Theor Appl Genet (1992) 84:682–687

Ribosomal DNA polymorphisms and the Oriental-Occidental genetic differentiation in cultivated barley

Qifa Zhang [1], M. A. Saghai Maroof [2], and P. G. Yang [1]

[1] Biotechnology Center, Huazhong Agricultural University, Wuhan 430070, China
[2] Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

Received January 9, 1992; Accepted January 18, 1992
Communicated by P. M. A. Tigerstedt

Summary. A total of 289 accessions of cultivated barley were assayed for ribosomal DNA (rDNA) polymorphisms. These accessions comprised four independent samples: (1) 79 entries from China, (2) 59 accessions from Ethiopia, (3) 59 entries from Tibet and (4) 92 entries representing 36 barley growing countries of the world (referred to as "world sample"). In all, 17 rDNA phenotypes (genotypes) were observed, which were composed of 10 alleles at two rDNA loci, *Rrn1* and *Rrn2*. The world sample contained the largest number of phenotypes and alleles and also demonstrated the highest level of diversity. Ribosomal DNA phenotypes 104, 112 and 107, 112 occurred at high frequencies worldwide. Allele 112 was the predominant allele of *Rrn1* in all four samples, and 104 and 107 were the two major alleles of *Rrn2* worldwide. The distributions of rDNA genotypes and alleles demonstrated a clear differentiation of two distinct barley groups: an Oriental group represented by the samples from China and Tibet, which is characterized by allele 107 at the *Rrn2* locus (rDNA phenotype 107, 112); and an Occidental group, represented by Ethiopian and world samples, which is comprised mostly of allele 104 at the *Rrn2* locus (rDNA phenotype 104, 112). The results also raised new questions concerning the phylogeny and evolution of cultivated barley.

Key words: rDNA polymorphism – *Hordeum vulgare* ssp. *vulgare* – Geographical differentiation – Evolution – Phylogeny

Introduction

The ribosomal DNA of plants is arranged in tandemly repeating arrays varying from several hundred to several thousand copies per haploid genome. In barley, the average copy number of rDNA is 1880 per haploid genome (Zhang et al. 1990a). Each repeat consists of a transcription unit and an intergenic spacer (IGS) region that separates adjacent transcription units. The IGS region includes an array of repeated sequences referred to as subrepeats. In barley, the restriction enzyme *Sst*I cleaves each of the rDNA units twice, once on each side of the transcription unit, thus generating two rDNA fragments from each repeat unit. One fragment, representing the major portion of the transcription unit, is invariant (≈3880 bp), whereas the other fragment, made up primarily of the IGS region, varies in length as a result of differences in the number of subrepeats present in the IGS region. Twenty rDNA spacer-length variants (slvs) have been identified as alleles at two Mendelian loci, *Rrn1* and *Rrn2*, associated with the nucleolar organizer regions of chromosomes 6 and 7, respectively (Saghai Maroof et al. 1984; Allard et al. 1990). The 20 slvs are organized into two families, one comprised of a regularly complete 8-step ladder, 4625–5430 bp long, arranged in the nucleolar organizer region of chromosome 7. The other is a 12-step ladder, 5545-6695 bp long, located in the nucleolar organizer region of chromosome 6. All slvs differ from their immediately adjacent neighbors by a 115-bp subrepeat except for slv 108a (the shortest variant in *Rrn1*), which is about 42 bp shorter than slv108 and 73 bp longer than slv107. The 8 shorter slvs (100–107) segregate and serve as markers for the 8 rDNA alleles of the locus *Rrn2*, whereas the 12 longer slvs (108a–118) segregate and serve as markers for the 12 alleles of *Rrn1*. Thus, in barley, an rDNA allele is comprised of a large number of tandem repeats of a transcriptional unit plus an associated IGS region that is variable in length.

It has been demonstrated in an experimental population of cultivated barley (*Hordeum vulgare* ssp. *vulgare*)

Correspondence to: M. A. Saghai Maroof

grown in large plots without conscious selection that the dynamic changes in the rDNA allele frequencies are under the influence of natural selection. It has also been established, in a study of rDNA polymorphisms in 18 populations of *H. vulgare* ssp. *spontaneum* (Saghai Maroof et al. 1990), that rDNA alleles and genotypes vary widely from habitat to habitat and that the occurrence of specific alleles and two-locus genotypes is closely associated with factors of the physical environments. These studies suggest that natural selection acting on the portions of the genome marked by the alleles of the two rDNA loci plays an important role in the development and maintenance of the observed patterns of molecular and genetic organization of rDNA variability in barley.

In the paper presented here we extend our investigation of rDNA variability to barley samples representing many barley growing areas of the world and to three specific regions that have been postulated as playing important roles in the evolution and variation of cultivated barley. The results clearly show that there is extensive variation in rDNA alleles and genotypes in cultivated barley. The study reveals two distinct barley groups with respect to rDNA genotypes: an Oriental group that is likely to be centered at Tibet, and an Occidental group that appears to be diversified from Ethiopia. The results also raise new questions concerning the phylogeny and evolution of cultivated barley.

Materials and methods

The genetic materials used in this study were 289 accessions of cultivated barley (*H. vulgare* ssp. *vulgare*), grouped in four independent samples. The first sample, henceforth referred to as "world sample", comprised 92 accessions from 36 countries representing all the major barley growing areas of the world. The second sample consisted of 59 landraces from Ethiopia. These two samples were obtained from the barley collection maintained by the United States Department of Agriculture. The third sample included 79 landraces and cultivars collected from barley growing areas all over China. The fourth sample consisted of 59 accessions of Tibetan barley drawn from the *Hordeum* collection maintained at the Agricultural Institute of Tibet, Lhasa.

DNA was prepared from single plants for the 92 accessions of the world sample and from bulked seedlings for each entry of the other three samples following the procedures of Saghai Maroof et al. (1984). About 1 μg of total cellular DNA from each accession was digested with the restriction endonuclease *SstI* and fractionated in an 1.2% agarose gel. Electrophoresis, blotting and hybridization closely followed described procedures (Saghai Maroof et al. 1984). A genomic clone, pTA71 (Gerlach and Bedbrook 1979), containing the entire wheat rDNA repeat unit was used as the hybridization probe.

Scoring of the rDNA spacer-length variants and designations of alleles at the two rDNA loci were essentially the same as described by Saghai Maroof et al. (1984) and modified by Allard et al. (1990).

Results

Variation of rDNA spacer-length phenotypes (genotypes)

A total of 17 rDNA spacer-length phenotypes (genotypes can be deduced from these phenotypes) were observed in the entire sample of 289 accessions (Table 1). Two of the 17 phenotypes (fragment patterns) were one banded, 6 were two banded, 7 were three banded, and 1 phenotype each was four and five banded, respectively. Each of

Table 1. Ribosomal DNA phenotypes (genotypes) and their frequencies in each of the four samples of cultivated barley

slv phenotype	World[a]		China		Ethiopia		Tibet		Total	
	n	f	*n*	f	*n*	f	*n*	f	*n*	f
101, 104, 112	3	0.033							3	0.010
101, 104, 105, 107, 112					1	0.017			1	0.003
102, 112	1	0.011							1	0.003
104, 105, 112					1	0.017			1	0.003
104, 107, 112	1	0.011	2	0.025					3	0.010
104, 108					3	0.051			3	0.010
104, 111, 112					1	0.017			1	0.003
104, 112	56	0.609	26	0.329	51	0.864			133	0.460
106, 107, 111, 112	2	0.022							2	0.007
106, 107, 112			1	0.013			2	0.034	1	0.003
107	2	0.022					1	0.017	3	0.010
107, 108, 112							1	0.017	1	0.003
107, 110	2	0.022							2	0.007
107, 111							1	0.017	1	0.003
107, 111, 112	1	0.011					4	0.068	5	0.017
107, 112	19	0.207	50	0.633	2	0.034	50	0.847	121	0.419
112	5	0.054							5	0.017
Total	92		79		59		59		289	
h_s	1.317		0.803		0.601		0.645			1.234

[a] A sample of 92 accessions from 36 countries representing all the major barley growing areas of the world

these 17 phenotypes contained at least 1 of the 3 most frequent spacer-length variants: slv104, slv107 or slv112. Two phenotypes, namely 104, 112 and 107, 112, were predominant (f > 0.80) both in the 92 entries of the world sample and in all samples of the 289 accessions as a whole. All other phenotypes were very rare, except for 2 that were locally frequent: 104, 108 in Ethiopia (f = 0.051), and 107, 111, 112 in Tibet (f = 0.068).

Ten phenotypes were observed in the world sample, 6 each in Ethiopian and Tibetan samples, and only 4 in the sample from China. Thus, the world sample contained a larger number of rDNA phenotypes than any of the regional samples. However, a number of phenotypes appeared in regional samples that were not observed in the world sample. Examples include: phenotypes 104, 108, 104, 111, 112 and 101, 104, 105, 107, 112 in the Ethiopian sample; phenotypes 107, 111 and 107, 108, 112 in the Tibetan sample; and 106, 107, 112 in both the Tibetan and Chinese samples.

Variation of rDNA alleles

Previous studies (Saghai Maroof et al. 1984; Allard et al. 1990) established that each of the slvs represents an allele at one of the two rDNA loci, *Rrn1* and *Rrn2*. Table 2 gives the numbers and frequencies of alleles observed in these four samples. Four conventions (Saghai Maroof et al. 1990) were followed in assigning slvs to alleles of individual loci and in enumerating numbers of alleles. (1) Each individual was considered to have two rDNA loci with 2 alleles per locus. (2) All variants in the series slv 101–107 were assigned as alleles of *Rrn2* and enumerated as alleles 101–107, respectively, and all variants in the series slv 108–112 were assigned as alleles of *Rrn1* and enumerated as alleles 108–112, respectively. (3) All plants with 3 slvs were considered to be heterozygous at one locus and homozygous at the other locus, although there was evidence indicating that some of the three-banded types are homozygotes for a "compound allele" consisting of 2 different slvs (Allard et al. 1990). This convention also applies to four-banded phenotypes. (4) A one-banded plant was scored as homozygous for one locus, and null for the other locus. Our previous data indicated that these four conventions are not necessarily always correct.

It can be seen from Table 2 that, in total, 10 alleles (not including deduced null alleles) were observed in these four samples. Allele 112 was predominate at the *Rrn1* locus in all four samples and was the only allele at that locus in the sample from China. Allele 108 appeared to be locally frequent (f = 0.051) in the Ethiopian samples, while allele 111 was locally frequent (f = 0.052) in the Tibetan samples. Both collection regions represent environmental extremes. At the *Rrn2* locus, alleles 104 and 107 occurred at a high frequency in the world sample as well as in all the samples as a whole. There were also alleles that appeared in the three regional samples but

Table 2. Alleles of the two rDNA loci, *Rrn1* and *Rrn2*, observed in each of the four samples. See text for the conventions of allele enumeration

Allele	World[b]		China		Ethiopia[c]		Tibet		Total	
	n	f	*n*	f	*n*	f	*n*	f	*n*	f
Rrn2										
101	3	0.016							3	0.005
102	2	0.011							2	0.003
104	116	0.630	54	0.342	111	0.957			281	0.488
105					1	0.009			1	0.002
106	2	0.011	1	0.006			2	0.017	5	0.009
107	51	0.277	103	0.652	4	0.034	116	0.983	274	0.476
Null[a]	10	0.054							10	0.017
Total	184		158		116		118		576	
Rrn1										
108					6	0.051	1	0.008	7	0.012
110	4	0.022							4	0.007
111	3	0.016			1	0.008	6	0.051	10	0.017
112	173	0.940	158	1.00	111	0.941	109	0.924	551	0.953
Null[a]	4	0.022					2	0.017	6	0.010
Total	184		158		118		118		578	

[a] Null alleles were deduced according to the conventions we followed in this paper
[b] A sample of 92 accessions from 36 countries representing all the major barley growing areas of the world
[c] The five-banded individual, 101, 104, 105, 107, 112, was not enumerated in this table

were not observed in the world sample. Examples include: allele 105 in Ethiopia and allele 108 in both Ethiopia and Tibet.

The level of diversity

Shannon's information statistic ($h_s = -\Sigma p_i \ln p_i$, where p_i is the frequency of the i^{th} rDNA phenotype) was calculated to evaluate the phenotypic diversity level of rDNA in barley from each region as represented by the samples. These calculations (Table 1) indicate that the world sample has the highest diversity value ($h_s = 1.317$), followed by the sample from China (0.803). The information statistic value for the Tibetan sample is not significantly higher than that of the Ethiopian sample (0.645 versus 0.601).

Contrasting patterns of rDNA variation in samples from different regions

It can be seen from Table 1 that there are two sharply contrasting patterns of rDNA genotypes in these four samples as illustrated by the frequencies of phenotypes 104, 112 and 107, 112. Tibet and Ethiopia appear to represent the two extremes of such contrasting patterns: phenotype 107, 112 is predominant in Tibetan barley ($f = 0.840$), but very rarely observed in the Ethiopian sample ($f = 0.034$); whereas phenotype 104, 112 occurs at a very high frequency in Ethiopian barley ($f = 0.864$), but is not observed at all in the Tibetan sample. The rDNA pattern of the world sample, characterized by a high frequency of the 104, 112 phenotype, is more similar to that of the Ethiopian barley sample, whereas the rDNA pattern of the Chinese sample, with 107, 112 as the most frequent phenotype, more closely resembles that of the Tibetan barley sample.

This is also the case for the distribution of the rDNA alleles among the four samples. As can be seen clearly from Table 2, the contrasting patterns of the rDNA genotypes discussed above are attributable to drastic differences in the frequencies of alleles 104 and 107 in the *Rrn2* locus among the four samples. Allele 104 is observed at an extremely high frequency in Ethiopian barley ($f = 0.957$), while 107 is the predominant allele of *Rrn2* in Tibetan barley ($f = 0.983$). These represent the extremes in allele frequency distribution. Allele 104 is the most frequent allele of this locus in the world sample ($f = 0{,}630$), and 107 occurs most commonly in the sample from China. Thus, the rDNA allelic compositions of these samples also supports the conclusion of greater similarity between the world and Ethiopian barley samples and between the Chinese and Tibetan barley samples.

Discussion

The 289 accessions included in the present study are likely to represent the widest range of genetic variation in cultivated barley. The results indicate that although a total of 17 rDNA phenotypes are present in these four samples, only 2, namely 104, 112 and 107, 112, occur at high frequencies. Two additional phenotypes, 104, 108 and 107, 111, 112, are observed at a substantial frequency only in environmentally extreme conditions (104, 108 in Ethiopia and 107, 111, 112 in Tibet). When the occurrence of individual alleles is followed, it is found that allele 112 predominates at the *Rrn1* locus worldwide, while at the *Rrn2* locus, allele 104 is most frequent in the world sample and in barley from Ethiopia, but 107 is the most common allele in barley from China and Tibet. Alleles 108 and 111 occur frequently in some localities of Ethiopia and Tibet, respectively.

In this connection, it is interesting to note that in a study of rDNA diversity in a large sample of 267 accessions of the wild barley *H. vulgare* ssp. *spontaneum* from Southwest Asia, Saghai Maroof et al. (1990) detected a total of 36 phenotypes composed of 17 alleles of the two rDNA loci. Phenotype 107, 112 was the most frequent phenotype ($f = 0.38$) in their sample; 4 additional phenotypes were also observed at substantial frequencies ($f \geq 0.05$). When alleles at each of these two loci were counted, it was found that allele 112 was the most frequent allele of the *Rrn1* locus ($f = 0.53$). However, 4 other alleles also occurred at frequencies between 0.05 and 0.19. Allele 107 was the most common allele of the *Rrn2* locus; 2 additional alleles were also observed at substantial frequencies. Thus, it is clear that this wild barley is much more variable with respect to rDNA genotypes than cultivated barley. It is also interesting to note that phenotype 104, 112 and allele 104, which occurred at high frequencies in cultivated barley, were rarely observed in this wild barley ($f < 0.01$ for phenotype 104, 112, and $f = 0.02$ for allele 104).

Vavilov (1926) and Freisleben (1940) suggested the existence of two distinct groups in world barley: an Oriental group and an Occidental group. Takahashi (1955) provided evidence of genetic differentiation between these two barley groups based on his genetic analyses of ear rachis brittleness and powdery mildew resistance. He placed barleys grown in China proper, southern Japan and South Korea in the Oriental group, and barleys from the rest of the world in the Occidental group. Similar but less clear-cut differentiation has also been observed in patterns of multilocus associations among enzyme loci (Kahler and Allard 1981; Zhang et al. 1990b). The present study shows that rDNA alleles and genotypes, as marked by the spacer-length variants, are clearly differentiated into these two groups: barley from China and Tibet, characterized by allele 107 or phenotype 107, 112,

falls into the Oriental group, while accessions from the rest of the world, with Ethiopia as the most extreme example, belong to the Occidental group.

An inspection of the sampling allocation of the accessions included in the present study indicate, that 11 accessions in the world sample are from China, Japan and Korea. Ten of these 11 accessions are phenotypically 107, 112. Allele 104 in the Chinese barley sample is mostly from two sources: spring barley landraces from Northeast China (formerly known as Manchuria), which Takahashi (1955) placed in the Occidental group, and cultivars from the Southeast coastal area where germ plasm from Western countries has been incorporated into barley breeding programs. When those entries are discounted, the rDNA pattern of the world sample is even more similar to that of Ethiopian barley, and allele 104 becomes very rare in the Chinese sample. These results suggest that Tibet be the center of Oriental barley and Ethiopia the center of Occidental barley.

Several issues have emerged from the results of the present study regarding the phylogeny and evolution of cultivated barley. There have been a number of hypotheses concerning the origin of cultivated barley. For example, Åberg (1940) postulated that the six-rowed wild barley, *H. vulgare* ssp. *agriocrithon*, found in Tibet was the progenitor of cultivated barley. Harlan (1976) maintained an opposite opinion that barley was first domesticated in Southwest Asia from a two-rowed wild barley *H. vulgare* ssp. *spontaneum*. Freisleben (1940), on the other hand, proposed a diphyletic hypothesis that six-rowed cultivated barley in the Oriental region was derived from *H. vulgare* ssp. *agriocrithon*, while two-rowed cultivated barley in Southwest Asia originated from *H. vulgare* ssp. *spontaneum*. The numerous other forms of cultivated barley are considered to be either direct descendants from one or the other ancestral forms, or derived from hybridizations between the two ancestral forms.

However, results from our previous study (Saghai Maroof et al. 1990) showed that at the *Rrn2* locus 107 is the most frequent allele (f=0.78) and 104 is very rare (f=0.02) among 267 accessions of *H. vulgare* ssp. *spontaneum* representing almost the entire distribution range of this wild barley in Southwest Asia. Thus, with any of the above hypotheses, either singly or in combination, it is difficult, in two respects, to explain the distribution patterns of the rDNA polymorphisms observed in these studies. First, if *H. vulgare* ssp. *spontaneum* in Southwest Asia was the progenitor of the cultivated barley, 107 should be expected to occur at the highest frequency in world barley, and allele 104 would have very little chance to be brought into cultivation. This, however, is not the case. Instead, allele 104, which occurred very rarely (f=0.02) in *H. vulgare* ssp. *spontaneum*, is the most frequent allele in the world sample of cultivated barley. It should be pointed out that the generation of one rDNA allele from another involves a series of molecular events including unequal crossing-over and the homogenization of several hundred to several thousand repeating units that would probably take thousands of years to accomplish unless there was extremely strong selection for such turn-over. It may also be difficult for allele 104 to replace allele 107 once the latter is established, since 107 seems to be favored under both favorable (e.g., Davis, California, Saghai Maroof et al. 1984) and unfavorable (e.g., Tibet in the present study) conditions. Thus, it is unlikely that the widely occurring 104 allele was derived from allele 107 after barley was domesticated. This observation appears to be at odds with the proposition that *H. vulgare* ssp. *spontaneum* is the immediate progenitor of the cultivated barley.

The second difficulty is that it is not apparent why the rDNA patterns of *H. vulgare* ssp. *spontaneum* are more similar to those of barley in the Oriental region than to those of cultivated barley in its immediate surroundings. Detailed information on the distribution of rDNA genotypes in cultivated barley across the continent of Asia and in wild barley of Tibet will certainly be helpful, and studies are currently in progress toward resolving such difficulties.

Acknowledgements. The authors thank Dr. S. VanScoyoc for critical reading of the manuscript. This work was supported in part by a grant from the International Foundation for Science (No. C/1654-1). MASM acknowledges contributions from the College of Agriculture and Life Sciences, VPI&SU and from the North American Barley Genome Mapping Program coordinated by Professor R.A. Nilan.

References

Åberg E (1940) The taxonomy and phylogeny of *Hordeum* L Sect. Cerealia Ands., with special reference to Tibetan barleys. Symb Bot Ups 4:1–156

Allard RW, Saghai Maroof MA, Zhang Q, Jorgensen RA (1990) Genetic and molecular organization of ribosomal DNA (rDNA) variants in wild and cultivated barley. Genetics 126:743–751

Freisleben R (1940) Die phylogenetische Bedeutung asiatischer Gersten. Zuchter 12:257–272

Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucleic Acids Res 7:1869–1885

Harlan JR (1976) Barley, pp 93–98 in *Evolution of Crop Plants*, edited by Simmonds NW. Longman, London and New York

Kahler AL, Allard RW (1981) Worldwide patterns of genetic variation among four esterase loci in barley (*Hordeum vulgare* L.). Theor Appl Genet 59:101–111

Saghai Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Saghai Maroof MA, Allard RW, Zhang Q (1990) Genetic diversity and ecogeographical differentiation among ribosomal DNA alleles in wild and cultivated barley. Proc Natl Acad Sci USA 87:8486–8490

Takahashi R (1955) The origin and evolution of cultivated barley. Adv Genet 7:227–266

Vavilov NI (1926) Studies on the origin of cultivated plants. Bull Appl Bot Genet Plant Breed USSR 16:1–248 (Russian with English summary)

Zhang Q, Saghai Maroof MA, Allard RW (1990a) Effects on adaptedness of variations in ribosomal DNA copy number in populations of wild barley (*Hordeum vulgare* ssp. *spontaneum*). Proc Natl Acad Sci USA 87:8741–8745

Zhang Q, Saghai Maroof MA, Allard RW (1990b) Worldwide pattern of multilocus structure in barley determined by discrete log-linear multivariate analyses. Theor Appl Genet 80:121–128