Romanian Biotechnological Letters *Vol. 18, No.1, 2013*

ORIGINAL PAPER

Developing, linkage mapping and phylogenetic analyses of AP2-EREBP type transcription factor markers in citrus

Received for publication, July 5, 2011
Accepted, November 5, 2011

OSMAN GULSEN[1*], CENGIZ IKTEN[2], AYDIN UZUN[1], UBEYİT SEDAY[3], GUCER KAFA[3]
[1]*Erciyes University, Faculty of Agriculture, Department of Horticulture, 38039, Kayseri, Turkey*
[2]*Akdeniz University, Faculty of Agriculture, Department of Plant Protection, Antalya, Turkey*
[3]*Alata Horticultural Research Institute, Erdemli 33740, Mersin, Turkey*
**Corresponding author, e-mail: ogulsen@erciyes.edu.tr,*
Phone: +90-352-4374901 #38927, Fax: +90-352-4376902

Abstract

Ethylene responsive AP2/EREBP type transcription factors (TF) play major roles such as growth, development, and tolerance to biotic and abiotic stresses in plants. Forward and reverse AP2/EREBP type TF-specific primers were designed, sequenced, and linkage mapped in a population of 164 F_1 individuals derived between 'Clementine' mandarin (Citrus reticulata Blanco 'Clementine') and 'Orlando' tangelo' (C. paradisi Macf. 'Duncan' x C. reticulata Blanco 'Dancy'). A total of 26 pairs of primers were designed for PCR reactions using Primique software available in TF database (DATFAP) based on default parameters using available dicot's AP2/EREBP sequences. These primers included 17 to 35 bases, and produced a total of 21 polymorphic markers. Bright 13 markers were excised, sequenced, deposited in the NCBI web site, and BLAST-analyzed for homology. Of the 21 markers, 13 were linkage mapped in a previous citrus map. Seven and five markers were mapped in 'Clementine' and 'Orlando' map, respectively. Based on Maximum Parsimony algorithm nested in MEGA 4 evolutionary genetic analysis software, the 13 TF sequences obtained in this study were found to be closely related to known TFs of Arabidopsis thaliana. Few of TF markers were found to be closely linked in existing linkage map of citrus, suggesting possible ancestral origin. These AP2/EREBP primers helped identification of citrus AP2/EREBP type transcription factor genes and can be used in other dicots such as tomato and cotton may have potential in understanding evolutionary relationships, establishing linkage map, and estimating diversity among other dicots since these TFs may reflect adaptability of plants.

Keywords: *Citrus reticulata*, DATFAP, F_1 mapping, EREBP, transcription factor, maximum parsimony

Introduction

Regulation of gene expression is critical for a variety of essential processes in plants, such as growth, development, differentiation, metabolic regulation, and stress tolerance (VERMA 1992 [1]). Transcription, the first step in gene expression, plays a central role in the regulation of the expression of genes. Therefore, the analysis of transcription factors (TF) is essential for an understanding of mechanisms of gene expression and their genome organization.

There are 68 TF gene families as reviewed by (SHIGYO & al. 2006 [2]). Plant specific AP2/EREBP type TF genes are widely distributed in plants and form a gene family. The AP2/EREBP genes are divided into two subfamilies: AP2 genes with two AP2 domains and EREBP genes with a single AP2/ERF (ethylene responsive element binding factor) domain. The AP2/EREBP gene family has 146 AP2 members listed in: http://arabtfdb.bio.uni-potsdam.de/v1.1/fam_mem.php?family_id=AP2-EREBP. Members of the AP2/EREBP superfamily are believed to play important roles in a variety of biological processes. For example, TINY and APETALA2 are involved in developmental and growth processes (WILSON & al. 1996 [3]; OKAMURO & al. 1997 [4]). Pti4, Pti5, and Pti6 are three members of the EREBP family in tomato and have been shown to interact with disease resistance gene products of Pto (ZHOU & al. 1997 [5]). Three other EREBPs, CBF1,

DREB1A and DREB2A, bind to a CRT/DRE (C-repeat/dehydration-responsive element) that is involved in gene expression in response to drought and cold stress (STOCKINGER & al. 1997 [6]; LIU & al. 1998 [7]).

These ethylene-responsive transcription factors have a conserved basic 58 or 59-amino acid DNA-binding domain and function as regulators of the GCC box mediated transcription of genes (OHME-TAKAGI & SHINSHI 1995 [8]; LIU & al. 1998 [7]). Analyses of promoter regions of several basic TF genes indicate that an 11-bp consensus sequence (TAAGAGCCGCC with the core sequence of AGCCGCC known as GCC box) is required for the ethylene responses (OHME-TAKAGI & SHINSHI 1995 [8]; LIU & al. 2006 [9]). These conserved sequences may allow us to develop new tools study genetic analysis of plants such as locus specific DNA markers for construction of genetic linkage maps and estimation diversity of plants.

Citrus is suitable for genetic linkage mapping studies because of its simple ploidy level, diploid with only 9 haploid chromosomes, has a relatively small genome of about 1C = 0.62 pg, and known cultivars are highly polymorphic (GUERRA 1984 [10]). In addition, interspecific and intergeneric hybrids can be produced (BARRETT 1985 [85]). Many citrus whole or partial genetic maps have been developed over the past decade (CHEN & al. 2007 [12]; ROOSE 2007 [13]). Each has a different mapping population type and size, genome coverage, and marker systems. Most of these maps were covered by a majority of randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), and/or inter-simple sequence repeat (ISSR) markers, and a limited number of restriction fragment length polymorphism (RFLP), simple sequence repeats (SSR), sequence characterized amplified region (SCAR), cleaved amplified polymorphic sequence (CAPS), sequence tagged site (STS), or single/simple nucleotide polymorphism (SNP) markers. The linkage groups of the current maps are not all uniform in length and marker saturation (ROOSE 2007 [13]), whereas the cytogenetic evidence suggests that the chromosomes of citrus are similar in length or show a continuous gradation in size (YAMAMOTO & TOMINAGA 2004 [14]). The reason is that map distance strictly depends on level of polymorphism between two parents and heterozygosity in parental genotypes. However, an increased number of polymorphic markers and progenies would be a distinct advantage for marker-assisted selection and for cloning quantitative trait loci (QTLs) and other loci responsible for traits of horticultural importance. Data on citrus genome is accumulating, which will help better understanding its genome organization.

Developing new tools may increase our understanding of the AP2/EREBP type TF gene superfamily. In addition, combining different marker systems with a sufficient number of marker loci should saturate each of the linkage groups. Using locus-specific markers would be distinct advantage in constructing a linkage map of plants. With these in mind, we previously employed various markers systems in combination to achieve more genome coverage, including sequence related amplified polymorphism (SRAP), SSR, ISSR, RAPD, peroxidase gene polymorphism (POGP), and resistance gene analogs (RGA) (GULSEN & al. 2010 [15]). The objectives of this study were to develop AP2/EREBP type TF markers for dicots, and integrate these markers into an available genetic linkage map of citrus in this study. This is the first comprehensive report where specific AP2/EREBP type TF primers were used to study genetic linkage mapping in citrus.

Materials and Methods

Plant materials

This segregating population was previously used for genetic linkage map construction of citrus (GULSEN & al. 2010a [15]). ‘Clementine’ mandarin and ‘Orlando’ tangelo were

selected as parents based on previous molecular study (DALKILIC & al. 2005 [16]) which showed that each has a high level of heterozygosity and is therefore likely to be useful for mapping a large number of markers. Along with 'Clementine' mandarin and 'Orlando' tangelo, their 164 progeny plants were used in this study. The plants are located at Alata Horticultural Research Institute, Erdemli, Mersin, Turkey.

Seeds obtained from 'Clementine' X 'Orlando' tangelo (derived from 'Duncan' x 'Dancy') in 2002 were treated with benomyl to eliminate pest development, and germinated in boxes containing a mixture of 1 sand: 1 organic matter: 1 soil in a semi-controlled greenhouse. The progeny plants were randomly selected from the population to minimize selection that might affect segregation ratio. The plants were also treated with benomyl according to manufacturer and transferred to 8-lt pots for further plant growth. When plants had hardened sufficiently, two buds from each seedling were grafted onto one-year old sour orange rootstocks. Those, which failed to grow, were regrafted onto fresh sour orange rootstocks. Special care was given to those seedlings weakened due to inbreeding depression, and immature seeds used in this study. Fresh leaves were selected from the plants maintained in the greenhouse, cleaned with moist paper towels, and stored at –80°C.

Designing TF specific primers

AP2/EREBP type TF-specific primers were designed according to FREDSLUND (2008 [17]). The DATFAP database offers a considerable number (more than 55.000) of EST sequences of TFs of 11 plant species including 8 dicots such as tomato (*Solanum lycopersicum* L.), potato (*Solanum tuberosum* L.), poplar (*Populus trichocarpa (Torr. & Gray*), soybean (*Glycine max* L.), *Lotus japonicus*, cotton (*Gossypium hirsutum* L.), *Medicago truncatula* Gaertn., and *Arabidopsis thaliana*, and the following web site allows primer design (http://cgi-www.daimi.au.dk/cgi-chili/datfap/frontdoor.py). The DATFAP database uses EST sequences from several databases: the Institute of Genomic Research, the National Center for Biotechnology Information, the Database of Arabidopsis Transcription Factors. For designing primers, default parameters of *Primique* program were used. The AP2/EREBP TF genes were selected and aligned, and, separate primers were designed for each group (Table 1).

DNA extraction

The total genomic DNA was extracted from young leaves by a modified CTAB method from DOYLE & DOYLE (1990 [18]). Plant tissues were placed between two rollers of a sap extraction apparatus (Ravenel Specialities Co., Seneca, SC) and 1.2 ml of extraction buffer [2% (W/v) CTAB, 100 mM Tris-HCl pH 8.0, 1.4 M NaCl, 20 mM Na EDTA pH 8.0, 0.2% mercaptoethanol] were added onto rolling press. Homogenate was collected from the bottom of the roller into sterile 2.0-ml Eppendorf and immediately incubated for 45-60 min at 65 ^{o}C. A 24 chloroform: 1 octanol extraction was performed and mixed by inverting 100 times, and centrifuging 14.000 rpm for 5 min. Supernatant (500-700 ul) were transferred to sterile 1.5-ul Eppendorf tubes and chilled isopropanol (2/3 of the volume) was added, followed by gently inverting 2 times. DNA pellets were precipitated at –20 ^{o}C for 30 min, recovered by centrifugation at 14.000 rpm for 2 min, washed with 500 ul of washing buffer (76% ethanol; 10 mM ammonium acetate), and centrifuged at 14.000 rpm for 1 min. Supernatant was removed and the DNA pellet was rinsed with 70% ethanol, dried at room temperature, and re-suspended in 100 ul buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA). RNAse (Sigma-Aldrich Co., St. Louis, MO, U. S. A.) was added (1 ul of 500 ng/ml), and the samples were incubated at 37 °C for 1 h. DNA was recovered by adding 200 ul buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA) plus 15 ul of ammonium acetate (10 M, pH 7.7), 600 ul of ethanol, and centrifuging at 14.000 rpm for 1 min. Supernatant was removed and the pellet were air dried at RT for 15 min, dissolved in 200 ul of buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA). DNA concentrations were measured with spectrophotometer and double-checked with 1% agarose

gels. A portion of the DNA was diluted to 10 ng/ul with sterile double-distilled water for use in analyses, and both the stock and diluted portions were stored at –20°C.

PCR conditions and electrophoresis

PCR was carried out using Sensoquest thermal cycler (SensoQuest GmbH, Göttingen, Germany) with modifications of the annealing temperature to optimize the reaction conditions for individual primers. Reaction components were put in a 0.2 ml microtube. An overlay of mineral oil (1–2 drops) was used to minimize evaporation. The contents of the reaction mix were as follows: 1.5 ul of 10× Taq buffer (Bioron, Ludwigshafen, Germany), 2.5 mM of $MgCl_2$, 333 uM of each of dNTPs, 1.3 uM of each of forward and reverse primers, 15 ng of DNA template, 0.2 ul of Taq polymerase of 5 u/ul (Bioron, Ludwigshafen, Germany), 160 ng/ul of BSA, sterile H_2O to 15 ul. Cycling parameters were as described in relevant papers above. Few amplifications were optimized by visual assessment of the banding patterns. Repeatability was determined between experiments and between parental and a few or whole progeny samples within experiments. PCR products were separated on 2.5 % agarose gel at 110 volt for 3 or 5 h, and visualized under UV light.

Sequencing and phylogenetic analysis of AP2/EREBP type TF markers

Clear sharp bands produced with the primers designed in this study were excised using special pipette tips, and DNAs were isolated and purified using QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA). The Sanger sequencing of direct PCR product was used for sequencing, the sequences were deposited as genome survey sequences (GSS) in the following web site: www.ncbi.gov. Sequence similarities of TFs identified in this study with 13 known TF sequences (5 AP2, 4 ERF, 2 DREB, 2 RAW) and one *60S* ribosomal protein cDNA sequence (as control) of *A. thaliana* was estimated Maximum Parsimony (MP) method by using Mega version 4.0 genetic analysis software (KUMAR & al., 2004 [19]). FENG & al. (2005 [20]) assigned *A. thaliana* TFs into subgroups such as AP2, ERF, RAW, ERF, and others based on their amino acid sequences, which were used to estimate grouping of TFs in this study. Default parameters were used for constructing phylogenetic tree.

Marker scoring, grouping, and ordering

Marker scoring, grouping and ordering were made according to GULSEN & al. (2010 [15]). The progeny population was genotyped for polymorphic markers by scoring for their presence or absence in parental and progeny samples in the gel profile. Scored data for TF markers were formatted for the JoinMap 4.0 program and analyzed to produce a genetic linkage map. The segregation ratios observed among the progeny for putative markers were tested for goodness of fit to hypothesized Mendelian ratios using the Chi-square test with the single locus analysis module of JoinMap 4.0 (VAN OOIJEN 2006 [21]). The independence LOD score based on the G2 statistics was calculated by JoinMap for the recombination frequency to determine linkage groups. Linkages for segregation types in the population were established using log-likelihood of the odds (LOD) score at 4.0 to 6.0. Among four options, the Independence LOD score was used because it is not affected by distorted segregation, therefore causes less incidence of spurious linkage as suggested by VAN OOIJEN (2006 [21]). Assembled linkage groups were ordered with the JoinMap mapping module using the Kosambi function of Regression mapping algorithm. Then, ungrouped markers listed in GROUPING node were re-assigned to LGs based on their strongest cross link information, but only for those having the strongest cross link LOD score higher than 2.0 or higher. Once the framework maps for each parent were constructed based on testcross markers, they were used as fixed order and intercross markers were added to these. The final linkage map was drawn using MapChart 2.2 (VOORRIPS 2002 [22]). In addition, each marker band was visually scored as present (1) or absent (0), and data were analyzed with the Numerical Taxonomy Multivariate Analysis System (NTSYS-pc) version 2.1 software package (Exeter

Software, Setauket, New York, USA) (Rohlf 1993 [23]) to estimate linkage relationships among the TF markers produced. A genetic similarity (GS) matrix was constructed based on correlation coefficients within the SIMGEND module. The GS matrix was used to construct a dendrogram using the unweighted pair group method arithmetic average (UPGMA) to estimate linkage relationships among the TFs studied.

Table 1. The list of 26 pairs of primers designed using EST sequences of 8 plant species available in the DATFAP web site.

Primer name	Forward primer	Reverse primer	Number of polymorphic bands	Segregation types
TF3	5'-CGACATTCATCAAACCAACGCCTC-3'	5'-GGYGARGCTTTCACAGTGGCG-3'	NP[a]	
TF10	5'-CCGGGCAGCAGCAACATTGGTAGCGGAGGCGG-3'	5'-TGAGAGGAGGTTGGAAGCCATTTGTCTGCAGCC	1	ab x aa
TF13	5'-CAAAAAGGAATCGYCTCTACTTCTAC-3'	5'-GCRGCTTGATCAGAGTGAGGG-3'	1	aa x ab
TF21	5'-GCTGCAGCATCATCAGGATTCCCCCAGCAG-3'	5'-TGGTCAAGGTCTCATAAAATAATGATGGAAGCC-3'	NP	
TF22	5'-TGGCTTCCATCATTATTTTATGAGACCTTGACC-3'	5'-ACCAATATAATCTGTTTGTTTATGTTTTGAG-3'	NP	
TF23	5'-TGATGATGTTGATGATGTTTTGGCTGATG-3'	5'-CGGATCTCAGCAGCCCATTTACCCC-3'	NP	
TF27	5'-GGAAGTGGGGCAAATGGGCTGC-3'	5'-GGGCAAGCAATGTTGAGGGGGG-3'	2	ab x aa, aa x ab
TF30	5'-CAYTGGGGHTCWTGGGT-3'	5'-GCDGCTTCATCRTAKGC-3'	2	ab x aa, aa x ab
TF31	5'-GCAGCGAGCGGCTCGAGCTCTG-3'	5'-CGACTTTCACGGCTGGAAACTCTTCCTCAGCCG-3'	NP	
TF34	5'-GCTTCCATCATTATTTTATGAGACCTTGACCATTC-3'	5'-TCTGTTTGTTTATGTTTTGAG-3'	NP	
TF43	5'-GCTGCAGCATCATCAGGATTCCCCCAGCAG-3'	5'-TGGTCAAGGTCTCATAAAATAATGATGGAAGCC-3'	NP	
TF54	5'-TGCTGAAGGTGGACGTGGAGGTGTCG-3'	5'-ACAACTCCGGCACCGACACCGG-3'	NP	
TF61	5'-TKCTYATAAACAGAAACGGCC-3'	5'-TWTKTATATCTCACGTCCA-3'	2	ab x aa, aa x ab
TF62	5'-GGTGGATTTGACACTGCACATGC-3'	5'-GTACTTTGGCGGCGCAG-3'	NP	
TF66	5'-CGTCCRTGGGGMAAATGGGC-3'	5'-GATTGTCCATCATAGTACGGAATCTG-3'	NP	
TF68	5'-GCGGCTCCGCCACCTGGATGTGGCGGAGCCGC-3'	5'-GCTGCCCTGGTCTGGGTCCTTTGC-3'	3	ab x aa (2), ab x ab
TF71	5'-AGAGCMTATGATGAAGCAGCWAGGCT-3'	5'-CACTTTCTTGAACATKGGTTGC-3'	NP	
TF73	5'-AAGGCAGCGGCAGCGCCGGCGC-3'	5'-TTCCTCGCCGCCGCCGCCGCCG-3'	NP	
TF74	5'-CCAATTCAAGCTCCTCTGCGGTGG-3'	5'-CGGGTGGGTGGTCACTGTCATC-3'	1	ab x ab
TF77	5'-GCATATGATAGGGCAGCAAT-3'	5'-CCCAAATAAACATACTTTTTGCC-3'	5	ab x aa (2), aa x ab (1), ab x ab (2)
TF78	5'-AGAGATCTATCGCTTGCGGCGACGAG-3'	5'CTGCTGAGGGAATCCTGATGATGCTGCAGCAGC-3'	NP	
TF79	5'-TKCTYATAAACAGAAACGGCC-3'	5'-TWTKTATATCTCACGTCCA-3'	NP	
TF81	5'-CAYTGGGGHTCWTGGGT-3'	5'-GCDGCTTCATCRTAKGC-3'	3	ab x aa, aa x ab, ab x ab
TF82	5'-GCTTTTMAACCAAGCGCGATG-3'	5'-CCGTCGTGAGACAGGTTAGTTTTACCCTACTG-3'	NP	
TF84	5'-CCMAGGTCWAGAAGCTCWCA-3'	5'-CTGCTGRGGGAATCCTGATGATGCTGCAGCAG-3'	1	ab x aa
TF86	5'-CTCACTTTCAATTCCCTCACTCTGATCAAGC-3'	5'-AGARCAGATGATATGYTCYGGAGTC-3'	NP	

Table 1. The list of 26 pairs of primers designed using EST sequences of 8 plant species available in the DATFAP web site.
[a] NP, no polymorphism or amplified product

Results

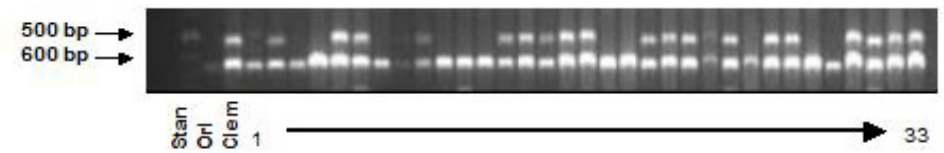
Designing and applying AP2/EREBP type TF-specific primers

Twenty-six pairs of AP2/EREBP type TF-specific primers were designed using DATFAP transcription factor database in this study (Table 1). Conserved domains of the 8 genera listed in M&M allowed us to design these primers. Nucleotide sizes of these primers varied between 17 and 35 bp. Reannealing temperatures ranged from 45 to 54 ^{0}C.

We initially tested 26 primer combinations using two replications, then using 10 progenies along with parental genotypes 'Clementine' and 'Orlando' (Table 1). Among all, 10 gave polymorphic markers and were applied to all 164 segregating progenies. Total number of markers scored was 21. The TF77 primer combination was the most productive (5 markers per amplification) followed by the TF68 (3), and TF81 (3). Overall, scorable marker productivity was 1.6 markers per amplification.

Of the 21 markers scored, 16 (76%) were testcross markers (aa x ab or ab x aa) and 5 (24%) were intercross markers (Table 1; Fig. 1). Out of the 16 testcross markers generated, 10 (62%) and 6 (38%) markers were heterozygous in 'Clementine' and 'Orlando', respectively. Out of the 21 testcross markers, 6 (24%) showed significant distorted segregation, 2 markers from intercross, 2 markers from 'Clementine', and 2 markers from 'Orlando'. Of the 5 intercross markers, all were scored as dominant markers. The two primers, TF77 and TF81, gave all three types of segregations (segregating in female, male, or both). In this segregating population, we previously detected 4-allel segregation of the six molecular marker systems. However, no codominant AP2/EREBP type TF marker was detected based on parameters in this study.

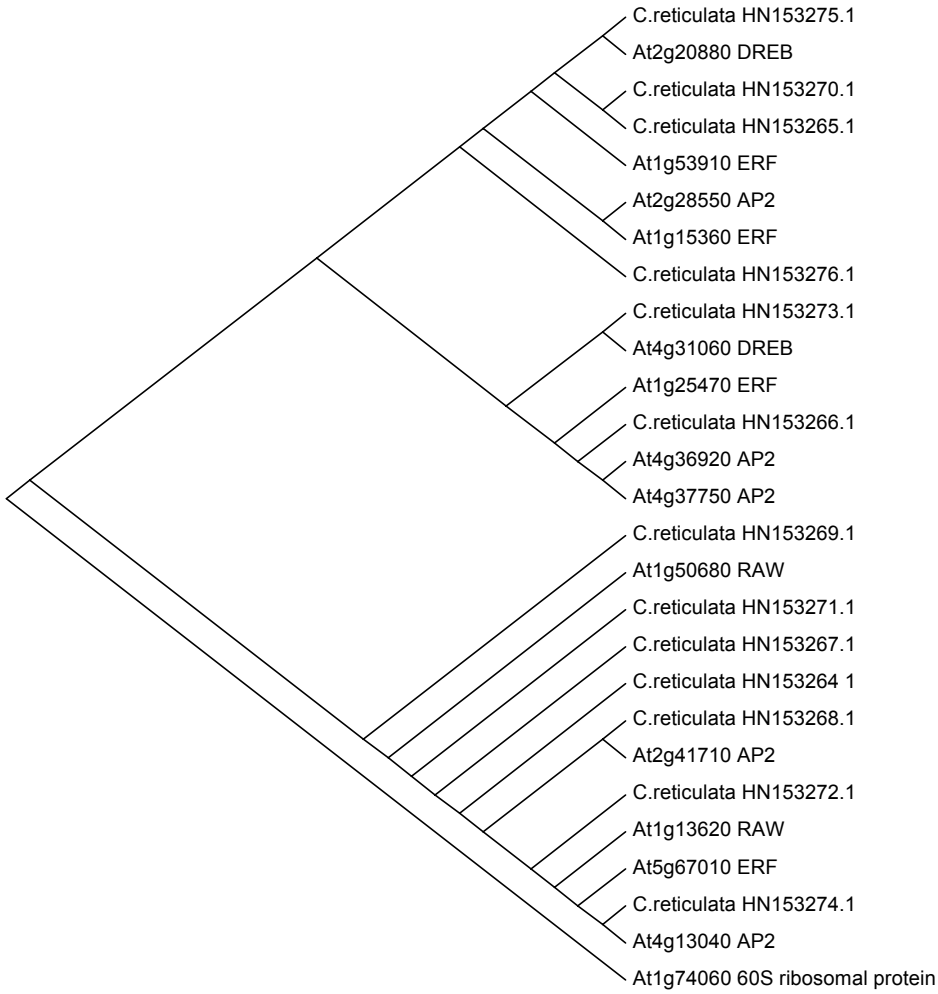
Fig. 1. Gel images of PCR products amplified with the AP2/EREBP type TF27 primers designed using the 8 dicotyledon plant species based on the DATFAP data base. The genotypes are as follow: 1, 'Orlando'; 2, 'Clementine', …….segregating progenies.



Sequencing and phylogenetic analysis of TFs identified in this study

From the sequencing study, 13 putative TFs were identified in this study, deposited into the www.ncbi.gov as GSS, and their gene bank numbers are HN153276.1, HN153275.1, HN153274.1, HN153273.1, HN153272.1, HN153271.1, HN153270.1, HN153269.1, HN153268.1, HN153267.1, HN153266.1, HN153265.1, and HN153264.1 (Table 2). Evolutionary relationships among the 13 TF sequences identified in this study, 13 known *A. thaliana* TFs, and one *A.thaliana* 60S ribosomal protein coding sequence (control) was inferred using the Maximum Parsimony method nested in MEGA 4 software (NEI & KUMAR 2000 [24]). Total tree length of most parsimonious tree was 235 (Fig. 2). The consistency index is 0.222, the retention index is 0.266, and the composite index is 0.059 for all sites and parsimony-

Fig. 2. Evolutionary relationships estimated by Maximum Parsimony tree among the 13 stress related TF sequences identified in this study and 13 known *A. thaliana* AP2/EREBP type TF family and one *60S* ribosomal protein as outgroup. The tree was constructed by using Mega version 4.0 genetic analysis software.



informative sites. The MP tree was obtained using the Close-Neighbor-Interchange algorithm with search level 2 (NEI & KUMAR 2000 [25]), in which the initial trees were obtained with the random addition of sequences (10 replicates). The codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated from the dataset (Complete Deletion option). There were a total of 175 positions in the final dataset, out of which 175 were parsimony informative.

The MP tree detected three main cluster of the sequences obtained in this study and known *A. thaliana* genes. At1g74060 60S ribosomal protein cDNA was a typical outgroup (bottom of Fig. 2). Cluster 1 had four *C. reticulata* TF genes (HN153275.1, HN153270.1, HN153265.1, HN153276.1) and four *A. thaliana* (At2g20880 DREB, At1g53910 ERF, At2g28550 AP2, At1g15360 ERF). Cluster 2 consisted 2 *C. reticulate* TFs (HN253173.1, HN153266.1) and four *A. thaliana* TFs (At4g31060 DREB, At1g25470 ERF, At4g36920 AP2, At4g37750 AP2). Cluster 3 included seven *C. reticulata* TFs (HN153269.1, HN153271.1, HN153267.1, HN153264.1, HN153268.1, HN153272.1, HN153274.1) and five *A. thaliana* TFs (At1g50680 RAW, At2g41710 AP2, At1g13620 RAW, At5g67010 ERF, At4g13040 AP2). In general we identified diverse range of AP2, ERF, RAW, and DREB type TFs in citrus.

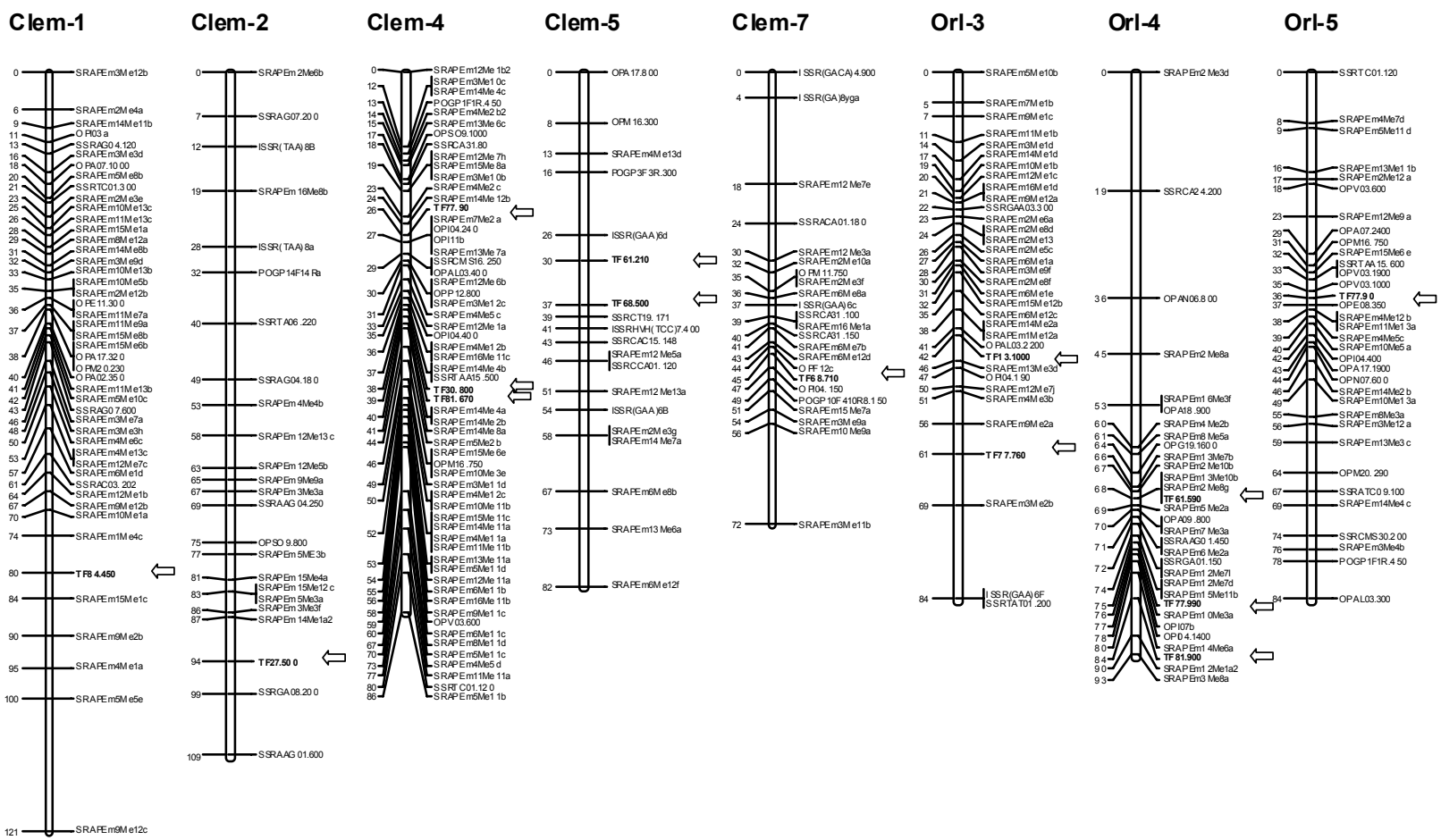


Fig. 3. Linkage groups of 'Clementine' mandarin and 'Orlando' tangelo consisting of TF, SRAP, SSR, ISSR, POGP, RGA, RAPD markers. Map distances in centiMorgans are indicated to the left and loci to the right of each linkage group. The 13 TF markers are indicated in bold-italics with and open arrows.

Map construction and distribution of segregating markers on linkage map

Of all bands observed, 50% were polymorphic, 13 were mapped in current map (Fig. 3). TF markers on the linkage map were indicated as TFx.xxx. Of these TF markers, 8 and 6 were mapped in the 'Clementine' and 'Orlando' maps, respectively. TF markers were located in 1., 2., 4., 5. and 7. linkage groups (LG) of 'Clementine' map, and 3., 4. and 5. LGs of 'Orlando' map. The 4. and 5. 'Clementine' LGs contained 2 TFs markers, and the 4. LG of 'Orlando' map included 2 TF markers. The other LGs consist of one TF markers.

The GS matrix was calculated based on Pearson correlation coefficiencies and the tree was constructed by using UPGMA procedure to estimate associations among the TF markers produced. In general no or little correlations were observed among TFs identified in this study Fig. 4). Few pairs indicated very strong correlation, which are TF30.800-TF81.670 and TF30.100-TF81.1100.

Discussion

The AP2/EREBP type TF primers were developed, applied, and their PCR targets were sequenced and the TF markers were linkage mapped in this study (Table 1 and 2, Fig. 3). The 12 of the TF markers were mapped in a previously constructed map of GULSEN & al. (2010 [15]). DENG & al. (2000 [25]) identified and cloned another stress related genes (NBS-LRR class resistance gene analogs) of citrus. GULSEN & al. (2010 [15]) mapped the POGP and RGA markers in citrus. This was the first comprehensive report where specific TF primers were used to study genetic linkage mapping in citrus. We previously construct a new linkage mapped based on six different marker systems (SRAP, SSR, ISSR, RGA, POGP, RAPD) (GULSEN & al. 2010 [15]). Now, new set of gene based marker system with TF-specific primers was developed for genetic analyses for dicotyledon plants. Overall, scorable marker productivity was 1.6 markers per amplification in this study, being very similar to RGA and POGP markers.

The 13 TF genes were similar to the known *A. thaliana* TF sequences as estimated by using the MP algorithm. Although AP2/EREBP type TFs were targeted by using specific primers in this study, the identified sequences were also related to DREB, RAW and ERF type TFs (Fig. 2), which were all related (FENG & al. 2005 [20]). Based on DNA sequences they have high level of sequence and functional similarities (RIECHMANN & al. 2000 [26]). This was consistent with our results.

Saturated genetic linkage maps with equal distribution of molecular markers are required and gaps in genetic linkage maps are not favorable. The gaps are caused by lack of polymorphism or insufficient number of markers. The genetic linkage map used in this study contains gaps despite utilization of high number of markers. For example, 1., 2., 5. and 7. LGs of 'Clementine' map and 3. and 4. LGs of 'Orlando' maps have relatively large size of gaps (>10 cM) (Fig. 3). The TF markers developed in this study little helped filling some of gaps on the current map. For instance, the TF84.450 marker filled the gap of 10 cM in the 'Clem-1'. Similarly, TF markers filled 12 cM in 'Clem-2, and 13 cM in Clem-5. Similarly, TF77.760 marker filled the gap of 13 cM in the 'Orlando's map. These results concluded that the TF markers could be successful in filling gaps in genetic citrus linkage maps.

In addition, some TF markers were mapped in the same region of LGs. In example, TF30.800 and TF81.670 markers were mapped within 1 cM in Clem-4 (Fig. 3). Similarly, two TF markers were mapped within 7 cM in the 'Clem-5' LG. These results indicated that ethylene responsive AP2/EREBP type TFs might be found in close distance, perhaps, originating from the same ancestral TFs by gene duplication as observed in peroxidase gene family. On the other hand, TF61.590, SRAPEm13Me10b, SRAPEm2Me8g were mapped in the same locus. Probably, the TF61.590 marker corresponds the same locus as these two SRAP markers. In

addition, Pearson correlation analyses were performed to estimate associations among the TF markers. Few TF markers were highly associated ($r \cong 1.00$; $P > 0.0001$) such as TF81.670-TF30.800 and TF81.900-TF30.1100 (Fig. 4). Few other pairs of the TF markers indicated relatively lower level of associations ($r \cong 0.5$). These TFs are probably linked in the LGs. The similar results were observed in the other stress related gene family, peroxidase (GULSEN & al. 2010 [15] & 2010 [27]).

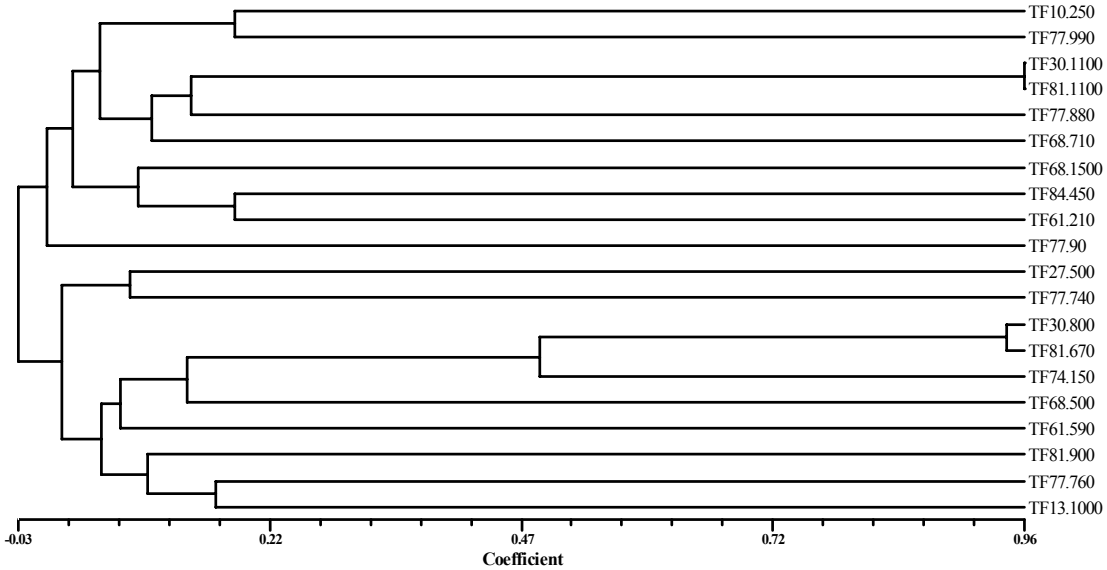


Fig. 4. UPGMA tree of 21 TF markers based on correlation coefficient as scored in 164 full-sib segregating progeny.

The TF primers were designed based on the 8 dicots such as tomato, cotton, potato, and *A. thaliana* TF sequences, and tested in this mapping population of citrus. The results also indicated that these TF primers are readily transferable to the other dicot plant species as peroxidase primers designed in rice cDNA sequences and transferred to citrus (GULSEN & al. 2007 [28]). Annealing temperatures were as low as 45 ^{o}C for some AP2/EREBP type TF primers in this study. It is probably due to relative distance between citrus genome and genomes of tomato, potato, and *A. thaliana*). Application of these new primers to the other dicots may likely vary in reannealing temperatures.

The transcription factor markers might be used to construct linkage maps and define relationships among plant genotypes from different geographic locations. Plant species from different locations are exposed to varying biotic or abiotic stresses that could enhance developmental variations in the TF genes due to varying selection pressure. Therefore, TF marker diversity and relationships likely reflect exposure to stress factors. There are 68 TF gene families (SHIGYO & al. 2006 [2]). As one of them, the AP2/EREBP type TF genes contain 146 members in a single plant species that vary in sequence and function (FENG & al. 2005 [20]). The 26 AP2/EREBP type TF primers designed in this study may be helpful in identifying the AP2/EREBP TF genes in citrus and the other dicots, and may also be helpful in studying gene expression profiles in the presence of abiotic or biotic stresses (Table 1). More information is needed regarding TF gene expression (Table 1). We sequenced the 13 markers developed in this study, extending the number of the TF sequences of citrus in gene bank (Table 2). We added 13 TF genomic sequences to the 24 TF sequences available in the NCBI. As more TF genes are identified, microarray technology may have a greater potential to improve our understanding of TF gene expression due to conserved sequences among genes.

Of the 13 deposited gene sequences in this study, only one (HN153276.1) was similar to the TF sequence (EU883665.1) of *Poncirus trifoliata,* a closely related species of *Citrus*. This implies that most of the TF sequences obtained in this were different from the citrus genomic TF sequences available in the NCBI.

In this study, PCR-based TF markers were used to saturate genetic linkage map of citrus. Furthermore, few of the TF markers developed filled relatively large gaps (≥10cM) in the previous map of citrus. We targeted the AP2/EREBP type TFs because of their importance in developmental physiology and stress tolerance. The targeted gene family approach provides advantages over the use of random primers or anonymous loci to study evolution, mapping and diversity. The targeted loci reveal polymorphism in genes that have a characterized biological role in the plant. This approach may be applied to the study of other gene families that carry conserved motifs similar to the peroxidases.

The 26 AP2/EREBP type TF primers were designed, 10 of which produced scorable polymorphic fragments in this study. The genus *Citrus* is relatively distant to the plant species (listed in M&M) used for primer design. The other 16 primers may have potential in other dicot plant species.

To our knowledge, this is the first description of designing AP2/EREBP type TF primers and of mapping these important genes of citrus. We aimed the AP2/EREBP genes responsive to ethylene concentrations usually modified with biotic and abiotic stress factors such as low or high temperatures, salt, insect, fungus, and seed mass (RIECHMANN & MEYEROWITZ 1998 [29]; KASUGA & al. 1999 [30]; HASEGAWA & al. 2000 [31]; OHTO & al. 2005 [32]; LIN & al. 2007 [33]). Therefore these primers may reflect environmental constrains, and have potential in estimating diversity and evolutionary relationships among plant genotypes, constructing genetic linkage maps, identifying new TFs in other plant species, and designing single nucleotide polymorphism markers.

Conclusion

Regulation of gene expression is critical for a variety of essential processes in plants, such as growth, development, differentiation, metabolic regulation, and stress tolerance (VERMA 1992 [1]). Transcription, the first step in gene expression, plays a central role in the regulation of the expression of genes. Therefore, the analysis of transcription factors (TF) is essential for an understanding of mechanisms of gene expression and their genome organization. We developed forward and reverse AP2/EREBP type TF-specific primers in this study, sequenced some of these TFs for verification, and linkage mapped in citrus. Of the 21 markers, 13 were linkage mapped in a previous citrus map. The 13 TF sequences obtained in this study were found to be closely related to known TFs of Arabidopsis thaliana, and few of TF markers were found to be closely linked in existing linkage map of citrus. These AP2/EREBP primers helped identification of citrus AP2/EREBP type transcription factor genes and can be used in other dicots such as tomato and cotton may have potential in understanding evolutionary relationships, establishing linkage map, and estimating diversity among other dicots since these TFs may reflect adaptability of plants.

Acknowledgements

This study was funded by the Scientific Research Commission of Erciyes University (ERU-BAB) with the project code FBA08-419 and the General Directorate of Agricultural Research of Ministry of Agriculture and Rural Affairs, Turkey.

References

1. VERMA, D.P.S., Control of Plant Gene Expression. The CRC Press, Boca Raton, Florida, USA, (1992).
2. SHIGYO, M., HASEBE, M., ITO, M., Molecular evolution of the AP2 subfamily. Gene 366, 256-265, (2006).
3. WILSON, K., LONG, d., SWINBURNE, J., COUPLAND, G., A dissociation insertion causes a semidominant mutation that increases expression of TINY, an Arabidopsis gene related to APETALA2. Plant Cell. 8, 659–671, (1996).
4. OKAMURO, J.K., CASTER, B., VILLARROEL, R., VAN MONTAGU, M., JOFUKU, K.D., The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in Arabidopsis. Proc. Nat. Acad. Sci. 94, 7076-7081, (1997).
5. ZHOU, J., TANG, X., MARTIN, G.B., The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that bind a cis-element of pathogenesis-related genes. EMBO J. 16, 3207–3218, (1997).
6. STOCKINGER, E.J., GILMOR, S.J., THOMASOW, M.F., *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. Proc. Nat. Acad. Sci. 94, 1035–1040, (1997).
7. LIU, Q., KASUGA, M., SAKUMA, Y., ABE, H., MIURA, S., YAMAGUCHI-SHINOZAKI, K., SHINOZAKI, K., Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. Plant Cell. 10, 1391-1406, (1998).
8. OHME-TAKAGI, M, SHINSHI, H., Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. Plant Cell. 7, 173-182 (1995).
9. LIU, Y., ZHAO, T.J., LIU, J.M., LIU, W.Q., YAN, Y.B., ZHOU, H.M., The conserved Ala37 in the ERF/AP2 domain is essential for binding with the DRE element and the GCC box. FEBS Lett. 580, 1303-1308, (2006).
10. GUERRA, M.S. Cytogenetics of *Rutaceae*. II. Nuclear DNA content. Caryologia 37, 219-226, (1984).
11. BARRETT, H.C., Hybridization of *Citrus* and related genera. Fruit Var. J. 39, 11-16, (1985).
12. CHEN, C., BOWMAN, K.D., CHOI, Y.A., DANG, P.M., RAO, M.N., HUANG, S., SONEJI, J.R., MCCOLLUM, T.G., GMITTER JR, F.G., EST-SSR genetic maps for *Citrus sinensis* and *Poncirus trifoliata*. Tree Genet. Genom. 4, 1-10, (2007).
13. ROOSE, M.L., Mapping and marker assisted selection in Citrus. In: Khan, I.A. (ed). Citrus Genetics, Breeding and Biotechnology. CAB International Wallingford UK pp., 275-286 (2007).
14. YAMAMOTO M., TOMINAGA, S., Chromosome identification in haploid clementine (*Citrus clementine* hort. Ex Tanaka) by fluorescent staining. Scientia Horticulturae 101, 201-206, (2004).
15. GULSEN, O., UZUN, A., CANAN, I., SEDAY, U., CANIHOS, E., A new citrus linkage map based on SRAP, SSR, ISSR, POGP, RGA and RAPD markers. Euphytica 173, 265-277, (2010)
16. DALKILIC, Z., TIMMER, L.W., GMITTER, JR. F.G., Linkage of alternaria disease resistance gene in mandarin hybrids with RAPD fragments. J. Amer. Soc. Hort. Sci. 130, 191-195 (2005).
17. FREDSLUND, J. DATFAP: A database of primers and homology alignments for transcription factors from 13 plant species. BMC Genomics 9, 140-149 (2008).
18. DOYLE, J.J., DOYLE, J.L., Isolation of plant DNA from fresh tissue. Focus 12, 13-15, (1990).
19. KUMAR, S., TAMURA, K., NEI, M., MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. Briefings in Bioinformatics 5, 150-163, (2004).
20. FENG, J.X., LIU, D., PAN, Y., GONG, W., MA, L.G., LUO, J.C., DENG, X.W., ZHU, Y.Z., An annotation update via cDNA sequence analysis and comprehensive profiling of developmental, hormonal or environmental responsiveness of the Arabidopsis AP2/EREBP transcription factor gene family. Plant Mol. Biol. 59, 853-68, (2005).
21. VAN OOIJEN, J.W., JoinMap ® 4.0, Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen, Netherlands, (2006).
22. VOORRIPS, R.E., MapChart: Software for the graphical presentation of linkage maps and QTLs. Heredity 93, 77-78, (2002).
23. ROHLF, F.J., NTSYS-PC, numerical taxonomy and multivariate analysis system. Version 1.8. Exeter Software, Setauket, New York, (1993).
24. NEI, M., KUMAR, S., Molecular Evolution and Phylogenetics. Oxford University Press, New York. (2000).

25. DENG, Z., HUANG, S., LING, P., CHEN, C., YU, C., WEBER, C.A., MOORE, G.A., GMITTER, JR., F.G., Cloning and characterization of NBS-LRR class resistance-gene candidate sequences in citrus. Theor. Appl. Genet. 101, 814-822, (2000).
26. RIECHMANN, J.L., HEARD, J., MARTIN, G., REUBER, L., JIANG, J.,Z., KEDDIE, J., ADAM, L., PINEDA, O., RATCLIFFE, O.J., SAMADA, R.,R., CREELMANN, R., PILGRIM, M., BROUN, P., ZHANG, J.Z., GHANDEHARI, D., SHERMAN, B.K., YU, G.L., Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290, 2105–2110, (2000).
27. GULSEN, O., KAYMAK, S., OZONGUN, S., UZUN, A., Genetic analysis of Turkish apple germplasm using peroxidase gene-based markers. Scientia Horticulturae 125, 368–373. (2010)
28. GULSEN, O. SHEARMAN, R.C., HENG-MOSS, T.M., MUTLU, N., LEE, D.J., SARATH, G., Peroxidase gene polymorphism in buffalograss and other grasses. Crop Sci. 47, 767-772. (2007).
29. RIECHMANN, J.L., MEYEROWITZ, E.M., The AP2/EREBP family of plant transcription factors. Biol. Chem. 379, 633-646 (1998).
30. KASUGA, M., LIU, Q., MIURA, S., YAMAGUCHI, K., SHINOZAKI, K., Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-induced transcription factor. Nat. Biotechnol. 17, 287–291, (1999).
31. HASEGAWA, P.M., BRESSAN, R.A., ZHU, J.K., BOHNERT, H.J., Plant cellular and molecular responses to high salinity. Ann. Rev. Plant Physiol. Plant. Mol. Biol. 51, 463-499, (2000).
32. OHTO, M., FISHER, R.L., GOLDBERG, R.B., NAKAMURA, K., HARADA, J.J., Control of seed mass by APETALA2. Proc. Nat. Acad. Science 102, 3123-3128, (2005).
33. LIN, R., ZHAO, Z., MENG, X., PENG, P.L., Molecular cloning and characterization of a rice gene encoding AP2/EREBP-type transcription factor and its expression in response to infection with blast fungus and abiotic stresses. Physiol. Mol. Plant Pathol. 70, 60-68, (2007).