Theor Appl Genet (1998) 97: 308–315

D. Struss · J. Plieske

The use of microsatellite markers for detection of genetic diversity in barley populations

Received: 26 November 1997 / Accepted: 19 January 1998

Abstract A barley lambda-phage library was screened with (GA)*n* and (GT)*n* probes for developing microsatellite markers. The number of repeats ranged from 2 to 58 for GA and from 2 to 24 for GT. Fifteen selected microsatellite markers were highly polymorphic for barley. These microsatellite markers were used to estimate the genetic diversity among 163 barley genotypes chosen from the collection of the IPK Genebank, Germany. A total of 130 alleles were detected by 15 barley microsatellite markers. The number of alleles per microsatellite marker varied from 5 to 15. On average 8.6 alleles per locus were observed. Except for GMS004 all other barley microsatellite markers showed on average a high value of gene diversity ranging from 0.64 to 0.88. The mean value of gene diversity in the wild forms and landraces was 0.74, and even among the cultivars the gene diversity ranged from 0.30 to 0.86 with a mean of 0.72. No significant differences in polymorphism were detected by the GA and GT microsatellite markers. The estimated genetic distances revealed by the microsatellite markers were, on average , 0.75 for the wild forms, 0.72 for landraces and 0.70 among cultivars. The microsatellite markers were able to distinguish between different barley genotypes. The high degree of polymorphisms of microsatellite markers allows a rapid and efficient identification of barley genotypes.

Key words Microsatellite markers · Barley · Gene diversity · Genetic distances

Communicated by G. Wenzel

D. Struss (✉) · J. Plieske
Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstrasse 3, 06466 Gatersleben, Germany

Introduction

Simple-sequence repeats (SSRs), called microsatellites, are randomly interspersed in eukaryotic genomes. They are highly variable in the number of repeats they contain and are co-dominantly inherited (Johansson et al. 1992). The variation can be detected by primers flanking the microsatellite locus in the conserved DNA sequences. The microsatellites are very often multiallelic and the number of alleles representing a microsatellite locus is highly variable, e.g. for the SAT1 locus in soybean 26 alleles were found (Rongwen et al. 1995).

Microsatellites have been increasingly used as molecular markers. Their polymorphisms have shown high efficiency for many studies. They can be used for genomic mapping, population and evolutionary studies, as well as for fingerprinting and pedigree analyses (Hazan et al. 1992; Plaschke et al. 1995; Rongwen et al. 1995; Guilford et al. 1997). Microsatellites were found to be more polymorphic than RFLP markers in soybean, rice and *Arabidopsis* (Akkaya et al. 1992; Wu and Tanksley 1993; Bell and Ecker 1994).

Barley is a genetically well-studied crop. Cultivated barley, *Hordeum vulgare*, including its wild relative *H. vulgare* ssp. *spontaneum*, belongs to the primary gene pool of *Hordeum*. The total number of barley accessions in the genebanks, including redundant material, is estimated to be about 280 000 (Plucknett et al. 1987). It is the intensive collection of landraces and the creation of genetic stocks that has led to the accumulation of such a large number of accessions in genebanks and breeder collections. Studies of the extent of genetic diversity in barley have been performed using RFLP markers (Graner et al. 1994; Melchinger et al. 1994). Microsatellites as a marker system have a high potential for use in the genome and genetic analyses of self-pollinating crops like wheat because of their high degree of polymorphism (Plaschke et al. 1995; Röder et al. 1995). Screening and evaluating the available

genetic diversity in barley germplasm could both optimize and facilitate the breeding process. The assessment of genetic diversity in barley germplasm has previously been based on pedigree records, phenotypic traits, biochemical and DNA markers (Tolbert et al. 1979; Doebley 1989; Melchinger et al. 1994).

The objectives of the present study were: (1) to develop a set of microsatellite markers, (2) to determine their variability, and (3) to use them for genome analysis, distinguishing between genotypes, and the estimation of genetic diversity in a barley germplasm collection comprising wild forms, landraces and cultivars.

Material and methods

Plant material and DNA-extraction

Three sets of materials comprising 163 different barley lines from various barley growing regions were provided by the genebank of IPK Gatersleben. These include a set of 46 wild forms (*H. vulgare* ssp. *spontaneum* and *H. vulgare* ssp. *agriocrithon*), a second set of 52 landraces, and a third set consisting of 65 mainly old cultivars and landraces which have been selected for breeding purposes, and referred to here as cultivars/landraces (a detailed list of the surveyed material will be provided upon request). The selection of accessions was done using a genebank based on the highest possible morphological and geographical diversity.

Approximately 10 g of leaves from 20 plants of each accession were pooled for the extraction of genomic DNA. The leaves were frozen in liquid N_2, ground in 25 ml of pre-heated extraction buffer (100 mM TrisHCl pH 8, 500 mM NaCl, 50 mM EDTA and 1.25% SDS) and incubated at 60°C for 30 min. An equal volume of chloroform/isoamyl alcohol (24 : 1) was added and mixed well. The mix was centrifuged at 3000 g for 15 min. The supernatant was precipitated with an equal volume of ice-cold ethanol. The precipitate was hooked out by sterile pipettes, washed in 70% ethanol and air-dried. The DNA was suspended in 3 ml of 1 × TE buffer, followed by another chloroform/isoamylcohol extraction. The mix was centrifuged at 3000 g for 15 min. The supernatant was transferred into a new tube and incubated with a 1/10 vol of 3 M sodium acetate pH 5.2, and 2 vol of 96% ethanol for 1 h at −20°C. The DNA was recovered by centrifugation at 3000 g for 5 min. The pellet was washed in 70% ethanol, air-dried and suspended in 1 ml of 1 × TE buffer.

Isolation of microsatellites

The isolation of microsatellites was accomplished according to Plaschke et al. (1995). A phage library was constructed with genomic DNA of the barley cultivar Shin Ebisu 16. The genomic DNA was digested by the methyl-sensitive restriction enzyme *Pst*I. The low-molecular-weight fraction was isolated from an agarose gel, followed by subsequent digestion with *Mbo*I, and cloned into the *Bam*HI site of the lambda phage vector ZAP Express (Stratagene). The library was screened for the presence of repeats by hybridization with $(GA)_n$ and $(GT)_n$ which were labelled with ^{32}P-γATP.

Primer design and PCR

Primers flanking microsatellite loci were designed using the computer program PRIMER 0.5 (E. Lander, Cambridge, Mass., USA). The forward primers were labelled with 5′ fluorescein. PCR was carried out under the following conditions: 100–150 ng of template DNA, 250 nM of each primer, 200 µM of dNTPs, 1 U of *Taq* Polymerase, 1.5–2 mM of $MgCl_2$. The reaction, depending on the primer pair, was run for 35–45 cycles (denaturing at 94°C for 1 min, annealing at 55°C or 60°C for 1 min, with a 2-min extension at 72°C), followed by a single extension at 72°C for 60 min. The amplification products were detected on polyacrylamide gels by 4 µl of loading buffer (formamide containing 0.5% w/v of bromophenol blue) and 1.5 µl of PCR products.

Data analysis

Fragments amplified by microsatellite primers were scored as present (1) or absent (0). Genetic distances were estimated according to Nei and Li (1979) based on the probability that the amplified fragment from one genotype will be present in another genotype, GD = 1–2a/2a + b + c, (a: number of shared fragments, b: number of fragments in line A, and c: number of fragments in line B).

Cluster analysis was performed with the Syntax 5.02 (Podani 1990) package based on average linkage (UPGMA, unweighted pair-group method with arithmetic average).

Gene diversity (heterozygosity) was calculated according to Weir (1990). Thus:

$$\text{gene diversity} = 1 - \sum P_{ij}^2,$$

P_{ij} is the frequency of the jth pattern for microsatellite marker i and is summed across n patterns. Anderson et al. (1993) suggest that gene diversity is the same as the polymorphism information content (PIC).

Results

Isolation of microsatellite sequences in barley

The frequencies of dinucleotide repeats in the barley genome was estimated among 39 000 plaques from a lambda genomic library by hybridization under stringent conditions with the dinucleotide repeats $(GA)_n$ and $(GT)_n$. A total of 482 clones (1.2%) hybridized with $(GA)_n$ and 285 clones (0.7%) hybridized with $(GT)_n$ (Table 1). Only a few AT and trinucleotide GAA sequences were observed. The average repeat number was 20 for GA loci and 13 for GT loci. The number of repeats ranged from 2 to 58 for GA and from 2 to 24 for GT (Table 1). While 68% of the detected GA sequences were perfect repeats, the majority (65%) of GT repeats were imperfect.

Characterization of selected microsatellite markers

From a large number of microsatellite sequences isolated by screening a lambda-phage library, 15 were

Table 1 Frequency of microsatellite sequences in the genome of barley cultivar Shin Ebisu 16

Repeats	Positive plaque numbers (%)	Number of repeats	
		Average	Range
GA	485 (1.2)	20	2–58
GT	285 (0.7)	13	2–24

Table 2 Primer sequences, fragment size and repeat types of 15 barley microsatellite markers

Markers	Primer sequences (5′ → 3′)	Annealing temperature °C	Predicted product size in bp	Repeats
GMS001	CTGACCCTTTGCTTAACATGC TCAGCGTGACAAACAATAAAGG	60	134	$(CT)_7TTT(CT)_2$
GMS002	CCGACAACATGCTATGAAGC CTGCAGCAAATACCCATGTG	60	131	$(CT)_{18}$
GMS003	TTTCAGCATCACACGAAAGC TTGCATGCATGCATACCC	60	144	$(GT)_{15}$
GMS004	AGCGCAACCTACACAGCC ATTGCCTTCTAATTTCTGCAGG	60	138	$(CA)_7T(CA)_{10}$
GMS005	ACTACGTCCAGTCGTTTCC TGAACACCACGGGTTCATC	55	170	$(GT)_{19}T(CT)_2(GT)_{13}$
GMS006	TGACCAGTAGGGGCAGTTTC TTCTTCTCCCTCCCCCAC	55	154	$(GA)_2ATA(GA)_{19}$
GMS021	CTATCACACGACGCAACATG CCTGAGAAAGAAAGCGCAAC	55	169	$(GA)_{17}(GA)_7$
GMS027	CTTTTTCTTTGACGATGCACC TGAGTTTGTGAGAACTGGATGG	60	154	$(GT)_5(CT)_2(GT)_{27}$
GMS032	ATTTGACGAGAAGGACCGG CACTAAAGGGAACAAAAGCTGG	60	166	$(CT)_{13}(GT)_3GA$ $(AT)_3(CT)_3$
GMS046	ATGTATTTATCACCCACCCAGC AAGGCATTAGAACCGGCAC	60	156	$(GA)_{13}$
GMS056	GAGAAACGCAGCTGTGGC GTCACCGAGGCCTTCCTC	60	137	$(GA)_{11}$
GMS061	CACCTGTTCCGTCCCGTC AACCTCTTTTTTATCCCTCGC	60	145	$(GA)_{16}$
GMS089	TGAAGTGGAAGGCTTCGC GCTCTCGTTGTGCGGAG	55	154	$(GA)_{21}$
GMS114	AACCAGTGGGTTTTAACCCC TGCCACCACATGCATACAC	55	152	$(GT)_{11}$
GMS116	GAAAGACTGACAGGCGGAAG TTTCTTTGTTGTGTGTGCAGTG	55	138	$(CA)_{16}CTG(CA)_4$

selected to profile the genetic diversity among the barley genotypes (Table 2). The corresponding primer pairs were tested for amplification of the expected fragments in the donor cultivar Shin Ebisu 16. The amplified fragments ranged from 131 to 170 bp. All primers were also tested for their efficiency to reveal polymorphisms in different barley genotypes. Of the 15 microsatellites, eight were perfect repeats, four were imperfect repeats and three showed a compound composition (Table 2).

PCR reactions were subsequently performed with 163 barley accessions. All primer pairs amplified the expected fragments with a variation of $\pm$ 10 base pairs in all 163 accessions. The amplification products were highly polymorphic in the material surveyed. A total of 130 alleles was detected by the 15 barley microsatellite markers. The alleles of marker GMS006 with 38 different barley genotypes are shown in Fig. 1. The size of the 14 alleles ranged from 128 to 153 bp. The number of alleles per microsatellite marker varied from 5 to 15 (Table 3). On average, 8.6 alleles per locus were observed. The length of alleles in the 163 surveyed genotypes ranged from 113 to 202 bp. Except for GMS004, all other barley microsatellite markers showed a high value of gene diversity ranging from 0.64 to 0.88 (Table 3). The mean value of gene diversity in wild forms and landraces was 0.74. Even among the cultivars the gene diversity ranged from 0.30 to 0.86 with a mean of 0.72. There was no significant difference in polymorphism between GA and GT microsatellite markers.

Efficiency of microsatellite markers for identification of barley lines

To assess the efficiency of microsatellite markers for the evaluation of barley genotypes and to characterize the barley accessions of the genebank of IPK Gatersleben, estimates of genetic distances were obtained for all genotypes of three groups, 46 wild forms, 52 landraces and 65 old barley cultivars, using 130 polymorphisms detected by 15 barley microsatellites. The estimated genetic distances revealed by microsatellite markers were on average 0.75 among the wild forms, 0.72 in landraces and 0.70 among cultivars.

The lowest genetic distance was observed between two wild forms of *H. vulgare* ssp. *spontaneum* (AHOR 4862 and AHOR 4863) from the same origin (Turkmenistan). The lowest genetic distances in the other two groups were also found between accessions from the same geographical region. Genetic relationships among the accessions were further studied by cluster analysis.

The dendrogram discriminated between all 163 barley genotypes. In the sets of landraces and wild forms most of the subgroups were related by their origins (Fig. 2).

Discussion

The di-nucleotide microsatellites isolated from the barley genome occurred at a frequency comparable with other crops. GA repeats occur in barley at a higher frequency than GT repeats, confirming the data obtained with rice (Wu and Tanksley 1993), apple (Guilford et al. 1997) and *Brassica napus* (Lagercrantz et al. 1993; Szewc-McFadden et al. 1996). Becker and Heun (1995) were unable to find any predominance of GA sites, or other repeat sites, by searching the database of barley sequences in Genebank and EMBL. By contrast, in mammals the CA motif predominates (Beckmann and Weber 1992). The occurrence of only a few AT motifs in barley was consistent with the data of Röder et al. (1995), who failed to find any AT repeats in wheat. Similar data were reported in *Arabidopsis* (Bell and Ecker 1994). The repeat numbers of GA motifs were higher than those of GT, as was reported by Plaschke et al. (1995) in wheat.

On average, $(GA)_n$ repeats detected more alleles (10.25) than the GT motif (eight). In contrast to the finding of Saghai Maroof et al. (1994), the number of alleles detected at a locus did not correspond to the number of nucleotide repeats in the microsatellite sequences. The highest number of alleles was detected by

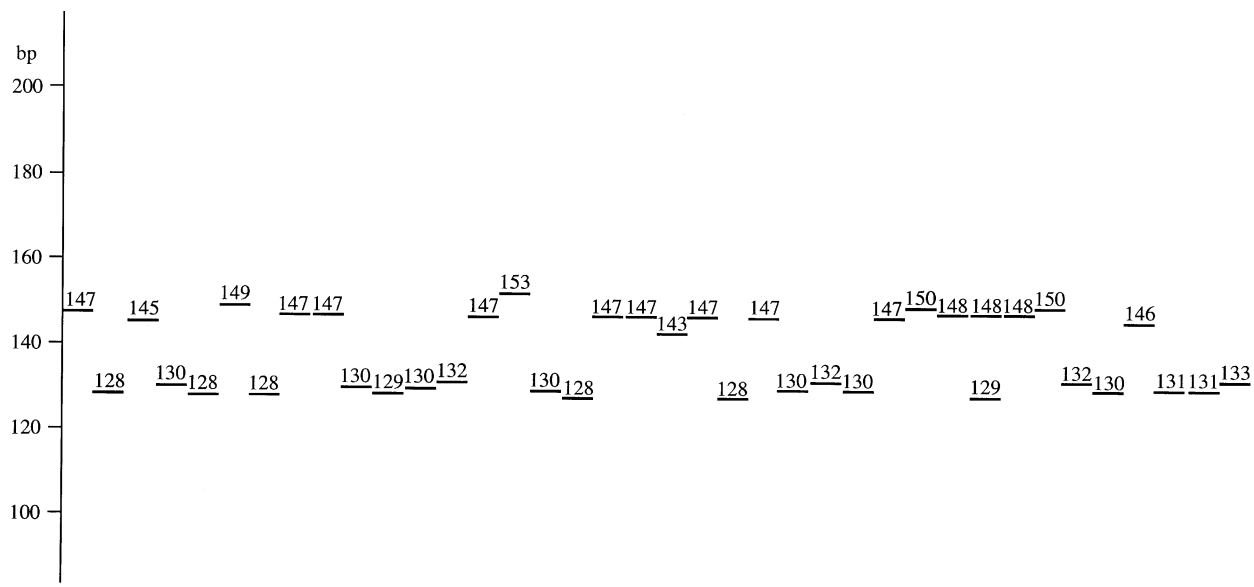

Fig. 1 PCR products of microsatellite marker GMS006 showing 14 alleles among 38 different barley genotypes. The allele length is in bp

Table 3 Gene diversity, number of alleles and allele length of 15 microsatellite markers assayed in 163 barley genotypes

Markers	Gene diversity				Number of alleles	Allele length in bp
	Wild forms	Landraces	Cultivars	Average		
GMS001	0.76	0.72	0.74	0.74	5	123–137
GMS002	0.78	0.71	0.64	0.71	6	119–137
GMS003	0.85	0.81	0.80	0.82	8	124–152
GMS004	0.31	0.44	0.30	0.35	5	136–216
GMS005	0.61	0.77	0.83	0.73	10	152–205
GMS006	0.59	0.62	0.74	0.65	14	134–162
GMS021	0.75	0.70	0.76	0.74	6	150–176
GMS027	0.91	0.87	0.86	0.88	5	147–160
GMS032	0.67	0.70	0.56	0.64	9	113–158
GMS046	0.79	0.63	0.58	0.67	15	113–185
GMS056	0.90	0.87	0.86	0.88	13	127–188
GMS061	0.75	0.83	0.80	0.79	8	139–160
GMS089	0.90	0.77	0.84	0.84	8	124–142
GMS114	0.74	0.79	0.76	0.77	11	114–159
GMS116	0.86	0.86	0.79	0.83	9	149–202

microsatellite marker GMS046 (15) with 13 GA repeats, whereas the GMS002 with 18 GT repeats detected only six alleles. Rongwen et al. (1995), Szewc-McFadden et al. (1996) and Guilford et al. (1997) working with apple, *Brassica* and soybean, respectively, also found no correlation between the number of repeats in microsatellite loci and the number of alleles detected at each locus.

The microsatellite markers were able to distinguish between different barley genotypes. The high degree of polymorphism of microsatellite markers allows a rapid and efficient identification of barley genotypes. The high gene-diversity value of barley microsatellite markers (with an average of over 0.73) makes them ideal markers for differentiating between barley

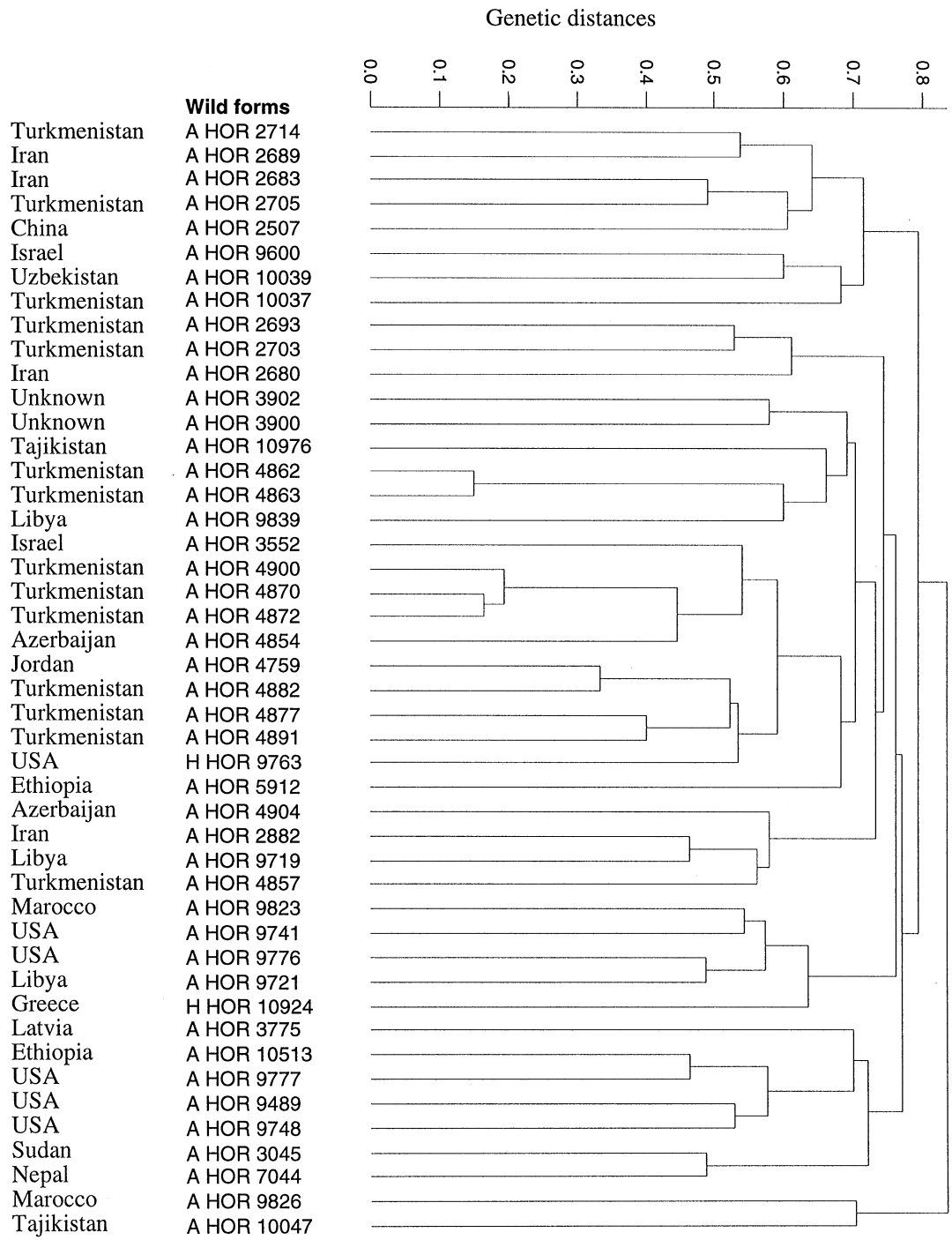

Fig. 2 Cluster analysis of 163 different barley genotypes using 15 microsatellite markers. The distance matrix was estimated by the average linkage method. *I* wild forms, *II* landraces and *III* cultivars/landraces. *A* = spring form, *H* = winter form

genotypes. Similarly high values of gene diversity were reported for other plant species (Saghai Maroof et al. 1994; Rongwen et al. 1995; Guilford 1997). We failed to find any differences between the gene diversity of microsatellites with a GA composition and those with GT.

The number of alleles detected by microsatellite markers did not correlate with the gene diversity. For

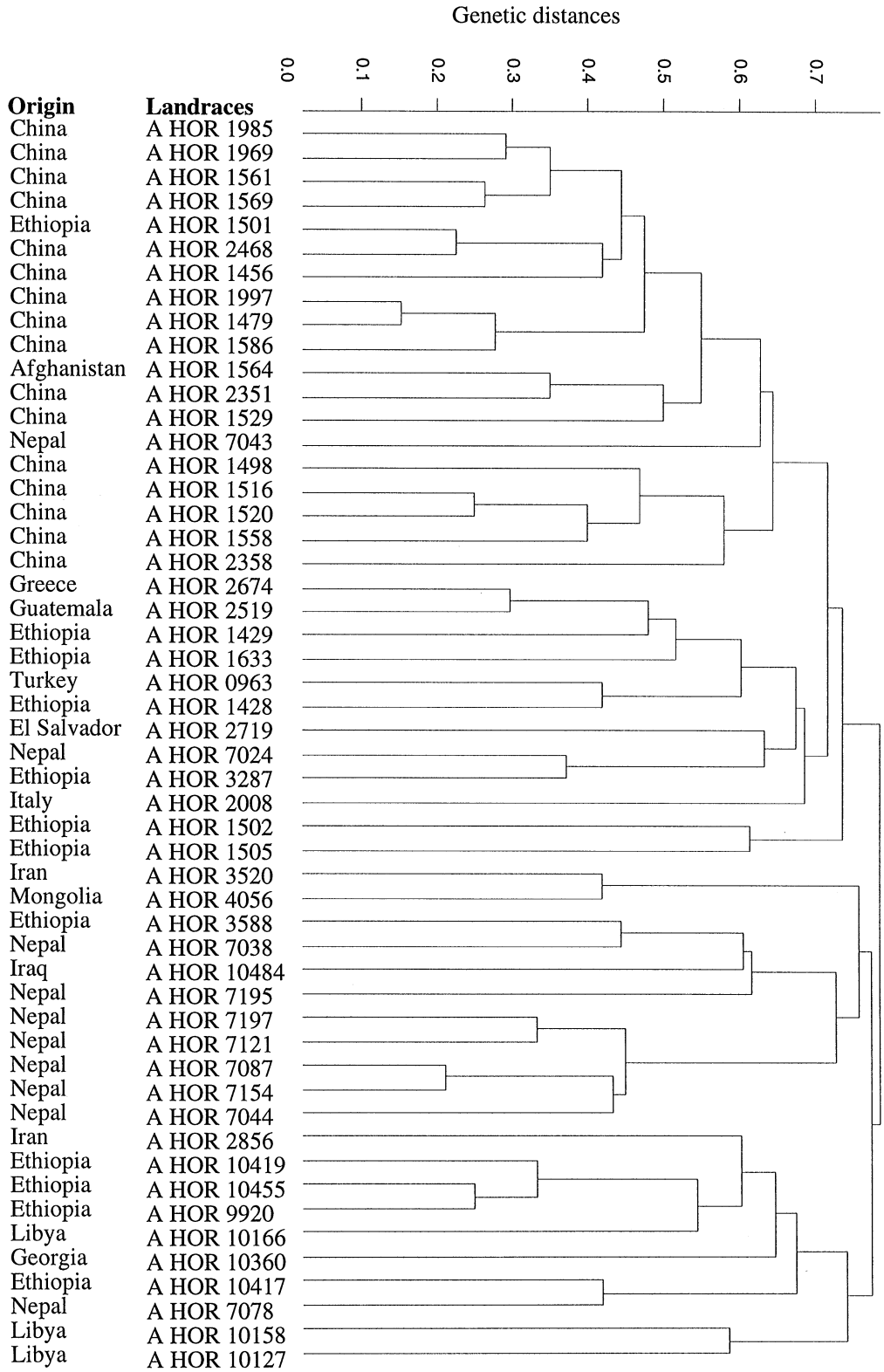

Fig. 2 See page 312 for legend

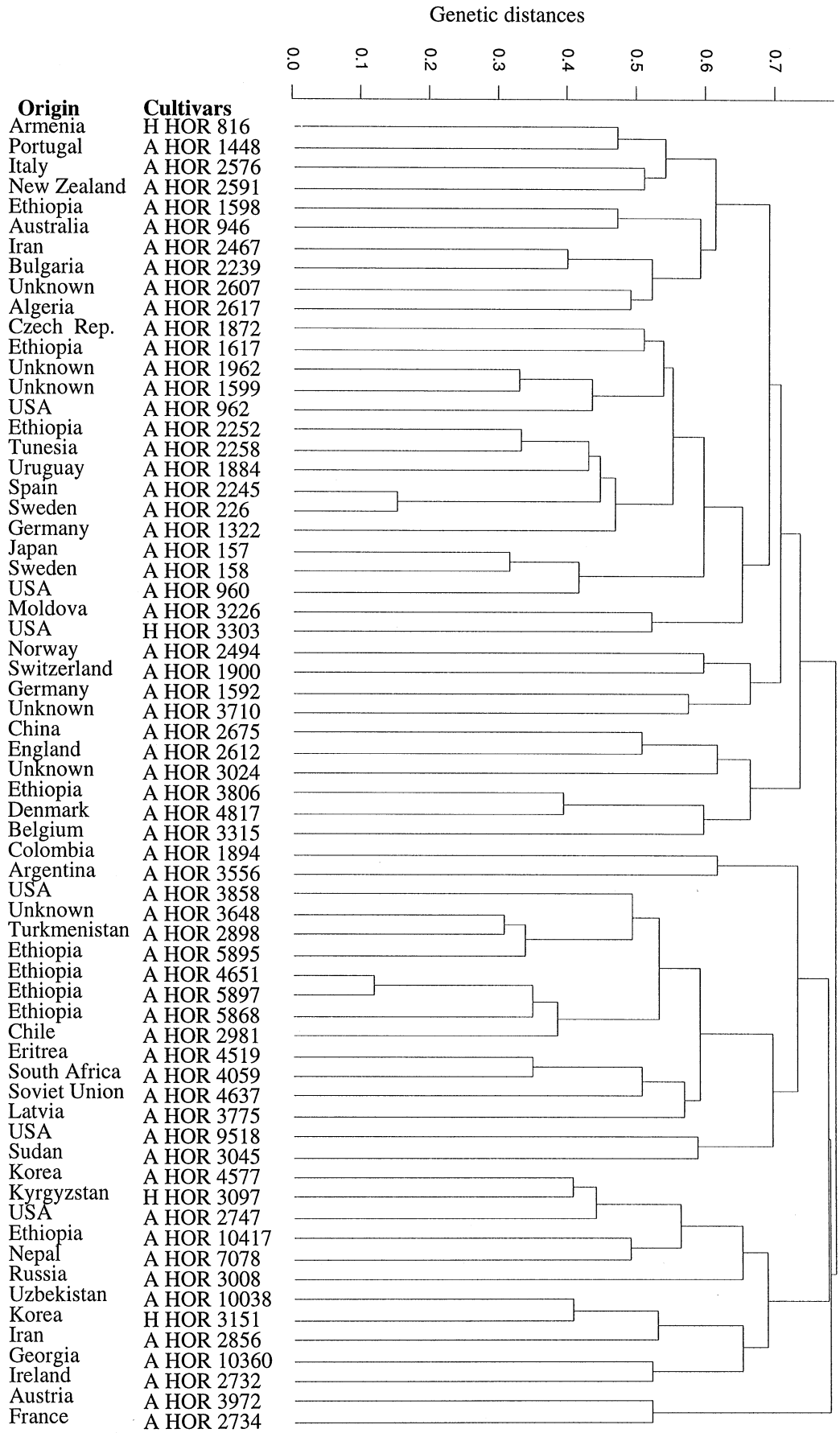

Fig. 2 See page 312 for legend

example, the marker GMS046 showed an average gene diversity of 0.67, detecting 15 alleles within all three barley genotype sets, whereas GMS001 detected five alleles and showed a gene diversity value of 0.74. Obviously, most of the alleles detected by GMS046 were rare alleles with a low frequency. GMS046 exhibited a higher gene diversity (0.79) among the wild forms, corresponding to a higher probability of occurrence of rare alleles than among the cultivars/landraces (0.58).

Graner et al. (1990) observed a relatively low RFLP diversity within advanced barley germplasm. The high degree of polymorphisms in our microsatellite markers is not only due to the materials tested but is also based on the hypervariable character of microsatellites. We found more or less the same diversity in wild forms as in advanced group cultivars/landraces.

The cluster analysis could clearly identify the genetic relationship of the barley genotypes examined in each set and demonstrates the potential and ability of microsatellite markers for genome analysis in barley. The lowest genetic distances have been estimated between genotypes derived from the same origin in all three populations. In the landrace and wild form populations, most of the classified groups consist of genotypes from the same origin. In the cultivar/landrace population no correlation between geographical origin and classification was observed. The intensive exchange of seed material may have caused a mis-alignment of accessions to a particular origin and very often it is not possible to identify the real origin of cultivars or advanced landraces. The examined wild forms, landraces and cultivars/landraces were represented by single homozygous lines. It can be assumed that these lines were not always identical to the original landraces or wild forms. Our intention was to provide a molecular tool to identify and characterize barley accessions, estimating the genetic diversity, and providing a set of genetically defined barley materials for experimental and breeding purposes. Undoubtedly, a higher precision for monitoring the extent of genetic diversity of the barley collection can be achieved by using more markers. Nevertheless, the set of microsatellite markers employed in this study demonstrates the usefulness of microsatellite markers for the identification of barley genotypes. They provide an optimal system with which to increase the efficiency of germplasm evaluation and to identify duplicate accessions in the barley germ-plasm collection.

Acknowledgments We thank H. Knüpfer for providing the barley genotypes, I. Schubert, A. Graner and K. Bachmann for constructive reviewing of the manuscript. We also thank Margit Hantschmann for excellent assistance and Susanne König and Andreas Polley for sequencing the clones.

References

Akkaya SM, Bhagwat AA, Cregan PB (1992) Length polymorphism of simple-sequence repeat DNA in soybean. Genetics 132: 1131–1139

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

Becker J, Heun M (1995) Barley microsatellites: allele variation and mapping. Plant Mol Biol 27: 835–845

Beckmann JS, Weber JL (1992) Survey of human and rat microsatellites. Genomics 12: 627–631

Bell CJ, Ecker JH (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19: 137–144

Doebley J (1989) Isozyme evidence and the evolution of crop plants. In: Soltis DE, Soltis, PS (eds) Isozymes in plant biology. Chapman and Hall, London, (pp 165–191)

Graner A, Siedler H, Jahoor A, Herrmann RG, Wenzel G (1990) Assessment of the degree and the type of restriction fragment length polymorphism in barley (*Hordeum vulgare*). Theor Appl Genet 80: 826–832

Graner A, Ludwig WF, Melchinger AE (1994) Relationship among European barley germplasm. II. Comparison of RFLP and pedigree data. Crop Sci 34: 1199–1205

Guilford P, Prakash S, Zhu JM, Rikkerink E, Gardiner S, Bassett H, Forster R (1997) Microsatellites in *Malus* × *domestica* (apple): abundance, polymorphism and cultivar identification. Theor Appl Genet 94: 249–245

Hazan J, Dubay C, Pankowiak MC, Becuwe N, Weissbach J (1992) A genetic linkage map of human chromosome 20 composed entirely of microsatellite markers. Genomics 12: 183–189

Johansson M, Ellegren H, Andersson L (1992) Cloning and characterization of highly polymorphic porcine microsatellites. J Hered 83: 196–198

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

Melchinger AE, Graner A, Singh M, Messmer MM (1994) Relationship among European barley germplasm. I. Genetic diversity among winter and spring cultivars revealed by RFLPs. Crop Sci: 34: 1191–1199

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

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

Plucknett DL, Smith NJH, Williams JT, Murthi Anishetty N (1987) Genebanks and the world's food. Princeton University Press, Princeton, New Jersey, USA

Podani J (1990) SYN_TAX-pc-supplement 3: Macintosh version. Abstr Bot 14: 23–29

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

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Gregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor Appl Genet 90: 43–48

Saghai Maroof MA, Biyashev RM, Yang GP, Zhang Q, Allard RW (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc Natl Acad Sci USA 91: 5466–5470

Szewc-McFadden AK, Kresovich S, Bliek SM, Mitchell SE, McFerson JR (1996) Identification of polymorphic, conserved simple-sequence repeats (SSRs) in cultivated *Brassica* species. Theor Appl Genet 93: 534–538

Tolbert DM, Qualset CO, Jain SK, Craddock JC (1979) A diversity analysis of a world collection of barley. Crop Sci 19: 789–794

Weir BS (1990) Genetic-data analysis methods for discrete genetic data. Sinauer Assoc Inc, Sunderland, Massachusetts, USA

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241: 225–235