules for the four genetic groups that could be readily distinguished in the hybrid swarm on the basis of flower morphology (*G. urbanum*, *G. rivale*, F1 hybrid and backcrosses to *G. rivale*). Although the nine traits used in the PCA analysis did not completely distinguish *G. rivale* from backcrosses to *G. rivale*, the overall 'Gestalt' of early backcrosses, including not easily quantifiable measurements such as openness of flowers, hairiness and shades of petal colour, was usually distinctive (Supplementary Figure 1d). The fifth 'uncertain' category contains individuals, which were either *G. rivale* or backcrosses to *G. rivale*.

Peak flowering in *G. rivale* occurred ~4 weeks before peak flowering in *G. urbanum*, but there was significant overlap in flowering time between the two species over a period of ~3 weeks. The flowering time in F1 hybrids was intermediate between that of the two parents, and showed considerable overlap with both. Backcrosses to *G. rivale* genotypes had a similar but slightly extended flowering phenology compared with those of *G. rivale*. Although the data collection was not designed to provide an unbiased census, it does give a rough indication of the relative frequencies of the genetic groups that are present in the hybrid swarm. It is clear that hybrid genotypes represented approximately one-quarter of all individuals, and of these, 73% were F1 and 16% were backcrosses to *G. rivale* hybrids.

Discussion

Our molecular analysis of a recent hybrid swarm between self-fertilizing *Geum urbanum* and predominantly outcrossing *G. rivale* has documented the presence

of a high proportion of hybrids (28%), three-quarters of which are F1s and a quarter of which are backcrosses to *G. rivale*. No F2s and no backcrosses to *G. urbanum* have been detected. This situation indicates asymmetric introgression from inbreeder to outcrosser and is broadly in line with our expectations based solely on the genetic attributes of outcrossing and inbreeding plant lineages. The hybrid swarm is sufficiently young, so that variation falls into distinct parental and hybrid classes, rather than forming a continuum, and there exists a strong association between genotypic class and morphology of individuals.

Although our controlled crosses have shown that the F1 seed can be derived from crosses in either direction, a number of lines of evidence suggest that *G. urbanum* is likely to be the maternal parent of F1s in the hybrid swarm, in line with expectations. First, the absence of the species-limited cpDNA haplotype from *G. rivale* in F1 hybrids is consistent with this prediction. Second, the production of F1s from *G. urbanum* maternal parents is favoured by the flowering phenology. Both parents are protogynous. Pollen release at the end of the flowering period from earlier flowering *G. rivale* is likely to coincide with stigma receptivity at the beginning of the flowering period in later flowering *G. urbanum*, favouring pollen transfer from *G. rivale* to *G. urbanum* rather than *vice versa*. Finally, although seed is produced in controlled crosses in either direction, we have also demonstrated that the viability of the F1 seed is much lower when *G. rivale* is the mother. Difficulty with production of the viable F1 seed from *G. rivale* mothers has been noted before by a number of authors (Marsden-Jones, 1930; Prywer, 1932). A possible reason is that in the sexual conflict between pollen and maternal genomes during provisioning of the seed, the pollen genotype derived from the inbreeder is less able to elicit resources from the outbreeding maternal genome than *vice versa* (Wright *et al.*, 2008). This phenomenon has been noted as a possible barrier to hybridization in a range of other plant genera (Brandvain and Haig, 2005).

Although no genotypes resulting from selfing of the F1 were detected in the hybrid zone, our study indicates that the conditions necessary for their production are present. All F1 individuals were self-compatible and had high pollen fertility (63%). Moreover, when flowers were bagged, 47% were able to set seed by auto-pollination. There are a number of reasons that may be offered for failure to detect F2s. The first is that their frequency is low, and the sample size used was insufficient to detect their presence. The second is that outcross pollen from *G. rivale* outcompetes selfed pollen. The third is that selfed offspring are unfit as a consequence of the expression of genetic load present in the genome of *G. rivale*. Inbreeding depression in *G. rivale*, although low for an outcrossing species ($\delta = 0.33$), is much higher than in *G. urbanum* ($\delta = 0.0$) (Ruhsam *et al.*, 2010).

If an F2 generation was produced by selfing, inbred lines carrying varying proportions of the two parent genomes would be formed. Where inbreeding depression is high in the outcrosser, the most likely inbreeding lines to establish would be those carrying a limited contribution from the outcrossing parent. Inbreeding lines of this type would largely resemble *G. urbanum* and may be difficult to distinguish from pure *G. urbanum* using the limited number of species-specific markers

used in this study. Thus, a further possible reason for failing to detect F2s in the hybrid swarm is the lack of discriminatory power based on the genetic markers available. It is interesting to note that in previous studies of *Geum* hybrid swarms, individuals have been found resembling *G. urbanum* morphologically but displaying tolerance of waterlogging and manganese, adaptive characters normally associated with *G. rivale* (Waldren *et al.*, 1987, 1988). Such individuals could represent inbreeding lines derived from F1s carrying limited genetic contributions from *G. rivale*, which confer adaptation to waterlogged conditions.

The most striking finding from our study is the high prevalence of backcross genotypes to *G. rivale*, and the absence of backcrosses to *G. urbanum*. The data from cpDNA analysis, which show no evidence of the species-limited marker from *G. rivale* in the backcrosses to *G. rivale*, are consistent with the F1 being the female parent for these backcrosses. These findings are as predicted on the basis of the relative male fitness of the parental and hybrid genotypes and effectively result in asymmetric introgression of cpDNA and nuclear genes from *G. urbanum* into *G. rivale*. If this pattern of fertilization of maternal hybrids by *G. rivale* pollen continues in future generations, individuals with nuclear genomes that are predominantly from *G. rivale*, but possessing cpDNA from *G. urbanum* will be formed. The chloroplast haplotype of the inbreeder will be 'captured' by the outcrossing species, but not *vice versa*. It is interesting to note that populations of *G. rivale* from high elevation (where *G. urbanum* does not occur) tend to lack the chloroplast haplotype that is fixed in *G. urbanum*. In contrast, the frequency of this haplotype is higher in lower-elevation populations of *G. rivale* in which it co-occurs with *G. urbanum* (unpublished data). This may reflect the evolutionary effects of different levels of hybridization between the two species in the past.

Our general prediction of asymmetric introgression from inbreeding to outcrossing lineages as a consequence of hybridization between them is upheld in two other studies of homoploid hybridization, involving *Cyclamen* and *Mimulus*, in which genetic markers have been used to study the process and proceed beyond the first-generation hybrid (Sweigart and Willis, 2003; Thompson *et al.*, 2010). However, in the case of *Senecio*, introgression has been shown to occur in the opposite direction (Chapman and Abbott, 2010). In *Senecio*, hybridization occurs between taxa at different ploidy levels and involves the formation of a largely sterile triploid hybrid. Introgression relies on the rare formation of balanced gametes from this triploid, which then fertilize the tetraploid parent. Thus, this situation is quite different from that pertaining in the present case in which hybridization occurs between parents with the same chromosome number, the F1 hybrid of which is highly fertile. Therefore, we would not expect our predictions about the consequences of hybridization between outcrossing and inbreeding species to hold for situations such as *Senecio* in which hybridization occurs across ploidy levels.

Although our predictions about early evolution through recurrent pollination of both *G. urbanum* and the F1 by *G. rivale* have been upheld, the situation is predicted to become more complicated in the future. First-generation backcross genotypes are likely to enjoy

high male fitness and be successful in many outcross pollinations. These backcrosses are fully pollen fertile, have a distinct and larger flower morphology than do *G. rivale* (an expression of transgressive segregation) and show an extended flowering period. Thus, we expect these individuals to act as pollen donors in the future. To understand how the presence of backcross genotypes affects future evolution in the hybrid swarm, further work is required to analyze the patterns of mating within and among backcrosses, F1s and parental genotypes in the hybrid swarm, and to document the range of phenotypes and genotypes that are thereby generated.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This research forms a part of the PhD thesis of Markus Ruhsam, which was funded by the Torrance Bequest (University of Edinburgh). The chloroplast work was funded by the BSBI. We thank the Edinburgh City Council, the Ben Lawers National Nature Reserve and various landowners for the support of our fieldwork.

References

- Anderson EC, Thompson EA (2002). A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**: 1217–1229.
- Balfour DH (1863). *Flora of Edinburgh—Being a List of Plants Found in the Vicinity of Edinburgh*. Adam & Charles Black: Edinburgh.
- Bonin A, Bellemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P (2004). How to track and assess genotyping errors in population genetics studies. *Mol Ecol* **13**: 3261–3273.
- Brandvain Y, Haig D (2005). Divergent mating systems and parental conflict as a barrier to hybridization in flowering plants. *Am Nat* **166**: 330–338.
- Chapman MA, Abbott RJ (2010). Introgression of fitness genes across a ploidy barrier. *New Phytol* **186**: 63–71.
- Charlesworth D, Charlesworth B (1987). Inbreeding depression and its evolutionary consequences. *Annu Rev Ecol Syst* **18**: 237–268.
- Charlesworth D, Morgan MT (1991). Allocation of resources to sex functions in flowering plants. *Philos T Roy Soc B* **332**: 91–102.
- Diaz A, Macnair MR (1999). Pollen tube competition as a mechanism of prezygotic reproductive isolation between *Mimulus nasutus* and its presumed progenitor *M. guttatus*. *New Phytol* **144**: 471–478.
- Donnelly K (2010). *Assessing the Efficiency of NEWHYBRIDS to Analyse Hybridization between Outcrossing and Inbreeding Plant Species*, Unpublished MSc Thesis, University of Edinburgh.
- Doyle JJ, Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* **12**: 13–15.
- Ennos RA, French GC, Hollingsworth PM (2005). Conserving taxonomic complexity. *Trends Ecol Evol* **20**: 164–168.
- Ferguson CJ, Levin DA, Jansen RK (1999). Natural hybridization between an outcrossing and a selfing *Phlox* (Polemoniaceae): the maternal species of F1 hybrids. *Plant Syst Evol* **218**: 153–158.
- Fishman L, Aagaard J, Tuthill JC (2008). Toward the evolutionary genomics of gametophytic divergence: patterns of transmission ratio distortion in monkeyflower (*Mimulus*) hybrids reveal a complex genetic basis for conspecific pollen precedence. *Evolution* **62**: 2958–2970.

- French GC, Hollingsworth PM, Silverside AJ, Ennos RA (2008). Genetics, taxonomy and the conservation of British *Euphrasia*. *Conserv Genet* **9**: 1547–1562.
- Glemin S, Bazin E, Charlesworth D (2006). Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proc R Soc Lond B Biol Sci* **273**: 3011–3019.
- Goldberg EE, Kohn JR, Lande R, Robertson KA, Smith SA, Igic B (2010). Species selection maintains self-incompatibility. *Science* **330**: 493–495.
- Hammer Ø, Harper DAT, Ryan PD (2006). *PAST—Paleontological Statistics, version 1.42*. Published at: <http://folk.uio.no/ohammer/past>.
- Hollingsworth PM, Squirrell J, Hollingsworth ML, Richards AJ, Bateman RM (2007). Taxonomic complexity, conservation and recurrent origins of self-pollination in *Epipactis* (Orchidaceae). In: Bailey J, Ellis RG (eds). *Current Taxonomic Research on the British and European Flora*. BSBI: London, pp 27–44.
- Husband BC, Schemske DW (1996). Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution* **50**: 54–70.
- Igic B, Lande R, Kohn J (2008). Loss of self-incompatibility and its evolutionary consequences. *Int J Plant Sci* **169**: 93–104.
- Knuth P (1898). *Handbuch der Blütenbiologie*. Vol. 2, part 1 Verlag Wilhelm Engelmann: Leipzig.
- Lloyd DG (1987). Allocations to pollen, seeds and pollination mechanisms in self-fertilizing plants. *Funct Ecol* **1**: 83–89.
- Marsden-Jones EM (1930). The genetics of *Geum intermedium* Willd. haud Ehrh., and its backcrosses. *J Genet* **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. *Am J Bot* **80**: 93–103.
- McKellar MA, Quesenberry KH (1992). Chromosome pairing and pollen viability in *Desmodium ovalifolium* Wall × *Desmodium heterocarpon* (L) Dc hybrids. *Aust J Bot* **40**: 243–247.
- 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 Soc Bot Pol* **9**: 87–114.
- Ruhsam M (2009). *Evolutionary Consequences of Hybridization between Geum urbanum and G. rivale*. Unpublished PhD thesis, University of Edinburgh.
- Ruhsam M, Hollingsworth PM, Squirrell J, Ennos RA (2010). Significant differences in outcrossing rate, self-incompatibility, and inbreeding depression between two widely hybridizing species of *Geum*. *Biol J Linn Soc* **101**: 977–990.
- Schemske DW, Lande R (1985). The evolution of self-fertilization and inbreeding depression in plants. 2. Empirical observations. *Evolution* **39**: 41–52.
- Schoen D, Busch J (2008). On the evolution of self-fertilization in a metapopulation. *Int J Plant Sci* **169**: 119–127.
- Smedmark JE, Eriksson T, Evans RC, Campbell CS (2003). Ancient allopolyploid speciation in Geinae (Rosaceae): evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. *Syst Biol* **52**: 374–385.
- Smith PM, Dixon ROD, Cochran MP (2002). *Plant Life of Edinburgh and the Lothians*. Edinburgh University Press: Edinburgh. p 416.
- Swiegiart 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.
- Taylor K (1997a). Biological flora of the British Isles: *Geum rivale*. *J Ecol* **85**: 721–731.
- Taylor K (1997b). Biological flora of the British Isles: *Geum urbanum*. *J Ecol* **85**: 705–720.
- Thompson JD, Gaudet M, Debussche M (2010). Conservation value of sites of hybridization in peripheral populations of rare plant species. *Cons Biol* **24**: 236–245.
- 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. *Mol Ecol* **16**: 4171–4179.
- Vos P, Hogers R, Bleeker M, Reijans M, Vandele T, Hornes M et al. (1995). AFLP—a new technique for DNA-fingerprinting. *Nuc Acids Res* **23**: 4407–4414.
- Waldren S, Davies MS, Etherington JR (1987). The effect of manganese on root extension of *Geum rivale* L., *G. urbanum* L. and their hybrids. *New Phytol* **106**: 679–688.
- 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 Phytol* **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. *Bot J Linn Soc* **101**: 199–211.
- Wright SI, Ness R, Foxe J, Barrett SCH (2008). Genomic consequences of outcrossing and selfing in plants. *Int J Plant Sci* **169**: 105–118.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)