Cite this poster as:
Baldo A.M., W.F Lamboy, L.D. Robertson, S.M. Sheffer, and J.A. Labate. 2007. The distribution of genetic variation in cultivated tomato. Plant and Animal Genome XV, San Diego, CA.

Abstract

Cultivated tomato (*Solanum lycopersicum* L.) is known to have a narrow genetic base. COSII, EST-based, and several loci related to fruit quality traits were resequenced in a diverse panel of 30 Plant Genetic Resources Unit (PGRU) tomato accessions and line TA496. The majority of sampled tomato accessions represented the primary center of diversity (Peru, Chile, and Ecuador), and countries contiguous with the primary center. These were the same accessions studied by Villand et. al. (1998) using RAPDS.

Original collections were made between 1932 and 1976. Evidence of historical introgression and the population-level distribution of genetic variation reveal relationships between tomato landraces. There was the most genetic variation among the samples collected in the primary center of diversity and the least from secondary centers of diversity, as expected.

Poster 170

PAG-XV
San Diego, CA
January 2007

The Distribution Of Genetic Variation In Cultivated Tomato

Angela M Baldo, Warren F Lamboy, Larry D Robertson, Susan M Sheffer, Joanne A Labate
USDA - ARS Plant Genetic Resources Unit, 630 W. North St., Geneva, NY 14456
http://www.ars.usda.gov/NAA/Geneva

USDA Agricultural Research Service

Abstract / Introduction

Cultivated tomato (*Solanum lycopersicum* L.) is known to have a narrow genetic base. COSII, EST-based, and several loci related to fruit quality traits were resequenced in a diverse panel of 30 Plant Genetic Resources Unit (PGRU) tomato accessions and line TA496. The majority of sampled tomato accessions represented the primary center of diversity (Peru, Chile, and Ecuador), and countries contiguous with the primary center. These were the same accessions studied by Villand et. al. (1998) using RAPDS.

Original collections were made between 1932 and 1976. Evidence of historical introgression and the population-level distribution of genetic variation reveal relationships between tomato landraces. There was the most genetic variation among the samples collected in the primary center of diversity and the least from secondary centers of diversity, as expected.

Results

Resequencing fragments from 9 fruit quality genes, 11 Cos II markers, and 20 EST-based markers yielded 100 polymorphic sites which were coded for principal components analysis (Fig 1). The samples sequenced represented plants from a range of locations, from the South America to Europe and Asia (Fig 2). Fruits were assayed for cherry (cerasiforme) morphology according to the number of locules (2), size (1.5-2.5cm), and shape (spherical). Only one accession (#25, PI 127825) fit this profile.

Accessions were plotted along the first 3 Principal Components (PCs, Fig 3). As expected, the accessions from South America showed the greatest diversity, and the accessions from Europe and Asia the least. A breeding line (TA496) with introgressions from Peru Wild, *S. pimpinellifolium*, and *S. peruvianum* in its pedigree is the most divergent of all.

The contribution of markers to the PCs were small. Polymorphisms suspected to be due to historical introgressions did not appear to contribute much to the first PC, and mostly to the second PC if at all (Fig 1).

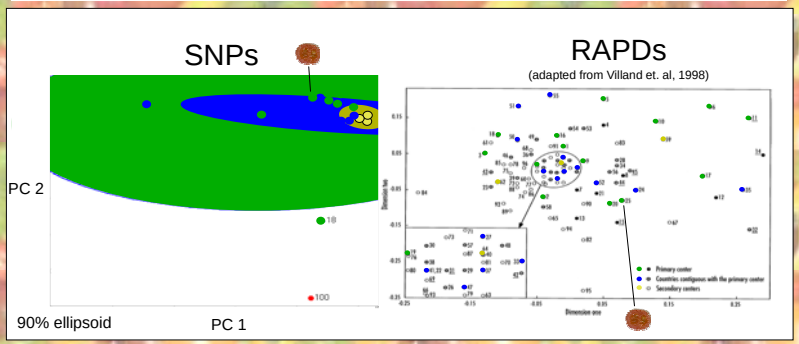
1. PC Loadings by Marker

Fruit Quality		PC1	PC2	PC3
11_pds		-0.06	-0.01	0.02
12_psy1	+	-0.06	-0.01	0.02
14_hp2		-0.04	-0.02	-0.15
14_hp2_ex2		0.01	-0.01	0.07
15_fw2.2		0.01	-0.07	0.00
15_fw2.2.1		-0.04	0.00	0.10
15_fw2.2.2		-0.04	0.00	0.10
15_fw2.2.3		0.01	-0.07	0.00
16_TG11		-0.05	-0.02	-0.19
16_TG11.1		-0.26	0.02	0.00
16_TG11.2		0.01	0.02	-0.02
16_TG11.3		-0.26	0.02	0.00
17_nor		-0.06	-0.01	0.02
17_nor.1		-0.16	0.04	0.10
17_nor.2		-0.04	0.02	0.01
18_CRTISO		-0.07	0.00	-0.22
18_CRTISO.1		-0.02	0.01	-0.12
18_CRTISO.2		-0.07	0.00	-0.22
19_rin		-0.26	0.04	-0.07
19_rin.1		-0.22	0.05	0.03
20_PTOX		-0.05	0.03	0.06
20_PTOX.1		-0.16	0.04	0.12
20_PTOX.2		-0.07	0.02	0.07
20_PTOX.3		-0.19	0.02	0.06
COS II				
10_1		-0.16	0.01	0.11
11_1		-0.06	-0.03	0.04
11_2		-0.04	-0.02	-0.15
11_3		-0.10	-0.05	-0.12
12_1		0.00	0.13	-0.02
12_2		0.13	0.10	-0.12
12_4		-0.09	0.05	0.07
12_5		0.00	0.13	-0.02
12_6		0.04	0.11	-0.03
12_7		0.00	0.13	-0.02
12_8		-0.07	0.01	0.02
15_1		-0.06	-0.01	0.02
16_1		-0.01	-0.04	0.12
16_2		-0.17	0.00	0.24
18_1		0.00	0.02	-0.04
18_2		-0.19	0.06	0.16
18_3		-0.24	0.03	-0.02
18_4		-0.04	-0.02	-0.15
18_5		-0.23	0.04	0.01
18_6		0.00	0.02	-0.04
2_1		0.01	0.01	0.00
3_1		-0.11	0.04	0.14
4_1	+	-0.19	0.01	0.15
5_1		0.01	-0.01	0.07
7_1		0.01	0.01	0.02
7_3		-0.02	0.05	0.18
7_4		-0.04	0.04	0.08
7_5		-0.03	0.03	0.16
7_6		-0.04	0.04	0.08

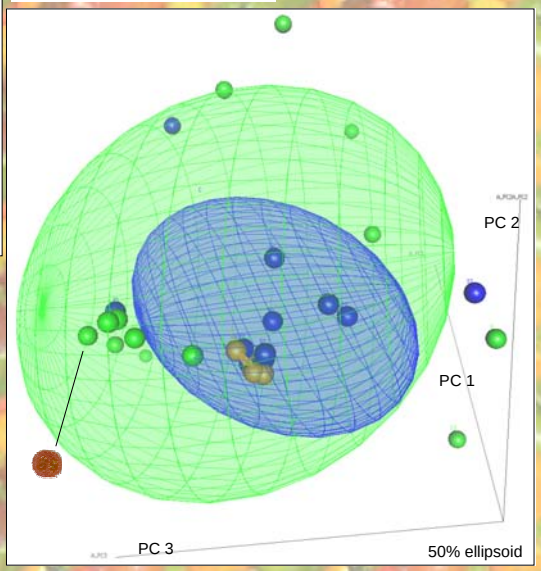
EST-Based		PC1	PC2	PC3
1		0.00	0.13	-0.02
10		0.04	0.15	0.13
13		0.13	0.13	0.04
19		-0.16	0.03	-0.13
20		-0.07	0.11	0.17
20.1		0.07	0.08	-0.07
21		-0.19	0.07	0.11
21.1		-0.04	-0.02	-0.15
23		-0.07	0.12	-0.06
24		0.00	0.10	-0.04
24.1		-0.21	0.14	-0.20
24.2		-0.21	0.14	-0.20
25	+	0.00	0.13	-0.02
28b		-0.17	0.01	-0.11
28b.1		-0.05	-0.01	-0.23
3		0.03	0.16	0.05
30	+	0.00	0.13	-0.02
30.1	+	0.07	0.23	0.12
30.2	+	0.00	0.13	-0.02
30.3	+	-0.07	-0.09	-0.15
30.4	+	0.00	0.13	-0.02
34	+	-0.04	-0.02	-0.15
35b		-0.06	-0.01	0.02
35b.1		0.01	0.14	-0.03
35b.2		-0.06	-0.01	0.02
35b.3		0.00	0.21	-0.02
35b.4		0.18	0.21	0.06
35b.5		0.07	0.22	-0.04
35b.6		0.00	0.21	-0.02
35b.7		-0.08	-0.02	0.04
35b.8		0.00	0.21	-0.02
4		-0.01	0.01	-0.08
40b		0.05	0.12	-0.10
40b.1		0.06	0.13	-0.11
5		0.01	0.01	0.00
5.1		-0.02	0.22	-0.05
5.2		0.04	0.23	-0.08
5.3		-0.02	0.22	-0.05
7		-0.01	0.01	-0.08
7.1		-0.05	-0.01	-0.23
7.2		0.07	0.06	0.03
9b	+	0.00	0.13	-0.02
9b.1	+	0.07	0.23	0.12
9b.2	+	0.00	0.13	-0.02
9b.3	+	0.07	0.26	0.01
9b.4	+	0.00	0.13	-0.02

+ Possible introgression

3A. Landrace Relationships



3B. Location of cerasiforme accession in 3D



2. Tomato Landraces

Primary Center of Domestication

1	PI	124037	Chile	1937
2	PI	128586	Chile	1938
3	PI	128592	Chile	1938
5	PI	129026	Ecuador	1938
6	PI	129033	Ecuador	1938
9	PI	129142	Ecuador	1938
10	PI	258474	Ecuador	1959
11	PI	390510	Ecuador	1974
16	PI	99782	Peru	1932
17	PI	124035	Peru	1937
18	PI	155372	Peru	
19	PI	159009	Peru	1947
20	PI	258478	Peru	1959
25	PI	127825	Peru	1938

AVRDC says "Bolivia"
(Contiguous in Villand et. al)
The only Cherry (cerasiforme) in the dataset

Contiguous with Primary Center

22	PI	97538	Argentina	1932
24	PI	127820	Bolivia	1938
27	PI	117563	Brazil	1936
33	PI	129084	Colombia	1938
35	PI	212062	Costa Rica	1954
37	PI	272703	Guatemala	1961
47	PI	270408	Mexico	1960
50	PI	270430	Mexico	1960
51	PI	196297	Nicaragua	1951
52	PI	406952	Nicaragua	1976
55	PI	129128	Panama	1938
58	PI	118783	Venezuela	1996

Secondary center of Domestication

59	PI	125831	Afghanistan	1937
62	PI	158760	China	1947
64	PI	98097	Cuba	1932
97	PI	262995	Netherlands	1960

Introgressed breeding line

100	TA 496		1998

Discussion

There is debate whether cherry tomatoes carry more of the representative genetic diversity of ancestral *S. lycopersicum* than the usual "esculentum" type plants. The only cherry tomato in this dataset (#25, Fig 2), is less divergent than esculentum-type accessions from the primary center of domestication, #5 and #20 (Fig 3).

In comparison with the Multidimensional Scaling (MDS) of RAPD data published by Villand et. al. (1998) using the same accessions, SNP data show a more marked similarity among accessions from secondary centers of diversity. Likewise, the accessions from the primary center of domestication are more scattered.

There is a question about the location where PI 127825 was collected. The passport data in the USDA-ARS Germplasm Resources Information Network (GRIN) shows this accession is from Peru, while the Asian Vegetable Research and Development Center (AVRDC) lists the sample as from Bolivia. Reexamination of the original source data confirms this accession is from Peru. Given how diverse this sample is compared with the cloud of other accessions from contiguous countries, it seems likely that it may have been collected in a primary center of domestication.

Materials and Methods

Polymorphisms (SNPs and indels) were scored as either homozygote or heterozygote (intermediate). Adjacent polymorphisms displaying the same haplotype were scored as one to prevent large haplotype blocks from contributing disproportionately. The R Statistical Computing Language (www.r-project.org) was used to conduct Q-type PCA and generate the PCA plots in Fig 3.

References

Labate JA, and Baldo AM. 2005. **Tomato SNP discovery by EST mining and resequencing**. Mol Breed 16:343-349.
Mueller LA et. al. 2005. **The SOL Genomics Network. A Comparative Resource for Solanaceae Biology and Beyond**. Plant Phys 138:1310-1317
Villand J et. al. 1998. **Genetic Variation among Tomato Accessions from Primary and Secondary Centers of Diversity**. Crop Sci 38:1339-1347.

Acknowledgements

The authors gratefully thank Jim Giovannoni, Steve Tanksley and Feinan Wu for sharing primer sequences, David Spooner, David Francis, and Ed Buckler for stimulating discussion. We thank Paul Kisly for assistance with growing the plants and Teri Ferrin for fruit photos.