

Genetic analysis of microsatellite markers for salt stress in two contrasting maize parental lines and their RIL population

Ayşen Yumurtacı^{1*}, Hülya Sipahi², Li Zhao³

¹ Marmara University, Faculty of Science and Letters, Department of Biology, 34722, Istanbul, Turkey

² Sinop University, Faculty of the Arts and Sciences, Department of Biology, 57000, Sinop, Turkey

³ The State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China

Abstract – Salt stress considerably hinders the growth and productivity of maize (*Zea mays* L.). Identification of salt tolerant genotypes and integration of alternative molecular markers have important roles in enhancing breeding processes. In this study, 3308 maize expressed sequence tags (ESTs) from salt stress-related libraries were assembled to mine repetitive sequences for development of applicable markers. In this core EST data, 208 simple and 18 non-simple repetitive regions were detected in 312 contigs and 1121 singletons. The di-nucleotide repeats were the most abundant type and accounted for 79.3%, followed by tri (19.7%), and tetra-nucleotide (1%). Among 59 EST-simple sequence repeats (SSRs), a total of 55 were screened for polymorphism between F35 (salt sensitive) and F63 (salt tolerant) parents and 48 out of 55 were detected as monomorphic. Significantly, seven of them (12.7%) were found to be polymorphic and were used for genotyping of 158 F₅ derived recombinant inbred maize lines, and four of them were located on chromosome 1 and 3. Using *in silico* mapping, 44 out of 59 EST-SSR markers were mapped on 10 maize chromosomes. Analysis of sequence homology revealed different functional groups such as: membrane transport, cell defense, cell division, signaling components, photosynthesis and cell metabolism. These EST-SSRs might be used as new functional molecular markers in the diversity analysis, identification of quantitative trait loci (QTLs) and comparative genomic studies in maize in the future.

Keywords: EST, maize molecular markers, salinity, SSR

Introduction

Soil salinization is an important external stress factor for plants except halophytes, and has proved to be a serious drawback for plant growth from seed germination to adult plant stage (Singh 2015). Nearly 50% of the world's territorial areas may be under pressure from salinity in the next forty years, due to the rapid shifts that are occurring in global weather conditions (Bray et al. 2000). Therefore, improvement of salt tolerance in plants has gained importance because of the need to provide sufficient food for the world's increasing population (Bita and Gerats 2013, Yumurtacı 2015).

Maize (*Zea mays* L.) is the third most cultivated crop after wheat and provides many commercial benefits for clean energy production, food and livestock feed (Klopfenstein et al. 2013). Since detrimental declines in yield and agro-morphological traits have seen under saline condi-

tions, maize is accepted as a salt sensitive crop (Schubert et al. 2001). In this perspective, investigation of the genetic control and identification of genome regions associated with salt tolerance are of great significance.

Implementation of various DNA markers and quantitative trait loci (QTL) mapping techniques have contributed to an improved knowledge of the genetic bases of agriculturally significant traits and assisted the progress of plant breeding (Lee 2007, Xue et al. 2010). The development and mapping of DNA markers are essential for using QTL mapping of salt tolerance and marker-assisted selection of this trait in maize genotypes. Maize is one of the first major crop species to have had maps of different molecular markers (Davis et al. 1999, Sharopova et al. 2002, Sibov et al. 2003, Lima et al. 2009, Xu et al. 2013, Zhou et al. 2011). Recently, research has focused on genetic dissection in maize using QTL mapping of recombinant inbred lines

* Corresponding author, e-mail: yumurtaciay@yahoo.com

(RILs) derived from a cross between two maize inbred lines with contrasting salinity tolerance. A high density genetic map based on single nucleotide polymorphism (SNP) markers identified 20 QTLs on seven maize chromosomes (Cui et al. 2015). Significantly, three QTLs had only additive effects, while 12 had both additive and additive x treatment interaction effects. These QTLs were mainly clustered on maize chromosomes 1, 3 and 5. The five unconditional and three conditional QTLs could individually explain more than 20% of the phenotypic variation. Xiang et al. (2015) identified six sequence-related amplified polymorphism (SRAP) markers linked to salt tolerance using bulk segregant analysis of DNA pools from two salt-tolerant and salt-sensitive maize genotypes.

Compared with other types of molecular markers, simple sequence repeats (SSRs) have many advantages because of their multi-allelic nature and co-dominant inheritance, as well as simple and inexpensive developmental methodology. Expressed sequence tag SSRs (EST-SSRs) are derived from expressed genome parts, which are more evolutionarily conserved than non-coding sequences and therefore have transferability. Also, EST-SSRs have close associations with gene/QTL regions that offer a practical strategy for molecular plant breeding. Thus, EST mining can be accepted as a reconstruction of genome-scale analyses. Many maize ESTs contained simple sequence repeats and could be readily converted to functional markers (Lee 2007). In recent years, a number of maize EST-SSRs have been developed (Xu et al. 2013) and applied to construct linkage maps (Sharopova et al. 2002, Zhou et al. 2011, Orsini et al. 2012, Sa et al. 2012), QTL mapping and marker-assisted breeding (Orsini et al. 2012, Cui et al. 2015). In addition, Banisetti et al. (2012) identified candidate gene-based SSR markers for lysine, tryptophan pathway and opaque2 modifiers. A total of twenty-four SSR loci were developed and found to be useful markers for fine mapping and high-density mapping of opaque2 modifiers. Despite an increasing number of DNA markers, the identification of comprehensive markers for screening of salinity specific regions in maize is lagging behind and needs to be further developed.

In the present study, motif structures and density estimations of SSR regions were identified through *in silico* analysis of publicly available maize EST libraries which were constructed under salt stress at seedling stage. A set of primer pairs flanking repetitive regions was developed from these analyses and validated using two contrasting parental genotypes and their recombinant inbred lines (F35-salt sensitive and F63-salt tolerant) for the genetic basis of salt tolerance. Also, relative *in silico* map positions were defined and functional annotations of these SSRs were carried out to detect the functional marker source for maize.

Materials and methods

Plant material

Maize parental lines F63 (salt tolerant) and F35 (salt sensitive) and their segregating population of 158 F₅ RILs were used in this study. These materials were previously

developed by Cui et al. (2015) in the State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China.

SSR detection

Two publicly available cDNA libraries, constructed by Wang/Bohnert lab under salinity stress in maize B73 line, were used; library ZM14 from root and library ZM13 from shoot tissue. This study was performed at four main computational steps. At first, a total of 3308 EST sequences were downloaded in fasta format from NCBI GenBank (<http://www.ncbi.nlm.nih.gov>). For scanning vector contaminated sites in ESTs, NCBI VecScreen module and DNADragon (<http://www.sequentix.de/download/dnadragon.zip>) were used. After trimming of contaminated sites, the DNADragon software “sequence assemble” module was used to cluster all ESTs into condensed non-redundant groups. Non-clustered (singletons) and clustered ESTs (contigs) were analyzed on eTRA 1.0 software, developed by Bilgen et al. (2004) and Karaca et al. (2005), under the following parameters: up and down filtering was selected as 5%, minimum and maximum motif length was 2 and 10, respectively. Repeat index and repeat percentages were automatically calculated on eTRA 1.0 software.

PCR conditions

Genomic DNA was extracted from young leaves of each RIL using cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1987). Polymerase chain reaction (PCR) was conducted in 15 µl volume containing 5 ng of genomic DNA template, final concentrations of 0.25 pM of both forward and reverse primers, 1.5 mM of MgCl₂, 0.2 mM of dNTPs, 1× PCR buffer, and 0.1 U of Taq polymerase enzymes. The cycling conditions for PCR were used to an initial denaturation of 2 min at 95 °C, followed by 33 cycles of 94 °C for 30 sec, annealing temperatures at 59 °C (for ZM13-AI977889, ZM14-Contig119), 56 °C (for the remaining primers) for 30 sec, and extension at 72 °C for 30 sec. This was followed by final extension of 7 min at 72 °C. PCR products were separated electrophoretically on polyacrylamide gels (12%) for detailed fragment separation and results were scored manually. A total of 55 primers were generated and used in the validation studies.

Data analysis and linkage map construction

Salt tolerant (F63) and sensitive (F35) parents were screened with newly developed markers. Subsequently, polymorphic ones were also tested in a segregating population of 158 F₅ RILs derived from a cross between F63 and F35. Chi-square analysis was performed at a significance threshold of 5% in order to detect deviations from the expected Mendelian segregation ratio of 1:1. Linkage analysis was performed using the program QTL IciMapping 3.3 (Li et al. 2007, 2008). The linkage groups were created with LOD threshold of 3.0. Map distances (centimorgan) were calculated according to the Kosambi mapping function (Kosambi 1943).

Chromosomal localization and annotation studies of SSRs

For nucleotide-protein similarities in contig and singleton groups, Blastx (<http://www.ncbi.nlm.nih.gov>) search was retained for the best hit match score with the following criteria; E-value<10⁻⁵ and 90% accepted as for minimum identity. EST-based three hundred and twenty-six SSR contained sequences were aligned with several linkage maps of maize genomes, using e-PCR module (Rotmistrovsky et al. 2004) under default stringency parameters (<http://www.ncbi.nlm.nih.gov/projects/e-pcr/>). An *in silico* consensus map was drawn on MapChart 2.0 software according to Voorrips (2002).

Results

Based on the *in silico* analysis, 3308 ESTs consisting of 477119 base pairs were analyzed to find SSR loci. Redundancy profiles revealed 312 contigs and 1121 singletons. Average lengths of contig sequences were 714 bp and 632 bp for root and shoot tissues, respectively, but these values decreased to 483bp and 373bp for singleton sequences in root and shoot tissue. Among contig sequences, repeat index was 0.314 and 0.266 for shoot and root tissue, respectively. Additionally, singleton sequences had lower repeat indices for shoot (0.158) and root (0.156) tissues. Overall statistics for all clustered sequences indicated a repeat index as low as 0.186.

Here, 226 repeats, 208 simple and 18 non-simple, were detected in a total of 1433 contigs and singletons. Di-nucleotide repeats were the most abundant type and accounted for 79.3%, followed by tri- (19.7%), and tetra-nucleotide (1%). The vast majority of Class I type dinucleotide repeats were found in shoot tissue, in the percentage of 84.9% and root tissue specific Class I type trinucleotide repeats (30.1%) marked as a higher level than shoot (13.5%) (Tab. 1). Among di-nucleotide motifs, AA (18.75%) and TT (13%) had the highest frequency, while GG had the lowest frequency, 7.2%. Thirty different tri-nucleotide motif repeats were identified. The most frequent tri-nucleotide motif was CCG (7.3%). Some tri-nucleotide repeats (CTC, CTG, CTT, TTA) that are responsible for leucine synthesis were found only in root tissue-specific sequences. Only two tetra-nucleotide motifs (ATAA, ACCA) specific to shoot tissue were identified. A total of 55 primer pairs were designed from 59 EST-SSR containing regions (Tabs. 2, 3). Of the 55 tested maize EST-SSRs, 48 generated amplification of monomorphic products and 7 others (12.7%) produced clear and consistent polymorphic banding patterns between the parental lines F35 and F63. These seven polymorphic markers (ZM14-Contig119, ZM13-AI964577, ZM13-AI977889, ZM13-Contig37, ZM13-Contig83, ZM14-AI855336, ZM13-AI966933) were also genotyped across the 158 individuals of the RIL population; four of them were assigned to two maize chromosomes. The loci ZM14Contig119 and ZM13AI964577 were located on the same

Tab. 1. Distribution of root and shoot tissue specific expressed sequence tag (EST) based simple sequence repeats (SSR) and sequence tagged site (STS) marker densities and number of ClassI and ClassII type repeats.

EST	Shoot tissue		Root tissue	
	Number of contigs	Number of singletons	Number of contigs	Number of singletons
Total	188	647	124	474
Only SSR included	46	90	25	65
Only STS included	40	66	32	69
Both (SSR + STS)	14	18	9	18
ClassI (di/tri/tetra repeats)	11/2/2		4/2/-	
ClassII (di/tri/tetra repeats)	95/15/-		54/23/-	

Tab. 2. Detailed blastx similarity matches and motifs with repeat types of both simple sequence repeats (SSR) and sequence tagged site STS contained sequences in maize shoot tissues. “P” indicates perfect repeats, “IP” imperfect repeats and “*” shows unmapped markers. ZM13 is the library name.

Sequence ID	Motifs/Repeat Type	Organism	Protein	E-value	Primer pairs (5'-3')	Product size (bp)
ZM13-Contig13	(TA) ₆ /P	<i>Z. mays</i>	Glyceraldehyde-3-dehydrogenase	2e-91	CAGATTATTCGACGAAAGAGA GCATGGTTGATGAAACAAATA	148
ZM13-Contig19	(GAT) ₆ /P	<i>Z. mays</i>	Hypothetical protein	8e-36	CAACCAAGTCCTAAATTGTCA TTAAGCAGAGCTCAAAAAC TG	153
ZM13-Contig33*	(CA) ₅ (GT) ₅ /P-(GT) ₃ (TA) ₅ IP	<i>Z. mays</i>	Jacalin like lectin protein	3e-100	AGGTTTCGACGAGAGCTTGAC CGTTTTGTTCGTGCTGAAGA	800
ZM13-Contig37	(TC) ₁₀ /P	<i>Z. mays</i>	Phospholipid transfer protein precursor	1e-38	AATAAACCAACAGCCAAAA AATAAAATCTCTCCGTGTGTG	163
ZM13-Contig 60*	(TA) ₆ /P	<i>Z. mays</i>	Unknown protein	1e-11	GGAGGAACAGAGTTTTGTTAT ACTTTC AAGGTGGTGAAG	335

Tab. 2. – continued

Sequence ID	Motifs/Repeat Type	Organism	Protein	E-value	Primer pairs (5'-3')	Product size (bp)
ZM13-Contig63	(GCG) ₆ /P	<i>Z. mays</i>	Unknown protein	3e-43	TTTGATACACGGAGGAAGATA AACCAGAGATTCATCAGAGTG	427
ZM13-Contig65	(GA) ₆ (GCAG) ₄ /IP	<i>Z. mays</i>	Hypothetical protein	5e-96	GTTCCCTTCACACAGACACAGT CTTTTACATGGTGGACGAC	508
ZM13-Contig72	(GC) ₆ /P	<i>Z. mays</i>	Unknown protein	3e-19	AACAACCTGGCTGCATCTATAA ACAGAAGCAGAACCAGCA	504
ZM13-Contig83	(TA) ₅ /P	<i>Z. mays</i>	Nonspecific lipid transfer protein	8e-04	GCAGGCCACACATACATAATA GACCAAGCTAATAAGCCTACC	169
ZM13-Contig99*	(TA) ₆ /P	<i>Z. mays</i>	Unknown protein	2e-29	GCCATAGCAAAGCTCAAATCC CTCCCATCAGCCGACTCTAC	482
ZM13-Contig111*	(GAG) ₅ (ACG) ₄ /IP	<i>Z. mays</i>	Chlorophyll ab binding protein	1e-107	AGTATCCCCTTTTACTACTGG CGTCGTCGTCGTTGTT	941
ZM13-Contig136	(AA) ₆ /P	<i>Z. mays</i>	Putative Glutathion peroxidase	2e-97	CATACAGAAAGGGCGAAACA ACACGCTTCAAGGCTGAGTA	501
ZM13-Contig148	(ACCA) ₅ /P	<i>Z. mays</i>	Putative Protein kinase	1e-18	AAATATACGGCCCCAAGAAAA CAAGCAAGATCGGTGGAAAC	327
ZM13-Contig175	(CCA) ₇ /P-(TA) ₅ /P	<i>Z. mays</i>	Unknown protein	5e-46	ATGAGTACAGCGTCGGAGT ATGGATGGGATACAAATACAA	900
ZM13-AI649785*	(CA) ₅ /P	<i>Z. mays</i>	Unknown protein	9e-33	GAAACAAGAAAGCAGGAACTC GCGCTACAACCTCCAACAC	552
ZM13-AI649789	(TT) ₁₀ /P	<i>Z. mays</i>	Penta tricopeptid repeat	1e-89	GCCACTCTGTTTGGTGGATAG CGGCCTTGTTAATTACCTTGA	537
ZM13-AI649800*	(AC) ₇ (CA) ₃ /P	<i>Z. mays</i>	Transcription factor	5e-32	CTAAGCACAAATTCAGAAGAC TAGATAGCCTCGATCACTC	406
ZM13-AI964452	(CT) ₈ (AA) ₅ /P	<i>Z. mays</i>	Hypothetical protein	2e-37	GAAGTTGCAGTTGTTGTTTTT TAGGCAAGGTGTTATTACTGG	207
ZM13-AI964488	(AT) ₅ (TA) ₄ /IP	<i>Z. mays</i>	Carbonic anhydrase	4e-38	CTCAAGACCTACCCCTTCGTC ACTCCTCGCATTACATATCG	215
ZM13-AI964426	(CT) ₅ (CA) ₅ /P	<i>Z. mays</i>	Phospholipid transfer protein	1e-13	TAAATAAAGACCGCATGAAC AAGCACTTGCCATCTACC	399
ZM13-AI964559*	(AC) ₃ (AA) ₅ /IP	<i>Z. mays</i>	NADH ubiquinon oxidoreductase	6e-59	GAGAGGGGTATTTGCTGTATT CCTACTCTGTTGTACGTGGAT	397
ZM13-AI964577	(CT) ₁₃ /P	<i>G. max</i>	Anaphase entrance complex	0.44	CCACAAAGGCATCAAAAAGAAT ATTGTCCGCCTTAACCAGAAG	356
ZM13-AI977807*	(CG) ₅ /P	<i>Z. mays</i>	Hypothetical protein	7e-108	CTCATCATCTCGTTGTCTC GACGCACCTGTACCTGAA	453
ZM13-AI977889	(GAGAAGG) ₄ (GG) ₃ (GA) ₅ /IP	<i>T. versicolor</i>	Unknown protein	8.2	CGCCAATGGTGAGTGAAAGAA GCAAACGACACCTCCAACCAC	281
ZM13-AI977895	(TA) ₅ /P	<i>Z. mays</i>	Nucleotide sugar transporter	3e-32	GCTGCACATGGAAACACTTCA ACGCAGGTGGCGTTAGAAAGG	463
ZM13-AI977907	(CGT) ₅ /P	<i>Z. mays</i>	Hypothetical protein	3e-09	ATCTCGATCTGCTTCATCAT GTAACCTCAACGTCAGCTTG	253
ZM13-AI977935	(GCT) ₅ /P	No similarity	–	–	CAAACAAAAGAGACATCAGGA CCCAAGCATTCAAACACTACTA	303
ZM13-AI966834	(AA) ₅ /P	<i>Z. mays</i>	40S ribosomal protein	2e-48	TCCATTTTGCTCTGTTTCTT AGCCTTCTTCTGCCACAG	285
ZM13-AI966924*	(TT) ₆ /P	<i>Z. mays</i>	Phosphoglycomutase	1e-35	ATCCGATACCTTTCGGAAAT CAAAGGCTCAAACAAAACCAG	505
ZM13-AI966933	(AATA) ₅ /P-(GGT) ₅ /IP,(CGG) ₃ /IP	<i>Z. mays</i>	Unknown protein	2e-21	ATTTGAGAACGGAAGCAAGTA ACCACCACCACCGACAAG	268
ZM13-AI966957*	(AA) ₅ (AG) ₅ /P	<i>Z. mays</i>	Unknown protein	5e-30	CATCCACCGTGTGTGTGGAA GTGGTTATGGCAGTCGGCAAC	281
ZM13-AI967009*	(GC) ₅ /P	<i>Z. mays</i>	Alcohol dehydrogenase	1e-56	TCAACTGCGTGCCCCCTATG TGGAATGTTCAATCATCTGTTGG	400

Tab. 3. Detailed blastx similarity matches and motifs with repeat types of both simple sequence repeats (SSR) and sequence tagged site STS contained sequences in maize root tissues. “P” indicates perfect repeats, “IP” imperfect repeats and “*” shows unmapped markers. ZM14 is the library name.

Sequence ID	Motifs/Repeat Type	Organism	Protein	E-value	Primer pairs (5'-3')	Product size (bp)
ZM14-Contig13	(GT) ₆ / P-(GCG) ₆ , (CG) ₃ / IP	<i>Z. mays</i>	Putative ring zinc finger protein	7e-38	ATACATTTTTACGTCCAC GTTTCGTTTGGAGGGTGTCTT	214
ZM14-Contig24*	(CA) ₆ / P	<i>Z. mays</i>	Unknown protein	3e-26	TTTGTCCAATCAAGCGAGATA TTGTGTCCGTGCAAATTGGTG	488
ZM14-Contig27*	(TA) ₉ , (AC) ₃ / IP	<i>Z. mays</i>	Jacalin like lectin protein	3e-93	CGGTAGCGAAAATACAGT TCCATCTCCTTCATCTACATC	504
ZM14-Contig53	(CCG) ₆ / P	<i>Z. mays</i>	Unknown protein	3e-140	Not suitable for primer design	–
ZM14-Contig68*	(AT) ₅ / P	<i>Z. mays</i>	Aquaporin plasma membran protein	1e-32	AAACAGGGAGTCTTCTCTTAAC AAACAGGGAGTCTTCTCTTAAC	382
ZM14-Contig79	(CG) ₅ / P	<i>Z. mays</i>	Unknown protein	3e-166	Not suitable for primer design	–
ZM14-Contig106*	(TT) ₆ / P	<i>Z. mays</i>	Hypothetical protein	1e-48	GCCCATATTACATACAAAAG CAACAAGAAGGTTTATTCTG	708
ZM14-Contig114	(GG) ₅ / P	<i>Z. mays</i>	Unknown protein	5e-48	TGGACCAAACATCGGTTGAGC GAATGGAAGGAAGAGGGGTGGT	530
ZM14-Contig119	(ACA) ₅ , (CCA) ₆ / P(CCG) ₅ , (CCA) ₃ /IP	<i>Z. mays</i>	Hypothetical protein	6e-39	TGGCCCCACCTGATGAAATAA GGAGTGGAGGCGGAGGATCT	615
ZM14-AI649544	(TA) ₃ , (TGGA) ₅ / IP	<i>Z. mays</i>	Oxin transporter protein	1e-81	CGATGCCGAAAACCCATTCTT GCATCTGCTCGTGGAGGAAAA	303
ZM14-AI649556	(CTG) ₅ / P	<i>Z. mays</i>	Lysin decarboxylase	3e-07	Not suitable for primer design	–
ZM14-AI855154	(GAG) ₃ , (CTG) ₅ / IP	<i>Z. mays</i>	Hypothetical protein	0.002	TGGTCTGCTGCTTCTCTTGC CAAACGGTCCACCTCCACCTC	316
ZM14-AI855158	(AT) ₅ / P	<i>Z. mays</i>	Ribosome biogenesis protein	1e-94	Not suitable for primer design	–
ZM14-AI855164	(AT) ₅ / P	<i>Z. mays</i>	Unknown protein	4e-08	GTGCAGAGACGGACAGCGAGT TTTTGAGTTTTGCCGGAGTGG	346
ZM14-AI855182	(AT) ₆ / P	<i>Z. mays</i>	Glyceraldehyde-3-dehydrogenase	1e-47	CGACTCCCAACAGGAAATGGA GTCGCCTGGTACGACAACGAG	349
ZM14-AI855214	(TA) ₅ / P	<i>Z. mays</i>	Unknown protein	5e-61	TGGATCGAAGCAACTCGCACT CAACAGCGTCAAGAGCGTCAA	310
ZM14-AI855258	(TC) ₁₃ / P- (TA) ₃ , (CACAG) ₆ / IP	<i>Z. mays</i>	Unknown protein	3e-19	GAGATAGCGGAGGAAGGTGAG ATCGTCATCATTCGAGCAGAG	180
ZM14-AI855272	(AGG) ₆ / P	<i>Z. mays</i>	Putative MYB binding protein	2e-43	GACCTCAACCTCGACCTCTG GGCTTCGTTCACTTCATCTTG	275
ZM14-AI855286	(CTG) ₉ / P	<i>Z. mays</i>	Unknown protein	3e-34	CCCTCTCTTTTCTCAGCCCTA ATCCAGAGGACGAGGGTTTT	288
ZM14-AI855336	(CGC) ₅ / P	<i>Z. mays</i>	Hypothetical protein	2e-28	GAGAAGCCGAGAACAGTAGCA CACGAGGCAGAGTCGTAGTTT	365
ZM14-AI855352	(GC) ₅ / P- (CGC) ₆ , (CGA) ₃ / IP	<i>Z. mays</i>	Hypothetical protein	8e-11	CGAGAGCGCCATTAGAAGTCG TTTTGTGGAACGAAGCGATGG	401
ZM14-AI855361	(ACC) ₉ , (TA) ₅ / P	<i>Z. mays</i>	Unknown protein	2e-47	GACCTGGAGTGGTGGTTC GATGGGATACAAATACAATAC	481
ZM14-AI649556	(CTG) ₅ / P	<i>Z. mays</i>	Lysin decarboxylase	3e-07	CCGGGTTTTGGACTTTGGAGA TGTTGTGGCTCTTGCCTGTG	301
ZM14-AI861145	(TT) ₅ , (GTC) ₅ / P	<i>Z. mays</i>	Unknown protein	4e-56	TTGTTTTTGCCTTTCCTTGAA TGCTGTACCCAAATCCTTCTG	447
ZM14-AI857227	(GCC) ₆ / P	<i>Z. mays</i>	putative cytochrome P450 superfamily protein	5e-43	GCCGTGCCAACTTTTAATTC CTGGAGGTGGAAGGAGAGG	367
ZM14-AI855418	(AG) ₅ / P	<i>Z. mays</i>	Putative phototropic-responsive NPH3 family	0.082	GCCTGCCTGTCCATCAATCAA CATCCACCTCCCACCCAGAAC	242
ZM14- AI857233	(ATG) ₆ / P	<i>Z. mays</i>	Protein transport protein sec31-like		ATCTGCACCTCAACCTGAATG ATTGGATGGTTCTTGTGTTGG	236

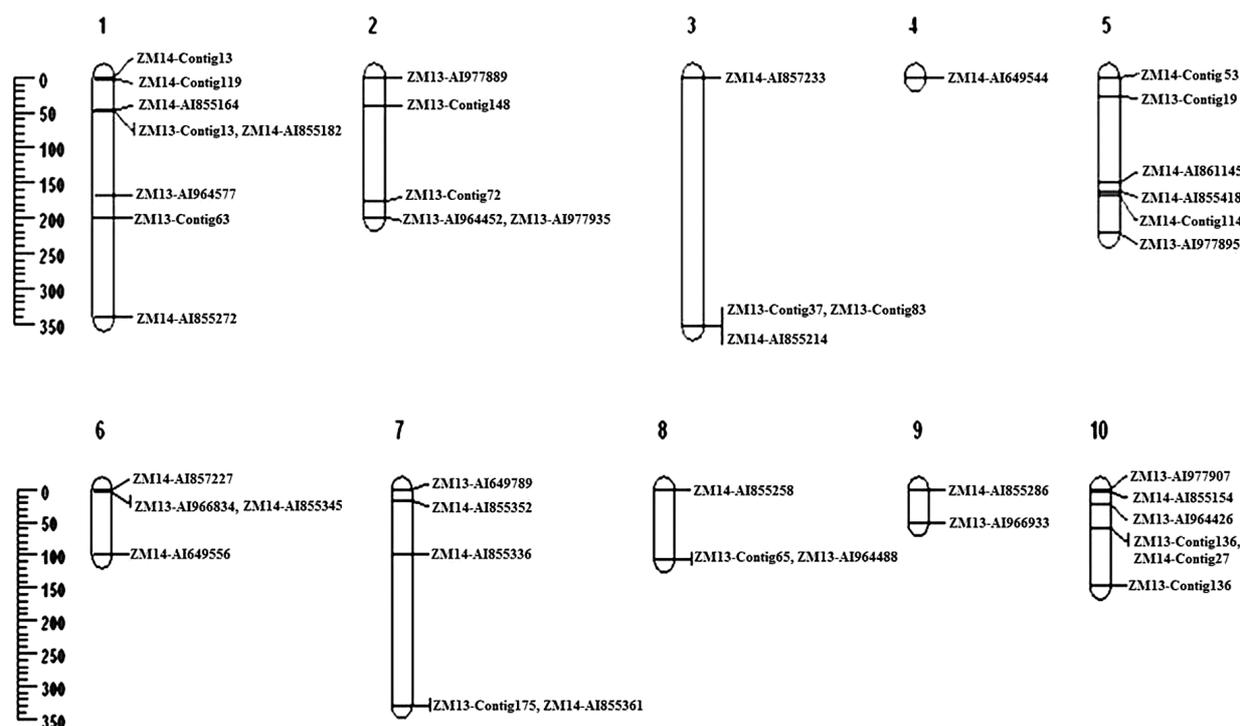


Fig. 1. *In silico* relative map locations of 44 expressed sequence tagged-simple sequence repeats (EST-SSRs) on haploid maize chromosomes constructed on MapChart 2.0 software and mapping unit represented as centimorgan (cM).

chromosome 1, but they were located far from each other (50.58 cM). On the other hand, the loci ZM13-Contig37 and ZM13-Contig83 were located together on the other chromosome 2, with 154.58 cM. The *in silico* mapping of 44 EST-SSR markers in comparison to maize genome assembly is given in Fig. 1. All primer pairs were assigned to ten maize chromosomes by *in silico* mapping while only 4 polymorphic primer pairs were assigned to two chromosomes by classical genetic mapping.

The putative functions of SSR contained sequences were assigned in the GenBank database using similarity search of BLASTX (Tabs. 2, 3). The data indicated that 26 (44.06%) of 59 loci have shown similarity to known genes and have a range of functions, such as metabolic enzymes, structural proteins, disease signaling and transcription factors. More than half (52.7%) of the 55 primer pairs did not yield any significant annotation.

Discussion

DNA sequences retrieved from maize cDNA libraries in NCBI data-base were searched for repetitive regions. The repeat percentage was found to be 1.46% under salt stress in maize cDNA libraries from NCBI GenBank. This frequency is similar to that previously reported by Kantety et al. (2002) (1.5% percentage of maize EST-SSRs). The most abundant repeats were di-nucleotide (78.8%). This result is consistent with those obtained by Wang et al. (1994) in temperate maize and Sibov et al. (2003) in tropical maize. The most repeated di-nucleotide motifs were (TT)₃₁, (CT)₂₇, (CT)₃₂ and (GA)₂₆, (GA)₃₄. Jayashree et al. (2006) observed (CT) and (GA) repeats to be the most common di-nucleotide motifs in cereals and legume plants. Four tri-nucleotide

repeats (CTC, CTG, CTT, TTA) responsible for leucine synthesis were found only in root tissue specific sequences.

The triplet codons for proline and arginine were mostly detected in contigs and singletons from root and shoot tissue. Proline containing repeats were detected in ZM14-Contig119 and ZM13-Contig175 sequences for root and shoot tissue respectively, while ZM14-AI855272, ZM14-AI855352 and ZM14-AI855336 sequences contained triplets for arginine synthesis. Proline is a well-known osmoprotectant molecule and responsible for balancing the cell water potential in plants (Hu et al. 1992). In case of decreased photosynthesis rate, proline has also functions in protein deamination to provide extra energy from proteins (Mattioli et al. 2009, Hayat et al. 2012). Thus, it has been pointed that these ESTs may especially have the potential to develop markers for salt stress.

Although many DNA markers have been developed from maize, SSR markers have not been developed specifically from ESTs constructed under salt stress condition. In this research, 55 primer pairs were specifically developed by using salt stress-induced maize ESTs. Fragments with expected sizes were clearly amplified in all primer pairs. Molecular markers can be mapped by using classical genetic mapping or *in silico* mapping. In this study, many markers have not been mapped classically due to the lack of polymorphism between the parents of the mapping population. On the other hand, according to *in silico* mapping, 44 markers were assigned to ten maize chromosomes (Fig. 1). *In silico* mapping position of four markers (ZM14-Contig119, ZM13-AI964577, ZM13-Contig37, ZM13-Contig83) were matched with the conventional linkage map, in terms of both chromosome arm location and order.

In classical genetic mapping analysis, segregation deviation from expected Mendelian segregation ratios was detected for two markers (ZM14-Contig119 and ZM13-AI964577). These markers mapped to chromosomes 1 were skewed towards the F63 (tolerant) parent. Deviation from the expected Mendelian ratios was previously reported in maize (Pereira and Lee 1995, Sharopova et al. 2002). Segregation distortions are often due to differential gametophytic selection (Lu et al. 2002), chromosomal rearrangements (Lu et al. 2002). Sa et al. (2012) indicated that segregation distortions might be influenced by the mapping population of F_2 , a backcross or recombinant inbred line, and by sizes of the mapping population.

Expansion of the knowledge of the function of ESTs will increase the likelihood that EST based markers for salt-tolerance in maize molecular breeding will be identified. In this study, there are many ESTs exhibiting sequence homology in sequence databases (Tabs. 2, 3). Two contig sequences (Contig33, Contig27) relating to maize root and shoot tissues were matched with jacalin-like plant stress proteins. Xiang et al. (2011) have proved that jacalin is a core compound for plant disease resistance mechanism. In addition, two contigs (Zm13-Contig37 and Zm13-Contig83) assigned to chromosome 3 showed homology to nonspecific lipid transfer proteins (nsLTP). LTPs in maize (ZmLTPs) have critical roles in resistance to biotic and abiotic stress. They provide high salinity resistance with decreasing solute permeability of cell membrane (Liu et al. 2015). 63 nsLTP genes identified and differentially expressed under drought, salt and cold stresses were unevenly assigned to ten maize chromosomes by *in silico* mapping (Wei and Zhong 2014). Six ZmLTPs (ZmLTPd6, ZmLTPd7, ZmLTPd8, ZmLTP1.1, ZmLTP1.2, ZmLTP1.3) were placed on chromosome 3 (Wei and Zhong 2014). Sharopova et al. (2002) used an IBM population B73xMo17 and mapped the microsatellite marker p-umc1010, as an anchor marker for phospholipid transfer protein homolog2 (*plt2*), on maize chromosome 3. This marker region amplified a (GA) motif with ten repeats. In our study, we have identified two different alternative marker regions associated to phospholipid transfer protein in maize F63xF35 RILs. These marker loci (ZM13-Contig37 and ZM13-Contig83) were placed on chromosome 3 and amplified the repeat regions for (TC)₁₀ and (TA)₅. In addition, Zm14-Contig119, which covered three different types of triple motifs (ACA, CCG and CCA) and was located on maize chromosome 1, matched with hypothetical proteins. In *Sorghum*, Ngara et al. (2012) observed 22 hypothetical protein inductions after salt stress application in moderately salt tolerant plants. Similarly, Zahra et al. (2013) observed hypothetical protein induction after salt stress application. Shinozaki et al. (2005) reported that hypothetical proteins with uncharacterized domains were relevant to salt tolerance. Another EST-SSR tagged as ZM13-AI964577 showed similarity to *Glycine max* anaphase entrance complex. However, this sequence showed a polymorphic fragment pattern, and Blastx *E*-value resulting was out of the accepted score limits.

Parental genotypes and their 158 F_5 derived RILs which were used in our study were previously tested with 3072 SNP markers (Cui et al. 2015) and 81 of these SNP markers clustered on chromosomes 1, 3 and 5. Results suggested that some QTLs were related to the traits of shoot sodium and potassium concentration in maize. In our study, we have identified four EST-SSR markers that were annotated to two different maize chromosomes (1 and 3). Significantly, annotation analysis of these EST-SSRs showed a close relatedness to some salt tolerance proteins.

Another matched EST (AI964488) displayed homology with carbonic anhydrase protein. Yu et al. (2007) demonstrated that carbonic anhydrase gene was expressed in *Oryza* under environmental stress such as salt stress. Further, the root specific Contig68 showed a similarity with Aquaporin plasma membrane protein. Plasma membrane proteins control osmotic pressure of the cell and they are closely correlated to specific transport proteins such as SYP121. This protein was identified as vesicle transport protein and controls potassium traffic in plant cells (Besserer et al. 2012). Potassium is an essential, and the most abundant, cation in plants and it has protective effects for plants under salinity (Cakmak 2005). Lastly, EST (AI855272) derived from root tissue displayed high similarity with the MYB binding protein. Genes coding MYB transcription factors were first discovered from maize (Paz-Ares et al. 1987). These proteins have various regulatory functions in gene expression mechanism under salinity (Hasegawa et al. 2000).

For screening of high potential parts of plant genomes, there are different types of molecular marker techniques such as the PCR based; SSR, ISSR (inter simple sequence repeat) and hybridization based; DArT (diversity array technology) and sequencing based markers such as SNPs (Xie et al. 2006, Kalia et al. 2010). ESTs are attractive tools for marker development since they represent coding regions of the genome. There are a number of advantages for markers developed from EST-derived sequences (Davis et al. 1999). When an EST marker is found to be genetically associated with a trait of interest, the corresponding mapped gene can directly affect the trait. Also, genetic mapping with ESTs would provide a more rapid transfer of linkage information between species (Cato et al. 2001). SSR markers are abundantly distributed throughout the maize genome and they are cost effective markers for screening QTLs (Xu et al. 2013). In this study, both advantages of ESTs and SSRs were merged to characterize the new marker sources for maize breeding.

In terms of overall evaluation of homology analysis, the new primers that were designed from computationally extracted EST-SSRs and tested in tolerant and sensitive maize genotypes might be used for scanning other maize germplasm sources. Furthermore, integration of molecular validation of these SSRs may serve cost effective molecular markers for improvement of salt tolerant maize. With the improvement of feasible molecular markers, these candidate EST-SSRs might have the ability to broaden the genetic base of maize.

Acknowledgements

The authors would like to thank Professor Huabang Chen, State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmen-

tal Biology, Chinese Academy of Sciences, for his great support during experiments. This work was financially supported by a Scientific Research Projects (BAP project no. FEF-1901.14-02) sponsored by the University of Sınop, Turkey.

References

- Banisetti, K., Agrawal, P. K., Gupta, H. S., Kumar, A., Bhatt, J. C., 2012: Identification of candidate gene based SSR markers for lysine and tryptophan metabolic pathways in maize. *Plant Breeding* 131, 20–27.
- Besserer, A., Burnotte, E., Bienert, G. P., Chevalier, A. S., Errachid, A., Grefen, C., Blatt, M. R., 2012: Selective regulation of maize plasma membrane aquaporin trafficking and activity by the Snare SYP121. *Plant Cell* 24, 3463–3481.
- Bilgen, M., Karaca, M., Onus, A. N., Ince, A. G., 2004: A software program combining sequence motif searches with keywords for finding repeats containing DNA sequences. *Bioinformatics* 20, 3379–3386.
- Bitá, C. E., Gerats, T., 2013: Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. *Frontiers in Plant Science* 4, 273.
- Bray, E. A., Bailey-Serres, J., Weretilnyk, E., 2000: Responses to abiotic stresses. Chapter 22. In: Grissem W., Buchannan B., Jones R. (eds.) *American Society of Plant Physiologists*, Rockville, MD, pp 1158–1240.
- Cakmak, I., 2005: The role of potassium in alleviating detrimental effects of abiotic stresses in plants. *Journal of Plant Nutrition and Soil Science* 168, 521–530.
- Cato, S., Gardner, R., Kent, J., Richardson, T., 2001: A rapid PCR based method for genetically mapping ESTs. *Theoretical and Applied Genetics* 102, 296–306.
- Cui, D., Wu, D., Somarathna, Y., Xu, C., Li, S., Li, P., Zhang, H., Chen, H., Zhao, L., 2015: QTL mapping for salt tolerance based on snp markers at the seedling stage in maize (*Zea mays* L.). *Euphytica* 203, 273–283.
- Davis, G. L., McMullen, M. D., Baysdorfer, C., Musket, T., Grant, D., Staebell, M., Xu, G., Polacco, M., Koster, L., Melia-Hancock, S., Houchins, K., Chao, S., Coe, E. H., 1999: A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map. *Genetics* 152, 1137–1172.
- Doyle, J. J., Doyle, J. L., 1987: A rapid DNA isolation procedure for small quantities of fresh tissue. *Phytochemical Bulletin* 19, 11–15.
- Hasegawa, P. M., Bressan, R. A., Zhu, J. K., Bohnert, H. J., 2000: Plant cellular and molecular responses to high salinity. *Annual Review of Plant Physiology and Plant Molecular Biology* 51, 463–499.
- Hayat, S., Hayat, Q., Alyemini, M. N., Wani, A. S., Pichtel, J., Ahmad, A., 2012: Role of proline under changing environments: A review. *Plant Signaling & Behavior* 7, 1456–1466.
- Hu, C. A., Delauney, A. J., Verma, D. P., 1992: A bifunctional enzyme (delta 1-pyrroline-5-carboxylate synthetase) catalyzes the first two steps in proline biosynthesis in plants. *Proceedings of the National Academy Sciences of the United States of America* 89, 9354–9358.
- Jayashree, B., Punna, R., Prasad, P., Bantte, K., Hash, C., Chandra, T., Hoisington, D. A., Varshney, R. K., 2006: A database of simple sequence repeats from cereal and legume expressed sequence tags mined in silico: survey and evaluation. *In Silico Biology* 6, 0054.
- Kalia, K. R., Rai, M. K., Kalia, S., Singh, R., Dhawan, A. K., 2010: Microsatellite markers: an overview of the recent progress in plants. *Euphytica* 177, 309–334.
- Kantety, R. V., Rota, M. L., Matthews, D. E., Sorrells, M. E., 2002: Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum, and wheat. *Plant Molecular Biology* 48, 501–510.
- Karaca, M., Bilgen, M., Onus, A. N., Ince, A. G., Elmasulu, S. Y., 2005: Exact tandem repeats analyzer (eTRA): a new program for DNA sequence mining. *Journal of Genetics* 84, 49–54.
- Klopfenstein, T. J., Erickson, G. E., Berger, L. L. 2013: Maize is a critically important source of food, feed, energy and forage in the USA. *Field Crops Research* 153, 5–11.
- Kosambi, D. D., 1943: The estimation of map distances from recombination values. *Annals of Human Genetics* 12, 172–175.
- Lee, M. (eds.) 2007: *Genomics-assisted crop improvement: Vol 2: Genomics applications in crops*. Springer Science & Business Media.
- Li, H. H., Ye, G. Y., Wang, J. K., 2007: A modified algorithm for the improvement of composite interval mapping. *Genetics* 175, 361–374.
- Li, H. H., Ribaut, J. M., Li, Z. L., Wang, J. K., 2008: Inclusive composite interval mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations. *Theoretical and Applied Genetics* 116, 243–260.
- Lima, M. D., Souza, L. A., Jr, C. L. D., Souza, A. P. D., 2009: Microsatellite-dense genetic map: towards genome coverage in a tropical maize (*Zea mays* L.) population. *Brazilian Journal of Botany* 32, 499–508.
- Liu, F., Zhang, X., Lu, C., Zeng, X., Li, Y., Fu, D., Wu, G., 2015: Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. *Journal of Experimental Botany* 66, 5663–5681.
- Lu, H., Romero-Severson, J., Brnard, R., 2002: Chromosomal regions associated with segregation distortion in maize. *Theoretical and Applied Genetics* 105, 622–628.
- Mattioli, R., Costantino, P., Trovato, M., 2009: Proline accumulation in plants: Not only stress. *Plant Signaling and Behavior* 4, 1016–1018.
- Ngara, R., Ndimba, R., Borch-Jensen, J., Jensen, O. N., Ndimba, B., 2012: Identification and profiling of salinity stress-responsive proteins in *Sorghum bicolor* seedlings. *Journal of Proteomics* 75, 4139–4150.
- Orsini, E., Krchov, L. M., Uphaus, J., Melchinger, A. E., 2012: Mapping of QTL for resistance to first and second generation of European corn borer using an integrated SNP and SSR linkage map. *Euphytica* 183, 197–206.
- Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P. A., Saedler, H., 1987: The regulatory c1 locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *EMBO Journal* 6, 3553–3558.
- Pereira, M. G., Lee, M., 1995: Identification of genomic regions affecting plant height in sorghum and maize. *Theoretical Applied Genetics* 90, 380–388.

- Rotmistrovsky, K., Jang, W., Schuler, G. D., 2004: A web server for performing electronic PCR. *Nucleic Acids Research* 32, 108–112.
- Sa, K. J., Park, J. Y., Park, K. C., Lee, J. K., 2012: Analysis of genetic mapping in a waxy/dent maize RIL population using SSR and SNP markers. *Genes and Genomics* 34, 157–164.
- Schubert, S., Zörb, C., Sümer, A., 2001: Salt resistance of maize: recent developments. In: Horst, W.J. et al. (eds.), *Plant nutrition. Food security and sustainability of agro-ecosystems*. Kluwer Academic Publisher, Dordrecht, The Netherlands.
- Sharopova, N., McMullen, M. D., Schultz, L., Schroeder, S., Sanchez-Villeda, H., Gardiner, J., et al., 2002: Development and mapping of SSR markers for maize. *Plant Molecular Biology* 48, 463–481.
- Shinozaki, N., Yamada, M., Yoshida, Y., 2005: Analysis of salt stress-inducible ESTs isolated by PCR-subtraction in salt tolerant rice. *Theoretical and Applied Genetics* 110, 1177–1186.
- Sibov, S. T., Lopes De Souza, Jr. C., Garcia, A. A., Garcia, A., Silva, A. R., Mangolin, C. A., Benchimol, L. L., De Souza, A. P., 2003: Molecular mapping in tropical maize (*Zea mays* L.) using microsatellite markers. 1. Map construction and localization of loci showing distorted segregation. *Hereditas* 139, 96–106.
- Singh, A., 2015: Soil salinization and waterlogging: A threat to environment and agricultural sustainability. *Ecological Indicators* 57, 128–130.
- Voorrips, R. E., 2002: MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity* 93, 77–78.
- Wang, Z., Weber, J. L., Zhong, G., Tanksley, S. D., 1994: Survey of plant short tandem repeats. *Theoretical and Applied Genetics* 88, 1–6.
- Wei, K., Zhong, X., 2014: Non-specific lipid transfer proteins in maize. *BMC Plant Biology* 14, 281.
- Xiang, Y., Song, M., Wei, Z., Tong, J., Zhang, L., Xiao, L., Ma, Z., Wang, Y., 2011: A jacalin-related lectin-like gene in wheat is a component of the plant defence system. *Journal of Experimental Botany* 62, 5471–5483.
- Xiang, C., Du, J., Zhang, P., Cao, G., Wang, D., 2015: Preliminary study on salt resistance seedling trait in maize by SRAP molecular markers. *Advances in Applied Biotechnology*. Springer, Berlin, Heidelberg, pp. 11–18.
- Xie, Y., McNally, K., Li, C. Y., Leung, H., Zhu, Y. Y., 2006: A high-throughput genomic tool: diversity array technology complementary for rice genotyping. *Journal of Integrative Plant Biology* 48, 1069–1076.
- Xu, J., Liu, L., Xu, Y., Chen, C., Rong, T., Ali, F., Shufeng, Z., Fengkai, W., Yaxi, L., Jing, W., Moju, C., Lu, Y., 2013: Development and characterization of simple sequence repeat markers providing genome-wide coverage and high resolution in maize. *DNA Research* 20, 497–509.
- Xue, S., Li, G., Jia, H., Lin, F., Cao, Y., Xu, F., Tang, M., Wang, Y., Wu, X., Zhang, Z., Zhang, L., Kong, Z., Ma, Z., 2010: Marker-assisted development and evaluation of near-isogenic lines for scab resistance QTLs of wheat. *Molecular Breeding* 25, 397–405.
- Yu, S., Zhang, X., Guan, Q., Takano, T., Liu, S., 2007: Expression of a carbonic anhydrase gene is induced by environmental stresses in rice (*Oryza sativa* L.). *Biotechnology Letters* 1, 89–94.
- Yumurtaci, A., 2015: Utilization of wild relatives of wheat, barley, maize and oat in developing abiotic and biotic stress tolerant new varieties. *Emirates Journal of Food and Agriculture* 27, 1–23.
- Zahra, A. R. F., De Costa, D. M., De Costa W. A. J M., 2013: Identification of differentially-expressed genes in response to salt stress in the salt-tolerant Sri Lankan rice variety At354. *J National Sci Foundation Sri Lanka* 41, 93–112.
- Zhou, J. Q., Guo, Y. Q., Gao, Y. F., Li, J. S., Yan, J. B., 2011: A SSR linkage map of Maize×Teosinte F₂ population and analysis of segregation distortion. *Agricultural Sciences in China* 10, 166–174.