© NIAB 2008
ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization **6**(2); 167–174
doi:10.1017/S1479262108993187

SSR and SNP diversity in a barley germplasm collection

Rajeev K. Varshney[1*†], Khaled F. M. Salem[1‡], Michael Baum[2], Marion S. Roder[1], Andreas Graner[1] and Andreas Börner[1]

[1] *Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstrasse 3, Gatersleben D-06466, Germany and* [2] *International Centre for Agricultural Research in Dryland Areas (ICARDA), PO Box 5466, Aleppo, Syria*

Received 30 August 2007; Accepted 3 January 2008; First published online 14 May 2008

Abstract

Sets of microsatellites extracted from both a genomic library (gSSRs) and from expressed sequence tag sequence (eSSRs), and single nucleotide polymorphisms (SNPs) were applied to assess the levels of genetic diversity in a sample of 70 barley accessions, originating from 28 countries in Asia, Africa, the Middle East and Europe. The eSSR assays detected a mean of 9.5 alleles per locus, and the gSSRs only 5.7 alleles per locus, but the polymorphism information content values for the two assay types were indistinguishable. Strong and statistically significant correlations were observed between the eSSR and gSSR ($r = 0.86$, $P < 0.05$), the eSSR and SNP ($r = 0.74$, $P < 0.05$) and the gSSR and SNP genotypes ($r = 0.67$, $P < 0.05$). Accessions originating from the Middle East and Asia had the highest levels of genetic diversity. Pairwise genetic similarity ranged from 0.16 to 0.87 (mean 0.43), indicating that the sample was genetically diverse. When clustered on the basis of genotype, Asian and African accessions tended to be grouped together, but those originating from the Middle East were not concentrated in any particular cluster.

Keywords: barley; eSSR; gSSR; molecular markers; SNP

Introduction

The detection and characterization of genetic variation present in germplasm collections is important for plant breeding programmes seeking to widen the genetic base of breeding populations. Barley (*Hordeum vulgare* L.) is globally the fourth most important cereal crop. It is grown across a wide range of temperate and semi-arid environments, primarily for animal feed and as a feedstock for beer production. A major public barley germplasm collection is curated by International Center for Agricultural Research in the Dry Areas (ICARDA), but the molecular diversity (commonly measured using one or a combination of molecular marker systems) present in this collection has been rather poorly characterized to date.

Microsatellites (or simple sequence repeats, SSRs) are distributed throughout eukaryotic genomes. Some are present within coding sequence, but the majority, particularly in large genome species such as barley, is located with the non-coding fraction of the genome. A number of attempts have been made to use SSRs to measure genetic diversity in barley, and most have exploited allelic variation in microsatellites derived from genomic DNA (gSSRs) (Liu *et al.*, 1996; Russell *et al.*, 1997; Struss and Plieske, 1998). Few gSSRs have been characterized as to whether or not they assay a stretch of transcribed DNA. However, the wide availability of extensive amounts of barley expressed sequence tag (EST) sequence has now allowed for the systematic

* Corresponding author. E-mail: r.k.varshney@cgiar.org
† International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Andhra Pradesh 502 324, India.
‡ Genetic Engineering and Biotechnology Research Institute (GEBRI), Menoufia University, PO Box 79, Sadat, Egypt.

development of both expressed SSRs (eSSRs; Varshney *et al.*, 2005) and genic single nucleotide polymorphisms (SNPs; Kota *et al.*, 2001a). Because the latter generate a simple binary output, which is well suited to automatic data collection systems, their use is gaining momentum (Kota *et al.*, 2001a, b; Kanazin *et al.*, 2002; Russell *et al.*, 2004; Rostoks *et al.*, 2006). However, SNP genotyping in large germplasm collections is a capital-intensive process, unless the markers can be converted into cleaved amplified polymorphic sequence (CAPS) assays (Kota *et al.*, 2007; Varshney *et al.*, 2007a, c).

The present study was undertaken with the following objectives: (1) to generate a comparative assessment of the potential of gSSRs, eSSRs and SNPs to estimate genetic diversity; (2) to investigate the relationship between genetic diversity and the provenance; (3) to analyse the genetic structure of the ICARDA barley germplasm collection.

Material and methods

Genotypes and genotyping

The sample set of the ICARDA barley collection (containing 223 genotypes) consisting of 70 accessions was chosen to maximize diversity on the basis of provenance. In all, seven African (21 accessions), ten Asian (21 accessions), nine Middle Eastern (26 lines) and two European (two accessions) countries were represented (further details given in Table 1). DNA was isolated as described by Thiel *et al.* (2003).

The set of 20 gSSRs was selected from those developed by Liu *et al.* (1996), Struss and Plieske (1998), Ramsay *et al.* (2000) and Li *et al.* (2003) based on high-quality amplification profile and distribution on all the linkage groups. PCR and fragment detection followed methods described by Röder *et al.* (1998). The set of 20 eSSRs was assembled from those developed by Thiel *et al.* (2003) and Varshney *et al.* (2006), and form part of the core set defined by Varshney *et al.* (2007c). PCR conditions for the eSSRs followed Thiel *et al.* (2003), except that fluorescent-dye-labelled primer pairs were used for amplification, the amplicons were separated on an ABI377 device and alleles were scored using GenoTyper 3.7 (Applied Biosystems). SNP primer pairs were selected from those derived by Kota *et al.* (2007), and PCR conditions followed Kota *et al.* (2001a). A CAPS assay was applied to 16 out of the 20 SNPs, while the remaining four were assayed using pyrosequencing (for experimental details of these assays, see Varshney *et al.*, 2007c). The 60 marker loci were distributed across all seven linkage groups of barley.

Diversity analysis

While SSR allelic data were scored using GenoTyper, the SNP marker profiles were scored manually: each allele was scored as present (1) or absent (0) at each SSR and SNP locus. Polymorphism information content (PIC) values were calculated following Anderson *et al.* (1993).

Genetic similarities (GSs) were calculated in the NTSYS-pc 2.11 software package (Biostatistics Inc., USA: Rohlf, 1998) for each marker pair, using Dice's similarity coefficient. Sequential, agglomerative, hierarchical and nested clustering was employed for the construction of unweighted pair group method with arithmetic mean (UPGMA) dendrograms. The correlations between matrices were calculated using the Mantel (1967) test, employing 10,000 random iterations in the non-parametric test calculator (Mantel version 2.0, Liedloff, 1999).

Results

Genotypic variation

Allele number among the gSSR loci ranged from four (Bmac0040) to eight (Ebmac0788), giving a mean of 5.7 alleles per locus. The PIC values ranged from 0.59 (GBMS192) to 0.82 (Ebmac0788) (mean 0.74 ± 0.06) (Table 2). The 20 eSSR markers generated a mean of 9.5 alleles per locus, with allele number varied between 4 (GBM1212) and 17 (GBM1461). The mean eSSR PIC value was 0.73 ± 0.08 [ranging from 0.56 (GBM1029) to 0.88 (GBM1461)] (Table 3). Among the SNP markers provided a total of 23 SNP datapoints as two SNP markers namely GBS0461 and GBS0576 in pyrosequencing assays yielded two and three SNP datapoints, respectively. The PIC values of assayed SNPs was observed in the range of 0.09 (GBS0136) to a maximum of 0.50 (GBS0526) with an average of 0.39 ± 0.10 per SNP (Table 4).

Genetic diversity

Accessions derived from the Middle East were characterized by the highest PIC values within each marker type (Fig. 1). To measure the comparative diversity in genotypes of different geographic regions, the numbers of alleles at a particular microsatellite locus were counted in each of four geographic regions and these were relatively higher in the Asian region (Fig. 2). The eSSRs tended to display more alleles per locus than did the gSSRs across all the geographical regions (Fig. 2). However, the eSSR and gSSR PIC values were indistinguishable from one another within each geographical region (except Europe). The SNP loci had lower PIC values than the SSRs. When the gSSR, eSSR and SNP GS

Table 1. Provenance and classification of the barley accessions sampled from the ICARDA collection

S. No.	ICARDA accession number	Description of genotype[a]	Country of origin	Provenance
Asia				
1	IG20900	*Hordeum vulgare* convar. *distichon*	Afghanistan	Samangan
2	IG20905	*H. vulgare* convar. *vulgare*	Afghanistan	Herat
3	IG27683	*H. vulgare* convar. *vulgare*	Afghanistan	Jowzjan
4	IG24745	*H. vulgare* convar. *vulgare*	Azerbaijan	Baku
5	IG125827	*H. vulgare* convar. *vulgare*	Azerbaijan	Naxcivan
6	IG33608	*H. vulgare* convar. *vulgare*	China	Gansu
7	IG33649	*H. vulgare* convar. *vulgare*	China	Jiangsu
8	IG138244	cv (Mari/Aths × 2)	Cyprus	–
9	IG25710	*H. vulgare* convar. *distichon*	India	Gujarat
10	IG19453	*H. vulgare* convar. *vulgare*	Iran	Khuzestan
11	IG27649	*H. vulgare* convar. *vulgare*	Iran	Yazd
12	IG27653	*H. vulgare* convar. *vulgare*	Iran	Fars
13	IG27794	*H. vulgare* convar. *vulgare*	Iran	Khorasan
14	IG113120	*H. vulgare* convar. *vulgare*	Iran	Kerman
15	IG27630	*H. vulgare* convar. *vulgare*	Pakistan	Balochistan
16	IG32608	*H. vulgare* convar. *vulgare*	Pakistan	Punjab
17	IG36058	*H. vulgare* convar. *distichon*	Tajikistan	Khudzhand
18	IG16981	*H. vulgare* convar. *vulgare*	Turkmenistan	Ashgabat
19	IG135624	*Hordeum spontaneum*	Turkmenistan	Kazanjik
20	IG123901	*H. vulgare* convar. *vulgare*	Uzbekistan	Andizhan
21	IG123923	*H. vulgare* convar. *vulgare*	Uzbekistan	Fergana
Africa				
22	IG24634	*H. vulgare* convar. *vulgare*	Algeria	Batna
23	IG35385	*H. vulgare* convar. *vulgare*	Algeria	Saida
24	IG138230	cv (Tichedrett)	Algeria	–
25	IG25839	*H. vulgare* convar. *vulgare*	Egypt	Alexandria
26	IG32475	*H. vulgare* convar. *distichon*	Egypt	North Sinai
27	IG32812	*H. vulgare* convar. *distichon*	Egypt	Marsa Matruh
28	IG138243	cv (CalM)	Egypt	–
29	IG138263	cv (Demhay)	Eritrea	–
30	IG23515	*H. vulgare* convar. *vulgare*	Ethiopia	Tigray
31	IG138262	cv (Shege)	Ethiopia	–
32	IG24953	*H. vulgare* convar. *vulgare*	Libya	Tarabulus
33	IG37576	*H. vulgare* convar. *vulgare*	Libya	Birak
34	IG138254	cv (Barjouj)	Libya	–
35	IG31876	*H. vulgare* convar. *vulgare*	Morocco	Tiznit
36	IG31938	*H. vulgare* convar. *vulgare*	Morocco	Bouarfa
37	IG138235	cv (Arig8)	Morocco	–
38	IG138221	Moroc 9-75	Morocco	–
39	IG37726	*H. vulgare* convar. *vulgare*	Tunisia	Sidi Bouzid
40	IG37784	*H. vulgare* convar. *vulgare*	Tunisia	Medenine
41	IG138228	cv (Martin)	Tunisia	–
42	IG138255	cv (M126/CM67//As/Pro/3/Alanda)	Tunisia	–
Middle East				
43	IG108911	*H. vulgare* convar. *vulgare*	Iraq	Ar Ramadi
44	IG138234	cv (IPA7)	Iraq	–
45	IG107427	*H. spontaneum*	Iraq	Mosul
46	IG128199	*H. vulgare* convar. *vulgare*	Jordan	Zarqa
47	IG128202	*H. vulgare* convar. *distichon*	Jordan	Ma'an
48	IG138236	cv (Rum)	Jordan	–
49	IG40022	*H. spontaneum*	Jordan	Mafraq
50	IG40039	*H. spontaneum*	Jordan	Karak
51	IG138257	cv (Litani)	Lebanon	–
52	IG32814	*H. vulgare* convar. *distichon*	Oman	North Batinah
53	IG32954	*H. vulgare* convar. *vulgare*	Oman	Interior
54	IG32971	*H. vulgare* convar. *vulgare*	Oman	Western Hajar
55	IG39126	*H. spontaneum*	Palestine	Hadarom
56	IG138211	*H. spontaneum* (H.spont.41-1)	Palestine	–
57	IG27892	*H. vulgare* convar. *distichon*	Saudi Arabia	Asir

Table 1. *Continued*

S. No.	ICARDA accession number	Description of genotype[a]	Country of origin	Provenance
58	IG113084	*H. vulgare* convar. *vulgare*	Saudi Arabia	Taif
59	IG138218	cv (WI2291)	Syria	ICARDA
60	IG31396	*H. vulgare* convar. *distichon*	Syria	Homs
61	IG31410	*H. vulgare* convar. *vulgare*	Syria	Damascus
62	IG33094	*H. vulgare* convar. *distichon*	Syria	Dayr Az Zawr
63	IG138217	cv (ER/Apm)	Syria	–
64	IG138237	cv (Tokak)	Turkey	–
65	IG38672	*H. spontaneum*	Turkey	Urfa
66	IG28677	*H. vulgare* convar. *distichon*	Turkey	Kars
67	IG37608	*H. vulgare* convar. *distichon*	Yemen	Al Bayda
68	IG37612	*H. vulgare* convar. *distichon*	Yemen	Sa'dah
Europe				
69	IG138233	cv (Express)	France	–
70	IG138238	cv (Radical)	Russia	–

[a] cv, cultivars belonging to *H. vulgare* convar. *vulgare* and their names are written in parenthesis.

matrices were correlated, the highest correlation was between the gSSRs and eSSRs ($r = 0.86$, $P < 0.05$), followed by that between the eSSRs and the SNPs ($r = 0.74$, $P < 0.05$). The correlation between the gSSRs and the SNPs was nevertheless still statistically significant ($r = 0.67$, $P < 0.05$).

Genetic structure of the genotype set

Since a significant correlation existed between the three GS matrices, the allelic data (349 alleles) were pooled to examine the genetic structure of the sample. The overall mean GS was 0.43, ranging from 0.16 (IG138235–IG27683) to 0.87 (IG137726–IG138228). The resulting UPGMA dendrogram suggested the presence of two major clusters (I and II), both of which could be classified into sub-clusters. Cluster I was formed from two sub-clusters (IA and IB), comprising sub-sub-clusters IA-i, IA-ii, IB-i and IB-ii. Similarly, cluster II comprised three sub-clusters (IIA-i, IIA-ii and IIB). Six out of these seven sub-clusters were further divisible into the sub-sub-clusters designated p and q (Fig. 3, Supplementary Table 1, available online only at http://journals.cambridge.org). Cluster IA-i p included 75% of the cluster I Asian accessions, while the 48% of the African accessions fell into sub-cluster IIA-ii p. In contrast, the Middle East-derived materials were distributed across clusters I and II, with 31% of the accessions in IA-i q and 23% in IIA-i q.

Table 2. Diversity analysis based on gSSR markers

Linkage group	Marker name	SSR motif	Number of alleles	PIC value
1H	Bmac0032	(AC)7T(CA)15(AT)9	5	0.65
	Bmag0579	(AC)6(AG)15	5	0.65
	HVM20	–	6	0.71
2H	Bmag0749	(AG)11	5	0.73
	GBMS160	(GA)14	5	0.72
	GBMS247	(GT)9	6	0.81
3H	Bmag0603	(AG)24	7	0.80
	GBMS189	(GA)32	7	0.79
4H	EBmac0788	(TG)23	8	0.82
	HVM40	(GA)6(GT)4(GA)7	6	0.78
	HVM67	(GA)11	5	0.72
5H	EBmac0684	(TA)7(TG)11,(TG)11 (TTTG)5	5	0.70
	GMS001	(CT)7n(CT)2	5	0.71
	GBMS119	(CA)22	7	0.78
	GBMS032	(GT)5n(GT)11	6	0.80
6H	Bmac0040	(AC)20	4	0.73
	HVM65	(GA)10	5	0.77
7H	EBmac0755	(AC)16	6	0.80
	GBMS192	(GT)14	4	0.59
	GBMS035	(GA)12	7	0.75

Table 3. Diversity analysis based on eSSR markers

Linkage group	Marker name	SSR motif	Number of alleles	PIC value
1H	GBM1461	(CA)6n(CA)18	17	0.88
	GBM1029	(AG)10	7	0.56
	GBM1007	(AC)11	14	0.74
2H	GBM1047	(AGC)5	7	0.69
	GBM1459	(AC)7	11	0.69
	GBM1208	(AG)6	7	0.72
3H	GBM1405	(CGCA)5	5	0.76
	GBM1059	(GGT)5	10	0.79
	GBM1413	(TCATA)6	6	0.64
4H	GBM1221	(AC)10	15	0.78
	GBM1323	(GCC)8	7	0.75
5H	GBM1483	(GCG)7	14	0.81
	GBM1054	(CCG)5	7	0.73
	GBM1176	(AT)8	6	0.67
6H	GBM1021	(AC)8	13	0.83
	GBM1212	(AGG)5	4	0.60
	GBM1063	(ACAT)7	6	0.73
7H	GBM1464	(CAG)8n(CAG)5	17	0.85
	GBM1516	(CT)9	8	0.79
	GBM1326	(CTT)8	8	0.70

Discussion

Comparison of marker systems

SSR markers are inherently multi-allelic, and the range of allele number across the 40 loci assayed ranged from 4 to 17. This multi-allelism has established SSRs as the effective marker platform in the current crop diversity studies (Gupta and Varshney, 2000). The expectation is that the allele number at gSSR loci is likely to be superior to that at eSSR loci, as non-coding DNA tends to be the more tolerant of sequence expansion (Russell *et al.*, 2004; Varshney *et al.*, 2005). Surprisingly, this expectation was not borne out in the present study, where the mean allele

Table 4. SNP-based genetic diversity analysis

Linkage group	Marker name	Genotyping assay[a]	PIC value per SNP
1H	GBS0554	CAPS (*Hha*I)	0.35
	GBS0361	CAPS (*Hha*I)	0.30
	GBS0546	CAPS (*Sml*I)	0.37
2H	GBS0379	CAPS (*Dra*I)	0.44
	GBS0535	CAPS (*Mse*I)	0.48
	GBS0705	PS	0.34
3H	GBS0526	CAPS (*Psi*I)	0.50
	GBS0431	CAPS (*Rsa*I)	0.49
	GBS0667	CAPS (*Cac8*I)	0.24
4H	GBS0461	PS (pos1)	0.45
		PS (pos2)	0.30
	GBS0288	CAPS (*Hha*I)	0.47
	GBS0192	CAPS (*Rsa*I)	0.43
5H	GBS0576_pos1	PS (pos1)	0.48
		PS (pos2)	0.46
		PS (pos 3)	0.44
	GBS0577	CAPS (*Dde*I)	0.35
	GBS0527	CAPS (*Eco*RV)	0.43
6H	GBS0157	CAPS (*Sal*I)	0.48
	GBS0136	CAPS (*Taq*I)	0.09
7H	GBS0591	PS	0.30
	GBS0291	CAPS (*Hin*fI)	0.33
	GBS0317	CAPS (*Hha*I)	0.47

[a] In parentheses, the identity of the restriction enzyme (CAPS assay), or the SNP position (pos) numbers (pyrosequencing assay).

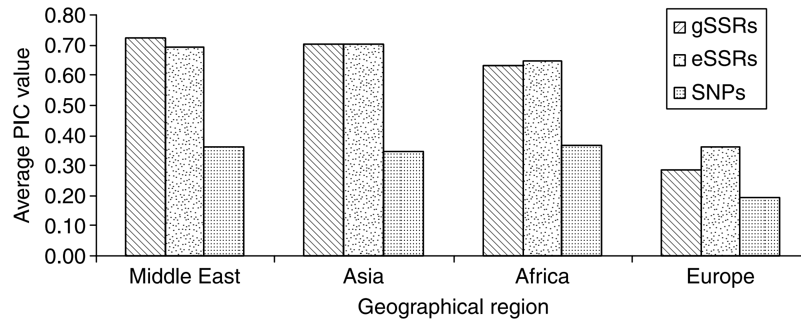

Fig. 1. gSSR-, eSSR- and SNP-based PIC values for accessions originating from the Middle East, Asia, Africa and Europe.

number at the eSSR loci was 9.5, while at the gSSR loci it was only 5.7. We suppose that this apparent inconsistency arose because, while the gSSR set was a random sample (Varshney *et al.*, 2007b), the eSSR set represents a 'core set', pre-selected on the basis of a high PIC (Varshney *et al.*, 2007c). Unlike for the number of alleles detected, no significant difference was observed between the gSSR and the eSSR PIC values (0.74 ± 0.06 and 0.73 ± 0.08, respectively) markers. This suggests that although a greater number of alleles was present at the eSSR loci, the allele frequencies of the rarer alleles was too low to have any marked effect on the informativeness of the loci.

In terms of generating the comparable genotyping datasets, the highest correlation ($r = 0.86$, $P < 0.05$) was observed between gSSR and eSSR marker data. This is possible as both markers are multi-allelic and represent a random distribution in genome. The second strongest correlation ($r = 0.74$, $P < 0.05$) between eSSR and SNP marker datasets can be explained based on the similar origin (transcribed part of genome) of these two markers, as shown in Kota *et al.* (2001a) and Varshney *et al.* (2007c).

Germplasm diversity and relationships

Accessions originating from the Middle East and Asia were more genetically diverse than those from Africa or Europe. Specifically, the former displayed a higher PIC value, while the latter tended to show a greater number of alleles. These tendencies agree well with outcomes reported by Malysheva-Otto *et al.* (2006), and are predictable, given that the Middle East and Central Asia have been identified as major centres of origin and/or diversification of barley (Pozzi *et al.*, 2004). The germplasm collection showed an overall mean GS of 0.43. However, the sampling strategy for the 70 accessions was not random, so caution needs to be exercised in extrapolating the sample mean GS levels to the collection as a whole.

The clustering exercise identified a substantial degree of association between provenance and genotype. Thus, subclusters IA-i p and IIA-ii p were formed from accessions sharing, respectively, Asian and African provenance. The small clusters IB-i, IB-ii, IIA-i q, IIA-ii q and IIB were all dominated by accessions of similar provenance. However, two moderately sized clusters (IA-i q and IIA-i p) were geographically heterogeneous. These patterns are suggestive of a narrowing of the genetic base of barley due to breeding for adaptation in certain geographical regions. This applies less to accessions from the Middle East, which group with both Asian and African materials, consistent with the accepted view that the centre of origin of barley is in the Fertile Crescent (Pozzi *et al.*, 2004), from where it spread to both Africa (Orabi *et al.*, 2007) and Asia (Pandey *et al.*, 2006). There does therefore seem to be scope for barley

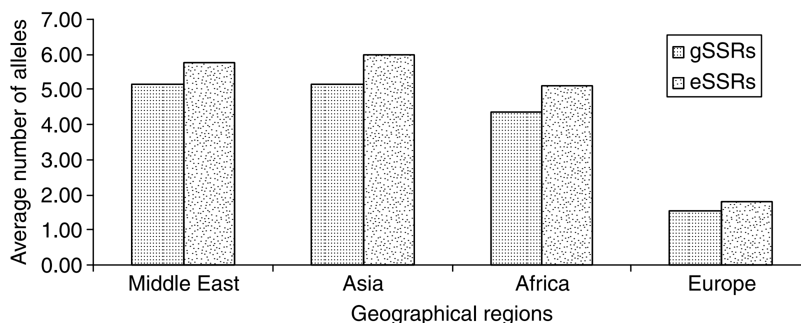

Fig. 2. The distribution of gSSR and eSSR alleles among accessions originating from the Middle East, Asia, Africa and Europe.

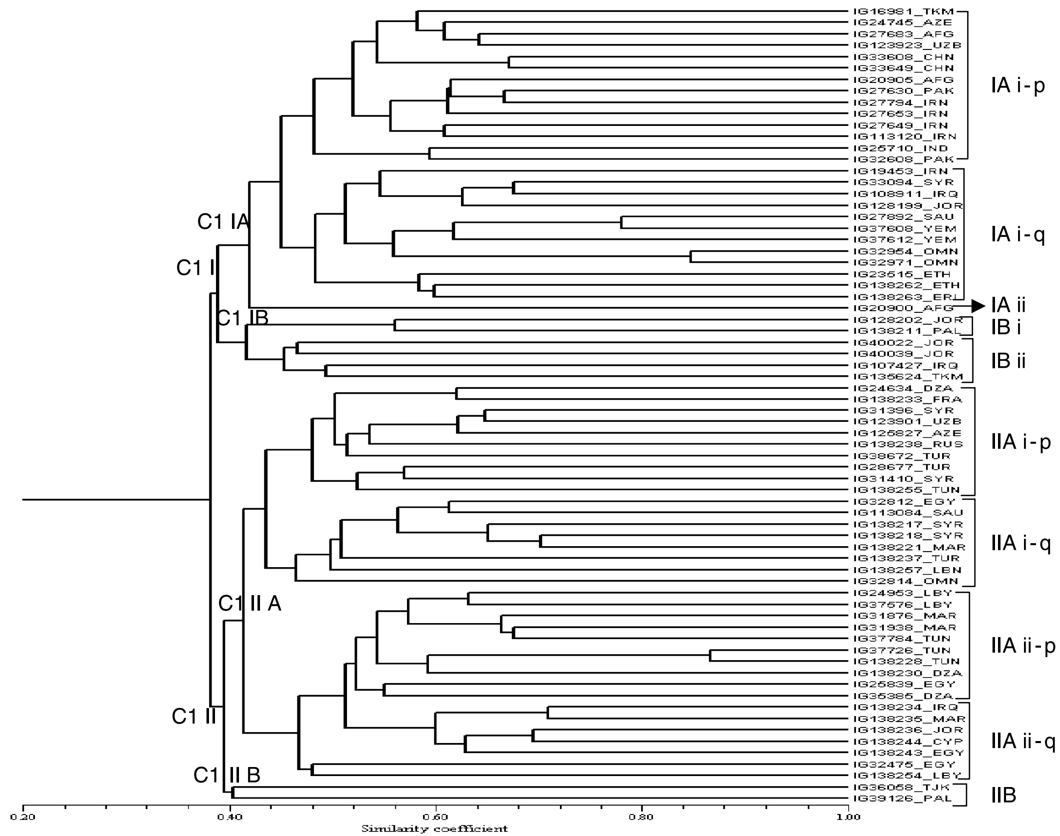

Fig. 3. Dendrogam analysis of genetic relationships, based on combined genotypes (gSSRs, eSSRs and SNPs).

improvement via the exploitation of geographically exotic materials.

In summary, the present study has shown that the core set eSSRs display a comparable level of polymorphism as do gSSRs and SNPs. It appears that more genetic diversity is present among accessions from the Middle East or Asia than among those from Africa, so that this diversity could be of benefit for the improvement of barley in less diversity-rich regions.

Acknowledgements

We thank Jan Valkoun (ICARDA) for providing seed of the barley accessions. This research was financially supported in part by the GTZ project 2002.7860.6-001.00 and Contract 1060503, provided by BMZ (Bundesministerium für Wirtschaftliche Zusammenarbeit und Entwicklung), Germany.

References

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

Gupta PK and Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica* 113: 163–185.

Kanazin V, Talbert H, See D, DeCamp P, Nevo E and Blake T (2002) Discovery and assay of single nucleotide polymorphsims in barley (*Hordeum vulgare*). *Plant Molecular Biology* 48: 529–537.

Kota R, Varshney RK, Thiel T, Dehmer KJ and Graner A (2001a) Generation and comparison of EST-derived SSRs and SNPs in barley (*Hordeum vulgare* L.). *Hereditas* 135: 145–151.

Kota R, Wolf M, Michalek W and Graner A (2001b) Application of DHPLC for mapping of single nucleotide polymorphisms (SNPs) in barley (*Hordeum vulgare* L.). *Genome* 44: 523–528.

Kota R, Varshney RK, Prasad M, Zhang H, Stein N and Graner A (2007) EST-derived single nucleotide polymorphism (SNP) markers for assembling genetic and physical maps of the barley genome. *Functional and Integrative Genomics*, doi: 10.1007/s10142-007-0060-9.

Li JZ, Sjakste TG, Röder MS and Ganal MW (2003) Development and genetic mapping of 127 new microsatellite markers in barley. *Theoretical and Applied Genetics* 107: 1021–1027.

Liedloff A (1999) Mantel non Parametric Test Calculator, Version 2.0. School of Natural Resource Sciences, Queensland University of Technology, Queensland, Australia.

Liu ZW, Biyashev RM and Saghai Maroof MA (1996) Development of simple sequence repeat DNA markers and their integration into a barley linkage map. *Theoretical and Applied Genetics* 93: 869–876.

Malysheva-Otto LV, Ganal MW and Röder MS (2006) Analysis of molecular diversity, population structure and linkage disequilibrium in a worldwide survey of cultivated barley germplasm (*Hordeum vulgare* L.). *BMC Genetics* 7: 6.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209–220.

Orabi J, Backes G, Wolday A, Yahyaoui A and Jahoor A (2007) The Horn of Africa as a centre of barley diversification and a potential domestication site. *Theoretical and Applied Genetics* 114: 1117–1127.

Pandey M, Wagner C, Friedt W and Ordon F (2006) Genetic relatedness and population differentiation of Himalayan hulless barley (*Hordeum vulgare* L.) landraces inferred with SSRs. *Theoretical and Applied Genetics* 113: 715–729.

Pozzi C, Rossini L, Vecchietti A and Salamini F (2004) Genes and genome changes during domestication of cereals. In: Gupta PK and Varshney RK (eds) *Cereal Genomics*. Dordrecht: Kluwer Academic Publishers, pp. 165–198.

Ramsay L, Macaulay M, Ivanissevich DS, MacLean K, Cardle L, Fuller J, Edwards KJ, Tuvesson S, Morgante M, Massari A, Maestri E, Marmiroli N, Sjakste T, Ganal M, Powell W and Waugh R (2000) A simple sequence repeat-based linkage map of barley. *Genetics* 156: 1997–2005.

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

Rohlf FJ (1998) NTSYS-pc numerical taxonomy and multivariate analysis system. Version 2.02. Exeter Publications Setauket, New York.

Rostoks N, Ramsay L, MacKenzie K, Cardle L, Bhat PR, Roose ML, Svensson JT, Stein N, Varshney RK, Marshall DF, Graner A, Close TJ and Waugh R (2006) Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. *Proceedings of the National Academy of Sciences USA* 103: 18656–18661.

Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor A, Powell W and Waugh R (1997) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical and Applied Genetics* 95: 714–722.

Russell J, Booth A, Fuller J, Harrower B, Hedley P, Machray G and Powell W (2004) A comparison of sequence-based polymorphism and haplotype content in transcribed and anonymous regions of the barley genome. *Genome* 47: 389–398.

Struss P and Plieske J (1998) The use of microsatellite markers for detection of genetic diversity in barley populations. *Theoretical and Applied Genetics* 97: 308–315.

Thiel T, Michalek W, Varshney RK and Graner A (2003) Exploiting EST databases for the development of cDNA derived microsatellite markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 106: 411–422.

Varshney RK, Graner A and Sorrells ME (2005) Genic microsatellite markers: features and applications. *Trends in Biotechnology* 23: 48–55.

Varshney RK, Grosse I, Hahnel U, Thiel T, Rudd S, Zhang H, Prasad M, Stein N, Langridge P and Graner A (2006) Genetic mapping and BAC assignment of EST-derived SSR markers shows non-uniform distribution of genes in the barley genome. *Theoretical and Applied Genetics* 113: 239–250.

Varshney RK, Beier U, Khlestkina EK, Kota R, Korzun V, Graner A and Börner A (2007a) Single nucleotide polymorphisms in rye (*Secale cereale* L.): discovery, frequency, and applications for genome mapping and diversity studies. *Theoretical and Applied Genetics* 114: 1105–1116.

Varshney RK, Marcel TC, Ramsay L, Russell J, Röder M, Stein N, Waugh R, Langridge P, Niks RE and Graner A (2007b) A high density barley microsatellite map with 775 SSR loci. *Theoretical and Applied Genetics* 114: 1091–1103.

Varshney RK, Thiel T, Sretenovic-Rajicic T, Baum M, Valkoun J, Guo P, Grando S, Ceccarelli S and Graner A (2007c) Identification and validation of a core set of informative genic SSR and SNP markers for assaying functional diversity in barley. *Molecular Breeding*, doi: 10.1007/s11032-007-9151-5.