academicJournals

Vol. 15(46), pp. 2603-2612, 16 November, 2016
DOI: 10.5897/AJB2015.15118
Article Number: 03BDE3A61673
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Linkage disequilibrium and association mapping of drought tolerance in cotton (*Gossypium hirsutum* L.) germplasm population from diverse regions of Pakistan

Abdelhafiz Adam Dahab[1*], Muhammad Saeed[2], Nada Babiker Hamza[3], Bahaeldeen Babiker Mohamed[1] and Tayyab Husnain[4]

[1]Environmental and Natural Resources Research Institute (ENRRI), National Centre for Research (NCR), Khartoum, Sudan.
[2]Department of Botany, Government College University, Faisalabad, Pakistan.
[3]Commission for Biotechnology and Genetic Engineering, National Centre for Research, Khartoum, Sudan.
[4]Centre of Excellence in Molecular Biology (CEMB), University of the Punjab, Lahore, Pakistan.

Received 22 November, 2015; Accepted 1 February, 2016

Drought stress is a major abiotic stress that limits crop production. Molecular association mapping techniques through linkage disequilibrium (LD) can be effectively used to tag genomic regions involved in drought stress tolerance. With the association mapping approach, 90 genotypes of cotton *Gossypium hirsutum,* from diverse regions of Pakistan were used. The morpho-physiological traits of all genotypes were evaluated in greenhouse under well-watered and drought stress conditions. Mean squares from analysis of variance for all morpho-physiological traits revealed highly significant variations (P≤0.05) between water levels and genotypes. Cotton varieties were screened for polymorphism with 180 simple sequence repeat (SSR) markers. Out of these 180 SSR markers, 95 were polymorphic. Genotyping of the selected 95 SSR primer pairs generated 57.5% polymorphism, and the number of polymorphic alleles per primer was 2.10. Population structure, linkage disequilibrium, and association mapping between pairs of SSR marker loci were studied. The significance of pairwise LD (P≤0.005) among all possible SSR loci was evaluated at significant threshold values (R^2≥0.05); 7.1% of the SSR marker pairs showed significant pairwise LD in 90 accessions of *G. hirsutum.* Also we observed a significant (R^2) LD between 13 pairs of SSR loci; each pair within the same chromosome in a range of 180 cM between NAU1230 and NAU3095 loci in chromosome (D5) and 1.612 cM between NAU462 and NAU3414 in chromosome A9. This indicates tight linkage between two alleles on the same chromosome. Markers, NAU3414, NAU2691, NAU1141 and NAU1190 were associated with more than single traits under drought treatments. Highest phenotypic variance explaining (R^2) was ascribed to NAU3011 chromosome D13 significantly (p≤0.001) associated with root length under drought treatment.

Key words: Cotton *Gossypium hirsutum*, drought stress, association mapping, linkage disequilibrium (LD), simple sequence repeats (SSRs).

INTRODUCTION

Global climate change poses serious problems for sustained crop production. Due to the continuous water deficit for agricultural production, development of drought tolerant crop to meet the food and fibre demand has

become a necessity (Saeed et al., 2012). Increasing aridity of semi-arid regions and limited water resources have led to a crucial necessity for improving crop drought resistance (Passioura, 2007). There is extensive research on genetic and breeding programs of cotton and it has a long history of improvement through conventional breeding and selection with frequent long-term yield achievements. The identification of drought-related QTLs plays an essential role in crop improvement through marker assisted selection. DNA marker studies have laid basis for revealing the molecular basis for the traits related to drought tolerance (YongSheng et al., 2009). Among the variety of genetic markers, SSR markers have shown high potential to detect polymorphism (Dongre et al., 2011; Dahab et al., 2013) and have been used extensively for cotton genome mapping and marker assisted selection (He et al., 2007). Researchers have mapped QTLs for morphological traits (Liang et al., 2014), physiological traits (Saeed et al., 2011), earliness (Li et al., 2013), yield (LiFang et al., 2010) and fibre traits (Islam et al., 2014). Abdelraheem et al. (2015) assessed a backcross inbred line (BILs) population derived from a cross between Pima cotton, 'Pima S-7' (*Gossypium barbadense*) and Upland cotton 'Sure- Grow 747' (*Gossypium hirsutum*) for their drought tolerance morphological traits using PEG induced osmotic stress. A total of six QTLs were detected for plant height (PH), fresh shoot and root weight (RW), explaining 10.9 to 19.2% of the phenotypic variation (PV). Rodriguez-Uribe et al. (2014), studying drought tolerance of cotton, stated that, of a total of 110 drought-responsive genes identified through a microarray analysis, there was 79% expression. Saeed et al. (2011) evaluated F2, F2:3 and F3:4 populations derived from an intraspecific cross between two *G. hirsutum* lines for morphological, physiological and yield traits. Seven QTLs were detected for osmotic potential, osmotic adjustment, plant height, and seed cotton yield. Association mapping, based on linkage disequilibrium (LD), is a new methodology which examines thousands of polymorphisms for assessing QTL effect. It is more effective compared to linkage analysis since it does not require generation of segregating populations/large numbers of progeny. Association mapping has three advantages: increased mapping resolution, reduced research time, and greater allele numbers. It is a powerful technique used to identify genomic regions linked to specific variants of a phenotypic trait (Saeed et al., 2014). Genome-wide association in plants has wide range of use and there are many reports of association studies on many crops such as barley (Gutiérrez et al., 2011; Cockram et al., 2010), rice (Shao et al., 2011), bread wheat (Yu et al., 2012), maize (Poland et al., 2011), triticale (Niedziela et al., 2012), bean (Shi et al., 2011), sugar beet (Würschum et al., 2011), cotton fiber quality traits (Abdurakhmonov et al., 2009) and cotton salinity stress tolerance (Saeed et al., 2014). Cotton *Gossypium* spp. is widely used for natural fiber in the textile industry. The worldwide commercial effect of cotton production is ~$500 billion per annum with consumption of ~115-million bales or ~27-million metric tons (MT) of fibre (Chen et al., 2007). Cotton is one of the main warm-season cultivars, grown during summer in arid and semiarid regions where water is limited (Singh et al., 2007). The adaptive traits of plants to unfavourable environmental conditions include numbers of physiological, morphological and biochemical features of whole plant (Saeed et al., 2012). Genes involved in molecular mechanisms can be tagged with the help of molecular mapping approaches. In our present study, we assessed extent of LD in the *G. hirsutum* germplasm from diverse regions of Pakistan. The aims of this study were to assess the population structure, linkage disequilibrium (LD), and association of molecular markers with drought stress tolerance of cotton (*G. hirsutum* L.) in a collection of 90 elite cotton germplasm accessions.

MATERIALS AND METHODS

Plant

The plant material consisted of 90 genotypes of cotton *G. hirsutum* (Table 1) collected from diverse regions of Pakistan. The cotton varieties were grown in green house during March-April 2010. Four seeds of each genotype were sown after soaking overnight at field capacity of 2-3 cm^3 depth in 8 polythene bags of 25×5 cm^2 , filled with ~ 1.15 kg of compost soil (peat, sand, soil, 1:1:1). They were arranged in a randomized complete block design with three replications and two treatments. One set consists of 4 bags kept as a control (W1) and the other as water stressed (W2). After germination, only one plant/bag was kept for data recording. Standard pH (6.5), temperature (25 ± 2°C), humidity (50%) and light requirements (13 h photoperiod) for cotton growth were maintained throughout the total duration of experiment. Seedlings grown under both stress and non-stress conditions were irrigated and fertilized till the development of the first true leaf, and thereafter, seedlings grown under control condition (W_1) were watered daily to keep the soil at field capacity. Water stress condition was developed by withholding water supply to the seedlings grown under water stress condition (W_2), and the effect of water stress was monitored visually and with soil moisture meter (HH2 Theta Probe Type, Delta-T device, Cambridge, England). At initial wilting stage (observed visually), when soil had 14 to 16% soil moisture content, the stressed plants were watered to relieve the sign of wilting but not enough to reach field capacity. The experiment was continued for 45 days from the date of emergence till the 3rd main stem leaf was fully expanded. The plants grown under normal water supply and stressed conditions were measured for morphological and physiological parameters.

*Corresponding author. E-mail: dahababdelhafiz@yahoo.com.

Table 1. List of cultivars used in the study.

S/N	Cultivar name	S/N	Cultivar name	S/N	Cultivar name
1	CIM-496	31	MNH-806	61	XU-ZHOU-142
2	CIM-482	32	MNH-802	62	XING TAI-68-71
3	CIM-473	33	MNH-770	63	XIAO-VEMIAN
4	CIM-448	34	MNH-720	64	VH-148
5	CIM-446	35	MNH-638	65	UA-73
6	CIM-4/99	36	MNH-636	66	UA-7-25/46
7	CIM-1100	37	MNH-6070	67	UA-31-102
8	CIM-109	38	MNH-554	68	U-4(5143)
9	CHINA	39	MNH-552	69	U-4
10	BT-3701	40	MNH-147	70	U-276
11	BT-2009	41	MG-66	71	TH-35/99
12	BT-1573	42	M-944-0243	72	Stoneville-825
13	BS-1	43	M-944-00-0030	73	Stoneville-213
14	BH-118	44	L-S-S	74	Stonevilla
15	B-557	45	LA-208	75	STAMP-82
16	AYT-85094	46	IS-7-F1	76	SLS-1
17	AYT-85094	47	IRMA-1480	77	SLH-284
18	Australia-407721	48	GR-156-	78	S-14
19	4 F	49	FH-901	79	S-12
20	362-F	50	FH-900	80	S-11
21	299F	51	FH-682	81	Rehmani
22	268-F	52	FH-125	82	N-KARSHOMG
23	268 F	53	FH-113	83	NIBGE-4
24	1027 ALF	54	FH-1000	84	NIBGE-4
25	1021(KIVI)	55	841/52	85	NIAB--78
26	65090	56	CRSM-38	86	MS-93
27	6040	57	CRSM-38	87	MS-40
28	3996	58	CRS-2009	88	MNH-814
29	2616	59	CIM-534	89	MNH-812
30	814	60	CIM-506	90	MNH-807

Phenotyping of plant materials

On 26th April, 2010 green-house experiment was completed and the following parameters were measured. First, plant length (PH) and plant fresh weight (PFW) were recorded after the plants were separated into shoot and root parts, and data were recorded for shoot length (SL), root length (RL), fresh shoot weight (FSW) and fresh root weight (FRW). The respective shoots and roots of all plants were then oven-dried at 70°C till a constant dry weight was reached. The dry weight of shoot and root of respective plants were recorded and summed up to get the dry plant weight (DPW). The root shoot ratio (RSR) was calculated using the formula: RSR= DRW/DSW. Relative water contents (RWC) were calculated using the following formula: RWC = [(Fresh weight - Dry weight) / (Turgid weight – Dry weight)] × 100.

SSR genotyping

For extraction of the genomic DNA from each accession group, 4 to 5 young fully expanded leaves from each plant were collected and stored at −80°C. The genomic DNA was isolated from the frozen leaf tissues using the cetyltrimethyl ammonium bromide (CTAB) method described (Zhang et al., 2000). The DNA samples were stored at -20°C until further use. The DNA quality was evaluated with 1% agarose gel electrophoresis prepared using 0.5X TAE buffer, and ethidium bromide (10 ng/100 ml) was added to the gel to stain the DNA bands. The samples were electrophoresed for approximately 30 min after which the products were viewed using an ultraviolet trans- illuminator and photographed using the Syngene Gel Documentation System. The DNA concentration was estimated by the absorbance at 260 nm. The working DNA samples (containing 50 ng μL^{-1}) were stored at 4°C for genotyping. SSR primer pairs used were from different sources: NAU from Nanjing Agricultural University, Nanjing, China (Han et al., 2006); BNL primers from Research Genetics Co. (Huntsville, AL, USA, http://www.resgen.com); JESPR from sequences of Reddy et al. (2001); TM from Dr. John Yu, USDA-ARS, Crops Germplasm Research Unit, TE, USA; CIR from Nguyen et al. (2004) (Table 2). Details about these markers can be found at www.cottonmarker.org. All 90 accessions were genotyped using a 180 core set of SSR marker primers. These chromosome-specific primer pairs were selected using the results of different laboratories and published papers (Siu et al., 2000; Han et al., 2006; Shen et al., 2005; Abdurakhmonov et al., 2007). They were based on information related to important QTLs and chromosome distribution.

Table 2. List of SSR markers used in the study.

S/N	Primers	chr.0	Position	S/N	Primers	chr.0	Position
1	BNL1053	-	-	49	NAU2691-275	D3(ch.17)	81.063
2	BNL1421	A1(ch.1)	163.5	50	NAU2697-165	D13(ch.18)	87.399
3	BNL1606-180	D3(ch.17)	50.727	51	NAU2714-450	D6(ch.25)	112.068
4	BNL1672-110	A9(ch.9)	60.015	52	NAU2974_195	D7(ch.16)	101.7
5	BNL226	D2(ch.14)	120	53	NAU2980-270	D13(ch.18)	-
6	BNL2569-170	A6(ch.6)	111.001	54	NAU2995-195	A7(ch.7)	70.732
7	BNL3089-125	A4(ch.4)	3.399	55	NAU3011-330	D13(ch.18)	98.587
8	BNL3280-195	D10(ch.20)	107.214	56	NAU3053-190	D7(ch.16)	30.704
9	BNL3359-210	A6(ch.6)	134.565	57	NAU3084-235	D12(ch.26)	130.319
10	BNL3590-180	A2(ch.2)	61.01	58	NAU3092-245	D5(ch.19)	9.968
11	BNL409-100	A13(Chr.13)	127.092	59	NAU3095-225	D5(ch.19)	180.972
12	BNL448	D4(ch.22)	22.7	60	NAU3100-195	D9(Chr.23)	25.886
13	BNL946-335	D10(ch.20)	61.17	61	NAU3206-500	A6(ch.6)	36.353
14	JESPR153-130	A13(ch.13)	62.641	62	NAU3207-275	A8(ch.8)	70.619
15	JESPR220-200	D4(ch.22)	102.872	63	NAU3209-265	D2(ch.14)	104.477
16	JESPR274-105	A9(ch.9)	42.802	64	NAU3254-1600	A1(ch.1)	102.873
17	JESPR291-175	D8(ch.24)	152.036	65	NAU3306-2400	D6(ch.25)	39.547
18	JESPR92	D12(ch.26)	132.1	66	NAU3385-500	A1(ch.1)	74.005
19	NAU1070-165	D2(ch.14)	13.512	67	NAU3414-225	A9(ch.9)	119.156
20	NAU1102	D5(ch.19)	0	68	NAU3424-210	D7(ch.16)	-
21	NAU1103	D11(ch.21)	7.5	69	NAU3519-215	A12(ch.12)	146.714
22	NAU1125	D8(ch.24)	47.7	70	NAU3529-295	A5(ch.5)	117.035
23	NAU1141-200	A13(ch.13)	31.193	71	NAU3606	-	-
24	NAU1167-200	A3(ch.3)	105.325	72	NAU3608-220	D7(ch.16)	121.531
25	NAU1190_400	A3(ch.3)	62.9	73	NAU3654-350	A7(ch.7)	23.067
26	NAU1200	A5(ch.5)	16.6	74	NAU3695_280	A11(ch.11)	125.6
27	NAU1218-150	A6(ch.6)	57.568	75	NAU3703_660	A11(ch.11)	67.3
28	NAU1230	D5(ch.19)	0.2	76	NAU3860_250	A12(ch.12)	85
29	NAU1233	A10(ch.10)	40.2	77	NAU3901_280	D1(ch.15)	117.3
30	NAU1254-305	A8(ch.8)	21.557	78	NAU3903_300	D2(ch.14)	110.1
31	NAU1266	-	-	79	NAU3916-320	D10(ch.20)	30.947
32	NAU1350-305	D8(ch.24)	107.342	80	NAU3961-210	D12(ch.26)	48.569
33	NAU1366-700	D11(ch.21)	142.709	81	NAU4073-160	A1(ch.1)	53.3
34	NAU1369-410	A8(ch.8)	43.809	82	NAU418	D9(ch.23)	99.6
35	NAU2016-250	A11(ch.11)	18.128	83	NAU437-290	A2(ch.2)	52.833
36	NAU2169-700	D8(ch.24)	77.052	84	NAU4516	D13	
37	NAU2190-360	D2(ch.14)	24.1	85	NAU453-180	D10(ch.20)	-
38	NAU2265-380	A2(ch.2)	5.902	86	NAU4565	-	-
39	NAU2306-380	D8(ch.24)	50.968	87	NAU458-190	D1(ch.15)	12.422
40	NAU2317-220	A10(ch.10)	2.554	88	NAU462-650	A9(ch.9)	117.544
41	NAU2336-215	D2(ch.14)	54.308	89	NAU483-250	A3(ch.3)	27.986
42	NAU2355b	-	-	90	NAU4956-240	D7(ch.16)	50.973
43	NAU2439-150	D8(ch.24)	70.516	91	NAU5091-160	D11(ch.21)	66.492
44	NAU2443-150	D13(ch.18)	61.754	92	NAU5189-280	D9(ch.23)	140.07
45	NAU2477-200	A4(ch.4)	82.172	93	NAU6106	-	-
46	NAU2503-250	D5(ch.19)	175.169	94	NAU6191	-	-
47	NAU2527	A10(ch.10)	31.7	95	TMP20-205	A11(ch.11)	72.758
48	NAU2540_160	D10(ch.20)	25.4				

base pairs (bp) for gels. DNA bands of SSRs were developed with silver staining, and recorded.

Table 3. Mean squares of the ANOVA of morpho-physiological traits.

Trait/Source of variation	Block	Water Levels (WL)	Error	Genotypes (G)	Water Levels × Genotypes	Coeff Var. %
Shoot Length (cm)	445.7589	21182.3**	12.91	67.8**	18.7*	14.1
Root Length (cm)	56.1126	3463.6**	18	44.47**	17.73n.s.	22.2
Fresh Shoot Weight (g)	13.8	2409.6**	1.46	3.93**	2.25**	29.7
Fresh Root Weight (g)	0.29	34.0**	0.04	0.07**	0.048n.s.	39.4
Dry shoot weight (g)	2.1758	60.04**	0.089	0.25**	0.14**	34
Dry root weight (g)	0.0136	0.0072**	0.0044	0.0101**	0.0064.n.s.	51
Relative Water Contents %	151	14307**	29.11	96.25**	71.13**	12

*= P≤0.05; ** = P ≤0.01; *** = P ≤ 0.001

The PCR amplifications were performed in a 10 µl reaction mix containing 1 µl 10× PCR buffer, 0.2 µl dNTPs (5 mM each), 0.1 µl 25 mM $MgCl_2$, 0.1 µl Taq DNA polymerase, and 1 µl (50 ng) genomic DNA. The microsatellites were amplified using the standard polymerase chain reaction (PCR) procedures described by Zhang et al. (2000). Two millilitre of PCR products was separated vertically on denaturing 16% polyacrylamide gels with 5× TBE buffer at 180 V for 45 min and stained with silver (Bassam et al., 1991). A 50 bp DNA ladder was used to estimate allele sizes in

Silver staining and development of bands

After a specific migration of the band on the gel, the gels were placed in fixing solution (40% ethanol and 10% glacial acetic acid) for 20 min. They were washed three times with distilled water and stained with silver staining solution (0.2% $AgNO_3$) for 20 min. After staining, the gels were again washed three times with distilled water for 20 s, and the developing solution (3% NaOH and 0.05% formaldehyde) was applied for 3 to 5 min.

Inference of population structure for association mapping

The STRUCTURE software is a DOS, Windows, UNIX (Solaris) and Linux based database that performs a model-based clustering method for gathering the occurrence of population structure, finding diverse genetic populations, allocating individuals to populations, and classifying migrants and admixed individuals. Complementary studies on genotypic data for evaluating the population structure before continuing with LD analysis were performed by a model-based approach; they were implemented in the software package STRUCTURE (Pritchard et al., 2000) to identify subgroups in cotton cultivars under study. Admixture model under independent allele frequencies using the burn-in time of 50,000, and a number of MCMC repeats at 100,000 were used (Pritchard et al., 2000), with the K ranging from 2 to 15.

Extent of linkage disequilibrium and marker-trait association analysis

The genome-wide LD between pairs of SSR marker loci was studied using the software package TASSEL ver. 2.1. Linkage disequilibrium was estimated by a weighted average of squared allele-frequency correlations (R^2) between SSR loci. The significance of pairwise LD (P-values≤0.005) among all possible SSR loci was evaluated using TASSEL. The LD values between all pairs of SSR loci were plotted as LD plots using TASSEL to estimate the general view of genome-wide LD patterns and to evaluate LD structures. The marker-trait associations were calculated by GLM association test incorporating Q matrices from STRUCTURE2.2 into TASSEL software package (Bradbury et al., 2007). To assess significant marker-trait associations P-marker ≤ 0.05 was used.

RESULTS

Phenotypic variation

The cotton varieties under study revealed a wide range of phenotypic variation in morpho-physiological traits under both control and drought treatments (Table 3). Mean squares from analysis of variance for all morpho-physiological traits revealed highly significant variation (P≤0.05) with respect to water levels and genotypes. However, the interaction between water levels and genotypes was significant for shoot length, fresh shoot weight and dry shoot weight (Table 1). Mean value of all 90 genotypes indicated significant reduction in all seedling traits. The considerable amount of genotypic variance apparent in all traits shows that variance under water stress conditions is genetically determined and selection of varieties\lines for drought tolerance on the basis of seedling traits is possible. Correlation coefficients between means in the stressed and non-stressed conditions were positive and highly significant (P ≤ 0.01) for most of the traits (Table 4). The frequency distribution of morpho-physiological traits under control (W1) and water stressed (W2) treatments indicated considerable genotypic variance (Figure 1).

SSR genotyping, population structure, pairwise linkage disequilibrium and LD decay

The selected SSR primer pairs generated a total of 241 SSR alleles, of which 147 (60%) were polymorphic, resulting in 57.5% polymorphism. The average number of polymorphic alleles per primer was 2.10. To determine the population structure and number of subgroups in cotton cultivars under study, a model-based approach, implemented in the software package STRUCTURE (Pritchard et al., 2000), the distribution of log probability of data, LnP(D), did not show a clear peak against any value of K, but by the use of parameter ΔK, rate of

Table 4. Correlation coefficients of morpho-physiological traits under well-watered (W) and drought-stress (W2) treatment.

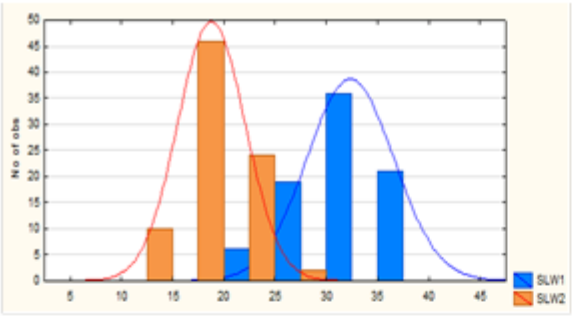
Traits		PDW	PFW	PH	RDW	RFW	RFW	RSR	SDW	SFW
PFW	W1	0.89958**								
	W2	0.6296**								
PH	W1	0.7083**	0.7527**							
	W2	0.4677**	0.5698**							
RDW	W1	0.7779**	0.7261**	0.5532**						
	W2	0.3963**	0.1202NS	0.1075NS						
RFW	W1	0.6880**	0.6861**	0.4361**	0.7831**					
	W2	0.2841*	0.5756**	0.2715*	0.1734NS					
RL	W1	0.4815**	0.5279**	0.7717**	0.4702**	0.3889**				
	W2	0.1851NS	0.1836NS	0.7498**	0.0888NS	0.1043NS				
RSR	W1	-0.3833**	-0.3175*	-0.3001*	0.2084NS	0.0163NS	-0.0793NS			
	W2	-0.1579NS	-0.2843*	-0.2024NS	0.8251**	-0.0110NS	-0.0482NS			
SDW	W1	0.9963**	0.8927**	0.7066**	0.7221**	0.6512**	0.4689**	0.4522**		
	W2	0.9692**	0.6499**	0.4753**	0.1597NS	0.2585**	0.1740NS	-0.390**		
SFW	W1	0.8862**	0.9954**	0.7596**	0.6854**	0.6135**	0.5216**	-0.3470*	0.8837**	
	W2	0.6367**	0.9897**	0.5733**	0.1001NS	0.4530*	0.1818NS	-0.3085*	0.6636**	
SL	W1	0.6569**	0.6852**	0.8517**	0.4349**	0.3295*	0.3241*	-0.381**	0.6648**	0.7007**
	W2	0.5268**	0.6762**	0.8041**	0.0795NS	0.3097**	0.2097NS	-0.2562*	0.5481**	0.6825**

PH, Plant length; PFW, Plant fresh weight ;SL, shoot length; RL, Root length; SFW, shoot fresh weight; shoot dry weight; SDW, RDW, root dry weight; RFW, Root fresh weight; RSR, root shoot ratio; RWC, relative water contents.

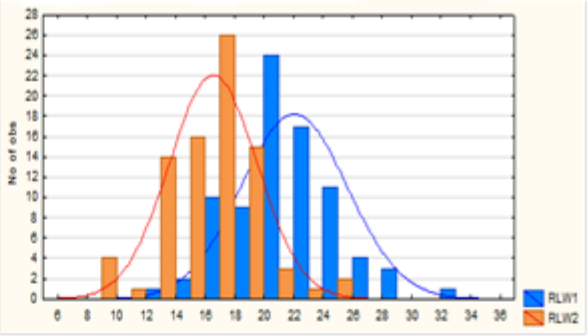
change in the log probability of the data, graph peaked against a value of K = 5. This confirmed 5 subpopulations in the germplasm at significant threshold values (R^2≥0.05). 7.1% of the SSR marker pairs showed significant pairwise LD in 90 accessions of *G. hirsutum* in our study. Plots for pairwise LD between SSR markers demonstrated significant LD blocks in the genome-wide LD analysis. We observed a significant (R2) LD between 13 pairs of SSR loci within the same chromosome in range of 180 cM between NAU1230 and NAU3095 loci in chromosome D5 and 1.612 cM and between NAU462 and NAU3414 in chromosome A9. Triangle plots for pairwise LD between SSR markers demonstrated significant LD blocks in the genome-wide LD analysis. Genome-wide LD decay was assessed by plotting r2 LD values as a function of genetic distance in cM. Two long stretches of LD blocks were observed on chromosomes A3 and D9, extending to a distance of 180 and 77 cM respectively (Table 5). Genome-wide LD at r^2 > 0.1 rapidly decayed within ~1.61 to 11 cM, indicating a strong potential for association mapping (Saeed et al., 2014). The percentage of SSR loci pairs in LD observed in 90 *G. hirsutum* (7.1%) was comparable with reports in cotton (11 and 12%) (Abdurakhmonov and Addukarimov, 2008) maize (10%) (Remington et al., 2001) and sorghum (8.7%) (Hamblin et al., 2004). A high recombination rate in allopolyploid cottons was reported (Brubaker et al., 1999) and it might be one of the factors explaining the observed low level of pairwise LD in cotton, along with mutation, selection, and genetic drift that occurred in the domestication of *G. hirsutum* germplasm.

Marker-trait association for morpho-physiological traits

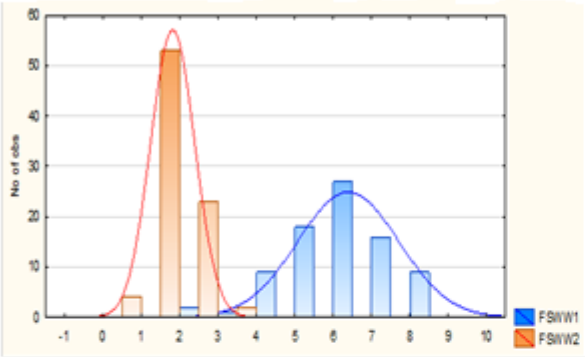
A total of 21 marker loci identified by GLM analysis were significantly associated (P≤0.001) with phenotyped traits under both control and drought treatment (Table 6). Out of these 21 markers, NAU3414, NAU2691, NAU1141 and NAU1190 were associated with more than one morpho-physiological trait under drought treatments (Table 4). Phenotypic variance explained (R^2) value ranging from 9.91 to 19%. Highest phenotypic variance explaining (R^2) was ascribed to NAU3011 chromosome D13 significantly (P≤0.001) associated with root length under drought treatment. This locus appeared to be a major locus as it is associated with highest phenotypic variance. NAU3414 located on chromosome A9 was associated with



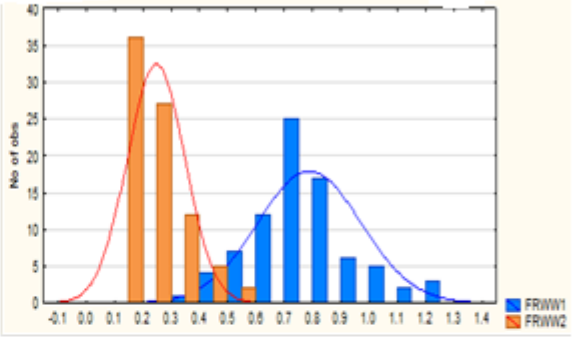
Shoot length under control (W1) and water stressed (W2) treatments



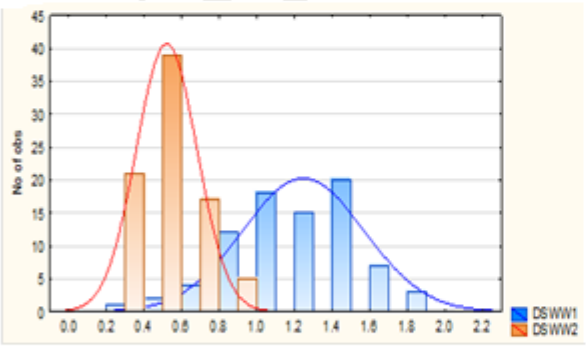
Root length under control (W1) and water stressed (W2) treatments



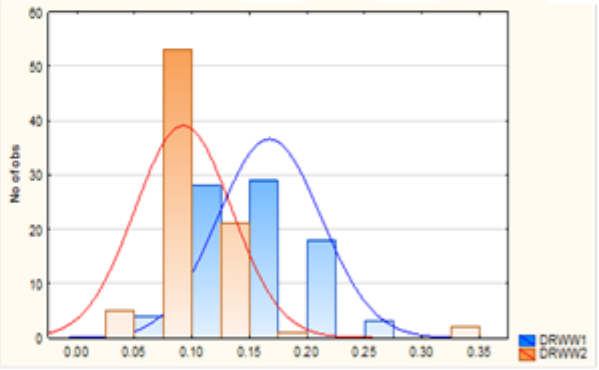
Fresh Shoot weight under control (W1) and water stressed (W2) treatments



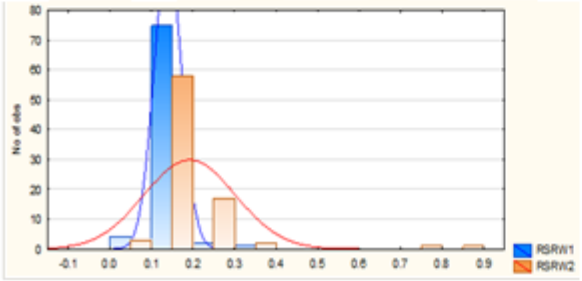
Fresh Root weight under control (W1) and water stressed (W2) treatments



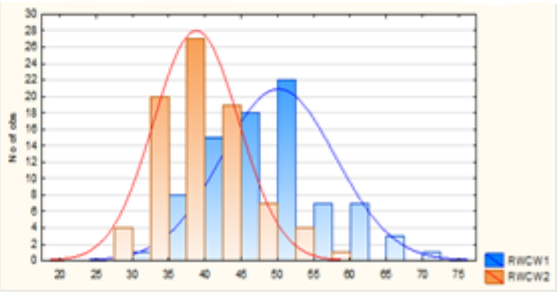
Dry Shoot weight under control (W1) and water stressed (W2) treatments



Dry Root weight under control (W1) and water stressed (W2) treatments



Root Shoot Ratio under control (W1) and water stressed (W2) treatments



Relative Water content under control (W1) and water stressed (W2) treatments

Figure 1. Frequency distribution of morpho-physiological traits under control (W1) and water stressed (W2) treatments.

Table 5. The pairwise genome linkage disequilibrium (LD) between pairs of SSR markers in same chromosomes and their distances.

Locus Name[1]	Locus Name[2]	R^2	DPrime	pDiseq	Chr	cM
NAU3095	NAU1230	0.181971	0.827381	0.000386	D5	180.772
NAU2691	BNL1606	0.107744	0.444444	0.047173	D3	30.336
NAU3011	NAU2697	0.098475	0.373626	0.02466	D13	11.188
NAU2336	NAU2190	0.083965	0.576803	0.014481	D2	30.208
NAU462	NAU3414	0.083056	0.288194	0.020138	A9	1.612
NAU4105	NAU2697	0.080887	0.355656	0.013277	D13	29.299
NAU483	NAU1167	0.076643	0.835045	0.003996	A3	77.339
NAU1230	NAU1102	0.065643	0.367246	0.036619	D5	-
NAU462	BNL1672	0.064808	0.280702	0.036295	A9	57.529
NAU3703	NAU3695	0.0625	0.4375	0.022808	A11	58.3
NAU3254	NAU2083	0.061728	0.333333	0.03903	A1	75.02
NAU3011	NAU2980	0.060445	0.340909	0.078648	D13	-
NAU2336	NAU1070	0.048902	0.511364	0.102041	D2	40.796

Table 6. Marker-trait associations assessed by GLM analysis with their phenotypic variance explained (R^2) values ($P \leq 0.001$).

Traits	Marker	Chr.	W1	W2	Relative value
SL	NAU3414	A9	0.113	-	-
	BNL1606	D3	-	-	0.118
	BNL1421	A1	-	-	0.1232
RL	NAU2691	D3	0.0793	-	-
	NAU2697	D13	0.1021	-	-
	NAU3011	D13	-	0.1904	-
PL	NAU5091	D11	-	0.1529	-
FSW	NAU1190	A3	0.1357	-	-
	NAU3414	A9	-	0.1791	-
	NAU3903	D2	0.0996	-	-
	JESPR220	D4	0.1687	-	-
FRW	NAU6191	A6	-	0.102	-
	NAU1141	A13	0.0942	-	-
FPW	NAU3100	D9	-	-	0.1133
	NAU1190	A3	0.1263	-	-
	NAU2974	D7	-	-	0.1096
DSW	NAU1190	A3	0.1226	-	-
	NAU1070	D2	-	-	0.1219
	NAU3414	A9	-	0.1129	-
DRW	NAU2974	D7	-	-	0.0991
	NAU2265	A2	0.1068	-	-
RSR	NAU1167	A3	0.1193	-	-
	NAU1369	A8	0.1193	-	-
RWC	NAU1141	A13	0.1096	-	-
	NAU2016	A11	-	0.1068	-

maximum number of traits (shoot length, shoot fresh weight and dry shoot weight whose value ranged from 11.3, 17.9 and 11.29% respectively). Six markers were exclusively associated with drought treatment (W_2).

DISCUSSION

Currently, association mapping methods has been used in diverse plant species such as bread wheat (Yu et al.,

2012; Phumichai et al., 2012), barley (Cockram et al., 2008), triticale (Niedziela et al., 2012) and bean (Shi et al., 2011). For molecular studies, there should be a reasonable degree of variability present among the organism of interest; only then the molecular approaches can identify the genetic case underlying this variability. As there was a significant variability shown in our experimental material under greenhouse conditions, thus our molecular study findings are of future significance. In our cotton germplasm, the number of polymorphic alleles detected per primer pair ranged from one to eight, with 2.10 alleles per primer pair. The SSR markers revealed a considerable amount of variation in the sampled genome, even though the overall polymorphism detected for these cotton cultivars was relatively low. The narrow genetic base of cotton has been mentioned in many studies using such molecular markers as SSRs (Bertini et al., 2006; Zhang et al., 2011; Kalivas et al., 2011), within Upland cultivars, which generally reveal a low level of genetic variety. There was little variation in the estimation of the molecular diversity among the upland cultivars (*G. hirsutum*). However, Abdurakhmonov et al. (2007) reported that the genetic distance for the Upland cultivars was in 0.01 to 0.28 range. The simple sequence repeats (SSR) allelic diversity found in our population for association analysis is approximately the same as the total diversity presented in more extended studies. The same mean number of alleles per locus as in our study was found in a collection of 106 accessions, with 2.13 SSR alleles (Guo et al., 2006). In this study at significant threshold values ($R^2 \geq 0.05$), 7.1% of the SSR marker pairs showed significant pairwise LD in a total of 90 accessions of *G. hirsutum*. This is comparable with previous reports on cotton: 11 to 12% of SSR loci pairs in the exotic *G. hirsutum* accessions (Abdurakhmonov and Abdukarimov, 2008), 4% SSR markers in *G. hirsutum* variety accessions (Abdurakhmonov et al., 2009) and 3% cotton germplasm from China and USA (Saeed et al., 2014). In our study, cotton germplasms used are from Pakistan; whereas in previous reports, the cotton germplasm used were of African, Australian, and Latin American, Mexican, Uzbek, China and USA ecotypes. The identification of QTLs' morpho-physiological traits related to drought tolerance has been reported in many studies: they include drought-related QTLs for morphological traits (Liang et al., 2014), physiological traits (Saeed et al., 2011), earliness (Li et al., 2013), yield (LiFang et al., 2010) and fibre traits (Islam et al., 2014). In our study, significant marker-trait associations were found in a total of 21 marker loci which were significantly associated ($P \leq 0.001$) with phenotyped data under both control and drought treatment. Markers, NAU3414, NAU2691, NAU1141 and NAU1190 were associated with more than one traits under drought treatments. Highest phenotypic variance explaining (R^2) was ascribed to NAU3011 loci significantly ($P \leq 0.001$) associated with root length under drought treatment. This locus appeared to be a major locus as it is associated with highest phenotypic variance. NAU3414 located on chromosome A9 was associated with maximum number of traits (shoot length, shoot fresh weight and dry shoot weight whose value ranged from 11.3, 17.9 and 11.29% respectively). Six markers were exclusively associated with drought treatment (W_2). This study also proved that association mapping approach has strong potential to assess significant marker-trait associations, save much time and cost compared to linkage mapping approach.

Conflict of interests

The authors have not declared any conflict of interests.

Abbreviations

QTL, Quantitative trait loci; **SSR,** simple sequence repeat; **TASSEL,** trait analysis by association, evolution and linkage; **GLM,** general linear model; **PL,** plant length, **PFW,** plant fresh weight, **SL,** shoot length, **RL,** root length, **SFW,** shoot fresh weight, **SDW,** shoot dry weight, **RDW,** root dry weight, **RFW,** root fresh weight, **RSR,** root shoot ratio; **RWC,** relative water contents; **LD,** linkage disequilibrium.

REFFERENCES

Abdelraheem A, Hughs SE, Jones DC, Zhang JF (2015). Genetic analysis and quantitative trait locus mapping of PEG-induced osmotic stress in cotton. Plant Breed. 134:111-120

Abdurakhmonov IY, Abdukarimov A (2008). Application of association mapping to understanding the genetic diversity of plant germplasm resources. Int. J. Plant Genom. 2008:18.

Abdurakhmonov IY, Saha S, Jenkins JN, Buriev ZT, Shermatov SE, Scheffler BE, Pepper AE, John ZY, Kohel RJ, Abdukarimov A (2009). Linkage disequilibrium based association mapping of fiber quality traits in *G. hirsutum* L. variety germplasm. Genetica 136:401-417.

Abdurakhmonov IY, Kohel R, Saha S, Pepper A, Yu J, Buriev Z, Abdullaev A, Shermatov S, Jenkins J, Scheffler B (2007). Genomewide linkage disequilibrium revealed by microsatellite markers and association study of fiber quality traits in cotton, Proceedings of the 15th Plant and Animal Genome Conference. San Diego, California, USA.

Bertini CH, Schuster I, Sediyama T, Barros EGD, Moreira MA (2006). Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genet. Mol. Biol. 29:321-329.

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

Brubaker C, Paterson A, Wendel J (1999). Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. Genome 42:184-203.

Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang T, Guo W, Chen X, Stelly DM, Rabinowicz PD, Town CD (2007). Toward sequencing cotton (*Gossypium*) genomes. Plant Physiol. 145:1303-1310.

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

Cockram J, White J, Zuluaga DL, Smith D, Comadran J, Macaulay M, Luo Z, Kearsey MJ, Werner P, Harrap D (2010). Genome-wide association mapping to candidate polymorphism resolution in the

unsequenced barley genome. Proc. Natl. Acad. Sci. 107:21611-21616.
Dahab AA, Saeed M, Mohamed BB, Ashraf MA, Puspito AN, Shahid KSBAA, Husnain T (2013). Genetic diversity assessment of cotton (*Gossypium hirsutum* L.) genotypes from Pakistan using simple sequence repeat markers. Aust. J. Crop Sci. 7(2):261-267.
Dongre A, Raut M, Paikrao V, Kadam R, Ashtikar S (2011). Use of RAPD and ISSR markers for determining genetic diversity of cotton (Gossypium hirsutum L.) working germplasm. J. Cotton Res. Dev. 25:137-143.
Guo WZ, Zhou BL, Yang LM, Wang W, Zhang TZ (2006). Genetic diversity of landraces in *Gossypium arboreum* L. race sinense assessed with simple sequence repeat markers. J. Integr. Plant Biol. 48:1008-1017.
Gutiérrez L, Cuesta-Marcos A, Castro AJ, von Zitzewitz J, Schmitt M, Hayes PM (2011). Association mapping of malting quality quantitative trait loci in winter barley: positive signals from small germplasm arrays. Plant Genome 4:256-272.
Hamblin MT, Mitchell SE, White GM, Gallego J, Kukatla R, Wing RA, Paterson AH, Kresovich S (2004). Comparative population genetics of the panicoid grasses: sequence polymorphism, linkage disequilibrium and selection in a diverse sample of *Sorghum bicolor*. Genetics 167:471-483.
Han Z, Wang C, Song X, Guo W, Gou J, Li C, Chen X, Zhang T (2006). Characteristics, development and mapping of Gossypium hirsutum derived EST-SSRs in allotetraploid cotton. Theor. Appl. Genet. 112:430-439.
He D-H, Lin Z-X, Zhang X-L, Nie Y-C, Guo X-P, Zhang Y-X, Li W (2007). QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of Gossypium hirsutum× Gossypium barbadense. Euphytica 153:181-197.
Islam MS, Zeng L, Delhom CD, Song X, Kim HJ, Li P, Fang DD (2014). Identification of cotton fiber quality quantitative trait loci using intraspecific crosses derived from two near-isogenic lines differing in fiber bundle strength. Mol. Breed. 34:373-384.
Kalivas A, Xanthopoulos F, Kehagia O, Tsaftaris A (2011). Agronomic characterization, genetic diversity and association analysis of cotton cultivars using simple sequence repeat molecular markers. Genet. Mol. Res. 10:208-217.
Li C, Wang X, Dong N, Zhao H, Xia Z, Wang R, Converse RL, Wang Q (2013). QTL analysis for early-maturing traits in cotton using two upland cotton (Gossypium hirsutum L.) crosses. Breed. Sci. 63:154.
Liang Q, Li P, Hu C, Hua H, Li Z, Rong Y, Wang K, Hua J (2014). Dynamic QTL and epistasis analysis on seedling root traits in upland cotton. J. Genet. 93:63-78.
LiFang S, Lei H, BaoMin H, Ling C, PeiZheng W (2010). QTL mapping of yield and agronomic traits of interspecific hybrid cotton. Xinjiang Agric. Sci. 47:67-72.
Nguyen T-B, Giband M, Brottier P, Risterucci A-M, Lacape J-M (2004). Wide coverage of the tetraploid cotton genome using newly developed microsatellite markers. Theor. Appl. Genet.109:167-175.
Niedziela A, Bednarek PT, Cichy H, Budzianowski G, Kilian A, Anioł A (2012). Aluminum tolerance association mapping in triticale. BMC Genomics 13:67.
Passioura J (2007). The drought environment: physical, biological and agricultural perspectives. J. Exp. Bot. 58:113-117.
Phumichai C, Chunwongse J, Jampatong S, Grudloyma P, Pulam T, Doungchan W, Wongkaew A, Kongsiri N (2012). Detection and integration of gene mapping of downy mildew resistance in maize inbred lines though linkage and association. Euphytica 187:369-379.
Poland JA, Bradbury PJ, Buckler ES, Nelson RJ (2011). Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. Proc. Natl. Acad. Sci. 108:6893-6898.
Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155:945-959.
Reddy OUK, Pepper AE, Abdurakhmonov I, Saha S, Jenkins JN, Brooks T, Bolek Y, El-Zik KM (2001). New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research. J. Cotton Sci. 5:103-113
Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc. Natl. Acad. Sci. 98:11479-11484.
Rodriguez-Uribe L, Abdelraheem A, Tiwari R, Sengupta-Gopalan C, Hughs S, Zhang J (2014). Identification of drought-responsive genes in a drought-tolerant cotton (Gossypium hirsutum L.) cultivar under reduced irrigation field conditions and development of candidate gene markers for drought tolerance. Mol. Breed. 34:1777-1796.
Saeed M, Dahab AhA, Wangzhen G, Tianzhen Z (2012). A cascade of recently discovered molecular mechanisms involved in abiotic stress tolerance of plants. Omics: J. Integr. Biol. 16:188-199.
Saeed M, Guo W, Ullah I, Tabbasam N, Zafar Y, Zhang T (2011). QTL mapping for physiology, yield and plant architecture traits in cotton (Gossypium hirsutum L.) grown under well-watered versus water-stress conditions. Electronic J. Biotechnol. 14:3-3.
Saeed M, Wangzhen G, Tianzhen Z (2014). Association mapping for salinity tolerance in cotton ('Gossypium hirsutum'L.) germplasm from US and diverse regions of China. Aust. J. Crop Sci. 8(3):338-346.
Shao Y, Jin L, Zhang G, Lu Y, Shen Y, Bao J (2011). Association mapping of grain color, phenolic content, flavonoid content and antioxidant capacity in dehulled rice. Theor. Appl. Genet. 122:1005-1016.
Shen X, Guo W, Zhu X, Yuan Y, Yu JZ, Kohel RJ, Zhang T (2005). Molecular mapping of QTLs for fiber qualities in three diverse lines in upland cotton using SSR markers. Mol. Breed. 15:169-181.
Shi C, Navabi A, Yu K (2011). Association mapping of common bacterial blight resistance QTL in Ontario bean breeding populations. BMC Plant Biol. 11:52.
Singh RP, Prasad P, Sunita K, Giri S, Reddy KR (2007). Influence of high temperature and breeding for heat tolerance in cotton: a review. Adv. Agron. 93:313-385.
Siu L, Saha S, Stelly D, Burr B, Cantrell R (2000). Chromosomal assignment of microsatellite loci in cotton. J. Heredity 91:326-332.
Würschum T, Maurer HP, Kraft T, Janssen G, Nilsson C, Reif JC (2011). Genome-wide association mapping of agronomic traits in sugar beet. Theor. Appl. Genet. 123:1121-1131.
YongSheng Q, RenZhong L, HongXian M, TianZhen Z, WangZhen G (2009). QTL mapping for yield traits in Upland cotton (Gossypium hirsutum L.). Acta Agron. Sinica 35:1812-1821
Yu L-X, Morgounov A, Wanyera R, Keser M, Singh SK, Sorrells M (2012). Identification of Ug99 stem rust resistance loci in winter wheat germplasm using genome-wide association analysis. Theor. Appl. Genet. 125:749-758.
Zhang Y, Muyrers JPP, Testa G, Stewart AF (2000). DNA cloning by homologous recombination in Escherichia coli. Nat. Biotechnol. 18:1314-1317.
Zhang Y, Wang X, Li Z, Zhang G, Ma Z (2011). Assessing genetic diversity of cotton cultivars using genomic and newly developed expressed sequence tag-derived microsatellite markers. Genet. Mol. Res. 10:1462-1470.