

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

The SolCAP project is supported by the Agriculture and Food Research Initiative Applied Plant Genomics CAP Program of the USDA's Cooperative State Research, Education and Extension Service



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

Breeding:

- 90% of potato breeding is still conducted in the US public sector
- 8 University programs develop commercial tomato varieties

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

Innovative:

- Utilizing high throughput sequencing and bioinformatics
- Illumina Infinium and Illumina OPA based Platforms

Extension:

SolCAP's extension component involves creation of a web community within the eXtension.org Community of Practice through which plant breeders, basic scientists, seed industry professionals, 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

SolCAP has created a project website which is a centralized resource for SolCAP participants. The site includes personnel information, project newsletters, meetings and workshop information, databases, tools, resources, project objectives and downloads. Visit us at <http://solcap.msu.edu>

Within the first 9 months we have published 3 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 480 potato lines currently used by the community and they are being SNP genotyped and phenotyped across 3 different environments. Phenotypic and genetic data will be a community resource.

Tomato:

The core collection of tomato germplasm has been assembled which includes 288 inbred lines from fresh market and processing tomato breeding programs. This includes 48 landraces representing geographic range, diversity of fruit shapes and genetic variation. 48 Heifrom 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.

SolCAP cDNA Libraries Sequenced (Using an Illumina Genome Analyzer –Solexa)

SolCAP has obtained 1.8-2.1Gb of potato sequence data from normalized cDNA of cultivated germplasm (Fig. 1). For the potato transcriptome, two paired 61 bp runs have been generated from three different varieties (Atlantic, Snowden, and Premier Russet). Assembly of these sequences generated contigs totaling ~38 Mb of sequence that are currently being used for SNP discovery.

For the tomato transcriptome, SolCAP is generating sequence data from normalized cDNA and it is currently near completion.

High Throughput SNP Analysis:

From existing sequence databases, we have identified potato and tomato sequences with candidate Single Nucleotide Polymorphisms (SNPs). To identify additional SNPs, a panel of elite germplasm and mapping populations will be genotyped using Illumina, Luminex or Infinium platforms.



Figure 1. Tomato and potato varieties and tissue samples used for cDNA sequencing.

SNP Development:

Illumina Infinium Platform for 7600 potato SNPs
Illumina OPA Platform for 1536 tomato SNPs

We have identified simple sequence repeats from tomato and potato and designed primers for use by the community that can be searched, viewed and downloaded through the project website (<http://solcap.msu.edu/tools.shtml>; Figure 2). We have also identified putative SNPs in Sanger-derived potato and tomato EST collections which can be accessed through the project website (<http://solcap.msu.edu/tools.shtml>) which includes a graphical view of the multiple sequence alignment of the EST assembly to confirm depth of support for the polymorphism (Figure 3).

With respect to high confidence eSNPs, from the Sanger-derived potato ESTs, 9,620 high confidence candidate SNPs (> 4x coverage of alternate allele, no other SNP within 25 bp) could be identified in 2,332 contigs. For the Sanger-derived tomato ESTs, 1,849 high confidence SNP positions in 903 contigs were identified. Using known, validated SNPs from tomato, we were able to identify 32 contigs with 92 eSNPs in our high confidence tomato SNP data set, suggesting a sensitivity of 42% (32/77) and are currently experimentally validating our eSNP prediction pipeline from Sanger-derived ESTs. Lack of detection of



Figure 2. Simple Sequence Repeat Search Tool. A. Users can search potato and tomato sequences for Simple Sequence Repeats (SSRs) via identifier, key word, or by class of SSR. B. Sample output using "sequipota" as the annotation query. Primers flanking the SSR are predicted using Primer3. All data can be downloaded.

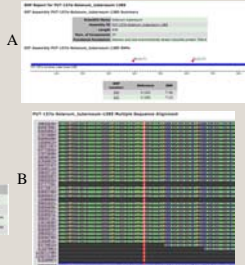


Figure 3. Single Nucleotide Polymorphism Search and Visualization Tool. A. Users can search potato and tomato sequences for SNPs via identifier or annotation. Output of a search of SNPs within a potato PlantGDB Unique Transcript (PUT; EST assembly) is shown. B. Multiple sequence alignments of the EST assemblies allow for visualization of the depth of support for the SNP.

all validated SNPs can be attributed to our high threshold for calling a SNP; the validated SNP occurs within an intron rather than the coding sequence; and/or the lack of the alternate allele in the public EST data set as not all validated SNPs were designed using only publicly available EST data.

We have modified our SNP detection pipeline to accommodate Illumina-generated ESTs and are currently identifying additional SNPs in potato and tomato in preparation for design of the SolCAP tomato and potato genotyping platforms. Target genes to be included in the genotyping platform include genes responsible for synthesis, degradation, and transport of carbohydrates and vitamins along with other established tomato/potato genetic markers to cross-integrate maps. Researchers can nominate genes to include on the genotyping platforms through the project website (<http://solcap.msu.edu/candidates.shtml>).

Education and Extension

SolCAP has developed **PBGworks** as a collaborative workspace developed using an open source software. <http://pbgworks.hort.oregonstate.edu/>

Target audience: The practicing plant breeder.

Our long-term goal is to market:

- Start-to-finish examples of marker-assisted selection applications
- Resource pages including protocols, software tutorials, and up-to-date contact information for companies offering genetic services
- Improved access to genetic resources through the "breeder's toolbox"

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

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

Application of Genomic Resources to Tomato Breeding Workshop.

SolCAP provided a workshop designed specifically for breeders in conjunction with The 42nd Tomato Breeders Round Table held June 28, 2009.

Application of Genomic Resources to Potato Breeding Workshop.

SolCAP is providing a workshop designed specifically for breeders in conjunction with the Potato Association of America 93rd Annual Meeting on August 9, 2009 in Fredericton, Canada.

