African Journal of Biotechnology Vol. 10(45), pp. 9023-9031, 17 August, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.1029
ISSN 1684–5315 © 2011 Academic Journals

*Full Length Research Paper*

# Characterization of novel developed expressed sequence tag (*EST*)-derived simple sequence repeat (*SSR*) markers and their application in diversity analysis of eggplant

**Haiyan Ge, Huaizhi Li, Yang Liu, Xiang Li and Huoying Chen***

School of Agriculture and Biology, Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang District, Shanghai 200240, P. R. China.

Accepted 17 June, 2011

**A total of 101,270 eggplant expressed sequence tag (*EST*) sequences at public databases were used to search for simple sequence repeats (*SSR*s) and 405 potential SSR loci were identified from 388 sequences. The highest proportion (34.07%, 138) was represented by trinucleotide, followed by dinucleotide (19.51%, 79) and hexanucleotide (15.8%, 64). Among the dinucleotide repeats, AG/CT was the most common (55.69%), followed by AT/AT (31.64%) and AC/GT (12.66%). Further, 288 pairs of primers were developed from these sequences. A random set of 100 EST-SSR primers were amplified in 12 eggplant accessions and 88 successfully amplified expect PCR products. 32 markers revealed 83 polymorphic alleles among the 42 cultivated accessions and the number of allelles per locus varied between 2 and 6 (mean 2.6). Polymorphism information content (PIC) values among the 42 cultivated types were calculated and varied from 0.045 to 0.701 (mean 0.289). The markers showed low frequency transferability in Solanaceae. The 32 SSRs were used to evaluate genetic diversity. These SSRs will be valuable markers for future genetic study, such as genetic diversity estimation, linkage mapping, association mapping and molecular breeding.**

**Key words:** Expressed sequence tags, transferability, diversity, dendrogram.

## INTRODUCTION

Microsatellites (simple sequence repeats, SSRs) are short tandemly repeated motifs of 1-6 nucleotide units, and their value for genetic analysis lies in their multi-allelic nature, codominance, relative abundance, extensive genome coverage, high resolution and easy detection by PCR with small amount of genomic DNA template (Powell et al., 1996; Stagel et al., 2008). SSRs have been widely used for studies of genetic variability, linkage mapping, gene tagging and map-based gene cloning (Huang et al., 2010). About 1.5 to 4.7% of the ESTs in different plant species were reported to contain SSRs suitable for marker development (Kantety et al., 2002). The EST-based markers have been successfully used in many species, such as barley (Castillo et al., 2008; Emebiri, 2009), wheat (Ercan et al., 2010; Varshney et al., 2005), rice (Varshney et al., 2005), soybean (Liu et al., 2010; Mulato et al., 2010), cotton (Lu et al., 2010), sugarcane (Oliveira et al., 2007), grape (Huang et al., 2010), coffee (Aggarwal et al., 2007), peanut (Liang et al., 2009; Song et al., 2010), cucumber (Hu et al., 2010), safflower (Chapman et al., 2009), eggplant (Stagel et al., 2008; Tumbilen et al., 2011) and pepper (Huang et al., 2001; Lee et al., 2004; Nagy et al., 2007; Portis et al., 2007).

Eggplant (*Solanum melongena* L.), a member of Solanaceae, is an important vegetable in many countries. It is a good source of minerals and vitamins, and some polyphenols which show potent antioxidant activity (Nisha et al., 2009; Sudheesh et al., 1999). Despite the wide-spread cultivation and economic importance, its molecular genetics studies were behind those of tomato,

*Corresponding author. E-mail: chhy@sjtu.edu.cn. Tel: +86(0)21 3420 6934. Fax: +86(0)21 3420 6934.

**Table 1.** Fifty-four accessions included in the present study.

| Code | Name | Scientific name | Source |
|---|---|---|---|
| **Sample panel 1** | | | |
| 1 | Suanjie Qie | *S. sisymbriifolium* | |
| 2 | Hongqie72 | *S. aethiopicum* | China |
| 3 | Huangguo Qie | *S. xanthocarpum* | Guizhou,China |
| 4 | Congo Qie | *S. aethiopicum* | Congo |
| 5 | Guanshang Qie | *S. aethiopicum* | China |
| 6 | Jiló | *S. aethiopicum* | China |
| 7 | Guangdong Qing Qie | *S. melongena* | Guangdong,China |
| 8 | Zi Chang Qie | *S. melongena* | China |
| 9 | Liangshui Qie | *S. melongena* | Jiangxi,China |
| 10 | Ep 143 | *S. melongena* | India |
| 11 | Dian Qie | *S. melongena* | Yunnan,China |
| 12 | Bai Qie | *S. melongena* | Yunnan,China |
| **Sample panel 2** | | | |
| 13 | 96-2 | *S. melongena* | China |
| 14 | Malaysia Qie | *S. melongena* | Malaysia |
| 15 | Shenxian Yuan qie | *S. melongena* | Shandong,China |
| 16 | Zi Hebao Qie | *S. melongena* | China |
| 17 | Hexian Qie | *S. melongena* | China |
| 18 | Rubai Qie | *S. melongena* | China |
| 19 | Xiao Bai Qie | *S. melongena* | China |
| 20 | Pinghu Bai Chang Qie | *S. melongena* | Zhejiang,China |
| 21 | Songjiang Qie | *S. melongena* | Shanghai,China |
| 22 | Pinghu Hong Qie | *S. melongena* | Zhejiang,China |
| 23 | Hangzhou Hong Qie | *S. melongena* | Zhejiang,China |
| 24 | Meiguo Da Chang Qie | *S. melongena* | USA |
| 25 | Zi Hei Chang Qie | *S. melongena* | China |
| 26 | Xian Qie | *S. melongena* | China |
| 27 | Hangzhou Tiao Qie | *S. melongena* | Zhejiang,China |
| 28 | Liu Tiao Qie | *S. melongena* | Liaoning,China |
| 29 | Qing Yangjiao Qie | *S. melongena* | China |
| 30 | Pingdong Chang Qie | *S. melongena* | Taiwan,China |
| 31 | Arka Keshav | *S. melongena* | India |
| 32 | Local-1 | *S. melongena* | India |
| 33 | Dunhe Qie | *S. melongena* | China |
| 34 | Chang Zi Qie | *S. melongena* | China |
| 35 | Qing Qie | *S. melongena* | Hainan,China |
| 36 | Nantong Qing Qie | *S. melongena* | Zhejiang,China |
| 37 | Fushe No.1 | *S. melongena* | China |
| 38 | Chang Qie | *S. melongena* | Fujian,China |
| 39 | Lanzhou Chang Qie | *S. melongena* | Gansu,China |
| 40 | Xiao Hongpao | *S. melongena* | China |
| 41 | Niujiao Qie | *S. melongena* | China |
| 42 | Benxi Zao Zi Qie | *S. melongena* | Liaoning,China |
| 43 | Longmenchi Zi Qie | *S. melongena* | Guangdong,China |
| 44 | Lv Qie | *S. melongena* | Yunnan,China |
| 45 | Diana Qie | *S. melongena* | China |
| 46 | Baicuo Qie | *S. melongena* | China |
| 47 | Hu Qie | *S. melongena* | Shanghai,China |
| 48 | Dian Xian Bai Qie | *S. melongena* | China |

**Table 1.** Continue.

| | | | |
|---|---|---|---|
| **Sample panel 3** | | | |
| 49 | Tomato Y43 | *S. pimpinellifolium* | America |
| 50 | Tomato Y5 | *S. chiele* | America |
| 51 | Tomato C1 | *S. lycopersicon* | Shanghai, China |
| 52 | potato | *S. tuberosum* | China |
| 53 | Ornamental Pepper | *C. annuum* | China |
| 54 | pepper | *C. annuum* | China |

Sample panel 1 was used to test PCR amplification and polymorphisms for the 100 primers. The primers that detected polymorphism in sample panel 1 were tested further using DNAs from sample panel 1 and sample panel 2. The transferability of the 100 markers was evaluated with the accessions in sample panel 3.

potato, and pepper, especially in the aspect of high density linkage map construction (Nunome et al., 2009). Several linkage maps had been reported in eggplant (Barchi et al., 2010; Cao et al., 2006; Doganlar et al., 2002a, b; Nunome et al., 2001, 2003, 2009; Sunseri et al., 2003; Wu et al., 2009), however, the populations for mapping were limited. Up till now, all the interspecific $F_2$ populations were developed from the cross between *Solanum linnaeanum* and *S. melongena* (Doganlar et al., 2002a, b; Sunseri et al., 2003; Wu et al., 2009). The other maps were all constructed on the intraspecific populations. This was because of the hybrid obstacle between interspecific species. SSR markers had been developed and used for mapping in eggplant (Nunome et al., 2009), however, the level of intraspecific DNA marker polymorphism were rather limited. To construct a high density linkage map, much more markers are needed. Because the eggplant genetic research lagged behind other crops, little sequences were available and used from public data banks. Fukuoka et al. (2009) submitted volumes of EST sequences to the public databases, including 98,086 pieces of EST containing 50,438,137 bp nucleotides, which enriched the eggplant databases and gave much information for SSR searching and developing.

In this study, we reported the characterization of novel eggplant EST-SSR markers developed from public databases and their application in genetic diversity evaluation.

## MATERIALS AND METHODS

### Plant materials and DNA extraction

Forty-eight (48) morphologically different eggplant accessions were selected in this study (Table 1). The 1st to 6th eggplants were wild relatives, while the 7th to 48th eggplants were cultivated accessions. The accessions in sample panel 1 were used to test PCR amplification and detect polymorphisms. SSRs displaying polymorphism in six cultivated accessions (in panel 1) were subsequently tested against panel 1 and 2 for data acquisition and analysis. Accessions in panel 3 including 3 tomato accessions, 2 pepper accessions and 1 potato accession were used to study the transferability of eggplant EST-SSRs (Table 1). DNA extraction followed an improved procedure (Paterson et al., 1993), however without DIECA addition in the DNA extraction buffer.

### Eggplant EST data retrieval and SSR detection

A total of 101,270 eggplant EST sequences were retrieved from the NCBI and Solanaceae Network database (SGN). Of all these ESTs, 98,089 ESTs were gotten from the NCBI website. The other 3,181 ESTs were retrieved from the SGN (http://solgenomics.net). PolyA and polyT tracts were removed using the EST-trimmer software (http://pgrc.ipk-gatersleben.de/misa/), by applying the criterion that no 50 bp window contain a run of five A's or five T's. The ESTs were assembled using the CAP3 assemble software (Huang and Madan, 1999). Identification and localization of microsatellites were carried out by MISA software (http://pgrc.ipk-gatersleben.de/misa/) (Thiel et al., 2003; Zhang et al., 2002). SSR motifs were searched with the criteria as follows: 20 repeats for mononucleotide, 10 repeats for dinucleotide, 7 repeats for trinucleotide, 5 repeats for tetranucleotide, 4 repeats for penta- and hexanucleotide, and 3 repeats for heptanucleotide. Primer pairs were designed from the flanking sequences, using PRIMER3 software (Rozen and Skaletsky, 2000) in batch mode via the p3_in.pl and p3_out.pl Perl5 scripts within the MISA package. The target amplicon size was set as 100 to 400 bp. Melting temperatures ranging from 55 to 59°C were tested, and the optimal temperature was found to be 57°C. The redundant primer pairs were analyzed using the BLAST (Altschul et al., 1990) software. The main parameters were "-p blastn -m 8 –F F –e 0.32". The result was analyzed using script of PERL. It extracted primer pairs located in the same sequence. The match had no gaps. If several primer pairs were located in the same sequence with the same SSR site, we thought these primer pairs were redundant.

### PCR amplification and polyacrylamide gel analyses

PCR was carried out in a 10 µl reaction mixture containing 20 ng template DNA, 0.1 µM of forward and reverse primers each, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 1×*Taq* buffer and 1 U *Taq* DNA polymerase (Shanghai Promega). Amplification was performed in a 96 well thermocycler (Eppendorf AG 6321). Cycles were programmed as follows: one cycle of 95°C for 2 min, 30 cycles of 94°C for 45 s, 55°C for 45 s and 72°C for 60 s, and one cycle of 72°C for 7 min, stored at 4°C. The PCR products were separated on 5% polyacrylamide gel and visualized by silver staining (Zhang et al., 2002).

**Table 2.** Number and percent of major repeat motifs of EST-SSR in eggplant.

| EST-SSR type | Motif | Number | Frequency (%) | EST-SSR type | Motif | Number | Frequency (%) |
|---|---|---|---|---|---|---|---|
| Mononucleotide | A/T | 29 | 7.16 | Pentanucleotide | AAAAT/ATTTT | 6 | 1.48 |
| Dinucleotide | AG/CT | 44 | 10.86 | | AAAAG/CTTTT | 3 | 0.74 |
| | AT/AT | 25 | 6.17 | | AAACC/GGTTT | 3 | 0.74 |
| | AC/GT | 10 | 2.47 | | AAAGC/CTTTG | 3 | 0.74 |
| Trinucleotide | AAG/CTT | 53 | 13.09 | | AATAT/ATATT | 3 | 0.74 |
| | AAT/ATT | 29 | 7.16 | Hexanucleotide | AAGAGG/CCTCTT | 6 | 1.48 |
| | AAC/GTT | 16 | 3.95 | | ACCTCC/AGGTGG | 6 | 1.48 |
| | AGC/CTG | 15 | 3.70 | | AGGCGG/CCGCCT | 4 | 0.99 |
| | ATC/ATG | 8 | 1.98 | | AAAAAT/ATTTTT | 4 | 0.99 |
| | ACC/GGT | 7 | 1.73 | | AAGGAG/CCTTCT | 3 | 0.74 |
| | AGG/CCT | 3 | 0.74 | HEPTANUCLEOTIDE | AAAAAAT/ATTTTTT | 5 | 1.23 |
| | CCG/CGG | 3 | 0.74 | | AAAAAAG/CTTTTTT | 3 | 0.74 |
| Tetranucleotide | AAAG/CTTT | 10 | 2.47 | | AAATCTC/AGATTTG | 3 | 0.74 |
| | AAAT/ATTT | 10 | 2.47 | | | | |
| | AAAC/GTTT | 5 | 1.23 | | | | |

The motifs with a frequency of <0.5% were not listed in the table.

**Data acquisition and statistical analysis**

The amplified bands of EST-SSRs were recorded as present (1) or absent (0), thus generating a binary data matrix. Polymorphism information content (PIC) value of each EST-SSR marker was calculated by Anderson et al. (1993). PIC indices were calculated for the information content in the 42 cultivated genotypes. Cluster analysis was performed using the unweighted pair-group method with arithmetic averages (UPGMA) and a dendrogram was constructed using the NTSYS-pc software version 2.10t (Rohlf, 2000).

## RESULTS

### Characterization of EST-SSRs

In this study, all the 101,270 eggplant EST sequences were retrieved from the public databases. The non-redundant sequence pool contained 23,313 sequences. A total of 405 potential SSR loci were identified from 388 sequences, which represented 1.66% of all ESTs. Among the 405 EST-SSRs, most (397, 98.02%) consisted of simple repeats, whereas only few (8, 1.98%) were of the compound type. The mean distance of the SSRs was 39.46 kb. The different SSR unit size was not evenly distributed. Of the total 405 SSRs, the highest proportion (34.07%) was represented by trinucleotide with the number of 138, followed by dinucleotide (19.51%, 79) and hexanucleotide (15.80%, 64). The frequency of tetranucleotide and mononucleotide was low for 6.67 and 7.16%. So, trinucleotide, dinucleotide and hexanucleotide repeats represented majority of EST-SSRs in eggplant. Among the dinucleotide repeats, AG/CT was the most common (55.69%), followed by AT/AT (31.64%), and AC/GT (12.66%), whereas GC/GC did not appear. As for trinucleotide repeats, AAG/CTT, AAT/ATT, AGC/GCT, ATC/GAT and AAC/GTT were common motifs (87.69%). In hexanucleotide repeats, AAGAGG/CCTCTT, ACCTCC/AGGTGG, AAAAAT/ATTTTT and AGGCGG/CCGCCT were the most common repeats (35.94%) (Table 2). In all the repeat motifs, most of the SSR repeat motifs derived from the ESTs were AAG/CTT (13.09%), followed by AG/CT (10.86%), AAT/ATT (7.16%) and A/T (7.16%). In the 1-7 repeat types, the most frequent repeat motifs were A/T, AG/CT, AAG/CTT, AAAG/CTTT and AAAT/ATTT, AAAAT/ATTTT, AAGAGG/CCTCTT and ACCTCC/AGGTGG, and AAAAAAT/ATTTTTT, which accounted for 100, 55.69, 38.41, 37.02, 18.75, 9.38 and 13.16% of all types, respectively

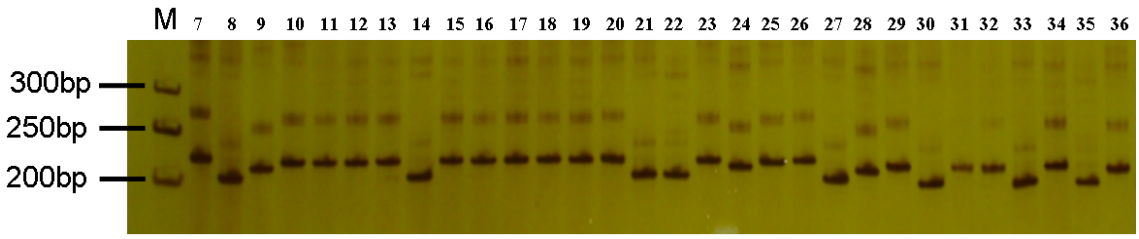


**Figure 1.** Genotypes of the 30 accessions at EES066 loci. M, marker. Genotypes and their order are shown in Table 1.

(Table 2).

### Development of EST-SSR markers and their polymorphisms

All the 388 sequences were used for primer design by the software PRIMER3 and only 289 (74.48%) EST-SSR primers were obtained from these sequences. Among the 289 primers, only one pair of primers was redundant, so 288 non-repeated primers were produced and these 288 primers were different from the eggplant SSR primers reported before (Nunome et al., 2003, 2009; Stagel et al., 2008; Tumbilen et al., 2011). Out of the 288 primer pairs, a random set of 100 EST-SSR primers was selected for PCR optimization, characterization and amplification with 48 selected eggplant accessions. The result showed that 88 (88%) primer pairs were successfully amplified expect the PCR products (Figure 1). The remaining 12 primer pairs (12%) failed to amplify or were amplified weakly. Among the 88 primers, 9 had no polymorphism among 12 accessions in panel 1, while the remaining 79 markers revealed 323 polymorphic alleles. As for the wild relatives, 64 of the 79 markers revealed 155 polymorphic alleles; as for the cultivated accessions, 32 of 79 markers revealed 83 polymorphic alleles. The details of the 32 EST-SSR primers are listed in Table 3. Of the 32 primers, 2 (EES019 and EES033) could not amplify any product in the wild relatives but could reveal polymorphisms among the cultivated accessions. Another 4 primers (EES020, EES035, EES076 and EES084) showed no polymorphism among the wild relatives but did in the cultivated accessions. A total of 83 alleles were amplified from 42 cultivated accessions, with the number of alleles per locus varying between 2 and 6 (mean 2.6). PIC values among the 42 cultivated types were calculated and varied between 0.045 and 0.701 (mean 0.289). Primer EES038 had the highest PIC, while EES033 and EES067 had the lowest. The correlation coefficient between PIC and SSR length was 0.03. Our study showed that 32% of the EST-SSRs were polymorphic.

### Transferability of EST-SSRs in Solanaceae

The transferability of the developed EST-SSR markers was evaluated with three Solanaceous crops including tomato, pepper and potato. Out of the 100 selected primer pairs, 31 (31%) primers could amplify PCR products from at least one of the three species successfully, 27 (27%) in tomato, 27 (27%) in potato and 24 (24%) in pepper. Of the 27 primers that could produce amplicons in tomato, 3 primers (EES043, EES080 and EES081) showed polymorphism in the three tomato accessions. In this study, low transferability was obtained for tomato, potato and pepper.

### Diversity analysis

A dendrogram based on the similarity coefficients of the 48 accessions was constructed (Figure 2). The dendrogram scale varied from 0.21 to 0.91. The dendrogram indicated a clear separation between the cultivated species and the wild relatives. Hong Qie, Congo Qie, Guangshang Qie and Jiló were grouped into cluster I, which belonged to *Solanum aethiopicum*. Cluster II contained two wild species, Suanjie Qie (*Solanum sisymbriifolium*) and Huangguo Qie (*Solanum xanthocarpum*). Cluster III consisted of the cultivated species (*S. melongena*). Cluster III-1 consisted of the accession of Bai Qie, which is a prickly plant that bears small, green, striped and round fruit. Thus, Bai Qie had a close relationship with the wild accessions. Cluster III-2 consisted of the other cultivated accessions with a mean similarity of 0.62. The cluster distant to the cultivated group was the *S. aethiopicum* accessions with a mean similarity of 0.25. The cluster closest to the cultivated group contained both *S. sisymbriifolium* and *S. xanthocarpum*, with a mean similarity of 0.30.

## DISCUSSION

### Characterization of EST-SSRs

The result that trinucleotide, dinucleotide and hexanucleotide repeats represented majority of EST- SSRs was in agreement with the previous observations of SSR repeat units in barley, maize, rice, sorghum, wheat and grape (Huang et al., 2010; Kantety et al., 2002). The

**Table 3.** Characterization of 32 polymorphic EST-SSRs in eggplant.

| Name | Repeat motif | Primer sequence (5'-3') forward | Primer sequence (5'-3') reverse | Product size (bp) | Number of allele | PIC |
|---|---|---|---|---|---|---|
| EES08 | (CAG)7 | TCTTGGGTTTTTCCTTTTTG | TCAGAAATTCTAGCAAGGGG | 194 | 2 | 0.371 |
| EES015 | (TGCAGG)4 | TGGAAACAACGATAACATGG | GAGGGTGAGACAGAAGGTTG | 108 | 3 | 0.587 |
| EES019 | (AAT)16 | TTGTCTCATTGTTGGTATGGA | GCCCATTGTTGAGGTGATTA | 265 | 2 | 0.375 |
| EES021 | (GCC)8 | AAAAATCCCCAAATCCATCT | ACGCTCTCTCACAACAACAA | 352 | 2 | 0.305 |
| EES022 | (ATT)7 | CAAAGTACCTTCCATTTATCCAG | CAGGTGCAGGTATCATCGTA | 219 | 2 | 0.374 |
| EES026 | (TTTGC)4 | GATGGAATTCAACAGTTACACAA | GGTCAATCCTGGTAAAGGTG | 305 | 2 | 0.157 |
| EES028 | (CT)11 | ACCGTTCTCGTCTCTTTGTC | CAACAACAGTTCAACCCAAA | 266 | 2 | 0.319 |
| EES029 | (TAGTGC)4 | GCCAATCACAAAATGTTGAA | GATCTTGAACAACTCCAGGG | 396 | 2 | 0.067 |
| EES030 | (ACA)7 | CATTCTACCGTCTCCAAACC | AAACAGCCGCTCTACCTCTA | 289 | 2 | 0.087 |
| EES031 | (AT)12 | AGAGGAGAAAGCGCTAGACA | TGATCAATCTTTGCATCCAC | 226 | 6 | 0.636 |
| EES033 | (TTC)7 | CCCTATTGGTTTCTTCCAAA | GAGCAAGGAGGATCAGAGAA | 129 | 2 | 0.045 |
| EES038 | (T)20 | ATACTGGGCTGGACCACTTA | GAAATAACAAAATCCGTCCG | 240 | 4 | 0.701 |
| EES040 | (CCA)7 | CTCGAGTACAAAAACATGCG | CTGCTGAATTTCTTGGCTTT | 200 | 2 | 0.36 |
| EES043 | (CTC)7 | AATGCCAGGACATCTGAAAT | AAACGGAAACGATGAAGAAG | 271 | 5 | 0.626 |
| EES045 | (A)20 | CAAATCAAAGAATGTGCTGC | AATTCCTTAGTTGCTCCGTG | 203 | 2 | 0.297 |
| EES049 | (TAC)7 | CAAATCTCCACCAACATCAA | TTGATGAAACCCCAAATCTT | 302 | 3 | 0.435 |
| EES050 | (CCTGCA)4 | CTCCAGAATCTGCTCCTGTT | CCACCACCCATATCAAGAAT | 166 | 2 | 0.124 |
| EES051 | (ATA)8 | CATCCACAATTTCAAAACAAA | TGAAAGCCATGAGATGCTAA | 365 | 3 | 0.316 |
| EES054 | (TCT)14 | TCCCTTAACTTTTCCCTCCT | TTCCCATCCAAGAACAAACT | 316 | 2 | 0.188 |
| EES062 | (TTTG)5 | GAACAACCCAACAAAAATCC | GGACAACGACGAAGAAGAGT | 336 | 2 | 0.087 |
| EES063 | (TC)17 | AGCAAACATTACAAAAGCAGTT | TCAGGCATCAGTATCACCAC | 258 | 4 | 0.512 |
| EES064 | (TTTA)6 | CAGCCGAAGTGATAAAGGTG | CCGAGATTAAACGAAAATGC | 205 | 2 | 0.363 |
| EES065 | (GCT)7 | CATCAGACATATTCGGAGCA | AAGAGAGATGCAGAACCCTG | 382 | 3 | 0.409 |
| EES066 | (TAT)7 | CAATTTGTAAACACTAGTACCATCA | CAGCTGATTCAAGGAGGAGT | 200 | 3 | 0.354 |
| EES067 | (AC)11 | GGCCCTGCTTTGTTATATTT | CTCACAGTGCTGATCGTAGG | 375 | 2 | 0.045 |
| EES068 | (AT)10 | GATTTCCGATACTTTCCCCT | ACGATCCAGATGCCATACTT | 175 | 4 | 0.091 |
| EES071 | (AAT)8 | ACACAAACTGGCAACTTCAA | ATGCTTCGAGGACTTTTGTC | 184 | 3 | 0.069 |
| EES075 | (TC)16 | TTAATTTCGTCTGGACGTTG | TTCAAGCAAGCGACTGATTA | 232 | 2 | 0.239 |
| EES080 | (ACCCAT)6 | GCATCTGATATCCTTGACCC | CCAAACCAAATGGTAGGTTC | 217 | 2 | 0.188 |
| EES085 | (AC)11 | ACCTCATCTTCCCTTTCCTT | TTAAACGTCCGTTGCTTGTA | 176 | 2 | 0.086 |
| EES091 | (TTC)7 | GAAGGTTGGTTTTCCATGAG | TCATCCCAAGAATTCCAGTT | 353 | 2 | 0.124 |
| EES094 | (AAAAAAT)3 | CGTTTTAGTCACCGTTGATG | TCGCACGAGAAAATCTCAC | 216 | 2 | 0.312 |

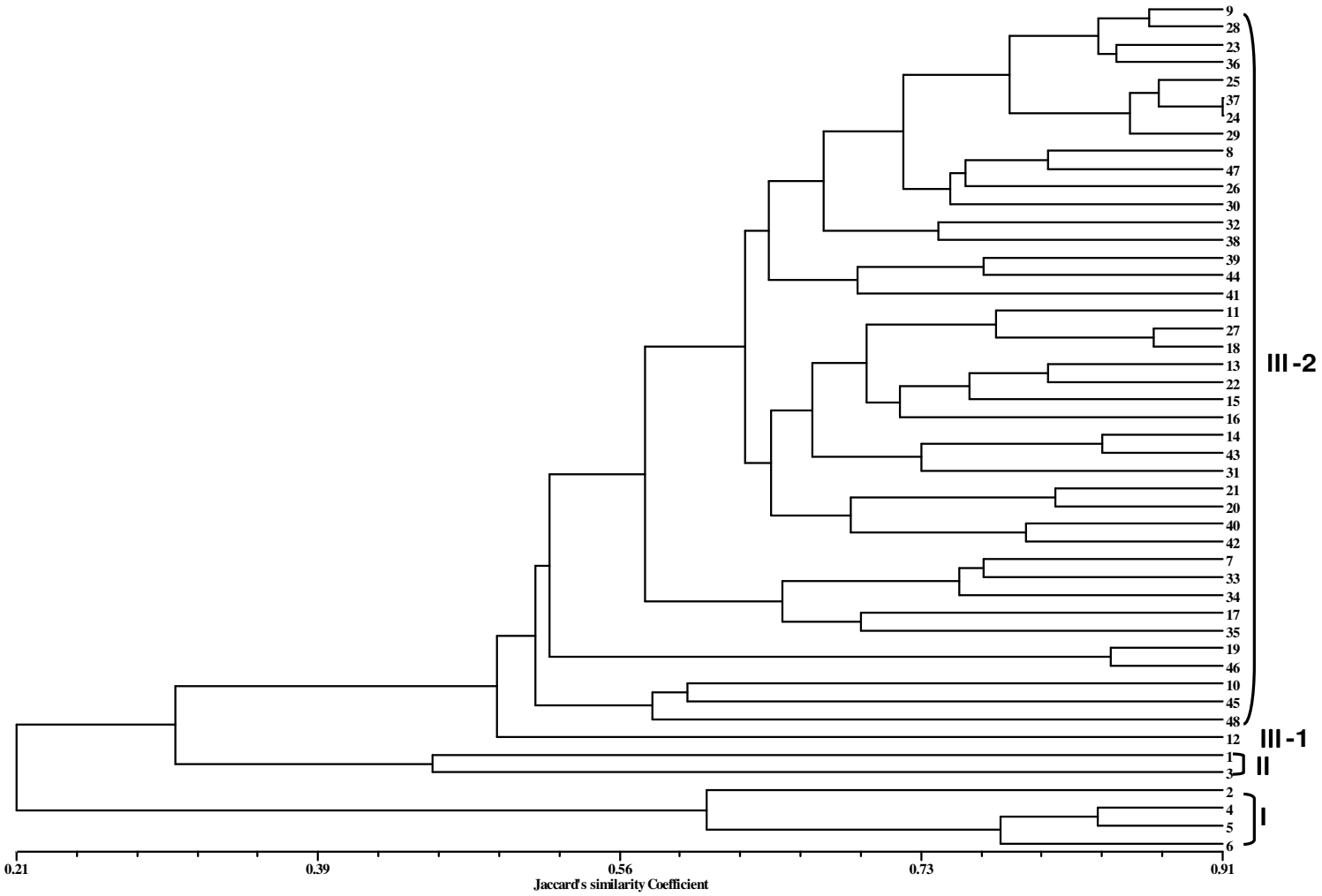


**Figure 2.** A dendrogram constructed based on Jaccard's similarity coefficient and UPGMA clustering. The samples are labeled with the codes liste

trinucleotide motifs represented the most common class in the expressed sequences (Stagel et al., 2008). The dominance of trinucleotide and hexanucleotide SSRs was viewed as the result of the frame shift in the size of an amino acid read, or the three nucleotides, a selection against possible frame shift mutations (Huang et al., 2010). Fukuoka et al. (2010) also reported that AG/CT was most common, followed by AT/AT and AC/GT among the dinucleotide repeats. The same result agreed with that of pepper. The result that AG/CT ranked first agreed with that of grape (Huang et al., 2010) and cotton (Lu et al., 2010). In eggplant, Stagel et al. (2008) also reported that trinucleotide represented the most common repeat, with AAG/CTT the most frequent. In summary, many researchers had gotten similar results in Solanaceous crops (Fukuoka et al., 2010; Nunome et al., 2003, 2009; Stagel et al., 2008). But, the results were not the same, perhaps because the criteria and the size of the sequence dataset were different.

## Development of EST-SSR markers and their polymorphisms

The correlation coefficient between PIC and SSR length indicated that there was no clear correlation between SSR length and informativeness. Similar observations had been reported in pepper and some other species (Nagy et al., 2007). However, in other studies, researchers had found correlation between repeat length and informativeness (Frary et al., 2005; Stagel et al., 2008; Tumbilen et al., 2011). Studies have reported that of the 25 studies across a variety of plant species, on average, 17.7% of loci producing PCR products were monomorphic (Squirrell et al., 2003). Low frequency of DNA polymorphism of most of the SSR markers had been observed in cultivated eggplants. Nunome et al. (2003, 2009) had reported that 56.7% of genomic SSRs and 30.3% of EST-SSRs were polymorphic. Studies by Stagel et al. (2008) showed that only 28.2% SSRs were

informative among the cultivated eggplant. Our result also indicated that only 32 EST-SSRs were informative among the cultivated eggplant, which was in agreement with the earlier mentioned researches. The reason might be because of the intensive breeding efforts and a narrow genetic background (Nunome et al., 2003).

### Transferability of EST-SSRs in Solanaceae

In this study, low transferability was obtained for tomato, potato and pepper. In Solanaceous plants, transferability between potato and tomato, and from tomato to eggplant had been confirmed (Nunome et al., 2003; Stagel et al., 2008) and also it had been reported that only few of tomato SSRs (15/600) can be applied to eggplant (Li et al., 2010), which was in agreement with this study. The transferability frequency was low when compared with that of cucumber (Hu et al., 2010) and previous studies in Solanaceae (Nunome et al., 2003, 2009; Stagel et al., 2008), all of which reached 50%. The reasons might be because of the far genetic relationships between eggplant and the other Solanaceous plants and the larger amount of EST numbers used than before. Though the transferability was not high enough, it can be exploited as anchor markers for Solanaceous comparative mapping.

### Diversity analysis

The genetic relationships between the 48 eggplant genotypes as displayed by genetic similarity at the SSR level were in good agreement with prior taxonomic classification based on AFLP markers (Furini and Wunder, 2004), SRAP markers (Li et al., 2010) and SSR markers (Stagel et al., 2008). The result that *S. aethiopicum* accessions were not the closest to cultivated accessions agreed with the study by Stagel et al. (2008) but not with the research by Furini and Wunder (2004) and Tumbilen et al. (2011). Tumbilen (2011) had reported that genetic diversity analysis based on molecular data was highly dependent on the number and type of marker chosen and the plant accessions tested. Also, the interpretation of the genetic relationships 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.

## ACKNOWLEDGEMENTS

The authors thank Dingshi Zha at the Shanghai Academy of Agricultural Sciences for providing some materials. This research was supported in part by the 863 program of China (no. 2008AA10Z150-4), Shanghai Municipal Agricultural Commission (2007 # 5-1) and the Shanghai Leading Academic Discipline Project (B209).

## REFERENCES

Aggarwal RK, Hendre PS, Varshney RK, Bhat PR, Krishnakumar V, Singh L (2007). Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. Theor. Appl. Genet. 114: 359-372.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

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

Barchi L, Lanteri S, Portis E, Stagel A, Vale G, Toppino L, Rotino GL (2010). Segregation distortion and linkage analysis in eggplant (*Solanum melongena* L.). Genome, 53: 805-815.

Cao BH, Lei JJ, Sun XD, Chen GJ, Meng CM (2006). Construction of RAPD markers linkage map for eggplant. Acta Horticulturae Sinica, 33:1092-1092.

Castillo A, Budak H, Varshney RK, Dorado G, Graner A, Hernandez P (2008). Transferability and polymorphism of barley EST-SSR markers used for phylogenetic analysis in Hordeum chilense. BMC Plant Biol.. 8: 97-105.

Chapman MA, Hvala J, Strever J, Matvienko M, Kozik A, Michelmore RW, Tang SX, Knapp SJ, Burke JM (2009). Development, polymorphism, and cross-taxon utility of EST-SSR markers from safflower (*Carthamus tinctorius* L.). Theor. Appl. Genet. 120: 85-91.

Doganlar S, Frary A, Daunay MC, Lester RN, Tanksley SD (2002a). A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the Solanaceae. Genetics, 161: 1697-1711.

Doganlar S, Frary A, Daunay MC, Lester RN, Tanksley SD (2002b). Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. Genetics, 161: 1713-1726.

Emebiri LC (2009). EST-SSR markers derived from an elite barley cultivar (*Hordeum vulgare* L. 'Morex'): polymorphism and genetic marker potential. Genome, 52: 665-676.

Ercan S, Ertugrul F, Aydin Y, Akfirat FS, Hasancebi S, Akan K, Mert Z, Bolat N, Yorgancilar O, Altinkut-Uncuoglu A (2010). An EST-SSR marker linked with yellow rust resistance in wheat. Biol. Plantarum, 54: 691-696.

Frary A, Xu Y, Liu J, Mitchell S, Tedeschi E, Tanksley S (2005). Development of a set of PCR-based anchor markers encompassing the tomato genome and evaluation of their usefulness for. Appl. Genet. 111: 291-312.

Fukuoka H, Yamaguchi H, Nunome T, Negoro S, Miyatake K, Ohyama A (2010). Accumulation, functional annotation, and comparative analysis of expressed sequence tags in eggplant (*Solanum melongena* L.), the third pole of the genus Solanum species after tomato and potato. Gene. 450: 76-84.

Furini A, Wunder J (2004). Analysis of eggplant (*Solanum melongena*)-related germplasm: morphological and AFLP data contribute to phylogenetic interpretations and germplasm utilization. Theor. Appl. Genet. 108: 197-208.

Hu JB, Zhou XY, Li JW (2010). Development of novel EST-SSR markers for cucumber (*Cucumis sativus*) and their transferability to related species. Sci Hortic-Amsterdam, pp. 534-538.

Huang H, Lu J, Ren ZB, Hunter W, Dowd SE, Dang P (2010). Mining and validating grape (*Vitis* L.) ESTs to develop EST-SSR markers for genotyping and mapping. Mol. Breeding. 10.1007/s11032-010-9477-2

Huang SW, Zhang BX, Dan M, Cardle L, Yang GM, Guo JZ (2001). Development of pepper SSR markers from sequence databases. Euphytica, 117: 163-167.

Huang X, Madan A (1999). CAP3: A DNA sequence assembly program. Genome Res. 9: 868-877.

Kantety RV, Rota LM, Matthews DE, Sorrells ME (2002). Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. Plant Mol. Biol. 48: 501-510.

Lee JM, Nahm SH, Kim YM, Kim BD (2004). Characterization and molecular genetic mapping of microsatellite loci in pepper. Theor. Ap. Genet. 108: 619-627.

Li HZ, Chen HY, Zhuang TM, Chen J (2010). Analysis of genetic

variation in eggplant and related Solanum species using sequence-related amplified polymorphism markers. Sci. Hortic-Amsterdam, 125: 19-24.
Liang XQ, Chen XP, Hong YB, Liu HY, Zhou GY, Li SX, Guo BZ (2009). Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and Arachis wild species. BMC Plant Biol. 9: 35-43.
Liu YL, Li YH, Zhou GA, Uzokwe N, Chang RZ, Chen SY, Qiu LJ (2010). Development of soybean EST-SSR markers and their use to assess genetic diversity in the subgenus Soja. Agric. Sci. China, 9: 1423-1429.
Lu YD, Cai CP, Wang L, Lin SY, Zhao LA, Tian LL, Lu JH, Zhang TZ, Guo WZ (2010). Mining, characterization, and exploitation of EST-derived microsatellites in *Gossypium barbadense*. Chinese Sci. Bull. 55: 1889-1893.
Mulato BM, Moller M, Zucchi MI, Quecini V, Pinheiro JB (2010). Genetic diversity in soybean germplasm identified by SSR and EST-SSR markers. Pesqui Agropecu Bras. 45: 76-283.
Nagy I, Stagel A, Sasvari Z, Roder M, Ganal M (2007). Development, characterization, and transferability to other Solanaceae of microsatellite markers in pepper (*Capsicum annuum* L.). Genome, 50: 668-688.
Nisha P, Nazar PA, Jayamurthy P (2009). A comparative study on antioxidant activities of different varieties of Solanum melongena. Food Chem. Toxicol. 47: 2640-2644.
Nunome T, Ishiguro K, Yoshida T, Hirai M (2001). Mapping of fruit shape and color development traits in eggplant (*Solanum melongena* L.) based on RAPD and AFLP markers. Breed. Sci. 51: 19-26.
Nunome T, Suwabe K, Iketani H, Hirai M (2003). Identification and characterization of microsatellites in eggplant. Plant Breed. 122: 256-262.
Nunome T, Negoro S, Kono I, Kanamori H, Miyatake K, Yamaguchi H, Ohyama A, Fukuoka H (2009). Development of SSR markers derived from SSR-enriched genomic library of eggplant (*Solanum melongena* L.). Theor. Appl. Genet. 119: 143-1153.
Oliveira KM, Pinto LR, Marconi TG, Margarido GRA, Pastina MM, Teixeira LHM, Figueira AV, Ulian EC, Garcia AAF, Souza AP (2007). Functional integrated genetic linkage map based on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross. Mol. Breed. 20:189-208.
Paterson AH, Brubaker CL, Wendel JF (1993). A rapid method for extraction of cotton ( *Gossypium* Spp. ) genomic DNA suitable for RFLP or PCR analysis. Plant Mol. Biol. Rep. 11: 122-127.
Portis E, Nagy I, Sasvari Z, Stagel A, Barchi L, Lanteri S (2007). The design of Capsicum spp. SSR assays via analysis of in silico DNA sequence, and their potential utility for genetic mapping. Plant Sci. 172: 640-648.
Powell W, Machray GC, Provan J (1996). Polymorphism revealed by simple sequence repeats. Trends Plant Sci. 1: 215-222.
Rohlf FJ (2000). NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, Version 2.1. Exeter Software, Setauket, NY.
Rozen S, Skaletsky H (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132: 365-386.
Song GQ, Li MJ, Xiao H, Wang XJ, Tang RH, Xia H, Zhao CZ, Bi YP (2010). EST sequencing and SSR marker development from cultivated peanut (*Arachis hypogaea* L.). Electron J. Biotechnol. 13: 1-9.
Squirrell J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M, Powell W (2003). How much effort is required to isolate nuclear microsatellites from plants? Mol. Ecol. 12: 1339-1348.
Stagel A, Portis E, Toppino L, Rotino GL, Lanteri S (2008). Gene-based microsatellite development for mapping and phylogeny studies in eggplant. BMC Genomics, 9: 357-370.
Sudheesh S, Sandhya C, Koshy AS, Vijayalakshmi, NR (1999). Antioxidant activity of flavonoids from *Solanum melongena*. Phytother. Res. 13: 393-396.
Sunseri F, Sciancalepore A, Martelli G, Acciarri N, Rotino GL, Valentino D, Tamietti G (2003). Development of RAPD-AFLP map of eggplant and improvement of tolerance to Verticillium wilt. Acta-Horticulturae, 625: 107-110.
Thiel T, Michalek W, Varshney RK, Graner A (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.
Tumbilen Y, Frary A, Daunay MC, Doganlar S (2011). Application of EST-SSRs to examine genetic diversity in eggplant and its close relatives. Turk J. Biol. 35:125-136.
Varshney RK, Sigmund R, Borner A, Korzun V, Stein N, Sorrells ME, Langridge P, Graner A (2005). Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. Plant Sci. 168: 195-202.
Wu FN, Eannetta NT, Xu YM, Tanksley SD (2009). A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. Theor. Appl. Genet. 118: 927-935.
Zhang J, Guo W, Zhang T (2002). Molecular linkage map of allotetraploid cotton (*Gossypium hirsutum* L. x *Gossypium barbadense* L.) with a haploid population. Theor. Appl. Genet. 105: 1166-1174.
Stagel A, Portis E, Toppino L, Rotino GL, Lanteri S (2008). Gene-based microsatellite development for mapping and phylogeny studies in eggplant. BMC Genomics, 9: 357-370.
Sudheesh S, Sandhya C, Koshy AS, Vijayalakshmi, NR (1999). Antioxidant activity of flavonoids from *Solanum melongena*. Phytother. Res. 13:393-396.
Sunseri F, Sciancalepore A, Martelli G, Acciarri N, Rotino GL, Valentino D, Tamietti G (2003). Development of RAPD-AFLP map of eggplant and improvement of tolerance to Verticillium wilt. Acta-Horticulturae, 625:107-110.
Thiel T, Michalek W, Varshney RK, Graner A (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.
Tumbilen Y, Frary A, Daunay MC, Doganlar S (2011). Application of EST-SSRs to examine genetic diversity in eggplant and its close relatives. Turk J. Biol. 35:125-136.
Varshney RK, Sigmund R, Borner A, Korzun V, Stein N, Sorrells ME, Langridge P, Graner A (2005). Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. Plant Sci. 168:195-202.
Wu FN, Eannetta NT, Xu YM, Tanksley SD (2009). A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. Theor. Appl. Genet. 118: 927-935.
Zhang J, Guo W, Zhang T (2002). Molecular linkage map of allotetraploid cotton ( *Gossypium hirsutum* L. x *Gossypium barbadense* L.) with a haploid population. Theor. Appl. Genet. 105: 1166-1174.