Euphytica (2007) 153:73–85
DOI 10.1007/s10681-006-9243-z

Characterization of dinucleotide and trinucleotide EST-derived microsatellites in the wheat genome

Agata Gadaleta · Giacomo Mangini · Giuseppina Mulè · Antonio Blanco

Received: 14 March 2006 / Accepted: 14 July 2006 / Published online: 30 August 2006

Abstract Over the past decade microsatellites or simple sequence repeats (SSRs) have attracted a considerable amount of attention from researchers. The aim of the present paper was to analyse expressed sequence tag-derived SSR (EST-SSR) marker variability in wheat and to investigate the relationships between the number and type of repeat units and the level of microsatellite polymorphism. Two hundred and forty-one new EST-SSR markers available in a public database (http//wheat.pw.usda.gov) were characterized in eight durum wheat cultivars (Svevo, Ciccio, Primadur, Duilio, Meridiano, Claudio, Latino, Messapia), two accessions of *Triticum turgidum* var. *dicoccoides* (MG4343, MG29896), one accession of *T. turgidum* var. dicoccum (MG5323) and in the common wheat cv. Chinese Spring. Of these, 201 primer pairs (83.4%) amplified PCR products successfully, while the remaining 40 (16.6%) failed to amplify any product. Of the EST-SSRs analysed, 45.2% of the primer pairs amplified one or two PCR products. Multiple discrete PCR products were observed among both di- and trinucleotide EST-SSR markers (31.2 and 40.5%, respectively). Markers based on dinucleotide microsatellites were more polymorphic than those based on trinucleotide SSRs in the 12 wheat genotypes tested (68.9 and 52.7%, respectively). An average of 2.5 alleles for dinucleotide and 2.0 alleles for trinucleotide SSRs was observed. The data reported in the present work indicate the presence of a significant relationship between motif sequence types and polymorphism. The primer set based on the AG repeat motif showed the lowest percentage of polymorphism (55.0%), while the primer set based on the AC repeat motif showed t he highest percentage (85.0%). Among trinucleotide SSRs, the AGG microsatellite markers showed the highest percentage of polymorphism (70.0%), and the ACG motif the lowest value (25.0%). The characterization of these new EST-SSR markers and the results of our studyon the effect of repeat number and type of motifs could have important applications in the genetic analysis of agronomically important traits, quantitative trait locus discovery and marker-assisted selection.

Keywords EST-SSR · Microsatellites · Molecular markers · Wheat

A. Gadaleta (✉) · G. Mangini · A. Blanco
Department of Agro-Forestry and Environmental Biology and Chemistry, University of Bari, Via Amendola 165/A, 70126 Bari, Italy
e-mail: agata.gadaleta@agr.uniba.it

G. Mulè
Institute of Sciences of Food Production, National Research Council (CNR), 70126 Bari, Italy

Abbreviations
EST Expressed sequence tag
SSR Simple sequence repeat

Introduction

The level of molecular polymorphism in cultivated hexaploid and tetraploid wheats has been found to be low compared to many other species (Song et al. 2002). The use of molecular markers in genome analysis, the systematic mapping of agricultural important traits and marker-assisted selection have been greatly advanced by the development of reliable PCR-based markers, such as amplified fragment length polymorphism (AFLP) and microsatellite or simple sequence repeat (SSR) markers. Microsatellites represent a valid alternative marker system because of their abundance in plant genomes and easy detection as PCR-based molecular markers on polyacrylamide or high-resolution agarose gel. The advantages of SSR markers include high information content, codominant inheritance, locus specificity and reproducibility. Microsatellites are short (1–6 bp long) tandemly repeated DNA sequences that are highly polymorphic as a result of freque nt variations in the number of times the core sequence is repeated. The observed polymorphism may originate from unequal crossing-over or from slipped strand mispairing during replication. Microsatellites are ubiquitous in that they are found both in coding and non-coding regions of eukaryotic genomes, and they show high allelic variation (Turpeinen et al. 2001).

During the past 10 years a great deal of effort has been expended on the development of SSR markers based on genomic library screening. This process requires the construction and screening of a genomic library of size-selected DNA fragments with SSR-specific probes, followed by DNA sequencing of isolated positive clones, PCR primer synthesis and, finally, testing (Ostrander et al. 1992; Edwards et al. 1996). A major limitation in SSR development with this procedure is the time and cost required to isolate and characterize each locus by sequencing. A new source of SSRs has recently been found in the form of expressed sequences tags (ESTs). mRNA transcripts contain repeat motifs, and the abundance of microsatellites in the expressed sequences of many species make these markers very interesting because of a possible role in gene expression or function.

The overall frequency of microsatellites among species has been found to be inversely related to genome size and to the proportion of repetitive DNA but to remain constant in the transcribed portion of the genomes (Morgante et al. 2002). In addition, bioinformatic analysis indicated that the frequency of microsatellites is significantly higher in ESTs than in genomic DNA across all species (Morgante et al. 2002). At present over three million sequences from approximately 200 plant species have been deposited in publicly available plant EST databases. Many of the ESTs have been sequenced as an alternative to complete genome sequencing, thereby creating a formidable resource for microsatellite marker development (Rudd 2003). In several studies the genetic variation detected with EST-SSR markers was comparable to that revealed by genomicSSR markers (Cho et al. 2000; Leigh et al. 2003). Although EST-SSR markers have been shown to be less polymorphic than genomic SSR markers (Becker and Heun 1996; Eujayl et al. 2001), the use of markers from genic regions may be functionally more informative than SSRs from unexpressed chromosome regions. Moreover, EST-SSRs represent a unique opportunity: sequence information is readily available, and the need of genomic library screening can be avoided. The most important feature of the EST-SSR approach in studies on genetic variation is that primer pairs designed from EST-SSRs are more likely to function in distantly related species than primer pairs derived from genomic libraries (Kanteny et al. 2002).

Among SSRs the dinucleotide repeats are reported to be more polymorphic among plant species (Lagercrantz et al. 1993; Morgante and Olivieri 1993; Li et al. 2002), while trinucleotide repeats are over-represented in coding sequences, but less frequent than mono- and dinucleotide repeats in non-coding regions (Tòth et al. 2000; Gao et al. 2003). Furthermore, microsatellite distribution seems to differ between intergenic and intronic sequences (Morgante et al. 2002).

Microsatellites gain and lose repeat units at high rates, the rate depending on SSR repeat number and on the repeat motif (Schlötterrer and Tautz 1992).

The selection of suitable markers is one of the key factors for the success of molecular breeding programmes. Knowledge of a relationship between a specific class of SSR and the level of polymorphism among wheat cultivars may positively influence the selection of SSR-markers to be used in breeding programmes and may significantly increase the efficiency of a specific research strategy or the identification of markers linked to agronomically important traits.

In the present study we analysed the variability of EST-SSR markers and investigated relationships between the number of repeat units, type of motifs and level of polymorphism in wheat genotypes.

Materials and methods

Genetic material

A total of eight durum wheat (*Triticum turgidum* var. *durum*) cultivars (Svevo, Ciccio, Primadur, Duilio, Meridiano, Claudio, Latino, Messapia), two accessions of *T. turgidum* var. *dicoccoides* (MG4343, MG29896), one accession of *T. turgidum* var. *dicoccum* (MG5323) and cv. Chinese Spring of hexaploid wheat (*Triticum aestivum*) were used to screen for SSR polymorphism. The selected tetraploid wheat genotypes are parents of different mapping populations developed at the section of Genetics and Plant Breeding, DIBCA, University of Bari, Italy.

Genomic DNA was isolated from fresh leaves using the method described by Sharp et al. (1988), and subsequently purified by phenol-chloroform method. The purity of the DNA was determined by the A260/A280 ratio (1.6–1.8) and by restriction enzyme digestibility and agarose gel electrophoresis.

Microsatellite source

Polymorphism in the 12 wheat genotypes was investigated using 241 new EST-SSR primer pairs developed by Mauricio La Rota and available in the public database (http//wheat.pw.usda.gov). In general, EST-SSR markers produced high-quality patterns (Nicot et al. 2004). A total of 61 dinucleotide EST-SSR markers (20 each of classes AG and AC, 21 of the class AT) and 180 trinucleotide EST-SSRs (19 each of classes AAG, AAT, 20 each of classes ACC, ACG, ACT, AGC, AGG, and 21 for AAC and CCG) were examined. Primer pairs were chosen to produce PCR products ranging from 100 to 250 bp in length. Primer sequences, annealing temperature, and expected PCR product size are reported in the web site http//wheat.pw.usda.gov.

SSR detection

DNA amplifications were carried out in 25-µl reaction mixtures, each containing 100 ng template DNA, 2 µM of each microsatellite primer, 200 µM of each dNTP, 2.5 mM $MgCl_2$, 1× PCR buffer (10 mM Tris–HCl, pH 8.3, 10 mM KCl) and 1 U *Taq* DNA polymerase. The following PCR profile in a Perkin-Elmer DNA Thermal Cycler (Foster City, Calif.) was used: an initial denaturing step at 94°C for 3 min, followed by 35 cycles of 94°C for 1 min, 50/65°C for 1 min, 72°C for 2 min, with a final extension step at 72°C for 10 min. The amplification products were resolved on 6% polyacrylamide gels, stained with ethidium bromide (0.25 ng/µl) and visualized under UV light using a DIG-DOC (Bio-Rad, Hercules, Calif.) system.

Statistical analysis

The informativeness of any given DNA marker was measured by the polymorphism information content (PIC). According to Weber (1990) and Anderson et al. (1993) the PIC-value is calculated as follows:

$$\mathrm{PIC} = 1 - \sum_{i=1}^{k} P_i^2$$

where k is the total number of alleles detected for a microsatellite, and P_I is the frequency of the Ith allele in the set of 12 wheat genotypes investigated.

The null allele was not included to compute PIC value.

Results

In total, 241 primer pairs were selected from a collection of 5424 wheat EST-derived SSR markers (available at the web site http//wheat.pw.usda.gov) and tested on the genomic DNA of 12 wheat genotypes in order to assess amplification products and polymorphism level. The results of all 2892 combinations (241 SSRs × 12 wheat genotypes) were carefully examined. Only 201 (83.4%) of the primer pairs led to at least one PCR product; the other (16.6%) markers produced no product (Table 1). When these 201 amplifying primer pairs were assessed more than half gave one or two PCR products (55.7% as dinucleotide and 41.7% as trinucleotide SSRs, respectively). The other primer pairs amplified multiple PCR discrete products, including bands of expected and unexpected size. Multiple discrete PCR products were observed in both dinu cleotide and trinucleotide EST-SSRs (31.2% and 40.5%, respectively). Among the dinucleotides, SSRs with AG as the repeat motif showed the lowest value (20.0%), and those with AC as the repeat motif, the highest (45.0%); among the trinucleotides, SSRs with AGG and ACT as the repeat motif showed the lowest value (30.0%), while those with AAT as the repeat motif, the highest (52.6%).

Some differences between the expected and the observed size of the amplification products were observed. Of the markers amplifying bands, 8% amplified bands of a higher size than expected, suggesting the presence of introns in the corresponding genomic DNA. Less than 1.0% of the primers gave smaller products than expected.

Of the 2412 combinations (201 × 12 wheat genotypes) a total of 503 alleles were identified: 150 were detected with dinucleotide EST-SSR markers (2.5 alleles per primer pair) and 353 with trinucleotide EST-SSR markers (2.0 alleles per primer pair) (Table 2). The most informative microsatellite markers were TC66874 and BQ281574 with seven alleles within the dinucleotide SSRs and TC63424 and BF483209 with six alleles within trinucleotide SSRs.

The mean values of polymorphism were 68.9% for dinucleotide and 52.7% for trinucleotide SSR markers. Polymorphism was calculated as the number of primers that gave at least two alleles among the 12 genotypes analysed divided by the total number of primers amplifying PCR products. Among dinucleotide SSRs, the lowest value (55.0%) was found among the primer set based on the AG repeat motif, and the highest one (85.0%) was found among the AC motif primer

Table 1 Percentage of primer pairs that did not amplify a PCR product and amplified one or two PCR products, multiple PCR products and high-molecular weight bands, on the 12 wheat genotypes used in this study

Repeat motif	Primer pairs screened (no.)	PCR products			High-molecular weight bands (%)
		Number of products (%)	One or two discrete bands (%)	Multiple discrete bands (%)	
AC	20	0.0	55.0	45.0	75.0
AG	20	25.0	55.0	20.0	40.0
AT	21	14.3	57.1	28.6	18.8
AAC	21	19.0	33.4	47.6	33.3
AAG	19	26.3	42.1	31.6	64.3
AAT	19	10.5	36.8	52.6	64.7
ACC	20	5.0	35.0	60.0	65.0
ACG	20	45.0	20.0	35.0	72.7
ACT	20	10.0	60.0	30.0	50.0
AGC	20	10.0	50.0	40.0	77.8
AGG	20	10.0	60.0	30.0	27.8
CCG	21	23.8	38.1	38.1	37.5
Dimeric	61	13.1	55.7	31.2	47.1
Trimeric	180	17.8	41.7	40.5	54.0

Table 2 Number of primer pairs that amplified PCR products on the 12 wheat genotypes used in this study, number of alleles and PIC

Repeat motif	Primer pairs (no.)	Total number of alleles (no.)	Number of alleles per primer pair[a]	Mean value of PIC[a]
AC	20	57	2.9 (1–4)	0.48 (0–0.73)
AG	15	46	3.1 (1–7)	0.42 (0–0.82)
AT	18	47	2.6 (1–5)	0.35 (0–0.76)
AAC	17	48	2.8 (1–4)	0.26 (0–0.79)
AAG	14	26	1.9 (1–5)	0.05 (0–0.28)
AAT	17	39	2.3 (1–4)	0.42 (0–0.78)
ACC	19	56	2.9 (1–6)	0.35 (0–0.78)
ACG	11	22	2.0 (1–3)	0.10 (0–0.43)
ACT	18	38	2.1 (1–3)	0.30 (0–0.74)
AGC	18	44	2.4 (1–5)	0.29 (0–0.70)
AGG	18	43	2.4 (1–4)	0.30 (0–0.71)
CCG	16	37	2.3 (1–6)	0.10 (0–0.28)
Dimeric	53	150	2.5	0.41
Trimeric	148	353	2.0	0.24

[a] Range is given in parenthesis

set. Among the trinucleotide SSRs, AGG microsatellite markers showed the highest polymorphism (70.0%) and ACG microsatellite markers, the lowest value (25.0%) (Fig. 1).

The informativeness of EST-derived SSRs was also measured using the PIC. The PIC value was calculated for each of the polymorphic markers, and each band in each of the eight durum cultivars (Svevo, Ciccio, Primadur, Duilio, Meridiano, Claudio, Latino, Messapia) examined was considered to be an allele of a genetic locus for each EST-SSR marker. The PIC value indicates the usefulness of DNA markers for gene mapping, molecular breeding and germplasm evaluation. A PIC mean value of 0.41 (range: 0–0.82) for dinucleotide SSRs and of 0.24 (range: 0–0.78) for trinucleotide SSRs (Table 2) were observed.

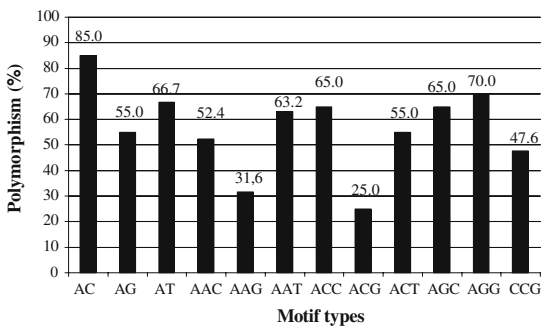

Fig. 1 Polymorphism of expressed sequence tag (EST)-derived simple sequence repeats (SSRs) based on different motif sequence types

Discussion

ESTs are currently the most widely sequenced nucleotide commodity from plant genomes in terms of number of sequences and total nucleotide count. During the past few years a great deal of attention has been directed towards discovering and characterizing the range of protein-coding genes existing within the genome of plant species with large genomes. *Arabidopsis thaliana* and *Oryza sativa*, model plant species with fully sequenced genomes, have the smallest known genomes (125 Mbp and 430 Mbp, respectively), while hexaploid wheat has agenome size of 17 Gb (Rudd 2003). The larger size of the wheat genome is a result of polyploidy and the presence of regions with repeat motifs, both of which make it difficult to sequence the complete genome. One possible method that could be used to investigate genome coding regions is cDNA sequencing, which may be considere d to be an alternative to the complete sequencing of the genome in those plants with large genomes. A considerable number of ESTs are currently available for several crop species (*Triticum aestivum*, *Hordeum vulgare*, *Glycine max*, *Zea mays*, *Lycopersicon esculentum*, etc.) (Kanteny et al. 2002).

The availability of ESTs in public databases provides the opportunity to identify SSRs and to develop molecular markers. Consequently, the large number of wheat EST-SSRs available in

public databases is an important research resource which can be used to analyse the functional portion of the genomes.

In the present study, 241 EST-SSR markers were analysed in 12 wheat genotypes, of which 16.6% did not amplify PCR products. This value is much lower than the rate (30.0%) obtained by Nicot et al. (2004) and by Holton et al. (2002), although a similar value (15%) was observed with genomic SSRs by Guyomarc'h et al. (2002). The lack of amplification observed in the present study may be due to the presence of introns in the genomic DNA.

Microsatellites are commonly considered to be codominant markers, and the polymorphism detected among genotypes is the result of variation in the number of tandem repeat units. However, the polymorphism detected in the present work was only in part due to differences in the number tandem repeat units, as several primers showed allelic differences as the presence/absence of bands. Allelic variation at some EST-SSR loci may be due to deletion and base substitution events in the intervening sequences between the primer sites and the SSRs. The variability within a flanking sequence is sufficient to be able to distinguish even between two duplicate loci in the same genome (Decroocq et al. 2003). This could be particularly useful in wheats for differentiating loci located on homoeologous chromosomes.

EST-SSRs generally generate high-quality amplification products (Eujayl et al. 2002; Nicot et al. 2004), suggesting that ESTs are more suited for specific primer pair design than genomic sequences. The lower quality obtained for genomic sequences could also be due to a large number of low-complexity repeated regions framing the microsatellites. In support of this possibility, Ramsay et al. (1999) showed that microsatellites are associated with cereal retrotransposon-like sequences. In the present study more than one-half of the primer pairs tested gave one or two discrete PCR bands. However, EST-SSRs that amplified multiple discrete PCR products were observed in both dinucleotide and trinucleotide motifs, with AG, AGG and ATT motif-based SSRs showing the lowest values (20.0, 30.0, 30.0%, respectively). Although the presence of non-speci fic amplification products are attributed to PCR conditions, such as annealing temperature and $MgCl_2$ concentration, multiple bands could also be due to the occurrence of more than one primer binding site (Fisher et al. 1998; Westman and Kresovich 1998). In some cases we observed differences between the expected and the observed size of the amplification product, with the PCR products being either smaller or larger than the expected products, as determined from EST sequences analysis. PCR products smaller than expected may have resulted from a small deletion within the sequence framed by the two primers or from non-specific annealing of the primers because of the low complexity of the flanking sequences, or they could be caused by duplications of EST sequences on the same or another chromosome (Nicot et al. 2004). High -molecular -eight bands (bands >500 bp) were relived in both dinucleotide and trinucleotide motifs. Such bands are due to the amplification of introns and may cause problems during genome analyses since fragments above 500 bp cannot be scored accurately for small differences in fragment size (Thiel et al. 2003).

A relatively low level of polymorphism has been reported in a number of different plant species when EST-SSR markers were used (Weising et al. 1996; Smulders et al. 1997; Bandopadhyay et al. 2004). The level of EST-SSR polymorphism found in wheat in the present study (56.8%) is similar to that reported by Nicot et al. (2004) (53.0%) and Yu et al. (2004) (51.0%) and comparable that found in a genomic SSR study (53.0%) (Röder et al. 1998). However, Eujayl et al. (2002) found EST-SSRs to be significantly less polymorphic than genomic SSR markers.

The polymorphism level among di- and trinucleotide EST-SSRs was found to be different. Markers based on dinucleotide repeats were more polymorphic (68.9%) than trinucleotide SSR markers (52.7%) in the 12 wheat genotypes tested. Similar results were reported by Nicot et al. (2004): in their experiments dinucleotide SSRs showed a much higher polymorphism level (83%) than both trinucleotide SSRs (46%) and other SSR classes. Various researchers have reported that the dinucleotide repeats are the most polymorphic of the genome markers among plants

Table 3 List of EST-SSR markers (http//wheat.pw.usda.gov) analysed on the 12 wheat genotypes used in this study

S. No	Accession no.	Repeat motif	Annealing temperature (°C)	Presence of PCR product[a]	Polymorphism[b]
1	TC87342	AC	55	+	+
2	CA723284	AC	55	+	+
3	CA723396	AC	55	+	+
4	TC73356	AC	55	+	+
5	TC100265	AC	55	+	+
6	TC82211	AC	55	+	+
7	CA694538	AC	55	+	–
8	TC73378	AC	55	+	+
9	BE430392	AC	55	+	+
10	CA612754	AC	55	+	+
11	CA646285	AC	55	+	+
12	BJ227727	AC	55	+	+
13	CA698471	AC	55	+	+
14	BJ296546	AC	50	+	+
15	CA596934	AC	50	+	+
16	BQ251631	AC	50	+	–
17	TC81688	AC	50	+	+
18	TC66497	AC	50	+	+
19	TC83420	AC	50	+	+
20	CA669680	AC	50	+	–
21	TC94553	AG	55	+	–
22	TC67482	AG	55	+	+
23	CA614288	AG	55	–	
24	TC101123	AG	55	+	–
25	TC82680	AG	55	–	
26	TC64267	AG	55	–	
27	CA741546	AG	55	+	+
28	CA653131	AG	55	–	
29	TC83596	AG	55	+	–
30	CA484639	AG	55	+	+
31	TC66874	AG	55	+	+
32	BQ282789	AG	55	+	+
33	BQ281574	AG	55	+	+
34	CA486384	AG	55	+	+
35	TC83694	AG	55	+	+
36	AL825137	AG	50	+	–
37	TC94800	AG	50	+	+
38	TC86475	AG	50	+	+
39	TC80973	AG	50	+	+
40	TC72728	AG	50	–	
41	CA632473	AT	55	+	+
42	BE419757	AT	55	+	+
43	BE401607	AT	55	+	+
44	BQ160961	AT	55	+	–
45	TC82246	AT	55	–	
46	CA626219	AT	55	+	+
47	CA610925	AT	55	+	–
48	BQ243113	AT	55	–	
49	CA502606	AT	55	+	+
50	TC100999	AT	55	+	+
51	CA681433	AT	55	–	
52	BJ301149	AT	55	+	+
53	CA697200	AT	55	+	+
54	TC77994	AT	55	+	+
55	CA613690	AT	55	+	+

Table 3 Continued

S. No	Accession no.	Repeat motif	Annealing temperature (°C)	Presence of PCR product[a]	Polymorphism[b]
56	CA621576	AT	55	+	+
57	CA680079	AT	55	+	+
58	TC92484	AT	50	+	–
59	CA627342	AT	50	+	+
60	TC65815	AT	50	+	+
61	BJ274672	AT	50	+	+
62	TC66737	AAC	55	+	–
63	CA715581	AAC	55	+	–
64	BQ237352	AAC	55	–	
65	CA709844	AAC	55	+	–
66	NP234579 |K02069.1| AAA34284.1	AAC	55	+	+
67	TC65969	AAC	55	–	
68	TC67050	AAC	55	+	+
69	TC84566	AAC	55	–	
70	TC74127	AAC	55	+	–
71	CA717120	AAC	55	+	+
72	CA681143	AAC	55	–	
73	TC71237	AAC	55	+	+
74	TC65956	AAC	55	+	+
75	TC85050	AAC	55	+	+
76	CA485812	AAC	50	+	+
77	BJ237020	AAC	50	+	+
78	TC66133	AAC	50	+	–
79	TC66370	AAC	50	+	+
80	CA717892	AAC	50	+	+
81	TC84551	AAC	50	+	+
82	BQ607256	AAC	50	+	+
83	CA737007	AAG	55	+	–
84	BJ313172	AAG	55	+	–
85	CA613743	AAG	55	+	–
86	TC85953	AAG	55	–	
87	TC94359	AAG	55	+	–
88	BF202588	AAG	55	+	–
89	TC68217	AAG	55	–	
90	TC63221	AAG	55	–	
91	CA715722	AAG	55	+	+
92	TC86610	AAG	50	+	+
93	CA605638	AAG	50	+	+
94	TC66918	AAG	50	+	+
95	TC92024	AAG	50	–	
96	BQ483099	AAG	50	+	–
97	BQ241171	AAG	50	–	
98	TC71017	AAG	50	+	–
99	TC88757	AAG	50	+	–
100	BG904656	AAG	50	+	+
101	CA620394	AAG	50	+	+
102	BG905242	AAT	55	–	
103	BJ239878	AAT	55	+	+
104	CA704477	AAT	55	+	+
105	BQ620347	AAT	55	+	+
106	BF483631	AAT	55	+	+
107	CA709708	AAT	55	+	–
108	CA655819	AAT	55	+	–

Table 3 Continued

S. No	Accession no.	Repeat motif	Annealing temperature (°C)	Presence of PCR product[a]	Polymorphism[b]
109	CA741577	AAT	55	+	+
110	BQ252012	AAT	55	+	+
111	TC100598	AAT	50	–	
112	TC89954	AAT	50	+	–
113	TC90707	AAT	50	+	+
114	TC90706	AAT	50	+	+
115	TC89955	AAT	50	+	+
116	TC64794	AAT	50	+	–
117	TC64796	AAT	50	+	+
118	CA593727	AAT	50	+	+
119	BQ620347	AAT	50	+	+
120	CA485678	AAT	50	+	+
121	CA732895	ACC	55	+	+
122	TC63424	ACC	55	+	+
123	CA486296	ACC	55	+	+
124	BE400563	ACC	55	+	–
125	TC94376	ACC	55	+	+
126	CA618377	ACC	55	+	+
127	TC100823	ACC	55	+	+
128	CA730301	ACC	55	+	–
129	CA618656	ACC	55	+	–
130	TC100719	ACC	55	+	–
131	CA690172	ACC	50	+	–
132	TC82866	ACC	50	+	+
133	TC87112	ACC	50	+	+
134	TC91785	ACC	50	+	+
135	BJ235341	ACC	50	+	+
136	CA732895	ACC	50	+	–
137	TC63424	ACC	50	+	+
138	CA718858	ACC	50	+	+
139	TC97410	ACC	50	+	+
140	CA701016	ACC	50	+	+
141	CA677802	ACG	55	+	–
142	TC71445	ACG	55	+	+
143	TC83809	ACG	55	+	–
144	TC90829	ACG	55	+	+
145	CA744169	ACG	55	–	
146	BQ608325	ACG	55	+	–
147	CA484813	ACG	55	+	–
148	CA695607	ACG	55	–	
149	CA605496	ACG	55	–	
150	CA596953	ACG	55	–	
151	TC67093	ACG	50	–	
152	TC78251	ACG	50	+	+
153	CA696865	ACG	50	–	
154	TC69937	ACG	50	+	+
155	TC83809	ACG	50	+	+
156	BF483389	ACG	50	–	
157	CA729823	ACG	50	–	
158	TC73262	ACG	50	–	
159	BG313615	ACG	50	+	+
160	CA650776	ACG	50	+	+
161	CA718646	ACT	55	+	+
162	TC98867	ACT	55	+	–
163	CA664300	ACT	55	+	+
164	CA661581	ACT	55	+	+

Table 3 Continued

S. No	Accession no.	Repeat motif	Annealing temperature (°C)	Presence of PCR product[a]	Polymorphism[b]
165	TC68633	ACT	55	+	+
166	CA732258	ACT	55	+	+
167	CA662775	ACT	55	+	–
168	CA704353	ACT	55	+	–
169	BQ282800	ACT	55	–	
170	TC77993	ACT	55	+	+
171	CA618403	ACT	50	+	+
172	CA693702	ACT	50	+	–
173	BQ162419	ACT	50	+	+
174	TC63295	ACT	50	+	+
175	TC64141	ACT	50	–	
176	TC79888	ACT	50	+	+
177	TC91851	ACT	50	+	+
178	TC65216	ACT	50	+	+
179	TC65170	ACT	50	+	–
180	TC84150	ACT	50	+	+
181	TC64360	AGC	55	+	+
182	TC71106	AGC	55	+	+
183	BJ318987	AGC	55	+	+
184	TC94880	AGC	55	+	–
185	TC67049	AGC	55	+	–
186	NP234852 |U50984.1| AAA96276.1	AGC	55	+	+
187	TC89813	AGC	55	–	
188	BE515440	AGC	55	+	–
189	BQ245951	AGC	55	+	+
190	CA501135	AGC	55	+	+
191	BQ171047	AGC	55	+	+
192	TC96817	AGC	55	+	+
193	CA728820	AGC	55	+	–
194	CA642505	AGC	55	+	+
195	CA728856	AGC	55	+	–
196	TC93568	AGC	55	–	
197	TC81607	AGC	55	+	+
198	TC91141	AGC	50	+	+
199	TC88840	AGC	50	+	+
200	CA648994	AGC	50	+	+
201	CA656642	AGG	50	+	+
202	AL829536	AGG	50	+	+
203	TC98928	AGG	50	–	
204	TC86249	AGG	50	+	+
205	BJ249131	AGG	50	+	–
206	BQ743416	AGG	50	+	+
207	BE438417	AGG	50	+	+
208	TC82726	AGG	50	+	+
209	TC68249	AGG	50	–	
210	CA499013	AGG	50	+	+
211	TC69043	AGG	50	+	+
212	CA598248	AGG	50	+	–
213	CA622856	AGG	50	+	–
214	TC91009	AGG	50	+	+
215	TC71366	AGG	50	+	+
216	TC75438	AGG	50	+	+
217	CA711628	AGG	50	+	+

Table 3 Continued

S. No	Accession no.	Repeat motif	Annealing temperature (°C)	Presence of PCR product[a]	Polymorphism[b]
218	TC89764	AGG	50	+	+
219	TC70636	AGG	50	+	+
220	TC71885	AGG	50	+	+
221	TC63543	CCG	50	+	+
222	TC98561	CCG	50	+	+
223	TC71130	CCG	50	+	+
224	TC91144	CCG	50	+	–
225	TC100243	CCG	50	–	
226	TC101394	CCG	50	–	
227	BJ271437	CCG	50	+	+
228	CA666539	CCG	50	+	+
229	CA666539	CCG	50	+	–
230	BJ225377	CCG	50	+	+
231	CA607284	CCG	50	+	–
232	CA731845	CCG	50	–	
233	BF483209	CCG	50	+	+
234	TC68120	CCG	50	+	+
235	TC70186	CCG	50	+	+
236	TC73864	CCG	50	+	+
237	TC85928	CCG	50	+	–
238	BE471206	CCG	50	–	
239	BJ278437	CCG	50	+	+
240	CA634443	CCG	50	–	
241	TC90852	CCG	50	+	+

[a] +, Indicates the presence of amplified bands; –, indicates the absence of amplified bands

[b] + and – indicate polymorphic and monomorphic markers, respectively

(Lagercrantz et al. 1993; Morgante and Olivieri 1993; Li et al. 2002; Baek et al. 2003), which is surprising since increasing or decreasing the number of dinucleotide motifs would result in the modificationof the reading frame of the gene. One possible explanation for this finding would be that dinucleotides follow steps of three motifs which do not modify a gene's reading frame (Nicot et al. 2004).

The analysis of SSR type sets reported here revealed different polymorphisms within dinucleotide and trinucleotide SSR types. A complete list of the EST-SSRs analysed in the present paper is provided in Table 3. The data reported here indicate the existence of a significant relationship between motif sequence type and polymorphism, with the mean PIC values indicating that a SSR based on the AG repeat motif is the least informative among the dinucleotide SSRs and that an AT motif primer set is the most informative. Interestingly, AT repeats seem to be typical of non-transcribed regions, whereas AG repeats were abundant in transcribed regions (Morgante et al. 2002). The low polymorphism detected among the 12 wheat genotypes using AG-based SSRs primers may suggest the presence of these microsatellites in classes of genes highly conserved in plants, while the high polymorphism detected byAT-based SSRs primers may suggest that non-transcribed regions of genes easily accumulate mutations. It would be interesting to investigate whether monomorphic markers among di- or trinucleotide EST-SSRs are more transferable to related species, such as rice or barley, than others. Although several studies (Weising et al. 1996; Smulders et al. 1997; Bandopadhyay et al. 2004; Röder et al. 1998; Nicot et al. 2004; Peng et al. 2005) have found different levels of polymorphism among di- and tri- or tetranucleotide EST-SSRs, there seems to be a lack of detailed analyses on the relationship between polymorphism and EST-SSR types to date. As a result of the continually increasing number of ESTs being depositedin public

databases, hundreds of new EST-SSR markers have been developed in recent years. EST-derived SSR markers constitute a novel source of markers that are physically associated with coding regions of the genome and which enhance the role of genetic markers in germplasm evaluation by enabling the variation in expressed genes to be assayed (Peng et al. 2005). Polymorphic EST-SSR markers directly sample variations in transcribed regions of the genome, and this may enhance their value in wheat genetic diversity assessment and marker-assisted selection. The results of the present investigation indicate that some SSR types are able to discriminate among cultivars much more than others, a finding which may help researchers choose the appropriate markers to be used in breeding programmes.

Acknowledgements This research project was supported by grants from Ministero delle Politiche Agricole, Alimentari e Forestali, project 'FRUMISIS'.

References

Anderson JA, Churchill GA, Sutrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36:181–186

Baek HJ, Beharav A, Nevo E (2003) Ecological-genomic diversity of microsatellites in wild barley *Hordeum spontaneum*, population in Jordan. Theor Appl Genet 106:397–410

Bandopadhyay R, Sharma S, Rustgi S, Singh R, Kumar A, Balyan HS, Gupta PK (2004) DNA polymorphic among 18 species of *Triticum-Aegilops* complex using wheat EST-SSRs. Plant Sci 166:349–356

Becker J, Heun M (1996) Barley microsatellites: allele variation and mapping. Plant Mol Biol 27:835–845

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

Decroocq V, Fave MG, Hagen L, Bordenave L (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor Appl Genet 106:912–922

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. Biotechniques 20:759–760

Eujayl I, Sorrells ME, Baum M, Wolters P, Powell W (2001) Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. Euphytica 119:39–43

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

Fisher PJ, Richardson TE, Gardener RC (1998) Characteristics of single- and multi-copy microsatellites in *Pinus radiate*. Theor Appl Genet 96:969–979

Gao L, Tang J, Li H, Jia J (2003) Analysis of microsatellites in major crops assessed by computational and experimental approaches. Mol Breed 12:245–261

Guyomarc'h H, Sourdille P, Charmet G, Edwards KJ, Bernard M (2002) Characterization of polymorphic microsatellite markers from *Aegilops tauschii* and transferability to the D-genome of bread wheat. Theor Appl Genet 104:1164–1172

Holton TA, Christopher JT, McClure L, Harker N, Henry RJ (2002) Identification and mapping of polymorphic SSR markers from expressed gene sequences of barley and wheat. Mol Breed 9:63–71

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

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res 21:1111–1115

Leigh F, Lea V, Law J, Wolters P, Powell W, Donini P (2003). Assessment of EST- and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. Euphytica 133:359–366

Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions, and mutational mechanisms (a review). Mol Ecol 11:2543–2565

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–183

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with non repetitive DNA in plant genomes. Nat Genet 30:194–200

Nicot N, Chiquet V, Gandon B, Amilhat L, Legeai F, Leroy P, Bernard M, Sourdille P (2004) Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). Theor Appl Genet 109:800–805

Ostrander EA, Pong PM, Rine J, Duyk G (1992) Construction of small-insert genomic DNA libraries highly enriched for microsatellite sequences. Proc Natl Acad Sci USA 89:3419–3423

Peng JH, Nora L, Capitan V (2005) Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. Funct Integra Genomics 5:80–96

Ramsay L, Macaulay M, Cardle L, Morgante M, Degli Ivanissevich S, Maestri E, Powell W, Waugh R (1999) Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plant J 17:415–425

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier NH, Laroy P, Ganal MV (1998). A microsatellite map of wheat. Genetics 149:2007–2023

Rudd S (2003) Expressed sequence tags: alternative or complement to whole genome sequences? Trends Plant Sci 8:321–329

Schlötterrer C, Tautz D (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Res 20:211–215

Sharp PJ, Kreis M, Shewry PR, Gale MD (1988) Location of β-amylases sequences in wheat and its relatives. Theor Appl Genet 75:286–290

Smulders MJM, Bredemeijer G, Rus-Kortekass W, Arens P, Vosman B (1997) Use of short microsatellites from data base sequences to generate polymorphism among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. Theor Appl Genet 97:264–272

Song QJ, Fickus EW, Cregan PB (2002) Characterization of trinucleotide SSR motifs in wheat. Theor Appl Genet 104:286–293

Thiel T, Michaelek 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

Tòth G, Gáspári Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res 10:967–981

Turpeinen T, Tenhola T, Manninen O, Nevo E, Nissila E (2001) Microsatellite diversity associated with ecology factors in *Hordeum spontaneum* populations in Israel. Mol Ecol 10:1577–1591

Weber JL (1990) Informativeness of human $(dC\text{-}dA)_n \times (dG\text{-}dT)_n$ polymorphisms. Genomics 7:524–530

Weising K, Fing RWM, Keeling RG, Atkinson RG, Gardner RC (1996) Characterization of microsatellites from *Actinidia chinensis*. Mol Breed 2:117–131

Westman AL, Kresovich S (1998) The potential of cross-taxa simple sequence repeat (SSR) amplification between *Arabidopsis thaliana* L. and crop brassicas. Theor Appl Genet 96:272–278

Yu JK, La Rota M, Kanteny RV, Sorrels ME (2004) EST derived SSR markers for comparative mapping in wheat and rice. Mol Gen Genomics 271:742–751

Zang LY, 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