

SolCAP Solanaceae Coordinated Agricultural Project

Translating Solanaceae Sequence Diversity and Trait Variation into Applied Outcomes Through Integrative Research, Education, and Extension : Initiation Date October 2008

Project Co-Directors: David Douches¹, Robin Buell¹, David Francis², Allen Van Deynze³, Walter De Jong⁴, Lukas Mueller⁴ and Alexandra Stone⁵, Project Assistant: Kelly Zarka¹

1: Michigan State University, 2: Ohio State University, 3: University of California–Davis, 4: Cornell University, 5: Oregon State University



SolCAP PROJECT OVERVIEW

The Basis of SolCAP:

Leverage knowledge and resources from the *Solanaceae* species, potato and tomato, to positively impact applied breeding and crop quality across traditional commodity boundaries.

Potato and Tomato:

- Are the two most economically important species in *Solanaceae*
- Annually account for \$6.3 Billion in farm value
- The United States is one of the world leading producers

SolCAP Research:

High Value Traits Targeted:

- Carbohydrate
- Sugar Content
- Vitamin Content



Primary Research Objective:

- Focusing on elite breeding material
- Provide infrastructure to link allelic variation in genes to valuable traits in cultivated germplasm of potato and tomato
- Illumina Infinium and Illumina OPA based Platforms

Education and Extension:

SolCAP's extension component involves creation of a web community within the eXtension.org Community of Practice through which plant breeders, basic scientists, agricultural professionals, extension specialists and others can publish content and network.

Education:

- Scientific workshops
- Development of curricula for graduate students
- Breeder practicum

Solanaceae Genome Network:

- Integrated genomic and phenotypic databases
- Solanaceae Genome Network (GN) website database. <http://sgn.cornell.edu/>

SolCAP PROJECT UPDATE

The SolCAP project website includes personnel information, project newsletters, meetings and workshop information, databases, tools, resources, project objectives and downloads. Visit us at <http://solcap.msu.edu>.

To date in the project, we have published 7 newsletters that have reached over 400 people in the *Solanaceae* community.



SolCAP Germplasm Panels

Potato:

Elite potato germplasm contributions were made from 16 programs across the U.S. and 6 international programs. The panel consists of 320 potato lines currently used by the community and are being SNP genotyped and phenotyped across 3 different environments. Phenotypic and genotypic data will be a community resource. The list of germplasm and initial phenotype data can be found on the project website at: <http://solcap.msu.edu/potato.shtml>

Tomato:

The core collection of tomato germplasm has been assembled which includes 478 inbred lines from fresh market and processing tomato breeding programs as well as landraces, heirloom varieties, and wild accessions. This includes 48 landraces representing geographic range, diversity of fruit shapes and genetic variation. 48 Heirloom varieties have been selected representing the major fruit morphologies. Wild accessions have also been added to provide insight into loci that have been introgressed into cultivated backgrounds. The list of germplasm and initial phenotype data can be found on the project website at: <http://solcap.msu.edu/tomato.shtml>

SNP Discovery in Elite Tomato and Potato Germplasm

From existing sequence databases, we have identified potato and tomato sequences with candidate Single Nucleotide Polymorphisms (SNPs). These can be accessed through the project website under the tomato and potato pages.

To expand our SNP discovery, we have generated 1.8-2.1 Gb of potato sequence data from normalized cDNA from three different varieties (Table 1). For the tomato, SolCAP has generated ~3 Gb of sequence from each of six varieties (Table 1).

Using normalized cDNA libraries and the Illumina Genome Analyzer 2 (GA2) platform, we were able to generate 19.6 and 7.0 Gb of sequence for tomato and potato, respectively, increasing by 165- and 60-fold the amount of transcriptome sequence for these two species. To reduce the redundancy in the GA2 derived transcript sequences, we performed *de novo* assembly of the quality filtered reads. For the GA2-generated transcript data, *de novo* assembly resulted in a combined total of 195 and 86.9 Mb of contigs for tomato and potato, respectively. For tomato, the six accessions had assembled transcriptome size ranging from 30.6 to 34.9 Mb. For potato, the three accessions had assembled transcriptome size ranging from 28.6 to 29.4. The narrow range of assembled transcriptome sizes within the potato and tomato GA2-generated datasets suggest that the underlying cDNA populations and the sequencing and assembly process were similar within the potato and tomato samples.

Infinium Platform Development for Tomato and Potato

The GA2 sequences were used in conjunction with the potato and tomato reference genomes (made available by the Potato Genome Sequencing Consortium and the Tomato Genome Initiative, respectively) to identify SNPs within the sequenced

Table 1. Elite potato and tomato germplasm sequenced.

Species	Cultivar	Market Class
<i>S. tuberosum</i>	Atlantic	Chip processing
<i>S. tuberosum</i>	Premier Russet	Processing-French Fry
<i>S. tuberosum</i>	Snowden	Chip processing
<i>S. lycopersicon</i>	FL7600	Fresh market
<i>S. lycopersicon</i>	NC84173	Fresh market
<i>S. lycopersicon</i>	OH9242	Processing
<i>S. lycopersicon</i>	OH08-6405	Fresh market
<i>S. lycopersicon</i>	PI114490	Cherry
<i>S. pimpinellifolium</i>	PI212816	Wild

accessions. To design the potato and tomato Infinium SNP platforms, we applied a stringent set of filters to remove sequencing errors, alignment errors, low coverage issues, closely spaced SNPs, tri-allelic/tetra-allelic SNPs and those SNPs with low Infinium SNP scores. After filtering, 69,011 total candidate SNPs from the three potato transcriptomes and 28,830 total candidate SNPs from the six tomato transcriptomes were available for design of the potato and tomato Infinium platforms. A set of 96 SNPs from both the potato and tomato candidate Infinium SNP sets were used in a BeadXpress assay to validate our SNP predictions. The validation rate for potato was ~85% and 97% for tomato.

Candidate genes and existing SNPs in use within the tomato and potato communities were requested for inclusion on the Infinium platforms to ensure cross-comparison with existing maps and datasets. Using these, plus SNPs identified from the SolCAP GA2 datasets, and existing tomato/potato Sanger ESTs, we collated a set of SNPs for tomato and potato that represent high priority genes of interest, leverage existing marker data and datasets, and provide wide coverage of the respective genomes.

The final design for the potato array includes 520 SNPs within a known genetic marker, 3,243 SNPs in candidate genes from the community, and 93 SNPs within a known genetic marker and candidate gene. The remainder of the SNPs are distributed over 650 Mb of the potato genome. After validation, the final potato Infinium SNP platform provides data on 8,303 markers.

The final design for the tomato array includes 1,470 SNPs from the community, 567 SNPs in candidate genes, and 4,548 SNPs from the six GA2 transcriptomes. The remainder of the SNPs that provide coverage of the tomato genome. After validation, the final tomato Infinium SNP platform provides data on 7,527 markers.

Calling SNPs in Tetraploid Potato Using the Infinium Platform

The Infinium platform is designed for diploid species yet most cultivated potato is tetraploid in nature. To assess and permit calling of all five genotypes when present in tetraploid potato (Figure 1), we genotyped three potato populations (4X mapping population, 2X mapping population, and germplasm panel) with the SolCAP Infinium array and scored the lines manually to generate not only genotypes but also the cluster files for use by the community.

In total, 7,412 (89.3%) of the potato Infinium markers could be called. The failures (1,187, 10.8%) were due to bad markers (341, 4.1%), questionable (296, 3.6%), and questionable segregation (254, 4.1%). The genotype call rate across the samples within our three benchmarking populations was high: for 7,661 (92.3%) of the "good markers", >90% of the samples in the three benchmarking populations could be genotyped. Cluster files for the SolCAP Potato Infinium Array are available at: http://solcap.msu.edu/potato_infinium.shtml

Tetraploid Segregation

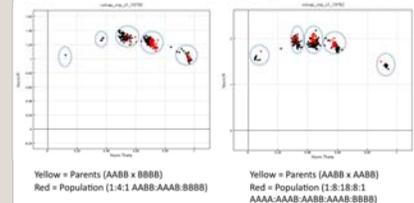


Figure 1. Examples of SNP genotyping in tetraploid potato for SolCAP Infinium Potato Array. The five clusters for two sample SNPs are shown thereby identification of the 5 possible genotypes in a tetraploid population.

Education and Extension

SolCAP has led an effort bringing researchers and educators from America's land-grant universities, government agencies and industry together to provide a comprehensive resource aimed at putting basic research on crop genomes into practice through traditional plant breeding to improve crops. The resource is at eXtension (pronounced e-Extension), http://www.extension.org/plant_breeding_genomics and includes articles, videos, tutorials, and more, provided free with 24/7 access.

SolCAP has also developed PBGworks as an online hub for the eXtension effort. It allows groups to form and online tools for collaboration. You can find it at <http://pbgworks.org/>.



Additionally, SolCAP has completed tomato and potato community surveys and the information obtained was used to determine the baseline knowledge in breeding and marker technologies. This information helped to formulate annual SolCAP workshops.

SolCAP Workshops/Webinars

Topics included are:

- 1) Marker technologies
- 2) Analyzing quantitative trait loci
- 3) Marker assisted selection in tomato or potato
- 4) Effects of population structure on genetic analysis



SolCAP provided workshops/webinars designed specifically for breeders in conjunction with The Potato Association of America 94rd Meeting and the 2010 Tomato Disease Workshop. The webinars were recorded and can be downloaded and viewed at our SolCAP website: <http://solcap.msu.edu/meetingsworkshops.shtml>



United States
Department of
Agriculture

National Institute of
Food and
Agriculture



The SolCAP project is supported by the Agriculture and Food Research Initiative Applied Plant Genomics CAP Program of the USDA's National Institute of Food and Agriculture