

Cite this poster as:

Baldo AM, Wan Y, Lamboy WF, Simon CJ, Labate JA, Sheffer SM. 2007. SNP validation and genetic diversity in cultivated tomatoes and grapes. Plant and Animal Genome XV, San Diego, CA.

Abstract

Cultivated grapes and tomatoes have very different reproductive and propagation systems. While grapes are typically outcrossed and are grafted once a variety is defined, tomatoes are generally selfed and are propagated by seed. Large-scale public EST datasets were used in both crops to predict SNPs and PCR primers flanking these SNPs. Genomic DNA was amplified and the gene fragments were resequenced. These results were analyzed in each crop for the distribution of genetic diversity among cultivars and within genes. Frequency of intron discovery, SNP validation rates, and proportions of rare alleles may reflect the different propagation histories of these crops.

SNP Validation And Genetic Diversity In Cultivated Tomatoes And Grapes



PAG-XV
San Diego, CA
January 2007

Angela M Baldo^{1,2}, Yizhen Wan^{2,3}, Warren F Lamboy¹, Charles J Simon^{1,2}, Joanne A Labate^{1,2}, Susan M Sheffer¹

1.USDA-ARS Plant Genetic Resources Unit 630 W. North St. Geneva, NY 14456 USA
2.Cornell University Department of Horticultural Sciences 630 W. North St. Geneva, NY 14456 USA
3.Northwest Sci-Tech University of Agriculture and Forestry Shaanxi Province, Yangling, China
<http://www.ars.usda.gov/NAA/Geneva>

30 Tomato Accessions

PI 124037	Chile	1937	
PI 128586	Chile	1938	
PI 128592	Chile	1938	
PI 129026	Ecuador	1938	
PI 129033	Ecuador	1938	
PI 129142	Ecuador	1938	
PI 258474	Ecuador	1959	
PI 390510	Ecuador	1974	
PI 99782	Peru	1932	
PI 124035	Peru	1937	
PI 155372	Peru		
PI 159009	Peru	1947	
PI 258478	Peru	1959	
PI 127825	Peru	1938	cerasiforme
PI 97538	Argentina	1932	
PI 127820	Bolivia	1938	
PI 117563	Brazil	1936	
PI 129084	Colombia	1938	
PI 212062	Costa Rica	1954	
PI 272703	Guatemala	1961	
PI 270408	Mexico	1960	
PI 270430	Mexico	1960	
PI 196297	Nicaragua	1951	
PI 406952	Nicaragua	1976	
PI 129128	Panama	1938	
PI 118783	Venezuela	1996	
PI 125831	Afghanistan	1937	
PI 158760	China	1947	
PI 98097	Cuba	1932	
PI 262995	Netherlands	1960	Ailsa Craig

Abstract

Cultivated grapes and tomatoes have very different reproductive and propagation systems. While grapes are typically outcrossed and are grafted once a variety is defined, tomatoes are generally selfed and are propagated by seed. Large-scale public EST datasets were used in both crops to predict SNPs and PCR primers flanking these SNPs. Genomic DNA was amplified and the gene fragments were resequenced. These results were analyzed in each crop for the distribution of genetic diversity among cultivars and within genes. Frequency of intron discovery, SNP validation rates, and proportions of rare alleles may reflect the different propagation histories of these crops.

Results

In both crops gene fragments were selected where publically available ESTs showed SNPs between two or more cultivars. Primers were designed to bracket the predicted SNPs and were used to amplify that region of the genome. In a number of cases introns were discovered between the primers.

The datasets are similar in that tomato landraces correspond to cultivated *S. lycopersicum*, while *Vitis sylvestris* is considered to be the wild form of *Vitis vinifera*. In both cases the accessions used represent the range of wild and cultivated genetic diversity. As expected, overall grapes show a greater amount of genetic diversity than the tomatoes, as tomato is known for its lack of sequence diversity.

Avg per 100 bp	Tomato	Grape	Significance (one-way Anova)
SNPs, Total	0.87	2.12	1.49E-05
SNPs, Exon	0.60	1.33	5.83E-03
SNPs, Nonsynonymous	0.29	0.43	3.37E-01
SNPs, Intron	1.74	3.23	9.47E-02
SNPs, 3' UTR	0.27	2.59	3.77E-03
Indels, Total	0.14	0.24	2.25E-01
Indels, Exon	0.04	0.02	6.97E-01
Indels, Frameshift	0.00	0.02	2.25E-01
Indels, Intron	0.34	0.57	5.68E-01
Indels, 3' UTR	0.06	0.35	3.12E-02

Discussion

While roughly the same number of genes were sequenced, yielding a similar number of nucleotides, there were a number of differences in the results.

Many more introns were found among the tomato amplicons (42%) relative to those of grape (13%). Much more of the grape sequence was comprised of 3' UTR (32%) than the tomato sequence (12%). The frequency of tomato introns is similar to what is expected using the same methodology in *Arabidopsis* (data not shown).

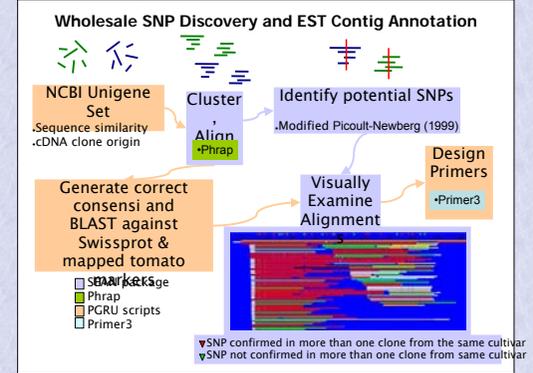
The lower frequency of introns in grapes may be due to a difference in the composition of the ESTs from which the primers were designed. If the grape ESTs were derived mostly from the 3' portion of cDNA, this might explain why there would be more 3' UTR and fewer interrupted exons represented.

The overall frequency of SNPs was much higher in grape, as expected. This may have been the result of the higher SNP frequencies found in exons and 3'UTR. There may not have been enough intron sequence available in the grape ESTs to establish significance.

Indels are an order of magnitude less abundant than SNPs. The only significant difference in frequency between the crops is in the 3' UTR, perhaps due to UTR rich bias of the grape ESTs.

Each crop is at the far end of the continuum between low and high genetic diversity, yet in both cases this EST-based method for assaying diversity was accomplished.

	Tomato		Grape	
No. Accessions	30		31	
	<i>S. lycopersicum</i> (29 esculentum, 1 cerasiforme)		(12 <i>V. vinifera</i> , 19 <i>V. sylvestris</i>)	
No. Gene fragments sequenced	26		30	
No. Gene fragments containing exons	21	81%	28	93%
No. Gene fragments containing introns	11	42%	4	13%
No. Gene fragments containing 3' UTR	10	38%	20	67%
No. Gene fragments, SNPs confirmed	20	77%	30	100%
No. Gene fragments, Indels found	8	31%	16	53%
Total No. SNPs found	103		273	
No. SNPs in exon	24	23%	104	38%
No. Exon SNPs that are Synonymous	14	58%	71	68%
No. SNPs in intron	76	74%	36	13%
No. SNPs in 3' UTR	3	3%	133	49%
Total No. Indels found	18		30	
Exons	1	6%	2	7%
Introns	16	89%	7	23%
3' UTR	1	6%	21	70%
Total Nucleotides	9335		12624	
Exon	3775	40%	7471	59%
Intron	4455	48%	1131	9%
3' UTR	1105	12%	4022	32%



31 Grape Accessions

DVIT2009	V sylvestris	1993	France	
DVIT2011	V sylvestris	1993	France	
DVIT2012	V sylvestris	1993	France	
DVIT2014	V sylvestris	1993	France	
DVIT2015	V sylvestris	1993	France	
DVIT2016	V sylvestris	1993	France	
DVIT2017	V sylvestris	1993	France	
DVIT2018	V sylvestris	1993	France	
DVIT2019	V sylvestris	1993	France	
DVIT2020	V sylvestris	1993	France	
DVIT2021	V sylvestris	1993	France	
DVIT2022	V sylvestris	1993	France	
DVIT2023	V sylvestris	1993	France	
DVIT2426-1	V sylvestris	1992	Tunisia	
DVIT2426-2	V sylvestris	1992	Tunisia	
DVIT2426-4	V sylvestris	1992	Tunisia	
DVIT2426-5	V sylvestris	1992	Tunisia	
DVIT2426-6	V sylvestris	1992	Tunisia	
DVIT2426-7	V sylvestris	1992	Tunisia	
PI 173295	V vinifera	1949	Italy	Syrach
	V vinifera		USA	Red_Globe
DVIT1095	V vinifera	1983	USA	Flame_Seedless
DVIT1307	V vinifera	1983	USA	Midget Thompson Seedless
DVIT2339	V vinifera	1989	Germany	Rotberger
DVIT3067	V vinifera	2002	Switzerland	Bondola
DVIT376	V vinifera	1983	Europe	Chasselas Rose
DVIT677	V vinifera	1983	France	Cabernet_Sauvignon
DVIT688	V vinifera	1983	France	Chardonnay
DVIT826	V vinifera	1983	France	Merlot
DVIT915	V vinifera	1983	France	Pinot_Noir
DVIT947	V vinifera	1983	France	Sauvignon_Blanc

References

Labate JA, and Baldo AM. 2005. **Tomato SNP discovery by EST mining and resequencing.** Mol Breed 16:343-349.
Huntley D, Baldo AM, Johri S, Sergot M. 2006. **SEAN: SNP prediction and display program utilizing EST sequence clusters.** Bioinformatics 22(4):495-496

Acknowledgements

We would like to thank the USDA-ARS germplasm repository in Davis, CA for grape material, Larry Robertson at the USDA-ARS germplasm repository in Geneva, NY for tomato material, and Peter Cousins for helpful discussions.