Theor Appl Genet (2006) 113:1449–1458
DOI 10.1007/s00122-006-0391-2

ORIGINAL PAPER

Tall fescue genomic SSR markers: development and transferability across multiple grass species

Malay C. Saha · John D. Cooper · M. A. Rouf Mian · Konstantin Chekhovskiy · Gregory D. May

Received: 21 March 2006 / Accepted: 4 August 2006 / Published online: 1 September 2006
© Springer-Verlag 2006

Abstract Simple sequence repeat (SSR) markers are highly informative and widely used for genetic and breeding studies. Currently, a very limited number of SSR markers are available for tall fescue (*Festuca arundinacea* Schreb.) and other forage grass species. A tall fescue genomic library enriched in $(GA/CT)_n$ repeats was used to develop primer pairs (PPs) flanking SSRs and assess PP functionality across different forage, cereal, and turf grass species. A total of 511 PPs were developed and assessed for their utility in six different grass species. The parents and a subset of a tall fescue mapping population were used to select PPs for mapping in tall fescue. Survey results revealed that 48% (in rice) to 66% (in tall fescue) of the PPs produced clean SSR-type amplification products in different grass species. Polymorphism rates were higher in tall fescue (68%) compared to other species (46% ryegrass, 39% wheat, and 34% rice). A set of 194 SSR loci (38%) were identified which amplified across all six species. Loci segregating in the tall fescue mapping population were grouped as loci segregating from the female parent (HD28-56, 37%), the male parent (R43-64, 37%), and both parents (26%). Three percent of the loci that were polymorphic between parents were monomorphic in the pseudo F1 mapping population and the remaining loci segregated. Sequencing of amplified products obtained from PP NFFAG428 revealed a very high level of sequence similarity among the grass species under study. Our results are the first report of genomic SSR marker development from tall fescue and they demonstrate the usefulness of these SSRs for genetic linkage mapping in tall fescue and cross-species amplification.

Communicated by F. J. Muehlbauer.

Electronic supplementary material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00122-006-0391-2 and is accessible for authorized users.

M. C. Saha (✉) · M. A. R. Mian · K. Chekhovskiy
Forage Improvement Division, The Samuel Roberts Noble Foundation, Inc., 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
e-mail: mcsaha@noble.org

J. D. Cooper · G. D. May
Plant Biology Division, The Samuel Roberts Noble Foundation, Inc, Ardmore, OK 73401, USA

M. A. R. Mian
Corn and Soybean Research Unit, USDA-ARS, Ohio State University, OARDC, Wooster, OH 44691, USA

G. D. May
National Center for Genome Resources, 2935 Rodeo Park Drive East, Santa Fe, NM 87505, USA

Introduction

Molecular markers are powerful tools for genetic mapping, genotype fingerprinting, population structure, and genetic diversity studies. Simple sequence repeat (SSR) markers have become the marker class of choice because they are mostly co-dominant, abundant in genomes, highly reproducible, and some have high rates of transferability across species (Gaitán-Solís et al. 2002; Thiel et al. 2003; Saha et al. 2004). Regular use of SSR markers for breeding and other applied research in a plant species depend on development of a large number of SSRs for the species of interest. The

number of SSR markers currently available for tall fescue and related forage grasses are very limited (Saha et al. 2004).

The primary disadvantage of SSR markers is that the development of these markers is time consuming and expensive. However, once developed, the SSR markers can be used for a variety of purposes and often across different species. Besides, recently developed high-throughput techniques facilitate marker development and reduce cost. SSR markers can be developed from either genomic or expressed sequence tag (EST) libraries. The rapidly growing EST databases are becoming an important source for SSRs. In species where EST databases are not well established, genomic libraries are considered as an important source for SSRs. EST-SSRs are derived from transcribed regions of DNA and are more conserved which limit their polymorphism rates when compared to genomic SSRs (Cho et al. 2000; Thiel et al. 2003). EST-SSRs are associated with expressed genes and are usually concentrated in the gene rich regions of the genome. In contrast, the genomic SSRs are highly polymorphic and tend to be widely distributed throughout the genome resulting in better map coverage (Taramino et al. 1997; Warnke et al. 2004; La Rota et al. 2005; Saha et al. 2005).

A number of different repeat motifs are found in the eukaryotic organisms. Di-nucleotide repeats are the most abundant in genomic SSRs (Lee et al. 2004) whereas tri-nucleotide motifs are the most abundant in EST-SSRs (Saha et al. 2004; La Rota et al. 2005). Among the di-nucleotide motifs, GA/CT is the most common and important repeat in different grasses (Kantety et al. 2002; Cai et al. 2003) and other species such as paspalum (*Paspalum vaginatum* Swartz—Liu et al. 1995), grape (*Vitis* spp.—Sefc et al. 1999), and olive (*Olea europaea* L.—Rallo et al. 2000).

Comparative genomics has become an important strategy for utilizing genetic information across different species. Comparative genetics revealed that gene content and order are generally conserved among some closely related species (e.g. Van Deynze et al. 1995b; Gale and Devos 1998). Comparative mapping has been used for extending genetic information from model organisms to more complicated species (Paterson et al. 1995). Co-linearity of common markers among species suggests that knowledge gained through molecular analysis in one species will be useful in related species; however, recent studies have revealed considerable complexity in genome structure between species (Sorrells et al. 2003) as well as within species (Fu and Dooner 2002; Brunner et al. 2005). Sequence analyses of SSR loci indicated high homology in SSR flanking regions of several grass species (Saha et al. 2004). Thus SSR primer pairs (PPs) developed from one species could be used to detect SSRs in related species (Dirlewanger et al. 2002; Kuleung et al. 2004; Yu et al. 2004b).

High rates of SSR marker transferability across species within a genus (Gaitán-Solís et al. 2002) and sometimes across genera (Kuleung et al. 2004; Varshney et al. 2005) have been reported. The transferability of genomic SSR markers across genera and beyond is generally low (Peakall et al. 1998; Roa et al. 2000). However, the relatively high sequence similarity among members of the Poaceae family (Kantety et al. 2002) and transferability of EST-SSRs across species (Saha et al. 2004; Yu et al. 2004b; Varshney et al. 2005) suggest that the transferability of grass genomic SSRs across different grass genera should be further investigated. These markers can be useful for increasing the resolution of comparative maps among the related species.

Tall fescue (*Festuca arundinacea* Schreb.) is a major perennial forage crop in the temperate regions of the world. Genetic improvement is impaired due to features such as large genome size, polyploidy, obligate cross-pollination, and severe inbreeding depression. The availability of SSR markers in tall fescue is very limited. The percent of ESTs containing SSR is fairly low in tall fescue compared to other grass species (Kantety et al. 2002; Saha et al. 2004). Only 157 PPs were obtained from approximately 20,000 tall fescue ESTs. A large number of SSR markers would augment the genetic mapping of tall fescue and potentially other related forage and turf grass species. The current study was undertaken to develop a set of SSR markers from an enriched genomic libraries of tall fescue.

Materials and methods

Development of $(GA/CT)_n$ enriched library

DNA was isolated from a pooled population of tall fescue (*F. arundinacea* Schreb.) cv. Kentucky-31 plants based on a modified protocol of Sharma et al. (2002). Leaf tissues were grounded in liquid nitrogen and extracted in 100 mM Tris–HCl (pH 8.0), 20 mM EDTA, 1.4 M NaCl, 2% CTAB, and 1% β-mercaptoethanol. Methods used in the construction of the library were base on a protocol reported in Hamilton et al. (1999) with minor modifications. About 9.16 μg of genomic DNA was digested with 20 U each of the restriction endonuclease *AluI, Hae III, Nhe I*, and *Rsa*

I. Following digestion, single-stranded overhangs were removed with the addition of 10 U of mung bean exonuclease. A biotinylated $(GA)_n$ oligonucleotide stretch was used to capture and enrich for DNA fragments containing $(GA/CT)_n$ repeats. A detailed protocol entitled "Microsatellite-Enhanced Genomic Library Construction Streptavidin-Biotin Bench Version, SNX Linker" (version 1.1 December, 2002) is available online at: http://www.bioserver.georgetown.edu/faculty/hamilton/. Sequence analysis was performed on an ABI 3730 DNA Sequence Analyzer. Sequences with desired quality were grouped together and screened for SSRs using the PERL program, simple sequence repeat identification tool (SSRIT), downloaded from the Cornell University web site http://www.gramene.org/gramene/searches/ssrtool. Sequences with a minimum of nine repeats and SSR positions not less than 25 bases from the start and end points of sequences were grouped together to design primers.

Primer design

Primers were designed from non-redundant sequences containing SSRs using the Primer3 software (http://www.genome.wi.mit.edu/genome_software/other/primer3.html). The major parameters for designing primers were: primer length from 18 to 24 bases with 22 as the optimum, PCR product size from 125 to 325 bp, optimum annealing temperature 60°C, and GC contents from 40 to 70% with 50% as optimum. PPs were custom synthesized by Qiagen/Operon Technologies (Alameda, CA, USA).

Plant materials

Ten genotypes from six important cereal, forage, and turf grass species (Table 1) were screened for amplification of the genomic SSRs. These genotypes represented diploid to hexaploids, annual to perennial growth habit, and cross-pollinated to self-pollinated species (Table 1). Two genotypes, parents of respective mapping populations were taken each from tall fescue (Saha et al. 2005), ryegrass (Wranke et al. 2004), rice (Temnykh et al. 2000), and wheat (Van Deynze et al. 1995a). The rice and wheat seeds were kindly provided by Dr. Mark Sorrells, Cornell University, Ithaca, NY, USA. The meadow fescue (PIWF92-109-2) and the tetraploid fescue (W-2) were received from Dr. Tim Phillips, University of Kentucky, Lexington, KY, USA. The inheritance of SSR loci were tested on six genotypes of the tall fescue mapping population (Saha et al. 2005).

DNA isolation, PCR amplification, and detection of SSR bands

The DNA was extracted from freshly frozen leaf tissue using Qiagen DNeasy® Plant Mini kit (Qiagen, Valencia, CA, USA) with minor modifications (Saha et al. 2004). The DNA concentrations were quantified using a Hoefer® Dyna Quant® 200 DNA fluorometer (Amersham Biosciences, Piscataway, NJ, USA). The PCR reactions include 20 ng of template DNA, one unit of AmpliTaq Gold® with GeneAmp PCR bufferII (Applied Biosysytems, Branchburg, NJ, USA), 3.25 mM $MgCl_2$, 240 µM of dNTPs, and 0.2 µM of each primer in a 10 µl reaction. To reduce non-specific amplifications, a second round of PCR was performed with selected PPs using 13 ng of template DNA, 0.65 U of AmpliTaq Gold 10X PCR bufferII, 2.5 mM $MgCl_2$, 200 µM of dNTPs, and 0.13 µM of each primer in a 10 µl reaction. The PCR protocols included: initial denaturation of 10 min at 95°C followed by 40 cycles with 50 s at 95°C, 50 s at 58–64°C (the optimum annealing temperature of respective PPs), 90 s at 72°C, and a final extension step of 10 min at 72°C. The PCR products were resolved on 6% polyacrylamide denaturing gels and visualized by silver staining (Protocol suggested by Promega, Madison, WI, USA). The band size is reported using a 10 bp DNA ladder (Invitrogen Life Technologies, Carlsbad, CA, USA) as the reference point. The polymorphism was determined according to the presence or absence of the SSR band.

Table 1 Grass species and genotypes used for screening of tall fescue genomic SSR primers

Common name	Scientific name	Genotype/cultivars	Genome	Mode of pollination
Tall fescue	*Festuca arundinacea*	HD28-56 R43-64	$2n = 6x = 42$	Cross
Meadow fescue	*F. pratensis*	PIWF92-109-2	$2n = 2x = 14$	Cross
Tetraploid fescue	*F. arundinacea* var. *glaucescens*	W-2	$2n = 4x = 28$	Cross
Ryegrass	*Lolium perenne/multiflorum*	MFA-4 MFB-2	$2n = 2x = 14$	Cross
Rice	*Oryza sativa*	IR-64 Azucena	$2n = 2x = 24$	Self
Wheat	*Triticum aestivum*	Opata W7984	$2n = 6x = 42$	Self

Sequencing of PCR bands

PCR products from the polyacrylamide gels were excised using pipette tips and dipped in 10 μl of nuclease free water for 30 min at room temperature. Seven micro liter of it was used as template for a second PCR following the same protocol in a 25 μl reaction volume. The PCR products were resolved in a 2% SeaKem® LE agarose (FMC BioProducts, Rodkland, ME, USA) gel. Gels were stained with ethidium bromide and bands were excised under UV light and put into the Nebulizer cap of Montage DNA Gel Extraction Device (Millipore Corporation, Bedford, MA, USA). DNA was extracted following their standard protocol and concentration was measured against a Molecular Mass Standard (Bio-Rad Laboratories, Hercules, CA, USA). Sequencing was performed using an ABI 3730 in 20 μl of sequencing reaction consisted of 100 ng of template, 4.0 pmol of primer, and 8 μl of sequencing reaction. The sequence data were analyzed with DNASTAR software. Consensus sequences from each species were aligned using Clustal W as the alignment tool.

Results

Isolation and characterization of microsatellites

A total of 5,712 sequences were characterized from the tall fescue small insert genomic library highly enriched for $(GA/CT)_n$ repeats. High quality sequences (5,380) were screened for SSR motifs and approximately 20% were positive. Nine hundred and six of these were candidate SSR sequences that had repeat motifs >18 bases and at least 25 bases of 5′ and 3′ flanking DNA sequences. Five hundred and eleven PPs were developed from these SSR sequences according to the criteria described in the Materials and methods section. Four hundred and twenty five of these PPs were singleton SSRs. Thus PPs were developed from 56% of the SSR sequences. Remaining sequences did not satisfy the criteria for primer design. The sequences of the 511 PPs, along with the marker name, sequence ID, primer sequences, Tm, and expected sizes are included in the ESM.

Primer evaluation

A majority of the PPs amplified characteristic SSR-type bands in different species (Fig. 1). In the first round of evaluation, approximately 34% of the PPs generated non-specific amplification products. These PPs were re-evaluated in a second round of PCR experiments under more stringent conditions to obtain the desired SSR amplification products. An estimated 27% of the PPs still amplified more bands than expected, and were subsequently grouped separately (Table 2). As expected, the success rate was highest in tall fescue and only 7% of the PPs did not generate amplification products in any of the species tested. The portion of PPs that produced clean SSR products ranged from 66% (tall fescue) to 48% (rice). When non-specific amplifications are included, however, amplification success rates were as high as 93 and 74% for the same two species, respectively (Table 2). High success rate of SSR loci amplification among different crop species has been reported earlier (Cipriani et al. 1999). Alleles per PP varied between two to four with the lowest in diploid meadow fescue and the highest in hexaploid tall fescue.

Cross-species amplification

To evaluate the cross-species amplification, all 511 of tall fescue genomic SSR PPs were screened against ten genotypes representing six different grass species, representative of three important tribes and two subfamilies of the Poaceae family. Most PPs amplified characteristic bands in a set of genetically diverse grass species. Cross species amplification potential of tall fescue genomic SSR PPs was fairly high. Thirty-eight percent of the PPs amplified across all six species, while six and 27% of PPs amplified across five and four of the six species, respectively. Only 4% of the PPs were species-specific. A vast majority of the PPs amplified across all forage and turf grass species rather than across the cereal species (data not shown). These findings suggest a high level of sequence conservation among the forage species examined.

Polymorphism of SSR loci

Simple sequence repeat polymorphism rates were tested using the parents of mapping populations and rates varied from 68% (tall fescue) to 34% (rice). Higher polymorphism rates were observed in forage and turf grass species than in cereal grass species. Variation in polymorphism rates may be related to the reproductive strategies of these species. High polymorphism rates for outcrossing compared to self-pollinating species observed in this study are supported by earlier findings (Saha et al. 2004). Inheritance of polymorphic SSR loci was tested with parents and a subset of progeny of the tall fescue mapping population (Fig. 2). A total of 1,347 alleles amplified by 338 PPs

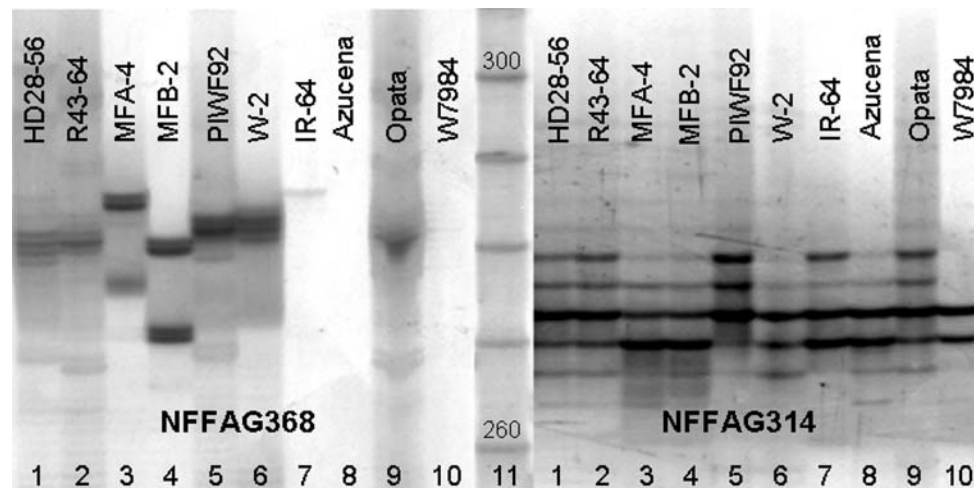

Fig. 1 Polyacrylamide gel image of SSR bands amplified by primer pairs (PPs) NFFAG314 and NFFAG368 in ten genotypes from six different grass species. *Lanes 1*, *2* tall fescue; *3*, *4* ryegrass; *5* meadow fescue; *6* tetraploid fescue; *7*, *8* rice; and *9*, *10* wheat genotypes. *Lane 11* shows the size standard (10 bp)

Table 2 Number and percent of primer pairs (PPs) that amplified PCR products in different grass species

Species	Amplification				Alleles/primer
	PPs with expected PCR bands		PPs with expected plus non-specific bands		
	Number	Percent[a]	Number	Percent	
Tall fescue	338	66.2	476	93.2	4.01
Meadow	282	55.2	419	82.0	1.94
Tetraploid	299	58.5	434	84.9	2.04
Ryegrass	290	56.8	426	83.4	3.37
Rice	244	47.8	379	74.2	2.90
Wheat	296	57.9	434	84.9	3.09

[a] Percent of total 511 PPs tested in this study

were segregating in the tall fescue mapping population. Four hundred and ninety two (37%) of the alleles were segregating from the HD28-56 parent of which 445 were polymorphic while the remaining 47 were monomorphic. Similar marker segregation was also observed in the R43-64 parent. Of the 357 (26%) alleles segregating from both parents, 112 were polymorphic and 245 were monomorphic. Polymorphic markers segregating from both parents are important for identifying and joining of homologous linkage groups (Saha et al. 2005).

Sequence comparison of SSR locus

PCR products obtained from an SSR locus amplified across all six species were sequenced to determine homoplasy among species. A DNA sequence alignment for the monomorphic SSR locus NFFAG428 reveals a high level of sequence similarities in eight genotypes of all six species (Fig. 3). Sequences obtained from tall fescue, ryegrass, rice, and wheat were identical (Fig. 3). Sequences of meadow fescue and tetraploid fescue were identical to each other, but differed from other species by one insertion and five substitutions.

Discussion

Development of genomic SSRs

A pre-requisite for marker-assisted breeding is a robust set of informative markers for the species of interest. Tall fescue is an important cool-season perennial forage grass species with limited microsatellite markers (Saha et al. 2004, 2005). Construction and screening of partial genomic libraries and sequencing of SSR-positive clones are considered effective methods for microsatellite development (Rafalski et al. 1996). However, these approaches are expensive and labor intensive process. Highly-enriched genomic libraries significantly reduced the cost and effort for

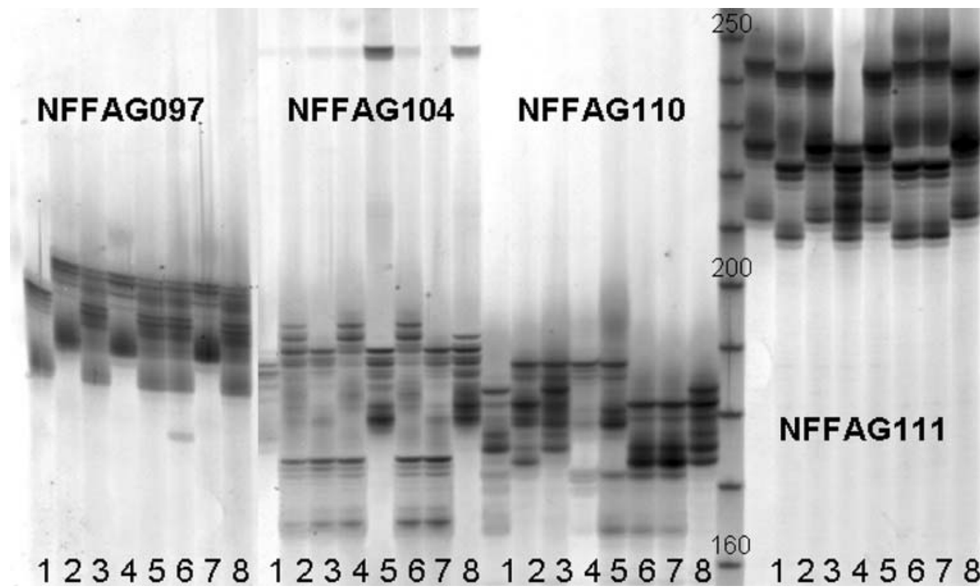

Fig. 2 Amplification of four genomic SSR PPs (NFFAG097, NFFAG104, NFFAG110, and NFFAG111) in a subset of tall fescue mapping population. *Lanes 1* and *2* represent the two parents. *Lanes 3* thru *8* represent the six progeny. The lane before the right most panel (NFFAG111) shows the 10 bp size standard

```
    TTGTAAAACGACGGCCAGTGAGATTGGAGAGGAAGAGCAAAGAGGAAATCAACAAATGACTCCAAGATCTAGATCCCAAGAGTTC Majority
    ---------+---------+---------+---------+---------+---------+---------+---------+-----
             10        20        30        40        50        60        70        80
    ---------+---------+---------+---------+---------+---------+---------+---------+-----
  1 TTGTAAAACGACGGCCAGTGAGATTGGAGAGGAAGAGCAAAGAGGAAATCAACAAATGACTCCAAGATCTAGATCCCAAGAGTTC HD28-56
  1 TTGTAAAACGACGGCCAGTGAGATTGGAGAGGAAGAGCAAAGAGGAAATCAACAAATGACTCCAAGATCTAGATCCCAAGAGTTC R43-64
  1 TTGTAAAACGACGGCCAGTGAGATTGGAGAGGAAGAGCAAAGAGGAAATCAACAAATGACTCCAAGATCTAGATCCCAAGGGTTC PIWF92
  1 TTGTAAAACGACGGCCAGTGAGATTGGAGAGGAAGAGCAAAGAGGAAATCAACAAATGACTCCAAGATCTAGATCCCAAGGGTTC W-2
  1 TTGTAAAACGACGGCCAGTGAGATTGGAGAGGAAGAGCAAAGAGGAAATCAACAAATGACTCCAAGATCTAGATCCCAAGAGTTC MFB-2
  1 TTGTAAAACGACGGCCAGTGAGATTGGAGAGGAAGAGCAAAGAGGAAATCAACAAATGACTCCAAGATCTAGATCCCAAGAGTTC IR-64
  1 TTGTAAAACGACGGCCAGTGAGATTGGAGAGGAAGAGCAAAGAGGAAATCAACAAATGACTCCAAGATCTAGATCCCAAGAGTTC Opata
  1 TTGTAAAACGACGGCCAGTGAGATTGGAGAGGAAGAGCAAAGAGGAAATCAACAAATGACTCCAAGATCTAGATCCCAAGAGTTC W7984

    CCCTCACAAAGAGGAGGAATGGATTGGTGGAG-TTGTAGATCTAGATCTCCTCTCTTAGATCCCTCAAGAATGAGCAAGAATCAT Majority
    ----+---------+---------+---------+---------+---------+---------+---------+---------+
        90       100       110       120       130       140       150       160       170
    ----+---------+---------+---------+---------+---------+---------+---------+---------+
 86 CCCTCACAAAGAGGAGGAATGGATTGGTGGAG-TTGTAGATCTAGATCTCCTCTCTTAGATCCCTCAAGAATGAGCAAGAATCAT HD28-56
 86 CCCTCACAAAGAGGAGGAATGGATTGGTGGAG-TTGTAGATCTAGATCTCCTCTCTTAGATCCCTCAAGAATGAGCAAGAATCAT R43-64
 86 CCCTCACTTAGAGGAGAAATGGATTGGTGGGGATTGTAGATCTAGATCTCCTCTCTTAGATCCCTCAAGAATGAGCAAGAATCAT PIWF92
 86 CCCTCACTTAGAGGAGAAATGGATTGGTGGGGATTGTAGATCTAGATCTCCTCTCTTAGATCCCTCAAGAATGAGCAAGAATCAT W-2
 86 CCCTCACAAAGAGGAGGAATGGATTGGTGGAG-TTGTAGATCTAGATCTCCTCTCTTAGATCCCTCAAGAATGAGCAAGAATCAT MFB-2
 86 CCCTCACAAAGAGGAGGAATGGATTGGTGGAG-TTGTAGATCTAGATCTCCTCTCTTAGATCCCTCAAGAATGAGCAAGAATCAT IR-64
 86 CCCTCACAAAGAGGAGGAATGGATTGGTGGAG-TTGTAGATCTAGATCTCCTCTCTTAGATCCCTCAAGAATGAGCAAGAATCAT Opata
 86 CCCTCACAAAGAGGAGGAATGGATTGGTGGAG-TTGTAGATCTAGATCTCCTCTCTTAGATCCCTCAAGAATGAGCAAGAATCAT W7984

    GGGGGGAATCAAGAGATAGGGCAAGTTCTTCAAAGATGACAATGGAGGAGAGAGAGAGGAAGAACTTATCTTGCCCAAGGTGA   Majority
    ---------+---------+---------+---------+---------+---------+---------+---------+---
            180       190       200       210       220       230       240       250
    ---------+---------+---------+---------+---------+---------+---------+---------+---
170 GGGGGGAATCAAGAGATAGGGCAAGTTCTTCAAAGATGACAATGGAGGAGAGAGAGAGGAAGAACTTATCTTGCCCAAGGTGA   HD28-56
170 GGGGGGAATCAAGAGATAGGGCAAGTTCTTCAAAGATGACAATGGAGGAGAGAGAGAGGAAGAACTTATCTTGCCCAAGGTGA   R43-64
171 GGGGGGAATCAAGAGATAGGGCAAGTTCTTCAAAGATGACAATGGAGGAGAGAGAGAGGAAGAACTTATCTTGCCCAAGGTGA   PIWF92
171 GGGGGGAATCAAGAGATAGGGCAAGTTCTTCAAAGATGACAATGGAGGAGAGAGAGAGGAAGAACTTATCTTGCCCAAGGTGA   W-2
170 GGGGGGAATCAAGAGATAGGGCAAGTTCTTCAAAGATGACAATGGAGGAGAGAGAGAGGAAGAACTTATCTTGCCCAAGGTGA   MFB-2
170 GGGGGGAATCAAGAGATAGGGCAAGTTCTTCAAAGATGACAATGGAGGAGAGAGAGAGGAAGAACTTATCTTGCCCAAGGTGA   IR-64
170 GGGGGGAATCAAGAGATAGGGCAAGTTCTTCAAAGATGACAATGGAGGAGAGAGAGAGGAAGAACTTATCTTGCCCAAGGTGA   Opata
170 GGGGGGAATCAAGAGATAGGGCAAGTTCTTCAAAGATGACAATGGAGGAGAGAGAGAGGAAGAACTTATCTTGCCCAAGGTGA   W7984
```

Fig. 3 Alignment of sequences obtained from PP NFFAG428 in two tall fescue (HD28-56 and R43-64), a meadow fescue (PIWF92), a tetraploid fescue (W-2), a ryegrass (MFB-2), a rice (IR-64), and two wheat (Opata and W7984) genotypes. *Top row* represent the consensus sequence. Nucleotide substitutions and insertion are marked in *bold letters*

microsatellite development (Kijas et al. 1994; Edwards et al. 1996). The enrichment procedure we employed was effective and enabled the development of SSR loci from GA/CT enriched libraries. Seventy percent of the sequenced clones contained SSR motifs. Our results represent a substantial improvement over the non-enriched genomic libraries where only 0.1–0.3% of the clones contained SSRs (Liu et al. 1995; Kubik et al. 1999). In this study we developed 511 microsatellite loci from 5,712 $(GA/CT)_n$ enriched tall fescue genomic

sequences, approximately 9% of the total. Thus, the enrichment procedure was an efficient way to generate a large number of SSR loci. Similar percentages of PPs development was reported in $(CA)_n$ enriched genomic libraries of white clover (Kölliker et al. 2001). Fifty six percent of 906 candidate SSR containing sequences were converted into SSR primers. The rest (44%) of those sequences were not developed as primers due lack of adequate flanking sequences. Similar results were also reported for ryegrass (Jones et al. 2001) and sugarcane (Cordeiro et al. 1999).

The rapidly growing EST databases are considered important resources for SSR development. However, only 1.5–4.6% of the ESTs contained SSRs in various cereal species (Kantety et al. 2002; La Rota et al. 2005). Analysis of 20,000 tall fescue ESTs, resulted in the development of 157 PPs, approximately 0.8% of the total sequences (Saha et al. 2004). Furthermore, EST-SSR loci have a tendency to be clustered in euchromatic regions of the genome which limit the potential for genome-wide coverage (Warnke et al. 2004; Yu et al. 2004a, La Rota et al. 2005; Saha et al. 2005). EST resources are mainly restricted to certain well-studied species. In most forage and turf grass species, where EST sequences are limited or do not exist at all, this enrichment procedure is an attractive way to generate SSRs. Libraries developed through enrichment can provide a large number of SSR-containing sequences and thus substantially reduce time and cost per marker.

Microsatellite motifs

Microsatellites are an important class of markers widely used for molecular dissection of many species. Microsatellite motifs, number, length, composition, and distribution in a genome vary greatly among taxa. GA/CT is one of the most abundant and variable classes of repeat motifs found in several plant species (La Rota et al. 2005). The highest occurrence of GA/CT motif was reported in five cereal species (Kantety et al. 2002), in the forage grass timothy (*Phleum pratense*) (Cai et al. 2003), and also in other species like pepper (*Capsicum annuum*) (Lee et al. 2004). Though CA/GT repeat is an abundant class of microsatellites in mammalian genomes, it is generally less frequent in plant genomes (Depeiges et al. 1995; Powell et al. 1996). Besides, GA/CT repeats are also functionally associated with gene expression. The waxy gene in rice contains $(CT)_n$ repeats and their length polymorphism is associated with amylose content (Ayers et al. 1997). These results further justify the development of $(GA/CT)_n$ enriched genomic libraries and the characterization of microsatellites in the major forage grass species, tall fescue.

Amplification of SSR loci

A total of 511 genomic SSRs were tested for functionality in ten genotypes representing six different grass species. Genomic SSRs developed in this study led to a high rate of success (Table 2). Though the overall amplification rate was very high (74–93%), the percentage of PPs that produced clean products (expected amplification) was 26–27% less than the overall amplification. The high rate of non-specific amplifications observed in this study might be associated with high number of lengthy PCR cycles, elevated concentration of PCR reagents, and unraveling the PCR products in 6% polyacrylamide gels. Pfeiffer et al. (1997) reported that in conifer genome many genomic SSR PPs yielded complex banding patterns that cannot be genetically interpreted. Separating PPs producing non-specific amplification products from those that generate expected products will further simplify the use of these SSR markers in mapping and genetic analyses. As anticipated, the most success was achieved in tall fescue, while the rate of success was related to the taxonomic relationship of the species to tall fescue. The number of bands obtained from each species corresponded to their genome constitution (Table 2). However, more amplification products than expected were noted in tall fescue, ryegrass, and rice which may be attributed to gene duplication and/or duplicated genomic regions. A similar trend was observed in an analysis of EST-SSRs (Saha et al. 2004). Gene duplication and amplification of multiple bands in different grass species were previously reported in wheat (Yu et al. 2004a) and tall fescue (Saha et al. 2005). Experiments are currently in progress to map these SSR markers on a tall fescue genetic linkage map.

Cross-species amplification of SSR loci

Several studies have shown that SSRs developed for a species could be used in related plant species (Dayanandan et al. 1997). The success of cross-species amplification of SSRs depends on the evolutionary relatedness of the species sampled (Dayanandan et al. 1997). Amplification of similar genomic fragments and resolution of the microsatellites in different grass species indicated that the PPs generated in this study could potentially be used for genetic studies in related grass species. Rice and wheat are rich in molecular information. On the contrary, molecular dissection of most forage and turf grass species is limited by the amount of available data. Comparative mapping can make use of the vast amount of genomic information developed in rice and wheat by applying this knowledge to the

less-studied forage species. The high transferability of tall fescue EST-SSRs across grass species was reported by Saha et al. (2004). The genomic SSRs developed from tall fescue demonstrated a high degree of cross-species amplification among a variety of forage, turf, and cereal grass species. This suggests that the SSR flanking regions of the sequences identified in this tall fescue library are well conserved among the grass species surveyed. Hernández et al. (2001) reported high level of maize genomic SSRs transferability (74.5%) to sugarcane. High transferability of SSR markers was also reported in peach (*Prunus* spp) species by Dirlewanger et al. (2002).

Polymorphism rate

SSR primer polymorphism was first examined in the parents of tall fescue, ryegrass, rice, and wheat mapping populations. High polymorphism rates were observed in tall fescue (68%), a lower rate of polymorphism (≤46%) was reported in other species. These results indicate that SSR loci are highly variable within tall fescue. PPs that amplified in other species must come from a more conserved region and thus less polymorphic. Our results indicate that genomic SSRs exhibit 2% higher level of polymorphism than the rates observed for EST-SSRs in tall fescue (Saha et al. 2004). However, polymorphism rate of genomic SSRs in other grass species (46% in ryegrass, 39% in wheat, and 34% in rice) was much lower than the rates noticed for EST-SSRs (66% in ryegrass, 43% in wheat, and 38% in rice) examined on the same parental panel. A significant portion of SSR markers that segregate from both parents of a tall fescue mapping population were identified. These markers will be of value to improve the resolution of the existing genetic map (Saha et al. 2005). In addition, these markers will also be useful for cross referencing maps between grass species. Markers present in both parents but segregating in the population will be useful to identify the homologous linkage groups.

Sequence comparison of SSR locus

Homoplasy is a limitation of SSR marker use in evolutionary and genetic diversity studies (Thiel et al. 2003). Though the SSR primer sequences are highly conserved, polymorphisms in the forms of insertion, deletion, and substitutions are fairly common within SSRs and their flanking regions (Saha et al. 2004). However, genomic SSR amplification products that were sequenced from a monomorphic SSR locus were highly conserved across different grass species. If these results are substantiated with additional sequencing from other PPs, it would suggest that the genomes of grass species examined are conserved not only in the genic regions but also in the non-genic regions of the genome. Similar homoplasy between tall fescue and tetraploid fescue were also reported in earlier studies (Saha et al. 2004). Insertion/deletions (indels) and substitutions were more frequent among more distantly related species when polymorphic SSR loci were sequenced from a number of grass species (Saha et al. 2004; Mian et al. 2005).

We have for the first time isolated and characterized microsatellite loci developed from enriched genomic libraries of the forage grass species tall fescue. A high percentage of the PPs amplified characteristic SSR-type products and were also determined to be highly polymorphic not only in tall fescue, but also in other important forage, turf, and cereal grass species. Many of these markers also have good potential for use in comparative genomic studies among a number of grass species.

References

Ayers NM, McClung AM, Larkin PD, Bligh HFJ, Jones CA, Park WD (1997) Microsatellite and single nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germplasm. Theor Appl Genet 94:773–781

Brunner S, Fengler K, Morgante M, Tingey S, Rafalski A (2005) Evolution of DNA sequence nonhomologies among maize inbreds. Plant Cell 17:343–360

Cai H-W, Yuyama N, Tamaki H, Yoshizawa A (2003) Isolation and characterization of simple sequence repeat markers in the hexaploid forage grass timothy (*Phleum pretense* L). Theor Appl Genet 107:1337–1349

Cho YG, Ishii T, Temnykh S, Chen X, Lopovich L, McCouch SR, Park WD, Ayres N, Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:713–722

Cipriani G, Lot G, Huang W-G, Marrazzo MT, Peterlunger E, Testolin R (1999) AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterization and cross-species amplification in *Prunus.* Theor Appl Genet 99:65–72

Cordeiro GM, Maguire TL, Edwards KJ, Henry RJ (1999) Optimisation of a microsatellite enrichment technique in *Saccharum* spp. Plant Mol Biol Rep 17:225–229

Dayanandan S, Bawa KS, Kesseli RV (1997) Conservation of microsatellites among tropical trees (*Leguminosae*). Am J Bot 84:1658–1663

Depeiges A, Goubely C, Lenoir A, Cocherel S, Picard G, Raynal M, Grellet F, Delseny M (1995) Identification of the most represented repeated motifs in *Arabidopsis thaliana* microsatellite loci. Theor Appl Genet 91:160–168

Dirlewanger E, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanetto A, Arús P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their

use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L). Theor Appl Genet 105:127–138

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. Biotechniques 20:758–759

Fu H, Dooner HK (2002) Intraspecific violation of genetic colinearity and its implications in maize. Proc Natl Acad Sci USA 99:9573–9578

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

Gaitán-Solís E, Duque MC, Edwards KJ, Tohme J (2002) Microsatellite repeats in common bean (*Phaseolus vulgaris*): isolation, characterization, and cross-species amplification in *Phaseolus* ssp. Crop Sci 42(6):2128–2136

Hamilton MB, Pincus EL, DiR=Fiore A, Fleischer RC (1999) A universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. Biotechniques 27(3):500–507

Hernández P, Oorado G, Laurie D, Martin A, Snape J (2001) Microsatellites and RFLP probes from maize are efficient sources of molecular markers for the biomass energy crop *Miscanthus*. Theor Appl Genet 102:616–622

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

Kantety RV, Rota ML, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. Plant Mol Biol 48:501–510

Kijas JMH, Flower JCS, Garbett CA, Thomas MR (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. Biotechniques 16:656–662

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

Kubik C, Meyer WA, Gaut BS (1999) Assessing the abundance and polymorphism of simple sequence repeats in perennial ryegrass. Crop Sci 39:1136–1141

Kuleung C, Baenziger PS, Dweikat I (2004) Transferability of SSR markers among wheat, rye, and triticale. Theor Appl Genet 108:1147–1150

La Rota M, Kantety RV, Yu JK, Sorrells ME (2005) Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice wheat and barley. BMC Genomics 6:23–35

Lee JM, Nahm SH, Kim YM, Kim BD (2004) Characterization and molecular genetic mapping of microsatellite loci in pepper. Theor Appl Genet 108:619–627

Liu Z-W, Jarret RL, Kresovich S, Duncan RR (1995) Characterization and analysis of simple sequence repeat (SSR) loci in seashore paspalum (*Paspalum vaginatum* Swartz). Theor Appl Genet 91:47–52

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

Paterson AH, Lin YR, Li S, Schertz KF, Doebley JF, Pinson SRM, Liu SC, Stansel JW, Irvine JE (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science 269:1714–1717

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross species amplification of soybean (*Glycine max*) simple sequence repeat (SSRs) within the genus and other legume genera: Implication for transferability of SSRs in plants. Mol Biol Evol 15(10):1275–1287

Pfeiffer A, Olivieri AM, Morgante M (1997) Identification and characterization of microsatellites in Norway spruce (*Picea abies* K). Genome 40:411–419

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2:225–238

Rafalski JA, Vogel JM, Morgante M, Powell W, Andre C, Tingey SV (1996) Generating and using DNA markers in plants. In: Birren B, Lai E (eds) Non-mammalian genomic analysis: a practical guide. Academic, San Diego, pp 75–135

Rallo P, Dorado G, Martin A (2000) Development of simple sequence repeats (SSRs) in olive tree (*Olea europaea* L). Theor Appl Genet 101:984–989

Roa AC, Chavarriaga-Aguirre P, Duque MC, Maya MM, Bonierbale MW, 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

Saha MC, Mian MAR, Eujayl I, Zwonitzer JC, Wang L, May GD (2004) Tall fescue EST-SSR markers with transferability across several grass species. Theor Appl Genet 109:783–791

Saha MC, Mian MAR, Zwonitzer JC, Chekhovskiy K, Hopkins AA (2005) An SSR- and AFLP-based genetic linkage map of tall fescue (*Festuca arundinacea* Schreb.). Theor Appl Genet 110:323–336

Sefc KM, Regner F, Turetschek E, Glössl J, Steinkellner H (1999) Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. Genome 42:367–373

Sharma AD, Gill PK, Singh P (2002) DNA isolation from dry and fresh samples for polysaccharide-rich plants. Plant Mol Biol Rep 20:415a–415f

Sorrells ME, La Rota M, Bermudez-Kandianis CE, Greene RA, Kantety RV, Munkvold JD, Miftahudin , Mahmoud A, Ma X, Gustafson PJ, Qi LL, Echalier B, Gill BS, Matthews DE, Lazo GR, Chao S, Anderson OD, Edwards H, Linkiewicz AM, Dubcovsky J, Akhunov ED, Dvorak J, Zhang D, Nguyen HT, Peng J, Lapitan NLV, Gonzalez-Hernandez JL, Anderson JA, Hossain K, Kalavacharla V, Kianian SF, Choi DW, Close TJ, Dilbirligi M, Gill KS, Steber C, Walker-Simmons MK, McGuire PE, Qualset CO (2003) Comparative DNA sequence analysis of wheat and rice genomes. Genome Res 13:1818–27

Taramino G, Tarchini R, Ferrario S, Lee M, Pe' ME (1997) Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor.* Theor Appl Genet 95:66–72

Temnykh S, Park WD, Ayers N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000) Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:697–712

Thiel T, Michalek W, Varshney RK, 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

Van Deynze AE, Ducovsky J, Gill KS, Nelson JC, Sorrells ME, Dvorak J, Gill BS, Lagudah ES, McCouch SR, Apples R (1995a) Molecular-genetic maps for chromosome 1 in *Triticeae* species and their relation to chromosomes in rice and oats. Genome 38:45–59

Van Deynze AE, Nelson JC, O'Donoughue LS, Ahn SN, Siripoonwiwat W, Harrington SE, Yglesias ES, Braga DP, McCouch SR, Sorrells ME (1995b) Comparative mapping in grasses. Oat relationships. Mol Gen Genet 249:349–356

Varshney RK, Sigmund R, Boerner A, Korzun V, Stein N, Sorrells ME, Langridge P, Graner A (2005) Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. Plant Sci 168:195–202

Warnke SE, Barker RE, Jung G, Sim S-C, Mian MAR, Saha MC, Brilman LA, Dupal MP, Forster JW (2004) Genetic linkage mapping of an annual x perennial ryegrass population. Thero Appl Genet 109:294–304

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 (SSR) markers for hexaploid wheat. Genome 47:805–818

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