IDENTIFICATION OF VALUABLE TRAITS THROUGH MOLECULAR AND MORPHOLOGICAL MARKERS IN DIPLOID WHEAT

Zareen Sarfraz1*, Mohammad Maroof Shah1, Muhammad Shahid Iqbal2, Mian Faisal Nazir2, Syeda Akash Fatima3

1Department of Biotechnology, COMSATS University Islamabad, Abbottabad Campus, Abbottabad, Pakistan, 2Ayub Agricultural Research Institute, Faisalabad, Pakistan, 3Department of Plant Breeding and Genetics, MNS University of Agriculture, Multan, Pakistan. E.mail: *zskpbg@hotmail.com

Article received 15.4.2020, Revised 18.5.2020, Accepted 29.5.2020

ABSTRACT

Bread wheat (Triticum aestivum L.) harbors hexaploid genome with a complex of three diploid genomes (A, B and D). Major goals of today’s research in the wheat world count basically on improvement of grain quality and high yielding genotypes. It is relatively easier to target its diploid lineages for identification of high yielding genomic segments as compared to hexaploid with least polymorphism. The current study was aimed at the detection of genes/QTLs of high economic interest on ‘A’ genome of diploid progenitors. Segregating population from cross of Triticum monococcum and Triticum boeoticum was utilized for the purpose. 98 RILs along with parents were analyzed for various quantitative and qualitative traits. The molecular data obtained from RAPD markers was analyzed along with morphological traits data to identify association among them. Single factor ANOVA gave association (a= 0.01) of days to booting, days to heading and days to flowering with plant type and leaf orientation, flag leaf width with crown color, number of spikes per experimental unit with leaf orientation, 1000 kernel weight with aleurone color and RAPD-11 (locus b).

Keywords: A-genome donor, Wheat, RILs, RAPDs, QTLs

INTRODUCTION

Wheat (Triticum aestivum L.) is a member of Grass family i.e. Poaceae, belongs to the genus Triticum. It is a major cereal crop in the world and is used as a staple food worldwide. Due to cultivation on large acreages and thus high production it holds a major position in the International food trade market hence, called the ‘King of cereals’. Ethiopian and Mediterranean regions are world renowned centers of its origins however, central Asia and Near East (Kundu and Nagarajan, 1996) Hindukush regions are described as the center of diversity for Hexaploid wheat (Kundu and Nagarajan, 1996). The evolutionary history of wheat began about ten thousand years ago, comprising following steps (i) Aegilops speltoides (Goat grass 1) naturally crossed with T. urartu (Wild einkorn) and T. dicoccoides (Wild Emmer) was produced in result. (ii) Durum Wheat and T. spelta got produced from further hybridizations of other Aegilops species. (iii) It is reported that group of tetraploid emmer had been evolved from natural crossing of T. monococcum (Einkorn) with unknown donor of B-genome. (iv) Ae. Squarrosa then hybridized with Emmer wheat and chromosomes of the triploid got spontaneously doubled to produce the hexaploid wheat (Feldmann, 2001). Cytogenetics studies have shown that Triticum species consist of three genome groups regarding their chromosome number such as Hexaploid (2n=6x =42), Tetraploid (2n= 4x=28) and Diploid (2n=2x =14) keeping the basic chromosome number x=7.

Major hurdle in the mapping of Triticum aestivum genome is existence of low level of polymorphism especially in cultivated genotypes. Also T. aestivum has huge genome i.e. 16× 109bp (Gupta et al., 1991), with abundant amount of repetitive DNA that accounts for 80% of its genome; making it arduous for genomic studies. On the other hand, T. monococcum (donor of A genome) genotypes either cultivated or wild have comparatively higher level of polymorphism hence, widely used for genetic maps construction to identify important loci also existent in T. aestivum genome (Castañga et al., 1994). Similarly, T. tauschii which had donated the D genome to the T. aestivum can be used for the linkage map construction (Lagudah et al., 1991). T. monococcum and its subspecies viz; T. aegilopoides are closely related to the A genome of T. aestivum (Johnson and Dhaliwal, 1976). T. boeoticum; a subspecies of T. monococcum got hybridized with T. monococcum and viable hybrid seeds were produced as a reslut after the successful formation of seven bivalents (Johnson and Dhaliwal, 1976). Triticum boeoticum is very similar to Triticum urartu for most of its genetic make-up (Johnson and Dhaliwal, 1976) and possesses vast amount of variability regarding stresses either biotic or abiotic and for various yield cont-
ributing traits (Nazir et al., 2019). It mainly includes resistances against many diseases and insect pests such as powdery mildew, leaf (Hussien et al., 1997) and stripe rust (Harjit Singh et al., 1998), karnal bunt (Vasu et al., 2000), aphids (Migui and Lamb, 2004) and cereal cyst nematodes. *T. monococcum* (Einkorn) is different from *T. boeoticum* for three important traits viz; bigger round shaped seeds, easily threshable and hard rachis to avoid spike damage (Salamini et al., 2002).

The A genome of Hexaploid wheat is not well conserved especially chromosomes 4A and 5A undergo translocation mutations; also, chromosome 4A of Hexaploid wheat is different from 4Am of *T. monococcum* by having inversion mutations (Mickelson-Young et al., 1995). Some traits are controlled by many genes and are more sensitive to the environmental changes; larger sized populations and standardized statistical methods are utilised for their study to determine their genetic architecture. The parts of genome which are controlling the expression of such traits are termed as Quantitative trait loci, abbreviated as QTL ( Tanksley, 1993). To determine the influence and genomic position of QTLs with the help of molecular markers is called as QTL mapping (Tanksley, 1993, Sarfraz et al., 2018). Mapping defines the placement of molecular markers in correct order with respect to their genetic distance on basis of their recombination frequencies. Mapping population is derived from cross between individuals of same species differing for traits under study. RILs are the progeny individuals of self-pollinated plants obtained from continuous selfing of F2 population utilized for the mapping of genes related to traits under investigations. Homozygosity is achieved in the self-pollinated plants after six to seven generations only whereas in cross-pollinated individuals almost 20 generations are needed for gene fixation. As at this stage, recombination between chromosomes produces no alteration of genetic make-up and no segregation of genes results (Kahl, 2004). Also, the map generated from RILs population is of high resolution as compared to F2 population (Burr and Burr, 1991). The foremost purposes of QTL mapping in plant species constitute (a) Identification of hidden molecular markers existing in the germplasm which further can be utilized as indirect method of selection. (b) Upgradation in the level of our knowledge regarding plant genomes for the sake of appropriate genes arrangement (Bernardo, 2008). RAPDs were the first DNA markers to be used in the genomic studies for the sake of genotyping as well as QTL mapping via PCR as they are easy to handle, take less time and don’t require sequence information. It has been reported that with RAPDs, many genomic portions get amplified in a single PCR by chance when a sequence of 10 nucleotides in single primer finds complementary sequence there (Williams et al., 1990, Raheel et al., 2015). RAPD markers namely Pr11 (230bp), Pr19 (240bp), OPU06 (340bp) and OPH13 (450bp) were discovered to be linked with gene of flag leaf senescence and they are indicators of tolerance to water scarcity condition (Sana et al., 2011).

**MATERIALS AND METHODS**

**Plant material:** The germplasm used was a set of 98 recombinant inbred lines produced from a cross of two diploid parents viz: *T. boeoticum* acc. pau5088 × *T. monococcum* acc. pau14087. Dr. BL Johnson, University of California, Riverside, USA actually provided the seeds of both the species. These segregating lines were at F7 generation through single seed descent method of breeding. The two ‘A’ genome species were utilised for the experiment specifically because they hold a wide range of variability regarding numerous agronomic traits prominently as well as resistance to biotic or abiotic stresses.

**DNA extraction:** DNA was extracted from 100 samples including 2 parents and 98 RILs following the CTAB method. 4-5 young leaves for each sample were collected and kept at -80°C overnight. Autoclaved Pestle and Mortar were also kept at -80°C. 2-3 leaves were ground into fine powder which was then put into autoclaved Eppendorf with sterilized spatula with great care so that no powder remains behind because there are chances that only the cell debris gets picked up and the nuclear portion remains behind. Then 700 µL of pre-heated CTAB buffer was added up into the Eppendorfs. Then the samples were incubated in water-bath for 30-45min which was already set at 65°C. During incubation gentle inversion was done at an interval of 15min. After incubation, 700µL of Chloroform isooamyl-alcohol was put in samples. Then the samples were centrifuged at 13000rpm (rotations per minute) for 15min. Three layers were formed in the supernatant and the top clear one was collected. Equal amount of 95% chilled Ethanol was added up in samples and were kept overnight at -20°C. Then samples were centrifuged at 13000rpm for 10min. A pellet was formed at the bottom of Eppendorf after centrifugation. Supernatant was discarded and the pellet was then washed with 70% chilled Ethanol. Air-drying of pellet was done which was followed by its dissolution in 100µL of 1X TE buffer.
DNA purification: After DNA extraction, purification was done before running on gel. 2.5 volumes of 95% Ethanol and 1/10th volume of Sodium acetate were added together in extracted DNA sample with gentle inversion. Then the samples were kept at -20°C for half an hour and then centrifuged at 13000rpm for 10min which gave clear pellet at bottom. Supernatant was discarded-off and 100µL of 1X TE buffer was added to dissolve the pellet which was then followed by its washing with 70% Ethanol and air-drying.

**RNase treatment:** 4µL of diluted RNase-A solution (double distilled autoclaved water and RNase-A in a ratio of 9:1) was put in extracted DNA sample and centrifuged at 3000rpm for 30 sec. Then the sample was incubated at 37°C for an hour. After incubation, sample was again centrifuged at 3000rpm at 30sec.

**Gel electrophoresis:** 0.8% agarose gel was used for running of DNA sample. Ethidium bromide (5µL/100mL) was added in solution to increase the molecular weight of DNA. Loading of samples was done with 1X loading dye mixed in a ratio of 1:1 before loading. 100V were given for 25min for running of samples.

**Polymerase chain reaction:** 18 Gene Link RAPD Decamer primers were used on both Parents to check polymorphism. 5 of them were polymer- phic which were then applied on 98 RILs. The quantities of reagents used in optimized polymerase chain reaction are given in Table 1 which was the same as given by Röder et al., (1998).

Table 1: Amounts of PCR reagents used for the amplification of template DNA

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Reagents</th>
<th>Amount (µL/reaction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10X Taq Buffer</td>
<td>2.5</td>
</tr>
<tr>
<td>2</td>
<td>dNTPs (20mM)</td>
<td>0.5</td>
</tr>
<tr>
<td>3</td>
<td>MgCl₂ (25mM)</td>
<td>3.75</td>
</tr>
<tr>
<td>4</td>
<td>Primer (10pmol/µL)</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>DNA (50ng)</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>Taq Polymerase (5U/µL)</td>
<td>0.5</td>
</tr>
<tr>
<td>7</td>
<td>PCR water</td>
<td>11.25</td>
</tr>
</tbody>
</table>

Table 2: PCR Cycling Profile used to make copies of template DNA

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Steps</th>
<th>Temperature (°C)</th>
<th>Time Duration</th>
<th>Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initial Denaturation</td>
<td>94</td>
<td>5 min</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Denaturation</td>
<td>94</td>
<td>30 sec</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Annealing</td>
<td>50</td>
<td>1 min</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>Extension</td>
<td>72</td>
<td>1 min</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Final Extension</td>
<td>72</td>
<td>7 min</td>
<td>1</td>
</tr>
</tbody>
</table>

**Gel analysis:** PCR amplified products were analyzed on 1.5% agarose gel prepared in 1X TBE buffer. A gel run of 85min at 75V was given. The loading of samples was done in a way that 1µL of 6X Bromophenol Blue was mixed with 5µL of PCR product through micro-pipette and was loaded up in well. Polymorphic bands were scored as 1 for the presence and 0 for the absence for each polymorphic band that resulted in a matrix of 1 and 0.

**Field experiment:** 98 RILs and 2 Parents were sown in field in a triplicated trial according to Randomized Complete Block Design (RCBD). Data for different morphological traits both quantitative i.e. germination, flag leaf area, flag leaf sheath length, peduncle length, tillering, spike length, days to booting, days to heading, dry biomass, plant height, seed weight, 1000 kernel weight and grain yield and qualitative i.e. plant type, leaf orientation, aleurone color, coleoptile color and crown color were taken.

**Analysis of variance:** All the phenotypic traits scored were analyzed by Statistical Analysis System (SAS) software with version 9.2 PROC GLM command (SAS Institute, Cary, NC). Analysis of variance (ANOVA) for germination percentage, flag leaf area, flag leaf sheath length, peduncle length, pillaring, spike length, days to booting, days to heading, days to flowering, plant height, dry biomass, seed weight, 1000 kernel weight and grain yield was carried out at α=0.05 probability level by following the method given by Steel et al. (1997). Single-factor analysis of variance was then carried out at α=0.05 probability level using SAS software with version 9.2 PROC GLM command (SAS Institute, Cary, NC). This analysis was done to find out the association of phenotypic traits with molecular and phenotypic markers. The associated phenotypic traits and markers information was used as input data for stepwise regression analysis. The stepwise regression analysis was also carried out using SAS software version 9.2 PROC GLM command (SAS Institute, Cary, NC). The analysis was done at α=0.01 probability level, to find out the markers which are tightly linked to the phenotypic traits.
RESULTS AND DISCUSSION

High quality DNA extracted from parents and RILs is given in figure 1.

![Gel image of DNA of Parents and RILs resolved on 0.8% agarose gel](image)

Figure 1: Gel image of DNA of Parents and RILs resolved on 0.8% agarose gel

High molecular weight intact DNA of Parents and 98 recombinant inbred lines was used for polymerase chain reaction. Firstly, the high-quality DNA of *Triticum boeoticum* and *Triticum monococcum* was used for polymerase chain reaction of 18 Gene Link Decamer (GLD) RAPD primers for polymorphism survey. The polymorphism survey showed 5 GLD RAPDs polymorphic which were then applied on 98 recombinant inbred lines.

![PCR results of RAPD A-19](image)


![PCR results of RAPD A-19 M](image)


**Field data evaluation:** Analysis of variance of *T. boeoticum*, *T. monococcum* and 48 recombinant inbred lines was carried out for flag leaf area, flag leaf sheath length, peduncle length, tillering, spike length, number for spikes per experimental unit, seed weight, 1000 kernel weight, and grain yield by using SAS software version 9.2 with PROC GLM command which gave highly significant differences among all the genotypes at \( \alpha = 0.01 \) probability level (Sarfraz et al., 2016a; 2016b; Sarfraz et al., 2020). Different statistical parameters were calculated from this analysis viz; sum of squares, mean sun of squares, coefficient of variation and R-square values.

**QTLs identification:** One to three significant (P ≤0.05) marker locus-trait associations were determined from the results of single-factor ANOVA for the phenotypic traits and marker associations. Two markers were observed to be significantly associated with germination, three markers with days to booting, two with days to flowering, two with biomass, two with 1000 kernel weight, three with spike number per experimental unit, three markers were observed to be associated with days...
to heading, two markers with flag leaf length and
two with flag leaf sheath length. None of the mar-
ker association was observed with tillering, pedu-
ncle length, plant height, spike length, seed weight
and yield. It might be possible that QTLs with
smaller effects were responsible for the traits,
which were not found to be associated with any of
the markers. There is also a chance that those
QTLs were more vulnerable to the environmental
effects that is why they remained unidentified
with the used markers.

Table 3: List of significantly associated markers with Quantitative traits viz; germination, days to booting, days to
flowering, biomass with and without spikes, 1000 kernel weight, spike number per experimental unit, days to
heading, flag leaf length and flag leaf width calculated from single factor ANOVA from SAS software version 9.2
with PROC GLM command (SAS Institute, Cary, NC) for the RILs of T. booeicum and T. monococcum cross.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Markers significantly (p&gt;0.01, 0.05) associated with trait</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germination (ger)</td>
<td>Crown color, Leaf orientation</td>
</tr>
<tr>
<td>Days to Booting (DB)</td>
<td>Plant type, Leaf Orientation, RAPD-9(a)</td>
</tr>
<tr>
<td>Days to Flowering (DF)</td>
<td>Plant type, Leaf Orientation</td>
</tr>
<tr>
<td>Biomass (Bss)</td>
<td>Leaf Orientation, RAPD-11(b)</td>
</tr>
<tr>
<td>1000 kernel weight (TKW)</td>
<td>Aleurone color, RAPD-11(b)</td>
</tr>
<tr>
<td>Spike no. per experimental unit (SPEU)</td>
<td>Coleoptile color, Leaf Orientation, RAPD-19</td>
</tr>
<tr>
<td>Days Heading (DH)</td>
<td>Plant type, Leaf orientation, RAPD-19</td>
</tr>
<tr>
<td>Flag leaf length (FLL)</td>
<td>Coleoptile color, RAPD-11(b)</td>
</tr>
<tr>
<td>Flag leaf width (FLD)</td>
<td>RAPD-11(a), Crown color</td>
</tr>
</tbody>
</table>

It is the summary of the single factor ANOVA
and stepwise regression analysis of the molecular
and phenotypic markers for the identification of
QTLs associated with above mentioned morpho-
logical traits of diploid wheat. These above-men-
tioned traits were estimated to be highly associated
with the markers. Single factor ANOVA was
followed by stepwise regression analysis for the
determination of association of complex traits
with phenotypic and molecular data. Significant
associations among flag leaf parameters, coleo-
ptile color, crown color, plant type, 1000 kernel
weight, weight of seeds, aleurone color, number
of spikes, RAPD-09, RAPD-11 (Máric et al., 2004)
have been observed from single factor ANOVA
(Khaled et al., 2015), depicting that they are sharing
same chromosomal regions which can be
detected by QTL mapping of them using Mapping
softwares like MAPMAKER, MAPL and TESSEL
etc. As associated or correlated traits often share
same regions on a chromosome (Abler et al.,
1991). These significantly associated markers were
used as input data for stepwise regression analy-
sis. A range of 2.5% to 49.2% of total phenotypic
variation was shown by individual loci. Single
marker locus (leaf orientation) was observed to be
significantly associated (P < 0.01) with number of
spikes per experimental unit with 2.5% of overall
phenotypic variation (Li et al., 2007). Single locus
(crown color) was estimated to be significantly
associated with flag leaf area exhibiting 14.45% of
total phenotypic variation. Two loci (plant type
and leaf orientation) with days to booting (50.2
%), days to heading (73.1%) (Khaled and Hamam,
2016) and days to flowering (56.7%) were found
to be significantly linked and accounted above
mentioned phenotypic variation. Single locus (RAPD-11) with 13.5% of overall phenotypic variation was observed to be significantly associated with 1000 kernel weight (Galande et al., 2001). These percentages of phenotypic variation indicated that most of genotypic variation in these traits is due to these loci or chromosomal regions. Our findings for linkage of morphological traits of wheat with RAPDs are in congruence with those Khaled et al., (2015). Their locations on the chromosomes were not determined due to the availability of few markers which can be further worked out easily by applications of more molecular markers on samples.

The chromosomal location of these putative QTLs can be further detected with the help of some more markers linked with them and by calculating their distances (centimorgan) from each other and ultimately their order on chromosome. The identified genes can now be transferred to the bread wheat after their localization on the chromosomes for the betterment of wheat after their localization on the chromosomes.


Migui, S.M. and R.J. Lamb, Seedling and adult plant resistance to Sitobion avenae (Hemiptera: Aphididae) in Triticum monococcum
Identification of valuable ----


 Sarfraz, Z., M.M. Shah and M.S. Iqbal, Estimation of components of variation of morpho-


