Theor Appl Genet (2010) 121:567–576
DOI 10.1007/s00122-010-1330-9

ORIGINAL PAPER

The genetic location of the self-incompatibility locus in white clover (*Trifolium repens* L.)

Nora M. Casey · Dan Milbourne · Susanne Barth · Melanie Febrer · Glyn Jenkins · Michael T. Abberton · Charlotte Jones · Daniel Thorogood

Received: 27 January 2010 / Accepted: 26 March 2010 / Published online: 10 April 2010

Abstract White clover (*Trifolium repens* L.) is a forage legume of considerable economic importance in temperate agricultural systems. It has a strong self-incompatibility system. The molecular basis of self-incompatibility in *T. repens* is unknown, but it is under the control of a single locus, which is expressed gametophytically. To locate the self-incompatibility locus (*S* locus) in *T. repens*, we carried out cross-pollination experiments in an F_1 mapping population and constructed a genetic linkage map using amplified fragment length polymorphism and simple sequence repeat markers. As the first step in a map-based cloning strategy, we locate for the first time the *S* locus in *T. repens* on a genetic linkage map, on the homoeologous linkage group pair 1 (E), which is broadly syntenic to *Medicago truncatula* L. chromosome 1. On the basis of this syntenic relationship, the possibility that the *S* locus may or may not possess an *S-RNase* gene is discussed.

Communicated by Y. Xue.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-010-1330-9) contains supplementary material, which is available to authorized users.

N. M. Casey · D. Milbourne · S. Barth · M. Febrer
Teagasc Crops Research Centre, Oak Park,
County Carlow, Ireland

N. M. Casey · M. Febrer · G. Jenkins · M. T. Abberton · C. Jones · D. Thorogood (✉)
Institute of Biological, Environmental and Rural Sciences (IBERS), Aberystwyth University, Penglais, Aberystwyth, Ceredigion, Wales SY23 3DA, UK
e-mail: dnt@aber.ac.uk

Present Address:
M. Febrer
John Innes Centre, Norwich Research Park, Colney,
Norwich NR4 7UH, UK

Introduction

White clover (*Trifolium repens* L.) is a herbaceous perennial legume and is important as a pasture species in temperate regions, where it is generally grown in conjunction with a companion grass (Frame et al. 1998; Abberton and Marshall 2005).

T. repens is an allotetraploid (2n = 4x = 32) with a 1C genome size of 956 Mbp (Atwood and Hill 1940; Williams et al. 1982; Bennett and Leitch 2003). Chromosome pairing is bivalent and inheritance is disomic (Atwood and Hill 1940; Williams et al. 1982). The karyotype formula is eight metacentric plus eight median centromeres (Salimpour et al. 2008). Allopolyploidy in this species is thought to have arisen from the hybridisation of two ancestral diploid species and subsequent chromosome doubling. The parent species of *T. repens* are posited to be *Trifolium occidentale* Coombe and *Trifolium pallescens* Schreb. (Ellison et al. 2006), although it has also been suggested that only one of these (*T. occidentale*) can be a direct ancestor (Hand et al. 2008).

T. repens is obligately outcrossing due to its strong self-incompatibility (SI) system, by which means a plant is prevented from being fertilised by its own pollen, and in some cases by that of a close relative (Wright 1939; Atwood 1940). SI has been reported in about 60% of all angiosperms (Allen and Hiscock 2008), and it has been reviewed recently by several authors (Silva and Goring 2001; Takayama and Isogai 2005; Rea and Nasrallah 2008). Of the two hypothetical ancestors of *T. repens*, *T. occidentale* is self-compatible (Williams et al. 2008),

whilst *T. pallescens* is self-incompatible (Abberton 2007). Another close relative of *T. repens*, *Trifolium nigrescens* Viv., also proposed as a parent species (Badr et al. 2002), is also self-incompatible (Brewbaker 1953).

A self-incompatibility locus contains a male-expressed gene and at least one female-expressed gene, which are tightly linked and control SI specificity (Ushijima et al. 2003; Ikeda et al. 2005). SI can be controlled by one or more such loci. White clover is a single-locus system (Atwood 1942). In single-locus systems, the self-incompatibility locus is called the *S* locus. In *T. repens*, the male SI gene is expressed gametophytically (gametophytic SI = GSI) (Atwood 1940).

Two unrelated mechanisms of homomorphic single-locus-controlled GSI are known to exist: one in Papaveraceae and one in Solanaceae, Rosaceae and Plantaginaceae (*S*-RNase based) (Steinbachs and Holsinger 2002). The *S*-RNase-based system, which has a common origin in all three families studied (Igic and Kohn 2001; Steinbachs and Holsinger 2002; Vieira et al. 2008), is most like that of *T. repens* based on current evidence, in that pollen tube arrest takes place late in the pollination process about two-thirds of the way down the style (Williams 1931; Silow 1931; Atwood 1941; Leduc et al. 1990; Xue et al. 1996). The SI reaction takes place on the stigma in Papaveraceae (Lawrence 1975). The families in which the *S*-RNase-based system is known are also more closely related to Leguminosae than is Papaveraceae (Magallón et al. 1999).

To date, very little is known about SI in legumes. Studies conducted on *T. repens* and its close relatives provide strong evidence for gametophytic expression encoded by a single locus (Arroyo 1981) on the basis of cross-pollination experiments (Atwood 1940), pollen and stigma characteristics (Brewbaker 1957; Heslop-Harrison and Shivanna 1977) and competitive interaction in chromosome-doubled *T. repens* where self-compatibility is induced (Brewbaker 1954). There is also evidence that the incompatible response takes place in the style, although the stigma may be involved in the case of *T. repens* (Williams 1931; Silow 1931; Atwood 1941; Leduc et al. 1990). There is some evidence for *S*-RNase-based SI in *Trifolium*. Incompatible pollen tubes in styles of *Trifolium pratense* L. appear to grow in a similar way to incompatible pollen tubes in *S*-RNase reactions (Silow 1931; East 1934). A glycoprotein found specifically in stylar extracts of *T. pratense* was of a similar size (about 24 kDa) to *S*-RNase proteins from Solanaceae and Rosaceae; however, this has not been further characterised (Heslop-Harrison and Heslop-Harrison 1982; Anderson et al. 1986; Sassa et al. 1993).

S loci are represented by a poly-allelic series maintained in natural outcrossing populations by frequency-dependent selection (reviewed in Schierup and Vekemans 2008). Because the *S* locus is a complex of several genes inherited as one unit, *S* alleles are known as haplotypes (Takayama and Isogai, 2005). *T. repens* and its close relative *T. pratense* have a higher estimated number of *S* haplotypes than other species investigated, based on analysis of data from cross-pollination studies (Atwood 1942, 1944; Lawrence 1996). For *T. repens*, the estimates range from 74 to 139, compared to figures in the mid-40s for several species of Solanaceae (Richman et al. 1996; Lu 2006). For *T. pratense*, Lawrence's estimates were even higher (Williams and Williams 1947; Lawrence 1996).

Three genetic linkage maps of *T. repens* based on SSRs have been published in recent years (Jones et al. 2003, Barrett et al. 2004, Zhang et al. 2007). The first map (Jones et al., 2003) referred to each linkage group (LG) using a number from 1 to 18; Barrett et al. (2004) introduced a method of calling each homoeologue group (HG) by the letters A–H and identified homoeologues by shared marker loci. Zhang et al. (2007) adopted this system, but inverted HGs E and F relative to Barrett et al.'s orientation to reflect synteny with *Medicago sativa* L. and *T. pratense*. By using some markers from Jones et al. (2003), the map of Zhang et al. enabled some of Jones et al.'s LGs to be identified tentatively with HGs from Barrett et al. (2004)'s system.

Recently, George et al. (2008) have proposed a change in the order and nomenclature of *T. repens* HGs to reflect synteny with the model legume *Medicago truncatula* L. This agrees with the synteny that Zhang et al. (2007) identified between *T. repens* HGs and *M. sativa* LGs (Sledge et al., 2005). The orientation of the HGs can also be fixed with regard to *M. truncatula*: HG E best matches the *M. truncatula* alignment when it is in Barrett et al. (2004) orientation; HGs A, B, C and D are inverted relative to Barrett et al.; there is too little information as yet regarding HGs F, G and H to tell which orientation should be used, so the orientation of Barrett et al. has been conserved for now (George et al. 2008).

In this study, the *S* locus in *T. repens* was located on a genetic linkage map for the first time, using a combination of AFLP and SSR markers. The single-parent genetic linkage maps which were developed were compared to previously published linkage maps for *T. repens* in terms of the position of shared markers.

Materials and methods

Plant material

An F_1 mapping population consisting of 94 progeny individuals was created by crossing two highly divergent white clover genotypes referred to as S1S4 (female parent) and R3R4 (male parent). S1S4 (in this case S for susceptible,

not for self-incompatible) had been identified as being susceptible to the stem nematode *Ditylenchus dipsaci* Kuhn; the male parent, R3R4, as resistant to the same nematode.

Genomic DNA was extracted from the parents and the progeny individuals using a modified 2× CTAB extraction (Doyle and Doyle 1987). The modifications were that samples were larger (about 2 g); liquid nitrogen was used in grinding; extraction buffer did not contain sodium chloride and contained EDTA at 50 mM; incubation was at 65°C; wash solution was 70% ethanol; pellets were dried overnight; TE buffer volume was 500 µL and its pH was 8.0.

Pollination experiments

Plants in the mapping population, identified by numbers 1 through 94, were cross-pollinated by hand to determine incompatibility relationships within the population. A piece of card was cut to form a folded triangle at one end and used to 'trip' and collect pollen from unfertilised florets. It was roughened on the inside of the fold to collect pollen more readily. The card was used to transfer pollen from one flower to another as follows: the inside of the fold was rubbed against the anthers of one flower and then brushed against the stigma of another when it had collected sufficient pollen. In the majority of cases, petals and anthers were removed from the female pollen recipients primarily to allow observation of pollen clumps being deposited on styles ensuring that successful pollination had taken place. Pollen transferred could be seen with the naked eye. A new card was used for each pollination and plants were kept in a closed chamber throughout, so the chance of unintended pollen transfer was slight. Several flowers (ca. 5–30) were pollinated in each cross-pollination. Some crosses were made reciprocally. Temperature and daylength were constant (16 h; ca. 22°C). Cross-pollinations were scored as compatible or incompatible after ca. 7 days by checking the seed pods of pollinated flowers visually and by tactile examination to determine whether or not seeds were developing. In some cases, this was verified by dissection using a scalpel and a stereo microscope. An absence of seeds was taken to indicate a self-incompatible reaction resulting from the parental plants sharing the same *S* genotype. Presence of seeds indicated a compatible reaction, indicating that the parents possessed at least one different *S* haplotype.

AFLP procedure

A total of 30 AFLP primer combinations were applied to the mapping population. AFLP analysis was performed essentially as in Vos et al. (1995) except for the use of *Pst*I instead of *Eco*RI as the rare cutter in most of the assays (Paglia and Morgante 1998). The frequent cutter was always *Mse*I. PCR reactions were performed in a Peltier Thermocycler (PTC) 2000 (MJ Research, Massachusetts) or a Biometra TProfessional Basic Thermocycler (gradient model).

Selective amplification of pre-amplified DNA was carried out using a variety of selective primer combinations. The following three base extension E primers (E + AAC, ACA, AAG, ACG, TCG) were used in conjunction with the primer MCCA where *Eco*RI was used as the rare cutter. When *Pst*I was used as the rare cutter, all possible combinations of the following two base extended P primers and three base extended M primers were used: P + AC, CG, TA, AT, GC and M + AAC, AAG, AAT, AGA, ATC, CCA). The E or P primer in each reaction was end labelled with 1 µCi γ-[^{33}P]-ATP, in 1× T4 buffer using 2 U/µL of T4 polynucleotide kinase (NEB) according to the manufacturer's instructions. PCR reactions were performed as in Griffin et al. (2009).

To separate the labelled fragments, loading dye was added to each reaction product in an equal volume [loading dye contained 98% deionised formamide, 10 mM EDTA pH 8.0, bromophenol blue (1 mg/mL), xylene cyanol (1 mg/mL)]. Amplification products were electrophoresed as in Altinkut et al. (2003). Gels were dried on 3MM Chr filter paper (Whatman) and exposed to storage phosphor screens for 1–3 days at room temperature. The results were scored visually. Markers were named according to primer combination and order of occurrence on the gel.

Microsatellites (SSRs)

A total of 64 SSR markers were used in this study. Published SSR primer sequences were obtained from several sources: 6 TRSSR markers from Jones et al. (2003); 26 ats and prs markers from Barrett et al. (2004); 3 RCS and 1 BG markers from Zhang et al. (2007), originally from Sato et al. (2005) (RCS) and Sledge et al. (2005) (BG). Four further SSRs (TRagr) were derived from publicly available white clover sequences, which contained microsatellite motifs (Genbank: Benson et al. 2008). These sequences were identified using tandem repeats finder (Benson 1999). Primers were designed using Primer3 (Rozen and Skaletsky 2000) and tested subsequently for amplification (Barth et al. 2004). A total of 24 microsatellites (16 WCBE, 8 C27 and C28) were derived from work on a white clover BAC library (Febrer et al. 2007). The markers were analysed using either radioactive labelling (SSRs from Barrett et al. 2004 except ats072; BAC-derived SSRs) or fluorescent labelling (all other SSRs), as described by Febrer et al. (2007).

Leaf marking

In addition to the molecular markers and SI phenotypes described above, the white "V"-shaped leaf mark also segregated in the population and was scored visually as a dominant marker (Williams 1987). This morphological marker was present on the S1S4 parent and absent on the R3R4 parent.

Linkage analysis and mapping

Markers with alleles segregating in one or both parents were analysed in the entire population of 94 F_1 progeny individuals. Both the AFLP markers and individual alleles of the SSR markers were scored as dominant. Individual parental maps were created using JoinMap 3.0 (Van Ooijen and Voorrips 2001). Construction of the linkage map was accomplished by treating the segregating data as a cross-pollinator. Genetic map distances were estimated using Kosambi's mapping function. For construction of single-parent maps, locus grouping and ordering probability minima were set to an LOD of 3. Linkage groups were drawn using MapChart 2.2 (Voorrips 2002).

Results

S *locus* genotyping

The relatively divergent nature of the parents meant that the population also exhibited variation in flowering time. As a result, different combinations of plants were present at different times over a 3-month period, and this factor determined which plants were intercrossed to determine the intra-incompatible groups in the population. By crossing plants strategically as they became available, data about the *S* genotypes were gathered for the entire population. For example, plant 77 was incompatible with plants 16 and 23, which demonstrated that they had the same incompatibility genotype. However, plant 77 was compatible with plants 33, 62, 70 and 81. Of these, 81, 62 and 33 were incompatible with one another, showing that they expressed a single incompatibility genotype, whilst plant 70 was compatible with plant 33, showing that 70 expressed a further incompatibility genotype. A full summary of the crosses made to determine the intra-incompatible groups is shown as electronic supplementary material (Supplementary Table 1).

Four plants, 7, 10, 20 and 56, were found to have low fertility relative to other genotypes. In crosses involving these plants, negative (incompatible) results were obtained with plants that had already been shown to belong to different incompatibility groupings. However, compatible crossings were identified with at least one plant in three of the groupings, but not in the fourth, and we therefore allocated the plant to this fourth wholly intra-incompatible grouping. In any case, a linkage analysis excluding these plants did not change the map position of the *S* locus (results not shown).

All of the progeny individuals were genotyped, and four different intra-incompatible groups (1–4) were clearly identified in the population. The numbers of plants per intra-incompatible group were 24, 24, 17 and 29. Plant frequencies in each intra-incompatible group were not significantly different from a 1:1:1:1 ratio (df 3; χ^2 3.11; $P = 0.376$), indicating that a single *S* locus with four different *S* haplotypes is segregating in the population with haplotypes from both parents being transmitted in equal proportions. *S* incompatibility genotypes were allocated to each group as follows based on the assumption that the male parental genotype at the *S* locus was S_1S_2 and that of the female parent was S_3S_4. Early on in the crossing programme, a number of AFLP markers that produced an electrophoretic gel band in the male parent, but not in the female (markers used to generate the male parental map), were found to associate positively either with intra-incompatible groups 1 and 2 or with intra-incompatible groups 3 and 4 (Table 1a). In markers for which there was generally a band in plants from groups 1 and 2, there was generally no band in groups 3 and 4, and vice versa. The intra-incompatible groups were therefore allocated haplotypes S_1 (groups 1 and 2) and S_2 (groups 3 and 4), respectively, derived from the male parent. Markers that were present in the female parent but not in the male (markers used to generate the female parental map) were similarly found to associate with groups 1 and 3 or with groups 2 and 4 (Table 1b). These groups were therefore allocated haplotypes S_3 (groups 1 and 3) and S_4 (groups 2 and 4), respectively. By combining the data from both parental maps, it was therefore possible to allocate the genotypes S_1S_3 (group 1), S_1S_4 (group 2), S_2S_3 (group 3) and S_2S_4 (group 4). Alleles derived from the male parent are transmitted in equal proportions: group 1 and 2 individuals versus group 3 and 4 individuals = 48:46 (df 1, χ^2 0.04, $P = 0.837$), as are alleles derived from the female parent: group 1 and 3 individuals versus group 2 and 4 individuals = 41:53 (df 1, χ^2 1.53, $P = 0.216$).

Construction of the genetic linkage map

To identify the genetic location of the *S* locus, a genetic linkage map of each parent was constructed using 64 white clover SSRs and 339 AFLP markers. The S1S4 map comprised 17 distinct linkage groups, spanning 1,477 centimorgans (cM) and ranging in size from 44 to 145 cM. This map contained alleles of 48 SSR markers and 196

Table 1 Association of four AFLP markers (PatMaat7, PacMaac10, PgaMaat8, PgaMaac9) with the four intra-incompatible groups

(a) Markers with band present in male parent																
Group 1 (S_1S_3)																
Plant no.	9	12	14	16	21	23	24	36	40	42	51	55	56	63	77	92
PatMaat7	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−
PacMaac10	+	+	+	+	+	+	+	+	+	+	+	+	+	−	+	+
Group 2 (S_1S_4)																
Plant no.	10	14	30	45	50	58	70	71	73	79	80	84	86	93		
PatMaat7	−	−	−	−	−	−	−	−	−	−	−	−	−	−		
PacMaac10	+	+	+	+	+	+	+	+	+	+	+	+	+	+		
Group 3 (S_2S_3)																
Plant no.	1	8	11	20	25	28	32	39	46	48	54	66	85			
PatMaat7	+	+	+	+	+	+	+	+	+	+	+	+	+			
PacMaac10	−	−	−	−	−	−	−	−	−	−	−	−	−			
Group 4 (S_2S_4)																
Plant no.	7	26	33	37	41	44	47	49	53	57	61	62	65	76	81	94
PatMaat7	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+
PacMaac10	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−
(b) Markers with band present in female parent																
Group 1 (S_1S_3)																
Plant no.	9	12	14	16	21	23	24	36	40	42	51	55	56	63	77	92
PgaMaat8	−	−	−	+	+	−	−	−	−	−	−	−	−	+	+	−
PgaMaac9	−	−	−	+	−	−	−	−	−	−	−	−	−	+	+	−
Group 2 (S_1S_4)																
Plant no.	10	14	30	45	50	58	70	71	73	79	80	84	86	93		
PgaMaat8	−	+	+	+	+	+	+	+	+	*	+	+	+	+		
PgaMaac9	+	+	+	+	+	+	+	+	−	+	−	+	+	+		
Group 3 (S_2S_3)																
Plant no.	1	8	11	20	25	28	32	39	46	48	54	66	85			
PgaMaat8	−	−	−	−	−	−	+	−	−	−	−	−	−			
PgaMaac9	+	−	−	−	−	−	−	−	−	−	−	−	−			
Group 4 (S_2S_4)																
Plant no.	7	26	33	37	41	44	47	49	53	57	61	62	65	76	81	94
PgaMaat8	+	+	+	+	−	+	+	+	−	+	+	+	−	+	−	+
PgaMaac9	+	−	+	+	−	+	+	+	−	+	−	+	+	+	−	+

See text for explanation of allocation of *S* locus genotypes to intra-incompatible groups

+ band present on gel, − band absent on gel, * unclassified

AFLP markers, resulting in an average map density of 5.7 markers per cM. A single morphological marker (the white V-shaped leaf mark) was also placed on the female parental map at a distal region of homoeologue group 5 (G) (see Supplementary Figure 1). Approximately, 24% of the markers on the S1S4 map exhibited segregation distortion (declared at $P < 0.05$). The R3R4 map comprised 14 linkage groups spanning 1135 cM, ranging in size from 37 to 128 cM. This map contained alleles of 43 SSR markers and 146 AFLP markers with an average map density of 5.7 markers per cM. Approximately 14% of markers exhibited segregation distortion on the paternal map; 23 marker loci, 22 SSRs and one AFLP, were shared between the two parental maps.

The presence of SSR markers from previously published maps of white clover on the maps of S1S4 and R3R4 allowed the identification of linkage groups (LGs) belonging to the eight homoeologue groups (HGs) of white clover. For the S1S4 parental map, it was possible to assign 15 of the 17 LGs to one of the eight white clover HGs A-H (as per the nomenclature of Barrett et al. 2004) or 1–8 (George et al. 2008), each HG being represented by two LGs with the exception of HG 2 (F) that was represented by a single LG. Two further LGs remained unassigned

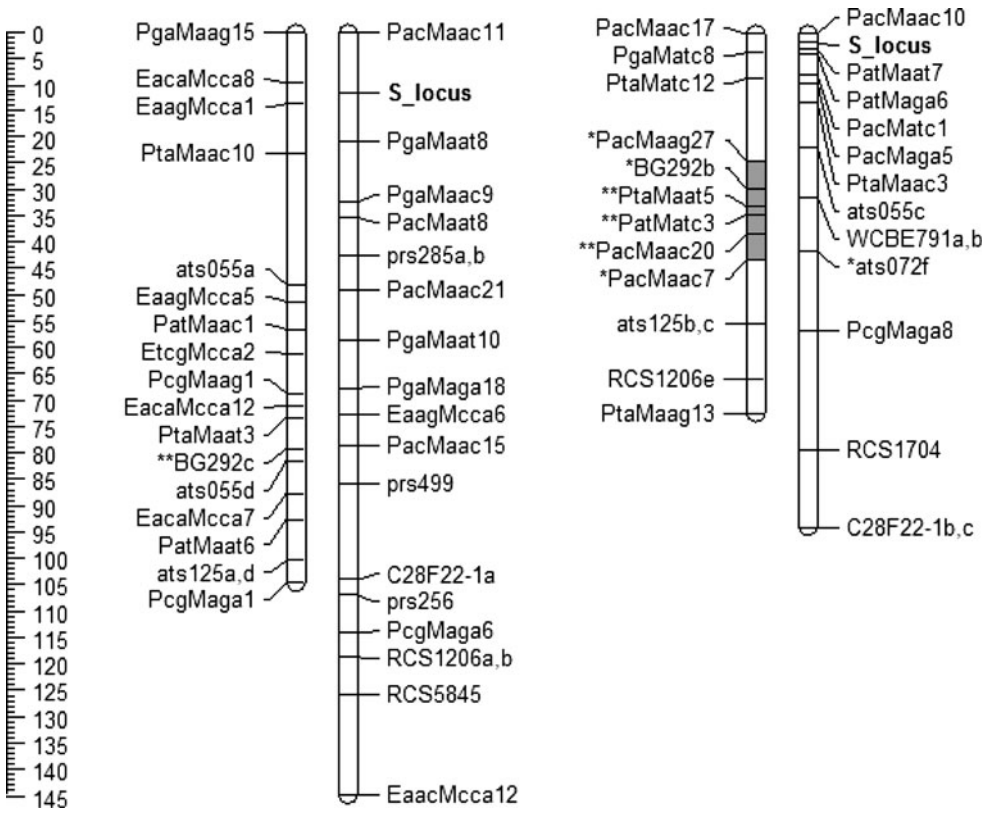

Fig. 1 Linkage maps of homologue 1 (E), S1S4 (*left*) and R3R4 (*right*). Scale (at *left*) is in centimorgans (cM). *Asterisks* segregation distortion (declared at $P < 0.05$). *Grey shading* area is defined by distorted markers. The *S* locus is shown in *bold*

because they consisted entirely of AFLPs. For the R3R4 parental map, 13 LGs were assigned to HGs, with HGs 5 (G), 6 (H) and 8 (B) being represented by a single identifiable linkage group. One further LG, consisting entirely of AFLPs, remained unassigned.

The parental maps are presented as supplementary material (Supplementary Figure 1).

The genetic location of the *S* locus in white clover

The *S* locus was mapped in both parental maps to individual linkage group members of HG 1 (E) (Fig. 1). When an interim round of analysis revealed HG 1 (E) was the likely genetic location for the *S* locus, a concerted effort was made to include a reasonable number of HG 1-specific SSRs in the largely AFLP-based map to order and orient these groups comprehensively. As a result, alleles of 9 HG 1 (E)-specific SSR markers in total have been used to identify and orient the individual parental LGs constituting HG 1. The orientation of the linkage maps shown in Fig. 1 (in which the *S* locus is mapped to the distal end of HG 1) is presented as in Barrett et al. (2004).

Discussion

This study provides molecular evidence for the first time that the control of GSI in *T. repens* involves a single functioning *S* locus. This result is in agreement with previous data from cross-pollination studies showing a pattern of cross-incompatibility consistent with a single locus under gametophytic control (Atwood 1940).

At the outset of the experiment, a number of assumptions about the probable nature of the inheritance of self-incompatibility in the population were made on the basis of the available information. First, despite the allotetraploid nature of white clover, on the basis of previous work (Atwood 1942; Brewbaker 1954) it was assumed that the SI response in white clover would be governed by a single *S* locus located on only one of the two homoeologous genomes of the species. In this case, it was assumed that the *S* locus would exhibit simple disomic inheritance, with a maximum of two different haplotypes per parent. Furthermore, the nature of single-locus incompatibility systems precludes the possibility of homozygosity at the *S* locus, so each parent would have to be heterozygous (containing two different *S* haplotypes). Also, given the relatively genetically divergent nature of the parents and the fact that *T. repens* exhibits a very high number of *S* haplotypes (Lawrence 1996), it was considered likely that the parents would exhibit divergent *S* genotypes, resulting in the segregation of four different *S* haplotypes in the population (parental *S* genotypes S_1S_2 and S_3S_4). All of these assumptions were supported by the results obtained during the experiment.

Reading the linkage map

We have followed the HG order and alignment suggested by George et al. (2008). However, since the nomenclature of George et al. (2008) is different from any of the previous linkage maps of *T. repens*, we have retained the A–H nomenclature alongside the numerical nomenclature to aid comparison between maps. Because the SSRs used for HG identification and orientation are not ancestral-genome specific, we have not employed suffix numerals to label the linkage groups within an HG. However, as much as possible, we have retained the LG order within HGs of previous maps (Barrett et al. 2004; Zhang et al. 2007). Suffix numerals are likely to cause confusion if numbers are used to identify HGs. The work of Hand et al. (2008) suggests replacing this with parental genotype designators (proposed as O and P'); however, the data required for this are not yet available for the majority of *T. repens* HGs.

Self-incompatibility and polyploidy

The presence of only one functional *S* locus in an allotetraploid raises the question of how a single-locus system became established after the initial hybridisation event. SI is theoretically considered to be difficult to evolve (or re-evolve) from self-compatibility (Steinbachs and Holsinger 1999). Two functioning *S* loci would be likely to cause self-compatibility by competitive interaction (Lewis 1947), as has been shown in artificially chromosome-doubled *T. repens* (Brewbaker 1954). This suggests that, of the two parent species, one was self-incompatible whilst the other was self-compatible. This would be the situation with the posited parent species *T. occidentale* (SC) and either *T. pallescens* or *T. nigrescens* (both SI) (Brewbaker 1953, Badr et al. 2002; Ellison et al. 2006; Abberton 2007; Williams et al. 2008). However, rapid epigenetic changes causing gene silencing within a generation of polyploidisation have been reported in wheat (*Triticum*) (Feldman and Levy 2005). If these changes took place as rapidly in *T. repens*, self-compatibility due to competitive interaction might be avoidable even if both parents exhibited SI. Seeking insights into the parental genomes of *T. repens* should be done cautiously.

In other allopolyploids such as common wheat (*Triticum aestivum* L. = 6x), ancestral sub-genomes are relatively easily identified by virtue of sub-genome specific markers. However, the SSR markers used in this study are not capable of distinguishing between the proposed ancestral genomes of white clover; a progenitor comparison approach has been used successfully in single-nucleotide polymorphism (SNP) development in *T. repens* (Hand et al. 2008) and subsequent linkage maps in this species may be able to use SNPs to assign homoeologues to parental genotypes. Based on the current data, we cannot comment on the ancestral source of SI in this species. Ultimately, sequencing of the *S*-locus region and the equivalent region in the homoeologous chromosome would provide useful insight into how a single-locus system may have evolved after the initial hybridisation event.

The physical location of the *S* locus

The location of the *S* locus near one end of a homoeologue in *T. repens*, together with the fact that the centromeres in this species are not near chromosome ends (Salimpour et al. 2008), suggests that the *S* locus in *T. repens* may not be centromeric. *S*-RNase-based *S* loci are located near the centromere in Solanaceae (Entani et al. 1999; Yang et al. 2007), but not in *Antirrhinum* (Plantaginaceae) (Yang et al. 2007), and believed not to be in Rosaceae (Ushijima et al. 2001). We are not yet in a position to locate the *T. repens S* locus physically.

Synteny with *Medicago truncatula* and *Trifolium pratense*

M. truncatula has been shown to exhibit high levels of conserved macrosynteny with the genome of *T. repens* (George et al. 2008). Conserved synteny can allow translational genomic approaches for the discovery of genes in target species via better characterised model species. Given that *M. truncatula* is self-compatible, it would not be surprising if it had lost an *S* locus (originally present in an ancestral species) or components of it. However, other self-compatible relatives of SI species [e.g. *Arabidopsis thaliana* (L.) Heynh.] have been found to have non-functional, but recognisable copies of *S* locus genes (Kusaba et al. 2001). HG 1 (E), to which the *S* locus mapped in this study, is largely syntenic to *M. truncatula* chromosome 1. However, data from the study of George et al. (2008) on synteny between *M. truncatula* and *T. repens* suggest that there is a breakdown in synteny between HG 1 and chromosome 1 at the end of HG 1 that includes the *S* locus. Perhaps, unsurprisingly given this breakdown in synteny, BLAST-based comparisons of Rosaceae and Solanaceae *S*-RNase genes and proteins to the *M. truncatula* genome did not reveal any sequences on *M. truncatula* chromosome 1 exhibiting significant similarity to *S-RNases* (data not shown). Extending the above analysis to the rest of the *M. truncatula* genome revealed numerous *T2 RNase*-like sequences in the *M. truncatula* genome, some of which were found located near *F-box* genes (as might be expected for this type of locus), but there was no evidence to suggest that any of these sequences might have been derived from an *S-RNase*. Data on the synteny between *T. repens* and *T. pratense* given in Zhang et al. (2007) allow a prediction

to be made about the location of the *S* locus in *T. pratense* based on what is now known in *T. repens*. The most likely location is *T. pratense* chromosome 1 (Heathcliffe Riday, personal communication). Chromosome 7 also has some synteny to HG 1, but current data show this is mainly to the other end of the homologue.

Further work

This study is the first to assign the gametophytic *S* locus in *T. repens* to a single homoeologue group, 1 (E). Using the comparative map information between *T. repens* and *T. pratense*, together with the increasing genomics resource in *T. pratense* (Sato et al. 2005; Winters et al. 2009; ERA-NET Plant Genomics 2009), it will be possible to identify further markers linked to the *S* locus in *T. repens* and ultimately to locate these markers on *T. pratense* BAC clones to map the syntenic region physically and identify candidate genes. The *S* genotype information in the *T. repens* mapping family can then be used for candidate gene validation experiments. *T. pratense*, as a self-incompatible species closely related to *T. repens*, may well prove to be a more appropriate species than *M. truncatula* for comparative genomics studies aimed at elucidating the nature of the *S* locus in *T. repens.* Indeed, research effort towards map-based cloning of the *S* locus in this species is reported in this issue (Heathcliffe Riday, personal communication) strongly suggesting that the mechanistic basis of self-incompatibility is common to both *Trifolium* species.

Acknowledgments This project is financed by the National Development Plan of Ireland through Teagasc core funding. NC is funded by a Teagasc Walsh Fellowship. We offer our sincere thanks to Matt Lowe, Gerry Douglas, Adriana Ravagnani and Terry Michaelson-Yeates for their help in this work, and to the technicians at Teagasc Oak Park and IBERS. Thanks are due to the colleagues whose comments improved this manuscript.

References

Abberton MT (2007) Interspecific hybridization in the genus *Trifolium*. Plant Breed 126:337–342

Abberton MT, Marshall AH (2005) Progress in breeding perennial clovers for temperate agriculture. J Agr Sci 143:117–135

Allen AM, Hiscock SJ (2008) Evolution and phylogeny of self-incompatibility systems in angiosperms. In: Franklin-Tong VE (ed) Self-incompatibility in flowering plants: evolution, diversity and mechanisms. Springer Verlag, Berlin, pp 73–101

Altinkut A, Kazan K, Gozukirmizi N (2003) AFLP marker linked to water-stress-tolerant bulks in barley (*Hordeum vulgare* L.). Genet Mol Biol 26:77–82

Anderson MA, Cornish EC, Mau S-L, Williams EG, Hogart R, Atkinson A, Bonig I, Grego B, Simpson R, Roche PJ, Haley JD, Penschow JD, Niall HD, Tregar GW, Coghlan JP, Crawford RJ, Clarke AE (1986) Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature 321:38–44

Arroyo MTK (1981) Breeding systems and pollination biology in Leguminosae. In: Polhill RM, Raven PH (eds) Advances in legume systematics. Royal Botanic Gardens, Kew, pp 723–769

Atwood SS (1940) Genetics of cross-incompatibility among self-incompatible plants of *Trifolium repens*. J Am Soc Agron 32:955–968

Atwood SS (1941) Cytological basis for incompatibility in *Trifolium repens*. Am J Bot 28:551–557

Atwood SS (1942) Oppositional alleles causing cross-incompatibility in *Trifolium repens*. Genetics 27:333–338

Atwood SS (1944) Oppositional alleles in natural populations of *Trifolium repens*. Genetics 29:428–435

Atwood SS, Hill HD (1940) The regularity of meiosis in microsporocytes of *Trifolium repens*. Am J Bot 27:730–735

Badr A, Sayed-Ahmed H, El-Shanshouri A, Watson LE (2002) Ancestors of white clover (*Trifolium repens* L.), as revealed by isozyme polymorphisms. Theor Appl Genet 106:143–148

Barrett B, Griffiths A, Schreiber M, Ellison N, Mercer C, Bouton J, Ong B, Forster J, Sawbridge T, Spangenberg G, Bryan G, Woodfield D (2004) A microsatellite map of white clover. Theor Appl Genet 109:596–608

Barth S, Gonzales M, Febrer M, Connolly V (2004) Molecular genetic diversity within and among Irish ecotypes of perennial ryegrass and white clover collected from old pastures. In: Vollmann J, Grausgruber H, Ruckenbauer P (eds) Genetic variation for plant breeding: proceedings of the 17th Eucarpia General Congress. BOKU, Vienna p 162

Bennett MD, Leitch IJ (2003) Angiosperm DNA C-values database. Royal Botanic Gardens, Kew. http://data.kew.org/cvalues/CvalServlet?querytype=2. Accessed 15 January 2010

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573–580

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2008) GenBank. Nucleic Acids Res 36:D25–D30

Brewbaker JL (1953) Oppositional allelism in diploid and autotetraploid *Trifolium hybridum* L. Genetics 38:444–455

Brewbaker JL (1954) Incompatibility in autotetraploid *Trifolium repens* L. I. Competition and self-compatibility. Genetics 39:307–315

Brewbaker JL (1957) Pollen cytology and self-incompatibility systems in plants. J Hered 48:271–277

Doyle JJ, Doyle JL (1987) A rapid total DNA preparation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

East EM (1934) Norms of pollen tube growth in incompatible matings of self-sterile plants. P Natl Acad Sci USA 20:225–230

Ellison NW, Liston A, Steiner JJ, Williams WM, Taylor NL (2006) Molecular genetics of the clover genus (*Trifolium*: Leguminosae). Mol Phylogenet Evol 39:688–705

Entani T, Iwano M, Shiba H, Takayama S, Fukui K, Isogai A (1999) Centromeric localizaton of an *S-RNase* gene in *Petunia hybrida* Vilm. Theor Appl Genet 99:391–397

ERA-NET plant genomics (2009) Research programme report. http://www.erapg.org/downloadattachment/19018/ERA-NET%20PG%20Report%20LR.pdf. Accessed 15 January 2010

Febrer M, Cheung F, Town CD, Cannon SB, Young ND, Abberton MT, Jenkins G, Milbourne D (2007) Construction, characterization and preliminary BAC-end sequencing analysis of a bacterial artificial chromosome library of white clover (*Trifolium repens* L.). Genome 50:412–421

Feldman M, Levy AA (2005) Allopolyploidy: a shaping force in the evolution of wheat genomes. Cytogenet Genome Res 109:250–258

Frame J, Charlton JFL, Laidlaw AS (1998) Temperate forage legumes. CABI Publishing, Wallingford

George J, Sawbridge TI, Cogan NOI, Gendall AR, Smith KF, Spangenberg GC, Forster JW (2008) Comparison of genome structure between white clover and *Medicago truncatula* supports homoeologous group nomenclature based on conserved synteny. Genome 51:905–911

Griffin PC, Woodrow IE, Newbigin EJ (2009) Population genetics of *Ryparosa kurranigii* (Achariaceae), a rare lowland rainforest tree. Biochem Syst Ecol 37:334–340

Hand ML, Ponting RC, Drayton MC, Lawless KA, Cogan NOI, Brummer EC, Sawbridge TI, Spangenberg GC, Smith KF, Forster JW (2008) Identification of homologous, homoeologous, and paralogous sequence variants in an outbreeding allopolyploid species based on comparison with progenitor taxa. Mol Genet Genomics 280:293–304

Heslop-Harrison J, Heslop-Harrison Y (1982) Pollen–stigma interaction in the Leguminosae: constituents of the stylar fluid and stigma secretion of *Trifolium pratense* L. Ann Bot 49:729–735

Heslop-Harrison Y, Shivanna KR (1977) Receptive surface of angiosperm stigma. Ann Bot 41:1233–1258

Igic B, Kohn JR (2001) Evolutionary relationships among self-incompatibility RNases. Proc Natl Acad Sci USA 98:13167–13171

Ikeda K, Ushijima K, Yamane H, Tao R, Hauck NR, Sebolt AM, Iezzoni AF (2005) Linkage and physical distances between the *S*-haplotype *S-RNase* and *SFB* genes in sweet cherry. Sex Plant Reprod 17:289–296

Jones ES, Hughes LJ, Drayton MC, Abberton MT, Michaelson-Yeates TPT, Bowen C, Forster JW (2003) An SSR and AFLP molecular marker-based genetic map of white clover (*Trifolium repens* L.). Plant Sci 165:531–539

Kusaba M, Dwyer K, Hendershot J, Vrebalov J, Nasrallah JB, Nasrallah ME (2001) Self-incompatibility in the genus *Arabidopsis*: Characterization of the *S* locus in the outcrossing *A. lyrata* and its autogamous relative A. *thaliana*. Plant Cell 13:627–643

Lawrence MJ (1975) The genetics of self-incompatibility in *Papaver rhoeas*. P R Soc Lond B Bio 188:275–285

Lawrence MJ (1996) Number of incompatibility alleles in clover and other species. Heredity 76:610–615

Leduc N, Douglas GC, Monnier M, Connolly V (1990) Pollination in vitro: effects on the growth of pollen tubes, seed set and gametophytic self-incompatibility in *Trifolium pratense* L. and T. *repens* L. Theor Appl Genet 80:657–664

Lewis D (1947) Competition and dominance of incompatibility alleles in diploid pollen. Heredity 1:85–108

Lu Y (2006) Historical events and allelic polymorphism at the gametophytic self-incompatibility locus in Solanaceae. Heredity 96:22–28

Magallón SA, Crane PR, Herendeen PS (1999) Phylogenetic pattern, diversity, and diversification of eudicots. Ann Mo Bot Gard 86:297–372

Paglia G, Morgante M (1998) PCR-based multiplex DNA fingerprinting techniques for the analysis of conifer genomes. Mol Breed 4:173–177

Rea AC, Nasrallah JB (2008) Self-incompatibility systems: barriers to self-fertilization in flowering plants. Int J Dev Biol 52:627–636

Richman AD, Uyenoyama MK, Kohn JR (1996) *S*-allele diversity in a natural population of *Physalis crassifolia* (Solanaceae) (ground cherry) assessed by RT-PCR. Heredity 76:497–505

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365–386

Salimpour F, Sharifnia F, Mostafavi G, Hajarasoliha SH, Ukhneh E (2008) Chromosome counts and determination of ploid levels in Iranian species of *Trifolium*. Chromosome Bot 3:53–63

Sassa H, Hirano H, Ikehashi H (1993) Identification and characterization of stylar glycoproteins associated with self-incompatibility genes of Japanese pear, *Pyrus serotina* Rehd. Mol Gen Genet 241:17–25

Sato S, Isobe S, Asamizu E, Ohmido N, Nakamura Y, Kaneko T, Sakurai N, Okumura K, Klimenko I, Sasamoto S, Wada T, Watanabe A, Kohara M, Fujishiro T, Tabata S (2005) Comprehensive structural analysis of the genome of red clover (*Trifolium pratense* L.). DNA Res 12:301–364

Schierup MH, Vekemans X (2008) Genomic consequences of selection on self-incompatibility genes. Curr Opin Plant Biol 11:116–122

Silow RA (1931) A preliminary report on pollen-tube growth in red clover (*Trifolium pratense* L.). Welsh Plant Breed Stn Bull Series H: 228–233

Silva NF, Goring DR (2001) Mechanisms of self-incompatibility in flowering plants. Cell Mol Life Sci 58:1988–2007

Sledge MK, Ray IM, Jiang G (2005) An expressed sequence tag SSR map of tetraploid alfalfa (*Medicago sativa* L.). Theor Appl Genet 111:980–992

Steinbachs JE, Holsinger KE (1999) Pollen transfer dynamics and the evolution of gametophytic self-incompatibility. J Evol Biol 12:770–778

Steinbachs JE, Holsinger KE (2002) *S*-RNase-mediated gametophytic self-incompatibility is ancestral in eudicots. Mol Biol Evol 19:825–829

Takayama S, Isogai A (2005) Self-incompatibility in plants. Annu Rev Plant Biol 56:467–489

Ushijima K, Sassa H, Dandekar AM, Gradziel TM, Tao R, Hirano H (2003) Structural and transcriptional analysis of the self-incompatibility locus of almond: identification of a pollen-expressed F-box gene with haplotype-specific polymorphism. Plant Cell 15:771–781

Van Ooijen JW, Voorrips RE (2001) JoinMap® 3.0: Software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Vieira J, Fonesca NA, Vieira CP (2008) An *S-RNase*-based gametophytic self-incompatibility system evolved only once in eudicots. J Mol Evol 67:179–190

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. J Heredity 93:77–78

Vos P, Hogers R, Bleeker M, Reijans M, Vandelee T, Hornes M, Fritjers A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Williams RD (1931) Methods and technique of breeding red clover, white clover, and lucerne. Welsh Plant Breed Stn Bull Series H: 46–77

Williams WM (1987) Genetics and breeding. In: Baker MA, Williams WM (eds) White clover. CAB International, Wallingford, pp 343–419

Williams RD, Williams W (1947) Genetics of red clover (*Trifolium pratense* L.) compatibility, III: the frequency of incompatibility *S*-alleles in two non-pedigree populations of red clover. J Genet 48:67–79

Williams EG, Plummer J, Phung M (1982) Cytology and fertility of *Trifolium repens*, *T. ambiguum*, T. *hybridum*, and interspecific hybrids. NZ J Bot 20:115–120

Williams WM, Ansari HA, Hussain SW, Ellison NW, Williamson ML, Verry IM (2008) Hybridization and introgression between two diploid wild relatives of white clover, *Trifolium nigrescens* Viv. and *T. occidentale* Coombe. Crop Sci 48:139–148

Winters A, Heywood S, Farrar K, Donnison I, Thomas A, Webb KJ (2009) Identification of an extensive gene cluster among a family of PPOs in *Trifolium pratense* L. (red clover) using a large insert BAC library. BMC Plant Biol 9:94–105

Wright S (1939) The distribution of self-sterility alleles in populations. Genetics 24:538–552

Xue Y, Carpenter R, Dickinson HG, Coen ES (1996) Origin of allelic diversity in *Antirrhinum* S locus RNases. Plant Cell 8:805–814

Yang Q, Zhang D, Li Q, Cheng Z, Xue Y (2007) Heterochromatic and genetic features are consistent with recombination suppression of the self-incompatibility locus in *Antirrhinum*. Plant J 51:140–151

Zhang Y, Sledge MK, Bouton JH (2007) Genome mapping of white clover (*Trifolium repens* L.) and comparative analysis within the Trifolieae using cross-species SSR markers. Theor Appl Genet 114:1367–1378