Identifying profiles of SSR and SNP markers in cultivars of tetraploid wheat: physical and chemical analysis

Iman Yousefi Javan[a], Mario Augusto Pagnotta[b],*, Renato D′Ovidio[b]

[a]Assistant Professor, Department of Plant Production, Faculty of Agriculture, University of Torbat Heydarieh, Torbat Heydarieh, Iran

[b]Department of Agricultural and Forest sciences (DAFNE), Tuscia University, Via S. C. de Lellis, snc, 01100 Viterbo, Italy

Abstract

In this work, we studied the distribution of molecular markers in the chromosomes of tetraploid wheat. This distribution was drawn through the 192 line came from based on a cross between two durum wheat genotypes. The first parental line was a *Triticum turgidum* ssp. *durum* (Desf.). The second parental line was a durum wheat genotype derived from a cross between the *Triticum turgidum* ssp. *Durum,* (Omrabi 5). One hundred ninety-two F8 recombinant inbred lines (RILs) derived from the above mentioned cross by single-seed descent. A total of 254 markers were analyzed, including 216 microsatellites and 38 SNPs markers. Linkage analysis defined 14 linkage groups. Most markers (57.2%) were found to be located to the A genome, with an average of 12 markers per chromosome. The remaining (42.7%) were located to the B genome. To construct a stabilized (skeleton) map, markers interfering with map stability were removed. Efficient user-friendly methods for mapping plant genomes were highly desirable for the studies marker-assisted selection. SSR (microsatellite) markers are user-friendly and efficient in detecting polymorphism, but they detect few loci. The skeleton map consisted of 100 markers with a total length of 3170.29 cM and an average distance of 31.7 cM between adjacent markers. Majority of the markers showed a statistical significantly Mendelian segregation with 1:1 ratio (α=0.01). The highest percentage of markers was similar with the first parental. This SSR and SNP markers revealed a high proportion of clustering, which may be indicative of gene-rich regions. Some of the SSR, SNP markers were distributed for the first time on the current work. This project provided a useful groundwork for further genetic map, genetic analysis of important quantitative traits, positional cloning, and marker-assisted selection, as well as for genome comparative genomics and genome organization studies in wheat and other cereals.

Keywords: Genome; SSR; wheat; tetraploid.

Introduction

Simple sequence repeat (SSR) is a PCR-based method to detect polymorphisms. SSR has been widely used for producing 48 lines of wheat and its relatives. The objectives of the project are [1] to add genotype information to the stocks, and [2] to

*Corresponding author: Mario Augusto Pagnotta

Tel: +39 (3402779310), Fax: +39 (3402779310)
E-mail: pagnotta@unitus.it

find a set of SSR markers suitable for detection of diversity in wheat species. Durum wheat is a tetraploid species (2n=28, genomes AABB), with about 21.0 million hectares under cultivation (about 8% of the total wheat cultivated area) durum wheat ranks eighth among all cereals. One of the most widely grown food grain crops in the world, providing about one-fifth of the calories consumed by humans. Nowadays, several technologies are available to increase the abundance of DNA markers and to contribute in developing high resolution genetic maps suitable for genetic analysis and to perform marker assistance selection to breeding activities. Emmer wheat, *Triticum turgidum* L. subsp. *durum* cultivated in the past [1-2], mostly grown in modern organic farms especially in Italy and considered a candidate crop for sustainable agriculture [3]. The relevance of this minor cereal has been demonstrated by basic research on quality, adaptation, breeding and promotion of both locally adapted relict populations [4] and on gene bank collections [5]. During the last decade field evaluations of germplasm collections have been carried out to assess variation in agronomic [1,3] and quality traits [6-8]. Emmer wheat is the base population from which the founder genotypes of durum wheat populations were derived, it represents a genetic resource for durum wheat tivars, providing good genes for resistance to biotic and abiotic stress [1,9]. Pre-breeding and breeding work aimed at improving yield stability and quality traits [4-5]. Recently, two new cultivars have been selected and registered. Nonetheless, to complete the knowledge of the dicoccon gene pool genetic structure, a genetic study at molecular level is also required. At the inter-specific level, the greatest morphological difference between emmer and the closest domesticated relative, the ''durum wheat'' is due to hulledness which is mainly controlled by a single dominant gene, the Q factor, on chromosome 5A [10-11]. Cereal production and its sustainability are largely influenced by environmental stresses, such as drought, causing reduced crop yields. Limited water supply is a major constraint on wheat production worldwide, particularly in arid and semiarid regions. About 45% of the 120 million ha of land allocated in developing countries to wheat production is prone to drought [12]. The progress in improving drought tolerance of wheat cultivars with empirical breeding has been slow due to the poor definition of the target environment, the complexity and difficulty of drought evaluation procedures, the inconsistency of morpho-physiological traits as selection criteria for drought tolerance and the interaction between genotypes and environmental factors [13]. Selection and breeding for grain yield is the ultimate way for producing the stress tolerant crop plants [7]. However, due to the low heritability and complexity of grain yield, and its high dependence on environmental conditions, other traits such as yield components associated with drought resistance and with higher heritability than grain yield can be employed (analytical breeding). Also some traits are associated with plant adaptability to drought and improve the yield under drought stress environment. For instance, phenological and physiological traits, plant height, harvest index, and components of yield are selection criteria involved in improving yield under drought stress conditions; however, some of them

may affect yield indirectly. For example, an extended duration of the stem elongation phase without a change in the timing of anthesis has been proposed as a physiological determinant of an increase in the number of grains per m^2 without altering the amount of water used by the crop [14]. Meiosis, anthesis, and male and female fertility are all extremely susceptible to drought stress and their failure directly affects kernel number, thus leading to a significant yield penalty [15]. Molecular markers are efficient tools to speed up crop improvement and for the construction of molecular linkage maps, the first step in the genetic dissection of target traits [16].

Purpose that is brought forward in this important work includes:

1- The distribution of the markers, to know the proximity of relatives and between the lines came from these relatives.

2- Information on the sequences of the markers that will be able to have linkage with different chromosomes, and different traits.

Through the molecular markers, we can draw the genetic maps, that is a key step in convenient linkage analysis of biologically and agronomically important traits. Genetic linkage maps are a fundamental tool for several purposes, such as evolutionary genomics, understanding the biological basis of complex traits, dissection of genetic determinants underlying the expression of agronomically important traits, and facilitating marker-assisted selection (MAS) and map based cloning. Microsatellites (simple sequence repeats, SSRs) are PCR-based markers characterized using a high level of polymorphism that permits to discriminate among wheat breeding lines [17]. PCR-based

molecular markers such as microsatellites and simple sequence repeats (SSRs) are easy to be used, exhibiting a high degree of polymorphism. As of now, 150 microsatellite markers have been added to wheat genetic maps [1,18]. While SSRs are standard PCR based markers and can be considered as proven anchor markers, their suitability for high-throughput mapping is not favorably comparable to the new single nucleotide polymorphism (SNP)-based genotyping techniques [19]. Moreover, SSR-multiplexing requires an extensive and additional optimization [18].

Applications of molecular markers

Development of molecular techniques for genetic analysis increased our knowledge of cereal genetics and our understanding of the structure and behavior of the cereal genomes. Subsequently, it was possible to look at more subtle variation in the structure of polypeptides. Since 1980s, it has become possible to explore variation at the level of DNA itself. The recent progress in DNA analysis techniques and data analysis have greatly increased our ability to understand the genetic relationship among organisms at the molecular level.

The molecular markers have several advantages over the morphological markers: (1) Numerous markers can be identified in breeding materiel; (2) A relatively large number of alleles can be found; (3) Most molecular markers show codominant mode of inheritance; (4) Molecular markers are generally silent in their effect on the phenotype; (5) Genotypes of most molecular markers can be determined at a very early developmental stage. Among the different classes of molecular markers available, microsatellites (SSR, simple sequence repeats) include an important role. Up to 90% of the plant genome

consists of repetitive DNA. Tandemly repetitive DNA is classified into three major classes [2]:

(i). Satellite DNA, which shows repeat units with a length of up to 300 base pairs (bp).

(ii). Minisatellite comprised between 9 and 100 bp.

(iii). Microsatellite or simple sequences that exhibit repeats of 1-4 bp in length [3].

Materials and methods

Plant material

In this project 192 individuals of *Triticum turgidum* subsp. *durum* belonging to the recombinant inbred lines (RILs) utilized to generate a genetic linkage map. The RIL population was developed and cultivated at ICARDA, Aleppo, Syria. (Nachit pers.com) The population was in $F_{8:9}$ (seven generations of selfing) by single-seed decent (SSD) after a cross between 2 durum wheat.

First parental was the genotype 112 of the RIL from a previous cross between two durum wheat lines Jennah Khetifa and ChamI. It was chosen the RIL number 112, for its traits: (i) high quality in the production of pasta, couscous, and burghul; and (ii) resistance against the abiotic stresses, such as drought, cold and heat. The second parental was genotype 101 of the RIL from Backcross between two durum wheat lines (Omrabi5 × *T.dicoccoides*) × Omrabi5.

SSR markers

A total of 216 genomic SSR primer pairs were screened using two parental lines. Markers were prevalently chosen within the public SSRs. These markers consisted primarily of Gatersleben Wheat Microsatellites [20] and a few additional markers from Wheat Microsatellite Consortium [18], INRA Clermont-Ferrand [21-22], and Beltsville Agriculture Research Center [23]. These markers showed codominant alternative alleles between the two parental lines and covering all 14 chromosomes of tetraploid wheat. Some of the SSRs used in this study were mapped in previous published maps including, the durum wheat mapping population [14,23] Most of these primers generated SSR loci that were not previously mapped either [4]. We used PCR methods for goals, simplicity, speed, and a small amount of starting material. Simplicity and speed are absolutely essential for processing large numbers of individuals an obvious necessity when large populations of several hundred, or even thousands, of individuals need to be examined. DNA used for this segregation of markers should be highly purified. The SSRs markers were amplified using the PCR in the plates of 96, in 10 ul of reaction containing 3 ul genomic DNA with a concentration of 5 ng/ul, 5 μl of H_2O Mol Bio grade autoclaved, 1X PCR Buffer Promega (100 mM Tris-HCL – pH 8.3, 500 mM KCL, 1.5 mM MgCl2), 0.002 μl Forward primer (→) labeled with 0.02 μl colored fluorescent dyes or M13 tail (FAM, VIC or Hex, NED and PET) with a size respectively 494, 538, 535, 546 and nm (absorbance maximum), (M13 appointed a DNA fragment with the sequence complement to a tail added into 5' of the Forward primer to give the relevance the gene or interest fragment with the two primers), 0.02 μl unlabeled Reverse primer (←), 0.6 ul each dNTP 100 mM(25umol), e 0.05 μl di AmpliTaq-DNA Polymerase (Applied Biosystems, Foster City, USA). PCR amplifications were performed in a thermocycler Eppendorf (Master cycler) and Biometra using the following conditions according to

Röder *et al.* [24], 2 min at 94 °C, followed by 35 cycles of 30 s 94 °C, 30 s annealing (between 50 °C and 65 °C, depending on the optimal annealing temperature of the primers), 1 min 72°C, and a final extension of 10 min at 72 °C, and finally the samples were kept at 8 °C.

The reactions of SSR markers for the parental screening were grouped on the basis of the annealing temperature. This was verified for each marker through the formula.

69.3+ [0.41× (% GC) – 650/L], and the site of Grain Genes.

SNP Markers

Two different typologies of SNP mutations were analyzed: (i) unspecific SNPs on genes with different functions localized on chromosomes 2 and 4 and (ii) SNPs specific for genes involved in drought and salinity stresses tolerance. The first class of SNPs has been developed employing the SNP mutations available on database "snpdb/Haplotype Polymorphism in Polyploid Wheats and their Diploid Ancestors". An online SNP database (http://probes.pw.usda.gov:8080/snpworld/Search) was constructed. Utilizing the mutations localized on exon sequences on chromosomes 2 and 4. To identify the polymorphism of SNP markers we used high resolution melting (HRM) that is a novel and homogeneous, to analyze genetic variations (SNPs, mutations, methylations) in PCR amplicons. As this technique requires the employment of amplification fragment no longer than 100 bp, new primer pairs for all the SNPs were designed to obtain fragments with the right length.

Statistical analysis

To develop the second class of SNPs, several sequences deriving from UniGene (gene-oriented clusters of transcript sequences) and "mRNA complete cds" from wheat, Arabidopsis, barley and rice species were multi-aligned using ClustalW2 (Multiple Sequence Alignment, EMBL-EBI, website: http://www.ebi.ac.uk/Tools/msa/clustalw2) software and conserved portions (overlapping the substrate binding sites of proteins) were selected and loaded on Primer3 software (primer3_www.cgi) for primer design. Also in this case, primers were chosen in order to obtain amplification fragments no longer than 100 bp as requested in HRM procedure.

Result

Data analysis and linkage among markers

Progress in computer technology has been essential for developing the DNA marker maps. While the theory behind linkage mapping with DNA markers is identical to map with classical genetic markers, complexity of the problem has increased dramatically [25]. For each segregating marker, a Chi-square (X2) analysis was performed to assess the deviation from the 1:1 expected segregation ratio [26-28]. The first parental (J.K/C1) is a modern cultivar (released in 2000) and the second parental (O5d/O5) is a new (year of release 2001) high-yielding durum wheat variety selected by Dr. Miloudi Nachit (ICARDA) from a late-maturing pure line (Recombinant inbred lines N.101). Whenever possible, the larger is the mapping population much precise are the results.

Microsatellite (SSR) markers analysis

Sufficient DNA polymorphisms between parents must be present. This cannot be overemphasized in the absence of DNA polymorphism segregation analysis and linkage

mapping are impossible. The high level of polymorphism in the microsatellite markers, combined with their high interspersion rate, makes the SSR suitable markers for genetic mapping [17, 29-30]. We analyzed 254 molecular markers to have a total of 115 polymorphic loci which were used to assemble the genetic linkage map including 105 SSR and 10 SNP markers. Identification of the polymorphism and monomorphism was processed by screening all the markers on our two parents. We read and analyzed the peaks came from Genemapper software. The monomorphic markers depicted a single band with a size equal in both parents. The height of peaks could be different (y-axis); however, the size of the amplified fragments was the same (x-axis). In contrast to the polymorphic markers carrying peaks with the same shape had different sizes. The level of monomorphism in the chosen microsatellite markers was 51.4%. As expected, microsatellite markers were more polymorphic (48.6%) than SNP markers (26.3%). This percentage showed polymorphism between first and second parental (Figure 3.a). Looking to the SSR markers, origin come out that polymorphism was more present in WMS (41%) and WMC

(25%) SSR. Other SSR groups: BARC, CFA, CFD, and MST/GPW contained 19%, 8%, 5% and 2% respectively. Looking at the polymorphism distribution among the 7 homeologous groups, group 2 has the majority of polymorphism by 45% (Figure 1). They generated 115 polymorphic fragments, *i.e.* 1.13 fragment per each microsatellite, the vast majority of the microsatellite primer pairs (92%), amplified only one polymorphic locus (polymorphic fragment) in agreement with earlier results (Figure 2 and Table 1). Majority of microsatellite markers are genome-specific and usually amplify only a single locus [31-32]. Nevertheless, other SSRs generated 2 or 3 fragments, some amplifying orthologous loci. For example, WMS291 and WMC201 amplified two fragments, mapped to the two homoeologous sites on the A and B genome. All markers used in present research that had two fragments belonged to this group (homoeologous sites). But in other cases SSRs could amplified in non-homoeologous regions such as gwm554, gwm264, gwm131 and gwm537 [33] or gwm154 produced four polymorphic fragments two mapped in the O5d/O5 population, one on 3BL and one on 7AS [32].
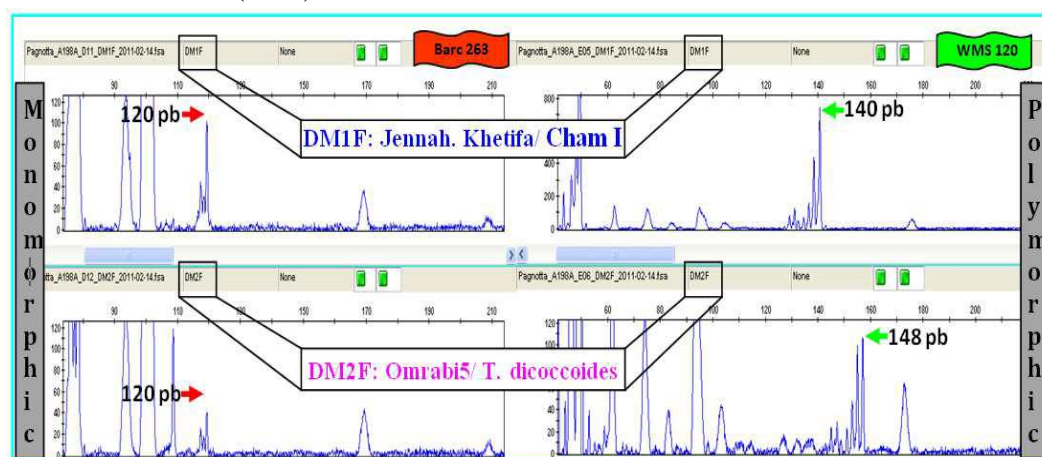
Figure 1. Graphs on the left came from genemapper V.4.0 evidenced by a peak with a size equal for both parents (Barc263). But the graphs on the right showed two different sizes for screening of relatives (WMS120).

Allele size differences in SSR markers

The allele size differences between parents were computed, excluding the M13 tail. To identify the difference in allele, the increase of SSR motifs in one allele was compared with each other. A comparison was made among the individual RILs with respect to the differences in parent allele size across all the mapped loci (Figure 2).
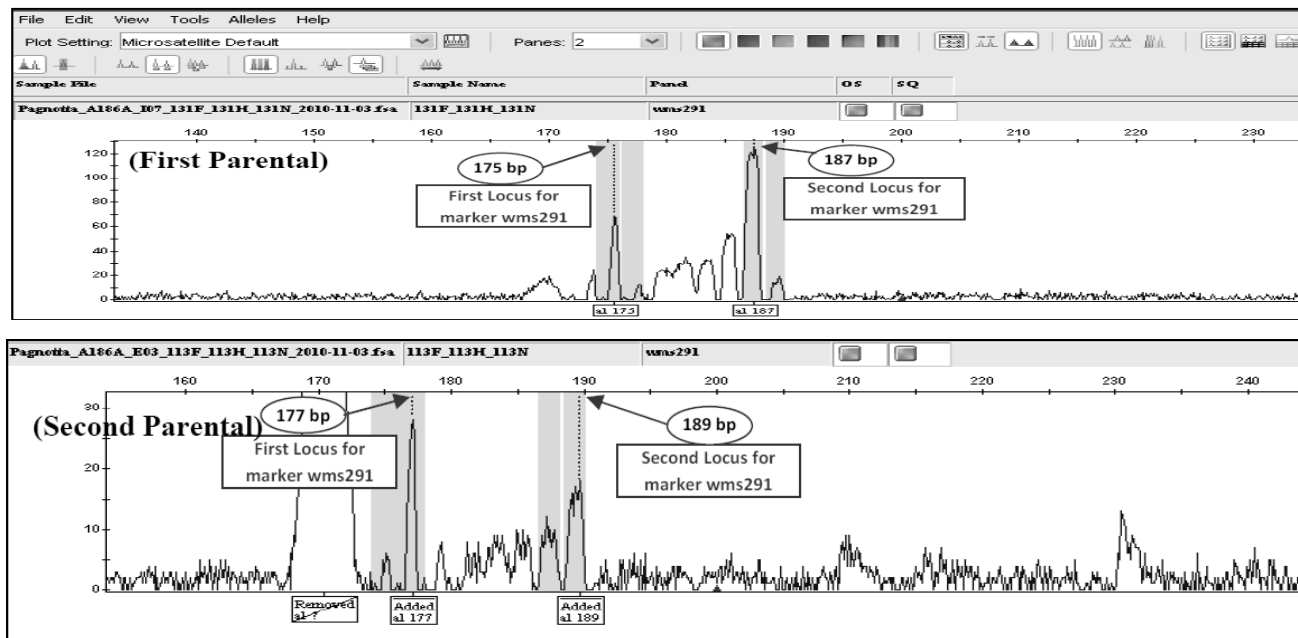

Figure 2. Analyzed WMS291 marker on the RIL n.131 and n.113 with FAM florescent trough of Genemapper V.4.0.

Table 1. Number of generated and mapped SSR fragments and their chromosomal assignment in the Omrabi5/*dicoccoides//*Omrabi5 × Jennah.Khetifa/ChamI population in comparison with wheat published maps

SSRs	Am.Frag	Mapped	Loc.Popu	Kno.Loc
Wms 136	2	1	1AS	1A
Wms 311	3	1	7AL	7A
Wms 95	2	1	2AS	2AS
Wms 249	2	1	2AS	2AS
Wms 113	2	1	4BS	4B
Barc 309	2	2	2AS	2AL
WMS 291	2	1	5AL	5A
Wmc 262	2	2	4AL	4A
WMC 201	2	2	6AS	6A
CFD 88	2	2	4A	4B

Am.Frag: amplified fragments
Loc.Popu: localization in J.K/C × Od/O population
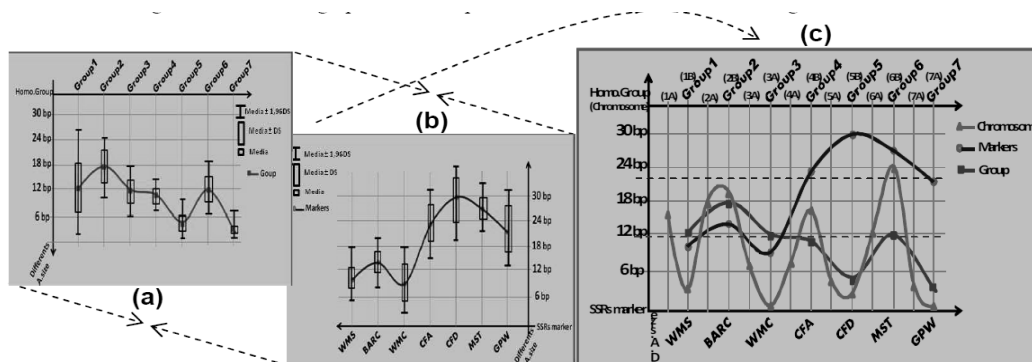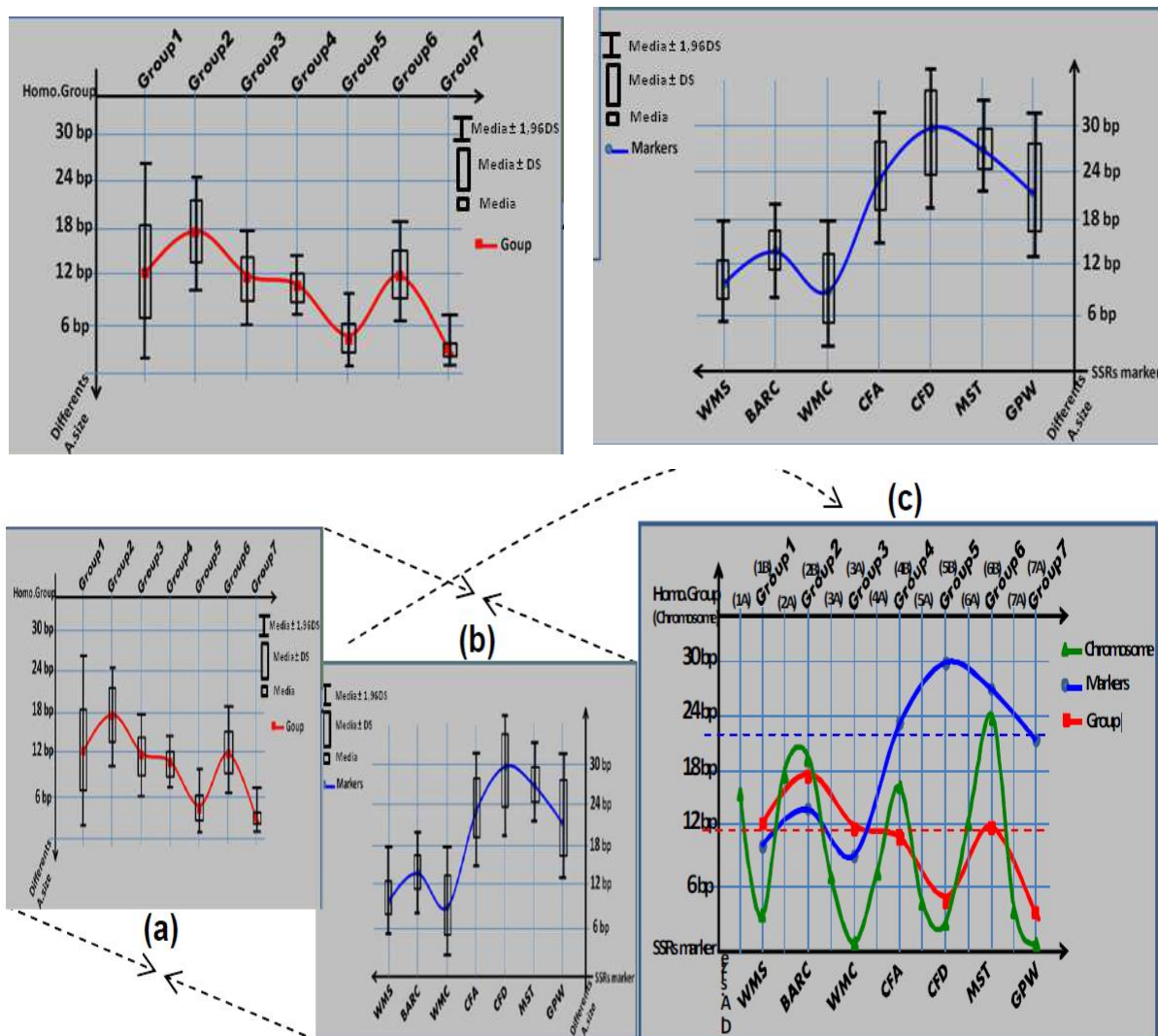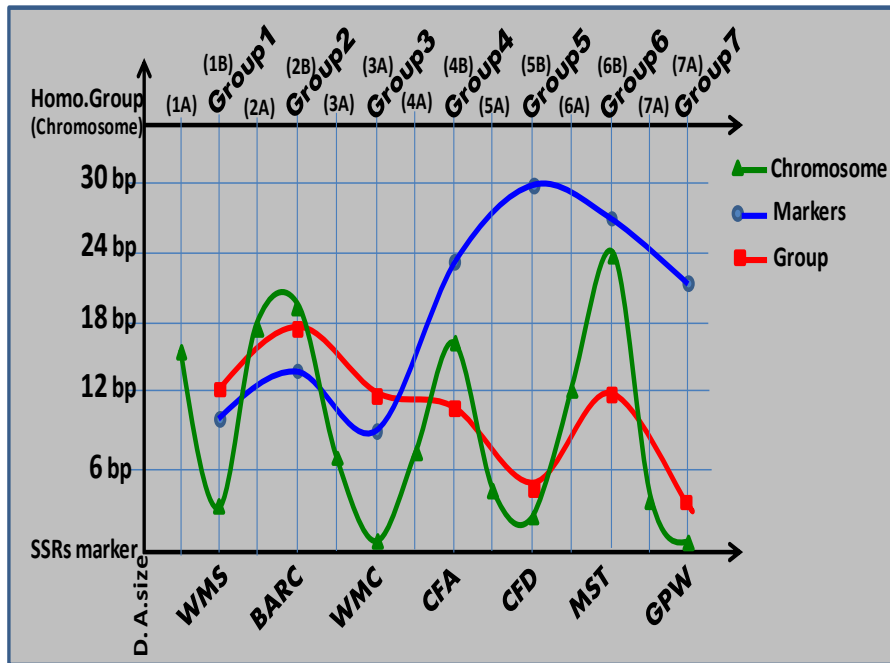Kno.Loc: known localization

Figure 3. The differences size between two alleles of homologous chromosomes

Figure 4. The average of differences between parents in allele size: (a) different homeologous groups, (b) different SSR markers group, (c) different SSR markers group. Homeologous groups and chromosoms brought together in the same graphical to compare their variability with their average.

CFD markers revealed the greatest parent allele size differences (28.2 bp), compared to WMC markers with the lowest parent allele-size difference (9.3 bp). The average parent allele size difference using SSRs markers was 12.4 bp. Allele pairs differing by 9 bp or less between mapping parents, occurred on average at 25% of the mapped loci with a tendency of genetically narrow crosses O5d/O5×J.k/C1 to have a larger fraction of parent allele pairs at 9 bp or less. The MST and CFD markers showed the largest fraction of markers with parent allele pairs of 25 bp and 28 bp. An allele pair of 10 bp or less between parents from different crosses was observed for 69 allele (66%) of the 102 microsatellite loci. The difference size of fragments identified by SSR markers between the two parents were longer in chromosome two and shorter in chromosome five compared to other. This low difference size between the two alleles in chromosome five indicates that alleles are similar (Figures 3 and 4). Difference in size between two alleles amplified through the markers of the WMC group is on average 9 nucleotides (Table 1). Studies for the differences allele size, also brings an understanding on the origin of the mapped markers in our map. The difference size between two alleles amplified through the markers of the WMC group, is an average of 9 nucleotides in the regions or genes that contain this fragments and alleles, showing a little differences (Table 1). Sometimes, mapping of inbreeding species requires that parents be as distantly related as possible, which can often be inferred from geographical, morphological, or isozyme diversity. In some cases, suitable wide crosses may already be available because a frequent goal in plant breeding in the past has been the introduction of desirable characters from wild relatives into cultivars. Comparing the markers and homology groups was observed that the great variability was between marker groups while the homology groups are

similar. Then there was a similarity nucleotide numbers of the differences size in the homology groups compared to markers group.

Distribution of SSR markers

The proportion of BARC, WMC, WMS, CFA and CFD markers amplifying paralogous loci ranged from 13% to 41%. The proportion of GPW/MST markers amplifying paralogous loci was 2% (Figure 5). In the analyzed RIL, the SSR markers BARC, WMC, WMS, CFA, CFD and GPW/MST markers contain 49%, 52%, 56%, 53%, 61%, and 62%, respectively, of alleles as the first parental (J.K/C1). The BARC markers,

compared to other SSR markers, were almost unique marker to be shared with same percentage between the two parents. Therefore these markers contain a maximum diversity and difference, between the first and second parental. A significant segregation distortion was found in 105 (48%) out of 216 markers analyzed on the 192 RILs. Fifty four markers (52.6%) showed distortion in favor of the first parental (J.K/C1) allele whereas 42 (40.5%) showed distortion in favor of the second parental (O5d/O5) allele (at $P \leq 0.05$). The rest of the markers (6) had a rate almost equal between the two parental.
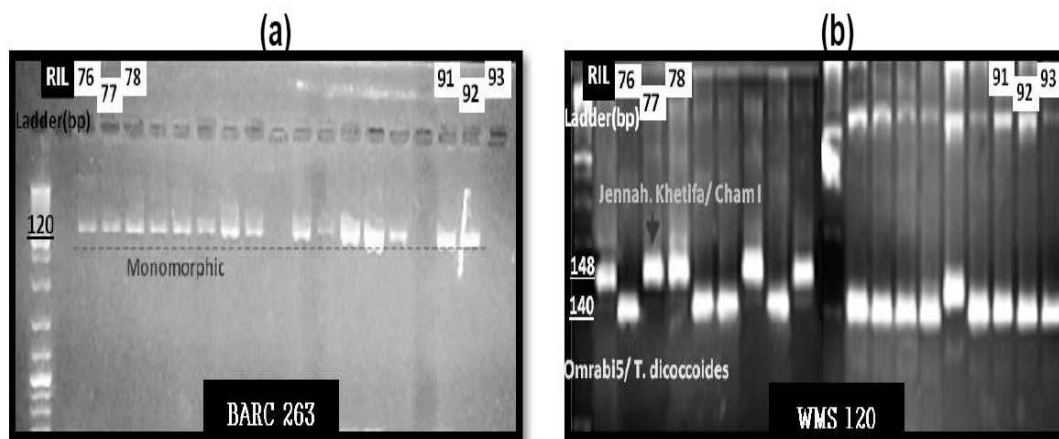
Figure 5. Distribution of the fragments amplified through (a) monomorphic markers, (b) polymorphic markers. For the similarity of parents in the 18 RILs (76-93).

Segregation distortion of SSR markers

Molecular markers representing skewed segregation have already been reported in other Triticeae species [34-37]. Chromosomal rearrangements [38], alleles inducing gametic, parental reproductive differences [39], presence of lethal genes [36], wide genetic background of the parents, and the single seed descend method [37] have been suggested as potential causes of distortion.

Most of the molecular markers, SSR and SNP (57.9%), showed Mendelian

segregation in the 192 RILs. X2 test was used to check whether the marker segregation in F2 fitted the Mendelian model (1:2:1 for codominant and 3:1 for dominant markers) in F7-8 fitted the Mendelian model (1:1 ratio; α=0.01). However in total (42.1%) revealed significant (P<0.05) deviation from the expected ratio which is close to the number one would expect to get by chance in a test including many loci even if no real distortion for all (Figure 6). These markers with skewed segregation occurred in all chromosomes except in 3B, 7A, and

7B. The chromosomes with the most mapped skewed markers are 1A, 2A, 2B, 4A, 6A and 6B. Distorted markers favoring the first parental were found on 1A, 1B, 2A, 2B, 4B and 6B. Those favoring the second parental only on 6A (Figure 6). Additionally, distortions favoring both first and second parental in the same chromosome were also found and therefore could not be segregation to one of the two parents (Figure 6). The 54 markers that showed distortion in favor of the first parental (J.K/C1) allele were distributed among seven chromosomes as follows: 1B (4), 2B (21), 3A (3), 4A (11), 5A (3), 5B (3) and 7B (1). The markers that showed distortion in favor of the second parental (O5d/O5) allele were distributed as follows: 6A (12) and 6B (4), Chromosome 1A showed segregation distortion in favor of the second parental allele in the long arm and in favor of the first parental allele in the short arm. Also chromosome 2B showed opposite patterns in the two chromosome arms: markers on 2AS were in favor of the second parental alleles while markers on 2AL were in favor of the second parental alleles (Figure 7). Notably, all markers mapped on homeologous groups of 2 (chromosomes 2A, 2B), showed skewed segregation in favor of the first allele parental, but the distorted segregation in chromosome 2A a frequency nearby the first Parental, compared distorted segregation showed in chromosome 2B. Out of the 14 chromosomes of genome A and B of tetraploid wheat, only three chromosomes (3B, 4B, and 7B) showed no segregation distortion (Figure 7). Most of our RILs had some similarity with the first parental (J.K/C1). The reason for this majority, could be containing the alleles or fragments, that belonged to the first parental.
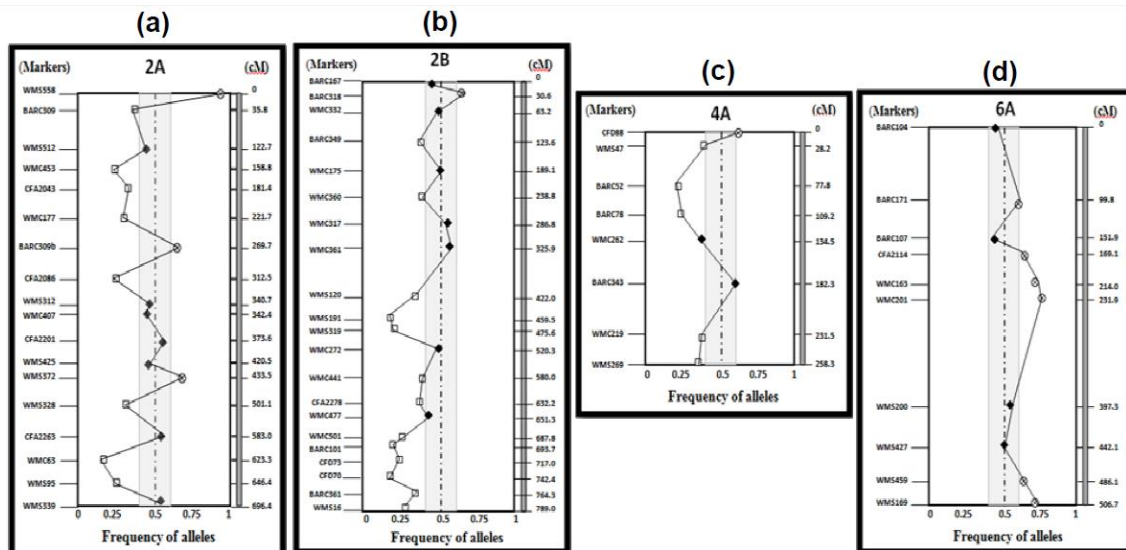
Figure 6. Allele frequencies as a function of the genetic linkage map along chromosomes 2A, 2B, 4A and 6A, The x-axis indicates the segregation distortion from the 1:1 ratio observed for each marker, and the y-axis corresponds to the genetic linkage map, Markers marked with crossed square indicate the significance threshold of $P < 0.05$, (a) chromosome 2A that the majority of markers are distributed to the first parental, (b) chromosome 2B that the majority of markers in the short arm are distributed to the first parental, and long arm to the second parental, (c) chromosome 4A that the majority of markers are distributed to the first parental, (d) chromosome 6A that the majority of markers are distributed to the second parental. Segregation distortion between 0.45- 0.55 (45%-55%) frequency of alleles, Frequency of alleles between first and second parental, Frequency of alleles nearby first parental, Frequency of alleles nearby first parental.
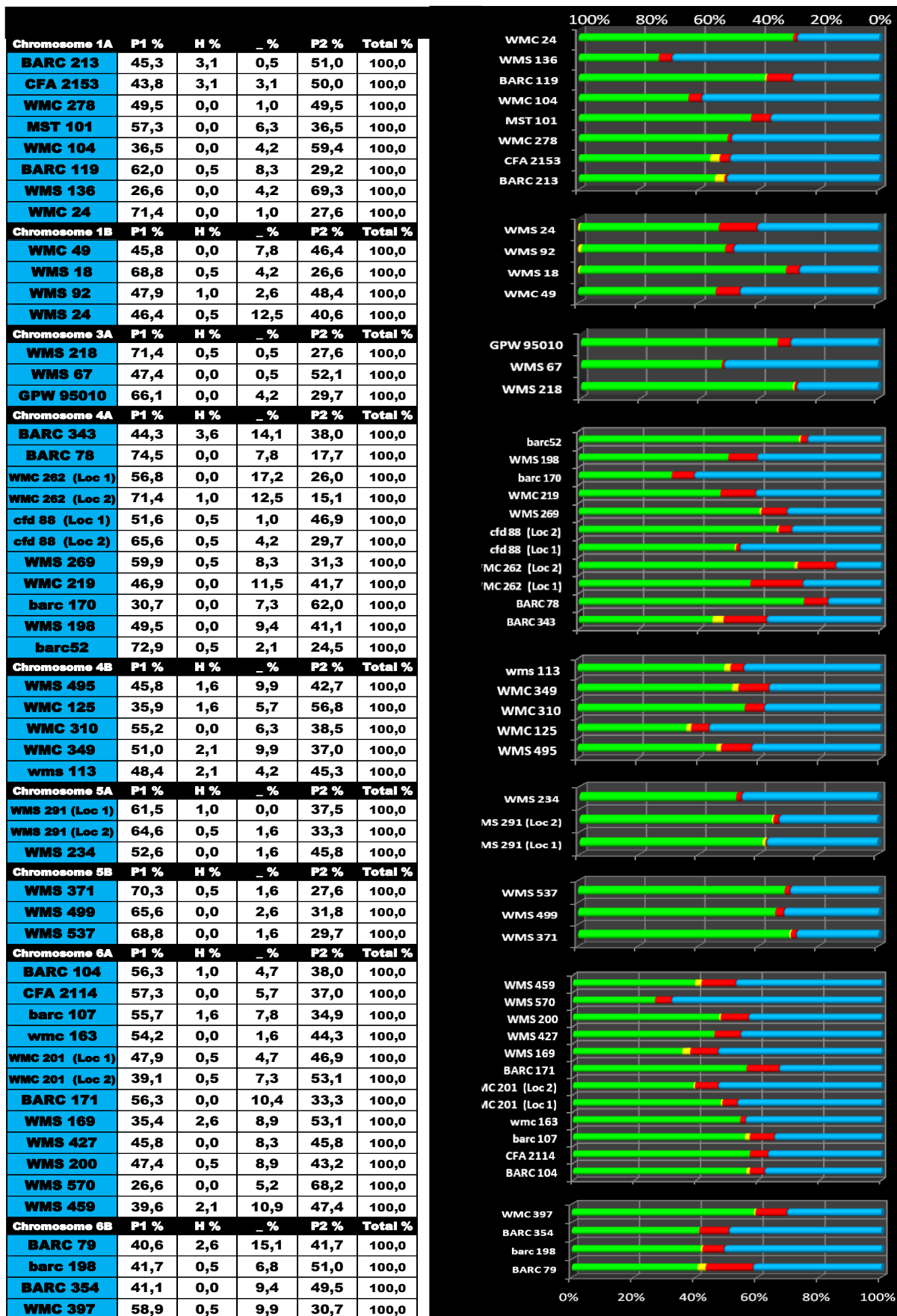
Chromosome 1A	P1 %	H %	_ %	P2 %	Total %
BARC 213	45,3	3,1	0,5	51,0	100,0
CFA 2153	43,8	3,1	3,1	50,0	100,0
WMC 278	49,5	0,0	1,0	49,5	100,0
MST 101	57,3	0,0	6,3	36,5	100,0
WMC 104	36,5	0,0	4,2	59,4	100,0
BARC 119	62,0	0,5	8,3	29,2	100,0
WMS 136	26,6	0,0	4,2	69,3	100,0
WMC 24	71,4	0,0	1,0	27,6	100,0
Chromosome 1B	P1 %	H %	_ %	P2 %	Total %
WMC 49	45,8	0,0	7,8	46,4	100,0
WMS 18	68,8	0,5	4,2	26,6	100,0
WMS 92	47,9	1,0	2,6	48,4	100,0
WMS 24	46,4	0,5	12,5	40,6	100,0
Chromosome 3A	P1 %	H %	_ %	P2 %	Total %
WMS 218	71,4	0,5	0,5	27,6	100,0
WMS 67	47,4	0,0	0,5	52,1	100,0
GPW 95010	66,1	0,0	4,2	29,7	100,0
Chromosome 4A	P1 %	H %	_ %	P2 %	Total %
BARC 343	44,3	3,6	14,1	38,0	100,0
BARC 78	74,5	0,0	7,8	17,7	100,0
WMC 262 (Loc 1)	56,8	0,0	17,2	26,0	100,0
WMC 262 (Loc 2)	71,4	1,0	12,5	15,1	100,0
cfd 88 (Loc 1)	51,6	0,5	1,0	46,9	100,0
cfd 88 (Loc 2)	65,6	0,5	4,2	29,7	100,0
WMS 269	59,9	0,5	8,3	31,3	100,0
WMC 219	46,9	0,0	11,5	41,7	100,0
barc 170	30,7	0,0	7,3	62,0	100,0
WMS 198	49,5	0,0	9,4	41,1	100,0
barc52	72,9	0,5	2,1	24,5	100,0
Chromosome 4B	P1 %	H %	_ %	P2 %	Total %
WMS 495	45,8	1,6	9,9	42,7	100,0
WMC 125	35,9	1,6	5,7	56,8	100,0
WMC 310	55,2	0,0	6,3	38,5	100,0
WMC 349	51,0	2,1	9,9	37,0	100,0
wms 113	48,4	2,1	4,2	45,3	100,0
Chromosome 5A	P1 %	H %	_ %	P2 %	Total %
WMS 291 (Loc 1)	61,5	1,0	0,0	37,5	100,0
WMS 291 (Loc 2)	64,6	0,5	1,6	33,3	100,0
WMS 234	52,6	0,0	1,6	45,8	100,0
Chromosome 5B	P1 %	H %	_ %	P2 %	Total %
WMS 371	70,3	0,5	1,6	27,6	100,0
WMS 499	65,6	0,0	2,6	31,8	100,0
WMS 537	68,8	0,0	1,6	29,7	100,0
Chromosome 6A	P1 %	H %	_ %	P2 %	Total %
BARC 104	56,3	1,0	4,7	38,0	100,0
CFA 2114	57,3	0,0	5,7	37,0	100,0
barc 107	55,7	1,6	7,8	34,9	100,0
wmc 163	54,2	0,0	1,6	44,3	100,0
WMC 201 (Loc 1)	47,9	0,5	4,7	46,9	100,0
WMC 201 (Loc 2)	39,1	0,5	7,3	53,1	100,0
BARC 171	56,3	0,0	10,4	33,3	100,0
WMS 169	35,4	2,6	8,9	53,1	100,0
WMS 427	45,8	0,0	8,3	45,8	100,0
WMS 200	47,4	0,5	8,9	43,2	100,0
WMS 570	26,6	0,0	5,2	68,2	100,0
WMS 459	39,6	2,1	10,9	47,4	100,0
Chromosome 6B	P1 %	H %	_ %	P2 %	Total %
BARC 79	40,6	2,6	15,1	41,7	100,0
barc 198	41,7	0,5	6,8	51,0	100,0
BARC 354	41,1	0,0	9,4	49,5	100,0
WMC 397	58,9	0,5	9,9	30,7	100,0

Figure 7. Distribution of SSRs markers in each of chromosome for similarity to parents.

Clustering of markers (SSR, SNP)

The genetic markers were distributed non-randomly (Px2 (df 114)≤0.0001) along the chromosomes. Clusters of markers were observed on most of the chromosomes of the A and B genomes. As said before, the X2 test was used to check whether the marker segregation in F2 fitted the Mendelian model (1:1 for dominant and codominant markers). The SSR markers used in the current study were selected, according to previously published maps, to cover all 14 chromosomes, whereas SNP markers scored in the mapping population were not targeted to specific genomic regions. Therefore, clustering of markers was tested for the two types of markers separately, revealing significant clusters of SNP markers on most chromosomes. In our map most of markers are located on the centromeric regions, while few markers far apart each other, and were located in the telomerich regions. The cluster of markers in the 11 chromosome from these 14 chromosomes (1A, 1B, 2B, 3A, 4A, 4B, 5A, 5B, 6A, 6B and 7A) was increased the number of markers with the length of the chromosomes up to centromeric regions and then in some chromosomes begins to increase and in some starts to decrease (Figure 8). In wheat and in many other organisms, recombination is unevenly distributed with "hot-spots" and "cold-spots" across chromosomes [40-43].

Clustering around centromeres is a well-known phenomenon with all types of markers, resulting from centromeric suppression of recombination [44-45].

Contrary to other wheat mapping populations [32,46]. In the current study, SSRs showed only a moderate tendency to cluster around centromeres, presumably due to the selection of markers based on their known location. However, a remarkable clustering of SNP markers was found in telomeric regions (Figure 8). Zhang *et al.* (2006) reported for durum wheat that SNP markers showed a stronger tendency than SSR markers in particular to map to gene-rich telomeric regions. High-density physical maps in wheat revealed that more than 85% of wheat genes are present in gene-rich regions, physically spanning only 5-10% of the genome [41-43]. These regions are strongly associated with recombination rate in wheat [42-43, 47] and are predominantly located in telomeres [48]. For example, the clusters of SNP markers in homoeologous group 4 (Figure 8) are associated with a reported gene-rich region near chromosome telomeric ends [42]. The high proportion of clustering SNP markers may, therefore, be indicative of gene-rich regions. If this is indeed the case, SNP markers may be unique for fine mapping of genes/QTLs residing in gene-rich.
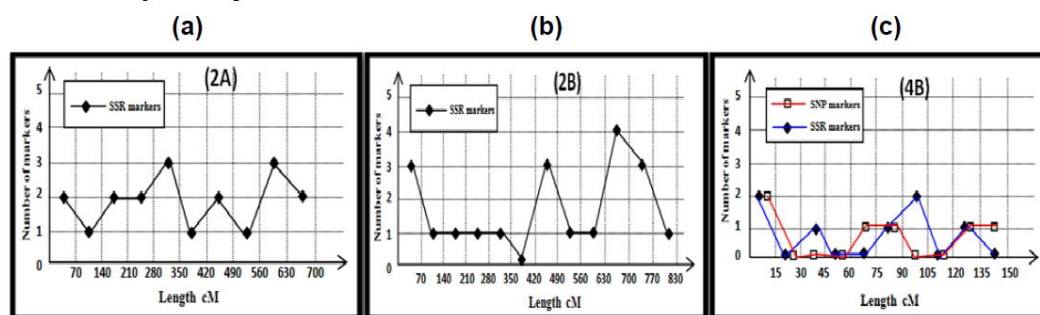
Figure 8. Distribution of SSR and SNP loci along chromosomes 2A, 2B, and 4B of the (J.K/C1) x (O5d/ O5) genetic map.
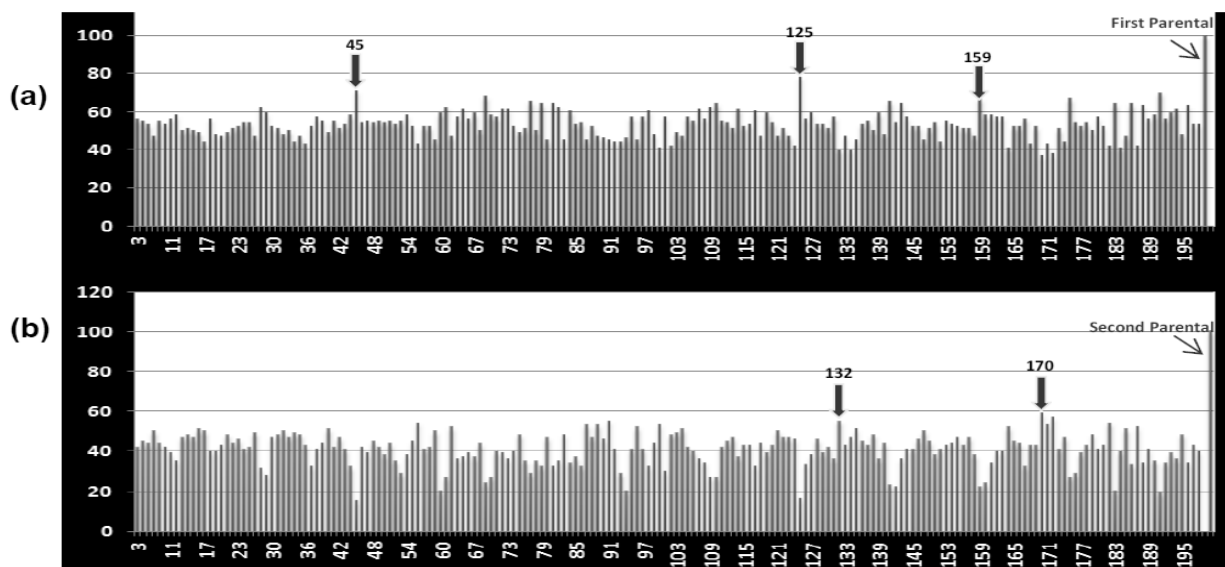
Figure 9. The distribution of the 192 RILs the percentage similarity with (a) First Parental J.K/C1, (b) Second Parental O5d/O5, ( ) with high similarity to the parental, (4, 5, 27, 62, 147 and 150 RILs number do not exist).

Frequency of RILs with parental (non-recombinant) chromosome

Out of the 2,688, RIL × chromosome combinations, 429 RIL × chromosomes (15.9%) remained non-recombinant (parental) and the number remained (2,259 RIL×chromosome) belonged to the recombined chromosomes even after seven cycles of recombination (Figure 9). The lowest frequency of parental chromosomes was observed in group 2 chromosomes (2 and 3 for 2A and 2B, respectively), while the highest frequency of parental chromosomes was observed in group 5 chromosomes (62 and 117 for 5A and 5B, respectively). The recovery rate of non-recombinant chromosomes correlated ($R^2 = 0.815$, $P = 0.001$;) with chromosome length (cM) but was unrelated ($R^2 = 0.43$, $P = 0.31$) to the number of markers per chromosome. The probability of recovering non-recombinant chromosomes was similar for both parental lines J.K/C1 and O5d/O5.

Non-recombinant chromosomes

A relatively higher frequency (15.9%) of non-recombinant chromosomes was observed in this RIL. The highest frequency of non-recombinant chromosomes was found in group 5 chromosomes (41.1%), which is close to the theoretical proportion expected for the estimated length of chromosome 5 (~ 40 cM). While, the highest frequency of RILs with non-recombinant chromosome in T. boeoticum × T. monococcum RIL population, was in chromosome 4A [50]. Chromosome 4A is known to be involved in cyclical interchange with chromosome 5A. Also in the study of [46], the highest frequency of non-recombinant chromosomes was found in group 4 chromosomes (34%), but in our case, only 14.3% of the non-recombinant chromosomes were found in group 4 chromosomes. The parental genotypes of our cross cannot be considered as highly divergent in evolutionary terms. The slightly reduced proportion observed for recombinant chromosomes compared to the theoretical expectation could explain as a result of the known general trend of decrease in recombination in hybrids.

The distribution of SNP markers

In the online SNP database (http://probes.pw.usda.gov:8080/snpworld/Search), 2114 putative SNPs had been found in 274 EST sequences. A total of 38 candidate SNPs from 32 loci were validated using direct sequencing of PCR products, The SNPs were chosen based on a range of redundancy and cosegregation scores and predicted expression of multiple genes. To identify the polymorphism of SNP markers we used HRM, identification of polymorphism is verified through the movement of a peaks in the graphics that determined the melting temperature of SNP markers. Our samples (38 markers) are discriminated on the basis of GC content and, according to their sequence, length or strand complementarily. Even single base changes such as SNPs can be readily identified. Mutations in PCR products are detectable by High Resolution Melting because they change the shape of DNA melting curves. A combination of new-generation DNA dyes, high-end instrumentation and sophisticated analysis software allows to detect these changes and to derive information about the underlying sequence constellation. Out of the 38 candidate SNPs, after sequencing them 10 (26%) were shown to be true polymorphic, and 28 were shown to be false. There were also some fragments that contained 2, 3 or 4 SNPs in the same sequence. The distribution of the 10 SNP markers similarity between parents in all 192 genotypes, have shown that 44.27% were similar to the first parental, and 40.52% to the second, while the remaining 15.22% were heterozigote. From these 10 SNP markers only 6 markers (2 markers from first classand 4 markers from second class) are mapped in chromosome 4B, and the remaining 4

markers, have not been mapped. Chromosome 4B has shown a high variability between the two parents of our population through the verification of SNP markers. In the second class of our SNPs the identification of multiple cosegregating SNPs within an alignment of EST sequences allows the accurate prediction of sequence haplotypes. Comparison of predicted haplotypes with known wheat lines from which the sequences were derived allows the identification of predicted orthologous genes. SNPs are becoming the marker to be chosen for molecular genetic analysis, even if their discovery and characterization are expensive and laborious. The methods of the mining sequence data sets applied to discovery of SNPs, should provide the cheapest source of abundant SNPs [51-54]. Although every effort has been made to produce and submit sequence the highest quality, the high-throughput nature of the sequencing programs inevitably leads to the submission of inaccuracies. The electronic filtering of these data to identify potentially biologically relevant polymorphisms is thereby hampered by the false calling of these bases. Previous methods used to identify SNPs in aligned sequence data has relied on the comparison of sequence trace files to filter out polymorphisms, where the base calling within one or more of the traces is of dubious quality and, therefore, likely to be due to sequence error rather than representative of a true polymorphism [55-57]. This method, although suitable for comparing genomic sequence, is limited by the requirement of sequence trace file data and does not distinguish errors incorporated during the reverse transcription of mRNA, For highly redundant data sets compiled from a variety of sources with a limited availability of sequence trace files, this

means of filtering sequence errors from true polymorphisms is not feasible. However, the redundant nature of these EST data sets does permit the selection of polymorphisms that occur multiple times within a set of aligned sequences. The frequency of occurrence of a polymorphism at a particular locus provides a measure of confidence in the SNP representing a true polymorphism and is referred to as the SNP redundancy score. By examining SNPs that have a redundancy score of two or greater, i.e. two or more of the aligned sequences represent the polymorphism, the vast majority of sequencing errors are removed. Although some true genetic variation is also ignored due to its presence only once within an alignment, the high degree of redundancy within the data permits the rapid identification of large numbers of SNPs from data collated from a variety of sources.

Discussion

We presented a number of molecular markers SSR, which may provide some clear information on the sequences of chromosomes. The linkage traits in the study was [58] 80% of the primer pairs successfully amplified products. The linkage rate was higher than that reported by Roder *et al.* [32] from wheat genomic DNA (30%) and that by Stack *et al.* [59] from wheat ESTs (50%). Most of the primer pairs produced clear and strong amplification products. Under similar PCR conditions (same thermocycler and Taq polymerase and buffer), there were no obvious differences between SSR markers designed from genomic sequences and ESTs. However, it appears that, EST-SSRs show fewer alleles compared with that of the SSRs designed from genomic DNAs. Functional constraints on ORFs may account for the lower percentage of

polymorphism (19.2%) between the parents of map population as compared to the polymorphism rate of 33% reported by Gupta et al. [18] using SSR designed from genomic DNA. On the other hand, we were able to map 65 markers (73.9% of all mapped primer sets) with putative functions to 20 chromosomes. Such results will be valuable for targeted traits selection in crop breeding. For instance, EST-SSRs associated with gliadin or glutenin will be helpful for evaluating bread-making quality, whereas markers related to stress responsive genes may facilitate selection for tolerance against biotic and abiotic stresses. Optimum conditions for PCR. First, we tested reproducibility of SSR amplification using the DNA samples of relatives as templates. 12 markers were randomly selected from Barc, Cfa, Cfd, Gdm, Gwm, and Wmc markers and amplified according to the conditions given by the original literature. However, in our hands, the number and size of amplified fragments were quite different from those in the literature.

Transferability of markers

The several SSR markers used in the present study were developed from hexaploid wheat (WMS and SWM), A genome of T. *urartu* (CFA), and the D genome of Ae. *Tauschii* (CFD, GDM). These markers helps us to know about our population the percentage of genomes transfer. T. urartu genome, that contributed A genome of hexaploid wheat [40], shows considerable differences from the A genome of T. *monococcum* or T. *boeoticum* [60]. About 65–70% of SSRs markers developed from hexaploid wheat and diploid Ae. *Tauschi*, showed amplification in T. durum from Jennah Ketifa/ChamI and T. durum Omrabi5/ dicoccoides. A high level of conservation exists between A and B

genomes of tetraploid wheat. As many as 46.3% of the D genome, 32% of the B genome specific, 8% of the A genome specific SSR markers showed amplification in our population. In agreement with the work of Bai *et al.* [12] the SSR markers isolated from hexaploid wheat are transferable for about 67%, in T. *monococcum*. Only about 50% of markers is transferable from the A genome of hexaploid wheat to tetraploid, A and B genome. The majority of the markers that did not show transferability in the study of Bai *et al.* [12], did not show transferability in the present study as well, with few exceptions like GWM136, GWM614 and GWM154 that did show amplification either in T. *monococcum*. On the other hand, Sourdille *et al.* [61] reported transferability of 93% of the SSR markers derived from hexaploid wheat on the corresponding ancestral A genome of diploid species. However, they used only 12 A genome specific primers, which is much less a number than used in the present study. Guyomarc'h et al. [21] reported about 50% the transferability of Ae. *tauschii* derived SSR markers in the A genome of diploid species. Genomic relationship among A, B and D genomes of hexaploid wheat has been established by studying chromosome pairing in hexaploids having the recessive mutant allele ph1b and nullihaploids that lack chromosome 5B [62].

Conclusion

In this research study, we obtained 260 amplification profiles. As we set the PCR conditions stringent (high annealing temperature) to avoid amplification of ambiguous bands, the SSR amplification tends to be genome-specific. The marker Gwm136 showed a highly polymorphic amplification pattern and Gwm148 revealed less polymorphic pattern. For basic genetic studies, development of primary genetic linkage maps is of paramount importance. Furthermore, saturated genetic maps provided geneticists and breeders with powerful tools for quantitative trait mapping, positional gene cloning, and marker-assisted selection. Due to the large genetic distance between the parents of this population (first, J.K/CI) and second (O5d/O5, parental), a high level of polymorphism was detected (>45%), the mapped population was a RIL. The (J.K/CI)×(O5d/O5) map was constructed with 100 microsatellites and 10 SNPs as framework markers. The length of the map was 3170.293 cm, with an average of marker per 31.70 cm. The map depicted a high synteny with former published durum and bread wheat maps. The SSR markers were evenly mapped across the whole genome. The microsatellites proved to be good as anchor probes and were very useful in arms and chromosomes identification. SNPs proved also to be good markers for map saturation in one of chromosome 4B. Most of the combinations were mapped in different genomic regions, which amplified numerous fragments that were mapped in 13 different linkage groups. More genetic markers were mapped in the A genome comparatively to the B genome. In conclusion, Wms and Wmc from Bread and durum wheat microsatellites amplified and mapped perfectly in durum/durum genetic background. Indeed, the mapping population showed a fixed genetic structure, as it was developed using single seed descent (SSD) method. This makes Jennah Ketifa/ChamI × Omrabi5/*T. dicoccoides600545*//Omrabi5 population ideal for assessing the environmental impact on trait

expression. In the future this population will be used for identification of QTLs linked to agronomic, physiologic, and biotic traits.

References

[1] A.B. Damania, M. Tahir, B.H. Somaroo, *ICARDA*., **1992**.

[2] M. Nesbitt, D. Samuel, *IPGRI, Rome*, **1995**.

[3] G. Laghetti, A.R. Piergiovanni, N. Volpe, D. Semeraro, M. Falivene, M. Basile, *Info Agrario*., **1997**. *40, 105–108*.

[4] P. Perrino, S. Infantino, P. Basso, A. Di Marzo, N. Volpe, G. Laghetti, *Inform Agrario*., **1993**, *41–45*.

[5]G. Laghetti, A.R. Piergiovanni, N. Volpe, P. Perrino, *Agr Med*., **1999**, *129, 199–211*.

[6] R. Cubadda, E. Marconi, *IPGRI, Rome*, **1995**, *203–221*.

[7] G. Galletti, P. Bocchini, L.F. D'Antuono, *J Agric food chem*., **1996**, *44, 31–33*.

[8] G. Galterio, M. Cappelloni, E. Desiderio, N.E. Pogna, *J Genet Breed*., **1994**, *48, 391–398*.

[9] L. Corazza, M. Pasquini, P. Perrino, *Genet Agrar.,* **1986**, *40, 243–254*.

[10] M.C. Luo Z.L. Yang, J. Dvorak, *Theor Appl Genet*., **2000**, *100, 602–606*.

[11] J. MacKey, *Hereditas*., **1954**, *40, 65–180*.

[12] J. Bai, K. Liu, X. Jia, D. Wang, *Plant Sci*., **2004**, *166, 341-347*.

[13] A. Blanco, M.P. Bellomo, A. Cenci, C. De Giovanni, R. D'Ovidio, E. Iacono, B. Laddomada, M.A.

[14] D.J. Somers, P. Isaac, K. Edwards, *Theor Appl Genet*., **2004**, *109, 1105-1114*.

[15] G.J. Bryan, A.J. Collins, P. Stephenson, A. Orry, J.B. Smith, M.D. Gale. *Theoret App Gen*., **1997**, *94, 557-563*.

[16] Pagnotta, E. Porceddu, A. Sciancalepore, R. Simeone, O.A. Tanzarella, *TAG*., **1998**, *97, 721-728*.

[17] J. Plaschke, M.W. Ganal, M.S. Röder, *Theor Appl Genet*., **1995**, *92, 1078–1084*.

[18] P.K. Gupta, H.S. Balyan, K.J. Edwards, P. Isaac, V. Korzun, M. Röder, M.F. Gautier, P. Joudrier, A.R. Schlatter, J. Dubcovsky, R.C. De la Pena, M. Khairallah, G. Penner, M.J. Hayden, P. Sharp, B. Keller, R.C.C. Wang, J.P. Hardouin, P. Jack, P. Leroy, *Theor Appl Genet*., **2002**, *105, 413-422*.

[19] W. Karl, T. Broman, *Johns Hopkins University, Baltimore, Maryland 21205, the Genetics Society of America*., **2005**.

[20] M.W. Ganal, M.S. Röder, *Genom applic in crops. Springer, Berlin.* **2007**.

[21] H. Guyomarc'h, P. Sourdille, K.J. Edwards, M. Bernard, *Theor Appl Genet*., **2002**, *105, 736-744*.

[22] P. Sourdille, S. Singh, T. Cadalen, G. Brown-Guedira, G. Gay, L. Qi, B. Gill, P. Dufour, A. Murigneux, M. Bernard, *Funct Integr Genom*., **2004**, *4, 12-25*.

[23] Q.J. Song, J.R. Shi, S. Singh, E.W. Fickus, J.M. Costa, J. Lewis, B.S. Gill, R. Ward, P.B. Cregan, *Theor Appl Genet*., **2005**, *110, 550-560*.

[24] M.S. Röder, V. Korzun, K. Wendehake, B.S. Gill, M.W. Ganal, *Genome*., **1998a**, *41, 278-283*.

[25] N.D. Young, Phillips R.L, Vasil J.K. (eds.)., **2000**, *32, 3445-3457*.

[26] D.I. Mester, Y.I. Ronin, Y. Hu, J. Peng, E. Nevo, A.B. Korol, *Theor Appl Genet*., **2003a**, *107, 1002-1112*.

[27] D.I. Mester, Y.I. Ronin, D. Minkov, E. Nevo, A.B. Korol, *Genet*., **2003b**, *165, 2269-2282*.

[28] D.I. Mester, Y.I. Ronin, E. Nevo, A.B. Korol, *Comput Biol Chem*., **2004**, *28, 281-290*.

[29] Z.Q. Ma, M. Röder, M.E. Sorrells, *Genome*., **1996**, *39, 123-130*.

[30] M.S. Röder, J. Plaschke, S.U. Konig, A. Borner, M.E. Sorrells, *Mol. Gen. Genet*., **1995**, *246, 327-333*.

[31] V. Korzun, M.S. Röder, K. Wendekake, A. Pasqualone, C. Lotti, M.W. Ganal, P.Y. Kwok, C. Carlson, T.D. Yager, W. Ankener, D.A. Nickerson, *Genom*., **1999**. *23, 138-144.*

[32] M.S. Röder, V. Korzun, K. Wendehake, J. Plaschke, M.H. Tixier, P. Leroy, M.W. Ganal, *Genetics*., **1998b**, *149, 2007-2023.*

[33] I. Elouafi, M.M. Nachit, *Theor Appl Genet*., **2004**, *108, 401-413*.

[34] M. Heun, A.E. Kennedy, J.A. Anderson, N.L.V. Lapitan, M.E. Sorrells, S.D Tanksley, *Genome*., **1991**, *34, 437-447.*

[35] Y.G. Liu, K. Tsunewaki, *Jap J Genet*., **1991**, *66, 617-633*.

[36] A. Blanco, *TAG*., **1999**, *98, 1202-1207*.

[37] M.M. Nachit, I. El ouafi, M.A. Pagnotta, A. El saleh, E. lacono, M. Labhilili, A.P. Asbati, M. Azrak, H. Hazzam, D. Benscher, M. Khairallah, J.M. Ribaut, O.A. Tanzarella, E. Porceddu, M.E. Sorrels, *TAG*., **2001**, *102, 177-186*.

[38] S.D. Tanksley, *Chrom*., **1984**, *89, 352-360*.

[39] M.R. Foolad, S. Arulsekar, V. Becerra, F.A. Bliss, *TAG*., **1995**, *91, 262-269*.

[40] J. Dvorak, P.E. Mc Guire, B. Cassidy, *Genome*., **1988**, *30, 680-689.*

[41] J.D. Faris, K.M. Haen, B.S. Gill, *Genetics*., **2000**, *154, 823-835*.

[42] K.S. Gill, B.S. Gill, T.R. Endo, E.V. Boyko, *Gen*., **1996a**, *143, 1001-1012*.

[43] K.S. Gill, B.S. Gill, T.R. Endo, T. Taylor, *Gen*., **1996b**, *144, 1883-1891*.

[44] A.B Korol, I.A. Preygel, S.I. Preygel, *Chapman and Hall, London, UK*, **1994**.

[45] S.D. Tanksley, M.W. Ganal, J.P. Prince, M.C. De Vicente, M.W. Bonierbale, P. Broun, T.M. Fulton, J.J. Giovannoni, S. Grandillo, G.B. Martin, R. Messeguer, J.C. Miller, L. Miller, A.H. Paterson, O. Pineda, M.S. Röder, R.A. Wing, W. Wu, N.D. Young, *Genet Soc Americ*., **1992**, *132, 1141-1160*.

[46] J.H. Peng, T. Fahima, M.S. Röder, Y.C. Li, A. Grama, Y.I. Ronin, A.B. Korol, E. Nevo, *Genome Res*., **2000**, *10, 1059-1031*.

[47] Y. Weng, M.D. Lazar, *Theor Appl Genet*., **2002**, *104, 1078-1085*.

[48] L.L Qi, B. Echalier, S. Chao, G.R. Lazo, G.E. Butler, O. Anderson, O.D. Akhunov, E.D. Dvolák, J. Linkiewicz, A.M. Ratnasiri, A. Dubcovsky, J. Bermudez-Kandianis, C.E. Greene, R.A. Kantety, R. La Rota, M. Munkvold, J.D. Sorrells, S.F. Sorrells, M.E. Dilbirligi, M. Sidhu, D. Erayman, M. Randhawa, H.S. Sandhu, D. Bondareva, S. Gill, K.S. Mahmoud, A.A. Ma, X.F. Miftahudin, J.P. Conley, E.J. Nduati, V. Gonzalez-Hernandez, J.L. Anderson, J.A. Peng, J.H. Lapitan, N.L.V. Hossain, K.G. Kalavacharla, V. Kianian, S.F. Pathan, M.S. Zhang, D. Nguyen, H.T. Choi, D.W. Close, T.J. McGuire, P.E. Qualset, B.S. Gill, *Genet*., **2004**. *168, 701-712*.

[49] G. Wricke, W.E. Weber, *Walter de Gruyter*, **1986**, *Berlin. New York*.

[50] Q.S. Zhang, L. Manche, R.M. Xu, A.R. Krainer, *RNA*., **2006**, *12, 1116-1128*.

[51] P. Taillon-Miller, Z.J. Gu, Q. Li, L. Hillier, P.Y. Kwok, *Genome Res*., **1998**, *8, 748-754*.

[52] L. Picoult-Newberg, T.E. Ideker, M.G. Pohl, S.L. Taylor, M.A. Donaldson, D.A. Nickerson, M. Boyce-Jacino, *Genome Res*., **1999**, *9,167- 174*.

[53] Z. Gu, L. Hillier, P.Y. Kwok, *Hum Mutat*., **1998**, *12, 221–225*.

[54] K.H. Buetow, M.N. Edmonson, A.B. Cassidy, *Nat Genet*., **1999**, *21, 323-325*.

[55] P.Y. Kwok, C. Carlson, T.D. Yager, W. Ankener, D.A. Nickerson, *Genom.*, **1994**, *23, 138-144*.

[56] K. Garg, P. Green, D.A. Nickerson, *Gen Res*., **1999**, *9, 1087-1092*.

[57] G.T. Marth, I. Korf, M.D. Yandell, R.T. Yeh, Z.J. Gu, H. Zakeri, N.O. Stitziel, L. Hillier, P.Y. Kwok, W.R. Gish, *Nat Genet*., **1999**. *23, 452-456*.

[58] L.F. Gao, R.L. Jing, N.X. Huo, Y. Li, X.P. Li, H.R. Zhou, X.P. Chang, J.F. Tang, Z.Y. Ma, J.Z. Jia, *Theor App Gen*., **2004**, *108, 1392-1400*.

[59] C.M. Stack, S.G. Easwaramoorthy, U.K. Metha, M.J. Downes, C.T. Griffin, A.M. Burnell, *Nemat*., **2000**, *2, 477-487*.

[60] B.L. Johnson, H.S. Dhaliwal, *Am J Bot*., **1976**, *63, 1088-1094*.

[61] P. Sourdille, M. Tavaud, G. Charmet, M. Bernard, *Theor Appl Genet*., **2001**, *103, 346-352*.

[62] P.P. Jauhar, O. Riera-Lizarazu, W.G. Dewey, B.S. Gill, C.E. Crane, J.H. Bennett, *Theor Appl Genet*., **1991**, *82, 441-449*.