Field Crops Research 118 (2010) 28–35

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Cross-amplification of EST-derived markers among 16 grass species

M. Zeid[a], J.K. Yu[a,b], I. Goldowitz[a], M.E. Denton[d], Denise E. Costich[d,e], C.T. Jayasuriya[a], M. Saha[c], R. Elshire[d], D. Benscher[a], F. Breseghello[f], J. Munkvold[a,g], R.K. Varshney[h], G. Belay[i], M.E. Sorrells[a,*]

[a] Dept. of Plant Breeding and Genetics, Cornell University, Ithaca, NY 14853, USA
[b] Syngenta Seeds Inc., 317 330th Street, Stanton, MN 55018, USA
[c] Forage Improvement Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
[d] Institute for Genomic Diversity, Cornell University, Ithaca, NY 14853, USA
[e] USDA-ARS, Robert W. Holley Center for Agriculture and Health, Ithaca, NY, USA
[f] Embrapa Rice and Beans, P.O. Box 179, Santo Antônio de Goiás, GO 75375-000, Brazil
[g] Boyce Thompson Institute for Plant Research, Tower Road, Ithaca, NY 14853, USA
[h] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India
[i] Ethiopian Institute of Agricultural Research, Debre Zeit Agricultural Research Center, P.O. Box 32, Debre Zeit, Ethiopia

ARTICLE INFO

Article history:
Received 3 February 2010
Received in revised form 26 March 2010
Accepted 27 March 2010

Keywords:
Poaceae
Expressed sequence tag (EST)
EST-SSR
CISP
Comparative mapping
Homology
Sequence conservation

ABSTRACT

The availability of a large number of expressed sequence tags (ESTs) has facilitated the development of molecular markers in members of the grass family. As these markers are derived from coding sequences, cross-species amplification and transferability is higher than for markers designed from genomic DNA sequences. In this study, 919 EST-based primers developed from seven grass species were assessed for their amplification across a diverse panel of 16 grass species including cereal, turf and forage crops. Out of the 919 primers tested, 89 successfully amplified DNA from one or more species and 340 primers generated PCR amplicons from at least half of the species in the panel. Only 5.2% of the primers tested produced clear amplicons in all 16 species. The majority of the primers (66.9%) were developed from tall fescue and rice and these two species showed amplification rate of 41.6% and 19.0% across the panel, respectively. The highest amplification rate was found for conserved-intron scanning primers (CISP) developed from pearl millet (91%) and sorghum (75%) EST sequences that aligned to rice sequences. The primers with successful amplification identified in this study showed promise in other grass species as demonstrated in differentiating a set of 13 clones of reed canary grass, a species for which very little genomic research has been done. Sequences from the amplified PCR fragments indicated the potential for the transferable CISP markers for comparative mapping purposes. These primer sets can be immediately used for within and across species mapping and will be especially useful for minor grass species with few or no available molecular markers.

1. Introduction

The grass family, *Poaceae*, is one of the largest families of flowering plants, with approximately 10,000 species in 700 genera. *Poaceae* surpasses all other botanical families in economic importance. Three grain crops, wheat (*Triticum aestivum*), rice (*Oryza sativa*) and corn (*Zea mays*), are the world's predominant food sources, but the family also includes several other less-researched crops. Tef (*Eragrostis tef*), for example, is a major staple food in Ethiopia, but almost unknown elsewhere. Turf and forage crops such as tall fescue (*Lolium arundinaceum*), Kentucky bluegrass (*Poa pratensis*) and bermudagrass (*Cynodon dactylon*), are vital to the rangeland management and lawn care industries generating millions of dollars in seed sales, but have limited genetic resources available. In addition, future biofuel crops such as switch grass (*Panicum virgatum*) and reed canary grass (*Phalaris arundinacea*) are also examples of crops that are definitely in need of more research to bring about the required improvements.

Genetic studies of these minor crops are hindered because of the scarcity of molecular markers available. Because marker development is laborious, time-consuming, and expensive, given the limited resources and researchers available for minor crops such as turf/forage species, it has lagged behind that of major and well-researched crops. Microsatellite markers (SSR) developed from genomic libraries (gSSR) have been widely used for mapping and population genetic analysis. This can be mainly attributed to their high level of polymorphism, abundance and dispersion throughout the genome, besides their codominant nature of inheritance and

* Corresponding author at: Dept. of Plant Breeding and Genetics, 240 Emerson Hall, Cornell University, Ithaca, NY 14853-1902, USA. Tel.: +1 607 255 2180; fax: +1 607 255 6683.
E-mail address: mes12@cornell.edu (M.E. Sorrells).

0378-4290/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.fcr.2010.03.014

reproducibility (Gupta and Varshney, 2000; Squirrell et al., 2003). The disadvantage of gSSR markers is the high initial cost of development and their low transferability across genera and beyond (Roa et al., 2000; Kindiger, 2006).

A large amount of coding sequence information has been generated by EST (expressed sequence tag) projects for gene discovery in several crop species, and deposited in the National Center for Biotechnology Information (NCBI) database. EST-based markers are derived from transcribed regions of the genome, which are more constrained with respect to sequence diversity since they code for functional proteins. For this reason, EST-derived markers are more likely to produce amplicons in multiple species than those designed from non-coding sequences (Yu et al., 2004; Zhang et al., 2005; Parida et al., 2006). By December 21st, 2009, NCBI had more than six million readily accessible ESTs in members of the *Poaceae*, 83% of which were derived from rice, maize or wheat. The large amount of genetic information available on the major grain crops, including maize and sorghum (Zhu and Buell, 2007), and the full genome sequences for rice (Yu et al., 2005) and brachypodium (Opanowicz et al., 2008) are useful resources that can be extended to less well-funded grasses using comparative genomics tools (Varshney et al., 2005; Feltus et al., 2006). Based on this information, one can search for variation in EST sequences to develop markers flanking SSRs, insertions and deletions (INDEL), and single nucleotide polymorphisms (SNP).

It has become widely accepted to screen EST-based markers derived from one species with other species in the same genus and even across genera within the same family. Gupta et al. (2003) reported that 24 out of 59 wheat EST-SSR markers amplified fragments in five species including barley, maize, oat, rice, and rye. Similarly, Wang et al. (2005), demonstrated the transferability of EST-SSR markers from maize, sorghum, rice and wheat to minor grass species (finger millet, seashore paspalum and bermudagrass) and observed the correlation between the transferability rate of markers and the phylogenetic relationship of the species tested. Numerous studies, however, have used the term "transferability of markers", which implies amplification of orthologous loci, to describe amplification of an amplicon regardless of orthology. While many studies have suggested that EST-SSR are most interesting because of their amplification of conserved (orthologous) sequences across different grass species (Varshney et al., 2005; Feltus et al., 2006), others have observed loss of sequence homology when markers developed from one species were screened on distantly related species (Asp et al., 2007; Sim et al., 2009).

The conserved-intron scanning primers (CISP) designed by Feltus et al. (2006) to conserved exonic regions flanking introns from sorghum/pearl millet ESTs and aligned to the rice genome, successfully amplified in barley, maize, tef and wheat. Those markers and others, such as the PCR-based landmark unique genes (PLUG) described by Ishikawa et al. (2009), are much more conserved than EST-SSR markers and could provide better resources for comparative mapping studies, provided they amplify orthologous sequences that are polymorphic. Thus, more research is needed on the level of transferability of molecular markers from well-researched cereal crops to distantly related, minor grass species and also on the nature of the products of those markers. It is crucial to understand whether those markers will only add novel markers to less-researched crops or will also provide the basis for comparative mapping work.

The objectives of this research were: (i) to evaluate the cross-amplification of 919 primers developed using EST sequences derived from wheat, rice, tef, sorghum, pearl millet, tall fescue, and rye on a panel of 16 grass species, (ii) to evaluate the utility of some of those markers in discriminating reed canary grass accessions, and (iii) to evaluate the transferability of markers for comparative mapping work of less-researched grass species.

2. Materials and methods

2.1. Plant materials

Sixteen grass (*Poaceae*) species, including maize (*Z. mays*), sorghum (*Sorghum bicolor*), pearl millet (*Pennisetum glaucum*), bermudagrass (*C. dactylon*), tef (*E. tef*), creeping bentgrass (*Agrostis stolonifera*), harding grass (*Phalaris aquatica*), oat (*Avena sativa*), brachypodium (*Brachypodium distachyon*), smooth bromegrass (*Bromus inermis*), barley (*Hordeum vulgare*), western wheatgrass (*Pascopyrum smithii*), wheat (*T. aestivum*), Kentucky bluegrass (*P. pratensis*), tall fescue (*Festuca arundinacea*) and rice (*O. sativa*), were selected to represent 10 tribes from four of the six subfamilies within the *Poaceae* family (Table 1). The subfamily *Pooideae* was represented by 11 species of grain, turf and forage crops in this panel. To evaluate the level of polymorphism transferable markers provide, and investigate the nature of the amplified fragments, 13 reed canary grass clones and Indian lovegrass (*Eragrostis pilosa* accession 30-5) were employed. Those clones represent cultivars and accessions from northeast and north central United States, Canada and one European cultivar (Table 2).

Table 1
A panel of 16 grass species used for evaluation of 919 primers developed from seven species (boldfaced).

Common name	Variety	Scientific name	Tribe	Subfamily
Maize	B73	*Zea mays*	*Andropogoneae*	*Panicoideae*
Sorghum	BTx623	*Sorghum bicolor*	*Andropogoneae*	*Panicoideae*
Pearl millet	Titft23A	*Pennisetum glaucum*	*Paniceae*	*Panicoideae*
Bermudagrass	Midland 99	*Cynodon dactylon*	*Cynodonteae*	*Chloridoideae*
Tef	Kaye Murri	*Eragrostis tef*	*Eragrostideae*	*Chloridoideae*
Creeping bentgrass	AA61	*Agrostis stolonifera*	*Aveneae*	*Pooideae*
Harding grass	Maru 20-2	*Phalaris aquatica*	*Aveneae*	*Pooideae*
Oat	Ogle	*Avena sativa*	*Aveneae*	*Pooideae*
Brachypodium	Bd3-1	*Brachypodium distachyon*	*Brachypodieae*	*Pooideae*
Smooth bromegrass	Lincoln 8-7	*Bromus inermis*	*Bromeae*	*Pooideae*
Barley	Morex	*Hordeum vulgare*	*Triticeae*	*Pooideae*
Western wheatgrass	Barton	*Pascopyrum smithii*	*Triticeae*	*Pooideae*
Wheat	Chinese Spring	*Triticum aestivum*	*Triticeae*	*Pooideae*
Rye[a]	–	*Secale cereale*	*Triticeae*	*Pooideae*
Kentucky bluegrass	SR2394	*Poa pratensis*	*Poeae*	*Pooideae*
Tall fescue	KY31	*Festuca arundinacea*	*Poeae*	*Pooideae*
Rice	IR64	*Oryza sativa*	*Oryzeae*	*Ehrhartoideae*

[a] Rye was only used as a source of markers in this study.

Table 2
Reed canary grass clones used for verification of amplification results and sequence analysis.

Accession/cultivar name[a]	Geographic origin
Brummer (cultivar)	IA, USA
Marathon (cultivar)	IA, USA
Vantage	IA, USA
Denton	NY, USA
Rensselaer Falls	NY, USA
Watkins Glen	NY, USA
State College	PA, USA
Juneau Ditch	WI, USA
Poygan Marsh	WI, USA
Rabbit Lane	WI, USA
Bellevue	Quebec, Canada
Rival (cultivar)	Manitoba, Canada
SW Bamse (cultivar)	Sweden

[a]Germplasm (except Denton) was provided by Dr. Michael D. Casler, USDA-ARS, U.S. Dairy Forage Research Center, Madison, WI, USA. Denton clone was collected at Aurora, NY, USA.

2.2. Primer resources

Markers developed from four subfamilies of the *Poaceae* family were used in this study. A total of 919 primers derived from seven grass species developed by four research institutes were tested. Almost 90% of the primers screened were developed from EST sequences harboring SSR repeats, while 10% were CISP markers (Table 3). A total of 602 primers (RM, REMS, CNL, PRSC/SRSC and CNLT), were tested at Cornell University, while the remaining 364 primers (NFFA and NFFS) were tested at the Noble Foundation (Table 3 and Supplemental Table I).

2.3. Primer screening and data analyses

DNA was isolated from fresh tissues as described in Tai and Tanksley (1990). A sample of 20 ng of template DNA was used for PCR amplification in a PTC-225 thermocycler (MJ Research) with an initial denaturation at 94 °C for 3 min, 35 cycles of 1 min at 94 °C, 1 min at 53 °C, and 2 min at 72 °C, and a final elongation for 10 min at 72 °C. A single annealing temperature was employed to accommodate all primers tested in this study. Reactions were carried out in 20 μL volumes containing 1× PCR buffer, 2.5 mM $MgCl_2$, 0.125 mM of each dNTP, 0.8 Units *Taq* DNA polymerase (Promega), and 5.5 μM of each primer pair. PCR amplicons were examined on 1.5–2% agarose gels with ethidium bromide staining. Amplification was evaluated as a binary score: clear, strong bands were scored as 1; weak bands and those with multiple bands were considered non-specific amplifications and were scored as 0. Scoring results are presented in Supplemental Table II. Primer cross-amplification was calculated as the average number of primers amplifying a product divided by the total number of primers screened (Table 4). Genetic similarity estimates between species on the panel were calculated according to Nei and Li (1979). Similarity coefficients were used to construct a UPGMA dendrogram using NTSYS-pc version 2.20h (Rohlf, 2002). Similarity coefficients between the six source species (excluding rye) and the 16 species on the panel were correlated with their corresponding cross-amplification rate values of EST markers ($N = (6 \times 16) - 6 = 90$ comparisons) using Pearson's correlation.

2.4. Length variation and sequence analysis of amplified fragments

Accessions of reed canary grass (Table 2) were screened with those primers that amplified a clear single band on agarose gels in harding grass (Supplemental Table II). Amplicons were then run on 4% denaturing polyacrylamide gels and stained with silver nitrate after replicating the PCR reaction mentioned above with a faster PCR program characterized by a higher annealing temperature and fewer cycles (30 cycles of 45 s at 94 °C, 1 min at 54 °C, and 1 min at 72 °C, and a final elongation for 7 min at 72 °C), to reduce the number of non-specific amplicons.

PCR products from 24 primers that amplified in reed canary grass and 7 that were previously mapped in tef (Yu et al., 2006) were cloned using TOPO TA cloning kit (Invitrogen). Recombinant clones were directly amplified using vector primers and 10 μL of the PCR products were screened on a 1% agarose gel for presence of inserts of the correct size before being sequenced using an ABI 3730XL automated DNA sequencer (Applied Biosystems, Foster City, CA, USA). In other instances, only the desired amplicon of the correct size was excised from the polyacrylamide gel (Chalhoub et al., 1997) and used as a template in another round of PCR before cloning and sequencing. Sequences were edited and aligned using the Chromas pro software version 1.22 (Technelysium Pty, Ltd.). The online tool ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) was used to align sequenced fragments to the reference sequence from which the markers were designed.

3. Results and discussion

A total of 919 primers from seven grass species were screened for their amplification across a diverse panel of 16 grass species (Table 1). Of those, 891 (97%) amplified a product with at least one species, 340 primers (35%) amplified in at least half of the species tested and 50 primers (5.2%) amplified across all 16 species on the panel (details are listed in Supplemental Table II). Only 119 primers (13%) amplified in a single species. On average, 314.7 primers amplified in each species, varying from 19.0% for rice primers to 91.1% for pearl millet CISP primers (Table 4), however, only 9 primers of the latter were tested. Gupta et al. (2003) observed similar average cross-amplification rates (40.7%) between bread wheat and five other cereals (barley, maize, oats, rice and rye). Similarly, Zhang et al. (2005) reported an average of 50% cross-amplification and transferability of markers between wheat and four cereals (*Agropyrum*, barley, rice and rye). In those studies, only cereal crops were included in the screening panel. The study presented here used a more diverse panel of grass species and found comparable cross-species amplification.

Table 3
Groups of primers screened in this study indicating the species for which the primers were originally designed, type of marker, and the institution that developed each marker.

Marker name	No. of primers	Species of origin	Marker type	Institution	Source
CNL	344	Rice, wheat	EST-SSR	Cornell Univ., USA	Yu et al. (2004)
CNLT	107	Tef	EST-SSR, SNP/INDEL, IFLP	Cornell Univ., USA	Yu et al. (2006)
RM	20	Rice	EST-SSR	Cornell Univ., USA	Temnykh et al. (2000)
SRSC	87	Sorghum/rice	CISP	Univ. Georgia, USA	Feltus et al. (2006)
PRSC	9	Pearl millet/rice	CISP	Univ. Georgia, USA	Feltus et al. (2006)
NFFA/NFFS	317	Tall fescue	EST-SSR, STS	Noble Found., USA	Saha et al. (2004)
REMS	35	Rye	EST-SSR	IPK, Germany	Khlestkina et al. (2004)

Table 4
Number of primers amplifying clear, strong bands and percent cross-amplification for seven grass species across a panel of 16 species.

Species of origin and marker type	No. of markers	*Pooideae*									*Panicoideae*			*Chloridoideae*		*Ehrhartoideae*	Average successful primers	Cross-amplification (%)
		Barley	Western wheatgrass	Wheat	Creeping bentgrass	Harding grass	Oat	Kentucky bluegrass	Tall fescue	Smooth bromegrass	Brachypodium	Maize	Sorghum	Pearl millet	Bermudagrass	Tef	Rice		
Wheat (EST)	81	34	48	–	22	22	27	22	23	27	27	16	15	21	20	11	21	23.73	29.30
Tef (EST)	107	15	29	12	17	22	20	19	24	20	24	13	34	22	34	–	13	21.20	19.81
Rice (EST)	282	50	75	35	46	68	74	74	69	55	58	26	33	61	53	27	–	53.60	19.01
Sorghum (CISP)	87	55	66	65	63	58	58	63	73	55	57	75	–	69	73	73	77	65.33	75.10
Pearl millet (CISP)	9	9	8	9	6	8	9	9	9	7	5	9	9	–	9	8	9	8.20	91.11
Tall fescue (STS and EST)	317	139	174	152	202	174	204	217	–	164	147	100	114	100	117	113	130	149.80	47.26
Rye (EST)	35	15	25	22	17	14	15	16	15	13	1	7	5	13	15	10	4	12.93	36.96
Total	919	317	425	295	373	366	407	420	213	341	319	246	210	286	321	242	254	314.69	

Genetic studies on forage and turf crops are limited by the number of markers available. One possible approach to acquiring new markers and facilitating comparative mapping in less-researched species is to screen markers already developed from better studied and related species. Limitations to this approach include the availability of such markers, their transferability, polymorphism, and nature of the amplified fragments in the less-researched species.

3.1. Marker availability

The rice genome sequence provides a readily available, rich source of molecular information. Out of the 282 rice EST-SSRs screened in this study, 220 (CNL160 to 403) were selected from 4726 rice EST-SSRs by La Rota et al. (2005) to cover the entire rice genome using in silico mapping against the rice genome sequence (Supplemental Table I). Rice primers tested here showed an average of only 19% amplification across the other 15 species, with no clear difference among subfamilies. Despite this low cross-amplification, the large number and precise sequence anchorage of rice EST markers make them a powerful resource for breeders and geneticists working on less-researched crops. Rice markers could provide a set of conserved orthologous markers to facilitate comparative analyses for cross-referencing genes and genomes (Varshney et al., 2005; Feltus et al., 2006).

Sorghum is another crop with a large amount of genetic information available in databases. All sorghum primers screened here were based on low-copy exons flanking intron regions conserved between sorghum and rice, which ensures higher amplification rate and transferability across grass species, as compared to EST-SSR and gSSR-based markers (Feltus et al., 2006). The sorghum/rice CISP primers showed a very high cross-amplification rate of 75% in this study. These markers appear to be limited in number and require considerable in silico and lab work to develop and verify as compared to EST-SSR primers, however, recently more of those intron flanking marker are becoming available (Ishikawa et al., 2009).

On the other hand, EST-derived primers from tall fescue, with only about 100,000 EST sequences available in Genbank (accessed December 22nd, 2009) showed 47.3% amplification across 15 species, which is comparable to the 57% reported by Saha et al. (2004) on five other grass species screened with only half of the tall fescue primers screened here, indicating that this set of primers amplify across a wide range of species and genera as compared to other tested EST-SSR primers (Table 4).

Marker-derived primers from tef, with nearly 4000 EST sequences currently available, showed 20% cross-amplification, comparable to the results for rice and wheat, as shown in Table 4. Sim et al. (2009) screened 22 of the tef primers tested here on ryegrass and found that 14 of these primers cross-amplified in ryegrass and six were included in their map. These results and the ones presented here suggest that all sources of markers available for less-researched species should be exploited to cope with the lag in marker information for those crops. In the present study, we identified an average of 314.7 (varying from 210 for sorghum to 425 for western wheatgrass) that can be immediately used for studies on a variety of turf and forage species (Table 4).

3.2. Primer cross-amplification and taxonomic relationship

The 317 primers from tall fescue tested for cross-amplification in this work showed 51% amplification in the nine species of its own subfamily (*Pooideae*). A large number of those primers amplified DNA of the closest relative Kentucky bluegrass (62%), which is from the same tribe, *Poeae*. The next highest amplification rate was observed for members of the *Aveneae* tribe. This supports the molecular analysis based on internal transcript space (ITS) sequences from nuclear rDNA that indicated a sister group

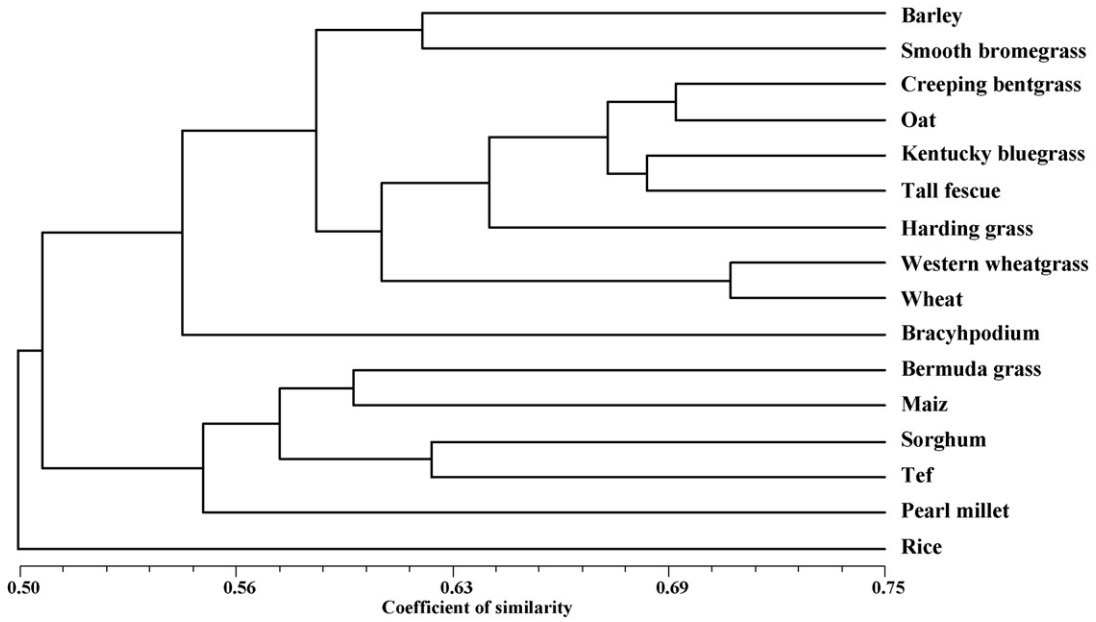

Fig. 1. Dendrogram of the 16 grass species based on Dice coefficient of similarity calculated from 841 polymorphic markers.

relationship between *Poeae* and *Aveneae* (Hsiao et al., 1995). As expected from their phylogenetic distance (Kellogg, 2001; Gaut, 2002), fewer tall fescue primers were functional on species from the *Panicoideae*, *Chloridoideae* and *Ehrhartoideae* subfamilies. The same trend was observed for primers from wheat, showing the highest cross-amplification to members of the *Triticeae* tribe (barley and western wheatgrass). On the other hand, both rice and tef showed the lowest rate of amplification across the panel (Table 4). Rice primers in this study showed the least frequent amplification in members of the *Andropogoneae* tribe, represented by maize and sorghum.

Studies have indicated that the number of SSRs amplified in a given species was positively correlated with the phylogenetic relatedness of that species and the species from which the marker was designed (Saha et al., 2004; Mian et al., 2005; Wang et al., 2005; Jensen et al., 2007). Sequence alignment of PCR products from EST-SSR markers (Saha et al., 2004; Gimenes et al., 2007) have indicated that sequences flanking the repeats are usually conserved and variation appears more in the repeat motif length. However, for less related species the polymorphisms that were detected included more INDEL and substitutions in the SSR flanking regions. We used the data provided by the 841 functional markers (excluding the 50 markers that amplified in all 16 species on the panel) in a cluster analysis of the 16 species in our panel. Using the Nei and Li (1979) coefficient of similarity to construct a dendrogram based on the UPGMA method, members of the *Pooideae* grouped together in one cluster (Fig. 1). Rice, from subfamily *Ehrhartoideae* was clearly separated from all other groups. No clear separation was apparent between members of subfamily *Panicoideae* (sorghum, maize and pearl millet) and those of subfamily *Chloridoideae* (bermudagrass and tef). This clustering is in agreement with previously established taxonomic relationships among grass species (Kellogg, 2001; Gaut, 2002). The correlation between the genetic similarity coefficients and their corresponding cross-amplification values of EST primers ($N = 90$ comparisons) was surprisingly low ($r = 0.26$). This correlation was rather weak considering the expected positive correlation between the phylogenetic relatedness of the species and cross-amplification of primers. However, when CISP primers, were excluded from the correlation matrix ($N = 60$ comparisons), a significant correlation with $r = 0.80$ ($p < 0.01$) was observed. Since CISP primers were designed to target conserved low-copy exons flanking intron regions across sorghum and rice or pearl millet and rice (Feltus et al., 2006), those primers are expected to amplify products with most of the species on the panel, thus weakening the correlation between cross-amplification and relatedness. The high rate of cross-amplification of sorghum/rice (74%) and pearl millet/rice (91%) CISP primers, as shown in Table 4, was due to the type of marker screened, rather than their phylogenetic relationship to other species on the panel. Consequently, the remainder of the data supported a positive correlation between the phylogenetic relatedness and EST-SSR primer cross-amplification for the tested species.

3.3. Polymorphism of amplified fragments as observed on reed canary grass

Reed canary grass is a perennial grass native to temperate zones of the Northern Hemisphere and mainly represented by an allotetraploid cytotype ($2n = 4\times = 28$) and a hexaploid form ($2n = 6\times = 42$) (Lavergne and Molofsky, 2004). Although it was not included in our screening panel, it is a close relative of harding grass (*P. aquatica*), which was screened in this study. Both species are not yet well characterized on the molecular level (Mian et al., 2005; Casler et al., 2009). DNA from 13 accessions of the reed canary grass was amplified using 91 primers (56 CISP and 35 EST-SSR primers) out of the 192 primers that amplified a PCR product in harding grass. The PCR products were visualized on polyacrylamide gels and scored for polymorphisms. Fifteen of the CISP primers showed polymorphism (27%) in at least one of the 13 accessions screened, three failed to amplify and 38 were monomorphic. On the other hand, 33 of the 35 EST-SSR primers (94%) were polymorphic with only two primers (CNLT 110 and RM 109) scored as monomorphic. CISP primers were able to discriminate all 13 clones except Marathon and Rival. However, including the rest of the polymorphic primers all clones were discriminated with Rensselaer and Rival showing the highest genetic similarity estimates (data not shown).

These results clearly indicated that the primers selected in this study cross-amplified even across genera, but CISP primers are less polymorphic as compared to EST-SSR primers based on fragment size differentiation as observed on polyacrylamide gels. It is worth

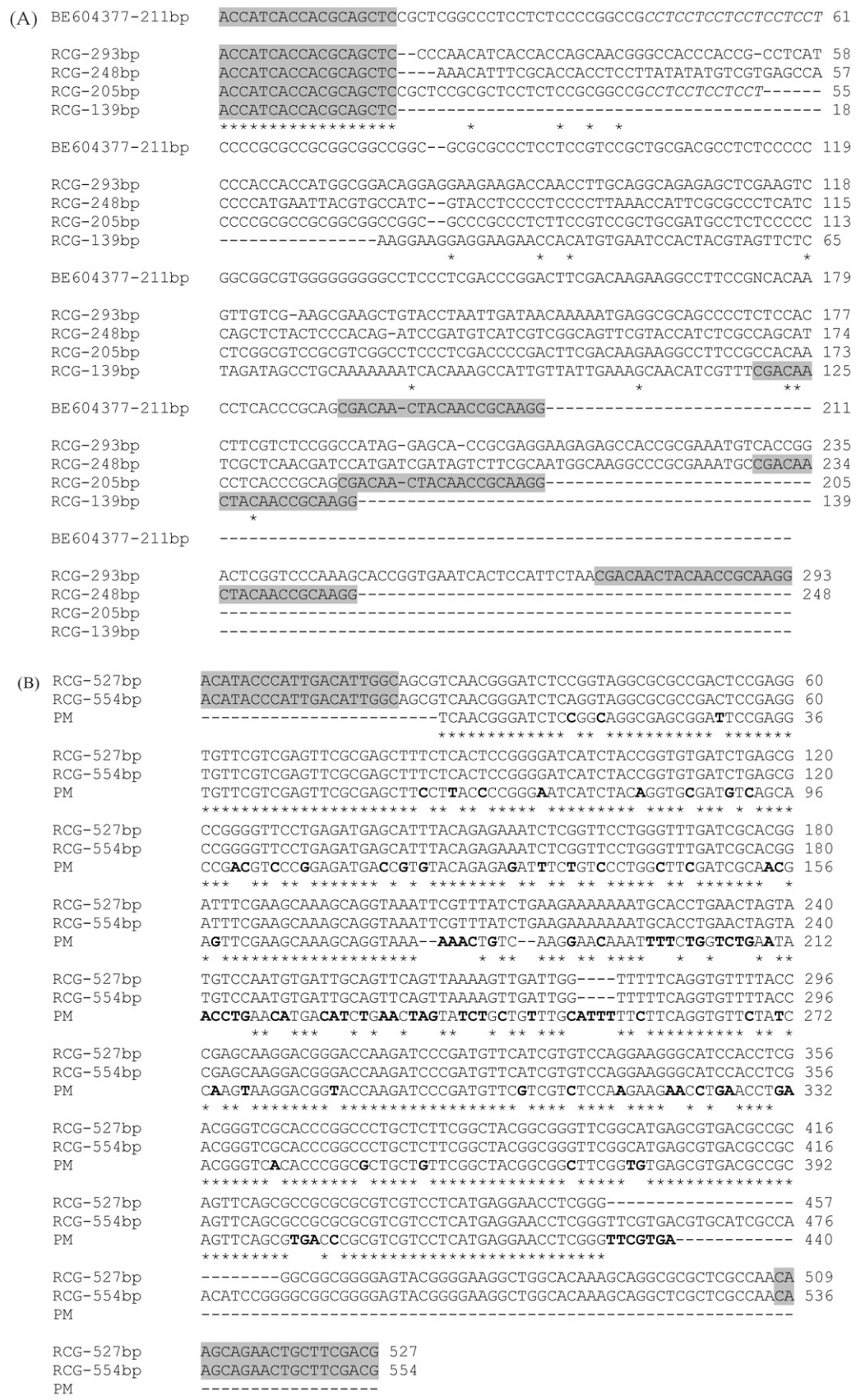

Fig. 2. ClustalW2 alignment of sequences obtained from: (A) four fragments of the EST-SSR primer CNL98 amplified in reed canary grass (RCG) clone Bellevue as compared to the reference sequence from wheat (BE604377), (B) two fragments of the CISP primer SRC4-031 amplified in RCG clone Bellevue as compared to the reference sequence from pearl millet (PM), and (C) two fragments of the CISP primer PRSC1-22 from *Eragrostis pilosa* (*E. pilosa*). Primer sequences are shaded, single nucleotide polymorphisms (SNP) are in bold and SSR repeats are in italics.

```
(C) E.pilosa-283bp      CCATTGCTCCATAATCCAATCCTAAAATGCAAAAAAAAAGAAAGAAAGAAAGAAAGAAGA 60
    E.pilosa-263bp      CCATTGCTCCATAATCCAATCCTAAAATGCAAAAAAAAA-------------------A 40
                        ************************************                     *

    E.pilosa-283bp      GTAACTTAGGACATCACGATATAGCCTAGCTTAGAGAAAGTATTATGCTTGAGACCGTAC 120
    E.pilosa-263bp      GTAACTTAGGACATCATGATATAGCCTAGCTTAGAAAAAGCATTATGCTTGAGACCGTAC 100
                        *************** ***************** **** ******************

    E.pilosa-283bp      CTTATGAAGGAATATGGGTGAAAGTTCATAAACTTGGGATGTAAAACCAGTTCCAGCGTC 180
    E.pilosa-263bp      CTTGTGAAGGAATATGGGTGAAAGTTCATAAACTTGGGATGTAAAACCAGTTCCAGCATC 160
                        *** ********************************************** **

    E.pilosa-283bp      CATAATGAGGGCCCACAGATTTGTTGAAGATGCAACTGAGCTAATGAACAATCCATCCTC 239
    E.pilosa-263bp      CATAATGAGGGCCCACAGATTTGTTGAAGATGCAACTGAGCTAATGAACAATCCATCCTC 219
                        ************************************************************

    E.pilosa-283bp      ATTACCCTTCTCTATGTGTTGATGAAGTCTGGCATCTGCTACA 283
    E.pilosa-263bp      ATTACCCTTCTCTATGTGTTGATGAAGTCTGGCATCTGCTACA 263
                        *******************************************
```

Fig. 2. (*Continued*).

mentioning that CISP primers were originally developed to detect SNP variation through sequencing the amplified fragments (Feltus et al., 2006) and not based on size variation as done in this work. Thus screening for SNP in CISP amplicons across clones is another venue to better utilize those markers. The recently developed PLUG markers (Ishikawa et al., 2009) provide a cheaper method for SNP detection that relies on restricting the amplified fragments with various restriction enzymes and observing the variation on agarose gels. This approach is less demanding for equipment and better suited for the resources available in less-researched crops.

The vast majority of EST-SSR primers tested amplified multiple fragments in reed canary grass. Increasing the annealing temperature to 54 °C and reducing the number of cycles to 30 instead of 35 cycles has reduced the number of non-specific fragments, however, many obvious fragments were still scoreable. Amplification of more than one fragment using EST-SSR markers has been reported in previous studies and was attributed to the possible amplification of both orthologous and paralogous copies of the target region (Varshney et al., 2005; Sim et al., 2009), or homoloci from different genomes, as in the case of hexaploid wheat (Gupta et al., 2003; Saha et al., 2004). Although reed canary grass is a polyploid and amplification of homoloci is rather expected, these assumptions could only be verified by sequencing the observed fragments. For the CISP primers, single fragments and rarely two or more fragments were observed, adding another distinctive character to those markers, besides their remarkably high cross-amplification rate as compared to the EST-SSR primers tested here.

3.4. Nature of the amplified products in reed canary grass and Eragrostis species

Some of the concerns regarding amplification of EST-SSR primers across species as pointed by Thiel et al. (2003), were that markers producing identical fragment sizes may not be identical by descent (homoplasy). In addition, two alleles of different size might be the result of INDEL in the SSR motif or the flanking region rather than variation in the SSR repeat length. These issues would be of special interest for less-researched species as in case of tef, where some of the markers tested here were already mapped (Yu et al., 2006). We sequenced the fragment/s obtained from various markers tested on tef, *E. pilosa* and reed canary grass clones in order to: (i) understand the nature of the multi-fragments observed with EST-SSR primers as opposed to those fragments observed with CISP primers and (ii) determine the genetic basis of the polymorphisms observed (variation in SSR repeat length, INDEL and/or SNP).

All fragments in the range of 120 bp and 400 bp produced by the primers CNL56 and RM176 tested on the line Kaye Murri (*E. tef*), and CNL98 and CNL239 from the clone Bellevue of the reed canary grass were sequenced and vector sequences were trimmed. From these four primers 16 fragments were recovered. Only one fragment (CNL98-204 bp) showed homology to the reference sequence including the SSR repeat motif $(CCT)_4$ as compared to $(CCT)_6$ in the reference sequence as seen in Fig. 2A. Except for the primer sequences, the other three fragments from the primer CNL98 and the 12 fragments from the other primers showed very low or no homology to the reference sequence in addition, the targeted repeat sequence was not recovered in any of those fragments. Since, the sequence of any set of fragments within each primer did not align to each other, it could be concluded that those fragments were amplifications of different loci. The screened CISP primers on the other hand, were more conserved as compared to EST-SSR primers in terms of the number of amplicons amplified per primer. We sequenced the fragments amplified from each of the primers (PRSC1-22 and SRSC3-5) in both *Eragrostis* species and (SRSC4-31, SRSC7-6, SRSC9-2 and SRSC5-19) on Bellevue. Sequencing results have indicated that for each primer, the amplified fragments aligned to each other with very high homology and the differences in sizes were mainly caused by INDEL, although numerous SNP were also observed (Fig. 2B). In one case the cause of the two fragments observed was the absence of the entire SSR motif $(GAAA)_5$ in the shorter fragment (263 bp) as compared to 283 bp in the longer fragment for the primer (PRSC1-22) with *E. pilosa* (Fig. 2C). Six amplicons (25%) from 24 CISP primers sequenced from reed canary grass showed homology to the reference sequence.

It could thus be concluded that the multi-fragment amplification observed with the EST-SSR primers tested here were due to loss of specificity of the primers to the targeted sequence in distantly related species. These results are in general agreement with the results of Sim et al. (2009), where for 83% of the sequenced fragments, from various EST-SSR primers tested on ryegrass, lacked homology to the reference sequence. The results also suggest that although an EST-SSR marker targeting a specific sequence in one species can amplify the orthologous sequence in a distantly related species, the uniqueness of this marker is no longer valid because other non-targeted loci co-amplify. On the contrary the CISP fragments amplified in this study from a single primer are probably the outcome of amplification from two loci on the homologous chromosomes from the polyploid species used in this study.

In conclusion, CISP primers showed higher levels of both cross-amplification and transferability than EST-SSR primers. This may be attributed to the nature of the sequence targeted for primer design and also the sequence flanked by the two primers. EST-SSR primers were designed to either flank SSR repeats observed in a single species (CNLT, RM, REMS, and NFFA/NFFS) or in multiple species, as in case of the CNL markers (Yu et al., 2004; La Rota et al., 2005). These features of the EST-SSR markers allow for detection of polymorphism based on variation in the SSR repeat length and/or INDEL (Varshney et al., 2005). The CISP markers on the other hand, were designed to be positioned on exon sequences conserved across either sorghum and rice or pearl millet and rice, flanking at

least one intron region and rarely harboring an SSR repeat motif (Feltus et al., 2006). While both marker types should be useful for less-researched crops, CISP primers are more promising than EST-SSR primers for comparative mapping across distantly related species.

This study assessed the cross-amplification of 919 EST primers developed from seven different grass species on a total of 16 grasses. An average of 315 primers amplified per grass species (excluding self amplification) and could be considered a readily available resource for genetic mapping and diversity analysis in a wide range of minor, less-researched grass species from different subfamilies of the *Poaceae*. The results also confirmed the expectation of higher level of cross-amplification of EST-SSR primers to more closely related species, extending previous results to several other grass and forage species. While cross-amplification of different types of EST-based primers even across genera was possible, more attention should be given to the nature of the amplified fragments before inferring synteny or orthology. As new sequencing technologies emerge, longer sequence reads at much lower costs than what is available right now are anticipated, making genome sequencing the method of choice for many crop species. However, until this is realized future studies on minor crops and less-researched grass species will benefit directly from the primers described here and from similar studies.

Acknowledgements

The present work was partially supported by grants from the USDA/NRI project numbers 2001-35301-10612, 2005-35301-15728, and 2006-55606-16629, the DOE-USDA Plant Feedstock Genomics for Bioenergy Program (DE-A102-07ER64454), the Hatch Project 149419 and the McKnight Foundation's Collaborative Crop Research Program – African Chloridoid Cereals project (Grant No. 06-448). We would like to thank the anonymous reviewers for their useful criticism and helpful suggestions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fcr.2010.03.014.

References

Asp, T., Frei, U.K., Didion, T., Nielsen, K.K., Lubberstedt, T., 2007. Frequency, type, and distribution of EST-SSRs from three genotypes of *Lolium perenne*, and their conservation across orthologous sequences of *Festuca arundinacea*, *Brachypodium distachyon*, and *Oryza sativa*. BMC Plant Biol. 7, 36.

Casler, M.D., Phillips, M.M., Krohn, A.L., 2009. DNA polymorphisms reveal geographic races of reed canarygrass. Crop Sci. 49, 2139–2148.

Chalhoub, B.A., Thibault, S., Laucou, V., Rameau, C., Hofte, H., Cousin, R., 1997. Silver staining and recovery of AFLP (TM) amplification products on large denaturing polyacrylamide gels. Biotechniques 22, 216–220.

Feltus, F.A., Singh, H.P., Lohithaswa, H.C., Schulze, S.R., Silva, T.D., Paterson, A.H., 2006. A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. Plant Physiol. 140, 1183–1191.

Gaut, B.S., 2002. Evolutionary dynamics of grass genomes. New Phytologist 154, 15–28.

Gimenes, M.A., Hoshino, A.A., Barbosa, A.V.G., Palmieri, D.A., Lopes, C.R., 2007. Characterization and transferability of microsatellite markers of the cultivated peanut (*Arachis hypogaea*). BMC Plant Biol. 7, 1471–2229.

Gupta, P.K., Rustgi, S., Sharma, S., Singh, R., Kumar, N., Balyan, H.S., 2003. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. Mol. Genet. Genom. 270, 315–323.

Gupta, P.K., Varshney, R.K., 2000. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113, 163–185.

Hsiao, C., Chatterton, N.J., Asay, K.H., Jensen, K.B., 1995. Molecular phylogeny of the *Pooideae* (*Poaceae*) based on nuclear rDNA (ITS) sequences. Theor. Appl. Genet. 90, 389–398.

Ishikawa, G., Nakamura, T., Ashida, T., Saito, M., Nasuda, S., Endo, T.R., Wu, J., Matsumoto, T., 2009. Localization of anchor loci representing five hundred annotated rice genes to wheat chromosomes using PLUG markers. Theor. Appl. Genet. 118, 499–514.

Jensen, L.B., Holm, P.B., Lubberstedt, T., 2007. Cross-species amplification of 105 *Lolium perenne* SSR loci in 23 species within the *Poaceae*. Mol. Ecol. Notes 7, 1155–1161.

Kellogg, E.A., 2001. Evolutionary history of the grasses. Am. Soc. Plant Biol., 1198–1205.

Khlestkina, E.K., Than, M.H.M., Pestsova, E.G., Röder, M.S., Malyshev, S.V., Korzun, V., Börner, A., 2004. Mapping of 99 new microsatellite-derived loci in rye (*Secale cereale* L.) including 39 expressed sequence tags. Theor. Appl. Genet. 109, 725–732.

Kindiger, B., 2006. Cross-species amplification of *Lolium* microsatellites in *Poa* ssp. Grassland Sci. 52, 105–115.

La Rota, M., Kantety, R.V., Yu, J.K., Sorrells, M.E., 2005. Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat, and barley. BMC Genom. 6, 23.

Lavergne, S., Molofsky, J., 2004. Reed canary grass (*Phalaris arundinacea*) as a biological model in the study of plant invasions. Critical Rev. Plant Sci. 23, 415–429.

Mian, M.A.R., Saha, M.C., Hopkins, A.A., Wang, Z.Y., 2005. Use of tall fescue EST-SSR markers in phylogenetic analysis of cool-season forage grasses. Genome 48, 637–647.

Nei, M., Li, W.H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. 76, 5269–5273.

Opanowicz, M., Vain, P., Draper, J., Parker, D., Doonan, J.H., 2008. *Brachypodium distachyon*: making hay with a wild grass. Trends Plant Sci. 13, 172–177.

Parida, S.K., Anand Raj Kumar, K., Dalal, V., Singh, N.K., Mohapatra, T., 2006. Unigene derived microsatellite markers for the cereal genomes. Theor. Appl. Genet. 112, 808–817.

Roa, A.C., Chavarriaga-Aguirre, P., Duque, M.C., Maya, M.M., Bonierbale, M.W., Iglesias, C., Tohme, J., 2000. Cross-species amplification of cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationship. Am. J. Bot. 87, 1647–1655.

Rohlf, F.J., 2002. NTSYS-pc, Version 2.10 z. Exeter Software, Setauket, New York.

Saha, M.C., Mian, M.A.R., Eujayl, I., Zwonitzer, J.C., Wang, L., May, G.D., 2004. Tall fescue EST-SSR markers with transferability across several grass species. Theor. Appl. Genet. 109, 783–791.

Sim, S.C., Yu, J.K., Jo, Y., Sorrells, M.E., Jung, G., 2009. Transferability of cereal EST-SSR markers to ryegrass. Genome 52, 431–437.

Squirrell, J., Hollingsworth, P.M., Woodhead, M., Russell, J., Lowe, A.J., Gibby, M., Powell, W., 2003. How much effort is required to isolate nuclear microsatellites from plants? Mol. Ecol. 12, 1339–1348.

Tai, T.H., Tanksley, S.D., 1990. A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. Plant Mol. Biol. Rep. 8, 297–303.

Temnykh, S., Park, W.D., Ayres, N., Cartinhour, S., Hauck, N., Lipovich, L., Cho, Y.G., Ishii, T., McCouch, S.R., 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theor. Appl. Genet. 100, 697–712.

Thiel, T., Michalek, W., Varshney, R., Graner, A., 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 106, 411–422.

Varshney, R.K., Graner, A., Sorrells, M.E., 2005. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23, 48–55.

Wang, M.L., Barkley, N.A., Yu, J.K., Dean, R.E., Newman, M.L., Sorrells, M.E., Pederson, G.A., 2005. Transfer of simple sequence repeat (SSR) markers from major cereal crops to minor grass species for germplasm characterization and evaluation. Plant Genet. Resour. 3, 45–57.

Yu, J., Wang, J., Lin, W., Li, S., Li, H., Zhou, J., Ni, P., Dong, W., Hu, S., Zeng, C., et al., 2005. The genomes of *Oryza sativa*: a history of duplications. PLoS Biol. 3, e38.

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

Yu, J.K., Kantety, R.V., Graznak, E., Benscher, D., Tefera, H., Sorrells, M.E., 2006. A genetic linkage map for tef [*Eragrostis tef* (Zucc.) Trotter]. Theor. Appl. Genet. 113, 1093–1102.

Zhang, L.Y., Bernard, M., Leroy, P., Feuillet, C., Sourdille, P., 2005. High transferability of bread wheat EST-derived SSRs to other cereals. Theor. Appl. Genet. 111, 677–687.

Zhu, W., Buell, C.R., 2007. Improvement of whole-genome annotation of cereals through comparative analyses. Genome Res. 17, 299–310.