Theor Appl Genet (2003) 107:413–421
DOI 10.1007/s00122-003-1261-9

J. R. Russell · A. Booth · J. D. Fuller · M. Baum · S. Ceccarelli · S. Grando · W. Powell

Patterns of polymorphism detected in the chloroplast and nuclear genomes of barley landraces sampled from Syria and Jordan

Received: 30 July 2002 / Accepted: 20 December 2002 / Published online: 3 July 2003

Abstract In order to examine how molecular polymorphism in barley landraces, sampled from five different ecogeographical regions of Syria and Jordan, is organised and partitioned, genetic variability at 21 nuclear and 10 chloroplast microsatellite loci were examined. Chloroplast polymorphism was detected, with most variation being ascribed to differences between the five regions (F_{st} 0.45) and to within sites within each region (F_{st} 0.44). Moreover, the distribution of chloroplast polymorphism is structured and not distributed randomly across the barley landraces sampled. From a total of 125 landrace accessions (five lines from each of five sites from each of five regions) genotyped with 21 SSRs a total of 244 alleles were detected, of which 38 were common to the five regions sampled. Most nuclear variation was detected within sites. Significant differentiation between sites (F_{st} 0.29) was detected with nuclear SSRs and this partially mirrored polymorphism in the chloroplast genome. Strong statistical associations/interaction was also detected between the chloroplast and nuclear SSRs, together with non-random association (linkage disequilibrium) of alleles at both linked and unlinked SSR loci. These results are discussed in the context of adaptation of landraces to the extreme environment, the concept of 'adapted gene complexes' and the exploitation of landraces in breeding programmes.

Keywords Barley · Landraces · Microsatellites · Linkage disequilibrium

Communicated by P. Langridge

J. R. Russell · A. Booth · J. D. Fuller · W. Powell (✉)
Scottish Crop Research Institute (SCRI), Invergowrie,
Dundee DD2 5DA, Scotland, UK
e-mail: wpowel@scri.sari.ac.uk
Tel.: +44-1382-568512
Fax: +44-1382-562426

M. Baum · S. Ceccarelli · S. Grando
International Center for Agricultural Research in the Dry Areas (ICARDA), Aleppo, Syria

Introduction

Over the thousands of years of domestication and spread from their centres of origin, crops have become increasingly adapted to a wide range of environments, responding to various selection pressures, including, biotic, abiotic and human intervention. In a natural system a population's survival is affected by the population structure, mating system, life history characteristics and natural selection. In an agricultural system, human selection and management plays a major role in the generation and maintenance of intraspecific diversity and differentiation (Teshome et al. 2001). This is particularly significant in countries where traditional cultivars of major crops continue to be grown. Locally adapted barley cultivars or landraces are one such example and are one of the most important cereal crops in the Fertile Crescent, exhibiting the ability to survive in a range of ecological situations. Such landraces cover approximately five million hectares of marginal, drought-stressed environments (Ceccarelli 1994) and represent a major resource for germplasm conservation, breeding and targeted allele discovery (Ceccarelli et al. 1987). Under traditional agro-ecosystems, farmers often select plants with preferred agro-morphological characters and choose specific genotypes for planting in a certain microenvironment. There is now an increasing recognition of the importance of on-farm maintenance of locally adapted genotypes as a component of conservation efforts (UNEP 1994), together with the recognition of the role of farmers and biophysical heterogeneity as significant factors influencing the maintenance of genetic diversity. However, very little is known about the genetic structure of landrace populations, the amount of genetic variation they possess and how this relates to the success of landraces in marginal environments. Such information will not only play an important role in guiding breeding programmes for stressed environments but may also provide a better understanding of genomic diversity for the mode of adaptation of landraces to ecological factors and environmental stresses.

A large collection of barley landraces was made in 1981 in Syria and Jordan from farmers who had been using their own seed for generations (Weltzien 1988). Preliminary morphological evaluation revealed considerable variation between and within collection sites for many agronomically important characters (Ceccarelli et al. 1987) and disease reactions (van Leur et al. 1989). The exploitation of the diversity within barley landraces has proved to be a powerful means to improve barley yields in marginal environments (Ceccarelli 1996). However, despite the large effort devoted to the exploitation of barley landraces, little is known about the geographical partitioning of genetic diversity in this important genetic resource. This material also provides an opportunity to examine the impact of farmers' selection and agronomic practices on the generation of diversity in situ.

During the last decade, methods for detecting and analysing diversity have progressed from analysing discrete morphological characters, and pedigree data to molecular examination of nuclear and chloroplast DNA sequence variation and the interaction between both genomes (Koebner et al. 2001; Rafalski 2002a, b). There is at present a large body of data pertaining to nuclear and to a lesser-extent of chloroplast genetic diversity within barley (Clegg et al. 1984; Neale et al. 1988; Struss and Plieske 1998; Doebley et al. 1992; Graner et al. 1994; Russell et al. 1997, 2000; Schut et al. 1997; Marmiroli et al. 1999; Provan et al. 1999; Maestri et al. 2002); however, data is sparse concerning the level and pattern of diversity in natural populations (Jarvis and Hodgkin 2000). With the development of chloroplast and nuclear SSRs, we have an opportunity to update our knowledge on the distribution of genetic diversity between and within populations of barley, and the interactions between the two genomes. A total of 21 mapped SSR markers represented by five gene-specific and 16 anonymous genomic SSR markers spanning the seven linkage groups of barley were selected for study. Recent studies (Morgante et al. 2002) have shown that microsatellite frequency is higher in transcribed and low-copy regions of plant genomes, making them an even more attractive marker class for genetic analysis in plants.

The overall objectives of this study are twofold: to examine the level of diversity present in a collection of landraces adapted and collected from different ecogeographical areas with both chloroplast and nuclear SSRs, and to study the patterns of non-random association of alleles or linkage disequilibrium (LD) detected in these samples. Such an analysis will allow the overall global genetic pattern of polymorphism to be studied and considered, in relation to sites of geographical origin and previous agricultural and farmers' selection practises.

Materials and methods

Genetic material

The genetic material consisted of lines which were selected from a collection of barley landraces maintained at the International Center for Agricultural Research in the Dry Areas (ICARDA), Aleppo, Syria, originally collected by Dr. Weltzien (Weltzien 1988). The sampling strategy and evaluation details are given in Ceccarelli et al. (1987); and Weltzien (1988). For this particular study, five distinct regions were identified and 20 lines were selected from five sites within each region, with a total of 500 lines sampled (Fig. 1). The regions are designated as South Jordan (SJ), North Jordan and South Syria (NJSS), West Syria (WS), Central Syria (CS) and Northeast Syria (NES) (Salvatore Ceccarelli, personal communication).

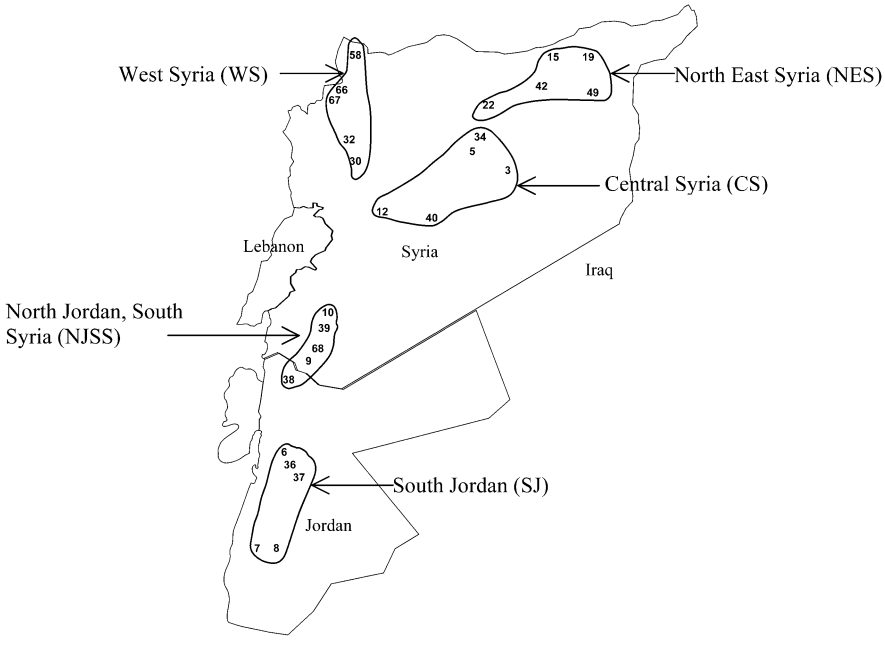

Fig. 1 Geographic location of five regions and the five sampling sites within each region

SSR analysis

DNA extraction, PCR, allele visiualisation and analysis were carried out as described in Russell et al. (1997). The twenty one nuclear SSR loci sampled in this study and their approximate location on a reference map are given in Ramsay et al. (2000). The ten chloroplast SSRs are described in Provan et al. (1999).

Sequence analysis

The hvcptrnLF locus was amplified using primers and conditions described in Taberlet et al. (1991). PCR products were purified using Exonuclease I and Shrimp Alkaline Phosphateae (USB corporation, Cleveland, Ohio, USA), and sequenced directly using the forward and reverse trnLF primers. Sequencing reactions were performed using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction kit (PE Applied Biosystems, Foster City, Calif., USA) and analysed on ABI 377 (PE Applied Biosystems, Foster City, Calif., USA).

Data analysis

Diversity values based on allele frequencies were calculated for each SSR locus using Nei's unbiased statistic (Nei 1987). Population differentiation was estimated using F_{ST} (Wright 1965) and partitioning of genetic diversity was obtained using the analysis of the molecular variance framework (AMOVA, Excoffier et al. 1992) on allelic frequencies. AMOVA analysis was performed using Arlequin software (Version 1.1, Scneider et al. 1997). In order to examine linkage disequilibrium (nonrandom association between alleles at different loci) between linked and unlinked loci across the genome, the standardised disequilibrium coeffecient (D′) for all pairs of polymorphic loci (Hedrick 1987) was estimated using Arlequin. Alleles were classified as common or non-common for each locus as described by Innan et al. (1997). This scaled-measure of LD is less affected by allele frequencies between sites than other measures of LD (Hedrick 1987; Jorde 2000; Remington et al. 2001).

Results

Only one (hvcptrnLF) of the ten chloroplast primer pairs tested were polymorphic, resulting in two length variants. Both variants were sequenced in a range of accessions from the five different regions using the universal primers designed by Taberlet et al. (1991), and the sequence alignments are shown in Fig. 2. Within the 850 bp that was sequenced in this region the mononucleotide length polymorphism correlated with the size of the amplification product generated and no other sequence polymorphism was observed. The distribution of these variants across regions is shown in Fig. 3, which highlights the structured regional variation associated with polymorphism in the chloroplast genome. Overall the data reveals a chloroplast cline characterised by a predominance of the 101-bp variant in Jordan and South Syria (SJ, NJSS). The 100-bp variant is predominant in Central, West and Northeast Syria (CS, WS, NES). These observations indicate that the distribution of chloroplast haplotypes in barley landraces sampled is structured and not distributed randomly.

Twenty one mapped nuclear SSRs were also used to examine the levels and patterns of variation in these five regions. A sample of 125 individuals were genotyped and a total of 244 alleles were detected, ranging from 2 (HVHVA1; HvLEU) to 34 (Bmac0067). Table 1 summarises the number of alleles and levels of diversity for each of the five regions. The total number of alleles per region differed ranging from 94 (South Jordan) to 136 (Northeast Syria) and the number of common alleles observed in all five regions was only 38. The level of diversity also varied between populations, ranging from 0.49 (South Jordan) to 0.63 (North Jordan and South Syria). However, the number of alleles and the level of diversity varied depending on the locus (the highest diversity values at each locus are shown in bold, Table 1). At locus Bmac0399 the number of alleles varied from 4 to 17 with an uneven allelic distribution,the dominance of smaller-sized alleles and less polymorphism in South Jordan (SJ), and North Jordan and South Syria (NJSS), compared to a wider distribution of a larger number of alleles in the Northeast and Central regions of Syria. A more marked example at locus Bmac0156 is highlighted in Table 1, with the virtual exclusion of alleles between South Jordan (SJ) and North East Syria (NES) regions; only two alleles from 19 are shared between the two genepools. In contrast, several loci show a similar number of alleles and allelic distributions with most regions sharing the most-frequent allele (e.g. HVM54; HVHVA1 and HvLEU) (Table 1). In general, the frequency distributions of alleles among the 21 loci fell into two patterns. The first pattern, in which the most-frequent allele was the same in all regions, was at locus HVHA1, HVM54, HvLTPPB, Bmac0029, Bmag0005, HvLEU, Bmac0040 and HvCMA. The second pattern observed was in which the most-frequent allele at each locus differed from one population to another, as displayed by locus Bamc0399, HVM20, Bmac0134, Bmag0125,

Fig. 2 Sequence variants in the hvcptrnLF region (850 bp) of the chloroplast genome; examples of sequence variants identified in accessions sampled from the five regions

Table 1 Summary of allelic and diversity values detected across all accessions from five different eco-geographical regions of Jordan and Syria

Accession	South Jordan			N. Jordan, S. Syria			West Syria			Central Syria			N.E. Syria			Total no. of alleles	Overall diversity	RST
	No. of alleles	Diversity	SD	No. of alleles	Diversity	SD	No. of alleles	Diversity	SD	No. of alleles	Diversity	SD	No. of alleles	Diversity	SD			
Bmac0 399	4	0.2906	0.0811	8	0.8392	0.0305	12	0.8784	0.0245	11	0.885024	0.0205	**17**	**0.9502**	**0.0104**	**24**	**0.8984**	0.19614
HVM20	2	0.2743	0.0709	**4**	**0.6661**	**0.0331**	2	0.085	0.055	2	0.0888	0.0572	2	0.1765	0.0736	**5**	**0.5618**	0.63054
HVHVA1	1	0.000	0.000	2	0.3722	0.0637	2	0.3722	0.0637	**2**	**0.383**	**0.0409**	2	0.2743	0.0709	**2**	**0.3032**	0.05696
Bmac0134	8	0.7536	0.0436	9	0.8196	0.0323	8	0.7804	0.0372	8	0.8121	0.0323	**10**	**0.8774**	**0.0197**	**17**	**0.8947**	0.08095
Bmag0125	3	0.4963	0.0515	6	0.7347	0.0447	6	0.7771	0.0306	**6**	**0.7553**	**0.0419**	5	0.7412	0.0366	**9**	**0.7953**	0.32304
HVM54	3	0.6269	0.0344	**5**	**0.7118**	**0.0499**	3	0.4016	0.0778	3	0.5496	0.0621	4	0.4255	0.0837	**5**	**0.5860**	0.0221
HvLTPPB	2	0.0974	0.0621	**4**	**0.5874**	**0.0687**	2	0.0784	0.051	2	0.0816	0.0529	1	0.000	0.000	**4**	**0.2867**	0.00018
Bmac0067	14	0.9391	0.0132	14	0.9082	0.0281	10	0.8972	0.0194	**13**	**0.9516**	**0.0153**	14	0.9429	0.0128	**34**	**0.9514**	0.16168
Bmag0013	6	0.6694	0.0469	7	0.7771	0.0407	5	0.7543	0.0292	**8**	**0.8327**	**0.0269**	6	0.7804	0.0274	**11**	**0.8042**	0.03517
Bmac0029	5	0.7184	0.0354	7	0.5747	0.0774	5	0.4669	0.0811	5	0.5649	0.0738	**7**	**0.7576**	**0.0477**	**9**	**0.6595**	0.15435
HvOLE	2	0.5029	0.022	**5**	**0.5745**	**0.0712**	3	0.2305	0.0771	2	0.1560	0.0666	2	0.1502	0.0646	**6**	**0.4319**	0.20907
HVM3	11	0.8584	0.0357	11	0.8359	0.04	8	0.8288	0.0338	**12**	**0.9237**	**0.0129**	10	0.9078	0.0146	**24**	**0.9396**	0.29711
HVM67	3	0.431	0.0663	**4**	**0.6694**	**0.0377**	5	0.500	0.0665	3	0.5682	0.056	3	0.5714	0.0567	**5**	**0.6528**	0.14198
Bmag0005	**5**	**0.6498**	**0.0517**	6	0.5682	0.0751	4	0.2906	0.0811	4	0.3787	0.0861	5	0.6241	0.0677	**8**	**0.5457**	0.11355
Bmac0113	6	0.7935	0.0329	7	0.7872	0.0262	5	0.6312	0.0525	6	0.7652	0.0343	**7**	**0.8077**	**0.0285**	**9**	**0.8040**	0.12606
HvLEU	1	0.000	0.000	**2**	**0.0784**	**0.051**	**2**	**0.0784**	**0.051**	1	0.0000	0.000	1	0.000	0.000	**2**	**0.0315**	0.01051
Bmac0316	4	0.7641	0.0134	9	0.6694	0.0698	4	0.4865	0.067	6	0.6531	0.0649	**9**	**0.8294**	**0.0315**	**13**	**0.7588**	0.09233
Bmac0018	2	0.1502	0.0646	2	0.3265	0.0685	**4**	**0.6237**	**0.0456**	3	0.5426	0.0604	4	0.5709	0.0697	**5**	**0.6026**	0.04605
Bmac0040	7	0.6563	0.0663	10	0.8511	0.0288	9	0.8457	0.0261	**14**	**0.9306**	**0.0125**	10	0.8457	0.0294	**18**	**0.8794**	0.03246
HvCMA	1	0.000	0.000	2	0.0784	0.051	2	0.1502	0.0646	**3**	**0.5284**	**0.0415**	2	0.422	0.056	**3**	**0.3923**	0.37278
Bmac0156	4	0.6171	0.0368	9	0.7865	0.0477	11	0.8522	0.0334	14	0.9184	0.0178	**15**	**0.9371**	**0.0121**	**31**	**0.9132**	0.47221
Unique genotypes	**25**			**24**			**25**			**25**			**25**					
Total	**94**			**133**			**112**			**128**			**136**			**244**		**0.32048**
Mean	**4.48**	**0.49**		**6.33**	**0.63**		**5.33**	**0.52**		**6.10**	**0.58**		**6.48**	**0.60**			**0.6520**	

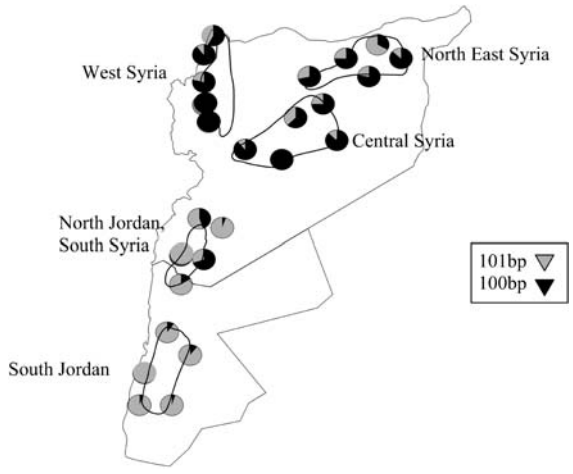

Fig. 3 Chloroplast haplotype distribution throughout the landrace growing regions of Jordan and Syria. Each 'pie' represents up to 20 lines, the total number of lines in 448

Bmag0013, HvOLE, HVM3, HVM67, Bmac0113, Bmac0316, Bmac0018 and Bmac0156.

Estimates of population differentiation for the chloroplast genome (Table 2a) are highly significant ($P < 0.001$) for all items included in the analysis. Most variation can be ascribed to differences between regions (F_{st} 0.45) and to within-populations (F_{st} 0.44). A smaller but significant F_{st} value (0.12) can be ascribed to differences between sampling sites within regions. Similarly the population differentiation for the nuclear SSRs was highly significant ($P < 0.001$), with an F_{st} value of 0.29 between the regions (Table 2b). For both chloroplast and nuclear genomes a significant level of variation was partitioned between individuals within sites. The highest level of differentiation for the nuclear genome was between South Jordan (SJ) and Northeast Syria (NES), with an F_{st} value of 0.6778; and the lowest was the comparison between North Jordan, South Syria (NJSS) and West Syria (WS), with a F_{st} value of 0.027. The highest F_{st} values are commonly between South Jordan (SJ), North Jordan and South Syria (NJSS), West Syria (WS) and Central Syria (CS), and Northeast Syria (NES), splitting the regions into two large zones.

This geographical differentiation at both the nuclear and the chloroplast level can be illustrated by plotting a genetic-distance tree generated by the nuclear markers with the distribution of chloroplast haplotypes superimposed (Fig. 4). There is a clear differentiation between the zones based on all of the nuclear data and this is partially mirrored with the distribution of the chloroplast haplotype data.

In order to investigate the extent of linkage disequilibrium using multiallelic SSRs, alleles were classified as common or non-common at each locus. Linkage disequilibrium between pairs of loci was tested using chi-square. Of 231 pairwise comparisons 75 were significant at the 5% level, this value was higher than expected by chance, which is not unexpected in a selfing plant with low rates of recombination. Interestingly, when we examined the distribution of significant associations between loci, we observed that many of these were between loci, which

Table 2a Analysis of molecular variance (AMOVA) and pairwise F_{st} values, for chloroplast haplotypes

Source of variation	df	Sum of squares	Variance	% Variation	
Between regions	4	47.14	0.12417	44.93	
Between sites within regions	20	13.83	0.03201	11.58	
Within sites	423	50.83	0.12018	43.49	
Total	447	111.81	0.27635		
Populations	SJ	NJ, SS	WS	CS	NES
SJ	–				
NJ, SS	0.14	–			
WS	0.77	0.49	–		
CS	0.75	0.46	–0.01	–	
NES	0.61	0.29	0.06	0.04	–

Table 2b Analysis of molecular variance (AMOVA) and pairwise F_{st} values, for nuclear SSRs

Source of variation	df	Sum of squares	Variance	% Variation	
Between regions	4	133,600.58	590.51264	29.53	
Between sites within regions	20	77,490.24	273.9344	13.7	
Within sites	225	255,412.8	1,135.168	56.77	
Total	249	466,503.62	1,999.615		
Populations	SJ	NJ, SS	WS	CS	NES
SJ	–				
NJ, SS	0.187	–			
WS	0.187	0.027	–		
CS	0.397	0.134	0.081	–	
NES	0.678	0.469	0.393	0.170	–

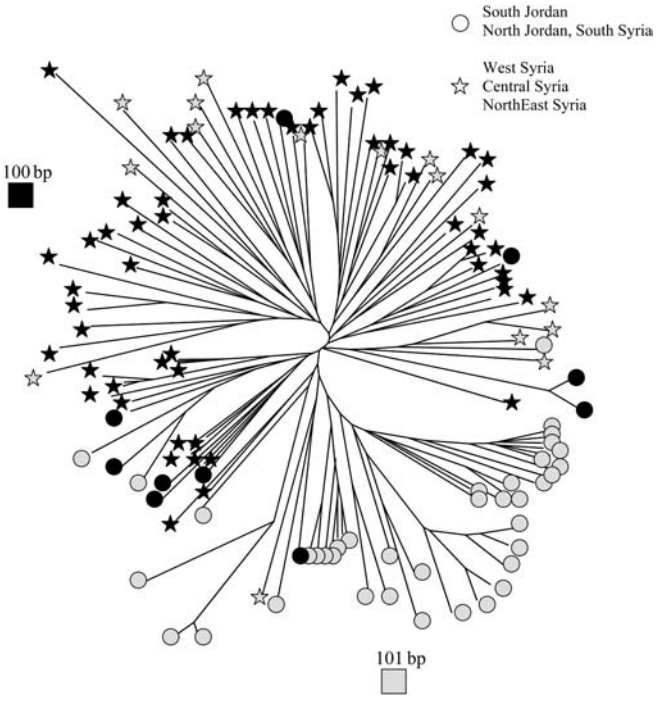

Fig. 4 Cluster analysis of 21 nuclear SSRs split into two groups; *circles* representing accessions from Jordan and South Syria (SJ, NJSS) and *stars* representing West, Central and Northeast Syria (WS, CS, NES), with chloroplast variants superimposed in *black* (100-bp variant) and *grey* (101-bp variant)

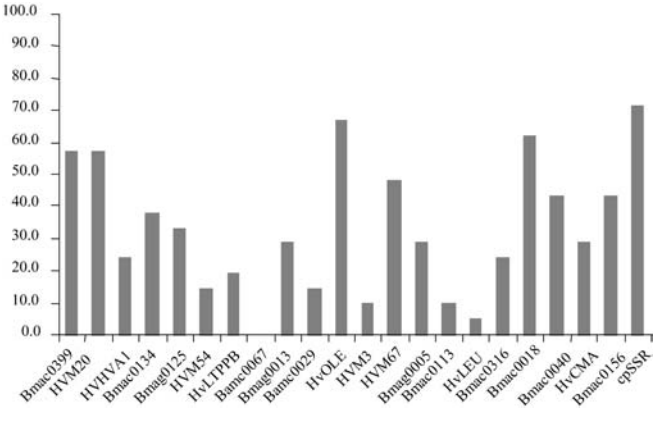

Fig. 5 Percent of significant linkage disequilibrium values detected between loci across the genome

were not physically linked, and conversely loci on the same chromosome did not show a significant association. Figure 5 shows the percentage of significant LD detected for each locus; for example the frequency of the significant pairwise comparison for HvOLE is 66.7%, with 14 of 20 comparisons showing a significant association; in contrast, Bmac0067 shows no significant pairwise LD at the 5% level. The cpSSR showed significant association with 15 of the 20 loci studied; those loci for which no significant association was observed included four loci on 3H and two on 5H. Linkage disequilibrium

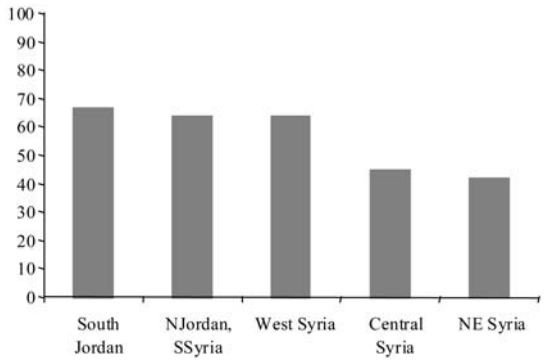

Fig. 6 Comparison of the number of significant pairwise LD association detected between the five regions (SJ = 67, NJSS = 64, WS = 64, CS = 45, NES = 42)

was also measured within each ecogeographic region and Fig. 6 highlights the extent of associations detected at each locus across the five different regions. Overall, the greatest number of significant disequilibria were within SJ, NJSS and WS regions (67, 64 and 64 respectively), in comparison to 45 and 42 significant associations in CS and NES regions.

Discussion

Since techniques for detecting polymorphism at the DNA level have been developed, many have been used to examine genetic diversity within important agricultural species and their wild relatives (reviewed by Buckler and Thornsberry 2002). Such surveys provide information on genomic diversity, domestication and evolution, identify geographic regions, which contain high levels of diversity, and discriminate between groups of similar accessions. We have analysed variation in a range of barley landraces from different ecogeographical regions in Syria and Jordan using previously mapped nuclear SSR loci and a chloroplast SSR locus. These data clearly illustrate a number of important features concerning genetic variation within and between these ecogeographically diverse regions.

There is a substantial amount of genetic variation within all of the regions included in this study. The majority of the loci are highly polymorphic, with diversity indices ranging from 0.0315 to 0.9514. Zhang et al. (1993) using RFLPs found similar high levels of polymorphism between and within populations of wild barley (*Hordeum vulgare* ssp. *spontaneum*) sampled from populations in Iran and Israel, with levels of diversity ranging from 0.153 to 0.665. Brown and Munday (1982) using isozyme loci found Iranian landrace variation intermediate between wild barley and cultivated barley. In a recent study on European spring barley cultivars, using a similar set of mapped SSRs (12 SSRs in common between the two studies), the levels of diversity were similar, with values ranging between 0.270 and 0.860 (Russell et al. 2000). Although the levels of diversity between the landraces and the spring genepools were similar, the allelic composition was very different for

those 12 common SSRs; from a total of 172 alleles only 16 were shared between the two genepools (Russell, unpublished data). Bjornstad et al. (1997), compared cultivated barley from Europe, North America and Japan with Ethiopian landraces, and found that the Ethiopian germplasm was significantly less diverse than the cultivated germplasm, but that it was also genetically more distinct.

Our results indicate that the five regions differ, considerably from each other both in genetic composition and in the amount of genetic variability detected with nuclear and chloroplast loci. A remarkable feature of allelic distribution is the sharp contrast between allelic frequencies throughout the regions, and the overall mosaic pattern of allele occurrence. This is evident in the low number of shared alleles; out of 244 only 38 are shared between all regions. Schoen and Brown (1991) view the variation among populations in a predominately selfing species, such as barley, as being composed of a number of historically and genetically diverse populations from which marginal populations of limited genetic diversity have been derived. In contrast, Papa et al. (1998) found greater variation within populations than between, which they explained by the exchange of seed between farmers in Sardinia. Zhang et al. (1990) compared genepools from different geographical regions of the world and observed that both the multilocus genotypes and the structure of marker associations differed markedly in different regions. Enjalbert et al. (1999) have used RFLP markers to compare allele frequencies in wheat populations that had been under natural selection and found strong differentiation between populations, leading to the conclusion that selection greatly influences the evolution of the populations. Similarly, Perez de la Vega et al. (1994) observed different multilocus genotypes of *Avena sativa* in different eco-geographical regions of Spain. Zhang et al. (1993) suggest that such a degree of differentiation between regions and taxa (*H. vulgare* and *H. vulgare* ssp. *spontaneum*) would indicate that domestication and subsequent adaptation have occurred not only as a process of selection of alleles at individual loci, but that these processes have reorganised the genome into multilocus genotypes adapted to different environments.

According to Allard (1999), inbreeding is expected to be efficient in organising favourable alleles, many located on different chromosomes, into combinations that permit superior performance. With regional germplasm pools undergoing selection and or drift independently, disequilibria could persist indefinitely. Self-pollination predominates in cereals and their wild relatives, and tend to maintain disequilibria among unlinked loci even within individual populations. (Allard et al. 1972). These so called "adapted gene complexes" and the assembly of favourable interacting alleles into synergistic complexes arising from gametic-phase disequilibrium is highly relevant to the evolution of landraces. The percentage of significant associations (32.2%) was higher in this study, than expected by random occurrence, and SSRs on the same chromosomes did not necessarily appear to be in linkage disequilibrium. Similarly, Innan et al. (1997) detected a higher than expected linkage disequilibrium between loci, approximately 12% in natural populations of *Arabidopsis thaliana*, and they also observed a significant LD between pairs of unlinked loci. Kraft et al. (2000) noted a low but significant level of linkage disequilibrium between non-linked markers in sugar beet, using AFLPs and RFLPs. With both types of markers they observed a similar pattern of LD, with an increase in LD only at distances of less than 3 cM. Although the SSR loci used in this study are all mapped to specific locations in the barley genome, they are too-far apart to examine LD to the same extent as the sugar beet study (the closest markers are HvLTPPB and Bmac0067 on 3H which are 29 cM apart and show no significant LD). In a recent review by Buckler and Thornsberry (2002), the authors stated that the main obstacle to successful association studies in plants is population structure, more specifically subgroups within populations. By analysing LD within each region, we have been able to examine associations without population structure or, in this case, regional structure which we observed with both nuclear and chloroplast SSRs. The number of significant LDs varied across the regions ranging from 67 (South Jordan) to 42 (North East Syria). Different loci showed significant associations at different regions, for example on 1H Bmac0399 showed a significant LD with HVM20 in South Jordan and in North East Syria, but not in the other three regions. Although beyond the scope of this study it may be possible to evaluate sequence-based diversity across natural populations and correlate polymorphism with phenotypic variation. Recently co-adapted gene complexes were described in rice landraces (Ford-Lloyd et al. 2001). These authors conclude that although migration and inbreeding has given rise to linkage disequilibrium it could not fully account for the extent and nature of the associations detected. Owour et al. (1997) also found that multilocus natural selection in wild barley populations in Israel maintained a high degree of linkage disequilibria. Similar proportions of significant linkage disequilibria were observed in a study of AFLP variation in *Arabidopsis thaliana* populations (Miyashita et al. 1999). The high degree of selfing in *Arabidopsis* can also lead to LD extending over 250 kb or approximately 1 cM (Nordborg et al. 2002). Recent studies of LD in livestock (Farnir et al. 2000; McRae et al. 2002) have identified substantial levels of LD between pairs of unlinked markers. If such associations were caused by linkage of the unlinked markers to one or more relevant traits, then this information would be valuable in identifying sequences useful as markers for quantitative agronomic traits. However, determining and validating the pattern of association between loci on different chromosomes is also important for a better understanding of the evolution of landraces and for the conservation and exploitation of genetic resources.

Chloroplast variation was found both within and among populations of landraces from Syria and Jordan. The chloroplast lineage does not appear to be randomly

distributed, the 101-bp variant is predominant in Jordan and South and West Syria and the 100-bp variant in Central and Northeast Syria. Neale et al. (1988) also found a non-random distribution of chloroplast haplotypes within and among populations of wild barley from Israel and Iran. Saghai-Maroof et al. (1992) observed significant associations between nuclear and cytoplasmic loci in wild barley populations. These associations were observed between populations, with complete equilibrium observed between chloroplast and nuclear markers within populations. These authors suggest that genes favoured in the nuclear and chloroplast genomes can be selected simultaneously with changes in environmental conditions, although they suggest that other forces such as genetic drift and founder effects may be responsible for the strong association observed.

Schoen and Brown (1991) have suggested that self-pollinating species, which are related to crop plants, consist of hot spots of genetic variation and as such these hot spots may be relevant to the conservation of genetic resources. By examining molecular patterns of diversity in landrace populations we can begin to facilitate how such populations can be represented in germplasm collections and also be exploited. The second important feature of this study is the need to consider genome-wide patterns of polymorphism and their interactions. The concept of adapted gene complexes (Allard et al. 1972) is particularly relevant to the evolution of landraces and the adaptation of populations to local environments. It must be stressed that LD can also be caused by population admixture and structure attributable to coancestry. Although the multi allelic D′ estimator of LD (Lewontin 1964) is assumed to be independent of allele frequency (Zapata 2000), in practice such an assumption is difficult to accommodate with highly polymorphic microsatellites. Despite these caveats, the so called 'non-syntenic marker pairs' exhibiting LD, and identified in this study, may have biological relevance since they could represent historical selection pressure for multiple unlinked traits. This area of research is of particular relevance to inbreeders, and validating the concept of adapted gene complexes is an important challenge for future research. Further efforts combining computer simulation and empirical studies focussing on the use of mapped molecular polymorphisms, together with sequence-based haplotyping, will provide an opportunity to validate the nature, biological basis and significance of gametic-phase disequilibrium in crop plants.

References

Allard RW (1999) History of plant population genetics. Annu Rev Genet 33:1–27

Allard RW, Babbel GR, Clegg MT, Kahler AL (1972) Evidence for coadaptation in *Avena barbata*. Proc Natl Acad Sci USA 69:3043–3048

Bjornstad A, Demissie A, Kilian A, Kleinhofs A (1997) The distinctness and diversity of Ethiopian barleys. Theor Appl Genet 94:514–521

Brown AHD, Munday J (1982) Population-genetic structure and optimal sampling of land races of barley from Iran. Genetica 58:85–96

Buckler ES, Thornsberry JM, (2002) Plant molecular diversity and applications to genomics. Curr Opin Plant Biol 5:107–111

Ceccarelli S (1994) Specific adaptation and breeding for marginal conditions. Euphytica 77:205–219

Ceccarelli S (1996) Adaptation to low high-input cultivation. Euphytica 92:203–214

Ceccarelli S, Grando S, van Leur JAG (1987) Genetic diversity in barley landraces from Syria and Jordan. Euphytica 36:389–405

Clegg M, Brown AHD, Whitfeld PR (1984) Chloroplast DNA diversity in the wild and cultivated barley: implications for genetic conservation. Genet Res 43:339–343

Doebley J, Von Bothmer R, Larson S (1992) Chloroplast DNA variation and phylogeny of *Hordeum* (*Poaceae*). Am J Bot 79:576–584

Enjalbert J, Goldringer I, Paillard S, Brabant P (1999) Molecular markers to study genetic drift and selection in wheat populations. J Exp Bot 50:283–290

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Farnir F, Coppieters W, Arranz JJ, Berzi P, Cambisano N, Grisart B, Karim L, Marcq F, Moreau L, Mni M, Nezer C, Simon P, Vanmanshoven P, Wagenaar D, Georges M (2000) Extensive genome-wide linkage disequilibrium in cattle. Genome Res 10:220–227

Ford-Lloyd BV, Newbury HJ, Jackson MT, Virk PS (2001) Genetic basis for co-adaptive gene complexes in rice (*Oryza sativa* L.) landraces. Heredity 87:530–536

Graner A, Ludwig WF, Melchinger AE (1994) Relationships among European barley germplasm. 2. Comparison of RFLP and pedigree data. Crop Sci 34:1199–1205

Hedrick PW (1987) Gametic disequilibrium measures: proceed with caution. Genetics 117:331–341

Innan H, Terauchi R, Miyashita NT (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics 146:1441–1452

Jarvis DI, Hodgkin T (2000) Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agro-ecosystems. Mol Ecol 8:5159–5173

Jorde LB (2000) Linkage disequlibrium and the search for complex disease genes. Genome Res 10:1435–1444

Koebner RMD, Powell W, Donini P (2001) Contributions of DNA molecular marker technologies to the genetics and breeding of wheat and barley. In: Janick J (ed) Plant breeding reviews. John Wiley and Sons. Vol 21:181–212, ISBN 0-471-41847-1

Kraft T, Hansen M, Nilsson NO (2000) Linkage disequilibrium and fingerprinting in sugar beet. Theor Appl Genet 101:323–326

Leur van JAG, Ceccarelli S, Grando S (1989) Diversity for disease resistance in barley landraces from Syria and Jordan. Plant Breed 103:324–335

Lewontin RC (1964) The interaction of selection and linkage. I. General considerations: heterotic models. Genetics 49:49–67

Maestri E, Malcevschi A, Massari A, Marmiroli N (2002) Genomic analysis of cultivated barley (*Hordeum vulgare*) using sequence-tagged molecular markers. Estimates of divergence based on RFLP and PCR markers derived from stress-responsive genes, and simple-sequence repeats (SSRs). Mol Gen Genomics 267:186–201

Marmiroli N, Maestri E, Liviero L, Massari A, Malcevschi A, Monciardini P (1999) Application of genomics in assessing biodiversity in wild and cultivated barley. Mol Ecol 8:S95–S106

McRae AF, McEwan JC, Dodds KG, Wilson T, Crawford AM, Slate J (2002) Linkage disequilibrium in domestic sheep. Genetics 160:1113–1122

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with the non-repetitive DNA in plant genomes. Nature Genet 30:194–200

Miyashita NT, Kawabe A, Innan H (1999) DNA variation in the wild plant *Arabidopsis thaliana* revealed by amplified fragment length polymorphism analysis. Genetics 152:1723–1731

Neale DB, Saghai-Maroof MA, Allard RW, Zhang Q, Jorgensen RA (1988) Chloroplast DNA diversity in populations of wild and cultivated barley. Genetics 120:1105–1110

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA, Weigel D (2002). Nature Genet 30:190–193

Owuor E, Fahima T, Beiles A, Korol A, Nevo E (1997) Population genetic response to microsite ecological stress in wild barley, *Hordeum spontaneum*. Mol Ecol 6:117–1187

Papa R, Attene G, Barcaccia G, Ohgata A, Konishi T (1998) Genetic diversity in landrace populations of *Hordeum vulgare* L. from Sardinia, Italy, as revealed by RAPDs, isozymes and morphophenological traits. Plant Breed 117:523–530

Perez de la Vega M, Saenz-de-Miera LE, Allard RW (1994) Ecogeographical distribution and differential adaptedness of multilocus allelic associations in Spanish *Avena sativa* L. Theor Appl Genet 88:56–64

Provan J, Russell JR, Booth A, Powell W (1999) Polymorphic chloroplast simple sequence repeat primers for systematic and population studies in the genus *Hordeum*. Mol Ecol 8:505–511

Rafalski AJ (2002a) Application of single nucleotide polymorphism in crop genetics. Curr Opin Plant Biol 5:94–100

Rafalski AJ (2002b) Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Sci 162:329–333

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

Remington DL, Thornsberry JM, Matsouka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES IV (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc Natl Acad Sci USA 98:11,479–11,484

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

Russell JR, Ellis RP, Thomas WTB, Waugh R, Provan J, Booth A, Fuller JD, Lawrence P, Young G, Powell W (2000) A retrospective analysis of spring barley germplasm development from 'foundation genotypes' to currently successful cultivars. Mol Breed 6:553–568

Saghai Maroof MA, Zhang Q, Neale DB, Allard RW (1992) Associations between nuclear loci and chloroplast DNA genotypes in wild barley. Genetics 131:225–231

Schoen DJ, Brown AH (1991) Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. Proc Natl Acad Sci USA 88:4494–4497

Schut JW, Qi X, Stam P (1997) Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. Theor Appl Genet 95:1161–1168

Scneider SJ, Kueffer M, Roessili D, Excoffier L (1997) ARLEQUIN ver 1.1: a software for population genetic data analysis. Genetic and Biometry Laboratory. University of Geneva, Geneva

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

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109

Teshome A, Brown AHD, Hodgkin T (2001) Diversity in land races of cereal and legume crops. In: Janick J (ed) Plant breeding reviews, vol 21, pp 221–261

UNEP (1994) Convention on biological diversity. UNEP, Geneva, Switzerland

Weltzien E (1988) Evaluation of barley (*Hordeum vulgare* L.) landrace populations originating from different growing regions in the Near East. Plant Breed 101:95–106

Wright S (1965) The interpretation of population structure by *F*-statistics with special regard to systems of mating. Evolution 19:395–420

Zapata C (2000) The D′ measure of overall gametic disequilibrium between pairs of multiallelic loci. Evolution 54:1809–1812

Zhang QF, Saghai Maroof MA, Allard RW (1990) 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 QF, Saghai-Maroof MA, Klienhofs A (1993) Comparative diversity analysis of RFLPs and isozymes within and among populations of *Hordeum vulgare* ssp. *spontaneum*. Genetics 134:909–916