

Association mapping for salinity tolerance in cotton (*Gossypium hirsutum* L.) germplasm from US and diverse regions of China

Muhammad Saeed^{1,*}, Guo Wangzhen² and Zhang Tianzhen²

¹Department of Botany, Government College University, Faisalabad, Pakistan

²Cotton Research Institute, State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China

*Corresponding author: saeed242@hotmail.com

Abstract

Salinity affects cotton production worldwide. In our study, we assessed marker-trait associations for salinity tolerance in cotton using a set of 109 cotton variety germplasm (mainly from China and USA). Cotton varieties were screened for polymorphism with 250 SSR markers. Out of these 250 SSR markers, 98 were found to be polymorphic. Plant material was grown under normal versus saline (100mM and 200mM NaCl) conditions in greenhouse and data was collected for morpho-physiological traits at seedling stage. SSR markers linked with T₁, T₂, relative value of T₁, and relative value of T₂ treatments only were considered associated with salinity tolerance. On the basis of overall performance of cultivars judged by relative values, cultivars Jian mian 13, Si mian 4 and Gan mian 8 were found to be salt tolerant, whereas, Ke yi 2, Yan mian 48 and Zhong mian suo 49 were found to be salt sensitive. STRUCTURE software identified 5 sub-populations in this cotton germplasm. These sub-populations consisted of 10-30 varieties. At $r^2 \geq 0.05$, 3% SSR marker pairs showed significant pairwise linkage disequilibrium (LD). At the highly significant threshold of $r^2 \geq 0.1$, 1.82% of SSR marker pairs were remained in LD. Genome-wide LD at $r^2 \geq 0.1$ was reduced to ~4 – 7 cM, indicating a strong potential for association mapping. Markers BNL3103 (D6), NAU478 (D8) and BNL3140 (D9) were associated with salt treatment. These markers can be utilized in molecular breeding of cotton for the release of salt tolerant cultivars.

Keywords: Linkage disequilibrium (LD); Morpho-physiological traits; Simple sequence repeats (SSRs); STRUCTURE; TASSEL.

Abbreviations: DPW_dry plant weight; DRW_dry root weight; DSW_dry shoot weight; FPW_fresh plant weight; FRW_fresh root weight; FSW_fresh shoot weight; GLM_general linear model; LD_linkage disequilibrium; MAF_minor allele frequency; MCMC_Markov chain Monte Carlo; Max_maximum; Min_minimum; MLM_mixed linear model; PL_plant length; QTL_quantitative trait loci; RL_root length; RSR_root-shoot ratio; SD_standard deviation; SL_shoot length; SSR_simple sequence repeat; TASSEL_trait analysis by association, evolution and linkage; WC_water content.

Introduction

Genome-wide association mapping, based on linkage disequilibrium (LD), is a powerful technique to identify genomic regions linked to specific variants of a phenotypic trait. Compared to traditional QTL mapping using biparental populations, LD-based association mapping approach, using natural populations for mapping purposes, is a high resolution method (Abdurakhmonov et al., 2009). Genome-wide association studies have been extensively used in human genetics to find genomic regions linked to susceptibility to various diseases (Jorde, 2000; Weiss and Clark, 2002). In plants also, it is gaining wide spread use and there are reports of association studies in many crops like bread wheat (Brescghello and Sorrells, 2006; Reif et al., 2011; Yu et al., 2012; Hao et al., 2012), rice (Wen et al., 2009; Yan et al., 2009a; Shao et al., 2011), maize (Li et al., 2011; Lu et al., 2012; Phumichai et al., 2012), barley (Ivandic et al., 2002; Ivandic et al., 2003; Cockram et al., 2008; Roy et al., 2010), triticale (Niedziela et al., 2012); rape (Rezaeizad et al., 2011), and bean (Shi et al., 2011). In cotton (*Gossypium hirsutum* L.), there are limited reports of association mapping and that are on fiber quality traits (Abdurakhmonov et al., 2008, 2009). Association mapping for abiotic stress tolerance in cotton have not been attempted yet.

Cotton (*Gossypium* spp.) is the most important fiber and oilseed crop in the world, grown in more than 80 countries with a worldwide production of 123 million bales (480 pounds per bale) during the 2011/2012 growing season (United States Department of Agriculture, 2012). World cotton production is affected by a number of abiotic stresses (Saeed et al., 2011). Out of these, salinity is a major abiotic stress limiting cotton growth and development at the germination and seedling stage (Ashraf and Ahmad, 2000). There are a number of molecular mechanisms which are involved in tolerance to abiotic stresses in plants (Saeed et al., 2012). These molecular mechanisms encompass stress tolerance or stress avoidance phenomenon. Genes involved in these molecular mechanisms can be tagged with the help of molecular mapping approaches. In our present research, we assessed extent of LD in the *G. hirsutum* germplasm from USA and diverse regions of China. This is the first report of extent of LD in the cotton germplasm from an important cotton growing region of the world. Marker-trait associations for salinity tolerance were also identified. The objectives of this study were to (i) estimate extent of LD in the cotton variety germplasm (ii) assess power of association mapping

Table 1. List of cultivars used in the study.

S. No	Cultivar Name	Origin	S. No	Cultivar Name	Origin	S. No	Cultivar Name	Origin
1	140007	China	38	Ji 668	Hebei, China	75	Yu mian 5	Henan, China
2	Shan mian 1	Shaanxi, China	39	DPL14 A	America	76	Zhong mian suo 15	Henan, China
3	XiangSC-24	Hunan, China	40	TM-1	America	77	Yan mian 1	Jiangsu, China
4	Su mian 9	Jiangsu, China	41	Zhong mian suo 9409	Henan, China	78	Chuan 414	Sichuan, China
5	Lu mian 4	Shandong, China	42	Lu mian 6	Shandong, China	79	DPL 16	America
6	Xu zhou 142	Jiangsu, China	43	E mian 14	Hubei, China	80	Zhong mian suo 7	Henan, China
7	Ji mian 8	Hebei, China	44	Zhong mian suo 41	Henan, China	81	Yan mian 48	Jiangsu, China
8	57-681	Sichuan, China	45	Yun an 1	China	82	Jin mian 6	Shanxi, China
9	Liao mian 4	Liaoning, China	46	King	America	83	Liao mian 10	Liaoning, China
10	Ejing 1	Hubei, China	47	Shi duan 5	Hebei, China	84	Ji gan 3	China
11	Dai hong dai	Hunan, China	48	140005	China	85	Jin mian 19	Shanxi, China
12	MD51ne	America	49	Su mian 5	Jiangsu, China	86	Jin mian 23	Shanxi, China
13	Su mian 6	Jiangsu, China	50	Zhong mian suo 5	Henan, China	87	Dai 61	China
14	Lu mian 1	Shandong, China	51	Su mian 3	Jiangsu, China	88	Deng en 118	China
15	Ji mian i	Hebei, China	52	Xu zhou 1818	Jiangsu, China	89	Han dan 428	Hebei, China
16	Su mian 1	Jiangsu, China	53	Ke yi 2	Beijing, China	90	Zhong mian suo 23	Henan, China
17	Zhong mian suo 3	Henan, China	54	Xiang mian 10	Hunan, China	91	Gan mian 6	Jiangxi, China
18	Wan mian 2	China	55	DPL 15	America	92	Handi mian 289	China
19	Zhong mian suo 16	Henan, China	56	Lu mian 2	Shandong, China	93	Liao mian 5	Liaoning, China
20	Si mian 3	Jiangsu, China	57	Dong ting 1	Hunan, China	94	Dai 62	China
21	Esha 28	Hubei, China	58	Chuan mian 56	Sichuan, China	95	Zhong mian suo 49	Henan, China
22	Ji mian 12	Hebei, China	59	86-6	Henan, China	96	Jun mian 1	Xingjiang, China
23	Jian mian 13	China	60	Stoneville 4	America	97	Lu mian yan 18	Shandong, China
24	Yu mian 21	Henan, China	61	Jing simian	China	98	Lu mian 14	Shandong, China
25	Su mian 12	Jiangsu, China	62	Yu mian 1	Henan, China	99	Zhi mian 3	China
26	Shan 1155	Shaanxi, China	63	Su mian 16	Jiangsu, China	100	Dai xu mian	China
27	Zhong mian suo 34	Henan, China	64	Si mian 4	Jiangsu, China	101	Ji feng 106	China
28	Shi yuan 321	Hebei, China	65	Stoneville 2B	America	102	Lu mian 12	Shandong, China
29	I40006	China	66	Hua 101	Hubei, China	103	Liao mian 17	Liaoning, China
30	Lu mian 5	Shandong, China	67	Su mian 2	Jiangsu, China	104	Zhong mian suo 17	Henan, China
31	Zhong mian suo 12	Henan, China	68	Gan mian 8	Jiangxi, China	105	Zhong mian suo 44	Henan, China
32	Xiang mian 16	Hunan, China	69	Zhong mian suo 4133	Henan, China	106	Shan mian 4080	Shaanxi, China
33	Si mian 2	Jiangsu, China	70	86-1	Henan, China	107	Shang qiu 24	Henan, China
34	Xua hou 514	Jiangsu, China	71	Shan 401	Shaanxi, China	108	Foster 6	America
35	Zhong mian suo 19	Henan, China	72	Ji mian 7	Hebei, China	109	Shan 6192	Shaanxi, China
36	52-128	Sichuan, China	73	Zhong mian suo 45	Henan, China			
37	Zhong mian suo 25	Henan, China	74	Ejing 92	Hubei, China			

to detect reliable QTLs, and (iii) identify markers linked to salt tolerant traits in cotton.

Results

Phenotypic variation

Cotton varieties revealed a wide range of phenotypic variation in morpho-physiological traits (SL, RL, PL, FSW, FRW, FPW, DSW, DRW, DPW, RSR, and WC) under both control and salt treatments (Table 2). The growth of cotton cultivars was severely affected at 200mM NaCl treatment. There were significant differences for salt treatments, genotypes and salt \times genotype interactions (Table 3). Under T_0 treatment, the traits SL, RL, PL, FSW, FRW, FPW, DSW and DPW showed significant positive correlation ($P \leq 0.05$) (Table 4). FRW, DSW, DRW and DPW had significant negative correlation ($P \leq 0.05$) with WC. FPW, DSW and DRW had significant negative correlation ($P \leq 0.05$) with RSR. Under T_1 treatment, the traits SL, RL, PL, FSW, FRW, FPW, DSW, DRW and DPW had significant positive correlation ($P \leq 0.05$). FRW had significant positive correlation with RSR ($P \leq 0.05$). Under T_2 treatment, the traits SL, RL, PL, FSW, FRW, FPW, DSW, DRW and DPW had positive correlation ($P \leq 0.05$). There was significant negative correlation ($P \leq 0.05$) between SL and RSR; RL and RSR; PL and RSR; DSW with WC and RSR; DRW and DPW with WC. Quite large number of individuals had increase in RL (37 cultivars), FRW (68 cultivars), DRW (46 cultivars) and RSR (74 cultivars) under T_1 treatment (Fig 1). There was higher WC in more number of individuals under

T_2 treatment (91 cultivars) (Figure 1). On the basis of overall performance of cultivars judged by relative values, Jian mian 13, Si mian 4 and Gan mian 8 were found to be salt tolerant cultivars, whereas, Ke yi 2, Yan mian 48 and Zhong mian suo 49 were found to be salt sensitive cultivars.

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

SSR genotyping yielded a total of 217 amplicons or alleles from 98 primer pairs, with an average of 2.21 alleles/primer pair (a range of 2-7 alleles per primer pair). Every chromosome had 3- 4 primer pairs.

For determination of population structure, the distribution of log probability of data, $\ln P(D)$, did not show a clear peak against any value of K , but by the use of parameter ΔK , rate of change in the log probability of the data, graph peaked against a value of $K = 5$ (Evanno et al. 2005). This confirmed 5 subpopulations in the germplasm. Number of varieties in each subpopulation ranged from 10-30. Subpopulation 2 consisted of 10 varieties, whereas subpopulation 3 consisted of 30 varieties.

At significant threshold values of $r^2 \geq 0.05$, 3% SSR marker pairs showed a significant pairwise LD in a total of 109 cotton varieties (in a total of 4,560 pairwise comparisons). At the highly significant threshold of $r^2 \geq 0.1$, only 1.82% of SSR marker pairs were remained in LD. r^2 values ranged from 0.0 to 0.63. 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 r^2 LD values as a function of

Table 2. Phenotypic variation of cotton cultivars for morpho-physiological traits under T₀, T₁ and T₂ treatments.

Trait	Units	Treatment	Mean	Min	Max	SD	Kurtosis	Skewness
SL	cm	T ₀	21.00	14.83	24.75	2.00	0.02	-0.29
		T ₁	17.83	12.75	23.33	2.22	-0.41	0.25
		T ₂	14.90	10.00	19.83	1.78	0.46	0.25
RL	cm	T ₀	21.64	13.25	29.00	2.36	1.36	-0.03
		T ₁	20.50	12.00	25.25	2.32	1.31	-0.58
		T ₂	19.49	12.50	25.00	2.48	0.31	-0.57
PL	cm	T ₀	42.64	30.25	51.00	3.52	1.59	-0.57
		T ₁	38.33	26.13	46.33	3.66	0.50	-0.44
		T ₂	34.41	22.50	42.33	3.46	0.52	-0.48
FSW	g	T ₀	1.59	0.72	2.64	0.35	-0.01	0.10
		T ₁	1.36	0.71	2.03	0.28	-0.49	0.29
		T ₂	1.09	0.52	1.84	0.21	1.11	0.31
FRW	g	T ₀	0.26	0.12	0.46	0.07	0.18	0.40
		T ₁	0.30	0.16	0.50	0.08	-0.30	0.48
		T ₂	0.21	0.12	0.39	0.04	1.16	0.54
FPW	g	T ₀	1.86	0.86	2.93	0.39	-0.06	0.06
		T ₁	1.66	0.91	2.43	0.32	-0.59	0.24
		T ₂	1.30	0.66	2.11	0.23	1.00	0.30
DSW	g	T ₀	110.97	44.67	270.50	41.07	1.47	0.91
		T ₁	80.67	29.33	186.75	28.31	2.43	1.24
		T ₂	56.47	25.00	135.75	15.57	6.21	1.48
DRW	g	T ₀	44.22	10.33	138.00	19.91	4.21	1.46
		T ₁	41.83	16.00	93.50	17.32	0.65	0.96
		T ₂	17.76	5.50	37.33	5.89	0.95	0.89
DPW	g	T ₀	155.19	61.67	339.17	53.65	0.86	0.86
		T ₁	122.50	50.67	233.08	37.19	0.31	0.71
		T ₂	74.23	38.00	167.75	19.10	5.89	1.43
RSR		T ₀	0.42	0.18	1.14	0.19	2.03	1.39
		T ₁	0.56	0.16	1.42	0.25	2.80	1.46
		T ₂	0.32	0.12	0.68	0.10	1.02	0.65
WC		T ₀	12.23	5.30	25.46	3.49	1.93	1.04
		T ₁	14.03	7.53	25.90	3.84	0.46	0.68
		T ₂	17.45	9.36	28.23	3.95	-0.06	0.38

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

Trait/Source of variation	Block	Salt	Error	Genotype	Salt × Genotype	Model	Coeff Var
Shoot length	78.40***	2918.49***	5.15	19.58***	7.84***	29.89***	12.60
Root length	2.88 NS	269.13***	7.32	18.12***	11.48***	15.19***	13.13
Plant length	154.09***	4458.12***	13.43	47.96***	24.31***	60.04***	9.49
Fresh shoot weight	2.09***	14.99***	0.09	0.26***	0.13***	0.28***	22.19
Fresh root weight	0.05***	0.50***	0.01	0.01***	0.01***	0.01***	26.28
Fresh plant weight	2.29***	19.55***	0.07	0.42***	0.16***	0.38***	16.87
Dry shoot weight	3744.67**	183911.18***	634.12	3117.65***	1682.84***	3283.88***	30.97
Dry root weight	549.56*	59683.31***	137.44	973.73***	478.49***	1004.6***	33.86
Dry plant weight	3803.91*	416341.57***	854.08	5513.24***	2676.15***	6150.13***	25.35
Root-shoot ratio	0.64***	3.95***	0.04	0.13***	0.1***	0.14***	45.24
Water content	166.53***	1214.02***	18.87	29.16**	23.71*	33.69***	29.78

* = $P \leq 0.05$; ** = $P \leq 0.01$; *** = $P \leq 0.001$

genetic distance in cM. Two long stretches of LD blocks were observed on chromosomes D4 and D11, extending to a distance of 81.6 cM and 138.4 cM respectively (Table 4). Genome-wide LD at $r^2 \geq 0.1$ rapidly decayed within ~4 – 7 cM, indicating a strong potential for association mapping (Abdurakhmonov et al., 2008; 2009). There were a number of unlinked markers showing significant LD between pairs of loci. This shows that there are factors other than linkage generating LD in the cotton genome. Extent of LD varied on different chromosomes i.e., on chromosome D8, LD extended to 4.5 cM, whereas on chromosome D11, LD extended to 138.4 cM (Table 5).

Marker-trait associations

There were 16 significant ($P \leq 0.001$) marker-trait associations identified by MLM analysis (Table 6). Out of these 16 associations identified by MLM analysis, 11 were also confirmed by GLM analysis (Table 7). Phenotypic variance explained values (R^2) for these associations ranged

from 6% to 10%. Markers BNL3103 (D6), NAU478 (D8) and BNL3140 (D9) were associated with salt treatment. Markers NAU478 and BNL3140 were associated with more than one morpho-physiological trait under salt treatments. Marker NAU478 (D8) was associated with DRW and RSR. BNL3140 (D9) was associated with DRW and RSR.

Discussion

Linkage disequilibrium in cotton

For association mapping studies, occurrence of significant LD in the population is a pre-requisite. In our cotton germplasm, about 3% SSR marker pairs showed a significant pairwise LD at $r^2 \geq 0.05$. In the previous reports on cotton, 11–12% of SSR loci pairs in the exotic *G. hirsutum* accessions (Abdurakhmonov et al., 2008) and 4% SSR markers in *G. hirsutum* variety accessions (Abdurakhmonov et al., 2009) were in significant LD at $r^2 \geq 0.05$. Our cotton germplasm included varieties from China and USA; whereas

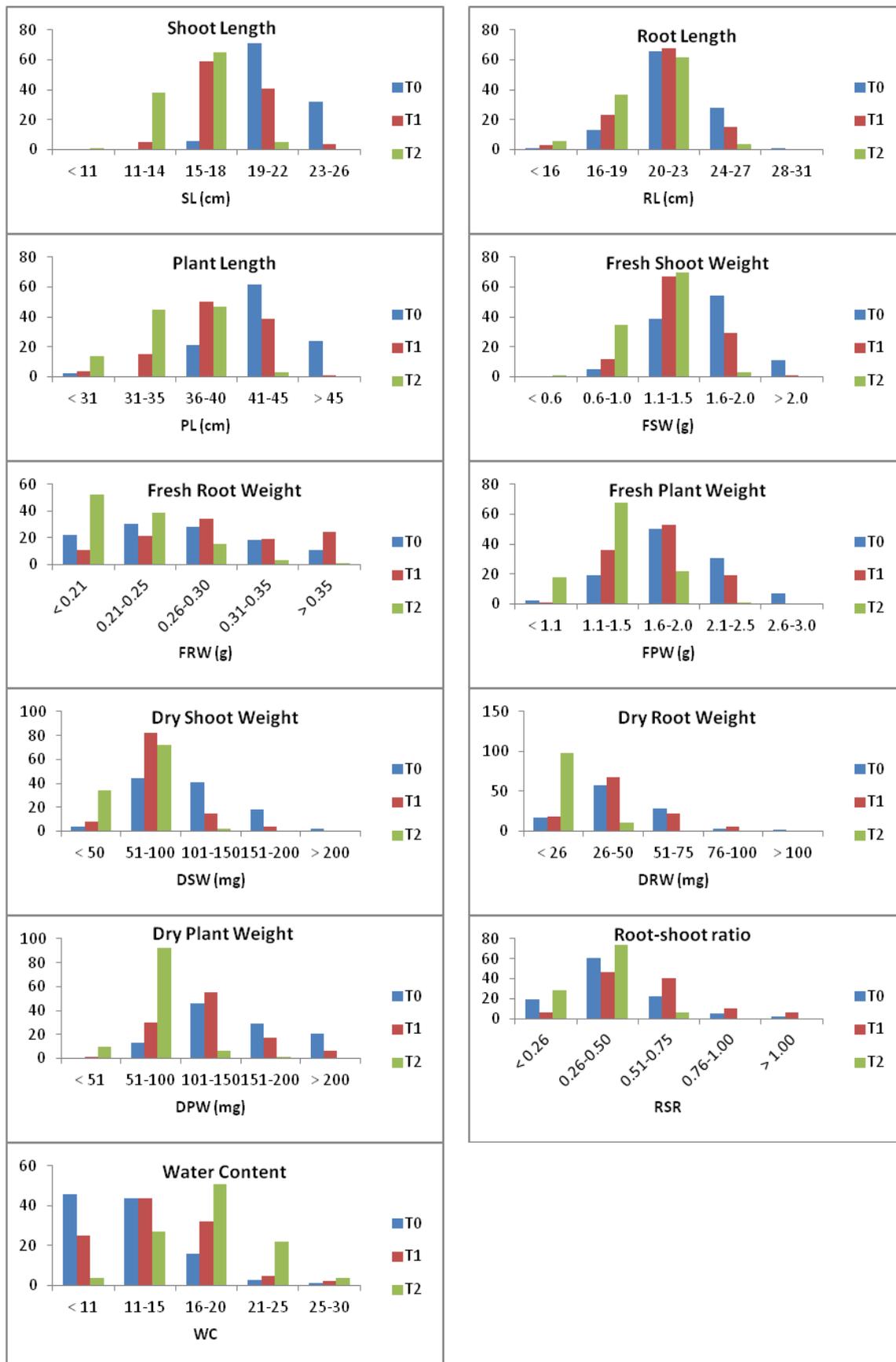


Fig 1. Frequency distribution of morpho-physiological traits under T₀, T₁ and T₂ treatments.

Table 4. Correlation coefficients of morpho-physiological traits under T₀, T₁ and T₂ treatments.

	SL	RL	PL	FSW	FRW	FPW	DSW	DRW	DPW	WC	RSR
SL	1	0.3098*** 0.2952**	0.8163*** 0.6883***	0.6344*** 0.6503***	0.1386NS 0.2518**	0.5927*** 0.5606***	0.4219*** 0.4000***	0.1016NS 0.1404NS	0.3619*** 0.2909**	0.0580NS 0.1559NS	-0.1982* -0.2205*
RL	0.3026**	1	0.7774*** 0.8663***	0.2889** 0.3297***	0.3187*** 0.3615***	0.3299*** 0.3283***	0.1979* 0.2166*	0.2295* 0.1283NS	0.2564** 0.1803NS	-0.0349NS 0.0408NS	0.0746NS -0.0314NS
PL	0.7646***	0.8178***	1	0.5851*** 0.5284***	0.2689** 0.4010***	0.5785*** 0.4754***	0.3698*** 0.3491***	0.1849NS 0.1787NS	0.3603*** 0.2807**	0.0448NS 0.0771NS	-0.0828NS -0.1096NS
FSW	0.4892***	0.4876***	0.6307***	1	0.3872*** 0.5037***	0.9646*** 0.9485***	0.6113*** 0.6353***	0.2845** 0.4046***	0.5827*** 0.5945***	0.0449NS 0.1480NS	-0.1422NS -0.1441NS
FRW	0.3644***	0.3377***	0.4372***	0.5564***	1	0.5709*** 0.6128***	0.2198* 0.5055***	0.6598*** 0.5393***	0.4810*** 0.5685***	-0.1764NS -0.1379NS	0.4649*** 0.1324NS
FPW	0.5199***	0.4622***	0.6288***	0.9689***	0.6796***	1	0.5849*** 0.6375***	0.4152*** 0.4325***	0.6294*** 0.6015***	-0.0039NS 0.1730NS	-0.0039NS -0.1110NS
DSW	0.3750***	0.3675***	0.4752***	0.6878***	0.5382***	0.7172***	1	0.2912** 0.4778***	0.8754*** 0.8706***	-0.5590*** -0.4045***	-0.4110*** -0.3538***
DRW	0.2778**	0.0596NS	0.1978*	0.2863**	0.5589***	0.3761***	0.4861***	1	0.6800*** 0.7143***	-0.5269*** -0.4073***	0.6763*** 0.6229***
DPW	0.3782***	0.2844**	0.4190***	0.6360***	0.6406***	0.6991***	0.9349***	0.7461***	1	-0.6705*** -0.5189***	0.0141NS -0.0290NS
WC	-0.0614NS	0.0286NS	-0.0255NS	-0.0120NS	-0.2245*	-0.0605NS	-0.5126***	-0.5592***	-0.6079***	1	-0.0701NS -0.0650NS
RSR	-0.0270NS	-0.1813NS	-0.1546NS	-0.3068**	0.1121NS	-0.2442*	-0.3523***	0.5608***	-0.0516NS	-0.1142NS	1

Note: Lower values of the diagonal are for T₀, upper values of the diagonal for T₁ and bold face values are for T₂ treatments, respectively.

Table 5. Extent of linkage disequilibrium in the cotton variety germplasm used.

Locus Name1	Locus Name2	r ²	D'	P-Diseq	Chr.	cM
TME03	NAU2083	0.05	0.30	0.006	A1	30
BNL3590	NAU437	0.09	0.59	0.003	A2	9
NAU3214	NAU2190	0.13	0.47	0.001	D2	8
JESPR220	BNL448	0.12	0.44	0.006	D4	81.57
NAU1042	NAU3269	0.08	0.35	0.000	A5	30.2
BNL1604	NAU3654	0.13	0.87	0.001	A7	50.7
NAU478	NAU2306	0.09	0.48	0.092	D8	15
NAU2439	NAU478	0.10	1.00	0.005	D8	4.53
NAU1350	NAU2169	0.08	0.69	0.016	D8	30
JESPR291	NAU1350	0.11	0.88	0.005	D8	45
NAU462	BNL1414	0.07	0.66	0.005	A9	20.5
NAU5166	NAU2317	0.49	1.00	0.001	A10	20.8
TML05	BNL946	0.12	0.71	0.000	D10	7
TML05	NAU2549	0.05	0.26	0.029	D10	42.5
NAU1366	TMH05	0.07	0.41	0.010	D11	138.4

previous reports were of the cotton germplasm including African, Australian, Latin American, Mexican, and Uzbek ecotypes. It suggests that the *G. hirsutum* germplasm from diverse sources have undergone similar factors contributing to LD. Distance for decay of LD varied in different crop plants as in *Arabidopsis thaliana*, 50 kb (Nordborg et al., 2005); wheat, <1 cM to ~5cM (Bressegello and Sorrells, 2006); maize, 1-10 kb (Yan et al., 2009b); barley, 5-10 cM (Pasam et al., 2012); tobacco, 1-75 cM (Fricano et al., 2012); and sunflower, 100 kb (Fusari et al., 2008). Reports of linkage disequilibrium in other crop plants indicate that extent of LD varies in different organisms and it depends on factors involved in specific mode of breeding and selection pressure. In our study, long haplotypic blocks of LD were observed on some chromosomes (Table 4). This may be the result of selection pressure for some specific traits in cotton.

Salinity tolerance and marker-trait associations

Salt tolerance in plants is a complex phenomenon involving a large number of biochemical, morphological and physiolo-

gical processes (Flowers et al., 1977; Greenway and Munns, 1980). Cotton is most sensitive to salinity at seedling stage (Pessaraki and Tucker, 1985; Khorsandi and Anaghali, 2009) and effect can be quantified by measuring morpho-physiological traits (Munns, 2007). Clear and significant differences were found in cotton genotypes for all the measured morpho-physiological traits under salt treatments (Table 3). For molecular studies, there should be a fair degree of variability present among the organism of interest, only then the molecular approaches can identify the genetic cause underlying this variability. As there was a significant variability shown in our experimental plant material, so the results of our molecular findings are of future significance. In this study, significant marker-trait associations were found. Marker NAU2679 (A6) had significant associations both under control and salt treatments. This marker will be helpful in future endeavors for developing cotton cultivars best suited under both control and salt stress conditions. Markers BNL3103 (D6), NAU478 (D8) and BNL3140 (D9) were associated with salt treatments only. This finding highlights the contribution of D subgenome of tetraploid cotton in

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

Trait	Marker	Chr.	T ₀	T ₁	T ₂	Rel-T ₁	Rel-T ₂
SL	NAU2679	A6				0.07	
PL	JESPR135	A11					0.06
FSW	NAU2679	A6	0.08			0.08	
	JESPR135	A11					0.06
FPW	NAU2679	A6	0.08			0.09	
DRW	BNL3103	D6			0.06		
	NAU478	D8			0.06		
	BNL3140	D9				0.10	
	TMH05	D11	0.06				
DPW	BNL3103	D6			0.06		
RSR	NAU478	D8			0.06		
	BNL3140	D9		0.06			
	JESPR135	A11	0.10			0.05	

abiotic stress tolerance. Previous reports have identified QTLs controlling fiber quality and yield located on the D subgenome (Jiang et al., 1998). Our results suggest, complement to the previous findings, that improved fiber quality, yield and abiotic stress tolerance can be combined in the same variety simultaneously. Identified markers can be utilized for molecular breeding of cotton for the release of salt tolerant varieties.

Materials and methods

Plant materials

The plant material consisted of 109 cotton (*G. hirsutum* L.) varieties. Out of these 109 varieties, 9 were from USA and the remaining 100 were from China, originated from diverse regions of China (Table 1).

Sowing of plant materials

These cotton varieties were grown in the green house in polythene bags (30 cm × 5 cm) containing vermiculite (500 g) of pH 6.5, arranged according to a randomized complete block design with three replications and three treatments (T_0 , control with ordinary tap water application; T_1 , 100mM NaCl solution application; and T_2 , 200mM NaCl solution application).

Each replication contained 5 bags. On December 24, 2008, two days before sowing the seeds, the polythene bags containing vermiculite were given ordinary water to saturation. On December 26, 2008, 5 seeds/bag were sown for each cultivar at a depth of 3cm. After germination, only 1 plant/bag was kept. 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. All three treatments were also applied the nutrient solution for proper cotton seedlings growth. On January 4, 2009 the control treatment was given the ordinary tap water, while T_1 and T_2 treatments were given the salt solution treatment (100mM each). T_2 treatment (200mM NaCl) was applied in two doses, the second dose on

January 13, 2009. At that time control and T_1 treatment was given ordinary tap water.

Phenotyping of plant materials

On January 26, 2009 green-house experiment was completed and all the plants were made free of vermiculite carefully and following parameters measured. First, PL (cm) and PFW (g) were recorded. After that plants were separated into shoot and root parts and data were recorded for SL (cm), RL (cm), FSW (g) and FRW (g). 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 DPW (g). The RSR was calculated using the formula:

$$RSR = \frac{DRW}{DSW}$$

Salt stress tolerance indices of genotypes were counted by determining the relative values for every trait:

$$\text{Relative value} = \frac{\text{Value under stress treatment (S)}}{\text{Value under control treatment (C)}}$$

SSR Genotyping

For extraction of genomic DNA, from each variety 4-5 young, fully expanded, leaves were collected, and stored at -80 °C. Genomic DNA was extracted from these leaf tissues following the method of Paterson et al. (1993). Cotton germplasm was genotyped for polymorphism with 250 SSR markers. Out of these 250 SSR markers, 98 were found to be polymorphic in this cotton germplasm (Table S1). SSR primer pairs used were from different sources as NAU from Nanjing Agricultural University, Nanjing, China (Han et al., 2004; 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). Details about these markers can be found at www.cottonmarker.org and www.cottongen.org. Microsatellites were amplified by

Table 7. Marker-trait associations assessed by GLM analysis with their phenotypic variance explained (R^2) values ($P \leq 0.001$).

Trait	Marker	Chr.	T0	T1	T2	Rel-T1	Rel-T2
SL	NAU2679	A6				0.14	
RL	JESPR135	A11			0.09		
PL	NAU437	A2	0.10				
	NAU483	A3			0.10		
	NAU3016	D3		0.11			
FSW	BNL3255	A8				0.09	
	BNL3590	A2		0.13			
	NAU2679	A6				0.15	
FRW	NAU1254	A8	0.10				
FPW	BNL3590	A2		0.14			
	NAU2679	A6				0.17	
DSW	JESPR135	A11					0.10
DRW	BNL3103	D6			0.10		
	NAU478	D8			0.11		
	NAU462	A9				0.10	
	BNL3140	D9		0.10		0.17	
	TMH05	D11	0.13				
RSR	NAU478	D8			0.10		
	BNL3140	D9		0.10			
	NAU5189	D9				0.13	
	JESPR135	A11	0.19				
	JESPR204	D13				0.13	

standard PCR procedures described by Zhang et al. (2000). DNA bands of amplification products were developed with silver staining and recorded with SX-image system (Sixing Biological Technology Co. Shanghai, China). Every cotton chromosome contained ~3-5 SSR markers. These chromosome-specific markers were selected based on findings of Han et al. (2004; 2006). These SSR markers spanned approximately 2,468 cM distance (48% of cotton genome coverage). This genome coverage was based on previous mapping experiments of Han et al. (2004; 2006). SSR bands were scored on their base pair sizes and the missing bands were scored either as “?” or “-9” depending on the software requirements. SSR markers linked with T₁, T₂, relative value of T₁, and relative value of T₂ treatments only were considered associated with salt tolerance.

Inference of population structure

Number of subpopulations in the cotton variety germplasm was determined by using software package STRUCTURE (Pritchard et al., 2000). Admixture model under independent allele frequencies using the burn-in time of 50,000 and number of MCMC repeats at 100,000 was used (Pritchard and Wen, 2004) with the K ranging from 2 to 10.

Pairwise linkage disequilibrium and LD decay

The genome-wide LD between pairs of SSR loci was studied according to Witt and Buckler (2003). For determination of LD and marker-trait associations software package TASSEL ver. 2.1 (Bradbury et al., 2007) (<http://www.maizegenetics.net>) was used. MAF filtered datasets were used to do this analysis, because minor alleles are usually problematic and biased for LD estimates between pairs of loci (Mohlke et al., 2001). The MAF removal was performed using the TASSEL site filtration function. LD 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 with the rapid permutation test in 10,000 shuffles. The LD values between all pairs of SSR loci were plotted as triangle LD plots to estimate the general view of genome-wide LD patterns and evaluate ‘block-like’ LD structures. LD decay (at $r^2 < 0.1$) was estimated by plotting r^2 values for pairs of SSR loci plotted as a function of map distances (cM).

Analysis of marker-trait associations

Marker-trait associations were calculated by MLM association test incorporating Q (structure relatedness) + K (kinship) matrices into TASSEL software package (Bradbury et al., 2007). Marker-trait associations were also confirmed by GLM association test incorporating Q matrices. For association mapping, the 5% MAF filtered SSR datasets were used. To assess significant marker-trait associations P -marker ≤ 0.001 was used.

Conclusion

From the findings of this study, it is concluded that D subgenome of cotton contains genomic regions involved in abiotic stress tolerance along with improved yield and fiber quality. It is an important finding with respect to the molecular breeding efforts for development of elite cotton varieties with improved abiotic stress tolerance characteristics in view of climate change paradigm. This study also elaborated that association mapping approach has strong potential to assess significant marker-trait associations by utilizing the commercial varieties which can save much time and cost as compared to traditional linkage mapping approach. MLM analysis can remove most of the false

positives and thus significance of marker-trait associations can be improved by incorporation of both MLM and GLM analyses in the association mapping approach.

Acknowledgements

We highly appreciate and acknowledge the financial assistance provided by the Higher Education Commission (HEC), Pakistan to conduct this research work.

References

- Abdurakhmonov IY, Kohel RJ, Yu JZ, Pepper AE, Abdullaev AA, Kushanov FN, Salakhutdinov IB, Buriev ZT, Saha S, Scheffler BE, Jenkins JN, Abdugarimov A (2008) Molecular diversity and association mapping of fiber quality traits in exotic *G. hirsutum* L. germplasm. *Genomics*. 92: 478-487.
- Abdurakhmonov IY, Saha S, Jenkins JN, Buriev ZT, Shermatov SE, Scheffler BE, Pepper AE, Yu JZ, Kohel RJ, Abdugarimov A (2009) Linkage disequilibrium based association mapping of fiber quality traits in *G. hirsutum* L. variety germplasm. *Genetica*. 136: 401-417.
- Ashraf M, Ahmad S (2000) Influence of sodium chloride on ion accumulation, yield components and fiber characteristics in salt-tolerant and salt-sensitive lines of cotton (*Gossypium hirsutum* L.). *Field Crops Res*. 66: 115-127.
- 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.
- Breseghele F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics*. 172: 1165-1177.
- 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 Genetics* 9: 16. doi: 10.1186/1471-2156-9-16.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 14: 2611-2620.
- Flowers TJ, Troke PF, Yeo AR (1977) Mechanisms of salt tolerance in halophytes. *Annu Rev Plant Physiol*. 28: 89-121.
- Fricano A, Bakaher N, Corvo MD, Piffanelli P, Donini P, Stella A, Ivanov NV, Pozzi C (2012) Molecular diversity, population structure, and linkage disequilibrium in a worldwide collection of tobacco (*Nicotiana tabacum* L.) germplasm. *BMC Genetics*. 13 (18) doi:10.1186/1471-2156-13-18
- Fusari CM, Lia VV, Hopp HE, Heinz RA, Paniego NB (2008) Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach. *BMC Plant Biol*. 8 (7) doi:10.1186/1471-2229-8-7
- Greenway H, Munns R (1980) Mechanisms of salt tolerance in nonhalophytes. *Annu Rev Plant Physiol*. 31: 149-190.
- Han ZG, Guo WZ, Song XL, Zhang TZ (2004) Genetic mapping of EST-derived microsatellites from the diploid *Gossypium arboreum* in allotetraploid cotton. *Mol Genet Genomics*. 272: 308-327.
- Han ZG, Wang CB, Song XL, Guo WZ, Gou JY, Li CH, Chen XY, Zhang TZ (2006) Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. *Theor Appl Genet*. 112: 430-439.
- Hao C, Wang Y, Hou J, Feuillet C, Balfourier F, Zhang X (2012) Association mapping and haplotype analysis of a 3.1-Mb genomic region involved in *Fusarium* head blight resistance on wheat chromosome 3BS. *PLoS One* 7, e46444. doi:10.1371/journal.pone.0046444.
- Ivancic V, Hackett CA, Nevo E, Keith R, Thomas WTB, Forster BP (2002) Analysis of simple sequence repeats (SSRs) in wild barley from the Fertile Crescent: associations with ecology, geography and flowering time. *Plant Mol Biol*. 48: 511-527.
- Ivancic V, Thomas WTB, Nevo E, Zhang Z, Forster BP (2003) Associations of simple sequence repeats with quantitative trait variation including biotic and abiotic stress tolerance in *Hordeum spontaneum*. *Plant Breeding*. 122: 300-304.
- Jiang CX, Wright RJ, El-Zik KM, Paterson AH (1998) Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). *Proc Natl Acad Sci USA*. 95: 4419-4424.
- Jorde LB (2000) Linkage disequilibrium and the search for complex disease genes. *Genome Res*. 10: 1435-1444.
- Khorsandi F, Anaghali A (2009) Reproductive compensation of cotton after salt stress relief at different growth stages. *J Agron Crop Sci*. 195: 278-283.
- Li L, Hao Z, Li X, Xie C, Li M, Zhang D, Weng J, Su Z, Liang X, Zhang S (2011) An analysis of the polymorphisms in a gene for being involved in drought tolerance in maize. *Genetica*. 139: 479-487.
- Lu Y, Xu J, Yuan Z, Hao Z, Xie C, Li X, Shah T, Lan H, Zhang S, Rong T, Xu Y (2012) Comparative LD mapping using single SNPs and haplotypes identifies QTL for plant height and biomass as secondary traits of drought tolerance in maize. *Mol Breeding*. 30: 407-418.
- Mohlke KL, Lange EM, Valle TT, Ghosh S, Magnuson VL, Silander K, Watanabe RM, Chines PS, Bergman RN, Tuomilehto J, Collins FS, Boehnke M (2001) Linkage disequilibrium between microsatellite markers extends beyond 1 cM on chromosome 20 in Finns. *Genome Res*. 11: 1221-1226.
- Munns R (2007) Utilizing genetic resources to enhance productivity of salt-prone land. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources* 2, No. 9.
- Nguyen TB, Giband M, Brottier P, Risterucci AM, Lacape JM (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. <http://www.biomedcentral.com/1471-2164/13/67>
- Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H, Bakker E, Calabrese P, Gladstone J, Goyall R, Jakobsson M, Kim S, Morozov Y, Padhukasahasram B, Plagnol V, Rosenberg NA, Shah C, Wall JD, Wang J, Zhao K, Kalbfleisch T, Schulz V, Kreitman M, Bergelson J (2005) The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol*. 3 (7), e196.
- Pasam RK, Sharma R, Malosetti M, van Eeuwijk FA, Haseneyer G, Kilian B, Graner A (2012) Genome-wide association studies for agronomical traits in a worldwide spring barley collection. *BMC Plant Biol*. 12, 16. doi:10.1186/1471-2229-12-16
- Paterson AH, Brubaker CL, Wendel JF (1993) A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol Biol*

- Rep. 11: 122-127.
- Pessaraki M, Tucker TC (1985) Uptake of nitrogen-15 by cotton under salt stress. *Soil Sci Soc Am J.* 49: 149–152.
- 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 through linkage and association. *Euphytica.* 187: 369–379.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multi-locus genotype data. *Genetics.* 155: 945–959.
- Pritchard JK, Wen W (2004) Documentation for Structure software. The University of Chicago Press, Chicago
- Reddy OUK, Pepper AE, Abdurakhmonov I, Saha S, Jenkins JN, Brooks T, Bolek Y, El-Zik KM (2001) New dinucleotide and trinucleotide microsatellite markers resources for cotton genome research. *J Cotton Sci.* 5: 103–113.
- Reif JC, Gowda M, Maurer HP, Longin CFH, Korzun V, Ebmeyer E, Bothe R, Pietsch C, Würschum T (2011) Association mapping for quality traits in soft winter wheat. *Theor Appl Genet.* 122: 961–970.
- Rezaeizad A, Wittkop B, Snowdon R, Hasan M, Mohammadi V, Zali A, Friedt W (2011) Identification of QTLs for phenolic compounds in oilseed rape (*Brassica napus* L.) by association mapping using SSR markers. *Euphytica.* 177: 335–342.
- Roy JK, Smith KP, Muehlbauer GJ, Chao S, Close TJ, Steffenson BJ (2010) Association mapping of spot blotch resistance in wild barley. *Mol Breeding.* 26: 243–256.
- Saeed M, Guo WZ, Ullah I, Tabbasam N, Zafar Y, Rahman M, Zhang TZ (2011) QTL mapping for physiology, yield and plant architecture traits in cotton (*Gossypium hirsutum* L.) grown under well-watered versus water-stress conditions. *Electron J Biotechnol* 14 (3). doi:10.2225/vol14-issue3-fulltext-3
- Saeed M, Dahab AHA, Guo WZ, Zhang TZ (2012) A cascade of recently discovered molecular mechanisms involved in abiotic stress tolerance of plants. *OMICS: J Integr Biol.* 16: 188-199.
- 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.
- 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 doi:10.1186/1471-2229-11-52
- United States Department of Agriculture (2012) Cotton: World Markets and Trade. <http://www.fas.usda.gov/cotton/Current/> (June 8, 2012)
- Weiss KM, Clark AG (2002) Linkage disequilibrium and mapping of human traits. *Trends Genet.* 18: 19–24.
- Wen W, Mei H, Feng F, Yu S, Huang Z, Wu J, Chen L, Xu X, Luo L (2009) Population structure and association mapping on chromosome 7 using a diverse panel of Chinese germplasm of rice (*Oryza sativa* L.). *Theor Appl Genet.* 119: 459–470.
- Witt SR, Buckler ES (2003) Using natural allelic diversity to evaluate gene function. *Methods Mol Biol.* 236: 123–139.
- Yan J, Shah T, Warburton ML, Buckler ES, McMullen MD, Crouch J (2009b) Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PLoS One* 4 (12): e8451. doi:10.1371/journal.pone.0008451
- Yan WG, Li Y, Agrama HA, Luo D, Gao F, Lu X, Ren G (2009a) Association mapping of stigma and spikelet characteristics in rice (*Oryza sativa* L.). *Mol Breeding* 24: 277–292.
- 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 J, Wu YT, Guo WZ, Zhang TZ (2000) Fast screening of SSR markers in cotton with PAGE/silver staining. *Cotton Sci Sin.* 12: 267-269.