Theor Appl Genet (2004) 109: 783–791
DOI 10.1007/s00122-004-1681-1

ORIGINAL PAPER

Malay C. Saha · M. A. Rouf Mian · Imad Eujayl · John C. Zwonitzer · Liangjiang Wang · Gregory D. May

Tall fescue EST-SSR markers with transferability across several grass species

Received: 7 January 2004 / Accepted: 28 March 2004 / Published online: 16 June 2004
© Springer-Verlag 2004

Abstract Tall fescue (*Festuca arundinacea* Schreb.) is a major cool season forage and turf grass in the temperate regions of the world. It is also a close relative of other important forage and turf grasses, including meadow fescue and the cultivated ryegrass species. Until now, no SSR markers have been developed from the tall fescue genome. We designed 157 EST-SSR primer pairs from tall fescue ESTs and tested them on 11 genotypes representing seven grass species. Nearly 92% of the primer pairs produced characteristic simple sequence repeat (SSR) bands in at least one species. A large proportion of the primer pairs produced clear reproducible bands in other grass species, with most success in the close taxonomic relatives of tall fescue. A high level of marker polymorphism was observed in the outcrossing species tall fescue and ryegrass (66%). The marker polymorphism in the self-pollinated species rice and wheat was low (43% and 38%, respectively). These SSR markers were useful in the evaluation of genetic relationships among the *Festuca* and *Lolium* species. Sequencing of selected PCR bands revealed that the nucleotide sequences of the forage grass genotypes were highly conserved. The two cereal species, particularly rice, had significantly different nucleotide sequences compared to the forage grasses. Our results indicate that the tall fescue EST-SSR markers are valuable genetic markers for the *Festuca* and *Lolium* genera. These are also potentially useful markers for comparative genomics among several grass species.

Communicated by P. Langridge

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00122-004-1681-1.

M. C. Saha · M. A. R. Mian (✉) · I. Eujayl · J. C. Zwonitzer · L. Wang · G. D. May
The Samuel Roberts Noble Foundation,
2510 Sam Noble Parkway,
Ardmore, OK 73401, USA
e-mail: rmmian@noble.org
Tel.: +1-580-2246840
Fax: +1-580-2246802

I. Eujayl
Department of Food Science and Human Nutrition, Washington State University,
E202 Food Quality Building, WWQL,
Pullman, WA 99164, USA

Introduction

Microsatellites or simple sequence repeats (SSRs) have become one of the most useful molecular marker systems in plant breeding. They are widely used in cultivar fingerprinting, genetic diversity assessment, molecular mapping, and marker assisted breeding. The development of SSR markers from genomic libraries is expensive and inefficient (Squirrell et al. 2003). However, with the availability of large numbers of expressed sequence tags (ESTs) and other DNA sequence data, development of SSR markers through data mining has become an efficient and low cost option for many plant species. About 1–5% of the ESTs in different plant species have been found to have SSRs (20 bp or more in length) for marker development (Kantety et al. 2002). Development of SSR markers from ESTs has been reported for a number of plant species, including grape (Scott et al. 2000), sugarcane (Cordeiro et al. 2001), durum wheat (Eujayl et al. 2002), rye (Hackauf and Wehling 2002), barley (Thiel et al. 2003), and the barrel medic *Medicago truncatula* (Eujayl et al. 2004).

The scope of EST-SSR marker development is limited to the species for which sequence databases already exist. The SSR marker development for plant species lacking a sequence database can still be expensive and time consuming. An alternative approach for SSR marker development in those species could be the utilization of SSR markers from related species. Evidences suggest high rates of transferability of SSR loci across species (>50%) within a genus (Peakall et al. 1998; Gaitàn-Solís et al. 2002; Dirlewanger et al. 2002; Thiel et al. 2003; Eujayl et al. 2004). However, the transferability of SSR loci across genera and beyond seems to be low (White and Powell

1997; Peakall et al. 1998; Roa et al. 2000; Thiel et al. 2003). As the EST-SSR markers are derived from transcribed regions of DNA, they are expected to be more conserved and have a higher rate of transferability than genomic SSR markers (Scott et al. 2000). On the other hand, the conserved nature of the EST-SSRs may limit their polymorphism. The EST-SSRs also have a higher probability of being functionally associated with differences in gene expression than the genomic DNA-derived SSRs (Ayers et al. 1997). The function of many of the transcripts containing SSRs can be predicted through homology searches from the genomic databases. The SSR markers developed from DNA sequences with putative functions can be evaluated for association with phenotypes.

Tall fescue is a major cool season forage and turf grass in the temperate regions of the world and it belongs to the grass family Poaceae, subfamily Pooideae, and tribe Poeae (Soreng and Davis 1998). Tall fescue is a hexaploid ($2n=6\times=42$) consisting of three genomes (PG1G2) and has a genome size of approximately $5.27–5.83\times10^6$ kb (Seal 1983). The P ($2n=2\times=14$) genome is derived from *Festuca pratensis* while the G1G2 ($2n=4\times=28$) genome is derived from *F. arundinacea* var 'glaucescens' (Sleper 1985). The genus *Lolium*, which is closely related to the *Festuca* genus, contains several diploid species of interest to researchers ($2n=2\times=14$) including the outcrossing perennial *L. perenne*, the annual *L. multiflorum*, and the self-pollinating annual *L. temulentum*, all of which are cross-compatible (Terrell 1966).

The conservation of grass genomes has been comprehensively documented. Comparative genomics has become an important strategy for extending genetic information from model species to more complicated species (Gale and Devos 1998). Comparative mapping with DNA markers can reveal important information to understand the extent of genetic conservation among the genomes of various grass species. RFLP marker-based comparative maps have been constructed for several grass genera (Van Deynze et al. 1995a, 1995b). Comparative maps were also constructed within the Panicoideae subfamily with maize, sorghum, and sugarcane (Whitkus et al. 1992; Melake Berhan et al. 1993; Grivet et al. 1994) and Pooideae subfamily using wheat, barley, and rye (Naranjo et al. 1987; Devos et al. 1993). Jones et al. (2002) conducted the first comparative mapping study between a forage grass species, *Lolium perenne*, and two Triticeae grasses, oat and rice. Recently, Alm et al. (2003) have reported that meadow fescue genome was highly orthologous and colinear with those of ryegrass, oat, maize, and sorghum. Genome conservation has also been revealed among other distantly related grass species (Ahn and Tanksley 1993; Kurata et al. 1994; Devos et al. 1998). Until recently, the comparative genomics efforts have relied heavily on the hybridization based RFLP technique, and the resolution of the comparative maps is generally too low for determination of microsynteny (Kilian et al. 1997). However, recently, sequence-based comparative maps have been developed for rice-wheat (Sorrells et al. 2003) and sorghum-rice (Klein et al. 2003) that enhance the resolution considerably. The application of a PCR-based co-dominant marker system for comparative genomics would be highly desirable, because such a marker system can increase the efficiency of transferring genetic information across species. Due to their transferability, the EST-SSR markers have good potential for application in comparative genomics (Kantety et al. 2002; Thiel et al 2003; Eujayl et al. 2004).

An EST project was started on tall fescue at the Samuel Roberts Noble Foundation (Mian et al. 2002). This project was initiated to (1) develop a large set of SSR primer pairs from tall fescue ESTs, and (2) examine the ability of the primer pairs to amplify reliable PCR products in a range of grass species.

Materials and methods

Mining of ESTs for SSRs

Twenty thousand tall fescue EST sequences developed at The Noble Foundation were 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. The SSRIT program was run on a local computer and the parameters were set for detection of perfect di-, tri-, tetra-, and pentanucleotide motifs with a minimum of ten, six, five, and four repeats, respectively. The following information was extracted from the SSRIT output and the tall fescue EST database using PERL programs: EST identifiers, repeat motifs, number of repeats, start and end positions of the SSRs within the respective ESTs, and gene functional annotations.

Primer design

One hundred fifty seven primer pairs were designed from non-redundant EST sequences with SSRs using the Primer3 software (freely available at http://www.genome.wi.mit.edu/genome_software/other/primer3.html). The major parameters for designing the primers were: primer length from 18 to 24 nucleotides, 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. The primer pairs were custom synthesized by Qiagen/Operon Technologies (Alameda, Calif., USA).

Plant materials

Eleven genotypes representing seven species of four grass genera were used for the initial screening of the SSR primers (Table 1). These genotypes exemplify different ploidy levels (from diploids to hexaploids), variation in growth habit from annual to perennial, and differences in

reproductive strategy from cross-pollination to self-pollination. The three tall fescue genotypes are the parents of two mapping populations (Xu et al. 1995; Saha et al. 2003). 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 two genotypes of ryegrass are from a *Lolium perenne* × *Lolium multiflorum* cross. These two genotypes are the parents of a ryegrass mapping population (Warnke et al. 2003). The two rice genotypes are the parents of a mapping population (Temnykh et al. 2000) and the two wheat genotypes are the parents of a mapping population (Van Deynze et al. 1995a). The inheritance of a number of SSR markers were tested on a tall fescue (Saha et al. 2003) and a ryegrass mapping population (Warnke et al. 2003).

DNA isolation, PCR amplification, and detection of SSR bands

Approximately 200 mg of tissue from young leaves of each plant was collected in a 2.0 ml Eppendorf tube, immediately frozen in liquid nitrogen and ground to fine powder using a Mixer Mill Type MM 300 (Retsch, Hannover, Germany). The DNA was extracted using Qiagen DNeasy Plant Mini kit (Qiagen, Valencia, Calif., USA). The DNeasy protocol was used with the following modifications: 500 µl of AP1 buffer, 5 µl of RNase A, 165 µl AP2 buffer, and 90 µl of AE buffer was added to each tube. The DNA concentrations were quantified using a HOEFER Dyna Quant 200 (Amersham Biosciences, Piscataway, N.J., USA) DNA fluorometer. Twenty ng of DNA was used as template for each PCR reaction. The PCR reactions were run under standard conditions for all primers for the screening using one unit of *Taq* DNA polymerase with 10× PCR reaction buffer (Invitrogen Life Technologies, Carlsbad, Calif., USA) or one unit of AmpliTaq Gold with GeneAmp PCR bufferII (Applied Biosysytems/Roche, Branchburg, N.J., USA), 3 mM $MgCl_2$, 200 µM of dNTPs, and 0.2 mM of each primers in a 10 µl reaction. After 10 min at 95°C, 40 cycles were performed with 50 s at 95°C, 50 s at a temperature between 58°C and 64°C (the optimum annealing temperature for the respective primer pair was used for each PCR reaction), 90 s at 72°C, and a final extension step of 10 min at 72°C. All primers were initially screened using *Taq* DNA polymerase. AmpliTaq Gold was used for SSR primers that amplified non-specific bands with the use of *Taq* DNA polymerase. Each of the primer pairs were screened twice to confirm the repeatability of the observed bands in each genotype. The PCR products were resolved on 6% polyacrylamide denaturing gels (Gel Mix 6, Invitrogen LifeTechnologies). The gels were silver stained using Silver Sequence Kit (Promega, Madison, Wis., USA) for SSR bands detection.

Allele scoring and evaluation of polymorphism

The band size is reported for the most intensely amplified band for each SSR or the average of the stutter if the intensity was the same using a 10 bp DNA ladder (Invitrogen Life Technologies) as the reference point. Null alleles were assigned to genotypes with confirmed no amplification product under the standard conditions. The polymorphism was determined according to the presence or absence of the SSR locus.

Sequencing of PCR bands

PCR products were separated in a 6% polyacrylamide gel and selected bands were excised and dipped in 10 µl of nuclease free water for 30 min. Another round of PCR was made following the same protocol with extracted DNA as template. The PCR products were separated in a 3% agarose gel and DNA was extracted using the DNA Gel Extraction Kit (Millipore Corporation, Bedford, Mass., USA). Samples were then dried to a 20 µl volume using a Speed Vac Plus SC110A (Thermo Electron Corporate, Waltham, Mass., USA). DNA concentration was measured against a Molecular Mass Standard (Bio-Rad Laboratories, Hercules, Calif., USA) using the Quantity One-4.4.0 software. Each 20 µl of sequencing reaction consisted of 100 ng of DNA, 4.0 pmol of primer, 8 µl of sequencing reaction. The PCR conditions were: an initial hold at 94°C for 4 min followed by 25 cycles of 96°C for 10 s, 60°C for 5 s and 62.5°C for 4 min. Sequencing was performed

Table 1 The grass genotypes and species used for initial screening of tall fescue EST-SSR primers

Common name	Scientific name	Genotype/cultivars	Genome
Tall fescue	*Festuca arundinacea*	HD28-56	$2n=6\times=42$
		R43-64	
		Kentucky-31	
Meadow fescue	*F. pratensis*	PIWF92-109-2	$2n=2\times=14$
Tetraploid fescue	*F. arundinacea* var. *glaucescens*	W-2	$2n=4\times=28$
Ryegrass	*Lolium perenne/multiflorum*	MFA-4	$2n=2\times=14$
		MFB-2	
Rice	*Oryza sativa*	IR-64	$2n=2\times=24$
		Azucena	
Wheat	*Triticum aestivum*	Opata	$2n=6\times=42$
		W7984	

Table 2 Screening results for 157 tall fescue EST-SSR primers on different forage and cereal species. The percentage of the total number (157 primer pairs) is given. The polymorphism percentage is given as the percentage of the useful primer pairs for each species

Species	Amplification		Average no. of bands/primer	Polymorphism (%)
	Number	Percent		
Tall fescue	145	92	2.78	66
Ryegrass	135	86	2.30	66
Meadow fescue	130	83	1.88	–
Tetraploid fescue	128	82	2.71	–
Rice	93	59	1.92	43
Wheat	112	71	2.52	38

using an ABI 3100 sequencer with stdSeq50_P0P6Def. as the run module and BC-3100-SeqOffFtOff.saz as the analysis module. The sequence data were analyzed with DNASTAR software using Clustal W as the alignment tool.

Determination of genetic relationship among genotypes

For genetic diversity evaluation, bands were scored as present (1) or absent (0) for each of the 11 genotypes on the panel. The genetic similarity among the genotypes were calculated according to DICE (qualitative data module) coefficients (Dice 1945) of the NTSYS-pc software package (version 2.1; Rohlf 2002). The DICE coefficients are same as the Nei and Li (1979) similarity coefficients. The similarity matrix of DICE coefficients were used to construct UPGMA dendograms using the SAHN module of the NTSYS-pc. The FIND module of the NTSYS-pc software was used to identify all trees that could result from different choices of tied similarity or dissimilarity values. The reliability of clustering was tested by Mantel test statistics for the comparison of the similarity matrix and the cophenetic matrix (Rohlf 2002).

Results

Distribution of EST-SSRs and marker development

Two hundred sixty one (1.3%) of the 20,000 ESTs contained the SSRs specified in the search. The trinucleotide motifs were the most abundant type of SSRs (70%), followed by di- (20%), tetra- (5%) and pentanucleotide (5%) motifs in the database. The GA/CT repeats were the most abundant dinucleotide repeats, while CCG/GGC were the most abundant trinucleotide repeat motifs found in the ESTs (data not shown).

Of the 261 SSR containing ESTs, primers were designed for 157 (60%) ESTs. The remaining ESTs were not used because: (1) the EST was redundant with another one from which primer pairs were already developed, (2) the DNA sequences flanking the SSRs (at least on one side of the SSRs) were too short (generally <40 nucleotides), or (3) the flanking sequences were inappropriate for designing high quality primer pairs (e.g., low GC content). Out of 157 primer pairs, 145 (92%) produced clear SSR type bands in at least one of the 11 grass species tested [see the Electronic Supplementary Material (ESM)]. The sequences of these 145 primer pairs, along with the marker name, EST id, expected band size, annealing temperature, SSR repeat motif and number, observed band size range, plant species with clean repeatable bands, and gene functional annotation are presented in the ESM. The remaining 12 primer pairs either had no amplification products or produced a number of faint bands indicative of non-specific amplifications. Amplification failure can result from the presence of an intron within the primer sequences that prevents primer annealing or a large intron in the flanking region that disrupts PCR extension. These 12 primers were not investigated further. The percentages of primer pairs that produced clean repeatable bands were 92, 86, 83, 82, 59, 71% for tall fescue, ryegrass, meadow fescue, tetraploid fescue, rice, and wheat, respectively (Table 2).

Sequence homology searches

A BLAST search (BLASTN and/or BLASTX) for the ESTs (from which the 145 useful primer pairs were designed) was performed against the GenBank database (on December 29, 2003). One hundred and three (71%) ESTs had significant homology with the sequences in the database. The best putative gene functions for these ESTs are listed in the ESM .

Polymorphism and allele frequency of SSR markers

The polymorphism of the SSR markers was tested on the parents of the tall fescue, ryegrass, wheat, and rice mapping populations (Table 1). The marker polymorphism between the parents of the tall fescue (HD28-56 × R43-64), ryegrass, rice, and wheat was 66, 66, 43 and 38%, respectively (Table 2). The average number of SSR bands ranged from 2.78 for tall fescue to 1.88 for meadow fescue. Across all seven species, the number of different bands detected by these markers varied from 2 to 15 with an average of 12 (data not shown).

Table 3 Percent similarity in nucleotide sequences of different forage and cereal grass species amplified with the primer pair NFFA147. A longer (R43-64-1) and a shorter (R43-64-2) band were sequenced from tall fescue genotype R 43-64 while only one band was sequenced from all other genotypes

	HD28-56	R43-64-2	W-2	MFA-4	MFB-2	IR64	Azucena	Opata	W7984
R43-64-1	100.0	100.0	99.0	97.6	97.6	89.5	89.5	92.7	92.3
HD28-56		100.0	99.0	97.6	97.6	89.5	89.5	92.7	92.3
R43-64-2			99.0	98.5	98.5	90.8	90.8	92.7	92.7
W-2				99.5	99.5	90.3	90.3	93.7	93.7
MFA-4					100.0	86.1	86.1	93.2	92.8
MFB-2						86.1	86.1	93.2	92.8
IR64							100.0	85.4	84.6
Azucena								85.4	84.6
Opata									99.5

Sequence comparison of SSR bands

The sequence similarities of the PCR products amplified using a single EST-SSR primer pair (NFFA147) from ten genotypes across different grass species were very high (Table 3). The sequences of three PCR bands (one from HD28-56 and two from R43-64) from tall fescue were identical. The sequence identities across all genotypes of the *Festuca* and *Lolium* species (tall fescue, tetraploid fescue, and perennial ryegrass) were ≥97.6%. The sequence identity of wheat with the forage grass species was also high (≥92.3%). Rice had the least sequence identity with the other species in the panel (between 84.6 and 90.8%) and this was reflected in the frequency of marker amplification in different species.

The alignment of the sequences is shown in Fig. 1. All three bands from tall fescue had identical flanking sequences and they also contained the (CTG)*n* repeat motif. The two monomorphic bands from two tall fescue genotypes both had seven CTG repeats, while shorter band had four CTG repeats. Thus the marker polymorphism was due to the variation in the number of repeats. The tetraploid fescue (W-2) had four CTG repeats and an identical length of the flanking region with tall fescue, but there were two single base substitutions in the 3’ end of the flanking region. The *Lolium* genotypes had four CTG repeats and one more repeat was interrupted by the substitution of a C with a T. There were also three single base substitutions and an insertion of an A in the 3’ end of the flanking region. Wheat genotype W7984 had two CTG repeats and Opata had three CTG repeats. There were also a more than a dozen base substitutions and insertion of two bases in the flanking regions when compared to tall fescue. Neither of the two rice genotypes had any CTG repeat. Among all species, rice had the highest number of substitutions and insertions of bases (26 bases) in the flanking regions when compared to tall fescue.

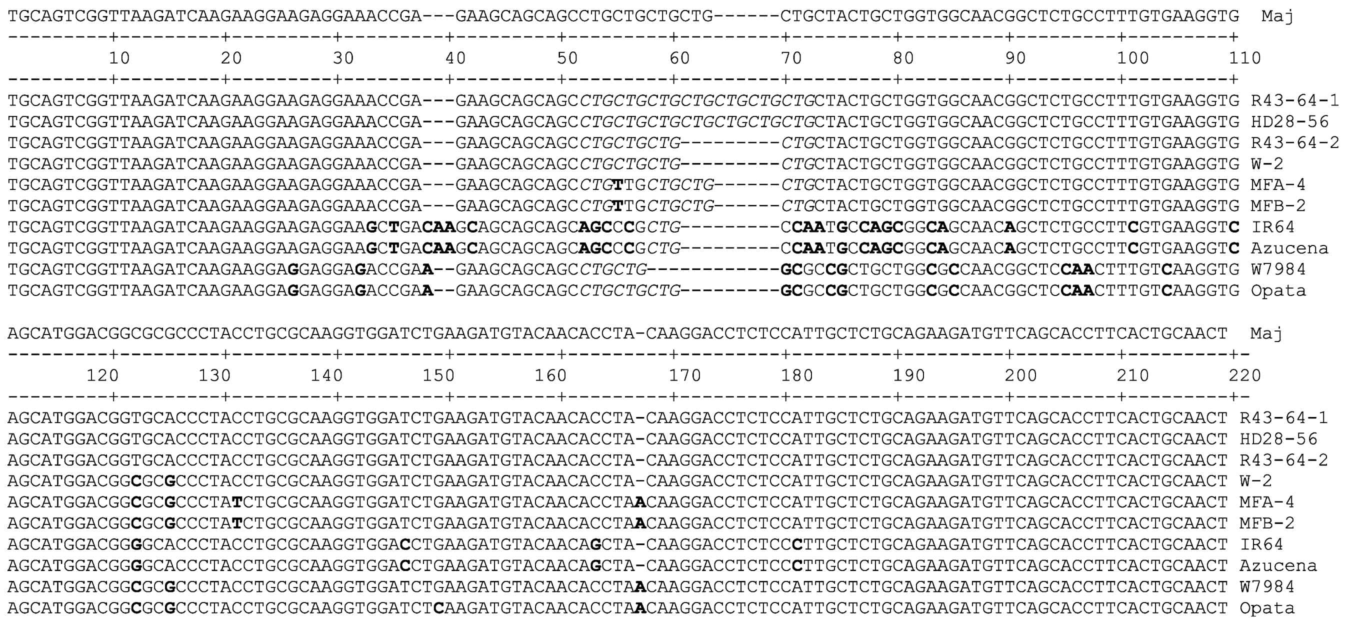

Fig. 1 CLUSTAL-W alignment of sequences obtained from selected PCR bands amplified by primer pair NFFA147 in two tall fescues (HD28-56 and R43-64), a tetraploid fescue (W-2), two ryegrass (MFA-4 and MFB-2), two rice (IR64 and Azucena) and two wheat (W7984 and Opata) genotypes. One band was sequenced from each genotype except for R43-64, from which a longer band (R43-64-1) and a shorter band (R43-64-2) were sequenced. The tall fescue sequences were aligned first and then the other species were aligned in relation to the tall fescue sequences. The changes in sequences compared to tall fescue are marked in *bold letters*. The (CTG)*n* repeats are shown in *italics*. The primer pair was designed from a tall fescue EST containing seven CTG repeats

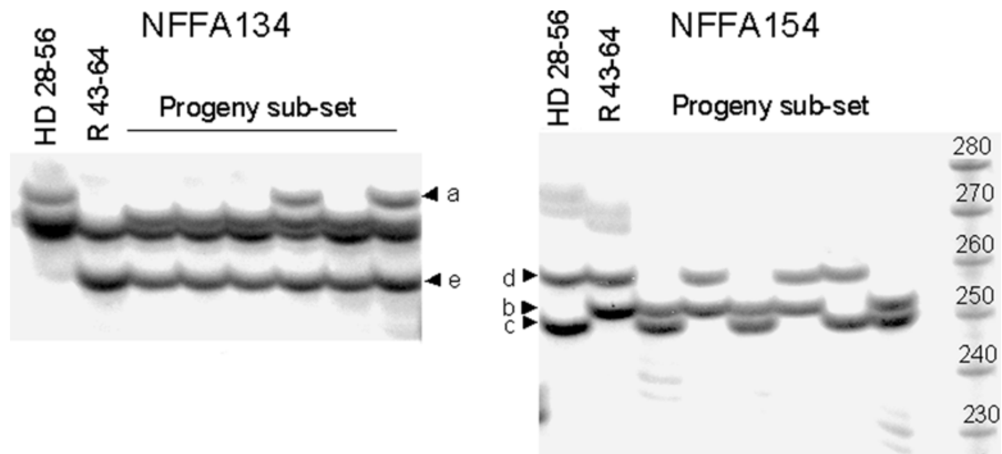

Fig. 2 The polyacrylamide gel images of SSR bands amplified by primer pairs NFFA134 and 154 in the parents and six F_1 progeny of the HD28-56 × R43-64 tall fescue mapping population. The *right*-most lane shows the size standard (10 bp)

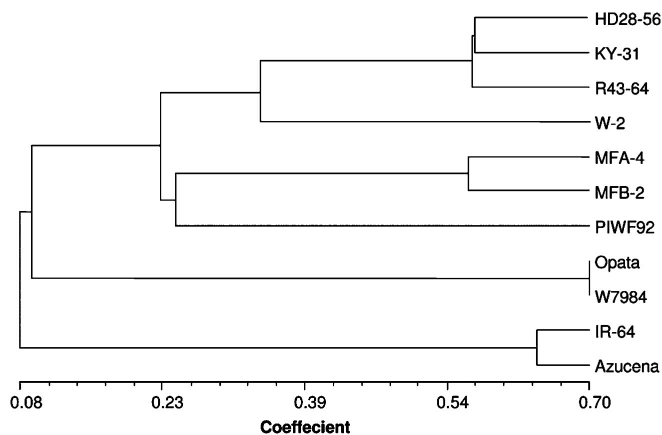

Fig. 3 The UPGMA tree of the 11 grass genotypes based on 1,007 EST-SSR markers. The genotypes include three tall fescue (*HD28-56, KY-31, and R43-64*), one tetraploid fescue (*W2*), two ryegrass (*MFA-4 and MFB-2*), one meadow fescue (*PIWF92*), two wheat (*Opata and W7984*), and two rice (*IR-64 and Azucena*)

Inheritance and segregation of the polymorphic markers

The inheritance and segregation of all polymorphic markers were tested on a subset of six progeny lines of each of the HD28-56 × R43-64 tall fescue and the ryegrass mapping populations. All bands present in the parents were detected in the subset of progeny in each population (data not shown). Figure 2 shows the amplification and inheritance of SSR alleles in tall fescue amplified by two primer pairs. Three alleles (a, b, c) were present in one parent but segregated in the F_1 progeny. Such alleles are present in a heterozygous state (Aa) in the parent and segregated in a 1:1 ratio in the F_1 progeny. Allele "d" was present in both parents, yet segregated in the progeny. This allele is present in a heterozygous state (Aa × Aa) in both parents and segregated in a 3:1 ratio in the progeny. Allele "e" was present in one parent and it was monomorphic in the progeny. This is an allele that is in a homozygous (AA) state in the parent and thus will not segregate in the F_1 progeny.

Amplification of markers across species

All of the 145 primer pairs amplified clean repeatable PCR bands in tall fescue. One hundred and twenty one (83%) primer pairs worked across all three *Festuca* species while 118 (81%) primer pairs worked across the *Festuca-Lolium* species (see the ESM). Fifty seven percent of the primer pairs produced clean PCR products in all six grass species. Only eight primer pairs were specific to tall fescue, indicating that these sequences were not conserved across species.

Evaluation of genetic relationships among species

One thousand seven SSR bands from 80 primer pairs were used for the evaluation of genetic diversity among the 11 genotypes from the seven grass species. The SSR markers showed the discriminatory powers to clearly separate the grass species into distinct clusters (Fig. 3). The forage species of the *Festuca-Lolium* complex were clustered together separately from rice and wheat. The FIND module revealed a single tree as shown in Fig. 3. The Mantel test statistics indicated that the goodness-of-fit for the cluster analysis was very high (r=0.99).

Discussion

Distribution of EST-SSRs and marker development

The percent of ESTs in the database (1.3%) is low compared to barley (3.4%), wheat (3.2%), rice (4.7%), and sorghum (3.6%) (Kantety et al. 2002) and *Medicago truncatula* (3.0%) (Eujayl et al. 2004) but comparable to maize (1.5%) (Kantety et al. 2002). The tall fescue genome appears to be GC rich rather than AT rich as has been reported for some plant genomes, e.g., soybean and other legumes (Brown-Guedira et al. 2000). Our findings are similar to those reported by Scott et al. (2000) and Kantety et al. (2002). Kantety et al. (2002) reported that GA/CT was the most abundant di-nucleotide repeats in all five grass species. They also reported that CCG/GGC was

the most abundant tri-nucleotide repeat motif in four of the five grass species, wheat being the exception. The distribution of di-, tri-, tetra-, and pentanucleotide repeats were similar to that reported by earlier investigators (Eujayl et al. 2004; Thiel et al. 2003; Kantety et al. 2002).

The low rates of polymorphism of EST-SSR markers may limit the value of these markers in some species (Eujayl et al. 2002). However, this problem seems to be more important for self-pollinated species. The high rates of polymorphism reported for the outcrossing species in this study is supported by earlier findings in *Medicago* spp. (Eujayl et al. 2004).

Non-specific amplifications were a problem for some (almost 30%) of the tall fescue EST-SSR markers. The use of AmpliTaq GOLD and higher annealing temperatures was effective in reducing the non-specific amplifications for some of these primer pairs. A significant number of primer pairs produced more bands than expected in some species, but the bands were repeatable. This may be due to amplifications of loci from the duplicated genomic regions of each species in question. Yu et al. (2004) reported that 39% of EST-SSR markers detected multiple loci in wheat probably due to the high rate of conservation of EST-SSRs as well as 25–30% gene duplication (Anderson et al. 1992).

Amplification and frequency of EST-SSR markers across species

A high percentage (57%) of the primer pairs produced clear bands across a very diverse set of seven grass species, indicating a high level of sequence conservation among these species. Thiel et al. (2003) reported transferability of 40% of barley EST-SSR markers to rice. Almost 80% of the primer pairs were conserved across all four *Festuca-Lolium* species. The *Lolium* species in this study are closely related to *Festuca*, and are compatible for intergeneric sexual hybridization (Terrell 1966). In addition, meadow fescue and the tetraploid *Festuca arundinacea* var. *glaucescens* are the two putative progenitors of tall fescue.

EST-SSR markers for evaluation of genetic relationships among grasses

The results from our survey of the limited number of genotypes from six grass species indicate that EST-SSR markers could be used for evaluation of both within and among species genetic relationships. The SSR marker-based genetic relationship among the *Festuca-Lolium* species reported in this study is in agreement to their RFLP marker-based relationship reported by Xu and Sleper (1994). Thiel et al. (2003) also found that RFLP and EST-SSR markers yielded a similar pattern of genetic diversity among 54 barley cultivars. The tetraploid fescue showed a closer relationship with hexaploid tall fescue compared to the relationship between tall fescue and meadow fescue. The closer relationship of tall fescue to tetraploid fescue rather than to meadow fescue may be because tetraploid fescue shares two genomes (G1G2) whereas meadow fescue shares only one (P) genome with tall fescue (PG1G2).

Thiel et al. (2003) pointed to the limitation of interspecific transfer of SSR markers for diversity studies, because (1) SSR markers of identical band sizes may not be identical by descent (homoplasy), and (2) alleles of different size may occur from complex mutational events (e.g., insertion or deletion in both SSR regions and flanking regions) rather than from simple variation in the number of SSR repeats. Peakall et al. (1998) stated that size homoplasy due to complex mutations within and beyond the SSR region, and genetic relationships among taxa and beyond were likely to be underestimated. An example of size homoplasy between R43-64-2 (tall fescue) and W-2 (tetraploid fescue) was observed in this study as well as many examples of insertion/deletions (indels) and substitutions in the flanking regions. Such events were more frequent among the more distantly related species (e.g., fescue vs rice or wheat). Also in this study, the separation of the taxa is based on small samples of intraspecific genetic variation that may reduce the reliability of the phylogenetic relationships of the outcrossing species. Thus, the usefulness of the EST-SSR markers for determination of phylogenetic relationships among grass species will need to be tested on a larger set of species with more entries from each species.

The tall fescue EST-SSR primers developed in this study will be useful for comparative mapping among grass species, particularly among the forage grass species belonging to the *Festuca-Lolium* complex because of the limited number of cross-species markers available and the limited knowledge of their genome relationships. Relatively understudied forage grass species such as tall fescue will gain the most from comparative mapping among grass species. High levels of synteny and co-linearity of *Lolium perenne* (Jones et al. 2002) and *Festuca pratensis* (Alm et al. 2003) genomes with the genomes of a number major grass species have been established. Thus comparison of tall fescue genetic map with those of ryegrass, meadow fescue and major cereal crops (e.g., wheat, oat, barley, and rice) will facilitate transfer of genetic information among these species.

Acknowledgements We thank Dr. Mark E. Sorrells for critically reviewing this manuscript. We also thank Konstantin Chekhovskiy, Jennifer Black, Ann Harris, and Jarrod Steele for their technical support. This research work was funded by The Samuel Roberts Noble Foundation.

References

Ahn SN, Tanksley SD (1993) Comparative linkage maps of the rice and maize genomes. Proc Natl Acad Sci USA 90:7980–7984

Alm V, Fang C, Busso CS, Devos KM, Vollan K, Grieg Z, Rognli OA (2003) A linkage map of meadow fescue (*Festuca pratensis* Huds.) and comparative mapping with other Poaceae species. Theor Appl Genet 108:25–40

Anderson JA, Ogihara Y, Sorrells ME, Tanksley SD (1992) Development of a chromosomal arm map for wheat based on RFLP markers. Theor Appl Genet 83:1035–1043

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

Brown-Guedira GL, Thomson JA, Nelson RL, Warburton ML (2000) Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. Crop Sci 40:815–823

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

Devos KN, Millan T, Gale MD (1993) Comparative RFLP maps of homologous group 2 chromosomes of wheat, rye and barley. Theor Appl Genet 85:784–792

Devos KM, Wang ZM, Beales J, Sasaki T, Gale MD (1998) Comparative genetic maps of foxtail millet (*Setaria italica*) and rice (*Oryza sativa*). Theor Appl Genet 96:63–68

Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302

Dirlewanger E, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanettto 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

Eujayl I, Sorrells ME, Wolters P, Baum M, Powell W (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor Appl Genet 104:399–407

Eujayl I, Sledge MK, Wang L, May GD, Chekhovskiy K, Zwonitzer JC, Mian MAR (2004) *Medicago truncatula* EST-SSRs reveal cross-species genetic markers for *Medicago* spp. Theor Appl Genet 108:414–422

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 42:2128–2136

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

Grivet L, D'Hont A, Dufour P, Roques D, Glaszmann JC (1994) Comparative mapping of sugar cane with other species within the Andropogoneae tribe. Heredity 73:500–508

Hackauf B, Wehling P (2002) Identification of microsatellite polymorphisms in an expressed portion of the rye genome. Plant Breed 121:17–25

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

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

Kilian A, Chen J, Han F, Steffenson B, Kleinhofs A (1997) Towards map-based cloning of the barley stem rust resistance genes *Rpg1* and *rpg4* using rice as an intergenomic cloning vehicle. Plant Mol Biol 35:187–195

Klein PE, Klein RR, Vrebalov J, Mullet JE (2003) Sequence based alignment of sorghum chromosome 3 and rice chromosome 1 reveals extensive conservation of gene order and one major chromosomal rearrangement. Plant J 34:605–621

Kurata N, Moore G, Nagumara Y, Foote T, Yano M, Minobe Y, Gale M (1994) Conservation of genome structure between rice and wheat. Biotechnology 12:276–278

Melake Berhan A, Hulbert SH, Bulter LG, Bennetzen JL (1993) Structure and evolution of the genomes of *Sorghum bicolor*and *Zea mays*. Theor Appl Genet 86:598–604

Mian, MAR, Wang Z, Eujayl I, Zhang J, Chen L, Scott AD, Gonzales RA, Hopkins AA, May GD (2002) EST and expression analysis in tall fescue (abstract). In: Plant and animal genome X, the international conference on the status of plant and animal genome research, 12–16 January 2002, San Diego, p 92

Naranjo T, Roca P, Goicoechea PG, Giradlez R (1987) Arm homology of wheat and rye chromosomes. Genome 29:873–882

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

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 15:1275–1287

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

Rohlf FJ (2002) NTSYS-pc. Numerical taxonomy and multivariate analysis system, Version 2.10. Exeter Software, New York

Saha MC, Zwonitzer JC, Hopkins AA, Mian MAR (2003) A molecular linkage map of a tall fescue population segregating for forage quality traits. In: Molecular breeding of forage and turf, 3rd international symposium, 18–22 May 2003, Dallas, p 86

Scott KD, Eggler P, Seaton G, Rossetto M, Ablett EM, Lee LS, Henry RJ (2000) Analysis of SSRs derived from grape ESTs. Theor Appl Genet 100:723–726

Seal AG (1983) DNA variation in *Festuca*. Heredity 50:225–236

Sleper DA (1985) Breeding tall fescue. J Plant Breed Rev 3:313–342

Soreng RJ, Davis JI (1998) Phylogenetic and character evolution in the grass family (Poaceae): simultaneous analysis of morphology and chloroplast DNA restriction site character sets. Bot Rev 64:1–85

Sorrells ME, La Rota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD, Miftahudin M, 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. Genet Res 13:1818–1827

Squirrell J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M, Powell W (2003) How much effort is required to isolate nuclear microsatellites from plants? Mol Ecol 12:1339–1348

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

Terrell EE (1966) Taxonomic implications of genetics in ryegrass (*Lolium*). Bot Rev 32:138–164

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, Dubcovsky J, Gill KS, Nelson JC, Sorrells ME, Dvorak J, Gill BS, Lagudah ES, McCouch SR, Appels 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

Warnke S, Barker R, Sim S, Jung G, Mian MAR (2003) Identification of flowering time QTLs in an annual x perennial ryegrass mapping population (abstract). In: Molecular breeding of forage and turf, 3rd international symposium, 18–22 May 2003, Dallas, p 122

White G, Powell W (1997) Isolation and characterization of microsatellite loci in *Swietenia humilis* (Meliaceae): an endangered tropical hardwood species. Mol Ecol 6:851–860

Whitkus R, Doebley J, Lee M (1992) Comparative genome mapping of sorghum and maize. Genetics 132:1119–1130

Xu WW, Sleper DA (1994) Phylogeny of tall fescue and related species using RFLPs. Theor Appl Genet 88:685–690

Xu WW, Sleper DA, Chao S (1995) Genome mapping of polyploid tall fescue (*Festuca arundinacea* Schreb.) with RFLP markers. Theor Appl Genet 91:947–955

Yu J, Dake TM, Singh S, Benscher D, Li W, Gill B, Sorrells ME (2004) Development and mapping of EST-derived simple sequence repeat (SSR) markers for hexaploid wheat. Genome Res (in press)