Pak. J. Bot., 49(5): 1857-1865, 2017.

EVALUATION OF CULTIVATED TOMATO GERMPLASM RESOURCES

YUTING WANG[1,2], WENZHEN LI[1,2], CHEN LU[1,2], SHAOZHU FAN[3], CHAOBIN FU[1,2], MINGSUO CHEN[1,2] AND LINGXIA ZHAO[1,2,*]

[1]Joint Tomato Research Institute, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China
[2]Plant Biotechnology Research Center, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China;
[3]Harbin Academy of Agricultural Science, Harbin, 150029, China
**Corresponding author's: email: lxzhao@sjtu.edu.cn; Phone: 86-21-34205775; Fax: 86-21-34205775*

Abstract

Lack of germplasm resources has severely limited genetic improvement of tomato (*Solanum lycopersicum*) in China. To potentially solve this issue, a total of 127 cultivated tomato accessions were introduced from the United States, Department of Agriculture (Geneva, NY, USA). These accessions have been disseminated to North America from Europe by a different route than the cultivated tomatoes in China, and have a different genetic background. A phylogenetic tree was drawn using 47 morphological markers, and a core germplasm collection comprising 20 tomato accessions was identified. Important quality traits such as fruit size, carotenoid levels, total soluble solids (TSS), fruit color and fruit softnesswere further examined in this core tomato germplasm collection. The results provide valuable information about this breeding material for genetic improvement of tomato in China. In order to save time and labor during the evaluation of the tomato germplasm resources, principal component analysis (PCA) was used to reduce dimensionalities, and it was found that the first 14 principal components contributed to 72.18% of the 47 phenotypes in the 127 tomato accessions. If the analysis of the core germplasm collection and the PCA analysis were used to evaluate other tomato germplasm resources, it could enhance breeding, and in addition it could also provide an important reference for evaluation of germplasm resources in other crops.

Key words: Phylogenetic tree; Genetic improvement; Tomato quality; Principal component analysis.

Introduction

Cultivated tomato (*Solanum lycopersicum*) originated from progenitors growing on the western side of the Andes Mountains in South America, close to the Pacific coast, and was domesticated in Mesoamerica by the year 7,000 BCE. Tomato cultivars had arisen by 3,500 BCE, which were cultivated in Mexico and other areas of Mesoamericaby the year 500 BCE. These tomato cultivars are thought to have been brought to Europe from Mexico by Hernan Cortez, a Spanish explorer in 1521 (yellow fruited tomato), or by Christopher Columbus, an Italian explorer, as early as 1493 (both red and yellow fruited tomato). Subsequently, the Spanish distributed tomato cultivars throughout their colonies in the Caribbean, after which it was introduced into North America. In parallel, tomato spread throughout Southeast Asia via the Philippines**,** and was then cultivated widely across Asia (Jenkins, 1948; Rodrı´guez *et al*., 2011; https://en.wikipedia.org/wiki/Tomato; https://en.wikipedia.org/wiki/Mesoamerica).

Originally, tomato was cultivated in Europe as an ornamental plant as the fruit were thought to be poisonous. However, after the 1540s, tomato was extensively grown in the Mediterranean area, reflecting the suitable growth climate, and by the early 17th century the fruit were consumed in countries including Italy, Spain and England (Rodrı´guez *et al*., 2011; Parisi *et al*., 2016; https://en.wikipedia.org/wiki/Tomato; https://en.wikipedia.org/wiki/Mesoamerica).

Alexander W. Livingston, an early tomato breeder in North America, developed different breeding methods and helped popularize tomato as a commercial crop in the 1870s, with different cultivars being used as a fresh fruit, canning and processing. Today, tomato has become the fourth commercially most important crop, with a value of more than $50 billion per annum (Lin *et al*., 2014; Uluisik *et al*., 2016; http://faostat3.fao.org/home/E). However, it is becoming more difficult to breed new high quality tomato varietiesusing European tomato germplasm due to a deficiency in essential genetic diversity (Jenkins, 1948; Zamir, 2001). In addition, to satisfy demands from customers, breeders have focused on elevating yield, increasing resistance to biotic/abiotic stresses, and extending shelf life, which has resulted in a further narrowing of the genetic background (Rick & Chetelat, 1995; Zamir, 2001; Rodrı´guez *et al*., 2011; Casals *et al*., 2012; Ercolano *et al*., 2012; Ghiani *et al*., 2016; Lin *et al*., 2016; Ohlson & Foolad, 2016; Parisi *et al*., 2016; Zeinab Ibrahim, 2016).

Substantial genetic diversity exists in the wild relatives of tomato collected from the center of origin of tomato in South America, which collectively represent a potential gene bank for tomato genetic improvement (Rick, 1986; Rick & Chetelat, 1995; Qu *et al*., 2015). Many of these wild relatives and various genotypes were collected by the tomato research pioneer, Charles M. Rick and his colleagues, and are currently conserved in the Tomato Genetic Resource Center (TGRC, UC Davis)(http://tgrc.ucdavis.edu/).

Genetic crossing with wild relatives has provided an effective strategy for improving cultivated tomato, and has resulted in numerous cultivated tomato varieties with traits such as resistance to biotic stresses (Martin *et al*., 1994; Yaghoobi *et al*., 2005; Ercolano *et al*., 2012; Ohlson & Fooald, 2016), abiotic stresses (Fischer *et al*., 2011; Zeinab Ibrahim, 2016), and improved fruit quality (Chetelat *et al*., 1995a; 1995b). However, both crossing incompatibility (CI) and unilateral incompatibility (UI) reproductive barriers exist between cultivated tomato and certain wild relatives, and these have proven difficult to overcome (Li *et al*., 2010; Li & Chetelat, 2010; Bedinger *et al*., 2011). Furthermore, the hybrid progeny exhibit considerable genetic segregation, even compared to its direct ancestor,

Solanum pimpinellifolium (Rick, 1951; Gao *et al*., 2016) and eliminating linkage drag through conventional breeding is time-consuming and laborious. There are therefore limitations in the application of wild tomato species to improve cultivated tomato (Zamir, 2001).

Current breeding goals for fresh tomato focus on specific quality traits, increased yield, shelf life and resistance to biotic/abiotic stresses. An example of a fruit quality trait that has been targeted for improvement is the accumulation of carotenoids and flavonoids, which has been of great interest to both breeders and consumers (Zamir, 2001; Schauer *et al*., 2005; Giovannoni, 2006; Gonzali *et al*., 2009). With the rapid development of genetic engineering techniques, genetic modification has also been extensively used to improve tomato quality (Delannay *et al*., 1989; Butelli *et al*., 2008; Lim *et al*., 2016; Sagor *et al*., 2016). However, the products of transgenic plants face consumer resistance in many countries, including China and Japan and in the European Union. This has promoted fundamental research into the biosynthetic and regulatory pathways that govern the accumulation of secondary metabolites using tomato breeding (Gao *et al*., 2016).

Tomato is one of the leading vegetable crops worldwide, with a production of 164 million tons in China, which corresponds to a third of global production (http://faostat3.fao.org/home/E). It is extensively used as a fresh vegetable crop or in salads, in processed foods, such as like tomato ketchup and canned tomato, and in the production of tomato juice. Fresh tomato qualities such as fruit color, texture, taste, aroma and levels of nutrient, including sugars, organic acids, carotenoids, are all valued, but it is difficult to optimize all these parameters to satisfy consumer demands.Essential genetic resources to improve tomato qualities are not available in China, and the genetic diversity has become narrow during the spread of the crop into China and the subsequent selection bottleneck resulting from modern breeding programs (Jenkins, 1948; Zamir, 2001; https://en.wikipedia.org/wiki/Tomato). The acquisition of new genetic resources therefore has the potential to provide a path for improving tomato quality. S. *lycopersicum* heirloom or landrace genotypes are particularly attractive in this regard as they have no reproductive barriers with the cultivated tomatoes, and have not yet been employed in any tomato breeding programs in China.

We acquired 127 tomato heirloom or landrace accessions from the United States Department of Agriculture collection in Geneva, NY, USA. These tomato accessions were spread from Europe via different routes from those currently grown in China, and have not yet been employed extensively in modern tomato breeding programs. In this current study, we evaluated their anatomical and growth characteristics, as well as a range of fruit quality traits. This research provides important information for the future improvement of tomato quality in China.

Materials and Methods

Plant materials and growth conditions: Tomato seeds from a total of 127 tomato accessions (Supplementary Table 1) were obtained from Professors Gan-yuan Zhong and Larry Robertson (Agricultural Research Service, Plant Genetic Resources Research, United States Department of Agriculture, Geneva, NY, USA) via Professor Shi-heng Wang (Hangzhou Academy of Agricultural Sciences, Hangzhou, China). The seeds were sown separately in 60 cell breeding plug trays (Taizhou Sophia Import & Export Co., Ltd, Zhejiang, China) in humid peat pellets and germinated at a temperature of 26/20°C(day/night) in a standard greenhouse at the Pujiang experimental farm in the School of Agriculture and Biology in Shanghai Jiao Tong University (Shanghai, China). Seedlings with four expanded leaves were planted in a natural light polycarbonate greenhouse (10 seedlings per tomato accession) at the same location.

Tomato phenotypes: A total of 47 phenotypes were measured or characterized for the 127 tomato accessions according tocriteria described in the section of descriptions for tomato (*Solanum* spp.) in the International Plant Genetic Resource Institute (Anon., 1996).

Hypocotyl color and the color of the primary leaf vein were recorded at the seedling stage, when the primary leaves were fully opened and the terminal bud was approximately 5 mm in size. Plant characteristics, including growth type and size, were recorded when the fruits on the second to third truss ripened, as were leaf characteristics (called **l**) including leaf posture, type, color, and coloration of the leaf veins due to anthocyanin accumulation.

All observations of the inflorescence and fruit were conducted using the third fruit of the second and/or third truss at the fully mature stage. Flower characteristics (named **f**, which comprised **f1**, inflorescence type; **f2**, inflorescence after leaf; **f3**, fascicle type; **f4**, corolla color; **f5**, style length; **f6**, anther number per flower; **f7**, petal number per flower; **f8**, sepal number per flower; **f9**, flower number per inflorescence type; **f10**, sepal length; **f11**, petal length; **f12**, anther length and **f13**, style length). The fruit characteristics (named **fr**) included **fr1**, immature fruit color; **fr2**, fruit cross-sectional shade; **fr3**, fruit apex; **fr4**, fruit shoulder shape; **fr5**, fruit shoulder color; **fr6**, fruit shape; **fr7**, flesh color of pericarp; **fr8**, skin color of ripe fruit; **fr9**, fruit shoulder ribbing; **fr10**, pubescence; **fr11**, green fruit shoulder; **fr12**, ventricle number per fruit; **fr13**, pedicel length from abscission layer; **fr14**, longitudinal diameter; **fr15**,transverse diameter; **fr16**, longitudinal/ transverse diameter ratio; **fr17**, thickness of pericarp; and **fr18**, fruit weight. The tomato seeds (named **s**) were removed from fruit at the red ripe stage, and then seed length (**s1**), width (**s2**), thickness (**s3**) and dry weight of 1,000 seeds (**s4**) were recorded.

The lengths of all organs including flower, fruit and seed were measured using calipers (Mitutoyo CD-15CPX, Japan, 0.01 mm) (five biological replicates), while width and length of the largest leaves were measured with a measuring tape (L19-50, Shanghai Pengxing, Shanghai, China, 0.1cm). The weights of the fruit and 1,000 seeds were determined using an electronic balance (JE3001, Shanghai, China, precision, 0.1g) and an analytical electronic balance (HZY-A120, Zhengzhou Mingyi Instrument Equipment Co., Ltd, Zhengzhou, China, precision, 0.001g), respectively. Other characteristics not mentioned above were recorded according to standards in Descriptors for Tomato (1996).

Supplementary Table 1. The origin of the tomato accessions in this study.

*Accession number	Original accession ID	Accession number	Original accession ID
001	P19753870A1	065	PL63920804G1
002	P19809706G1	066	G3300910G1
003	P19978275A1	067	PL63920804G1
004	P110983406G1	068	PL64488511G1
005	PL11756384A1	069	G3304611G1
006	PL64751399G1	070	PL63627703G1
007	PL58445607G1	071	G3304711G1
008	PL11878306G1	072	G3304811G1
009	PL12403787G1	073	G3304911G1
010	PL12782008G1	074	G3305011G1
011	PL12782508G1	075	PL43887797G1
012	PL12859208G1	076	G3303811G1
013	PL12902608G1	077	G3304511G1
014	PL12903308G1	078	G3304011G1
015	PL12908408G1	079	PL44173997G1
016	PL12912806G1	080	PL64753397G1
017	PL12914208G1	081	G3306311G1
018	PL15537208G1	082	G3307711G1
019	PL15799368A1	083	G3307811G1
020	PL15876006G1	084	PL30381168A1
021	PL15900970A1	085	PL27021263A1
022	PL15919806G1	086	PL45201897G1
023	PL21206269A1	087	PL26595597G1
024	PL25847806G1	088	PL27023663A1
025	PL26299507G1	089	PL27023999G1
026	PL27020606G1	090	PL27956562G1
027	PL27040861A1	091	PL30374965A1
028	PL27043096G1	092	PL30967272A1
029	PL27270306G1	093	G3300811G1
030	PL28155506G1	094	PL30966981A1
031	PL29133706G1	095	PL33991470A1
032	PL29463806G1	096	PL34112498G1
033	PL34113406G1	097	PL34113296G1
034	PL39051075A1	098	PL34113396G1
035	PL40695276A1	099	PL64521411G1
036	PL45202606G1	100	PL64712284A1
037	PL45202706G1	101	PL63630203G1
038	PL50531706G1	102	PL64537011G1
039	PL64744505G1	103	PL64538910G1
040	PL647447	104	PL64539009G1
041	PL64755601G1	105	Pl64539109G1
042	PL64756602G1	106	PL64539811G1
043	PL3301011G1	107	PL64731698G1
044	PL45199379A1	108	PL60090611G1
045	G3301111G1	109	PL60090711G1
046	PL63921104G1	110	PL60092006G1
047	G3301311G11	111	PL60113605G1
048	G3301410G1	112	PL60116511G1
049	PL27018601G1	113	PL60117711G1
050	PL23425473A1	114	PL60117811G1
051	G3301711G1	115	Pl60119207G1
052	PL2701989061	116	PL60141187ll0
053	PL27020270A1	117	PL55991294G1
054	PL45199079A1	118	Pl60134209G1
055	PL63921504G1	119	Pl60139610G1
056	PL29085705G1	120	PL60144910G1
057	PL12899001G1	121	PL60145011G1
058	G3301911G1	122	PL60151211G1
059	PL25043604G1	123	PL60162910G1
060	G3302511G11	124	PL28625504G1
061	G3302010G1	125	PL64730510G1
062	PL33993896G1	126	PL63626203G1
063	PL64504811G11	127	PL63921304G1
064	PL30381004G1		

***** All accessions were obtained from the Agricultural Research Service, Plant Genetic Resources Research, United States Department of Agriculture, Geneva, NY, USA.

Phylogenetic tree using morphological markers: A total of 47 phenotypes/morphological markers from the 127 tomato accessions were measured or characterized according to criteria described in the section of descriptions for tomato (Anon., 1996), and data were analyzed using gplots (R 3.2.2 version) (http://mirror.bjtu.edu.cn/cran/ and HelpFilehttp://docs.ggplot2.org/current/index.html) software. Cluster analysis was conducted using the heatmap.2 function to draw a heat-map of the phylogenetic tree derived from the phenotypic data. The core germplasm resource was determined by phylogenetic trees based on both morphological markers and RAPD (Random Amplified Polymorphic DNA) markers.

Quantity traits of tomato fruit in the core germplasm resource: Carotenoid content was measured as described by Gao *et al*. (2015), with a α-carotene standard purchased from Santa Cruz Biotechnology (Santa Cruz, CA, USA), and lutein, zeaxanthin, β-carotene and lycopene standards from Sigma Chemicals (St. Louis, MO, USA). Vitamin C content was determined using the 2,6-dichloro-indophenol titration method described as Jones & Hughes (1983). Total soluble solids (TSS) were measured using a Sugar Refractometer (Bellingham Stanley DRl03L, Britain). Fruit acidity was determined using the neutralization titration method described as Jakmunee *et al*. (2006), and chlorophyll content as described by Lichtenthaler (1987). Three biologic replicate samples were analyzed for all quality traits mentioned above. Fruit firmness and thickness of the pericarp were measured using a Fruit hardness tester (GY-3, Zhejiang, China) and calipers (Mitutoyo CD-15CPX, Japan, 0.01 mm), respectively.

Principal component analysis: A total of 47 tomato phenotypic traits were assayed and a principal component analysis (PCA) was performed using gplot (Ri386 3.3.0 version, http://mirror.bjtu.edu.cn/cran/ and Help File http://docs.ggplot2.org/current/index.html) software. Dimensionality reduction was achieved using the princomp function in PCA, and the scatter plot was drawn with a gplot function (Bro & Smilde, 2014).

Results

Tomato phenotypic characters: A total of 47 tomato phenotypic traits were examined in 127 accessions, comprising 2 traitsfor plant growth habits, 9 traits for leaf growth, 1 trait for glandular hair growth, 13 flower traits, 18 fruit traitsand 4 seed traits. Two types of plant growth habits were noted: indeterminate and determinate (dwarf).

Four-leaf types: standard, potato leaf, broad leaf and *pimpinellifolium* were observed in the tomato accessions, as well as three leaf postures: semi-erect, horizontal and drooping. Two leaf shapes, odd-pinnately compound leaf and even-pinnately compound leaf, were observed, as were four leaf colors: yellow green, pale green, green, and deep green (Fig. 1a).

Glandular hairs were observed on both leaves and stems. Leaf length and leaf width varied from 20.20 cm (accession 007) to 63.63 cm (accession 085), and 12.63 cm (accession 007) to 55.43 cm (accession 087), respectively.

Three inflorescence types were noted: partly uniparous, multiparous and generally multiparous and solitary flowers were not observed. The corolla colors were light yellow or yellow to orange, with yellow corolla accessions constituting 83% (106 out of 127) (Fig. 1b). Long, short and nearly equal to antheral length style lengths were observed, with the short and nearly equal to antheral predominating. The 16 tomato accessions with longer styles represented 13% of the total accessions (Fig. 1c). The tomato accessions with longer styles have the potential to facilitate crossing between accessions and increase genetic diversity. Anther number per flower varied from 5 (common) to 13 (accession 078), and the number of petals and sepals was generally equal to that of anthers in the same flower, while it occasionally varied between floral organs.

Fruit cross-sectional shapes ranged from round to angular and irregular, and 78% of the accessions were round fruited (99 out of 127). Three types of fruit apices were observed, indented, slightly indented, and flat, as well as three fruit shoulder shapes: flat, depressed, and strongly depressed. We also saw three different shoulder colors, absent, light green and green, with 90% of the accessions (114 out of 127) being in the absent category.

Nine different fruit shapes were observed: flat, oblate, round, high round, prolate round, ovate, peach, pear and prolate pear-shaped. The oblate shaped fruits were most common, representing 32% (Fig. 1d). Of the six flesh pericarp colors (yellowish white, light yellow, yellow, pinkish red, red and green), the red and green represented 47% (59/127) and 42% (53/127), respectively (Fig. 1e). The locule number varied from 2 (accessions 005, 006, 032 and more) to 19 (accession 073), and accessions with two and three locules constituted 24% and 17%, respectively (Fig. 1f). The ongitudinal diameter and transverse diameter, respectively, ranged from 22.15 mm (accession 083) to 150.93 mm (accession 107), and from 19.94 mm (accession 034) to 175.24 mm (accession 107) in fruit, while pericarp thickness ranged from 1.77 mm (accession 034) to 8.98 mm (accession 113). Single fruit weight varied from 4.86 g (accession 034) to 426.42 g (accession 073), and the accessions of that more than 150 g only had 14% (Fig. 1g). The 1,000 seed weight parameter varied from 1.52 g (accession 0.58) to 4.24 g (accession 119), with most between 2.00 g and 4.00 g (Fig. 1h).

Phylogenetic trees and construction of a tomato core germplasm resource: Based on the 47 phenotypic/ morphological markers derived from the 127 tomato accessions, a phylogenetic tree was generated, using the heatmap.2 function in the gplot software package (Fig. 2). The tomato accessions formed four distinct groups. Accession 100 and accession 007 grouped independently from each other and the others accessions, and were assigned group 1 and group 3, respectively. The third group consisted of accessions 046, 066, 067, 073, 078 and 107, while the other 119 tomato accessions were in group 2, which could be further divided into subgroups 1 and 2, comprising 59 and 60 tomato accessions, respectively (Fig. 2).

Based on the results of the phylogenetic tree, combined with the phylogenetic tree drawn using RAPD markers (data not shown), we found more diversity in tomato accessions from the USDA than from indigenous Chinese tomato genotypes. Thus, the introduced tomatoes represent a potentially valuable source of germplasm for tomato breeding in China.

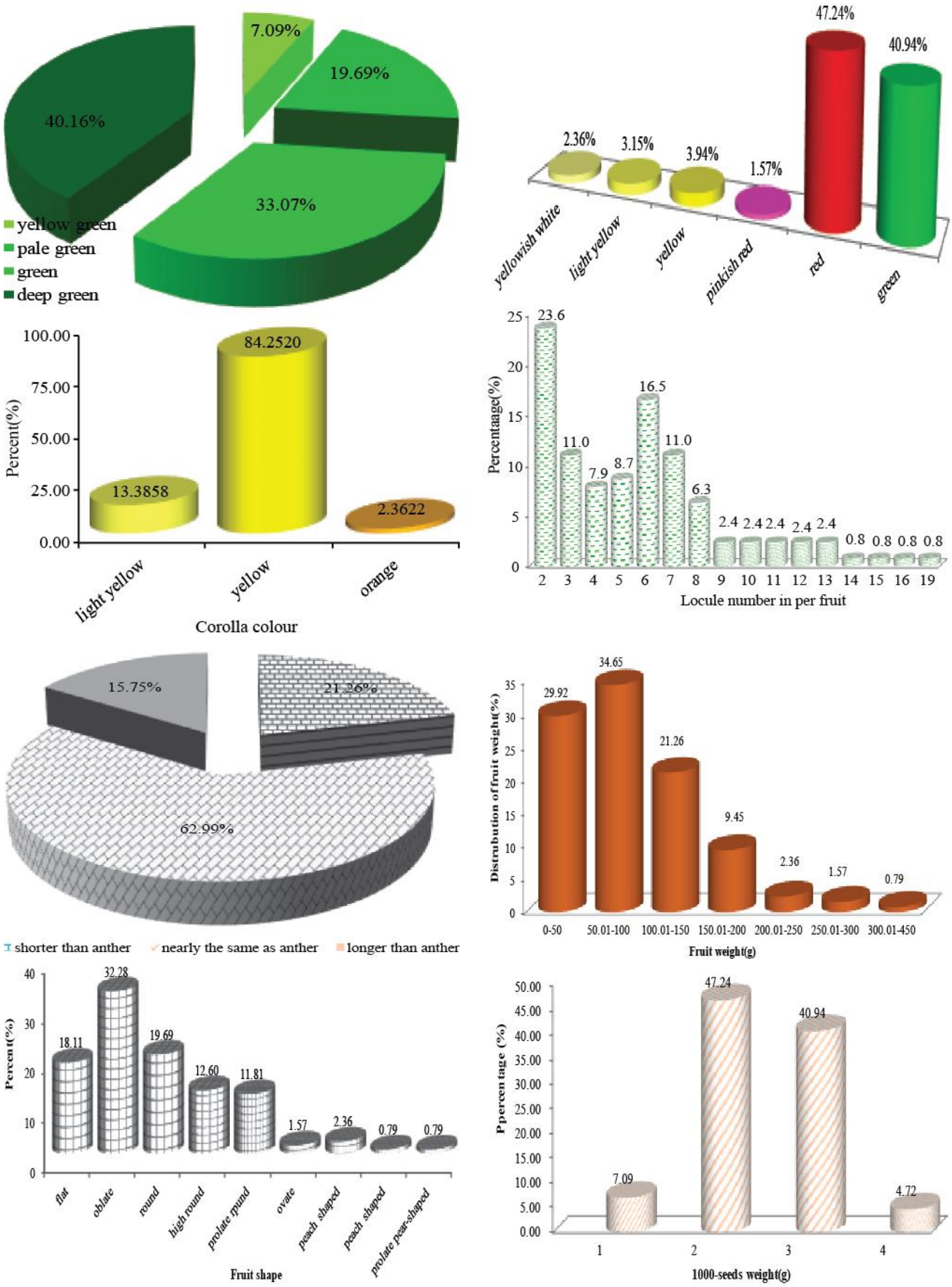

Fig. 1. Distribution of phenotypes in 127 tomato accessions obtained from USA
(a, Distribution of leaf color**; b,** Distribution of corolla color**; c,** Distribution of style length; **d,** Percentage of different fruit shapes; **e**, Distribution of flesh color at the red ripe stage; **f**,Percentage of locule number**; g,** Distribution of single fruit weight; h, Percentage of 1000-seed weight**.**

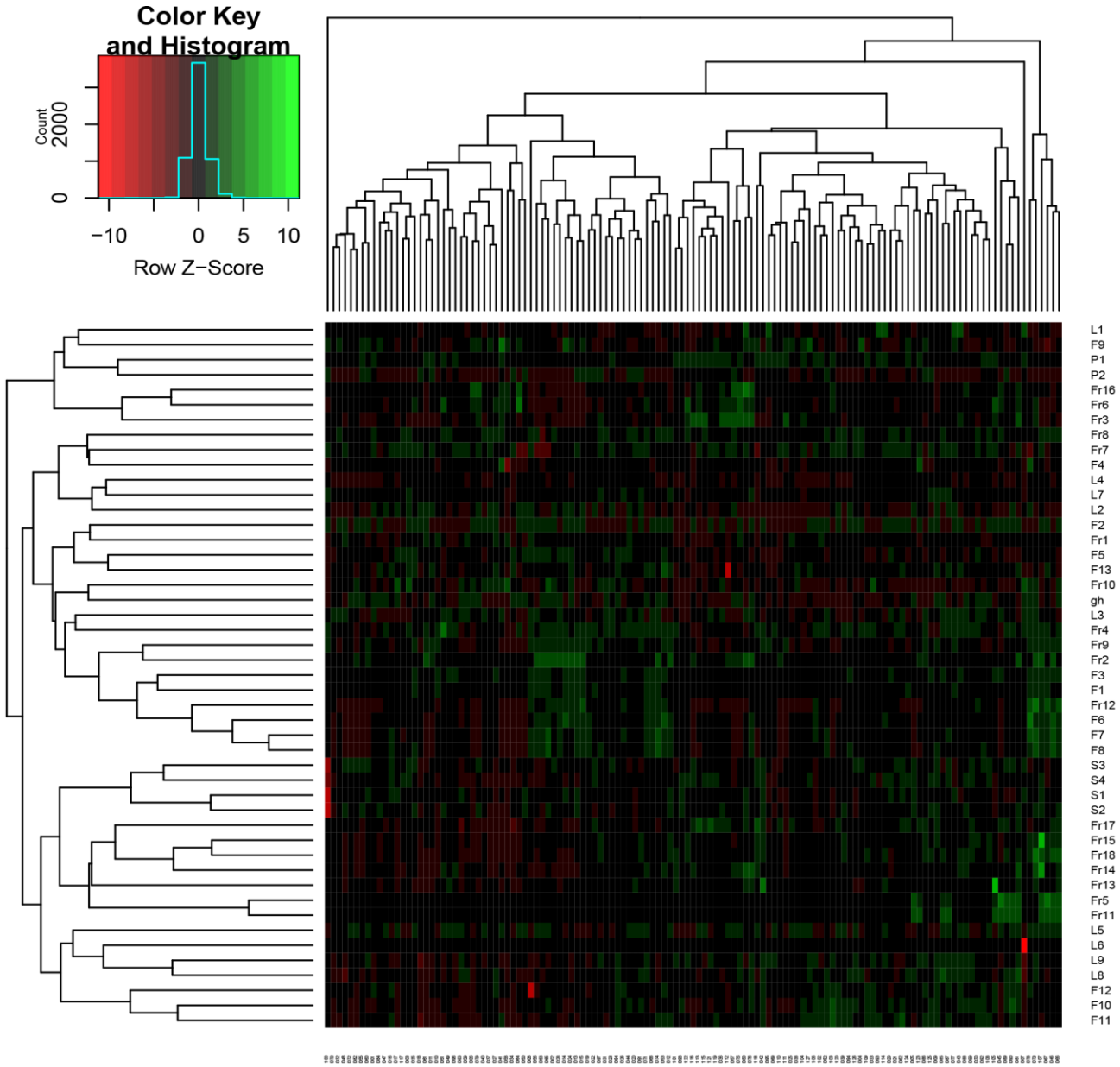

Fig. 2. Heat-map showing data from a phylogenetic tree of 127 tomato accessions based on morphological markers.
The numbers 001 to 127 at bottom of Phylogenetic tree indicate accessions of tomato derived from USDA, and the codes at right side of the Phylogenetic tree indicate total of 47 phenotypic features in tomato.
P, plant; p1, growth type, 0- indeterminate, 1- dwarf, determinate; **p2**, tomato plant size, 0- intermediate, 1 -small and 2-large. **L, leaves;L1**, colors of primary leaf vein, 0-clear, 1-purple, 2-green; **L2**, leaf type, 1-standard, 2- potato leaf, 3-broad leaf, 4-pimpinellifolium; **L3**,leaf morphology, 1-semierect, 2-horizontal, 3-drooping; **L4**, **l**eaf shade, 1- odd-pinnately compound leaf, 2-even-pinnately compound leaf; **L5**, leaf color, 1-yellow green, 2-pale green, 3-green, 4-deep green; **L6**, anthocyanin coloration of leaf veins, 1-purple, 2- green; **L7**, leaf lobes/ degree of leaf dissection, 1-no, 2- low, 3- intermediate, high. **gh**: glandular hairs on the leaves or steams, 0-none, 1-thin short glandular hairs, 2- thick short glandular hairs, 3- thin long glandular hairs, 4-thick long glandular hairs. **F, flower; f1**, inflorescence type, 1-solitary flower, 2-generally uniparous, 3-both partly uniparous and multiparous, 4- generally multiparous; **f2,** inflorescence after leaf, 0- absent, 1-present; **f3**, fascicle type, 0-absent, 1-present; **f4**, corolla color, 1-light yellow, 2-yellow, 3-orange; **f5**, style length, 1-shorter than stamen, 2-nearly the same level as stamen, 3-longer than stamen; **f6**, anther number per flower; **f7**, petal number per flower; **f8**, sepal number per flower; **f9**, flower number per inflorescence type; **f10**, sepal length (mm); **f11**, petal length (mm) ; **f12**, anther length (mm); **f13**, style length (mm). **Fr, fruit**;**fr1**, immature fruit color, 1-greenish white, 2-light green, 3-green, 4-dark green; **fr2**, fruit cross-sectional shade, 1- round, 2- angular, 3- irregular; **fr3**, fruit apex, 1-indented, 2-slightly indented, 3-flat, 4-salient, 5-pointed; **fr4**, fruit shoulder shape, 1-flat, 2-depressed, 3-strongly depressed; **fr5**, fruit shoulder color, 0-absent, 1-light green,2-green; **fr6**, fruit shape, 1-flat, 2-oblate, 3-round, 4-high round, 5-prolate round, 6-ovate, 7-peach-shaped, 8-pear-shaped, 9-prolate pear-shaped; **fr7**, flesh color of pericarp, 1-yellowish white, 2-light yellow, 3-yellow, 4-pinkish red, 5-red, 6-green; **fr8**, skin color of ripe fruit, 1-transparent, 2-yellow, 3-red; **fr9**, fruit shoulder ribbing, 0-none, 1-little, 2-intermediate, 3-prominent; **fr10**, pubescence, 0-none, 1-sparse, 2-intermediate, 3-dense; **fr11**, green fruit shoulder, 0-absent, 1-present; **fr12**, ventricle number per fruit; **fr13**, pedicel length from abscission layer (mm); **fr14**, longitudinal diameter (mm); **fr15,** transverse diameter (mm); **fr16**, longitudinal / transverse diameter ratio; **fr17**, thickness of pericarp; **fr18**, fruit weight (g). **S, seeds**; **s1**, length of seed (mm); **s2**, width of seed (mm); **s3**, thickness of seed; **s4**, 1000-seed weight (g).

To conserve and utilize this tomato germplasm resource, a core germplasm resource consisting of 20 accessions (005, 007, 012, 016, 019, 043, 045, 051, 052, 056, 062, 077, 082, 088, 098, 101, 104, 107, 109 and 125) was created based on the phylogenetic analysis of both the morphological markers and the RAPD markers.

Qualitative characters of the tomato core germplasm: Carotenoids are antioxidants that accumulate in tomato fruit and petals. The carotenoid pigment lycopene is one of the most abundant, and we detected between 6.31 μg/g fresh weight (FW, accession 007) to 1,745.00 μg/g FW (accession 098) in pericarp at ripe-red stage from the core germplasm resource. The β-carotene levels varied from 32.00 μg/g FW (accession 109) to 156.30 μg/g FW (accession 062), while the levels of lutein, α-carotene and zeaxanthin were too low to be detected by the HPLC assay (Table 1).

The content of ascorbic acid varied from 0.1 mg/100g FW (accession 125) to 0.2 mg/100 g FW (accession 077), while the TSS ranged from 3.70 % (accession 101) to 5.27 % (accessions 052 and 088), and the acid content from 262 mg/100g FW (accession 005) to 638 mg/100 g FW (accession 056). The longitudinal firmness of the ripe-red fruit differed from that of the transverse firmness, with the former varying from 3 Newtons (N, accession 056) to 6.13 N (accession 007), and the latter from 2.59 N (accession 077) to 5.84 N (accession 007). The mean value of the longitudinal firmness was 4.61 N, which was greater than the mean transverse value (3.68 N) (Table 1).

PCA of phenotypic traits: PCA analysis was used to reduce the complex data set consisting of 47 phenotypic features to a lower dimension, to identify hidden and simplified structures, which often underlie complex data sets. Forty-seven phenotypic traits derived from the 127 tomato accessions were analyzed and the contribution of each principal component was calculated. The first twenty-two principal components contributed 85.86% (Table 2), and the single fruit weight could be distinctly divided into five classes (0 to 50g, 51to 100g, 101 to 150g, 151 to 200g and over 201g, Fig. 3). For the floral characteristics, the first seven principal components contributed 88.98%, while the first nine principal components contributed 85.51% for the eighteen fruit phenotypes and 89.37% for the twelve leaf phenotypes.

Discussion

The ancestors of cultivated tomato were native to a long and narrow area of the western Andes in an area of that is currently in Peru and Ecuador. In pre-Columbian times, tomato was possibly treated as a weed that spread north, and was not extensively domesticated until it reached Mexico. It is believed that the cultivated forms of tomato were spread worldwide via two routes from the center of domestication in Mexico. Firstly, it was originally brought to Spain by European explorers in the fifteenth century. These genotypes then spread into East Asia and China *via* the Philippines. Secondly, the cultivated tomato in Europe was then introduced into Northern America *via* the Caribbean (Jenkins, 1948; https://en.wikipedia.org/wiki/Tomato). It can therefore be inferred that the tomato was introduced into China from Spain by a different route than the cultivars that are grown in Northern America. Currently, China is the biggest tomato producer worldwide; however, tomato genetic improvement has been hampered by a lack of tomato germplasm resources, consumer resistance to the use of genetic modification, and limitations in the introgression potential of traits from wild tomato species (Zamir, 2001; Li & Chetelat, 2010; Lim *et al*., 2016; Sagor *et al*., 2016). To elevate the level of genetic diversity, we evaluated a collection of tomato accessions from the USDA that we hypothesized had a distinct genetic background to existing Chinese accessions due to their different origins. In addition, these tomato accessions have no reproductive barriers with the cultivated tomato, and have properties such as high carotenoid content (1745.00 μg/ g FW, accession 098) (Table 1), large fruit weight (426.42 g, accession 073) and leaf morphologies, which have potential for improving tomato varieties in China.

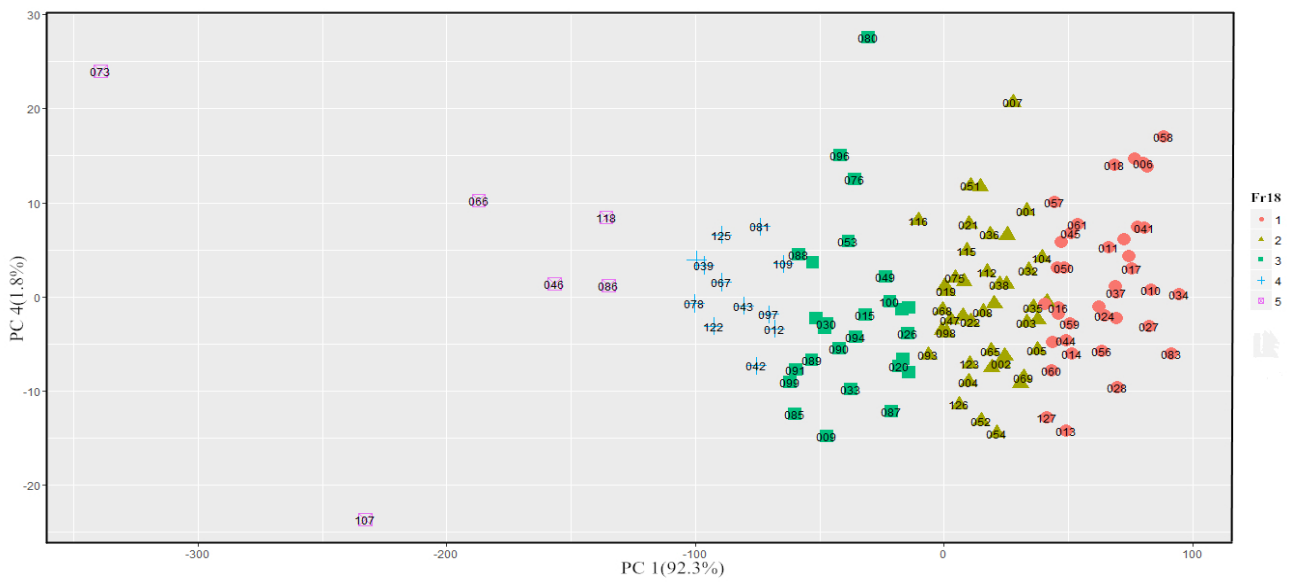

Fig. 3. PCA scores of 47 phenotypic characters on PC1 and PC4 for the 127 tested tomato accessions.
Note, numbers in the figure corresponding to the tested tomato accessions are shown in Supplementary Table 1. Fr18 indicates single fruit weight, and fruit weight 0 to 50g assigned to 1, 51 to 100g assigned to 2, 101 to 150g assigned to 3, 151 to 200g assigned to 4 and over 201g assigned to 5.

Table 2. Principal component analysis of the forty-seven phenotypic features in total of 127 tomato accessions.

	Com.1	Com.2	Com.3	Com.4	Com.5	Com.6	Com.7	Com.8
PV*	0.1612	0.1150	0.0593	0.0581	0.0513	0.0434	0.0399	0.0366
CP	0.1612	0.2757	0.3350	0.3931	0.4444	0.4878	0.5278	0.5643
	Com.9	**Com.10**	**Com.11**	**Com.12**	**Com.13**	**Com.14**	**Com.15**	**Com.16**
PV	0.0315	0.0289	0.0272	0.0253	0.227	0.0219	0.0212	0.0201
CP	0.5958	0.6247	0.6520	0.6772	0.6999	0.7218	0.7430	0.7631
	Com.17	**Com.18**	**Com.19**	**Com.20**	**Com.21**	**Com.22**	**Com.23**	**Com.24**
PV	0.0183	0.0175	0.0163	0.0152	0.0145	0.0138	0.0129	0.0119
CP	0.7813	0.7989	0.8151	0.8303	0.8449	0.8586	0.8716	0.8835
	Com.25	**Com.26**	**Com.27**	**Com.28**	**Com.29**	**Com.30**	**Com.31**	**Com.32**
PV	0.0111	0.0108	0.0101	0.0097	0.0086	0.0079	0.0077	0.0071
CP	0.8950	0.9054	0.9155	0.9251	0.9337	0.9416	0.9493	0.9564
	Com.33	**Com.34**	**Com.35**	**Com.36**	**Com.37**	**Com.38**	**Com.39**	**Com.40**
PV	0.0062	0.0052	0.0051	0.0046	0.0040	0.0037	0.0032	0.0028
CP	0.9626	0.9678	0.9729	0.9775	0.9815	0.9852	0.9884	0.9912
	Com.41	**Com.42**	**Com.43**	**Com.44**	**Com.45**	**Com.46**	**Com.47**	
PV	0.0027	0.0020	0.0016	0.0011	0.0008	0.0004	0.0002	
CP	0.9939	0.9959	0.9975	0.9986	0.9994	0.9998	1.0000	

*PV: Proportion of Variance; CP: Cumulative Proportion; Com. n: Component serial number.

We systematically evaluated the heirloom or landrace resources introduced from the USA and defineda representative germplasm resource of 20 accessions from the total of 127 tomato accessions, reflecting phylogenetic trees that were constructed based on RAPD analysis and 47 phenotypic features (Fig. 2). The 20 accessions collectively covered the diversity of the 127 tomato accessions. Importantly, accessions in this core germplasm collection have no reproductive barriers with cultivated tomato, so the best candidate genes controlling important agricultural traits can be easily introduced into parental varieties for tomato breeding in China. The established core germplasm collection can easily be conserved in a limited space, such as the national germplasm bank.

The PCA analysis, involving 5969 phenotypic traits derived from 127 accessions,revealed that the first 22 principal components contributed 85.86%, and that even the first 14 principal components contributed 72.18% (Table 2). The contribution of the first four principal components comprised 70% of 13 floral characteristics, while the first nine principal components contributed 85.51% of the eighteen fruit phenotypes. We could conclude that the PCA method was effective in evaluating the tomato diversity and will help facilitate conservation of major tomato germplasm resources for research (Saeed *et al*., 2017). This research not only highlights resources for tomato genetic improvement, but also provides a pipeline for the evaluation of other crop resources.

Contributions

YTW performed phenotypic evaluation and PCA, WZL conducted the RAPD analysis, CL participated in polygenetic analysis of the phenotypes, FSZ conducted the work in the field; CBF and MSC participated in the phenotypic trait assessment; LXZ designed the experiment and wrote the manuscript.

Acknowledgements

The authors are grateful to Dr. Ganyuan Zhong and Dr. Larry Robertson (Agricultural Research Service, Plant Genetic Resources Research, United States Department of Agriculture, Geneva, NY, USA), and Professor Shi-heng Wang (Hangzhou Academy of Agricultural Sciences, Hangzhou, China) for generously providing tomato seeds. We would like to thank Dr. Qishan Wang (School of Agriculture and Biology, Shanghai Jiaotong University) and Dr. Yu Dong (School of Life science and Biotechnology, Shanghai Jiaotong University) for providing assistance with PCA and phylogenetic analyses of morphological markers, respectively. We thank PlantScribe (www.plantscribe.com) for editing this manuscript. This work was supported by the Key Technology Research and Development Program of Shanghai Technology Committee (no.14JC1403400 and no. 16391900900) and National Key Research and Development Project (no.2016YFD0100204-28).

References

Anonymous. 1996. International Plant Genetic Resources Institute (IPGRI). Descriptors for tomato (*Lycopersicon* spp) Rome: pp. 1-46.

Bedinger, P., R.T. Cheteleat, B. McClure, L. Moyle, J.K.C. Rose, S. Stack, E. van der Knaap, Y. Baek, G. Lopez-Casado, P. Covey, A. Kumar, W. Li, R. Nunez, F. Cruz-Garcia and S. Royer. 2011. Interspecific reproductive barriers in the tomato clade: opportunities to decipher mechanisms of reproductive isolation. *Sex. Plant Reprod.*, 24(3): 171-187.

Bro, R. and A.K. Smilde. 2014. Principal component analysis. *Anal. Methods.*, 6: 2812-2831.

Butelli, E., L. Titta, M. Giorgio, H.P. Mock, A. Matros, S. Peterek , E.G. Schijlen , R.D. Hall, A.G. Bovy, J. Luo and C. Martin. 2008. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat. Biotechnol.*, 26: 1301-1308.

Casals, J., L. Pascual, J. Can˜ izares, J. Cebolla-Cornejo, F. Casan˜as and F. Nuez. 2012. Genetic basis of long shelf life

and variability into penjar tomato. *Genet. Resour. Crop Evol.*, 59: 219-229.

Chetelat, R.T., J.W. Deverna and A.B. Bennett. 1995a. Introgression into tomato of the *L. chmielewskii* sucrose accumulator gene controlling fruit sugar composition. *Theor. Appl. Genet.*, 91(2): 327-333.

Chetelat, R.T., J.W. DeVerna and A.B. Bennett. 1995b. Effects of the *Lycopersicon chmielewskii* sucrose accumulator gene (*sucr*) on fruit yield and quality parameters following introgression into tomato. *Theor. Appl. Genet*., 91(2): 334-339.

Delannay, X., B.J. LaVallee, R.K. Proksch, R.L. Fuchs, S.R. Sims, J.T. Greenplate, P.G. Marrone, R.B. Dodson, J.J. Augustine, J.G. Layton and D.A. Fischhoff. 1989. Field performance of transgenic tomato plants expressing the *Bacillus Thuringiensis* var. *Kurstaki* insect control protein. *Nat. Biotechnol*., 7: 1265-1269.

Ercolano, M.R.,W. Sanseverino, P. Carli, F. Ferriello and L. Frusciante. 2012. Genetic and genomic approaches for R-gene mediated disease resistance in tomato: retrospects and prospects. *Plant Cell Rep*, 31: 973-985.

Fischer, I., L. Camus-Kulandaivelu, F. Allal and W. Stephan. 2011. Adaptation to drought in two wild tomato species: the evolution of the *Asr* gene family. *New Phytol*, 190: 1032-1044.

Gao, L., W.H. Zhao, H.O. Qu, M.H. Wang, L.D. Zhang, Q.S. Wang and L.X. Zhao. 2016. The *yellow fruited tomato* (*yft1*) mutant has altered fruit carotenoid accumulation and reduced ethylene production as a result of a genetic lesion in *Ethylene intensitive* 2. *Theor. Appl. Genet.*, 129(4): 717-728.

Gao, M., H. Qu, L. Gao, L. Chen, R.S.J. Sebastian and L. Zhao. 2015. Dissecting the mechanism of *Solanum lycopersicum*and *Solanum chilense* flower colour formation. *Plant Biol.*, 17: 1-8.

Ghiani, A., N.D'Agostino, S. Citterio, A. Raiola, R. Asero, A. Barone and M.M. Rigano. 2016. Impact of wild loci on the allergenic potential of cultivated tomato fruits. *PLoS ONE*, 11(5): e0155803.

Giovannoni, J.J. 2006. Breeding new life into plant metabolism. *Nat. Biotechnol*, 24: 418-419.

Gonzali, S., A. Mazzucato and P. Perata. 2009. Purple as a tomato: towards high anthocyanin tomatoes. *Trends Plant Sci*., 14: 237-241.

Harmut, A. and K .Lichtenthaler. 1987. Chlorophylls and carotenoids-pigments of photosynthetic biomembranes. *Method Enzymol*, 148: 350-382.

Jakmunee, J., T. Rujiralai and K. Grudpan. 2006. Sequential injection titration with spectrophotomertric detection for the assay of acidity in fruit juices. *Anal. Sci*., 22: 157-159.

Jenkins, J.A. 1948. The origin of the cultivated tomato. *Econ. Bot.*, 2: 379-392.

Jones, E. and R.E. Hughes. 1983. Foliar ascorbic acid in some angiosperms. *Phytochemistry*, 22(11): 2493-2499.

Li, W. and R. Chetelat. 2010. A pollen factor linking inter- and intraspecific pollen rejection in tomato. Science, 330: 1827-1830.

Li, W., S. Royer and R. Chetelat. 2010. Fine mapping of ui6.1, a gametophytic factor controlling pollen-side unilateral incompatibility in interspecific *Solanum* hybrids. *Genetics*, 185(3): 069-1080.

Lim, M.Y., B.R. Jeong, M. Jung and C.H. Harn. 2016. Transgenic tomato plants expressing strawberry D-galacturonic acid reductasegene display enhanced tolerance to abiotic stresses. *Plant Biotechnol. Rep.*, 10(2): 105-116.

Lin, D., Y. Xiang, Z. Xian and Z. Li.2016. Ectopic expression of *SlAGO7* alters leaf pattern and inflorescence architecture and increases fruit yield in tomato. *Physiol. Plantarum doi*: 10.1111/ppl.12425.

Lin, T., G. Zhu, J. Zhang, X. Xu, Q. Yu, Z. Zheng, Z. Zhang, Y. Lun, S. Li, X. Wang, Z. Huang, J. Li, C. Zhang, T. Wang, Y. Zhang, A. Wang, Y. Zhang, K. Lin, C. Li, G. Xiong, Y. Xue, A. Mazzucato, M. Causse, Z. Fei, J.J. Giovannoni, R.T. Chetelat, D. Zamir, T. Städler, J. Li, Z. Ye, Y. Du and S. Huang. 2014. Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.*, 46: 1220-1226.

Martin, G.B., A. Frary, T. Wu, S. Brommonschenkel, J. Chunwongse, E.D. Earle and S.D.Tanksley. 1994. A member of the tomato *Pto* gene family confers sensitivity to fenthion resulting in rapid cell death. *The Plant Cell*, 6: 1543-1552.

Ohlson, E.W. and M.R.Foolad. 2016. Genetic analysis of resistance to tomato late blight in *Solanum pimpinellifolium* accession PI 163245. *Plant Breed.*, doi: 10.1111/pbr.12366.

Parisi, M., R. Aversano, G. Graziani, V. Ruggieri, V. Senape, L. Sigillo and A. Barone. 2016. Phenotypic and molecular diversity in a collection of 'Pomodoro diSorrento' Italian tomato landrace. *Scientia Horticulturae*, 203: 143-151.

Qu, H., J.S. Rylosona and L. Zhao. 2015. Cytological observation of *Solanum pimpinellifolium* L. microspore development. *Pak. J. Bot*., 47(4): 1459-1465.

Rick, C.M. 1986. Germplasm resources in the wild tomato species. *Acta Hortic*., 190: 39-47.

Rick, C.R. 1951. Hybrids between *Lycopersicon esculentum* Mill. and *Solanum lycopersicoides* Dun. *Proc. Natl. Acad. Sci.*, 37: 741-745.

Rick, C.R. and R.T. Chetelat. 1995. Utilization of related wild species for tomato improvement. *Acta Hortic.*, 412: 21-35.

Rodríguez, G.R., S. Muños, C. Anderson, S.C. Sim, A. Michel, M. Causse, B.B. Gardener, D. Francis and E. van der Knaap. 2011. Distribution of *SUN*, *OVATE*, *LC*, and *FAS* in the tomato germplasm and the relationship to fruit shape diversity. *Plant Physiol*., 156: 275-285.

Saeed, S., M.Y.K. Barozai, A. Ahmed, R.B.Tareen and S.B.G.M. Ali. 2017. Impact of ecological diversity on genetic and phytochemical variation injuniperusexcelsa from high elevation zones of quetta valley, Pakistan. *Pak. J. Bot*., 49(1): 201-206.

Sagor, G.H.M., T. Berberich, S. Tanaka, M. Nishiyama, Y. Kanayama, S. Kojima, K. Muramoto and T. Kusano. 2016. A novel strategy to produce sweeter tomato fruits with high sugar contents by fruit-specific expression of a single bZIP transcription factor gene. *Plant Biotechnol. J.*, 14: 1116-1126.

Schauer, N., D. Zamir and A.R. Fernie. 2005. Metabolic profiling of leaves and fruit of wild species tomato: A survey of the *Solanum lycopersicum*complex. *J. Exp. Bot.*, 56: 297-307.

Uluisik, S., N.H. Chapman, R. Smith, M. Poole, G. Adams, R.B. Gillis, T.M.D. Besong, J. Sheldon, S. Stiegelmeyer, L. Perez, N. Samsulrizal, D. Wang, I.D. Fisk, N. Yang, C. Baxter, D. Rickett, R. Fray, B. Blanco-Ulate, A.L.T. Powell, S.E. Harding, J. Craigon, J.K.C. Rose, E.A. Fich, L. Sun, D.S. Domozych, P.D. Fraser, G.A. Tucker, D. Grierson and G.B. Seymour. 2016. Genetic improvement of tomato by targeted control of fruit softening. *Nat. Biotech*., doi:10.1038/nbt.3602.

Yaghoobi, J., J.L. Yates and V.M. Williamson. 2005. Fine mapping of the nematode resistance gene *Mi-3* in *Solanum peruvianum* and construction of a *S. lycopersicum* DNA contig spanning the locus. *Mol. Gen. Genomics*, 274: 60-69.

Zamir, D. 2001. Improving plant breeding with exotic genetic libraries. *Nat. Rev. Genet.*, 2: 983-989.

Zeinab Ibrahim, E.S. 2016. Tomato breeding for heat stress conditions. *E. J. A. E.*, 3(2): 87-93.

(Received for publication 13 September 2016)