Indian Journal of Agricultural Sciences **88** (1): 74–8, January 2018/Article

Genetic diversity analysis in tomato (*Solanum lycopersicum*) using microsatellite markers

MANISH KUMAR[1], R K YADAV[2], T K BEHERA[3], AKSHAY TALUKDAR[4] and AJAY ARORA[5]

ICAR–Indian Agricultural Research Institute, New Delhi 110 012

Received: 19 April 2017; Accepted: 24 August 2017

ABSTRACT

Present study was conducted to determine the genetic diversity at molecular level among 21 thermo-tolerant tomato accessions. All the accessions were screened by using 18 simple sequence repeats (SSR) primers to estimate genetic diversity and genetic relationships among these accessions. Among the 18 SSRs, only 5 primers (SSR47, TES628, TGS149, TGS230 and TGS330) were polymorphic which were used for accessing the genetic distance to find out the phylogenetic relationship among the different tomato accessions. Polymorphism information content (PIC) ranged from 0.17 (TGS230) to 0.36 (TGS149) with average PIC as 0.30. Unweighted pair group method with arithmetic mean (UPGMA) clustering grouped the accessions into two main clusters. The information related to genetic diversity presented in this study might be useful to breeders for planning crosses among these accessions for future tomato breeding programmes.

Key words: Cluster analysis, Diversity, Microsatellite, Polymerase chain reaction, Tomato

The cultivated tomato (*Solanum lycopersicum* L.) is one of the most extensively studied plant species grown all over the world (Foolad 2007). The major growing states in India are Bihar, Karnataka, Uttar Pradesh, Odisha, Andhra Pradesh, Maharashtra Madhya Pradesh and Asom. Tomato production is highly influenced by environmental factors such as temperature, light, relative humidity and carbon dioxide level in the atmosphere. In tropical and subtropical regions, heat stress is a major limiting factor for growth, reproduction and yield of the crop (Hall 2011). Lack of tolerance to high temperature in most tomato cultivars is one of the major limitations for growing an economic crop in regions where the temperature during the growing season, even for a short duration, reaches 38ºC or higher (Stevens and Rudich 1978). These problems can be minimized by the improvement of cultural practices and breeding approaches. Since thermo tolerance is a complex trait and affected by environmental factors, selection based on phenotypic traits may result in a high level of misleading effects on breeding programs.

Diversity analysis based on molecular markers is one of the most effective approaches for crop improvement. Molecular marker techniques with high polymorphism can be used to identify similarity and differences between varieties and to identify diverse parents, which can be used in breeding programs (Terzopoulos and Bebeli 2008). SSR or microsatellites are widely used in molecular research in different plant species and have been developed in many crops including tomato (Stich *et al.* 2006, Cadalen *et al.* 2010, Shah *et al.* 2013). Microsatellites are having high reproducibility, co-dominant inheritance, multi-allelic nature, high abundance and wide genome coverage (El-Awady *et al.* 2012). Many studies have indicated the efficacy of microsatellite markers for determination of genetic divergence in the genus *Solanum* (Alvarez *et al.* 2001, He *et al.* 2003, Frary *et al.* 2005, Garcia-Martinez *et al.* 2006 and Hu *et al.* 2012). Keeping above in view, the present study was carried out for the assessment of genetic diversity among different thermo tolerant tomato accessions using SSR markers.

MATERIALS AND METHODS

Twenty one tomato accessions were grown and maintained at the Research Farm of Division of Vegetable Science, Indian Agricultural Research Institute, New Delhi. The plant material for the study consisted of various germplasm lines (Chikko, LP 2, PSH 3, TH 348-T2, Balkan, TH 348-4-R, TH 348-4-2 and TH-348-4-5-1), standard released varieties (Pusa Sadabahar, Pusa Ruby, Pusa 120, Pusa Rohini, Pusa Gaurav and Pusa Sheetal) and wild species [Spr-1 and Spr-2 (*S. peruvianum*), Spm, SPM 1, SPM 2, SPM 3 and SPM 4 (*S. pimpinellifolium*)] of tomato.

[1]Ph D scholar (e mail: imanishkumar91@gmail.com), [2]Principal Scientist (rkyadavneh@gmail.com), [3]Professor and Principal Scientist (tusar@rediffmail.com), Division of Vegetable Science; [4]Principal Scientist (akshayassam@yahoo.co.in), Division of Genetics and Plant Breeding; [5]Principal Scientist (romiarora@yahoo.com), Division of Plant Physiology.

18 SSR markers were used for the genotyping assays (Table 1). These markers were randomly selected from the SOL Genomics Network (http://www.sgn.cornell.edu/) or as the description in Fray *et al.* (2005).

Table 1 List of SSR markers with their sequence and PIC values

Primer	Primer sequence (5′-3′)	No. of alleles	PIC
SSR43	f: CTCCAAATTGGGCAATAACA	1	
	r: TTAGGAAGTTGCATTAGGCCA		
SSR47	f: TCCTCAAGAAATGAAGCTCTGA	2	0.32
	r: CCTTGGAGATAACAACCACAA		
SSR48	f: ATCTCCTTGGCCTCCTGTTT	1	
	r: GTCATGGCCACATGAATACG		
SSR66	f: TGCAACAACTGGATAGGTCG	1	
	r: TGGATGAAACGGATGTTGAA		
SSR96	f: GGGTTATCAATGATGCAATGG	1	
	r: CCTTTATGTCAGCCGGTGTT		
SSR111	f: TTCTTCCCTTCCATCAGTTCT	1	
	r: TTTGCTGCTATACTGCTGACA		
SSR146	f: TATGGCCATGGCTGAACC	1	
	r: CGAACGCCACCACTATACCT		
SSR601	f: TCTGCATCTGGTGAAGCAAG	1	
	r: CTGGATTGCCTGGTTGATTT		
TES111	f: ATCTCCTTGGCCTCCTGTTT	1	
	r: GTCATGGCCACATGAATACG		
TES428	f: GAGGGGGATGAAGTAGAGGC	1	
	r: TCCGACAGTGCAAAGTTCAG		
TES628	f: GTCTCCTCGTTTATCCACGCT	2	0.34
	r: TTCTCCACTTATGTGATTATACT-GGG		
TES1873	f: GTGTTCAAATTTGGTTTGGGC	1	
	r: AAAACCGCCAGGATATAGGC		
TGS149	f:GTTTCGTCAGTTGTTAAAAGTT-GAAA	2	0.36
	r: TGAAGCATTGGCTCAAAGAA		
TGS230	f: GAAGCTGATGAACCCAGCAAT	2	0.17
	r: CACATGTTTTGCATTTTTGTTG		
TGS330	f: GCAATGCGATAGTCTTCATGTCA	2	0.34
	r: AAGTTTGTATTCGATTCACCCA		
TGS919	f: TACTTATGTTCAAGGGGCCG	1	
	r: GGCAATTAGTGCATTCCGAT		
TGS2108	f: GTGTGTGTGGGCTGCTTACTC	1	
	r: TTGGGCAATGAAGAAGGAAG		
TGS2204	f: GCTTGACTTTTGCAGCCAACA	1	
	r: AAATAGCAAACAACTTACTC-GAAAA		

Genomic DNA of twenty one accessions of tomato was isolated using the CTAB (Cetyl trimethyl ammonium bromide) method (Murray and Thompson 1980) with few modifications. Briefly, 1g of leaves was ground in liquid nitrogen to a fine powder. The powder was added to 3 ml of extraction buffer (100 mM Tris-HCl pH 8.0, 20 mM EDTA, 1.4 M NaCl, 2% (w/v) CTAB, 2–mercaptoethanol 2%) and incubated at 65°C for 30 min. The DNA was extracted with chloroform-isoamyl alcohol (24:1). The DNA was washed with 70% ethanol and dissolved in 100-400 µl of T E (10 mM Tris–HCl pH 8.0, 1 mM EDTA and 0.2-1 mg/ml RNAase). DNA is quantified by spectrophotometer by using a comparison of the optical density values of the solution at A260/A280 wavelengths. Stock DNA samples were stored at -20°C and diluted to 20 ng/µl when used.

PCR amplification was performed in 20 µl of reaction mixture, containing 11 µl distilled water, 2.5 µl of 10× assay buffer with 15 mM $MgCl_2$, 2 µl of 40 ng template DNA, 1 µl of each forward and reverse primer (1 µm), 2 µl (200 mM of each dNTP) and 0.5 µl (1 U) *Taq* DNA Polymerase. Reactions were conducted in a thermal cycler system (Eppendorf, Germany). The PCR profile starts with initial denaturation at 95°C for 4 min, followed by 35 cycles of denaturation at 95°C for 1 min, annealing as per the primers for 1 min, extension at 72°C for 2 min followed by a final extension at 72°C for 5 min. The products were size-separated on 3% agarose gel and visualized on gel documentation system.

After separation of amplified products of each reaction on 3% agarose gel, it was photographed using a Gel Documentation System, under equal magnification. Scoring was done manually for each of the gel sections. The size of a allele was determined based on the position of the bands relative to the ladder. Band patterns for each of the microsatellite markers were recorded for each genotype by assigning size of band in term of base pairs based on the ladder of minimum 50 bp size. Finally allelic differences in accessions were indicated by scoring (1) for presence or (0) for absence of band. Any band thought to be an artifact, or diffused was considered as missing data. These missing data were designated as '9' (in comparison with '1' for presence and '0' for absence of a band) in the data matrix, the missing data was not considered while analyzing the genetic similarities. 'Null' allele for any specific markers in a genotype was again considered as absence of band (designated as '0') clearly indicating the absence of SSR primer binding site, after re-runs with specific check. The re-runs facilitated the confirmation of allele scoring in various accessions. Monomorphic markers were not excluded from data analysis, since they all have equal effects during dendrogram construction and cluster analysis. Any marker with more than or equal to 30% missing data across various accessions was excluded from analysis.

For each polymorphic SSR marker, 'polymorphism information content' (PIC) was estimated as described

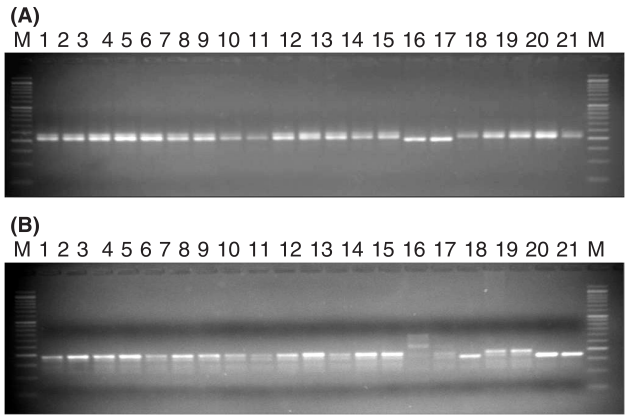

Fig 1 Amplification profiling of 21 tomato accessions using SSR primers. A and B show the allelic segregation of the SSR markers TGS230 and SSR47 respectively. Lane M is standard marker (50 bp). Lanes 1 to 21 represent tomato accessions (1=Pusa Sadabahar, 2=Pusa Ruby, 3=Pusa 120, 4=Pusa Rohini, 5=Pusa Gaurav, 6=Pusa Sheetal, 7=Chikko, 8=LP 2, 9=PSH 3, 10=TH 348-T2, 11=Balkan, 12=TH 348-4-R, 13=TH 348-4-2, 14=TH 348-4-5-1, 15=Spr 1, 16=Spr 2, 17= Spm, 18=SPM 1, 19=SPM 2, 20=SPM 3 and 21=SPM 4).

by Senior *et al*. (1998). The PIC values ranging from '0' (monomorphic) to '1' (highly discriminative, with many alleles in equal frequencies) provide an estimate of the discriminatory power of a marker by taking into account not only the number of alleles at locus, but also the relative frequencies of those alleles in the accessions under study. Mean PIC value was calculated based on binary matrix obtained across various SSRs loci, excluding a few primers that revealed a higher frequency of nulls, as latter could inflate the PIC value.

Similarity matrix using the similarity coefficient of Jaccard (1908) was constructed from the whole SSR data. There are several agglomerative hierarchical methods which are commonly used in clustering techniques, among them the UPGMA (Sneath and Sokal 1973) is the most common method used in this study. Clusters were analyzed using the computer program NTSYS-PC, version 2.11s (Rohlf 2000).

RESULTS AND DISCUSSION

Different types of genetic markers have been used to assess genetic diversity in tomato including Isozymes (Terzopoulos and Bebeli 2008), RFLPs, AFLP and SSRs (Powell *et al.* 1996 and Tam *et al.* 2005). Simple sequence repeats being the species/crop specific markers, are most reliable and reproducible and their application in varietal identification in tomato (Smulders *et al.* 1997, Hokanson *et al.* 1998, He *et al*. 2003, Rajput *et al.* 2006 and Pritesh *et al.* 2010) have been well explored. Further being co-dominant able to distinguish between homozygote as well as heterozygote. All the 18 SSR primers were able to amplify the DNA of 21 diverse thermotolerant tomato accessions, but 13 primers were found to be monomorphic. Intraspecific diversity in *S. lycopersicum* has been reported to be very low, even the genome sequencing in tomato has been done. The low frequency of polymorphism among cultivars and intraspecific crosses is probably due to its autogamous nature (Smulders *et al.* 1997, Nunome *et al.* 2003, Stagel *et al.* 2008). In addition, cultivated tomato is known to be highly monomorphic at the molecular level (Labate and Roberts

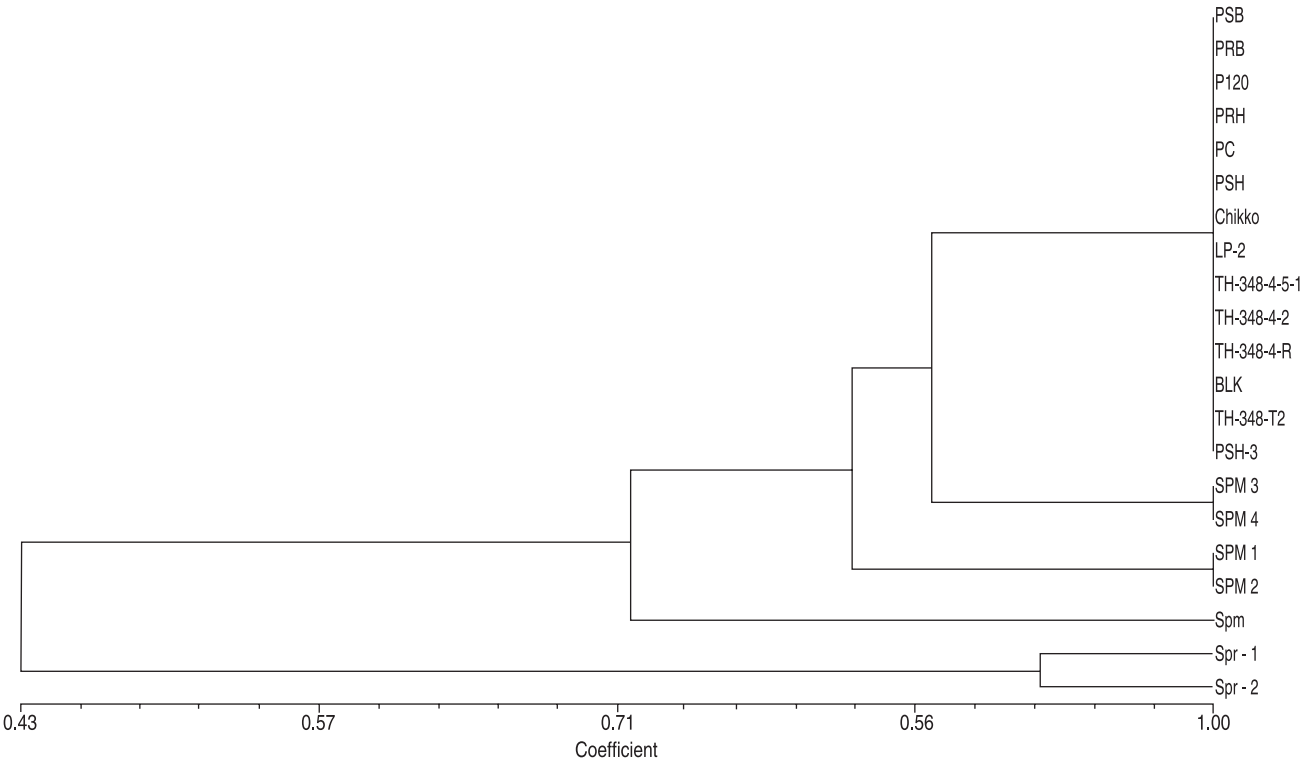

Fig 2 Dendrogram showing clustering of 21 tomato accessions based on Jaccard's similarity coefficient

2002). Out of 18 SSR markers, five SSRs (SSR47, TES628, TGS149, TGS230 and TGS330) showed polymorphism and were used to assess the genetic diversity. The SSR primers could amplify 1 to 2 loci/primer pair. They could amplify a total of 10 alleles giving on an average 2 alleles/primer pair (Fig 1). PIC value ranged from 0.17 (TGS230) to 0.36 (TGS149). The average PIC was 0.30, which confirms that these SSR markers are highly informative. These results were in close conformity with the findings of He *et al.* (2003) and Benor *et al.* (2008). Use of more number of SSR markers will be useful in further analysis.

The Jaccard's similarity coefficient values ranged from 0.43 to 0.71 indicating a significant genetic diversity. The dendrogram based on these values was constructed using UPGMA based clustering method (Fig 2). At coefficient value of 0.43, the accessions were grouped in to two main clusters. The cluster I comprised of two accessions of *S. peruvianum* (Spr-1 and Spr-2). At value of 0.90, the cluster II could be further grouped into two sub clusters IIA and IIB, whereas IIA carrying the accessions of *S. pimpinellifolium* and all the tomato varieties namely, Pusa Sadabahar, Pusa Ruby, Pusa 120, Pusa Rohini, Pusa Gaurav, Pusa Sheetal, Chikko, LP 2, Balkan, PSH 3, TH 348-T2, TH 348-4-5-1, TH 348-4-2,TH 348-4-R. This is expected as the accessions including three species were diverse at phylogenetic level also. Although the numbers of SSR primers were less, however, these were able to differentiate the accessions at species level. The SSR markers were not able to discriminate among the tomato varieties. One of the reasons may be for absence of any introgression in these varieties for heat tolerance traits from two wild species, i.e. *S. peruvianum* and *S. pimpinellifolium*.

In the present study, among the 18 SSR markers used for estimating the genetic diversity of 21 diverse thermo tolerant tomato accessions, five markers (TGS230, TGS330, TGS149, SSR47 and TES628) showed polymorphism and were most significant as they were able to recognize all 21 accessions. The present investigation suggests that primers used in this experiment were appropriate for the analysis of genetic diversity in closely related accessions but for better germplasm characterization and development of potential markers for agronomically important traits more SSR markers need to be screened. Genotyping coupled with phenotyping can present a starting point for MAS (Marker Assisted Selection) in tomato breeding.

REFERENCES

Alvarez A E, Van de Wiel C C M, Smulders M J M and Vosman B. 2001. Use of microsatellites to evaluate genetic diversity and species relationships in the genus *Lycopersicon*. *Theoretical and Applied Genetics* **103(8)**: 1283–92.

Benor S, Zhang M, Wang Z and Zhang H. 2008. Assessment of genetic variation in tomato (*Solanum lycopersicum* L.) inbred lines using SSR molecular markers. *Journal of Genetics and Genomics* **35(6)**: 373–9.

Cadalen T, Morchen M, Blassiau C, Clabaut A, Scheer I, Hilbert J L, Hendriks T and Quillet M C. 2010. Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.). *Molecular Breeding* **25(4)**: 699–22.

El-Awady M A, El-Tarras A A and Hassan M M. 2012. Genetic diversity and DNA fingerprint study in tomato (*Solanum lycopersicum* L) cultivars grown in Egypt using simple sequence repeats (SSR) markers. *African Journal of Biotechnology* **11(96)**: 16233–40.

Foolad M R. 2007. Genome mapping and molecular breeding of tomato. *International Journal of Plant Genomics* **1**: 1–52.

Frary A, Xu Y, Liu J, Mitchell S, Tedeschi E and Tanksley S. 2005. Development of a set of PCR-based anchor markers encompassing the tomato genome and evaluation of their usefulness for genetics and breeding experiments. *Theoretical and Applied Genetics* **111(2)**: 291–312.

Garcia-Martinez S, Andreani L, Garcia-Gusano M, Geuna F and Ruiz J J. 2006. Evaluation of amplified fragment length polymorphism and simple sequence repeats for tomato germplasm fingerprinting: utility for grouping closely related traditional cultivars. *Genome* **49(6)**: 648–56.

Hall A E. 2011. Crop Responses to Environment, p. 227. CRC Press, Boca Raton, Florida.

He C, Poysa V and Yu K. 2003. Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. *Theoretical and Applied Genetics* **106(2)**: 363–73.

Hokanson S C, Szewc-McFadden A K, Lamboy W F and McFerson J R. 1998. Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a Malus× domestica Borkh. core subset collection. *Theoretical and Applied Genetics* **97(5)**: 671–83.

Hu X R, Wang H, Chen J and Yang W C. 2012. Genetic diversity of Argentina tomato varieties revealed by morphological traits, simple sequence repeat, and single nucleotide polymorphism markers. *Pakistan Journal of Botany* **44(2)**: 485–92.

Jaccard P. 1908. Nouvelles researches surla distribution florale. *Bulletin Society Vaudoise Sciences Naturerelles* **44**: 223–70.

Labate J A and Robertson L D. 2002. Genetic variation in heirloom versus modern tomato (*Lycopersicon esculentum*) cultivars. *Program for the 43rd Annual Meeting of the Society for Economic Botany,* New York, p 27.

Murray M G and Thompson W F. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* **8(19)**: 4321–6.

Nunome T, Suwabe K, Iketani H, Hirai M and Wricke G. 2003. Identification and characterization of microsatellites in eggplant. *Plant Breeding* **122(3)**: 256–62.

Parmar P, Oza V P, Chauhan V, Patel A, Kathiria K, Subramanian R, Belur P D, Mugeraya G, Kuppalu N R, El-Banna S G and Attia A M. 2010. Genetic diversity and DNA fingerprint study of tomato discerned by SSR markers. *International Journal of Biotechnology and Biochemistry* **6(5)**: 657–66.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalski A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* **2(3)**: 225–38.

Rajput S G, Wable K J, Sharma K M, Kubde P D, Mulay S A. 2006. Reproducibility testing of RAPD and SSR markers in Tomato. *African Journal of Biotechnology* **5(2)**: 108–12.

Rohlf F J. 2000. NTSYS-pc: Numerical Taxonomy System. Ver.2.1. Exeter Publishing, Ltd.: Setauket, NY.

Senior M L and Heun M. 1998. Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT Primer. *Genome* **36**: 884–9.

Shah S M, Naveed S A and Arif M. 2013. Genetic diversity in basmati and non-basmati rice varieties based on microsatellite markers. *Pakistan Journal of Botany* **45**: 423–31.

Smulders M J M, Bredemeijer G, Rus-Kortekaas W, Arens P and Vosman B. 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theoretical and Applied Genetics* **94(2)**: 264–72.

Sneath P H A and Sokal R R. 1973. *Numerical Taxonomy-The Principles and Practice of Numerical Classification*, pp 230-4. Freeman W H Company, San Francisco, California, USA.

Stagel A, Portis E, Toppino L and Rotino G L. 2008. Gene-based microsatellite development for mapping and phylogeny studies in eggplant. *BMC Genomics* **9**: 1–14.

Stevens M A and Rudich J. 1978. Genetic potential for overcoming physiological limitations on adaptability, yield, and quality in tomato. *Horticulture Science* **13**: 673–78.

Stich B, Maurer H P, Melchinger A E, Frisch M, Heckenberger M, van der Voort J R, Peleman J, Sorensen A P and Reif J C. 2006. Comparison of linkage disequilibrium in elite European maize inbred lines using AFLP and SSR markers. *Molecular Breeding* **17(3)**: 217–26.

Tam S M, Mhiri C, Vogelaar A, Kerkveld M, Pearce S R and Grandbastien M A. 2005. Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. *Theoretical and Applied Genetics* **110(5)**: 819–31.

Terzopoulos P J and Bebeli P J. 2008. DNA and morphological diversity of selected Greek tomato (*Solanum lycopersicum* L.) landraces. *Scientia Horticulturae* **116(4)**: 354–61.