African Journal of Biotechnology Vol. x(xx), pp. xxx-xxx, xx May, 2012
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 ©2012 Academic Journals
DOI

Full Length Research Paper

Tomato landraces are less sensitive to assessment of *in silico* BAC-based simple sequence repeat (SSR) marker development for tomato (*Solanum lycopersicum* L.)

Lana W. Al-Qadumii[1,2], Monther T. Sadder[1,3] and Hussein Migdadi[3*]

[1]Department of Horticulture and Crop Science, Faculty of Agriculture, University of Jordan, Amman 1142, Jordan.
[2]Faculty of Science, Philadelphia University, Jerash, Jordan.
[3]Department of Plant Production, College of Food and Agricultural Sciences, King Saud University, P. O. Box 2460, Riyadh 11451, Saudi Arabia.

Accepted 23 May, 2012

Tomato landraces are less sensitive to environmental stresses and grown mainly under rain fed conditions. They are still grown in small farms due to quality and special demand of consumers. These landraces are valuable sources of genetic traits, and plant breeders can use breeding programs for crop improvement. One of the primary needs of the crop improvement is the estimation of genetic diversity. Development of microsatellite (SSR) markers from map-referenced BAC clones is a very effective means of targeting markers to marker scarce positions in the genome. This study was aimed at developing a set of functional SSR markers via *in silico* analysis of publicly available tomato DNA sequences. As a result, 17 SSR markers were developed and tested on one tomato commercial cultivar and eight local landraces. 12 loci (27 alleles) were scored and showed 100% polymorphic patterns. The calculated PIC values for the SSR markers developed ranged from 0.62 to 0.97 (mean 0.89). The SSR motifs CT(26) AT(27) and TTC(6) TTA(4) had the highest PIC value (0.97), while CAA(5)A(8) had the lowest PIC value (0.62). According to tomato expressed sequence tag (EST) analysis, some of these developed SSR markers, such as mono and di-nucleotide are related to some genes. The T(16) motif is related to hydroxyproline-rich glycoprotein, which is a family protein from *Arabidopsis thaliana*. On the other hand, the SSR with tri-nucleotide repeat motif AAC(4)A(11) was related to a putative homologous protein to A7Q2S4 from *Vitis vinifera.*

Key words: Tomato landraces, *in silico* simple sequence repeat (SSR) markers, DNA markers, genetic diversity.

INTRODUCTION

Tomato (*Solanum lycopersicum* L.) is one of the most widespread vegetable crops around the world. It is an annual herbaceous plant belonging to the Solanaceae family. Worldwide, tomato is considered as the second most consumed vegetable after potato (FAO, 2008). Wild tomato species are a rich source of desirable genes. They have been utilized in breeding programs to improve the cultivated tomatoes. One of the primary needs of the crop improvement is the estimation of genetic diversity among and within genotypes. To this end, various marker techniques have been applied, either individually or in combinations to study the genetic diversity of various plant species, especially tomato (Tam et al., 2005).

In the last decade, several molecular markers have been developed in tomato such as restriction fragment length polymorphisms (RFLP) (Stevens et al., 1995), random amplification of polymorphic DNA (RAPD) (Smeiech et al., 2000), inter simple sequence repeat

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

(ISSR) (Zietkiewicz et al., 1994) and amplified fragment length polymorphism (AFLP) (Vos et al., 1995).

Among the wide variety of available marker systems, simple sequence repeats (SSRs, also called microsatellites), which are short tandem repeat units of between 1 and 6 bp in length (Tautz, 1989), offer a number of advantages which have made them increasingly popular in plant and animal studies. SSRs are highly polymorphic due to variation in repeat number and are co-dominantly inherited (Rafalski et al., 1996). Most SSRs are single-locus markers, and many SSR loci are multi-allelic. These characteristics made SSRs ideal markers not only for creating genetic maps, but also as unambiguous means of defining linkage group homology across mapping populations (Song et al., 2004). SSR markers are becoming the preferred molecular marker for variety identification in tomato. The presence of SSRs in the transcripts of genes suggests that they might have a role in gene expression or function; however, it remains to be seen whether any unusual phenotypic variation might be associated with the length of SSRs in coding regions (Varshney et al., 2005).

Isolation of microsatellite clones by the enrichment procedure (Ostrander et al., 1992; Edwards et al., 1996) makes development of markers more efficient than before. However, *in silico* mining of SSRs from sequence databases (Temnykh et al., 2001) provides a promising alternative to the molecular approaches. Development of SSR markers from map-referenced clones of bacterial artificial chromosomes (BAC) was a very effective means of targeting markers to marker scarce positions in the genome (Song et al., 2004). The *in silico* approach is not only time and cost effective, but also it allows for the discovery of SSRs from expressed sequence tags (ESTs) that represent the coding region of the genome. EST-SSR markers are thus potential candidates for gene tagging and comparative genetic studies in related species. This study was aimed at developing and characterizing new set of SSR markers for *S. lycopersicum* L. and to assess the genetic variation among tomato landraces using the new developed SSR primers.

MATERIALS AND METHODS

SSR marker development

SSR markers were developed from BAC end sequences from tomato BAC libraries randomly chosen from SGN (SGN, 2009). For each chromosome, one BAC was picked randomly. The identification of SSR motifs was achieved using vector NTI program which were T(n), CT(n), AT(n), CAA(n), GAA(n), TTC(n), TTA(n), AAG(n), TCT(n) and CTT(n). Primers were designed according to the vector NTI default parameters with regards to spacing to amplify fragments of 100 to 300 bp, and high and constant annealing temperature (40 to 50°C). Primers were longer than 20 nucleotides (nt) and had at least 50% GC content (for stability), variation in the number, types (mono, di, and tri) and length of SSR motif. Oligonucleotide primers were synthesized by the Midland Certified Reagent Company INC.

Plant and DNA sources

Seeds of nine tomato landraces representing the geographical distribution in Jordan kindly provided by National Center for Agricultural Research and Extension (NCARE) Genebank and one commercial cultivar (Gardenia) were used. Seeds were grown under *in vitro* conditions. DNA isolations were conducted using leaf tissue (0.35 g) according to Saghai-Maroof et al. (1984).

PCR amplification

All PCR reactions were performed in 10 µl PCR reactions containing 20 ng genomic DNA, 0.2 mM of each dNTP, 2.5 mM $MgCl_2$, 1X PCR buffer, 0.1 U *Taq* polymerase (Biobasic) and 0.4 mM for forward and reverse primer.

To screen the SSR primers for optimal annealing temperature, thermal cycling profiles were as follows: denaturation was at 94°C for 2 min, followed by 35 cycles consisting of denaturation at 94°C for 30 s, followed by annealing at 40.1, 40.7, 41.5, 42.7, 44.0, 45.3, 46.7, 48.0, 49.1, 49.9 and 50.4°C for 30 s, and extension at 72°C for 30 s. For SSRs validation PCR, profiles were as follows: denaturation was at 94°C for 2 min, followed by 35 cycles consisting of denaturation at 94°C for 30 s, followed by annealing at 43.3°C for 30 s and extension at 72°C for 30 s with final extension at 72°C for 6 min.

PCR products were separated on 2% agarose A (Biobasic) at 100 V for 100 min; all gels were run in 1X TAE buffer and were visualized using ethidium bromide staining and with UV transillumination (Alpha Imager).

SSR marker screening and data analysis

All previous primers were screened on a panel of nine tomato landraces accessions and one commercial cultivar was used as the control. Polymorphism information content (PIC) value for each SSR was estimated using Anderson et al. (1993) equation: PIC = 1- $\sum_{j=1}^{n} Pij2$, Where Pij is the frequency of the ith allele for marker j and the summation extends over n alleles, calculated for each SSR locus.

The SSR products from the tested landraces and commercial cultivar were scored as band presence (1) and absence (0), thus generating a binary matrix. The binary data matrix was used to compute Jaccard pair-wise similarity coefficients (Jaccard, 1908). The similarity matrices obtained were utilized to construct a UPGMA-based dendrogram (Sneath and Sokal, 1973). The analyses were performed using NTsys software package version 2.10 (Rohlf, 1998). The used BACs were screened for putative genes, close to the selected *in-silico* SSRs. If there were no putative genes available, the BACs were blasted using the available tomato ESTs (SGN, 2009). In addition, the data were presented as aligned along the BAC scheme showing the position of selected SSRs using vector NTI software. (Invitrogen, USA).

RESULTS

In silico screening and analysis

Information from BAC libraries were obtained directly from SGN GenBank and randomly chosen (SGN, 2009).

Table 1. Selected chromosome, BAC accession, repeats (motif) and types of repeat.

Chromosome number	BAC SGN number (GenBank number) [Size kbp]	Repeat	Type of repeat
1	C01HBa0003D15.1 (AC193776) [97.174]	T(16) AAC(4) A(10)	Mono- nucleotide Tri - mono - nucleotide
3	C03HBa0030F10 (EU124744) [120.802]	CT(26) AT(27) TA(15)	Di- nucleotide Di- nucleotide
5	C05HBa0189E17 (AC212302) [102.6]	T(18)A(7) CT(14)	Mono- nucleotide Di- nucleotide
6	C06HBa0203N09 (AC219216) [108.754]	(CAA)5A8 (GAA)4	Tri- nucleotide Tri- nucleotide
7	C07HBa0002M15 (AC210346) [138.275]	AT(6)GT(7) GC(7)	Di- nucleotide Di- nucleotide
8	C08SLm0058F09 (AP009607) [110.461]	C(7)A(7)	Mono- nucleotide
9	C09SLe0103M07 (EF647597) [90.836]	(TTC)6(TTA)4	Tri- nucleotide
10	C10SLe0045H11 (AC193781) [109.965]	(AAG)4 A7 (TCT)5	Tri- nucleotide Tri- nucleotide
11	C11SLe0053P22 (AC212315) [105.178]	AT(29) AT(20)(ATT)10T(7)	Di- nucleotide Di and Tri- nucleotide
12	LE_HBa-165B12 (AC212767) [121.526]	TA(7)A(9) (CTT)5	Di- nucleotide Tri- nucleotide

For each chromosome, one BAC was randomly picked, and searched along the genome for simple sequence repeats. Repeats features are represented in Tables 1 and 2.

Gene scan for the BAC C01HBa0003D15.1, showed gene in the T (16) repeat SSR that is related to hydroxyproline-rich glycoprotein family protein from *Arabidopsis thaliana*, containing InterPro domains: IPR008997: Ricin B-related lectin. In addition, it contains a gene within the AAC(4) A(11) repeat SSR that is a putative homologous protein to A7Q2S4 from *Vitis vinifera*; containing InterPro domains: IPR007524: Pectate lyase, N-terminal; IPR012334: pectin lyase fold. Gene scan for the BAC C03HBa0030F10, showed gene in the CT (26) AT (27) repeat SSR which is related to DNA-binding family protein from *A. thaliana*; containing InterPro domains: IPR005175: protein of unknown function DUF296. Gene scan for the BAC C08SLm0058F09, showed gene within the C (7) A (7) repeat SSR that is related to a hypothetical protein; containing InterPro domains: IPR001878: zinc finger, CCHC-type; IPR005162: retrotransposon gag protein; IPR013242: retroviral aspartyl protease. Gene scan for the BAC C09SLe0103M07, showed gene close to the C (7) A (7) repeat SSR which is related to MCE1_MOUSE mRNA capping enzyme from *Mus musculus*; containing InterPro domains: IPR012310: ATP dependent DNA ligase, central; IPR017074: mRNA capping enzyme, bifunctional. Gene scan for the BAC C10SLe0045H11, showed two gene bordering the (AAG) 4 A7 repeat SSR, which are related to a putative homologous protein to CNGC4_ARATH Cyclic nucleotide gated ion channel 4 from *A. thaliana*; containing InterPro domains: IPR000595: cyclic nucleotide-binding; IPR014710: RmlC-like jelly roll fold and in addition, a putative homologous protein to B0CN63_NICBE Myosin VIII 2 from *Nicotiana benthamiana*; contains InterPro domains: IPR000048: IQ calmodulin-binding region. The (TCT) 5 repeat SSR was found in a hypothetical protein. Gene scan for the BAC C11SLe0053P22, showed a gene within the AT (29) repeat SSR, which is a hypothetical protein and gene scan for the BAC LE_HBa-165B12, showing a few putative genes around the TA (7) A (9) repeat SSR. Gene scan for the BAC C05HBa0189E17, BAC

Table 2. The repetitive units (motif), 5'-3' primer and the optimum annealing temperature for each forward primer.

Repeat motif	Forward 5' primer sequence	Reverse 3' primer sequence	Annealing temperature (°C)
(T)16	ACTGAAACTTCTTTGCACTT	GTTATAAAATTTGCGATAAATT	43.3
(AAC)4 (A)10	AAACACAATGTTTGAACCGA	TGGGACTAATGAAGCTAACC	43.3
(CT)26 (AT)27	ACTGATTTACCTTTCACCAC	GGGAAAGAAACAAAAGTACA	43.3
(TA)15	AAGTGTCTAATAGTAAGAGTCTCAG	TGATGATCAGATTGAGAAGA	52.3
(T)18(A)7	GCTATCTTTTATCCAAGAGA	GTTGATTATTTTATATTAAAAAGT	43.3
(CT)14	TGATTTATTAGCTCAGGTATGC	GCATACCTGAGCTAATAAATCA	44.3
(CAA)5(A)8	TAACTCACCCAACTATGATT	AGGTACTCACATACTAATAGTTTT	43.3
(GAA)4	ATACAACATGCAGATATGAGTG	GTTGCTTGATTTCAGGAGAA	43.3
(AT)6(GT)7(GC)7	AGTCATTCTAACGTATCGAT	GCATTTTCCTTACACTTTAG	43.3
(C)7(A)7	GTGGCTCTGCCTTAATCATT	AGGCATGGTAAGAGAGAAAA	43.3
(TTC)6(TTA)4	TAGTTAGGTAATTGTAGGGG	GTAGCTCCAAAACTCTAAAT	43.3
(AAG)4(A)7	GGTGCAGAGGTTATGTCATA	TCGAAAATCTTTAGGTGATC	43.3
(TCT)5	ACAATTAATCTTGTTTAACTGC	AAAGTTGTTGTCGAAGAAGA	43.3
(AT)29	TTGGCAGCTCAGGACTGTTT	CACGCGCATACCAGTCTCTA	53.0
(AT)20(ATT)10(T)7	TTGTAGCTGACGTACTTTCG	AATGGAAGGAGGGTTGATTT	43.3
(TA)7(A)9	GAACATCATTGCACTCATCT	CACGTGAATCAGATTTATGA	46.3
(CTT)5	TCAATTAGTTGCGATTAAGG	AGGTTGAGAATTATACCCTC	45.6

C06HBa0203N09 and BAC C07HBa0002M15 showed no detectable genes within the selected SSRs, however, several ESTs were detected along the BACs.

PCR analysis

Genomic DNA was extracted from 20 tomato landraces and cultivar and was screened via gel electrophoresis after 5S PCR test. Only eight landraces and the commercial cultivar (6) showed better quantity and quality after 5S PCR test and were used in validation of developed SSR primers. The results show that di-nucleotide repeats were found to be the most abundant (44%) followed by tri-nucleotide (37%) and mononucleotide repeats (19%). The most common repeats observed in our study were: AT (47%), CT (35%), CA (17.6%), GA (12%), T (12%) and GT (6%).

To standardize the optimal temperatures for primer annealing, a set of temperatures were screened for each marker using gradient PCR and run on gel electrophoresis. The SSR motives that gave a positive results were: (CAA)5(A)8 , (TA)15 (GAA)4, (AT)6(GT)7(GC)7, (AT)29, (AT)20(ATT)10(t)7, (AAC)4 A10, (AAGA)7, T16, T18 A7, (CT)26 (AT)27, C7 A7 and (TCT)5. The (CT)14, (TTC) 6(TTA) 4, (TA) 7 A9 and (CTT) 5 failed to amplify the tested cultivar and the optimal annealing temperature recorded was 43.3°C. Out of the 17 SSR primers tested, 14 were successfully amplified in the tested cultivar. Seven showed polymorphic pattern (50%), and the remaining primers were monomorphic (50%). Examples of amplification profiles are presented in Figures 1 and 2.

Validation of 12 microsatellite markers for identity test in tomato landraces

The number of alleles and the PIC value for each SSRs type (motif) are presented in Table 3. Number of alleles varied from 1 to 6 (mean 2.25). PIC values ranged from 0.62 to 0.97 (mean 0.89). The SSRs type (CT) 26 (AT) 27 and (TTC) 6(TTA) 4 had the highest PIC value (0.97), while (CAA) 5(A) 8, had the lowest PIC value (0.62). The usefulness of SSR markers was investigated by assessing genetic similarity (GS) among the tomato landraces and commercial cultivar for diversity analysis. A total of 12 loci (27 alleles) were scored. All these loci (100%) were polymorphic (Table 3). Assessed genetic variation between tomato landraces and commercial cultivars was based on Jaccard genetic similarity and unweighted pair group method with arithmetic averages (UPGMA). GS varied from 0.0 between landrace 19 with both landraces 14 and 38 to 0.86 between landraces 22 and 29 (Table 4). At the average genetic similarity coefficient (0.56), the landraces were grouped into five separated clusters, where landraces 19, 24 and 38 formed a single cluster, while landraces 14 and 21 grouped into a separated cluster and the third cluster contained the remaining landraces with commercial cultivar (Figure 3).

DISCUSSION

The recent innovation and development in sequencing technology have fostered a new genomics era. Several consortiums have been initialized to crack the code of

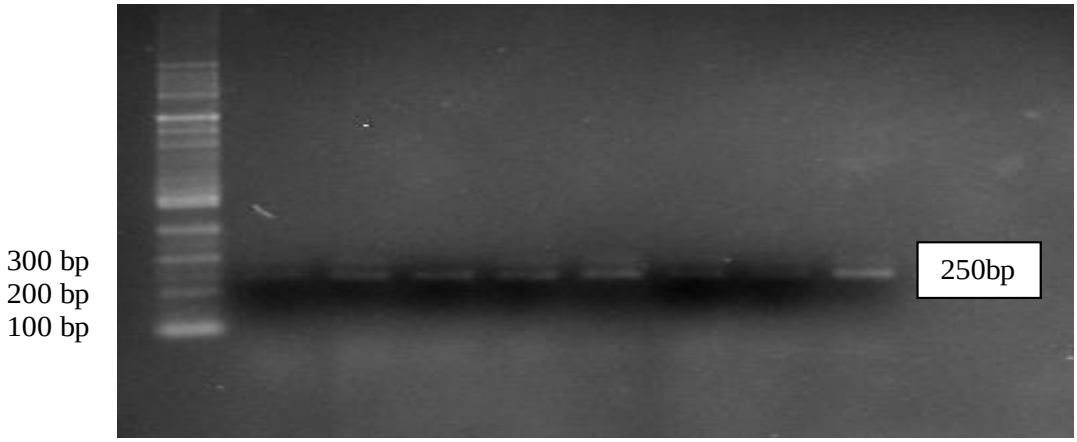

Figure 1. Detecting the presence and the polymorphism of the marker (AAC)4 A10 among the tomato landraces (19, 21, 22, 24, 31, 33a) and the commercial tomato cultivar (6). PCR reaction was run at 43.3°C. Agarose (2%) was run at 100 V for 100 min. L = Mass-ruler marker.

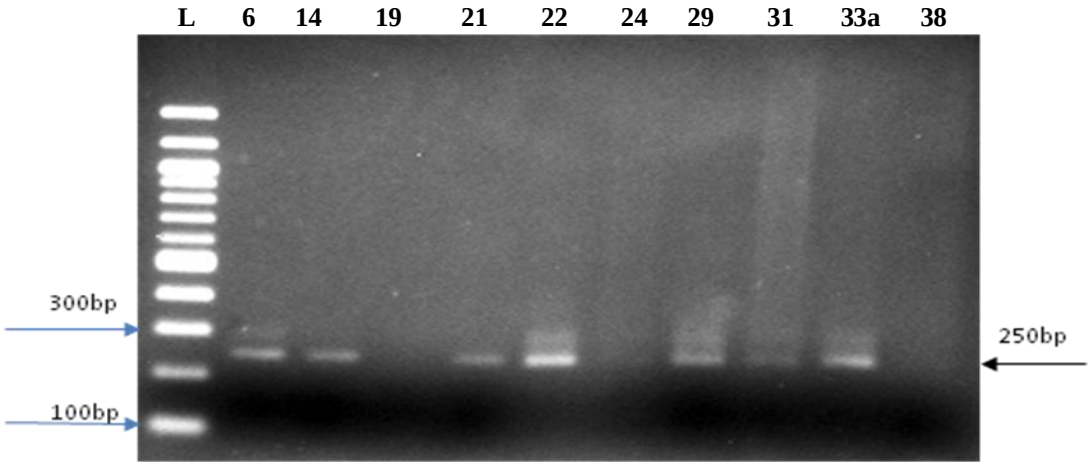

Figure 2. Detecting the presence and polymorphism of the marker GAA(4) among tomato landraces (14, 19, 21, 22, 24, 29, 31, 33a, 38) and the commercial tomato cultivar (6). PCR reaction was run at 43.3°C. Agarose (2%) was run at 100 V for 100 min. L = Mass-ruler marker.

Table 3. Allelic variation in 12 informative SSR loci.

SSR type	Number of alleles	PIC
(T)16	2	0.96
(AAC)4 (A)10	2	0.86
(CT)26 (AT)27	1	0.97
(T)18(A)7	3	0.93
(CAA)5(A)8	6	0.62
(GAA)4	2	0.87
(AT)6(GT)7(GC)7	2	0.90
(C)7(A)7	1	0.92
(TTC)6(TTA)4	1	0.97
(AAG)4(A)7	1	0.93
(TCT)5	2	0.88
(AT)20(ATT)10(T)7	4	0.93
Mean	2.25	0.895

Table 4. Average genetic similarity calculated as Jaccard coefficient for tomato landraces and commercial cultivar based on 27 SSRs alleles.

	6	14	19	21	22	24	29	31	33	38
6	1.00									
14	0.44	1.00								
19	0.13	0.00	1.00							
21	0.53	0.64	0.07	1.00						
22	0.71	0.36	0.10	0.50	1.00					
24	0.41	0.38	0.25	0.40	0.40	1.00				
29	0.75	0.38	0.11	0.45	0.86	0.42	1.00			
31	0.45	0.30	0.06	0.32	0.57	0.41	0.52	1.00		
33	0.67	0.33	0.14	0.35	0.70	0.57	0.74	0.67	1.00	
38	0.05	0.07	0.00	0.12	0.18	0.08	0.09	0.22	0.11	1.00

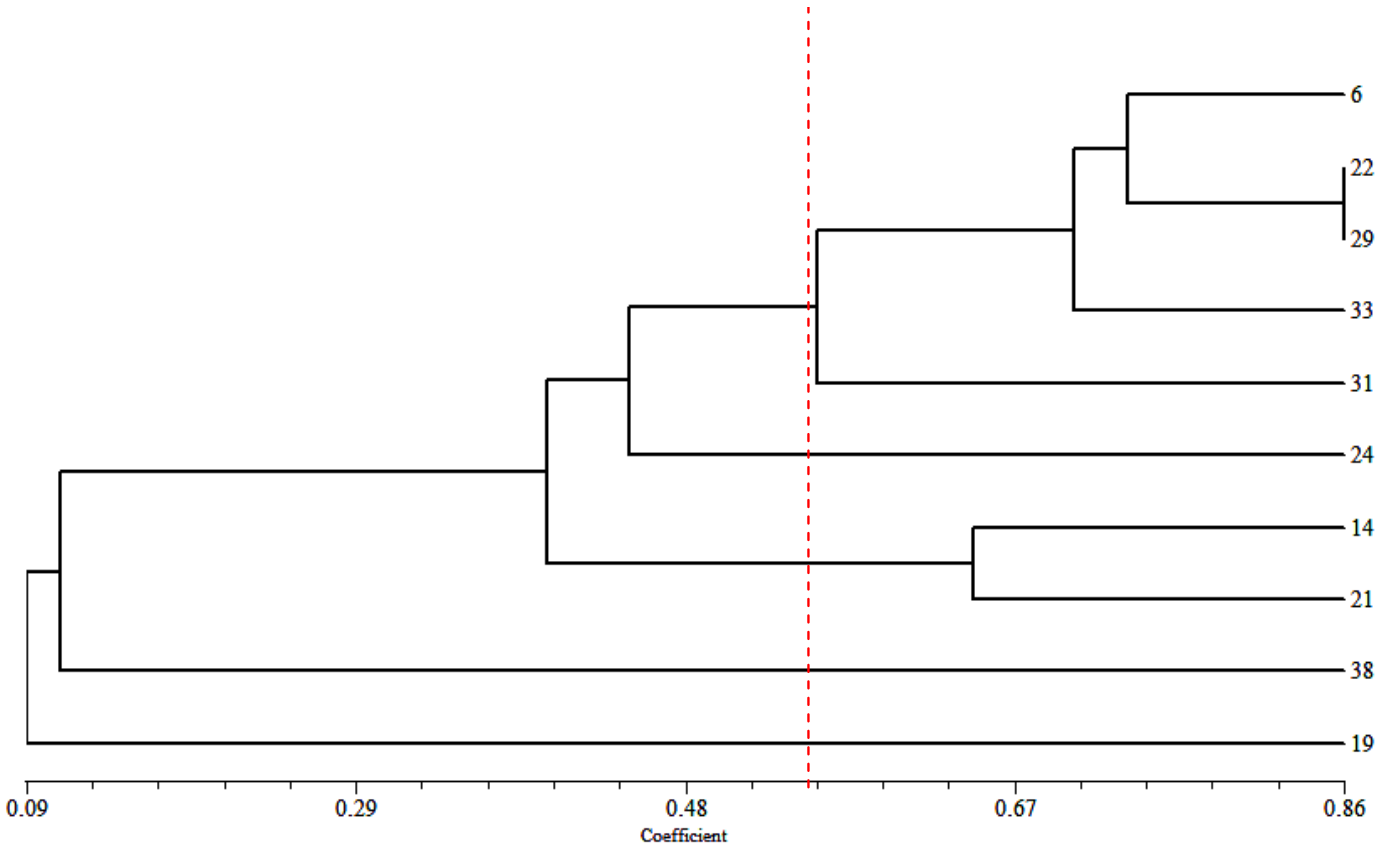

Figure 3. UPGMA dendrogram following Jaccard similarity coefficient of tomato landraces and commercial cultivars based on 27 SSRs alleles. Dotted line refers to the average of genetic similarity coefficient (0.56).

important organisms for human health and food resources. Tomato is one of these organisms, which is the mandate of the SGN. Huge data are being streamed into the tomato database; tomato is the centerpiece for genetic and molecular research of the Solanaceae, which is attributed in part to inherent features of the species, including diploidy, modestly sized genome (950 Mb), tolerance of inbreeding, amenability to genetic transformation, and the availability of well-characterized genetic resources (SGN, 2009).

Tomato chromosomes, like those of other species in the Solanaceae family, are composed of centromeric heterochromatin with more distal euchromatic regions. The centromeric heterochromatic regions of tomato

constitute about 77% of the chromosomal DNA and based on deletion studies, contain few genes (Paterson et al., 1996). However, in our random selection of the BAC clones, they all contained putative genes and several EST sequences. This is likely to be the case as the SGN selected the BAC clones for sequencing, first using available DNA molecular markers mapped on developed recombination genetic maps, and it is well documented that genetic maps cannot map the centromeres for any species. This means that the selected BACs by SGN (SGN, 2009) should map to the euchromatin areas of tomato chromosomes. If 7 kb/gene is characteristic of euchromatin, which constitutes 23% of the tomato genome (Paterson et al, 1996), then the euchromatic portions of the tomato chromosomes should contain about 31,000 genes [(950,000 kb/7 kb) x 0.23)], which is remarkably similar to the 35,000 genes estimated for the entire genome, and hence sequencing the euchromatic regions of the tomato genome would reveal the majority of these genes (Van der Hooven et al., 2002). Therefore, our selected SSR potentials would be highly valuable to breeding studies as they represent at least a portion of these 31,000 genes of tomato.

In addition, as available BAC sequences would probably map to euchromatin regions and, therefore, rich in active genes, the development of SSR markers *in silico* offers an unprecedented opportunity in recombination genetics. Analysis of polymorphism followed by PIC calculations for several SSR along each BAC clone or contig of group of BAC clones, would certainly, offer a revolutionary nano insights. This would represent an indirect molecular scope to investigate recombination frequencies along tiny stretches of tomato chromosomes. In turn, such data would deliver hotspots and correlated sequence stability factors. Such measures would benefit both molecular researchers and breeders concerned with desired traits finely mapped to similar tiny pins.

The PCR conditions were optimized for each primer pair, which would be very laborious when screening batches of putative SSR sequences. This was achieved by first testing different cycling conditions and then by varying the amount of DNA template, the concentration of primer, the use of different annealing temperature for optimum band detection and to detect informative bands in each primer with maximum polymorphism. Future investigation of scaled up nature would preferably involve some mechanization level to handle the sequential procedure. This requires an advanced genomic facility hosting semi automated and fully automated robotics for liquid handling and genotyping steps.

As several of our developed SSR markers were mapped to informative sequences, they are called EST-SSR putatives. The set of EST-SSR markers would be informative for phylogenetics analysis and genetic mapping. Moreover, since EST-SSRs lie within expressed sequence, they have the potential to serve as perfect markers for genes determining variation in phenotype. *In silico* analysis of the limited quantity which is publicly available has enabled the development of a set of functional SSR markers. Because these sequences are derived from the expressed portion of the genome, they are relevant for assaying functional diversity in populations or germplasm collections. The properties of the 17 SSR loci identified here were classified on the basis of repeat motif and the number of repeat units, and dinucleotide and trinucleotide were the most frequent (40%) for each and for mononucleotide, it was 20%. The most frequent motifs were AT and GA, followed by AC, GA was less frequent and the least were A and C as mononucleotides.

An analysis of the association between SSR motifs and the rate of polymorphism is important for the development of effective SSR markers (Yonemaru et al., 2009). Mc Couch et al. (2002) reported that AT-rich motifs in rice have a larger number of repeats and longer repeat tracks than other dinucleotide motifs and were associated with high rates of polymorphism. Yonemaru et al. (2009) also reported that the most repetitive SSR motif with a perfect repeat was (TAT) 65. In our study, AT-rich motif was the most repetitive sequence and the longest among our repeats and it also shows informative signs of polymorphism. This contradicts the euchromatin nature of the selected tomato BAC clones, as heterochromatin contain AT-rich sequences. This is not strange as it is important to differentiate between short stretches of AT repeats similar to SSRs and between highly repeated stretches of AT tandem repeats as documented for plant telomeres.

SSR polymorphisms are influenced by slippage mutation rates than repeat length, with SSR evolutionary age being a key factor for SSR diversity (Lai and Sun, 2003). We cannot however, rule out the possibility that insertions and deletions (INDEL) at regions other than the SSR motifs may account for some polymorphisms. AT and AG combinations of base pair motif types either in di- and tri- nucleotide configuration are the most abundant type in plants genomes (Cardel et al., 2000).

Previous studies conducted by Poysa et al. (2003), found PIC ranging from 0.09 to 0.67 and 0.98 as recorded by ALAbadi (2007). Tam et al. (2005) reported that the average polymorphism information content was 0.39 in collections of tomato industrial lines. These results are in agreement with that of our study where PIC values ranged from 0.62 to 0.97.

Thiel et al. (2003) have stressed some limitations of the application of SSR markers for diversity studies, emphasizing the possibility of homoplasy (identical allele sizes may not be identical by descent), and pointing out that alleles of different size can be generated by indel events, as well as by variation in the number of SSR repeats. However, the SSR marker-based genetic relationships among the tomato landraces used in this study agrees with those generated by other markers such as RAPD (Qian et al., 2001), AFLP (Suliman et al., 2002)

and SSRs (Poysa et al., 2003; Areshchenkova and Ganal, 2002; Mazzucato et al., 2008; Al-Abadi, 2007). With the exception of SSRs, limited information was obtained due to lack of variability which was ascribed to the self-pollinating nature of modern tomato cultivars combined with their narrow genetic base (Alvarez et al., 2001). A high level of polymorphism (95%) was revealed among 39 tomato landraces (Al-Abadi, 2007), which is almost similar (100%) to that found among tomato (Tam et al., 2005). Previous study conducted by Alvarez et al. (2001), found that polymorphism percentage (94%) in tomato was recorded. Accordingly, the results of the current study are in agreement with previous results and reveal that the variation among tomato landraces is attributable to genetic causes.

Conclusion

Map-referenced BAC clone is a very effective means of targeting SSR markers and *in silico* analysis of publicly available DNA sequence is an easy and very important means. We developed 17 SSR markers and they were tested on one tomato commercial cultivar and eight local landraces showed 100% polymorphic patterns. The SSRs motif CT(26) AT(27) and TTC(6) TTA(4) had the highest PIC value (0.97). Some of our SSR markers developed are related to some genes such as T(16) motif which is related to hydroxyproline-rich glycoprotein, a family protein from *A. thaliana*. On the other hand, some tri-nucleotide SSR repeats are coding for proteins such as AAC(4)A(11) motif which is putative homologous protein to A7Q2S4 from *V. vinifera.*

ACKNOWLEDGEMENT

We are grateful to the National Center for Agricultural Research and Extension (NCARE) for providing the tomato landraces seeds.

REFERENCES

Al-Abadi B (2007). Genetic variations among Jordanian tomato landraces using simple sequence repeat (SSR) and performance of their seedling to salinity stress. MSc. Thesis. Faculty of Graduate Studies. Jordan University of Science and Technology. Irbid-Jordan

Alvarez AE, van de Wie CCM, Smulders MJM, Vosman B (2001). Use of microsatellites to evaluate genetic diversity and species relationships in the genus *Lycopersicon*. Theor. Appl. Genet. 103: 1283-1292.

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993). Optimizing parental selection for genetic linkage maps. Genome. 36: 181-186.

Areshchenkova T, Ganal MW (2002). Comparative analysis of polymorphism and chromosomal location of tomato microsatellite markers isolated from different sources. Theor. Appl. Genet. 104: 229-235.

Cardel L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Wangh R (2000). Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics. 156: 847-854.

Edwards KJ, Barker JHA, Daly AC, Jones, Karp A (1996). Microsatellite libraries enriched for several microsatellite sequences in plants. Biotechniques. 20: 758-760

[FAO] (2008) <http://faostat.fao.org>.

Jaccard P (1908). Nouvelles reserches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. 44: 223-270

Lai Y, Sun F (2003). The relationship between microsatellite slippage mutation rate and the number of repeat units. Mol. Biol. Evol., 20: 2123-2131.

Mazzucato A, Papa R, Bitocchi E, Mosconi P, Nanni L, Negri V, Picarella ME, Siligato F, Soressi G P, Tiranti B, Veronesi F (2008). Genetic Diversity, Structure and Market- Trait Association in a Collection of Italian Tomato *(Solanum lycopersicum L.)* Landraces. Theor. Appl. Genet. 5: 657-669.

Mc Couch SR, Teytelman L, Xu Y (2002). Development and mapping of 2240 new SSR markers for rice (*Oryza sativa L.*). DNA Research. 9:199-207.

Ostrander EA, Jong PM, Rine J, Duyk G (1992). Construction of small-insert genomic libraries highly enriched for microsatellite repeat sequences. Proc Natl. Acad. Sci. USA. 89: 3419-3423.

Paterson AH, Lan TH, Reischmann KP, Chang C, Lin YR , Liu SC , Burow MD, Kowalski SP , Katsar CS , DeMonte (1996). Toward a unified genetic map of higher plants, transcending the monocot-dicot divergence. Nat. Genet. 14: 380-382.

Poysa V, He C, Yu K (2003). Development and characterization of simple repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. . Theor. Appl. Genet. 106: 363-373.

Qian W, Ge S, Houng DY (2001). Genetic variation within and among populations of a wild rice *Oryza granulata* from China detected by RAPD and ISSR markers. Theor. Appl. Genet. 102: 440-449.

Rafalski JA, Vogel JM, Morgante M, Powell W, Andre C, Tingey SV (1996). Generating and using DNA markers in plants. In: B. Birren and E. Lai, Editors, Non-mammalian Genomic Analysis. A Practical Guide, Academic Press, New York, pp. 75-134.

Rohlf FJ (1998). NTSYS-pc Numerical Taxonomy and Multivariate Analysis System version 2.02 User Guide.

Saghai-maroof SMA, Jorgensen KM, Allard RW (1984). Ribosomal DNA spacer- length polymorphisms in barley; Mendelian inheritance, chromosomal location, and population dynamics. Proceed. Natl. Acad. Sci. 81: 8014-8018.

SGN (2009) .Solanaceae Genomic Network. <http://sgn.cornell.edu>.

Śmiech M, Rusinowski Z, Malepszy S, Niemirowicz-Szczytt K (2000). New RAPD markers of tomato spotted wilt virus (TSWV) resistance in *Lycopersicon esculentum.* Acta Physiologiae Plantarum 22(3): 299-303.

Sneath PHA, Sokal RR (1973). Numerical Taxonomy, WH Freeman, San Francisco.

Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, and Cregan PB (2004). A new integrated genetic linkage map of the soybean. Theor. Appl. Genet. 109: 122-128.

Stevens MR, Lamb EM, Rhoads DD (1995). Mapping the *Sw.5* locus for tomato spotted wilt virus resistance in tomatoes using RAPD and RFLP analyse. Theor. Appl. Genet. 90:451 456.

Suliman-Pollatschek S, Kashkush K, Shats H, Hille J, Lavi U (2002). Generation and mapping of AFLP, SSRs and SNPs in *Lycopersicon esculentum*. Cell Mol. Biol. Lett. 7: 583-579.

Tam SM, Vogelaar CM, Kerkveld M, Pearce SR, and Grandbastien M (2005). Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. Theor. Appl. Genet. 110: 819-831.

Tautz D (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res. 17: 6463-6471.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001). Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.) Frequency, length variation, transposon associations, and genetic marker potential. Genome. Res. 11: 1441-1452.

Thiel TW, Michalek RK,Varshney (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.) Theor. Appl. Genet. 106: 411-422.

Van der Hooven R, Ronning C, Giovannoni J, Martin G, Tanksley S (2002). Deductions about the Number, Organization, and Evolution of Genes in the Tomato Genome Based on Analysis of a Large Expressed Sequence Tag Collection and Selective Genomic Sequencing. Plant Cell. 14: 1441-1456.

Varshney RK , Graner A, Sorrells ME (2005). Genic microsatellite markers: their characteristics, development and application to plant breeding and genetics. Trends Biotech. 23: 48-55.

Vos P, Hogers R, Bleeker M, Reijans M, Lee van de T, Hornes M, Frijters A,Pot J, Peleman J, Kuiper M, Zabeau M (1995). AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23: 4407-4414.

Yonemaru JI, Ando T, Mizubayashi T, Kasuga S, Matsumoto T, Yano M (2009). Development of Genome_ wide Simple Sequence Repeat Markers Using Whole- genome Shotgun Sequences of Sorghum (*Sorghum bicolor* (L.) Moench). DNA Research. 1-7.

Zietkiewicz E, Rafalski A, Labuda D (1994). Polymerase chain reaction amplification. Genomics. 20(22): 176-183.