CrossMark

RESEARCH ARTICLE

# Molecular and phenotypic diversity of ICARDA spring barley (*Hordeum vulgare* L.) collection

Reda Amezrou · Sanjaya Gyawali · Loubna Belqadi · Shiaoman Chao · Mustapha Arbaoui · Sujan Mamidi · Sajid Rehman · Avinash Sreedasyam · Ramesh Pal Singh Verma

**Abstract** Plant breeders are interested in using diverse genotypes in hybridization that can segregate for traits of importance with possibility of selection and genetic gain. Information on molecular and agro-morphological diversity helps the breeders reduce the effort for parental selection and helps the advancement of generations. A phenotypic and molecular diversity study, using 24 traits (agronomic and disease) and 6519 SNPs in a diverse collection of 336 spring barley genotypes, was carried out at Marchouch and Jemma Shiam research stations in Morocco. Based on structure and multivariate analyses, strong differentiation between the two- and six-row types were observed. The linkage disequilibrium (LD) decay of the current collection (for the combined population) was up to 3.58 cM ($r^2 = 0.15$) while LD decay were estimated 3.91 and 2.36 cM for two- and six-row barley, respectively. PCA of agro-morphological traits revealed grain per spike, net form of net blotch (NFNB), spot form of net blotch (SFNB), and 1000 kernel weight were the most discriminatory traits in the current collection. Association mapping in the two independent populations will be ideal for identification of markers, and QTL related to traits. The generated information on relatedness between individuals will help identify diverse genotypes for breeding programs.

**Keywords** Barley · Diversity · Genetic · *Hordeum vulgare* · Phenotypic · SNP · Structure

R. Amezrou · S. Gyawali (✉) · S. Rehman · R. P. S. Verma
BIGM Program, International Center for Agricultural Research in the Dry Areas (ICARDA), Rabat, Morocco
e-mail: gyawalisanjaya@gmail.com

L. Belqadi · M. Arbaoui
IAV Hassan II, Rabat, Morocco

S. Chao
USDA-ARS, Fargo, ND, USA

S. Mamidi · A. Sreedasyam
Hudson Alpha Institute for Biotechnology, Huntsville, AL, USA

*Present Address:*
S. Gyawali
Department of Plant Sciences, University of Manitoba, 222 Agric. Building, 66 Dafoe Road, Winnipeg, MB R3T 2N2, Canada

## Introduction

Barley (*Hordeum vulgare* L.) is one of the most important cereal crops in the world with nearly 50 million hectares (ha) of harvested area and 145 million

🖄 Springer

tons (t) production worldwide (FAOSTAT 2015). This crop in particular was domesticated from its wild relative *Hordeum vulgare* subsp. s*pontaneum* (K. Koch) around 10,000 years ago in the Fertile Crescent (Badr et al. 2000; Zohary and Hopf 2000). New evidence based on *RBP2* gene shows that barley was domesticated both in the Fertile Crescent and Tibetan Plateau (Wang et al. 2016a, b). Barley is mainly used for animal feed, brewing malts and human consumption (Munoz-Amatriain et al. 2014; Hayes et al. 2002) and is considered a staple food in several regions of the world, including the North and East Africa (Shewayrga and Sopade 2011).

The worldwide distribution of barley is due to its wide adaptation to diverse agro-ecology and different abiotic stresses (drought, cold, heat, and salinity). Recent studies indicated that polymorphism in flowering time genes *HvCO1*, *HvFT1*, *Ppd-H1*, and *VRN1-H1* has contributed to the adaptation of barley towards diverse agro-ecology (Aslan et al. 2015). With changes in climate, food productivity has to be increased to meet the global food demands. Barley can be considered as a model species due to its ability to grow in different environments which shaped its diversity, accumulating a rich pool of genes as a result of adaptation to wide environments and survival in harsh conditions (Grando et al. 2001). In fact, an extensive amount of data has been generated from genetic diversity surveys in wild and cultivated barley over the past decade (Munoz-Amatriain et al. 2014; Comadran et al. 2009; Orabi et al. 2007; Brantestam et al. 2006; Feng et al. 2006; Malysheva-Otto et al. 2006; Pandey et al. 2006; Chabane et al. 2005; Hou et al. 2005; Hamza et al. 2004; Baek et al. 2003; Matus and Hayes 2002; Struss and Plieske 1998; William et al. 1997). Genetic diversity studies are important tools that help crop improvement by identification of diverse parental lines for hybridization and to introgress desirable genes into elite germplasm (Chakravorty et al. 2013; Gyawali et al. 2013). Such studies can provide information about the resource allocation that affect the long-term maintenance of diverse germplasm collections (McClean et al. 2012). An understanding of diversity and genetic structure is also important for association mapping since population structure can lead to spurious associations and a control can be used to reduce false positives (Gyawali et al. 2016). High-throughput genotyping platforms and candidate gene studies have promoted association mapping as a viable approach for quantitative trait locus (QTL) mapping. It is an alternative to the traditional QTL mapping by using the recombination events from multiple lineages and to exploit the natural variation in large samples. Genotyping a diverse collection will help identify genomic regions of interest that control phenotypic variation.

The success of association mapping depends on the extent and patterns of linkage disequilibrium (LD). The extent of LD in a given population determines the density of markers required for whole genome scan that have implications for the identification of candidate genes associated with traits of interest (Szalma et al. 2005). Patterns of LD help discern the regions of low LD that has implications for breeder's selection. The overall LD facilitates in the understanding of the population genetic processes involved in shaping the present diversity of plants (Iqbal et al. 2012; Gurung et al. 2011; Mackay and Powell 2007; Malysheva-Otto et al. 2006; Gupta et al. 2005; Flint-Garcia et al. 2003), because the LD is affected by mating systems, recombination, selection, and genetic bottlenecks (Hamblin et al. 2011; Flint-Garcia et al. 2003). Therefore, it is important to know the population structure and the diversity of the population that can be used for association mapping.

High-throughput SNP genotyping platforms have revolutionized the gene mapping and genome-wide association studies (GWAS) in plants (Tian et al. 2011). The barley 9 K iSelect Illumina SNP platform gives whole genome coverage and an adequate genetic characterization of germplasm collections, which will make the diversity contained in a given collection efficiently accessible to barley breeders (Munoz-Amatriain et al. 2014; Comadran et al. 2009). This 9 K SNP chip has been effective in the identification of QTL in several studies, including Turuspekov et al. (2016), Tamang et al. (2015) and Mamo and Steffenson (2015).

The International Center for Agricultural Research in the Dry Areas (ICARDA) has the global mandate for barley improvement among the Consultative Group for International Agricultural Research (CGIAR) centers and holds one of the largest barley accessions (more than 30,000 barley accessions including wild relatives, landraces, and cultivars) in its gene banks across the world. In order to conduct GWAS for multiple traits of interest, a collection of 336 genotypes consisting of elite lines from multiple agro-ecological environments, released cultivars, landraces, and differentials, was assembled representing

much diversity present in the ICARDA's spring barley breeding gene pool adapted to variable environments. The objectives of the current study were to explore genetic and phenotypic diversity of the collection and to determine the patterns of population structure and LD within this collection.

## Materials and methods

### Plant materials

A total of 336 barley genotypes that includes advanced breeding lines, cultivars, and landraces from ICARDA and other sources such as barley genotypes introduced from different countries into ICARDA's barley breeding program were used for this study (Supplemental Table S1). Barley genotypes were selected for this study specifically representing tolerances to abiotic stress (drought and heat) and biotic (foliar diseases including rust, net blotch, spot blotch, powdery mildew) stresses. Further, genotypes were selected from low input barley breeding programs (stressed conditions for moisture and fertility), high input barley breeding programs (favorable conditions) of ICARDA. Barley genotypes selected for this study also represented feed, food, and malt barley programs of ICARDA. While selecting genotypes, appropriate consideration was given to select representative samples from both two- and six-row barley. All genotypes are of spring growth habit, out of which 199 are six-row and 137 are two-row barley. Any genotypes showing winter or facultative growth habit were removed from collection. Furthermore, the collection can be classified as hulled (276), primarily used for feed and malting purposes, and hulless (60) barley for food. The collection consisted of 230 barley genotypes from low input barley breeding programs (genotypes adapted to abiotic stresses), 82 from high input breeding program (adapted to the favorable production conditions) and rest 24 genotypes being frequently used by both programs. All available information on this collection is presented in the Supplemental Table S1.

### Field experiment and phenotyping

Evaluations of agronomic traits and screening of disease resistances were carried out at two research stations in Morocco. The experiments were carried out in alpha-lattice design with two replications during 2014–15 season in Marchouch (MCH) (33°33′38.2″N 6°41′24.7″W), and Jemma-Shiam (JS) (32°21′09.3″N 8°50′32.0″W) stations. Marchouch research station has been considered a high production potential with no stresses of moisture and soil fertility. In contrast to MCH, JS research station lack water supply for crop growth and is dependent on rain fed condition, therefore growing conditions in JS is considered to be moisture and nutrient stressed is considered. Data was recorded at both locations for agro-morphological and yield components, including days to heading (DH), days to maturity (DM), plant height (PH), spike length (SL), grains per spike (G/S), biological yield $ha^{-1}$ (BY) grain yield $ha^{-1}$ (GY), harvest index (HI = GY/BY), 1000 kernel weight (TKW) and hectoliter (test) weight in kg/hectoliter (HW). The genotypes were also screened for adult plant resistance (APR) to spot form of net blotch (SFNB), net form of net blotch (NFNB), and powdery mildew (PM) under natural conditions. In JS, PM resistance was evaluated at Zadoks growth stage 19–29 using 1–5 scale. At adult stage (Zadoks GS 77–87), disease rating was visually recorded using double digit scale (00–99) where the first digit indicates vertical disease progress on the plant and the second digit refers to severity measured in the infected leaf area (Saari and Prescott 1975).

The statistical analyses for all traits, in each location (MCH or JS), were taken up using Genstat v18 (VSN international, GenStat.co.uk). Multivariate analysis was performed on the measured qualitative and quantitative traits by using the principal component analysis (PCA) implemented in Genstat v18. The ANOVA was performed to evaluate the effects of genotypes (G), environment (E) and G × E interaction. In addition, each trait was investigated to determine relatedness of traits using Pearson's correlation coefficients. In multivariate analysis of agronomic traits, data from only 326 genotypes were considered that met the criteria of no missing data. The remaining 10 genotypes had at least one trait data missing and were excluded from PCA. For further investigation, a dendrogram based on mean traits from both locations was generated using hierarchical cluster analysis with the group average linkage method in Genstat v18.

SNP genotyping and diversity

Single plants of each line were grown in a greenhouse and the leaf tissue was lyophilized. Genomic DNA was extracted using the method described in Slotta et al. (2008). The barley genotypes were genotyped using 9 K iSelect SNP array based on Illumina's Infinium Assay (Illumina, San Diego, CA, USA) at Cereal Crop Research Unit, USDA-ARS, Fargo, ND. The obtained SNP data were further filtered for (a) a minor allele frequency of 0.05, (b) rate of missing values above 10%.

Diversity statistics including genetic diversity, major allele frequency and Polymorphic Information Content (PIC) were analyzed using PowerMarker v3.25 (Liu and Muse 2005). The phylogenetic analysis was conducted using Nei distance matrix (Nei 1972), computed by PowerMarker and used as input to generate the Unweighted Pair-Group Method using Arithmetic averages (UPGMA) dendrogram, viewed in TreeView X v0.5 (Page 1996). The genetic distance (D) among the genotypes was estimated by Unbiased Measures of genetic distance (Nei 1972). The genetic relationships between genotypes were further investigated by principal coordinate analysis (PCoA) based on the Nei genetic distance matrix in NTSYSpc 2.02i (Rohlf 2000).

Population structure analysis

Analysis of the population structure among barley genotypes was performed using the Bayesian model-based analysis implemented in the STRUCTURE v2.3.4 (Hubisz et al. 2009; Falush et al. 2003; Pritchard et al. 2000). Each individual is assigned to different groups according to a membership coefficient ($q_i$; $\Sigma q_i = 1.0$). The posterior probabilities were estimated using the Markov Chain Monte Carlo (MCMC) method. The number of hypothetical populations (K) tested was from 1 to 7. For each K, 5 runs were set and the MCMC chains were run with a 100,000 burn-in period, followed by 100,000 iterations using the admixture model with correlated allele frequencies. The most likely number of sub populations was determined using the $\Delta$K (Evanno et al. 2005) implemented in Structure Harvester (Earl and vonHoldt 2012).

Linkage disequilibrium

The estimates of the linkage disequilibrium (LD) of SNPs were determined for pairs of loci using the software package Tassel 3.0 (Bradbury et al. 2007) using SNPs of known marker positions only. The squared allele-frequency correlations ($r^2$) (Weir 1979) was calculated for each intra chromosomal combination. The distribution and extent of LD were visualized by plotting intra-chromosomal $r^2$ values against the genetic distance in cM for all inter-chromosomal marker pairs using nonlinear regression as described in Remington et al. (2001) and implemented in SAS 9.3.

## Results

Phenotypic diversity

The phenotypic stats (minimum, maximum, mean, standard error of the mean, and range) of 24 traits are presented in Table 1. The agro-morphological traits of individual genotypes are presented in Supplemental Table S1. The agronomic data of two- and six-row types are presented in Supplemental Tables S2 and S3, respectively. Using data for the quantitative and qualitative traits from both locations, the first three principal components (PCs) accounted for 66.4% of the total variability. The first PC explained 25.45% of the total variation (Fig. 1). Particularly, G/S, NFNB and SFNB resistance in both locations were the variables with high positive loadings, while TKW had the largest negative loading. The second PC explained 21.79% of the total variation. In second PC, SFNB resistance in JS and TKW had the highest positive loadings while G/S and NFNB resistance in JS were the variables with the largest negative loadings. The third component that explained 19.16% of the total variation was associated with high positive loadings of NFNB resistance in both locations and TKW while the largest negative loading was associated with SFNB resistance in JS (Table 1). The PCA of agronomic traits measured in MCH and JS is presented in Supplemental Figs. 1a and 1b. The ANOVA of agronomic traits are presented in Tables 2 and 3. A highly significant ($P < 0.01$) effect of genotypes (G) was found for DH, PH, SL, NFNB, SFNB, G/S, BY, GY and HI (Table 2). A highly significant ($P < 0.05$) effect of environments (MCH and JS) was found for DH, PH, SL, NFNB, GY, BY, and HI index. Likewise, highly significant ($P < 0.01$) effect of G × E interaction was observed for DH, SL, NFNB, SFNM, and GY. A highly significant

**Table 1** Descriptive statistics and principal component analysis (PCA) of studied phenotypic traits

| Trait[a] | Mean ± SEM | Max | Min | Range | PC1 | PC2 | PC3 |
|---|---|---|---|---|---|---|---|
| DH-MCH | 103.3 ± 0.30 | 130 | 90.5 | 39.5 | 0.004 | 0.03 | 0.001 |
| DH-JS | 115.7 ± 0.38 | 147 | 89 | 58 | 0.0007 | −0.02 | −0.03 |
| DM-MCH | 154.2 ± 029 | 172 | 104.5 | 67.5 | 0.013 | 0.015 | 0.002 |
| PH-MCH | 100.2 ± 0.53 | 123.8 | 63.58 | 60.17 | 0.065 | −0.05 | −0.003 |
| PH-JS | 67.18 ± 0.40 | 85 | 46.5 | 38.5 | 0.034 | −0.006 | −0.013 |
| SL-MCH | 7.717 ± 0.08 | 12.22 | 4.867 | 7.35 | −0.01 | 0.01 | 0.006 |
| SL-JS | 7.128 ± 0.07 | 16.88 | 4.625 | 12.25 | −0.006 | 0.005 | 0.003 |
| G/S-MCH | 48.66 ± 1.04 | 83 | 23.33 | 59.67 | **0.246**[b] | **−0.399** | −0.09 |
| G/S-JS | 48.52 ± 0.80 | 90 | 22.5 | 67.5 | **0.146** | **−0.231** | −0.06 |
| BY-MCH | 9.567 ± 0.09 | 14.84 | 4.444 | 10.4 | 0.001 | 0.019 | 0.006 |
| BY-JS | 9.117 ± 0.08 | 12.92 | 5.583 | 7.333 | 0.0006 | 0.003 | 0.003 |
| GY-MCH | 4.026 ± 0.05 | 6.567 | 1.522 | 5.044 | −0.002 | 0.01 | 0.001 |
| GY-JS | 2.873 ± 0.03 | 4.458 | 1.048 | 3.41 | −0.006 | 0.002 | 0.0006 |
| HI-MCH | 0.425 ± 0.003 | 0.668 | 0.242 | 0.425 | −0.0007 | 0.0001 | −0.0001 |
| HI-JS | 0.32 ± 0.002 | 0.59 | 0.1 | 0.49 | 0.002 | 0.01 | 0.002 |
| TKW- MCH | 45.32 ± 0.28 | 60 | 27.5 | 32.5 | **−0.118** | **0.802** | **0.181** |
| NFNB-MCH | 66.51 ± 1.71 | 93 | 00 | 93 | **0.258** | 0.068 | **0.561** |
| SFNB-MCH | 58.75 ± 2.05 | 97 | 00 | 97 | **0.219** | 0.092 | 0.007 |
| PM-MCH | 76.26 ± 1.43 | 50 | 00 | 3 | 0.001 | −0.002 | 0.001 |
| NFNB-JS | 68.24 ± 1.85 | 95 | 00 | 95 | **0.426** | **−0.13** | **0.643** |
| SFNB-JS | 71.23 ± 1.62 | 95 | 00 | 95 | **0.774** | **0.323** | **−0.413** |
| PM-JS | 1.375 ± 0.04 | 4 | 10 | 3 | −0.001 | −0.003 | 0.001 |
| HW-MCH | 68.33 ± 0.34 | 83 | 56.5 | 26.5 | −0.043 | 0.028 | 0.011 |

[a] *DH* Days to heading, *PH* plant height, *SL* spike length, *G/S* grain per spike, *BY* biomass yield, *GY* grain yield, *HI* harvest index, *TKW* thousand kernel weight, *NFNB* net form of net blotch, *SFNB* spot form of net blotch, *PM* powdery mildew, *HW* hectoliter (test) weight, *MCH* Marchouch, *JS* Jemma Shiam

[b] Highlighted in bold are relevant characteristics that explained respective components

($P < 0.01$) effect of genotypes was found for DM, TKW, HI, and PM-Adult in MCH while PM-Seedling was non-significant in JS (Table 3).

*Correlation between phenotypic traits*

Correlation coefficients ($r^2$) were highly significant ($P < 0.001$) in 46 of the 276 trait combinations, where $r^2$ ranged from 0.01 to 0.96 (Fig. 2). The correlation coefficient and $P$ values are presented in a correlation matrix in Supplemental Table S4. High positive correlations ($r^2 \geq 0.5$) were found between BY and GY at both locations (BY-JS and GY-JS; BY-MCH and GY-MCH); DH-MCH and DM-MCH; G/S-JS and G/S-MCH with row type. High biomass imply high grain yield and similarly for days to heading and days to maturity. Highly significant negative correlations were found for G/S with, TKW and HW at both locations; HI-MCH with PH-MCH; and row type with SL, TKW and HW. The row type appears to play a key role on the number of grains per spike, TKW and HW

where the two-row types have less grains compared to the six-row types and tend to have heavier and larger grains which determine the TKW and HW. Significant positive correlations between NFNB resistance at both locations was observed ($r^2 = 0.39$). Similarly, the correlation was positively significant for resistance to SFNB at two locations ($r^2 = 0.33$). This indicates that the resistance/susceptibility was mainly governed by genetic factors, while the environment has very little impact.

*Cluster analysis of phenotypic traits*

Hierarchical cluster analysis clearly classified barley genotypes into two main groups (Figs. 3a, c). Without exceptions, the two clusters separated two-row from six-row barley genotypes. Within a given cluster, genotypes aggregated into small groups based on their disease resistance/susceptibility, agronomic performance, and morphological traits. Net blotch resistance, earliness, biomass, yield and plant height were
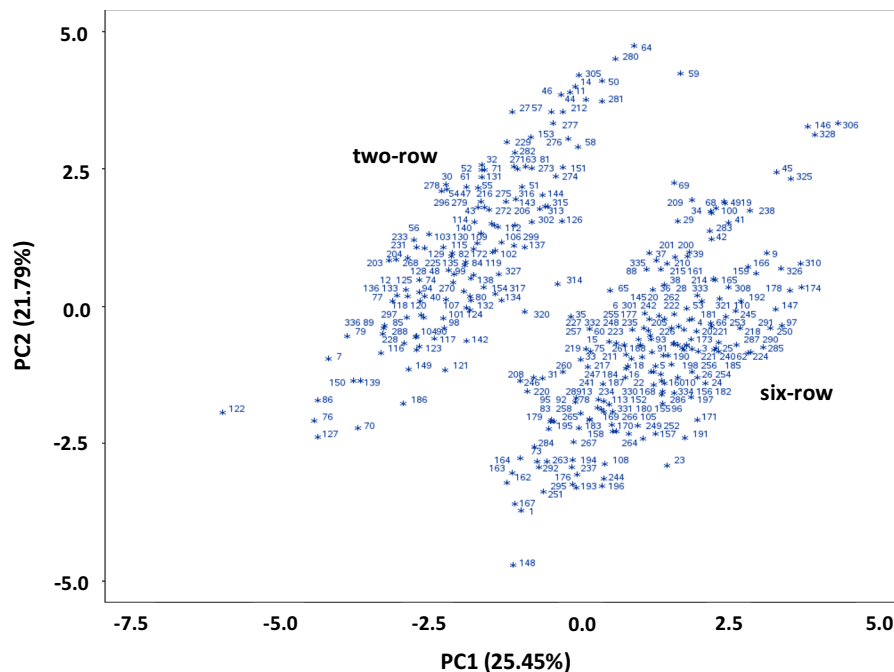
**Fig. 1** Principal component analysis based on genetic distance computed for 336 individuals using 24 agronomic traits measured for 336 barley genotypes in Marchouch and Jemma Shiam

the main traits of discrimination within the six-row and two-row clusters (Supplemental Table 1). The lowest similarity (71.3%) was found between AM-1 (Alanda/5/Aths/4/Pro/TolI//Cer*2/TolI/3/5106/6/ Baca'S'/3/AC253//CI08887/CI05761) and AM-304 (CI3576) while the highest (99.7%) was observed between AM-156 (Aths/Lignee686/4/Avt/Attiki// Aths/3/Giza121/Pue) and AM-18 (Massine/Arig8).

Genetic diversity and cluster analysis

A subset of 6940 genome-wide SNPs was used to assess genetic diversity in the collection. Out of 6940, 1982 did not have a known chromosome position and the remaining 4958 were distributed over all seven chromosomes. The subset was further filtered for minor allele frequencies (MAF $\leq$ 0.05) and missing SNPs (>10%), and a final set of 6519 SNPs were used for further analyses (Table 4). Gene diversity and polymorphism information content (PIC) values on different chromosomes varied from 0.005 to 0.500 and 0.006 to 0.375, with average values of 0.366 and 0.290, respectively (Table 4).

The genetic similarity between genotypes quantified using Nei genetic distance (Nei 1972) resulted

into two main clusters of significant size corresponding to row type. Furthermore, within the same cluster, genotypes were grouped depending on their adaptation (high-input barley, low-input barley, landrace). The largest distance (D = 0.89) was found between AM-27 (LIMON/BICHY2000//DEFRA/DESCONO-CIDA-BAR) and AM-300 (Arimont). The smallest genetic distance (D = 0.00) was observed between a seventeen pairs of genotypes, all sister lines originated from the same crosses. In order to demonstrate the phylogenetic relationships of the 336 barley genotypes studied, an Unweighted Pair-Group Method using Arithmetic averages (UPGMA) dendrogram was generated (Fig. 3a) and all genotypes were assigned to two major groups (two- and six-row barley genotypes) and three sub-groups (high-input barley, low-input barley, landraces).

Population structure analysis

The break point of $\Delta K$ in the current study was $K = 2$ (Fig. 3b). As described by Evanno et al. (2005), the true value of $K$ is when $\Delta K$, an ad hoc quantity the second order rate of change of the likelihood function with respect to $K$, reached its peak. Out of 336

**Table 2** Analysis of variance of agronomic traits measured in Marchouch and Jemma Shiam, Morocco in 2014–2015

| Sources of variation | df | DH[a] | PH | SL | NFNB | SFNB | G/S | BY | GY | HI |
|---|---|---|---|---|---|---|---|---|---|---|
| Genotype (G) | 335 | 72.4** | 217.6** | 4.151** | 2366.4** | 2707.4** | 781.5** | 3.456** | 0.76674** | 0.0070** |
| Environment (E) | 1 | 48,681.5** | 357,290.3** | 83.560** | 80,272.8** | 2443.6$^{ns}$ | 16.0$^{ns}$ | 15,363.6** | 3342.8** | 4.249** |
| G × E | 334 | 61.8** | 60.6$^{ns}$ | 2.340** | 1025.5** | 1409.9** | 228.3$^{ns}$ | 2.596$^{ns}$ | 0.6446** | 0.0051$^{ns}$ |
| Error | 649 | 2.6 | 82.2 | 1.52 | 812.2 | 781.0 | 201.7 | 2.66843 | 0.493 | 0.005 |

[a] *PH* Plant height, *SL* spike length, *NFNB* net form of net blotch, *SFNB* spot form of net blotch, *G/S* grain per spike, *BY* biomass yield, *GY* grain yield, *HI* harvest index

*, ** are significant at 0.05 and 0.01 probability levels

genotypes, 138 (41.08%) were assigned to Q1 and 84 (25%) were assigned to Q2, while the remaining 114 genotypes (33.92%) were admixed (membership coefficient, $q_i \leq 0.8$). The genetic structure of the collection was also analyzed by using Principal Coordinate Analysis (PCoA). The PCoA of genetic distance revealed a clear differentiation between two and six-row barley sub-populations (Fig. 4). The first and second axes explained 45.49 and 18.05% variations, respectively, and separated genotypes in different clusters corresponding to the row type. One of the clusters mostly contained two-row while another cluster contained six-row barley genotypes. However, some overlaps between two- and six-row clusters were also observed.

Linkage disequilibrium (LD)

The extent of LD was assessed among all chromosomes as well as for the two sub-populations separately. For all genotypes, 16.27% of the total SNP pairs were in LD ($P < 0.001$) and 26.53% at $P < 0.05$ significance. In our samples, the genome-wide LD decay was 3.58 cM at $r^2 > 0.15$ (Fig. 5). However, for the two row genotypes, the number of SNP pairs that are in LD is 29.62% ($P < 0.05$) and 19.65 ($P < 0.001$) and for six row genotypes it is 32.78% ($P < 0.05$) and 21.96 ($P < 0.001$). The decay values are about 3.91 for two-row barley and 2.26 for six-row barley (Supplemental Fig. 2a and 2b).

## Discussion

### Phenotypic diversity

Descriptive statistics (mean, range and standard error of the means) of 24 agronomic traits showed high levels of variation in barley genotypes. For example, the number of QTL for yield reported were about 60 (Wang et al. 2016a, b; Xue et al. 2010; Pillen et al. 2003; Marquez-Cedillo et al. 2001; Teulat et al. 2001) and for disease resistance there were 31 QTL for leaf rust (Kertho et al. 2015), between 8 and 13 for various strains of SFNB in barley (Tamang et al. 2015). Based on PCA of the phenotypic traits, this barley collection was mainly clustered with respect to their disease resistance (SFNB and NFNB), number of grains per spike, and TKW. This clustering was quite evident

**Table 3** Analysis of variance of agronomic traits measured either in Marchouch or Jemmas Shiam, Morocco in 2014–2015
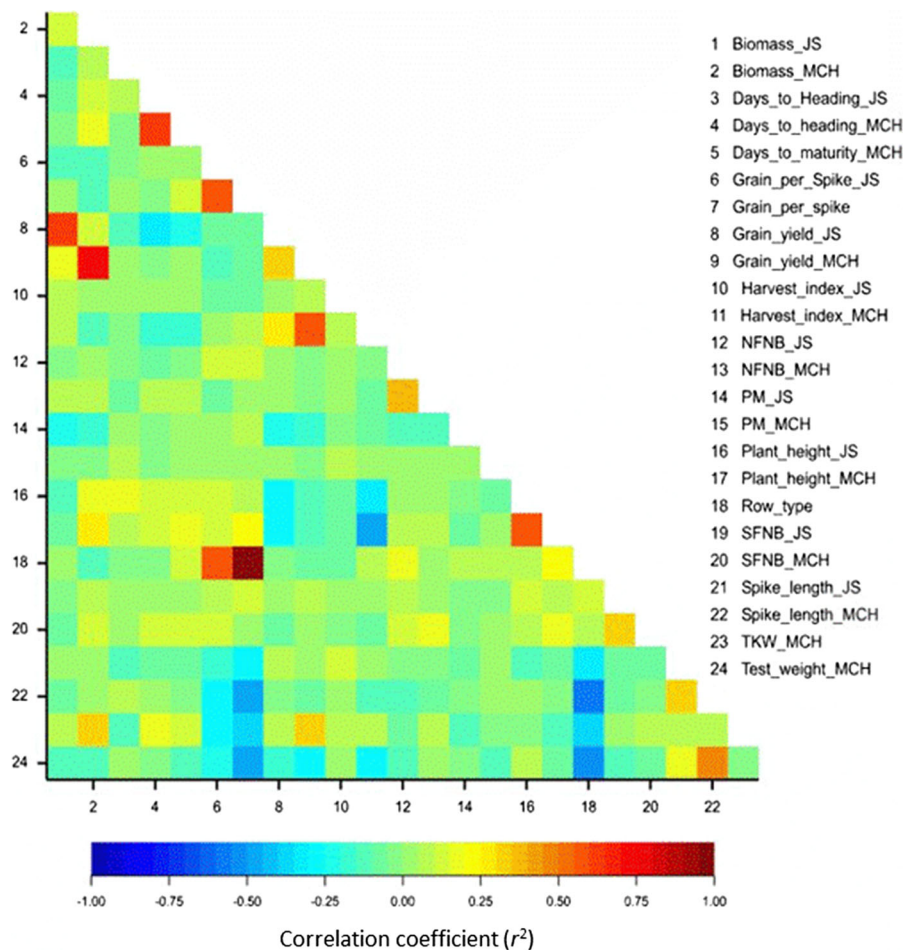
| Source of variation | df | DM[a] | TKW | HW | PM-Adult | PM-Seedling[b] |
|---|---|---|---|---|---|---|
| Replication (R) | 1 | 181.2** | 120.9** | 4.34** | 0.0554 | 2.93* |
| Genotype (G) | 335 | 47.8** | 45.7** | 67.71** | 1292.6** | 0.614[ns] |
| Error | 335 | 20.1 | 9.15412 | 0.936 | 0.0276 | 0.6555688 |

[a] DM-Days to maturity, TKW-1000 kernel weight, HW-Hectoliter (test) weight, PM-Adult-Powdery mildew severity recorded at adult stage using double digit in Marchouch

[b] PM-Seedling-Powdery mildew recorded at seedling stage using 1–5 scale in Jemma Shiam where 1 is resistance response and 5 is susceptible

*, ** are significant at 0.05 and 0.01 probability levels



**Fig. 2** Pearson's correlations plot based on correlation coefficients ($r^2$) of the 24 phenotypic traits using 326 barley genotypes

1 Biomass_JS
2 Biomass_MCH
3 Days_to_Heading_JS
4 Days_to_heading_MCH
5 Days_to_maturity_MCH
6 Grain_per_Spike_JS
7 Grain_per_spike
8 Grain_yield_JS
9 Grain_yield_MCH
10 Harvest_index_JS
11 Harvest_index_MCH
12 NFNB_JS
13 NFNB_MCH
14 PM_JS
15 PM_MCH
16 Plant_height_JS
17 Plant_height_MCH
18 Row_type
19 SFNB_JS
20 SFNB_MCH
21 Spike_length_JS
22 Spike_length_MCH
23 TKW_MCH
24 Test_weight_MCH

Correlation coefficient ($r^2$)

since there was strong variability in terms of net blotch (both NFNB and SFNB) response among the genotypes. The variation in the number of grains per spike reflects the row type, but was not enough to separate our population into two groups as revealed by the SNPs markers. Thousand kernel weight (TKW) had the highest PC2 positive loading compared to rest of the traits which show there was high variation in TKW in this collection. Most of two-row genotypes had larger grain than six-row in this study (Supplemental Table 2) which was in agreement with previous reports (Ayoub et al. 2002; Marquez-Cedillo et al.

**Fig. 3 a** UPGMA dendrogram of 336 barley genotypes using SNP markers, **b** inferred population structure based on 6519 SNPs markers and 336 barley genotypes. Each individual is represented by *Q1* and *Q2* sub-populations (the estimated membership fraction to each sub-population defined by STRUCTURE), **c** hierarchical cluster based on phenotypic traits of 326 barley genotypes
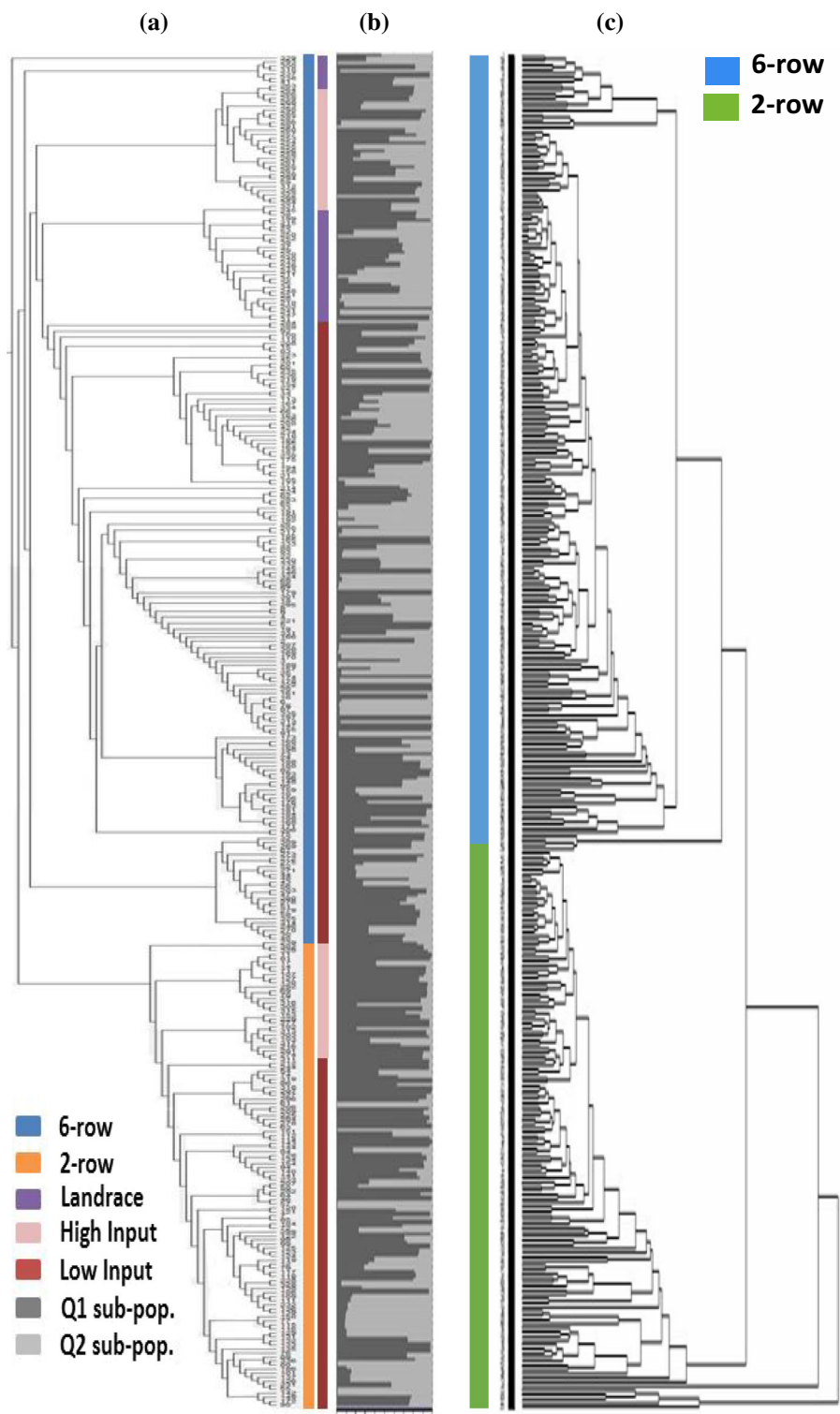
**Table 4** Marker information and diversity statistics of markers mapped in individual chromosomes

| Chromosome | No. of SNP used | No. of SNP used (after filtering) | Average distance per SNP (cM) | Major allele frequency | Gene diversity | PIC[a] |
|---|---|---|---|---|---|---|
| 1H | 499 | 475 | 0.292 | 0.722 | 0.369 | 0.293 |
| 2H | 866 | 804 | 0.207 | 0.742 | 0.346 | 0.277 |
| 3H | 765 | 713 | 0.215 | 0.712 | 0.370 | 0.292 |
| 4H | 502 | 472 | 0.258 | 0.706 | 0.381 | 0.300 |
| 5H | 975 | 941 | 0.190 | 0.720 | 0.366 | 0.291 |
| 6H | 670 | 633 | 0.208 | 0.721 | 0.366 | 0.291 |
| 7H | 681 | 640 | 0.248 | 0.704 | 0.375 | 0.295 |
| Unknown | 1982 | 1841 | – | – | – | – |
| Overall | 6940 | 6519 | 0.231 | 0.719 | 0.366 | 0.290 |

[a] PIC-Polymorphic information content
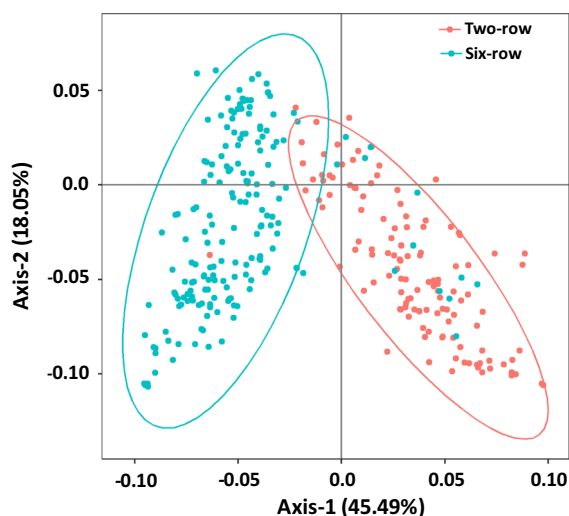


**Fig. 4** Principal coordinate analysis based on genetic distance computed for 336 barley genotypes using 6519 SNP markers genotyped for AM-2014 panel

2000; Kjaer and Jensen 1996). Grain weight compensates for early stages of environmental stresses if favorable conditions prevail during the period of grain filling. In dry areas, moisture stress is prevalent at all stages, especially grain filling, and ICARDA barley breeders tend to select material based on grain weight within level of inputs. Although in many cases, the coefficients ($r^2$) were low, there were significant correlations among different traits. Hence a trade-off of key traits should be taken into consideration during selection and breeding. The disease reactions for NFNB and SFNB at two locations indicated a similar response to either of the disease at both locations, indicating that the pathotypes at both locations might be similar.

The classification of genotypes based on hierarchical clustering using Euclidean distance resulted in two main groups, six-row and two-row types. This is in support of the classification of the SNP markers. However, subgroups within a given cluster gathered with a contrasting expression of agronomic traits. Based on the agronomic merit of each subgroup, the genotypes can be classified according to their disease resistance/susceptibility, biomass, yield, height and earliness. No specific differentiations can be made based on other traits. The maximum distance was found between AM-1 (Alanda/5/Aths/4/Pro/TolI//Cer*2/TolI/3/5106/6/Baca'S'/3/AC253//CI08887/CI05761), a six-row accession, highly susceptible to NFNB, semi-dwarf with short spikes and AM-304 (CI3576) which is a two-row landrace highly resistant to NFNB, tall with long spikes. This amplitude of agronomic traits and disease resistance in barley genotypes reflects the wide genetic variability present in our collection, which is a fundamental condition for the genetic improvement. Similar observations were reported earlier by Shakhatreh et al. (2010) and Manjunatha et al. (2007) in barley collections according to agro-morphological traits.

Genetic diversity

The current study is amongst the first in ICARDA to deliberately assemble and analyze a specific
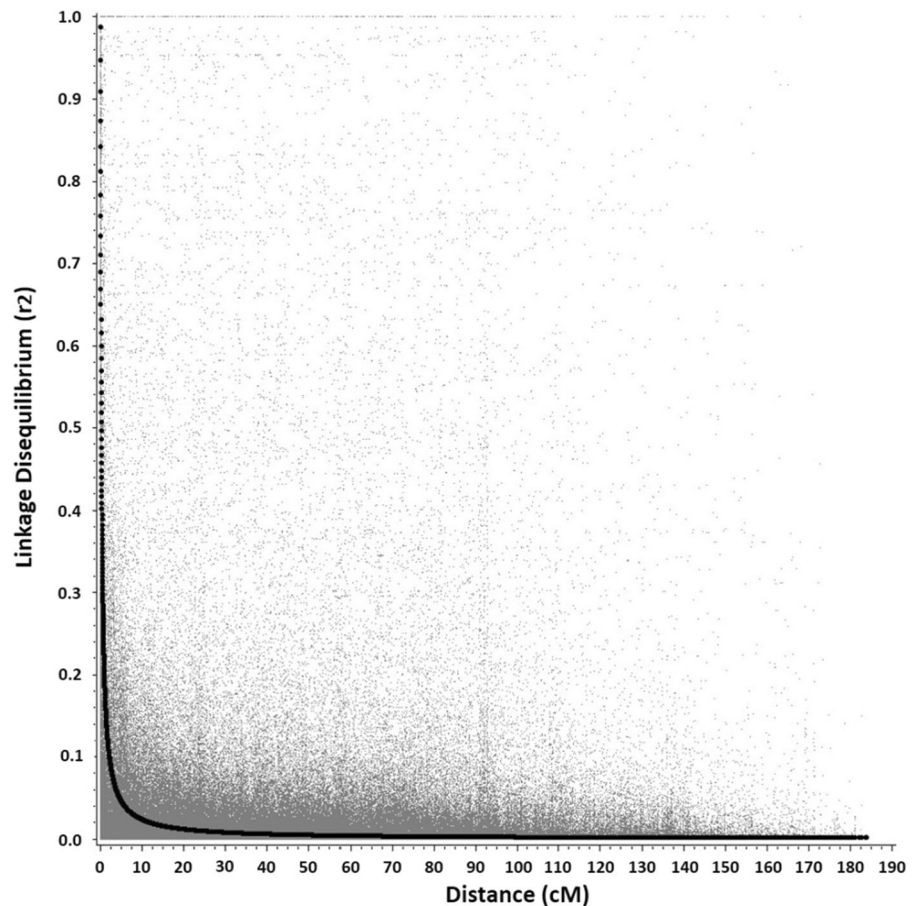
population representing very diverse cultivated barley from ICARDA germplasm to provide a platform of GWAS for several important traits. We used SNP markers because it offers a highly polymorphic, co-dominant, and high-throughput marker system which can be used in germplasm characterization and selection of desirable alleles in breeding programs (Lombardi et al. 2014). Minor allele frequency and expected heterozygosity are directly correlated. This additional measure can determine the proportion of rare alleles (MAF < 0.2), which in turn determines the diversity of the population. In our study, we found an average expected heterozygosity of 0.29 which is comparable to that observed in other studies (Lombardi et al. 2014; Emanuelli et al. 2013; Jones et al. 2007; Ching et al. 2002). Furthermore, the average gene diversity in our sample was 0.366, which is slightly higher than that reported by Rodriguez et al. (2012) and Sun et al. (2011) using SSR markers, 0.298 in barley landraces from Sardinia and 0.338 in a

worldwide barley genotypes, respectively. Higher genetic diversity is generally expected in the current mapping panel because of the diverse nature of genotypes used in the current study, which were originated from different barley breeding programs across the globe, landraces collected from diverse geographical regions. Therefore, by selecting SNPs based on their high polymorphism levels, the discriminating power of the SNP can be considerably increased (Jones et al. 2007).

Many of the ICARDA's breeding lines, analyzed in this study, share common parents. As genetic distance is based on the principal that shared alleles are identical by descent, this measure of discrimination power is meaningful in our population. The maximum distance was found between Arimont, an American six-row, naked genotype and LIMON/BICHY2000// DEFRA/DESCONOCIDA-BAR, a two-row malt barley cross derived from highly separated localities and breeding programs and inversely, the lowest distance

was found between pairs of sister lines from the ICARDA breeding programs, which is evident as they had same parentage.

## Population structure and linkage disequilibrium

Cluster analysis based on Nei (1972) distances separated, with some exceptions, the genotypes according to their row type. Our results correlate with previous studies showing a clear separation between two- and six-row types (Usubaliev et al. 2013; Chaabane et al. 2009; Chen et al. 2009; Lasa and Igartua 2001; Franckowiak and Lundqvist 1997). Historically, in ICARDA, breeders had made several two-by-six row crosses which was evident in this study by the identification of admixtures (Fig. 2b). This admixture was clearly shown from the pedigree of ICARDA barley breeding lines where both two- and six-row genotypes were included in particular crosses (Supplemental Table S1). Hence, both structure and PCA analyses support the hypothesis of genetic admixture of two- and six-row barley in ICARDA germplasm. Despite that the optimum number of subpopulations was two ($K = 2$), genotypes tend to cluster (based on their coefficient of membership; $Q_i$) according to their adaptation mode (high-input barley, low-input barley), regardless of their row-type. This is evident since ICARDA had two distinct barley breeding programs, in the past, located in Syria and Mexico based on target countries and end uses. The one in Syria was the low-input breeding program where the developed genotypes are more adapted to stressed environments (poor crop management, cold and drought conditions), and are bred for feed and food purposes. Whereas the genotypes developed in Mexico under the high input breeding program are more adapted to favorable conditions (high rainfall/irrigated and appropriate crop management) and mainly bred for malt or feed. However, in the current study, the structural tendencies may not be absolute as 34% of the genotypes were admixed, and can be derived from the crosses of different parents and may be suitable for both environments.

In our study, LD at $P < 0.001$ was observed in 16.27% of loci pairs and in 26.53% at $P < 0.05$ significance level, where 74.4% are linked (<40 cM). Our results considerably exceeded LD reported by Rodriguez et al. 2012 using S-SAP markers where 25 genotypes of *Hordeum spontaneum* with 15% of loci

pairs at $P < 0.05$ and 13% of loci pairs at $P < 0.01$ in a landrace population of Sardinia were observed. Our results were lower than the proportion reported by Malysheva-Otto et al. (2006), where 42% of loci pairs at $P < 0.05$ in 207 European two-row spring barley using SSR markers were observed. The most plausible explanations for the moderately low LD in our collections compared to Malysheva-Otto et al. (2006) are, the use of bi-allelic SNP markers and secondly, nature of barley germplasm used in this study. Our panel includes a considerable number of landraces while breeding lines used in the current study were generated by frequently including landraces in the ICARDA's barley breeding programs. The number of detected loci pairs in LD is greater in multi-allelic markers such as SSR compared to bi-allelic markers such as SNPs. Also, the level of LD is higher in cultivated barley compared to landraces and wild genetic resources (Flint-Garcia et al. 2003). In the current study, we used bi-allelic SNP markers and nearly 12% of our population consisted of landraces or cultivars with a background of wild barleys, therefore an average low level of LD was expected (Massman et al. 2011; Cockram et al. 2008; Malysheva-Otto et al. 2006).

Mean $r^2$ LD values higher than 0.15 extended up to 3.58 cM in our study and we argued that the current marker density (0.231 cM/SNP) was sufficient for genome wide association studies in barley. In the case of bi-allelic markers, previous studies have reported successful association mapping in barley using a marker density of 1 DArT marker per 1.5 cM (Comadran et al. 2009) and 1 SNP marker per 0.72 cM (Pasam et al. 2012; Massman et al. 2011; Cockram et al. 2008). In this study, the barley 9 K Ilumina SNP array (6519 SNP markers) gave an approximate coverage of 1 SNP marker per 0.231 cM. The 9 K SNP platform was successfully used for various GWAS of different traits in barley (Tamang et al. 2015; Munoz-Amatriain et al. 2014).

## Conclusions

This study provided a detailed description of a population, that represents a wide range and historical survey of barley diversity within ICARDA germplasm and comprised a considerable proportion of the genetic and phenotypic variation underlying the

different strategies for adaptation to different environments. We have demonstrated that the barley genotypes studied were genetically and phenotypically diverse, and strongly structured. The marker coverage, population stratification and the level of LD in our germplasm set was appropriate to run different GWAS studies for key traits in barley. For detecting the most confident QTLs and avoid spurious associations, it is important to consider association mapping using combined and independently in the two subpopulations i.e. two- and six-row barley.

## References

Aslan S, Forsberg NEG, Hagenblad J, Leino MW (2015) Molecular genotyping of historical barley landraces reveals novel candidate regions for local adaption. Crop Sci 55:2766–2776. doi:10.2135/cropsci2015.02.0119

Ayoub M, Symons SJ, Edney MJ, Mather DE (2002) QTLs affecting kernel size and shape in a two-rowed by six-rowed barley cross. Theor Appl Genet 105:237–247. doi:10.1007/s00122-002-0941-1

Badr A, Muller K, Schaefer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, Pozzi C, Rohde W, Salamani F (2000) On the origin and domestication history of barley (*Hordeum vulgare*). Mol Biol Evol 17:499–510. doi:10.1093/oxfordjournals.molbev.a026330

Baek HJ, Beharav A, Nevo E (2003) Ecological-genomic diversity of microsatellites in wild barley, *Hordeum spontaneum*, populations in Jordan. Theor Appl Genet 106:397–410

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Brantestam AK, von Bothmer R, Dayteg C, Rashal I, Tuvesson S, Weibull J (2006) Genetic diversity changes and relationships in spring barley (*Hordeum vulgare* L.) germplasm of Nordic and Baltic areas as shown by SSR markers. Genet Resour Crop Evol 54:749–758. doi:10.1007/s10722-006-9159-4

Chaabane R, El-Felah M, Salah HB, Abdelly C, Ramla D, Nada A, Saker M (2009) Molecular characterization of Tunisian barley (*Hordeum vulgare* L.) genotypes using microsatellites (SSRs) markers. Eur J Sci Res 36(1):6–15

Chabane K, Ablett GA, Cordeiro GM, Valkoun J, Henry RJ (2005) EST versus genomic derived microsatellite markers for genotyping wild and cultivated barely. Genet Resour Crop Evol 52:903–909. doi:10.1007/s10722-003-6112-7

Chakravorty A, Ghosh PD, Sahu PK (2013) Multivariate analysis of phenotypic diversity of landraces of rice of West Bengal. Am J Exp Agric 3(1):110–123

Chen F, Chen D, Valles MP, Gao Z, Chen X (2009) Analysis of diversity in Chinese cultivated barley with simple sequence repeats: differences between eco-geographic populations. Biochem Genet 48(1):44–56

Ching A, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski AJ (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genet 3:19

Cockram J, White J, Leigh FJ, Lea VJ, Chiapparino E, Laurie DA, Mackay IJ, Powell W, O'Sullivan DM (2008) Association mapping of partitioning loci in barley. BMC Genet 9:16. doi:10.1186/1471-2156-9-16

Comadran J, Thomas WT, van Eeuwijk FA, Ceccarelli S, Grando S, Stanca AM, Pecchioni N, Akar T, Al-Yassin A, Benbelkacem A, Ouabbou H, Bort J, Romagosa I, Hackett CA, Russell JR (2009) Patterns of genetic diversity and linkage disequilibrium in a highly structured *Hordeum vulgare* association-mapping population for the Mediterranean basin. Theor Appl Genet 119:175–187. doi:10.1007/s00122-009-1027-0

Earl DA, vonHoldt BM (2012) STRUCURE HARVERSTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4(2):359–361

Emanuelli F, Lorenzi S, Grzeskowiak L, Catalano V, Stefanini M, Troggio M, Myles S, Martinez-Zapater JM, Zyprian E, Moreira FM, Grando MS (2013) Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. BMC Plant Biol 13:39. doi:10.1186/1471-2229-13-39

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587

FAOSTAT (2015). http://faostat3.fao.org (Verified in Feb 2, 2017)

Feng ZY, Zhang LL, Zhang YZ, Ling HQ (2006) Genetic diversity and geographical differentiation of cultivated six-rowed naked barley landraces from the Qinghai-Tibet of China detected by SSR analysis. Genet Mol Biol 29:330–338

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54:357–374

Franckowiak JD, Lundqvist U (1997) Barley genetics symposium 6, six-rowed spike 1, *vrs1*. Barley Genet Newsl 26:49–50

Grando S, von Bothmer R, Ceccarelli S (2001) Genetic diversity of barley: use of locally adapted germplasm to enhance yield and yield stability of barley in dry areas. In: Cooper

HD, Spillane C, Hodgkin T (eds) Broadening the genetic base of crop production. IPGRI/FAO/CABI Publishing, CAB Interantional, UK, pp 351–372

Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol Biol 57:461–485

Gurung S, Mamidi S, Bonman JM, Jackson EW, Del Rio LE, Acevedo M, Mergoum M, Adhikari TB (2011) Identification of novel genomic regions associated with resistance to *Pyrenophora tritici-repentis* races 1 and 5 in spring wheat landraces using association analysis. Theor Appl Genet 123(6):1029–1041

Gyawali S, Hegedus DD, Parkin IAP, Poon J, Higgins E, Horner K, Bekkaoui D, Coutu C, Buchwaldt L (2013) Genetic diversity and population structure in a world collection of *Brassica napus* accessions with emphasis on South Korea, Japan, and Pakistan. Crop Sci 53(4):1537–1545. doi:10.2135/cropsci2012.10.0614

Gyawali S, Harrington M, Durkin J, Horner K, Parkin IAP, Hegedus DD, Bekkaoui D, Buchwaldt L (2016) Microsatellite markers used for genome-wide association mapping of partial resistance to *Sclerotinia sclerotiorum* in a world collection of *Brassica napus*. Mol Breed 36:72. doi:10.1007/s11032-016-0496-5

Hamblin MT, Buckler ES, Jannink JL (2011) Population genetics of genomics-based crop improvement methods. Trends Genet 27(3):98–106

Hamza S, Ben Hamida W, Rebai A, Harrabi M (2004) SSR-based genetic diversity assessment among Tunisian winter barley and relationship with morphological traits. Euphytica 135:107–118. doi:10.1023/B:EUPH.0000009547.65808.bf

Hayes PM, Castro A, Marquez-Cedillo L, Corey A, Henson C, Jones BL, Kling J, Mather D, Matu I, Rossi C, Sato K (2002) Genetic diversity for quantitatively inherited agronomic and malting quality traits. In: Von Bothmer R, Knüpffer H, van Hintum T, Sato K (eds) Diversity barley. Elsevier, Amsterdam

Hou YC, Yan ZH, Wei YM, Zheng YL (2005) Genetic diversity in barley from west China based on RAPD and ISSR analysis. Barley Genet Newsl 35:9–22

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332

Iqbal MJ, Mamidi S, Ahsan R, Kianian SF, Coyne CJ, Hamama AA, Narina SS, Bhardwaj HL (2012) Population structure and linkage disequilibrium in *Lupinus albus* L. germplasm and its implication for association mapping. Theor Appl Genet 125(3):517–530

Jones E, Sullivan H, Bhattramakki D, Smith JSC (2007) A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). Theor Appl Genet 115:361–371

Kertho A, Mamidi S, Bonman JM, McClean PE, Acevedo M (2015) Genome-wide association mapping for resistance to leaf and stripe rust in winter-habit hexaploid wheat landraces. PLoS ONE 10(6):e0129580. doi:10.1371/journal.pone.0129580

Kjaer B, Jensen J (1996) Quantitative trait loci for grain yield and yield components in a cross between a six-rowed and a two rowed barley. Euphytica 90:39–48

Lasa J, Igartua E (2001) Morphological and agronomical diversity patterns in the Spanish barley core collection. Hereditas 135(23):217–225

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Lombardi M, Materne M, Cogan NOI, Rodda M, Daetwyler HD, Slater AT, Forster JW, Kaur S (2014) Assessment of genetic variation within a global collection of lentil (*Lens culinaris* Medik.) cultivars and landraces using SNP markers. BMC Genet 15:150. doi:10.1186/s12863-014-0150-3

Mackay I, Powell W (2007) Methods for linkage disequilibrium mapping in crops. Trends Plant Sci 12:57–63

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

Mamo BE, Steffenson BJ (2015) Genome-wide association mapping of fusarium head blight resistance and agromorphological traits in barley landraces from Ethiopia and Eritrea. Crop Sci 55(4):1494–1512

Manjunatha T, Bisht IS, Bhat KV, Singh BP (2007) Genetic diversity in barley (*Hordeum vulgare* L. ssp. *vulgare*) landraces from Uttaranchal Himalaya of India. Genet Resour Crop Evol 54:55–65. doi:10.1007/s10722-005-1884-6

Marquez-Cedillo LA, Hayes PM, Jones BL, Kleinhofs A, Legge WG, Rossnagel BG, Sato K, Ullrich SE, Wesenberg DM (2000) QTL analysis of malting quality in barley based on the doubled-haploid progeny of two elite North American varieties representing different germplasm groups. Theor Appl Genet 101:173–184

Marquez-Cedillo LA, Hayes PM, Kleinhofs A, Legge WG, Rossnagel BG, Sato K, Ullrich SM, Wesenberg DM (2001) The North American barley genome mapping project. QTL analysis of agronomic traits in barley based on the doubled haploid progeny of two elite North American varieties representing different germplasm groups. Theor Appl Genet 103:625–637

Massman J, Cooper B, Horsley R, Neate SM, Dill-Macky R, Chao S, Dong Y, Schwarz P, Muehlbauer GJ, Smith KP (2011) Genome-wide association mapping of *Fusarium* head blight resistance in contemporary barley breeding germplasm. Mol Breed 27:439–454

Matus IA, Hayes PM (2002) Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats. Genome 45:1095–1106

McClean PE, Terpstra J, McConnell M, White C, Lee R, Mamidi S (2012) Population structure and genetic differentiation among the USDA common bean (*Phaseolus vulgaris* L.) core collection. Genet Resour Crop Evol 59(4):499–515

Munoz-Amatriain M, Cuesta-Marcos A, Endelman JB, Comadran J, Bonman JM, Bockelman HE, Chao S, Russel J, Waugh R, Hayes PM, Muehlbauer GJ (2014) The USDA barley core collection: genetic diversity, population structure, and potential for genome-wide association studies. PLoS ONE 9(4):e94688. doi:10.1371/journal.pone.0094688

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292

Orabi J, Backes G, Wolday A, Yahyaoui A, Jahoor A (2007) The horn of Africa as a centre of barley diversification and a potential domestication site. Theor Appl Genet 114:1117–1127. doi:10.1007/s00122-007-0505-5

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358

Pandey M, Wagner C, Friedt W, Ordon F (2006) Genetic relatedness and population differentiation of Himalayan Hulless barley (Hordeum vulgare L.) landraces inferred with SSRs. Theor Appl Genet 113:715–729. doi:10.1007/s00122-006-0340-0

Pasam RK, Sharma R, Malosetti M, van Eeuwijk FA, Haseneyer G, Kilian B, Graner A (2012) Genome-wide association studies for agronomical traits in a worldwide spring barley collection. BMC Plant Biol 12:16. doi:10.1186/1471-2229-12-16

Pillen K, Zacharias A, Lon L (2003) Advanced backcross QTL analysis in barley (Hordeum vulgare L.). Theor Appl Genet 107:340–352. doi:10.1007/s00122-003-1253-9

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

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt WR, Doebley J, Kresovich S, Goodman MM, Buckler ES (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. PNAS 98(20):1479–11484

Rodriguez M, Rau D, Sullivan D, Brown AHD, Papa R, Attene G (2012) Genetic structure and linkage disequilibrium in landrace populations of barley in Sardinia. Theor Appl Genet 125:171–184. doi:10.1007/s00122-012-1824-8

Rohlf FJ (2000) NTSYS-pc. Numerical taxonomy and multivariate analysis system, version 2.1. Exeter Software, Setauket

Saari EE, Prescott JM (1975) A scale for appraising the foliar intensity of wheat disease. Plant Dis Report 59:377–380

Shakhatreh Y, Haddad N, Alravavah M, Gando S, Ceccarelli S (2010) Phenotypic diversity in wild barley (Hordeum vulgare L. ssp. spontaneum (C. Koch) Thell.) accessions collected in Jordan. Genet Resour Crop Evol 57:131–146. doi:10.1007/s10722-009-9457-8

Shewayrga H, Sopade PA (2011) Ethnobotany, diverse food uses, claimed health benefits and implications on conservation of barley landraces in North Eastern Ethiopia highlands. J Ethnobiol Ethnomed 7:19. doi:10.1186/1746-4269-7-19

Slotta TAB, Brady L, Chao S (2008) High throughput tissue preparation for large-scale genotyping experiments. Mol Ecol Resour 8:83–87. doi:10.1111/j.1471-8286.2007.01907.x

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

Sun D, Ren W, Sun G, Peng J (2011) Molecular diversity and association mapping of quantitative traits in Tibetan wild and worldwide originated barley (Hordeum vulgare L.) germplasm. Euphytica 178:31–43. doi:10.1007/s10681-010-0260-6

Szalma SJ, Buckler ESIV, Snook ME, McMullen MD (2005) Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. Theor Appl Genet 110:1324–1333

Tamang P, Neupane A, Mamidi S, Friesen T, Brueggeman R (2015) Association mapping of seedling resistance to spot form net blotch in a worldwide collection of barley. Phytopathology 105(4):500–508

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. doi:10.1007/s001220100619

Tian F, Bradbury P, Brown P, Hung H, Sun Q, Flint-Garcia S, Rocheford T, McMullen M, Holland J, Buckler ES (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. Nat Genet 43:159–162. doi:10.1038/ng.746

Turuspekov Yerlan, Ormanbekova Danara, Rsaliev Aralbek, Abugalieva Saule (2016) Genome-wide association study on stem rust resistance in Kazakh spring barley lines. BMC Plant Biol 16(1):13

Usubaliev B, Brantestam AK, Salomon B, Garkava-Gustavson L, Bothmer R (2013) Genetic diversity in farmer grown spring barley material from Kyrgyzstan. Genet Resour Crop Evol 60:1843–1858. doi:10.1007/s10722-013-9959-2

Wang L, Sun G, Ren X, Li C, Liu L, Wang Q, Du B, Sun D (2016a) QTL underlying some agronomic traits in barley detected by SNP markers. BMC Genet 17:103. doi:10.1186/s12863-016-0409-y

Wang Y, Ren X, Sun D, Sun G (2016b) Molecular evidence of RNA polymerase II gene reveals the origin of worldwide cultivated barley. Sci Rep 6:36122. doi:10.1038/srep36122

Weir BS (1979) Inferences about linkage disequilibrium. Biometrics 35:235–254

William M, Dorocicz I, Kasha KJ (1997) Use of microsatellite DNA to distinguish malting and non-malting barley cultivars. J Am Soc Brew Chem 55:107–111

Xue DW, Zhou MX, Zhang XQ, Chen S, Wei K, Zeng FR, Mao Y, Wu FB, Zhang GP (2010) Identification of QTLs for yield and yield components of barley under different growth conditions. J Zhejiang Univ Sci B 11(3):169–176. doi:10.1631/jzus.B0900332

Zohary D, Hopf M (2000) Domestication of plants in the old world: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley, 3rd edn. Oxford University Press, New York