Theor Appl Genet (2002) 105:699–707
DOI 10.1007/s00122-002-0959-4

X. Q. Huang · A. Börner · M. S. Röder · M. W. Ganal

Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers

Received: 8 November 2001 / Accepted: 4 January 2002 / Published online: 19 June 2002

Abstract A set of 24 wheat microsatellite markers, representing at least one marker from each chromosome, was used for the assessment of genetic diversity in 998 accessions of hexaploid bread wheat (*Triticum aestivum* L.) which originated from 68 countries of five continents. A total of 470 alleles were detected with an average allele number of 18.1 per locus. The highest number of alleles per locus was detected in the B genome with 19.9, compared to 17.4 and 16.5 for genomes A and D, respectively. The lowest allele number per locus among the seven homoeologous groups was observed in group 4. Greater genetic variation exists in the non-centromeric regions than in the centromeric regions of chromosomes. Allele numbers increased with the repeat number of the microsatellites used and their relative distance from the centromere, and was not dependent on the motif of microsatellites. Gene diversity was correlated with the number of alleles. Gene diversity according to Nei for the 26 microsatellite loci varied from 0.43 to 0.94 with an average of 0.77, and was 0.78, 0.81 and 0.73 for three genomes A, B and D, respectively. Alleles for each locus were present in regular two or three base-pair steps, indicating that the genetic variation during the wheat evolution occurred step by step in a continuous manner. In most cases, allele frequencies showed a normal distribution. Comparative analysis of microsatellite diversity among the eight geographical regions revealed that the accessions from the Near East and the Middle East exhibited more genetic diversity than those from the other regions. Greater diversity was found in Southeast Europe than in North and Southwest Europe. The present study also indicates that microsatellite markers permit the fast and high throughput fingerprinting of large numbers of accessions from a germplasm collection in order to assess genetic diversity.

Keywords Genetic diversity · Germplasm · Microsatellites · *Triticum aestivum*

Communicated by G. Wenzel

X.Q. Huang (✉) · A. Börner · M.S. Röder · M.W. Ganal
Institut für Pflanzengenetik und Kulturpflanzenforschung,
Corrensstrasse 3, D-06466 Gatersleben, Germany
e-mail: huang@ipk-gatersleben.de
Tel.: +49-39482-5589, Fax: +49-39482-5137

Present address:
M.W. Ganal, TraitGenetics GmbH, Am Schwabeplan 1b,
06466 Gatersleben, Germany

Introduction

Genetic diversity is the basis for genetic improvement. Knowledge of germplasm diversity has a significant impact on the improvement of crop plants. Due to modern breeding, it has been suggested that genetic diversity in wheat has been increasingly narrowed. Narrow genetic diversity is problematic in breeding for adaptation to biotic stresses, like diseases, and abiotic stresses, such as drought or salt tolerance. Therefore, it is necessary to investigate the genetic diversity in wheat germplasm in order to broaden the genetic variation in future wheat breeding. Morphological traits can be used for assessing genetic diversity but are often influenced by the environment. The use of molecular markers for the evaluation of genetic diversity is receiving much attention. Many wheat scientists have studied genetic diversity in common wheat using different molecular markers such as RAPDs (Joshi and Nguyen 1993), RFLPs (Siedler et al. 1994; Kim and Ward 2000), AFLPs (Barrett and Kidwell 1998; Burkhamer et al. 1998), STS (Chen et al. 1994) and ISSRs (Nagaoka and Ogihara 1997). However, most of these marker systems (Chao et al. 1989; Devos and Gale 1992) show a low level of polymorphism in wheat, especially among cultivated lines and/or cultivars.

Microsatellites, also termed simple sequence repeats (SSRs), have been proposed as one of the most-suitable markers for the assessment of genetic variation and diversity among wheat varieties/lines, because they are multiallelic, chromosome-specific and evenly distributed along chromosomes (Röder et al. 1998a, b). Microsatellite markers have been applied widely for tagging resistance genes (Peng et al. 1999; Börner et al. 2000b), iden-

tifying QTLs (Parker et al. 1998), marker-assisted selection in wheat (Korzun et al. 1998; Huang et al. 2000a) and verifying the integrity and genetic stability of gene bank accessions (Börner et al. 2000a). Such markers also revealed a high level of polymorphism among diploid species (Hammer et al. 2000), in the accessions of tetraploid wild wheat *Triticum dicoccoides* (Fahima et al. 2002) and of *Aegilops tauschii*, the D-genome donor of bread wheat (Pestsova et al. 2000), and as well as in hexaploid wheat varieties (Plaschke et al. 1995; Donini et al. 1998; Prasad et al. 2000; Stachel et al. 2000).

The objectives of the present study were to examine the genetic factors that affected microsatellite diversity and to use wheat microsatellite markers for the characterization and assessment of the genetic diversity of a large number of wheat germplasms from the gene bank at the IPK, Gatersleben.

Materials and methods

Plant materials and DNA isolation

At Gatersleben, the germplasm collection of hexaploid bread wheat consists of more than 10,000 accessions. Nine hundred and ninety eight accessions (approximately 10% of the total collection) which are representative for the different geographical origins (66 countries, Table 1) were chosen for microsatellite analysis. Total genomic DNA was extracted from five grains of each accession according to Plaschke et al. (1995).

Microsatellite markers

Twenty four wheat microsatellite markers for 26 loci representing at least one microsatellite marker from each chromosome (Table 2) were selected for genotyping (Röder et al. 1998b; unpublished results). The primer sequence of taglgap was described by Devos et al. (1995). All Gatersleben Wheat Microsatellites (gwm) used were dinucleotide repeats, whereas taglgap has a trinucleotide motif.

Polymerase chain reaction and fragment analysis

PCR reactions were performed according to Röder et al. (1998b) and Devos et al. (1995), respectively. Multiplex PCR (two or three microsatellite primer pairs with the same annealing temperature in one PCR reaction) and multiple sample loading (loading more than one PCR reaction in each well) were used in this study. Microsatellite fragments were detected on an automated laser fluorescence (A.L.F.) sequencer and analysed using the computer program Fragment Analyser Version 1.02 (Pharmacia) by comparison with internal size standards (Röder et al. 1998b). The varieties 'Chinese Spring' and '-Aztec' were used as controls in each run to

Table 1 Origin and distribution of 998 wheat accessions from the IPK gene bank

Continent	Geographical region	Country/ region	Number of accessions		Total
Africa	East Africa	Ethiopia	31	37	113
		Kenya	1		
		Mozambique	1		
		South Africa	1		
		Sudan	3		
	Northwest Africa	Algeria	1	76	
		Canary Islands	2		
		Egypt	1		
		Libya	65		
		Tunisia	7		
America	North America	USA	10	10	38
	Central America	Mexico	3	4	
		Unkonwn	1		
	South America	Argentina	4	24	
		Bolivia	1		
		Brazil	2		
		Chile	11		
		Columbia	2		
		Unknown	2		
		Uruguay	2		
Asia	The Near East	Azerbaijan	1	159	391
		Georgia	20		
		Iran	113		
		Iraq	10		
		Jordan	1		
		Palestine	1		
		Syria	2		
		Turkey/Anatolia	11		
	The Middle East	Afghanistan	24	204	
		Hindukusch	40		
		(North) India	12		
		Kazakhstan	1		
		Nepal	19		
		Pakistan	24		
Asia		Tajikistan	1		
		Tibet	83		
	The Far East	China	13	28	
		Japan	4		
		Korea	6		
		Mongolia	5		
Europe	North	CSSR	76	139	450
		England	3		
		Estonia	2		
		Finland	4		
		Germany	6		
		Holland	1		
		Poland	34		
		Sweden	9		
		USSR	3		
		Ukraine	1		
	Southeast	Albania	11	189	
		Austria	112		
		Bulgaria	18		
		Crete	4		
		Croatia	1		
		Hungary	1		
		North Greece	25		
		Peloponnesos	12		
		Romania	3		
		Yugoslavia	2		
	Southwest	France	5	122	
		Italy	99		
		Portugal	7		
		Spain	9		
		Switzerland	2		
Oceania		Australia	1	1	1
Unknown		Unknown	5	5	5

Table 2 Description of 24 wheat microsatellites, their chromosomal location and motif

Microsatellite	Chromosomal location	Motif	Fragment size in 'CS' (bp)
Xgwm3	3DL	(CA)18	79
Xgwm18	1BS	(CA)17GA(TA)4	183
Xgwm46	7B (C)	(GA)2GC(GA)33	179
Xgwm95	2AS	(AC)16	122
Xgwm155	3AL	(CT)19	144
Xgwm160	4AL	(GA)21	182
Xgwm186	5AL	(GA)26	135
Xgwm190	5DS	(CT)22	209
Xgwm192	4AS, 4BL, 4DL	(CT)46	135,189,141
Xgwm261	2DS	(CT)21	189
Xgwm325	6DS	(CT)16	137
Xgwm337	1DS	(CT)5(CACT)6(CA)43	177
Xgwm357	1A (C)	(GA)18	122
Xgwm389	3BS	(CT)14(GT)16	129
Xgwm408	5BL	(CA)>22(TA)(CA)7(TA)9	173
Xgwm437	7DL	(CT)24	107
Xgwm458	1D (C)	(CA)13	112
Xgwm459	6AS	(GA)>28	149
Xgwm513	4BL	(CA)12	140
Xgwm577	7BL	(CA)14(TA)6	131
Xgwm619	2BL	(CT)19	148
Xgwm631	7AS	(GT)23	196
Xgwm680	6BS	(GT)7(GA)24imp	121
Xtaglgap	1BS	(CAA)31	282

ensure size accuracy and to minimize run-to-run or gel-to-gel variation.

Data analysis

Fragments amplified by microsatellite primers were scored as presence (1) or absence (0). The genetic similarities (GS) were calculated for each pair of lines using the Dice similarity index (Dice 1945). Cluster analyses were performed with the NTSYS-pc package (Rohlf 1998) based on the unweighted pair-group method with arithmetic average (UPGMA).

Gene diversity was calculated according to formula of Nei (1973):

$$\text{Gene diversity} = 1 - \sum \text{Pij}^2,$$

where Pij is the frequency of the jth allele for ith locus summed across all alleles for the locus. Anderson et al. (1993) referred to gene diversity as the polymorphic information content (PIC).

Because genetic distances of microsatellites from the centromere correspond to different physical distances on different wheat chromosomes, relative distances (%) from the centromere were employed in this study. The relative distance (%) was a percentage of the genetic distance of a marker from the centromere to the genetic distance relative to the most distal marker on the same chromosome arm. The repeat number of the main allele was calculated according to the repeat number of the allele in 'CS', and by comparing the fragment sizes between 'CS' and the main allele. The data were used to examine the following relationships: (1) the number of alleles vs the gene diversity; (2) the number of alleles vs the relative distance (%) from the centromere; (3) the number of alleles vs the repeat number of the main allele. Statistical analyses were performed using the computer software Statview version 4.02.

Results

Microsatellite polymorphism

Twenty four microsatellite markers for 26 loci were used to characterize and evaluate the genetic diversity of 998 wheat accessions. A total of 470 alleles were detected. The number of alleles per locus ranged from four for *Xgwm192a* to 46 for *Xgwm459* with an average number of 18.1 alleles per locus (Table 3). The largest number of alleles per locus occurred in the B genome with 19.9, compared to 17.4 and 16.5 for genomes A and D, respectively (Table 4). The lowest number of alleles per locus among the seven homoeologous groups was observed in homoeologous group 4 with 10.0. Two microsatellite markers from chromosomes 1B, 4B, 7B and 1D were used for analysis (Table 5). In each case, one microsatellite marker was close to the centromere, whereas the other was located distantly from the centromere. The allele numbers per marker revealed that less genetic variation occurs in the centromeric regions compared to the non-centromeric regions of chromosomes. When the results from all 26 microsatellite loci on the 21 chromosomes were combined, allele numbers were also correlated to the relative distance of the locus from the centromere with $r = 0.48$ ($P < 0.05$).

The number of alleles was related to the motif and repeat number of the microsatellites. Because of the normal distribution of allele frequencies at most microsatellite loci, the analysis was based on the repeat number of the main allele for each locus. The (GA)*n* microsatellites produced an average of 20.2 alleles per locus, whereas microsatellites with the (GT)*n* motif had only 10.6 (Ta-

Table 3 Chromosomal location, relative distance (%) from the centromere, size range and number of alleles as well as gene diversity detected for 998 wheat accessions from the IPK gene bank

Locus	Location	Relative distance (%) from the centromere	Size range of alleles (bp)	Number of alleles	Gene diversity
Xgwm357	1A (C)	0	106–138	12	0.69
Xgwm95	2AS	2.09	104–130	13	0.78
Xgwm155	3AL	57.31	126–156	15	0.79
Xgwm160	4AL	100	170–192	12	0.70
Xgwm192c	4AS	35.77	130–137	4	0.43
Xgwm186	5AL	22.66	101–157	25	0.90
Xgwm459	6AS	100	105–207	46	0.94
Xgwm631	7AS	4.02	180–208	12	0.76
Xgwm18	1BS	5.2	171–203	13	0.77
Xtaglgap	1BS	76.92	209–304	26	0.82
Xgwm619	2BL	67.79	136–178	21	0.86
Xgwm389	3BS	98.91	103–149	19	0.88
Xgwm513	4BL	12.26	136–150	8	0.65
Xgwm192a	4BL	18.75	165–213	18	0.84
Xgwm408	5BL	63.13	145–199	20	0.84
Xgwm680	6BS	9.52	107–141	14	0.66
Xgwm46	7B (C)	0	141–201	24	0.88
Xgwm577	7BL	75.45	127–227	36	0.90
Xgwm337	1DS	1.35	159–211	24	0.83
Xgwm458	1D (C)	0	108–120	7	0.62
Xgwm261	2DS	51.39	161–215	22	0.56
Xgwm3	3DL	42.89	71–89	10	0.67
Xgwm192b	4DL	9.3	126–147	8	0.63
Xgwm190	5DS	100	183–245	23	0.82
Xgwm325	6DS	5.47	123–173	15	0.82
Xgwm437	7DL	13.94	89–133	23	0.90

Table 4 Genetic diversity in different genomes, chromosomes and motifs across 26 loci in 998 wheat accessions

	Number of alleles	Gene diversity
Genome		
A	17.4	0.78
B	19.9	0.81
D	16.5	0.73
Chromosome		
1	16.4	0.74
2	18.7	0.73
3	14.7	0.78
4	10.0	0.65
5	22.7	0.85
6	25.0	0.81
7	23.8	0.85
Motif		
(GT)*n*	10.6	0.71
(GA)*n*	20.2	0.78
Mean	15.4	0.75
(CA)j(TA)k	23.0	0.84
(CT)j(GT)k	19.0	0.88
(CT)j(CA)k	24.0	0.83
(GT)j(GA)k	14.0	0.66
Mean	20.0	0.80
(NN)*n* < 20	12.6	0.73
(NN)*n* > 20	19.3	0.75

Table 5 Comparison of allele numbers and gene diversity between centromeric regions and non-centromeric regions of chromosomes

Locus	Chromosomal location	Genetic distance from the centromere (cM)	Number of alleles	Gene diversity
Xgwm18	1BS	5.2	13	0.77
Xtaglgap	1BS	50	26	0.82
Xgwm513	4BL	6.1	8	0.65
Xgwm192a	4BL	23.1	18	0.84
Xgwm46	7B	0	24	0.88
Xgwm577	7BL	185	36	0.90
Xgwm458	1D	0	7	0.62
Xgwm337	1DS	16.5	24	0.83

ble 4). An average allele number of 12.6 and 19.3 per locus was found for the dinucleotide repeat microsatellites (NN)*n* < 20 and (NN)*n* > 20, respectively. Compared to simple microsatellites such as (GT)*n* and (GA)*n*, compound microsatellites such as (CA)j(TA)k and (CT)j(GT)k generated on average 20.0 alleles. The correlation coefficient for the number of alleles versus the repeat number of the main allele was $r = 0.49$ ($P < 0.05$). The loci *Xgwm192b* and *Xgwm192c* were excluded in this comparison, because the primer pair WMS192 detected three loci and the fragment with the expected size was assigned to chromosome 4B at the locus *Xgwm192a*.

Gene diversity for 26 microsatellite loci varied from 0.43 to 0.94 with an average of 0.77. Gene diversity for the three genomes A, B and D was 0.78, 0.81 and 0.73, respectively (Table 4). The correlation coefficient between gene diversity and the number of alleles was high, $r = 0.73$ ($P < 0.01$). The linear relationship between them is shown in Fig. 1.

The most polymorphic microsatellite marker was *Xgwm459* with 46 alleles. The allele sizes ranged from 105 bp to 207 bp. A fragment of 149 bp was amplified in 'Chinese Spring', where the repeat number of dimeric GA

units was 28. Hence it is assumed that the repeat number in the 998 investigated accessions ranged from 6 to 43. According to the one-step stepwise mutation model (Shriver et al. 1993), in which alleles change in size by only one repeat unit, a total of 52 alleles for a size range from 105 bp to 207 bp at the locus *Xgwm459* should be detected. In fact, 43 of 46 alleles ranging from 115 bp to 199 bp were found. There was one allele each with two base pairs except for six alleles, namely four alleles in the range of 105 bp to 115 bp and two alleles in the range of 199 bp to 205 bp in the 998 accessions studied. Similarly, one allele was found for each two or three base-pair steps for all microsatellites with dimeric units and *Xtaglgap* with a trimeric unit, respectively. Figure 2 shows some reference alleles that were present in three base-pairs steps at the locus *Xtaglgap*. Allele frequencies of 15 microsatellite loci showed a normal distribution (Fig. 3A). The distribution of allele frequencies at 11 loci were irregular, showing bimodality at *Xtaglgap*, *Xgwm46*, *Xgwm186*, *Xgwm389*, *Xgwm459*, *Xgwm577* and *Xgwm631*, trimodality at *Xgwm95*, *Xgwm261* and *Xgwm408*, and a tetramodal distribution at *Xgwm619* (Fig. 3B).

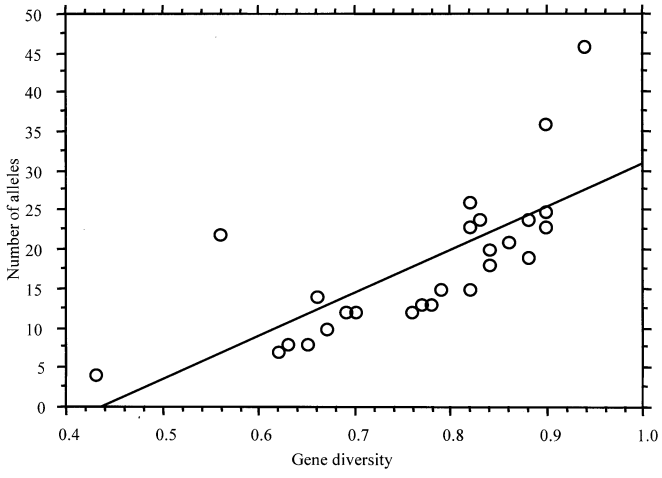

Fig. 1 Relationship between gene diversity and the number of alleles detected at 26 microsatellite loci, described by the function Y = –23.933 + 54.777X (R^2 = 0.535)

Genetic relationship and diversity among different geographical regions

A dendrogram derived from UPGMA cluster analysis based on the GS coefficient matrix for the 998 accessions was constructed. Basically all accessions could be distinguished. The genetic similarity coefficient for all accessions ranged from 0.15 to 0.96 and averaged 0.52. Because all 998 investigated accessions were from hexaploid wheat and belong to one species, 70% of the accessions from the same geographical origin could be clustered in specific groups. The groups were found in Libya, Iran and Iraq, Nepal and Tibet, CSSR and Poland, Austria, and Italy and Spain, respectively. The accessions from Afghanistan, Hindukusch, Pakistan and Iran were located in one cluster.

A comparison of the genetic diversity of wheat accessions was performed among eight germplasm pools, namely Africa, America, the Near East, the Middle East, North Europe, Southeast Europe and Southwest Europe. The number of alleles per locus, the number of region-specific alleles, the number of accessions carrying rare al-

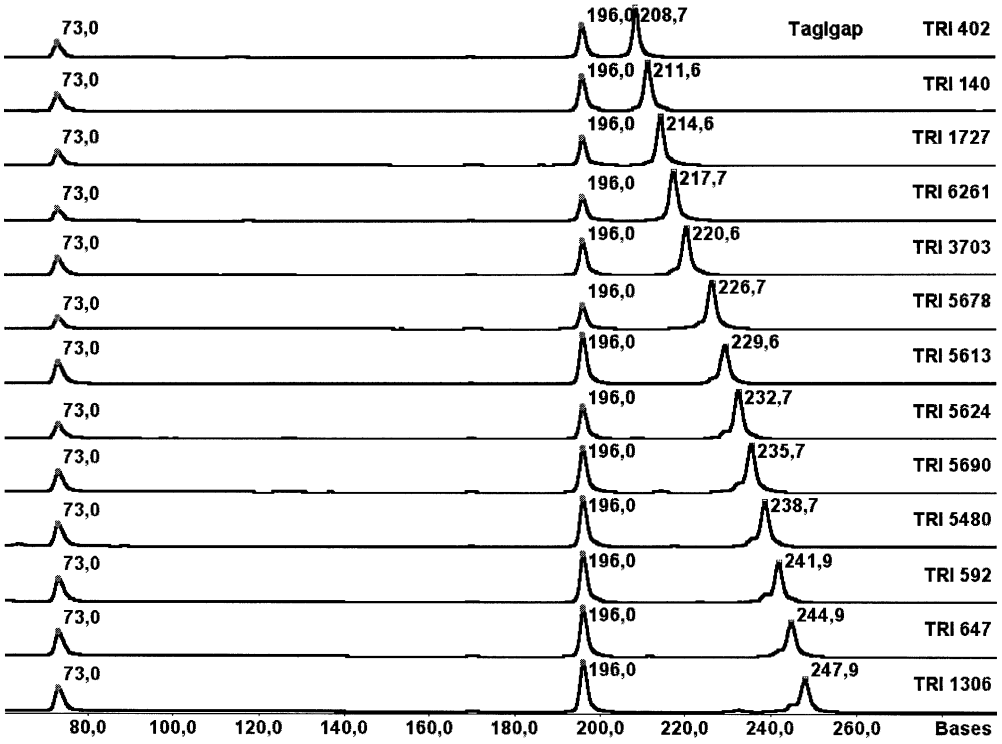

Fig. 2 Electropherograms analysed using Fragment Analyser version 1.02 show some reference alleles at the locus *Xtaglgap*. The *peaks* represent fragments of different alleles, whereas the *horizontal scale* indicates fragment sizes in base pairs calculated from internal standards (73 bp and 196 bp)

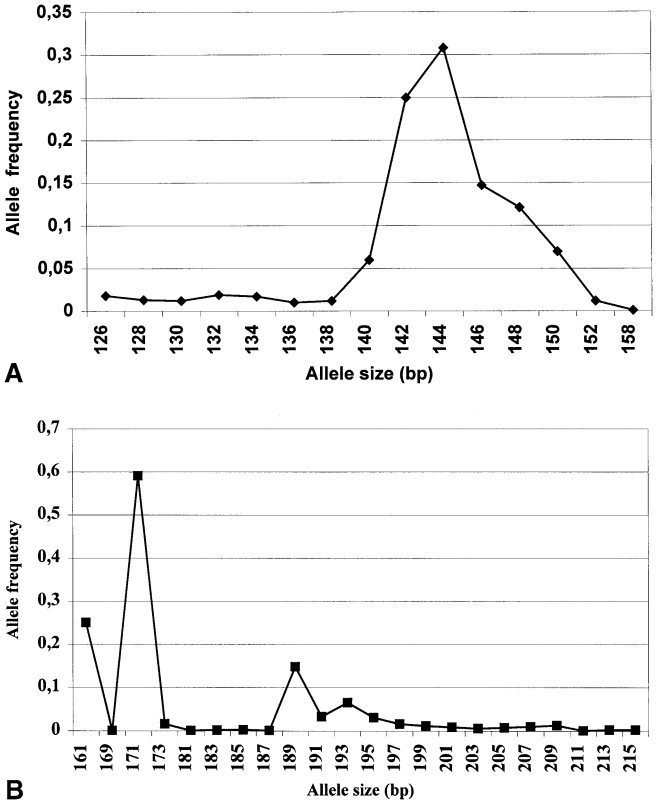

Fig. 3 The distribution of allele frequencies at the *Xgwm155* locus (**A**) and *Xgwm261* locus (**B**)

leles, the mean gene diversity and the mean genetic similarity within a region were used to compare genetic diversity among the regions (Table 6). The accessions from the Near East and the Middle East revealed more microsatellite diversity compared to other regions. The number of alleles per locus was 13.92 and 13.88 for the Near East and the Middle East, respectively, and higher than that of the other regions. Similarly, the number of region-specific alleles and the number of accessions carrying rare alleles were larger in the Near East and the Middle East than in the other regions. Forty six and 52 accessions were found to carry rare alleles in the Near East and the Middle East, respectively. The mean gene diversity in the Near East and the Middle East was 0.741 and 0.732 compared with values from 0.645 in North Europe to 0.731 in Africa (Table 6). The Near East and the Middle East accessions had a relatively low mean genetic similarity (GS) with values of 0.446 and 0.464, respectively. The lowest mean GSs were found in America and the Far East with 0.408 and 0.367. This can be explained by the fact that the accessions from both regions originated from geographically largely different countries. These results suggested that the Near East and the Middle East exhibited greater genetic diversity than other regions. Both regions were the presumed center of origin for hexaploid wheat (Yen et al. 1983; Dvorak et al. 1998).

Four hundred and fifty one accessions from Europe were divided into three different regions, namely North Europe, Southeast Europe and Southwest Europe. It is very evident that the largest genetic variation exists in Southeast Europe, followed by Southwest Europe and North Europe, based on the allele number per locus, the number of region-specific alleles, the number of accessions carrying rare alleles, the mean gene diversity and the mean genetic similarity within a region. Compared to 65 accessions from Libya, Northwest Africa, 31 accessions from Ethiopia, East Africa, revealed greater genetic diversity (data not shown).

Discussion

In the present study, a set of 24 microsatellite markers detecting 26 loci was used to characterize 998 accessions of hexaploid wheat germplasm. These microsatellites produced a large number of alleles. Alleles for each locus were present in regular two or three base-pair steps, indicating that genetic variation during wheat evolution occurred in a continuous step by step manner. However, gaps with one or more than alleles were detected at 18 loci. These gaps were observed between the alleles with the lowest allele frequencies. This finding supports the two-phase mutation model for SSR loci (Di Rienzo et al. 1994), which assumed that most mutational changes result in an increase or decrease of one repeat unit and that large jumps in repeat number occur only infrequently. Fifteen of 26 loci displayed the regular normal distributions of allele frequencies that were consistent with the stepwise mutation model (Ohta and Kimura 1973). The distribution of allele frequencies at 11 loci were irregular, showing a bimodal, trimodal or tetramodal distribution (Fig. 3B). It appears that natural selection plays a role in creating allelic variation at these loci. The alleles

Table 6 Comparison of genetic diversity of wheat accessions among the different regions. Note: AFR = Africa; AME = America; NET = Near East; MET = Middle East; FET = Far East; NEU = North Europe; SEE = Southeast Europe; SWE = Southwest Europe

Region	AFR	AME	NET	MET	FET	NEU	SEE	SWE
No. of accessions	113	38	159	204	28	139	189	122
No. of total alleles	303	233	362	361	224	273	330	307
No. of alleles per locus	11.65	8.96	13.92	13.88	8.62	10.5	12.69	11.81
No. of region-specific alleles	6	1	12	15	3	3	11	2
No. of accessions carrying rare alleles	25	4	46	52	4	9	35	15
Mean gene diversity	0.731	0.700	0.741	0.732	0.724	0.645	0.721	0.685
Mean GS within a region	0.531	0.408	0.446	0.464	0.367	0.589	0.509	0.483

with higher frequencies might be selected and kept for adaptational reasons. For example, the microsatellite marker *Xgwm261*, which mapped 0.6 cM distal to the dwarfing gene *Rht8* on the short arm of chromosome 2D, produced three main alleles that are widespread in varieties of wheat. This could be explained by the selection for these different alleles in different geographic regions (Korzun et al. 1998). Similarly, the taglgap marker resides in a storage protein gene cluster which might have an influence on bread-making quality.

The value of gene diversity increased with the number of alleles at a given locus. There was significant correlation between gene diversity and the number of alleles ($r = 0.73$, $P < 0.01$). Therefore the number of alleles can be used for the evaluation of genetic diversity. Our results did not agree with those of Prasad et al. (2000) who reported that the PIC value was not correlated with the number of alleles. A large sample size is necessary for the characterization of a reliable correlation coefficient. The sample size in their study was only 55 accessions.

The number of alleles was also correlated with the repeat number of microsatellite DNA and its relative distance from the centromere. It has been suggested that the three mechanisms for creating a new allele at SSR loci are replication slippage (Tachida and Iizuka 1992), unequal crossing-over (Harding et al. 1992) and genetic recombination. Replication slippage is considered to be a major factor affecting the repeat number for short tandem repeat sequences, whereas unequal crossing-over is thought to result in a very large number of alleles for long tandem repeat arrays. All microsatellites used in the present study were di- or tri-nucleotide repeats, hence replication slippage probably plays a major role in creating new alleles at these SSR loci. The number of alleles does not seem to be correlated with the motifs of microsatellites. The lower number of alleles detected for microsatellites with a (GT) motif is because these four microsatellites have shorter repeat units ranging from 13 to 18 with a mean of 14.7. Compound microsatellites produced a very large number of alleles, because they possess very long repeat units (Table 3). Genetic diversity is related to recombination frequency. Lukaszewski and Curtis (1993) investigated the physical distribution of recombination in B-genome chromosomes of tetraploid wheat and found that recombination was absent in proximal regions of all chromosome arms; its frequency increased exponentially with the distance from the centromere. In the present study, a lower number of alleles was detected in the centromeric regions than in the non-centromeric regions of chromosome 1B, 4B, 7B and 1D (Table 5). With respect to 26 microsatellite loci of 21 chromosomes, there was a significant correlation between the number of alleles and the relative distance of the locus from the centromere ($r = 0.48$, $P < 0.05$). This indicates that a higher number of alleles and a higher value of gene diversity reflect a higher recombination frequency. Therefore, the repeat length of microsatellites and their location on chromosomes should be considered for the selection of microsatellites in the evaluation of genetic diversity and variety identification.

The different contribution of three genomes to genetic variation within *Triticum aestivum* was confirmed in this study. The greatest genetic variation was found in the B-genome with 19.9 alleles per locus, followed by the A-genome and D-genome. This is consistent with the distribution of microsatellite markers and AFLP markers in the three genomes, respectively (Röder et al. 1998b; Huang et al. 2000b). It is interesting to note that the lowest genetic variation exists in the chromosomes of homoeologous group 4 based on five loci (Tables 3 and 4). Although *Xgwm160* is located in the distal region of chromosome 4AL, the number of alleles is only 12. Similar results were reported by Boyko et al. (1999) and Ma et al. (2001) who found that the lowest number of molecular markers was mapped to chromosome 4D of *Aegilops tauschii* and the short arm of chromosome 4R of rye, respectively. Moreover, the lowest number of RFLP markers was also mapped to chromosome 4H of barley (Graner et al. 1991; Kleinhofs et al. 1993). It is most likely that the chromosome structure of homoeologous group 4 is conserved or that some factors for important characters are located on chromosomes of homoeologous group 4. Any large variation or mutation occurring in these chromosomes might lead to the death of plants. It is known that one recessive gene and one dominant gene for male sterility were located on the short arm of chromosomes 4B and 4D of wheat, respectively (Barlow and Driscoll 1981; Liu and Deng 1986).

Microsatellites displayed a high level of polymorphism in the present study. The average number of alleles per microsatellite locus and mean gene diversity were 18.1 and 0.77, respectively. This contrasts sharply with the level of polymorphism detected by RFLPs and RAPDs in wheat. Kim and Ward (2000) employed 30 RFLP probes to estimate the genetic diversity of 292 accessions of common wheat and detected lower mean PIC values. According to Siedler et al. (1994), RFLP analysis revealed 4.7 polymorphisms per probe/enzyme combination among 81 European cultivars, whereas RAPDs primers generated only 1.8 polymorphisms/primer among 15 wheat cultivars (Joshi and Nguyen 1993). Our results indicate that microsatellite markers are much more informative than RFLPs and RAPDs in wheat.

A comparison with the results from previous studies on the genetic diversity of wheat cultivars indicates that the average number of alleles per locus in the present study is much higher than that in the previous studies. The mean number of alleles was 6.2 and 4.8 reported by Plaschke et al. (1995) and Stachel et al. (2000), respectively. Similarly, Prasad et al. (2000) assessed the genetic diversity of 55 genotypes of wheat originating from 29 countries representing six continents using microsatellite markers and found the average number of alleles with 7.4. The same set of microsatellite markers was used for the identification of 500 European wheat varieties, but the mean number of alleles was only 10.5 (Röder et al. 2002). It is evident that much greater variation exists in accessions of wheat germplasm from the gene bank than the wheat cultivars.

Table 7 Comparison of allele numbers between accessions of common wheat, accessions of *T. dicoccoides* and accessions of *Ae. tauschii*

Locus	Location	998 accessions of *T. aestivum* (AABBDD)	135 accessions of *T. dicoccoides* (AABB)[a]	113 accessions of *Ae. tauschii* (DD)[a]
Xgwm95	2AS	13	14	
Xgwm155	3AL	15	21	
Xgwm186	5AL	25	20	
Xgwm18	1BS	12	17	
Xgwm389	3BS	19	24	
Xgwm408	5BL	20	22	
Xgwm577	7BL	36	18	
Mean		20	19.4	
Xgwm261	2DS	22		22
Xgwm3	3DL	10		12
Xgwm190	5DS	23		22
Xgwm437	7DL	23		22
Mean		19.5		19.5

[a] The results from Fahima et al. (2002) and Pestsova et al. (2000), respectively

Wild emmer wheat, *Triticum dicoccoides* Korn (2n = 28, AABB), and diploid *Aegilops tauschii* (2n = 14, DD) are the tetraploid progenitors and the D-genome donors of cultivated wheat (2n = 42, AABBDD), respectively (Kihara 1944; Zohary 1970). Comparing the number of alleles detected in the accessions of common wheat in our study of *T. dicoccoides* (Fahima et al. 2002) and of *Ae. tauschii* (Pestsova et al. 2000) for the commonly used microsatellite markers (Table 7), it is evident that almost the same allele distribution and number for single loci was found in the accessions of common wheat, of wild emmer wheat and of *Ae. tauschii*, respectively, though the number of investigated accessions was different. This suggests that the accessions of hexaploid bread wheat possess nearly the same genetic diversity as wild emmer wheats and *Ae. tauschii*.

The 998 *T. aestivum* accessions collected in the gene bank could be distinguished by 24 microsatellite markers. However, not all accessions originating from the same geographic region were clustered in the same group, indicating that the genetic diversity of *T. aestivum* is not completely related to geographic distribution. The reason for this might be that similar genetic variation occurred independently in the different geographic regions or that artificial transfer of the accessions from one region to another region resulted in a false determination of the geographic origin.

Nishikawa et al. (1980) and Dvorak et al. (1998) suggested that the regions around the Caspian sea, such as SE and SW Caspian Iran and Armenia, were the geographical place of origin of *T. aestivum*. Large genetic variation should exist in the origin sites. Comparative analysis of microsatellite diversity among the regions of Africa, America, the Near East, the Middle East, North Europe, Southeast Europe and Southwest Europe indicated that the greatest genetic diversity was found in the Near East (Table 4). Twelve of 14 region-specific alleles in the Near East were found in 14 Iranian accessions, with one in Georgia. Our data thus support the above suggestions.

The present study indicates that microsatellites are very effective molecular markers for the assessment of genetic diversity in wheat. Combined with multiplex PCR or multiple sample loading, microsatellite markers permit the fast and high-throughput fingerprinting of large numbers of accessions from a germplasm collection. A relatively large genetic diversity was detected in the accessions of wheat germplasm from the gene bank at Gatersleben. Some accessions possessing rare alleles for many different loci could be used for broadening the allelic diversity and the potential for cultivar improvement of wheat.

Acknowledgements This research was supported by a grant from Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK). We thank Dr. R. Sobeck for selecting the accessions of wheat germplasm from the gene bank at Gatersleben, Dr. V. Korzun for his help in the early phase of this study and Dr. F. Blattner for assistance with the computer software NTSYS to construct the dendrogram.

References

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

Barlow KK, Driscoll CJ (1981) Linkage studies involving two chromosomal male-sterility mutants in hexaploid wheat. Genetics 98:791–799

Barrett BA, Kidwell KK (1998) AFLP-based genetic diversity assessment among wheat cultivars from the Pacific Northwest. Crop Sci 38:1261–1271

Börner A, Chebotar S, Korzun V (2000a) Molecular characterization of the genetic integrity of wheat (*Triticum aestivum* L.) germplasm after long-term maintenance. Theor Appl Genet 100:494–497

Börner A, Röder MS, Unger O, Meinel A (2000b) The detection and molecular mapping of a major gene for non-specific adult-plant disease resistance against stripe rust (*Puccinia striiformis*) in wheat. Theor Appl Genet 100:1095–1099

Boyko EV, Gill KS, Mickelson-Young L, Nasuda S, Raupp WJ, Ziegle JN, Singh S, Hassawi DS, Fritz AK, Namuth D, Lapitan NLV, Gill BS (1999) A high-density genetic linkage map of *Aegilops tauschii*, the D-genome progenitor of bread wheat. Theor Appl Genet 99:16–26

Burkhamer RL, Lanning SP, Martens RJ, Martin JM, Talbert LE (1998) Predicting progeny variance from parental divergence in hard red spring wheat. Crop Sci 38:243–248

Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. Theor Appl Genet 78:495–504

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 34:1628–1632

Devos KM, Gale MD (1992) The use of random amplified polymorphic DNA markers in wheat. Theor Appl Genet 84:567–572
Devos KM, Bryan GJ, Collins AJ, Gale MD (1995) Application of two microsatellite sequences in wheat storage proteins as molecular markers. Theor Appl Genet 90:247–252
Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302
Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci USA 91:3166–3170
Donini P, Stephenson P, Bryan GJ, Koebner RMD (1998) The potential of microsatellites for high throughput genetic diversity assessment in wheat and barley. Genet Res Crop Evol 45:415–421
Dvorak J, Luo M-C, Yang Z-L, Zhang H-B (1998) The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. Theor Appl Genet 97:657–670
Fahima T, Röder MS, Wendehake K, Kirzhner VM, Nevo E (2002) Microsatellite polymorphism in natural populations of wild emmer wheat, *Triticum dicoccoides*, in Israel. Theor Appl Genet 104:17–29
Graner A, Jahoor A, Schondelmaier J, Siedler H, Pillen K, Fischbeck G, Wenzel G, Herrmann RG (1991) Construction of an RFLP map of barley. Theor Appl Genet 83:250–256
Hammer K, Filatenko AA, Korzun V (2000) Microsatellite markers: a new tool for distinguishing diploid wheat species. Genet Res Crop Evol 47:497–505
Harding RM, Boyce AJ, Clegg JB (1992) The evolution of tandemly repetitive DNA: recombination rules. Genetics 132:847–859
Huang XQ, Hsam SLK, Zeller FJ, Wenzel G, Mohler V (2000a) Molecular mapping of the wheat powdery mildew resistance gene *Pm24* and marker validation for molecular breeding. Theor Appl Genet 101:407–414
Huang XQ, Zeller FJ, Hsam SLK, Wenzel G, Mohler V (2000b) Chromosomal location of AFLP markers in common wheat (*Triticum aestivum* L.) utilizing nulli-tetrasomic stocks. Genome 43:298–305
Joshi CP, Nguyen HT (1993) RAPD (random amplied polymorphic DNA) analysis based intervarietal genetic relationships among hexaploid wheats. Plant Sci 93:95–103
Kihara H (1944) Die Entdeckung des DD Analysators beim Weizen. Agric Hort 19:889–890
Kim HS, Ward RW (2000) Patterns of RFLP-based genetic diversity in germplasm pools of common wheat with different geographical or breeding program origins. Euphytica 115:197–208
Kleinhofs A, Kilian A, Saghai Maroof MA, Biyashev RM, Hayes P, Chen FQ, Lapitan N, Fenwick A, Blake TK, Kanazin V, Ananiev E, Dahleen L, Kudrna D, Bollinger J, Knapp SJ, Liu B, Sorrells M, Heun M, Franckowiak JD, Hoffman D, Skadsen R, Steffenson BJ (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theor Appl Genet 86:705–712
Korzun V, Röder MS, Ganal MW, Worland AJ, Law CN (1998) Genetic analysis of the dwarfing gene (*Rht8*) in wheat. Part I. Molecular mapping of *Rht8* on the short arm of chromosome 2D of bread wheat (*Triticum aestivum* L.).Theor Appl Genet 96:1104–1109
Liu BH, Deng JY (1986) A dominant gene for male sterility in wheat. Plant Breed 97:204–209
Lukaszewski AJ, Curtis CA (1993) Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. Theor Appl Genet 86:121–127
Ma XF, Wanous MK, Houchins K, Rodriguez Milla MA, Goicoechea PG, Wang Z, Xie M, Gustafson JP (2001) Molecular linkage mapping in rye (*Secale cereale* L.). Theor Appl Genet 102:517–523
Nagaoka T, Ogihara Y (1997) Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. Theor Appl Genet 94:597–602
Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323
Nishikawa K, Furuta Y, Wada T (1980) Genetic studies on alpha-amylase isoenzymes in wheat. III. Intraspecific variation in *Aegilops squarrosa* and birthplace of hexaploid wheat. Jpn J Genet 55:325–336
Ohta T, Kimura K (1973) The model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. Genet Res 22:201–204
Parker GD, Chalmers KJ, Rathjen AJ, Langride P (1998) Mapping loci associated with flour colour in wheat (*Triticum aestivum* L.). Theor Appl Genet 97:238–245
Peng JH, Fahima T, Röder MS, Li YC, Dahan A, Grama A, Ronin YI, Korol AB, Nevo E (1999) Microsatellite tagging of the stripe-rust resistance gene *YrH52* derived from wild emmer wheat, *Triticum dicoccoides*, and suggestive negative crossover interference on chromosome 1B. Theor Appl Genet 98:862–872
Pestsova E, Korzun V, Goncharov NP, Hammer K, Ganal MW, Röder MS (2000) Microsatellite analysis of *Aegilops tauschii* germplasm. Theor Appl Genet 101:100–106
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
Prasad M, Varshney RK, Roy JK, Balyan HS, Gupta PK (2000) The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. Theor Appl Genet 100:584–592
Röder MS, Korzun V, Gill BS, Ganal MW (1998a) The physical mapping of microsatellite markers in wheat. Genome 41:278–283
Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998b) A microsatellite map of wheat. Genetics 149:2007–2023
Röder MS, Wendehake K, Korzun V, Bredemeijer G, Laborie D, Bertrand L, Isaac P, Rendell S, Jackson J, Cooke RJ, Vosmann B, Ganal MW (2002) Construction and analysis of a microsatellite-based database of European wheat cultivars. Theor Appl Genet (in press)
Rohlf FJ (1998) NTSYS-pc: numerical taxonomy and multivariate analysis system, vers. 2.0. Applied Biostatistics Inc. New York
Shriver MD, Jin L, Chakraborty R, Boerwinkle E (1993) VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. Genetics 134:983–993
Siedler H, Messmer MM, Schachermayr GM, Winzeler H, Winzeler M, Keller B (1994) Genetic diversity in European wheat and spelt breeding material based on RFLP data. Theor Appl Genet 88:994–1003
Stachel M, Lelly T, Grausgruber H, Vollmann (2000) Application of microsatellites in wheat (*Triticum aestivum* L.) for studying genetic differentiation caused by selection for adaptation and use. Theor Appl Genet 100:242–248
Tachida W, Iizuka M (1992) Persistence of repeated sequences that evolve by replication slippage. Genetics 131:471–478
Yen C, Yang JL, Liu XD, Li LR (1983) The distribution of *Aegilops tauschii* Cosson in China and with reference to the origin of the Chinese common wheat. In: Sakamoto S (ed) Proc 6th Int Wheat Genet Symp, Kyoto, Japan, pp 55–58
Zohary D (1970) Centers of diversity and centers of origin. In: Frankel OH, Bonnet E (eds) Genetic resources in plants – their exploration and conservation. Blackwell, Oxford, pp 33–42

Note in proof A dendrogram of 998 wheat accessions from the gene bank Gatersleben was constructed using 24 microsatellite markers based on the genetic similarity coefficient matrix and is available of the WWW at "http://pgrc.ipk-gatersleben.de/dendro/".