Theor Appl Genet (2002) 105:577–584
DOI 10.1007/s00122-002-0907-3

**E. S. Jones · M. P. Dupal · J. L. Dumsday**
**L. J. Hughes · J. W. Forster**

# An SSR-based genetic linkage map for perennial ryegrass (*Lolium perenne* L.)

Received: 19 November 2001 / Accepted: 17 December 2001 / Published online: 23 May 2002


**Abstract** A simple sequence repeat (SSR)-based linkage map has been constructed for perennial ryegrass (*Lolium perenne* L.) using a one-way pseudo-testcross reference population. A total of 309 unique perennial ryegrass SSR (LPSSR) primer pairs showing efficient amplification were evaluated for genetic polymorphism, with 31% detecting segregating alleles. Ninety-three loci have been assigned to positions on seven linkage groups. The majority of the mapped loci are derived from cloned sequences containing $(CA)_n$-type dinucleotide SSR arrays. A small number (7%) of primer pairs amplified fragments that mapped to more than one locus. The SSR locus data has been integrated with selected data for RFLP, AFLP and other loci mapped in the same population to produce a composite map containing 258 loci. The SSR loci cover 54% of the genetic map and show significant clustering around putative centromeric regions. BLASTN and BLASTX analysis of the sequences flanking mapped SSRs indicated that a majority (84%) are derived from non-genic sequences, with a small proportion corresponding to either known repetitive DNA sequence families or predicted genes. The mapped LPSSR loci provide the basis for linkage group assignment across multiple mapping populations.

**Keywords** Molecular marker · Simple sequence repeats · *Lolium perenne* · Genetic map · Linkage group · DNA sequence analysis

Communicated by G. Wenzel

E.S. Jones · M.P. Dupal · J.L. Dumsday · L.J. Hughes
J.W. Forster (✉)
Plant Biotechnology Centre, Agriculture Victoria,
Department of Natural Resources and Environment,
La Trobe University, Bundoora, Victoria 3083,
and CRC for Molecular Plant Breeding, Australia
e-mail: john.forster@nre.vic.gov.au
Fax: +61-3-94793618

## Introduction

A number of molecular genetic marker systems have been developed for implementation in the breeding of the important pasture and turfgrass species perennial ryegrass (*Lolium perenne* L.) (Forster et al. 2001). An important objective for the improvement of this self-incompatible, outbreeding species is the development and mapping of reproducible, co-dominant genetic markers for quantitative trait locus (QTL) detection and gene tagging. A number of co-dominant isoenzyme, restriction fragment length polymorphism (RFLP) and sequence tagged site (STS) marker loci and dominant amplified fragment length polymorphism (AFLP), loci have been assigned to genetic maps constructed using two different one-way pseudo-testcross populations (Hayward et al. 1998; Bert et al. 1999). An enhanced molecular marker-based linkage map has recently been constructed using the p150/112 population (Jones et al. 2002), containing 109 RFLP loci detected by heterologous probes from other Poaceae species. Comparative genetic mapping has allowed the alignment of the perennial ryegrass genetic map with those of wheat, rice and oat, revealing substantial conserved synteny with the genomes of Triticeae species (Jones et al. 2002).

The enrichment of the perennial ryegrass genetic map with co-dominant RFLP loci has provided a set of framework markers which are capable of transfer across different mapping populations, allowing cross-referencing of map positions. However, RFLP technology is relatively expensive, time-consuming and labour intensive, thereby limiting the applicability of this approach. The ideal marker system for this purpose is simple sequence repeat (SSR) polymorphism. SSRs are reliable, polymerase chain reaction (PCR)-based markers that detect stable co-dominant multiallelic loci (Rafalski et al. 1996). SSR detection is capable of automation, allowing high-throughput genotyping and map construction (Mitchell et al. 1997). SSR-based linkage maps have been constructed for a range of important crops such as rice (Chen et al. 1997; Temnykh et al. 2000), soybean (Cregan et al.

1999), potato (Milbourne et al. 1998), sugar beet (Rae et al. 2000), sorghum (Kong et al. 2000), maize (Taramino and Tingey 1996) and wheat (Röder et al. 1998; Stephenson et al. 1998). SSR map development in forage species has so far only been reported for alfalfa (Diwan et al. 2000).

We have recently reported the characterisation of a set of 366 unique SSR clones from perennial ryegrass (Jones et al. 2001). The perennial ryegrass SSR (LPSSR) loci were evaluated for efficiency of amplification, polymorphism in a small panel of diverse genotypes and cross-amplification in other closely related taxa. In order to further enhance the reference map of perennial ryegrass, LPSSR loci have been assigned to positions on the current *L. perenne* reference map. A framework set of 93 loci are distributed across the seven linkage groups, providing the basis for future mapping in trait-specific populations. The cloned sequences from mapped LPSSR loci have been subjected to DNA sequence analysis in order to evaluate possible associations with genic regions.

## Materials and methods

### Plant material

The p150/112 reference mapping population was derived from the cross between a multiple heterozygous parent of complex descent as pollinator and a doubled haploid as female parent (Bert et al. 1999; Jones et al. 2002). Plant material was provided by Dr. Mervyn Humphreys and Mr. Gilbert Jones, IGER, UK. Genomic DNA was extracted from 155 $F_1$ progeny individuals by the 1× CTAB method of Fulton et al. (1995).

### Evaluation of SSR polymorphism

Assessment of genetic polymorphism was performed by separation of radiolabelled PCR products using vertical denaturing acrylamide gel electrophoresis and phosphorimager detection as described by Jones et al. (2001). The screening panel consisted of the multiple heterozygous parent and six $F_1$ progeny from the p150/112 population.

### Genetic mapping of SSR loci

Segregation data for up to 155 $F_1$ individuals was either obtained as described for polymorphism screening or by automated capillary electrophoresis of fluorochrome-labelled PCR products. LPSSR forward primers were modified with one of three fluorochrome moieties – 6-carboxyfluorescein (FAM), hexachloro-6-carboxyfluorescein (HEX), NED (Applied Biosystems, Foster City, Calif.). PCR amplifications were performed using one of four touchdown programs (Jones et al. 2001), depending on primer pair $T_m$ value, in ABI9700 thermocyclers (Applied Biosystems). Products labelled with one of each of the three dyes were pooled using a BioRobot 9600 liquid handling system (QIAGEN, Valencia, Calif.). The triplexed products were separated using an ABI3700 96-channel DNA sequencer (Applied Biosystems) and sized using the GeneScan–500 ROX size standard (Applied Biosystems; ROX, 6-carboxy-X-rhodamine) and the programme GENESCAN (Applied Biosystems). GENOTYPER 3.5 (Applied Biosystems) was used to score SSR alleles.

SSR marker data were mapped separately and in combination with the RFLP and AFLP data described by Jones et al. (2002) using MAPMAKER 3.0 (Lander et al. 1987). Linkage groups were assigned at LOD 5.0, with one linkage group being split at LOD 9.0. RFLP and AFLP markers that did not map at LOD>2.0 were excluded form the combined map data. Map distances were calculated using the Kosambi mapping function (Kosambi 1944).

The RFLP and AFLP dataset for the p150/112 reference genetic mapping family (Jones et al. 2002) is available at http://ukcrop.net/perl/ace/search/FoggDB.

### DNA sequence analysis of SSR loci

Cloned insert sequences were analysed using the BLASTN nucleotide sequence alignment algorithm provided as part of the suite of molecular biology tools at the National Centre for Biotechnology Information (NCBI) Internet site (http://www.ncbi.nlm.nih.gov). Sequence searches were conducted against the GenBank database. The E value of the highest sequence match (probability of sequence matching by chance) was determined for each query sequence, along with the length of the aligned segment and accession identity. The BLASTX protein alignment algorithm was used for further analysis of cloned sequences matching known genic regions.

## Results

### Evaluation of SSR polymorphism in reference population

In addition to the 366 unique non-truncated LPSSR clones previously reported (Jones et al. 2001), a further 960 clones were sequenced from the LPSSRH library, which was enriched primarily for $(CA)_n$ repeats and produced more polymorphic loci than the multiplex-enriched LPSSRK library (Jones et al. 2001). A total of 288 extra SSR-containing clones were identified, of which 84 were unique and non-truncated. Therefore, a total of 450 LPSSR clones suitable for locus characterisation were identified. Primer pairs were successfully designed for 391 unique clones, of which 328 (84%) produced clear and simple amplification products in the expected size range. The amplification efficiency varied from 78% to 100%, depending on SSR class (Table 1). Of the primer pairs showing efficient amplification, 31% detected genetic polymorphism in the p150/112 reference population (Table 1). Lower levels of genetic polymorphism were detected with interrupted/imperfect SSRs, trinucleotide SSRs and $n<10$ repeats, all of which are prevalent in the LPSSRK library.

### Molecular mapping data

The genetic mapping dataset was constructed using data from 89 of the 101 primer pairs that detected genetic polymorphism. Data generated by the remaining 12 LPSSR primer pairs was excluded on the basis of weak or equivocal amplification signals or inefficient allele discrimination. Of the 89 primer pairs, 83 produced single-locus amplification patterns, while six detected more than one locus, with five detecting duplicate loci and one detecting a triplicated locus, giving a total of 96 segregating markers. A majority of the data was generated using the automated fluorescence detection system (84 primer pairs

**Table 1** Amplification and polymorphism data for the p150/112 reference population obtained with LPSSR primer pairs

| SSR category | Number of primer pairs | Number (and percentage) of primer pairs with efficient amplification | Number (and percentage) of polymorphic primer pairs |
|---|---|---|---|
| Perfect SSRs | 191 | 154 (81%) | 61 (40%) |
| Interrupted SSRs | 114 | 89 (78%) | 23 (26%) |
| Imperfect SSRs | 59 | 53 (90%) | 12 (23%) |
| Compound SSRs | 27 | 22 (81%) | 5 (23%) |
| Dinucleotide repeats | 263 | 215 (82%) | 74 (34%) |
| Trinucleotide repeats | 124 | 99 (80%) | 26 (26%) |
| Tetranucleotide or higher repeats | 4 | 4 (100%) | 1 (25%) |
| $n < 6$ repeats | 32 | 27 (84%) | 6 (22%) |
| $6 \leq n < 10$ repeats | 194 | 163 (84%) | 47 (29%) |
| $n \geq 10$ repeats | 165 | 128 (78%) | 48 (38%) |
| LPSSRH library | 106 | 90 (85%) | 33 (37%) |
| LPSSRK library | 283 | 226 (80%) | 67 (30%) |
| All SSRs | 391 | 328 (84%) | 101 (31%) |

detecting 91 loci), while a minority (five primer pairs detecting five loci) was produced using the isotopic detection system.

Co-dominant genetic markers are expected to show either AB × BB or AB × CC segregation structures in the p150/112 family. A total of 61 AB × BB and 21 AB × CC patterns were observed. In addition, a number of patterns consistent with the detection of null alleles were observed (7 A0 × 00; 6 A0 × BB; 1 AB × 00). Loci with the first two segregation patterns were treated as dominant markers in the p150/112 $F_1$ progeny set, while the locus in the third class was mapped as a co-dominant marker.

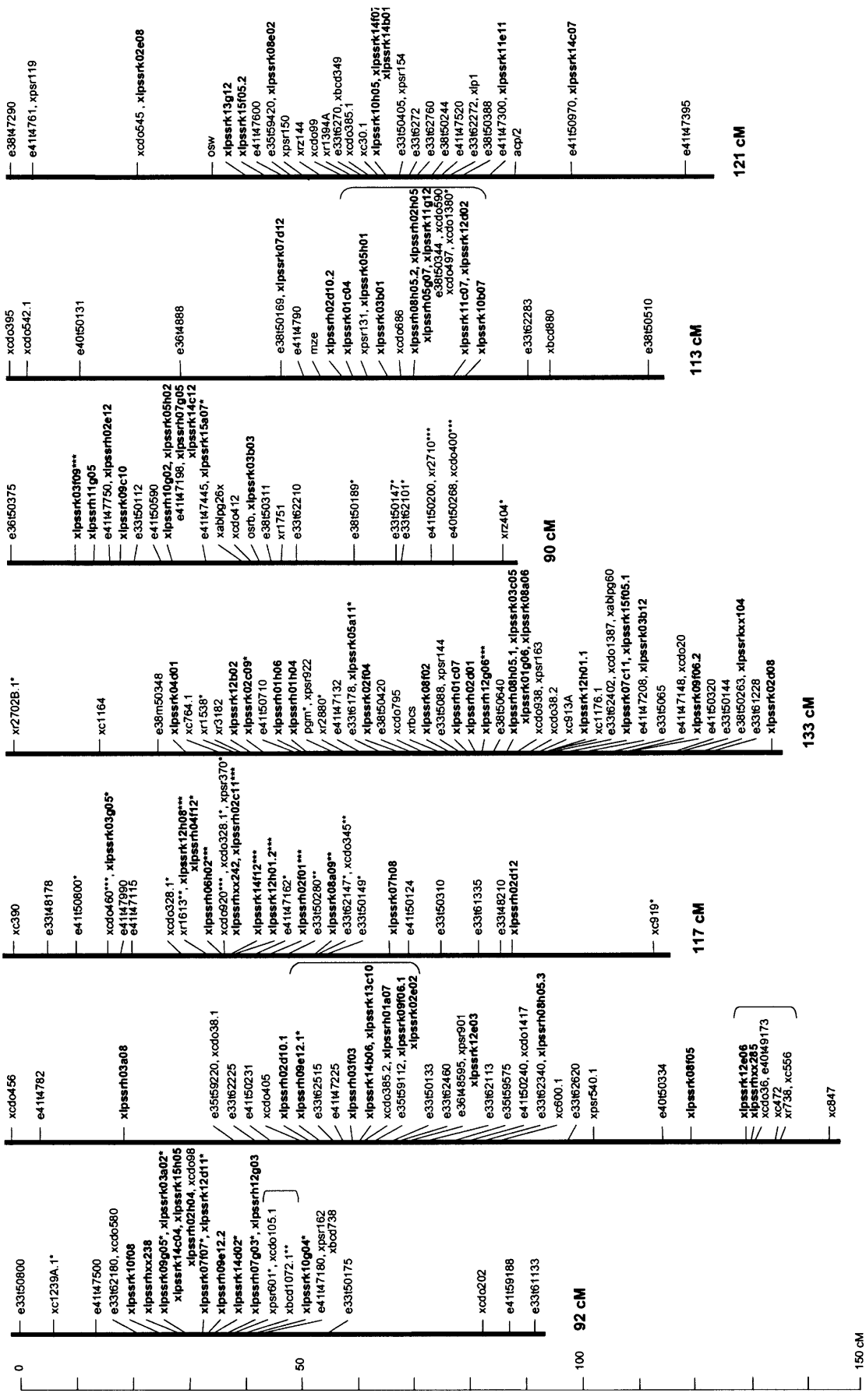
## Genetic map construction

The 96 LPSSR loci were mapped in combination with the RFLP, STS, isoenzyme and AFLP data described by Jones et al. (2002). To generate a more robust map, we excluded RFLP and AFLP loci that did not fit at LOD>2.0 from the map, resulting in the total number of these markers being reduced from 240 to 165 (30 AFLP loci and 45 co-dominant loci were eliminated). Most of the excluded markers were located around putative centromeres in regions of locus clustering, and their removal did not affect map coverage. There was no change in the order of markers common to this study and that of Jones et al. (2002).

Ninety-three LPSSR loci were assigned to seven linkage groups (LGs), with between 9 and 22 loci per group (Fig. 1). Three primer pairs detected single loci that could not be assigned to the genetic map. Two of these were scored as dominant markers and were relatively difficult to score accurately, while the third showed a high level of segregation distortion. The composite map contains 74 RFLP loci, five isoenzyme and STS loci and 86 AFLP loci in addition to the 93 SSR loci, giving a total of 172 markers of the generally co-dominant class. These loci cover 814 cM, as compared to 811 cM for the map of Jones et al. (2002). Relatively little map expansion has therefore occurred by addition of the SSR markers. The use of only the most robust RFLP and AFLP markers resulted in a reduction in the number of unordered regions compared with the study of Jones et al. (2002). Only four regions on LGs 1, 2 and 6 remained unordered at LOD>2.0, compared with 15 regions in the map of Jones et al. (2002). These four unordered regions covered distances of between 2 cM and 17 cM. Twenty-four percent (22/93) of all mapped LPSSR markers showed distorted segregation at $P<0.05$, and 10% (9/93) at $P<0.01$. These values are similar to the levels of skewness (22% at $P<0.05$ and 8% at $P<0.01$) described by Jones et al. (2002). The majority of LPSSR markers skewed at $P<0.01$ mapped to LG3 in a region previously shown to be associated with high levels of segregation distortion (Jones et al. 2002).

The LPSSR markers were also mapped independently of the RFLP and AFLP data. The same linkage groups were formed, except for LG2, in which one terminal locus (xlpssrh03a08) was unlinked and three terminal loci (xlpssrk08f05, xlpssrk12e06, xlpssrhxx285) formed a separate linkage group. LPSSR marker orders were unchanged. The total map distance using only LPSSR loci was 325 cM, with an average marker interval of 4 cM and a range of 0–20 cM.

The distribution of LPSSR loci within linkage groups was variable, with substantial clustering around the putative centromeric regions, especially on linkage groups 1, 2 and 6. The LPSSR loci are distributed across 54% of the total map distance, compared with 85% for AFLPs and 82% for RFLPs, STSs and isoenzymes. The distribution of different classes of SSR across the genetic map was examined in terms of map coverage and tendency to cluster. SSRs were classed by motif type (64 primer pairs were designed to dinucleotide repeat SSR loci, 21 to trinucleotide repeat loci and one to a tetranucleotide repeat locus), structure (55 corresponded to perfect repeat types, four to compound perfect repeats, 17 to interrupted SSRs and ten to imperfect repeats) and repeat number (varying between 5 and 31). Although the number of mapped trinucleotide repeat loci (21) is considerably


LG1
LG2
LG3
LG4
LG5
LG6
LG7
0
50
100
150 cM
92 cM
148 cM
117 cM
133 cM
90 cM
113 cM
121 cM
e33t50800
xc1239A.1*
e41t47500
e33t62180, xcdo580
xlpssrk10f08
xlpssrhxx238
xlpssrk09g05*, xlpssrk03a02*
xlpssrk14c04, xlpssrk15h05
xlpssrh02h04, xcdo98
xlpssrk07f07*, xlpssrk12d11*
xlpssrh09e12.2
xlpssrk14d02*
xlpssrh07g03*, xlpssrh12g03
xpsr601*, xcdo105.1
xbcd1072.1**
xlpssrk10g04*
e41t47180, xpsr162
xbcd738
e33t50175
xcdo202
e41t59188
e33t61133
xcdo456
e41t4782
xlpssrh03a08
e35t59220, xcdo38.1
e33t62225
e41t50231
xcdo405
xlpssrh02d10.1
xlpssrh09e12.1*
e33t62515
e41t47225
xlpssrh03f03
xlpssrk14b06, xlpssrk13c10
xcdo385.2, xlpssrh01a07
e35t59112, xlpssrk09f06.1
xlpssrk02e02
e33t50133
e33t62460
e36t48595, xpsr901
xlpssrk12e03
e33t62113
e35t59575
e41t50240, xcdo1417
e33t62340, xlpssrh08h05.3
xc600.1
e33t62620
xpsr540.1
e40t50334
xlpssrk08f05
xlpssrk12e06
xlpssrhxx285
xc472
xr738, xc556
xc847
xc390
e33t48178
e41t50800*
xcdo460***, xlpssrk03g05*
e41t47990
e41t47115
xcdo328.1*
xr1613**, xlpssrk12h08***
xlpssrh04f12*
xlpssrh06h02***
xcdo920***, xcdo328.1*, xpsr370*
xlpssrhxx242, xlpssrh02c11***
xlpssrk14f12***
xlpssrk12h01.2***
e41t47162*
xlpssrh02f01***
e33t50280**
xlpssrk08a09**
e33t62147*, xcdo345**
e33t50149*
xlpssrk07h08
e41t50124
e33t50310
e33t61335
e33t48210
xlpssrh02d12
xc919*
xr27028.1*
xc1164
e38m50348
xlpssrk04d01
xc764.1
xr1538*
xr3182
xlpssrk12b02
xlpssrk02c09*
e41t50710
xlpssrh01h06
xlpssrh01h04
pgm*, xpsr922
xr2880*
e41t47132
e33t6178, xlpssrk05a11*
xlpssrk02f04
e38t50420
xcdo795
xrbcs
xlpssrk08f02
e33t5088, xpsr144
xlpssrh01c07
xlpssrh02d01
xlpssrh12g06***
e38t50640
xlpssrh08h05.1, xlpssrk03c05
xlpssrk01g06, xlpssrk08a06
xcdo938, xpsr163
xcdo38.2
xc913A
xlpssrk12h01.1
xc1176.1
e33t62402, xcdo1387, xablpg60
xlpssrk07c11, xlpssrk15f05.1
e41t47208, xlpssrk03b12
e33t5065
e41t47148, xcdo20
xlpssrk09f06.2
e41t50320
e33t50144
e38t50263, xlpssrkxx104
e33t61228
xlpssrk02d08
e36t50375
xlpssrk03f09***
xlpssrh11g05
e41t47750, xlpssrh02e12
xlpssrk09c10
e33t50112
e41t50590
xlpssrh10g02, xlpssrk05h02
e41t47198, xlpssrh07g05
xlpssrk14c12
e41t47445, xlpssrk15a07*
xablpg26x
xcdo412
osrb, xlpssrk03b03
e38t50311
xr1751
e33t62210
e38t50189*
e33t50147*
e33t62101*
e41t50200, xr2710***
e40t50268, xcdo400***
xr2404*
xcdo395
xcdo542.1
e40t50131
e36t4888
e38t50169, xlpssrk07d12
e41t4790
mze
xlpssrh02d10.2
xlpssrk01c04
xpsr131, xlpssrk05h01
xlpssrk03b01
xcdo686
xlpssrh08h05.2, xlpssrh02h05
xlpssrh05g07, xlpssrk11g12
e38t50344, xcdo590
xcdo497, xcdo1380*
xlpssrk11c07, xlpssrk12d02
xlpssrk10b07
e33t62283
xbcd880
e38t50510
e38t47290
e41t4761, xpsr119
xcdo545, xlpssrk02e08
osw
xlpssrk13g12
xlpssrk15f05.2
e41t47600
e35t59420, xlpssrk08e02
xpsr150
xrz144
xcdo99
xr13944
e33t62270, xbcd349
xcdo385.1
xc30.1
xlpssrk10h05, xlpssrk14f07
xlpssrk14b01
e33t50405, xpsr154
e33t62272
e33t62760
e38t50244
e41t47520
e33t62272, xlp1
e38t50388
e41t47300, xlpssrk11e11
acp/2
e41t50970, xlpssrk14c07
e41t47395

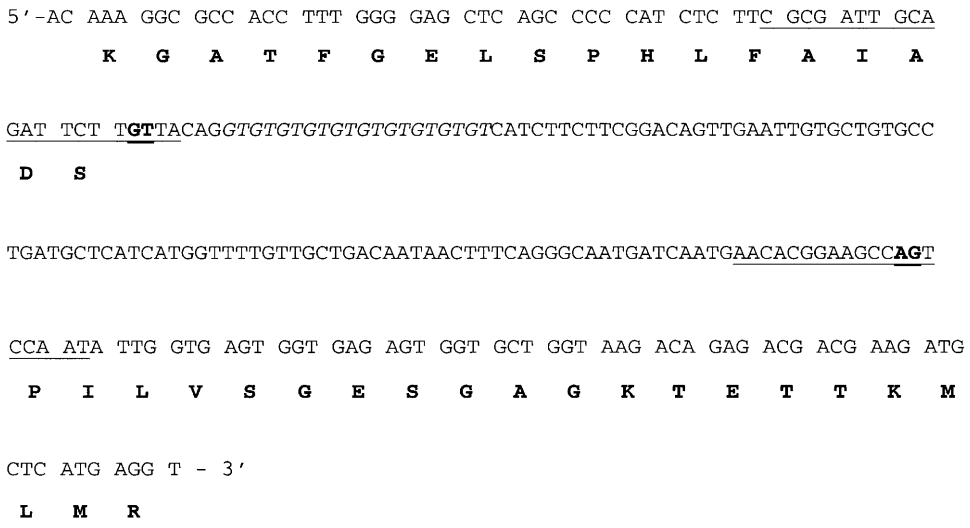



**Fig. 2** Location of the SSR array [$(GT)_{10}$] of the cloned sequence LPSSRH01H06 in a predicted intron of a heavy-chain myosin gene. The SSR array is shown in *italics*. The conserved splice junction dinucleotide motifs are shown in *bold type*. The positions of amplification primer sequences for detection of the corresponding SSR locus are *underlined*. The predicted primary structure of the corresponding 5′- and 3′-exon sequences is shown in standard amino acid code in *bold type below* the sequence

lower than the number of dinucleotide repeat loci (64), the map coverage was not proportionally less (33% and 45%, respectively). The intervals between trinucleotide repeat loci were larger (17 cM compared to 6 cM) and markers were less clustered than for dinucleotide repeat loci. To examine the distribution of repeat length, we divided the genetic map into 10 cM segments and calculated the average value for repeat numbers of loci in each segment, but no correlation with map location was observed. However, out of the 16 SSRs with repeat numbers greater than 15, 13 were found to map close to putative centromeric regions. No particular association between SSR structure (perfect, imperfect, interrupted or compound) and genetic map coverage was observed.

### DNA sequence analysis of SSR loci

The LPSSR clones were derived from genomic enrichment libraries and may hence be derived from both genic and non-genic regions of the perennial ryegrass genome. Sequence alignments determined that the majority (84%) of clones corresponding to mapped loci produced low level ($E>10^{-2}$) matches of short (<20 bp) sequences to current database accessions. No significant matches to chloroplast DNA sequences were obtained. A single match ($E = 3\times10^{-6}$) was found to a known repetitive DNA sequence class, corresponding to a portion of the long terminal repeat (LTR) of the barley BARE-1 retroelement (Manninen and Schulman 1993). Six (7%) cloned sequences produced high-level matches ($E<10^{-8}$). Of these, three clones matched genomic sequences derived from rice PAC clones. Further analysis with BLASTX revealed no current gene annotation for one of these sequences, while the other two sequences matched a hypothetical *Arabidopsis thaliana* L. protein and a putative *A. thaliana* GDSL-motif lipase/hydroxylase-like protein. The other three clone sequences were matched to an unknown predicted *A. thaliana* protein ($E = 4\times10^{-19}$), an oat (*Avena sativa* L.) gliadin-like avenin protein ($E = 2\times10^{-10}$) and an *A. thaliana* heavy-chain myosin gene ($E = 2\times10^{-9}$). BLASTX analysis demonstrated that the $(GT)_n$ SSR in this cloned sequence is located within a small intron, close to the boundary with the 5′-exon (Fig. 2).

The RFLP- and AFLP-based reference genetic map of perennial ryegrass has allowed the prediction of conserved syntenic relationships with other Poaceae species such as wheat, barley and rice (Jones et al. 2002). The three rice sequence accessions with high matches to LPSSR clone sequences were derived from genomic clones of known chromosomal origin, providing a means of testing predicted locus locations. Locus xlpssrk12d11 is located on perennial ryegrass linkage group (LG) 1 in a region of conserved synteny with rice chromosome 5. The matching rice PAC clone is also derived from chromosome 5. The same situation applies to xlpssrh06h02, on LG3 in a region of conserved synteny with rice chromosome 1. However, xlpssrh07g03 on LG1 corresponds to rice chromosome 10 instead of 5.

◀ **Fig. 1** Distribution of 93 loci detected by perennial ryegrass SSR (LPSSR) primer pairs (in *bold type*) across the seven linkage groups of perennial ryegrass for the cross p150/112. Segregation data is based on LPSSR loci and a selected sub-set of RFLP and AFLP loci from the reference map of Jones et al. (2002). Locus nomenclature for LPSSR loci is shown in the form xlpssrabbccc, where a indicates the enrichment library of origin (*h* or *k*), bb indicates sequencing block number and ccc indicates block coordinate. Multiplicated loci are shown with a *numerical extension* – e.g. lpssrk09f06.1 – with *.1* indicating the locus was mapped with fragments in the expected size range, and *.2* and *.3* with fragments not in the expected size range. Markers showing segregation distortion are indicated by * = significant distortion at $P < 0.05$, ** = $P < 0.01$ and *** = $P < 0.001$. The *bracketed* regions indicate markers that could not be ordered at LOD>2

## Discussion

The results presented here show the construction of a framework genetic linkage map of perennial ryegrass

based on SSR loci, relating the distribution of these loci to the previously generated RFLP- and AFLP-based reference map of Jones et al. (2002). The p150/112 $F_1$ population (Bert et al. 1999) provided the genetic resource for reference map development (Forster et al. 2001), and the further mapping of SSR loci in this population provides a significant enhancement of map information, with over 150 genetically co-dominant markers distributed across the seven linkage groups.

The quantification of SSR polymorphism in the p150/112 population allowed an estimate of amplification efficiency following primer design over a larger number of LPSSR clones (391) than the 101 clones evaluated by Jones et al. (2001). A comparable amplification efficiency (84% compared to 81%) was observed in this study. The proportion of SSR primer pairs showing efficient amplification that detected segregating alleles was 31%, with variation according to SSR structural categories, repeat motif lengths, array lengths and library of origin. As observed for a panel of diverse perennial ryegrass genotypes (Jones et al. 2001), higher levels of polymorphism were observed for long, perfect and dinucleotide repeat types compared to shorter, interrupted, imperfect and trinucleotide repeat types. Although both diallelic (AB×BB) and triallelic (AB×CC) segregation patterns were observed, the first class was in excess. This finding is perhaps surprising, as the p150/112 population was designed to maximise genetic polymorphism, and the parentage of the multiply heterozygous parent is quite distinct from the doubled haploid parent DH290 (derived from the anther culture-responsive Danish cultivar Verna). The limited allelic variation detected here may indicate a lower level of genetic variability within perennial ryegrass than anticipated, perhaps correlated with patterns of gene flow during the colonisation of Europe by the species (Balfourier et al. 2000) or restricted use of germplasm during varietal development.

A small proportion of the mapped loci showed segregation patterns consistent with the presence of null alleles. Such alleles have been reported in other studies (Uzunova and Ecke 1999; Kong et al. 2000) and may arise due to hemizygosity (large-scale deletion of sequences flanking the SSR locus) or, more likely, from point mutations or small insertion/deletion events in the region matching one or other amplification primers. Many of the LPSSR clones showed a high level of cryptic simplicity (Tautz et al. 1986) in the 5′- and 3′-flanking regions, suggesting that rapid sequence divergence is prevalent in these areas. The regions directly adjacent to SSR arrays may show elevated levels of base mutation and may be suitable for the identification of single nucleotide polymorphism markers (Brohede and Ellergren 1999).

A small number of LPSSR primer pairs (6) detected multiple loci, generally duplicated, but in one instance triplicated. The locations of these multiplicated loci are not consistent with large-scale internal duplications of the perennial ryegrass genome. A similar observation was made for duplicated RFLP loci detected by heterologous anchor probes (Jones et al. 2002). However, smaller scale duplications cannot be excluded at the current level of marker coverage. Although none of the multiplicated loci show significant sequence matches to known genes, it is possible that they are associated with paralogous sequences corresponding to small multigene families. The presence of multiplicated loci enhances the number of marker positions generated by the current set of SSR primer pairs but may produce inconsistencies between linkage group assignments when mapping is performed in other segregating populations. For this reason, the primer pairs detecting multiplicated loci should be preferentially excluded from framework mapping in subsequent populations.

The SSR loci show irregular distribution across and within the seven linkage groups. The largest (22) and best distributed set of loci are on LG4, while the 14 loci on LG1 are highly clustered around the putative centrometic location. Clustering on genetic maps of SSR loci isolated from genomic libraries has been reported for a number of species including barley (Ramsay et al. 2000) and tomato (Areshchenkova and Ganal 1999) but is not so prevalent in other species such as sorghum (Bhattramaki et al. 2000) or rice (Chen et al. 1997). Centromeric clustering of dinucleotide repeat SSRs in tomato was characteristic of loci containing large arrays. For the current perennial ryegrass SSR-based map, the average repeat number of the cloned SSR arrays for the most clustered group of loci on LG1 ($n$ = 11.8 for 14 loci over 20 cM) was similar to that of the least clustered group of loci on LG4 ($n$ = 10.1 for 22 loci over 80 cM; data not shown), and the average repeat number across 10 cM segments did not significantly vary across the genetic map. However, the largest SSRs with repeat numbers greater than 15 did tend to map to putative centromeric regions, suggesting that the trend observed in tomato may occur to some extent in perennial ryegrass. Trinucleotide repeat loci may be less likely to cluster as they are more commonly detected in expressed sequence tags and therefore genic regions (Scott et al. 2000; Cordeiro et al. 2001). In this study, trinucleotide repeat loci showed a lower level of clustering around the putative centromeres than dinucleotide repeat loci and gave a similar level of genetic map coverage, despite fewer being mapped.

Although the currently mapped SSR loci provide a powerful means for map alignment across different populations, the irregular distribution and total map coverage of approximately 54% are sub-optimal for whole genome scanning. Map coverage will be expanded through the use of additional LPSSR loci that are polymorphic in other populations.

The majority (84%) of the mapped genomic LPSSR loci produce sequence matches of low significance when screened against genomic databases. The failure to identify chloroplast-derived sequences is not unexpected as such loci would show maternal inheritance and show low levels of within-species polymorphism (Weising and Gardner 1999). There was only one match to a known repetitive DNA sequence (the barley BARE-1 element). SSR loci have been demonstrated to frequently co-locate

with dispersed repetitive sequences in species such as barley (Ramsay et al. 1999), but the LPSSR sequences do not show an association with known repeat families at higher levels than expected by chance.

BLASTN and BLASTX analysis has demonstrated that a small number of the LPSSR clone sequences are associated with either hypothetical or actual gene sequences. The SSRs are likely to be located in the untranslated regions of these genes, or, as for LPSSRH01H06 which matches the myosin heavy chain gene sequences, in an intron. The SSR structure for this clone has previously been shown to be conserved across species boundaries as far as oat (Jones et al. 2001), suggesting a relatively ancient evolutionary origin.

We have also used sequence similarity with rice genomic clones to test inferences of conserved synteny. Although two of three annotations are consistent with the current comparative map (Jones et al. 2002), a third does not conform to this pattern. Similar interruptions of conserved synteny have been observed for RFLP loci and may be related to small-scale genomic rearrangements or polymorphism in paralogous gene sequences in different chromosomal locations. However, one of the basic rearrangements which define the relationship between the Triticeae and rice genomes involves the insertion of rice chromosome 10 into chromosome 5 to form the Triticeae group 1 chromosomes (Devos and Gale 1997). Locus xlpssrh07go3 may provide evidence for such a rearrangement in perennial ryegrass LG1, along with a single RFLP marker (xcdo98).

In summary, we have presented here for the first time a framework linkage map of perennial ryegrass based on an efficient, PCR-based, locus-specific, genetically codominant and multiallelic marker system. The LPSSR loci have been mapped along with RFLP and AFLP markers using the reference population for this species, providing a ready means for map alignment and chromosome assignment across multiple mapping populations.

**Acknowledgements** This work was supported by the CRC for Molecular Plant Breeding and the Department of Natural Resources and Environment, Victoria, Australia. The scientific advice and support of Prof. German Spangenberg is gratefully acknowledged. All experiments conducted during this study comply with current Australian laws.

## References

Areshchenkova T, Ganal MW (1999) Long tomato microsatellites are predominantly associated with centromeric regions. Genome 42:536–544

Balfourier F, Imbert C, Charmet G (2000) Evidence for phylogeographic structure in *Lolium* species related to the spread of agriculture in Europe. A cpDNA study. Theor Appl Genet 101:131–138

Bert PF, Charmet G, Sourdille P, Hayward MD, Balfourier F (1999) A high-density molecular map for ryegrass (*Lolium perenne*) using AFLP markers. Theor Appl Genet 99:445–452

Bhattramaki D, Dong J, Chhabra AK, Hart GE (2000) An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. Genome 43:988–1002

Brohede J, Ellegren H (1999) Microsatellite evolution: polarity of substitutions within repeats and neutrality of flanking sequences. Proc R Soc London Ser B 266:825–833

Chen X, Temnykh S, Xu Y, Cho YG, McCouch SR (1997) Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). Theor Appl Genet 95:553–567

Cordeiro GM, Casu R, McIntyre CL, Manners JM, Henry RJ (2001) Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. Plant Sci 160:1115–1123

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of soybean genome. Crop Sci 39:1464–1490

Devos KM, Gale MD (1997) Comparative genetics in the grasses. Plant Mol Biol 35:3–15

Diwan N, Bouton JH, Kochert G, Cregan PB (2000) Mapping of simple sequence repeat (SSR) DNA markers in diploid and tetraploid alfalfa. Theor Appl Genet 101:165–172

Forster JW, Jones ES, Kölliker R, Drayton MC, Dumsday J, Dupal MP, Guthridge KM, Mahoney NL, van Zijll de Jong E, Smith KF (2001) Development and Implementation of Molecular Markers for Forage Crop Improvement. In: Spangenberg G (ed) Molecular breeding of forage crops. Kluwer, Dordrecht, pp 101–133

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13:207–209

Hayward MD, Forster JW, Jones JG, Dolstra O, Evans C, McAdam NJ, Hossain KG, Stammers M, Will JAK, Humphreys MO, Evans GM (1998) Genetic analysis of *Lolium*. I. Identification of linkage groups and the establishment of a genetic map. Plant Breed 117:451–455

Jones ES, Dupal MP, Kölliker R, Drayton MC, Forster JW (2001) Development and characterisation of simple sequence repeat (SSR) markers for perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 102:405–415

Jones ES, Mahoney NL, Hayward MD, Armstead IP, Jones JG, Humphreys MO, Kishida T, Yamada T, Balfourier F, Charmet G, Forster JW (2002) An enhanced molecular marker-based genetic map of perennial ryegrass (*Lolium perenne* L.) reveals comparative relationships with other Poaceae genomes. Genome 45:282–295

Kong L, Dong J, Hart GE (2000) Characteristics, linkage map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs). Theor Appl Genet 101:438–448

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary linkage maps of experimental and natural populations. Genomics 1:174–181

Manninen I, Schulman HA (1993) BARE-1, a copia-like retroelement in barley (*Hordeum vulgare* L.). Plant Mol Biol 22: 829–846

Milbourne D, Meyer RC, Collins AJ, Ramsey LD, Gebhardt C, Waugh R (1998) Isolation, characterisation and mapping of simple sequence repeat loci in potato. Mol Gen Genet 259: 233–245

Mitchell SE, Kresovich S, Jester CA, Hernandez CJ, Szewc-McFadden AK (1997) Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. Crop Sci 37:617–624

Rae SJ, Aldam C, Dominguez I, Hoebrechts M, Barnes SR, Edwards KJ (2000) Development and incorporation of microsatellite markers into the linkage map of sugar beet (*Beta vulgaris* spp.). Theor Appl Genet 100:1240–1248

Rafalski JA, Vogel JM, Morgante M, Powell W, Andre C, Tingey SV (1996) Generating and using DNA markers in plants. In: Birren B, Lai E (eds) Non-mammalian genomic

analysis: a practical guide. Academic Press, San Diego, pp 75–135
Ramsay L, Macaulay M, Cardle L, Morgante M, degli Ivanissevich S, Maestri E, Powell W, Waugh R (1999) Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plant J 17:415–425
Ramsay L, Macaulay M, McLean K, Fuller J, Edwards K, Turesson S, Morgante M, degli-Ivanissivich S, Marmiroli N, Maestri E, Massari A, Powell W, Waugh R (2000) A simple sequence repeat-based linkage map of barley. Genetics 156: 1997–2005
Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149:2007–2023
Scott KD, Eggler P, Seaton G, Rossetto M, Ablett EM, Lee LS, Henry RJ (2000) Analysis of SSRs derived from grape ESTs. Theor Appl Genet 100:723–726
Stephenson P, Bryan G, Kirby J, Collins A, Devos K, Busso C, Gale M (1998) Fifty new microsatellite loci for the wheat genetic map. Theor Appl Genet 97:946–949
Taramino G, Tingey S (1996) Simple sequence repeats for germplasm analysis and mapping in maize. Genome 39:277–287
Tautz D, Trick M, Dover GA (1986) Cryptic simplicity in DNA is a major source of genetic variation. Nature 322:652–656
Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000) Mapping and genome organisation of microsatellite sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:697–712
Uzunova MI, Ecke W (1999) Abundance, polymorphism and genetic mapping of microsatellites in oilseed rape (*Brassica napus* L.). Plant Breed 118:323–326
Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledenous angiosperms. Genome 42:9–19