

Genetic analysis of wheat grains using digital imaging and their relationship to enhance grain weight

Ahmad Ali^{1*}, Zahid Ullah¹, Naveed Alam², S.M. Saqlan Naqvi³, Muhammad Jamil⁴, Hadi Bux⁵, Hassan Sher¹

¹University of Swat/Center for Plant Sciences and Biodiversity, P.O. Box 19130 – Swat – Khyber Pakhtunkhwa – Pakistan.

²University of Swat/Center for Agriculture Sciences and Forestry, P.O. Box 19130 – Swat – Khyber Pakhtunkhwa – Pakistan.

³Arid Agriculture University Rawalpindi/Faculty of Science – Dept. of Biochemistry, Muree Road Rawalpindi – Rawalpindi, Punjab – Pakistan.

⁴University of Sargodha/Faculty of Science – Dept. of Botany, University Road – 40100 – Sargodha – Pakistan.

⁵University of Sindh/Institute of Plant Sciences, Allama I.I. Kazi Campus – 76080 – Jamshoro – Pakistan.

*Corresponding author <aali_swat@yahoo.com>

Edited by: Leonardo Oliveira Medici

Received March 22, 2019

Accepted June 06, 2019

ABSTRACT: Phenomic characterization through digital imaging (DI) can capture the three dimensional variation in wheat grain size and shape using different image orientations. Digital imaging may help identifying genomic regions controlling grain morphology using association mapping with simple sequence repeats (SSRs) markers. Accordingly, seed shape phenotypic data of a core collection of 55 wheat genotypes, previously characterized for osmotic and drought tolerance, were produced using computer based Smart grain software. Measured dimensions included seed volume, area, perimeters, length, width, length to width ratio, circularity, horizontal deviation from ellipse (HDEV), vertical deviation from ellipse (VDEV), factor form density (FFD) etc. The thousand grain weight (TGW) was positively correlated with grain size direct measurements; however, VDEV, FFD and other derived grain attributes showed no or negative correlation with TGW. Digital imaging divided the genotypes correctly into well-defined clusters. The wheat genotypes studied were further grouped into two sub-clusters by the Bayesian structure analysis using unlinked SSR markers. A number of loci over various chromosomal regions were found associated to grain morphology by the genome wide analysis using mixed linear model (MLM) approach. A considerable number of marker-trait associations (MTAs) on chromosomes 1D and 2D may carry new alleles with potential to enhance grain weight due to the use of untapped wild accessions of *Aegilops tauschii*. Conclusively, we demonstrated the application of multiple approaches including high throughput phenotyping using DI complemented with genome wide association studies to identify candidate genomic regions underlying these traits, which allows a better understanding on molecular genetics of wheat grain weight.

Keywords: synthetic derived bread wheats, seed digital imaging, path analysis, marker-trait association

Introduction

Wheat grain size and shape are important characteristics, influenced by genetic and environmental factors causing variations in these attributes. Various grading methods using diverse morphological attributes for sorting different cereal grains and varieties through image processing techniques have been reported in the literature (Majumdar and Jayas, 2000; Visen et al., 2001). The digital imaging analysis (DIA) is the process of converting digital images into quantitative measurements based on pixel counts. This can result in changing the generation of large quantitative data and has previously been used in many works (Kwack et al., 2005; Dana and Ivo, 2008).

Genetic improvement in yield related attributes has become possible through the introgression of dwarfing genes and recently the use of synthetic hexaploid wheats (Mujeeb-Kazi et al., 2013). From agronomic perspectives, wheat grain yield is the most important trait underpinned by two numerical components including grains per square meter and grain weight (Calderini and Reynolds, 2000). In addition, improvement in TGW is considered a promising approach to improve wheat yield potential and is considered an important area of wheat genetic and breeding studies (Peng et al., 2003; Su et al., 2011). Seed quality is mainly determined by

interactions between the environment and the genome during seed development and maturation (Groos et al., 2003; Tsilo et al., 2010; Jamil et al., 2017).

Exploring quantitative trait loci (QTLs) for grain weight and its related component are important steps to deploy favorable alleles through marker-assisted selection. However, the relative disadvantages of linkage mapping studies over linkage disequilibrium (LD) or association mapping (AM) for the underlying trait control mechanisms suggest the use of mapping technique to be more appropriate for diverse germplasm (Breseghello and Sorrells, 2006; Huang and Han, 2013). Since several studies in wheat have reported QTLs for grain size and weight (Breseghello and Sorrells, 2007; Sun et al., 2009; Ramya et al., 2010; Rasheed et al., 2014). However, few studies have reported on QTLs for grain shape (Gegas et al., 2010; Williams et al., 2013; Williams and Sorrells, 2014) and studies under drought stress conditions are very rare (Nezhad et al., 2012). This study therefore aimed to i) characterize a core collection of wheat germplasm (Ali et al., 2015; Ali et al., 2017) for grain weight and explore its relationship to size and shape using digital imaging for high-throughput phenotyping, ii) manipulate digital imaging data for genotype discrimination, and iii) to identify SSR markers associated with grain phenotypes using the association mapping analysis.

Materials and Methods

The main population consisted of conventional and synthetic-derived wheat lines developed for drought tolerance. Derived synthetics were produced earlier by crossing primary synthetic hexaploid wheats from crosses of durum wheat (*T. turgidum*)/*Ae. tauschii* with susceptible bread wheat cultivars and originated from CIMMYT through Wheat Wide Crosses and Cytogenetics program in Pakistan. From this main population, a core collection of 55 wheat genotypes was previously characterized for drought stress tolerance under field condition for three consecutive years and under osmotic stress tolerance in lab conditions. Details of the germplasm, its pedigree, experimental layout and the studied attributes are given in Ali et al. (2015 and 2017). Spikes were harvested individually by hand and threshed with a mechanical belt thresher.

Digital imaging analysis of wheat seeds

Twenty-five intact randomly selected seeds of each genotype were used for the imaging analysis using the procedure discussed in Williams et al. (2013), Tanabata et al. (2012). Vertical and horizontal photography of seeds was accomplished on black sheet in a 5 × 5 grid spaced 1.5 cm apart with a digital camera. The two photographs for each genotype were named horizontal (H) and vertical (V) image (Figures 1A and 1B). The procedure followed was according to Williams et al. (2013). Digital photos were transferred to a photographic expert group (JPEG) files and were further processed by using the software package SmartGrain version 1.1, which was initially designed for rice seed images (Tanabata et al., 2012). Photograph files were renamed according to concerned genotype. In SmartGrain software, quantitative measures were directly derived from JPEG image, by setting a uniform scale in each photograph as seen in Figures 1A and 1B, to allow estimation of the pixel based actual distance. Seed color and background color commands were selected after loading the image into the software. Then, the morphology of selected seeds (25) was analyzed, outlines detected, longitudinal and horizontal axes were

drawn automatically to calculate seed length, seed width, seed area, seed perimeter, length to width ratio, circularity and distance from the gravity point (Figures 2A and 2B). The result file was transferred as CSV (comma-separated values) output file that could be opened in spread software (Tanabata et al., 2012) for further statistical analysis. Then, Elliptic Fourier Descriptors (EFDs) analysis was performed to assess the variations in grain shape (Iwata and Ukai, 2002; Iwata et al., 2010). For that purpose, bitmap images of wheat seeds were taken with a digital camera, processed through software SHAPE that alters the shapes of objects (wheat seeds) into a data apposite to record the EFDs and calculate the PCA score for each object. The PCA score simplifies a large number of EFD coefficients produced for each shape and minimizes them to quantitative values (Williams et al., 2013).

Molecular analysis

The same germplasm was previously genotyped with 101 simple sequence repeat (SSR) markers. Details of the allele number, polymorphic information content (PIC), gene frequency, SSR genetic diversity, and population structure analysis carried through STRUcture software (v. 2.1; Bradbury et al., 2007) were discussed in Ali et al. (2015). Observed Q matrix for maximum ΔK was used for the association mapping analysis carried out on phenotypic and genotypic data of the studied germplasm with TASSEL 2.0.1 according to Bradbury et al. (2007). Due to the small population size, a marker-trait association study was conducted based on the mixed linear model (MLM) approach. The MLM approach is relatively strict because it takes into account both population structure coefficients and kinship matrices (K), hence, it is highly recommended in such cases in order to avoid false positives or spurious marker trait associations. A significant threshold considered was based on Bonferroni correction (Bland and Altman, 1995). Associated markers were assigned to wheat chromosomes based on their position on a wheat consensus map (Somers et al., 2004) and data in the GrainGenes database (www.graingenes.org).

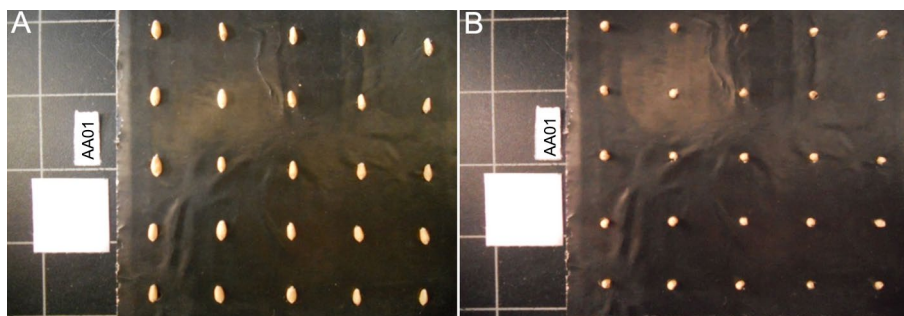


Figure 1 – Showing A) horizontal seed image, B) vertical seed image of genotype AA01 on a 5 × 5 grid. A paper cube of 30 × 30 mm is placed in each photograph to calibrate the scale.

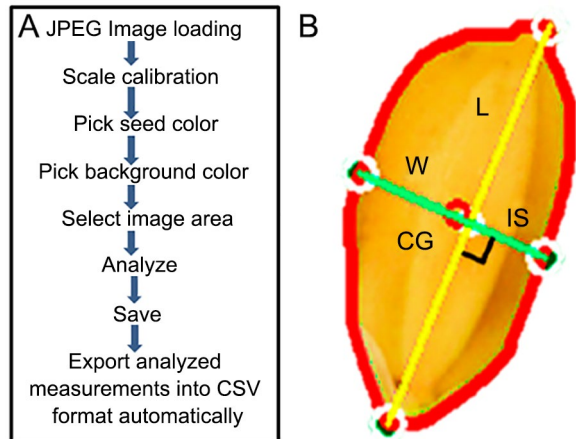


Figure 2 – (A) Steps involved in the digital image analysis using SmartGrain; (B) seed image analyzed by SmartGrain showing measured dimensions (L = length; W = width; CG = center of gravity; IS = intersection).

Statistical analysis

Relationships between variables were determined using the Pearson correlation test with the STATISTICA software (Statsoft v. 7.0, 2004). Based on the correlation approach, the data obtained were also subjected to the principal component analysis (PCA) in order to generalize and characterize the germplasm more comprehensively by using the multivariate option in PAST 2.12 (Hammer et al., 2001). The cluster analysis was performed to find similarity matrix with UPGMA (unweighted pair-group method with arithmetic mean) by using the same software PAST 2.12.

The multiple regression analysis was performed to check the variation in TGW as dependent variable affected by the seed shape parameters taken as independent variables using Lisrel 9.1 software package. Moreover, the covariance matrix and structure equation of the model was formulated. Path diagram with appropriate coefficients was drawn by the same software package elucidating the interaction of direct effects of dependent variables either positively or negatively to thousand-grain weight.

Results

Summary statistics for wheat grain attributes of the studied germplasm is given in Table 1. For TGW, the genotypes ranged from 29.5 g (AA02) to 51.9 g (AA24). There were seven synthetic derived bread wheats (SBWs) (AA17, AA18, AA19, AA24, AA28, AA32 and AA46) and four conventional bread wheats (CBWs) (AA03, AA15, AA25 and AA40) with TGW above 45.0 g. Grain length, width, and thickness are considered the direct measurement of TGW regarding its contribution. This was further confirmed by observing the same trends for these attributes between the genotypes as was observed for

Table 1 – Descriptive statistics of wheat grain attributes in the germplasm studied.

| Grain phenotype | Mean | Min | Max | Variance | CV |
|-----------------|-------|-------|-------|----------|------|
| | | | | | % |
| HA (mm) | 18.02 | 10.10 | 36.38 | 32.46 | 21.6 |
| HP (mm) | 17.96 | 13.99 | 25.55 | 5.98 | 13.6 |
| SL (mm) | 6.93 | 5.56 | 9.71 | 0.75 | 12.4 |
| HC | 0.68 | 0.587 | 0.77 | 0.002 | 7.1 |
| HDEV | 0.85 | 0.75 | 0.90 | 0.001 | 3.9 |
| AR | 2.12 | 1.80 | 2.75 | 0.06 | 11.9 |
| SW (mm) | 3.35 | 2.35 | 4.96 | 0.39 | 18.7 |
| SV | 333.3 | 107.4 | 954.6 | 33236.9 | 34.7 |
| ST (mm) | 3.19 | 1.52 | 5.11 | 0.48 | 21.7 |
| VP (mm) | 12.31 | 6.41 | 19.17 | 5.94 | 19.7 |
| VA (mm) | 9.54 | 2.47 | 22.47 | 16.87 | 23.1 |
| VC | 0.75 | 0.61 | 0.83 | 0.002 | 6.2 |
| VDEV | 0.68 | 0.38 | 1.08 | 0.019 | 20.3 |
| TGW (g) | 41.02 | 29.48 | 51.94 | 20.62 | 11.1 |
| FFD | 0.19 | 0.08 | 0.30 | 0.004 | 21.6 |
| DS | 0.64 | 0.35 | 0.99 | 0.019 | 21.7 |

HA = horizontal area; HP = Horizontal perimeter; SL = seed length; HC = Horizontal circularity; HDEV = horizontal deviation from ellipse; AR = aspect ratio; SW = seed width; SV = seed volume; ST = seed thickness; VP = vertical perimeter; VA = vertical area; VC = vertical circularity; VDEV = vertical deviation from ellipse; FFD = factor from density; DS = distance between IS (intersection of length and width) and CG (center of gravity); TGW = thousand grain weight.

TGW. For instance, maximum grain length (9.71 mm) and grain width (6.96 mm) were exhibited by genotype SBW10. The most consistent traits with reference to coefficient of variation (CV) were horizontal deviation from ellipse (HDEV, 9 %), vertical circularity (VC, 6 %), horizontal circularity (HC, 7 %), thousand grain weight (TGW, 11 %), followed by aspect ratio (AR, 12 %), seed length (SL, 12 %) and horizontal perimeter (HP, 13 %).

Pearson coefficient of correlation and path coefficient analysis for grain size descriptors

The Pearson coefficient of correlation (r) between the studied grain size descriptors are given in Table 2. The maximum positive correlation ($r = 0.98$) was recorded between horizontal area (HA) and horizontal perimeter (HP), between vertical perimeter (VP) and seed thickness (ST), between vertical area (VA) and seed thickness (ST), and, between vertical area (VA) and vertical perimeter (VP). Similarly, the maximum negative correlation ($r = -0.91$) was recorded between density (FFD) and seed weight (SW). Further, ST, SW, and SL had positive correlation with TGW ($r = 0.61$, $r = 0.46$ and $r = 0.62$ respectively). The FFD and VDEV (vertical deviation from ellipse), important derivatives, were negatively correlated with TGW, ($r = -0.29$ and $r = -0.31$ respectively). Similarly, the aspect ratio (AR) correlated negatively with most studied measurements.

For a further investigation of individual measurement on grain weight, the path coefficient analysis was conducted using TGW as dependent variable. The grain

Table 2 – The Pearson coefficient of correlation (r) and associated probabilities (*, ** and *** for $p \leq 0.05$, ≤ 0.01 , ≤ 0.001 respectively) between measured grain size descriptors (n = 55).

| | HA | HP | SL | HC | HDEV | AR | SW | SV | ST | VP | VA | VC | VDEV | TGW | FFD |
|------|---------------------|---------------------|--------------------|---------------------|---------------------|---------------------|----------|--------------------|---------------------|----------|---------------------|---------------------|---------------------|--------------------|---------|
| HP | 0.98*** | | | | | | | | | | | | | | |
| SL | 0.95*** | 0.97*** | | | | | | | | | | | | | |
| HC | 0.50** | 0.34* | 0.28* | | | | | | | | | | | | |
| HDEV | -0.01 ^{NS} | -0.14 ^{NS} | 0.02 ^{NS} | 0.47** | | | | | | | | | | | |
| AR | -0.59*** | -0.50** | -0.32* | -0.81*** | 0.10 ^{NS} | | | | | | | | | | |
| SW | 0.97*** | 0.93*** | 0.85*** | 0.63*** | -0.05 ^{NS} | -0.77*** | | | | | | | | | |
| SV | 0.96*** | 0.94*** | 0.93*** | 0.42** | 0.01 ^{NS} | -0.50** | 0.91*** | | | | | | | | |
| ST | 0.80*** | 0.80*** | 0.78*** | 0.34* | -0.06 ^{NS} | -0.44** | 0.77*** | 0.91*** | | | | | | | |
| VP | 0.79*** | 0.80*** | 0.77*** | 0.28* | -0.08 ^{NS} | -0.39** | 0.75*** | 0.89*** | 0.98*** | | | | | | |
| VA | 0.80*** | 0.80*** | 0.78*** | 0.31* | -0.04 ^{NS} | -0.39** | 0.76*** | 0.92*** | 0.98*** | 0.98*** | | | | | |
| VC | 0.29* | 0.24 ^{NS} | 0.22 ^{NS} | 0.34* | 0.16 ^{NS} | -0.29* | 0.32* | 0.34* | 0.43** | 0.29* | 0.40** | | | | |
| VDEV | 0.24 ^{NS} | 0.19 ^{NS} | 0.11 ^{NS} | 0.47*** | 0.03 ^{NS} | -0.52*** | 0.34* | 0.05 ^{NS} | -0.21 ^{NS} | -0.32* | -0.23 ^{NS} | 0.19 ^{NS} | | | |
| TGW | 0.56*** | 0.54*** | 0.62*** | 0.21 ^{NS} | 0.29* | -0.04 ^{NS} | 0.46** | 0.64*** | 0.61*** | 0.61*** | 0.63*** | 0.30* | -0.19 ^{NS} | | |
| FFD | -0.87*** | -0.89*** | -0.80*** | -0.48*** | 0.18 ^{NS} | 0.69*** | -0.91*** | -0.80*** | -0.70*** | -0.68*** | -0.68*** | -0.25 ^{NS} | -0.36** | -0.29* | |
| DS | 0.35* | 0.45** | 0.44** | -0.26 ^{NS} | -0.36** | 0.03 ^{NS} | 0.27* | 0.34* | 0.29* | 0.29* | 0.28* | 0.02 ^{NS} | -0.04 ^{NS} | 0.06 ^{NS} | -0.40** |

* and ** represents significance at $p < 0.05$ and $p < 0.01$, respectively and ^{NS}represent non-significant. Where, HA = horizontal area; HP = Horizontal perimeter; SL = seed length; HC = horizontal circularity; HDEV = horizontal deviation from ellipse; AR = aspect ratio; SW = seed width; SV = seed volume; ST = seed thickness; VP = vertical perimeter; VA = vertical area; VC = vertical circularity; VDEV = vertical deviation from ellipse; FFD = factor from density; DS = distance between IS (intersection of length and width) and CG (center of gravity); TGW = thousand grain weight.

size descriptors, taken as independent variables, comprised those describing the different aspects of grain size and shape as well as some miscellaneous derivatives (Figure 3). Horizontal and vertical deviations from the ellipse have an indirect positive effect on grain weight and both vertical and horizontal perimeters have direct positive effect on grain weight, because these are derivatives of grain length, width, and thickness. Grain thickness revealed a maximum direct effect (0.71) on TGW, followed by vertical area (VA, 0.68), vertical perimeter (VP, 0.61) and seed volume (SV, 0.41), while seed length (SL) exhibited the least direct effect of 0.05. Similarly, HDEV and VDEV, derivatives of seed thickness, width, and length, showed indirect positive effect on TGW. However, negative direct effect (-0.08) was observed for the derived variable FFD on TGW.

Principal component analysis (PCA)

In order to identify the most appropriate combination of the attributes studied for grain yield, the PCA and biplot analyses were conducted using mean values (Figure 4 and Table 3). Trait vectors displaying angles smaller than 90° have a positive association, while vectors with angles greater than 90° have a negative association. Further, correlation intensity increased in angles near 0° and 180° . The vector length shows the extent of variation explained by respective trait in the PCA. The first two axes, that is, PC1 (eigen value = 9.1) and PC2 (eigen value = 2.7), explained up to 74 % of the total variability. The attributes in order of their positive contribution to PC1 included HA (0.322), HP (0.316), SL (0.302), SW (0.318), SV (0.324), ST (0.303), and VP (0.295). Similarly, for PC2, the major contributing attributes were TGW (0.416) and DS (0.236). For PC3, only

Table 3 – The principal component analysis of all the traits studied measured in a set of 55 wheat genotypes.

| | PC1 | PC2 | PC3 |
|-------------------------|-------------|--------|--------|
| Eigenvalue | 9.1 | 2.7 | 1.3 |
| Variance (%) | 56.6 | 17.1 | 8.2 |
| Cumulative Variance (%) | 56.6 | 73.7 | 81.9 |
| | Eigenvector | | |
| HA | 0.322 | -0.002 | -0.085 |
| HP | 0.316 | 0.061 | -0.180 |
| SL | 0.302 | 0.146 | -0.133 |
| HC | 0.171 | -0.412 | 0.251 |
| HDEV | -0.007 | -0.370 | -0.087 |
| AR | -0.205 | 0.431 | 0.027 |
| SW | 0.318 | -0.128 | -0.085 |
| SV | 0.324 | 0.073 | 0.015 |
| ST | 0.303 | 0.087 | 0.135 |
| VP | 0.295 | 0.115 | 0.085 |
| VA | 0.300 | 0.119 | 0.148 |
| VC | 0.129 | -0.044 | 0.478 |
| VDEV | 0.069 | -0.551 | -0.142 |
| TGW | 0.198 | 0.416 | -0.116 |
| FFD | -0.292 | 0.062 | 0.240 |
| DS | 0.115 | 0.236 | -0.580 |

HA = horizontal area; HP = Horizontal perimeter; SL = seed length; HC = horizontal circularity; HDEV = horizontal deviation from ellipse; AR = aspect ratio; SW = seed width; SV = seed volume; ST = seed thickness; VP = vertical perimeter; VA = vertical area; VC = vertical circularity; VDEV = vertical deviation from ellipse; FFD = factor from density; DS = distance between IS (intersection of length and width) and CG (center of gravity); TGW = thousand grain weight.

FFD (0.624) was noted for its prominent contribution. Eigenvectors from the biplot analysis clearly indicated that HDEV, FFD, AR, and VDEV displayed a negative association.

Cluster analysis

All wheat genotypes were broadly divided into two major clusters (Figure 5). Major cluster 1 contained genotypes (AA31, AA28, AA18, AA39, AA32, AA17, AA24), all synthetic-derived bread wheats. Cluster 2 consisted of several sub-clusters designated as 2a, that is, 2a(i), 2a(ii), and 2b designated as 2b(i) and 2b(ii). The genotypes clustered in 2a(i) and 2a(ii) comprised conventional bread wheats (except AA36 and AA46 which are SBW). The check cultivars, which included AA53, AA54 and AA55, were also found confined to same cluster 2a(ii). Similar inferences were true for the rest of the genotypes in sub-cluster 2b(i) and 2b(ii).

Marker Trait Association

The scoring patterns of SSRs loci have shown 525 alleles across 55 wheat accessions. Number of alleles per locus ranged from two to 14 with an average of 5.2, indi-

cating that the diversity between wheat accessions was relatively high. The PIC values also confirmed these results. The population structure analysis was performed using STRUCTURE software version 2.1. The Q matrix was recorded by running structure at K = 2 and 7, where the highest value of ΔK occurred, demonstrating its maximum likelihood. The data obtained through the structure analysis was further validated by the clustering (UPGMA) method. For the AM analysis, K = 2 from population structure data was used. Significant markers ($p \leq 0.01$) are given in Table 4 along with their chromosomal locations. In total, 26 MTAs were found identified for 101 SSRs used in this study, of which, the number of multi-trait MTAs (markers associated with more than one trait) was 22, while trait-specific MTAs (marker associated with only one trait) were four in number. The specific trait MTAs included *Xwmc798-1BS* ($r^2 = 0.03$) for AR, *Xgdm35-2DS* ($r^2 = 0.06$) for HA, *Xgwm372-2AL* ($r^2 = 0.05$) for AR and *Xgwm544-5BS* ($r^2 = 0.03$) for distance between intersection of length and width (DS). The chromosomes with association included two from A-genome (2A and 3A), four from B-genome (1B, 3B, 5B and 7B), and two from D-genome (1D and 2D). Chromosome 2D had the largest number of MTAs (total 9 which included MTAs for HA, SL, HP, TGW, VA, VP and SW). Regarding homologous chromosomes, group 2 displayed

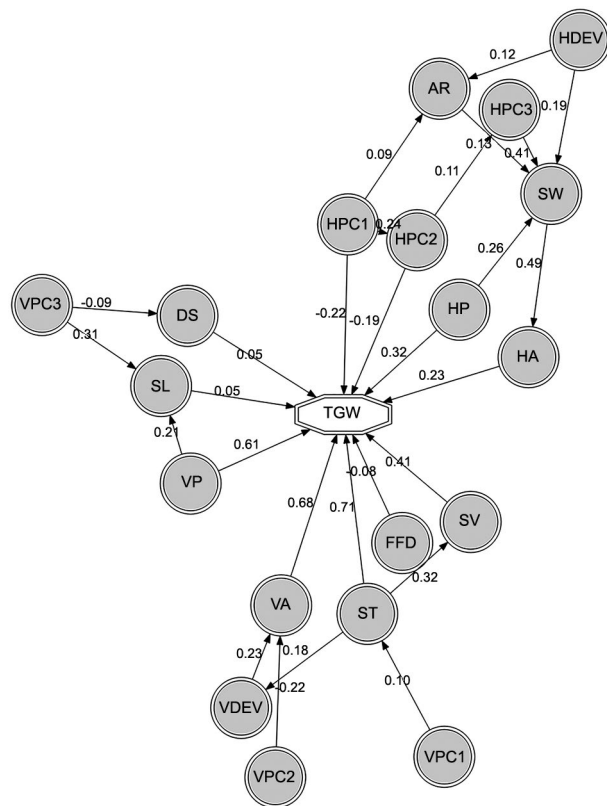


Figure 3 – The path analysis for direct and indirect effects of seed size and shape descriptors to grain weight. HA = horizontal area; HP = Horizontal perimeter; SL = seed length; HDEV = horizontal deviation from ellipse; HPC1,2,3 = horizontal principal component 1,2,3 respectively; AR = aspect ratio; SW = seed width; SV = seed volume; ST = seed thickness; VP = vertical perimeter; VA = vertical area; VDEV = vertical deviation from ellipse; VPC1,2,3 = vertical principal component 1,2,3 respectively; FFD = factor from density; DS = distance between IS (intersection of length and width) and CG (center of gravity); TGW = thousand grain weight.

Table 4 – Marker trait associations (MTAs) for the studied grain phenotypes at $p \leq 001$.

| Trait | Locus | Chr | cM | p | r ² |
|-------|---------|-----|-------|----------|----------------|
| HC | Barc147 | 3BS | 7.0 | 2.40E-03 | 0.103 |
| AR | Barc147 | 3BS | 7.0 | 5.00E-03 | 0.087 |
| SL | Barc149 | 1DS | 14.0 | 5.40E-03 | 0.039 |
| VA | Barc149 | 1DS | 14.0 | 3.49E-04 | 0.057 |
| SL | Barc149 | 1DS | 14.0 | 1.30E-03 | 0.048 |
| VP | Barc149 | 1DS | 14.0 | 1.20E-03 | 0.049 |
| SW | Barc149 | 1DS | 14.0 | 2.00E-03 | 0.045 |
| HDEV | Barc45 | 3AS | 37.0 | 1.70E-03 | 0.058 |
| VC | Barc45 | 3AS | 37.0 | 3.30E-03 | 0.052 |
| HDEV | WMC606 | 7BS | 0.0 | 9.80E-03 | 0.034 |
| DS | WMC606 | 7BS | 0.0 | 6.70E-03 | 0.036 |
| AR | WMC798 | 1BS | 19.0 | 9.80E-03 | 0.028 |
| HA | Xgdm35 | 2DS | 15.0 | 8.80E-03 | 0.055 |
| HA | Xgdm6 | 2DL | 141.7 | 4.10E-03 | 0.064 |
| SL | Xgdm6 | 2DL | 141.7 | 7.40E-03 | 0.056 |
| HP | Xgdm6 | 2DL | 141.7 | 8.60E-03 | 0.054 |
| TGW | Xgdm6 | 2DL | 141.7 | 1.90E-03 | 0.075 |
| AR | Xgwm372 | 2AL | 60.0 | 3.57E-04 | 0.053 |
| VA | Xgwm539 | 2DL | 91.0 | 1.85E-04 | 0.103 |
| SL | Xgwm539 | 2DL | 91.0 | 6.12E-04 | 0.087 |
| VP | Xgwm539 | 2DL | 91.0 | 1.70E-03 | 0.073 |
| SW | Xgwm539 | 2DL | 91.0 | 2.60E-03 | 0.067 |
| DS | Xgwm544 | 5BS | 137.0 | 4.90E-03 | 0.032 |

HA = horizontal area; HP = Horizontal perimeter; SL = seed length; HC = horizontal circularity; HDEV = horizontal deviation from ellipse; AR = aspect ratio; SW = seed width; VP = vertical perimeter; VA = vertical area; VC = vertical circularity; DS = distance between IS (intersection of length and width) and CG (center of gravity); TGW = thousand grain weight.

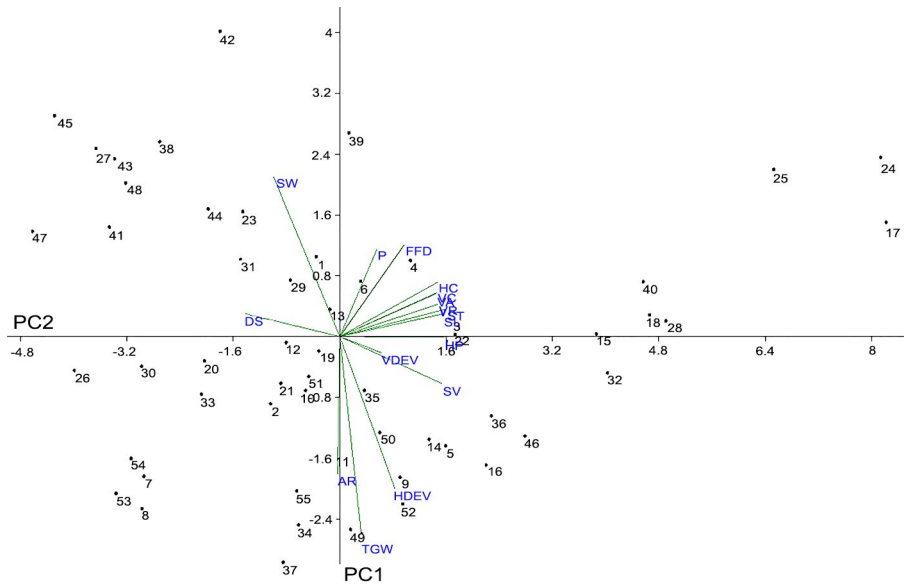


Figure 4 – The principal component and biplot analysis for PC1 and PC2 based on the trait means. Trait vectors displaying angles smaller than 90° have positive association, while those with angles greater than 90° have negative association (HP = Horizontal perimeter; HC = Horizontal circularity; SL = seed length; HDEV = horizontal deviation from ellipse; AR = aspect ratio; SW = seed width; SV = seed volume; ST = seed thickness; VP = vertical perimeter; VA = vertical area; VC = vertical circularity; FFD = factor from density; DS = distance between IS (intersection of length and width) and CG (center of gravity); TGW = thousand grain weight).

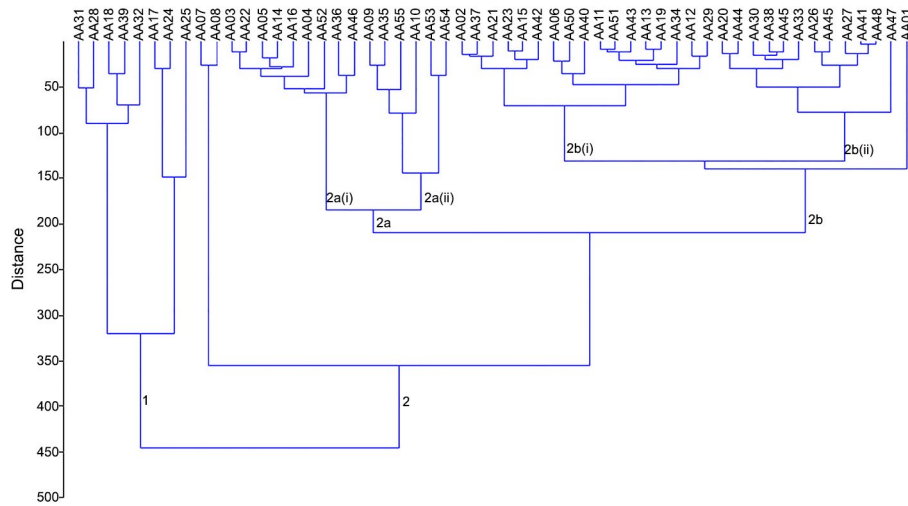


Figure 5 – Clustering analysis (UPGMA) of genotypes based on imaging data of wheat seed.

the largest number (10) of MTAs, followed by group 3 (7 MTAs) while the smallest number was observed in group 5 (with only one MTA for DS). While the homologous groups 4 and 6 showed no association with any of the studied grain phenotypes.

Discussion

Seed shape and size are vital agronomic traits because of their immense effects on yield and market value. Large collections of measurements are needed to

obtain precise seed size data because very small differences exist in size between seeds from a plant. Manual measurement methods have certain limitations, namely limited data, low quality measurements, and variety of shape data that could be gleaned. Therefore, a well-organized, consistent, high-throughput grain phenotyping method is needed to validate the genetic analysis and selection for seed shape in plant breeding (Brescghello and Sorrells, 2007; Gegas et al., 2010; Williams et al., 2013). Smartgrain software was originally used to analyze the rice seeds. However, here we are reporting digital im-

aging analysis of wheat grains to determine seed shape and size with some precise and updated attributes. This software identifies the seeds in digital image, recognizes seed outlines, and calculates seed length (SL), seed width (SW), length-to-width ratio (aspect ratio i.e. AR), seed thickness (ST, which is width of vertical image), seed area (SA), horizontal perimeter (HP), horizontal circularity (HC), distance of length to width intersection with the center of gravity (DS). As a preliminary examination, we analyzed the seeds with ImageJ software; however, SmartGrain was superior and user friendly and also described seed shape and size with more parameters as compared to ImageJ. Secondly, SmartGrain uses JPEG format, thus, there is no need to alter picture file formatting into bitmap, 8-bit or black and white as in ImageJ. Moreover, output files produced by SmartGrain are saved automatically, and could be transformed and analyzed by any spreadsheet software. Circularity and distance of L/W intersection with gravity are the additional parameters determined by SmartGrain as compared to ImageJ.

In this work, computational methods using DI imaging technique enabled to automatically measure a large quantitative dataset of robust size descriptors [AR, aspect ratio; DS, distance between IS (intersection of length and width) and CG (center of gravity); FFD, factor form density; HA, horizontal area; HC, horizontal circularity; HDEV, horizontal deviation from ellipse; HP, Horizontal perimeter; SL, seed length; TGW, thousand grain weight; ST, seed thickness; VA, vertical area; VDEV, vertical deviation from ellipse; VP, vertical perimeter; SV, seed volume; SW, seed width; TE, total effect].

Significant correlations between various seed morphological attributes provided new insights into the complex composition of grain size and shape components. For instance, vertical (VDEV) and horizontal (HDEV) deviations from ellipse were negatively correlated with grain length and width, meaning that deviation from the ellipse enhances grain length and width, and ultimately TGW. Similarly, the positive correlation between SL and SW indicated the possibility of finding some potential cultivars/genotypes possessing wider and lengthy grains simultaneously which may lead to enhanced TGW. The findings are in accordance to those reported recently by Zhang et al. (2013) and Rasheed et al. (2014); however, SL had more positive impact on TGW as compared to SW. Similar reports of mild to moderate correlations between grain weight, SL and SW with $r = 0.21-0.75$ were discussed in Shouche et al. (2001), Okamoto et al. (2013), and Abdipour et al. (2016). Similarly, the studies conducted by Gegas et al. (2010) and Williams et al. (2013) for seed shape variations targeted traits influencing grain size/weight and results are comparable to our findings. The positive direct and indirect effects of a trait on grain weight may allow its use in selection under specific conditions (Ramazani et al., 2017).

The principal component and cluster analysis revealed that variability in grain attributes could be captured quantitatively. In the PCA, differences in grain attributes were decomposed into mutually independent quantitative characteristics, that is, principal components (PCs). Using symmetric standardized coefficients for PCA, 74 % of the variation was attributed to the first two PCs. The trend observed from PCA paved a way for possible unsupervised classification algorithm development for grain type identification and classification. Similar inferences were also discussed in details by Mebatsion et al. (2012). With advances in imaging systems, identification and classification of cereal grain using one or a combination of morphological features has been attempted with different levels of success (Majumdar and Jayas, 2000; Choudhary et al., 2008). In this context, the present work is crucial as it aims at developing a consistent procedure for an objective and quantitative classification of cereal grains. General concerns in plant breeding are focused on the quantitative data generated per genotype and decrease in breeding cycle duration. This could be well achieved with the help of high throughput phenotyping techniques in breeding programs (Heffner et al., 2010). Unravelling morphological data through DIA accomplished in this study is consistent with these trends and has potential to allow direct use of quality and agronomic attributes. Thus, this study, which used grain morphometric parameters, could lead to wheat class/variety identification with useful implications in plant breeding. Precise clustering of the studied wheat genotypes based on diverse grain attributes obtained could further lead to genotype discrimination. This is supported by the fact that SBW and CBW were found clearly separated in different clusters. Further inferences could be made for individual genotypes within the sub-clusters based on its pedigree. Further studies are recommended for developing a universal statistical model to achieve more precise results (Zapotoczny, 2011). The DIA test stringency has potential for discerning genetic expression aspects of targeted diverse wheat genome accessions introgressed via interspecific hybridization and is under exploration.

For a better understanding of the effect of individual measurement on grain weight, the path coefficient analysis was conducted by having TGW as dependent variable, which showed the phenotypic model with greater precision. Grain thickness was found to reveal maximum direct effect, whereas SL exhibited the least direct effect of 0.05. Similarly, HDEV and VDEV, derivatives of seed thickness, width and length, exhibited indirect positive effect on TGW. However, negative direct effect was observed for the derived variable FFD on TGW, meaning that the loci harboring their control should undergo negative selection in order to get superior grain weight genotypes. The efficiency of indirect selection depends on heritability of the selected trait as well as correlation between a targeted and selected trait. The results further revealed that genes having pleiotropic effect or closely

linked genes might probably involve for correlations between these traits (Cooper et al., 2012; Abdipour et al., 2016). Gegas et al. (2010) and Rasheed et al. (2014) previously confirmed that kernel size and shape were largely independent traits in a study of six wheat populations.

The scoring patterns of SSRs loci have shown 525 alleles across 55 wheat accessions. Number of alleles per locus ranged from 2 to 14 with an average of 5.2, indicating that the diversity between wheat accessions was relatively high. The PIC values also confirmed these results. This indicated the importance of such diverse bread wheat germplasm in order to improve further wheat diversity and productivity in a scenario of climate change, which has posed serious problems in recent years. Henkrar et al. (2016) have also reported similar findings for enhancing genetic diversity and improving wheat productivity.

Since the experimental germplasm basically comprised synthetic derived (SBW) and conventional (CBW) bread wheat, along with check cultivars (CCT), a comparison between these groups with respect to marker trait associations revealed that most MTAs for grain shape and size variation were allocated in the D genome of allohexaploid wheat. This means that natural variations in *Ae. tauschii* populations are the most important sources to enhance the D-genome diversity in bread wheat (Jones et al., 2013). Derived-synthetic wheats are thus valuable sources in order to identify important loci from agricultural perspectives and investigate the expression of D-genome variation in different grain phenotypes that act in the hexaploid genetic background (Rasheed et al., 2014).

Genetic partition of the grain shape related attributes by means of association mapping and QTL mapping, followed by MAS (marker-assisted selection), is an active research area, and several QTLs have already been reported (Reif et al., 2011). Recently, Yan et al. (2017) identified a new QTL *QTgw.cau-2D* controlling grain weight from the synthetic allohexaploid wheat, which may play significant role in genetic improvement of wheat breeding. Similarly, Chen et al. (2019) reported on new QTLs that provided insights into the genetic basis of grain shape as well as additional genetic resources to develop elite rice varieties. In this study, 23 genomic loci were identified having association with wheat grain weight and other seed phenotypes for the germplasm grown under different water regimes. Jing-Lan et al. (2015) previously reported 208 MTAs in wheat under four different environments. They further identified SSR loci associated with some grain shape related attributes, which were common in hexaploid wheat and *Ae. tauschii*, and hexaploid wheat, suggesting that the presence of common alleles may elucidate selection between *Ae. tauschii* and hexaploid wheats, and the evolutionary process. Koebner and Summers (2003) stated that favorable allele identification could help select parents for breeding programs in order to ensure high level of favored alleles across different sets of loci. However, the probability of other genetic effects should not be ignored while considering the additive effects of

genes that resulted from linear association between grain shape phenotypes (including TGW, SL, SW and ST) and favorable alleles. Wu et al. (2012) reported on QTL related to TGW on chromosome 2D, which might be analogous to TGW-MTA found in this work with associated marker *Xgdm6-2DL*. The differences in chromosomal positions may be attributed to the different germplasm sources used, that is, mainly synthetic-derived wheats. The same MTA (*Xgdm6-2DL*) is also comparable to QTLs namely *QTKw.ncl-2D.2* (located between *Xwmc41* and *Xgwm349*) and *QTKw.ncl-2D.2* (located between *Xwmc601* and *Xwmc41*) reported by Ramya et al. (2010).

Conclusion

The integrated approach of using genomics with phenomics resulted in documentation of many genomic loci with its putative functions to enhance TGW in wheat. In this study, different alleles were identified from SSRs in the wheat germplasm studied that showed significant association with grain-shape related attributes and TGW. Some of these alleles exhibited positive effect on wheat genotypes, meaning the enhanced phenotypic value of grain-related attributes. As indirect selection indices, these traits could help deepen understanding on grain weight components in wheat. Further studies, including comparative genomics approaches, are recommended to investigate the association for sorting variations at loci with important grain shape related attributes.

Authors' Contributions

Conceptualization: Ali, A.; Ullah, Z. Data acquisitions: Ali, A.; Jamil, M. Data analysis: Ali, A.; Alam, N.; Jamil, M. Design of methodology: Ali, A.; Ullah, Z.; Alam, N. Writing and editing: Ali, A.; Bux, H.; Ullah, Z.; Naqvi, S.M.S.; Sher, H.

Acknowledgement

The manuscript is part of the Ph.D. research work and the funds provided by higher education commission (HEC) of Pakistan, are highly acknowledged.

References

- Abdipour, M.; Ebrahimi, M.; Izadi-Darbandi, A.; Mastrangelo, A.M.; Najafian, G.; Arshad, Y.; Mirniyam, G. 2016. Association between grain size and shape and quality traits, and path analysis of thousand grain weight in iranian bread wheat landraces from different geographic regions. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 44: 228-236.
- Ali, A.; Mastrangelo, A.M.; De Vita, P.; Laido, G.; Marone, D.; Arshad, M.; Naqvi, S.M.S.; Sher, H.; Rasheed, A.; Mujeeb-Kazi, A. 2017. Screening drought adaptive phenological traits in conventional and synthetic derived bread wheat germplasm and its implications for association mapping. *Cereal Research Communications* 45: 1-112.

- Ali, A.; Arshad, M.; Naqvi, S.M.S.; Rasheed, A.; Sher, H.; Kazi, A.G.; Mujeeb-Kazi, A. 2015. Comparative assessment of synthetic-derived and conventional bread wheat advanced lines under osmotic stress and implications for molecular analysis. *Plant Molecular Biology Reporter* 33: 1907-1917.
- Bland, J.M.; Altman, D.G. 1995. Multiple significance tests: the Bonferroni method. *BMJ* 310: 170.
- Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23: 2633-2635.
- Breseghello, F.; Sorrells, M.E. 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172: 1165-1177.
- Breseghello, F.; Sorrells, M.E. 2007. QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. *Field Crop Research* 101: 172-179.
- Calderini, D.F.; Reynolds, M.P. 2000. Changes in grain weight as a consequence of de-graining treatments at pre- and post-anthesis in synthetic hexaploid wheats. *Australian Journal of Plant Physiology* 27: 183-191.
- Chen, J.; Zhou, H.; Gu, Y.; Xia, D.; Wu, B.; Gao, G.; Zhang, Q.; He, Y. 2019. Mapping and verification of grain shape QTLs based on high-throughput SNP markers in rice. *Molecular Breeding* 39: 42.
- Choudhary, R.; Paliwal, J.; Jayas, D.S. 2008. Classification of cereal grains using wavelet, morphological, colour, and textural features of non-touching kernel images. *Biosystems Engineering* 99: 330-337.
- Cooper, J.K.; Ibrahim, A.M.; Rudd, J.; Malla, S.; Hays, D.B.; Baker, J. 2012. Increasing hard winter wheat yield potential via synthetic wheat. I. Path-coefficient analysis of yield and its components. *Crop Science* 52: 2014-2022.
- Dana, W.; Ivo, W. 2008. Computer image analysis of seed shape and seed color for flax cultivar description. *Computer and Electronics in Agriculture* 61: 126-135.
- Gegas, V.C.; Nazari, A.; Griffiths, S.; Simmonds, J.; Fish, L.; Orford, S.; Sayers, L. 2010. A genetic framework for grain size and shape variation in wheat. *Plant Cell* 22: 1046-1056.
- Groos, C.; Robert, N.; Bervas, E.; Charmet, G. 2003. Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theoretical and Applied Genetics* 106: 1032-1040.
- Hammer, O.; Harper, D.A.T.; Ryan, P.D. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 4: 1-9.
- Heffner, E.L.; Lorenz, A.J.; Jannink, J.; Sorrells, M.E. 2010. Plant breeding with genomic selection: gain per unit time and cost. *Crop Science* 50: 1681-1690.
- Henkrar, F.; El-Haddoury, J.; Ouabbou, H.; Nsarellah, N.; Iraqi, D.; Bendaou, N.; Udupa, S.M. 2016. Genetic diversity reduction in improved durum wheat cultivars of Morocco as revealed by microsatellite markers. *Scientia Agricola* 73: 134-141.
- Huang, X.H.; Han, B. 2013. Natural variations and genome-wide association studies in crop plants. *Annual Review of Plant Biology* 65: 410-421.
- Iwata, H.; Ebana, K.; Uga, Y.; Hayashi, T.; Jannink, J. 2010. Genome wide association study of grain shape variation among *Oryza sativa* L. germplasm based on elliptic fourier analysis. *Molecular Breeding* 25: 203-215.
- Iwata, H.; Ukai, Y. 2002. SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic fourier descriptors. *Journal of Heredity* 93: 384-385.
- Jamil, M.; Ali, A.; Ghafoor, A.; Akbar, K.F.; Napar, A.A.; Naveed, N.H.; Yasin, N.A.; Gul, A.; Mujeeb-Kazi, A. 2017. Digital image analysis of seed shape influenced by heat stress in diverse bread wheat germplasm. *Pakistan Journal of Botany* 49: 1279-1284.
- Jing-Lan, Z.; Hong-Wei, W.; Xiao-Cun, Z.; Xu-Ye, D.; An-Fei, L.; Ling-Rang, K. 2015. Association analysis of grain traits with SSR markers between *Aegilops tauschii* and hexaploid wheat (*Triticum aestivum* L.). *Journal of Integrative Agriculture* 14: 1936-1948.
- Jones, H.; Gosman, N.; Horsnell, R.; Rose, G.A.; Everst, L.A.; Bentley, A.R.; Tha, S.; Uauy, C.; Kowalski, A.; Novoselovic, D.; Simek, R.; Kobiljski, B.; Kondic-Spika, A.; Brbaklic, L.; Mitrofanova, O.; Chesnokov, Y.; Bonnett, D.; Greenland, A. 2013. Strategy for exploiting exotic germplasm using genetic, morphological and environmental diversity: the *Aegilops tauschii* Coss. example. *Theoretical and Applied Genetics* 126: 1793-1808.
- Koebner, R.; Summers, R.W. 2003. 21st century wheat breeding: plot selection or plate detection? *Trends in Biotechnology* 21: 59-63.
- Kwack, M.S.; Kim, E.N.; Lee, H.; Kim, J.; Chun, S.; Kim, K.D. 2005. Digital image analysis to measure lesion area of cucumber anthracnose by *Colletotrichum orbiculare*. *Journal of General Plant Pathology* 71: 418-421.
- Majumdar, S.; Jayas, D.S. 2000. Classification of cereal grains using machine vision. I. Morphology models. *American Society of Agricultural and Biological Engineering* 43: 1669-1675.
- Mebatsion, H.K.; Paliwal, J.; Jayas, D.S. 2012. Evaluation of variations in the shape of grain types using principal components analysis of the elliptic fourier descriptors. *Computers and Electronics in Agriculture* 80: 63-70.
- Mujeeb-Kazi, A.; Kazi, A.G.; Dundas, I.; Rasheed, A.; Ogbonnaya, F.; Kishi, M.; Bonnett, D.; Wang, R.R.R.; Xu, S.; Chen, P.; Mahmood, T.; Bux, H.; Farrakh, S. 2013. Genetic diversity for wheat improvement as conduit to food security. *Advances in Agronomy* 122: 179-258.
- Nezhad, K.Z.; Weber, W.E.; Roder, M.S.; Sharma, S.; Lohwasser, U.; Meyer, R.; Saal, B.; Börner, A. 2012. QTL analysis for thousand-grain weight under terminal drought stress in bread wheat (*Triticum aestivum* L.). *Euphytica* 186: 127-138.
- Okamoto, Y.; Nguyen, A.T.; Yoshioka, M.; Iehisa, J.C.; Takumi, S. 2013. Identification of quantitative trait loci controlling grain size and shape in the D genome of synthetic hexaploid wheat lines. *Breeding Science* 63: 423-429.
- Peng, J.H.; Ronin, Y.; Fahima, T.; Roder, M.S.; Li, Y.C.; Nevo, E.; Korol, A. 2003. Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proceeding of National Academy of Science USA* 100: 2489-2494.
- Ramazani, S.H.R.; Tajjali, H.; Ghaderi, M.G. 2017. Correlation and path coefficient analysis for determining interrelationships among grain yield and related characters in Iranian genotypes of triticale. *Bulgarian Journal of Crop Science* 54: 35-39.

- Ramya, P.; Chaubal, A.; Kulkarni, K.; Gupta, L.; Kadoo, N.; Dhaliwal, H.S.; Chhuneja, P. 2010. QTL mapping of 1000-kernel weight, kernel length, and kernel width in bread wheat (*Triticum aestivum* L.). *Journal of Applied Genetics* 51: 421-429.
- Rasheed A.; Xia, X.; Ogbonnaya, F.; Mahmood, T.; Zhang, Z.; Mujeeb-Kazi, A.; He, Z. 2014. Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. *BMC Plant Biology* 14:128
- Reif, J.C.; Maurer, H.P.; Korzun, V.; Ebmeyer, E.; Miedaner, T.; Würschum, T. 2011. Mapping QTLs with main and epistatic effects underlying grain yield and heading time in soft winter wheat. *Theoretical and Applied Genetics* 123: 283-292.
- Shouche, S.P.; Rastogi, R.; Bhagwat, S.G.; Sainis, J.K. 2001. Shape analysis of grains of Indiana wheat varieties. *Computers and Electronics in Agriculture* 33: 55-76.
- Somers, D.J.; Isaac, P.; Edwards, K. 2004. A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 109: 1105-1114.
- Su, Z.; Hao, C.; Wang, L.; Dong, Y.; Zhang, X. 2011. Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 122: 211-223.
- Sun, X.Y.; Wu, K.; Zhao, Y.; Kong, F.M.; Han, G.Z.; Jiang, H.M.; Huang, X.J.; Li, R.J.; Wang, H.G.; Li, S.S. 2009. QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. *Euphytica* 165: 615-624.
- Tanabata, T.; Shibaya, T.; Hori, K.; Ebana, K.; Yano, M. 2012. SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis. *Plant Physiology* 160: 1871-1880.
- Tsilo, T.J.; Hareland, G.A.; Simsek, S.; Chao, S.; Anderson, J.M. 2010. Genome mapping of kernel characteristics in hard red spring wheat breeding lines. *Theoretical and Applied Genetics* 121: 717-730.
- Visen N.S.; Paliwal, J.; Jayas, D.S.; White, N.D.G. 2001. Specialist neural networks for cereal grain classification. *Biosystems Engineering* 82: 151-159.
- Williams, K.; Munkvold, J.; Sorrells, M. 2013. Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (*Triticum aestivum* L.). *Euphytica* 190: 99-116.
- Williams, K.; Sorrells, M.E. 2014. Three-Dimensional seed size and shape QTL in hexaploid wheat (*Triticum aestivum* L.) populations. *Crop Science* 54: 98-110.
- Wu, X.; Chang, X.; Jing, R. 2012. Genetic insight into yield-associated traits of wheat grown in multiple rain-fed environments. *PLoS ONE* 7: e31249.
- Yan, L.; Liang, F.; Xu, H.; Zhang, X.; Zhai, H.; Sun, Q.; Ni, Z. 2017. Identification of QTL for grain size and shape on the d genome of natural and synthetic allohexaploid wheats with near-identical AABB genomes. *Frontiers in Plant Science* 8: 1705.
- Zapotoczny, P. 2011. Discrimination of wheat grain varieties using image analysis: morphological features. *European Food Research and Technology* 233: 769-779.
- Zhang K.; Wang, J.; Zhang, L.; Rong, C.; Zhao, F.; Peng, T.; Li, H.; Cheng, D.; Liu, X.; Qin, H.; Zhang, A.; Tong, Y.; Wang, D. 2013. Association analysis of genomic loci important for grain weight control in elite common wheat varieties cultivated with variable water and fertiliser supply. *PLoS ONE* 8: e57853.