Editura AcademicPres

Available online: www.notulaebotanicae.ro

Print ISSN 0255-965X; Electronic 1842-4309

Not Bot Horti Agrobo, 2017, 45(1):276-279. DOI:10.15835/nbha45110601

Notulae Botanicae Horti Agrobotanici Cluj-Napoca

Original Article

Evaluation of Genetic Diversity by DNA Barcoding of Local Tomato Populations from North-Western Romania

Marin CĂPRAR[1], Cristina M. COPACI[1], Diana M. CHENDE[1], Oana SICORA[1], Radu ȘUMĂLAN[2], Cosmin SICORA[1]*

[1]*Biological Research Center, Botanical Garden "Vasile Fati" Jibou, 16 Wesselenyi Miklos Str., 455200 Jibou, Romania;* caprarmarin@yahoo.com; cri441@netscape.net; chende.diana@yahoo.com; oanasavuro@yahoo.com; cosmin.sicora@gmail.com *(*corresponding author)*

[2]*Banat's University of Agricultural Sciences and Veterinary Medicine "King Michael I[st] of Romania" from Timisoara, Faculty of Horticulture and Forestry, 119 Aradului Str., 300645 Timisoara, Romania;* sumalanagro@yahoo.com

Abstract

Tomato is one of the most important crops worldwide. DNA barcoding is a molecular based method that has been successfully used for species identification, but a few studies have used this method for cultivated varieties identification. The aim of this study was to test the utility of DNA barcoding for the identification of five local salt tolerant tomato varieties and two commercial varieties. To assess the genetic diversity of tomato varieties, the non-coding plastid *trnH-psbA* intergenic spacer and three plastid regions (*rbcL*, *rpoC1, rpoB*) were used. Based on the sequence variation of the *trnH-psbA* barcode, three haplotypes were detected among the seven tomato varieties. A neighbor-joining tree was generated and separated the local tomato varieties from the commercial varieties into two distinct clusters. We found very low levels of variation in the chosen plastial markers, but additional markers could be tested in order to assess the utility of DNA barcodes in tomato varieties identification.

Keywords: molecular identification, plastidial markers, tomatoes varieties, *trnH-psbA*

Introduction

Tomato (*Solanum lycopersicum* L.) is an important vegetable from the Solanaceae family, cultivated worldwide due to its good flavor and rich source of nutrients (Sun *et al*., 2014). It is also a well-known model species for study fruit development and metabolite accumulation. To obtain tomato crops with desired agronomical traits requires a good understanding and management of tomato genetic resources diversity (Bauchet and Causse, 2012). Tomato landraces are highly heterogeneous as they were systemically selected for their performance in adverse agricultural environments (Ciulca *et al*., 2015)

For evaluating genetic variation and phylogenetic relationships among tomato varieties, different molecular methods have been used: RAPD (Carelli *et al*., 2006), RFLP (Asamizu and Ezura, 2009), AFLP and SSR (García-Martínez *et al*., 2006; Benor *et al*., 2008). In 2014, Sun *et al.* used the 5S rRNA region to discriminate tomato varieties, and sequence analysis of this region suggested that a large number of variable nucleotide sites exists among tomato varieties. SNP methodology reveals patterns of genetic variation between cultivated landraces and varieties of tomato (Sim *et al*., 2012; Corrado *et al*., 2014).

DNA barcoding is a method for taxonomic identification which uses a standard short genomic region that has sufficient sequence variation to distinguish among species. A DNA sequence from such a standardized gene region can be obtained from a small amount of tissue taken from an unidentified organism and then compared to a library of reference sequences from known species. If the sequence from the unknown organism match to one of reference sequences means that the organism is recognize, thus providing a rapid identification. An ideal DNA barcode should be present in all groups of land plants, it should be short (700-800 bp) and show enough sequence variation to discriminate among species, also it should be easy to amplify and sequenced with a single primer pair (Kress and Erickson, 2007). Different regions from the plastid genome, including *trnH-psbA* intergenic spacer, *rbcL, rpoC1* and *rpoB*, have been proposed and tested for DNA barcoding of land plants with different level of species identification success depending of the studied group taxa (Kress and Erickson, 2007; Singh *et al*., 2012). The purpose of this study was to test the utility of DNA barcoding for the identification of closely related tomato varieties. In a conservation project, tomato seeds were collected from local farmers of the Bihor County (North-Western Romania). The seeds were chosen

Received: 25 Oct 2016. Received in revised form: 24 Apr 2017. Accepted: 05 May 2017. Published online: 16 May 2017.

from heirloom tomato (varieties that has been passed through several generations of a family) due to high productivity, and moreover these tomato varieties are tolerant to salinity. In this study, we used the non-coding plastid *trnH-psbA* intergenic spacer region, and three plastid coding regions *rbcL, rpoC1,* and *rpoB*.

Materials and Methods

Plant materials

In this study, plant samples were collected from tomatoes grown in the "Vasile Fati" Botanical Garden, Jibou. The plant seeds were obtained in 2012 from gardens of local farmers from three villages, all located in the Bihor County (Table 1). Also, in this study, were included two varieties of commercial tomatoes, *Solanum lycopersicum* 'Marmande' and *Solanum lycopersicum* 'Kecskemeti Jubileum'. Four tomato varieties, cherry tomatoes and the commercial varieties were grown in pots and the remaining three varieties were cultivated in the field. Each tomato variety was represented by a single individual.

DNA extraction, PCR amplification and sequencing

Total genomic DNA was isolated from 80 mg of fresh young leaves from each individual using ISOLATE Plant DNA Mini Kit (Bioline USA Inc.) following a modified protocol as described in Căprar *et al*. (2014). The concentration and purity of each DNA sample was measured with Nanodrop 2000 UV-VIS Spectrophotometer (Thermo Fisher Scientific Inc., United States).

The four plant DNA barcodes, *rbcL*, *trnH-psbA* ,*rpoC1* and *rpoB* were amplified in a 25 µL reaction volume, using My Taq™ DNA Polymerase (Bioline Reagents Ltd, UK), 0.5 µL primers and 200 ng DNA template. PCR amplification was performed on a Bioer XP Thermal Cycler (Bioer Technology Co., Ltd.). Primers for PCR and sequencing (Kress and Erickson, 2007), and PCR cycling conditions used in this study are provided in Table 2.

To verify the success of PCR amplification, 5 µL of the PCR product were subjected to 2% agarose gel electrophoresis in TAE buffer and visualized under an UV trans-illuminator with G: BOX ChemiXR5 (Syngene, UK). The remaining PCR product was purified using the Favor Prep™ Gel/PCR Purification Kit (Favorgen Biotech Corp.). Purified PCR products were send to Macrogen Europe (Amsterdam, Netherlands) and sequenced in both directions with the same primers used for PCR.

Data analysis

Sequences for each region were assembled and edited using BioEdit v7.2.5 (Hall, 1999). Then, the edited sequences were aligned by Clustal W in MEGA 6 (Tamura *et al*., 2013). The genetic pair wise distance for *trnH-psbA* marker was calculated using MEGA 6 with the Kimura 2-parameter (K2-P) model. A neighbor-joining (NJ) tree was constructed based on the multiple sequence alignment of the *trnH-psbA* intergenic spacer in MEGA 6 with p-distance model. Bootstrap values were calculated over 1000 replications (Felsenstein, 1985). The barcode sequences were queried against Gen Bank database (NCBI) using Nucleotide BLAST algorithm.

Table 1. List of tomato varieties, location, fruit shape and color

No.	Sample	Species	Location and house number	Fruit shape and color
1	A	*Solanum lycopersicum* L.	Ateaş, 136	Oxheart, pink
2	AT	*Solanum lycopersicum* L.	Ateaş, 37	Round, red
3	C	*Solanum lycopersicum* L.	Cefa, 7	Flattened globe, red
4	CG	*Solanum lycopersicum* L.var.*cerasiforme*	Cefa, 7	Round, yellow
5	CR	*Solanum lycopersicum* L. var. *cerasiforme*	Mărţihaz, 7	Round, red
6	J	*Solanum lycopersicum* L. 'Kecskemeti Jubileum'	-	Round, red
7	M	*Solanum lycopersicum* L. 'Marmande'	-	Flattened globe ribbed, red

Table 2. Primers, their sequences and PCR conditions

Regions	Primer pairs	Sequence 5′→3′	PCR conditions
rbcL-a	a_f	ATGTCACCACAAACAGAGACTAAAGC	95 °C 3 min; [35 cycles: 95 °C 30 s; 50 °C 30 s; 72 °C 90 s]; 72 °C 2 min
	a_r	CTTCTGCTACAAATAAGAATCGATCTC	
trnH-psbA	f	ACTGCCTTGATCCACTTGGC	
	f	CGAAGCTCCATCTACAAATGG	
rpoC1	1f	GTGGATACACTTCTTGATAATGG	
	3r	TGAGAAAACATAAGTAAACGGGC	
rpoB	2f	ATGCAACGTCAAGCAGTTCC	
	4r	GATCCCAGCATCACAATTCC	

Results and Discussion

Sequence characteristics of the barcodes

The four barcodes, *rbcL*, *trnH-psbA*, *rpoC1* and *rpoB* showed high success rates for PCR amplification and sequencing using a single primer pair. The sequences characteristics of the four regions are presented in Table 3. Of the four barcodes, the *trnH-psbA* sequences had three variable sites among the seven tomato varieties, found in the commercial varieties ('K Jubileum and Marmande), and the *rbcL*, *rpoC1* and *rpoB* sequences did not show any variable sites, thus these sequences were 100% conserved within the species. The genetic distances for the *trnH-psbA* sequence ranged from 0 to 0.004.

BLAST Search

Each barcode sequence was compared against the NCBI database through a BLAST search. All sequences of the *rbcL* and *rpoB* loci identified the seven tomato varieties as *Solanum pimpinellifolium* with 99 or 100% identity. Sequences of *rpoC1* identified the seven tomato varieties as *Solanum tuberosum* with 100% identity. The lack of sequence variation did not allow to separate the samples into different tomato varieties, and after the BLAST search these loci were identified at genus level (*Solanum*). Only *trnH-psbA* sequences were correctly identified at species level, as *Solanum lycopersicum* with 99% identity.

In a study from The Tomato Genome Consortium (2012), the genome of cultivated tomato was compared with its closest wild relative, *Solanum pimpinellifolium,* and to the potato genome (*Solanum tuberosum*). The results revealed that the two tomato genomes have only 0.6% nucleotide divergence and evidence of recent admixture, but more than 8% divergence from potato.

Phylogenetic analysis

A neighbor-joining tree was constructed based on the sequence variation of the *trnH-psbA* region, and the cultivars were grouped into two distinct clusters (Fig. 1). The first cluster grouped all the tomato local populations; while the second cluster grouped the two commercial varieties. The tree topology is supported by a good bootstrap value. No differences between the five local tomato populations were found within the *trnH-psbA* barcode region. Although, the five local varieties have morphological different fruits, shared the same haplotype for *trnH-psbA* marker, which is considered one of the most variable non-coding regions of the plastid genome (Chase *et al*., 2007).

Studies of genetic diversity based on molecular markers in the section *Lycopersicon* revealed that wild species have a high level of genetic diversity compared to cultivated tomato (Stevens and Robbins, 2007). Domestication of tomatoes by selecting preferred traits has led to low genetic diversity among cultivated tomatoes. A high similarity coefficient was found among 29 cultivated tomatoes using SSR markers, as published by Zhou *et al*., 2015. In 2011, Sun *et al*. used three DNA markers to distinguish 26 tomato varieties, and found that nrDNA ITS region and rDNA 5S showed high nucleotide variation, whereas cpDNArbcL region was not suitable for tomato variety identification. Enan and Ahmed (2014) evaluated the potential of two DNA barcode markers, *matK* and *rpoC1*, for the authentication of 11 date cultivars, and *rpoC1* was less informative than *matK*. A study of Jarret (2008) showed that *trnH-psbA* could not discriminate among the members of the *Capsicum annum* complex, but this complex was separated from another *Capsicum* species. In a study that assessed the genetic diversity of seven taro cultivars (*Colocasia esculenta*), the *trnH-psbA* marker showed genetic variability among them, and grouped the cultivars according to their geographical origin, Midwest and Southeast of Brazil (Nunes *et al*., 2014).

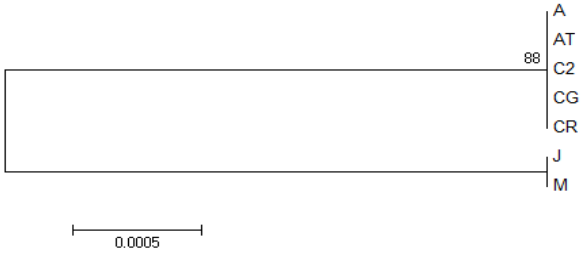

Fig. 1. The NJ tree based on *trnH-psbA* intergenic spacer sequences

Table 3. The characteristics of each single barcode

Marker	PCR success (%)	Sequencing success (%)	Aligned length (bp)	Variable sites (%)	Intraspecific distance (mean)
rbcL	100%	100%	631	0	0
trnH-psbA	100%	100%	475	0.63	0-0.004 (0.002)
rpoC1	100%	100%	510	0	0
rpoB	100%	100%	518	0	0

Table 4. BLAST search results

Region	Species identified	Ident.	E value	Accession
rbcL	*Solanum pimpinellifolium*	99%	0.0	KP117027.1
trnH-psbA	*Solanum lycopersicum*	99%	0.0	KP117024.1
rpoC1	*Solanum tuberosum*	100%	0.0	KM489056.2
rpoB	*Solanum pimpinellifolium*	100%	0.0	KP117027.1

Conclusions

In this study, four DNA barcodes were chosen to evaluate the genetic diversity of seven tomato varieties. Among these four barcodes, only *trnH-psbA* showed sequence variation, separating the commercial varieties from the local tomato populations. Sequence variability in *trnH-psbA* allowed three haplotypes to be distinguished among the seven tomato varieties. Our findings revealed that the chosen plastidial markers are not suitable to distinguish between the tomato varieties.

Acknowledgements

This study was supported by the Romanian National Authority for Scientific Research, CNDI-UEFISCDI (PN-II-PT-PCCA-2011-3.1-0965). We thank to all local farmers who were kind enough to let us visit their gardens and offered us the necessary seeds.

References

Asamizu E, Ezura H (2009). Inclusion of tomato in the genus Solanum as "*Solanum lycopersicum*" is evident from phylogenetic studies. Journal of the Japanese Society for Horticultural Science 78:3-5.

Bauchet G, Causse M (2012). Genetic Diversity in Tomato (*Solanum lycopersicum*) and its wild relatives. In: Çalişkan M (Ed.). Genetic diversity in plants. INTECH Riejka, Croatia pp 133-162.

Benor S, Zhang M, Wang Z, Zhang H (2008). Assessment of genetic variation in tomato (*Solanum lycopersicum* L.) inbred lines using SSR molecular markers. Journal of Genetics and Genomics 35:373-379.

Carelli BP, Gerald LTS, Grazziotin FG, Echeverrigaray S (2006). Genetic diversity among Brazilian cultivars and landraces of tomato *Lycopersicon esculentum* Mill. revealed by RAPD markers. Genetic Resources and Crop Evolution 53:395-400.

Căprar M, Cantor M, Sicora O, Copaci C, Sicora C (2014). Optimization of DNA isolation from four species of *Rhododendron* from Europe. Journal of Horticulture, Forestry and Biotechnology 18:117-122.

Sumalan R, Ciulca A, Botnarescu F (2015). Study of yield components for some Romanian tomato landraces under greenhouse conditions. Journal of Horticulture, Forestry and Biotechnology 19(3):36-41.

Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madriñán S, Petersen G, … Wilkinson M (2007). A proposal for a standardized protocol to barcode all land plants. Taxon 56:295-299.

Corrado G, Piffanelli P, Caramante M, Coppola M, Rao R (2013). SNP genotyping reveals genetic diversity between cultivated landraces and contemporary varieties of tomato. BMC Genomics 14:835-849.

Enan MR, Ahamed A (2014). DNA barcoding based on plastid matK and RNA polymerase for assessing the genetic identity of date (*Phoenix dactylifera* L.) cultivars. Genetics and Molecular Research 13:3527-3536.

Felsenstein J (1985). Phylogenies and the Comparative Method. The American Naturalist 125:1-15.

García-Martínez S, Andreani L, Garcia-Gusano M, Geuna F, Ruiz JJ (2006). Evaluation of amplified fragment length polymorphism and simple sequence repeats for tomato germplasm fingerprinting: utility for grouping closely related traditional cultivars. Genome 49:648-656.

Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program from Windows 95/98/NT. Nucleic Acids Symposium, Series No 41:95-98.

Jarret RL (2008). DNA Barcoding in a Crop Genebank: the *Capsicum annuum* species complex. The Open Biology Journal 1:35-42.

Kress WJ, Erickson DL (2007). A Two-Locus Global DNA Barcode for land plants: The Coding *rbcL* Gene Complements the Non-Coding *trnH-psbA* Spacer Region. Plos One 2, e508.

Nunes RSC, Del Aguila EM, Paschoalin VMF, da Silva JT (2014). DNA barcoding assessment of the genetic diversity of varieties of taro, *Colocasia esculenta* (L.) Schott in Brazil. In: Breeding and genetic engineering: the biology and biotechnology research. Ed 1st, iConcept Press Ltd.

Zhou R, Wu Z, Cao X, Jiang FL (2015). Genetic diversity of cultivated and wild tomatoes revealed by morphological traits and SSR markers. Genetic and Molecular Research 14 (4):13868-13879.

Sim SC, Van Deynze A, Stoffel K, Douches DS, Zarka D, Ganal MW, … Francis DM (2012). High-Density SNP genotyping of tomato (*Solanum lycopersicum* L.) reveals patterns of genetic variation due to breeding. Plos One 7, e45520.

Singh HK, Parveen I, Raghuvanshi S, Babbar SB (2012). The loci recommended as universal barcodes for plants on the basis of floristic studies may not work with congeneric species as exemplified by DNA barcoding of *Dendrobium* species. BMC Research Notes 5:42.

Stevens MR, Robbins MD (2007). Molecular markers in selection of tomato germplasm. In: Razdan MK, Mattoo AK (Eds). Genetic improvement of solanaceous crops, Vol 2, Tomato. Science Publishers, Enfield pp 239-260.

Sun YL, Kang HM, Kim YS, Baek JP, Zheng SL, Xiang JJ, Hong SK (2014). Tomato (*Solanum lycopersicum*) variety discrimination and hybridization analysis based on the 5S rRNA region. Biotechnology & Biotechnological Equipment 28:431-437.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30:2725-2729.

The Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485: 635-641.

Yan-Lin S, Jun PB, Mun-Haeng L, Hee-Kyoung L, Young-Sik K, Soon-Kwan H, … Ho-Min K (2013). Phylogenetic relationships of varieties of tomato (*Solanum lycopersicum*) using DNA Markers. Journal of Pure and Applied Microbiology 7:687-693.