GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Simple sequence repeat-based association analysis of fruit traits in eggplant (*Solanum melongena*)

H.Y. Ge[1], Y. Liu[1], J. Zhang[2], H.Q. Han[1], H.Z. Li[1], W.T. Shao[1] and H.Y. Chen[1]

[1]Department of Plant Science, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai, China
[2]Institute of Agricultural Science of Minhang District, Shanghai, China

Corresponding author: H.Y. Chen
E-mail: chhy@sjtu.edu.cn

Genet. Mol. Res. 12 (4): 5651-5663 (2013)
Received November 23, 2012
Accepted September 15, 2013
Published November 18, 2013
DOI http://dx.doi.org/10.4238/2013.November.18.14

ABSTRACT. Association mapping based on linkage disequilibrium (LD) provides a promising tool to identify quantitative trait loci (QTLs) in plant resources. A total of 141 eggplant (*Solanum melongena* L.) accessions were selected to detect simple sequence repeat (SSR) markers associated with nine fruit traits. Population structure analysis was performed with 105 SSR markers, which revealed that two subgroups were present in this population. LD analysis exhibited an extensive long-range LD of approximately 11 cM. A total of 49 marker associations related to eight phenotypic traits were identified to involve 24 different markers, although no association was found with the trait of fruit glossiness. To our knowledge, this is the 1st approach to use a genome-wide association study in eggplant with SSR markers. These results suggest that the association analysis approach could be a useful alternative to traditional linkage mapping to detect putative QTLs in eggplant.

Key words: SSR; Association mapping; Linkage disequilibrium (LD); Population structure; Eggplant

INTRODUCTION

Eggplant (*Solanum melongena* L.), a member of Solanaceae, is an economically and nutritionally important plant in many countries. Its fruit has a high content of antioxidant phenolic compounds that have been found to be beneficial to human health (Ge et al., 2011). Compared with the two *Solanum* model species, tomato and potato, eggplant has its unique traits with larger fruit size and different kinds of stress tolerances. Furthermore, it has a unique phylogenic aspect (Fukuoka et al., 2012). Despite the importance of this crop, studies of its molecular genetics fall behind those of tomato (*S. lycopersicum* L.), potato (*S. tuberosum* L.), and pepper (*Capsicum* spp L.), all of which belong to the Solanaceae family (Nunome et al., 2009; Fukuoka et al., 2012).

Molecular markers play a vital role in the enhancement of global food production and quantitative trait locus (QTL) mapping has been used as an important molecular tool for marker-assisted selection (MAS) in plant breeding. In recent years, several linkage maps (Doganlar et al., 2002b; Sunseri et al., 2003; Cao et al., 2006; Nunome et al., 2001, 2003, 2009; Wu et al., 2009; Barchi et al., 2010; Fukuoka et al., 2012) and fruit trait QTLs (Nunome et al., 2001; Doganlar et al., 2002a; Frary et al., 2003) have been reported in eggplant. However, most of the maps were mainly constructed with random amplified polymorphic DNA (RAPD) markers, restriction fragment length polymorphism (RFLP) markers, and amplified fragment length polymorphism (AFLP) markers (Doganlar et al., 2002b; Nunome et al., 2001, 2003; Sunseri et al., 2003; Cao et al., 2006; Barchi et al., 2010). The others were mainly constructed with conserved ortholog set II (cos II) markers (Wu et al., 2009), simple sequence repeat (SSR) markers (Nunome et al., 2009) and *Solanum* orthologous (SOL) markers (Fukuoka et al., 2012). The map constructed by Nunome et al. (2009) with 236 SSR markers contains the vast majority of SSR markers to date.

Association mapping, which is known as linkage disequilibrium (LD) mapping or association analysis, detects and locates QTLs based on the strength of the correlation between mapped genetic markers and traits (Flint-Garcia et al., 2003; MacKay and Powell, 2007). Association mapping has several advantages over traditional family-based linkage mapping, including shorter research time, higher mapping resolutions, and the ability to investigate a greater number of alleles (Flint-Garcia et al., 2005; Yu and Buckler, 2006; Yang et al., 2010). It was 1st applied to plant genetic mapping in maize (*Zea mays* L.) (Thornsberry et al., 2001), and LD mapping has now been conducted in many different plant species (Gupta et al., 2005), such as *Arabidopsis thaliana* Heynh. (Nordborg et al., 2002), soybean (*Glycine max* L. Merr.) (Wen et al., 2008; Hou et al., 2011), wheat (*Triticum* spp) (Maccaferri et al., 2005), barley (*Hordeum vulgare* L.) (Haseneyer et al., 2010), rice (*Oryza sativa* L.) (de Oliveira Borba et al., 2010), tomato (*S. lycopersicum* L.) (Mazzucato et al., 2008; Van Berloo et al., 2008), and potato (*S. tuberosum* L.) (D'hoop et al., 2008). However, association mapping has not been applied to eggplant genetic research to date. In 2009, Nunome et al. conducted an eggplant linkage map with a large number of SSR markers, which allowed for the detection of QTLs using association mapping with SSR markers. In this study, fruit traits were investigated within eggplant germplasms, and SSR markers closely related to these traits were determined by an association analysis.

MATERIAL AND METHODS

Plant material

A total of 141 eggplant accessions from the USA, India, Japan, Italy, Malaysia, the United Arab Emirates, Thailand, Korea, and China were used in this study. Of these, 128 accessions came from 22 different provinces in China, which largely represented the gene pool used by Chinese breeders. Details of accession names and origin are provided in Table 1.

Eggplant phenotype analysis

The eggplants were grown in a greenhouse in 2010 and 2011 in Shanghai, China. A total of nine fruit traits were scored: fruit weight (fw), fruit length (fl), fruit diameter (fd), fruit shape (fs), fruit calyx size (fcs), fruit anthocyanin presence (fap), fruit stripe (fst), fruit calyx prickle (ftcp), and fruit glossiness (fglo). The fw, fl, fd, fs, and fcs traits were evaluated in both growing years, and the other four traits were only examined in 2010. The survey of fruit traits followed the methods used by Frary et al. (2003) and Doganlar et al. (2002b). The fw was determined in grams for the five heaviest fruits. The fs was calculated as fl/fd. The fap was assessed in accordance with the national standards on eggplant germplasm from the Chinese Crop Germplasm Resources Information System (CGRIS) (http://icgr.caas.net.cn). This trait was scored on a scale from one to eight (one, white; two, whitish green; three, green; four, red; five, purplish red; six, purple; seven, dark purple; eight, other).

Marker analysis

The extraction of DNA followed the procedure by Ge et al. (2011). Based on their positions on the genetic map constructed by Nunome et al. (2009), 116 SSR markers were selected every 3 centimorgan (cM) (Figure 1). The markers were evenly distributed through the whole genome with 5-17 loci within each of the 12 linkage groups. Polymerase chain reaction (PCR) was carried out in a 10-μL reaction mixture containing 20 ng template DNA, 0.1 μM forward primer, 0.1 μM reverse primer, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 1 X PCR buffer, and 1 U *Taq* DNA polymerase (Promega, Shanghai, China). DNA amplification was performed in a 96-well thermocycler (Eppendorf AG 6321, Eppendorf, Hamburg, Germany), and cycles were carried out following the conditions of Nunome et al. (2009). The PCR products were then separated on a 5 to 8% polyacrylamide gel and visualized by silver staining (Ge et al., 2011). Each individual was genotyped at each locus by scoring the length of the amplified SSR band according to molecular ladders and the known allelic information.

Statistical analyses

Cluster analysis based on SSR markers was performed using the unweighted pair group method with arithmetic averages (UPGMA), and a dendrogram was constructed using the NTSYS-pc software, version 2.10t (Rohlf, 2000).

Table 1. A total list of the 141 eggplant genotypes used in this study.

Number	Name	Origin	Number	Name	Origin
YZ-3	YZ-3	The United Arab Emirates	79	Shouning Qie	Fujian, China
YZ-7	TG1	Thailand	80	Qing Qie	Hainan, China
5	96-2	China	81	Qingshan Chang Qie	China
6	Hong Qie	Shanghai, China	82	Qingniu Qie	China
7	Guangdong Qing Qie	Guangdong, China	83	Yu Qing Qie	Henan, China
8	Hong Qie	China	84	Nantong Qing Qie	Jiangsu, China
9	Malaysia Qie	Malaysia	86	Da Qing Qie	Jiangsu, China
10	Shenxian Yuan qie	Shandong, China	87	Qing Niutui Qie	China
12	Zi Hebao Qie	China	89	Hei Liu Tiao Qie	Heilongjiang, China
13	Hebao Qie	China	91	Fushe No.1	Hubei, China
14	Yaotou Qie	Jiangxi, China	92	Xin Cai Xian Qie	Henan, China
15	Hexian Qie	China	93	Qing Chang Qie	China
16	Rubai Qie	China	94	Xian Qie	Zhejiang, China
17	Wangbu Zi Chang Qie	Guangxi, China	95	Qingyang Qing Qie	Anhui, China
18	Yanzhi Qie	China	96	Lv Chang Qie	Shandong, China
19	Hanzhong Milk Qie	Shaanxi, China	97	Liao Qie No.1	Liaoning, China
20	Wugan Bai Qie	Hunan, China	98	Da Qing Qie	Shandong, China
21	Lueyang Milk Qie	Shaanxi, China	100	Er Hong Qie	Sichuan, China
22	Xiao Bai Qie	China	104	Hei Qie	Heilongjiang, China
23	Bai Milk Qie	China	105	Ai Gua	Jiangsu, China
24	Mei Qie	China	106	Chang Qie	Italy
25	Pinghu Bai Qie	Zhejiang, China	107	He Xian Qie	China
26	Pinghu Bai Chang Qie	Zhejiang, China	109	Jinan Zao Xiao Chang Qie	Shandong, China
27	Songjiang Qie	Shanghai, China	110	Diao Qie	Gansu, China
28	Pinghu Hong Qie	Zhejiang, China	111	Suzhou Niujiao Qie	Jiangsu, China
29	Shijiemei Qie	China	112	Min Chang Qie	Fujian, China
30	Guizhou Hong Qie	Guizhou, China	113	Yingzui Qie	Zhejiang, China
31	Haicheng Chang Qie	Liaoning, China	115	Lanzhou Chang Qie	Gansu, China
32	diFirenze	Italy	116	Xiao Hongpao	Shandong, China
33	Zhongsheng Zhen Hei	China	117	Niujiao Qie	Jiangsu, China
34	Hangzhou Hong Qie	Zhejiang, China	118	Zhusi Qie	Sichuan, China
35	Taiwan Qie	Taiwan, China	119	Dandong Zao Zi Qie	Liaoning, China
36	Mashu Chang Qie	China	120	Zi Tang Qie	Zhejiang, China
37	Xin Changqi	Japan	121	Xiangjiao Qie	Zhejiang, China
38	Zi Chang Qie	China	122	Shiqian Chang Qie	Guizhou, China
39	Min Qie No.1	Fujian, China	123	Benxi Zao Zi Qie	Liaoning, China
40	86-1	Heilongjiang, China	124	Longmenchi Zi Qie	Guangdong, China
41	Meiguo Da Chang Qie	USA	125	Wuzhua Xianbing Qie	Guangxi, China
42	Mo Chang Qie	China	126	Dian Qie	Yunnan, China
43	Zi Hei Chang Qie	China	127	Dian Qie No.2	Yunnan, China
45	Chang Qie	China	128	Jiajie Qie	China
46	Pearl Qie	China	129	Dazhong Qie	China
47	Xian Qie	China	131	Lv Qie	Yunnan, China
48	B75	China	132	Bai Qie	Yunnan, China
49	Hangzhou Tiao Qie	Zhejiang, China	133	Zi Qie	Yunnan, China
50	Weng'an Chang Qie	Guizhou, China	137	Lvziyuan Qie	Jiangsu, China
51	Changhong Zao Qie	Zhejiang, China	138	Diana Qie	Jiangsu, China
52	Qing Chang Qie	China	140	Baicuo Qie	Jiangsu, China
53	Liu Tiao Qie	Liaoning, China	141	Hu Qie	Shanghai, China
55	Eg-5	China	143	Dian Xian Bai Qie	Yunnan, China
56	Liangshui Qie	Jiangxi, China	YZ-1	Xi An lv Qie	Shaanxi, China
57	Zhu Qie	China	YZ-2	Mojiaolong	Jiangsu, China
58	Qing Yangjiao Qie	Fujian, China	YZ-4	Changguan Qie	Korea
59	Rong'an Chang Qing	Guangxi, China	YZ-5	YZ-5	Shandong, China
60	Mo qie	Sichuan, China	YZ-8	YZ-8	Guangdong, China
61	Sanyue Qie	Chongqing, China	YZ-9	YZ-9	Jiangsu, China
63	Pingdong Chang Qie	Taiwan, China	YZ-10	YZ-10	Henan, China
65	Arka Keshav	India	YZ-11	YZ-11	Shandong, China
66	Local-1	India	YZ-12	YZ-12	Henan, China
67	Ep 143	India	YZ-14	YZ-14	Hubei, China
68	Arka Nidhi	India	YZ-15	Heilong Qie	Jiangsu, China
69	Xian Chang Qie	Hubei, China	YZ-16	YZ-16	Liaoning, China

Continued on next page

Table 1. Continued.

Number	Name	Origin	Number	Name	Origin
70	Gouweiba Qie	Jiangsu, China	YZ-17	YZ-17	Shaanxi, China
71	Laiyang Qie	Shandong, China	YZ-18	Huangji Qie	Japan
72	Yushu Qie	China	YZ-19	YZ-19	Taiwan, China
73	Dunhe Qie	Guangdong, China	YZ-20	YZ-20	Jiangsu, China
74	Zao Er Hong Qie	Guangdong, China	YZ-21	YZ-21	China
75	Zhejiang Chang Qie	Zhejiang, China	YZ-22	YZ-22	Shandong, China
76	Chang Zi Qie	Guangdong, China	YZ-23	Yangzhou No.1	Jiangsu, China
77	Zao Zi Qie	China	YZ-24	YZ-24	Jiangsu, China
78	Chang Qie	Sichuan, China			

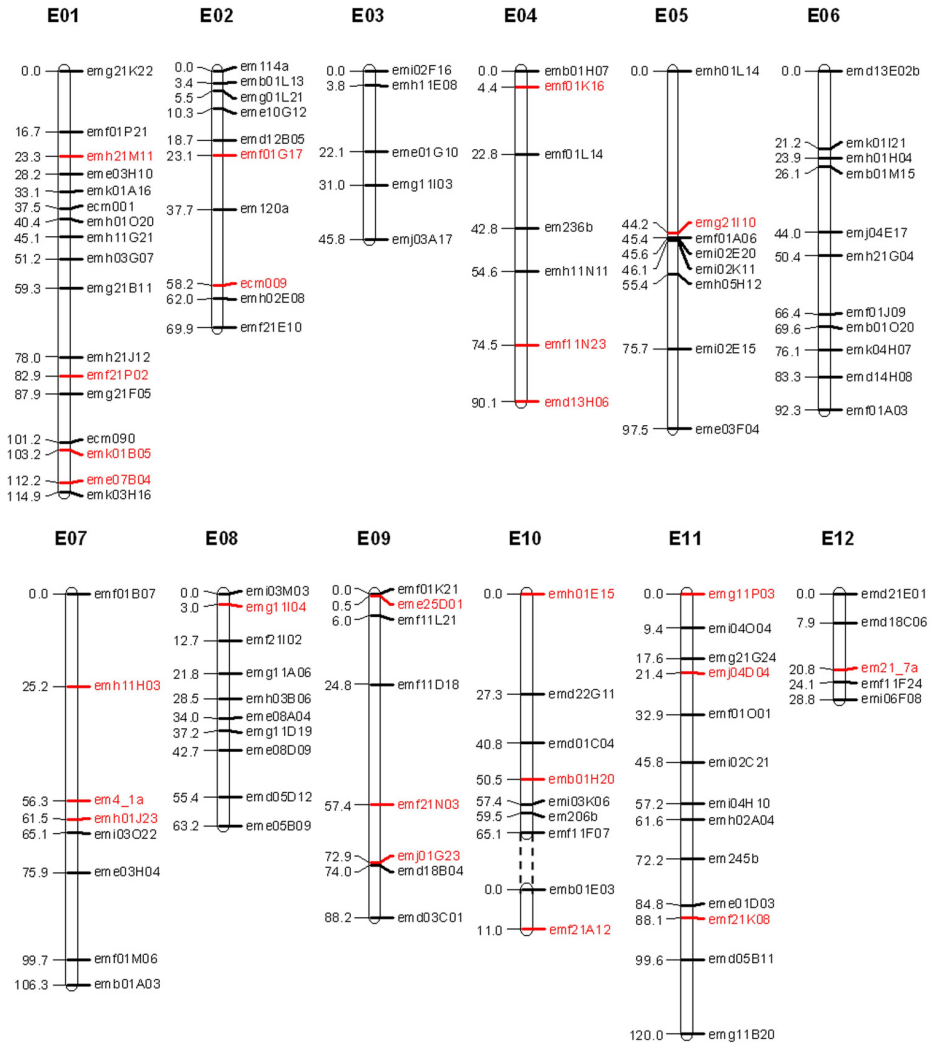

Figure 1. Eggplant SSR genetic linkage map showing the markers used in this study based on the linkage map by Nunome et al. (2009) and Fukuoka et al. (2012). Putative markers associated with fruit traits are indicated in red color.

The population structure was inferred through the software STRUCTURE 2.2 (http://pritch.bsd.uchicago.edu/software) (Pritchard et al., 2000) using the "admixture model" with a burn-in period of 10,000 followed by 100,000 iterations and a burn-in period of 100,000 followed by 1,000,000 iterations. Five independent runs were performed at each K level, which ranged from 1 to 15. The approach to maximum likelihood (Pritchard et al., 2000) and the calculation of ΔK by Evanno et al. (2005) were used to obtain the appropriate value for K.

LD analysis was performed by pairwise comparisons of the markers in the TASSEL 2.1 software (Bradbury et al., 2007). The pairs of loci were considered to be in significant LD if P was <0.05. LD decay was investigated by plotting r^2 against the marker distance in cM.

The putative associations between the markers and traits were calculated using the general linear model function in the TASSEL 2.1 software (Bradbury et al., 2007) with the population structure matrix (Q) as a cofactor. The P value calculated by TASSEL determined whether a QTL was associated with the marker.

RESULTS

In general, the traits with similar or related phenotypes showed significant positive correlations (Doganlar et al., 2002a). In this study, fw was positively correlated with fl ($r = 0.412$, $P \leq 0.01$) and fd ($r = 0.642$, $P \leq 0.01$). Strong correlations were also observed for fs with its components, fl ($r = 0.831$, $P \leq 0.01$) and fd ($r = -0.508$, $P \leq 0.01$). Fcs was negatively correlated with fl ($r = -0.780$, $P \leq 0.01$).

Of the 116 selected SSR markers, 11 markers showed monomorphisms in all of the genotypes studied. These 11 markers were excluded from subsequent study, and a final set of 105 SSR markers were selected for further analyses. A total of 373 alleles were detected by 105 pairs of SSR primers across 141 eggplant accessions, with the number of allelles per locus varying between 2 and 8 (mean 3.6). Polymorphic information content (PIC) values among the 141 cultivated types were calculated and varied from 0.014 to 0.657 (mean 0.30).

Genetic diversity and population structure

A dendrogram based on the similarity coefficients of the 141 accessions was constructed (Figure 2A). The dendrogram scale varied from 0.83 to 0.99, and the accessions clustered into two subgroups with distinct morphological characteristics. The 1st subgroup mostly contained the accessions with small, round fruits with a hard pulp, and members of this group were related to accessions that were less domesticated. The 2nd subgroup mostly represented varieties that have been cultivated extensively in recent years.

The structure of the population was considered to avoid false-positive associations. In order to complete this, two different burn-in periods were used and the likelihood values were obtained. True K was estimated in two ways; however, the distribution of LnP (D) did not show a clear trend (Figure 3A). ΔK was also calculated, and the result showed that ΔK reached the maximal value when K = 2 (Figure 3B). We also observed that the results from each of the two different burn-in periods were almost identical. The analysis of these data identified accessions into two subgroups as well, and the results were very similar to those of the clustering results (Figure 2B). The Q matrix outputs of the two subpopulations were used for the association analysis.

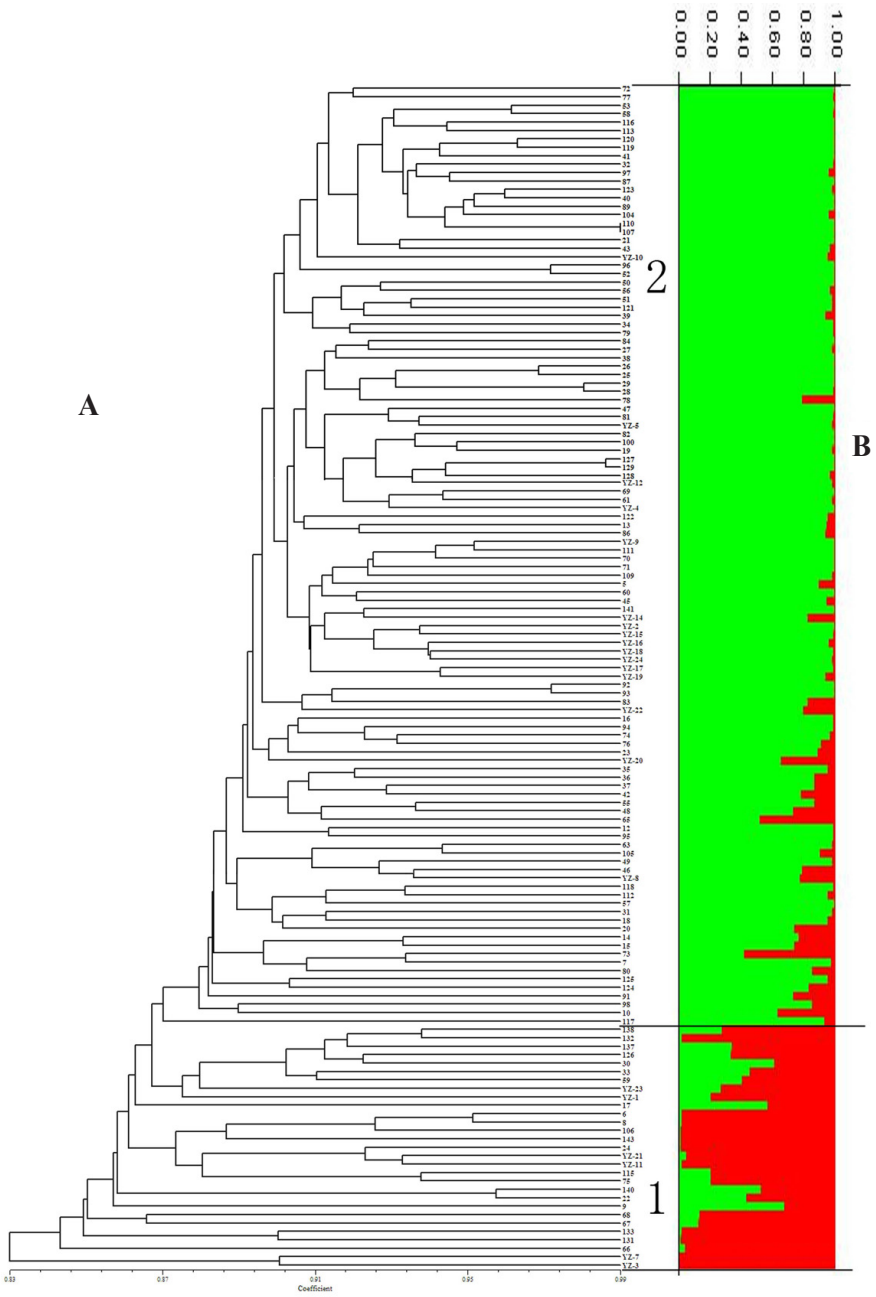

Figure 2. Comparison of a neighbor-joining tree and subpopulations detected by STRUCTURE 2.2. **A.** Dendrogram constructed based on Jaccard's similarity coefficient and UPGMA clustering. **B.** Each individual sample was represented by a single row broken into two-colored segments (green and red), with length proportions to each of the two inferred population subgroups. Each individual corresponded to the samples in the dendrogram. The samples are labeled with the codes listed in Table 1.

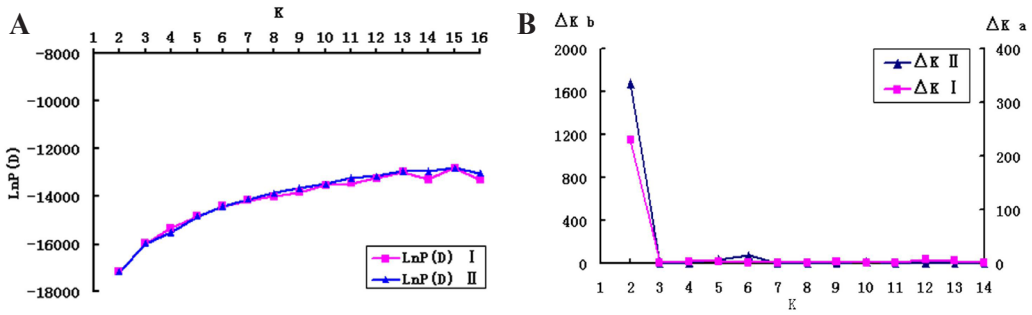

Figure 3. Comparison of two different burn-in lengths in estimation of the true population structure. **A.** Average logarithm of the probability of data likelihood, LnP(D) for 1 to 15 subpopulations (K). **B.** Values of ΔK, with its modal value detecting a true K of two groups (K = 2). I = burn-in period of 10,000 followed by 100,000 iterations; II = burn-in period of 100,000 followed by 1,000,000 iterations.

LD analyses

The pairwise LD between 105 SSR markers was estimated, and 5390 locus pairs were found. Of all of the identified locus pairs, 21.59% (1146/5309) were significantly in LD with an average r^2 value of 0.039. Of the 1146 assessed locus pairs, 23.39% (268) had r^2 levels above 0.05, and 8.1% (93) belonged to an intrachromosomal LD.

The scatter plot of the LD based on the r^2 values for the 141 genotypes was constructed, and it showed that the r^2 values decline with distance: r^2 values were down to 0.2 with approximately 5.5 cM and 0.1 with approximately 11 cM (Figure 4).

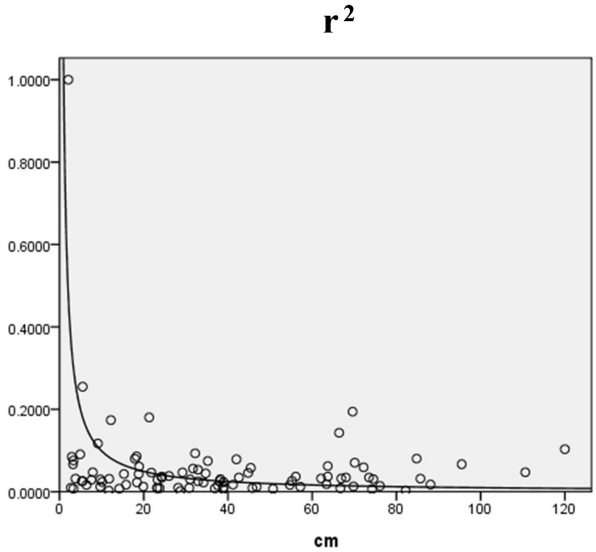

Figure 4. Linkage disequilibrium decay plot. The squared correlation (r^2) between paired marker intensities on the y-axis was plotted against the distance between pairs of markers in centiMorgan (cM). The line represents the fitted Loess curve.

Trait-marker associations

A total of 49 marker associations related to eight phenotypic traits were identified to involve 24 markers at the level of 0.01 (Table 2). The fd trait in 2011 and fglo trait did not identify any marker associations. The percentage of the total variation explained by each marker ranged from 4.5 (em4_1 associated with fw) to 22.8% (emf21A12 associated with fst). Among the 49 significant SSR-trait associations, six were found for fw, 10 for fl, four for fd, 10 for fs, seven for fcs, two for fap, seven for fst, and three for ftcp. The marker associations related to the eight phenotypic traits were located on 11 chromosomes, and no marker associations were found on E3 (Table 3). Associations related to fw, fl, fd, fs, and fcs were evaluated for 2 years. Seventeen markers were detected in 2010, 20 markers were found in 2011, and nine were found in both years. Among the 24 markers that were found to be associated with the fruit phenotypes studied, 12 markers were shared by two or more fruit traits.

Table 2. Marker loci associated with fruit traits ($P \leq 0.01$) and their explained phenotypic variation.

Marker	Location	fw		fl		fd	fs		fcs		fap	fst	ftcp
		2010	2011	2010	2011		2010	2011	2010	2011			
eme25D01	E9	0.098	0.154										
em4_1	E7	0.045	0.094			0.063							
eme07B04	E1		0.117										
emf01G17	E2		0.083	0.06								0.082	
emf01K16	E4			0.102			0.11	0.079					
emg21I10	E5			0.098	0.091		0.087	0.108	0.097		0.085		
emb01H20	E10			0.115	0.1		0.134	0.101	0.106				
emd13H06	E4			0.092	0.086			0.073					
emh21M11	E1				0.064			0.081	0.064	0.065			
emg11P03	E11				0.064					0.07			
emf11N23	E4					0.115						0.156	
emh01E15	E10					0.109		0.129					
em21_7	E12					0.118							
emf21K08	E11							0.087					
emk01B05	E1									0.11			
ecm009	E2									0.072			
emf21P02	E1										0.117		0.11
emf21A12	E10											0.228	
emj01G23	E9											0.154	
emg11I04	E8											0.142	
emf21N03	E9											0.117	
emj04D04	E11											0.062	
emh11H03	E7												0.131
emh01J23	E7												0.113

fw = fruit weight; fe = fruit length; fd = fruit diameter; fs = fruit shape; fcs = fruit calyx size; fap = fruit anthocyanin presence; fst = fruit stripe; ftcp = fruit calyx prickle

Table 3. Fruit trait QTLs found on each chromosome from different studies in eggplant.

References	Fruit traits	Previous studies	Present study
Doganlar et al. (2002)	Fruit weight (fw)	E2, E9, E11	E1, E2, E7, E9
	Fruit length (fl)	E2, E9, E11	E1, E2, E4, E5, E10, E11
	Fruit diameter (fd)	E1, E11	E4, E7, E10, E12
	Fruit shape (fs)	E2, E7	E1, E4, E5, E10, E11
	Fruit anthocyanin presence (fap)	E1, E8, E10, E12	E1, E5
	Fruit calyx prickle (ftcp)	E6, E9, E11	E1, E7
	Fruit stripe (fst)	E4, E10	E2, E4, E8, E9, E10, E11
Frary et al. (2003)	Fruit calyx size (fcs)	E2, E9	E1, E2, E5, E10, E11

DISCUSSION

Population structure

A prerequisite for association studies is a good estimation of the true population structure. In order to get accurate estimates of the parameters, two different length runs were conducted. The results demonstrated that a burn-in of 10,000-100,000 was adequate for this study; however, the burn-in of 100,000-1,000,000 was more accurate with lower variations between independent runs. In either case, both of the methods resulted in two subgroups, which was highly similar to the neighbor-joining dendrogram and morphological characteristics.

The association mapping that considered population structure had a lower false-positive rate (Yu et al., 2006). Studies have shown that association mapping that considered Q had fewer trait-marker associations and a different interpretation ratio to the phenotypic variation of the associated SSR loci (Wen et al., 2008; Hou et al., 2011). In this study, putative associations between the 105 SSR markers and fruit traits were calculated using the Q matrix of K = 2 as a cofactor.

Linkage disequilibrium analyses

For accurate association mapping based on LD, the LD in the genotypes was detected using the selected SSR markers. In this study, a certain degree of LD was detected in both syntenic markers and unlinked markers, and similar LD levels have been observed in soybean (Wen et al., 2008; Hou et al., 2011). Applying the commonly used reference value of 0.1 for r^2, we found that LD exhibits extensive long-range LD, approximately 11 cM. In tomato, stronger LD had been studied with an average extent of 15 cM (Van Berloo et al., 2008), and in potato, the LD decay was reported to drop from 0.3 to 10 cM (D'hoop et al., 2008). The difference in LD observed across species is a result of the interplay of many factors such as the recombination rate, mating system, selection, effective population size, and population structure (Rafalski and Morgante, 2004). LD differs between selfing and outcrossing species (Flint-Garcia et al., 2003). Eggplant is normally a highly self-pollinated crop, and it has been reported that the extent of natural outcrossing was from 0 to 8.2% in Asia (Chen, 2001). Tomato has a natural outcrossing rate that is less than 4% (Wehner, 1999), and potato is an outcrossing crop. The LD extent of the three Solanaceae crops was in accord with their mating system.

The number of markers needed to cover a genome is determined by the extent of LD. For example, the *Arabidopsis* genome may require 2000 markers, diverse maize landraces may require 750,000 markers, and elite maize lines may require only 50,000 markers (Flint-Garcia et al., 2003). Jun et al. (2008) considered that 150-300 markers were adequate to conduct a preliminary whole genome association study in soybean. In this study, 105 SSR markers were selected for an association analysis. Therefore, it was a preliminary study for the whole genome association mapping in eggplant and has laid the foundation for more thorough association mapping studies in the future.

Trait-marker associations

Tests for associations between the SSR markers using general linear methodology

revealed significant associations between fruit traits and markers in the germplasms studied here. The 49 associations related to eight phenotypic traits were identified to involve 40 SSR loci. Of those, nine loci that were associated with fw, fl, fs, and fcs were detected both years of our observations. The results indicated that these nine loci may be stably related to the traits. Of all of the studied markers, 12 were shared by two or more traits, which might be caused by the pleiotropic effects of linked genomic regions.

The identification of QTLs for eggplant fruit traits based on linkage mapping has been done in recent years. Nunome et al. (2001) reported two QTLs for fs and two QTLs for fruit color. However, the linkage map was constructed with RAPD and AFLP markers, and the linkage groups did not correspond to any chromosome. Doganlar et al. (2002b) and Frary et al. (2003) mapped several fruit traits to eggplant chromosomes. According to the recently constructed map by Fukuoka et al. (2012), we located the associated loci on each chromosome and compared them to the QTLs that were identified by the linkage analysis (Table 3). For example, observations of associations related to fcs on E2 and fruit stripe on E4 and E10 agreed with the those from linkage mapping studies (Frary et al., 2003).

Anthocyanin pigments had been studied in several solanaceous species. The genomes of tomato, potato, pepper, and eggplant have been shown, with the exception of several chromosomal rearrangements, to share extensive colinearity of gene order (De Jong et al., 2004). De Jong et al. (2004) suggested that a similar location of anthocyanin regulatory loci mapped to chromosome 10 had been subjected to parallel selection in the domestication of many solanaceous crops, such as tomato, pepper, eggplant, and potato. Doganlar et al. (2002b) reported that the anthocyanin locus *fap10.1* was located on chromosome E10, and the locus corresponded to the *af* locus on tomato chromosome T5 and the *af* locus on T10. Studies reported that several QTLs for anthocyanin such as *af*, *CHS3* (Tanksley et al., 1992), *chi* (De Jong et al., 2004), and *ec5.1* (Frary et al., 2004) were located on chromosome 5 in tomato. In this study, one *fap* QTL on E5 was detected, which may correspond to the relevant QTLs on tomato chromosome T5.

Fw and fs were very important traits in solanaceous species. In eggplant, fw QTLs were mapped on chromosome E2, E9, and E11 by Doganlar et al. (2002b), whereas our results indicated that four markers located on E1, E2, E7, and E9 were associated with fw. Thus, eme07B04 on E1 and em4_1 on E7 in this association analysis could be linked to novel QTLs for fw. In tomato, many QTLs for fw and fs were mapped to different chromosomes. Mazzucato et al. (2008) located fw QTLs on T1 and T12 and fs QTLs on T1, T7, T8, and T12. Grandillo et al. (1999) also mapped fw QTLs to T1 and T3.

In summary, we studied eggplant LD and detected 49 marker associations that were related to eight fruit traits in eggplant accessions with SSR markers. Because we analyzed a relatively small number of markers for 141 genotypes, we consider that the associations stated here are preliminary and require confirmation with additional markers or conventional QTL mapping. However, to our knowledge, this is the 1st report of a genome-wide association study in eggplant, and it provides a useful way of identifying candidate loci for important characteristics. Association mapping can be an effective method for QTL mapping and can help breeders to develop new strategies to improve plant productivity and quality.

ACKNOWLEDGMENTS

The authors thank Dingshi Zha at Shanghai Academy of Agricultural Sciences for

providing some materials. Research supported partially by the Shanghai Leading Academic Discipline Project (#B209) and the National High Technology Research and Development Program of China ("863" Program) (#2012AA100100). The authors thank the editor and the reviewers for their relevant comments on the manuscript.

REFERENCES

Barchi L, Lanteri S, Portis E, Stagel A, et al. (2010). Segregation distortion and linkage analysis in eggplant (*Solanum melongena* L.). *Genome* 53: 805-815.

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

Cao BH, Lei JJ, Sun XD, Chen GJ, et al. (2006). Construction of RAPD markers linkage map for eggplant. *Acta Hortic. Sin.* 33: 1092.

Chen NC (2001). Eggplant Seed Production. AVRDC Training Guide, Shanhua.

D'hoop BB, Paulo MJ, Mank RA, van Eck HJ, et al. (2008). Association mapping of quality traits in potato (*Solanum tuberosum* L.). *Euphytica* 161: 47-60.

De Jong WS, Eannetta NT, De Jong DM and Bodis M (2004). Candidate gene analysis of anthocyanin pigmentation loci in the Solanaceae. *Theor. Appl. Genet.* 108: 423-432.

de Oliveira Borba TC, Brondani RP, Breseghello F, Coelho AS, et al. (2010). Association mapping for yield and grain quality traits in rice (*Oryza sativa* L.). *Genet. Mol. Biol.* 33: 515-524.

Doganlar S, Frary A, Daunay MC, Lester RN, et al. (2002a). Conservation of gene function in the solanaceae as revealed by comparative mapping of domestication traits in eggplant. *Genetics* 161: 1713-1726.

Doganlar S, Frary A, Daunay MC, Lester RN, et al. (2002b). A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the Solanaceae. *Genetics* 161: 1697-1711.

Evanno G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620.

Flint-Garcia SA, Thornsberry JM and Buckler ES (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54: 357-374.

Flint-Garcia SA, Thuillet AC, Yu J, Pressoir G, et al. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* 44: 1054-1064.

Frary A, Doganlar S, Daunay MC and Tanksley SD (2003). QTL analysis of morphological traits in eggplant and implications for conservation of gene function during evolution of solanaceous species. *Theor. Appl. Genet.* 107: 359-370.

Frary A, Fulton TM, Zamir D and Tanksley SD (2004). Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *L. pennellii* cross and identification of possible orthologs in the Solanaceae. *Theor. Appl. Genet.* 108: 485-496.

Fukuoka H, Miyatake K, Nunome T, Negoro S, et al. (2012). Development of gene-based markers and construction of an integrated linkage map in eggplant by using *Solanum* orthologous (SOL) gene sets. *Theor. Appl. Genet.* 125: 47-56.

Ge H, Li H, Liu Y, Li X, et al. (2011). Characterization of novel developed expressed sequence tag (EST)-derived simple sequence repeat (SSR) markers and their application in diversity analysis of eggplant. *Afr. J. Biotechnol.* 10: 9023-9031.

Grandillo S, Ku HM and Tanksley SD (1999). Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theor. Appl. Genet.* 99: 978-987.

Gupta PK, Rustgi S and Kulwal PL (2005). Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Mol. Biol.* 57: 461-485.

Haseneyer G, Stracke S, Paul C, Einfeldt C, et al. (2010). Population structure and phenotypic variation of a spring barley world collection set up for association studies. *Plant Breed.* 129: 271-279.

Hou J, Wang C, Hong X, Zhao J, et al. (2011). Association analysis of vegetable soybean quality traits with SSR markers. *Plant Breed.* 130: 444-449.

Jun TH, Van K, Kim MY, Lee SH, et al. (2008). Association analysis using SSR markers to find QTL for seed protein content in soybean. *Euphytica* 162: 179-191.

Maccaferri M, Sanguineti MC, Noli E and Tuberosa R (2005). Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Mol. Breed.* 15: 271-290.

MacKay I and Powell W (2007). Methods for linkage disequilibrium mapping in crops. *Trends Plant Sci.* 12: 57-63.

Mazzucato A, Papa R, Bitocchi E, Mosconi P, et al. (2008). Genetic diversity, structure and marker-trait associations in a

collection of Italian tomato (*Solanum lycopersicum* L.) landraces. *Theor. Appl. Genet.* 116: 657-669.
Nordborg M, Borevitz JO, Bergelson J, Berry CC, et al. (2002). The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nat. Genet.* 30: 190-193.
Nunome T, Ishiguro K, Yoshida T and Hirai M (2001). Mapping of fruit shape and color development traits in eggplant (*Solanum melongena* L.) based on RAPD and AFLP markers. *Breed. Sci.* 51: 19-26.
Nunome T, Suwabe K, Iketani H, Hirai M, et al. (2003). Identification and characterization of microsatellites in eggplant. *Plant Breed.* 122: 256-262.
Nunome T, Negoro S, Kono I, Kanamori H, et al. (2009). Development of SSR markers derived from SSR-enriched genomic library of eggplant (*Solanum melongena* L.). *Theor. Appl. Genet.* 119: 1143-1153.
Pritchard JK, Stephens M, Rosenberg NA and Donnelly P (2000). Association mapping in structured populations. *Am. J. Hum. Genet.* 67: 170-181.
Rafalski A and Morgante M (2004). Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet.* 20: 103-111.
Rohlf FJ (2000). NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System. Version 2.1. Exeter Software, Setauket.
Sunseri F, Sciancalepore A, Martelli G, Rotino GL, et al. (2003). Development of RAPD-AFLP map of eggplant and improvement of tolerance to *Verticillium* wilt. *Acta Hortic.* 625: 107-115.
Thornsberry JM, Goodman MM, Doebley J, Kresovich S, et al. (2001). *Dwarf8* polymorphisms associate with variation in flowering time. *Nat. Genet.* 28: 286-289.
Van Berloo R, Zhu A, Ursem R, Verbakel H, et al. (2008). Diversity and linkage disequilibrium analysis within a selected set of cultivated tomatoes. *Theor. Appl. Genet.* 117: 89-101.
Wehner TC (1999). Heterosis in Vegetable Crops. ASA-CSSA-SSSA, Madison.
Wen ZX, Zhao TJ, Zheng YZ, Liu SH, et al. (2008). Association analysis of agronomic and quality traits with SSR markers in *Glycine max* and *Glycine soja* in China: I. population structure and associated markers. *Acta. Agron. Sin.* 34: 1169-1178.
Wu F, Eannetta NT, Xu Y and Tanksley SD (2009). A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. *Theor. Appl. Genet.* 118: 927-935.
Yang X, Yan J, Shah T, Warburton ML, et al. (2010). Genetic analysis and characterization of a new maize association mapping panel for quantitative trait loci dissection. *Theor. Appl. Genet.* 121: 417-431.
Yu J and Buckler ES (2006). Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* 17: 155-160.
Yu J, Pressoir G, Briggs WH, Vroh B, I, et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38: 203-208.