Molecular characterization of DNA sequences from the *Primula vulgaris* S-locus

Iain W. Manfield, Vassily K. Pavlov, Jinhong Li, Holly E. Cook, Florian Hummel and Philip M. Gilmartin*

Centre for Plant Sciences, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

Received 10 September 2004; Accepted 20 December 2004

Abstract

*Primula* species provide possibly the best known examples of heteromorphic flower development and this breeding system has attracted considerable attention, including that of Charles Darwin. However, despite considerable recent advances in molecular genetics, nothing is known about the molecular basis of floral heteromorphy. The first molecular marker for the *Primula* S-locus is reported here. This DNA sequence was identified by random amplification of polymorphic DNA (RAPD)-PCR, further defined as a sequence characterized amplified region (SCAR) marker, and subsequently shown to correspond to a restriction fragment length polymorphism (RFLP) that is linked to the thrum allele of the *Primula* S-locus. The sequence of 8.8 kb of genomic DNA encompassing this thrum-specific RFLP is presented. Analysis of this DNA reveals a highly repetitive sequence structure similar to that found at the S-locus in other species; it also contains sequences similar to elements of a Gypsy-like retrotransposon. The identification of a specific DNA sequence associated with the thrum allele of the *Primula* S-locus provides the first molecular probe with which to investigate the molecular basis of heteromorphic flower development in *Primula*.

Key words: Flower development, heteromorphy, *Primula*, S-locus.

Introduction

The existence of two forms of flower in species of *Primula* has long been documented; Darwin’s paper to the Linnean Society in 1862 (Darwin, 1862) and his subsequent book, *The different forms of flowers in plants of the same species* (Darwin, 1877), describe his observations on the structure and breeding behaviour of pin and thrum flowers of *Primula* species including the primrose, *Primula vulgaris*. Pin flowers are characterized by a long style with a stigmatic surface presented at the corolla mouth. In these flowers, the anthers are attached to the petals by short filaments half way down the corolla tube. By contrast, the stamens in thrum-form flowers are presented at the mouth of the flower with the stigma midway down the corolla.

Darwin’s correlation between the phenotype and the cross-compatibility of the two forms of flower provided the foundation for subsequent analyses of heterostyl in *Primula*. His observations on intermorph fertility demonstrated the presence of what we now recognize as an incompatibility system that reinforces the physical barriers to self-fertilization (Darwin, 1862, 1877). The physical differences between the two floral morphs also include pollen size and number (Darwin, 1862; Ornduff, 1979; Heslop-Harrison et al., 1981; Piper et al., 1986; Kurian and Richards, 1997), the length of stigmatic papillae, stigma shape, and style cell length (Ganders, 1979; Heslop-Harrison et al., 1981; Al Wadi and Richards, 1993; Richards, 1997).

The physical differences between pin and thrum plants are determined by the diallelic S-locus. Pin plants are homozygous for the recessive allele (*s/s*), and thrum plants are heterozygous with one dominant *S* allele (*S/s*). This locus comprises a co-adapted linkage group (Dowrick, 1956; Charlesworth and Charlesworth, 1979a, b; Richards, 1984) that controls development of the two different mating types. Observations by Ernst (Ernst, 1933, 1936, 1955) on self-fertile and homostyle *Primula* provided data from which Dowrick suggested that the S-locus comprises at least three genes referred to as *G*, which controls aspects of the gynaecium development and female incompatibility; *P*, which controls pollen size and male incompatibility, and *A*, which determines anther position within the flower (Dowrick, 1956). Pin plants are
homozygous recessive for all three loci (gpa/gpa), and thrum plants are heterozygous (GPA/gpa) (Richards, 1997).

Self-fertile long homostyle and short homostyle flowers that result from rare recombination events within the S-locus have been used to determine the gene order of G, P, and A within the S-locus (Emrst, 1955; Dowrick, 1956; Piper et al., 1984; Wedderburn and Richards, 1992; Lewis and Jones, 1993; Richards, 1997). Recombination between G and P yields gametes with the genotypes gPA and Gpa. Combination of such gametes with those from a pin plant (gpa) will produce progeny with genotypes gPA/gpa and Gpal/gpa, respectively. The former will develop a pin-form gynaecium, large pollen and anthers in the thrum position at the mouth of the corolla; this is referred to as a long homostyle. Plants that possess a dominant G allele in combination with recessive p and a alleles produce short homostyles that produce small pollen and develop both anthers and stigma mid-way down the corolla tube. Long and short homostyles represent valuable tools for the molecular genetic mapping of the Primula S-locus.

Observations of pollen size and self-incompatibility behaviour indicate sporophytic control of both pollen size and the male incompatibility response as all pollen produced by heterozygous thrum plants is of uniform size and incompatibility phenotype; homozygous thrums (S/S) are not found. The work of Kurian and Richards (1997) on the S-locus of P.tommasinii provided evidence of recombination with the S-locus that resulted in the loss of thrum-dominance for pollen size with recombinant P* plants producing two sizes of pollen corresponding to that found in both pin and thrum flowers. These studies suggest a gene controlling dominance for pollen size (Mpm) (Kurian and Richards, 1997). Furthermore, these data enabled the dissection of pollen size and male incompatibility functions previously referred to in combination as P into two distinct genetic functions, Pp and Pm, respectively. This study also provided evidence for the presence of a second determinant of gynoecium development (Gm) as well as a recessive thrum-linked lethal (l) which prevents survival of homozygous (S/S) thrum plants. In combination, these analyses suggest at least seven genes at the S-locus (Kurian and Richards, 1997). Of additional relevance is the finding that the floral homeotic mutant Hose-in-hose (Gerard, 1597; Parkinson, 1629; Webster and Gilmartin, 2003) is also linked to the thrum allele of the S-locus (Webster and Grant, 1990).

However, despite these classical genetic analyses, the molecular genetic structure of the Primula S-locus and identity of genes controlling heteromorphic flower developmental remain unknown. Heteromorphy and self-incompatability are widespread throughout the Primulaceae (Ganders, 1979; Wedderburn and Richards, 1992; Richards, 1993; Mast et al., 2001), and models have been proposed to explain the evolution of this breeding system (Charlesworth and Charlesworth, 1979a, b; Muenchow, 1981; Charlesworth, 1982; Barrett, 1990; Wedderburn and Richards, 1992; Al Wadi and Richards, 1993) which has a number of distinct features as compared with homomorphic self-incompatibility systems (de Nettancourt, 1997). An understanding of the molecular structure of the Primula S-locus will not only provide further insights into the evolution of self-incompatibility systems, but it will also lead to the identification of the molecular mechanisms underpinning heteromorphic flower development.

Based on this extensive background information, the identification of molecular genetic markers for the Primula S-locus was undertaken as the first step towards characterization of the genes that control heteromorphy in this plant. The approaches were based on the assumption that polymorphisms linked to the thrum allele of the S-locus (S) will be absent from homozygous s/s pin plants. In order to improve the chances of finding S-locus-linked DNA polymorphisms in the varied genetic background of an obligate out-breeder the commercially available diploid F1 hybrid P.vulgaris var. Blue Jeans derived from two inbred parental lines was used. The very first molecular marker for the Primula S-locus and the genomic sequence surrounding the region from which it is derived is reported here.

Materials and methods

Plant material

Primula vulgaris var. Blue Jeans seed was obtained from Thomson and Morgan seeds and from The Farmen Seed Company in Naples. Long and short homostyles were provided by Margaret Webster. Plants were pot grown in Levingtons plant protection compost and maintained in pots in an unheated greenhouse under natural daylight. During the summer plants were placed outside on hard standing.

Genomic DNA purification

Leaves were ground to a fine powder in liquid nitrogen with a mortar and pestle and resuspended in 25 ml CTAB extraction buffer (140 mM sorbitol, 220 mM TRIS-HCl pH 8.0, 22 mM EDTA, 800 mM NaCl, 1% sarcosyl and 0.8% CTAB) g⁻¹ of tissue. The suspension was incubated at 65 °C for 20 min before the addition of 10 ml chloroform:isoamyl alcohol (24:1, v:v) and gentle mixing for 20 min at room temperature. After centrifugation at 3000 g for 5 min, the aqueous phase was transferred to a clean tube, mixed with 17 ml isopropanol and placed on ice for 10 min. Centrifugation at 3000 g for 5 min collected the precipitate which was redissolved in 4 ml TE buffer (10 mM TRIS-HCl pH 8.0, and 1 mM EDTA). Lithium acetate (1 vol., 4 M) was added to the sample which was incubated on ice before centrifugation at 3000 g for 10 min. The supernatant was transferred to a clean tube, mixed with 2 vols of ice-cold ethanol and incubated on ice for 20 min. The precipitate was again collected by centrifugation and redissolved in 0.9 ml TE and 0.1 ml 3 M sodium acetate. The mixture was extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1 by vol.) and once with chloroform:isoamyl alcohol (24:1, v:v). DNA was precipitated from the aqueous phase by the addition of 2 vols of ice-cold ethanol and incubation on ice for 5 min before centrifugation at 13 000 g for 5 min. The pellet was redissolved...
in 250 µl TE and stored at −20 °C until use. Alternatively, fresh, young leaves were snap-frozen in liquid nitrogen immediately after harvesting, and used for genomic DNA purification using the Nucleon Phytopure DNA extraction kit following the manufacturer’s instructions (Amersham Pharmacia Biotech). DNA quality and concentration were estimated by the electrophoresis of samples through agarose gels containing ethidium bromide and comparison of band intensity to standards of known amount.

**RAPD PCR using 10-mer primers**

Amplification reactions were performed in a 25 µl volume containing 10 mM TRIS-HCl, pH 8.3, 50 mM KCl, 2.25 mM MgCl₂, 0.001% gelatine, 0.2 mM each dNTP and 1.75 units Taq (Super-Taq, HT Biochemicals). Each reaction contained 5–50 ng of genomic DNA template obtained using the CTAB method. Oligonucleotide primers were from Operon Technologies and 15 pmol of a 10-base primer was used per reaction. Samples were heated to 95 °C for 5 min before 25–35 cycles of 94 °C for 5 s, 36 °C for 30 s and 72 °C for 60 s followed by a final phase of 5 min at 72 °C. Reaction products were analysed by electrophoresis through 1–2% agarose gels and visualized by ethidium bromide staining after electrophoresis.

**Cloning differential RAPD products**

PCR cloning primers generated using Taq polymerase were ligated to the pCRII TA cloning vector (Invitrogen) and used to transform *E. coli* to ampicillin resistance. Colonies were checked for the presence of inserts in the plasmid and recombinant plasmids carrying the inserts of the anticipated size were sequenced by dye-terminator cycle sequencing using an ABI 373 sequencer. The sequence was used to design two 15-mer SCAR primers (P1; AGGCCCGATGAGCGT, P2; AGGGCCGATGATTCC) containing the original 10 bases plus an additional five nucleotides at the 3’ end.

**PCR using SCAR primers**

PCR using SCAR primers was performed in a 25 µl reaction volume containing buffer, 2.5 mM MgCl₂, 0.2 mM each dNTP, 0.2 µM each primer, 1.5 units Super-Taq (HT Biochemicals), and 50 ng *Primula* genomic DNA. After PCR thermal cycling (94 °C for 1 min, 30 cycles of 94 °C, 57 °C, and 72 °C for 30 s each followed by 10 min at 72 °C), loading buffer was added to each tube and samples analysed by electrophoresis through 1% agarose gels.

**Southern blotting**

Gels were washed for 30 min in denaturing buffer (1.5 M NaCl, and 0.5 M NaOH) then twice for 15 min in neutralizing buffer (1. M NaCl and 0.5 M TRIS-HCl pH 7.4) and DNA blotted onto Hybond-N membranes (Amersham Pharmacia Biotech) by capillary transfer in 20× SSC overnight. Membranes were rinsed in 2× SSC, dried between sheets of 3MM paper (Whatman) before UV cross-linking to the membrane (Stratalinker, Stratagene).

**RFLP analysis**

Genomic DNA (10 µg) was digested overnight with restriction enzymes in 200 µl of the manufacturer’s recommended buffer. The sample was extracted with 1:1 (v:v) phenol/chloroform and then with chloroform before ethanol precipitation. DNA pellets were washed with 70% ethanol, air-dried, and redissolved in 25 µl TE. Digestion products were resolved by electrophoresis through 0.7% agarose gels before blotting onto nylon membranes as described above. DNA probe fragments for Southern hybridization were generated either by restriction enzyme digestion of plasmids or by PCR using P1/P2 primers. Probe fragments were purified from agarose gels by electrophoresis onto dialysis membranes or using a gel extraction kit (Qiagen). DNA concentrations were estimated by comparison of fluorescence intensity against standards of known amount after agarose gel electrophoresis. Membranes were incubated in pre-hybridization solution (6× SSC, 0.5% SDS, 5× Denhardt’s solution, and 0.1 mg ml⁻¹ denatured, sonicated salmon-sperm DNA) at 65 °C for 2 h. Probe DNA (50 ng) was radiolabelled using ‘Ready-to-Go’ labelling beads (Amersham Pharmacia Biotech) with 50 µCi σ³²P-dCTP (ICN) and unincorporated isotope removed by chromatography through Nuc-trap columns (Stratagene). The sample was heated to 100 °C for 2 min and plunged into ice for 1 min before addition to the hybridization bottle. Membranes and radiolabelled probe were incubated together at 65 °C overnight with constant agitation. Unbound probe was removed by two 15 min washes with 2× SSC/0.1% SDS then two washes with 0.1× SSC/0.1% SDS at 65 °C. Autoradiography was performed using Kodak Biomax MS film at −70 °C with intensifying screens for the required length of time.

**Screening a phage lambda Primula genomic DNA library**

A phage lambda genomic DNA library was constructed in λ Fix (Stratagene) with partial Sau 3A digestion of *Primula* ‘Blue Jeans’ genomic DNA following the manufacturer’s instructions and screened using the 323 bp RAPD marker following the manufacturer’s instructions (Stratagene). Insert sequences were digested using a range of restriction enzymes, blotted onto Hybond membranes (Amersham Pharmacia Biotech), and probed with the 323 bp product to identify fragments which contained regions hybridizing to the probe. An approximately 9 kb BamHI fragment was sub-cloned into BamHI digested pBluescript following standard procedures. The sequence of this region was determined by sequencing overlapping subcloned restriction fragments. The sequence data was compiled into a single contig using the DNASTAR Seqman program. Primers P3 (CTCCGAGCCTC-TCCAGTTTTA), P4 (GGCACCTTACCCACGAGG), P5 (CCATAACGATGCCCAGGAGC), and P6 (CGCTAA-GAGCATAGGCCGGAG) were designed to the end sequences of the 8.8 kb BamHI fragment.

**Bioinformatic analysis**

Webcutter (www.firstmarket.com/cutter/cut2.html) was used to predict restriction enzyme sites within DNA sequence and sequence comparisons to DNA and protein databases were undertaken using NCBI-BLAST (www.ncbi.nlm.nih.gov/blast/). Self alignment of the genomic sequence was done using NCBI BLAST 2 (www.ncbi.nlm.nih.gov/blast) [Altschul et al., 1990] and repeat sequences defined using RepeatFinder tools (http://tandem.bu.edu/trf/trf.html) [Benson, 1999]. Prediction of putative genes structures was undertaken using gene prediction software using the default parameters (www.softberry.com/berry.phtml).

**Results**

**Identification of a thrum-specific RAPD marker**

The heterozygous nature of thrum plants (S:s) with dominant and recessive alleles as compared to homozygous recessive (s/s) pin plants, provides the opportunity to discriminate between these two floral morphs at the molecular level. It was reasoned that if it was possible to identify polymorphisms associated with the dominant S allele present in thrum plants, but absent from pin plants, these polymorphisms could be used as sequence tags for mapping these regions.
to initiate a molecular characterization of the *Primula S*-locus.

To initiate the search for thrum-specific DNA polymorphisms, an F₁ hybrid line *Primula vulgaris var. Blue Jeans* was selected. This line was developed by the Farmen Seed Company (Naples) from crosses between inbred parental lines. It was anticipated that this horticultural variety would provide a significant advantage over wild primrose as the inbred nature of the parental lines would reduce the polymorphic background normally associated with an obligate out-breeder. Therefore any polymorphism identified from thrum plants of the F₁ hybrid would arise either from differences between the two inbred parental lines, or differences in the two alleles of the S-locus. Segregation of polymorphisms between thrum and pin plants in the F₂ generation would distinguish between these two possibilities. The analysis of genomic DNA was chosen over techniques involving the analysis of differential transcript abundance to ensure that markers would be derived from the S-locus rather than represent genes which are differentially regulated in response to the S-locus.

Genomic DNA from thrum and pin plants was screened by PCR using 80 different decamer RAPD primers from the Operon W, X, Y, and Z primer sets. The majority of these primers showed no differential amplification products between the thrum and pin DNA samples (data not shown). However, some primers did show thrum-specific amplification products; examples of these are shown in Fig. 1. RAPD analysis of thrum and pin genomic DNA using the Operon W6 primer yielded over 20 amplification products, the majority of which are the same between both genomes. However, a fragment of approximately 300 bp consistently observed in DNA from thrum plants (marked with an arrow on Fig. 1) was consistently absent from pin DNA samples. A thrum-specific amplification product using the Operon Z9 primer (marked with an arrow on Fig. 1) was also obtained. In this case, the amplification product was approximately 1 kb. Subsequent analysis of the 1 kb Z9 amplification product using DNA samples from a larger population of thrum and pin plants revealed that it was not uniquely associated with the thrum phenotype (data not shown). However, further characterization of the 300 bp W6 amplification product revealed that it co-segregated exclusively with the thrum phenotype as described below.

In order to characterize the approximately 300 bp W6-derived thrum-specific amplification product, the DNA fragment was excised from an agarose gel and cloned into pBluescript. To confirm that the correct amplification product had been cloned, the DNA insert was isolated from the plasmid, radiolabelled and used as a hybridization probe against a Southern blot of RAPD products obtained by PCR amplification from pin and thrum genomic DNA using the W6 primer. Results of this analysis using seven thrum and seven pin plants are presented in Fig. 2. The presence of a hybridization signal at approximately 300 bp in the thrum plants, but not the pin plants, clearly reveals that the correct DNA sequence had been isolated from the original RAPD gel. Other hybridizing bands were also identified, some of which were present in both pin and thrum DNA (Fig. 2A).

**Development of a SCAR marker**

The DNA sequence of the subcloned thrum-specific amplification which was 323 bp in length (Fig. 2B) was determined: this sequence has been deposited in Genbank (Accession number AY854262). Comparison of this sequence with available DNA databases provided no clue to its function; BLAST searches revealed no significant similarity to other DNA sequences. Translation of the sequence predicted an open reading frame of 71 amino acid residues which shows no similarity to available protein databases. Northern analysis of total RNA from developing and mature pin and thrum flowers using the 323 bp product gave no indication that this sequence is expressed in flowers (data not shown). Furthermore, subsequent analysis of the genomic sequences flanking this sequence (see below) indicated that this sequence is not part of a transcription unit.

The ten base primer sequence (AGGCCCGATG) at either end of the 323 bp sequence is highlighted in bold. During PCR, any base mismatches between primer sequence and target sequence become obscured as the primer replaces the native DNA sequence during the amplification process. In order to increase the resolution
of these analyses and to remove background amplification products common to both pin and thrum RAPD profiles, two 15-base primers that contain the original RAPD primer sequence plus an additional five nucleotides at the 3' end as determined by the sequence of the 323 bp amplification product were synthesized. These primer sequences are shown in Fig. 2B.

These 15-mer oligonucleotides were used as primers to generate a sequence-characterized amplified region (SCAR) marker that discriminates between DNA isolated from thrum and pin plants. Figure 2C shows the amplification of DNA from three thrum and three pin plants. The use of different annealing temperatures made it possible to distinguish clearly between the two floral morphs. PCR amplification of thrum and pin genomic DNA samples with the SCAR primers at 54 °C generated a ladder of products from both samples with the predicted 323 bp thrum-specific product. In addition to the original 323 bp product, this analysis revealed another, approximately 3 kb thrum-specific product as well as an approximately 2 kb pin-specific product. Increasing the annealing temperature to 58 °C reduced the number of background bands, including the larger pin and thrum-specific products and generated a unique thrum-specific amplification product at 323 bp.

Identification of a thrum-specific RFLP

The RAPD amplification product was used to extend the analysis of pin and thrum samples by genomic Southern blot analysis. Genomic DNA was isolated from individual pin and thrum plants of the F1 hybrid Primula var. 'Blue Jeans', digested with a range of restriction enzymes and analysed on Southern blots using the 323 bp DNA fragment as a probe. Data obtained from ten pin and ten thrum plants are presented in Fig. 3. In this example, DNA was digested with KpnI. In DNA samples from both floral morphs hybridization to a large diffuse, approximately 20 kb, band as well as to DNA fragments at 4.2 kb and 3.5 kb was obtained. However, a dramatic difference between the thrum and pin DNA samples is revealed by the presence of two

![Fig. 2.](https://academic.oup.com/jxb/article-abstract/56/414/1177/551511) Characterization of the RAPD product and SCAR marker optimization. (A) Southern blot analysis of RAPD-PCR products. PCR products obtained with primer W6 from seven thrum and seven pin genomic DNA samples were separated by agarose gel electrophoresis and blotted onto nylon membranes. The 323 bp thrum-specific fragment was radiolabelled as a probe and hybridization detected by autoradiography. The thrum-specific product is highlighted with an arrow. (B) Sequence of the cloned 323 bp RAPD-PCR product (Genbank Accession number AY854261). The forward and reverse SCAR primers are highlighted in bold at the beginning and end of the sequence with the common 10-mer primer sequence underlined. (C) Optimization of SCAR-PCR conditions. SCAR primers of 15 nucleotides were used in PCRs with thrum and pin genomic DNA samples at two annealing temperatures as shown. Products were resolved by agarose gel electrophoresis and visualized using ethidium bromide. Thrum (T) and pin (P) samples are indicated and morph-specific products are marked with arrows.
additional \( KpnI \) bands in all the thrum samples. These two bands at 2.6 kb and 1.4 kb are totally absent from pin DNA samples. This finding reveals that not only does the SCAR marker also correspond to an RFLP marker, but that this region of the genome is duplicated in thrum plants.

**The RAPD, SCAR, and RFLP markers co-segregate with the thrum phenotype**

Following the definition of a thrum-specific RAPD amplification product, conversion of this sequence into a SCAR marker and identification of a thrum-specific RFLP suggested that a DNA fragment corresponding to the thrum allele of the \( S \)-locus had been identified. To extend these analyses to confirm linkage to the \( S \)-locus, a larger population of \( F_1 \) and \( F_2 \) individuals (from multiple independent families) was analysed using both SCAR and RFLP analysis. Data presented in Fig. 4A shows the analysis of a sample of 12 individual \( F_2 \) progeny. Amplification with the SCAR primers only yields the 323 bp PCR product with thrum and not pin DNA. No other amplification products are seen. PCR amplification of a product from the \( Primula Plena \) gene (Cook, 2002) that is unlinked to the \( S \)-locus was used as a positive control. Analyses of a larger number of \( F_1 \) and \( F_2 \) plants, 191 in total, provided further confirmation that this region of DNA is linked to the thrum allele of the \( S \)-locus (Table 1). The 323 bp amplification product was present in all 99 thrum plants analysed, but was not obtained from any of the 92 pin plants.

The RFLP analysis was also extended to 52 \( F_2 \) plants (Table 1) using the 323 bp RAPD fragment as a probe. A sample of these results are shown in Fig. 4B. These data confirm that the RFLP polymorphisms are associated with the thrum genome; all thrum plants analysed contain the two thrum-specific hybridization products. As observed for the original \( F_1 \) plants, the 4.2 kb and 3.5 kb hybridization bands are seen in both pin and thrum samples with the 1.4 kb and 2.6 kb bands only observed in thrum plants to reveal complete linkage of the polymorphisms with the thrum phenotype.

![Fig. 3. Southern blotting analysis of thrum and pin genomic DNA samples reveals an RFLP. Genomic DNA from 10 thrum (A) and 10 pin (B) plants was digested with \( KpnI \) digestion products were resolved on 0.7% agarose gels and blotting onto nylon membranes. Products homologous to the 323 bp thrum-specific probe were detected by hybridization and visualized by autoradiography. Sizes of hybridizing bands are indicated alongside each image with the thrum-specific RFLPs at 1.4 and 2.6 kb highlighted in bold.](https://academic.oup.com/jxb/article-abstract/56/414/1177/551511)

![Fig. 4. Co-segregation of the SCAR and RFLP markers with the thrum allele of the \( S \)-locus in \( F_2 \) plants. (A) Genomic DNA samples from thrum and pin \( F_2 \) plants were analysed by PCR using the SCAR primers and control primers derived from the \( Primula Plena \) gene that is unlinked to the \( S \)-locus. PCR products from six thrum and six pin samples, were resolved by agarose gel electrophoresis. (B) Genomic DNA from ten thrum and ten pin \( F_2 \) samples was digested with \( KpnI \) and products resolved by agarose gel electrophoresis, blotted onto nylon membranes, and probed with radiolabelled 323 bp thrum specific marker. Sizes of hybridizing bands are shown.](https://academic.oup.com/jxb/article-abstract/56/414/1177/551511)
Table 1. Segregation of the thrum-specific polymorphism in F1 plants and their progeny

<table>
<thead>
<tr>
<th></th>
<th>F1 generation</th>
<th>F2 generation</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of plants</td>
<td>No. of plants</td>
<td>No. of plants</td>
</tr>
<tr>
<td>Pin</td>
<td>0/10</td>
<td>0/16</td>
<td>0/26</td>
</tr>
<tr>
<td>Thrum</td>
<td>10/10</td>
<td>16/16</td>
<td>26/26</td>
</tr>
<tr>
<td>Totals</td>
<td>20 plants</td>
<td>32 plants</td>
<td>52 plants</td>
</tr>
</tbody>
</table>

Isolation of a genomic clone corresponding to the thrum-specific polymorphisms

Following the identification of SCAR and RFLP markers for the thrum allele of the S-locus, the characterization of the genomic region surrounding these markers was undertaken in a search for associated genes or open reading frames. The 323 bp probe was used to screen a *P. vulgaris* var. Blue Jeans phage λ genomic DNA library (Pavlov, 1997). Hybridizing clones were subsequently screened by PCR using the SCAR primers (data not shown), in order to identify which ones were derived from the thrum allele of the S-locus. From one such positive clone with an insert of around 20 kb, an approximately 9 kb *BamHI* fragment was identified which hybridized strongly to the probe. Furthermore, *KpnI* digestion of this fragment produced hybridization products of 1.4 kb and 2.6 kb corresponding to the thrum-specific RFLP bands. The sequence overlapping restriction fragments of this approximately 9 kb *BamHI* genomic clone was determined and these were assembled into a single contig of 8.8 kb. In silico predictive restriction analysis of this contig produced the same fragment sizes as restriction mapping of the 8.8 kb region (data not shown) confirming the correct assembly of the sequence contig. This sequence has been submitted to Genbank (Accession number AY854261)

Based on the sequence of the 8.8 kb fragment, primer pairs (P3+P4 and P5+P6; Fig. 5B) corresponding to either end of the genomic fragment were designed. These primers were used for PCR analysis of pin and thrum genomic DNA. Amplification with P3 and P4 only produced an amplification product with thrum, and not pin genomic DNA (data not shown). However, amplification with P5 and P6 amplified products from both pin and thrum DNA samples. BLAST analysis (Altschul et al., 1990) of the 323 bp sequence against the 8.8 kb fragment demonstrated a near-perfect match within the 1.4 kb *KpnI* fragment of the genomic clone; the only sequence differences being two bases within one of the 10-mer primer binding sites, and a one base mismatch within the other. In addition, both the 2.6 kb *KpnI* fragment and the *BamHI-KpnI* fragment at the end of the genomic clone also contained truncated versions of the 323 bp sequence. These observations demonstrate that this genomic clone contains the 2.6 and 1.4 kb thrum-specific RFLP bands observed in genomic Southern analysis and that these fragments contain duplications of the 323 bp marker sequences within tandem a duplication of an approximately 2 kb region (Fig. 5A, B).

Blast searches using the 8.8 kb sequence against nucleotide and protein sequence databases revealed only one region with any significant homology with other known sequences. Approximately 1 kb of sequence at one end of the genomic clone, which includes the P5 and P6 primer sequences, showed homology to regions of Ty3/gypsy-like retrotransposons. However, the predicted coding sequence of the retroviral polyprotein contains a number of deletions and point mutations. Similarly, gene structure prediction analysis of the entire 8.8 kb region indicated that the degenerated retrotransposon sequence represents the only similarity to a predicted gene within this fragment. The remainder of the genomic sequence comprises multiple short, repetitive elements with 2–5 copies of six different repeats of 50–104 bp and a single palindrome of 55 bp (Fig. 5A, B).

Analysis of S-locus recombinant homostyles

In a parallel project the development of a long homostyle and a short homostyle in the wild primrose genetic background has been characterized (MA Webster and PM Gilmartin, unpublished results). These homostyles
Fig. 5. (continued)
represent valuable tools for determining the position of the thrum-specific polymorphisms relative to the three loci that control gynoecium height, pollen size, and anther height. The long homostyle is true breeding with predicted genotype gPA/gPA, the predicted genotype of the short homostyle is Gpa/gpa. These plants were screened using the P1 and P2 SCAR primers in comparisons with genomic DNA from pin and thrum Primula var. Blue Jeans flowers (Fig. 6). Using these primers with genomic DNA from the long homostyle, a single PCR product was obtained that was the same size as the 323 bp product obtained from thrum DNA. No amplification products were obtained using these primers in combination with DNA extracted from the short homostyle. PCR amplification with primers derived from the Primula vulgaris Plena gene, which is not linked to the S-locus (Cook, 2002), yielded products from all four DNA samples (Fig. 6). These data suggest a possible position of the RAPD derived polymorphic marker relative to the previously defined genes that comprise S-locus.

Discussion

The phenomenon of floral heteromorphy in Primula has attracted considerable attention over the past 150 years, yet despite this interest, nothing is known about the basis of heteromorphic flower development, the nature of the genes responsible, their organization within the heteromorphic flower development, the nature of the genes despite this interest, nothing is known about the basis of The phenomenon of floral heteromorphy in

Discussion

Fig. 6. Analysis of long and short homostyle plants for the presence of PCR and RFLP markers. (A) Genomic DNA from normal thrum and pin plants and long and short homostyle plants was used in a PCR with the P1 and P2 primers and a control set of primers amplifying a product from a MADS box gene unlinked to the S-locus. Products were resolved by agarose gel electrophoresis.

Analyses were undertaken using an F1 hybrid commercial Primula line derived from inbred parental lines in order to minimize any polymorphic differences between individuals. Based on the assumption that any polymorphisms associated with the dominant S-allele present in heterozygous (S/s) thrum plants would be absent from homozygous recessive (s/s) pin plants, RAPD PCR was undertaken using 80 different 10 base Operon RAPD primers and two putative thrum-specific amplification products were identified with the W6 and Z9 primers (Fig. 1). Analysis of the 323 bp W6-derived product has shown that it co-segregates uniquely with the thrum phenotype, a finding which suggests that it is linked to the S-locus. The observation that a 1 kb polymorphic product identified using the Z9 primer (Fig. 1) does not co-segregate uniquely with the thrum phenotype suggests either that this polymorphism flanks the S-locus and is able to recombine, or that it represents an unlinked polymorphism originating from one of the two parental lines.

The analyses of available long and short homostyle plants have enabled the orientation of this marker with respect to the classically defined genes G, P, and A (Ernst, 1955; Dowrick, 1956; Wedderburn and Richards, 1992; Lewis and Jones, 1993; Richards, 1997). The presence of the 323 bp thrum-specific polymorphic amplification product within the true-breeding long homostyle (gPA/gPA) and its absence from the short homostyles (with predicted genotype Gpa/gpa) suggest that this thrum-specific fragment is positioned on the A side of the S-locus. Orientation of the thrum-specific DNA polymorphic marker with a specific side of the S-locus, provides additional valuable data for the molecular genetic dissection of the Primula S-locus.

Sequence analysis of the thrum-specific RAPD product indicated no similarity to available DNA databases, and expression analyses did not reveal any transcripts derived from this sequence. It was therefore concluded that this sequence does not represent an S-locus associated gene. However, these data made it possible to design primers with increased specificity by the addition of five bases at the 3' end of the W6 RAPD primer sequence to create the 15-base SCAR primers, P1 and P2, with predicted annealing temperatures of 50 °C and 48 °C, respectively. Three lines of evidence indicate that sequences contained within the RAPD fragment are not unique to the thrum genome: (i) Southern analysis of RAPD fragments generated with the original Operon W6 primer using the 323 bp thrum-specific fragment as a probe reveals homologous amplification products from both pin and thrum genomic DNA (Fig. 2A); (ii) PCR amplification with the 15 base SCAR primers, P1 and P2, at 54 °C generated amplification products from both pin and thrum genomic DNA (Fig. 2C); (iii) Southern analysis of pin and thrum genomic DNA using the thrum-specific amplification product reveals the presence of related restriction fragments within both thrum and pin genomes (Fig. 3). However, the thrum-specific nature of the 323 bp RAPD amplification product (Fig. 2A), the unique thrum-specific amplification product obtained using the P1 and P2 SCAR primers for PCR analysis at 58 °C, and the presence of 1.4 and 2.6 kb RFLP hybridization products uniquely in thrum DNA suggest that although similar sequences are present in both pin and thrum plants, their genomic environment is different, possibly due to an additional copy or copies of this sequence in
the thrum genome. Further analysis of additional F₁ plants as well as F₂ progeny derived from these plants demonstrated the complete association of the 323 bp amplification product with the thrum phenotype. Of 191 plants analysed by PCR and 52 analysed by Southern analysis no recombinants were identified (Table 1). It was therefore concluded that this sequence is tightly linked to the thrum allele of the *Primula S*-locus.

The use of RAPD analysis to screen for *S*-linked polymorphisms relied on the unpredictable presence of appropriate primer binding sites. These analyses used 80 random 10 base primers, each produced approximately 20 amplification products with an average length of 1 kb, giving a total length of 1.6 Mb. The genome of *Primula vulgaris* is approximately 480 Mb (Bennett and Smith, 1976), these analyses have therefore sampled less than 1% of the genome, suggesting that further such analyses could reveal additional thrum-specific polymorphic markers.

Previous analyses on recombination frequencies within the *S*-locus suggest either, that the different genes which comprise the locus are very close to each other or that mechanisms are in place to suppress recombination. Studies of other self-incompatibility loci have revealed that there is not a direct relationship between genetic and physical map distances. In Brassica, recombination suppression over large distances maintains tight linkage of *S*-locus genes for some *S*-haplotypes while their close proximity, within 13 kb, is responsible for preserving linkage of other haplotypes (Casselman et al., 2000). The availability of this first *S*-locus probe from *Primula* provides a unique tool to address the maintenance of *S*-locus integrity in this plant directly.

Due to the heterozygous nature of thrum plants, the thrum genomic DNA library contains both pin and thrum alleles of the *S*-locus and sequences corresponding to the 323 bp marker are present in both pin and thrum plants. The thrum-specific SCAR primers made it possible to identify those lambda genomic clones that contained the thrum-specific copy of the polymorphic sequence. Subsequent restriction analysis of one of these clones not only revealed that the 323 bp thrum-specific-amplification products resided on an 8.8 kb *BamHI* fragment, but that this same fragment contained the 1.4 kb and 2.6 kb thrum-specific *KpnI* fragments identified as thrum-specific RFLPs. This finding provided the first evidence that both the RAPD and RFLP polymorphism originate from the same region of the genome. These analyses using PCR primers designed to the extreme ends of the 8.8 kb *BamHI* fragment provided further evidence for the thrum-specific origin of this clone. Primers P3 and P4 only produced amplified products from thrum genomic DNA. The observation that primers P5 and P6 were not thrum-specific was explained by subsequent further sequence analysis that revealed the presence of sequences with similarity to the reverse transcriptase gene of a Gypsy-like retrotransposon (Kumar and Bennetzen, 1999). This sequence is degenerate and does not code for a functional protein and additional anticipated sequence similarities to other features of retrotransposons are absent, suggesting that this element has undergone considerable degeneration since its insertion close to the *S*-locus.

The presence of an approximately 2 kb tandem duplication that contains the 323 bp RAPD fragment, provided an explanation for an earlier observation that suggested the presence of more than one copy of this sequence at the *S*-locus. Earlier efforts to isolate sequences flanking the 323 bp sequence included inverse PCR using outward facing primers with self-ligated restriction fragments of thrum genomic DNA as template. In these experiments, the unligated control also yielded a product that could be cut with *KpnI* (Pavlov, 1997). The availability of the full sequence of the 8.8 kb *BamHI* fragment provides an explanation that is consistent with all these observations and also provides an explanation for both the 2.6 kb and 1.4 kb thrum-specific RFLP bands as each contains a copy of the 323 bp sequence. The presence of an additional copy of the 323 bp sequence within the 2.2 kb *BamHI-KpnI* fragment derived from the lambda clone suggests that genomic Southern analysis should have revealed an additional thrum-specific hybridization band. The absence of such a thrum-specific band suggests that this copy of the sequence is also located at the pin allele of the *S*-locus.

The presence of multiple short repeat elements, a stem loop structure, and tandem duplication within this region are all reminiscent of features previously identified as associated with the *S*-locus in other species (Coleman and Kao, 1992; Matton et al., 1995; Ushijima et al., 1998; Wheeler et al., 2003). Nothing is known about the origins or relation of the *Primula S*-locus with respect to homomorphic self-incompatibility (SI) systems. However, a common feature of all SI systems is the tight linkage of male and female incompatibility determinants which, in some cases, are associated with the presence of mechanisms to suppress recombination within the region of the *S*-locus (Cui et al., 1999; Casselman et al., 2000; Shiba et al., 2003; Wheeler et al., 2003). These findings provide the first evidence for the presence of sequence structures associated with the *Primula S*-locus that are also found within the *S*-loci of other unrelated species. Further analyses will be required to determine whether suppression of recombination over large physical distances, or the close proximity of the *S* locus genes, is responsible for maintaining the structure of this locus in *Primula*.

The presence of an *S*-linked degenerate retrotransposon in *Primula* raises the intriguing question of when this element was inserted into the *S*-locus. Studies are now possible to determine whether the sequences associated with the thrum allele of the *S*-locus in *Primula vulgaris* are also *S*-linked in other species within the *Primula* genus. Ninety per cent of *Primula* sp. demonstrate heteromorphy (Wedderburn and Richards, 1992; Richards, 1993; Mast et al., 2001) and future studies will help to determine whether the sequences that have been identified have evolved within the *S*-locus in
Primula vulgaris or whether they pre-date speciation and are ancestral S-locus-associated features.

Acknowledgements

Financial support for IWM and JL was provided by the BBSRC and HEC was supported by a PhD studentship from the BBSRC. We would also like to thank Margaret Webster for the long and short homostyle plants, Martin Lappage and Helen Martin for care of the Primula plants at the University of Leeds Experimental Gardens, and Denise Ashworth for DNA sequencing. We would also like to thank the Farmen Seed company for providing plant material and seeds during the course of this work.

References


Darwin C. 1862. On the two forms, or dimorphic condition, in the species of Primula, and on their remarkable sexual relationships. Proceedings of the Linnean Society 6, 105–139.

Darwin C. 1877. The different forms of flowers on plants of the same species. London: John Murray.


Webster MA, Grant CJ. 1990. The inheritance of calyx morph variants in Primula vulgaris (Huds.). Heredity 64, 121–124.

