Theor Appl Genet (2007) 114:1367–1378
DOI 10.1007/s00122-007-0523-3

ORIGINAL PAPER

# Genome mapping of white clover (*Trifolium repens* L.) and comparative analysis within the Trifolieae using cross-species SSR markers

**Yan Zhang · Mary K. Sledge · Joe H. Bouton**

Received: 25 July 2006 / Accepted: 2 February 2007 / Published online: 14 March 2007
© Springer-Verlag 2007

**Abstract** Allotetraploid white clover (*Trifolium repens* L.), a cool-season perennial legume used extensively as forage for livestock, is an important target for marker-assisted breeding. A genetic linkage map of white clover was constructed using simple sequence repeat (SSR) markers based on sequences from several Trifolieae species, including white clover, red clover (*T. pratense* L.), *Medicago truncatula* (Gaertn.) and soybean (*Glycine max* L.). An $F_1$ population consisting of 179 individuals, from a cross between two highly heterozygous genotypes, GA43 and Southern Regional Virus Resistant, was used for genetic mapping. A total of 1,571 SSR markers were screened for amplification and polymorphism using DNA from two parents and 14 $F_1$s of the mapping population. The map consists of 415 loci amplified from 343 SSR primer pairs, including 83 from white clover, 181 from red clover, 77 from *M. truncatula*, and two from soybean. Linkage groups for all eight homoeologous chromosome pairs of allotetraploid white clover were detected. Map length was estimated at 1,877 cM with 87% genome coverage. Map density was approximately 5 cM per locus. Segregation distortion was detected in six segments of the genome (homoeologous groups A1, A2, B1, B2, C1, and D1). A comparison of map locations of markers originating from white clover, red clover, and alfalfa (*M. sativa* L.) revealed putative macro-colinearity between the three Trifolieae species. This map can be used to link quantitative trait loci with SSR markers, and accelerate the improvement of white clover by marker-assisted selection and breeding.

Communicated by A. Kilian.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00122-007-0523-3) contains supplementary material, which is available to authorized users.

Y. Zhang (✉) · M. K. Sledge · J. H. Bouton
Forage Improvement Division,
The Samuel Roberts Noble Foundation, 2510,
Sam Noble Parkway, Ardmore, OK, 73401, USA
e-mail: yzhang@noble.org

## Introduction

White clover (*Trifolium repens* L.) is a major cool-season forage legume found throughout the world. It can grow in a broad range of soil and climatic conditions with proper management. Active growth from germinating seeds or existing plants begins with cooler temperatures and increased moisture, and growth continues until night temperatures reach freezing (Gibson and Cope 1985).White clover is more widely adapted within the United States than any other clover species (Gibson and Cope 1985). In addition to its nitrogen fixing ability, white clover is also one of the most nutritious and palatable legumes used in pastures to improve forage quality. White clover is hence a common companion species in cool-season, and to some extent warm-season, perennial grass pastures (Brink et al. 1999).

White clover is an allopolyploid ($2n = 4x = 32$) outcrossing species (Atwood and Hill 1940) with a 1C genome size of 956 Mbp (Bennett and Leitch 2003), and a common map length estimate of 1,200 cM (Barrett et al. 2004). Recently, Ellison et al. (2006) identified *T. occidentale* (Schreb.) and *T. pallescens* (Schreb.) as the two diploid progenitors of white clover

by phylogenetic analyses based on nuclear ribosomal DNA internal transcribed spacer and chloroplast *trnL* intron sequences. White clover belongs to the Fabaceae family and the Trifolieae tribe (Williams 1987), which also includes several well studied forage legumes such as barrel medic (*M. truncatula*), alfalfa (*M. sativa* L.), and red clover (*T. pratense* L.). Among all of these species, red clover is the most closely related to white clover within the Trifolieae.

Microsatellite or simple sequence repeat (SSR) markers are widely used for plant genome analysis. They are PCR-based, co-dominant markers, occurring at a high frequency in eukaryotic genomes (Li et al. 2002), and are generally associated with non-repetitive DNA regions (Morgante et al. 2002). These markers are highly reproducible across populations within the same species, and to some extent, across species and genera (Eujayl et al. 2004; Gaitán-Solís et al. 2002; Wang et al. 2004; Mian et al. 2005). SSR markers have been used to construct comprehensive genetic linkage maps of several Trifolieae species, including red clover (Sato et al. 2005) and alfalfa (Sledge et al. 2005; Julier et al. 2003).

Two white clover genetic maps have been previously published (Jones et al. 2003; Barrett et al. 2004). Jones et al. (2003) produced a map based on an Australian $F_2$ population of 125 individuals, generated at Institute of Grassland and Environmental Research (IGER) in Aberystwyth, UK. Eighteen linkage groups containing 135 loci were mapped with 78 SSR primers and 57 amplified fragment length polymorphism (AFLP) primers. A number of weak secondary loci were identified and suggested as homeoloci, but no homoeologous pairs of the tetraploid genome were identified. This genetic map, however, was useful for the detection of quantitative trait loci (QTLs) for vegetative morphogenesis traits, reproductive morphogenesis traits, and developmental traits (Cogan et al. 2006). Barrett et al. (2004) produced a white clover map based on a double pseudo-testcross (Grattapaglia and Sederoff 1994) mapping population developed in New Zealand. Even though fewer individuals (92 $F_1$ progeny) were used for construction of this map, the authors established a well defined, medium density map, which contained 493 loci detected by 365 SSR primer pairs, including homoeoloci from the two genomes of allotetraploid white clover (Barrett et al. 2004). QTLs linked to seed yield traits were detected using mapped markers (Barrett et al. 2005). Although these two genetic maps provided valuable information on white clover genomics, only a small number of the map-specific markers have been made available in the public domain. Kölliker et al. (2001a) used AFLP markers to determine levels of genetic variation in 52 white clover cultivars and accessions. Random Amplified Polymorphic DNA markers were used conducting a similar study for North American cultivars and germplasms (Gustine et al. 2002); however, the identification of additional PCR-based white clover markers for QTL detection is still needed. In order to map traits in germplasm adapted to North America, specially the southern Great Plains, construction of genetic linkage maps using locally adapted populations is critical.

Comparative genomics has become an important strategy for extending genetic information from model species to more genetically complex species (Gale and Devos 1998; Paterson et al. 2005). Studies have demonstrated that comparative genome analysis can reveal genetic conservation among the genomes of closely related species and can greatly facilitate gene discovery (Sorrells et al. 2003; Jaiswal et al. 2006; Herrmann et al. 2006). Comparisons of model plants with major crop genomes have been conducted based on the genomic sequence homology with *Arabidopsis* (Barnes 2002) and rice (Bennetzen and Ma 2003). Choi et al. (2004b) initiated comparison of the model legume *M. truncatula* with five crop legumes, including soybean, alfalfa, garden pea (*Pisum sativum* L.), mung bean (*Vigna radiate* L.), and common bean (*Phaseolus vulgaris* L.), using cross-species genetic markers, and reported high genome conservation among the six species. This was followed by an updated draft of the legume comparative map with eight species by adding *Lotus japonicus* (L.) and chickpea (*Cicer arietinum* L.) (Zhu et al. 2005). Recently, the Legume Information System (http://www.comparative-legumes.org) was developed (Gonzales et al. 2005) as a comparative legume resource that integrates genetic and molecular data from multiple legume species, enabling cross-species genomic, and transcript comparisons. These studies have established a framework of legume comparative genomics, and serve as a resource for the genomics of all legume species. Several comparative studies have focused on alfalfa (Kalo et al. 2004; Choi et al. 2004a), while the comparative mapping of other forage legumes has lagged behind. Macro-colinearity between the genomes of red clover and two model legumes, *L. japonicus and M. truncatula* has been reported (Sato et al. 2005). Although white clover has high agricultural importance, apparently there are no published reports on comparative mapping for white clover and the other major legume species.

In this study, a genetic linkage map was developed for an allotetraploid white clover population adapted to the southern Great Plains of the USA, using SSR markers from several legume species, including white clover, red clover, tetraploid alfalfa, *M. truncatula*, and

soybean. Common markers mapped in white clover, red clover (Sato et al. 2005), and alfalfa (Sledge et al. 2005) were compared to assess the colinearity of the genomes between these Trifolieae species.

## Materials and methods

### Plant materials

Two phenotypically divergent heterozygous white clover genotypes, GA43 and Southern Regional Virus Resistant (SRVR), were used as parents of the mapping population. GA43 is a genotype of the commercial cultivar 'Durana' (Bouton et al. 2005), which is an intermediate-type clover intended for use as a renovation legume for grass pastures in the southeastern USA. It has a high stolon density, a short plant height with prostrate growth habit, small leaflets, short petioles, an early heading date, and a high frequency of cyanogenesis. The germplasm SRVR (Gibson et al. 1989) is a ladino-type clover with medium-large leaves, long stolons, an upright growth habit, and is non-cyanogenic. A double pseudo-testcross population consisting of 179 $F_1$s was generated from reciprocal crosses between GA43 and SRVR.

### DNA isolation, SSR amplification, and detection of SSR fragments

DNA was purified from young leaf tissue using the Plant DNeasy Kit (Qiagen, Valencia, CA, USA). Forward and reverse primers, from a range of sources (Table 1), were synthesized by Qiagen/Operon Biotechnologies (Los Angeles, CA, USA) with an additional 18 nucleotides from the M13 universal primer appended to the 5′ end of the forward primer (Schuelke 2000). PCR reactions were prepared in a reaction volume of 10 μl which contained 20 ng of template DNA, 2.5 mM $MgCl_2$, 1 × PCR buffer II (Applied Biosystems, Foster City, CA, USA), 0.2 mM dNTPs, 1.0 pmol each of reverse and M13 universal primer, 0.25 pmol of the forward primer, and 0.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA). The M13 universal primer was labeled either with blue (6-FAM), green (VIC), yellow (NED), or red (PET) fluorescent tags (Applied Biosystems, Foster City, CA, USA). High-throughput robotic pipetting with the BioMek 2000 (Beckman Coulter, Inc., Fullerton, CA, USA) was used to set up the PCR reactions in 384-well plates (Corning Inc., Corning, NY, USA). PCR conditions were as follows: 95°C for 10 min; 30 cycles of 95°C for 30 s, the optimum annealing temperature of the reverse primer for 45 s, and 72°C for 45 s; 10 cycles of 95°C for 30 s, 53°C for 45 s and 72°C for 45 s, ending with a 72°C extension step for 10 min, followed by 4°C. After these 40 cycles, 3 μl of PCR products from each of four different fluorescent labeled reactions were pooled for detection. Pooled PCR products (3 μl) were combined with 10 μl of deionized formamide (Applied Biosystems, Foster City, CA, USA) and 0.5 μl of GeneScan-500 LIZ internal size standard (Applied Biosystems, Foster City, CA, USA). SSR fragments were analyzed on the ABI PRISM 3730 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and visualized and scored with GeneMapper 3.7 software (Applied Biosystems, Foster City, CA, USA).

### SSR linkage analysis

For each SSR, individual fragments were scored as dominant markers. Primers amplifying fragments that were polymorphic between the two parents and segregating among 14 progenies were used to screen the $F_1$ population. Segregating peaks were scored as a

**Table 1** Evaluation of SSR primer pairs for amplification efficiency and polymorphism detection between GA43 and SRVR parents of the mapping population

| Origin of PP | PP screened (no.) | Functional PP[a] (no.) | PP screened on $F_1$s (no.) | PP amplifying alleles (no.) | Transferability[b] | PP mapped (no.) |
|---|---|---|---|---|---|---|
| White clover SSR- enriched genomic | 78 | 77 | 75 | 73 | 94% | 64 |
| White clover EST | 32 | 30 | 28 | 23 | 72% | 19 |
| Red clover | 599 | 527 | 213 | 205 | 34% | 181 |
| *M. truncatula* EST | 772 | 640 | 236 | 166 | 22% | 70 |
| *M. truncatula* BAC | 40 | 12 | 8 | 7 | 18% | 7 |
| Soybean | 50 | 16 | 5 | 5 | 10% | 2 |
| Total | 1,571 | | 565 | 479 | | 343 |

*PP* primer pair

[a] Primer amplified white clover genomic DNA, and produced detectable amplicons

[b] Primer transferability: percentage of primer pairs amplifying at least one segregating allele vs. number of screened primers

0/1 data matrix and recoded as cross-pollinated loci (np × nn for SRVR, and ll × lm for GA43). SSR primer pairs with fragments that segregated approximately 1:1 in the $F_1$s were selected for construction of parental maps. The maps in each parent, and subsequent consensus maps of the genome, were constructed using the cross-pollinated population analysis by JoinMap 3.0 (Van Ooijen and Voorrips 2001). Genetic map distances were estimated using the Kosambi mapping function. For construction of single parent maps, loci grouping and ordering probability minima were LOD (logrithm of odds) = 3.

A bi-parental consensus map was created by joining the two single parent maps using markers mapped in both parents, with multiple segregating alleles per locus, as described by Barrett et al. (2004). Distorted alleles ($P < 0.05$) were included in the linkage analysis for consensus map construction. Homoeologous linkage groups were identified and aligned using putative homoeologous loci, which were detected by the same SSR primer pair and mapped at similar positions in the two homoeologues. Minor linkage groups were integrated into the major ones within a homologue by combining common loci present in both linkage groups or by comparing genetic distance in parental maps to construct sixteen linkage groups. The eight homoeologous pairs were named A–H and homoeologues within each pair were named one and two, respectively, based on the same loci detected on a previously published map of *T. repens* (Barrett et al. 2004).

## Results

### SSR primer pairs

A total of 1,571 SSR primer pairs from four target species were used to screen for amplification and polymorphism between the two parents and 14 progeny from the mapping population. These included primers pairs developed from DNA sequences of genomic, expressed sequence tags (ESTs) and bacterial artificial chromosome inserts (BACs) in white clover, red clover, soybean, and *M. truncatula* (Table 1). Five hundred and sixty-five primer pairs amplifying fragments that were polymorphic between the two parents were selected for genotyping the $F_1$ population (Table 1). Approximately 34% of red clover primer pairs were transferable, followed by 18–22% of *M. truncatula* primer pairs, and 10% transferability (the percentage of primers amplifying at least one segregating allele vs. number of screened primers) of soybean primer pairs.

Although multi-species markers were used in the construction of the current map, identity of the repeat array in white clover is unknown without sequencing the SSR. We found only primer pairs corresponding to SSR arrays with two to five nucleotide repeat motifs in the structure of the original species were mapped. Except for four unknowns, the 339 mapped primer pairs included 85 di-, 202 tri-, 41 tetra-, and 11 pentanucleotide motifs. The length of SSR arrays varied from 12 to 80 nucleotides (mean = 22).

### Genetic linkage map of *Trifolium repens*

A genetic linkage map was constructed using 179 $F_1$ genotypes from the white clover population GA43 × SRVR. Dominant marker data was used to initiate the linkage analysis. Alleles ranged in length from 85 to 497 bp, with a mean value of 235 bp. These single parent maps were subsequently assembled into a bi-parental consensus linkage map (Fig. 1) using co-dominant markers. A total of 415 loci were placed on this consensus map with 343 SSR primer pairs (Supplementary Table 1), including loci detected by 83 white clover primer pairs (Barrett et al. 2004; Jones et al. 2003; Kölliker et al. 2001b), 181 red clover SSR primer pairs (Sato et al. 2005), 70 EST–SSR and seven BAC–SSR primer pairs from *M. truncatula* (Sledge et al. 2005), and two soybean EST–SSR (derived from searching GenBank dbEST database) primer pairs. Of these 415 loci, 78 were mapped in GA43 only, 247 were mapped in SRVR only, and the remaining 90 were mapped in both parental genotypes.

Eight homoeologous pairs corresponding to the 16 chromosomes of white clover were established by assembling 25 linkage groups, which ranged from 11 to 155 cM in length, and contained 3–38 loci per group. Lengths of assembled eight homoeologous pairs were from 82cM (G2) to 155 cM (A1) (Table 2). Loci were not evenly distributed among homologues and homoeologous pairs, ranging from 12 (G2) to 41 (B1 and E1) loci per homologue and 27 (H) to 70 (B) per homoeologous pair (Table 2). Marker-locus densities also varied from 3.2 to 9.4 cM per locus (an overall mean of 5.0 cM per locus). Significant differences were present between homoeologous groups D (4.2 cM vs. 7.8 cM) and H (6.1 cM vs. 9.4 cM), although map density within other homoeologous pairs were fairly consistent (Table 2). Gap distances between loci ranged from 0.083 cM (B2) to 24.842 (H2). The average distance between loci in each homologue ranged from 3.27 cM on B2 to 10.14 cM on H2, with a mean of 5.3 cM (Table 2). The total map length was 1,877 cM, with 85 (G2) to 95% (A1,

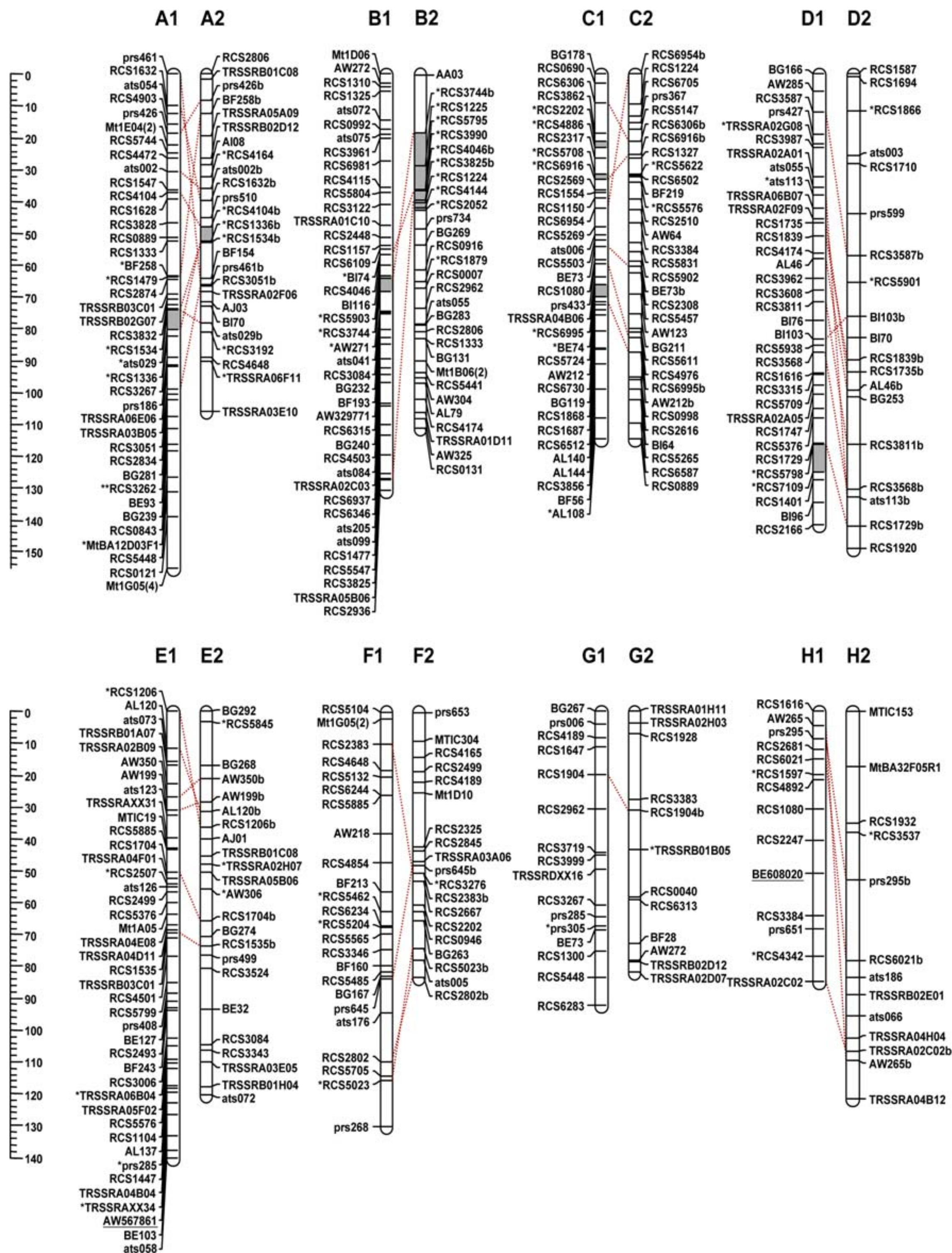


**Fig. 1** A genetic linkage map of the white clover (GA43 × SRVR). Eight homoeologous pairs of linkage groups are labeled A–H, and homoeologues within each pair are designated 1 and 2, respectively. Genetic length (cM) is indicated by the *ruler on the left side* of the map. Homoeologues loci were labeled with *b* at the end of primer names and are joined by *dashed lines*. Loci with segregation distortion ($P < 0.05$) are indicated by *astreisk* in front of primer names, and genomic segments encompassed by distorted loci are highlighted in *gray*. Origin of marker loci is indicated as follows: *RCS* red clover SSRs, *ats* white clover genomic SSRs, *prs* white clover EST–SSRs, *TRSSR* white clover SSR-enriched libraries, and *Mt M. truncatula* BAC-SSRs. Two soybean markers are *underlined*. All others markers are *M. truncatula* EST–SSRs

**Table 2** Simple sequence repeat marker distribution among the 16 linkage groups of the white clover genome

| Linkage group | Map length (cM) | No. of loci | Map density (cM) | Inter-locus gap distance | Genome coverage[a] (%) | Number of loci distorted |
|---|---|---|---|---|---|---|
| A1 | 155 | 39 | 4.0 | 4.1 | 95 | 7 |
| A2 | 106 | 25 | 4.2 | 4.4 | 92 | 6 |
| B1 | 131 | 41 | 3.2 | 3.3 | 95 | 4 |
| B2 | 111 | 29 | 3.8 | 4.0 | 93 | 10 |
| C1 | 115 | 34 | 3.4 | 3.5 | 94 | 7 |
| C2 | 115 | 32 | 3.6 | 3.7 | 94 | 2 |
| D1 | 141 | 34 | 4.2 | 4.3 | 94 | 4 |
| D2 | 149 | 19 | 7.8 | 8.3 | 90 | 2 |
| E1 | 140 | 41 | 3.4 | 3.5 | 95 | 5 |
| E2 | 120 | 23 | 5.2 | 5.5 | 92 | 3 |
| F1 | 130 | 24 | 5.4 | 5.7 | 92 | 3 |
| F2 | 83 | 19 | 4.4 | 4.6 | 90 | 1 |
| G1 | 92 | 16 | 5.8 | 6.1 | 88 | 1 |
| G2 | 82 | 12 | 6.8 | 7.4 | 85 | 1 |
| H1 | 85 | 14 | 6.1 | 6.5 | 87 | 2 |
| H2 | 122 | 13 | 9.4 | 10.1 | 86 | 1 |
| Total | 1,877 | 415 | – | | – | 59 |
| Mean | 117 | 26 | 5.0 | 5.3 | 87 | 3.7 |

Linkage data were developed using 179 $F_1$ progenies in the mapping population GA43 × SRVR. Locus segregation distortion was declared at $P < 0.05$

[a] Genome coverage = Map length/{map length × [(No. of loci + 1)/(No. of loci − 1)]}

B1, and E1) genome coverage (Chakravarti et al. 1991) in each linkage group.

Forty-three homoeologous loci were identified on linkage groups A–H (Fig. 1), which facilitated the detection of the eight homoeologous pairs in the allotetraploid *T. repens* genome. Ten of the loci were located on group A, followed by nine on group D, with only one homoeologous locus detected on group G. Fourteen percent of red clover, 12% of *M. truncatula*, and 10% of white clover primers detected homoeologous loci. Except for group A, marker order and inter-locus map distances were well conserved between homoeologues (Fig. 1). Twenty-nine (8.4%) primer pairs detected duplicate non-homoeologous loci between linkage groups (e.g., RCS1616 detected two loci, one on D1 and one on H1; TRSSRB01C08 detected two loci, one on A2 and one on E2).

A total of 59 distorted loci ($P < 0.05$) were detected (Table 2) in distinct clusters on linkage groups A1, A2, B1, B2, C1, and D1 (Fig. 1), which covered 78 cM of the genome. Seventeen and 16 distorted loci were detected only in GA43 or SRVR, respectively; the remaining 26 loci were detected in both parents. The degree of segregation distortion was compared among homologues using a $\log_{10}$ transformation of the $P$ value of each mapped locus (Fig. 2). Significantly higher distortion was detected for nine loci clustering at the top end of B2 ($7.1 < -\log_{10}P < 63.8$) compared to other distorted loci. Groups A1 and D1 had similar degrees of distortion, which was higher than the remaining groups (A2, B1, and C1) (Fig. 2).

## Comparative mapping of *T. repens*, *T. pratense*, and *M. sativa*

Two hundred and twenty-six markers from 182 primers previously mapped in diploid red clover (Sato et al. 2005) were placed on this white clover map. Markers were distributed unevenly between white clover homoeologous pairs, ranging from 12 markers on group H to 46 markers on group C (Table 3). Distribution of the markers originating from *T. pratense* indicated that portions of up to five linkage groups were colinear with some homoeologous groups in white clover (e.g., homoeologous group A contained markers from LG2, LG3, LG4, LG5, LG6, and LG7 in *T. pratense*). However, the numbers of markers were significantly different among linkage groups. Putative genome alignments between the two *Trifolium* species were proposed with 167 red clover markers (Table 3, Fig. 3). Among them, white clover homoeologous groups A, C, E, G, and H were well defined in the alignment with red clover linkage groups 7, 6, 1, 4, and 7 respectively, in which more than 50% of the mapped markers were common to linkage groups of both species.

The same approach was applied to 50 *M. truncatula* SSR markers previously mapped in autotetraploid *M. sativa* (Sledge et al. 2005). Thirty-seven SSRs were used to align the genomes of *T. repens* and *M. sativa* (Fig. 3). White clover homoeologous group C was closely related to LG7 of *M. sativa*, in which all nine markers mapped in both groups with no obvious difference in marker order or genetic distance between loci. Similar alignments were also present

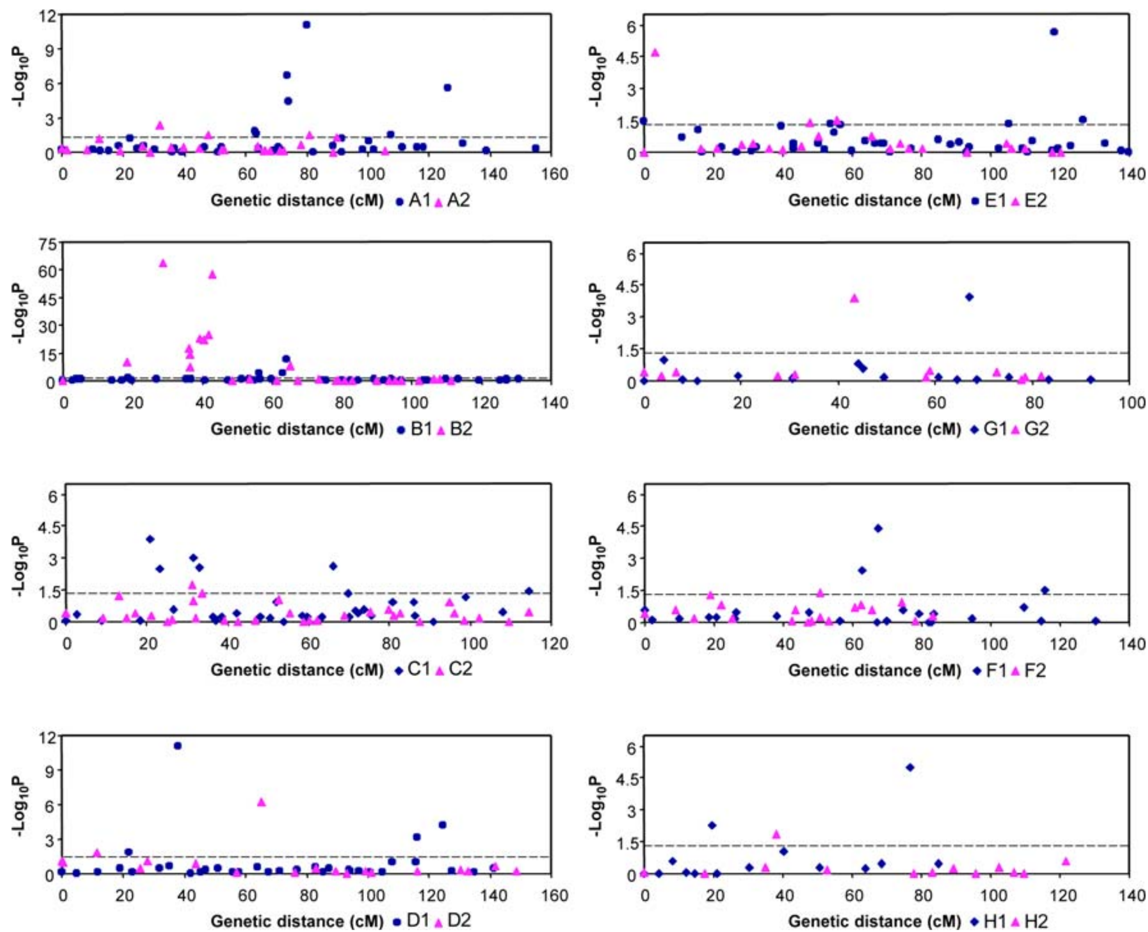


**Fig. 2** Segregation distortion of loci on linkage group A–H. *X axis* loci positions marked by genetic distance, *Y axis* $-\text{Log}_{10}P$. Locus segregation distortion was declared at $P < 0.05$, which is indicated by *dash line* on $-\text{Log}_{10}P > 1.3$. Homoeologous pairs were grouped in the same chart. Each chart was scaled based on $-\text{Log}_{10}P$ and genetic distance within *linkage groups*

**Table 3** *Trifolium pratense* marker distribution and genome alignments between *T. repens* and *T. pratense*

| *T. repens* homoeologous group | No. of *T. pratense* marker mapped | Distribution of *T. pratense* primers | | | | | | | Putative alignment to *T. pratense* |
|---|---|---|---|---|---|---|---|---|---|
| | | LG1 | LG2 | LG3 | LG4 | LG5 | LG6 | LG7 | |
| A | 31 | – | 1 (3%) | 3 (10%) | 2 (6%) | 4 (13%) | 2 (6%) | 19 (61%) | LG7 |
| B | 41 | 1 (2%)[a] | 10 (24%) | – | 7 (17%) | 15 (37%) | 3 (7%) | 5 (12%) | LG2, LG5 |
| C | 46 | – | 1 (2%) | – | 1 (2%) | 2 (4%) | 41 (89%) | 1 (2%) | LG6 |
| D | 32 | – | 1 (3%) | 14 (44%) | 13 (41%) | 1 (3%) | 1 (3%) | 2 (6%) | LG3, LG4 |
| E | 21 | 12 (57%) | 1 (5%) | – | 1 (5%) | 1 (5%) | 1 (5%) | 5 (24%) | LG1, LG7 |
| F | 28 | – | 2 (7%) | – | 6 (21%) | 12 (43%) | 4 (14%) | 4 (14%) | LG4, LG5 |
| G | 15 | – | 2 (13%) | – | 10 (67%) | 1 (7%) | 1 (7%) | 1 (7%) | LG4 |
| H | 12 | – | 1 (8%) | – | – | 1 (8%) | 4 (33%) | 6 (50%) | LG6, LG7 |
| Total | 226 | 13 | 19 | 17 | 40 | 37 | 57 | 43 | |

*LG* linkage group

[a] Number and percentage of markers originated from *T. pratense* linkage group

between white clover group B and *M. sativa* LG8, where seven markers were well aligned with minor rearrangements. Compared with other groups, homoeologous groups F, G, and H were less well defined due to a smaller number (<3 markers) of markers mapped in *M. sativa*.

## Discussion

### Genetic linkage mapping of white clover

A linkage map consisting 415 loci from 343 SSR markers was developed for *T. repens*. Although more

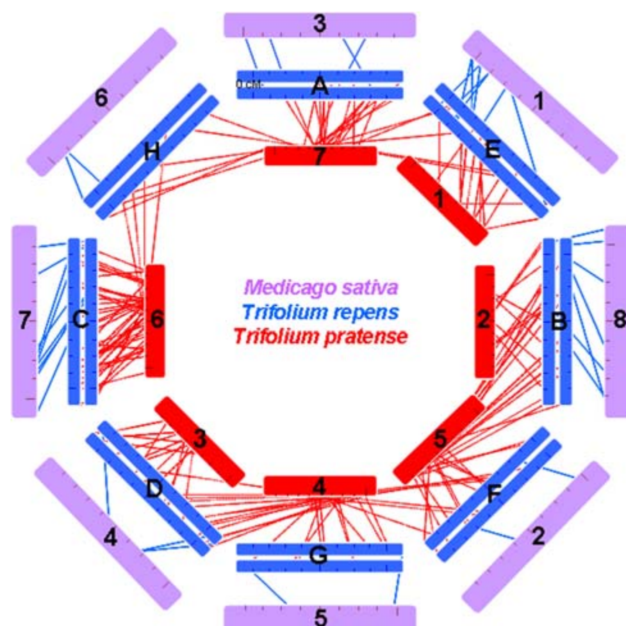


**Fig. 3** Consensus comparative map for *M. sativa*, *T. repens*, and *T. pratense*. *Three levels of circles* represent genomes of the three species, *outside circle* represents autotetraploid *M. sativa*, *middle circle* represents allotetraploid *T. repens* with two homoeologous pairs, *inside circle* represents *T. pratense*. Each *bar* represents one linkage group of the species, and is rescaled by the actual map length. The names of the *T. repens* homologue groups were labeled on the groups, with homoeologous *1* inside and *2* outside. Linkage group numbers were labeled on top of the *bars* for *M. sativa* and *T. pratense*. *Corresponding markers* between *T. repens* and *T. pratense*/*T. repens* and *M. sativa* were connected with *lines* based on map distance. Linkage groups were orientated with 0 cM on *left side* of homoeologues A of *T. repens*

then 58% of the markers were non EST-based SSRs, tri-nucleotide repeat markers formed the majority (59%) of the markers. This is in contrast to early findings that non-coding regions of eukaryotic genomes primarily contain di-nucleotide repeats (Li et al. 2002). Approximately 22% of mapped loci were polymorphic in both parental genotypes, and were used to assemble the two parental maps into a consensus map. This number is slightly lower than that reported in a previously published white clover map (Barrett et al. 2004), in which 28% of mapped loci were polymorphic in both parents. The total map length was calculated as 1,877 cM in our current study, which is longer than a previously reported map of 1,144 cM (Barrett et al. 2004). This could be due to large inter-locus gaps on D2, G1, G2, H1, and H2 (>6 cM in average). However, the estimated 87% genome coverage indicates that a high percentage of the genome is covered by the markers.

## Segregation distortion

Fourteen percent of the loci showed significant segregation distortion (59 of 415 loci) in the genetic map of this population. Similar numbers of distorted loci were detected in the GA43 and SRVR; therefore, no parental bias for segregation distortion was observed in this mapping population. The amount of segregation distortion is similar to that reported in a white clover pair-cross $F_2$ population (11%) (Jones et al. 2003), and higher than that reported for another white clover double-pseudo-testcross population (6%) (Barrett et al. 2004), but is comparable to the 5.5 and 10% segregation distortion reported in the backcross populations of tetraploid *M. sativa* (Brouwer and Osborn 1999; Sledge et al. 2005). A positional trend of segregation distortion was observed on groups A to D, especially on the upper end of B2, where highly distorted loci covered 39% (43 cM) of the linkage group. Segregation distortion encountered in homoeologous loci in both A and B, e.g., RCS 1336, RCS1534, and RCS3744, may affect the accuracy of the mapping in these groups. This distribution was also different from a previous study, in which distorted loci were clustered at the top of A1 and H2 (Barrett et al. 2004). The differences observed may be associated with different distributions of gametophytic or zygotic lethal or sub-lethal alleles within and between different populations (Williams et al. 1995).

## Transferability of SSR markers

Forty-nine loci (40 primer pairs) detected on the previously published map of *T. repens* (Barrett et al. 2004) were used to assign linkage groups in our study. Approximately 85% of the loci in each homologue were aligned in the same order between the two maps. Minor differences in map location of some markers between homoeologous pairs compared to previously published maps were identified. For example, ats002 detected homeoloci on both A1 and A2 in the current map, but only mapped on A1 in the previous map (Barrett et al. 2004); ats186 mapped on H2 only, but was detected on both H1 and H2 by Barrett et al. (2004). This effect is presumably due to the presence and absence of allelic variation at one, other or both homoeoloci between different parental genotypes. Two primers pairs, ats072 and ats055, are likely to be able to detect multiple loci in white clover: they were mapped to two homologues (B1 and D1, and B2 and E2, respectively) in the map of Barrett et al. (2004) also mapped two homologues on this map, although on different groups (B1 and E2, B2 and D1, respectively). Three of eight rearranged loci involved homologue G,

which was less well defined than other groups, with only 28 loci, suggesting that additional markers are necessary for clarification of the marker order. This same homologue was also the least well defined group in the map of Barrett et al. (2004). In addition, we mapped 43 markers designed from the sequences of SSR-enriched library clones (TRSSRs) that were used for developing the white clover linkage map presented by Jones et al. (2003). Although homeoloci between sub-genomes were not detected, markers from each linkage group were also mapped in similar positions in our mapping population. For example, markers from linkage groups 1, 3, 4, 6, 9, and 10 in Jones et al (2003) were placed on group E1, A1, A2, B1, D1, and H2 in the current map, respectively. This result suggests conservation of the location of loci detected by these TRSSRs primers in different populations of white clover. Considering all of the white clover markers used in this study, more than 86% of the SSR markers initially developed in other white clover populations produced amplification products in our mapping population. Therefore, white clover SSR markers appear to have a high rate of transferability between germplasms, as opposed to some species such as perennial ryegrass (*Lolium perenne* L.) which have only 61% transferability of markers between germplasms (Faville et al. 2004).

Simple sequence repeat markers can be transferred across species within a genus (Peakall et al. 1998; Aranzana et al. 2003). The transfer of SSR markers among species within some legume genera has already been reported in *Glycine* (Peakall et al. 1998), *Phaseolus* (Gaitán-Solís et al. 2002), and *Medicago* (Julier et al. 2003; Choi et al. 2004a; Sledge et al. 2005). The high level of SSR transferability allows the use of a large number of SSRs from publicly available red clover sequences to be used for mapping in white clover. We screened 599 red clover primers (Sato et al. 2005) for amplification and polymorphism in our mapping population. Approximately 88% of the red clover SSR primers screened generated reproducible cross-genus amplicons in white clover. This capacity to amplify SSR markers in these species is not surprising since they both belong to *Trifolium* genus. However, more then half (54%) of these amplicons, were monomorphic between the two parents, and therefore were not useful in map construction. Only 205 primers pairs amplified fragments that were polymorphic and segregating as a pseudo-testcross, which was 34% of red clover primers screened. This percentage is similar to that (30.78%) reported in a study on the transferability of SSR markers across the legume family (Wang et al. 2004), and higher than that (18.5%) reported of transferring five out of 27 red clover SSR into eight white clover genotypes (Kölliker et al. 2006). Red clover SSR sequences were generated from four different types of libraries, SSR-enriched genomic and cDNA libraries, a methyl-filtration genomic library, and a normalized cDNA library (Sato et al. 2005). Among them, primers originating from the SSR-enriched cDNA and genomic libraries gave the highest transfer rates (42 and 34%, respectively) in white clover. Therefore, in this study, primers from SSR-enriched sequences could be more successfully transferred both within species (94% from other white clover) and between species (average 38% from red clover) as compared to primers generated from other sources.

A previous study showed high transferability of *M. truncatula* EST–SSR primers between species within *Medicago* (Eujayl et al. 2004), in which high levels of polymorphism (>70%) were detected for 24 genotypes representing six *Medicago* spp. and subspecies, including diploid alfalfa (*M. sativa* ssp *caerulea*, *M. sativa* ssp *sativa*), tetraploid alfalfa (*M. sativa* ssp *falcata*, *M. sativa* ssp *sativa*), and annual medics (*M. polymorpha*, *M. rugosa*, *M. blancheana*, *M. truncatula*). A linkage map using these *M. truncatula* EST–SSR markers was constructed for tetraploid alfalfa (Sledge et al. 2005). We selected 772 *M. truncatula* EST–SSR primers and screened for fragment amplification in white clover. Approximately 83% produced amplicons in our mapping population. Two hundred and thirty-six primers that showed strong amplification were used to screen the population. Out of these, 166 primers amplified at least one mappable allele. The other 70 primers either did not give consistent amplification patterns for replicated parent samples (non-specific amplification), or amplified fragments that did not fit any cross-pollinated function structures, and therefore, were not used for map construction. Only 70 *M. truncatula* EST–SSR markers showed significant linkage (>95%) with alleles amplified by other primers in our current map. Gutierrez et al. (2005) also reported cross-genus amplification of *M. truncatula* microsatellites in three pulses (40% in faba bean, 36.3% in chickpea, and 37.6% in pea); however, none of the functional microsatellites showed polymorphism among the parental genotypes tested, consequently preventing their immediate use for mapping purposes (Gutierrez et al. 2005). Therefore, while these *M. truncatula* EST–SSRs are valuable for identifying linkage relationships within *Medicago*, they are less valuable for other legume species.

The difference of marker transferability, e.g., much higher between white clover germplasms compared to other species such as ryegrass, and much lower for cross-species transferring compared to *M. truncatula*

vs. to *M. sativa*, might associate with the fact that fertile interspecific hybrids of clover are difficult to achieve in nature, even within *Trifolium* (Taylor et al. 1980). It is likely that gene flow was transferred only between germplasms of conserved genome of white clover, but not to other species.

Genome alignments of Trifolieae species

A total of 167 red clover markers from a *T. pratense* map (Sato et al. 2005) and 37 *M. truncatula* SSR markers from an autotetraploid *M. sativa* map (Sledge et al. 2005) were used to compare genome structural relationships of *T. repens* with *T. pratense*, and *M. sativa*. A proposed alignment relationship between three species within Trifolieae is shown in Fig. 3. Although no markers from *T. pratense* were mapped in *M. sativa* or vice versa, to link two genomes, SSRs derived from these two species were mapped in *T. repens*, and suggest possible genome alignments between *T. pratense* and *M. sativa* (Fig. 3).

Red clover has only seven chromosome pairs ($2n = 14$) (Wipf and Cooper 1938); therefore, it could be expected that markers from two linkage groups of *T. pratense* would be aligned to one group of *T. repens*. Comparison of these two *Trifolium* species, however, revealed significant divergence in genome organization and rearrangements (Fig. 3), which may have been caused by breakage and reunion events (Kalo et al. 2004; Delseny 2004). However, multiple colinear blocks were evident through alignment of the two genomes, representing the junctions of such rearrangements. For example, markers mapped on the upper two-thirds of group D in *T. repens* were aligned with similar locations in LG 3 of *T. pratense*, and markers on the remaining section of D were aligned to the lower end of LG4.

Recently, an *in silico* comparative approach was applied to white clover, *M. truncatula* and *L. japonicus* sequences based on BLAST alignment of 700 SSRs, in which 269 matched sequences between species revealed conservation of genome structures and suggested a predominant 1:1 relationship between each of the homoeologous groups of white clover and a single *M. truncatula* chromosome (George et al. 2006). Because *M. truncatula* and *M. sativa* have very conserved genome structures with co-linearity of linkage groups (Choi et al. 2004b), it is not surprising to find that alignments of *T. repens* with *M. sativa* shown in this study match the *in silico* alignments of *T. repens* with *M. truncatula* (George et al. 2006). Another interesting finding was the similar genome alignments between *T. pratense* and *M. sativa* suggested in this study as compared to those between *T. pratense* and *M. truncatula* reported by Sato et al. (2005). These comparisons further confirmed the colinear relationship between the *M. sativa* and *M. truncatula* genomes, and validated the alignments between *T. repens* and two *Medicago* species. Although macro-colinearity was apparent, large numbers of gene duplications and rearrangements are present between genomes. FISH analysis with mapped SSRs (Choi et al. 2004a; Sato et al. 2005) could help to further define marker location and gene relationships, especially for homoeologous pairs in the white clover genome.

**Acknowledgments** We are grateful to Brent Barrett and Greig Cousins, AgResearch, New Zealand, for their helpful suggestions and mapping advice. Special thanks go to Dr Sachiko Isobe at Japan National Agricultural Research Center for providing primer sequences of *T. pratense* SSRs. We thank Dr Bob Gonzales, Ann Harris, and Jarrod Steele for analyzing the SSR fragments. Dr Brindha Narasimhamoorthy is appreciated for valuable discussion regarding linkage analysis. We thank Dr Mark Sorrells for critical review of the manuscript. This study was supported by The Samuel Roberts Noble Foundation, Inc.

## References

Aranzana MJ, Pineda A, Cosson P, Dirlewanger E, Ascasibar J, Cipriani G, Ryder CD, Testolin R, Abbott A, King GJ, Iezzoni AF, Arús P (2003) A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. Theor Appl Genet 106:819–825

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

Barnes S (2002) Comparing *Arabidopsis* to other flowering plants. Curr Opin Plant Biol 5:128–134

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

Barrett B, Baird I, Woodfield D (2005) A QTL analysis of white clover seed production. Crop Sci 45:1844–1850

Bennett MD, Leitch IJ (2003) Angiosperm DNA C-values database (release 4.0, January 2003). http://www.rbgkew.org.uk/cval/ homepage.html

Bennetzen JL, Ma J (2003) The genetic colinearity of rice and other cereals on the basis of genomic sequence analysis. Curr Opin Plant Biol 6:128–133

Bouton JH, Woodfield DR, Caradus JR, Wood DT (2005) Registration of 'Durana' white clover. Crop Sci 45:797

Brink GE, Pederson GA, Alison MW, Ball DM, Bouton JH, Rawls RC, Steudemann JA, Venuto BC (1999) Growth of white clover ecotypes, cultivars, and germplasms in the southeastern USA. Crop Sci 39:1809–1814

Brouwer DJ, Osborn TC (1999) A molecular marker linkage map of tetraploid alfalfa (*Medicago sativa* L.). Theor Appl Genet 99:1194–1200

Chakravarti A, Lasher LK, Reefer JE (1991) A maximum likelihood method for estimating genome length using genetic linkage data. Genetics 128:175–182

Choi HK, Kim D, Uhm T, Limpens E, Lim H, Mun JH, Kalo P, Penmetsa RV, Seres A, Kulikova O, Roe BA, Bisseling T,

Kiss GB, Cook DR (2004a) A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa*. Genetics 166:1463–1502

Choi HK, Mun JH, Kim DJ, Zhu H, Baek JM, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, Young ND, Cook DR (2004b) Estimating genome conservation between crop and model legume species. Proc Natl Acad Sci USA 101:15289–15294

Cogan NO, Abberton MT, Smith KF, Kearney G, Marshall AH, Williams A, Michaelson-Yeates TP, Bowen C, Jones ES, Vecchies AC, Forster JW (2006) Individual and multi-environment combined analyses identify QTLs for morphogenetic and reproductive development traits in white clover (*Trifolium repens* L.). Theor Appl Genet 112:1401–1415

Delseny M (2004) Re-evaluating the relevance of ancestral shared synteny as a tool for crop improvement. Curr Opin Plant Biol 7:126–131

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

Eujayl I, Sledge MK, Wang L, May GD, Chekhovskiy K, Zwonitzer JC, Mian MA (2004) *Medicago truncatula* EST–SSRs reveal cross-species genetic markers for *Medicago* spp. Theor Appl Genet 108:414–422

Faville MJ, Vecchies AC, Schreiber M, Drayton MC, Hughes LJ, Jones ES, Guthridge KM, Smith KF, Sawbridge T, Spangenberg GC, Bryan GT, Forster JW (2004) Functionally associated molecular genetic marker map construction in perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 110:12–32

Gaitán-Solís E, Duque MC, Edwards KJ, Tohme J (2002) Microsatellite repeats in common bean (*Phaseolus vulgaris*): isolation, characterization, and cross-species amplification in *Phaseolus* ssp. Crop Sci 42:2128–2136

Gale MD, Devos KM (1998) Comparative genetics in the grasses. Proc Natl Acad Sci USA 95:1971–1974

George J, Cogan NOC, Smith KF, Spangenberg GC, Forster JW (2006) Genetic map integration and comparative genome organization of white clover (*Trifolium repens* L.) with model legume species. Poster 452. In: Abstracts of plant and animal genome XI Conference, San Diego, CA, USA, 14–18 January 2006

Gibson PB, Cope WA (1985) White Clover. In: Taylor NL (ed) Clover Science and Technology. ASA, CSSA, SSSA, Madison, pp 471–490

Gibson PB, Barnett OW, Pederson GA, McLaughlin MR, Knight WE, Miller JD, Cope WA, Tolin SA (1989) Registration of southern regional virus resistant white clover germplasm. Crop Sci 29:241–242

Gonzales MD, Archuleta E, Farmer A, Gajendran K, Grant D, Shoemaker R, Beavis WD, Waugh ME (2005) The Legume Information System (LIS): an integrated information resource for comparative legume biology. Nucleic Acids Res 33(Database issue): D660–D665

Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137:1121–1137

Gustine DL, Voigt PW, Brummer EC, Papadopoulos YA (2002) Genetic variation of RAPD markers for North American white clover collections and cultivars. Crop Sci 42:343–347

Gutierrez MV, Vaz Patto MC, Huguet T, Cubero JI, Moreno MT, Torres AM (2005) Cross-species amplification of *Medicago truncatula* microsatellites across three major pulse crops. Theor Appl Genet 110:1210–1217

Herrmann D, Boller B, Studer B, Widmer F, Kolliker R (2006) QTL analysis of seed yield components in red clover (*Trifolium pratense* L.). Theor Appl Genet 112:536–545

Jaiswal P, Ni J, Yap I, Ware D, Spooner W, Youens-Clark K, Ren L, Liang C, Zhao W, Ratnapu K, Faga B, Canaran P, Fogleman M, Hebbard C, Avraham S, Schmidt S, Casstevens TM, Buckler ES, Stein L, McCouch S (2006) Gramene: a bird's eye view of cereal genomes. Nucleic Acids Res 34(Database issue):D717–D723

Jones E, Hughes L, Drayton M, Abberton M, Michaelson-Yeats T, Bowen C, Forster J (2003) An SSR and AFLP molecular marker-based genetic map of white clover (*Trifolium repens* L.). Plant Sci 165:531–539

Julier B, Flajoulot S, Barre P, Cardinet G, Santoni S, Huguet T, Huyghe C (2003) Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. BMC Plant Biol 3:9

Kalo P, Seres A, Taylor SA, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis TH, Kiss GB (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. Mol Genet Genomics 272:235–246

Kölliker R, Jones ES, Jahufer MZZ, Forster JW (2001a) Bulked AFLP analysis for the assessment of genetic diversity in white clover (*Trifolium repens* L.). Euphytica 121:305–315

Kölliker R, Jones ES, Drayton MC, Dupal MP, Forster JW (2001b) Development and characterisation of simple sequence repeat (SSR) markers for white clover (*Trifolium repens* L.). Theor Appl Genet 102:416–424

Kölliker R, Enkerli J, Widmer F (2006) Characterization of novel microsatellite loci for red clover (*Trifolium pratense* L.) from enriched genomic libraries. Mol Ecol Notes 6:50–53

Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol 11:2453–2465

Mian MA, Saha MC, Hopkins AA, Wang ZY (2005) Use of tall fescue EST–SSR markers in phylogenetic analysis of cool-season forage grasses. Genome 48:637–647

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet 30:194–200

Paterson AH, Freeling M, Sasaki T (2005) Grains of knowledge: genomics of model cereals. Genome Res 15:1643–1650

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. Mol Biol Evol 15:1275–1287

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

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

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

Sorrells ME, La Rota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD, Miftahudin, Mahmoud A, Ma X, Gustafson PJ, Qi LL, Echalier B, Gill BS, Matthews DE, Lazo GR, Chao S, Anderson OD, Edwards H, Linkiewicz AM, Dubcovsky J, Akhunov ED, Dvorak J, Zhang D, Nguyen HT, Peng J, Lapitan NL, Gonzalez-Hernandez JL, Anderson JA, Hossain K, Kalavacharla V, Kianian SF, Choi DW, Close TJ, Dilbirligi M, Gill KS, Steber C, Walker-Simmons MK, McGuire PE, Qualset CO (2003) Comparative DNA sequence analysis of wheat and rice genomes. Genome Res 13:1818–1827

Taylor NL, Quarles RF, Anderson MK (1980) Methods of overcoming interspecific barriers in *Trifolium*. Euphytica 29:441–450

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

Wang ML, Gillaspie AG, Newman ML, Dean RE, Pittman RN, Morris JB, Pederson GA (2004) Transfer of simple sequence repeat (SSR) markers across the legume family for germplasm characterization and evaluation. Plant Genet Res 2:107–119

Williams CG, Goodman MM, Stuber CW (1995) Comparative recombination distances among *Zea mays* L. inbreds, wide crosses and interspecific hybrids. Genetics 141:1573–1581

Williams WM (1987) White clover taxonomy and biosystematics. In: Williams WM (ed) White Clover. CAB International, Oxon, pp 323–342

Wipf L, Cooper DC (1938) Chromosome numbers in nodules and roots of red clover, common vetch and garden pea. Proc Natl Acad Sci USA 24:87–91

Zhu H, Choi HK, Cook DR, Shoemaker RC (2005) Bridging model and crop legumes through comparative genomics. Plant Physiol 137:1189–1196