Theor Appl Genet (2006) 113:1093–1102
DOI 10.1007/s00122-006-0369-0

ORIGINAL PAPER

A genetic linkage map for tef [*Eragrostis tef* (Zucc.) Trotter]

Ju-Kyung Yu · Ramesh V. Kantety · Elizabeth Graznak · David Benscher · Hailu Tefera · Mark E. Sorrells

Received: 10 March 2006 / Accepted: 7 July 2006 / Published online: 10 August 2006

Abstract Tef [*Eragrostis tef* (Zucc.) Trotter] is the major cereal crop in Ethiopia. Tef is an allotetraploid with a base chromosome number of 10 ($2n = 4x = 40$) and a genome size of 730 Mbp. Ninety-four F_9 recombinant inbred lines (RIL) derived from the interspecific cross, *Eragrostis tef* cv. Kaye Murri × *Eragrostis pilosa* (accession 30-5), were mapped using restriction fragment length polymorphisms (RFLP), simple sequence repeats derived from expressed sequence tags (EST–SSR), single nucleotide polymorphism/insertion and deletion (SNP/INDEL), intron fragment length polymorphism (IFLP) and inter-simple sequence repeat amplification (ISSR). A total of 156 loci from 121 markers was grouped into 21 linkage groups at LOD 4, and the map covered 2,081.5 cM with a mean density of 12.3 cM per locus. Three putative homoeologous groups were identified based on multi-locus markers. Sixteen percent of the loci deviated from normal segregation with a predominance of *E. tef* alleles, and a majority of the distorted loci were clustered on three linkage groups. This map will be useful for further genetic studies in tef including mapping of loci controlling quantitative traits (QTL), and comparative analysis with other cereal crops.

Keywords Cereal crop · Molecular marker · Genetic mapping · Segregation distortion

Communicated by C. Gebhardt.

Electronic Supplementary Material Supplementary material is available to authorised users in the online version of this article at http://dx.doi.org/10.1007/s00122-006-0369-0.

J.-K. YuR. V. Kantety · E. Graznak · D. Benscher · M. E. Sorrells (✉)
Department of Plant Breeding and Genetics,
Cornell University, 240 Emerson Hall, Ithaca,
NY 14853, USA
e-mail: mes12@cornell.edu

H. Tefera
Debre Zeit Agricultural Research Center,
P.O. Box 32, Debre Zeit, Ethiopia

Present Address:
R. V. Kantety
Department of Plant and Soil Science,
Alabama A&M University, 134 ARC Building,
P.O. Box 1927, Normal, AL 35762, USA

Introduction

Construction of genetic linkage maps provides information about the genome structure of a species, forms a foundation for studying simple and complex traits, and allows for comparative genetic studies. Thus, well-developed and broadly useful linkage maps are a valuable resource for studying economically important species such as cereals. A few examples of the maps that formed the basis for studying grass genomes at the molecular level include barley (*Hordeum vulgare*) (Heun et al. 1991), hexaploid wheat (*Triticum aestivum*) (Nelson et al. 1995a, b, c; Yu et al. 2004a), rice (*Oryza sativa*) (Causse et al. 1994; Temnykh et al. 2000; Yu et al. 2004b), maize (*Zea mays*) (Gardiner et al. 1993) and oat (*Avena sativa*) (O'Donoughue et al. 1995). Subsequently, comparative maps were developed for a number of grass species including rice, maize, oat and the *Triticeae* based on common molecular probes mapped in multiple species (reviewed in Gale and Devos 1998; Devos 2005).

Tef [*Eragrostis tef* (Zucc.) Trotter] is an allotetraploid with a base chromosome number of 10 ($2n = 4x = 40$) and it belongs to the subfamily *Chloridoideae*. The genus *Eragrostis* contains about 350 species and the closest relative of tef is *E. pilosa* (L.) P. Beauv. Tef has an estimated genome size of 730 Mbp and the smallest chromosome size reported for the *Poaceae* family members (Ayele et al. 1996). Tef is the most important cereal grain in Ethiopia occupying two million hectares in 2003–2004, which was 28% of eight cereal crops grown in the country (CSA 2004). Tef contains high levels of proteins with an excellent balance among the essential amino acids and is also rich in mineral content (Katema et al. 1993). In addition, it is adapted to a wide range of climatic and soil conditions and shows an exceptional level of resistance to biotic stresses.

PCR-based molecular linkage maps have been developed in a number of allotetraploid species such as cotton (*Gossypium hirsutum*) (Han et al. 2006), white clover (*Trifolium repens*) (Barrett et al. 2004) and durum wheat (*Triticum turgidum*) (Korzun et al. 1999). However, genetic information for tef especially at the molecular level is limited with one linkage map based on amplified fragment length polymorphism (AFLP) using an intraspecific population (Bai et al. 1999) and the other based on RFLP with an interspecific population (Zhang et al. 2001). A map using PCR-based markers is needed to facilitate future mapping studies and to improve the efficiency of crop breeding, especially for quantitative traits. The difficulty in cross-referencing AFLP markers between populations and the low resolution of the RFLP maps and high marker cost limits their utility for comparative genetics and marker-assisted selection (MAS).

The objective of this study was to construct a molecular linkage map of tef using markers based on heterologous and tef cDNA clones (RFLP), ISSR, heterologous primers derived from wheat, rice and finger millet, and primers derived from tef expressed sequence tags (EST–SSR), single nucleotide polymorphism/insertion and deletion (SNP/INDEL), intron fragment length polymorphism (IFLP) and targeted region amplification polymorphism (TRAP). This linkage map can be used for mapping quantitative trait loci (QTL) of agronomic importance and in marker-assisted selection for the tef improvement.

Materials and methods

Plant material

Ninety-four F_9 recombinant inbred lines (RIL) were randomly selected from a population derived from a cross of *E. tef* cv. Kaye Murri × *E. pilosa* (accession 30-5) that was developed at the Debre Zeit Agricultural Research Center in Ethiopia. The female parent, *E. tef*, is a tall, thick-culmed, late maturing, white-seeded cultivar with red lemmas and compact panicles. In contrast, *E. pilosa* is a short, thin-culmed, early maturing, brown-seeded accession with grayish lemmas and loose panicles.

DNA markers

Tef cDNA clones (TCD) were provided by Dr. Henry Nguyen at Texas Tech. University, Lubbock, TX, USA (Zhang et al. 2001). The heterologous ''anchor'' probes used in this study were derived from cDNA libraries of rice (RZ, Causse et al. 1994), barley (BCD, Heun et al. 1991), oat (CDO, O'Donoughue et al. 1995) and maize (CSU, Gardiner et al. 1993), and were extensively used for comparative genetic studies in grasses (Gale and Devos 1998: Van Deynze et al. 1995). In addition, genes involved in the lignin biosynthetic pathway (for lodging related trait) and other agronomically important genes (for plant height, branching and seed color) were amplified, cloned and used as RFLP probes. Primer pairs were designed for the following gene sequences: phenylalanine ammonia-lyase (*PAL*) from rice (X16099); cinnamic acid 4-hydroxylase (*C4H*) from sorghum (AY034143); cinnamoyl-CoA reductase (*CCR*) from maize (Y15069); coumarate O-methyltransferase (*COMT*) from maize (M73235); ferulate-5-hydroxylase (*F5H*) from Arabidopsis (AF068574); gibberellic acid insensitive dwarfing gene (*Rht-D1*) from wheat (AJ242531) and maize (dwarf8–*D8* gene from AF413162); teosinte branched1 (*Tb1*) from maize (U94494); and anthocyanin biosynthesis gene (*A1*, dihydroflavonol reductase) from maize (X05068).

Tef EST-derived molecular markers, described in Yu et al. (2006), including EST–SSR, SNP, IFLP and TRAP were mapped. The heterologous PCR primer pairs (CNL, DupW, KSUM, RM, SRSC/PRSC, inf/lfm) derived from ESTs of wheat, rice or finger millet (*Eleusine coracana*) were tested and mapped on tef (Temnykh et al. 2000; Feltus et al. 2006; Yu et al. 2004a). Eleven ISSR primers were designed based on known SSR motifs of UBC set #9 and were purchased from the biotechnology laboratory at University of British Columbia, Canada (Kantety et al. 1995). Overall marker information is summarized in Table 1 and the information regarding tef specific markers and candidate gene markers is in Supplementary Table 1.

DNA extraction and southern hybridization

The population, along with the parents, was planted in the greenhouse and 3 to 4 week-old leaf tissue was harvested for DNA extraction. Ten micrograms of DNA per digestion were used with each of the 11 restriction enzymes, *BamH*I, *Dra*I, *EcoR*I, *EcoR*V, *Hae*III, *Hind*III, *Kpn*I, *Pst*I, *Sca*I, *Xba*I and *Xho*I, for parental survey, and then the mapping filters were prepared with six restriction enzymes, *BamH*I, *Dra*I, *EcoR*I, *EcoR*V, *Hind*III and *Xba*I, which covered most of the polymorphisms. The DNA hybridization including the RFLP procedure used here was adapted from Nelson et al. (1995a).

PCR and gel electrophoresis

PCR reactions were carried out in a 25 μL solution containing 25 ng of genomic DNA, 5.5 pMol of each primer, 1.5 mM $MgCl_2$, 0.125 mM of each dNTP, 10× reaction buffer, 0.8 units of *Taq* polymerase. The PCR profile and the TRAP technique were adapted from Yu et al. (2006) and Hu and Vick (2003), respectively. The amplified products were fractionated, using 4% polyacrylamide gel electrophoresis (PAGE), under denaturing conditions and visualized by silver staining. In addition, the amplicons of SNP/INDEL were screened on a single-strand conformational polymorphism (SSCP) gel (Martins-Lopes et al. 2001).

A lower case letter was added as a suffix to the marker name when multiple bands were polymorphic and then mapped, from largest fragment size to the smallest fragment size.

Genetic mapping

A total of 192 loci from 142 molecular markers were scored on the population and tested for goodness of fit to 1:1 segregation ($P > 0.05$). Linkage analyses were performed using G-Mendel 3.0 (Holloway and Knapp 1993) and Map Manager QTX13 (Manly and Olson 1999). Loci were grouped and ordered using a likelihood odds (LOD) threshold of 4.0 and a recombination frequency threshold of 0.25 under G-Mendel and type-I error probability of 0.001 under Map Manager QTX13. The likelihoods of different locus orders were compared and the locus-order estimate with the highest likelihood was selected for each linkage group. Map distances (cM) were calculated using the Kosambi mapping function (Kosambi 1944). The length of the tef genome was estimated by $L + (2tL)/n$, as proposed by Fishman et al. (2001), and using $\sum L_i(k_i + 1/k_i - 1)$ as proposed by Chakravarti et al. (1991), where L is the length of the map (cM), $n = k - t$ is the number of marker loci intervals, k is the number of marker loci, k_i is the number of marker loci on the ith linkage group, and $i = 1, 2,...,t$, and $t = 21$ is the number of linkage groups generated. The proportion c of the genome within d cM of a marker locus, assuming that random distribution of polymorphic markers was estimated using relationship $c = 1 - e^{-2dk/L}$ (Lange and Boehnke 1982; Chakravarti et al. 1991).

Table 1 Summary of the different types and sources of DNA markers used for the construction of the *E. tef* × *E. pilosa* map (see Materials and methods)

Marker name	Originating species	Clone/sequence	Type	Reference	No. of markers	No. of loci
BCD	Barley	cDNA	RFLP	Heun et al. (1991)	6	9
CDO	Oat	cDNA	RFLP	O'Donoughue et al. (1995)	6	6
CSU	Maize	cDNA	RFLP	Gardiner et al. (1993)	4	6
RZ	Rice	cDNA	RFLP	Causse et al. (1994)	27	36
Genes	Various species	Coding region	RFLP	–	6	7
TCD	Tef	cDNA	RFLP	Zhang et al. (2001)	26	36
TCD	Tef	cDNA sequence	IFLP	–	2	2
CNLT	Tef	EST	EST–SSR	Yu et al. (2006)	15	17
CNLT	Tef	EST	SNP/INDEL	Yu et al. (2006)	2	2
CNLT	Tef	EST	IFLP	Yu et al. (2006)	5	5
CNLT	Tef	EST	TRAP	Yu et al. (2006)	6	7
CNL	Wheat	EST	EST–SSR	Yu et al. (2004a)	3	4
KSUM	Wheat	EST	EST–SSR	Yu et al. (2004a)	4	4
DupW	Wheat	EST	EST–SSR	–	3	3
RM	Rice	EST	SSR	Temnykh et al. (2000)	10	14
PRSC/SRSC	Rice	EST	IFLP	Feltus et al. (2006)	3	3
inf/lfm	Finger millet	EST	IFLP	–	3	3
ISSR	Universal		ISSR	Kantety et al. (1995)	11	28
Total					142	192

Results

Molecular marker polymorphisms

RFLP polymorphism

A total of 177 heterologous RFLP probes, including grass ''anchor'' markers was screened for polymorphism among the parental lines. Polymorphism varied for the probes originating from different species; maize probes identified the fewest polymorphisms (26.7%) while barley probes identified the most (68.8%) followed by rice (54.6%). The overall RFLP polymorphism of heterologous probes was 49.7% (88 markers), which is lower than the previously reported level of 64% among the same parental lines (Zhang et al. 2001) and may be due to the different restriction enzymes used in that study. Of nine candidate genes, *PAL*, *COMT*, *A1, D8, Rht1* and *Tb1* were polymorphic.

Zhang et al. (2001) screened 151 tef cDNA probes (TCD) on mapping parents and reported 105 polymorphic RFLP probes (69.5%). Of those polymorphic markers, 26 TCD markers were used for mapping in this study. Subsequently, 126 RFLP probes were polymorphic on tef mapping parents.

Tef EST-derived marker polymorphism

Yu et al. (2006) have developed markers derived from tef EST sequences including EST–SSR, SNP/INDEL and IFLP (Supplementary Table 1). In their study, 170 ESTs containing SSRs were identified, and 80 EST–SSR markers were developed and evaluated on tef mapping parents. Nineteen primer sets (23.8%) were polymorphic and could be mapped. This was relatively low polymorphism compared to RFLP (69.5%) for tef (Zhang et al. 2001). In addition, 18 SNP/INDEL and 34 IFLP markers were developed by Yu et al. (2006). Ten were validated as polymorphic markers by SSCP and six of the IFLPs showed intron length polymorphism between the mapping parents used in this study. To increase the efficiency of mapping, the remaining monomorphic IFLP markers were screened using an alternative strategy known as TRAP (Hu and Vick 2003); thus, six monomorphic IFLP markers were recovered as polymorphic TRAP markers.

Two TCD-IFLP markers were developed by sequencing cDNA clones of TCD308 and TCD424 and then aligning the tef cDNA and rice genomic DNA sequences to predict putative tef intron sites. As a result, 37 tef specific sequence-derived markers were polymorphic.

Heterologous EST-derived marker polymorphism

EST-derived markers derived from different species were screened on the tef parents, including wheat (Yu et al. 2004a), rice (Temnykh et al. 2000; Feltus et al. 2006) and finger millet (K. Devos, personal communication). A total of 236 wheat derived EST–SSRs (73 CNL, 130 KSUM and 33 DupW) was tested for amplification in tef and 84 wheat EST–SSRs (35.6%) produced tef PCR amplicons (17 CNL, 51 KSUM and 16 DupW). Of these, four CNL, eight KSUM and three DupW markers were polymorphic on tef parents (17.9%). Also, out of 76 rice EST–SSRs (RM) developed by Temnykh et al. (2000), 51 were functional (67.1%), and ten markers were polymorphic (19.6%).

We also used EST-derived markers that were developed for length or sequence differences in putative intron regions of rice (SRSC/PRSC) and finger millet (inf/lfm). Out of 92 functional rice EST markers (Feltus et al. 2006), three showed length polymorphism between the tef parents. Similarly, out of 87 finger millet primers provided (K. Devos), 67 were functional (77%), and three were polymorphic based on length differences. This proportion of marker functionality (77%) was higher than EST–SSRs for wheat (35.6%) or rice (67.1%), and comparable to tef specific markers (83.6%; Yu et al. 2006). Finger millet belongs to the same subfamily, the *Chloridoideae*, as tef and that may explain this high marker transferability. Comparative mapping between finger millet and rice showed that the *Eleusine* and rice genomes are highly colinear, and that rearrangements are specific either to finger millet or to the *Chloridoideae* tribe (Devos 2005). Subsequently, 31 heterologous EST-derived markers were polymorphic on tef mapping parents.

ISSR amplification and polymorphism

Screening of the two parental lines with 23 ISSR primers gave 5 to 32 distinct bands with an average of 12 fragments per reaction. Eleven of the 23 ISSR primers revealed clean banding patterns and polymorphism between the mapping parents and these were used to generate 129 amplification fragments on the population. Of these, 28 mappable polymorphic bands with an average of 2.6 polymorphic fragments per primer were generated.

Marker loci segregation

Of the 88 grass anchor probes and six candidate genes that were polymorphic for the parents, 49 were used for mapping, and these detected a total of 64 loci in this

study. In addition, 26 tef cDNAs previously mapped by Zhang et al. (2001) were used to map 36 loci. Out of 100 EST-derived, polymorphic markers, 28 tef and 26 heterologous EST-derived markers produced 31 and 31 segregating loci, respectively. In addition, two sequenced tef cDNAs were mapped using intron size difference detecting two loci. Eleven ISSR primers were scored and resulted in 28 mapped loci (Table 1). Overall, 75 restriction enzyme/probe combinations were used to detect 100 loci and 67 PCR-based markers were used to detect 92 loci (Table 1).

Of the 131 markers detecting 164 loci, not including 11 ISSRs, 103 markers generated a single polymorphic locus and 28 markers generated two or more polymorphic loci. Two PCR-based marker loci (DupW4 and CNLT7) were segregating as single null alleles that may result from primer annealing failures caused by sequence polymorphisms in primer sites. Of these 103 single locus markers, 45 were co-dominant and 58 were dominant. Subsequently, 86 polymorphic markers (65.6%) detected two or more loci.

Tef linkage map construction

Based on 192 segregating loci from 142 markers in 94 F_9 RILs, the genetic linkage map was constructed using 156 loci while 36 loci were unlinked (Table 2). The 156 loci from 121 markers coalesced into 21 linkage groups of two or more markers ranging in length from 6.5 to 305 cM (Table 2, Fig. 1). The linkage groups were numbered consecutively in descending order of the genetic distance covered by each linkage group (Fig. 1). The number of marker loci per linkage group ranged from two on LG21 to 25 on LG1. The linked loci covered 2,081.5 cM and the distance between two loci ranged from 0.6 to 39 cM with an average distance of 12.3 cM (Table 2). The density of loci varied within each linkage group; LG21 was the most dense (3.3 cM/locus), while LG5 was the least dense (19.4 cM/locus) (Table 2). The longest gap (39 cM) was between RZ460 and TCD197b on LG5 (Fig. 1). Loci were generally well distributed except for clusters in the central regions of LG2 and LG8, and in the distal region of LG6 (Fig. 1).

Table 2 Number of mapped loci, linkage group length and density for each linkage group in the tef genetic map

Linkage group	Loci	Map distance (cM)	Density (cM/locus)
1	25	305.0	12.2
2	16	257.3	16.1
3	13	179.6	13.8
4	10	158.4	15.8
5	8	155.4	19.4
6	11	135.4	12.3
7	10	130.0	13.0
8	10	107.3	10.7
9	6	91.2	15.2
10	7	88.1	12.6
11	5	87.8	17.6
12	5	85.4	17.1
13	5	76.6	15.3
14	5	49.3	9.9
15	4	46.7	11.7
16	3	34.6	11.5
17	4	32.6	8.2
18	2	22.8	11.4
19	3	13.8	4.6
20	2	17.7	8.9
21	2	6.5	3.3
Unlinked	36		
Total	192	2,081.5 cM	12.3 cM

This tef linkage map is represented by 21 linkage groups and 36 unlinked loci rather than the 20 chromosomes ($2n = 4x = 40$) expected (Fig. 1). An estimated 95% of the total genetic distance of the map is covered by LG1 to LG15 while the remaining 128 cM was covered by six small linkage groups (LG16 to LG21). The proportion of the genome within d cM of a marker locus, assuming a random distribution of polymorphic marker loci, was estimated to be $1 - e^{-2dk/L} = 1 - e^{-2d(156)/(2,081.5)} = 0.777$ on *E. tef* × *E. pilosa* map (Lange and Boehnke 1982) thus, roughly 77.7% of the genome is within 10.0 cM and 95% of the genome is covered at 20.0 cM intervals.

Of the 61 loci detected by 28 markers detecting multiple loci, 50 were mapped into linkage groups while 11 were unlinked. A majority of these loci (37) were interspersed in the different linkage groups. Fourteen loci from seven multi-locus markers were present on the same linkage groups but those were separated by at least 5 cM from each other. This may indicate the detection of gene families, intra-chromosomal duplications or rearrangements in the tef genome. By mapping multiple loci from the same markers, it was possible to identify putative homoeologous relationships between the linkage groups. Homoeologous chromosomes could be identified when multiple markers are mapped in the same linear order in different linkage groups. Three putative homoeologous pairs were identified; (i) LG2 and 4 connected by ISSR842 and CNLT146, (ii) LG3 and 6 connected by RM170 and ISSR549, and (iii) LG5 and 7 by ISSR811 and TCD230 (Fig. 1).

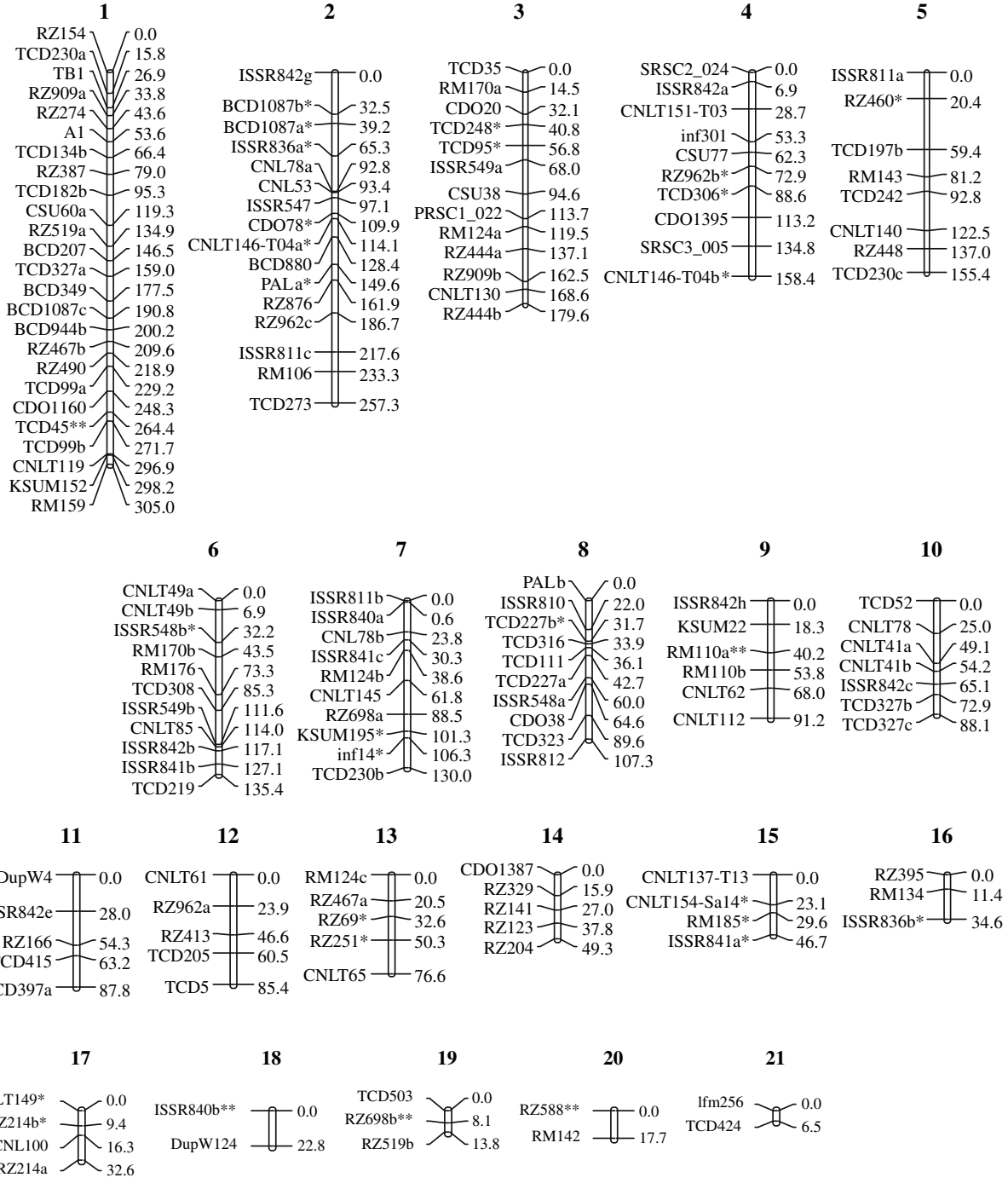

Fig. 1 Genetic linkage map of tef derived from 94 recombinant inbred lines from a cross between *E. tef* cv. Kaye Murri and *E. pilosa* (30-5). Linkage groups were denoted as 1 to 21. Loci names and map distance are shown to the left and right of each group, respectively. Loci names with an *asterisk* indicate significantly distorted segregation ($P < 0.05$), and *single asterisk* and *double asterisks* indicate the preferential transmission from the alleles of *E. tef*, or *E. pilosa*, respectively

Segregation distortion in tef map

The majority of the marker loci fit the expected 1:1 segregation ratio while 16% of the marker loci (31) exhibited segregation distortion based on a chi-square test ($P < 0.01$); among them, 26 loci favored the female parent. RFLP and PCR-based markers exhibited similar levels of distortion, with 17% for RFLP loci and 15% for PCR-based markers. This shows that segregation distortion was not due to the type of marker but

rather to an uneven segregation of the gametes or zygotes.

In order to evaluate the direction of segregation distortion, the frequency of *E. tef* genotypes with data from an average of RIL progeny was plotted for the 16, 13 and 10 loci on LG2, 3 and 4, respectively; which comprised 38% of all significantly distorted mapped loci (Fig. 2). The regions of segregation distortion for those three linkage groups resulted from an excess *E. tef* alleles and a deficiency of *E. pilosa* alleles. Ten marker loci spanning 117.1 cM on the short arm of LG2 had a deficiency of *E. pilosa* genotypes. The loci with the most segregation distortion in this region were BCD1087b (P = 2.7 e–06) and CNLT146-T04a (P = 6.2 e–06), suggesting that these are linked to genes causing segregation distortion loci (Fig. 2). As a result, on LG2, selection was against *E. pilosa* in the upper distal and central regions, and the frequency of *E. tef* alleles gradually decreased on each side of the peak. On LG3, there was an excess of *E. tef* alleles on the short arm up to 56.8 cM (TCD248 and TCD95); then, selection favored *E. pilosa*. However, in the long arm distal region on LG3, selection was against *E. pilosa* again (Fig. 2). On LG4, selection operated against *E. pilosa* alleles in the upper distal region and extended through the central region up to the locus TCD306. In LG4, CDO1395 had an excess of *E. pilosa* alleles but adjacent loci, TCD306 and SRSC3_005 have a predominance of *E. tef* alleles (Fig. 2). Similarly, TCD45 mapped in the distal region of LG1 had a significantly distorted segregation ratio (P = 2.9 e-05), resulting in a deficiency of *E. tef* alleles (27.6%) but adjacent marker loci were not skewed (Fig. 1). As a result, the significant distortion of TCD45 and CDO1395 loci might be by random chance. The overall patterns of segregation distortion on LG2, 3 and 4 suggested the presence of selection or segregation distorter genes flanking regions of high recombination (Lyttle 1991; Faris et al. 1998); such patterns are common in plants and differential transmission of gametes.

Discussion

Among all types of markers described herein (excluding ISSRs), 65.6% of polymorphic markers detected two or more loci. These markers were developed based on DNA sequences from coding regions, more than one third of which are typically duplicated genes (or gene families). Such gene duplication can arise from polyploidization events such as those believed to have preceded the origin of many plant species (reviewed in Wendel 2000). Thus, multi-locus markers derived from coding regions are common in plant genomes. Out of 142 markers, 45 segregated as co-dominant markers and the overall heterozygosity was 1.2%, which was not significantly higher ($P > 0.05$) than the expected

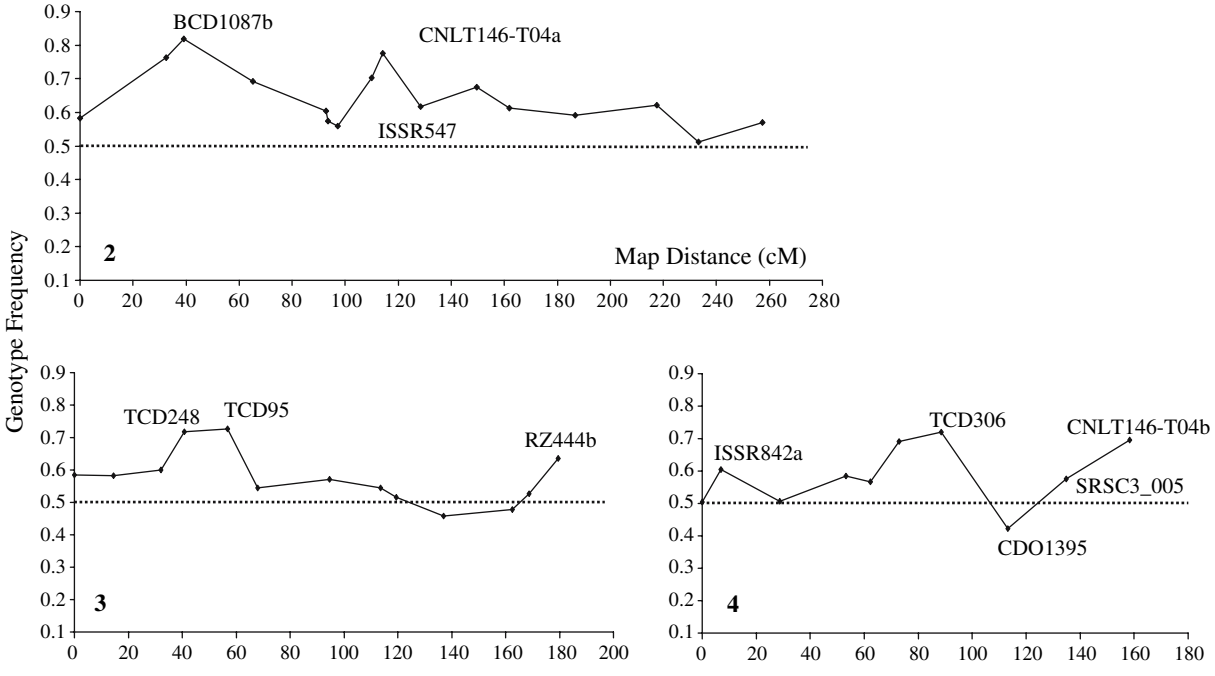

Fig. 2 Frequencies of *E. tef* genotypes in RILs derived from the cross of *E. tef* × *E. pilosa* are plotted along the genetic linkage group 2, 3 and 4 which demonstrated the most segregation distortion

heterozygosity of 0.39% in an F_9 generation. This heterozygosity, however, is reduced by half from the 2.5% heterozygosity in the F_8 map reported by Zhang et al. (2001).

The AFLP technique is similar to ISSR in that they both amplify a large number of fragments per primer, and hence, increase the chance of finding informative markers for mapping. Bai et al. (1999) reported that three AFLP polymorphic fragments per primer combination were detected from two *E. tef* cultivars. In addition, another tef linkage map (*E. tef*, cv. DZ-01-2785 × *E. pilosa*) currently under development using AFLP markers showed 2.9 polymorphic bands per primer combination (S. Chanyalew, personal communication). The ISSR markers described herein generated 2.6 polymorphic fragments per ISSR primer. Therefore, ISSR seems to be a comparable technique for generating tef molecular markers but has the advantage that DNA digestion is not required.

Several polymorphic markers were not used for mapping although they were evaluated on the mapping progeny because (i) they produced complicated polymorphic banding patterns, (ii) the intensity of polymorphic bands were significantly weaker than other bands, (iii) missing data were higher than 30%, or iv) no paternal type was recovered from the mapping population. Seven markers detected only homozygous female parental type (AA): no allele from the male parent, either as a homozygous or heterozygous type. This phenomenon may be explained if the markers were either a fragment from cytoplasmic DNA or monomorphic between the gametes that formed the mapping population. Unfortunately, DNA from the F1 plant was not available for testing these possibilities.

The development of a saturated genetic linkage map in tef is the first step in understanding the genetic control of traits of agronomic interest. The map described herein is comprised 156 loci from 121 markers that coalesced into 21 linkage groups. This genetic map totaled 2,081.5 cM and is estimated to cover approximately 79 or 76% of the tef genome (2,081.5/2,642.0 = 0.788 or 2,081.5/2741.2 = 0.759, Fishman et al. (2001) and Chakravarti et al. (1991), respectively). However, this estimate of genome coverage is likely to be over-estimated because of the large number of unlinked marker loci (36) compared to linked loci (156). Zhang et al. (2001) constructed a tef map based on RFLP using the same RIL population reported here. Their map comprised 149 loci on 20 linkage groups covering 1,489 cM. We estimate that their RFLP map covered about 88% of the tef genome by the same calculations (Chakravarti et al. 1991; Fishman et al. 2001). Similarly, the genome length of the tef linkage map for the intraspecfic cross F_5 RILs using AFLP, (Bai et al. 1999) covered 2,149 cM representing an estimated 81% of the genome (Chakravarti et al. 1991; Fishman et al. 2001).

The tef map reported here and the tef map reported by Zhang et al. (2001) share a total of 37 RFLP loci in common and, with the exception of seven loci, were in agreement with their grouping on linkage groups. Deviations for loci may be caused by use of different restriction enzymes resulting in the mapping of a different locus. In addition, the distances between loci were increased to the point that they were no longer linked, which may be caused by the different sizes of the mapping population. Of 20 linkage groups in their RFLP map, the first two groups contained 36% of total linked loci and covered 44% of the entire linkage map (Zhang et al. 2001). In their map, LG A and LG B have large gaps between marker loci (46 cM on LG A, and 41.8 cM and 31.5 cM on LG B). Based on the comparison of common RFLP loci, the tef map reported here indicated that the widest gaps on LG A and B of their map were broken at LOD 3 and corresponded to LG1 and LG13, and LG3 and LG4, respectively (Fig. 1). Our map data indicated that RZ490 on LG A and RZ962 on LG B in their RFLP map were mapped at low LOD, and thus, may result in different linkage groupings.

The advantage of mapping loci from cDNA probes or EST sequences is the potential for close association with loci controlling variation for the traits of interest. Among EST-derived markers of tef, CNLT125 to CNLT158, were developed for mapping candidate genes that may be related to lodging resistance (Yu et al. 2006). In order to identify candidate gene ESTs, tef ESTs were compared to the NCBI nr database using BLASTX with an expected value less than or equal to $1e^{-20}$, and > 80% similarity over a minimum of 100 amino acids. Nine loci of eight markers were mapped in tef linkage groups representing six candidate genes. In addition, the locations of two candidate gene loci for lignin biosynthesis and one locus for teosinte branched 1 using RFLP probes were also identified. The information on candidate gene loci could be useful for evaluation of co-location between QTLs and functionally associated genetic loci, as well as the comparative analyses of genes of interest between tef and other species, such as rice or wheat (Yu et al. 2004b).

In the current study, 16% segregation distortion is similar to the 15% segregation distortion observed by Zhang et al. (2001) in an F_8 tef inter-specific population; but higher than the 11.5% segregation distortion observed by Bai et al. (1999) in the F_5 intra-specific population at $P < 0.05$. Segregation distortion has been observed in mapping of many species, especially in

inter-specific crosses of rice (Xu et al. 1997), durum wheat (Blanco et al. 1998) and lentil (Duran et al. 2004). Our linkage analysis of distorted loci showed that 29 out of 31 were scattered on 16 linkage groups but 32% of the loci were clustered on LG2 (Fig. 1). Among those included in linkage groups, there was a tendency to form small clusters in the central (LG2, 3 and 4) or distal regions (LG2 and 7) of the groups (Fig. 1). The clustering of distorted loci has been reported for other cereal crops such as rice (Temnykh et al. 2000) and wheat (Faris et al. 1998). The other distorted loci (2 out of 31) were unlinked.

For 192 segregating loci, the mean segregation ratio across the genome was 1.18 *AA*: 0.82 *BB*, where *AA* is the *E. tef* homozygote and *BB* is the *E. pilosa* homozygote. In addition, loci from the *E. tef* exceeded those of the *E. pilosa* in 85% of the cases of segregation distortion, which was significantly greater than the expected ratio ($P < 0.01$). These results strongly suggest the presence of one of more genes affecting transmission from the maternal parent *E. tef* and this was also shown in the tef RFLP map derived using an earlier generation of the same cross (Zhang et al. 2001). However, distorted loci were derived almost equally from both parents in the intra-specific AFLP map constructed by Bai et al. (1999).

Mapping of additional markers may help to coalesce these linkage groups and unlinked loci into the expected number of 20 and help to identify homoeologous linkage groups. A comparative map between tef and other members of the *Poaceae* such as, rice, wheat and finger millet is under development. The comparative genetic information will benefit tef researchers working on the improvement of agronomic traits such as, lodging resistance, plant height, seed color, seed size, yield and yield components (Sorrells 2001).

Acknowledgments J.-K. Yu and R.V. Kantety contributed equally to the research described herein. The authors would like to acknowledge financial support from the McKnight Foundation Collaborative Crop Research Program (Grant # 00-933) and Hatch project 149419. We are grateful to Drs. Susan McCouch at Cornell University, Bikram Gill at Kansas State University, Andrew Paterson and Katrien Devos at University of Georgia for providing the primers of RM, KSUM, SRSC/PRSC and inf/lfm markers, respectively. We also wish to acknowledge R. Greene, C. Bermudez and C. T. Jayasuriya for their generous assistance.

References

Ayele M, Dolezel J, Van Duren M, Brunner H, Zapata-Arias FJ (1996) Flow cytometric analysis of nuclear genome of the Ethiopian cereal tef [*Eragrostis tef* (Zucc.) Trotter]. Genetics 98:211–215

Bai G, Tefera H, Ayele M, Nguyen HT (1999) A genetic linkage map of tef [*Eragrostis tef* (Zucc.) Trotter] based on amplified fragment length polymorphism. Theor Appl Genet 99:599–604

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

Blanco A, Bellomo MP, Cenci A, Giovanni CD, D'Ovidio R, Iacono E, Laddomada B, Pagnotta MA, Porceddu E, Sciancalepore A, Simeone R, Tanzarella OA (1998) A genetic linkage map of durum wheat. Theor Appl Genet 97:721–728

Causse MA, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu X, Ronald PC, Harrington SE, Second G, McCouch SR, Tanksley SD (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138:1251–1274

Central Statistical Authority (2004) Statistical abstract of Ethiopia 2004 (1996 EC). Addis Ababa, Ethiopia

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

Devos KM (2005) Updating the 'crop circle'. Curr Opin Plant Biol 8:155–162

Duran Y, Fratini R, Garcia P, Perez de la Vega M (2004) An intersubspecific genetic map of Lens. Theor Appl Genet 108:1265–1273

Faris JD, Laddomada B, Gill BS (1998) Molecular mapping of segregation distortion loci in *Aegilops tauschii*. Genetics 149:319–327

Feltus FA, Singh HP, Lohithaswa HC, Schulze SR, Silva TD, Paterson AH (2006) A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. Plant Physiol 140:1183–1191

Fishman L, Kelly AJ, Morgan E, Willis JH (2001) A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. Genetics 159:1701–1716

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

Gardiner GM, Melia-Hancock S, Hoisington DA, Chao S (1993) Development of a core RFLP map in maize using an immortalized F2 population. Genetics 134:917–930

Han Z, Wang C, Song X, Guo W, Gou J, Li C, Chen X, Zhang T (2006) Characteristics, development and mapping of *Gossypium hirsutum* derived EST–SSRs in allotetraploid cotton. Theor Appl Genet 112:430–439

Heun M, Kennedy AE, Anderson JA, Lapitan NL, Sorrells ME, Tanksley SD (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). Genome 34:437–447

Holloway JL, Knapp SJ (1993) Gmendel 3.0 users guide. Department of Crop and Soil Science, Oregon State University, OR

Hu J, Vick BA (2003) Targeted region amplification polymorphism: a novel marker technique for plant genotyping. Plant Mol Biol Rep 21:289–294

Kantety RV, Zeng X, Bennetzen JL, Zehr BE (1995) Assessment of genetic diversity in dent and popcorn (*Zea mays* L.) inbred lines using inter-simple sequence repeat (ISSR) amplification. Mol Breed 1:365–373

Ketema S (1993) Tef (*Eragrostis tef*): breeding, genetic resources, agronomy, utilization and role in Ethiopian agriculture. Institute of Agricultural Research, Addis Ababa, Ethiopia

Korzun V, Roder MS, Wendehake K, Pasqualone A, Lotti C, Ganal MW, Blanco A (1999) Integration of dinucleotide microsatellites from hexaploid bread wheat into a genetic linkage map of durum wheat. Theor Appl Genet 98:1202–1207

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

Lange K, Boehnke M (1982) How many polymorphic genes will it take to span the human genome? Am J Hum Genet 34:842–845

Lyttle TW (1991) Segregation distorters. Annu Rev Genet 25:511–557

Martin-Lopes P, Zhang H, Koebner R (2001) Detection of single nucleotide mutations in wheat using single strand conformation polymorphism gels. Plant Mol Biol Rep 19:159–162

Manly KF, Olson JM (1999) Overview of QTL mapping software and introduction to Map Manager QT. Mamm Genome 10:327–334

Nelson JC, Van Deynze AE, Autrique E, Sorrells ME, Liu YH, Merlino M, Atkinson M, Leroy P (1995a) Molecular mapping of wheat: homeologous group 2. Genome 38:517–524

Nelson JC, Van Deynze AE, Autrique E, Sorrells ME, Liu YH, Negre S, Bernard M, Leroy P (1995b) Molecular mapping of wheat: homeologous group 3. Genome 38:525–533

Nelson JC, Sorrells ME, Van Deynze AE, Liu YH, Autrique E, Atkinson M, Bernard M, Leroy P (1995c) Molecular mapping of wheat: major genes and rearrangements in homeologous groups 4, 5 and 7. Genetics 141:721–731

O'Donoughue LS, Kinian SF, Rayapati PJ, Panner GA, Sorrells M, Tanksley SD, Phillips RL, Rines HW, Lee M, Fedak G, Molnar SJ, Hoffman D, Salas CA, Wu B, Autrique E, Van Deynze A (1995) A molecular linkage map of cultivated oat. Genome 38:368–380

Sorrells ME (2001) Comparative genomics for tef improvement. In: Tefera H, Belay G, Sorrells M (eds). Narrowing the rift: tef research and development. Proceedings of the international workshop on tef genetics and improvement, 16–19 October 2000, Addis Ababa, Ethiopia, pp 109–119

Temnykh S, Park WD, Ayers N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000) Mapping and genome organization of microsatellites in rice (*Oryza sativa*). Theor Appl Genet 100:698–712

Van Deynze AE, Dubcovsky J, Gill KS, Nelson JC, Sorrells ME, Dvorak J, Gill BS, Lagudah ES, McCouch SR, Appels R (1995) Molecular-genetic maps for group 1 chromosomes of Triticeae species and their relation to chromosomes in rice and oat. Genome 38:45–59

Wendel JF (2000) Genome evolution in polyploids. Plant Mol Biol 42:225–249

Xu Y, Zhu L, Xiao J, Huang N, McCouch SR (1997) Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). Mol Gen Genet 253:535–545

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

Yu J-K, La Rota M, Kantety RV, Sorrells ME (2004b) EST derived SSR markers for comparative mapping in wheat and rice. Mol Genet Genomics 271:742–751

Yu J-K, Sun Q, La Rota M, Edwards H, Tefera H, Sorrells ME (2006) Expressed sequence tag analysis in tef [*Eragrostis tef* (Zucc.) Trotter]. Genome 49:365–372

Zhang D, Ayele M, Tefera H, Nguyen HT (2001) RFLP linkage map of the Ethiopian cereal tef [*Eragrostis tef* (Zucc.) Trotter]. Theor Appl Genet 102:957–964