



Solanaceae Coordinated Agricultural Project

SolCAP Solanaceae Coordinated Agricultural Project 2 nd Annual Meeting		
January 9, 2010	8:00AM – 10AM Open Session	Town and Country Resort, San Diego, CA
Meeting	2 nd Annual SolCAP Meeting	
Director	Dr. David Douches, MSU	
Executive Committee	David Francis, OSU	
	Robin Buell, MSU	
	Walter De Jong, Cornell Univ.	
	Allen Van Deynze, UC-Davis	
	Alexandra Stone, Oregon State Univ.	
	Lukas Mueller, Cornell	
Attendees	Kelly Zarka, MSU	
	Glenn Bryan, SCRI Scotland	
	David B. Neale, UC Davis	
	Deana Namuth Covert, University of Nebraska-Lincoln	
	Debby Lewis, Ohio State Univ.	
	Charles J. Rivara, California Tomato Research Institute	
	Michael Coe, Cedar Lake Research Group, SolCAP Assessment	
	Carl Braun, Monsanto (replaces Thomas Osborn)	
	James Giovannoni, Boyce Thompson Institute for Plant Research	
	Ed Kaleikau, NIFA-USDA	
	Joyce Van Eck, BTI Cornell University	
	Jay Scott, Univ. Florida	
	John McQueen, Oregon State Univ.	
	Hank Verbakel, Nunhems, a subsidiary of Bayer CropScience, Netherlands	
	Kent McCue, USDA/ARS	
	Barbara Liedl, West Virginia State Univ.	
	Nankui Tong, Campbell's	
	Mark Van Haaren, Netherlands	
	Richard Veilleux, Virginia Tech	
	Roger Leigh, Portland Oregon	
	Sam Hutton, University of Florida	
	Xiangyang Zheng	
	Matthew Robbins, Ohio State Univ.	
	Matthew Kinkade, Penn State	
	John Lindbo, Campbell's Seeds	
	Ann Marie Thro, USDA	
	Gera Jocham, NIFA, USDA	
	Cholani Weebadde, MSU RosBreed	

Welcome and Overview

8:00am Presenter: David Douches, SolCAP Director

Create a highly visible and innovative education program

8:30am Presenter: Walter DeJong, SolCAP Co-Director

1. Output: To create an innovative education program to train graduate students in genome-based breeding.

Outcomes: A graduate course in bioinformatics for plant breeding students was developed and delivered at Ohio State University.

Impact: Attended by 10 registered students, three visiting scientists, two staff. Student competence with the material (gauged in ten categories spanning operating systems, Unix/Linux, BLAST, Perl and BioPerl) increased 1.6 points on a 0-5 scale. This change was statistically significant.

2. Output: Surveys were designed and administered to the tomato and potato breeding communities to establish baseline levels of knowledge, attitude, skills, and level of use of genetic markers in breeding.

Outcome: The surveys reached 163 tomato and potato breeders, their staff, and allied professionals; response rate was 45.8% for the tomato community and 78.6% for the potato community.

Impact: Results will be used to evaluate future changes in breeder knowledge, attitude, skills, and use of markers, and to inform development and refinement of the content for future stakeholder extension workshops and online educational courses. The tomato and potato community displays market fragmentation between commodities and based on company resources. In response, SolCAP tailored separate extension training for the tomato and potato groups.

3. Output: Two workshops were developed for the tomato and potato breeding communities to initiate training in translational genomics.

Outcome: Over 120 breeders, staff and allied professionals received training at the workshops held at the 2009 Tomato Breeder's Roundtable and Potato Association of America meetings.

Impact: Pre- and post-workshop surveys were conducted to assess the effectiveness of the workshops. Participants returned very positive anonymous feedback on the quality, relevance and usefulness of the workshops, and reported significantly increased confidence in their ability to apply genetic marker technology and marker-based breeding, as well as their ability to communicate effectively to others about these topics. Comments on surveys indicated that some breeders will increase population size to improve breeding practices and breeders may alter practices in order to better leverage genomics resources. Participant survey responses and recommendations are being used to refine future extension and education events and information resources.

4. Output: SolCAP co-PDs have made over 20 presentations to audiences to the potato and tomato community to create awareness for the SolCAP project and its objectives.

Outcome: A diverse audience of community breeders, staff, extension specialists, allied professionals, industry executives, growers, and public attended the presentations.

Impact: These people were made aware of the SolCAP translational genomics project goals.

Amplify outreach efforts by developing an eXtension Community

8:35am Presenter: David Francis Co-Director, John McQueen and Alexandra Stone Co-Director

1. Output: A web community called Plant Breeding and Genomics (PBG) Community of Practice (CoP) was created to foster cross-disciplinary and cross commodity networking and collaboration to amplify outreach efforts targeting an eXtension Community of Interest (CoI) of plant breeders, seed industry professionals, and practitioners.

Outcomes: SolCAP held a webinar and an eXtension.org workshop to bring other plant translational genomics Coordinated Agricultural Projects (Rice, Wheat, Barley, Conifer and Roseaceae) into the CoP.

Impact: PBG CoP application was submitted, and PBG was listed as a new CoP at the eXtension National meeting.

2. Output: PBGworks was initiated as the collaborative Drupal-based workspace that is being used for the CoP. Content development for PBG CoP has been initiated.

Outcome: Four plant translational genomics CAPs have initiated their own workgroups within the PBG workspace, and three are active.

Impact: There are 78 active members in PBGWorks.

3. Output: Information on tomato or potato breeding was developed.

Outcome: A YouTube video on crossing tomato for breeding purposes was created:

<http://www.youtube.com/user/TomatoLab#p/a/u/2/acVHJBKIUIE>; A video on digital phenotyping has been submitted to the Journal of Visualized Research (JOVE).

Impact: The video has been viewed ~150 times per month since its release; YouTube statistics suggest it is viewed in North America, Central America, Africa, Europe, Asia, and Australia.

4. Output: Relative to our education and extension goals, we established a project web site (www.solcap.msu.edu) which is a centralized resource for SolCAP participants. Outcome: The site includes project news and personnel, databases, tools, resources, project objectives and downloads. A tri-fold brochure was also created for distribution at workshops and meetings.

Impact: We have published three newsletters that reach over 400 people in the Solanaceae community.

Collect standardized phenotypic data for tomato and potato.

8:40am Presenter: David Francis, SolCAP Co-Director

1. Output: Standardized phenotypic data was collected across multiple environments for tomato and potato.

Outcomes: A database of phenotypes for key traits across accessions of each commodity was initiated, accessible through SolCAP (<http://solcap.msu.edu/>) and Solanaceae Genome Network (SGN, <http://solgenomics.net/>). A cooperator's guide was developed for the collection of the phenotypic data. Seed increases for future phenotyping was also completed.

Impact: Data will be uploaded to SolCAP and SGN websites for public access.

2. Output: Field studies of tomato and potato lines from the germplasm panels were conducted in NC, NY, ID, FL, OH and CA.

Outcomes: The germplasm and intellectual property agreements for the crop panels were finalized, cooperator guides created and one field season of phenotypic evaluation completed. Flexible funds are being used to collect centralized sugar, chip color and vitamin C data (core traits for the project) in a 4x potato mapping population and also to increase the population size of the mapping population for 2010.

Impact: Phenotypic data will be maintained on the SolCAP and SGN databases for public access by the research community. Increasing the population size of the mapping population will increase our probability of identifying QTL for crucial traits.

Develop extensive sequence data.

8:45am Presenter: Robin Buell, SolCAP Co-Director

Illumina cDNA Sequencing and Design of Genotyping Platforms for Tomato and Potato

We have completed sequencing of the potato cDNA libraries using the Illumina GA2. We have processed the Illumina potato cDNA sequence data to identify intra- and intervarietal SNPs. Following quality filtering, 94 million reads were obtained with ~2 Gb of sequence per variety. We have assembled these reads into contigs representing ~38 Mb of assembled transcriptome per variety. We have aligned these contigs to the Doubled Monoploid potato genome assembly (*S. phureja* DM1-3 516R44) and identified high quality SNPs within and between cultivated potato varieties. For design of the potato Infinium genotyping platform, we have identified 69,011 high quality SNPs that meet the Infinium design specifications. Experimental validation of these predictions is underway.

For tomato, we have 323 million reads that passed quality filtering, representing ~2-3 Gb of sequence for each cultivar. Assembly of these reads on a per cultivar basis resulted in transcriptomes of 39-41 Mb per cultivar. Limited numbers of high quality SNPs were identified within cultivars. We anticipate delivery of the tomato genome sequence from the Tomato Genome Initiative by the end of 2009 and thus are waiting to identify high quality inter-varietal SNPs once we have access to the tomato genome sequence.

We have identified candidate genes for sugar, carbohydrate, and vitamin biosynthetic pathways in potato and tomato as suggested by the research community and have solicited the community for genes to include in the potato and tomato genotyping platforms. 1749 sequences were identified by project personnel and the community for this targeted SNP discovery

1. Output: Standardized phenotypic data was collected across multiple environments for tomato and potato. Outcomes: A database of phenotypes for key traits across accessions of each commodity was initiated, accessible through SolCAP (<http://solcap.msu.edu/>) and Solanaceae Genome Network (SGN, <http://solgenomics.net/>). A cooperator's guide was developed for the collection of the phenotypic data. Seed increases for future phenotyping was also completed.

Impact: Data will be uploaded to SolCAP and SGN websites for public access.

2. Output: To address emerging needs within the Solanaceae community a call for mapping populations was sent to the potato community.

Outcome: The research community is developing mapping populations for SNP genotyping to validate marker linkages to major loci influencing vitamin and sugar content and validate QTL.

Impact: These mapping populations will be used to link SNPs with segregating trait data. Manuscripts will be developed from each of the mapping populations that will be SNP genotyped in 2010.

Establish centralized facilities for genotyping a core set of SNPs

8:50am Presenter: Allen Van Deynze, SolCAP Co-Director

1. Output: Existing databases (publically available Sanger-derived ESTs) were mined for both potato and tomato SNPs.

Outcome: All SNPs identified within and between Sanger-derived potato and tomato ESTs are available for download at: http://solanaceae.plantbiology.msu.edu/analyses_snp.php. A webpage for querying the Heniz1706 vs TA486/MicroTom SNPs has been constructed and will be deployed shortly at the SolCAP website. Preliminary validation of tomato eSNPs has been completed using the Illumina GoldenGate assay.

Impact: SolCAP SNPs are available for the tomato and potato community through web-based interfaces and downloads. A manuscript will be written from this research in 2010.

Discussion:

SNPs within Sanger-derived ESTs:

Using publicly available Sanger-derived ESTs, we have identified putative SNPs (eSNPs) in the current potato EST collection; from the Sanger-derived 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. We benchmarked our computational pipeline by determining how many of the known, validated SNPs from tomato were present in our eSNP output. Using a set of 132 tomato sequences that were present within 77 contigs, 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). Lack of detection of 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. All SNPs identified within and between Sanger-derived potato and tomato ESTs can be seen at http://solanaceae.plantbiology.msu.edu/analyses_snp.php in which user query tools are made available.

For tomato, the majority of putative polymorphisms are of limited use as a number of the ESTs in the public tomato EST collection are derived from Micro-Tom, a tomato with seven wild-species in its pedigree. Thus, few data-mined tomato SNPs are useful within breeding populations. To identify polymorphisms among elite tomato germplasm, we developed a bioinformatic pipeline to identify SNPs between genomic sequence and assembled ESTs. Using the tomato TA496 ESTs in Genbank (108,685 ESTs), we cleaned and then clustered and assembled these sequences into 12,349 contigs. We then aligned the TA496 EST contigs to the emerging tomato genome sequence derived from Heinz 1706. Thirty-one percent (3,831) of the TA496 contigs aligned with the Heinz 1706 genomic sequence. Within these contigs, 86 high confidence SNPs were predicted within 32 contigs; an eSNP discovery rate of < 1% with 2-3 SNPs per contig. We have now expanded these analyses to include MicroTOM-derived ESTs vs. Heinz 1706. A webpage for querying the Heniz1706 vs. TA486/MicroTom SNPs has been constructed and will be deployed shortly at the SolCAP website. Validation assays have been conducted using the Illumina Golden-Gate Platform.

For potato, the above analyses include intra- as well as inter-varietal SNPs. We have also identified intra-varietal SNPs within Sanger EST collections for three potato varieties: Kennebec, Bintje, and Shepody. This will allow researchers to determine the allele distribution within varieties. A webpage for querying these SNPs has been constructed and will be deployed shortly at the SolCAP website.

2. Output: Extensive sequence data of expressed genes was generated from *de novo* next-generation sequencing of normalized cDNA libraries from 3 potato and 6 tomato varieties representing elite US germplasm using the Illumina GA2.

Outcome: We have identified high quality SNPs for development of the potato genotyping platform. We are poised to design the tomato platform in late 2009.

Impacts: Over 70,000 high quality SNPs were identified for potato. Currently, 7,600 SNP assays are being designed for the Illumina Infinium fixed array. Potato data has been shared with international collaborators and is being used to help anchor the potato genome sequence. A manuscript will be developed from this research in 2010.

Address regional, individual program and emerging needs through a small grants

8:55am Presenter: David Douches, SolCAP Co-Director

To increase breeder engagement with genomics, we will provide opportunities to genotype up to 20 populations and validate marker linkages to major QTL. This program provides a mechanism to reallocate resources and implement community-based strategies for validation and MAS. Small grant funds will be used primarily to genotype populations provided by breeders, thus leveraging SolCAP-developed SNPs to map traits (e.g., agronomic, disease resistance, insect resistance and quality) not addressed in our core germplasm panels. When requested, small grant funds will also be used to facilitate marker conversion in order to empower individual programs to pursue specific traits or loci and/or validate candidate QTL identified in our core germplasm panel. Many in the solanaceous community have begun to develop populations for SolCAP to genotype; where we know of such populations, we will genotype the parents at the same time we genotype the germplasm panels. We have set aside funds specifically to provide genotyping service in 2010, 2011 and 2012.

Create integrated, breeder-focused resources for genotypic and phenotypic analysis

9:10am Presenter: Lucas Mueller, SolCAP Co-Director

Output: An integrated, breeder-focused resource for genotypic and phenotypic analysis was initiated by leveraging existing databases and resources at SGN and MSU.

Outcome: The project website was established at MSU which is a central site for description, downloads, tools, newsletters and resources. A breeder's toolbox for SGN is being modified to serve the tomato and potato breeding community.

Impact: SolCAP and SGN websites will become key websites to access data for the tomato and potato communities. The SolCAP website has had a total of 3958 unique visitors from over 10,000 visits.

Evaluation of extension activities

9:15am Presenter: Michael Coe, Cedar Lake Research Group

At the January 2009 Plant and Animal Genome (PAG) meeting, we held our first advisory committee meeting and received verbal feedback and a written report from the Solanaceae Coordinated Agricultural Project (SolCAP) advisory group that was used by the executive committee to further refine deliverables.

External evaluation of extension and education activities by Dr. Michael Coe has proceeded according to plan. Institutional Review Board authorization for data collection from human subjects was obtained in

early 2009. An extensive survey for Solanaceae breeders was developed and administered to a comprehensive list of potato and tomato breeders during spring 2009. Data from this survey will provide a baseline against which to judge changes in breeder self-reported attitudes, knowledge, skills, and practices related to the use of genetic markers in their work. These data are also being used to aid in the design of education and extension materials and activities.

Additional surveys were developed for use at the first two extension workshops, at the Potato Association of America annual meeting and the Tomato Breeders Round Table. At each of these events, two evaluation instruments were used -- one asking direct questions about the quality, relevance, and utility of the workshop itself, and another (administered before and after the workshop) asking participants to self-assess particular knowledge and skills related to the application of genetic marker information in their work. Participants in these 4-hour workshops were generally very positive in their feedback about the workshops, and they provided helpful recommendations for additional extension sessions. Responses from 42 participants at TBRT and 28 participants at PAA indicated that the length and level of the workshops was appropriate; the presenters were well-informed, organized, and prepared; presenters were effective in style and delivery; participant self-confidence for applying markers in their work was significantly higher after the workshops; and participant self-confidence for communicating to others about the use of markers in plant breeding was significantly higher after the workshops. Findings from these questionnaires are being used in the design of future workshops, and in further development of instruments that will be used with education audiences.

A detailed report on these evaluation methods and findings will be available in 2010. Additional evaluation activities during the past year have included development of instruments for evaluating education activities; development and piloting of a usability testing protocol for the Breeder's Tool Box section of the SGN website; and planning for evaluation of online collaborative work involving various CAP plant breeding projects and eXtension.org.

SolCAP Future Outlooks/Issues and Open Discussion

9:30am- 10:00am Presenter: David Douches, SolCAP Director

We have almost complete participation of the US public breeding community for the germplasm panels. We expect the broader community of breeders to contribute valuable mapping populations for SNP genotyping.

In 2009, the potato genome sequence generated by the Potato Genome Sequencing Consortium was shared with SolCAP. This access allowed us to align the transcriptome sequences developed by SolCAP so that SNPs could be more robustly identified. In a reciprocal exchange, we provided the Potato Genome Sequencing Consortium our potato transcriptome data to facilitate development of SNP markers for anchoring of whole genome shotgun sequencing scaffolds to the genetic map.

We anticipate the tomato genome sequence to become available by the end of 2009. This will enable a robust identification of inter-varietal SNPs within tomato.

Strong interest from the international community: The International Potato Center, China, Canada, Scotland and Poland have asked to include germplasm on the SNP genotyping panel. The SNP array and germplasm panel will be a valuable community resource for at least the next 10 years.

Break

10:00am -10:15am

Advisory Board Meeting (Closed Door)

10:15-11:00am

SolCAP Executive Meeting with Advisory Board Meeting (Open Discussion)

11:00am-Noon

See SolCAP Advisory Board Report