A framework linkage map of perennial ryegrass based on SSR markers

G.P. Gill, P.L. Wilcox, D.J. Whittaker, R.A. Winz, P. Bickerstaff, C.E. Echt, J. Kent, M.O. Humphreys, K.M. Elborough, and R.C. Gardner

Abstract: A moderate-density linkage map for *Lolium perenne* L. has been constructed based on 376 simple sequence repeat (SSR) markers. Approximately one third (124) of the SSR markers were developed from GeneThresher® libraries that preferentially select genomic DNA clones from the gene-rich unmethylated portion of the genome. The remaining SSR marker loci were generated from either SSR-enriched genomic libraries (247) or ESTs (5). Forty-five percent of the GeneThresher SSRs were associated with an expressed gene. Unlike EST-derived SSR markers, GeneThresher SSRs were often associated with genes expressed at a low level, such as transcription factors. The map constructed here fulfills 2 definitions of a "framework map". Firstly, it is composed of codominant markers to ensure map transferability either within or among species. Secondly, it was constructed to achieve a level of statistical confidence in the support-for-order of marker loci. The map consists of 81 framework SSR markers spread over 7 linkage groups, the same as the haploid chromosome number. Most of the remaining 295 SSR markers have been placed into their most likely interval on the framework map. Nine RFLP markers and 1 SSR marker from another map constructed using the same pedigree were also incorporated to extend genome coverage at the terminal ends of 5 linkage groups. The final map provides a robust framework with which to conduct investigations into the genetic architecture of trait variation in this commercially important grass species.

Key words: Framework map, perennial ryegrass, SSR, simple sequence repeat, GeneThresher, *Lolium perenne*.

Résumé : Une carte génétique de densité moyenne a été produite pour le *Lolium perenne* L. à l'aide de 376 microsatellites (SSR). Environ un tiers (124) des microsatellites ont été développés à partir de banques GeneThresher qui sont enrichies en clones d'ADN génomique provenant des régions hypométhylées et riches en gènes au sein du génome. Les autres microsatellites ont été obtenus soit de banques génomiques enrichies en SSR (247) ou d'EST (5). Quarante-cinq pour cent des microsatellites GeneThresher étaient associés à un gène qui s'exprime. Contrairement aux microsatellites dérivés d'EST, ceux provenant de la banque GeneThresher étaient souvent associés à des gènes exprimés à un faible niveau tel que des facteurs de transcription. La carte produite répond à 2 critères définissant une carte de référence. D'abord, elle est composée de marqueurs codominants ce qui assure la transportabilité intra- et interspécifique. Deuxièmement, la carte a été assemblée de façon telle que l'ordre des gènes est appuyé statistiquement. La carte compte 81 marqueurs de référence étalés sur sept groupes de liaison ce qui correspond au nombre de chromosomes. La plupart des 285 autres marqueurs ont été assignés à l'intervalle le plus vraisemblable au sein de la carte de référence. Neuf marqueurs RFLP et 1 microsatellite d'une autre carte produite avec le même pedigree ont été incorporés de manière à accroître la couverture des extrémités de 5 groupes de liaison. La carte finale offre une carte de référence robuste avec laquelle il sera possible de réaliser des études de l'architecture génétique de la variation phénotypique chez cette graminée d'une grande importance commerciale.

Received 12 July 2005. Accepted 28 November 2005. Published on the NRC Research Press Web site at http://genome.nrc.ca on 28 April 2006.

Corresponding Editor: G.J. Scoles.

G.P. Gill,[1] D.J. Whittaker, R.A. Winz,[2] P. Bickerstaff,[3] and K.M. Elborough.[4] ViaLactia Biosciences, PO Box 109-185, Newmarket, Auckland, New Zealand.
P.L. Wilcox[1] and C.E. Echt.[5] Cell Wall Biotechnology Center, New Zealand Forest Research Institute Ltd., Private Bag 3020, Rotorua, New Zealand.
J. Kent.[6] SignaGen, New Zealand Forest Research Ltd., Private Bag 3020, Rotorua, New Zealand.
M.O. Humphreys. Institute of Grassland and Environmental Research, Plas Gogerddan, Aberystwyth, Wales SY23 3DA, United Kingdom.
R.C. Gardner.[1] School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand.

[1]Authors contributed equally to this work.
[2]Present address: Department of Chemistry, Wine Science, University of Auckland, Private Bag 92019, Auckland 1020, New Zealand.
[3]Present address: Infoics Limited, PO Box 83153, Edmonton, Auckland 1230, New Zealand.
[4]Corresponding author (e-mail: kieran.elborough@vialactia.com).
[5]Present address: USDA Forest Service, Southern Institute of Forest Genetics, 23332 Highway 67, Saucier, MS 39574-9344, USA.
[6]Present address: Applied Biosystems, 52 Rocco Drive, Scoresby, Victoria 3179, Australia.

 doi:10.1139/G05-120 © 2006 NRC Canada

Mots clés : carte génétique de référence, ivraie vivace, SSR, répétition de séquences simples, GeneThresher, *Lolium perenne*.

[Traduit par la Rédaction]

Introduction

Perennial ryegrass (*Lolium perenne* L.) is the most important temperate pasture species for providing forage to the dairy, wool, and meat industries and is considered a valuable turf and amenity grass. *Lolium perenne* is a diploid obligate outbreeder that maintains a high degree of genetic diversity in natural and agricultural populations (Roldan-Ruiz et al. 2000). Hybridization with other related species, such as *Lolium multiflorum* (Italian ryegrass), *Festuca arundinacea* (tall fescue), and *Festuca pratensis* (meadow fescue), provides a valuable resource for the introgression of commercially favourable traits.

As with other crops, there has been considerable interest in applying molecular marker technology for marker-assisted selection (MAS) to improve selection efficiency compared with conventional breeding methods. Perennial ryegrass displays continuous phenotypic variation for most target traits that are controlled by quantitative trait loci (QTL). The genome position and the number of QTL accounting for the majority of genetic variation for a trait may be determined through the use of genetic mapping to identify linked marker loci. An important prerequisite is a genetic map that provides adequate genome coverage and ideally uses markers that are transportable across different pedigrees. Several "comprehensive" genetic maps exist for perennial ryegrass that are predominantly based on a combination of restriction fragment length and amplified fragment length polymorphism (RFLP and AFLP) markers (Bert et al. 1999; Armstead et al. 2002; Jones et al. 2002*a*). Simple sequence repeats (SSRs) are favoured over other DNA marker types for the generation of genetic maps as they are easy to use, codominant, highly polymorphic, and are often transportable to other pedigrees. A more recent perennial ryegrass genetic map contains SSR markers derived from ESTs (Faville et al. 2004) that can potentially provide markers that are also functionally associated with trait variation.

The term "framework map" has been used in at least 2 contexts. One context is to describe maps constructed with codominant markers to ensure map transferability either within or among species. Another context that has been used is statistical: here, the order of marker loci on the final map achieves a predetermined level of statistical confidence in the final order of selected marker loci. This concept has been advocated for linkage mapping in humans by Keats et al. (1991), who used the term "support-for-order" to quantify the level of confidence in the final gene order relative to the next most-likely order for the same set of marker loci. A generally accepted figure is a log-likelihood of odds (LOD) difference of 3.0, which corresponds to a selected gene order being 1000 times more likely than the next most-likely order on a group-wise basis. The framework map concept has also been widely applied in the construction of linkage maps in forest tree species (Grattapaglia and Sederoff 1994; Echt and Nelson 1997; Remington et al. 1999; Wilcox et al. 2001). In contrast, comprehensive maps are those that include all loci without applying any support criteria. Since the quality of marker data can vary from one locus to another, inclusion of poorer-quality data can lead to artificial map expansion, and the overestimation of linkage group length and genome size. The ryegrass genetic maps described above are of the comprehensive type and do not meet all the statistical criteria of a framework map concept (Keats et al. 1991; Liu 1998).

Perennial ryegrass has an estimated genome size of 5.66 pg/2C (Arumuganathan et al. 1999). The repetitive DNA content is expected to be in the order of 75%, similar to that reported for maize (San Miguel et al. 1996), which has a comparable genome size. Repetitive DNA poses a challenge to complete genomic sequencing and the alternative EST-based approaches typically miss 40%–50% of genes as a result of their low expression level or cell-type specificity (Bonaldo et al. 1996). A targeted-sequencing strategy for gene-rich regions using methyl-filtration technology of plant genomic libraries (termed GeneThresher® by Orion Genomics, St. Louis, Mo.) removes methylated "junk DNA", leaving only the small unmethylated portion of the genome containing genes (Rabinowicz et al. 1999). GeneThresher libraries are slightly enriched for SSRs (Whitelaw et al. 2003; Bedell et al. 2005), thereby providing an excellent source of genetic markers that are often associated with genes.

This paper describes a robust genetic-linkage map for perennial ryegrass based on 376 SSR markers and meets the criteria of the framework map concept. Density and coverage of this map is such that it should be immediately useful for applications in molecular breeding, as well as for more basic applications such as comparative genomics.

Materials and methods

Mapping pedigree and SSR marker development

The F_2-mapping population used in this experiment was derived by selfing an F_1 hybrid between 2 partially inbred parental lines (produced from 4 generations of selfing). Parental lines were subject to opposing selection for carbohydrate content (Turner et al. 2001) and the same pedigree (including some of the same genotypes used in this study) was used to construct an RFLP-based map (Armstead et al. 2002).

SSR markers were developed from 3 sources. Genetic Identification Services (Chatsworth, Calif.) provided enriched genomic libraries for CA, GA, TAGA, AAG, and TACA repeats using the method described by Jones et al. (2000). Secondly, a perennial ryegrass genomic sequencing project using methyl-filtrated GeneThresher libraries (supplied by Orion Genomics) was mined for appropriate SSRs. Finally, an in-house EST database comprising 12 different libraries prepared from a range of tissue and treatment types was searched for SSRs that were suitable for marker development. SSRs from each library source were evaluated for redundancy and primer design using an in-house-developed bioinformatics pipeline. The cut-off threshold for the number of repeats was at least 8 for dinucleotides, 7 for tri-

© 2006 NRC Canada

nucleotides, 6 for tetranucleotides, and 5 for pentanucleotides and hexanucleotides. Academic research licenses are available for primer sequences of the 81 framework SSR markers (ViaLactia Biosciences, New Market, Auckland, New Zealand).

PCR amplification and product electrophoresis were performed by a contract genotyping provider (SignaGen, Rotorua, New Zealand) as outlined below. SSR primers were labelled using either 6-FAM, VIC, NED, or PET fluorochrome moieties (Applied Biosystems, Foster City, Calif.). PCR was performed in a total volume of 10 µL containing 10 ng of genomic DNA, 0.2 mmol/L dNTPs, 1.5 mmol/L MgCl, 0.25 µmol/L of each primer, 0.5 U of Ampli*Taq* Gold DNA polymerase (Applied Biosystems) and 1× GeneAmp PCR buffer I. A touchdown thermocycling protocol was used as follows: initial denaturation at 92 °C for 9 mins; 2 cycles of 94 °C for 1 min, 65 °C for 1 min, and 70 °C for 35 s; 18 cycles of 93 °C for 45 s and 64 °C for 45 sec (temperature was reduced by 0.5 °C/cycle until it reached 55.5 °C); 70 °C for 45 s; 20 cycles of 92 °C for 30 s, 55 °C for 30 s, and 70 °C for 60 s; and a final extension at 70 °C for 20 min. Products were separated using an ABI3100 capillary sequencer (Applied Biosystems) and sized using either the GS-500 ROX or GS-500 LIZ ladder (Applied Biosystems). Allele scoring was performed using Genotyper v. 3.7 (Applied Biosystems).

SSR markers were evaluated on a panel consisting of a single genotype from 6 New Zealand cultivars (data not shown) and DNA from the sibs of each grandparent of the mapping pedigree (grandparent DNA was not available). Data from Armstead et al. (2002) was used to extend map coverage (described below).

Framework map construction

To construct a suitably robust map of markers with a combination of segregation patterns, we chose a strategy where F_2 markers only were used to construct a preliminary framework. Since linkage phase was uncertain in a portion of the F_2 data (denoted "F_2 phase-unknown"), any unlinked markers (after the first analyses were conducted) were inverted by reversing homozygote categories and analysing the data with both phases. Here we assumed that the markers with the incorrect phase assignments would not be linked to any of the markers where the correct phases were known ("F_2 phase-known"). Marker groups were determined using a minimum LOD ratio of 6.0 and a maximum recombination rate of 0.20. For groups with more than 7 markers, optimal order was estimated using a matrix correlation method as implemented in Mapmaker (Lander et al. 1987) Macintosh v. 2.0 via the FIRST ORDER command. For groups with 7 markers or fewer, the COMPARE function was used to determine the best possible order based on a group-wise multipoint log-likelihood value. The most likely order for the group was next estimated by permuting or shuffling positions 3 markers at a time using the RIPPLE command. Orders with log-likelihood differences of 3.0 or more compared with the next most-likely order (support-for-order) were accepted for framework loci. Where the log-likelihood difference for a group of 3 markers was less than 3.0, markers were dropped one at a time from regions where support-for-orders were less than 3.0. After each marker was dropped, orders were re-evaluated and the RIPPLE analysis repeated until final support-for-order exceeded 3.0. Criteria for dropping markers from the framework included map contraction (determined using the DROP MARKER command) and consistency with triangular equality. The most-likely intervals where the dropped markers (denoted "accessory" markers) were located was estimated using the TRY command, and the nearest framework marker identified using the NEAR command.

Markers that amplified only 1 band and segregated in a 3:1 manner with the other allele having no discernable amplicon are referred to as "pseudointercross" (PI) markers. In these cases, the F_1 was assumed to be a heterozygote for null alleles. Pseudointercross (PI) markers were mapped by recoding the F_2 data to phase-unknown PI data (i.e., scored twice, by assigning different homozygous classes as dominant and recessive classes), and merging this with the 3:1 PI data.

Markers that were linked to framework markers, but that could not be placed into specific intervals, have been denoted as "floating" markers and were assigned to the chromosome showing strongest linkage. Map distances were calculated by Mapmaker using the Kosambi mapping function (Kosambi 1944). Segregation distortion of markers was checked using standard χ^2 goodness-of-fit measures for both the F_2 and PI data. A genome-wide threshold corresponding to a comparison-wise p value of 0.01 was used to approximate an experiment-wise p value of 0.05.

Once a map was constructed according to the above criteria, marker coverage was compared with that of an existing F_2 RFLP-based map constructed by Armstead et al. (2002). Gaps in genome coverage of the SSR framework map were identified, and markers from the RFLP-based map that were located in these gaps were incorporated. To do this, we used the TRY command to locate the most-likely interval the new markers should fall within, and re-ordered the group according to the framework map criteria described above. Marker placements were accepted where (*i*) map coverage was extended, (*ii*) the final order incorporating the additional marker achieved a minimum support-for-order of LOD 3.0, and (*iii*) the addition of the new marker did not change the pre-existing framework SSR marker order. A total of 20 markers from the F_2 map were initially selected for this analysis.

Results

Marker yields and sequence analysis

The proportion of markers developed and mapped from each library source is given in Table 1. The frequency of non-redundant SSRs present in the GeneThresher, enriched, and EST libraries was 0.83%, 25.4%, and 1.8%, respectively, when using the repeat-length criteria described in the Materials and methods section. The number of suitable SSR clones for marker development was severely reduced once primer-design criteria were met. The EST library source was most affected, with only 20.3% of non-redundant SSR clones passing primer-design criteria. Efficiency loss owing to primer design was mainly attributed to shortage of flanking sequence (truncated clones) adjacent to the SSR. All SSRs identified from the EST library source were confined

© 2006 NRC Canada

Table 1. Proportion of SSR markers developed from GeneThresher, EST, and SSR-enriched libraries.

Library	Clones sequenced	Non-redundant SSR yield	Primers designed[a]	Library yield[b] (%)	Polymorphic in screening panel[c]	Mapped[d]
GeneThresher	155 084	1287 (0.83%)	563 (43.7%)	0.36	258 (45.8%)	122 (21.7%)
SSR-enriched	6528	1658 (25.4%)	931 (56.2%)	14.3	355 (38.1%)	209 (22.4%)
EST	6596	118 (1.8%)	24 (20.3%)	0.36	9 (37.5%)	5 (20.8%)

[a]The proportion of non-redundant SSRs for which primers were designed (i.e., passed primer-design criteria).
[b]Percentage of SSR clones for which primers were designed out of the total number of clones sequenced in the library.
[c]Proportion of SSRs for which primers were designed that displayed polymorphism in a screening panel consisting of 8 *L. perenne* cultivars.
[d]Proportion of SSRs for which primers were designed that mapped in the F_2 population. An additional 39 loci contributed by the same SSRs amplifying more than one product are not included in the total.

to the 5′ or 3′ untranscribed region (UTR), which often limited the amount of potential sequence available for primer design. GeneThresher and EST libraries were of similar efficiency for yielding SSRs that passed primer-design criteria (0.36%) when the total number of clones sequenced in each library type was considered. As expected, SSR-enriched genomic libraries were the most efficient source (14.3%) based on the total number of clones sequenced in each library type. The proportion of SSRs that had primers designed and could be mapped in the population used for this study did not differ greatly depending on the library source (21–23%).

Our latest GeneThresher data set consists of 446 960 sequence reads, assembled into 80 162 contigs and 189 697 singletons. BLASTn analysis of the mapped GeneThresher SSR singleton or contig sequence against the GenBank nr and EST_other nucleotide databases determined that 44.6% of GeneThresher SSRs were most-likely associated with an expressed gene ($E < 7 \times 10^{-7}$). Thirty-eight percent of the GeneThresher SSR sequences mapped gave very high gene similarity BLASTn hits ($E < 2 \times 10^{-20}$). Given that GeneThresher sequences often contained non-coding sequences such as introns and promoters, BLASTn scores of $E < 7 \times 10^{-7}$ against the GenBank databases were significant. A portion of the total GeneThresher SSRs mapped were associated with a putative rice gene ortholog with clear annotation (18.5%) once aligned to the rice genome at Gramene (http:\\www.gramene.org) using BLASTn (Table 2). Regions of alignment can be viewed by searching for the rice TIGR locus ID in the genome browser and selecting the ryegrass_methylfilter option from the GSS feature menu. From the subset of mapped markers giving clear annotation, approximately 35% of these belonged to the transcription factor class of genes. Perennial ryegrass SSRs developed from enriched genomic library sources corresponded to genes only 6% of the time (BLASTn $E < 1 \times 10^{-7}$).

Data analysis and map construction

No DNA was available for genotyping from either grandparents (F_0), although DNA from a sib of each of the grandparents was available. A total of 94 F_2 progenies were genotyped for 381 SSR loci. Almost half (172) segregated in a 1:2:1 ratio consistent with a phase-known F_2 population, i.e., where the genotype of the grandparents' sib was consistent with those of the parents. In addition, another 63 loci were segregating in a 1:2:1 ratio consistent with an F_2 segregation pattern, but one or both of the grandparental sib genotypes were not consistent with the alleles segregating at the same markers. These were considered phase-unknown F_2 loci. Finally, 146 loci segregated in a 3:1 manner and are referred to as phase-unknown PI markers. Some of the primer pairs (31) amplified more than one segregating locus providing a total of 70 marker loci. A total of 52 F_2 and 17 PI markers showed segregation distortion (comparison-wise $\alpha < 0.01$). Of these markers, all but 10 mapped to either linkage group 5 or 7 (Fig. 1).

Using the procedure for framework map construction, as described in the Materials and methods section, all F_2 markers were tested for suitability as framework markers. Of these F_2 SSR markers, 81 met the criteria for acceptance as framework markers. For the phase-unknown F_2 markers, linkages were found for only 1 phase at the LOD threshold used to assign markers to groups. Properties of a subset of the framework SSR markers are described in Table 3. An additional 233 F_2 and PI markers were placed as accessory markers. A number of both F_2 and PI markers (57) were linked to framework markers, but could not be placed into specific intervals. These markers typically showed closest linkage to non-terminal framework markers, but showed more recombination to these markers than the immediately adjacent flanking framework markers and were hence denoted as "floating" markers. Ten markers were unlinked after this initial analysis that provided 7 linkage groups — the same number as the haploid chromosome number in perennial ryegrass.

To check genome coverage of the SSR markers, the framework map was aligned with a predominantly RFLP-based map constructed with 180 progenies (Armstead et al. 2002), which included the same 94 F_2 individuals that were used for this study. Alignment revealed a few regions located at the terminal ends of linkage groups that were not covered by the SSR-only framework map. To increase map coverage, we then added 10 markers (9 RFLPs and 1 SSR) from regions not covered, and incorporated these using the same analyses as the SSR markers. The addition of these markers to the framework map allowed 6 markers that were originally unlinked or floating to be placed as accessory markers. A final summary listing the proportion of markers derived from each category and the subsequent linkage group length is provided in Table 4. The final map (Fig. 1) shows framework loci, as well as accessory markers, with the interval that they are most likely to be placed within. Floating markers are shown below the group to which they show the strongest linkage. Markers showing segregation distortion are also denoted. The 7 linkage groups initially covered a total map length of 484.9 cM when only SSR markers were con-

© 2006 NRC Canada

Table 2. Putative gene function and map location of perennial ryegrass GeneThresher-SSR markers.

GeneThresher SSR marker	Putative gene function	Map location: Linkage group	Map location: Map distance (cM)	Rice TIGR locus ID
	Transcription factor class			
rv0904	Knotted homeodomain protein	7	6.7–20.2	LOC_Os06g43860
rv0922	Teosinte-branching 1 gene (TB1)	4	61.3–71.8	LOC_Os03g49880
rv1087	MADs box	1	53.6–82.4	LOC_Os03g54170
rv1159	Zinc-finger homeobox protein	5	68.6–71.9	LOC_Os11g03420
rv1190	YABBY transcription factor	4	50.3	LOC_Os03g11600
rv1316	NAM (no apical meristem) protein	3	39–43.6	LOC_Os07g37920
rv1385	GL2-type homeodomain protein	6	66.1	LOC_Os02g45250
rv1412	GT-2 transcription factor	4	0–20.3	LOC_Os03g02240
	Other			
rv0983	Cytochrome P450	6	62.2	LOC_Os05g12040
rv1063	Aluminium tolerance S222 gene	3	49.6	LOC_Os02g37590
rv1117	BEACH-domain-containing protein	2	—	LOC_Os04g46900
rv1139	Microtubule-associated protein	5	59–64.8	LOC_Os09g27700
rv1149	Integral membrane protein	2	57.6	LOC_Os04g58760
rv1131	Fasiclin-like arabinogalactan	3	57.3	LOC_Os05g38500
rv1175	Transducin protein	7	44.8–49.9	LOC_Os08g07960
rv1242	Glycosyl hydrolase	1	27.1	LOC_Os05g31140
rv1278	Pectin acetylesterase	6	—	LOC_Os10g40290
rv1282	Xyloglucan endotransglycosylase	2	—	LOC_Os04g51460
rv1307	DMC1 meiotic recombination	5	68.6–71.9	LOC_Os12g04980
rv1317	Ubiquitin carrier protein	4	—	LOC_Os01g16540
rv1338	Receptor-like protein kinase	7	50.2–52.3	LOC_Os04g44910
rv1384	Copper chaperone protein	2	—	LOC_Os03g26650
rv1438	Sugar transporter protein	2	54.9	LOC_Os04g41460

Note: Only GeneThresher-SSR markers giving clear annotation at BLASTn scores of $E < 7 \times 10^{-7}$ are listed. Map location for several accessory markers is given as an interval range, whereas floating markers were assigned only to the appropriate linkage group. The TIGR locus ID corresponds to the highest BLASTn match from rice obtained during a genome alignment of the GeneThresher perennial ryegrass sequence at Gramene (http:\\gramene.org). Regions of alignment can be viewed using the genome browser and selecting the ryegrass_methylfilter option from the GSS feature menu.

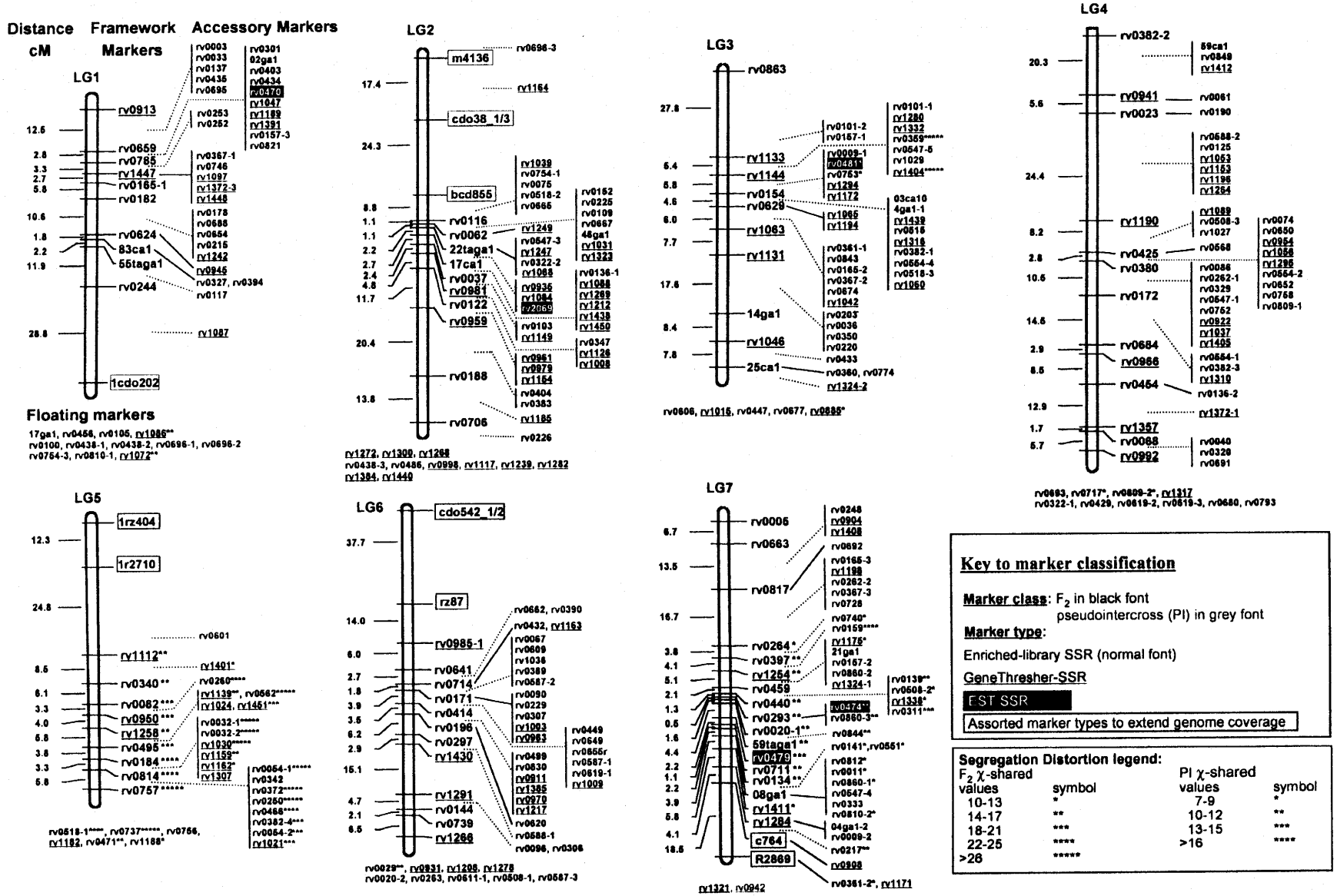
sidered and 675.6 cM once markers for extending terminal ends of linkage groups were incorporated.

Discussion

We have constructed a moderate-density map of the ryegrass genome using 376 SSR markers, complemented by 9 RFLP markers. The analytical approach we have used to construct linkage groups differs from that of other (published) ryegrass maps. Our approach is based on ordering markers with a given minimum level of confidence in gene order, and retaining in the resulting framework only those markers where order achieves a preset level of confidence in order (support-for-order > LOD 3.0). In so doing, we were able to identify markers thatwere problematic in their placement on the framework for a variety of reasons. This map should therefore provide a reasonably robust basis with which to apply to other pedigrees of this species and for comparing genomic rearrangements with closely related species. Such approaches have been used extensively in other species, mostly in constructing linkage maps of forest trees (Grattapaglia and Sederoff 1994; Echt and Nelson 1997; Remington et al. 1999; Wilcox et al. 2001). Comparison of framework maps constructed with independent progeny sets descended from the same parent in maritime pine (Plomion et al. 1995) using analytical methods similar to those described here showed that some differences in framework order do occur, but these were relatively infrequent and similar to that expected owing to chance. Indeed, while the support-for-order criterion of 3.0 is based on local orders, the multiple number of chromosomes effectively increases the probability of a spurious order, but not by a large amount for a genome with so few linkage groups. We calculate that the genome-wise support-for-order corresponding to group-wise LOD = 3 for 7 linkage groups is approximately LOD 2.15 (based on Bonferroni corrections for the number of linkage groups). This indicates that, on a genome-wide basis, the depicted orders are more than 100 times more likely than the next most-likely order, based on the genotypic data supplied. Slight changes in marker order that could result, albeit with low probability, may nonetheless be relatively trivial, depending of course on the specific application of the map. Preliminary marker data obtained in a second perennial ryegrass mapping population (data not shown) was sufficient to compare framework marker order of 2 linkage groups. All framework markers in common retained their order (8 from

© 2006 NRC Canada

Fig. 1. A framework genetic linkage map of perennial ryegrass. Framework markers are shown on the map itself, along with distance (in Kosambi centimorgans (cM)). Straight lines between accessory markers and framework markers indicate the accessory marker(s) is (are) linked without any observed recombination with the framework marker. Dashed lines from accessory markers to framework markers signify the most likely interval within which the accessory marker was placed. Floating markers are assigned to linkage groups. Legends for marker class, source, and segregation distortion are described in the figure key.



© 2006 NRC Canada

Table 3. Properties for a selection of SSR markers located on the framework map.

SSR marker	Linkage group	Location (cM)	SSR motif	Library source	Observed size range (bp)	Forward primer	Reverse primer
rv0913	1	0	$(CT)_9$	GeneThresher	135–176	GGCTTACAGCGAAGAAGACATT	CAGTTGAGCTCATTGGAGACAC
rv0659	1	12.5	$(CT)_{19}(CA)_{11}$	Enriched	103–148	TCTCTCCAAGTTCCTCTTCTGC	TCTGCTTCCATCAAGAAAGGTT
rv0244	1	53.6	$(CT)_{21}$	Enriched	197–236	CAGATTGCAACTCACGAGGA	ATTTTCCCGTCTGTATTGCG
rv0959	2	76.5	$(TAA)_9$	GeneThresher	123–176	GTAAACTGGGTTGATGGGATGT	CATTGTACGATTTATGCTTCTCCA
rv0706	2	101.7	$(CT)_{18}$	Enriched	203–225	TTCATGTTAGCCCTCCTGTTTT	AATCAAGCCATCTCGTCCTCTA
rv1133	3	27.8	$(CG)_8$	GeneThresher	131–137	CGGTTCTGAATTTTCTGCTTCT	AAAGCTGAGAACCGAATCCATA
14ga1	3	74.9	$(CT)_{25}$	Enriched	135–200	GTTTGCATTCCTTGTGCT	TCCTTGGGTGTGTCTACG
25ca1	3	91.1	$(GT)_{16}$	Enriched	280–300	ATCGAGTGCACATTTCGT	TGTGATCGATGGTGCATA
rv0941	4	20.3	$(CT)_9$	GeneThresher	199–211	ACTTGCATAGCGTACAAGCAAA	AAAAACTGTACACTGCGCTCAC
rv0454	4	97.7	$(CT)_{19}$	Enriched	224–263	AGATGTGAGCTGTGTGCCTT	CGATATATCAGCCTCACGCA
rv0992	4	118.0	$(TAA)_{11}$	GeneThresher	183–216	CAATGCCATCACCAACACTACT	TTTTGTAGCATGGAGAGATTACCA
rv1112	5	37.1	$(CT)_{11}$	GeneThresher	252–271	TTTGTTAGCCCGTCCTTACCTA	AACATGTGGAATGGAATGGAAT
rv0757	5	77.7	$(CT)_{27}$	Enriched	203–235	AGATGATGATGCAGTGGACAAG	CACCACTTCAATCACGCAGTAT
rv0985-1	6	51.7	$(CT)_{18}$	GeneThresher	212–248	CTTATTATTGCAAGCGATGGTG	GTATCGTCCCTCAGTCCTTTTG
rv0641	6	57.7	$(CT)_{10}$	Enriched	206–214	TGCATAACTTCACTGCAGCATA	AGAAACTCGGTAGAAGGACCTC
rv1266	6	107.1	$(CAG)_{10}$	GeneThresher	171–192	AAGAGGAAACAAATGGCAACAT	AGCTTCCTTACATGCTCTGTAT
rv0005	7	0	$(TAGA)_{22}$	Enriched	114–164	GGCAACTAGGAGAAAAATGAGAA	GATTGTAGTAGTGGGTAATCCT
rv0663	7	6.7	$(CT)_{15}$	Enriched	111–121	ATTAGTTCAAACACCCGCAAGT	CGTCTGCACCTCTACAGCATAG
rv1284	7	75.0	$(TA)_9$	GeneThresher	147–160	TTTGACTATTTCTCGCACTTCG	GAATTCCGCACATATCCAAACT

© 2006 NRC Canada

Table 4. Total number of marker loci per linkage group assigned to either "framework", "accessory", "floating", or "extension" marker categories and the length of genome coverage (in Kosambi centimorgans (cM)).

Linkage group	Framework SSR markers	Accessory F_2 and PI	Floating F_2 and PI	Markers for map extension[a]	Total[b]	Length (cM)
1	10	32	12	1	55	82.4
2	10	40	11	3	64	101.7
3	10	39	5	0	54	91.1
4	13	40	10	0	63	118.0
5	9	21	6	2	38	77.7
6	12	30	9	2	53	107.1
7	17	37	2	2	58	97.6
Total	81	239	55	10	385	675.6

[a]Nine RFLP markers and 1 SSR marker used to extend genome coverage (described in Armstead et al. 2002).
[b]Six additional unlinked SSR markers are not included in the total.

linkage group 3 and 11 from linkage group 7), providing further support for achieving a high-quality framework map in this study.

Threshold criteria for placement of markers into particular intervals was not applied, therefore the level of robustness regarding placement of accessory markers with regard to the framework markers is not nearly as high and should be treated with appropriate caution when using the accessory markers. We also found that within each linkage group there were floating markers that did not fit within the framework structure, but showed strong linkage to at least one (and usually more) of the framework markers. We rechecked genotypic assignments and re-genotyped a total of 18 of these markers and found that this improved placement of 12 of these markers, indicating some level of genotyping error. As a comparison, we genotyped 5 randomly chosen framework markers and found that we were able to repeat genotypic scores in more than 95% of the data (results not shown), indicating the approach we have used for framework map construction preferentially selects higher-quality marker data. We did not re-genotype all of the floating markers on the basis that, for most markers, they would not extend map coverage nor increase the density of the final map, as the latter threshold is limited more by sample size than number of markers. For the unlinked F_2 markers, we checked phase assignments by reversing homozygote categories and attempting to fit them into the framework using the same approach as for accessory markers. However, 6 markers remained unlinked. Possible causes include genotyping error, but these unlinked markers could also be due to genetic sampling in the distal regions of some linkage groups, resulting in some terminal markers appearing to be unlinked because of the sample used for genotyping.

The estimated map distance covered by the framework map of 675.6 cM was greater than that reported by the Armstead et al. (2002) RFLP-based maps. One of these maps was constructed from the same pedigree that included the 94 progeny used in this study. Their maps covered 515 cM in the same pedigree and 565 cM in a backcross pedigree, although they were not constructed using the same criteria as in this study. Alignment of our SSR map to the Armstead et al. (2002) RFLP-based map of the same pedigree identified regions where the different marker types complemented each other to increase overall genome coverage. Hayward et al. (1998) reported a comparable map length of 692 cM, whereas Bert et al. (1999) reported a map distance of 930 cM (based on 463 AFLP markers) and Jones et al. (2002*a*) reported a map distance of 811 cM. A more recent map (Faville et al. 2004), comprising SSR and RFLP markers, was 963 cM in one parent and 757 cM in the other. These larger distances were reported from maps with a high proportion of AFLP or RFLP markers, which may provide greater genome coverage than SSR markers alone. Different map length distances may also be a result of the various means used to construct the linkage maps and also differences in recombination frequency between contrasting genetic backgrounds (Barth et al. 2001).

The F_2 loci used to construct the initial framework map, excluding the floating SSR and RFLP-based markers, was also used to construct a comprehensive map (data not shown) using the program OutMap (CSIRO Forestry and Forest Products, Kingston, Australia). Map expansion of over 273 cM was observed in the comprehensive map compared with the framework SSR-only map, and is most likely due to the maximum-likelihood placement of the accessory loci for which statistical support for order is poor. This results in increased recombination fractions between loci with no or poor interval support. Even with the inclusion of such loci, the orders of the framework loci as defined by the support-for-order criteria implemented through Mapmaker Macintosh 2.0 are largely conserved. When only framework loci were analysed the OutMap method of map construction agreed precisely with the framework map constructed using MapMaker Macintosh 2.0. The placement of accessory markers into framework intervals was quite variable from the assignments on the framework map. This highlights the differences that can arise when different methods are used to place loci that have poor support-for-order, as inferences about map length and coverage can be erroneous. Furthermore, these results cast doubt over some of the published genome lengths and map coverage for this species, which indeed may be somewhat less owing to inclusion of markers in the final order for which there is no solid statistical evidence for doing so.

Explanations for map contraction include inbreeding depression leading to under-representation of some areas of the

© 2006 NRC Canada

genome with markers; however, this does not account for the differences in distance between framework and comprehensive mapping analyses, and would also be characterized by segregation distortion in flanking regions. While this may have occurred, particularly in LG 5 where there is an increasingly severe level of distortion toward one of the termini, the extent of distortion is limited to only 2 linkage groups. Of the 52 F_2 and 17 PI markers showing segregation distortion, all but 10 mapped to regions within either linkage group 5 or 7, indicating the presence of some heritable factor(s) contributing to segregation distortion. Previous maps of ryegrass have also found regions of segregation distortion. Armstead et al. (2002) also observed distorted marker segregation on LG5 and LG7 in the same mapping family. In a separate mapping population (ILGI reference population), segregation distortion was concentrated on linkage groups 3 and 4 (Armstead et al. 2002; Jones et al. 2002*b*). Consistent and significant segregation distortion of linked markers can indicate the location of gametophytic self-incompatibility loci. Thorogood et al. (2002) reported that the ryegrass *S* and *Z* incompatibility loci are located on linkage groups 1 and 2, respectively. The linkage group 1 (*S*) locus displayed strong genetic interaction with another locus on linkage group 3. Recently, the self-compatible *T* locus (analogous to the *S5* locus in rye) has been located on linkage group 5 in perennial ryegrass (Thorogood et al. 2005), which corresponds to the same region of marker-segregation distortion observed on linkage group 5 in this study. No incompatibility locus has been reported on linkage group 7, so marker-segregation distortion is more likely explained by the presence of lethal or semi-lethal loci that have not been purged during previous generations of inbreeding. Estimates of recombination fraction and therefore map distances are likely to be biased for the F_2 markers showing distorted segregation ratios. In these analyses, we have not attempted to correct for such bias, but recognise that estimates of map distances could be affected.

Significant clustering of SSR markers occurred around the putative centromeric region of each perennial ryegrass linkage group and was also observed by Jones et al. (2002*b*). SSR loci isolated from genomic libraries are also clustered near the centromeres on the genetic maps of both barley (Ramsay et al. 2000) and tomato (Areshchenkova and Ganal 1999). This may be attributed to the selection of longer SSR repeats (Areshchenkova and Ganal 1999) and non-random association of SSRs with retroelements (Ramsay et al. 1999) that are common near plant centromeres (Jiang et al. 2003). A major contributing factor of the observed clustering is the probable event of uneven distribution of recombination around the centromeric regions. This is evidenced in wheat (Weng et al. 2000), barley (Kunzel et al. 2000), and a 10-fold reduction in recombination was reported in the cases of potato and tomato (Tanksley et al. 1992). Chromosome 1 of perennial ryegrass displays very low recombination between the centromere and the nucleolar organizer region (King et al. 2002).

We found that 5 chromosomes were deficient of SSR markers near the distal ends, requiring the integration of RFLP-based markers to extend genome coverage. Similarly, Jones et al. (2002*b*) reported irregular distribution, finding that 93 SSR markers developed from enriched genomic libraries covered only 54% of the total map distance in *L. perenne*, without the inclusion of RFLP and AFLP markers. Given that ~6% of the perennial ryegrass SSRs derived from genomic libraries were associated with genes, a high proportion of these are likely to reside in either repetitive DNA regions or the gene-poor space between gene-rich regions. SSR marker-poor regions have been identified in other plant species and can be dependant on whether the markers are derived solely from genic or from non-genic sequence. For example, in sorghum, SSRs derived from genomic DNA were absent or at very low density for about 25% of the total map length (Bhattramakki et al. 2000). EST-SSR markers from wheat were more densely distributed in the distal chromosome regions (Yu et al. 2004). Similarly in rice, an EST transcript map identified gene-rich regions near the distal region of most chromosomes (Wu et al. 2002). In this study, the majority of SSR markers were derived from non-genic sequence and the marker deficient regions were located towards the distal chromosome ends where cDNA-RFLP markers were able to fill the gaps. The perennial ryegrass map of Faville et al. (2004) contains a high proportion of EST-SSR and cDNA-RFLP markers and appears to provide better coverage at the distal ends of linkage groups. Mapping a larger proportion of SSR markers derived from ESTs or gene-associated GeneThresher clones should therefore improve genome coverage in the SSR marker-deficient regions found in this study.

SSR-enriched genomic libraries obviously provide the most economical source for SSR marker development. Novel SSR yield from EST and GeneThresher libraries was low in perennial ryegrass (0.36%) once redundancy and primer-design criteria were met. Although the primary goal of EST and GeneThresher sequencing is for gene discovery, larger projects can also yield sufficient SSR markers for the construction of framework genetic maps. Gene associated markers are advantageous because of greater transportability to other pedigrees and species and also provide the possibility that the gene itself could be responsible for trait variance. Perennial ryegrass SSRs developed from SSR-enriched genomic library sources rarely corresponded to genes (~6%) as also found by Jones et al. (2002*b*). GeneThresher SSRs were often associated with genes (44.6%) and represent genes that are unbiased for expression level unlike EST-derived markers (Bonaldo et al. 1996). Genes with rare transcript levels such as the transcription factor class of genes were equally represented. Although it is reasonable to expect GeneThresher libraries may yield a higher proportion of SSR markers associated with transcription factors (~35% of clearly annotated genes in this study) compared with EST sources, this is difficult to accurately quantify and may be confounded by improved annotation of the transcription category of genes, since they share commonly occurring DNA-binding motifs. A contributing factor is also the high proportion of GeneThresher SSRs with significant BLASTn hits to the GenBank EST_other database, which correspond to genes encoding proteins without functional annotation (48%). Alignment of the mapped GeneThresher SSRs (associated with genes) to the rice genome at Gramene identified putative rice orthologs for 23 genes. These are potentially useful for comparing synteny between rice, perennial ryegrass, and other grass species.

© 2006 NRC Canada

Generation of a statistically robust genetic map consisting of reasonably solid gene orders is the first key step in applying such a map not only to existing applications such as quantitative trait dissection and marker-assisted breeding, but also to the developing field of association genetics (Neale et al. 2004, Gupta et al. 2005). Here, the ability to reliably locate expressed genes relative to each other will be important in detecting biologically meaningful associations between polymorphisms in DNA sequence, with trait variation in genetically unstructured populations. Having a set of expressed genes already mapped provides an additional advantage, as some of these genes may actually be expressed during formation of specific traits, and therefore be useful targets for testing whether or not they are associated with trait variation. Furthermore, the same genes can be used to determine expression-level differences among individual genotypes with microarrays or equivalent technologies, and can therefore be used to determine the genetic basis of variation in expression-level differences.

We have constructed a moderate-density framework linkage map of perennial ryegrass from a total of 376 SSR marker loci, complemented with 9 RFLP loci. Density and coverage of this map is such that it should be immediately useful, particularly for applications in molecular breeding programs, as well as for more basic science applications such as comparative genomics. Furthermore, because this map is constructed of mostly codominant markers, it should be transportable to other pedigrees in this species, as well as to closely related species where the primer sequences also amplify polymorphic loci. Currently in progress, our complete GeneThresher sequence data set is to be aligned at Gramene (www.gramene.org) with rice, sorghum, maize, barley, and wheat, which will enhance molecular breeding of perennial ryegrass through cross-species transfer of genetic information.

Acknowledgements

We gratefully acknowledge Tania Elder (Forest Research Institute, New Zealand) for assistance in data entry and SignaGen staff (Forest Research Institute, New Zealand) for data generation.

References

Areshchenkova, T., and Ganal, M.W. 1999. Long tomato microsatellites are predominantly associated with centromeric regions. Genome, **42**: 536–544.

Armstead, I.P., Turner, L.B., King, I.P., Cairns, A.J., and Humphreys, M.O. 2002. Comparison and integration of genetic maps generated from F_2 and BC_1-type mapping populations in perennial ryegrass. Plant Breed. **121**: 501–507.

Arumuganathan, K., Tallury, S.P., Fraser, M.L., Bruneau, A.H., and Qu, R. 1999. Nuclear DNA content of thirteen turfgrass species by flow cytometry. Crop Sci. **39**: 1518–1521.

Barth, S., Melchinger, A.E., Devezi-Savula, B., and Lubberstedt, T. 2001. Influence of genetic background and heterozygosity on meiotic recombination in *Arabidopsis thaliana*. Genome, **44**: 971–978.

Bedell, J.A, Budiman, M.A, Nunberg, A., Citek, R.W., Robbins, D. Jones, J., et al. 2005. Sorghum genome sequencing by methylation filtration. PLOS Biol. **3**(1): e13.

Bert, P.F., Charnet, G., Sourdille, P., Hayward, M.D., and Balourier, F. 1999. A high density map for ryegrass (*Lolium perenne* L.) using AFLP markers. Theor. Appl. Genet. **99**: 445–452.

Bhattramakki, D., Dong, J., Chhabra, A.K., and Hart, G.E. 2000. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. Genome, **43**(6): 988–1002.

Bonaldo, M.F., Lennon, G., and Soares, M.B. 1996. Normalization and subtraction: two approaches to facilitate gene discovery. Genome Res. **6**: 791–806.

Echt, C.S., and Nelson, C.D. 1997. Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.). Theor. Appl. Genet. **94**: 1031–1037.

Faville, M.J., Vecchies, A.C., Schreiber, M., Drayton, M.C., Hughes, L.J., Jones, E.S., et al. 2004. Functionally associated molecular genetic marker map construction in perennial ryegrass (*Lolium perenne* L.). Theor. Appl. Genet. **110**: 12–32.

Grattapaglia, D.G., and Sederoff, R.R. 1994. Genetic linkage maps of *Eucalyptus grandis* and *E. europhylla* using a pseudo-testcros strategy and RAPD markers. Genetics, **137**: 1121–1137.

Gupta, P.K., Rustgi, S., and Kulwal, P.L. 2005. Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol. Biol. **57**(4): 461–485.

Hayward, M.D., Forster, J.W., Jones, J.G., Dolstra, O., Evans, C., McAdam, N.J., et al. 1998. Genetic analysis of *Lolium* I. Identification of linkage groups and the establishment of a genetic map. Plant Breed. **117**: 451–455.

Jiang, J., Birchler, J.A., Parrott, W.A., and Dawe, R.K. 2003. A molecular view of plant centromeres. Trends Plant Sci. **8**(12):570–575.

Jones, K.C., Levine, K.F., and Banks, J.D. 2000. DNA-based genetic markers in black-tailed deer for forensic applications. Calif. Fish Game, **86**(2): 115–126.

Jones, E.S., Mahoney, N.L., Hayward, M.D., Armstead, I.P., Jones, J.G., Humphreys, M.O., et al. 2002*a*. An enhanced molecular marker based genetic map of perrenial ryegrass (*Lolium perenne*) reveals comparative relationships with other Poaceae genomes. Genome, **45**: 282–295.

Jones, E.S., Dupal, M.P., Dumsday, J.L., Hughes, L.J., and Forster, J.W. 2002*b*. An SSR-based genetic linkage map for perennial ryegrass (*Lolium perenne* L.). Theor. Appl. Genet. **105**: 577–584.

Keats, B.J.S., Sherman, S.L., Morton, N.E., Robson, E.B., Buetow, K.H., Cann, H.M., et al. 1991. Guidelines for human genetic linkage maps: an international system for human linkage maps (ISLM) 1990. Genomics, **9**: 557–560.

King, J., Armstead, I.P., Donnison, I.S., Thomas, H.M., Jones, R.N., Kearsey, M.J., et al. 2002. Physical and genetic mapping in the grasses *Lolium perenne* and *Festuca pratensis*. Genetics, **161**: 315–324.

Kosambi, D.D. 1944. The estimation of map distances from recombination values. Ann. Eugen. **12**: 172–175.

Kunzel, G., Korzun, L., and Meister, A. 2000. Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics, **154**: 397–412.

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E., and Newburg, L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics, **1**: 174–181.

Liu, B-H. 1998. Statistical genomics: linkage, mapping and QTL analysis. CRC Press, Boca Raton, Fla.

© 2006 NRC Canada

Neale, D.B., and Savolainen, O. 2004. Association genetics of complex traits in conifers. Trends Plant Sci. **9**(7): 325–330.

Plomion, C.O., Durel, C-E., and Malley, D.M. 1995. Genomic analysis in maritime pine (*Pinus pinaster*). Comparison of two RAPD maps using selfed and open-pollinated seeds of the same individual. Theor. Appl. Genet. **90**: 1028–1034.

Rabinowicz, P.D., Schutz, K., Dedhia, N., Yordan, C., Parnell, L.D., Stein, L., et al. 1999. Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. Nat. Genet. **23**: 305–308.

Ramsay, L., Macaulay, M., Cardle, L., Morgante, M., degli Ivanissevich, S., Maestri, E., et al. 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., et al. 2000. A simple sequence repeat-based linkage map of barley. Genetics, **156**: 1997–2005.

Remington, D.L., Whetten, R.W., Liu, B-H., and O'Malley, D.M. 1999. Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. Theor. Appl. Genet. **98**: 1279–1292.

Roldan-Ruiz, I., Dendauw, I., van Bockstaele, E., Depicker, A., and De Loose, M. 2000. AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). Mol. Breed. **6**: 125–134.

San Miguel, P., Tikhonov, A., Jin, Y.K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., et al. 1996. Nested retrotransposons in the intergenic regions of the maize genome. Science (Washington, D.C.), **274**: 765–768.

Tanksley, S.D., Ganal, M.W., Prince, J.P., deVicente, M.C., Bonierbale, M.W., Broun, P., et al. 1992. High density molecular linkage maps of the tomato and potato genomes. Genetics, **132**: 1141–1160.

Thorogood, D., Kaiser, W.J., Jones, J.G., and Armstead, I. 2002. Self-incompatibility in ryegrass 12. Genotyping and mapping the S and Z loci of *Lolium perenne* L. Heredity, **88**: 385–390.

Thorogood, D., Armstead, I.P., Turner, L.B., Humphreys, M.O., and Hayward, M.D. 2005. Identification and mode of action of self-compatibility loci in *Lolium perenne* L. Heredity, **94**: 356–363.

Turner, L.B., Humphreys, M.O., Cairns, A.J., and Pollock, C.J. 2001. Comparison of growth and carbohydrate accumulation in seedlings of two varieties of Lolium perenne. J. Plant Physiol. **158**: 891–897.

Whitelaw, C.A., Barbazuk, W.B., Pertea, G., Chan, A.P., Cheung, F., Lee, Y., et al. 2003. Enrichment of gene-coding sequences in maize by genome filtration. Science (Washington, D.C.), **302**: 2118–2120.

Weng, Y., Tulenn, N.A., and Hart, G.E. 2000. Extended physical maps and consensus physical map of the homoeologeous group 6 chromosome of wheat (*Triticum aestivum* L. em Thell). Theor. Appl. Genet. **100**: 519–52.

Wilcox, P.L., Richardson, T.E., Corbett, G.E., Lee, J.R., Djorovic, A., and Carson, S.D. 2001. Framework linkage maps of *Pinus radiata* D. Don based on pseudotestcross markers. Forest Genet. **8**(2):101–109.

Wu, J., Maehara, T., Shimokawa, T., Yamamoto, S., Harada, C., Takazaki, Y., et al. 2002. A comprehensive rice transcript map containing 6591 expressed sequence tag sites. Plant Cell, **14**: 525–535.

Yu, J.K., Dake, T.M., Singh, S., Benscher, D., Li, W., Gill, B., and Sorrells, M.E. 2004. Development and mapping of EST-derived simple sequence repeat markers for hexaploid wheat. Genome, **47**(5): 805–818.

© 2006 NRC Canada