Genet Resour Crop Evol (2010) 57:515–521
DOI 10.1007/s10722-009-9489-0

RESEARCH ARTICLE

EST-SSR based estimates on functional genetic variation in a barley (*Hordeum vulgare* L.) collection from Egypt

Khaled F. M. Salem · Rajeev K. Varshney · Marion S. Röder · Andreas Börner

Received: 20 December 2008 / Accepted: 24 September 2009 / Published online: 15 October 2009

Abstract Genetic diversity was investigated in a set of 27 barley genotypes, originated from Egypt, using 23 expressed sequence tag (EST)-derived simple sequence repeats (eSSRs), representatives of the seven barley chromosomes. Ninety-five alleles were detected among all the barley accessions. For 22 polymorphic eSSRs, the number of alleles per locus varied from 2 to 12, with a mean of 4.318 alleles per locus. The lowest and the largest number of alleles per locus among the seven homeologous groups was observed in homeologous group 5H and 4H with 3.00 and 6.33, respectively. The gene diversity increased as the number of alleles increased. Gene diversity for 22 ESTs loci varied from 0.137 for GBM1404 to 0.896 for GBM1015 with an average of 0.563. A significant correlation coefficient between gene diversity and the number of alleles was high, $r = 0.741$ ($P < 0.01$). Cluster analysis was conducted based on eSSRs data to group the barley genotypes and to construct a dendrogram. Four groups can be distinguished by truncating the dendrogram at gs value of 0.77.

Keywords Barley · eSSR · Genetic diversity · *Hordeum vulgare* · Molecular markers

K. F. M. Salem (✉)
Genetic Engineering and Biotechnology Research Institute (GEBRI), Menoufia University, P. O. Box 79, Sadat City, Egypt
e-mail: salemcairo@yahoo.com

R. K. Varshney
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru-502324, Greater Hyderabad, India

M. S. Röder · A. Börner
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstrasse 3, 06466 Gatersleben, Germany

Introduction

Barley (*Hordeum vulgare*) is the fourth important cereal crop in Egypt after wheat, maize and rice. Prior to national barley breeding programs cultivated barley represented essentially spring-type. Since the initiation of barley breeding program in Egypt, new cultivars were developed by (1) selection from local populations, (2) introduction of new varieties and (3) crossing and selection for yield and its components. Several high-yielding spring barley cultivars of two and six-row type were produced.

Molecular markers that reveal polymorphism at the DNA level have been shown to be a very powerful tool for genotype characterisation and estimation of genetic diversity. In this regard, microsatellites or simple sequence repeats (SSRs), due to their multi-allelic nature, have been extensively used in several crops (Gupta and Varshney 2000). In recent years, due to availability of expressed sequence tags (ESTs) and

progress in the area of bioinformatics, SSR markers have been derived from ESTs and such SSR markers have been referred as EST-SSRs or eSSRs (Varshney et al. 2005). These markers were shown to have a high potential for identification and estimation of functional genetic diversity in germplasm collections (Varshney et al. 2005, 2007b). Like other major crops, the availability of a large set of expressed sequence tags (ESTs) for barley, in recent years, has provided a resource for the systematic development of EST/gene based molecular markers including eSSRs (Varshney et al. 2005).

Availability of eSSR marker sequences for oligonucleotide synthesis, involvement of PCR amplification, the simplicity of protocol that produces reliable and highly detectable amplification products, their co-dominance and single localisation constitutes their advantages over AFLP, RFLP and RAPD markers. Molecular markers developed from gene or EST resources for crop plants have been popularly called as genic molecular markers (GMMs) (Varshney et al. 2007a). In case the polymorphism site in the marker determines the function of the marker, these have been referred as 'functional markers' (FMs, Anderson and Lübberstedt 2003). GMMs or FMs include SSR, SNP (single nucleotide polymorphism) and COS (conserved orthologous sequences) markers which are easy to develop utilizing the EST resources and mirrors the functional genomic component (Kota et al. 2003; Varshney et al. 2005, 2006). These markers, at present, are gaining momentum for estimating the functional genetic diversity in genebank collections, and natural as well as breeding populations of barley (Pillen et al. 2000; Khlestkina et al. 2006). The objectives of the present study were to: (1) evaluate polymorphism of EST-SSR markers in Egyptian barley and (2) survey the level of functional genetic diversity of Egyptian barley for the development of superior cultivars.

Materials and methods

Plant material

A total of 27 diverse barley (*Hordeum vulgare* L. subsp. *vulgare*) genotypes originated from Egypt but obtained from Agricultural Research Center (ARC), Giza, Egypt, Leibniz Institute for Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany and USDA, Fort Collins, USA, were cultivated at the experimental farm of Genetic Engineering and Biotechnology Research Institute (GEBRI), Menoufia University, Egypt during the barley growing season 2004/2005. In addition, the two varieties Morex and Barke were included as standards. Details on these genotypes are given in Table 1.

DNA isolation

Young leaves for 5 seedling from 8-weeks-old plants were cut as tissue samples for DNA extraction. An equal amount of tissue was taken from each single plant. Only one replication was sampled for DNA extraction. The tissues were transported the same day to the laboratory on liquid nitrogen and stored. DNA was isolated from these genotypes as described by Thiel et al. (2003).

PCR amplification

PCR reaction contained 50–100 ng template DNA, 250 nM forward primer, 250 nM reverse primer, 0.2 mM dNTPs, 2.5 µl PCR buffer (10×), 1.5 mM MgCl, 1U *Taq* DNA polymerase in a total volume of 25 µl. Amplifications were carried out as given in Thiel et al. (2003).

eSSR analysis

For eSSR analysis, 23 SSR markers were selected from Thiel et al. (2003) and Varshney et al. (2006). Indeed, the selected markers are part of the core set defined recently by Varshney et al. (2008) (Table 2). Fragment detection for eSSR markers was carried out as given in Thiel et al. (2003).

Data collection and diversity analysis

Gene diversity was calculated according to formula of Nei (1973) using the equation $\mathrm{PIC} = 1 - \sum_{i=1}^{k} P_i^2$, where k is the total number of alleles detected for a locus of a marker and P_i is the frequency of the ith allele in the set of 27 barley varieties investigated. Anderson et al. (1993) indicated that gene diversity is essentially the same as the polymorphism information content (PIC) as used by Botstein et al. (1980).

Table 1 Accession name, year of release, pedigree and seed source of the barley materials used in this study

No.	Accession name	Source of seed	Year of release	Pedigree
1	HOR 19704	IPK, Germany	No data	Selected landrace
2	HOR 17683	IPK, Germany	No data	Selected landrace
3	HOR 16107	IPK, Germany	No data	Selected landrace
4	HOR 16102	IPK, Germany	No data	Selected landrace
5	HOR 14411	IPK, Germany	No data	Selected landrace
6	HOR 16097	IPK, Germany	No data	Selected landrace
7	HOR 8255	IPK, Germany	No data	Selected landrace
8	HOR 8658	IPK, Germany	No data	Selected landrace
9	HOR 8659	IPK, Germany	No data	Selected landrace
10	HOR 8806	IPK, Germany	No data	Selected landrace
11	HOR 819	IPK, Germany	No data	Selected landrace
12	HOR 1938	IPK, Germany	No data	Selected landrace
13	HOR 2252	IPK, Germany	No data	Selected landrace
14	HOR 3711	IPK, Germany	No data	Selected landrace
15	HOR 7419	IPK, Germany	No data	Selected landrace
16	HOR 8212	IPK, Germany	No data	Selected landrace
17	HOR 19027	IPK, Germany	No data	Selected landrace
18	HOR 19308	IPK, Germany	No data	Selected landrace
19	HOR 19720	IPK, Germany	No data	Selected landrace
20	HOR 20117	IPK, Germany	No data	Selected landrace
21	Giza 123	ARC, Egypt	1988	Giza 117/FAO 86
22	Giza 125	USDA, USA	1995	No data
23	Giza 126	ARC, Egypt	1995	Baladi Bahteem/ SD 729-Por 12762-BC
24	Giza 127	USDA, USA	1993	No data
25	Giza 128	ARC, Egypt	1994	No data
26	Club Mariout	USDA, USA	No data	Selected landrace
27	MSS 25	USDA, USA	No data	No data

Genetic similarity estimation and cluster analysis

Each eSSR band was scored as present (1), absent (0) or as a missing observation for the different cultivars. Genetic similarity (gs) between two cultivars *i* and *j* was estimated following Nei and Li (1979) by the formula: $gsij = 2Nij/(Ni + Nj)$, where *Nij* is the number of bands present in genotypes *i* and *j*, *Ni* (resp. *Nj*) is the number of bands present in genotype I (resp. *j*). Markers with missing observations for genotype *i* and/or *j* were not included in the calculation of gs*ij*. Based on the genetic similarity matrix (denoted GS), Unweighted Pair Group Method of Arithmetic Average (UPGMA) cluster analysis were used to assess pattern of diversity among the barley entries. All calculations were performed using the NTSYS-pc version 2.1 software package (Biostatistics Inc., USA, Rohlf 2000).

Results

Characteristics of eSSR markers

Twenty-three eSSR markers dispersed across the genome were used to test the genetic diversity of 27 landraces/cultivars. Twenty-two eSSR markers generated polymorphic patterns and one (GBM1043) gave a monomorphic pattern, yielding a polymorphism rate of 95.65% (Table 2). A total of 95 alleles were detected. The number of alleles per locus varied from 2 to 12, with a mean of 4.318 alleles per locus

Table 2 Characteristics of eSSR markers used with the chromosomal location, marker name, eSSR motif, number of alleles per locus and polymorphism information content (PIC) calculated over a set of 27 barley genotypes

Marker name	eSSR motif	Chromosomal location	Number of alleles	PIC
GBM1461	(CA)6n(CA)18	1H	9	0.850
GBM1002	(CCT)7		4	0.331
GBM1029	(AG)10		3	0.447
GBM1047	(AGC)5	2H	4	0.647
GBM1218	(GA)11		3	0.615
GBM1208	(AG)6		6	0.625
GBM1035	(CT)8		5	0.775
GBM1031	(AG)15	3H	4	0.601
GBM1059	(GGT)5		4	0.528
GBM1280	(CTT)7		4	0.648
GBM1043	(AAC)5		1	0.000
GBM1221	(AC)10	4H	5	0.642
GBM1015	(ACAT)13		12	0.896
GBM1020	(AC)7		2	0.483
GBM1026	(AC)9	5H	3	0.322
GBM1054	(CCG)5		4	0.513
GBM1483	(GCG)7		2	0.483
GBM1021	(AC)8	6H	6	0.798
GBM1404	(TATG)5		2	0.137
GBM1464	(CAG)8n(CAG)5	7H	5	0.754
GBM1516	(CT)9		2	0.466
GBM1419	(CTCAT)5		3	0.203
GBM1060	(GGT)6		3	0.615
Total			96	12.379
Average			4.318	0.563

(Table 2). The lowest and the largest number of alleles per locus among the seven homeologous groups was observed in homeologous group 5H and 4H with 3.00 and 6.33, respectively (Table 3). Gene diversity for 22 ESTs loci varied from 0.137 for GBM1404 to 0.896 for GBM1015 with an average of 0.563. Gene diversity for the seven homeologous groups varied from 0.439 for homeologous group 5H to 0.674 for homeologous group 4H. The correlation coefficient between gene diversity and the number of alleles was high, $r = 0.741$ ($P < 0.01$). The linear relationship between them is shown in Fig. 1.

Table 3 Genetic diversity in different chromosomes across 22 loci in 27 barley genotypes

Chromosomal location	Number of alleles	PIC
1H	5.33	0.543
2H	4.50	0.665
3H	3.25	0.444
4H	6.33	0.674
5H	3.00	0.439
6H	4.00	0.468
7H	3.25	0.509
Total	29.67	3.742
Average	4.24	0.535

Genetic diversity

UPGMA cluster analysis of eSSR genetic similarity (gs) matrix resulted in the dendrogram presented in Fig. 2. Four groups can be distinguished by truncating the dendrogram at gs value of 0.77. The major group (denoted group I) consists of 9 genotypes and includes landraces HOR2252, HOR14411, HOR1938,

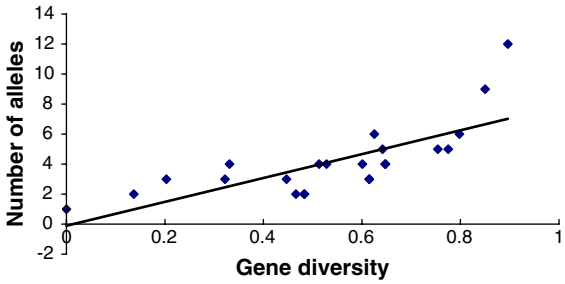

Fig. 1 Correlation between gene diversity and the number of alleles over 22 eSSR loci in 27 barley genotypes. Described by the function $y = -0.0991 + 7.9391x$, $R^2 = 0.549$

HOR20390 and HOR19027, the cultivars Giza 127 and Giza 128 as well as the two standards Barke and Morex. Another group (Group III) contains thirteen barley genotypes and includes landraces HOR19704, HOR8212, HOR7419, Mss25, HOR8255, HOR8658, HOR19308, HOR17683, HOR3711 and HOR20609 and the cultivars Giza 125, Giza 126 and Giza 123. These cultivars have the same origin but are very likely to be genetically unrelated (see Table 1) which may explain the high diversity level within this group. Interestingly, HOR 19720 is not clustered to any other landrace/cultivar and form a separate group (group II). The landraces HOR 16107, HOR 819, HOR 16097, Club Mariout, HOR 20117 and HOR 16102 are well separated from the other groups and form Group IV.

Discussion

Genetic diversity levels and patterns

The SSR and SNP markers derived from ESTs, due to their inexpensive developmental costs (Kota et al. 2003; Varshney et al. 2005) are increasingly being used for genotyping of natural or breeding populations. Together with these markers, AFLP markers are still considered good for fingerprinting or diversity analysis. Therefore, the present study documents the utility of eSSRs marker type for genetic diversity studies.

Marker polymorphism

Twenty-two eSSR markers generated polymorphic patterns and one (GBM1043) gave a monomorphic pattern, yielding a polymorphism rate of 95.65%. The high level of polymorphism associated with SSR is to be expected because of the unique mechanism responsible for generating SSR allelic diversity by replication slippage (Varshney et al. 2005; Gupta and Varshney 2000).

Barley eSSR markers showed an average PIC value of 0.563, which confirms that eSSR markers are highly informative. This value is higher than the PIC value (0.38) obtained by Pillen et al. (2000) who used 22 microsatellite markers and a set of 28 mainly German

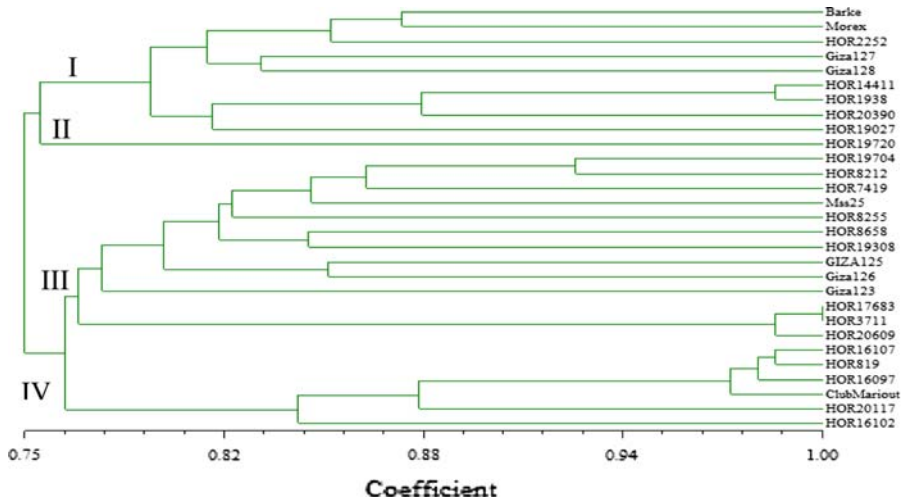

Fig. 2 UPGMA cluster analysis-based dendrogram depicting genetic relationships among 27 Egyptian barley genotypes and the two varieties Morex and Barke as standards and based on data of 22 eSSRs primer pairs

barley cultivars and two wild forms. This demonstrates that the genotypes used in this study are more diverse than German cultivars and could be related to different origin of the genotypes, local and introduced cultivars. In fact, utilisation of genotypes with the same origin implies a low genetic diversity, since these genotypes may have exchanged genetic material through breeding programs. However, a high level of PIC may also be due to the level of polymorphism of the eSSR markers. The choice of highly polymorphic markers contributes to the enhancement of PIC values (Struss and Plieske 1998). Using fewer genotypes (18) obtained from different parts of Europe and 13 microsatellite markers, a higher PIC value of 0.55 was obtained by Russell et al. (1997).

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 ($r = 0.741$, $P < 0.001$) Therefore, the number of alleles can be used to evaluate the genetic diversity. The obtained results agree with those of Huang et al. (2002) who reported that the PIC value was correlated with the number of alleles and did not agree with those of Prasad et al. (2000).

In this study, microsatellite markers were able to discriminate between all landraces/cultivars and allowed the identification of accessions based on the highest value of genetic similarity between these accessions and known cultivars.

This high discrimination between closely related landraces is also the result of the large heterogeneity of local landraces, which are defined in Egypt as geographically based populations. Each population has typical characteristics of the grain appreciated by farmers and local inhabitants for food and tasting properties. Grando et al. (2000) mentioned that considerable heterogeneity exists both between landraces collected in different farmers fields (if designated by the same name) and between individual plants within the same farmers field for several plant characteristics.

In summary, the present study highlights the advantages of eSSR markers system for diversity analyses in breeding and natural populations or genebank materials to exploit the genotyping data for crop improvement as well as *ex situ* conservation strategies of plant genetic resources. The finding of eSSR based polymorphism showed that eSSR markers are powerful to examine functional diversity. A broader study with a large collection of well characterised accessions collected from several areas of Egypt will be realised using EST-SSR markers to examine genetic diversity in relation to adaptive variation.

References

Anderson JR, Lübberstedt T (2003) Functional markers in plants. Trends Plant Sci 8:554–560

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

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

Grando S, Ceccarelli S, Tekle B (2000) Diversity in barley landraces from the Near East. In: 8th International barley genetics symposium. Adelaide, South Australia. Vol. II, pp 13–15

Gupta PK, 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

Huang XQ, Börner A, Röder MS, Ganal MW (2002) Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. Theor Appl Genet 105:699–707

Khlestkina E, Varshney RK, Röder M, Graner A, Börner A (2006) Comparative assessment of genetic diversity in cultivated barley collected at different periods of the last century in Austria, Albania and India by using genomic and genic SSR markers. Plant Genet Resour 4:125–133

Kota R, Rudd S, Facius A, Kolesov G, Thiel T, Zhang H, Stein N, Mayer K, Graner A (2003) Snipping polymorphisms from large EST collections in barley (*Hordeum vulgare* L.). Mol Genet Genomics 270:224–233

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

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

Pillen K, Binder A, Kreuzkam B, Ramsay L, Waugh R, Förster J, Leon J (2000) Mapping new EMBL-derived barley microsatellites and their use in differentiating German barley cultivars. Theor Appl Genet 101:652–660

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

Rohlf FJ (2000) NTSYS-pc: numerical taxonomy and multivariate analysis system. Version 2.1 Exeter, New York

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

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

Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and

characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor Appl Genet 106:411–422

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

Varshney RK, Grosse I, Hahnel U, Thiel T, Rudd S, Zhang H, Prasad M, Stein N, Langridge P, Graner A (2006) Genetic mapping and physical mapping (BAC-identification) of EST-derived microsatellite markers in barley (*Hordeum vulgare* L.). Theor Appl Genet 113:239–250

Varshney RK, Mahender T, Aggrawal RK, Börner A (2007a) Genic molecular markers in plants: development and applications. In: Varshney RK, Tuberosa R (eds) Genomics assisted crop improvement, vol I: genomics approaches and platforms. Springer, Dordrecht, pp 13–30

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

Varshney RK, Thiel T, Sretenovic-Rajicic T, Baum M, Valkoun J, Guo P, Grando S, Ceccarelli S, Graner A (2008) Identification and validation of a core set of informative genic SSR and SNP markers for assaying functional diversity in barley. Mol Breed 22:1–13