Euphytica (2008) 161:35–45
DOI 10.1007/s10681-007-9508-1

Mapping adaptation of barley to droughted environments

J. Comadran · J. R. Russell · F. A. van Eeuwijk · S. Ceccarelli · S. Grando · M. Baum · A. M. Stanca · N. Pecchioni · A. M. Mastrangelo · T. Akar · A. Al-Yassin · A. Benbelkacem · W. Choumane · H. Ouabbou · R. Dahan · J. Bort · J.-L. Araus · A. Pswarayi · I. Romagosa · C. A. Hackett · W. T. B. Thomas

Received: 13 March 2007 / Accepted: 10 July 2007 / Published online: 3 August 2007

Abstract Identifying barley genomic regions influencing the response of yield and its components to water deficits will aid in our understanding of the genetics of drought tolerance and the development of more drought tolerant cultivars. We assembled a population of 192 genotypes that represented landraces, old, and contemporary cultivars sampling key regions around the Mediterranean basin and the rest of Europe. The population was genotyped with a stratified set of 50 genomic and EST derived molecular markers, 49 of which were Simple Sequence Repeats (SSRs), which revealed an underlying population sub-structure that corresponded closely to the geographic regions in which the genotypes were grown. A more dense whole genome scan was generated by using Diversity Array

J. Comadran · J. R. Russell · W. T. B. Thomas (✉)
SCRI, Invergowrie, Dundee DD2 5DA, UK
e-mail: Bill.Thomas@scri.ac.uk

F. A. van Eeuwijk
Laboratory of Plant Breeding/Applied Statistics-Biometris, PO Box 16, 6700 A8 Wageningen, The Netherlands

S. Ceccarelli · S. Grando · M. Baum
International Center for Agricultural Research in the Dry Areas (ICARDA), PO Box 5466, Aleppo, Syria

A. M. Stanca
CRA–Istituto Sperimentale per la Cerealicoltura, I-29017 Fiorenzuola d'Arda (PC), Italy

N. Pecchioni
Dipartimento di Scienze Agrarie, Università di Modena e Reggio Emilia, I-42100 Reggio Emilia, Italy

A. M. Mastrangelo
CRA–Istituto Sperimentale per la Cerealicoltura, I–71100 Foggia, Italy

T. Akar
CRIFC, TR-06042 Ankara, Turkey

A. Al-Yassin
NCARTT, PO Box 639, Amman 19381, Jordan

A. Benbelkacem
ITGC, 25100 Constantine, Algeria

W. Choumane
Tishreen University, Latakia, Syria

H. Ouabbou · R. Dahan
INRA Morocco, 1000 Rabat, Morocco

J. Bort · J.-L.Araus
Department de Biologia Vegetal, Universitat de Barcelona, 08028 Barcelona, Spain

A. Pswarayi · I. Romagosa
Centre UdL-IRTA, Universitat de Lleida, 25196 Lleida, Spain

C. A. Hackett
BioSS, SCRI, Invergowrie, Dundee DD2 5DA, UK

Technology (DArT®) to generate 1130 biallelic markers for the population. The population was grown at two contrasting sites in each of seven Mediterranean countries for harvest 2004 and 2005 and grain yield data collected. Mean yield levels ranged from 0.3 to 6.2 t/ha, with highly significant genetic variation in low-yielding environments. Associations of yield with barley genomic regions were then detected by combining the DArT marker data with the yield data in mixed model analyses for the individual trials, followed by multiple regression of yield on markers to identify a multi-locus subset of significant markers/QTLs. QTLs exhibiting a predefined consistency across environments were detected in bins 4, 6, 6 and 7 on barley chromosomes 3H, 4H, 5H and 7H respectively.

Keywords Barley · Landraces · Association mapping · Drought · QTLs · Yield

Introduction

In terms of world production, barley is the fourth most important small-grained cereal crop but it is widely cultivated in marginal environments where it is less susceptible to the abiotic stresses, e.g. drought and cold, encountered. Nevertheless, abiotic stresses are a major factor limiting barley production in many Mediterranean environments (Ceccarelli and Grando 1996) and dissection of the genetic mechanisms underlying the response to such stresses is the key to a focused approach to improving the crop's resistance to them. This is likely to become an even more important area of research in view of climatic changes which, in a number of Mediterranean countries, are already causing farmers to change cropping from wheat to barley due to the latter's greater abiotic stress tolerance.

Drought is interpreted in different ways by different disciplines but it is clear that an integrated approach is required to design and interpret appropriate experiments to study drought response and that results be applicable to field-grown crops (Passioura 2007). Furthermore, the ability to include measurements of plant water status in such experiments enables plant physiologists to hypothesise what mechanisms are involved and thus identify some likely candidate genes (Jones 2007).

Yield in stressed Mediterranean environments has been the subject of several QTL studies that have each detected several QTLs. Three crosses (Tadmor × ER/APM (Teulat et al. 2001); Barke × HOR11508 (Talame et al. 2004) and Derkado × B83-12/21/5 (Ellis et al. 2002)) were studied as part of an EU funded project 'Stable yields in Mediterranean barley: application of molecular technologies in improving drought tolerance and mildew resistance' (ERBIC18-CT98-0311) but there was little evidence of QTLs being co-located across the separate populations. A complicating factor in these crosses was the segregation of major developmental genes in each cross, which either accounted for much of the phenotypic variation or were not included in the maps, making it difficult to determine whether or not the detected loci are robust or transient (Forster et al. 2004). In addition, although some physiological traits were also mapped in the Tadmor × ER/APM population, none of the QTLs for these traits were co-located with the yield and agronomic QTLs. More recently, association genetics has been used to identify marker QTL associations for yield in a range of adapted barley germplasm (Kraakman et al. 2004). This approach is of great interest to plant geneticists as it offers the prospects of identifying QTLs that are likely to be robust across a broad range of germplasm. Additionally the use of an appropriate model to account for any underlying population sub-structure (Pritchard et al. 2000) can largely account for major developmental genes and thus offer the prospect of identifying QTLs that are genuinely associated with yield in stressed environments.

This paper describes the analysis of phenotypic and genotypic data for a diverse set of germplasm, with the phenotypic data being collected from yield trials grown across a range of Mediterranean environments subject to varying levels of stress. The study formed part of an EU INCO-MED funded project called 'Mapping Adaptation of Barley to Droughted Environments' (ICA3-CT2002-10026), which combined the above data with environmental and physiological data in a detailed investigation of the response of four different barley populations (two bi-parental mapping populations and two germplasm collections) to stress. For the purposes of this paper, we have conducted analyses of the associations between molecular markers and the yield data from diverse barley germplasm collection.

Materials and methods

Germplasm

The 192 accessions used in this study comprised selected key landraces, old (pre 1980) and more modern (post 1980) cultivars from four distinct regions around the Mediterranean rim and elsewhere (Table 1) that formed a collection that we have called Diverse Barley Germplasm (DBG). We have used the phrase 'Germplasm Origin' to refer to the five groupings listed in Table 1. The genotypes were multiplied at the ICARDA field site in Tel Hadya, Syria for harvest year 2003 to produce sufficient seed for trialling in subsequent years. A further sample from the multiplication was sent to SCRI where each genotype was grown in a disease-free environment in a glasshouse and leaf tissue sampled for extraction of DNA using a TissueLyser with DNeasy kits (www1.qiagen.com).

Genotyping

Each entry was genotyped with a stratified set of 49 genomic and EST derived Simple Sequence Repeat (SSR) and one Single Nucleotide Polymorphism markers that gave good coverage of the barley genome (Russell et al. 2004). The genotypic data was then used to identify underlying sub-structure amongst the genotypes using STRUCTURE (Pritchard et al. 2000). Phenotypic and STRUCTURE analysis revealed that four accessions produced unexpected genotypes, probably due to selection of contaminants at some stage in the selection and multiplication process, and these were eliminated from subsequent genotyping analyses. Two other accessions were duplicate entries and these were also eliminated from subsequent genotyping analyses. The best grouping that also made biological sense corresponded closely to the original Germplasm Origin groupings in Table 1 (Comadran unpublished data) and therefore these groupings were used in subsequent analyses. DNA samples from the 186 correctly identified and unique accessions were also genotyped by Diversity Array Technology (DArT) (Wenzl et al. 2004) by Triticarte (www.triticarte.com.au). This yielded an additional 1,130 polymorphic markers, just over 800 of which were located on the DArT consensus map (www.biomedcentral.com/content/supplementary/1471-2164-7-206-s6.pdf). Overall map coverage was good but there were some gaps of greater than 20 cM and there was a notable lack of markers on chromosome 4H. The total length of the DArT consensus map is 1,158 cM, resulting in a marker density of one every 1.44 cM. Over the whole DBG, significant linkage disequilibrium persisted on average up to 3 cM (Comadran unpublished data) and we therefore considered that we had sufficient markers for a reasonably dense genome wide scan of marker trait associations.

Table 1 Composition of the population of Diverse Barley Genotypes based upon cultivation region and time of release

Germplasm origin	Landrace	Old	Modern	Totals[a]
East med	18	1	2	21
Elsewhere	8	16	14	38
North med	23	14	30	67
South med	21	4	10	35
Turkey	10	8	9	27
Totals	80	43	65	188

[a] Four wrongly identified genotypes omitted from totals

Phenotyping

Seed from the multiplication at ICARDA was used for sowing trials to estimate yield in a range of environments likely to be subject to varying levels of drought stress. All entries in the DBG were grown in trials for harvest years 2004 and 2005 at sites in Algeria (DZA), Spain (ESP), Italy (ITA), Jordan (JOR) Morocco (MOR), Syria (SYR) and Turkey (TUR). Two trials were grown in each country, either at sites selected from past meteorological data to be in comparatively low and high rainfall regions for the country or at the same site with one trial being rainfed and the other supplied with supplementary irrigation (Table 2). Each site in each country was therefore called "Dry" or "Wet" respectively.

Each trial was grown in an augmented design using rectangular incomplete blocks of size 60 with 4 checks repeated three times in each block laid out in a diagonal fashion at fixed distances. Four incomplete blocks thus comprised one full replicate of all DBG entries to which a fifth incomplete block, which contained a random selection of 48 DBG entries plus the same distribution of checks, was added to provide partial replication. In order to have a meaningful

Table 2 Trial sites of the Diverse Barley Gemplasm for harvest years 2004 and 2005

Code	Country	Site[a]	Location	Watering	Latitude	Longitude
DZA	Algeria	Dry	Ouled Hamla	Rainfed	36°5′ N	6°28′ E
DZA	Algeria	Wet	El Khroub	Rainfed	36°15′ N	6°‘42′ E
ESP	Spain	Dry	Foradada	Rainfed	41°39′ N	0°23′ W
ESP	Spain	Wet	Gimenells	Irrigated[b]	42°24′ N	0°21′ W
ITA	Italy	Dry	Foggia	Rainfed	41°28′ N	15°33′ E
ITA	Italy	Wet	Foggia	Irrigated[b]	41°28′ N	15°33′ E
JOR	Jordan	Dry	Ramtha	Rainfed	32°32′ N	36°2′ E
JOR	Jordan	Wet	Rabba	Rainfed	31°16′ N	35°44′ E
MOR	Morocco	Dry	Sidi El Aydi	Rainfed	33°04′ N	7°37′ W
MOR	Morocco	Wet	Sidi El Aydi	Irrigated[b]	33°04′ N	7°37′ W
SYR	Syria	Dry	Breda	Rainfed	35°56′ N	37°10′ E
SYR	Syria	Wet	Tel Hadya	Rainfed	36°1′ N	36°56′ E
TUR	Turkey	Dry	Hayamana	Rainfed	39°26′ N	32°30′ E
TUR	Turkey	Wet	Esenboga	Rainfed	40°8′ N	33°1′ E

[a] Dry and Wet sites classified according to previous meteorological data

[b] Wet site created artificially by supplementary irrigation supplied during the growing season

experimental error one we selected a local landrace, a local old variety and a local modern cultivar within those included in the 192 of DBG. The first check (cv Rihane) was common to all sites and the other three checks were a local landrace, a local old and a local new cultivar relevant to each country. This distribution of checks would therefore provide a more meaningful experimental error and also facilitate comparisons between sites and the performance of alien germplasm within a country. The 1.25 replicates of each trial were sown in a rectangular grid of 15 rows and 20 columns to estimate and correct for any spatial variation. Trials were sown in plots of 6 m^2 at each site and were grown according to local practise for sowing rate and other inputs.

Data analysis

Each trial was analysed by Restricted Maximum Likelihood (REML) in GENSTAT (Payne et al. 2006) with a mixed model with repeated checks as a fixed effect and rows, columns and entries as random effects. The variation between checks and the replicate entries was used to estimate an error variance to check for significant variation between the DBG entries. Estimates for the variance components for entries and error were then used to obtain estimates for repeatability (= entry variance/(entry variance + error variance)), a rough indicator for the amount of genetic variation, say broad sense heritability, in the DBG trials. The greater number of checks than test entries may, however, cause some bias in this estimate. From the mixed models for individual trials, Best Linear Unbiased Predictors (BLUPs) for entry performance were produced and used in subsequent analyses.

To explore adaptation of genotypes to trial sites, we compared average BLUPs of genotypes belonging to different Germplasm Origins. To simplify the procedure, we fitted one-way ANOVA models to the vectors of genotypic BLUPs per trial, treating the BLUPs as if they were observations. We then compared differences between means from Germplasm Origin groups with standard errors of differences. Note that this procedure should only be used as first and rough analysis of adaptational differences.

Correlations between environments were investigated as follows. A two-way table of genotype (entry) × environment (trial) BLUPs was formed. This table was then standardized per column (trial) and a principal components analysis was applied, treating the trials as variables and the genotypes as objects. Results were visualized in a bi-plot (Digby and Kempton 1987).

Marker-trait associations were assessed per trial by fitting mixed models using the REML directive in GENSTAT (Payne et al. 2006). The model that was fitted for each of 1,130 DArT markers was: yield = Germplasm Origin + marker, with Germplasm Origin a random factor with five levels that corrects for population sub-structure, and marker a binary indicator variable representing the scores for each DArT marker. The significance of the association of each DArT marker was determined by a Wald test, the standard test for testing fixed effects in mixed models (Payne et al. 2006). For each trial, the *P*-values obtained for the set of DArT markers were transformed to False Discovery Rates following (Benjamini and Hochberg 1995). The *P*-value equivalents of False Discovery Rates at 5% for the individual trials ranged from 0.0001 for both DZA_5D and ESP_4D to 0.018 for ITA_4D. For ease of presentation, we re-transformed the False Discovery Rates to a *P*-value threshold of 0.0001 across all markers and trials. Finally, we constructed multi-locus QTL models for individual trials, using a multiple regression approach combined with a variable subset selection procedure based on stepwise regression (function STEP in GENSTAT Payne et al. 2006), including Germplasm Origin in each model to correct for population sub-structure. The set of individually significant DArT markers within a particular trial was used as the set of predictors from which to select from in a forwards manner, with the criterion that each added marker must produce a significant reduction in the Residual Mean Square.

Results

Positive estimates for genetic (entry) variance were found in all trials apart from the Algeria Dry site in 2004. For the remaining trials, the broad sense heritabilities ranged from just 2% to over 80% but there was no significant association of heritability with site mean as some low yielding sites (<3 t/ha) had heritabilities of greater than 50% (Table 3). Yield performance differed markedly ranging from just under 0.3 t/ha to over 6.3 t/ha with the maximum cultivar yield being over 8 t/ha for the Turkish variety Tarm-92 when grown in the Wet site in Turkey in 2004. At some sites, notably the dry Jordan site in 2005, some entries failed to set any seed and thus recorded a zero yield. The yield BLUPs for these lines did, however, show small yield values, reflecting the shrinkage property of BLUP estimates. Therefore, overall, the minimum yield BLUPs ranged from just under 0.25 t/ha (effectively zero in reality) for the French variety Baraka at the dry site in Jordan 2004 to nearly 4.5 t/ha for a Jordanian landrace in the wet site in Spain in 2004 when considering only those trials where there was significant genetic variation.

Adaptation of germplasm to environments was tested by ANOVA of the BLUPs from each trial with Germplasm Origin as a factor but excluding DZA_4D as there was no significant genetic variation and also

Table 3 Broad sense heritability (repeatability) of yield and means, minima and maxima (based on BLUPs) of the Diverse Barley Germplasm population in 28 trials

Site	Heritability	Yield t/ha		
		Minimum	Mean	Maximum
DZA_4D	0	5.07	5.07	5.07
DZA_4W	0.22	4.34	5.16	6.08
DZA_5D	0.02	3.12	3.17	3.23
DZA_5W	0.49	1.85	3.27	4.42
ESP_4D	0.56	2.23	3.26	4.32
ESP_4W	0.75	4.48	6.32	7.93
ESP_5D	0.13	0.61	0.72	0.80
ESP_5W	0.53	2.46	3.07	3.64
ITA_4D	0.83	2.19	4.15	5.85
ITA_4W	0.37	3.36	4.73	5.96
ITA_5D	0.58	3.84	5.04	6.58
ITA_5W	0.74	3.01	5.24	7.51
JOR_4D	0.13	0.24	0.29	0.42
JOR_4W	0.62	0.55	1.18	2.52
JOR_5D	0.37	0.48	0.66	1.07
JOR_5W	0.09	1.00	1.10	1.25
MOR_4D	0.69	2.07	3.81	5.93
MOR_4W	0.64	1.92	3.90	5.90
MOR_5D	0.81	0.38	0.74	1.63
MOR_5W	0.55	0.69	1.55	2.72
SYR_4D	0.26	1.08	1.33	1.80
SYR_4W	0.30	3.24	3.99	4.67
SYR_5D	0.41	2.73	3.57	4.86
SYR_5W	0.48	3.74	4.78	5.90
TUR_4D	0.44	3.54	5.21	6.60
TUR_4W	0.62	3.39	6.21	8.27
TUR_5D	0.74	1.67	3.93	7.36
TUR_5W	0.59	2.11	4.63	6.40

Table 4 Mean yields of subsets of Diverse Barley Germplasm associated with five distinct regions of cultivation from data obtained from 27 trials where there was significant genetic variation

Site	Region mean[a]					SED[b]
	East_Med	Elsewhere	North_Med	South_Med	Turkey	
DZA_4W	4.959	**5.235**	**5.246**	5.097	5.093	0.069
DZA_5D	3.174	3.172	3.173	3.176	3.183	0.005
DZA_5W	**3.379**	2.946	3.177	**3.631**	3.336	0.130
ESP_4D	**3.442**	3.104	3.177	**3.453**	3.207	0.085
ESP_4W	5.673	**6.666**	**6.557**	6.131	6.018	0.154
ESP_5D	0.726	0.716	0.721	0.725	**0.742**	0.007
ESP_5W	2.899	**3.072**	**3.128**	**3.043**	**3.096**	0.053
ITA_4D	3.847	**4.531**	**4.422**	4.16	3.107	0.189
ITA_4W	**4.76**	**4.837**	**4.86**	4.599	4.329	0.126
ITA_5D	**5.224**	**5.09**	**5.072**	4.971	4.802	0.109
ITA_5W	4.951	**5.743**	**5.524**	5.024	4.284	0.215
JOR_4D	**0.335**	0.283	0.281	0.301	0.285	0.007
JOR_4W	**1.528**	1.134	1.128	1.267	1.027	0.078
JOR_5D	**0.8**	0.655	0.631	0.677	0.636	0.022
JOR_5W	1.114	1.099	1.103	1.106	1.093	0.011
MOR_4D	**4.425**	3.814	3.594	**4.414**	2.946	0.182
MOR_4W	**3.923**	**4.357**	3.893	**4.017**	2.951	0.225
MOR_5D	**1.222**	0.693	0.626	0.881	0.519	0.053
MOR_5W	**1.697**	1.472	1.557	**1.82**	1.156	0.097
SYR_4D	1.311	1.313	1.33	**1.401**	1.295	0.028
SYR_4W	**4.067**	3.831	**3.978**	**4.072**	**4.061**	0.054
SYR_5D	**4.034**	3.393	3.405	3.758	3.559	0.074
SYR_5W	4.657	**4.86**	**4.883**	**4.753**	4.483	0.108
TUR_4D	4.639	4.944	5.38	5.065	**5.821**	0.124
TUR_4W	5.342	6.156	6.314	5.961	**7.061**	0.171
TUR_5D	3.845	4.291	3.754	3.691	4.099	0.257
TUR_5W	4.341	**4.606**	4.541	4.757	**4.946**	0.175

[a] Figures in bold denote Germplasm Origin means that are greater than two standard deviations from the other means

[b] Approximate standard error – unequal numbers in groups

excluding the four aberrant genotypes noted above. Significant differences between the five regions of origin were present in all but three of the remaining 27 trials, the exceptions being DZA_5D, JOR_5W and TUR_5D Table 4). The mean performance of the "East Med" group is clearly superior to the other four means in three of the four trials grown in Jordan but this does not necessarily reflect their adaptation to low yielding environments per se as the "South Med" group had a higher mean in two of the other three trials where the mean yield was <3 t/ha. Lines from the "North Med" and "Elsewhere" groups generally had higher means than the other groups in the higher yielding environments but the adaptation of the Turkish lines to their own environments was also noticeable. Interestingly, the Turkish lines also had higher mean performances at ESP_5D, especially, and ESP_5W, sites that suffered severe cold stress during early vegetative growth, suggesting that this group of lines was a good source of cold hardiness.

Results of principal components analysis show that sites from the same country do, with some exceptions, tend to cluster together (Fig. 1). The lines sampled in the DBG tend to cluster around the sites in which they are cultivated and/or originated so that the lines in the "East Med" group tend to cluster around

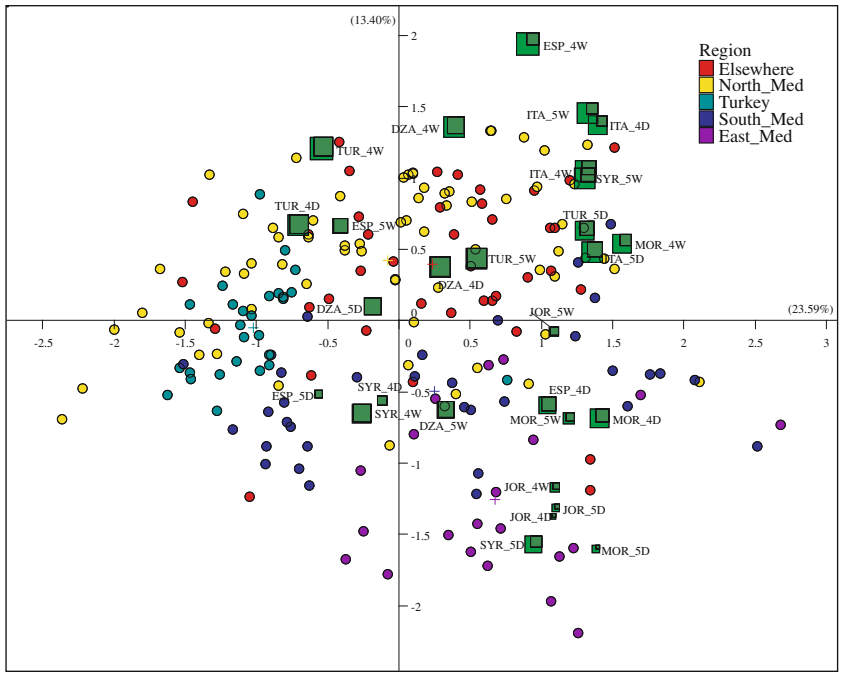

Fig. 1 GGE bi-plot of mean yields of 192 genotypes (coloured circles) from Diverse Barley Germplasm grown in 28 trials (green squares) classified according to their region of cultivation. Size of green squares is proportional to the trial mean and their cut-out portions reveal the proportion of the sum of squares for a site that is not represented by the displayed axes

the Jordanian and Syrian Sites, re-enforcing the results presented in Table 4. Given the diverse nature of this extensive data set, the fact that the first two components account for 36% of the variation is notable. Some caution should however, be exercised in interpreting this plot as the proportion of the sum of squares for some sites that was accounted for by the first two principal components was low, as displayed by the cut-outs in the squares representing the sites in Fig. 1, which applies especially to sites with lower loadings on PCA1

Single marker analysis of the yield BLUPs, including Germplasm Origin as a factor to account for population sub-structure, detected between 1 (DZA_5D) and 162 (TUR_5D) significant associations with the 1,130 DArT markers. Considering the individual DArT markers, three were significantly associated with yield BLUPs at 11 sites, one being located in Bin 4 on 3H and two unmapped but showing complete linkage disequilibrium with the marker in Bin 4 on chromosome 3H. Significant yield associations were found for 404 separate DArT markers but just 93 of these showed significant associations for five or more sites. Clearly, many of these markers are associated with yield because of linkage to other markers that may be more closely linked to the causal genes and it is not instructive to present more than a brief summary here.

From the 27 trials where genetic variation was detected, 96 marker trait associations were detected by stepwise regression resulting in multi-locus models, again including cultivation region as a factor to account for population sub-structure. These 96 associations were represented by 85 different DArT markers with a maximum of three significant associations for any one DArT marker in the multi-locus model. The number of markers fitted in each site ranged from one (DZA_4W and DZA_5D) to seven (MOR_4W) (Table 5). As expected, the percentage of variation accounted for by Germplasm Origin was low (<4%) in the sites where there were no clear differences between the means. For the remaining sites, Germplasm Origin was associated with a significant portion of the variation in yield (>20% at 11 sites with a maximum of nearly 50% at MOR_5D) emphasising the validity of accounting for population sub-structure. The significant marker associations in the multi-locus model also accounted for quite substantial portions of the phenotypic variation, ranging from just over 5% (DZA_4W) to over 50% (TUR_5D) with levels of >20% being detected at 15 sites (Table 5). A number of the significant DArT markers in the multi-locus models were closely linked so it is instructive to locate the 96 marker associations with yield on a bin map. We therefore utilised the DArT consensus map to define

Table 5 Numbers of QTLs detected in a multi-locus regression model when accounting for population sub-structure (Region) and percentages of variation accounted for by the sub-structure and all the markers fitted in the model for each of 27 trials where there was significant genetic variation

Site	QTLs	% Variation accounted for	
		Region	Marker
DZA_4W	1	11.2	5.4
DZA_5D	1	3.2	7.1
DZA_5W	2	15.8	11.8
ESP_4D	2	19.8	14.1
ESP_4W	4	18.5	27.5
ESP_5D	3	7.1	18.4
ESP_5W	4	8.1	28.5
ITA_4D	5	32.2	37.5
ITA_4W	6	10.8	44.3
ITA_5D	2	7.7	21.4
ITA_5W	6	24.5	40.0
JOR_4D	3	29.3	15.0
JOR_4W	4	22.4	26.7
JOR_5D	3	29.8	14.0
JOR_5W	2	2.1	15.3
MOR_4D	4	31.8	20.9
MOR_4W	7	21.4	44.9
MOR_5D	5	48.9	21.4
MOR_5W	5	18.3	26.5
SYR_4D	3	8.0	20.4
SYR_4W	3	14.3	17.4
SYR_5D	3	38.5	15.5
SYR_5W	3	10.6	24.3
TUR_4D	3	31.8	15.7
TUR_4W	2	33.0	9.7
TUR_5D	6	4.0	51.5
TUR_5W	4	5.7	34.0

bins according to the Steptoe × Morex Bin Map (Kleinhofs et al. 1998) (Fig. 2). Some DArT markers were not, however, present on the consensus map but we were able to assign definite and putative Bin positions to 27 of the 35 unmapped markers by using correlation thresholds of 0.7 and 0.5 respectively in linkage disequilibrium mapping of unknown markers (Rostoks et al. 2006) and they have been included in Fig. 2. The location of the remaining eight is currently unknown but they have been included in the multi-locus models as they accounted for considerable portions of the genetic variation. The 88 known marker trait associations are distributed across 42 bins but as many of these are adjacent, it is still possible that this number is partially inflated by linkage (and linkage disequilibrium). The greatest number of associations detected in a bin is 6 (chromosomes 3H and 4H, bins 4 and 6 respectively). Five significant associations were detected in chromosomes 5H and 7H bins 6 and 7 respectively.

Discussion

The division of genotypes into the different Germplasm Origin groupings largely reflected qualitative differences at major developmental gene loci such as spring/winter habit and 2/6 row ear type. Whilst these different developmental classes were not totally separated by the classification, including Germplasm Origin as a factor in the association analyses enabled the detection of yield associations under various degrees of drought stress without complications due to segregation of major developmental genes encountered in many bi-parental mapping studies (Forster et al. 2000). The only other major developmental gene polymorphisms that we were aware of in the DBG were *Ppd1* (Turner et al. 2005) for early flowering under long days and *sdw1* (syn *denso*) (Barua et al. 1993) for short straw. As most of the trials were autumn sown and all at latitudes under 43°N (Table 1), it is unlikely that *Ppd1* would be effective and QTLs from just two sites were located in the same Bin (Fig. 2). The semi-dwarf gene *sdw1* was present in a number of cultivars, largely from the region "Elsewhere" where the later maturity associated with it (Powell et al. 1985) would not be a handicap with more plentiful moisture supply. It could therefore be viewed as being in Linkage Equilibrium within groups and thus unlikely to bias results unduly. In fact, we did not detect any QTLs within the same Bin as *sdw1* (Fig. 2) so we can conclude that the presence of the gene is not biasing results. Thus the results obtained from this study do indeed represent the action of genes other than the familiar major developmental genes present in barley.

Grain yield has been the subject of numerous QTL studies in barley but most of these have been conducted in relatively high yielding environments. For comparison with our results, the most relevant studies have been conducted in two bi-parental

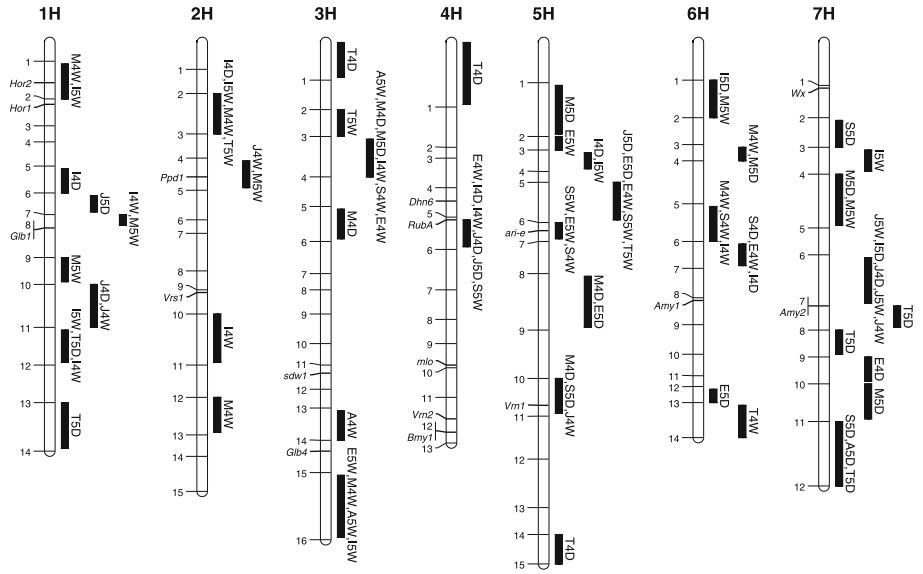

Fig. 2 Bin map for yield QTLs (solid vertical bars) detected in multi-locus models fitted for the Diverse Barley Germplasm population in each of 27 trials. First letter indicates country (A = Algeria, E = Spain, I = Italy, J = Jordan, M = Morocco, S = Syria, T = Turkey), number indicates harvest year (4 = 2004, 5 = 2005) and last letter indicates normal relative soil moisture status within the country (D = dry, W = wet). Lines denote bottom of bin segments

(Teulat et al. 2001); (Baum et al. 2003) and one Advanced Backcross (Talame et al. 2004) mapping populations. It is most interesting to determine whether or not these studies have highlighted the same genomic regions as those where we detected associations in five or more environments (Fig. 2). The regions in bins 6 and 5 on chromosomes 4H and 5H respectively were also found to harbour QTLs for grain yield in Tadmor × ER/APM and Barke × HOR 11508 (a *Hordeum spontaneum* accession). In addition, a QTL from Barke × HOR11508 was also detected in the region of bin 7 on 7H but no QTLs from the two populations were detected in the region of bin 4 on chromosome 3H. Despite the DBG being trialled at the same Syrian sites as a mapping population from Arta × *Hordeum spontaneum* 41–1, none of the QTLs detected for grain yield in that population were located in the same regions as the QTLs that showed some consistency across trials in the DBG. Comparison with the largely AFLP based map of Arta × *Hordeum spontaneum* 41–1 is, however, difficult but the fourth QTL on chromosome 5H appears to be in the same region as one detected in three environments (including two from Syria) in bin 7 on chromosome 5H in the DBG. The more consistent QTLs detected in the DBG can therefore, with the exception of bin 4 on chromosome 3H, be substantiated by results from other relevant studies in barley and provide candidate regions for more detailed future studies. Similarly, a previous association genetics study of yield in barley found that whilst many of the detected genomic regions were in bins that had also been identified in bi-parental QTL studies, some novel regions were also detected (Kraakman et al. 2004). The previous study was conducted exclusively upon adapted North–West European barley germplasm whereas the present study was conducted upon more diverse germplasm, some of which has not been the subject of previous QTL analysis so it is not surprising that we too have found some novel regions.

One of the encouraging findings in the present study is the detection of significant genetic variation for yield in the eight severely stressed environments where the mean yield was less than 2 t/ha. This is similar to a previous report of a lack of correlation between broad sense heritability and mean yield where heritabilities exceeding 50% were observed in environments yielding less than 1 t/ha (Al Yassin et al. 2005). Moreover, we were also able to detect

between two and five significant QTLs in environments with mean yields less than 2 t/ha and one of the most consistent genomic regions that we detected in the whole study was located in bin 7 on chromosome 7H, where four out of the five significant associations came from the four Jordanian sites with mean yield ranging from 0.3 to 1.2 t/ha. The detection of QTLs in the low yielding environments re-inforces the detection of genetic variation in such low yielding sites and offers the prospect of developing Marker Assisted Selection protocols for yield improvement in such situations. The fifth association detected in bin 7 on chromosome 7H was found in the 2005 dry site in Italy, where the mean yield was over 5 t/ha, suggesting that this QTL region might reflect a more general response rather than one that was only expressed under severe drought stress.

There remains the possibility that some of the yield associations that we have detected in the current study are due to the sites of origin not totally accounting for the underlying population sub-structure of the DBG. The fact that we can find results from other studies to substantiate our findings does not totally discount this possibility and further research is therefore required to choose and account for an appropriate model of the population sub-structure. This would then provide a sound basis for further research to identify and validate candidate genes underlying the QTLs that have been detected in the current study. Similarly, further research is necessary to investigate the role of physiological traits and also identify environmental factors that influenced QTL expression in the DBG but the data set assembled will allow these to be the feature of future studies.

Acknowledgements The above work was funded by the European Union-INCO-MED program (ICA3-CT2002-10026). SCRI receives grant in aid from the Scottish Executive Environment and Rural Affairs Department and Centre UdL-IRTA acknowledges partial funding from grant AGL2005-07195-C02-02 from the Spanish Ministry of Science and Education.

References

Al Yassin A, Grando S, Kafawin O, Tell A, Ceccarelli S (2005) Heritability estimates in contrasting environments as influenced by the adaptation level of barley germ plasm. Ann Appl Biol 147:235–244

Barua UM, Chalmers KJ, Thomas WTB, Hackett CA, Lea V, Jack P, Forster BP, Waugh R, Powell W (1993) Molecular mapping of genes determining height, time to heading, and growth habit in barley (Hordeum-vulgare). Genome 36:1080–1087

Baum M, Grando S, Backes G, Jahoor A, Sabbagh A, Ceccarelli S (2003) QTLs for agronomic traits in the Mediterranean environment identified in recombinant inbred lines of the cross 'Arta' x H-spontaneum 41-1. Theor Appl Genet 107:1215–1225

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate—a practical and powerful approach to multiple testing. J R Stat Soc B Methodol 57:289–300

Ceccarelli S, Grando S (1996) Drought as a challenge for the plant breeder. Plant Growth Regul 20:149–155

Digby PGN, Kempton RA (1987) Multivariate analysis of ecological communities. Chapman & Hall, London, UK

Ellis RP, Forster BP, Gordon DC, Handley LL, Keith RP, Lawrence P, Meyer R, Powell W, Robinson D, Scrimgeour CM, Young G, Thomas WTB (2002) Phenotype/genotype associations for yield and salt tolerance in a barley mapping population segregating for two dwarfing genes. J Exp Bot 53:1163–1176

Forster BP, Ellis RP, Thomas WTB, Newton AC, Tuberosa R, This D, El Enein RA, Bahri MH, Ben Salem M (2000) The development and application of molecular markers for abiotic stress tolerance in barley. J Exp Bot 51:19–27

Forster B, Ellis R, Moir J, Talame V, Sanguineti M, Tuberosa R, This D, Teulat-Merah B, Ahmed I, Mariy S, Bahri H, El-Ouahabi M, Zoumarou-Wallis N, El-Fellah M, Salem M (2004) Genotype and phenotype associations with drought tolerance in barley tested in North Africa. Ann Appl Biol 144:157–168

Jones HG (2007) Monitoring plant and soil water status: established and novel methods revisited and their relevance to studies of drought tolerance. J Exp Bot 58:119–130

Kleinhofs A, Kudrna DA, Matthews D (1998) Co-ordinators report: Integrating barley molecular and morphological/physiological marker maps. Barley Genet Newsl 28:89–91

Kraakman ATW, Niks RE, Van den Berg PMMM, Stam P, van Eeuwijk FA (2004) Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. Genetics 168:435–446

Passioura J (2007) The drought environment: physical, biological and agricultural perspectives. J Exp Bot 58:113–117

Payne RW, Harding SA, Murray DA, Soutar DM, Baird DB, Welham SJ, Kane AF, Gilmour AR, Thompson R, Webster R, Tunnicliffe Wilson G (2006) GenStat release 9 reference manual, part 2 directives. VSN International, Hemel Hempstead, UK

Powell W, Caligari PDS, Thomas WTB, Jinks JL (1985) The effects of major genes on quantitatively varying characters in Barley.2. The denso and day length response loci. Heredity 54:349–352

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rostoks N, Ramsay L, MacKenzie K, Cardle L, Bhat PR, Roose ML, Svensson JT, Stein N, Varshney RK, Marshall

DF, Graner A, Close TJ, Waugh R (2006) Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. Proc Natl Acad Sci U S A 103:18656–18661

Russell J, Booth A, Fuller J, Harrower B, Hedley P, Machray G, 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

Talame V, Sanguineti MC, Chiapparino E, Bahri H, Ben Salem M, Forster BP, Ellis RP, Rhouma S, Zoumarou W, Waugh R, Tuberosa R (2004) Identification of Hordeum spontaneum QTL alleles improving field performance of barley grown under rainfed conditions. Ann Appl Biol 144:309–319

Teulat B, Merah O, Souyris I, This D (2001) QTLs for agronomic traits from a Mediterranean barley progeny grown in several environments. Theor Appl Genet 103:774–787

Turner A, Beales J, Faure S, Dunford RP, Laurie DA (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. Science 310:1031–1034

Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, Kleinhofs A, Kilian A (2004) Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proc Natl Acad Sci U S A 101:9915–9920