SolCAP
Solanaceae Coordinated Agricultural Project
Dedicated to the Improvement of the Potato and Tomato

Executive Committee:
David Douches    Walter De Jong    Robin Buell    David Francis
Alexandra Stone    Lukas Mueller    Allen Van Deynze

This project is supported by the Agriculture and Food Research Initiative Applied Plant Genomics CAP Program of USDA’s National Institute of Food and Agriculture.
OSU
David Francis
Heather Merk
Sung-Chur Sim
Troy Aldrich
Oregon State
Alex Stone
John McQueen
Roger Leigh

Cornell
Walter De Jong
Lukas Mueller
Joyce van Eck

UCD
Allen Van Deynze
Kevin Stoffel
Alex Kozik
Jeanette Martins
Jamie Shattuck

MSU
David Douches
C Robin Buell
John Hamilton
Candice Hansey
Kelly Zarka
Kim Felcher

Cedar Lake
Research Group
Michael Coe

Funding
USDA/AFRI
This project is supported by the Agriculture and Food Research Initiative Applied Plant Genomics CAP Program of USDA’s National Institute of Food and Agriculture.
Broad Impacts

• Almost complete participation of US public potato and tomato breeding communities for the germplasm panels
• Use of draft genome sequence for transcriptome alignment
• Strong interaction with the international community
• Development of public SNP arrays for tomato and potato
  – Calling of allele dosage in tetraploid genotypes
• Launch of Plant Breeding and Genomics eXtension.org
Potato has numerous SNPs - how many are needed for routine mapping in 4x crosses?

A function of desired marker density, and allele frequency. If want one simplex SNP every 10 cM – 800 markers – theory predicts need 2000 SNPs or more.

<table>
<thead>
<tr>
<th>freq of A</th>
<th>P(simplex)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.28</td>
</tr>
<tr>
<td>0.10</td>
<td>0.38</td>
</tr>
<tr>
<td>0.125</td>
<td>0.39</td>
</tr>
<tr>
<td>0.15</td>
<td>0.38</td>
</tr>
<tr>
<td>0.20</td>
<td>0.34</td>
</tr>
<tr>
<td>0.25</td>
<td>0.27</td>
</tr>
<tr>
<td>0.30</td>
<td>0.20</td>
</tr>
<tr>
<td>0.40</td>
<td>0.10</td>
</tr>
<tr>
<td>0.50</td>
<td>0.06</td>
</tr>
</tbody>
</table>

\[
P(\text{simplex}) = P(\text{Aaaa x aaaa}) + P(\text{aaaa x Aaaa}) + P(\text{AAAa x AAAA}) + P(\text{AAAA x AAAa})
\]
## Pair-wise Comparison of SNPs

<table>
<thead>
<tr>
<th>Cross</th>
<th>Ploidy</th>
<th>Non-segregating SNPs (%)</th>
<th>Segregating SNPs (%)&lt;sup&gt;z&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>W2310-3 x Kalkaska</td>
<td>4X</td>
<td>22.4</td>
<td>37.6</td>
</tr>
<tr>
<td>MSG227-2 x Jacqueline Lee</td>
<td>4X</td>
<td>16.5</td>
<td>51.8</td>
</tr>
<tr>
<td>Atlantic x Superior</td>
<td>4X</td>
<td>5.9</td>
<td>51.8</td>
</tr>
<tr>
<td>Stirling x 12601ad1</td>
<td>4X</td>
<td>25.9</td>
<td>37.6</td>
</tr>
<tr>
<td>B1829-5 x Atlantic</td>
<td>4X</td>
<td>11.5</td>
<td>18.8</td>
</tr>
<tr>
<td>BER 63 x DM1-3</td>
<td>2X</td>
<td>79.3</td>
<td>20.7</td>
</tr>
<tr>
<td>BER 83 x DM1-3</td>
<td>2X</td>
<td>78.8</td>
<td>21.2</td>
</tr>
<tr>
<td>84SD22 x DM1-3</td>
<td>2X</td>
<td>46.0</td>
<td>54.0</td>
</tr>
<tr>
<td>MCR205 x DM1-3</td>
<td>2X</td>
<td>76.7</td>
<td>23.3</td>
</tr>
<tr>
<td>DI x DM1-3</td>
<td>2X</td>
<td>85</td>
<td>15</td>
</tr>
<tr>
<td>08675-21 x 09901-01</td>
<td>2X</td>
<td>53.8</td>
<td>46.2</td>
</tr>
<tr>
<td>RH x SH</td>
<td>2X</td>
<td>59</td>
<td>41</td>
</tr>
</tbody>
</table>

<sup>z</sup>I = segregation not dependent on scoring dosage; II = segregation dependent on scoring dosage
SolICAP Community Resources

- cDNA libraries sequenced and High Throughput SNP analysis
- Germplasm Panels
  - Phenotype and genotype data
- Public potato and tomato 10K SNP arrays
- Workshops (4)
- Website
- Databases
Databases and Resources

- Integrated, breeder-focused resources for genotypic and phenotypic analysis at SGN and MSU.
  - http://solcap.msu.edu
  - http://solanaceae.plantbiology.msu.edu/
  - http://solgenomics.net/
Objective 1. Create a highly visible and innovative education program to train plant breeders, graduate students, and industry cooperators in genome-based breeding

Walter De Jong
Cornell University

With special emphasis on Solanaceous breeders, in the theory and practice of translational genomics
SolCAP Education and Extension Objectives

- Team-taught distance-learning graduate level course in translational genomics at Cornell University

- Yearly workshops (at PAA) for breeders to integrate genotype-based breeding strategies with elite germplasm

- Use eXtension.org to develop a Community of Practice for plant breeders, called Plant Breeding and Genomics, across all CAPs (Barley, Wheat, Conifer, Roseaceae, Bean, Onion)
Target audience 1: Breeders

• Annual workshops will be held at Potato Association of America / Tomato Breeder’s Roundtable meetings
  – understanding marker detection platforms
  – developing skills to interpret data (mapping, population structure, QTL detection, and statistical analyses)
  – database usage and mining, especially SGN

• Week-long internships to provide practical experience will be offered for breeders and support staff in SolCAP labs
  – $$$ for travel and accommodations are budgeted
SolCAP PAA Workshop/Webinar

- Held August 15, 2010 Corvallis, Oregon
- Hands-on computer lab format and on-line audience
- Topics
  - Potato genome analysis: Robin Buell
  - Tetraploid QTL analysis: Christine Hackett
  - Use of Illumina Genome studio: Allen Van Deynze

Note: View entire workshop on-line via the SolCAP website and eXtension.org
SolCAP Tomato Disease Workshop /Webinar

- Held November 16, 2010 Balm, Florida
- Hands-on computer lab format and on-line audience
- Topics
  - Tomato sequence data and how to use
  - Analysis of Illumina Infinium array data
  - Leveraging the Illumina array data for breeding

Note: View entire workshop on-line via the SolCAP website (coming soon) and eXtension.org
A course in marker-assisted breeding was taught at both Cornell and Ohio State this past summer

- 6 students at Cornell, 8 in Ohio

To be followed by a new course in bioinformatics and R at Cornell this Spring

- We’re quickly moving to a time when students won’t have to know much about marker technology or marker use on a small scale – ability to implement huge datasets will be key
Objective 2. Amplify outreach efforts by developing an eXtension Plant Breeding and Genomics Community of Practice (PBGCoP) to develop continuing education material aimed at practicing plant breeders, their staff, and seed industry professionals.

David Francis
Ohio State University
Obj. 2: Extension outreach

SolCAP’s definition: continuing education aimed at professional plant breeders, their staff and allied professionals.

Teams actively developing information for growers (lead by Barley CAP) and educators (lead by Deana Namouth Covert)

Workspace
http://pbgworks.hort.oregonstate.edu

Publish
www.extension.org/plant_breeding_genomics
Plant breeders, basic scientists, seed industry professionals, agricultural professionals, extension specialists and others can publish content and network.

**Target audience**: The practicing plant breeder.

**Our long-term goal** is to provide:

- 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"
PBGWorks: central resource for storing, sharing, editing, reviewing
Obj. 2: Publish

These resources are brought to you by the Cooperative Extension System and your Local Institution

MICHIGAN STATE UNIVERSITY EXTENSION

Local Extension offices near you
Learn more about this institution

Select a different institution

Home | About | Resource Areas | News | Articles | Answers | Calendar | Learning Lessons

Search

Plant Breeding and Genomics

Here are some of our featured articles and activities...

Working with Infinium Genotype Data

This video focuses on highly parallel genotyping tools, moving away from scoring sequence polymorphism as a "band on a gel."

More...

In This Resource Area

Plant Breeding and Genomics Topics

- Data Sets
- Experimental Design
- Genotyping
- Phenotyping
- Plant Genetics

Answers from our Experts

January 07, 2011

Is it true that some companies that purchase barley for the purpose of malting or brewing pay...

January 07, 2011

In The News...

January 10, 2011

Tango Mandarins to Appear in Produce Aisle

January 04, 2011

University of Illinois Research Makes Plant Breeding Easier

January 03, 2011

'Food of the Gods' Genome Sequence Could Make Finest Chocolate Better

More...

Resource Area Feeds

- Track all new content
Obj. 2: content

Entry level to Advanced

SGN Tutorials
  Marker conversion SNPs > CAPs
  Genome browser

SolICAP Data
  Illumina data (moving from bands on a gel to SNP)
  Downstream analysis (e.g. STRUCTURE)

Mixed media
  Webinars, Video, Slide-share
New Technologies Revolutionize Sequencing

Working with the Potato Genome, C. Robin Buell, Michigan State University

Author: Kelly Zanka, Michigan State University

This page shows video of Dr. C. Robin Buell’s seminar “Working with the potato genome,” first presented at the ISCoP workshop at the Potato Association of America meeting in August 2010. This webinar will introduce you to the methods used to determine the potato genome sequence, quality issues associated with draft genome sequence and annotation, and tools to access the potato genome sequence.

Presented: C. Robin Buell, Michigan State University

This resource area was created by the Plant Breeding and Genomics community.

Resource Area Feeds:
- Track all new content

In This Resource Area:
- Plant Breeding and Genomics Topics:
  - Data Sets
  - Experimental Design
  - Genotyping
  - Genotyping
  - Phenotyping
  - Plant Genetics
  - Population Development
  - Selection
  - Short Review (Jewel)
  - Statistical Analysis
  - Tutorials and Case Studies
  - Videos

Crop Specific Breeding Information:
- Apple
- Cherry
Launch party tonight 6:00 PM – 8:00 PM
Tiki Pavilian
Objective 3. Collect standardized phenotypic data across multiple environments for tomato and potato.

David Douches

Michigan State University
Germplasm Panel: A community resource

• **Collect standardized phenotypic data across multiple environments.**
  
  – Through the “cooperators guide”, we will also promote a “genomics mentality” that emphasizes increased genotypic replication and the collection of standardized and objective phenotypic data that can be linked to ontologies.
  
  – Accessible through the SolCAP and Solanaceae Genome Network (SGN) websites.
Cooperators Guide

• Worksheets for different types of data
  Index, README, Field Passport, Plot data, Shape/Size data (from TomatoAnalyzer software), Color data (from TomatoAnalyzer software), optional quality data...

• Navigation through the guide is by links (Back to Index)

• Fields are filled in through a drop down menu with restricted entry (note what happens if you enter “o” instead of “0”)

• Linked to SGN Phenotype ontologies
Potato Germplasm Panel

- **Panel (325 clones)**
  - Top 50 N. American varieties
  - Historical varieties
  - **Advanced breeding clones**
    - from every US and Canadian program
  - Non-American germplasm
  - Genetic stocks

- **Phenotypic evaluation**
  - Key traits: specific gravity, sucrose, glucose, Vitamin C, maturity, tuber shape, tuber number, etc.
  - Additional traits determined by breeding community
  - Data curated at SGN

- **Analyses**
  - Association mapping
  - Parental selection
  - Resolve population structure
Potato Clonal Germplasm Panel

- Seed increase completed at Cornell
- Flower color, tuber measurements, tuber images for QC.
- To be field tested 2 years X 3 major environments for potato production.
- De Jong (NY), Jansky (WI), Yilma (OR)
Traits being evaluated within SolCAP
10 plant plots, two replicates

- specific gravity
- chip color after cold storage
- sucrose/glucose
- skin texture
- tuber shape (l/w/h)
- eye depth
- skin color, flower color
- flesh color
- vine maturity (95, 120 dap)
- growth habit (prostrate, erect, etc)
- total yield
- heat sprouts
- internal defects

“The key three”
Potato Clonal Germplasm Panel

- Genetic stocks from:
  - Veilleux – monoploid research
  - Wang-Pruski - after cooking darkening
  - Haynes – late blight
  - Yencho – internal heat necrosis
  - Brown – antioxidants, nematode
  - Halterman – disease resistance
  - Bradeen – Mexican species (DaRT)
  - De Jong – SNP mapping
  - Douches – disease resistance
  - Bryan (SCRI) – 4x trait mapping
  - Jansky – species-haploid hybrids

- Choose crosses based upon parent polymorphism
Russet 4x Mapping Population

- **Evaluate russet mapping population traits** (Yencho (NC), Novy (ID), Sowokinos, Thill (MN), Gupta, Douches)
  - Key traits: specific gravity, sucrose, glucose, Vitamin C, maturity, tuber shape, tuber number, tuber images, etc.
  - Data in QC

- **Genetic Mapping** (Van Deynze, De Jong, Douches)
  - Genotyping 7600 SNPs

- **QTL Analysis** (Haynes)
  - Identify markers associated with key traits
  - Marker conversion

- **MAS/MAB (Marker Assisted Selection / Breeding)**
  - Validation of QTL in additional mapping populations
  - Use markers in new breeding populations
Tomato Germplasm Evaluation

Tomato
FM tomato (CA, NC, FL)
PROC tomato (OH, CA)
Heirloom tomato (OH, NC, NY)
  - with assistance from Great Lakes Vegetable Working Group.

Outcomes:

- Data being used to address “grower” specific questions
- Relationship between pH, color and lycopene
- Engagement of larger community through GLVWG
- Genetic variation in traits of interest
- Association mapping/QTL mapping
Tomato Germplasm Panel

- M82 x LA716
- Hunt 100 x LA407
- NC 23E X LA1269
- Sun1642 x LA1589
- OH 88119 x PI128216
- OH 8245 x PI365914
- (Fla7600 xPI114490) xOH9242
- OH 88119 x Ha 7998
- OH 9242 x OH 8245
- OH 9241 x OH 8245
- NAM population

- 144 Fresh Market
- 144 Processing
- 48 cerasiformae
- 40 land races
- 40 heirloom
- 64 wild species
Field Evaluation
opportunities for UG training
Phenotyping as an Objective and Archived Resource:


CAP:
- Shape size
- Color
- pH
- Brix
- Titratable acids

Collaborations:
- Vit. C (Labatte)
- Lycopene, beta-carotene, sugars (Rodrigues-Saona)
Objective 4. Develop extensive sequence data of expressed genes, and identify Single Nucleotide Polymorphism (SNP) markers distributed across the genome and associated with specific candidate genes for sugar, carbohydrate, and vitamin biosynthetic pathways.

C. Robin Buell
Michigan State University
The Potato Genome Sequencing Consortium

• The Potato Genome Sequencing Consortium (PGSC) have collaborated to sequence the genomes of two species: *Solanum tuberosum (RH)* and *Solanum phureja (DM1-3 516 R44)*.

• Potato genome assembly  http://www.potatogenome.net
### Exploiting 100 Years of Potato Breeding & 10 Years of Genomics

<table>
<thead>
<tr>
<th>Species</th>
<th>Cultivar</th>
<th>Market Class</th>
<th>Platform</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. tuberosum</em> Group</td>
<td>Bintje</td>
<td>Fresh market (FM)</td>
<td>Sanger ESTs</td>
<td>1905 release</td>
</tr>
<tr>
<td>Tuberosum</td>
<td>Kennebec</td>
<td>Fresh market (FM)</td>
<td>Sanger ESTs</td>
<td>1948 release</td>
</tr>
<tr>
<td><em>S. tuberosum</em> Group</td>
<td>Shepody</td>
<td>Processing-FF (PF)</td>
<td>Sanger ESTs</td>
<td>1980 release</td>
</tr>
<tr>
<td>Tuberosum</td>
<td>Atlantic</td>
<td>Chip Processing- CP)</td>
<td>GA2 ESTs</td>
<td>1980 release</td>
</tr>
<tr>
<td><em>S. tuberosum</em> Group</td>
<td>Snowden</td>
<td>Chip processing (CP)</td>
<td>GA2 ESTs</td>
<td>1990 release</td>
</tr>
<tr>
<td>Tuberosum</td>
<td>Premier</td>
<td>Processing –FF(PF)</td>
<td>GA2 ESTs</td>
<td>2008 release</td>
</tr>
</tbody>
</table>

**Elite North American varieties**
## What did we really get from our 2 lanes of GA2?

<table>
<thead>
<tr>
<th></th>
<th>Sanger</th>
<th>GA2</th>
<th>Premier</th>
<th>Russet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bintje</td>
<td>Kennebec</td>
<td>Shepody</td>
<td>Atlantic</td>
</tr>
<tr>
<td>Total No. sequences</td>
<td>15,866</td>
<td>83,549</td>
<td>86,341</td>
<td>36,291,638</td>
</tr>
<tr>
<td>Total No. Gb sequences</td>
<td>0.0079</td>
<td>0.0544</td>
<td>0.0543</td>
<td>2.2</td>
</tr>
<tr>
<td>No. sequences passed quality filters</td>
<td>14,588</td>
<td>78,386</td>
<td>83,611</td>
<td>30,185,186</td>
</tr>
<tr>
<td>No. of Gb of sequences passed quality filters</td>
<td>0.0077</td>
<td>0.0533</td>
<td>0.053</td>
<td>1.8</td>
</tr>
<tr>
<td>Total No. contigs &amp; singletons</td>
<td>7,510</td>
<td>25,330</td>
<td>51,459</td>
<td>NA</td>
</tr>
<tr>
<td>No. contigs</td>
<td>2,332</td>
<td>10,318</td>
<td>10,716</td>
<td>45,214</td>
</tr>
<tr>
<td>No. singletons</td>
<td>5,178</td>
<td>15,012</td>
<td>40,743</td>
<td>NA</td>
</tr>
<tr>
<td>Total No. Mb contigs &amp; singletons</td>
<td>4.27</td>
<td>19.89</td>
<td>36.33</td>
<td>29.45</td>
</tr>
<tr>
<td>No. Mb contigs</td>
<td>1.61</td>
<td>10.6</td>
<td>8.68</td>
<td>29.45</td>
</tr>
<tr>
<td>No. Mb singletons</td>
<td>2.66</td>
<td>9.29</td>
<td>27.65</td>
<td>NA</td>
</tr>
<tr>
<td>N50 contig size (bp)</td>
<td>711</td>
<td>1,097</td>
<td>847</td>
<td>1,192</td>
</tr>
<tr>
<td>Max contig size (bp)</td>
<td>2,255</td>
<td>4,081</td>
<td>2,517</td>
<td>11,317</td>
</tr>
<tr>
<td>Min contig size (bp)</td>
<td>278</td>
<td>272</td>
<td>847</td>
<td>150</td>
</tr>
</tbody>
</table>
Finding the SNPs in GA2 Sequences

Atlantic/Atlantic

Atlantic/Premier

Atlantic/Snowden

Intravarietal SNP

No SNP

Intervarietal SNP

DM Draft Build

SNP Site on Genome

Potato SNP Location Relationships
Filtering out Low Confidence SNP Calls

SNP calling output MUST be filtered remove low confidence SNPs.

Maq SNPfilter criteria

- Minimum read depth required to call a SNP = 20
- Minimum adjacent consensus quality (3bp) = 20
- Required maximum mapping quality of reads covering the SNP = 60
- Minimum consensus quality = 30
- Minimum distance from indel = 50
- Max SNPs in Window = 2
- Size of Window = 100bp

GA2: 2.2M SNPs to 575K
Further filtering was performed to meet the design criteria for the Infinium platform using the following criteria:

- Biallelic based on all available sequence
- Within exons (mapped > 95% to DM1-3 draft genome sequence); specifically, 50 bp from exon/intron junction
- Max 1 SNP within 100 bp window of candidate SNP
- Passed Illumina Scoring
  - 69,011 SNPs passed all filtering steps
- Created a BeadXpress validation array
Validation

High Resolution Melting
- Tested: 48 primers
- Validation (75%)

GoldenGate BeadXpress
96 x 480 samples
Selected 32 SNPs total per variety (96 total):

-0.98 - 1.0 High Infinium design score (11 SNPs/variety)
-0.95-0.98 Medium conf of being designable (11 SNPs/variety)
-0.90-0.95 Low conf of being designable (11 SNPs/variety)

~85% validation
Spacing and gene region coverage

Final Potato Infinium SNP Design:
2,769 SNPs in candidate genes (disease resistance, CHO)
508 SNPs in genetic markers
6,723 SNPs throughout the genome (~650 Mb of the genome will be covered)

-After fabrication, testing by Illumina, this condenses down into SNPs for 8303 markers

->2304 samples lower price; $85 per sample

-Project 2010 genotyping planes: 1152 potato x 8303 markers
The Potato Infinium Array SNPs are on the Public Genome Browser : Potatogenome.net

And available for download from SolCAP
Tetraploid Segregation

Yellow = Parents (AABB x BBBB)
Red = Population
(1:4:1 AABB:AAAB:BBBB)

Yellow = Parents (AABB x AABB)
Red = Population
(1:8:18:8:1
AAAA:AAAB:AABB:AAAB:BBBB)
Bad Markers
3 Cluster Calling

Good - 7412 (89.3%)
Questionable - 296 (3.6%)
Segregation - 254 (3.1%)
Bad - 341 (4.1%)

Call Rate for only good markers (7,412)
>90% 7036 (94.9%)
>80% 228 (3.1%)
>70% 93 (1.3%)
<70% 55 (0.7%)
Calling SNPs with 8300 Infinium Chip

5 Cluster Calling From 200 SNPs (so far)

68 (34%) with 5 clusters
56 (28%) - 5 good clear clusters
12 (6%) - 5 ok clusters

27 (13%) only 4 of the clusters present so unable to clearly define all five positions

32 (16%) only 3 of the clusters present so unable to clearly define all five positions

73 (37%) bad (either only 2 of the 5 clusters or clusters on top of each other or just an overall bad marker because of low intensity or no clusters)
Objective 5. Establish centralized facilities for genotyping a core set of SNP markers in standard germplasm panels in tomato and potato.

Presenter:
Allen Van Deynze
UC Davis
Tomato RNA-Seq Alignment and Data Processing

- The RNA-Seq reads from each variety were mapped separately using Bowtie to the tomato scaffolds.
- Used Paired end mode in increase alignment accuracy.
- Only alignments where the reads mapped uniquely were kept.
- SAMTools was used to process the alignment output (SAM) to into BAM and a pileup alignment file.
Tomato SNP calling with SAMTools

- Used the SAMTools SNP caller (improved from Maq)

- SNP caller parameters:
  - Minimum read depth: 20 bp
  - Maximum read depth: 240 bp
  - Window size for max SNPs/window filter: 100 bp
  - Maximum SNPs in window: 2
  - Filter SNPs with 50bp of a gap or indel
  - Minimum consensus quality score: 20
  - Minimum SNP quality score: 20
  - Minimum mapping quality score: 60
  - Max Percentage of 3' end bases the alignment at that base: 10%
SNP Filters

**Allele Frequency**
- OH08-6405 (CA fresh-market)
- FL7600 (Florida fresh-market)
- NC84173 (NC fresh-market)
- OH9242 (Ohio processing)
- PI114490 (wild cherry tomato)
- PI128216 (wild current tomato)
Tomato Infinium Array Design Considerations

- Created a unified list of genomic positions with a tomato SNP and link the position back to SNP call of each variety.
- No tomato gene models available at the time
- Align the tomato velvet transcript contigs with GMAP
- Filter SNPs based on location in transcript contigs
- Exons had to map >95% with no gaps
- The remaining SNPs were sent off to Illumina to be scored for the Bead Express and Illumina platforms.
- Remove positions with Bead Express scores < 0.75 and/or Infinium scores < 0.9
Tomato SNP Classification

Instead of randomly picking SNPs for the Bead Express validation array, we classified the tomato SNPs by variety classes and SNP information across varieties. A sample from each group was placed on the validation array.

<table>
<thead>
<tr>
<th>Type of SNP</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. pimpinellifolium Only</td>
<td>18,631</td>
</tr>
<tr>
<td>1 Market + S. pimpinellifolium</td>
<td>795</td>
</tr>
<tr>
<td>2,3,4 Market or Cherry + S. pimpinellifolium</td>
<td>1,660</td>
</tr>
<tr>
<td>2,3,4 Market or Cherry Only</td>
<td>737</td>
</tr>
<tr>
<td>1 Market Only</td>
<td>1,083</td>
</tr>
<tr>
<td>1 Cherry + S. pimpinellifolium</td>
<td>5,474</td>
</tr>
<tr>
<td>Total Tomato SNPs</td>
<td>28,830</td>
</tr>
</tbody>
</table>
# Tomato Bead Express Validation

<table>
<thead>
<tr>
<th></th>
<th>PM</th>
<th>MM</th>
<th>PM w/only hetero</th>
<th>PM %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>94</td>
<td>2</td>
<td>0</td>
<td>0.98</td>
</tr>
<tr>
<td>Six lines + H1706</td>
<td>93</td>
<td>3</td>
<td>0</td>
<td>0.97</td>
</tr>
<tr>
<td>Cultivated</td>
<td>70</td>
<td>23</td>
<td>3</td>
<td>0.73</td>
</tr>
<tr>
<td>PM in PROC</td>
<td>55</td>
<td>39</td>
<td>2</td>
<td>0.57</td>
</tr>
<tr>
<td>PM in FM</td>
<td>50</td>
<td>43</td>
<td>3</td>
<td>0.52</td>
</tr>
<tr>
<td>PM in Vintage</td>
<td>45</td>
<td>51</td>
<td>2</td>
<td>0.47</td>
</tr>
</tbody>
</table>

97% validation rate
Results of validation Test-96 Bead express

After filters: 28,380 putative SNPs
Cultivated: 9200

<table>
<thead>
<tr>
<th>Category</th>
<th>% PM</th>
<th>Total Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number</td>
<td>97%</td>
<td>9,200</td>
</tr>
<tr>
<td>Cultivated</td>
<td>70%</td>
<td>6,500</td>
</tr>
<tr>
<td>Processing</td>
<td>60%</td>
<td>5,600</td>
</tr>
<tr>
<td>Fresh-Market</td>
<td>43%</td>
<td>4,000</td>
</tr>
<tr>
<td>Vintage</td>
<td>34%</td>
<td>3,500</td>
</tr>
<tr>
<td>S. pimpinellifolium</td>
<td>40%</td>
<td>3,750</td>
</tr>
<tr>
<td>M82 x LA716</td>
<td>30%</td>
<td>2,760</td>
</tr>
</tbody>
</table>
Infinium data

Potato

Tomato
10,000 Bead-Infinium Tomato Chip Design

SNPs from SolCAP, Trait Genetics and INRA

Designed by John Hamilton, Robin Buell, David Francis and Allen Van Deynze

<table>
<thead>
<tr>
<th>Description</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAII SNPs cultivated</td>
<td>4,548</td>
</tr>
<tr>
<td>Unique Community SNPs</td>
<td>1,387</td>
</tr>
<tr>
<td>Unique Community Gene SNPs</td>
<td>552</td>
</tr>
<tr>
<td>Scaffold Coverage SNPs</td>
<td>2,297</td>
</tr>
<tr>
<td>Total SNPs</td>
<td>8,784</td>
</tr>
<tr>
<td>Transversion SNPs</td>
<td>1,216</td>
</tr>
<tr>
<td>Total Probes</td>
<td>10,000</td>
</tr>
</tbody>
</table>
SNP Genotyping Consortium

- Potato 10K (8,303 SNPs) Infinium chip is publicly available
- So far 3000 chips have been ordered, world-wide; first were shipped in mid-August
- Anyone who placed order before June 15 pays $85/chip
- Anyone else can still order, although cost is about 25% more

- Tomato 10K (8,200 SNPs 7,620 passed manufacture and Infinium validation.
- ~3,000 Chips ordered
Objective 6. Address regional, individual program and emerging needs within the Solanaceae community through a small grants program.

David Francis

The Ohio State University
Small Grants Program

• Address regional, individual program and emerging needs within the Solanaceae community through a small grants program.

• First call for mapping populations (potato) October 2009
  – SolCAP funds used to support the SNP genotyping of the population (funding ‘in kind’)

• allocate resources to:
  – SNP genotyping mapping populations in the core facilities requested by the greater breeding community (need large populations)
  – Marker conversion – developing SNP markers linked to QTL into easily assayed (e.g. CAPs or dCAPs) markers that end-users can readily apply in their own research programs
  – QTL validation and MAS
  – New directions not envisioned at the time of proposal submission.
Criteria for Mapping populations

- SNP polymorphic parents
- Traits of importance
  - traits of national, regional or local interest
- Large population size
- Quality of collected phenotypic data
  - Replicated data with low CVs
Potato Mapping populations selected

- Atlantic x Superior (tuber calcium, reducing sugars, internal defects, specific gravity (starch))
- B1829-5 x Atlantic (chip color, internal heat necrosis, specific gravity, maturity)
- W2310-3 x Kalkaska (scab R, chip color, reducing sugars, specific gravity, asparagine, acrylamide)
Tomato Call for Proposals
Criteria –
• Phenotypic evaluation (exp. design, mean, variation expected in proposals)
• Population size
• PM based on SNP array
Small Grants Program

• **Approach(s)**
  – Genotyping by Sequencing, Infinium, BeadXpress

• **Potential consortium projects**
  – RIL populations
  – NAM populations
  – Directional selection populations
Small Grants Program

- Potential consortium projects
  - Rejected project: breeding for extreme weather
Objective 7. Create integrated, breeder-focused resources for genotypic and phenotypic analysis by leveraging existing databases and resources at SGN.

Lukas Mueller
Boyce Thompson Institute
Cornell University
Objective 7. Create integrated, breeder-focused resources for genotypic and phenotypic analysis by leveraging existing databases and resources at SGN.

Lukas Mueller
The Boyce Thompson Institute for Plant Research (BTI)
Cornell University
SolICAP data

- Panels loaded
- Trait measurements
  - Tomato: ~160,000 traits for 479 accessions
  - Potato: ~6000 traits for ~200 accessions
- SNPs & genotype information
  - Will be available soon
- Solanaceae Phenotype Ontology
  - Integration of traits
- Maps
  - Two tomato maps loaded
Breeder's Toolbox

SGN Breeders Toolbox

The purpose of this page is to give breeders direct links to breeder-relevant tools and data on SGN. It is a work in progress and your feedback or suggestions are welcome to build this into a comprehensive, easy to use and breeder-friendly resource.

We love feedback! Please contact Joyce van Eck with suggestions for improvements!

Go to “http://solgenomics.net/search/direct_search.pl?search=trait”
Search SGN – Sol Genomics Network

Search SGN

Browse traits with QTLs: D | E | F | H | N | O | P | R | S | V

QTL search

Search QTLs by trait name (help)

[Submit new QTL data]

Browse QTL populations:
- QTL Tomato Howard German x LA1589 F2
- QTL Tomato Sausage x LA1589 F2
- QTL Tomato Yellow Stuffer x LA1589 F2

Cite SGN using Mueller et al. (2005).
SGN is supported by the NSF (#0110076).
USDA CSREES, and hosted at the Boyce Thompson Institute.

Subscribe to the sgn-announce mailing list for updates
Send comments and feedback to sgn-feedback@solgenomics.net

Disclaimer.
Sol genomics network

SP:0000010 'fruit size'

Term id: SP:0000010
Term name: fruit size
Definition:
Comment:

Synonyms:

Definition dbxrefs:

Secondary IDs:
- PO:0000001
- PO:0007042
- PATO:00000117

Ontology browser

Phenotype data/QTLs (2 populations)
- QTL Tomato Sausage x LA1589 F2
- QTL Tomato Yellow Stuffer x LA1589 F2

Annotated loci (1)

Annotated accessions (226)
Chado ND DB schema
Accession: Heinz 1706

Stock details

[New] [Edit] [Delete]

Organism: Solanum lycopersicum
Stock type: accession
Stock name: Heinz 1706
Unique name: Heinz 1706
Description: Fruits are of medium size and ellipsoid in shape. The country of origin is USA. Seeds were obtained from Heinz. This accession belongs to the 'Contemporary (Modern) lines' varieties. It is a processing variety.

Synonyms

solcap number: SCT_0030
synonym: H1706, LA4345

Pedigree data

None

Additional information

variety: Processing
donor: TGRC
country: United States
state: OH
adaptation: Humid
## Associated loci

<table>
<thead>
<tr>
<th>Locus name</th>
<th>Allele symbol</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>ovate</td>
<td>ovate 1</td>
<td>Pear shaped fruits.</td>
</tr>
</tbody>
</table>

**Associate another Locus**

- type a locus name

## Experimental data

- None

## Stock history

- None

## Related stocks

### This accession is_member_of

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>population</td>
<td>Cultivars and heirloom lines</td>
</tr>
</tbody>
</table>

### Stocks that belong to this accession

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>plot</td>
<td>8532_1_2009,Fremont, Ohio</td>
</tr>
<tr>
<td>plot</td>
<td>8722_2_2009,Fremont, Ohio</td>
</tr>
</tbody>
</table>

## images (4)

[Add new image]

## Literature annotation (0)

- None

[Associate publication]

## Ontology annotation

- None

## Phenotype annotation

- None

## Genotype data

- None

## dbxrefs

- None

## User comments

- None

[Add comment]
Heinz1706 tomato fruit longitudinal section.

Heinz1706 tomato fruit longitudinal section.
Associated loci

<table>
<thead>
<tr>
<th>Locus name</th>
<th>Allele symbol</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>ovate</td>
<td>ovate 1</td>
<td>Pear shaped fruits.</td>
</tr>
</tbody>
</table>

Associate another Locus

- psy
  - gypsy-type retrotransposon PORE2 (pore2)
  - Phytoene synthase 1 (psy1) Allele: 1
  - Phytoene synthase 1 (psy1) Allele: prov4
  - Phytoene synthase 1 (psy1) Allele: prov5
  - Phytoene synthase 1 (psy1) Allele: y
  - Phytoene synthase 1 (psy1) Allele: (2s)
  - Phytoene synthase 1 (psy1) Allele: (1s)
  - phytoene synthase 1 (psy1)
  - Phytoene synthase 1 (psy1)
  - phytoene synthase 2 (psy2)
  - phytoene synthase (psy)
  - hydroxyproline-rich systemin precursor II (HypSys)

Name

- Cultivars and heirloom lines

Name

plot 8532 1 2009, Fremont, Ohio
Evaluation of Extension Activities

Michael Coe
Cedar Lake Research Group
2010 Project Evaluation Activities

Feedback from SolCAP Extension Workshops/Webinars:

- Potato Association of America
- Tomato Disease Workshop

Feedback from SolCAP-related graduate courses:

- Cornell University
- The Ohio State University
Feedback from SoICAP Extension Workshops/Webinars

- Workshops were rated as effective, timely, and highly relevant
- 80 - 90% of participants would recommend the workshop to others
- Some sessions were seen as needing more time
- Technical level was “just right” for a majority but “too advanced” for some participants
- There were requests for more opportunities such as a longer, more comprehensive hands-on workshop
- See full report for details and participant comments and recommendations
Feedback from SolICAP-related graduate courses

- Course presentations and materials were rated as clear and effective
- Students plan to use what they learned in further academic or professional work
- Some units needed more time
- Participants reported significantly increased confidence in their ability to apply marker techniques and communicate about them
- See full report for details and student comments and recommendations
2011 activities

- SNP genotyping on panels and populations
- Databases for genotypic and phenotypic data
- Mining SNP genotype data and sequence data - publications
- QTL analyses
- Hands on workshops for breeders
- eXtension.org
- SNP genotyping tomato mapping poplns
Further outlook

- Integration of SNPs with genome sequences and genetic maps
- LD estimates
- Breeding population structure
- Tomato-potato synteny
- Gene discovery
  - Causative SNPs
  - Tuber/fruit genes
- Translation to breeding
- Deep sequencing candidate genes?
- Genotype by Sequencing
Developing “Breeder Friendly” Tools

- Current SGN interfaces are aimed at the molecular biologist, with searches designed to facilitate molecular discovery
  - Need a portal that is trait and germplasm centered
- Ability to search by traits relevant to (and defined by) breeders
- When a marker is identified, a protocol for use in breeding will be provided.
- Search option that only yields polymorphic marker, phenotype, or QTL results from elite germplasm
- Ability to search for known parents or offspring of any given genotype
- Ability to generate a list of markers that are polymorphic between any two parents
- Detailed tutorial/definitions of terms and traits utilized within the database
Outcomes for Breeding from SolCAP

• A genome-wide set of markers and bioinformatic tools accessible by breeders

  – Breeders will access germplasm for crossing based upon SNP polymorphism and linked QTL of interest

  – design crosses complementary for QTL and traits, and then use MAB in early generation selection.
Outcomes for Breeding from SolCAP

• Better understanding of the allelic variation influencing CHOs

  – Design crosses to create improved sugar and starch levels and starch quality.

  – Crosses designed to manipulate and select variation within existing elite populations or introgress novel alleles from wild germplasm.

  – More predictable and directed breeding effort for processing and fresh market traits.
Visit us at http://solcap.msu.edu/

**SolCAP**

**Solanaceae Coordinated Agricultural Project**

**The SolCAP project** links together people from public institutions, private institutions and industries who are dedicated to the improvement of the Solanaceae crops: potato and tomato. Through innovative research, education and training the SolCAP project will focus on translating genomic advances to NRI tomato and potato breeding programs. This will lead to significantly improved varieties that benefit the processors, the consumer and the environment.

**Solanaceae Coordinated Agricultural Project**

Project Headquarters
Michigan State University
A372 Plant and Soil Science Building
East Lansing, Michigan 48824
Phone: 517-355-6271 x1111
solcap@msu.edu
In the next few months, SolCAP will genotype 1,152 potatoes with 8,303 SNPs

- potato germplasm panel: 325 clones
- 4x russet mapping population: 200 clones
- 2x mapping population: 96 clones
- Community SNP genotyping:
  - 2 populations: 350
Infinium data

Potato

Tomato