Application of EST-SSRs to examine genetic diversity in eggplant and its close relatives

Yeliz Tumbilen, Anne Frary, M-Christin Daunay, Sami Doganlar

Turk J Biol
35 (2011) 125-136

doi:10.3906/biy-0906-57

Application of EST-SSRs to examine genetic diversity in eggplant and its close relatives

Yeliz TÜMBİLEN[1], Anne FRARY[1], Marie Christine DAUNAY[2], Sami DOĞANLAR[1]

[1]İzmir Institute of Technology, Department of Molecular Biology and Genetics, 35430 Urla, İzmir - TURKEY

[2]INRA, UR1052, Unité de Génétique et d'Amélioration des Fruits et Légumes F-84140, Montfavet - FRANCE

Received: 23.06.2009

Abstract: Within the genus *Solanum*, the term 'eggplant' encompasses several cultivated species that are used for food and, to a lesser extent, for medicine. The use of one common name to describe more than one species and the existence of many related wild species have led to taxonomic confusion which, in turn, have complicated analyses of evolutionary relationships and genetic diversity within this groups of species. A further challenge for eggplant research is that, despite the fact that the use of molecular markers for phylogenetic studies is well-established, very few studies have described the development of new markers for eggplant. In our work, genic microsatellite (SSR) markers were identified from an expressed sequence tag library of *S. melongena* and used for analysis of 47 accessions of eggplant and closely related species. The markers had very good polymorphism in the 18 species tested including 8 *S. melongena* accessions. Moreover, genetic analysis performed with these markers showed concordance with previous research and knowledge of eggplant domestication. These markers are expected to be a valuable resource for studies of genetic relationships, fingerprinting, and gene mapping in eggplant.

Key words: Genic SSRs, microsatellites, *Solanum*, *S.* aethiopicum, *S. macrocarpon*, S. *melongena*

Patlıcan ve yakın akraba türlerinde genetik çeşitliliği incelemek için EST-SSRs işaretleyicilerinin uygulanması

Özet: *Solanum* cinsi içerisinde, 'patlıcan' terimi gıda ve daha az ölçüde tıp için kullanılan çeşitli kültür türlerini kapsamaktadır. Birden fazla türü tanımlamak için yaygın bir ismin kullanımı ve çok sayıda akraba yabani türlerin varlığı taksonomik karışıklığa yol açmıştır. Bu durumun bir sonucu olarak bu gruplar içerisindeki türlerin evrimsel ilişkilerin analizleri ve genetik çeşitliliğin ortaya çıkarılması karmaşık hale gelmiştir. Patlıcan araştırmaları için ortaya çıkan diğer bir sorun da, aslında filogenetik çalışmalar için moleküler işaretleyicilerin kullanımı iyi kurulmasına rağmen, çok az sayıdaki çalışmada patlıcan için yeni işaretleyicilerin geliştirilmesi nitelendirilmiştir. Bu çalışmada, genik mikrosatalit (SSR) işaretleri *Solanum melongena* türü için geliştirilmiş bir ifade edilmiş DNA dizi etiketi kütüphanesinden belirlenmiştir ve 47 adet patlıcan ve yakın akraba türlerine ait tohum örneğinin analizinde kullanılmıştır. Bu işaretleyiciler aralarında 8 adet *Solanum melongena* tohum örneği içeren 18 türde çok iyi polimorfizm vermiştir. Ayrıca, bu işaretleyiciler kullanılarak yapılan genetik analizler önceki araştırmalarla ve patlıcanın ıslahı ile ilgili bilinen bilgiler ile uyum göstermiştir. Bu işaretleyicilerin patlıcan da genetik ilişkilerin belirlenmesi, parmakizi analizleri ve genetik haritalama çalışmaları için çok değerli kaynak oluşturmaları beklenmektedir.

Anahtar sözcükler: Genik SSRs, mikrosatellitler, *Solanum*, *S. aethiopicum, S. macrocarpon, S. melongena*

Introduction

The Solanaceae is an important plant family that is distributed worldwide. The family contains many domesticated species including tomato, pepper, potato, eggplant, tobacco, and petunia. Solanaceous species are used in the human diet, for health purposes, as drugs and as ornamentals. As a result, the family ranks third among other plant families in terms of economic importance (1).

Solanum is the largest genus of the family; almost half of solanaceous plants belong to this genus (2-4) including the important crop plants, tomato, potato and eggplant, and lesser known cultivated species, such as pepino, lulo, tamarillo (tomato tree), and cocona. Among these crops, eggplant presents a scientific challenge for several reasons (5), including the taxonomic confusion in genus *Solanum* (6,7). *Solanum melongena* (brinjal eggplant), *S. aethiopicum* (scarlet eggplant), and *S. macrocarpon* (Gboma eggplant) are all commonly referred to as eggplant (5,7,8). Thus, in various regions of the world, people may refer to different species when they use the word eggplant. Among these species, *S. melongena* is of the most economic consequence, especially in Asian and Mediterranean countries. Recently, however, eggplant has become a globally cultivated crop and more scientific studies are examining its molecular genetics (e.g., 5, 9-13), its molecular intraspecific diversity, and its relationships with other *Solanum* species (e.g., 14-19).

Although solanaceous species are distributed worldwide, generic and species level diversity is concentrated in South America (2-4). An interesting exception consists of *S. melongena* and its relatives, which originate from Africa and Asia, and thus are Old World species (20). In addition to distributional diversity, there is also a great amount of morphological diversity in *Solanum* at both the species and cultivar level (4). This morphological diversity has been very helpful for classification and taxonomy of the genus. However, the exact number of species belonging to the genus *Solanum* is still indefinite. Weese and Bohs (21) estimate that 1250 to 1700 species comprise the genus. Several recent studies have used molecular markers in attempts to better understand the phylogeny of *Solanum* species (e.g., 14,16,21,22).

Among various techniques that can be used for such analyses, the SSR (simple sequence repeat) marker system has several advantages. SSRs or microsatellites are tandemly repeated short nucleotide units of 1 to 5 nucleotides. These repeats can be located in genes (genic SSRs) or non-coding regions (genomic SSRs) of the nuclear genome and also in cytoplasmic genomes (23,24). In the nuclear genome, genomic SSRs are reported to be collected around particular regions of the chromosomes, such as centromeric areas (25). Both types of SSR markers are easy to apply and have a high level of polymorphism, which make them ideal for mapping and diversity studies, fingerprinting, and population genetics (23,24,26,27). Moreover, once SSR primers have been designed, application of these markers is fairly inexpensive. Genomic SSRs for eggplant have been developed by Nunome et al. (15) while Stàgel et al. (19) have developed genic SSRs, which they tested primarily on eggplant cultivars.

The goal of the current study was to identify genic SSRs for eggplant from the publicly available expressed sequence tag (EST) database (http://sgn.cornell.edu) and to evaluate the use of these markers to examine genetic diversity and structure among *S. melongena* and its close domesticated and wild relatives. Thus, marker primers were designed, tested, and applied to 47 accessions representing 18 different species. The results showed that, although genic SSRs derived from ESTs may represent conserved regions of the genome, these markers have good polymorphism and are useful for the analysis of genetic relationships in *S. melongena* and related species.

Materials and methods

Plant material

Eggplant and its relatives were represented by 47 different accessions from the Institut National de la Recherche Agronomique (INRA, UR1052), Montfavet, France. The 47 accessions included 18 different species; and within these 18 species, *S. melongena*, *S. incanum,* and *S. aethiopicum* were represented by several accessions falling into different groups, as defined by Lester (28) and Lester and Hasan (29). The accessions for each species and cultigroup used in this study are listed in Table 1.

Table 1. Accessions of eggplant and its wild relatives used in this work. Genotype number refers to the sample number used in Figure 1.

Species Name	Group	Accession Number	Genotype Number	Number of Accessions
S. aculeastrum		MM 1169	38	1
S. aethiopicum	Kumba (cultigroup)	MM 0574	2	3
	Aculeatum (cultigroup)	MM 0134	12	
	Gilo (cultigroup)	MM 0232bis	16	
S. anguivi		MM 0982	32	2
		MM 1259	19	
S. burchellii		MM 1235	39	1
S. capsicoides		MM 0376	7	1
S. dasyphyllum		MM 1137	37	1
S. incanum	A	MM 0210	15	17
	A	MM 0661	1	
	A	MM 0700	24	
	A	MM 0702	25	
	A	MM 0707	26	
	A	MM 0712	27	
	B	MM 1244	40	
	B	MM 1426	22	
	C	MM 0577	5	
	C/D	MM 0672	42	
	C	MM 0677	46	
	C	MM 0715	29	
	D	RNL 0337	23	
	D	MM 0674	43	
	D	MM 0676	45	
	D	MM 0713	28	
	D	MM 1248	18	
S. lidii		MM 1005	33	1
S. linnaeanum		MM 0195	14	1
S. macrocarpon		MM 0132	11	4
		MM 0150	13	
		MM 1007	34	
		MM 1129	36	
S. marginatum		MM 0824	31	1
S. melanospermum		MM 1350	21	1
S. melongena	E (weedy)	MM 0498	6	8
	E (weedy)	MM 0669	41	
	E (weedy)	MM 0675	44	
	F (wild)	MM 0686	47	
	G (primitive cultivar)	MM 1010	35	
	H (advanced cultivar)	BIRM/S. 2458	9	
	H (advanced cultivar)	LF3.24	10	
	H (advanced cultivar)	MM 0738	30	
S. scabrum		MM 0373	8	1
S. sessilistellatum		MM 1269	20	1
S. sisymbriifolium		MM 0284	17	1
S. viarum		MM 0374	4	1
S. violaceum		MM 0497	3	1

DNA isolation

DNA from 10 individual plants of each of the 47 accessions was extracted from young leaves using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA). After rehydration, 5 µL of DNA for each individual was combined with the DNAs of the other individuals of the same accession. These pooled DNA samples were the material used for further experiments.

EST-SSR marker development

To design the SSR primers, a *S. melongena* EST library with 3181 sequences was accessed from Sol Genomics Network (http://www.sgn.cornell.edu/about/ about_solanaceae.pl). The SSR Discovery Input was used with the default parameters to find SSRs within the EST sequences and to design primers (http://hornbill.cspp.latrobe.edu.au/cgi-binpub/ssrprimer/indexssr.pl). The EST sequences containing SSRs were then analyzed for their uniqueness by seeing if multiple SSRs corresponded to the same unigene in the Sol Genomic Network (SGN). Based on this analysis and selection for microsatellite length in which only SSRs containing dinucleotides greater than 8, trinucleotides greater than 4, tetranucleotides greater than 3, and pentanucleotides greater than 2 units long were selected, 29 unique *S. melongena* SSR (smSSR) markers were selected for analysis on the *Solanum* species (Table 2). Primers were synthesized by Integrated DNA Technologies, Inc., IA, USA, and checked for amplification in PCR reactions with eggplant DNA. PCR reactions (25 µL) contained: 1× PCR buffer, 0.2 mM dNTP, 5 pmol of each F and R primer, 0.25 U Taq Polymerase, dH_2O, and 2 µL (50-100 ng) sample DNA. PCR conditions were a preliminary denaturation for 5 min at 94 °C; 35 cycles at 94 °C for 30 s, 50 °C for 1 min, 72 °C for 1 min; final extension for 5 min at 72 °C, and hold at 4 °C. Samples were electrophoresed for at least 4 h at 120 mA through 3% agarose, 1× TAE gels. All the primers amplified products.

For diversity analysis on the 47 accessions, a CEQ™ 8800 Genetic Analysis System (Beckman Coulter, Inc., Fullerton, CA, USA) was used. For more economical detection, a fluorescently labelled M13 (-21) primer was used and complementary M13 sequence was added to the 5' end of the smSSR forward primers as described by Schuelke (30). Fluorescently labelled primer was synthesized by Sigma-Proligo (Sigma-Aldrich Company, LTD Irvine, Ayrshire, UK). PCR reactions were 20 µL total for each sample and were composed of dH_2O, 1× PCR buffer, 0.2 mM dNTP, 0.2 U Taq Polymerase, 0.6 pmol F primer, 2.4 pmol of R and M13 primers, and 2 µL (~20 ng) sample DNA. The amplification profile was: 94 °C for 5 min; 94 °C for 30 s, 56 °C for 45 s, 72 °C for 45 s for 27 cycles; 94 °C for 30 s, 53 °C for 45 s, 72 °C for 45 s for 8 cycles; 72 °C for 10 min, hold at 4 °C. Before loading the samples for analysis in the sequencer, PCR products were diluted 1:10 with sample loading solution (SLS). Thus, for each sample, 3 µL PCR product was diluted with 27 µL SLS and 0.5 µL size standard–400 (Beckman Coulter, Inc., Fullerton, CA, USA) was added. The separation method was: capillary temperature 50 °C, denaturation temperature 90 °C for 120 s injection voltage 2.0 kV for 30 s with a separation voltage of 4.8 kV for 60 min.

Genetic analysis

For analysis of SSR data, each accession was genotyped for each smSSR based on the presence (1) and absence (0) of peaks (bands). Qualitative data were used to generate a matrix determining similarity among samples using Dice's method (31) and the similarity matrix was then used to draw a dendrogram with the clustering method UPGMA (Unweighted Pair Group Method with Arithmetic Mean) via the SHAN module of NTSYS-pc version 2.2j software, (Applied Biostatistics Inc, Setauket, NY, USA). To evaluate the efficiency of clustering, the cophenetic correlation coefficient was calculated with the Mantel method (32).

Results and discussion

Microsatellite identification and characterization

As a result of the analysis of the *S. melongena* EST library on SGN, 158 different sequences were identified as having at least one SSR. When the SSRs were individually counted, the total number of SSRs was found to be 168 as 9 of the sequences had 2 SSRs and 1 had 3 SSRs. Moreover, the 158 SSR-containing sequences were found to represent 110 unigenes based on analysis in SGN. In total, 7 compound repeats were identified. Compound repeats have 2 or

Table 2. SSR marker repeat motifs and sequences for the smSSRs designed from the eggplant EST database (smSSR = S. melongena SSR).

Marker Name	EST Identifier	Repeat Motif and Number	Forward Sequence	Reverse Sequence	Left TM	Right TM	Product Size (bp)
smSSR01	sgn|E513845	(ATT)21	GTGACTACGGTTTCACTGGT	GATGACGACGACGATAATAGA	55.0	55.3	310
smSSR03	sgn|E514601	(TA)9 (GA)8	ATTGAAAGTTGCTCTGCTTC	GATCGAACCCACATCATC	54.8	54.3	145
smSSR04	sgn|E514602	(TA)9 (GA)8	CTCTGCTTCACCTCTGTGTT	CCATGAAAGAGAAGATCGAG	55.5	55.0	320
smSSR09	sgn|E513913	(TTTGC)3	CACATGGGAACCTACTTACC	GACGACCATCAAACAAGAAT	54.5	55.0	344
smSSR11	sgn|E515884	(AGC)6	AAACAAACTGAAACCCATGT	AAGTTTGCTGTTGCTGCT	54.5	54.6	126
smSSR12	sgn|E516012	(ACCAA)3	AAACAGAAACCAGAGTACTTCA	CAGAAGAAGGTTCAGTTTGC	53.4	55.2	313
smSSR14	sgn|E517698	(ATTA)4	ATACCACATCAATCCAAAGC	CATCATCATCTTCACAGTGG	55.0	54.7	241
smSSR15	sgn|E518171	(CCTTT)3	CTGTGGTTGCCTTATCAGTA	TAGTCCAAGGGTTTGATGAC	53.8	55.0	116
smSSR16	sgn|E518867	(AGA)7	AAGAATTTGATGTTGAACCG	CTTTATCAGCCAATTTCTGG	55.2	55.1	390
smSSR17	sgn|E519219	(ATAC)4	TCTTGCCATTTAATTTCCTC	CTATGTCCCTATTATGCCCA	54.6	55.1	115
smSSR18	sgn|E519312	(TAAT)4	TTAGGCATTTGATTTAGCCT	TATGTCCCTAAGCATAACGG	54.4	55.4	342
smSSR19	sgn|E520513	(GAA)6	GAACAATGATTCATCGGATT	AGTTGATGTTGAATTTCCCA	54.9	55.5	241
smSSR21	sgn|E514329	(TAC)5	AAGTTTACATGACAGCACCA	TTGCCATCATCAATACCATA	54.1	54.8	249
smSSR22	sgn|E514434	(GCC)5	CTCCGTCAAATTCCTATCAA	GGGAGTCCACATAGAGCATA	55.3	55.2	276
smSSR24	sgn|E515827	(TCA)5	GATTTATGGCTTCTGATGGA	TCCTAACCCACTTGATGAAC	55.2	55.0	229
smSSR27	sgn|E516784	(TGT)5	ATACATTTGAGCCGAGAGTG	TAAATCTGAGAAGGTCGCAT	55.4	55.0	184
smSSR28	sgn|E517072	(TCA)5	CACACTCCTCAGAACTCCAT	CAGCAGTACCTCTTGGTCAT	55.1	55.3	301
smSSR29	sgn|E517168	(CTT)5	TCCACTTCAATTTCCAAGTC	GATCGCTTAGCAGAAGCC	55.2	56.2	188
smSSR31	sgn|E517356	(TCC)5	CTTCCTACCCACACTTCATC	TAGGCCGGAGATAGTTGTAA	54.6	55.1	225
smSSR35	sgn|E517795	(ATG)5	CACCACCAAAGAATTCCTAA	TTGCTAGAAATAGCAAAGGG	55.2	55.0	269
smSSR36	sgn|E517835	(CTG)5	AGCACCAGGACAATGAATAC	CCATTTCTTTCTCGACCTTA	55.1	54.6	231
smSSR37	sgn|E517892	(AAG)5	AAAGAAGCTTCCGACGAA	CACTTGTTTCAGCACTTTGA	56.1	55.0	115
smSSR40	sgn|E518161	(AAG)5	TTCTTTGATCTTCAATTCCAA	ATGAAGCTGTTCATGATTCC	55.0	55.1	283
smSSR41	sgn|E518430	(TCA)5	CTCCTCCTGGTAAGGAGTCT	GCAGTATAGAGACGCGAAAT	55.0	54.8	267
smSSR42	sgn|E518630	(CAC)5	ACAGTACACCAGAAACGGAA	GTTACAATGACGGTGGATCT	55.7	54.9	160
smSSR44	sgn|E519591	(CCA)5	TGCATTTCATACAGAAACCA	GCAAGGATATCACTGAGCTG	55.1	56.0	233
smSSR45	sgn|E519680	(TTC)5	TTTCTCAACCCAAACTGAAC	GCAGCTCTCGCATAGATAGT	55.3	55.0	172
smSSR46	sgn|E519853	(CAC)5	GGAAACCTTCATTCACTTCA	AGGTCACCGTTACAATTACG	55.2	55.2	272
smSSR47	sgn|E520160	(AGA)5	ACACGATGATCATAAGGGAG	ATCTAATCACTGTCGCTGCT	55.0	55.1	189

more different nucleotide motifs occurring in a single SSR. Overall, the AT repeat was the most common, representing 8.3% of the total SSRs identified. The AT repeat has also been frequently identified in other genic (19) and genomic (15) SSRs in eggplant. The longest simple SSR was a TAA SSR with 22 repeat units. Based on total length, the longest SSR was the compound repeat $(TAA)_{20}$ $(CGA)_{8,}$ 84 nucleotides long. When the repeat motifs were classified in terms of the number of bases in the repeat, it was observed that the most common ones were trinucleotide repeats, which represented 56.7% of the total. This result has also been observed in studies of genic SSRs in other species (33,34) and is expected because the variation in trinucleotide repeats does not cause frameshift mutations and is, therefore, more likely to survive negative selection (35). TCA and TTC/AAG were the 2 most frequently identified trinucleotide repeats with 8 SSRs identified for each. Stagel et al. (19) also searched the same EST library (http://www.sgn.cornell.edu/) for microsatellites. They found much fewer SSRs (only 70) than in our study (168), but this can be explained by their search criteria, which eliminated redundancies. They found, as we did, that the majority (51.4%) of genic SSRs were trinucleotide repeats with TCA and AAG as the most common. In contrast, ATC and AAC were found to be the most frequent trinucleotide repeats in eggplant genomic DNA (36).

As described in the materials and methods, 29 markers were selected for analysis of the 47 *Solanum* accessions. These smSSRs amplified a total of 307 alleles (Table 3). The allele number ranged from 3 for smSSR42 to 27 for smSSR3 with a mean of 10.6 alleles per marker. This mean value was higher than that obtained for eggplant genomic (6.7 alleles, 15) and genic (3.1 alleles, 19) SSRs; however, both of these previous studies used much fewer and less diverse eggplant relatives and focused more on polymorphism within *S. melongena*. Furthermore, when just the *S. melongena* accessions were considered, it was found that 79% of the smSSRs were polymorphic with an allele number ranging from 2 (smSSR21) to 12 (smSSR3). A total of 116 alleles were identified in the 8 *S. melongena* accessions (Table 3), which represent wild (group F), weedy (group E), primitive (group G), and advanced (group H) eggplant types. Thus, an average of 5.0 alleles per marker were detected in these accessions and this value is in better agreement with those obtained in previous studies of microsatellites in brinjal eggplant (15,19). Polymorphism was also examined in the other cultivated eggplants, *S. aethiopicum* and *S. macrocarpon,* and in *S. incanum*, the closest relative of *S. melongena* (29) (Table 3). Although only 3 *S. aethiopicum* and 4 *S. macrocarpon* accessions were assayed, the smSSRs showed considerable polymorphism. A total of 80 alleles were identified from the 25 markers that amplified products in *S. aethiopicum* and 93 alleles were identified for the 28 markers that amplified in *S. macrocarpon*. Thus, approximately 3.2 alleles were identified per marker in these 2 species of cultivated eggplants. Twice as many alleles were identified in the wild species *S. incanum* with an average of 6.2 alleles per marker. This greater polymorphism may reflect the greater genetic diversity of this species and/or the fact that many more accessions (17) were tested than for the other species. Overall, these results show that the smSSRs can be used for assessment of intraspecific genetic variation in the cultivated eggplants and their wild relatives.

When all accessions were examined, it was found that shorter SSR motifs (dinucleotides) and longer SSRs tended to be associated with a greater number of alleles. In other words, the number of alleles was negatively correlated with motif length ($r = -0.35$, $P = 0.07$) and positively correlated with total SSR length ($r = 0.34$, $P = 0.07$). Although these results were not statistically significant at $P = 0.05$, they demonstrate a general phenomenon that has been observed in previous studies in eggplant and related solanaceous species. Indeed, Stagel et al. (19) also found that longer genic SSRs were more informative and that dinucleotides were more variable than trinucleotides. For genomic trinucleotide SSRs, Nunome (15) observed that markers with more repeat units detected more polymorphism in *S. melongena* accessions, and in our work when only *S. melongena* alleles are considered, there are highly significant correlations between allele number and both motif length ($r = -0.49$, $P = 0.007$) and total length ($r = 0.68$. $P < 0.0001$). Similar results regarding SSR polymorphism have also been consistently obtained in tomato (25,37-39).

Table 3. Polymorphism of the SSRs in all accessions and different eggplant species. Nd= no data, marker did not amplify.

	# Alleles	**# Alleles**	**# Alleles**	**# Alleles**	**# Alleles**
SSR	**All Accessions**	***S.melongena***	***S.aethiopicum***	***S.macrocarpon***	***S.incanum***
smSSR1	12	11	4	1	9
smSSR3	27	12	11	15	15
smSSR4	23	5	2	9	15
smSSR9	13	3	2	3	8
smSSR11	6	5	5	3	5
smSSR12	5	1	nd	0	2
smSSR14	20	6	0	3	11
smSSR15	4	1	nd	1	2
smSSR16	10	3	3	3	7
smSSR17	7	1	1	2	1
smSSR18	8	1	0	1	4
smSSR19	8	3	3	2	2
smSSR21	4	2	1	2	2
smSSR22	10	3	2	2	6
smSSR24	25	4	3	2	12
smSSR27	6	5	3	3	5
smSSR28	16	4	3	3	13
smSSR29	7	5	5	4	6
smSSR31	10	3	1	1	7
smSSR35	11	1	2	1	6
smSSR36	12	4	3	1	5
smSSR37	5	5	3	3	4
smSSR40	5	3	3	1	4
smSSR41	5	4	3	5	5
smSSR42	3	1	1	1	2
smSSR44	11	4	4	4	4
smSSR45	11	4	3	5	6
smSSR46	11	5	4	6	4
smSSR47	12	7	5	6	9
TOTAL	**307**	**116**	**80**	**93**	**181**

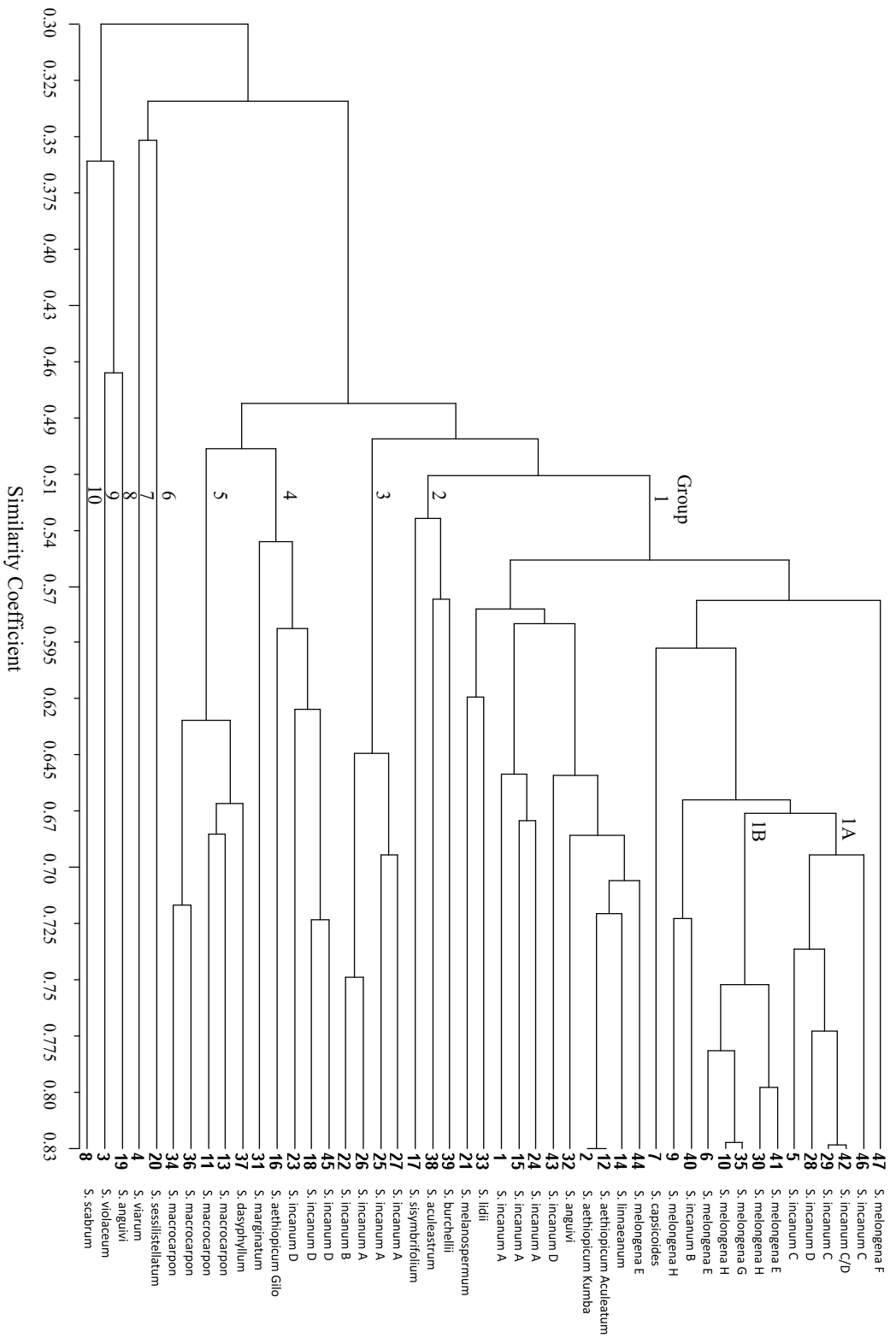

Figure. UPGMA dendrogram constructed from the smSSR data using the NTSYS software program. Samples are labelled with genotype number (from Table 1) and species name. Clusters 1A and 1B are described in the text.

The smSSRs were also tested for polymorphism on the parents of an F_2 mapping population (*S. melongena* MM738 and *S. linnaeanum* MM195, 9). Eight of the markers were polymorphic and mapped in the population on 5 linkage groups (data not shown). These results are partial confirmation that the smSSRs do not cluster to 1 or 2 genomic locations. This is an important criterion for use of these markers in mapping and also for their effective use in genetic diversity studies because they allow sampling of different regions of the genome.

Analysis of genetic relationships

According to the genetic analysis, the correlation between sample genotypic data and the dendrogram was quite high (r = 0.88) and indicated a good fit (40). The dendrogram scale varied from 0.30 to 0.83 with a mean similarity of 0.57 (Figure). Two accessions of *S. aethiopicum*, MM0134 (genotype number 12) and MM0574 (genotype 2), respectively from the Aculeatum and Kumba Groups (2 cultigroup with round, flat and fasciated fruits), showed the highest similarity (0.83). At a similarity of 0.53, the *Solanum* accessions fell into 10 groups (Figure). The largest group, group 1, contained 25 accessions, i.e. more than half of those tested, with a minimum similarity of ~0.56. Five of the other groups contained only 1 accession each with 4 of these groups representing species for which only 1 accession was analyzed: *S. viarum*, *S. violaceum*, *S. scabrum,* and *S. sessilistellatum*. These results indicate that these species were more distantly related to the rest of the accessions with a maximum similarity of only ~0.33 between these single-accession groups and the other material. Group 2 contained 3 single accession species: *S. burchellii*, *S. aculeastrum,* and *S. sisymbrifolium* with a mean similarity of ~0.54. In contrast, groups 1, 3, 4, and 5 consisted primarily of multiple accessions of *S. incanum* and the cultivated eggplants, *S. melongena*, *S. aethiopicum,* and *S. macrocarpon* (Figure). Within these groups, accessions of *S. melongena*, *S. aethiopicum*, *S. macrocarpon,* and *S. incanum* tended to fall into smaller species-specific clusters. For example, group 1A contained 5 of the 17 *S. incanum* accessions with a minimum similarity of 0.69 while group 1B contained 5 of the 8 *S. melongena* accessions used in this work with a minimum similarity of 0.75 (Figure). Despite this clustering, none of the species represented by multiple accessions formed strictly monophyletic groups when the smSSR data for analysis was used.

When the 3 cultivated eggplants and their closest wild relatives are considered, the dendrogram shows some interesting relationships. The putative progenitor of brinjal eggplant, *S. melongena*, is *S. incanum* (29) and in the dendrogram, more than half of its accessions clustered with *S. melongena* in group 1 (Figure). Several other molecular genetic studies of eggplant have confirmed the close genetic relationship between these 2 species using various marker systems including chloroplast DNA, isozymes, RAPDs, and AFLP (14,16,18,41-46). The *S. melongena* and *S. incanum* accessions had a minimum similarity of 0.49 (group 3) and a maximum similarity of 0.67 (groups 1A and 1B in Figure). For comparison, it should be noted that the minimum genetic similarity among *S. melongena* accessions was 0.56. Thus, some accessions of brinjal eggplant were more closely related to their immediate wild relative than to other cultivars. For example, *S. melongena* accession BIRM/S.2458 (genotype 9) was more similar to *S. incanum* MM1244 (genotype 40) than it was to the other *S. melongena* accessions. In addition, the *S. melongena* accessions in group 1B, which include weedy forms as well as primitive and advanced cultivars, were more closely related to the *S. incanum* accessions in group 1A than they were to a wild *S. melongena* accession (MM0686, genotype 47) and other weedy and cultivated types (genotypes 9 and 44). As expected, *S. incanum* had more genetic diversity than *S. melongena* with a minimum similarity between *S. incanum* accessions of only 0.47. This result and the placement of several *S. incanum* accessions outside of group 1 (in groups 3 and 4) may reflect the facts that: *S. incanum* is an aggregate of wild species (5) and that, unlike *S. melongena*, it has not been subjected to the selection pressures of domestication and breeding. Thus, *S. incanum* has maintained more genetic diversity than its cultivated relative.

S. aethiopicum, scarlet eggplant, is mainly cultivated in Africa for its fruits and leaves. *S. aethiopicum* was domesticated from *S. anguivi* (47). In the dendrogram, one of the *S. anguivi* accessions (MM0982, genotype

32) was found to cluster with *S. aethiopicum* accessions. However, *S. aethiopicum* Gilo (MM0232bis, genotype 16) and *S. anguivi* MM1259 (genotype 19) were found outside this cluster and very distant from each other indicating that both of these species maintain significant genetic variation for the tested markers. The close relationship between scarlet eggplant and *S. anguivi is* supported by DNA sequence analysis of 1 chloroplast and 2 nuclear regions (22) and by chloroplast DNA diversity (41). In our dendrogram, the high genetic similarity between *S. aethiopicum* Aculeatum (MM0134, genotype 12) and Kumba (MM0574, genotype 2) probably reflects the origin of Aculeatum, which, according to Lester and Niakan (28), was produced by selection of hybrids between *S. aethiopicum* Kumba and *S. anguivi*.

Like scarlet eggplant, *S. macrocarpon* or Gboma eggplant is cultivated in Africa for its fruits and leaves. *S. macrocarpon* was domesticated from *S. dasyphyllum* (28). Group 5 of the dendrogram indicated the close genetic relationship between these species (minimum similarity of 0.63) as this group consisted of the 4 *S. macrocarpon* accessions used in this work and *S. dasyphyllum*. Our finding confirms those of Mace et al. (14) and Levin et al. (22), who also found that *S. macrocarpon* and *S. dasyphyllum* were in the same cluster/clade with high similarity.

The dendrogram obtained with the eggplant genic SSRs has other similarities with previously published work. In general, the cultivated eggplants and their closest wild relatives were more closely related to each other than they were to the other wild species. Thus, *S. melongena* and *S. aethiopicum* had a maximum similarity of ~0.70 while *S. macrocarpon* had a maximum similarity of ~0.49 with these other 2 species (Figure). The close genetic relationship between the 3 cultivated eggplants has also been seen in phylogenetic studies of *Solanum* (16,21,42). *S. linnaeanum* showed a close genetic relationship with *S. aethiopicum* based on smSSR data (Figure). This was also observed by Furini and Wunder (16) using AFLP to examine *Solanum* phylogeny, but was not seen by Levin et al. (22) using nuclear and chloroplast DNA sequence data.

For plant breeders, cross compatibility is the ultimate proof of a close genetic relationship between species and molecular data do not always match compatibility data. For example, in our study, *S. sessilistellatum* had low similarity (0.33) with *S. melongena* accessions (Figure) despite the fact that the two species are easily cross-compatible (20). In contrast, *S. melongena* shares a 0.50 similarity with *S. aculeastrum* and *S. sisymbriifolium* but neither of these wild species is sexually compatible with brinjal eggplant (48). As a result, it is essential to remember that analyses based on molecular data are highly dependent on the number and type of marker chosen and the plant accessions tested. Interpretation of the genetic relationships among species/accessions will also depend on the point of view of the scientist (breeder vs. taxonomist vs. molecular geneticist) performing the analysis and, thus, should be performed with caution.

Conclusions

SSR marker systems are accepted as valuable molecular analysis tools (26). However, due to their conservative nature and expected low level of polymorphism, the usefulness of SSRs derived from ESTs for clustering analysis has been questioned (24,49). The results of the present study using genic microsatellite markers derived from an eggplant EST library show that such SSRs can be as polymorphic as genomic SSRs (15,35). In addition, the markers were found to be more informative in terms of the number of alleles revealed than RFLP (9) and RAPD (46) markers for analyses within *S. melongena*. AFLP markers have given both encouraging (14) and discouraging results (50) in intraspecific work; however, AFLP is generally considered to be more technically challenging than SSR analysis. The general agreement between the smSSR-derived dendrogram, the origins of domesticated eggplants, and the results obtained with other commonly used marker systems indicate that the eggplant microsatellite markers developed in this work are a reliable, simply applied, economical, and effective resource for investigating the genetic relationships between eggplants and their relatives. Thus, given the relative dearth of eggplant-specific markers as compared to those available for other solanaceous species, these SSR markers should be a valuable tool for eggplant breeders and germplasm conservationists who must perform their research with limited monetary resources and do not require the depth of phylogenetic information provided by nuclear or chloroplast DNA sequences.

Acknowledgments

The authors are grateful for financial support from the Scientific and Technological Research Council of Turkey (TÜBİTAK project no. 104T224). We also thank Dr. Amy Frary for helpful comments on the manuscript.

Corresponding author:
Sami DOĞANLAR
İzmir Institute of Technology,
Department of Molecular Biology and Genetics,
35430 Urla, İzmir - TURKEY
E-mail: samidoganlar@iyte.edu.tr

References

1. Mueller LA, Solow TH, Taylor N et al. The SOL genomics network. A comparative resource for Solanaceae biology and beyond. Plant Physiol 138: 1310-1317, 2005.
2. D'Arcy WG. The classification of the Solanaceae. In: JG Hawkes JG, RN Lester RN, AD Skelding AD ed. The Biology and the Taxonomy of the Solanaceae. Linnaean Society Symposium Series; 1979: pp. 3-47.
3. D'Arcy WG. The Solanaceae since 1976, with a review of its biogeography. In: JG Hawkes, RN Lester, M Nee, N Estrada-R ed. *Solanaceae* III, Taxonomy, Chemistry, Evolution. Royal Botanic Gardens; 1991: pp. 75-137.
4. Knapp S, Bohs L, Nee M et al. Solanaceae – a model for linking genomics with biodiversity. Comp Funct Genom 5: 285-291, 2004.
5. Daunay MC, Lester RN, Gebhardt CH et al. Genetic resources of eggplant (*Solanum melongena* L.) and allied species: a new challenge for molecular geneticists and eggplant breeders. In: RG Van Den Berg, GW Barendse, C Mariani ed. Solanaceae V, Nijmegen University Press; 2001: pp. 251–274.
6. Daunay MC, Lester RN, Laterrot H. The use of wild species for the genetic improvement of brinjal eggplant (*Solanum melongena*) and tomato (*Lycopersicon esculentum*). In: JG Hawkes, RN Lester, M Nee, N Estrada-R ed. *Solanaceae* III, Taxonomy, Chemistry, Evolution (Eds.). Royal Botanic Gardens; 1991: pp. 389-412.
7. Lester RN. Genetic Resources of Capsicum and Eggplants. In: X. EUCARPIAMeeting on Genetics and Breeding of Capsicum & Eggplant. Avignon France: pp. 25-

   30, 1998.
8. Daunay MC, Lester RN, Ano G. Eggplant. In: A Charrier, M Jacquot, S Hamon, D Nicolas ed. Tropical Plant Breeding. Science Publishers Inc.; 2001: pp. 199-222.
9. Doganlar S, Frary A, Daunay MC et al. A comparative genetic linkage map of eggplant *(Solanum melongena)* and its implication for genome evolution in the Solanaceae. Genetics 161: 1697-1711, 2002.
10. Doganlar S, Frary A, Daunay MC et al. Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. Genetics 161: 1713-1726, 2002.
11. Frary A, Doganlar S, Daunay MC et al. QTL analysis of morphological traits in eggplant and its implications for conservation of gene function during evolution of solanaceous species. Theor Appl Genet 107: 359-370, 2003.
12. Frary A, Doganlar S, Daunay MC. Eggplant. In: C Kole ed. Genome Mapping and Molecular Breeding in Plants. Springer Acad Press; 2007; pp. 287-314.
13. Mutlu N, Boyacı FH, Göçmen M et al. Development of SRAP, SRAP-RGA, RAPD and SCAR markers linked with a Fusarium wilt resistance gene in eggplant. Theor Appl Genet 117: 1303-1312, 2008.
14. Mace ES, Lester RN, Gebhardt CG. AFLP analysis of genetic relationships among the cultivated eggplant, *Solanum melongena* L., and wild relatives (Solanaceae). Theor Appl Genet 99: 626-633, 1999.
15. Nunome T, Suwabe K, Ohyama A et al. Characterization of trinucleotide microsatellites in eggplant. Breeding Science 53: 77-83, 2003.
16. Furini A, Wunder J. Analysis of eggplant (*Solanum melongena*)-related germplasm: morphological and AFLP data contribute to phylogenetic interpretations and germplasm utilization. Theor Appl Genet 108: 197-208, 2004.
17. Prohens J, Blanca JM, Nuez F. Morphological and molecular variation in a collection of eggplants from a secondary center of diversity: implications for conservation and breeding. J Am Soc Hort Sci 130: 54-63, 2005.
18. Singh AK, Singh M, Singh AK et al. Genetic diversity within the genus *Solanum* (Solanaceae) as revealed by RAPD markers. Curr Sci 90: 711-716, 2006.
19. Stàgel A, Portis E, Toppino L et al. Gene-based microsatellite development for mapping and phylogeny studies in eggplant. BMC Genomics 9: 357, 2008.
20. Daunay MC, Dalmon A, Lester RN. Management of a collection of *Solanum* species for eggplant (*Solanum melongena* L.) breeding purposes. In: M Nee, Lester RN, JP Jessop ed. *Solanaceae* IV, Advances in Biology and Utilization (Eds.). Royal Botanic Gardens, Kew; 1999: pp. 369-383.
21. Weese TL, Bohs L. A three gene phylogeny of the genus *Solanum* (Solanaceae). Systematic Botany 32: 445-463, 2007.

22. Levin RA, Myers NR, Bohs L. Phylogenetic relationships among the "Spiny Solanums" (*Solanum* subgenus *Leptostemonum*, Solanaceae). Amer J Bot 93: 157-169, 2006.

23. Jones N, Ougham H, Thomas H. Markers and mapping: we are all geneticists now. New Phytol 137: 165-177, 1997.

24. Varshney RK, Graner A, Sorrells ME. Genic microsatellite markers in plants: features and applications. Trends in Biotechnology 23: 48-55, 2005.

25. Areshchenkova T, Ganal MW. Long tomato microsatellites are predominantly associated with centromeric regions. Genome 42: 536-544, 1999.

26. Powell W, Machray GC, Provan J. Polymorphism revealed by simple sequence repeats. Trends in Plant Sci 1: 215-222, 1996.

27. Mohan M, Nair S, Bhagwat A et al. Genome mapping, molecular markers and marker-assisted selection in crop plants. Mol Breed 3: 87-103, 1997.

28. Lester RN, Niakan L. 1986 Origin and domestication of the scarlet eggplant, *Solanum aethiopicum*, from *S. anguivi* in Africa. In: Solanaceae: Biology and Systematics. (Ed. WG D'Arcy). Columbia Univ Press. New York: pp. 433-456, 1986.

29. Lester RN, Hasan SMZ. Origin and domestication of the brinjal-eggplant, *Solanum melongena*, from *S. incanum* in Africa and Asia. In: JG Hawkes, RN Lester, M Nee, N Estrada-R ed. *Solanaceae* III Taxonomy, Chemistry, Evolution. Royal Botanic Gardens, Kew; 1991: pp. 369-387.

30. Schuelke M. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18: 233-234, 2000.

31. Dice LR. Measures of the amount of ecological association between species. Ecology 26: 297-302, 1945.

32. Mantel N. The detection of disease clustering and a generalized regression approach. Cancer Res 27: 209-220, 1967.

33. Cho YG, Ishii T, Temnykh S et al. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa*). Theor Appl Genet 100: 713-722, 2000.

34. Thiel T, Michalek W, Varshney RK et al. Exploiting EST databases for the

development and characterization of gene-derived SSR markers in barley (*Hordeum*

vulgare L.). Theor Appl Genet 106: 411-422, 2003.

35. Metzgar D, Bytof J, Willis C. Selection against frameshift mutations limits microsatellite expansion in coding DNA. Genome Res 10: 72-80, 2000.

36. Nunome T, Suwabe K, Iketani H et al. Identification and characterization of microsatellites in eggplant. Plant Breeding 122: 256-262, 2003.

37. Smulders MJ, Bredemeijer G, Rus-Kortekaas W et al. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other Lycopersicon species. Theor Appl Genet 97: 264-272, 1997.

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

39. Frary A, Xu Y, Liu J et al. Development of a set of PCR-based anchor markers encompassing the tomato genome and evaluation of their usefulness for genetics and

breeding experiments. Theor Appl Genet 111: 291-312, 2005.

40. Mohammadi SA, Prasanna BM. Analysis of genetic diversity in crop plants-salient statistical tools and considerations. Crop Sci 43: 1235-1248, 2003.

41. Sakata Y, Lester RN. Chloroplast DNA diversity in eggplant (*Solanum melongena*) and its related species S. *incanum* and S. *marginatum*. Euphytica 80: 1-4, 1994.

42. Sakata Y, Lester RN. Chloroplast diversity in brinjal eggplant (*Solanum melongena* L.) and related species. Euphytica 97: 295-301, 1997.

43. Isshiki S, Okuba H, Fujieda K. Phylogeny of eggplant and related Solanum species constructed by allozyme variation. Scientia Horticulturae 59: 171-176, 1994.

44. Isshiki S, Okuba H, Fujieda K. Genetic control of isozymes in eggplant and its wild relatives. Euphytica 80: 145-150, 1994.

45. Isshiki S, Okuba H, Oda N et al. Isozyme variation in eggplant (*Solanum melongena* L.). J Japan Soc Hort Sci 63: 115-120, 1994.

46. Karihaloo JL, Brauner S, Gottlieb LD. Random amplified polymorphic DNA variation in the eggplant. Theor Appl Genet 90: 767-770, 1995.

47. Lester RN, Jaeger PML, Bleijendaal-Spierings BHM et al. African eggplants: a review

of collecting in West Africa. Plant Gen Res Newsl 81/82: 17-26, 1990.

48. Daunay MC. Eggplant. In: J Prohens, F Nuez ed. Handbook of Crop Breeding, Vol Vegetables: Fabaceae, Liliaceae, Umbelliferae and Solanaceae. Springer; 2008: pp. 163-220.

49. Rudd S. Expressed sequence tags: alternative or complement to whole genome sequences. Trends in Plant Sci 8: 321-329, 2003.

50. Nunome T, Ishiguro K, Yoshida T et al. Mapping of fruit shape and color development traits in eggplant (*Solanum melongena* L.) based on RAPD and AFLP markers. Breed Sci 51: 19-26, 2001.