*Pak. J. Bot.*, 44(2): 485-492, 2012.

# GENETIC DIVERSITY OF ARGENTINA TOMATO VARIETIES REVEALED BY MORPHOLOGICAL TRAITS, SIMPLE SEQUENCE REPEAT, AND SINGLE NUCLEOTIDE POLYMORPHISM MARKERS

**XIAORONG HU[1], HUI WANG[2], JIA CHEN[2] AND WENCAI YANG[2*]**

[1]*The National Key Facilities for Crop Genetic Resources and Improvement / Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China*
[2]*Department of Vegetable Science, College of Agronomy and Biotechnology, China Agricultural University, No. 2 Yuanmingyuan Xilu, Haidian District, Beijing 100193, China*
**Corresponding author: Tel: 86-10-62734136, Fax: 86-10-62733404, E-mail: yangwencai@cau.edu.cn*

**Abstract**

Twenty-six morphological traits as well as 47 single nucleotide polymorphism and simple sequence repeat markers were used to investigate genetic variation in 67 tomato (*Solanum lycopersicum* L.) varieties collected from Argentina between 1932 and 1974. Approximately 65.0% of the morphological traits and 55.3% of the molecular markers showed polymorphisms in the 67 varieties. Average taxonomic distance between any two varieties ranged from 0.6643 to 1.1776, while Nei's genetic distance varied from 0 to 0.2022. Cluster analysis indicated that 67 varieties could be grouped into three clusters at both morphological and molecular levels. The varieties collected before 1960 had larger genetic variation than those collected after 1960.

## Introduction

Knowledge of genetic variation has important implications for the conservation of genetic resources and breeding programs. The relative genetic diversity can be estimated using various approaches including pedigree information, morphological and molecular markers. In tomato (*Solanum lycopersicum* L.), one cultivated species and 12 wild relatives have been reported (Rick *et al*., 1990; Peralta *et al*., 2006). Large morphological variations have been observed and great genetic diversity has been revealed by molecular markers in wild species (McClean & Hanson, 1986; Rick *et al*., 1990; Miller & Tanksley, 1990; Egashira *et al*., 2000; Zhu *et al*., 2004). These variations provide great potential for crop improvement. However, genetic variation in modern cultivars or hybrids is limited (Sharma & Verma, 2001; Archak *et al*., 2002; Wang *et al*., 2006; Benor *et al*., 2008; Yi *et al*., 2008; Chen *et al*., 2009). It is estimated that cultivated tomato genome contains less than 5% of the genetic variation of the wild relatives (Miller & Tanksley, 1990).

Landraces and local varieties contain much more genetic diversity than modern cultivars or hybrids (Williams & St. Clair, 1993; Zeven, 1998; Zhu *et al*., 2004; Garcia-Martinez *et al*., 2005; Terzopoulos & Bebeli, 2008; Yi *et al*., 2008; Terzopoulos *et al*., 2009). Therefore they are among the most important sources of genetic variation for breeders. To date, a large number of landraces and local varieties have been collected (Robertson & Labate, 2007), which provides a potential for increasing the genetic variation in modern breeding. However, very few of them have been systematically evaluated. Single nucleotide polymorphism (SNP) is a newly developed marker system that shows nucleotide variations in the DNA sequences. SNPs are widely distributed and constitute the most abundant molecular markers in the genome. In cultivated tomato, more than 600 SNPs have been discovered (Yang *et al*., 2004; Labate & Baldo, 2005; Van Deynze *et al*., 2007; Wang *et al*., 2010). This provides a potential to characterize allelic variation in the whole genome of tomato.

In the present study, we investigated the variation of 67 Argentina tomato varieties collected during 1932 to 1974 using both molecular markers and phenotypic data. The results obtained here will provide some useful information for the conservation and the use of these Argentina varieties in breeding.

## Materials and Methods

A set of 67 tomato varieties from north part of Argentina collected from 1932 to 1974 were used for morphological and genetic variation analysis (Table 1). Most of them were originally collected from local markets (http://www.ars-grin.gov/npgs/index.html). Seeds of all varieties were kindly provided by Northeast Regional PI Station at Geneva, New York, USA. Twenty plants of each variety were grown in a protected field for morphological traits observation and DNA isolation.

Twenty six morphological traits were scored from 10 randomly selected plants for each variety using the description of Li & Du (2006). These traits included leaf color, leaf vein color, leaf shape, leaf state, leaf type, stem and leaf hairness, leaf division, corolla color, abscission layer, immature fruit color, color of mature fruit, fruit shoulder ribbing, pubescence, fruit apex, fruit shoulder, fruit shoulder shape, size of green shoulder, color of fruit shoulder, fruit shape, inflorescence type, fascicle, style length, style shape, style hairiness, growth habit, plant posture. A similarity matrix was generated using the Interval Data with DIST (average taxonomic distance) in the program of NTSYSpc 2.11a (Rohlf, 1998). UPGMA cluster analysis was performed to develop a dendrogram.

**Table 1. Sixty-seven tomato varieties from Argentina used in this study**

| Accession | Genotype | Locality | Year collected |
|---|---|---|---|
| PI119776 | Liso Colorado Argentino | unknown | 1936 |
| PI119777 | Grueso Liso Chemin | unknown | 1936 |
| PI119778 | Colorado Grueso | unknown | 1936 |
| PI128214 | 339 | Jujuy | 1938 |
| PI128272 | 327 | Jujuy | 1938 |
| PI128273 | 328 | Jujuy | 1938 |
| PI128274 | 335 | Salta | 1938 |
| PI128275 | 336 | Salta | 1938 |
| PI128276 | 337 | Salta | 1938 |
| PI128277 | 340 | Jujuy | 1938 |
| PI128278 | 341 | Jujuy | 1938 |
| PI128279 | 342 | Jujuy | 1938 |
| PI128280 | 343 | Jujuy | 1938 |
| PI128281 | 344 | Jujuy | 1938 |
| PI128282 | 345 | Jujuy | 1938 |
| PI128283 | 346 | Jujuy | 1938 |
| PI128285 | 349 | Jujuy | 1938 |
| PI128286 | 350 | San Juan/Tucuman | 1938 |
| PI128287 | 351 | San Juan/Tucuman | 1938 |
| PI128288 | 352 | Chaco/Tucuman | 1938 |
| PI128291 | 384 | Cordoba | 1938 |
| PI128292 | 385 | Cordoba | 1938 |
| PI128293 | 386 | Cordoba | 1938 |
| PI128294 | 383 | Cordoba | 1938 |
| PI128445 | 388 | Buenos Aires | 1938 |
| PI128990 | San Marzano | Buenos Aires | 1938 |
| PI129132 | 759 | Buenos Aires | 1938 |
| PI129133 | 760 | Buenos Aires | 1938 |
| PI129134 | 761 | Buenos Aires | 1938 |
| PI129135 | 762 | Buenos Aires | 1938 |
| PI129136 | 763 | Buenos Aires | 1938 |
| PI129137 | 764 | Buenos Aires | 1938 |
| PI129138 | 765 | La Plata | 1938 |
| PI129139 | 766 | La Plata | 1938 |
| PI129140 | 767 | La Plata | 1938 |
| PI129687 | Campana | Hudson | 1938 |
| PI129688 | Ciro | Villa Elisa | 1938 |
| PI129689 | Las Talas | Las Talas, Rio Santiago | 1938 |
| PI129690 | Palo Blanco | Palo Blanco, Rio Santiago | 1938 |
| PI129691 | Sino | Florencio Vasela | 1938 |
| PI129692 | Vasela | Florencio Vasela | 1938 |
| PI131877 | Campana | Hudson | 1939 |
| PI131878 | Los Talas | Rio Santiago | 1939 |
| PI131879 | Palo Blanco | Rio Santiago | 1939 |
| PI131880 | Rey Humberto | Buenos Aires | 1939 |
| PI131881 | San Marzano | Buenos Aires | 1939 |
| PI131882 | Varela | Florencio Varela | 1939 |
| PI162679 | Genova | Buenos Aires | 1948 |
| PI190858 | Rey de los Tempranos | unknown | 1950 |
| PI194561 | Morman 50 Day | Mendoza | 1951 |
| PI199016 | Juan Peron | Mendoza | 1952 |
| PI255955 | Piovano | Mendoza | 1959 |
| PI260395 | Magnit Potente | Instituto Nacional de Technologia Agropecuaria | 1959 |
| PI306211 | Blair Forcing | Instituto Nacional Technologia Agropecuaria | 1965 |
| PI306212 | El Naro | Instituto Nacional Technologia Agropecuaria | 1965 |
| PI306213 | Firesteel | Instituto Nacional Technologia Agropecuaria | 1965 |
| PI306214 | Grande Perfeicao | Instituto Nacional Technologia Agropecuaria | 1965 |
| PI306215 | Magnif Potente | Instituto Nacional Technologia Agropecuaria | 1965 |
| PI321040 | 4624b | unknown | 1967 |
| PI321041 | G 11704s | unknown | 1967 |
| PI386240 | Platense | Vigliola, M., Facultad de Agronomica y Veterinaria | 1974 |
| PI386241 | Roma Sel. La Consulta | Vigliola, M., Facultad de Agronomica y Veterinaria | 1974 |
| PI386242 | Ronita La Consulta | Vigliola, M., Facultad de Agronomica y Veterinaria | 1974 |
| PI386243 | Rossol Sel. La Consulta | Vigliola, M., Facultad de Agronomica y Veterinaria | 1974 |
| PI636277 | Coure Di Bue | unknown | 1963 |
| PI636296 | Argentine | unknown | 1966 |
| PI97538 | Cherry | Tucuman | 1932 |

Genomic DNA was isolated from young leaves collected from eight plants of each variety using the modified CTAB isolation method (Kabelka *et al*., 2002). Thirty-seven SNP and 10 simple sequence repeat (SSR) markers were used to genotype all varieties (Table 2). These markers were randomly selected from 12 chromosomes. Genotyping using SNP markers was conducted according to the method described in Yang *et al*., (2004). Restriction enzymes used for digesting PCR products of SNP markers can be found from the SOL Genomics Network (http://www.sgn.cornell.edu/) or Yang *et al*., (2004). SSR analysis was performed as the description in Chen *et al*., (2009). The presence or absence of each single fragment was coded by 1 or 0, respectively, and scored for a binary data matrix. Polymorphism information content (PIC) was calculated using the formula of PIC=1-$\sum pi^2$ (Weir, 1990), where *pi* is the frequency of $i^{th}$ allele for each marker locus. Nei's genetic distance (Nei, 1972) was calculated for each pair of varieties using the program in the software package NTSYSpc 2.11a. UPGMA cluster analysis was performed to develop a dendrogram.

**Table 2. Marker information, number of alleles, and polymorphism information content (PIC) for each marker in 67 Argentina tomato lines**

| Marker[a] | Chromosome | Forward primer (5'-3') | Reverse primer (5'-3') | No. of alleles | PIC |
|---|---|---|---|---|---|
| C2_At2g34860 | 1 | AGTTGAATATGAAGAAGAGGGTAGGG | ACAGCCAGGACTTTCATTTCCATC | 1 | 0.000 |
| C2_At2g38730 | 1 | AGCGGACCAAACACTAATGGATG | AGCCACATTCTCAATCTTCCTGAC | 1 | 0.000 |
| C2_At5g27620 | 1 | ATCTACAATGGTCCGTGATGGAAC | TTCCTCTGCCTTGCAAGCTGC | 2 | 0.455 |
| C2_At5g64350 | 1 | AGATCGGCCAAGGCAAAGTTATC | TGCATGCCCAGTACTCCTTCATCC | 2 | 0.430 |
| CosOH44 | 2 | TGCTTCTTGCACCACAAACT | TGTTGTCATGGTCCCTTTGA | 2 | 0.484 |
| SSR66 | 2 | TGCAACAACTGGATAGGTCG | TGGATGAAACGGATGTTGAA | 2 | 0.029 |
| SSR96 | 2 | GGGTTATCAATGATGCAATGG | CCTTTATGTCAGCCGGTGTT | 3 | 0.524 |
| SSR5 | 2 | TGGCCGGCTTCTAGAAATAA | TGAAATCACCCGTGACCTTT | 1 | 0.000 |
| C2_At5g67370 | 2 | TGAAACCAGTCATTAAAATGCTGAAG | AGTACTGTCCACCGGCCAATGC | 1 | 0.000 |
| C2_At1g67730 | 2 | TGGGATTGATGTGCAATGCCAGG | AGGGCAGCCCGAGCATAACC | 2 | 0.487 |
| C2_At5g23940 | 3 | TAGGCCTCTACTCGCCGTACAGC | TTAGTTCTTTCGAGGAAAGGTGGG | 1 | 0.000 |
| SSR111 | 3 | TTCTTCCCTTCCATCAGTTCT | TTTGCTGCTATACTGCTGACA | 2 | 0.284 |
| C2_At5g60160 | 3 | ACACAATGCTAATCAACGTTATGC | TCATCCACCGCGCACATTTC | 1 | 0.000 |
| SSR601 | 3 | TCTGCATCTGGTGAAGCAAG | CTGGATTGCCTGGTTGATTT | 3 | -0.111 |
| SSR43 | 4 | CTCCAAATTGGGCAATAACA | TTAGGAAGTTGCATTAGGCCA | 3 | 0.458 |
| C2_At3g54770 | 4 | ACCGGAAGATCCAAAGGCTATGG | AGGGACGCGAGATTACAGTTGGC | 2 | 0.029 |
| C2_At4g09010 | 4 | TAAGGGGCTTGATGCTGCTTTG | TAAAGGTCGATTTGACTGCACTTTG | 1 | 0.000 |
| SSR146 | 4 | TATGGCCATGGCTGAACC | CGAACGCCACCACTATACCT | 1 | 0.000 |
| C2_At5g42950 | 4 | AGCAATGGATTTCAGAGAATGGTGTG | ACATTTTTGGCACTTGCACCAGTGAC | 1 | 0.000 |
| C2_At1g60440 | 5 | TGCCCGGTCCCTCTTAAGGATG | TCCGCTTGAGCCCAAAACGAAG | 1 | 0.000 |
| C2_At5g14320 | 5 | TTCTCTTTCCCTTATCTGCAACAC | TTTGGAACTTCCACTCCTCCCAC | 2 | 0.481 |
| C2_At1g14300 | 5 | AGGCGCTAGAGGCTATTTATTTGC | TCACTGACCAAAATGCTCTTCTGCC | 2 | 0.044 |
| C2_At2g03510 | 5 | TGATACCCTGCTGAATTATGGGGTC | TGGTGCGCTCCTGTTCCATGTTCTC | 1 | 0.000 |
| TOM152 | 5 | ATTCAAGGAACTTTTAGCTCC | TGCATTAAGGTTCATAAATGA | 3 | 0.528 |
| C2_At5g26360 | 5 | TAGTTCCCGGTGGTGGTGCAAC | TCAAAAGCAATTGCAGCAGCTTC | 1 | 0.000 |
| SSR47 | 6 | TCCTCAAGAAATGAAGCTCTGA | CCTTGGAGATAACAACCACAA | 3 | 0.484 |
| C2_At3g10920 | 6 | TGGCTTGGTGTGGACAAAGAGC | TGCAAGTAGTATGCGTGTTCCC | 2 | 0.493 |
| C2_At5g05690 | 6 | ATGACCGTGTTTCAAAATACGGC | ATGGATCAAAACTCATCAGCTGCTTC | 1 | 0.000 |
| C2_At1g22850 | 6 | ATCATTGTTTCCATTGGTGGAACG | TGCAAGAAATTTCTTGTTCCCTTC | 1 | 0.000 |
| C2_At4g24820 | 7 | TGACTGAGAAAAAAACTGTTGCAGTTG | AGATCTGCTGCTTTCTTGAAGTTACG | 2 | 0.175 |
| C2_At3g14770 | 7 | TCAACTGAACAGTTCTCAGGGTTGCC | AACATTGATATCAAGGAAGCACAACTC | 2 | 0.425 |
| C2_At4g26750 | 7 | AAGGATAACGAACCAGCAAAGC | TTTGAGGAATCCTCAATCCTCG | 1 | 0.000 |
| C2_At5g56130 | 7 | ACATATAGCTGTTGGGAACAGGG | TAGGTTTAAACTTGCGAACATCC | 2 | 0.138 |
| LEOH343 | 8 | CAAATGGGTTTGGCTGAAAA | CGCAAACTGATTTGAACAGC | 2 | 0.500 |
| C2_At4g22670 | 8 | TGGGAGGCAGCTGCTAAGGATCTTC | TCTTTCTATCTTCTCGTTCTTTGCG | 2 | 0.058 |
| C2_At1g63770 | 8 | AGGTGGAAACGTTATGATGAAAC | ATGCGATTTCAAACACATTCTCTG | 2 | 0.386 |
| LEOH8 | 9 | CCACTGATCAATGTGGTGGA | CAACCACAAATGGCTCCTAAA | 1 | 0.000 |
| C2_At2g47590 | 9 | ACGAGCGTCGATTGTTTGGTTCC | ACTAGGATTGAGCCCCAAATCAACC | 1 | 0.000 |
| C2_At5g06430 | 10 | ATTGTTATGGCTGATGCAGAGAATG | ACGAAGCAAGGAACATACTTTATGTC | 1 | 0.000 |
| C2_At4g30220 | 10 | ACTGGAAAGCCTGTGATGGTAAAGC | TGCAAGTTCATATATGAATCCACAGAGAC | 1 | 0.000 |
| C2_At3g52220 | 11 | TGCTCGGGTGGATGGTCTTGG | TGATGGTGAACTTGGTTCTTCCC | 1 | 0.000 |
| C2_At4g22260 | 11 | TCCTCTAACGGTCTAGAGAAATGGG | AGGAACTCTTGCAATTGTTTCCAGAAC | 2 | 0.500 |
| C2_At5g59960 | 11 | TCCGATACTCATCAGCTCTTGTTC | ACGCCTTGTGTTTGTTTGGATGTC | 1 | 0.000 |
| C2_At4g03280 | 12 | TATGAATTTGCTTTTATTGGGTGC | ATCTTTGGCAGGGGTACCACCAC | 2 | 0.500 |
| SSR20 | 12 | GAGGACGACAACAACAACGA | GACATGCCACTTAGATCCACAA | 2 | 0.072 |
| C2_At4g16580 | 12 | TGTTACCTGCCTCATCCTGATAAAG | ATTTTGAAGACCTCTCCAGAACTTGG | 2 | 0.430 |
| LEOH301 | 12 | TGCTGTTTTGTTTGGCTCAC | TGTTCATATCTTTGATGGCATGT | 3 | -0.196 |

[a]Markers started with SSR are from Frary *et al*. (2005), markers started with Tom are from Suliman-Pollatschek *et al*. (2002), markers started with C2 are from http://www.sgn.cornell.edu, and markers started with LEOH are from Chen *et al*. (2009) or this study.

## Results

Seventeen of 26 traits had morphological variation in the 67 Argentina tomato varieties. Numbers of observed types for each trait ranged from 2 to 7. Nine traits (34.6%) had more than 2 types, of which fruit shape had the largest variation with seven types (flat, oblate, round, high round, prelate round, ovate, and pear-shaped). No obvious differences for 9 traits including leaf vein color, leaf shape, leaf state, stem and leaf hairiness, corolla color, abscission layer, fruit shoulder, inflorescence type, and plant posture were observed.

The average coefficient of taxonomic distance between any 2 varieties ranged from 0.6643 to 1.1776 with a mean of 0.7955. Most taxonomic distances were between 0.5001 and 1.1000 (Fig. 1). The largest distance was 1.6053 between PI131882 and PI128280, while the least was 0.1961 (PI129688 vs PI129689,

PI129133 vs PI128277, PI128286 vs PI128287). Lines collected from different regions at different years were randomly clustered into different groups. The 67 varieties formed 3 clusters at the average taxonomic distance of 0.88 (Fig. 2). Two small clusters I and II contained 3 and 8 varieties, respectively, while 83.6% varieties formed into a large cluster III. This group could be further divided into 2 sub-groups (IIIa and IIIb) at the average taxonomic distance of 0.788 (Fig. 2).

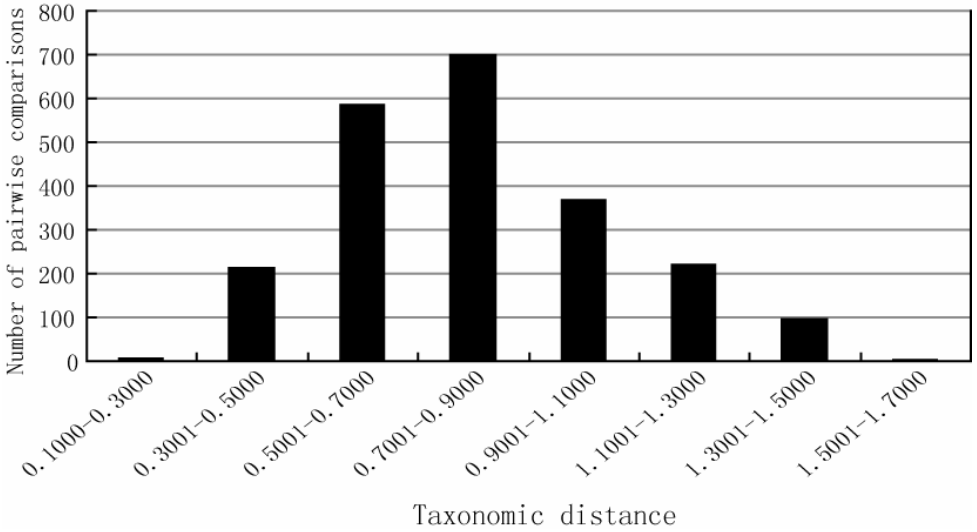


Fig. 1. Distribution of taxonomic distance values obtained from pair wise comparisons of 67 Argentina tomato varieties using 26 morphological trait data.

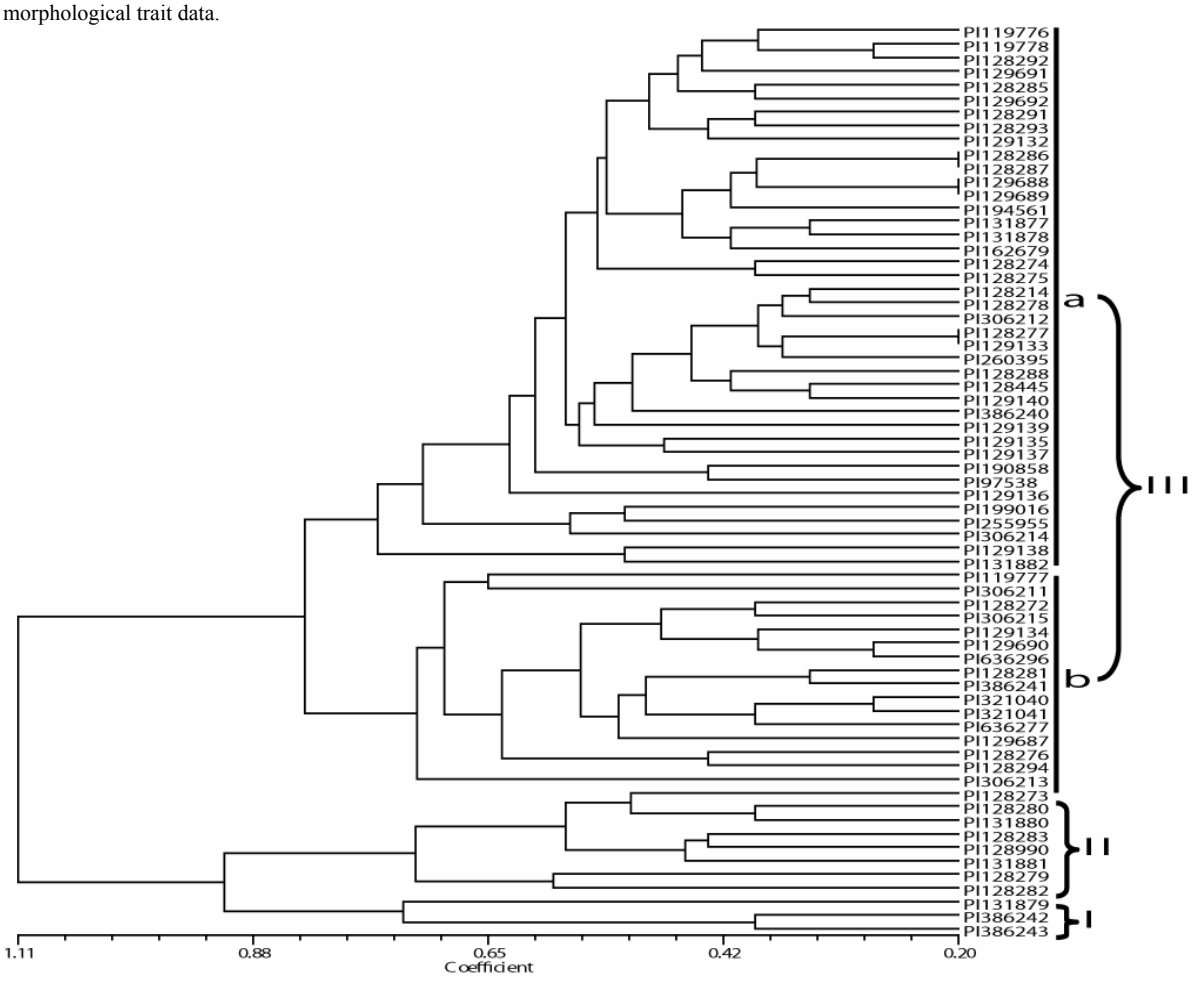


Fig. 2. Dendrogram of 67 Argentina tomato varieties, based on 26 morphological trait measurements, and generated from average taxonomic distance matrix by UPGMA in NTSYSpc 2.11a.

Among the 47 markers used in the study, 26 were polymorphic in the 67 Argentina tomato varieties (Table 2). Twenty polymorphic markers had 2 alleles and the remaining 6 markers had 3 alleles each. Only 3 markers generated alleles unique to 1 or 2 varieties. PIC for most polymorphic markers ranged from 0.029 to 0.528. However, 2 markers SSR601 and LEOH301 had negative PIC values because they detected heterozygous alleles in 67 varieties. In addition, a trend of alleles reduction in varieties collected after 1960 was observed. Among the 26 polymorphic markers, 7 had one allele lost and 6 had alleles fixed in the varieties collected after 1960 (Table 3). Allelic variation was reduced by one-third in varieties collected after 1960 compared with varieties collected before 1960.

**Table 3. Number of alleles from polymorphic markers in varieties collected before and after 1960**

| Marker | Number of alleles | | Marker | Number of alleles | |
|---|---|---|---|---|---|
| | Before 1960 | After 1960 | | Before 1960 | After 1960 |
| SSR20[a] | 2 | 1 | C2_At3g54770[a] | 2 | 1 |
| SSR111 | 2 | 2 | C2_At4g09010 | 2 | 2 |
| SSR43 | 3 | 3 | C2_At5g14320[b] | 2 | 2 |
| SSR47[ab] | 3 | 2 | C2_At1g14300[a] | 2 | 1 |
| SSR601 | 3 | 3 | C2_At3g10920 | 2 | 2 |
| LEOH301[a] | 3 | 2 | C2_At4g24820 | 2 | 2 |
| SSR66 | 1 | 2 | C2_At3g14770[b] | 2 | 2 |
| SSR96[a] | 3 | 2 | C2_At5g56130 | 2 | 2 |
| TOM152 | 3 | 3 | LEOH343 | 2 | 2 |
| CosOH44 | 2 | 2 | C2_At4g22670[a] | 2 | 1 |
| C2_At5g27620 | 2 | 2 | C2_At1g63770[b] | 2 | 2 |
| C2_At5g64350 | 2 | 2 | C2_At4g03280[b] | 2 | 2 |
| C2_At1g67730 | 2 | 2 | C2_At4g16580[b] | 2 | 2 |

[a] The marker had allele lost in varieties collected after 1960. [b] Alleles of the marker were fixed in the varieties collected after 1960.

The average Nei's genetic distance was 0.0899 with a range from 0.0679 (PI128273) to 0.2022 (PI128293) for each variety. The largest genetic distance (0.2022) was between varieties PI131878 and PI128281, while varieties PI255995 and PI321040 had the least genetic distance of 0. Sixty three percent of genetic distance between any 2 varieties was between 0.061 and 0.120, 17.8% were below 0.061, and only 4.7% were larger than 0.150 (Fig. 3). Genetic distances tended to decrease in varieties from 1930s to 1970s and decreased by approximately 14.0% for varieties collected in 1960s and 1970s.

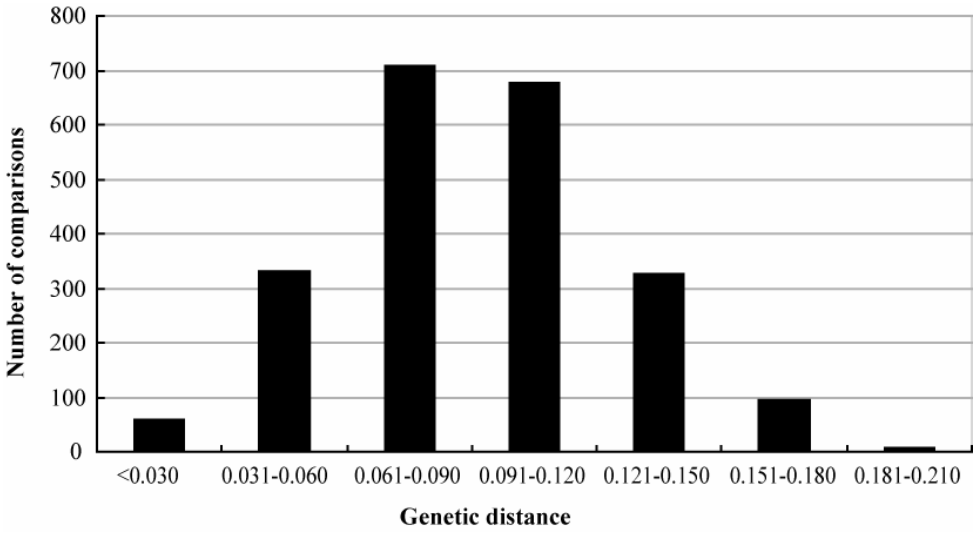


Fig. 3. Distribution of genetic distance values obtained from pair wise comparisons of 67 Argentina tomato varieties using SSR and SNP marker data.

Dendrograms were constructed from the pair wise distance matrices based on Nei's distance. Sixty seven varieties could also be grouped into 3 clusters at the genetic distance of 0.095, which was close to the maximum coefficient 0.11 (Fig. 4). Cluster I included 19 varieties from Jujuy, Salta, Cordoba, Tucuman, and unknown regions. Cluster II was the largest group containing 42 varieties from 9 regions. Almost all varieties from Buenos Aires region were found in this cluster. Cluster III was a small group with 6 varieties, of which 3 were from Jujuy region, 1 from Cordoba, 1 from Rio Santiago, and 1 from Buenos Aires. These results suggested that there was no relationship between the clustering pattern and the geographic origin of the material.

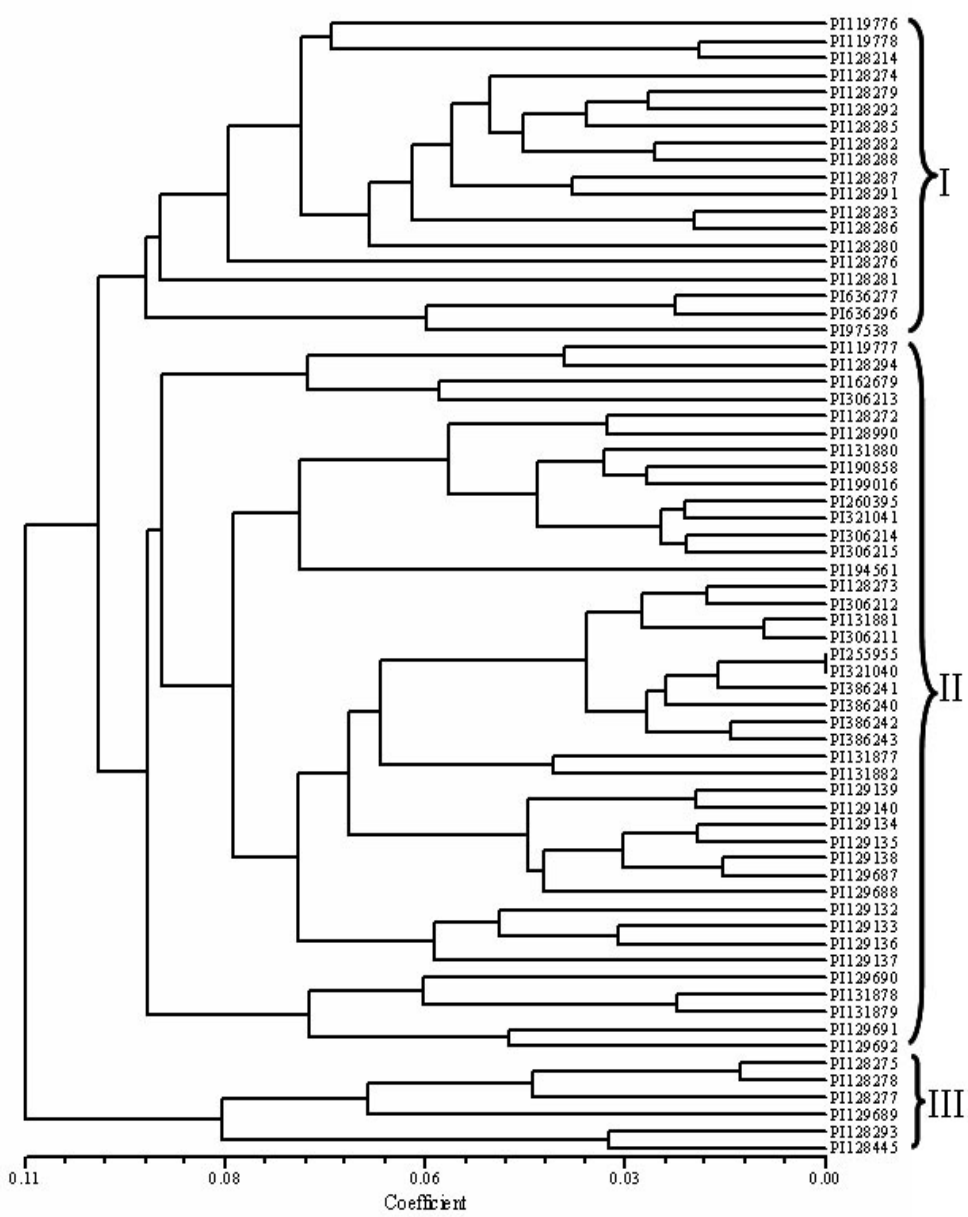


Fig. 4. Dendrogram of 67 Argentina tomato varieties, based on 47 SSR and SNP marker data, and generated from Nei's genetic distance matrix by UPGMA in NTSYSpc 2.11a.

## Discussion

The lack of genetic markers that detect differences between elite tomato breeding lines has prevented a detailed study of most traits of economic importance within genetic backgrounds that are relevant to plant breeders. Discovery of SNP markers in cultivated tomato (Yang *et al*., 2004, 2005; Labate & Baldo, 2005; Van Deynze *et al*., 2007; Wang *et al*., 2010) has provided a chance to characterize genome-wide allelic variation. In this study, 48.6% of the 37 randomly selected SNP markers showed polymorphisms in 67 varieties. The frequency of SNP markers detecting polymorphisms in cultivated tomatoes was higher than any other markers reported to date supporting that SNP markers was useful to characterize the genome-wide allelic variation in tomato.

Genetic diversity can be estimated using both morphological and molecular markers. Morphological trait measurements can provide a simple technique of quantifying genetic variation while simultaneously assessing genotype performance under relevant growing environments (Fufa *et al*., 2005; Shuaib *et al*., 2007). However, assessment of morphological traits is time-consuming and phenotypic characters are generally influenced by environments and plant developmental stages (Tatineni *et al*., 1996; Van Beuningen & Busch, 1997; Garcia, 1998). On the contrary, molecular markers are independent of environmental conditions and show higher levels of polymorphism. However, high morphological variability is not always reflected at the molecular level (Wang *et al*., 2006). In this study, none

of the markers used here was significantly associated with the 26 traits (data not shown). That might interpret why the clusters formed using morphological data were different from the one formed using SSR and SNP data.

It has been suggested that domestication and inbreeding dramatically reduced the genetic variation (Bai & Lindhout, 2007; Yi *et al*., 2008). Using RAPD markers to analyze 27 cultivars released in India, Archak *et al*., (2002) found that old introductions and locally developed cultivars of the 1970s exhibited significantly greater genetic variation than the ones released during the 1990s. This suggests that modern cultivars have less genetic variation than old ones. Same trend was observed in this study. Varieties collected in 1960s and 1970s had less genetic variation than varieties collected before 1960. Most varieties collected in 1960s and 1970s had determinate growth habit, no fruit shoulder ribbing, and low concentric cracking, while most varieties collected before 1960s had indeterminate growth habit, various fruit shoulder ribbing (from none to prominent), and diverse concentric cracking (from none to severe) (http://www.ars-grin.gov/npgs/index.html). Selection for these traits might be part of the reasons causing reduction of genetic variation in 1960s and 1970s.

## Acknowledgements

The authors thank Northeast Regional PI Station at Geneva, New York, USA, for providing tomato seeds. The work was partly supported by National Natural Science Foundation of China (30671425) and the Program for New Century Excellent Talents in University (NCET-08-0531).

### References

Archak, S., J.L. Karihaloo and A. Jain. 2002. RAPD markers reveal narrowing genetic base of Indian tomato cultivars. *Curr. Sci.*, 82:1139-1143.

Bai, Y.L. and P. Lindhout. 2007. Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? *Ann. Bot.*, 100:1085-1094.

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

Chen, J., H. Wang, H.L. Shen, M. Chai, J.S. Li, M.F. Qi and W.C Yang. 2009. Genetic variation in tomato populations from four breeding programs revealed by single nucleotide polymorphism and simple sequence repeat markers. *Sci. Hortic.*, 122: 6-16.

Egashira, H., H. Ishihara, T. Takashina and S. Imanishi. 2000. Genetic diversity of the 'peruvianum-complex' (*Lycopersicon peruvianum* (L.) Mill. and *L. chilense* Dun.) revealed by RAPD analysis. *Euphytica*, 116: 23-31.

Frary, A., Y. Xu, J. Liu, S. Mitchell, E. Tedeschi and S. Tanksley. 2005. 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.

Fufa, H., P.S. Baenziger, I. Beecher, V. Dweikat, R.A. Graybosch and K.M. Eskridge. 2005. Comparison of phenotypic and molecular marker-based classifications of hard red winter wheat cultivars. *Euphytica*, 145: 133-146.

Garcia, E., M. Jamilena, J.I. Alvarez, T. Arnedo, J.L. Oliver and R. Lozano. 1998. Genetic relationships among melon breeding lines revealed by RAPD markers and agronomic traits, *Theor. Appl. Genet.*, 96: 878-885.

Garcia-Martinez, S., L. Andreani, M. Garcia-Gusano, F. Geuna and J.J. Ruiz. 2005. Evolution of amplified length polymorphism and simple sequence repeats for tomato germplasm fingerprinting: utility for grouping closely related traditional cultivars. *Genome*, 49: 648-656.

Kabelka, E., B. Franchino and D.M. Francis. 2002. Two loci from *Lycopersicon hirsutum* LA407 confer resistance to strains of *Clavibacter michiganensis* subsp. *michiganensis*. *Phytopathology*, 92: 504-510.

Labate, J.A. and A.M. Baldo. 2005. Tomato SNP discovery by EST mining and resequencing. *Mol. Breed.*, 16: 343-349.

Li, X.X. and Y.C. Du.2006. *Descriptors and data standard for tomato (Lycopersicon esculentum Mill.)*. China Agricultural Press, Beijing.

McClean, P.E. and M.R. Hanson. 1986. Mitochondorial DNA sequence divergence among *Lycopersicon* and related *Solanum* species. *Genetics*, 112: 649-667.

Miller, J.C. and S.D. Tanksley. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theor. Appl. Genet.*, 80: 437-448.

Nei, M. 1972. Genetic distance between populations. *Am. Nat.*, 106: 283-292.

Peralta, I.E., S. Knapp and D.M. Spooner. 2006. Nomenclature for wild and cultivated tomatoes. *Rpt. Tomato. Genet. Coop.*, 56: 6-12.

Rick, C.M., H. Laterrot and J. Philouze. 1990. A revised key for the *Lycopersicon* species. *Rpt. Tomato. Genet. Coop.*, 40: 31.

Robertson, L.D. and J.A. Labate. 2007. Genetic resources of tomato (lycopersicum esculentum Mill) and wild relatives. In: *Genetic Improvement of Solanaceous Crops Vol. 2. Tomato*. (Eds.): M.K. Razdan and A.K. Mattoo. Science Publishers, New Hampshire. pp. 25-75.

Rohlf, F.J. 1998. *NTSYS-Numerical Taxonomy and Multivariate Analysis System*. Exeter Publisher, New York.

Sharma, K.C. and S. Verma. 2001. Analysis of genetic divergence in tomato. *Ann. Agric. Res.*, 22(1): 71-73.

Shuaib, M., Z. Alam, A. Zahir, A. Waqar, A. Taufiq and K. Ikhtiar. 2007. Characterization of wheat varieties by seed storage protein electrophoresis. *Afr. J. Biotechnol.*, 6: 497-500.

Suliman-Pollatschek, S., K. Kashkush, H. Shats, J. Hillel and U. Lavi. 2002. Generation and mapping of AFLP, SSRs and SNPs in *Lycopersicon esculentum*. *Cell. Mol. Biol. Lett.*, 7: 583-597.

Tatineni, V., R.G. Cantrell and D.D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. *Crop Sci.*, 36: 186-192.

Terzopoulos, P.J. and P.J. Bebeli. 2008. DNA and morphological diversity of selected Greek tomato (Solanum lycopersicum L.) landraces. *Sci. Hortic.*, 116: 354-361.

Terzopoulos, P.J., S.A. Walters and P.J. Bebeli. 2009. Evaluation of Greek tomato landrace populations for heterogeneity of horticultural traits. *Eur. J. Hortic. Sci.*, 74: 24-29.

Van Beuningen, L.T. and R.H. Busch. 1997. Genetic diversity among North American spring wheat cultivars: III. cluster analysis based on quantitative morphological traits. *Crop Sci.*, 37: 981-988.

Van Deynze, A., K. Stoffel, C.R. Buell, A. Kozik, J. Liu, E. van der Knaap and D. Francis. 2007. Diversity in conserved genes in tomato. *BMC Genomics*, 8: 465.

Wang, R.S., Y.R. Li, L.T. Yang, L.Z. Li, F.X. Fan and W.J. Li. 2006. Analysis of genetic diversity based on SSR and

morphological markers among tomato cultivars. *J. Trop. Subtrop. Bot.*, 14(2): 120-125.

Wang, Y.Y., J. Chen, D.M. Francis, H.L. Shen, T.T. Wu and W.C. Yang. 2010. Discovery of intron polymorphisms in cultivated tomato using genomic information of tomato and Arabidopsis. *Theor. Appl. Genet.*, 121: 1199-1207.

Weir, B.S. 1990. *Genetic data analysis-methods for discrete population genetics dData.* Sinauer Associates, ***Massachusetts***.

Williams, C.E. and D. A. St. Clair. 1993. Phenetic relationships and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accessions of *Lycopersicon esculentum*. *Genome*, 36: 619-630.

Yang, W.C., S.A. Miller, J.W. Scott, J.B. Jones and D.M. Francis. 2005. Mining tomato genome sequence databases for molecular markers: application to bacterial resistance and marker assisted selection. *Acta Hortic.*, 695: 241-250.

Yang, W.C., X.D. Bai, E. Kabelka, C. Eaton, S. Kamoun, E. van der Knaap and D. Francis. 2004. Discovery of single nucleotide polymorphisms in *Lycopersicon esculentum* by computer aided analysis of expressed sequence tags. *Mol. Breed.*, 14: 21-34.

Yi, S.S., S.A. Jatoi, T. Fujimura, S. Yamanaka and K.N. Watanabe. 2008. Potential loss of unique genetic diversity in tomato landraces by genetic colonization of modern cultivars at a non-center of origin. *Plant Breed.*, 127: 189-196.

Zeven, A.C. 1998. Landraces: a review of definitions and classifications. *Euphytica*, 104: 127-139.

Zhu, H.S., H. Zhang, K.M. Mao, H.T. Li and A.Y. Wang. 2004. The genetic diversity of Yunnan local varieties and wild species of tomato. *J. Yunnan Agric. Univ.*, 19: 373-377.

(Received for publication 25 June 2010)