Theor Appl Genet (2000) 100:584–592

M. Prasad · R.K. Varshney · J.K. Roy · H.S. Balyan
P.K. Gupta

The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat

Received: 15 May 1999 / Accepted: 27 July 1999

Abstract A set of 20 wheat microsatellite markers was used with 55 elite wheat genotypes to examine their utility (1) in detecting DNA polymorphism, (2) in the identifying genotypes and (3) in estimating genetic diversity among wheat genotypes. The 55 elite genotypes of wheat used in this study originated in 29 countries representing six continents. A total of 155 alleles were detected at 21 loci using the above microsatellite primer pairs (only 1 primer amplified 2 loci; all other primers amplified 1 locus each). Of the 20 primers amplifying 21 loci, 17 primers and their corresponding 18 loci were assigned to 13 different chromosomes (6 chromosomes of the A genome, 5 chromosomes of the B genome and 2 chromosomes of the D genome). The number of alleles per locus ranged from 1 to 13, with an average of 7.4 alleles per locus. The values of average polymorphic information content (PIC) and the marker index (MI) for these markers were estimated to be 0.71 and 0.70, respectively. The $(GT)_n$ microsatellites were found to be the most polymorphic. The genetic similarity (GS) coefficient for all possible 1485 pairs of genotypes ranged from 0.05 to 0.88 with an average of 0.23. The dendrogram, prepared on the basis of similarity matrix using the UPGMA algorithm, delineated the above genotypes into two major clusters (I and II), each with two subclusters (Ia, Ib and IIa, IIb). One of these subclusters (Ib) consisted of a solitary genotype (E3111) from Portugal, so that it was unique and diverse with respect to all other genotypes belonging to cluster I and placed in subcluster Ia. Using a set of only 12 primer pairs, we were able to distinguish a maximum of 48 of the above 55 wheat genotypes. The results demonstrate the utility of microsatellite markers for detecting polymorphism leading to genotype identification and for estimating genetic diversity.

Key words Microsatellite markers · Wheat · Genetic diversity · Genotype identification

Communicated by J. Beckmann

M. Prasad · R.K. Varshney · J.K. Roy · H.S. Balyan
P.K. Gupta (✉)
Molecular Biology Laboratory,
Department of Agricultural Botany, Ch. Charan Singh University,
Meerut-250004, India
Fax: +91–121—767018
e-mail: pkgupta@ndf.vsnl.net.in

Introduction

The availability of genetic variability in elite wheat material is a pre-requisite for any breeding programme aimed towards the improvement of wheat productivity. Wheat breeding through hybridization also requires the selection of diverse genotypes, irrespective of whether the end product is a pure line or a hybrid variety. In this respect, efforts have also been made to predict the prospects of developing superior genotypes from a cross by the measurement of genetic similarity (GS) or genetic distance (GD) between the parents, since the latter can be used as an estimation of expected genetic variances (σ^2_g) in different sets of segregating progenies derived, from different crosses (Burkhamer et al. 1998; Bohn et al. 1999). For this purpose, a number of diversity studies in the past have been undertaken in wheat using data on a variety of morphological traits that were subjected to D^2 analysis and clustering (Lee and Kaltsikes 1973). However, there are inherent problems with the use of data on morphological traits, the latter being limited in number and greatly influenced by the environment and by genotype × environment interactions. Molecular markers, therefore, provide a satisfactory alternative because they are almost unlimited in number and are not influenced by the environment. Despite this, studies on variability and diversity in wheat using molecular markers have been relatively few. This can be attributed to the detection of low levels of variability leading to the use of a limited number of polymorphic markers belonging to classes such as (1) restriction fragment length polymorphisms/sequence-tagged sites (RFLPs/STSs) for low-copy sequences (Vaccino et al. 1993; Chen et al. 1994; Talbert et al. 1994), (2) random amplified polymorphic

DNA (RAPDs) for random sequences (He et al. 1992; Dweikat et al. 1993) and (3) PAGE for such proteins as gliadins (Cox et al. 1985). In fact, among DNA markers, RFLPs and RAPDs, the two most commonly used markers in wheat, detect only low levels of polymorphism (Penner et al. 1995; Röder et al. 1998; Paull et al. 1998). In comparison, microsatellites or simple sequence repeats (SSRs) are more abundant, ubiquitous in presence, hypervariable in nature and have high polymorphic information content (PIC) (Gupta et al. 1996). Due to these properties, the microsatellites have recently been used to study genetic variability based on DNA polymorphism in a number of crop plants including wheat, soybean, maize, rice, sorghum, barley, etc. It has also been shown that the use of a limited number of microsatellites is adequate to discriminate even the most closely related wheat and barley genotypes (Plaschke et al. 1995; Russell et al. 1997; Struss and Plieske 1998).

In this communication, we report the results of a study involving the screening of 55 genotypes using 20 microsatellite primers. The study was undertaken with the following objectives: (1) to study the potential of microsatellite markers in general and specific SSRs in particular for detection of polymorphism and for genotype identification and (2) to assess the level of microsatellite-based genetic diversity among 55 elite wheat genotypes that were potentially useful in wheat breeding programmes.

Materials and methods

Seed material

Seed of 55 elite genotypes of wheat originating in 29 countries representing six continents was procured from the Directorate of Wheat Research, Karnal, India. The details of the genotypes, their pedigrees and the countries of their origin are given in Table 1.

DNA isolation and microsatellite primers

DNA was extracted from leaves of each of the 55 genotypes grown in the field using a modified CTAB method (Saghai Maroof et al. 1984). Twenty microsatellite primers were used, which were randomly chosen from a set of 232 primers, made available to us as a member of the Wheat Microsatellite Consortium (WMC) under an international collaborative project. The microsatellite primer pairs were designed using sequences of clones containing microsatellites. The clones bearing microsatellites belonged to a genomic library enriched for microsatellites (Edwards et al. 1996) and were sequenced by members of the WMC.

Polymerase chain reaction (PCR)

DNA amplifications were carried out in 25-µl reaction mixtures, each containing 100 ng template DNA, 2 µ*M* microsatellite primers, 200 µ*M* each of the dNTPs, 2.5 m*M* $MgCl_2$, 1 × PCR buffer and 2 U Stoffel fragment (Perkin Elmer), using the following PCR profile in a Perkin Elmer DNA Thermal Cycler: initial denaturation at 95°C for 5 min; 40 cycles at 95°C for 1 min, 51°C/61°C for 1 min, 72°C for 1 min, with a ramp at the rate of 0.5°C; a final extension at 72°C for 5 min. The amplification products were resolved on 10% polyacrylamide denaturing gels (PAGE) followed by silver staining (Tegelstrom 1992). Fragment sizes were calculated using the computer programme SEQAID II (Fragment Sizer) by comparing with fragments of 100-bp ladder marker DNA.

Chromosomal localization

PCR amplification using the above conditions was also carried out with target DNA samples from each of the 21 nulli – tetrasomic lines using the above 20 microsatellite primer pairs. This allowed chromosome localization of 18 out of the 21 loci sampled through the use of above microsatellite primer pairs. The different loci identified using individual primer pairs and assigned to specific chromosomes as above were given designations in accordance with the Rules of Nomenclature for DNA markers (McIntosh et al. 1998), as approved earlier at the 7th International Wheat Genetics Symposium, held at Cambridge (Hart and Gale 1988).

Statistical analysis

The fragment(s) sizes in 'Chinese Spring' were taken as standard, and the size differences of the fragments in other genotypes were considered to be the result of alterations in the repeat number of the simple sequences at the corresponding site(s). Allelic polymorphic information content (PIC) was calculated using the following formula.

$$PIC = 1 - \sum (P_i)^2$$

where P_i is the proportion of the population carrying i^{th} allele, calculated for each microsatellite locus (Botstein et al. 1980).

The marker index (MI) was calculated using the following formula (Powell et al. 1996). MI = Average polymorphic information content (PIC) × Proportion of polymorphic bands × Average number of loci per assay unit

For the purpose of assessing genetic diversity leading to the preparation of a dendrogram, gels were scored in binary format, with the presence of a band scored as unity and its absence scored as zero. The binary data were used to compute pair-wise similarity coefficients (Jaccard 1908), and the similarity matrix thus obtained was subjected to cluster analysis using the UPGMA (unweighted pair-group method with arithmetic average) algorithm on NTSYS-PC, version 1.70 (Rohlf 1992).

Results

The results of PCR amplification of a number of microsatellite loci in 55 wheat genotypes using 20 wheat microsatellite primer pairs are summarized in Table 2. The amplification profiles obtained with 20 individual microsatellite primer pairs could be grouped into three categories: (1) those showing one prominent fragment (Fig. 1a), (2) those showing two comigrating prominent fragments with equal spacing (Fig. 1b), and (3) those showing three fragments including the two comigrating prominent fragments with equal spacing (Fig. 1c). While scoring the above amplification profiles, we took due care to exclude the possibility of stuttering effect, particularly in the case of two comigrating prominent fragments. In each case two comigrating prominent fragments, the fragments were localized on the same chromosome (see later) and no recombinants were available, so that they were considered to belong to one locus, even though we do not have a suitable explanation for the comigration of the two fragments. Using the above criteria, we detected a total of 155 alleles at 21 loci.

Table 1 Fifty-five wheat genotypes, their pedigrees (as far as known) and countries of origin

S. no.	Genotype	Pedigree	Country of origin
1	E3249	Australian-144	Sweden
2	E3111	Tremez molle P.I. 56258-I (W/29)	Portugal
3	E3086	Exchange CI, 12635 (W-193)	USA
4	E3860	Khapli, C.I. 4013	Guatemala
5	E3896	Atacatzo no. 1, Kanred × Frondoso × Purcan 3138, P.I. 23185	Equador
6	E4813	P.I. No. 238396, 261.E	Kenya
7	E3877	Frontana-Kenya 58, 194–49 P.I. No. 206048	Brazil
8	E3901	Inca Rojo no. 1, P.I. 11890 × 41–6-7–3, 3332–3, P.I. 223189	Equador
9	E4229	Hibrido Lebrija	Spain
10	E4328	P.I. 94650, Dickson's No. 267	Czechoslovakia
11	E581	Reliance (Ex. USA) × Njoro hybrids-73.D.2.I.1.(L)	Kenya
12	E4205	356.A.2.A	E. Africa
13	E3275	Timstein	The Netherlands
14	E549	Redman CAN 3633	Canada
15	E3083	No. 43, P.I. 159106 (W-190)	S. Africa
16	E336	Khapli (*dicoccum* sp.)	Canada
17	E585	117.A	Kenya
18	E661	Barleta Benvenuto	Argentina
19	E288	Florence × Aurora A	Poland
20	E319	(Steinwedel × timophcevi × Pusa 4)^2	Australia
21	E3876	Frontana, C.I. 12470	Brazil
22	E271	Ostkachtopicka	Poland
23	E146	130 B.6.B	Kenya
24	E149	117.1.5.F(L)	Kenya
25	E680	Inca Benvenuto	Argentina
26	E677	General urquiza	Argentina
27	E784	No. 618 Sementes Dc Trigo var. Kenya 155 No.2	Brazil
28	E780	No. 618 Sementes Dc Trigo var. Florestana	Brazil
29	E1003	Pajbjerg Kinge II	Cyprus
30	E965	Rieti 11	Yugoslavia
31	E2336	Marfed	USA
32	E1000	B. × I.P.I.	Cyprus
33	E2724	N-Y.51	Israel
34	E2055	Ostkaskomoroska	Poland
35	E2161	Mentana	USA
36	E3066	Kentana 52 (W-172)	Mexico
37	E2990	1117/1, *T. vulgare* var. Ferruginew (W- 88) / (W.R.80)	Turkey
38	E2401	Rushmore	USA
39	E3068	Supremo 211 (W-174)	Mexico
40	E3070	Yaqui 50 (W-176)	Mexico
41	E3859	Vernal, C.I. 3686	Guatemala
42	E2602	R.16	Bologna
43	E3414	P.I. 194044. IRN. No. 330	Ethiopia
44	E3492	Tha × K338 Ac. A-5–501–3-1 ND.4. P.I. No. 13000	USA
45	E3387	Lignee 53 AB-3	Lebanon
46	E3389	Lignee 53 AB-5	Lebanon
47	E3280	Tammi	Finland
48	E3840	Sonop,P.I. No. 227060	Columbia
49	E3253	Australian-149	Sweden
50	E3250	Australian-145	Sweden
51	E3715	Llcafen, P.I. No. 231307	Chile
52	E3264	Fylgia, P.I. No. 174636	France
53	E3274	Tonko	The Netherlands
54	E3839	Kiein cometa × New thatch-mentana, II-1879–76–4b-2b-1b	Columbia
55	E3273	Vanhock	The Netherlands

Using nullisomic-tetrasomic lines we were able to assign 18 of the 21 loci detected as above to the following 13 chromosomes: 1A, 1B, 2A, 2B, 2D, 3A, 4B, 5A, 6A, 6B, 6D, 7A and 7B. This information was used for assigning appropriate symbols to the loci, which are also listed in Table 2. Each of the 20 primer pairs used in the present study amplified only 1 locus, except for primer pair WMC256, which amplified 2 loci located on 2 different homoeologous chromosomes (6A, 6D). The amplified products were available in all 55 genotypes with 18 of the 20 primer pairs. For the 2 remaining primer pairs, no amplified products were available in 2 of the 55 genotypes involving 3 loci. These primer pairs included WMC216, which gave no amplified products in 7 genotypes, and WMC267, with no amplified products in 14 genotypes. The cases with no amplified products were scored as null-alleles. Allelic variants were also observed at each locus except the locus *Xwmc256–6D* (see Table 2). The number of alleles per locus ranged from 1 to 13, with an average of 7.4 alleles per locus. The maxi-

Table 2 Details of microsatellite primers used for the PCR amplification of alleles at 21 loci and the values of polymorphic information content (PIC)

S. no.	Primer designation	Locus designation	Number of alleles	PIC
1	WMC24	*Xwmc24–2A*	12	0.90
2	WMC25	*Xwmc25–2D*	11	0.77
3	WMC35	*Xwmc35–4B*	5	0.75
4	WMC44	*Xwmc44–1B*	13	0.86
5	WMC47	*Xwmc47–7A*	4	0.73
6	WMC76	*Xwmc76–6B*	7	0.73
7	WMC83	*Xwmc83–7A*	9	0.85
8	WMC120	*Xwmc120–1A*	2	0.46
9	WMC149	*Xwmc149–2B*	9	0.77
10	WMC167	*Xwmc167–2D*	11	0.83
11	WMC169	*Xwmc169–3A*	9	0.79
12	WMC170	*Xwmc170–2A*	10	0.80
13	WMC177[a]	–	10	0.85
14	WMC216	*Xwmc216–7B*	13	0.87
15	WMC221[a]	–	2	0.30
16	WMC243[a]	–	3	0.64
17	WMC245	*Xwmc245–2D*	5	0.74
18	WMC254	*Xwmc254–4B*	12	0.85
19	WMC256	*Xwmc256–6A*	4	0.21
		Xwmc256–6D	1	–
20	WMC267	*Xwmc267–5A*	3	0.65

[a] Chromosomal assignment was not possible and hence locus designation is not given

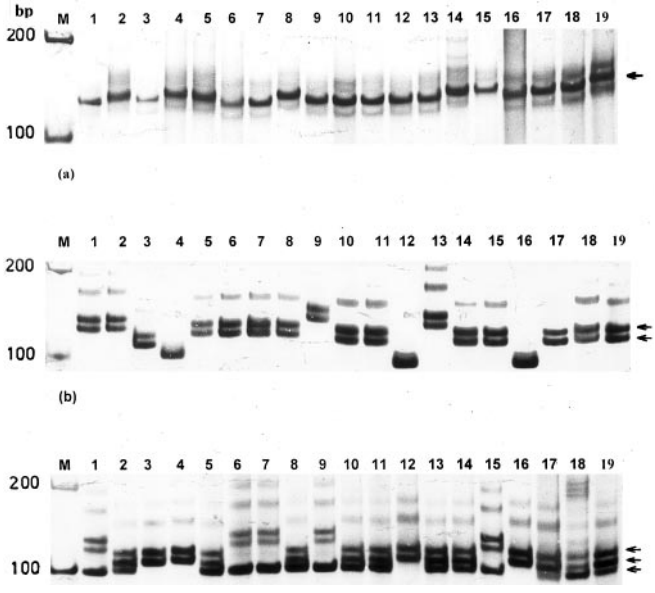

Fig. 1a–c Representative sample of amplification profiles in 19 wheat genotypes obtained using different microsatellite primer pairs: **a** showing one prominent fragment (product size range: 129—141 bp) amplified using WMC47 (*lanes 1 to 19* represent wheat genotypes E3249, E319, E3876, E271, E146, E149, E680, E677, E784, E780, E1003, E965, E2336, E1000, E2724, E2055, E2161, E3066 and E2990), **b** showing two comigrating prominent fragments (product size range: 106—150 bp and 98—137 bp) amplified using WMC24 (*lanes 1 to 19* represent wheat genotypes E3249, E3111, E3086, E3860, E3896, E4813, E3877, E3901, E4229, E4328, E581, E4205, E3275, E549, E3083, E336, E585, E661 and E288), **c** showing two comigrating prominent fragments (size range: 118—128 bp and 108—118 bp) and an additional fragment (size 100 bp) amplified using WMC256 (the order of wheat genotypes in *lanes 1–19* is same as in **Fig. 1a**). The *arrows* indicate the prominent fragments in each case

mum number of alleles was detected at *Xwmc44–1B* belonging to $(GT)_n$ and *Xwmc216–7B* belonging to $(CA)_n$ (Table 3). The size of alleles at *Xwmc44–1B* ranged from 200 to 312 bp, and those at *Xwmc216–7B* ranged from 142 to 220 bp. A similar variation was also present at other loci.

The PIC were also estimated, and it can be seen from the data of Table 2 that the highest values of 0.90 and 0.87 were recorded for WMC24 and WMC216, respectively. The MI value over all 20 microsatellite markers was 0.70. In 3 of the 55 genotypes, involving two of the 21 loci, 2 alleles at the same locus could be observed in the same genotype, suggesting heterozygosity. These 2 loci were *Xwmc44–1B* (genotypes E585 and E661) and *Xwmc254–4B* (genotype E4205).

Of the 20 microsatellite primer pairs used, information (supplied by WMC members) was also available on the repeat sequences of 13 of them (Table 3). In these cases, efforts were also made to estimate the number of repeats at individual microsatellite loci. For instance, at locus *Xwmc44–1B*, where the highest number of alleles was detected, the number of repeats was estimated to range from 3 to 59. This estimate is based on the information that in 'Chinese Spring', the product size is 264 bp with 35 GT repeats, suggesting that in the amplified product there is a sequence of 194 bp flanking the microsatellite $(GT)_n$, which is assumed to remain constant in different genotypes.

The genetic similarity (GS) coefficients for all the possible 1485 pairs of genotypes ranged from 0.05 to 0.88 and averaged 0.23. The dendrogram prepared through cluster analysis is shown in Fig. 2, suggesting a high level of diversity among the 55 genotypes. The genotypes could be grouped into two major clusters, cluster I with 23 genotypes and cluster II with 32 genotypes. Cluster I is further subdivided into two subclusters, sub-

Table 3 Details of microsatellite primer sequences, repeat motif and expected product size (*NA* not available)

S. no.	Primer designation	Primer sequence (5′ → 3′)	Microsatellite	Annealing temperature (°C)	Expected product size (bp) in Chinese Spring	Reference (personal communication)
1	WMC24	GTGAGCAATTTTGATTATACTG TACCCTGATGCTGTAATATGTG	$(GT)_{28}$	51	152	R.C. Wang (USA)
2	WMC25	TCTGGCCAGGATCAATATTACT TAAGATACATAGATCCAACACC	$(GT)_{26}$	51	166	R.C. Wang (USA)
3	WMC35	GCAGAGAGGCACAACTAGCGAG AACCTTAAGGGTCTGCCGGAAC	$(GT)_{25}$	61	265	R.C. Wang (USA)
4	WMC44	GGTCTTCTGGGCTTTGATCCTG TGTTGCTAGGGACCCGTAGTGG	$(GT)_{35}$	61	242	J. Dubcovsky (USA)
5	WMC47	GAAACAGGGTTAACCATGCCAA ATGGTGCTGCCAACAACATACA	$(CA)_8$	61	141	J. Dubcovsky (USA)
6	WMC120	GGAGATGAGAAGGGGGTCAGGA CCAGGAGACCAGGTTGCAGAAG	$(CA)_8$–$(GA)_{14}$–$(GT)_8$	61	160	G. Flore (France)
7	WMC149	ACAGACTTGGTTGGTGCCGAGC ATGGGCGGGGGTGTAGAGTTTG	$(CT)_8$–$(CT)_{24}$	61	230	S.G. Rogers (USA)
8	WMC167	AGTGGTAATGAGGTGAAAGAAG TCGGTCGTATATGCATGTAAAG	NA	51	185	M. Bernard (France)
9	WMC169	TACCCGAATCTGGAAAATCAAT TGGAAGCTTGCTAACTTTGGAG	NA	61	167	M. Bernard (France)
10	WMC170	ACATCCACGTTTATGTTGTTGC TTGGTTGCTCAACGTTTACTTC	NA	61	230	M. Bernard (France)
11	WMC177	AGGGCTCTCTTTAATTCTTGCT GGTCTATCGTAATCCACCTGTA	NA	51	184	M. Bernard (France)
12	WMC216	ACGTATCCAGACACTGTGGTAA TAATGGTGGATCCATGATAGCC	$(CA)_{22}$	51	123	P. Benoist (France)
13	WMC221	ACGATAATGCAGCGGGGAAT GCTGGGATCAAGGGATCAAT	$(CT)_9$–$(CA)_{16}$	61	209	P. Benoist (France)
14	WMC243	CGTCATTTCCTCAAACACACCT ACCGGCAGATGTTGACAATAGT	$(CA)_{15}$	61	177	P. Sharp (Australia)
15	WMC245	GCTCAGATCATCCACCAACTTC AGATGCTCTGGGAGAGTCCTTA	$(CA)_{10}$–$(CA)_{10}$–$(CA)_{13}$–$(CT)_8$	61	150	P. Sharp (Australia)
16	WMC254	AGTAATCTGGTCCTCTCTTCTTCT AGGTAATCTCCGAGTGCACTTCAT	$(AC)_{28}$–$(AC)_{27}$	51	193	P.K. Gupta (India)
17	WMC256	CCAAATCTTCGAACAAGAACCC ACCGATCGATGGTGTATACTGA	$(CA)_{12}$	61	117	P.K. Gupta (India)
18	WMC267	TCTTCACCCCATAATTGGAGAAGCCT TGCTTATTCTGCGCACTGGATGCCTA	NA	61	216	B. Keller (Switzerland)
19	WMC76[a]	–	NA	51	256	P. Devaux (France)
20	WMC83[a]	–	NA	61	160	P. Devaux (France)

[a] Permission to publish primer sequence was not obtained

cluster Ia containing 22 genotypes and subcluster Ib containing a solitary genotype, E3111. Thus, genotype E3111 from Portugal is unique and diverse with respect to all other genotypes. The remaining 32 genotypes belonging to cluster II were similarly grouped into two subclusters, subcluster IIa containing 14 genotypes and subcluster IIb containing 18 genotypes.

Discussion

DNA polymorphism and genotype identification

Microsatellite primer pairs are locus-specific and are, therefore, generally considered to be single-locus markers as against other molecular markers like RFLP probes and RAPD primers, which are multilocus in nature. Despite this general single-locus nature of microsatellites, in the present study 21 microsatellite loci were identified using 20 primer pairs: each of 19 primer pairs detected only 1 locus and only 1 primer pair, WMC256, detected 2 loci. Since microsatellite primers are locus specific, only 1 specific locus was expected to be amplified by each primer, and it was unexpected that this 1 primer amplified 2 loci. In two earlier studies also, more than 1 locus per microsatellite primer pair was detected and/or mapped in bread wheat (Röder et al. 1998; Stephenson et al. 1998). The microsatellite loci are also multiallelic (1–13 alleles per locus with a mean of 7.4 alleles/locus in the present study) and the alleles co-dominant, thus suggesting their relative superiority in detecting DNA polymorphism over some other markers (e.g. SNPs, single nucleotide polymorphisms) which are biallelic and dominant. The co-dominant nature of microsatellite primers is useful, as it also enables the detection of heterozygosity at specific loci. Another advantage of wheat

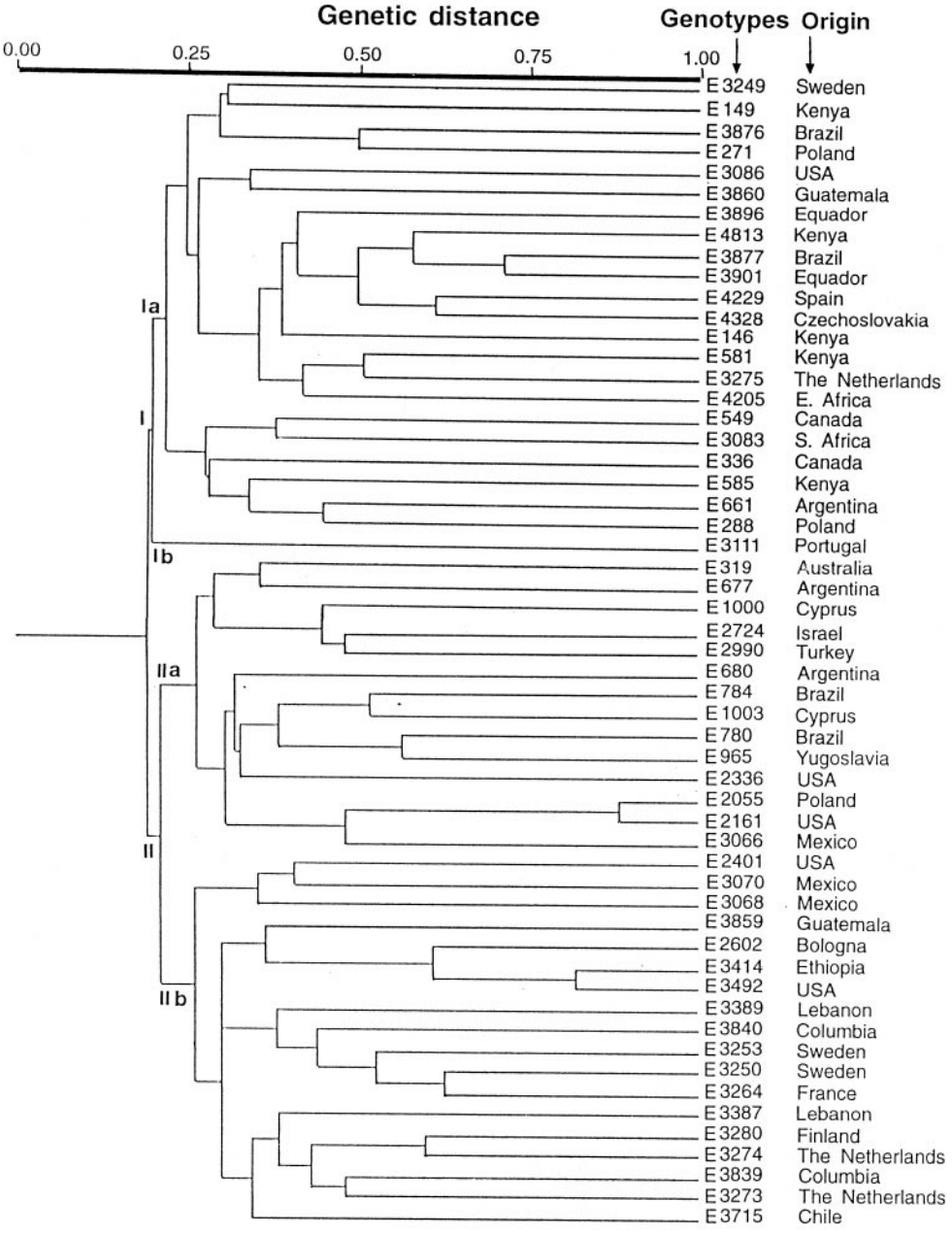

Fig. 2 Dendrogram of 55 wheat genotypes based on data on allelic profiles generated using 20 microsatellite primer pairs. *I* and *II* represent clusters, and *Ia*, *Ib*, *IIa*, *IIb* represent subclusters (for details, see Results)

microsatellites over other markers is their specificity, since microsatellite homoeoloci on 3 chromosomes of the same homoeologous group, if present, are generally not amplified with the same primer pair. In contrast to this, the same RFLP probe would generally identify three homoeoloci as RFLPs (Devos et al. 1992, 1993; Xie et al. 1993).

A comparison of the results obtained in the present study with those published earlier indicates that the average number of alleles per locus recorded during the present study was relatively higher than those earlier reported for several self-pollinated and/or annual crops including wheat, with estimates of 3.8, 4.6 and 6.2; tomato, with estimates of 1.5 and 3.1; sorghum, with an estimate of 2.3; cucumber and melons with estimates of 2.6 and 2.9, respectively and watermelons with an estimate of 2.0 alleles per locus (for references see Hokanson et al. 1998). Only in a recent study on barley, was a higher value of 8.6 alleles per locus reported (Struss and Plieske 1998). This high number of alleles per locus certainly contributed to the usefulness of these markers, although there seems to be no direct correlation between the number of alleles at a locus and the PIC value.

Keeping in mind their relative superiority in detecting DNA polymorphism as mentioned above, the utility of 20 specific microsatellite markers was also assessed for their power to uniquely discriminate the 55 elite wheat genotypes included in the present study. The results showed that a single microsatellite marker, WMC177, was sufficient to discriminate 16 genotypes; 2 microsat-

ellite markers (WMC177 and WMC169) discriminated 29 genotypes, 3 microsatellite markers (WMC177, WMC169 and WMC167) discriminated 36 genotypes and 4 microsatellite markers (WMC177, WMC169, WMC167 and WMC25) discriminated 41 genotypes. A further addition of as many as 8 more markers (WMC24, WMC44, WMC47, WMC83, WMC170, WMC216, WMC221 and WMC254), thus making a set of 12 microsatellite markers, also failed to discriminate all 55 genotypes; they distinguished only 48 genotypes and could not discriminate the remaining 7 genotypes (E2055, E2161, E2602, E3253, E3273, E3387 and E3414). However, supplementing the 12 markers with more markers did not improve the discriminating ability of the set, so that even 20 markers only discriminated the same 48 genotypes that could be discriminated by a set of 12 markers. In two earlier studies, a set of 4 microsatellite markers each discriminated all 24 genotypes used in barley and all 16 genotypes used in tomato (Russell et al. 1997; Bredemeijer 1998). Thus, the discriminating ability of microsatellite markers in these crops has been proven, and these markers can be used for a variety of crops as and when microsatellite markers specific for individual crops become available.

Since, in the present study, only 18 of the 21 loci that were assigned to specific chromosomes were restricted to only 13 of the 21 chromosomes and since only 48 of the 55 genotypes could be discriminated, the present set of 20 microsatellite primers, though useful for detecting polymorphism, is not entirely adequate for DNA fingerprinting of a large number of wheat genotypes. However, in another recent report, a set of microsatellites was prepared that could be used to discriminate each genotype in a set of more than 100 wheat genotypes (Manifesto et al. 1999). Therefore, one would have to select microsatellite primer pairs more judiciously to prepare a discriminatory set. Such a discriminatory set should also ensure the uniform distribution of the microsatellite primers of this set across the three genomes of bread wheat since in several studies, including our own, microsatellites have been shown to be more frequent in the A and B genomes, than in the D genome (Röder et al. 1998; Stephenson et al. 1998).

During the present study, it was also possible to assess the relative superiority of different SSRs in detecting DNA polymorphism. On the basis of the sequence data available for 13 of the 20 microsatellites (Table 3), for which the primers were used in the present study, these 13 microsatellites were classified into 8 simple-perfect, 2 simple-imperfect and 3 compound microsatellites. In our study, relative to other SSRs, more alleles (average of 10.3 alleles) were detected at the $(GT)_n$ loci (4 loci, 1 each identified by each of the 4 microsatellite primer pairs), followed by $(CA)_n$. Röder et al. (1995) also reported $(GT)_n$ repeats to be more polymorphic than other simple repeats such as $(GA)_n$ in wheat. However, recently, in barley more alleles were detected for $(GA)_n$ repeats than for $(GT)_n$ repeats (Struss and Plieske 1998).

Diversity analysis

We also used the data on microsatellite loci and their corresponding alleles to calculate the PIC and MI values in order to examine the extent of information on diversity that these markers can provide and to compare these results with those published earlier. The PIC values in the present study ranged from 0.21 to 0.90, with an average of 0.71; in two earlier studies on wheat, the PIC values ranged from 0.23 to 0.79 (Röder et al. 1995) and from 0.29 to 0.79 (Plaschke et al. 1995). Thus, the highest value (0.90) for PIC in our study is greater than the highest value (0.79) for PIC in two earlier studies on wheat (Röder et al. 1995; Plaschke et al. 1995). In another recent study, the PIC mean value (0.30) for SSRs was much lower than the mean value (0.71) observed in the present study (Bohn et al. 1999). The MI value was also calculated for the markers used in the present study and was higher (0.70) than the MI value (0.21) reported in a recent study on wheat using SSRs (Bohn et al. 1999). A higher number of alleles per locus and the higher values of PIC and MI for the markers used in the present study may be attributed to the use of more informative markers, and a bigger sample with relatively more diverse genotypes. Similarly, the low values of PIC and MI obtained in other earlier studies on wheat may be attributed either to a smaller population (18 genotypes – Röder et al. 1995; 11 genotypes – Bohn et al. 1999), to a set of closely related genotypes (Plaschke et al. 1995; Bohn et al. 1999) or to a set of fewer microsatellite primer pairs (15 markers – Röder et al. 1995). A drop in average PIC value from 0.63 to 0.54 on the omission of 6 synthetic wheat lines (out of a total of 18 genotypes) representing diverse material was actually reported in an earlier study where 15 microsatellite markers were used (Röder et al. 1995).

The information on microsatellite markers collected in the present study was also used to analyse genetic diversity among 55 wheat genotypes through cluster analysis. The estimates of a genetic similarity (GS) coefficient between pairs of genotypes ranged from 0.05 to 0.88. The average value of GS was as low as 0.23, suggesting that the 55 genotypes used in the present study were diverse. This GS value of 0.23 can be compared with those reported in three earlier studies, where SSR-based GS coefficient values of 0.31 (Plaschke et al. 1995) and 0.57 (Bohn et al. 1999) and a STS-based GS coefficient value of 0.81 (Chen et al. 1994) were reported. In these different studies on genetic diversity in bread wheat undertaken using a variety of molecular markers, the variation in GS coefficient values may be attributed either to the differences in number of genotypes and the probes/primers used (e.g. 119 RFLP probes were used by Paull et al. 1998) or to the relative superiority of microsatellites to detect DNA polymorphism. An unusually low value of RFLP-based GS (0.18) reported by Paull et al. (1998) is certainly due to the larger sample of 124 diverse genotypes and bigger set of 119 RFLP probes used in this study. In the present study, the highest GS value of 0.88

was observed between genotypes E2055 and E2161, suggesting that these 2 genotypes are closely related. Genotype E2055 originated from Poland, while the other genotype E2161 originated from USA. On the basis of available information on pedigree (data on pedigree was not adequate to work out coefficients of coancestry, or *f*), these 2 genotypes share no immediate common parents, but a high GS value suggests that they may be ancestrally related in the remote past. The pair with genotypes E336 (originating from Canada) and E965 (originating from Yugoslavia) had the lowest GS value (0.05); therefore, they represent members of a single pair of the most diverse genotypes among all the pair-wise combinations of the wheat genotypes used in the present study.

The cluster analysis enabled the grouping of all the genotypes used in the present study into two major clusters, I and II (Fig. 2). The 23 genotypes in cluster I were further grouped into subcluster Ia containing 22 genotypes originating from four different continents and subcluster Ib containing solitary genotype E3111 from Europe (Portugal), which is the most unique of all the genotypes studied. The remaining 32 genotypes were grouped into subclusters IIa and IIb with 14 and 18 genotypes, respectively, originating from five continents each (four of these continents were in common). This suggests that while genotypes from a number of continents may fall into the same subcluster, the genotypes from the same continent may also sometimes belong to two subclusters. From a closer look at the grouping pattern of the genotypes it was apparent that genotypes originating from a particular country were often grouped in the same subcluster such as Ia, IIa and IIb (Fig. 2). This suggests that genotypes developed in a particular country sometimes have a narrow genetic base, although the movement and use of germplasm over the countries/continents is a common feature for most wheat breeding programmes. Therefore, it appears that there is scope for improvement in national breeding programmes by exploiting more diverse genotypes, which could be identified following the approach used in the present study.

In a recent study an assessment was made to determine whether genetic distances calculated using molecular markers, as above, can be used for predicting the level of genetic variance(σ^2_g) among the progenies that would be derived from the crosses made between diverse genotypes (Bohn et al. 1999). It was shown that on the basis of molecular markers the prospects of predicting σ^2_g for each of seven agronomic traits examined was low, unless the markers linked to a trait were selected for calculating trait specific GS between parents. A similar inference had been drawn earlier by Hayes et al. (1997) while conducting amplified fragment length polymorphism (AFLP) analysis in barley. In our opinion, such an analysis may be necessary if a specific crossing programme is aimed towards the improvement of specific traits. However, if the breeding programme is aimed towards a general improvement of the crop involving a number of complex traits including yield, the genetic similarity (GS) estimated on the basis of molecular markers (SSRs) distributed over the whole genome may still be useful for selecting parents for the crossing programme. Moreover, if a random set of more markers is used, one would expect that these will be distributed over the whole genome and, therefore, will have a higher probability of being linked with quantitative trait loci (QTLs) for different traits, thus satisfying the requirement of QTL linked GS established in the above study. Therefore, further studies may need to be carried out to find out the utility of the GS worked out using random molecular markers rather than QTL-linked GS for predicting the σ^2_g in the progeny of a proposed cross.

The above discussion amply demonstrates the utility of microsatellites, which can be profitably utilized in wheat not only for detecting polymorphism and tagging genes (Prasad et al. 1999; Roy et al. 1999) but also for genotype identification and for estimation of genetic diversity. We conclude, therefore, that on the basis of microsatellite markers, diverse parents can be selected. With the availability of a rich collection of microsatellite primers recently made available through collaborative efforts under the Wheat Microsatellite Consortium, and through individual efforts elsewhere, microsatellites will certainly become the markers of choice in the future for a variety of studies.

Acknowledgements We are thankful to S. Nagarajan, DWR, Karnal for providing the seed materials along with details of pedigrees; to M. Röder, IPK, Gatersleben for supplying the DNA aliquots of nullisomic-tetrasomic lines; and to R.P. Sharma and T. Mohapatra, NRC on Plant Biotechnology, IARI, New Delhi for their help in preparing the dendrogram. Thanks are also due to several scientists (listed in Table 3) for granting permission to include the data on sequences of different WMC primer pairs and the corresponding microsatellites. We also gratefully acknowledge the financial support from the Department of Biotechnology (DBT), and Council of Scientific and Industrial Research (CSIR), Government of India, for carrying out this study.

References

Bohn M, Friedrich Utz H, Melchinger AE (1999) Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs and SSRs and their use for predicting progeny variance. Crop Sci 39: 228–237

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32: 314–331

Bredemeijer GMM, Arens P, Wouters D, Visser D, Vosman B (1998) The use of semi-automated fluorescent microsatellite analysis for tomato cultivar identification. Theor Appl Genet 97: 584–590

Burkhamer RL, Lanning SP, Martens RJ, Martin JN, Talbert LE (1998) Predicting progeny variance from parental divergence in hard red spring wheat. Crop Sci 38:234–248

Chen HB, Martin JM, Lavin M, Talbert LE (1994) Genetic diversity in hard red spring wheat based on sequence-tagged-site PCR markers. Crop Sci 34: 1629–1632

Cox TS, Lookhart GL, Walker DE, Harrell LG, Albers LD, Rodgers DM (1985) Genetic relationships among hard red winter wheat cultivars as evaluated by pedigree analysis and gliadin polyacrylamide-gel electrophoretic patterns. Crop Sci 25: 1058–1063

Devos KM, Atkinson MD, Chinoy CN, Liu CJ, Gale MD (1992) RFLP-based genetic map of the homoeologous group 3 chromosomes of wheat and rye. Theor Appl Genet 83: 931–939

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

Dweikat I, MacKenzie S, Levy M, Ohm H (1993) Pedigree assessment using RAPD-DGGE in cereal crop species. Theor Appl Genet 83: 497–505

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20: 758–760

Gupta PK, Balyan HS, Sharma PC, Ramesh B (1996) Microsatellites in plants: a new `class of molecular markers. Curr Sci 70: 45–54

Hart G, Gale MD (1988) Guidelines for nomenclature of biochemical/molecular loci in wheat. In: Miller TE, Koebner RMD (eds) Proc 7th Int Wheat Genet Symp Vol. 2. Bath Press, Bath, UK, pp 1215–1218

Hayes PM, Cerono J, Witsenhoer H, Kuiper M, Zabeau M, Sato K, Kleinofs A, Kudrna D, Kilian A, Saghai-Maroof M, Hoffman D (1997) Characterizing and exploiting genetic diversity and quantitative traits in barley (*Hordeum vulgare*) using AFLP markers. J Quantitative Trait Loci 3: http://probe.nalusda.gov:8000/other docs/jqtl

He S, Ohm H, MacKenzie S (1992) Detection of DNA sequence polymorphisms among wheat varieties. Theor Appl Genet 84: 573–578

Hokanson SC, Szewc-McFadden AK, Lamboy WF, McFerson JR (1998) Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in *Malus* × *domestica* borkh. Core subset collection. Theor Appl Genet 97: 671–683

Jaccard P (1908) Nouvelles recherches sur la distribution florale. Bull Soc Vaudoise Sci Nat 44: 223–270

Lee JA, Kaltsikes PJ (1973) The application of Mahalanobis's generalised distances to measure genetic divergence in durum wheat. Euphytica 22: 124–131

Manifesto MM, Schlatter AR, Hopp HE, Suarez EY, Dubcovsky J (1999) Bread wheat (*Triticum aestivum*) fingerprinting using microsatellites. In: Plant Anim Genome 7th Conf., p 371 (http://www.intl-pag.org/)

McIntosh RA, Hart GE, Devos KM, Gale MD, Rogers WJ (1998) Catalogue of gene symbols for wheat. In: Slinkard AE (ed) Proc 9th Int Wheat Genet Symp Vol. 5. University Extension Press, University of Saskatchewan, Canada, pp 1–12

Paull JG, Chalmers KJ, Karakousis A, Kretschmer JM, Manning S, Langridge P (1998) Genetic diversity in Australian wheat varieties and breeding material based on RFLP data. Theor Appl Genet 96: 435–446

Penner GA, Clarke J, Bezte LJ, Leisle D (1995) Identification of RAPD markers linked to a gene governing cadmium uptake in duram wheat. Genome 38: 543–547

Plaschke J, Ganal MW, Röder MS (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor Appl Genet 91: 1001–1007

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

Prasad M, Varshney RK, Kumar A, Balyan HS, Sharma PC, Edwards KJ, Singh H, Dhaliwal HS, Roy JK, Gupta PK (1999) A microsatellite marker associated with a QTL for grain protein content on chromosome arm 2DL of bread wheat. Theor Appl Genet 98: (in press)

Röder MS, Plaschke J, König SU, Börner A, Sorrells ME, Tanksley SD, Ganal MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol Gen Genet 246: 327–333

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149: 2007–2023

Rohlf FJ (1992) NTSYS-PC: Numerical taxonomy and multivariate analysis system. Exeter Software, New York

Roy JK, Prasad M, Varshney RK, Balyan HS, Blake TK, Dhaliwal HS, Singh H, Edwards KJ, Gupta PK (1999) Identification of a microsatellite on chromosome 6B and a STS on 7D of bread wheat showing association with preharvest sprouting tolerance. Theor Appl Genet 98: (in press).

Russell J, Fuller J, Young G, Thomas B, Taramino G, Macaulay M, Waugh R, Powell W (1997) Discriminating between barley genotypes using microsatellite markers. Genome 40: 442–450

Saghai Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81: 8014–8019

Stephenson P, Bryan G, Kirby J, Collins A, Devos K, Busso C, Gale M (1998) Fifty new microsatellite loci for the wheat genetic map. Theor Appl Genet 97: 946–949

Struss D, Plieske J (1998) The use of microsatellite markers for detection of genetic diversity in barley populations. Theor Appl Genet 97: 308–315

Talbert LE, Blake NK, Chee PW, Blake TK, Magyar GM (1994) Evaluation of "sequence-tagged-site" PCR products as molecular markers in wheat. Theor Appl Genet 87: 789–794

Tegelstrom H (1992) Detection of mitochondrial DNA fragments. In: Hoelzel AR (ed) Molecular genetic analysis of populations: a practical approach. IRL Press, Oxford, pp 89–114

Vaccino P, Accerbi M, Corbellini M (1993) Cultivar identification in *T. aestivum* using highly polymorphic RFLP probes. Theor Appl Genet 86: 833–836

Xie DX, Devos KM, Moore G, Gale MD (1993) RFLP-based genetic maps of the homoeologous group 5 chromosomes of bread wheat (*Triticum aestivum* L.). Theor Appl Genet 87: 70–74