

Project Title: Translating Solanaceae Sequence Diversity and Trait Variation into Applied Outcomes through Integrative Research, Education, and Extension

September 1, 2008 through November 30th, 2012.

Project Director: David Douches (douchesd@msu.edu), Michigan State University

Co-Directors: Robin Buell (buell@msu.edu), Michigan State University; Walter De Jong (wsd2@cornell.edu), Cornell University; David Francis (francis.77@osu.edu), Ohio State University; Allen Van Deynze (avandeynze@ucdavis.edu), University of California, Davis; Lukas Mueller (lam87@cornell.edu), Cornell University; Alexandra Stone (stonea@hort.oregonstate.edu), Oregon State University

Project Website: <http://solcap.msu.edu>

The SolCAP project was initiated September 2008 under grant no. 2008-55300-04757 of the USDA/NRI and is currently funded through USDA/NIFA/AFRI grant no. 2009-85606-05673.

The significance of SolCAP to Food and Agriculture

The Solanaceae Coordinated Agricultural Project (SolCAP) aimed to put research emerging from plant genome sequencing projects into the hands of applied breeding programs having direct impact to US Agriculture. SolCAP added significantly to the genomic resources available for potato and tomato through sequencing the gene space of 6 tomato and 4 potato lines. These sequence data represented the first comprehensive set of data for fresh market tomato and cultivated potato. Because the sequence resources were also based on currently relevant varieties rather than model accessions, the data led to the commercialization of high-throughput SNP genotyping tools that have become the industry standard. Combining these resources with trait data significantly advanced our understanding of germplasm in potato and tomato. Outreach materials developed to help translate these research tools into practice and accessible through eXtension.org, are viewed on average 9,000 times per month, while our workshops and webinars were attended by 750 individuals, and the recordings accessed more than 89,000 times. Breeding projects that leverage the SolCAP resources are addressing nutritional quality, resistance to abiotic stress, and resistance to important disease establishing a clear linkage to the goals of AFRI. SolCAP developed SNP markers will considerably facilitate the development of new potato and tomato varieties.

Broad Impacts

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 have validated the concordance of the SNPs with the draft genome and corrected the alignment of a few super-scaffolds. We have made our Illumina EST sequences available to the Tomato Genome Initiative to facilitate gene annotation. We have collaborated with Martin Ganai, Trait Genetics, Gaterslesben, Germany and Mathilde Causse, INRA, Avignon, France to incorporate some of their SNPs based on complementary germplasm into the community Tomato Infinium chip. We are co-publishing an integrated genetic and physical map of tomato based on the SolCAP Infinium chip.

We have completed the project objectives of SNP discovery, SNP platform development and deployment, and genotypic characterization of elite North American cultivated potato and tomato accessions as outlined in the initial grant proposal. Manuscripts describing SNP discovery as well as SNP platform development and validation are published or in press. A genotypic survey of 426 tomato accessions using the SolCAP 7720 SNP Infinium Array was completed and published. A genotypic characterization of a 250-potato clone diversity panel using the SolCAP 8303 Infinium SNP array has been completed and a manuscript describing these results is currently under review.

The SolCAP project developed two exceptionally powerful tools – potato and tomato SNP chips – that are now in widespread use around the world. These chips have dramatically reduced the time needed for investigators to make associations between markers and traits (weeks instead of years), which will have a profound impact on the development of new cultivars going forward. Because the chips are so easy to use, they have also greatly expanded the range of scientists who can now conduct genetic analyses in potato and tomato – no prior experience in molecular biology is needed. There has been very strong interest from the international community to use the SolCAP potato and tomato SNP arrays. Hundreds of the SolCAP developed potato/tomato arrays have been used, sold, or are in the queue at either user facilities or SolCAP project facilities available to the potato and tomato research community (MSU, UC-Davis, OSU). The SNP arrays have become a common marker set for which to evaluate and compare germplasm world-wide. The SNP arrays for potato and tomato as well as the germplasm panel will be a valuable community resource. We have been interacting with Illumina to continue to the manufacture of the potato and tomato SNP arrays. The impact of the arrays was emphasized at the International Solanaceae genomics meeting in Switzerland, August 28, 2012 with no less than 12 talks highlighting use of the arrays. The International Seed Federation is recommending the use of the tomato array for variety certification. It has 77 member countries including the US.

Use of the Infinium SNP genotyping system is allowing the calling of SNP allele dosage in tetraploid genotypes in over 5000 SNPs from the array. Scoring for dosage has increased our power to genetically analyze populations of cultivated tetraploid potato.

We have almost complete participation of the US public potato and tomato breeding communities for the germplasm panels. We expect the broader community of breeders to contribute valuable mapping populations for SNP genotyping. Through hands-on and

face-to-face interactions we have helped breeding and genetics programs handle the SNP array data.

In addition, the SolCAP project has created a highly educational website, housed at eXtension.org, to help train the next generation of breeders in marker-assisted breeding approaches and strategies. SolCAP extension accomplishments include: the establishment and certification of the Plant Breeding and Genomics Community of Practice (PGBCoP) on eXtension.org, the establishment of a shared DRUPAL workspace, the development of workshops addressing the application of genomic resources to plant breeding, and the development of PBG social network, primarily through membership to the PBG newsletter. The PGBCoP has published educational materials on a variety of topics important to the plant breeding community: experimental design, statistical inference, molecular biology, phenotyping, genotyping, mapping, QTL, and population development. In addition to traditional static tutorials, PGBCoP has worked to develop self-paced inquiry-based curriculum to connect breeders and students to open-access computational tools and genomic data through hands-on learning. PBG content has been viewed a quarter million times since launch of their CoP. The need for continuing education in technical breeding is highlighted by the quarter million views PBG has received on eXtension and YouTube since our launch, January 2010.

Accomplishments by Objective through November 2012

Education and Extension (Objectives 1 and 2)

Objective 1. Create an education program to train graduate students in genome-based breeding. (De Jong, Francis, Douches, Van Deynze, Liedl)

Specific outcomes and Timetable: We will educate graduate students in the theory and practice of translational genomics for vegetable crops. This objective will be achieved by developing and delivering curricula through a team-taught distance education course across four universities. **Year 2:** Graduate student curricula will be taught as a live pilot course at Cornell University. Course material will be added to the SolCAP website. **Year 3:** A course will be added to the plant breeding curriculum at the participating institutions as a live course. Course material will be updated integrating data and resources developed by SolCAP. The course will be team-taught as a distance learning course.

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

Outcomes: Graduate curriculum has been developed and delivered to students. Curriculum includes a graduate course in bioinformatics for plant breeding students <http://www.oardc.ohio-state.edu/tomato/HCS806/HCS806.htm>; data analysis with R (http://www.oardc.ohio-state.edu/tomato/hcs825/HCS806_R/HCS806_R_main.htm USERNAME: oardcwin\hcs825 PASSWORD: Genetics825); a graduate course “breeding with molecular markers” that was jointly delivered between Cornell and Ohio

State University; and “Introduction to Scripting and Statistics for Genetics Data Management” (PLBR 4092, Cornell).

The information delivered in these courses has been integrated into the plant breeding curriculum at The Ohio State University under “Advanced Plant Breeding” (HCS 7825), and at Cornell under PLBR 4092. The Ohio State University course description, syllabus, and learning objectives have been accepted as a permanent semester course.

In addition, SolCAP educational resources were used as the basis for a 30-hour course “Translating DNA sequence into new plant varieties” at the Universidad Nacional de Rosario, Facultad de Ciencias Agrarias, Rosario, Santa Fé, Argentina, November 3-9, 2010 (<http://www.fcagr.unr.edu.ar/escuelaposgrado/cursos-doc-2010/translatingdna.html>) and again October 23-25 (<http://www.fcagr.unr.edu.ar/blog/?p=246>). Similarly, SolCAP educational materials were used in the UC Davis Plant Breeding Academy (www.pba.ucdavis.edu) in the 1 week-long session in Barcelona, Spain and UC Davis and in the “Breeding with Molecular Marker” course (http://sbc.ucdavis.edu/education/Courses/breeding_with_molecular_markers_-_February_2012.html).

Impact: Methods in Bioinformatics for Plant Science was attended by 10 registered students, three visiting scientists, and two staff personnel. 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. Ten graduate students and seven postdocs/visiting scientists gained hands-on experience in basic bioinformatic scripting (taught by Mueller) and statistical analyses in R (taught by De Jong), to help them prepare for manipulating and analyzing large-scale marker datasets. SolCAP educational materials are being recognized as a tool to train international audiences.

Evaluation feedback from SolCAP-related courses at Cornell University and The Ohio State University revealed that after the experience, students reported significantly increased confidence (self-efficacy) in their ability to apply genetic markers in their work or teaching others about marker applications. The clarity and effectiveness of the course materials and presentations were rated as “very effective” or “OK” by 80 to 100 percent of participants. In addition, “Translating DNA Sequence into New Plant Varieties” was attended by 21 students from four universities and three companies and the Plant Breeding Academy was attended by 34 students from 12 countries as well as 65 professionals in the UC Davis marker course. In total, we reached 71 registered students, visiting scholars, and post-docs in these formal courses.

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 were 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 communities display market fragmentation between commodities and based on company resources. In response, SolCAP has developed separate extension training for the tomato and potato groups. In 2010/2011 learning from these two separate communities was shared. For example, David Francis communicated findings from tomato translational genomics in the potato workshop.

3. Output: To-date, seven 4-hour workshops were offered for the tomato and potato breeding communities in conjunction with the Potato Association of America, Tomato Breeders' Roundtable and Tomato Disease workshops. The workshops are designed to initiate training in translational genomics, including hands-on computer workshops that introduced plant breeders to evaluation of next generation sequencing data; how to browse the tomato and potato genomes; how to search and extract sequence for specific loci and DNA markers; how to manage, validate and export genotype data; and how to work with large genotyping datasets in breeding programs and populations as well as mapping in tetraploid populations. In 2011, workshops focused on working with phenotypic data, R programming and BLUPs, integrating molecular data generated with the Illumina chips and an introduction to SGN tools. Furthermore, workshops were broadcast as webinars to make them accessible to participants who were not able to travel to the meetings. These webinar sessions were recorded and are currently accessible at <http://solcap.msu.edu/meetingsworkshops.shtml>. Presentation slides and supporting documents (scripts for statistical analysis) are also available at this site and at <http://pbgworks.org/tomato-workshop>.

At the Plant Breeding Academics (which included 8 Solanaceae breeders and 26 breeders for non-Solanaceae species) topics included: Working with QTL, Marker quality, population structure, and bioinformatics in breeding.

In 2012 the last of four workshops was held at Denver, CO August 12. The workshop featured speakers from the community that presented SolCAP-generated SNP data. There were 16 online participants and 30 participants in Denver. The talks were as follows:

Candice Hirsch, Michigan State University

A Century of Potato Breeding: Improvement, Diversification, and Diversity

Richard Veilleux, Virginia Tech University

Seeking and Verifying Candidate Genes on the Infinium 8303 Potato SNP Array

Dave Douches, Michigan State University

SNP mapping and QTL analysis in a Tetraploid Mapping Population

Jaebum Park, Cornell University

A new user's experience with TetraploidMap and solCAP SNPs for QTL analysis in potato

Sarah Braun, University of Wisconsin

A case study: What to do when your SNP data set arrives.

Cinthya Zorilla, University of Wisconsin

Using SolCAP SNPs to map QTL for calcium and tuber quality in a tetraploid population derived from Atlantic x Superior

On August 1, 2012 a colloquium on Breeding in the Genomics Era: state of the art and new opportunities was organized by Dave Douches and Dave Francis and was held in Miami at the annual American Society of Horticulture Sciences meeting. The purpose of the colloquium was to help bridge the gap between progress with agronomic crops and vegetable/horticultural crops.

In addition, hands-on training with SNP software and mapping programs for six scientists and students took place at MSU the fall of 2011. In 2012 MSU trained lead scientists and staff at USDA/ARS Beltsville and University of Wisconsin, Madison.

Outcome: Over 200 breeders, staff and allied professionals from over 10 countries and 4 continents received training at the workshops. In 2011 workshops were in conjunction with the Tomato Disease Workshop (Ithaca, NY) and Potato Association of America (Wilmington, NC) meetings. These two workshops complemented previous ones as well as additional webinars sponsored by the project (See Output 4.). Both workshops were accessible online as webinars and recorded versions are now available as educational resources on eXtension.org and Youtube.

Impact: Among the 232 workshop participants who returned surveys (from the 2010, 2011, and 2012 SolCAP workshops at PAA, TDW, and ASHS), approximately 75 percent were professional breeders, technical staff members or related professionals; about 15 percent were students; and about 10 percent were postdocs.

In the 2011 potato workshop at PAA, 35 participants attended in person and 18 participated via webinar. Survey forms were completed by 58 percent of the workshop participants (49 percent of on-site participants and 78 percent of online participants). The survey sample included 31 participants with a focus on potato, two with a focus on tomato, and three non-Solanaceae scientists. Each of the workshop topics was viewed as relevant to their work by 97 to 100 percent of respondents; the clarity and effectiveness of each of the topical sessions was viewed as “very effective” or “OK” by 93 to 100 percent of respondents; 87 percent of participants moderately or strongly agreed that they gained new knowledge applicable to their work and 86 percent moderately or strongly

agreed that they planned to apply what they learned. Several comments highlighted access through the interactive webinar as very positive.

In the 2011 tomato workshop at TDW, there were 60 attending on site and 19 webinar participants. Feedback forms were collected from 18 of the onsite participants (30 percent) and from 11 of the online participants (58 percent) for a total response rate of 37 percent. Of those responding to surveys there were 29 participants with a focus on tomato, one with a focus on potato, and ten non-Solanaceae scientists. Each of the workshop topics was viewed as relevant to their work by 85 to 94 percent of respondents; the clarity and effectiveness of each of the topical sessions was viewed as “very effective” or “OK” by 68 to 81 percent of respondents; 70 percent of participants moderately or strongly agreed that they gained new knowledge applicable to their work and 63 percent moderately or strongly agreed that they planned to apply what they learned.

At the 2012 SolCAP colloquium at ASHS, there were 107 participants attending on site and 120 webinar participants. Feedback forms were collected from 15 of the onsite participants and from 67 of the online participants for a total response rate of 36 percent. Of those responding to surveys there were 10 participants with a focus on potato, 15 with a focus on tomato, and 65 who reported working on other crops. Each of the workshop topics was viewed as relevant to their work by 77 to 94 percent of respondents; the clarity and effectiveness of each of the topical sessions was viewed as “very effective” or “OK” by 93 to 98 percent of respondents; 72 percent of participants moderately or strongly agreed that they gained new knowledge applicable to their work and 68 percent moderately or strongly agreed that they planned to apply what they learned.

In the 2012 potato workshop at PAA, there were 30 participants attending on site and 16 webinar participants. Feedback forms were collected from 22 of the onsite participants and from 5 of the online participants for a total response rate of 59 percent. Of those responding to surveys there were 24 participants with a focus on potato and five who reported working on other crops. Each of the workshop topics was viewed as relevant to their work by 75 to 96 percent of respondents; the clarity and effectiveness of each of the topical sessions was viewed as “very effective” or “OK” by 92 to 100 percent of respondents; 65 percent of participants moderately or strongly agreed that they gained new knowledge applicable to their work and 61 percent moderately or strongly agreed that they planned to apply what they learned.

A detailed summary of findings from the workshop evaluation surveys including recommendations for future workshops was prepared by the external evaluator and is being used to help develop upcoming events.

Subsequent YouTube usage of the recorded webinar sessions as of the time of this report is as follows:

For the 2010 SolCAP workshop at Tomato Disease Workshop:

- 56,059 views; ~4671 views/month

For the 2011 SolCAP workshop at PAA
(<http://www.youtube.com/user/plantbreedgenomics#grid/user/FD1F4572812D3601>)

- 7522 views; ~578 views/month

For the 2012 SolCAP workshop at PAA

- 415 views; ~207 view/month

4. Output: A “how to” methods in plant breeding series was initiated (<http://www.extension.org/pages/60426>). The overall goal of the series is to help put basic research techniques, methods, and outcomes into practice through crop improvement. In addition, the series is intended to provide participants with hands-on experience using publicly available software and tools other resources.

Outcome: Nine one-hour webinars delivered by members of SolCAP and recruited experts were broadcast live to an online audience consisting of graduate students, post docs; educators, and plant breeding professionals who work with vegetables, grains, grasses, fruit trees, legumes, and even animals. In addition, the broadcasts were recorded and uploaded to YouTube. The presentation slides, supplementary software scripts and data files are available from <http://www.extension.org/pages/60426>. One of the webinars, “Breeding for Organic Production Systems” was a collaborative effort with the eOrganic Community of Practice (http://www.extension.org/organic_production). Collectively, the webinars have had 740 live attendees and been viewed on YouTube 84,945 times, averaging 6,067 views/month.

- Introduction to R Software for Plant Breeders
(<http://www.extension.org/pages/60427/introduction-to-r-statistical-software:-application-to-plant-breeding-webinar>)
- How to Align Sequences
(<http://www.extension.org/pages/60428/how-to-align-sequences-webinar>)
- Introduction to Augmented Design
(<http://www.extension.org/pages/60427/introduction-to-r-statistical-software:-application-to-plant-breeding-webinar>)
- Introduction to the rrBLUP Package in R for Genomewide Selection
(<http://pbgworks.org/node/1440>)
- How to Map Thousands of SNPs
(<http://www.extension.org/pages/63330/how-to-map-thousands-of-snps>)
- How to Develop SNP-based Tetraploid Maps for Potato
(<http://www.extension.org/pages/63187/how-to-develop-snp-based-tetraploid-maps-for-potato>)
- How to Use Double Haploids to Improve Winter Wheat
(<http://www.extension.org/pages/60429/how-to-use-doubled-haploids-to-improve-winter-wheat-webinar>)
- How to Design and Analyze Experiments Using an Augmented Design
(<http://www.extension.org/pages/60430/introduction-to-the-augmented-experimental-design-webinar>)
- How to Breed for Organic Production Systems

<http://www.extension.org/pages/60431/how-to-breed-for-organic-production-systems-webinar>)

Impact: Survey evaluations were distributed to all webinar attendees by an external evaluator. These surveys are designed to evaluate webinar impact. They are also a means by which the audience can provide input for future webinar topics and other online content. Survey comments from webinar participants indicate that the webinars are meeting the series objectives. For example, one attendee of the “Introduction to R Software” webinar reported:

“The benefit for me was to see some of the options available for ANOVA analyses with fixed and mixed models. I would like to understand the advantages and disadvantages of using R compared to other available software, and this webinar was a step in that direction. The online resources are very helpful and I expect to refer to them for data analysis and for preparation of classes that I teach.”

Response rates for the webinar evaluation questionnaires have ranged from 65 to 78 percent. When asked how much the webinar improved their understanding of the topic, 75 to 87 percent of respondents reported significant or moderate improvement; the remainder responded “A little improved” except for one individual who reported “not improved.” When asked about the technical difficulty of the presentations, 65 to 80 percent of respondents rated the difficulty as “just right,” while 0 to 30 percent rated the webinar as “too technical” and 5 to 20 percent rated the presentation as “too basic.” A detailed summary of findings from the aggregated webinar evaluation surveys including recommendations for future webinars is being prepared by the external evaluator; preliminary reports from individual webinars are currently being used to help develop upcoming events.

The R webinar was also featured in the OpenHelix blog (<http://blog.openhelix.eu/?p=9763>). OpenHelix is a well-known bioinformatics training company. The feature included the following:

“if you have been hearing about biologists who are using R for their work, or you’ve been seeing it mentioned in papers, and you’ve been wondering how and why to get started with R, this will be worth your time. It’s a gentle introduction to where to get R, how to start interacting with it, some tips on formatting your stuff, and links to additional resources for help using R.”

A significant impact is that through these webinars, we have extended our Community of Practice and Community of Interest well beyond Solanaceae with both speakers and audiences. For example, these webinars are viewed as highly informative by the National Association of Plant Breeders (over 1000 members) which has actively advertised these amongst their members and on their website (plantbreeding.org).

Discussion: Our progress suggests that there is an interest and need for continuing education in plant breeding and genomics: over 740 individuals participated in our live webinars and workshops and our video content has been viewed 84,945 times. Additionally, our video content has been used to develop two new university courses.

Resources developed by SolCAP, including eXtension webinars and tutorials and computer-based learning module developed at MSU to teach genomic skills; such as, manage and filter SNP data, create genetic maps, and assess genetic vs. physical maps. The MSU learning module has been previously used in CSS451, a plant biotechnology class, and at a molecular marker workshop at MSU. The course will benefit from our PBGworks platform, which provides a password-protected site for students and faculty to interact during classes, and the use of Multipoint, a Polycom software update that allows connection and interactivity for up to 4 sites in 2 modes (continuous presence with a split screen on monitor of all sites or voice activated where the last person that talks is on the screen and remains until another site takes over). When content is sent from a computer, all sites see content with an insert (PIP) of the presentation site. This software was accompanied by equipment upgrades on a 50% cost share available to academic departments supported by The Ohio Agricultural Research and Development Center. Ohio State University and Cornell have experience in shared teaching (2010), and now with Michigan State University a “video” linked course is now possible.

Objective 2. Amplify outreach efforts by developing an eXtension Plant Breeding Community of Practice (CoP) to develop continuing education material aimed at practicing plant breeders, their staff and seed industry professionals. (Stone, Francis, Van Deynze, Liedl and De Jong)

Specific Outcomes and Timetable: Our Community of Interest (CoI) will gain knowledge and ability to use genomic information in the context of applied plant breeding programs. The CoI will also have improved access to genomic data through the breeder’s toolbox (see Obj. 7). At the same time, genome scientists will gain a better understanding of breeding through engagement. Assessment of participant responses to these efforts will be measured through surveys conducted by Dr. Coe. **Year 2:** SolCAP eXtension content will include 12 featured articles, 100 FAQs, and 6 resource tutorials. These will be integrated through learning lessons into a single “start-to-finish” example of marker-assisted selection (MAS) focused on a high-heritability trait. Two workshops will be conducted for plant breeders and staff. **Year 3:** Additional public content will be launched. PBG content on eXtension.org will include contributions from other CAPs as well as news and interactive elements. Curricula will be expanded through learning lessons and database tutorials that encompass more complicated examples of MAS for quantitative trait loci. SolCAP will provide week-long breeder training in SolCAP labs. Two workshops will be conducted for plant breeders and staff. Content will be evaluated for its utility to agricultural professionals. The PBGWorks collaborative workspace at <http://pbgworks.org> will be evaluated for efficacy in fostering cross-disciplinary and cross-commodity networking and collaboration, as well as interaction among breeders, geneticists, and agricultural professionals.

Education and Extension

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

Outcome: SolCAP held a webinar series (discussed above) and annual eXtension.org workshop to bring other plant translational genomics Coordinated Agricultural Projects (Conifer, Barley, Roseaceae, Common Bean) into the CoP.

Impact: PBGCoP was accepted by eXtension as a new Community of Practice March 3, 2010. The CoP has a space on the People website of eXtension (login required to view) <http://www.extension.org/people/communities/363>.

2. Output: PBGworks.org is the collaborative workspace that is being used for the CoP. Content is authored and reviewed there and then transmitted to eXtension.org for publication.

Outcome: There are 5 CAP groups participating in the workspace, as well as a variety of lab groups and PBG eXtension content development and administrative groups. All SolCAP eXtension content is authored in the workspace.

Impact: The PBGWorkspace currently has 453 people registered as members who represent 30 universities and federal agencies, 11 educational institutions outside of the USA, and 5 industry groups. All eXtension content is managed and available to all CoP members 24-7, regardless of their location.

3. Output: The SolCAP project newsletter features timely updates and how-to sections. Upon release of the draft potato and tomato genome sequencing we provided information to the community on how to retrieve and use the data.

Outcome: We have published 2 of our 10 informative SolCAP newsletters in the past year.

Impact: Newsletters are sent out to over 400 people via e-mail and are available on our website for others to download. Feedback from the community regarding the newsletters has been very positive. Furthermore the PBGworks website is being used routinely to reach the Plant breeding community to inform members of external courses and meetings such as the National Association of Plant Breeders, UC Davis Seed Biotechnology Center, University of Minnesota.

4. Output: The PBG CoP publicly launched 124 pages of content on January 15, 2011 (www.eXtension.org/plant_breeding_genomics).

Outcome: The PBG CoP now has 165 pages of public content.

Impact: According to Google Analytics, the PBG eXtension website (www.eXtension.org/plant_breeding_genomics) has now had a total of unique 66,931 visits and 216,513 page views since launch on January 15, 2011, and is projected to receive 350,000 page views by the end of 2013. In the past quarter, the site has averaged 20,763 page views per month. Views of PBG eXtension content in the last quarter account for 2.6% of eXtension activity for the 60 publicly launched communities. Each visit to the PBG eXtension website results in an average of 3.22 page views with an average of 3:45 minutes spent on the eXtension site. In comparison, average visits for all of eXtension are 1.99 page views with 1:38 minutes spent on the eXtension site. These analytics suggest that an increasing number of visitors are spending time with the original content produced by SolCAP and our partners in the PBG CoP. Users report satisfaction with the content and multi-media approaches. One industry user posted the following comment:

"Good to see Extension leveraging IT and social networking to meet their mandate to transfer technical knowledge and expertise to the public. I love it when I SEE the value of my tax dollars at work. Thank you."

6. Output: The PBG CoP applied for certification by eXtension. The PBG CoP is required to be certified by eXtension within 18 months of public launch. Certification can be granted for one, two, or three years. After a CoP has been certified by eXtension, they become eligible to apply for eXtension leadership funds. Leadership funds for 2012 are \$12,500 per CoP.

Outcome: In September 2011 the PBG CoP was certified by eXtension for three years. The PBG CoP applied for 2012 leadership funds.

Impact: SolCAP has begun to lay the groundwork for sustainability of the PBG CoP beyond the period of funding. Leadership funds were awarded to PBG, and are being used to draw more projects into PBG in a meaningful way.

7. Output: The PBG CoP launched a YouTube channel (<http://www.youtube.com/user/plantbreedgenomics>) to host webinar, workshop, and educational videos.

Outcome: The YouTube channel currently has 122 videos uploaded.

Impact: The videos have been viewed 89,465 times. The channel currently has more than 300 subscribers.

8. Output: SolCAP has developed educational videos and uploaded them to YouTube.

Outcome: Additional crop-specific videos were developed and uploaded to YouTube.

Impact:

- DNA Extraction and Marker-Assisted Selection (<http://www.youtube.com/watch?v=yI8M9z4N4Y8>) - 5,189 views (~200 views/month)
- Tomato Cross Pollination (<http://www.youtube.com/watch?v=acVHJBKIUIE>) – 19,176 views (~450 views/month)
- Plant Breeding Solves Problems (<http://youtu.be/mtVtpO7FOaA>) – 3,395 views (~280 views/month)
- Tomato Seed Saving (<http://youtu.be/gg8FDRa-rBQ>) – 607 views (50 views/month)

9. Output: SolCAP collaborated with the Conifer Translational Genomics Network CAP to coordinate peer-review and publication of a multi-part series of learning modules that is focused on the fundamentals of quantitative genetics in the context of molecular tree improvement. These learning modules aim to provide a comprehensive introduction to the fundamental background science and technology required to understand and incorporate genetic markers in applied tree breeding and resource management.

Outcome: Fourteen learning modules were peer-reviewed, voice-over narration of the modules was recorded, and the modules were published to YouTube and eXtension.org (<http://www.extension.org/pages/60370/>). One additional module are currently in the peer-review process.

Impact: The learning module videos have been viewed 944 times (~410 times/per month) on YouTube. According to Google Analytics, the eXtension pages have been viewed 1,770 times.

10. Output: PBG expanded communication using PBGnews that is built into the <http://pbgworks.org> website.

Outcome: PBGnews is a great way for the PBG Community to keep in touch and communicate with other interested parties. Subscribers were initially account holders at PBGworks.org, but the general public is also permitted to sign up to receive the PBGnews. Subscribers can unsubscribe at any time via a link sent with each PBGnews post. In addition to email, the PBGnews is available for viewing at <http://pbgworks.org>. PBGnews is created and sent using the <http://pbgworks.org> website.

Impact: PBGnews has sent 50 posts containing information on upcoming webinars and new content published to eXtension.org to its subscribers. Currently, there are over 1000 subscribers to PBGnews, 453 of which are now PBG Community members.

11. Output: With additional funding of \$25,000, SolCAP organized an eXtension workshop for African plant breeders at the National Association of Plant Breeders. The meeting included 14 African plant breeders and Dave Frances, Allen Van Deynze and Dave Douches organized the content, and discussion was led by Greg Crosby.

Discussion: PBG CoP: The official recognition of PBG as a CoP by eXtension and the launch of PBG content on eXtension represent two important milestones. Other plant breeding, plant genetics, and plant genomics projects can therefore leverage our effort should they choose (see below).

NIFA eXtension policies and RFA language: As the result of the SolCAP report from 2010, Francis and Stone met with eXtension and NIFA leadership to discuss our executive committee's recommendations. A CoP leader advisory group was convened and that group developed new NIFA-eXtension policies and language for NIFA RFAs. This new language was included in NIFA SCRI and AFRI RFAs including language specific to partnering with eXtension.

PBGworks content and group management system: We believe that PBGworks has helped remove some barriers to participation in eXtension through the Drupal-based system created by eOrganic/Oregon State University. In addition, PBGworks has embedded group management and communication tools unavailable through eXtension. Due in part to the success of PBGworks and eOrganic.info, eXtension has now created a Drupal content management system. As documented above, PBGworks and NIFA are contributing to the productive development of eXtension.

a. Publication through eXtension

The potential advantages of publishing through eXtension are that 1) it inspired the creation of PBG CoP (described below), and 2) eXtension is viewed as the national cooperative extension web initiative, supported by land grant universities and NIFA.

Although we have found negotiating the eXtension system to be a frustrating experience in many ways, we also believe that progress is being made. The potential disadvantages of publishing through eXtension remain (1) hierarchical organization; (2) moving targets related to the expectations of a CoP; (3) barriers to participation that include added layers of approval before augmenting a community of practice through competitive grants; (4) lack of basic editorial guidelines, (5) the use of superficial evaluation criteria to assess impacts; (6) conflicting views between the NSF funded online projects and eXtension regarding duplicating and cloning content; (7) poor eXtension public site functionality; and (8) concerns about the costs and sustainability of eXtension delivery.

PBG has established editorial guidelines. Other plant breeding, plant genetics, and plant genomics projects can therefore leverage our effort should they choose (see below). During the 2010 PAG Advisory Committee meeting, the comment was made that eXtension's structure might inhibit our ability to deliver content in a timely manner. The official launch of PBG on eXtension.org in 2011 has removed this barrier and we are reaching a wider target audience than we initially imagined.

b. PBG Community of Practice

The potential advantages of PBG CoP are that 1) the plant breeding and genomics community and NIFA-supported PBG-related CAPs are brought together into one group

to work together over time; 2) individual projects can take leadership and then step back over time, but remain as members in the long term, allowing for an ebb and flow of leadership and participation – this will help the group endure over time, 3) this group can collaboratively develop and review outreach and training materials and support those materials over the long term (regardless of where that content is published), 4) this group could support and integrate other research, education, and outreach efforts.

The potential disadvantages of PBG CoP are 1) it requires a higher level of organization to coordinate this larger group and content; 2) barriers to participation related to eXtension management and functionality noted above; 3) it will require support from competitive grants beyond the funding period of SolCAP. This support will not happen unless RFAs require outreach through eXtension, and NIFA program managers insure that both core activities and content development are directly supported in budgets.

Barriers to participation remain to be overcome. Officially, eXtension retains review and approval rights to any augmentation to a CoP, irrespective of funding source. The system in place adds steps of approval to the submission process and is generally not well understood (or followed) by the scientists applying for competitive funds nor NIFA program managers. Our experience suggests that teams faced with the eXtension process often elect to use other delivery methods which are less confining. Progress has been noted in this area as eXtension is rapidly responding to requests for eXtension letters of support, facilitating inclusion of PBG CoP plans of work and budgets into NIFA proposals. At this time, it appears that NIFA support through the budgeting process may be a weak link in assuring sustainability. We have successfully partnered with other NIFA projects for PBG outreach, but these will account for only 15% of the funding needed to sustain the effort. We know of other projects that have claimed to leverage PBG and eXtension, but have not included in their budgets any support for the core PBG infrastructure that is necessary for those outreach methods.

c. PBGworks collaborative workspace

The potential advantages of the DRUPAL based collaborative workspace (PBGworks) are that 1) it gives PBG CoP a virtual place to interact and develop and archive content, 2) the personnel staffing provides support for and expertise in eXtension content development and publication, 3) it is easier to use than the eXtension wiki, and it also has group management tools not provided by eXtension, 4) it has the capacity to support networking, project management, collaboration, and co-learning, thereby supporting not only the outreach mission of PBG, but also the integration of research and education. Increasingly, researchers/educators/ extension personnel will network and collaborate through virtual tools and systems; PBGworks has the potential to be that system for the plant breeding and genomics community.

The potential disadvantages of PBGworks include that 1) it is fairly expensive to develop and support, and 2) it needs improved usability before CoP members can take advantage of its potential (Oregon SU personnel are continuously working with Shawn Yarnes and others to improve usability), and 3) it will require funding support over time.

We believe that PBG has helped remove some barriers to participation by adopting the DRUPAL system pioneered by eOrganic for our wiki. The functionality of this system within PBG has lead eXtension to transition to a DRUPAL system. PBG and NIFA are therefore contributing to an improvement in eXtension.

d. SolCAP Recommendation to NIFA regarding the NIFA/eXtension initiative:

We recommended in the 2010 report that NIFA create a **NIFA-directed advisory group for NIFA-supported CoP's**. This occurred, but as far as we know this group only met twice to develop the language in the current RFAs, and is not a formal group that meets regularly. We think this group should be convened at least twice per year to review NIFA RFA language and any other issues relevant to the NIFA/eXtension/CoP partnership; membership of this group should include at least one representative from every CoP receiving NIFA funding, as well as NIFA and eXtension administrators. SolCAP's recommendation is that RFAs contain language requiring or recommending that the eXtension components of appropriate proposals work with the PBGCoP, that their project budgets include support for the eXtension.org website, and that their project budgets include support for core activities of the PBG COP that they intend to leverage.

5. Output: SolCAP co-PDs have given numerous talks and have been invited speakers for audiences from the potato and tomato community at local, regional, national and international meetings to create awareness of the SolCAP project and its objectives while fostering research and education collaborations.

Outcome: A diverse audience of plant 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. Stakeholders had input on the relevance and impact of SolCAP to their programs.

Research Objectives

Objective 3. Collect standardized phenotypic data across multiple environments for tomato and potato. (Francis, Douches and De Jong)

Specific Outcomes and Timetable: phenotypic data, described above, for core germplasm sets in tomato and potato have been quality checked and are now integrated into the SGN database and are available as flat-files on the SolCAP website

(http://solcap.msu.edu/tomato_phenotype_data.shtml,

http://solcap.msu.edu/potato_phenotype_data.shtml). The germplasm panels will serve as a long-term breeding community resource. Breeders will be able to access data from these panels to assess parents of populations already developed for marker polymorphism as well as to facilitate the design of future crosses by allowing identification of parents with desirable combinations of marker and trait variation. **Year 2:** Seed multiplication; field evaluations by breeder-cooperators. Phenotypic data uploaded to the SolCAP website and made available to SGN. **Year 3:** Final phenotypic data uploaded to the SolCAP website and made available to SGN.

Research

Output: Field studies to collect standardized phenotypic data of tomato and potato lines from the germplasm panels were conducted in NC, NY, ID, MN, WA, WI, FL, OH and CA.

Outcomes: The germplasm and intellectual property agreements for the crop panels were finalized in the first year and three field seasons of phenotypic evaluation have now been completed. Some of the flexible funds were used to collect centralized sugar (sucrose, glucose and fructose), malic and citric acid and chip color data in a 4x potato mapping population and the potato diversity panel. Other funds were used to increase the population size of the mapping population to approximately 200 progeny for 2010. Cooperator's guides for potato and tomato were developed and refined for the collection of the 2010/11 phenotypic data. The proposed field trials were completed in 2011. The potato diversity panel was grown out in 2012 at Cornell University for further phenotyping (metabolites and glycoalkaloids) and providing tubers to be used in placing the panel in tissue culture. The potato diversity panel is a resource. The tissue culture lines will be maintained by MSU for future research by the potato community. The potato community measured additional traits (tuber late blight and starch quality) on the diversity panel in 2012. The russet mapping population field trials were completed in 2011 but a subsample of the population was grown out for further phenotyping (tuber asparagine and acrylamide in processing) in Michigan and Idaho through SCRI grant support. All post harvest sugar analysis has been completed. The 2010 data has been quality checked and posted on SGN. The 2011 sugar data is being quality checked and should be given to SGN this winter. Dr. Kathy Haynes is conducting the statistical analysis of the mapping population data as well as leading the QTL mapping. The James Hutton Institute in Scotland is revising their tetraploid mapping software to handle the larger data sets generated by the SNP array. The potato diversity panel manuscript was submitted in December 2012. The first of two tomato diversity panel manuscripts was published (Merk et al., 2012).

A database of phenotypes for key traits across accessions for both commodities is now in place, accessible through SolCAP (<http://solcap.msu.edu/>) and Solanaceae Genome Network (SGN, <http://solgenomics.net/>).

Impact: 2009 and 2010 data was uploaded to the SolCAP website for public access and is in the process of being incorporated into the SGN database and website. Phenotypic data will be maintained on the SolCAP and SGN databases for public access by the research community. For potato, the 2010 field data (400 plots per location) of the tetraploid mapping population from NC and MN was collected and tubers were sent to MSU for chip processing and sugar analysis sampling. Freeze-dried and processed tuber tissue samples were sent to the USDA/ARS Bethke lab in 2011 for HPLC analysis of sugars (glucose, fructose and sucrose) and additional metabolites (malic acid and citric acid). The ID fry and sugar data was collected in Idaho by Dr. Sanjay Gupta through the use of flex funding. Field data (440 plots) from the diversity panel was also collected from the NY, WI and WA sites. Tuber samples were sent to MSU and the same process was followed as with the tubers from the mapping population. The 2010 data was used for association analysis with the potato diversity panel and was presented at the 2011 potato workshop. The WA processing data was compromised by root knot nematode

damage and not used. A third year of tomato data was collected for yield on the processing sub-set in Ohio. Tomato data now contains 3 environments for all traits in the the fresh-market, processing, and vintage classes, with most data available for four environments. Agronomic, phenology and quality data were collected. Quality included titrateable acids, soluble solids and pH. Additionally, replicate samples were scanned and analyzed for over 20 objective digital fruit shape and color traits using the tomato analyzer. Field trials were conducted in CA, OH, NC and FL. Analysis of these data were presented at meetings, and manuscripts detailing trait heritability, genetic correlations, and estimated phenotypic breeding values of SolCAP lines are published or in preparation.

Discussion: The phenotyping for agronomic and quality traits was conducted by cooperating breeders. Germplasm coordinators and cooperators developed a “cooperators guide” to standardize data collection for both field performance and fruit/tuber quality. The guiding principles are to collect standardized objective data, to use consistent language, and order traits into ontologies where appropriate. Each cooperator evaluated germplasm for traits specific to their breeding program (e.g. resistance and quality traits). Buell, Douches and Francis led the phenotypic database design, descriptor language and rating scales. Final formatted data is posted on SolCAP and SGN. For tomato, descriptors consistent with UPOV/IPGR are collected to describe plant habit, inflorescence structure and leaf shape. Fruit size and shape data are collected from scanned images, as will color (L*, a*, b*, Hue, Chroma) and color uniformity data, using TomatoAnalyzer software. Standard measurements of total soluble solids, pH, and citric acid are also collected. These data have already been integrated into SGN phenotype ontologies. For potato, standard measures include glucose, sucrose, chip (or fry) color and specific gravity (an indirect measure of starch content). Other measurements include total yield, and vine maturity along with Plant Variety Protection (PVP) descriptors such as flower color, vine type, tuber skin color, tuber flesh color, tuber pigment pattern, tuber shape, skin brightness, and tuber eye depth. Sugar and vitamin HPLC measurements are being performed at the University of WI, Madison. Individual sites are scoring opportunistic disease and quality data such as frequency of heat sprouts, internal heat necrosis, reaction to common scab, etc. Phenotype data for potato and tomato) are available on the project web site (http://solcap.msu.edu/phenotype_data.shtml). The 2009 and 2010 SolCAP data is now searchable on SGN http://solgenomics.net/search/direct_search.pl?search=phenotypes . The SolCAP accessions and plots for tomato can be found at: http://solgenomics.net/stock/search?advanced=1&stock_type=56792&search_submitted=1&person=SolCAP%2C+project&submit=Search and http://solgenomics.net/stock/search?advanced=1&stock_type=56807&&search_submitted=1&person=SolCAP%2C+project&submit=Search. The potato accessions are also phenotyped and can be viewed at <http://solgenomics.net/stock/7313/view>.

Coordination and test sites are being managed through crop-specific germplasm committees headed by SolCAP PD/co-PDs. Cooperative tests are being grown in at least two locations for each crop under the supervision of experienced field breeders. Through the “cooperators guide and well-attended breeder workshops”, we also promote a

“genomics mentality” that emphasizes increased genotypic replication (perhaps at the expense of block replication) and the collection of standardized and objective phenotypic data. Regional cooperative trials already exist for potato, and are being leveraged in support of SolCAP. Regional cooperative trials (e.g. the Northern Tomato Exchange Program, NTEP) have not existed for tomato for over 15 years. SolCAP has provided a first step toward re-establishing cooperative trials for this crop in order to increase the ability to draw inference across genetic backgrounds and environments. Under SolCAP, we have performed the first coordinated multi-institutional and multi-location evaluation of processing, fresh-market, and vintage tomato germplasm to occur in over 15 years.

SNP genotyping of the potato and tomato panels is complete. With the panel being a community resource, we replicated the panel SNP genotyping to ensure that all SNP genotyping was correct and all lines in the panel were identified correctly. We are initiating association mapping, germplasm diversity and population structure analyses, estimation of marker-based breeding values, and cross validation for genome-wide selection. The clonal population for potato currently consists of 220 individuals. A recent publication (*Euphytica* (2008) 161:47-60) identified many quality trait-marker associations, using a panel of 221 genotypes and only 250 AFLP markers. Another published association study (*Genetics* (2007) 175:879-899) identified markers associated with resistance to late blight using a population of only 123 individuals. We expect to find associations and then examine and validate them more closely in bi-parental crosses. Additional potato lines have been genotyped with SolCAP flex funds. This additional genotyping includes diploid germplasm from Virginia and Wisconsin, cultivated germplasm from CIP and a core collection from the species core collection. In the US 10 potato populations have been SNP genotyped with the SolCAP 8303 Infinium SNP array.

QTL that span market classes will be tractable through association approaches if the associations explain >10% of the phenotypic variation. Not all QTL will span market classes, and a previous reviewer is correct that, in these cases, association mapping will not be successful. The ‘community grants’ program is providing a means for characterizing traits that are market-class specific using a more traditional bi-parental mapping approach. We believe that estimates of LD and population structure will have long-term impacts on how potato and tomato breeders conduct their programs. Our pilot study with tomato indicates that there is significant linkage phase disequilibrium within and between breeding programs. The graphical view of these patterns is already helping breeders choose parents to (1) preserve groups of co-adapted genes and (2) choose parents to introduce new variation into programs. Finally, our educational material is aimed at helping breeders understand the importance of increasing genotypic replication to facilitate identification and validation of QTL-marker associations.

Published QTL studies in potato range from 150-250 individuals for progeny sizes; we will be QTL mapping in populations greater than 150 individuals. Note that our mapping population and clonal potato panel are separate studies. Not all project resources are dedicated to association mapping. Genotyping panels that represent breeding germplasm both validate markers and assess allele frequency within breeding programs and market classes, providing valuable information for deployment of markers

in potato and tomato. The genotypes and phenotypes provide a guide for population structure and development of appropriate association studies for subclasses. The potato and tomato communities have submitted mapping populations for SNP genotyping (three 4x and one 2x for potato and five for tomato). The community populations were approved by the SolCAP Executive Committee and these populations will enhance our mapping and validation efforts. Further validation will be done outside of SolCAP with funds by the public and private sector investment in the SolCAP-based SNP platforms for potato and tomato. As of fall 2012, enough tomato and potato SNP arrays have been purchased to genotype over 10,000 lines each. Illumina has conducted a survey with the tomato and potato communities to determine the need to manufacture new bead pools for the SNP arrays. We expect that SNP arrays to continue to be a resource for the breeding communities in 2013 and beyond.

A previous reviewer expressed concern about experimental design, data flow and analysis pipeline. By necessity, details of field lay-out are left to cooperators. The experimental design is a randomized complete block design, with over-replication of checks within each block. Location weather data, management data, and field lay-out are stored as supplementary files. Phenotypic data are collected into a 'cooperators guide' format by cooperators, and passed on to the germplasm coordinators (Francis for tomato, Douches for potato) for quality control (QC). The cooperators guide uses defined vocabulary that is both consistent with historical descriptors and mapped to trait ontologies used in SGN. The guide is formatted such that data ranges are pre-determined in order to restrict opportunities for entry error. In the case of fruit and tuber morphology, objective data collected from scanned images are linked to those images which are available through SGN. QC data are posted to the wiki as dated files where they are available to post directly to the project web site and for SGN staff to upload. QC involves statistical analysis for outlier detection, cross checking for known phenotypes (e.g. yellow tomato fruit), and verification that data fall within expected ranges. Analysis of multi-location and multi-year data will first involve a mixed model analysis of variance to quantify variation across experimental environments. We expect that some of the data may be unbalanced when combined across sites. Data will therefore be combined across locations/years as estimated BLUPs.

Association analysis using variations of the Unified Mixed Model will follow, but we wish to stress that this analysis is not the principle goal of collecting phenotypic data. We acknowledge that for some traits, association analysis will not be feasible due to underlying population structure and the distribution of traits across genetic groups. Furthermore, some traits (e.g. purple fruit and/or long-shelf life in tomato) are conditioned in tomato germplasm by 3-5 alleles at the same locus. Our population sizes and marker density will not permit detection in such cases. On the other hand there are traits that span populations. Initial analysis focused on known traits and causal alleles (OVATE for fruit shape, the beta-cyclase gene for high lycopene/high beta carotene phenotypes). Markers will be analyzed both as single point and using groups of markers to define haplotypes. Models will be refined based on these traits prior to analysis of phenotypes with unknown genetic basis. Finally, we have already designed bi-parental crosses based on phenotypic distributions. In tomato, these bi-parental populations are

already at the F5 RIL stage with a nested structure. A multi-location experiment for QTL validation in processing tomato has been evaluated over two years in two locations each year. The population is structured as a nested RIL population (n = 288) with evaluation using an augmented design which can account for spatial variation through over-replicated checks. Genotyping was performed with a sub-set of 384 SNPs chosen to cover the genome and maximize polymorphism among the parents.

Objective 4. Develop extensive sequence data of expressed genes, and identify Single Nucleotide Polymorphisms (SNPs) markers distributed across the genome and associated with specific candidate genes for sugar, carbohydrate and vitamin biosynthetic pathways. (Francis, Van Deynze, Buell, De Jong)

Timeline and Specific Outcomes: Normalized cDNA libraries and 2 Gb sequence for each of 3 potato and 6 tomato genotypes was completed in 2009; SNPs discovery was completed in 2010. For both tomato and potato, an Illumina Infinium SNP genotyping array was commercialized in 2010, with 8,300 potato SNPs and 7,720 tomato SNPs. Genotyping for both potato and tomato was completed in 2011, and AFRI research dollars were leveraged through the Illumina consortium to generate many times the genotypic data points than originally budgeted.

THIS OBJECTIVE HAS BEEN COMPLETED

Research

1. **Output:** (Objective 4) Extensive sequence data of expressed genes was generated from *de novo* Illumina sequencing of normalized cDNA libraries from 3 potato and 6 tomato lines representing elite US germplasm.

Outcome: We have identified high quality SNPs for the development of the potato and tomato genotyping platforms. We have completed the design for the Infinium potato and tomato platforms. Both the potato and tomato SNP predictions have been validated using the Illumina BeadExpress platform. For tomato, we performed whole transcriptome sequencing with six accessions that span the market classes, cherry and wild tomato using the Illumina Genome Analyzer platform to identify single nucleotide polymorphisms (SNPs) that can be used for genome-guided breeding efforts. In total, we generated 291,915,037 quality filtered sequence reads representing 17 Gb of sequence. Assembly of the reads resulted in 30.6-34.9 Mb of sequence for the six accessions that provided representation of 55.3-59.6% of the predicted tomato gene set and a wide range of molecular function ontologies. We developed a computational pipeline to identify SNPs using transcript sequences in conjunction with a reference genome and when coupled with two Sanger-derived Expressed Sequence Tag transcript datasets, we identified 62,576 non-redundant SNPs in tomato. The SNPs within the contigs were present within all of the Gene Ontology molecular function categories suggesting broad coverage of genes involved in biological processes. Our computational pipeline was validated using the Illumina BeadXpress genotyping platforms with validation rates greater than 98.5%.

Impact: Over 69,000 high quality SNPs that meet the Illumina Infinium platform design criteria were identified for potato. Currently, ~8300 scoreable markers are represented on the potato Infinium array. Potato data has been shared with international collaborators (Scottish Crops Research Institute, SCRI) and the Potato Genome Sequencing Consortium to help anchor the potato genome sequence. Over 28,000 SNPs were identified in tomato, with 8,784 SNPs selected for development of an Illumina Infinium array. Of these, 7,720 SNPs passed production quality check and were subsequently scored on the tomato germplasm panel (n=489) representing cultivated varieties (processing, fresh market, vintage and landrace) and wild species. We obtained high quality genotype data (< 10% missing data) from 7,375 SNPs across the panel. Tomato data has been shared with collaborators, 3,700 SNPs have been placed on the *S. lycopersicum* x *S. pennellii* reference map, all markers have been placed on the physical map and have been integrated into SGN. Chromosome by chromosome patterns of linkage disequilibrium has been analyzed. The potato and tomato SNP manuscripts have been published.

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

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 identified ~69,000 high quality SNPs that meet the Infinium design specifications. The computational SNP prediction pipeline was validated using a 96 SNP BeadExpress assay with a validation rate of > 85%. Interest in the community and negotiations with Illumina has allowed us to create a consortium in potato to increase the number of samples being genotyped and reduce the price/sample. As a result, ~8303 validated SNP markers are being assayed with the Illumina Infinium potato array.

Tomato: For tomato, we have 323 million reads that passed quality filtering, representing ~2-3 Gb of sequence for each cultivar/accession. Assembly of these reads on a per cultivar basis resulted in transcriptomes of 39-41 Mb per cultivar. We have mapped the reads to the tomato genome draft assembly and have identified high quality SNPs. We identified >28,000 candidate SNPs for the Illumina Infinium platform. These SNPs were further filtered based on frequency in the accessions, and the computational pipeline was validated using 96 SNPs and 96 tomato accessions with the BeadExpress platform. We achieved a validation rate of 97%. Interest in the community and negotiations with Illumina has allowed us to create a consortium in tomato to increase the number of samples being genotyped and reduce the price/sample. As a result 8,784 SNPs were submitted to Illumina for fabrication of the final tomato platform, of these 7,720 SNPs

passed manufacture QC and are scorable. The original proposal budgeted for 737,280 data points. The Illumina consortium provided the world-wide community with resources for a minimum of 22,800,000 data points.

For both the potato and tomato arrays, SNPs were selected from within candidate genes for sugar, carbohydrate, and vitamin biosynthetic pathways as well as other genes of interest suggested by the research community. Almost 1800 sequences were identified by project personnel and the community for this targeted SNP discovery. In potato, over 500 candidate genes have at least one SNP. In summary, there are over 3000 SNPs in candidate genes, 500 SNPs in genetic markers and 4,700 SNPs selected to achieve maximum genome coverage. We estimate that we are covering about 650Mb of the estimated current reference genome of about 727 Mb). For tomato, 567 SNPs represent candidate genes, 1,470 SNPs were obtained from the community, and 270 scaffolds from the Heinz1706 reference genome assembly are represented in the final design.

The SolCAP team has extended an invitation to our fellow Solanaceae scientists to join us in a consortium through Illumina® to utilize the Infinium SNP array for potato and tomato for the interrogation of SNPs in the respective. After the first round of the consortium, there is strong interest in continuing the manufacture of the tomato and potato SNP arrays from Illumina.

Key benefits offered to participants:

1. *Early Access:* An opportunity to access a genotyping tool unlike what has been available previously in the tomato and potato communities.
2. *Technology Transfer Opportunity:* Part of the mandate of SolCAP is to facilitate the transfer of technology in the form of genotyping tools into the breeding effort in Solanaceae species. This consortium offers an opportunity to work with experienced SolCAP members to learn the methods associated with implementing high-throughput genotyping into characterizing genetic merit of elite potato and tomato germplasm.
3. *Reduced Pricing.* The price for consortium members will be lower than that available to contributors working alone. Per sample price of chips includes reagents necessary to run chips on an iScan or Bead Array Reader.

Objective 5. Establish centralized facilities for genotyping a core set of SNP markers in standard germplasm panels in tomato and potato. (Van Deynze, Douches and Francis)

Timetable and Specific Outcomes: With an anticipated SNP validation and assay success rate of $\geq 85\%$, we will deliver over 7M data points in both tomato and potato. Of the SNPs assayed, up to 25% will be from candidate genes. We used the draft genome sequences of potato and tomato to identify and use SNPs across as much of the genome scaffolds as possible. QTL analysis in the tetraploid potato mapping population should allow discovery of loci affecting $>10\%$ of the variation for a trait. Analysis of the core collections will define population structure, define LD decay, identify cryptic introgressions and describe patterns of variation and allelic diversity within and among market classes. We will complete genotyping of the approximately 480 individuals from

both the tomato and potato collections (approximately 960 total) in **year 1** and add genotypes from specific mapping populations to both crops in **years 2 and 3** through the small grants program.

Research

Output: Genotypes for over 2500 potato and over 680 tomato breeding lines and progeny from mapping populations with the Infinium SNP platforms have been assayed by SolCAP. In potato, over 5000 of the SNPs differentiate allele dosage in tetraploid potato. Germplasm panels for tomato and potato genotypes were assayed by SolCAP from the community mapping populations in 2011 and 2012.

Outcome: Preliminary population structure in potato has been defined with 82 SNPs with results published. A more extensive analysis with the 8303 potato SNPs are being worked on at this time and manuscripts for potato and tomato have been published in 2012. Analysis of tomato SNP data has been completed and published (Sim et al., 2012a and 2012b), establishing genetic map position for over 4,500 SNPs and revealing patterns of selection within the core collection. Contemporary tomato germplasm contains more genetic variation than vintage or heirloom varieties, suggesting that breeding has increased allelic richness and diversity in the tomato crop.

Impact: A common Illumina SNP array has been established for both potato and tomato. Arrays for over 10,000 samples for tomato and potato communities have been purchased worldwide.

Discussion: The potato and tomato Illumina Infinium arrays are available to assay by any facility. For SolCAP, UC Davis and MSU have been the primary facilities for genotyping the Infinium potato and tomato arrays. MSU is also providing service for outside parties for the potato and tomato arrays. Thus far, for potato, SolCAP has genotyping data for 100% of 1152 of the original planned samples. Thus far, for tomato, SolCAP has genotyping data for 680 samples.

Potato: DNA samples from three 4x mapping populations (160, 180 and 200 individuals), two 2x mapping populations (92 individuals each), and a set of diverse potato clones (250 individuals) were run on an Illumina iScan using the SolCAP designed 8,303 Infinium SNP chip, and subsequently analyzed using the Illumina GenomeStudio software. This software package has an auto-clustering option for three marker classes (AA, AB, and BB), however, due to variation in quality and positioning within each SNP assay, this auto-clustering option is not ideal. Rather, manual positioning of the clusters using a broad germplasm set provides the most accurate SNP genotype data. Personnel in the SolCAP project generated a custom three-cluster SNP file for the 8,303 Infinium chip by manually determining cluster positions for each marker within the Illumina GenomeStudio software. This custom cluster file is a coded GenomeStudio project file that can be used to view the clusters for the germplasm from which they were determined. This custom cluster file and comments relating to the quality of each SNP (good, questionable, bad, distorted segregation) were released on the SolCAP website

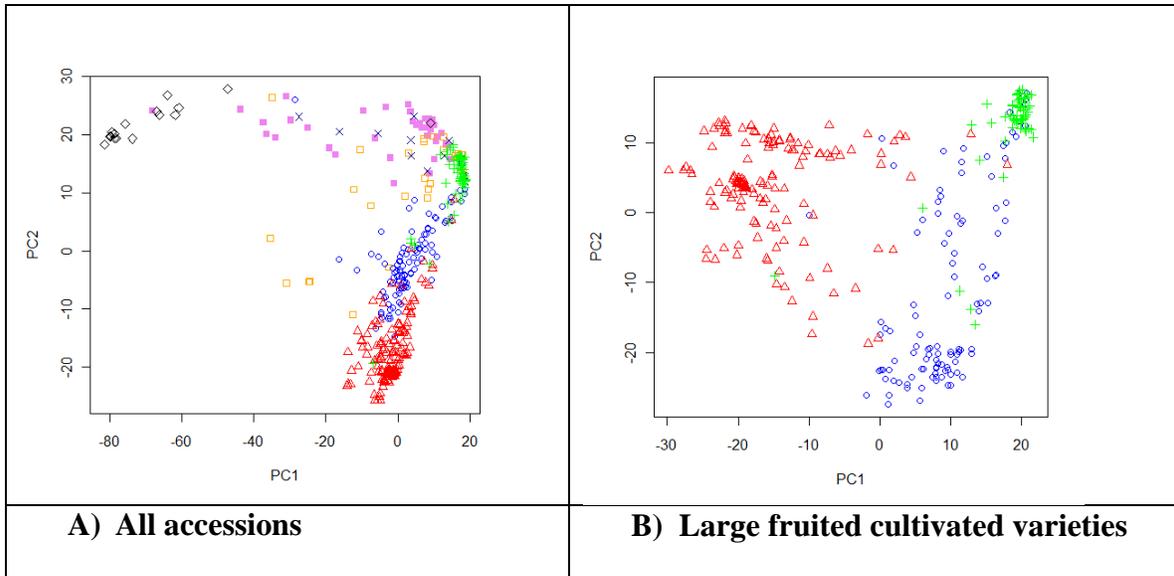
(http://solcap.msu.edu/potato_infinium.shtml) in December 2010. From this analysis it was determined that 7,412 SNPs were of high quality, and the remaining 891 SNPs were of lower quality, had distorted segregation based on the mapping populations, or were of unusable quality.

Many of the potato clones being SNP genotyped by the SolCAP project and those used in commercial production are tetraploid. The GenomeStudio software is not capable of generating clusters for all of the possible genotypes in a tetraploid (AAAA, AAAB, AABB, ABBB, and BBBB), and all heterozygous genotypes (AAAB, AABB, and ABBB) are all scored as AB in a three-cluster file. These heterozygous classes, however, can be resolved based on the theta-value, which is a function of the green/red light intensity that is emitted, for the clone for a given SNP. Using the mapping populations and diverse clones described above, SolCAP manually generated an additional five-cluster file that defines the theta boundaries for each marker class for each of the 8303 potato SNPs. Although, a diverse set of germplasm was used, not all possible marker classes were represented for each SNP, and thus theta boundaries could not be defined. The breakdown for number of marker classes that could be defined is: 2,645 SNPs with five clusters, 858 SNPs with four of the five clusters, 945 SNPs with three of the five clusters, 583 SNPs with two of the five clusters, and 3,272 SNPs either only had one cluster or clusters could not be defined due to loose clusters that could not be distinguished. SolCAP also generated a custom perl script to convert the theta positions as outputted by GenomeStudio into meaningful genotype scores using these theta boundaries. The custom five-cluster file was made publically available October 31, 2012.

Using the 8,303 Infinium SNP array, we genotyped a 250 line diversity panel that represents a broad range of genetic diversity, phenotypic diversity, geographic origin), ploidy level (1x, 2x, and 4x), wild species used for introgressions, release date, and market classes. Using the genotypic data, we characterized the panel for population structure, kinship, and alleles enriched in specific market classes. Using pairwise genetic distances, we were able to separate the sampled market classes within the panel suggesting that selection has occurred for traits specific to each market class. We also were able to clearly differentiate the wild species and genetic stocks from cultivated potato. Examination of biosynthetic pathways revealed enrichment of alleles in cultivated potato for genes involved in the biochemical pathways important for market class specific traits. A manuscript describing this work is in preparation and anticipated to be submitted in December.

Tomato: The 7,720 SNPs from the Infinium array have been mapped relative to both a standard genetic map (3,700 markers mapped through collaboration with Trait Genetics) and relative to the physical map via BLAST against the draft genome of H1706. The polymorphism rates of these SNPs were 63.3% for processing varieties, 86.8% for fresh market varieties, 81.4% for vintage varieties, and 96.9% for wild species. Graphing minor allele frequency (MAF) relative to genetic and physical positions revealed differences in haplotype blocks between market classes of cultivated tomato. Although the number of polymorphic markers within market classes and even within the breeding

programs represented in the SolCAP panel exceeded expectations, recombination will restrict the informative use of all polymorphic markers. Analysis of LD decay based on genetic and physical distance on a chromosome-by-chromosome basis reveals extensive haplotype blocks within and between market classes. The extent of linkage disequilibrium (LD) was examined across each chromosome with LD decay ranging between 1.3-12.2 cM within processing varieties, 3.4-12.2 cM within fresh market varieties, and 0.6-21.7 cM within vintage varieties. These results have been incorporated into the strategy for Objective 6 below.



Principle component analysis of A) 388 red-fruited tomato accessions and B) large-fruited cultivated accessions based on 7,720 markers. The first 3 PCA explain 26.6% of the cumulative variation. Processing varieties are coded red, Fresh Market are coded blue, Vintage varieties are coded green, landrace varieties are coded gold, cherry varieties are coded violet and wild species are coded black. Results are published in Sim et al., 2012. These results confirm our preliminary analysis on 96 lines separating market classes and germplasm within breeding classes (Robbins et al. 2011).

Objective 6. Address regional, individual program and emerging needs within the Solanaceae community through a small grants program. (Executive Committee)

Specific Outcomes and Timetable: An integrated Solanaceae genetic map based upon the 7720 (tomato) and 8,303 (potato) SNP genotyping arrays will be incorporated into SGN. Markers for CHO, vitamin and many other traits will be made available for individual breeding programs to use. **Year 1 and 2:** Small grants were awarded to genotype 3 potato mapping populations. Data will begin to be incorporated into SGN. **Years 2 and 3:** Small grant funding for SNP genotyping will continue with up to 8 mapping populations for

potato and 8 mapping populations for tomato; validation studies for CHO and vitamin QTLs will be initiated; marker conversion projects will be initiated and completed, QTL validation or other future national needs (as prioritized by the small grants funding) will be addressed.

Research

Output: To address emerging needs within the Solanaceae community, a call for mapping populations to be SNP genotyped with the Infinium arrays was sent to the potato and tomato communities in September 2010 and March 2011, respectively. The selection process was as follows: a request for proposals was distributed to the entire Solanaceae research community, through the newsletter, email lists, and the SolCAP website. Proposals were submitted to the PD, and peer evaluated through the EC. The EC ranked the proposals. The PD notified the PIs of each proposal's outcome.

SolCAP itself is focused on key cross-commodity quality traits, but there are many other traits of value in the Solanaceae. Six key areas we intend to allocate resources to are 1) genotyping mapping populations requested by the greater breeding community, 2) Marker conversion – developing SNP markers linked to QTL into easily assayed (e.g. CAPs or dCAPs) markers by individual research programs, 3) QTL validation and MAS, 4) population development to address emerging needs, 5) Extension or education special projects, and 6) new directions not envisioned at the time of proposal submission. Furthermore, the process of selecting which populations to genotype was based upon known or potential marker polymorphism, importance and number of traits segregating, and evidence of replicated phenotypic data and statistical analysis from the populations. This community genotyping effort seeks to leverage resources, and it is understood by the potato and tomato communities that this is a partnership. SolCAP will not carry the total cost of these studies. At the suggestion of our advisory board, our call for proposals explicitly sought populations for which phenotypic data are available and the parents for these populations must already be in the germplasm panel. In selecting populations, we have considered the number and distribution of markers and the quality of the phenotypic data (population size, distribution of data (mean and variation), replicates, etc.). Lastly, SNP polymorphism rate is not a problem, at all, in potato and the SNP polymorphism in tomato is greater than we expected.

Outcome: The potato community has developed six tetraploid and three diploid mapping populations for SNP genotyping. These populations are described as follows:

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)

DM x 84SD22 / DM x RH (SNP mapping)

Jacqueline Lee x MSG227-2 (specific gravity, late blight resistance, vine maturity)

W4 x 524-8 (specific gravity, chip color, disease resistance)

Premier Russet x Rio Grande (reducing sugars, FF, chips color, yield, tuber shape, etc.)

Waneta x Pike (chip color, specific gravity, scab resistance, Pink rot)

These populations are being used to evaluate concordance between SNP genetic positions and SNP physical position within the draft genome and to identify and validate marker linkages to major loci and QTL influencing sugar content and validate QTL. These bi-parental mapping populations should complement the diversity panel and russet mapping population that SolCAP has genotyped and phenotyped and should provide an opportunity for QTL validation. Each population has 1000s of SNPs segregating in the progeny. SNP Genotyping is now completed for potato. QTL analysis is being conducted on most of the tetraploid populations at this time as well as two diploid populations. Some of the preliminary results were presented at the PAA SolCAP workshop in 2012. In the Tundra x Kalkaska mapping population we have identified disease resistance QTL for scab and a major QTL for chip color from 45F storage. The tetraploid SNP data also provides an opportunity to examine double reduction in SNPs located distal to the centromere. Additional flex funds have been used to further phenotype the potato diversity panel and to put the potato lines into tissue culture.

There are limited funds associated with the community SNP genotyping program when one considers that two crops are being genotyped. We have a great challenge of prioritizing traits across the crops in these QTL studies. Therefore validation is driven by the individual breeding programs. We provide an example of how QTL-marker associations will be used in a potato breeding program: the QTL of largest effect for potato starch and reducing sugar levels will be used immediately by De Jong and Douches, and almost certainly by all other potato breeders. In many cases, the donor source of the QTL can be crossed with elite parents to create breeding populations with the expectation that advanced breeding lines or varieties may be selected from the populations. We foresee the breeders a) in the short term, select parents with complementary alleles for crossing and b) in the long term, study their interactions with other loci in the CHO pathways and drive breeding populations to fixation at these desirable loci. In potato, there will be a focus on resistance traits in addition to CHOs. We expect that validation will be carried out by programs that are interested in specific resistance traits. These studies may be regional in nature due to the environmental conditions necessary to promote disease development or insect pressure. Simply-inherited major genes for resistances should require little validation compared to more complex traits like yield and stability in potato (not the focus of SolCAP at this time). The ability to have simply-inherited markers linked to major resistances will reduce the cost of breeding by reducing the amount of germplasm that will be screened through costly and time-consuming *in vivo* assays. The impact of effectively introgressing resistance traits will be immense in potato since considerable amounts of insecticides and fungicides are applied to the seed, growing crop and harvested tubers.

For tomato, 12 populations were submitted by the community; six were selected for SNP genotyping. In assessing these populations, the SolCAP Executive Committee (EC) evaluated experimental design, data distribution, and data variability. These populations represent six public breeding programs and will assay yield, antioxidants, carbohydrates, late and early blight, virus, bacterial and fungal resistance, and flowering

and reproduction traits. One population is being genotyped with support from seed industry/genetic service providers, thus further leveraging SolCAP resources. Two populations (one potato and one tomato) will be genotyped by sequencing (GBS) with Illumina. The potato GBS procedure is being optimized this fall. The tomato GBS will be initiated as soon as the breeder is able to provide in writing that the population being used is publically available to the community. We felt that it was important that the small grants program was designed to vet community proposals. Through the review process, we have been able to direct translational research towards promising approaches. These represent a mixture of biparental mapping, introgression, MAS, and association mapping approaches.

Based on polymorphism of parents on the 7720 tomato Infinium chip and on observed recombination between markers, two 384 SNP panels were selected for processing and fresh market populations. SNPs in each panel were chosen based on genetic and physical locations in the genome. Three fresh market tomato populations from NC State, Florida and Cornell were genotyped with the 384 SNPs using the KASPAR assay. New marker and QTL were developed for virus resistance as a result. An interspecific RIL population of 143 lines was chosen for GBS up to 0.5X. Parental lines will be sequenced 10X. This population has been phenotyped for lycopene, early blight resistance and yield related traits.

Impact: The two diploid mapping populations for potato were used to map the SNPs to the pseudomolecules. Between the two populations, approximately 4500 SNPs from the potato array were accessed for this analysis. The study supported the concordance between the physical and genetic maps. The tetraploid mapping populations are being used to link SNPs with segregating trait data. We expect at least four manuscripts being written in 2013 from the initial SNP genotyping of potato and tomato. 384 SNP panels mapped to the genetic and physical maps have been developed for fresh market and processing tomatoes.

Discussion: In the previous comments from reviewers, there were questions regarding population size and the feasibility of association mapping given the diversity of germplasm. The first point is that the association mapping panel captures a wide range of genetic variation, which is useful for allele mining in exotic sources, the potential for inference to any one elite target population will be low because the number of members in each subpopulation or market class will be quite small. In brief, association mapping is only one of the translational approaches that we plan to pursue. The power of the approach and the strength of inference will depend on several factors for which we have limited or no quantitative data and that will vary between tomato and potato. Genotyping under Objective 5 will provide us data to make more informed decisions. These analyses are under way at this time. Our goal is to work toward a QTL-mapping emphasis that stresses both genome-wide approaches and introgression of traits across market classes and from exotic sources. We hope to learn from the design of our germplasm panels the potential of association mapping to map traits or QTL of importance. The germplasm panels are hierarchal in nature and the information gained from the initial genotyping

may direct our efforts for association mapping in the future, or may suggest other approaches.

In tomato mapping populations genotyped were chosen for their potential to identify QTL and to develop markers. Thus far novel QTL and markers for Tomato leaf curl virus have been identified with additional disease resistance genes being addressed in populations. By genotyping on the tomato Infinium chip, informative marker assays were designed to map in the genetically narrow germplasm of tomato, a task not feasible in the past.

Objective 7. Create integrated, breeder-focused resources for genotypic and phenotypic analysis by leveraging existing databases and resources at SGN and MSU. (Buell, Mueller, Francis)

Specific Outcomes and Timetable: Breeder-friendly interactive interfaces to query and display data from trait to gene including germplasm, phenotypic, genotypic, genomic data with links to protocols and references; marker analysis/assay/conversion resources will be developed; Breeder adoption of genomic tools and markers in their breeding programs will be measured by Dr. Coe. **Year 1:** Data will begin to be uploaded to the MSU SolCAP and SGN will initiate SolCAP data integration. **Years 2 and 3:** Year 2 efforts will continue with refinement based upon assessments of Dr. Coe.

Education and Extension

Output: We established a project web site (www.solcap.msu) which is a centralized resource for SolCAP participants and the Solanaceae breeding community. Modules to query phenotypic and genotypic data are being developed for SGN. These are cross-linked to eXtension Plant Breeding and Genomics hosting instructional modules and webinars on how to generate, analyze and integrate genomic data in breeding programs.

Outcome: The site includes project news and personnel, databases, tools, resources, project objectives and downloads.

Impact: According to Google Analytics the SolCAP website has had a total of 2,744 unique visitors from 4,423 visits in the last year with a total of 13,691 page views. Of these visits, 58.08% were new. Stakeholder feedback suggests the data downloads are helpful.

Research

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. Please see Output under Objective 2 Education and Extension for usage statistics.

The Tools section of the project website (<http://solcap.msu.edu/tools.shtml>) has links to SNPs and SSRs identified in publicly available EST data. The new toolbox design is up on SGN at <http://solgenomics.net/breeders/>. We have re-factored the database design for storing stocks, working with a number of other databases, to create a new module for the standard Chado database schema (publication in preparation). For the user interface, both an advanced search page and new detail pages have been implemented for this schema, using a new relational to object mapping model in Perl, called Bio::Chado::Schema, that is based on DBIx::Class. We have also implemented schemas for the storage of SNP data, but the data is still being generated and not yet available to SGN. In addition, SGN has loaded all currently submitted SolCAP phenotypic information. For example, using the "Advanced Search" options in the phenotypic search window (URL: http://solgenomics.net/search/direct_search.pl?search=phenotypes) and choosing 'Potato' from the organism box reveals 102 accessions with data for 20 traits each. Phenotypic data is available for over 7,000 tomato accessions and progeny.

SolCAP phenotype data were integrated into the Sol Genomics Network (SGN, <http://solgenomics.net/>) database. Tomato phenotype data from experiments in 2009 and 2010 are from four different locations with three different tomato categories (fresh market, vintage and processing) for a total of 13 experiments. A limited potato set has also been loaded (two 2009 experiments). Data is displayed in the form of pages for plant accessions, showing information on the origin, alternative synonyms, and other important metadata. Each accession has a list of member plots from all related field experiments, and a summary of the scored phenotypes. Data were first quality checked by SolCAP collaborators, transferred to SGN, and loaded into the SGN test database. In total, about 900,000 phenotypic observations were loaded from 3,459 plots. Additional quality control was performed on the data; we found that simple statistical analyses for identifying outliers were an efficient method for detecting data problems. At least a dozen problems were fixed in that way. The ways phenotypic measurements are displayed on the website were discussed among SolCAP collaborators, and a grouping of different traits according to measurement criteria was found to be the best solution for presenting the data. A limited genotype dataset is currently also being loaded. The availability of these data will allow the exploration of different breeder-friendly tools, such as downloading the raw data in certain formats and performing statistical analyses. A current concern for SGN is that potato phenotypic data has not been released to SGN. The delays are being actively addressed by the project.

The tomato Infinium chip sequences were mapped on the latest release of the tomato genome. All primer pairs mapped uniquely to this genome release using bwa mapping tool. The resulting genome coordinates were added to the tomato reference sequence map on SGN. Bulk download of the sequence information will be made available as soon as some additional quality checking of the data will be completed.

Genotyping data will be displayed on each accession page, and linked to the genetic map based on the SolCAP marker information. Our strategy ensures users will have access to the raw data based on different slicing criteria, and the ability to analyze phenotypes from the entire experimental dataset based on genotype variation.

Tutorials for using these tools have been or are under development and are linked through eXtension.org. Database development and utility are concerns. There is a group dedicated to the development of the breeder's toolbox and this group is interacting via conference calls.

Deliverables

Publications:

Ten SolCAP Newsletters (<http://solcap.msu.edu/news.shtml>)

Felcher KJ, Coombs JJ, Massa AN, Hansey CN, Hamilton JP, et al. (2012) Integration of Two Diploid Potato Linkage Maps with the Potato Genome Sequence. [PLoS ONE 7\(4\): e36347. doi:10.1371/journal.pone.0036347](https://doi.org/10.1371/journal.pone.0036347)

Merk, H. L.; Yarnes, S. C.; Van Deynze, A.; Tong, N.; Menda, N.; Mueller, L. A.; Mutschler, M. A.; Loewen, S. A.; Myers, J. R.; Francis, D. M. Reservoirs of Phenotypic and Genetic Variation Among Regionally Adapted Processing Tomato Germplasm. *Journal of the American Society for Horticultural Science* 137 (6) 427-43.

Sim, S. C.; Van Deynze, A.; Stoffel, K.; Douches, D. S.; Zarka, D.; Ganal, M. W.; Chetelat, R. T.; Hutton, S. F.; Scott, J. W.; Gardner, R. G.; Panthee, D. R.; Mutschler, M.; Myers, J.; Francis, D. M. 2012. High-density SNP genotyping of tomato (*Solanum lycopersicum* L.) reveals patterns of genetic variation due to breeding. *PLoS ONE* 7(9): e45520.

Sim, S.C.; Surstewitz, G.; Plieske, J.; Ganal, M.W.; Van Deynze, A.; Stoffel, K.; Hamilton, J.; Buell, C.R.; Zarka, D.; Douches, D.S.; Francis, D.M. 2012. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE* 7(7): e40563.

Hamilton, J. P.; Sim, S. C.; Stoffel, K.; Van Deynze, A.; Buell, C. R.; Francis, D. M. 2012. Single Nucleotide Polymorphism Discovery in Cultivated Tomato via Sequencing by Synthesis. *The Plant Genome*. Vol. 5, 17-29. 2012.

Hamilton, J.P., Hansey, C.N., Whitty, B.R., Stoffel, K., Massa, A.N., Van Deynze, A. De Jong, W.S., Douches, D.S., and Buell, C. R. 2011. Single Nucleotide Polymorphism discovery in elite North American potato germplasm. [BMC Genomics](https://doi.org/10.1186/1471-2164-12-302). 12:302

Pei, C., Wang, H., Zhang, J., Francis, D. M., Yang, W. 2011. Fine mapping and analysis of a candidate gene in tomato accession PI128216 conferring hypersensitive resistance to bacterial spot race T3. *Theor. Appl. Genet.* [Epub ahead of print]

Wang, H., Hutton, S.F., Robbins, M.D. Sim, S. C., Scott, J. W., Yang, W., Jones, J. B., Francis., D. M. 2011. Molecular mapping of hypersensitive resistance from tomato 'Hawaii 7981' to *Xanthomonas perforans* race T3. *Phytopathology.* 101(10):1217-1223.

Baldo, A. M., Francis, D. M., Caramante, M., Robertson, L.D., Labate, J. A. 2011. AlleleCoder: a PERL script for coding co-dominant polymorphism data for PCA. *Plant Genetic Resources* pp 1-3. Published online.
<http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=8333992>

Sim SC, Robbins MD, Van Deynze A, Michel AP, Francis DM. 2011. Population structure and genetic differentiation associated with breeding history and selection in tomato (*Solanum lycopersicum* L.). *Heredity.* 106: 927-935.

Robbins, M.D., Sim, S-C., Yang, W., Van Deynze, A., van der Knaap, E., Joobeur, T., and Francis, D.M. 2011. Mapping and linkage disequilibrium analysis with a genome-wide collection of SNPs that detect polymorphism in cultivated tomato. *Journal of Experimental Botany* 62(6):1831-1845 . [Epub ahead of print
<http://jxb.oxfordjournals.org/content/early/2010/12/30/jxb.erq367>]

Rodríguez GR, Muñoz, S. Anderson, C., Sim, S-C., Michel, A. Causse, M., McSpadden Gardener, B., Francis DM, van der Knaap E. 2011. Distribution of SUN, OVATE, LC and FAS in the Tomato Germplasm and the Relationship to Fruit Shape Diversity. *Plant Physiology.* 156: 275-285.

Hutton SF, Scott JW, Yang W, Sim SC, Francis DM, Jones JB. 2010. Identification of QTL associated with resistance to bacterial spot race T4 in tomato. *Theor Appl Genet.* Nov;121(7):1275-87.

Wang Y, Chen J, Francis DM, Shen H, Wu T, Yang W. 2010. Discovery of intron polymorphisms in cultivated tomato using both tomato and Arabidopsis genomic information. *Theor Appl Genet.* 121(7):1199-1207.

Teclé IY, Menda N, Buels RM, van der Knaap E, Mueller LA. solQTL: a tool for QTL analysis, visualization and linking to genomes at SGN database. *BMC Bioinformatics.* 2010 Oct 21;11:525. PubMed PMID: 20964836; PubMed Central PMCID: PMC2984588.

Oral/ Poster Presentations:

Douches: Plant Breeding and Genomics Webinar. April 2012.

Douches: ASHS Conference at the Colloquium on Breeding in the Genomics era. Miami, FL August 1, 2012.

Douches: National Association of Plant Breeders eXtension workshop. Indianapolis, IN, August 9, 2012

Douches: Potato Association of America (3 talks). Denver, CO August 12-15, 2012.

Douches: Sol 2012, Neuchatel, Switzerland August 28, 2012.

Douches: Michigan State University Marker Workshop, Sept 2, 2012.

Douches: Latin America Potato Association, Uberlandia, Brazil, September 18, 2012,.

Douches: Agronomy Society of America, Cincinnati, OH, October 23, 2012

Buell: University of Georgia Plant Science Center Retreat, October 2012

Buell: Brazilian Genomics Conference, Iguazu Falls, Brazil, September 2012

Allen Van Deynze: Plant Breeding and Genomics Community of Practice and eXtension. 2012. National Association of Plant Breeders. Indianapolis, IN, August 6-8, 2012.

David Douches, December 2010. SolCAP SNP development for potato. NCCC-84 Potato Breeding and Genetics Technical Committee meeting. Chicago, IL *40 potato breeders, students and technical staff.*

David Douches, January 2011. SolCAP Potato Translational Genomics. Michigan Potato Industry Commission Research Committee. East Lansing, MI. *20 industry members, scientists and technical staff.*

David Douches, April 2011. SolCAP Potato Translational Genomics. INIA, Lima, Peru. *30 technical staff and scientists.*

David Douches, April 2011, Chips, SNPs and GM: Potato Breeding in the 21st Century. International Potato Center, La Molina, Peru.

David Douches. August 2011. SNP development and genotyping in potato. WorldTap Marker workshop, East Lansing, MI.

Douches, David. January 15-19, 2011. Poster: SolCAP: Translating Solanaceae Sequence Diversity and Trait Variation into Applied Outcomes through Integrative Research, Education, and Extension. Plant and Animal Genome meeting XIX. San Diego, CA.

David Douches. August 2011, Wilmington, NC. SNP-based genetic maps: Linkage and QTL analysis. PAA Workshop.

David Francis. August 2011, Wilmington, NC. Breeding in a Genomics Era. PAA Workshop.

Walter De Jong. SolCAP phenotyping: a brief overview. August 14, 2011, Wilmington NC. SolCAP workshop at Potato Association of America Annual Meeting.

Walter De Jong. Developing SNP Markers in Elite Germplasm for Applied Potato Breeding. July 26, 2011, Xining, China. Annual Meeting of Chinese Potato Breeders.

Liabeuf, D., Sim, S. C., Francis, D. 2011. Screening of *Solanum pimpinellifolium* accessions for resistance to *Xanthomonas gardneri*. Proceedings of the 26th Annual Tomato Disease Workshop. Oct. 11-13, Cornell University, NY.

Francis, D. M, Van Deynze, A., Hamilton, J., Sim, S.C. De Jong, W. Douches, D., Buell, C. R. 2011. The SolCAP Project: filling the gap between genomic information and breeding through massive SNP discovery and intensive germplasm phenotyping. Invited lecture, XVII Eucarpia Meeting, Section Vegetables, Working Group Tomato, Malaga, Spain.

Francis, D.M., Monforte, A. 2011. Genetic markers development and utilization of germplasm resources in the genomics era. Chairpersons: David M. Francis (Ohio State University, USA) and Antonio Monforte (IBMCP-UPV-CSIC, Spain). XVII Eucarpia Meeting, Section Vegetables, Working Group Tomato, Malaga, Spain.

Francis, D. M. 2011. Quality is color, and color is lycopene, or not. 2011 Tomato Breeders' Roundtable Meeting, Oct. 9-10, Ithaca, NY.
http://tgc.ifas.ufl.edu/Presentations/6%20Francis%20Quality_TBRT2011_Final.pdf;
<http://tgc.ifas.ufl.edu/2011TBRTAbstracts.pdf>

Labate, J. A., Panthee, D.R., McGrath, M., Francis, D. M., Breksa, A., Robertson, L. D. 2011. Phenotypic and molecular variation in 44 vintage tomato varieties. 2011 Tomato Breeders' Roundtable Meeting, Oct. 9-10, Ithaca, NY.
<http://tgc.ifas.ufl.edu/2011TBRTAbstracts.pdf>;
<http://tgc.ifas.ufl.edu/Presentations/2%20Labate-2011-TBRT.pdf>

Buell, C. R, Hamilton, J. P., Hansey C., Whitty, B. R., Brett R. Whitty, Van Deynze, A., Francis, D. M., De Jong, W, Douches, D. 2011. SNP Development In Elite North American Potato Germplasm. P439. Plant & Animal Genomes XIX Conference

C. Robin Buell. August 2011. Using the potato genome sequence. Potato Association of America, Wilmington, NC.

Bradford, K.J. 2011. Plant Breeding education through eXtension. Plant Breeders and Biotechnology Committee, California Seed Association, Oct. 12th, 2011. Monterey, CA,

Van Deynze, A. Douches, D. Buell, C.R. Francis, D. Stone, A. Mueller, L. 2011. Allele Diversity in the Solanaceae as measured by whole transcriptome sequencing. Genomics, Based breeding in forest trees workshop. June 22-24, 2011, Davis, CA.

Van Deynze, A. 2011. Applying Genomics in Plant Breeding. PBI200A. University of California, Davis course. Oct. 21st, 2011. Davis, CA.

Van Deynze, A. 2011. Application of biotechnology in plant breeding. Gates Foundation. June 28th, 2011, Davis, CA.

Lukas Mueller. SGN Breeder's Toolbox and SolCAP. Eucarpia Meeting (tomato). Màlaga, Spain (April 11-13, 2011).

Merk, Heather, David Francis, Alex Stone, John McQueen, Roger Leigh, Allen Van Deynze, Jeannette Martins, Shawn Yarnes, Deana Namuth-Covert, David Douches, Kimberly Felcher, Kelly Zarka, and Walter De Jong 2011. The Plant Breeding and Genomics Community on extension: Putting Research into Practice. March 9, 2011. A poster presentation at the 58th Soft Wheat Quality Laboratory Research Review Conference, Wooster, OH, p24.

Merk, Heather L. Putting Research into practice: The Plant Breeding and Genomics. Community of Practice on Extension. National Association of Plant Breeders Conference. May 23-25, 2011. College Station, TX.

Van Deynze, Allen Theresa Hill, Jim Prince, Shawn Yarnes, Jareerat Chunthawodtiporn, William Rehrig, Sebastian Reyes Chin-Wo, Hamid Ashrafi and Alex Kozik. 2011 Development and Application of Genomic Tools in Pepper. Solanaceae resistance breeding technologies, genetics and genomics. Chiang Mai, Thailand. Feb. 17-18th, 2011.

C. Robin Buell. January 2011. Allele diversity in the *Solanaceae* as measured by whole transcriptome sequencing. Plant and Animal Genome Meeting, San Diego, CA.

C. Robin Buell. June 8, 2011. Allele diversity in the Solanaceae as measured by whole transcriptome sequencing. Plant Genomics: Applications to Plant Breeding, INIA-La Platina, Santiago, Chile.

David Douches, June 27, 2010. SNP Development for Elite Potato Germplasm. European Association of Potato Research. Wageningen, Netherlands. *120 potato breeders and geneticists*.

David Douches, August 4, 2010. SNP Development for Elite Potato Germplasm. Chinese Potato Association, Changbei, China. *150 potato scientists, breeders and geneticists*.

David Douches, September 4, 2010. Translational Genomics: SNP Development for Elite Potato Germplasm. WorldTap Molecular Marker Workshop. East Lansing, MI. 15 *international graduate students and scientists.*

Sim, S. C., Wang, H., Robbins, M. D., Yang, W., Francis, D. M. 2010. QTL Validation Of Bacterial Spot Resistance In Tomato Based On Association Analysis In A Complex Breeding Population. P445. Plant & Animal Genomes XVIII Conference

Robbins, M. D., Sim, S. c., Yang, W. van Deynze, A., van der Knaap, E. Joobeur, T., Francis, D. M. 2010. Genome-Wide Analysis Reveals Different Patterns Of Linkage Disequilibrium Among Market Classes Of Tomato P442. Plant & Animal Genomes XVIII Conference

Sim, S. C., Huarachi-Morejon, N., Van Deynze, A., Hamilton, J., Buell, C. R., Grass, A., Labuza, A., Francis, D. M. 2011. Genetic Association Of SP5 With Plant Height, Inflorescence Architecture, And Yield In Tomato. P434. Plant & Animal Genomes XIX Conference

Van Deynze, Allen. SolCAP 2010 Tomato Workshop, November 16, 2010 University of Florida, Wimauma, Florida, in conjunction with the Tomato Disease Workshop. Next Generation Sequencing.

Francis, David. SolCAP 2010 Tomato Workshop, November 16, 2010 University of Florida, Wimauma, Florida, in conjunction with the Tomato Disease Workshop. Accessing Sequence Resources.

Merk, Heather. SolCAP 2010 Tomato Workshop, November 16, 2010 University of Florida, Wimauma, Florida, in conjunction with the Tomato Disease Workshop. Tomato Genome Browser (GBrowse) for Plant Breeders.

Francis, David. SolCAP 2010 Tomato Workshop, November 16, 2010 University of Florida, Wimauma, Florida, in conjunction with the Tomato Disease Workshop. BioInformatics 101.

Van Denyze, Allen. SolCAP 2010 Tomato Workshop, November 16, 2010 University of Florida, Wimauma, Florida, in conjunction with the Tomato Disease Workshop. Working with Tomato Infinium Genotyping Data.

Sim, Sung-Chur. SolCAP 2010 Tomato Workshop, November 16, 2010 University of Florida, Wimauma, Florida, in conjunction with the Tomato Disease Workshop. Downstream Analysis with SNP Markers.

Van Denyze, Allen. SolCAP 2010 Potato Workshop, August 15, 2010, Oregon State University, Corvallis, Oregon, in conjunction with the Potato Association of America 94th Annual Meeting. Working with Infinium Genotype Data Presentation and Webinar

Hackett, Christine. SolCAP 2010 Potato Workshop, August 15, 2010, Oregon State University, Corvallis, Oregon, in conjunction with the Potato Association of America 94th Annual Meeting. Linkage Analysis and QTL Mapping in Tetraploids Slide Presentation and Webinar

John Hamilton. September, 2010. SNP Identification from Potato and Tomato RNA-Seq. SOL201, Dundee, Scotland

C. Robin Buell. August 15, 2010. Using the potato genome sequence. Potato Association of America, Corvallis, Oregon.

Van Deynze, Allen. June 23rd, 2010. Too much of a good thing: Challenges with high resolution genetic mapping. Plant Science Seminar, UC Davis. 30 Plant Scientists

Shattuck, Jamie California Seed Association. Biotech Committee. Research at the Seed Biotechnology Center. September 14th, 2010. ~30 plant scientists.

Van Deynze, Allen. Feb 18th, 2010. The use of markers in plant breeding. UC Davis Plant Breeding Academy. 21 Plant Breeders.

Bradford, Kent, May 26th, 2010. Research at the Seed Biotechnology Center. Board of Directors, Nunhems Inc.- *10 Seed industry professionals.*

Francis, David, Allen Van Deynze, John Hamilton, Matthew Robbins, Sung-Chur Sim, Walter De Jong, David Douches, Robin Buell. 2010. Next-Generation Sequencing of the Tomato Transcriptome: A resource for SNP discovery, high-throughput genotyping and translational research. Colloquium: Applications of Next-Generation Sequencing to Horticultural American Society of Horticultural Science (ASHS) Annual Conference, Palm Desert, CA.

Merk, Heather 2010. The Tomato Genome Browser for Plant Breeders. November 16, 2010. Tomato Disease Workshop, University of Florida IFAS Gulf Coast Research and Education Center, Wimauma, FL.

Francis, David, Allen Van Deynze, John Hamilton, Matthew Robbins, Sung-Chur Sim, Walter De Jong, David Douches, Robin Buell. 2010. Next Generation Sequencing of the Tomato Transcriptome. C04.3 International Society of Horticultural Science (ISHS), International Horticulture Congress (IHC), Lisbon, Portugal.

Lukas Mueller. SGN Breeder's Toolbox and SolCAP. ASPB meeting Montreal (July 31-August 4, 2010).

Lukas Mueller. SGN Breeder's Toolbox and SolCAP. SOL2010, Dundee (Sept 4-9, 2010).

Lukas Mueller. SGN Breeder's Toolbox; Computer demonstration at PAG (San Diego, January 15-20, 2011).

Community Resources Generated:

SolCAP cDNA Libraries Sequenced and High Throughput SNP analysis From existing potato and tomato sequence databases, we have identified candidate SNPs. To expand our SNP discovery, we generated 5 Gb of potato sequence data from normalized cDNA of cultivated germplasm. 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 from each of the three varieties; these have been used for SNP discovery. For the tomato transcriptome, SolCAP has generated ~15 Gb of sequence from six varieties. The potato sequence and SNP data are available in NCBI SRA and on the SolCAP project site, respectively. The tomato data will be publically available after the acceptance of the manuscript.

SolCAP Germplasm Panels

Potato: Elite potato germplasm was contributed from 16 programs across the U.S. as well as six international programs. The panel consists of 250 potato lines currently used by the community. The panel was genotyped with SNPs and 220 lines were phenotyped at three diverse production environments across the US in 2010 and 2011. In addition, a tetraploid mapping population of 200 progeny was both genotyped and phenotyped (three environments in 2010 and 2011, 2 environments in 2009) for key tuber and agronomic traits. 2009 and 2010 data can be found at SolCAP website. HPLC-based post-harvest sugar analysis of the 2010 and 2011 data has been. The 2011 data will be uploaded to the SolCAP website after final QC.

Tomato: A 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 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.

We have identified putative SNPs (eSNPs) 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. We have also 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 same website.

SNP Development for the Illumina Infinium Platform for potato and tomato The SolCAP team led a consortium with our fellow Solanaceae scientists to develop potato

and tomato Infinium SNP arrays through Illumina for the interrogation of SNPs in the respective genomes. These arrays are being used worldwide by the potato and tomato breeding and genetics communities. (See Obj. 4 and 5)

SolCAP workshops have been organized and conducted for the tomato and potato communities in 2009, 2010, 2011 and 2012 to provide a foundation for these communities to adopt genomics based analysis. SolCAP workshops generated educational content specifically designed for tomato and potato breeders on topics such as: use of the potato genome sequence, phenotypic data from association panels, linkage and QTL analysis and breeding in a genomics era, phenotyping to gain maximum benefit from genetic marker methods, use of software and online tools for genomics-based breeding, quantitative methods for managing and analyzing related data and working with SolCAP generated SNP data. Live webinars were conducted during the workshops to expand community outreach.

PBGworks SolCAP has developed PBGworks as a collaborative workspace developed using open source software. <http://pbgworks.hort.oregonstate.edu/>. The target audience is 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 and 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.

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. We continually update the site throughout the year. Visit us at <http://solcap.msu.edu>.

SGN website serves as a portal to SolCAP data and integrates it with extensive genomic and phenotypic data from other sources, and makes the SGN tools available to breeders through the dedicated breeder's toolbox (<http://solgenomics.net/breeders/>).

Training:

Alicia Massa, Ph.D. MSU Buell lab

Cory Hirsch, Ph.D. MSU Buell lab

Matt Zuelke, MSU undergraduate

Alison McKenna, technician, MSU potato breeding and genetics program

K. Haynes and L Wanner, USDA/ARS

Sarah Braun and Cinthya Zorilla, Ph.D. candidates at University of Wisconsin, Madison

Kim Felcher, Ph.D. MSU Potato Breeding and Genetics Program, works on marker database, germplasm panel and SNP data analysis in Douches lab.

Daniel Zarka, Post Doc MSU Potato Breeding and Genetics Program, works on SNP genotyping with the Infinium system.

Kelly Zarka, technical staff at MSU Potato Breeding and Genetics Program, was trained in SNP genotyping and genotype by sequencing.

Joe Coombs, technical staff at MSU Potato Breeding and Genetics Program, is participating in SNP data analysis.

Candice Hirsch, Post Doc, MSU Plant Biology, worked on bioinformatics analysis of tomato and potato sequences.

Amanda Labuza, Rensselaer Polytechnic Institute, Undergraduate intern in Francis lab (December 2010-January 2011). Co-funded by the OARDC Research Internship Program.

Gabriel Abud, College of Wooster, Wooster, Ohio, Undergraduate intern in Francis lab (June-Aug 2010). Co-funded by the OARDC Research Internship Program.

Caleb Orchard, Grove City College, Pennsylvania. Undergraduate intern in Francis lab (June-Aug 2010). Co-funded by the Summer Undergraduate Research Experience (SURE) internship program.

Jamie Shattuck, postdoc UC Davis manage and organize tomato genomics workshop.

Jacob Schleppi, Cleveland Institute of Art, Undergraduate intern in Francis lab, June-August 2009 and 2010.

Laura Bonefeld, Michigan State University, Undergraduate bioinformatics intern in Buell lab.

Andy Murley: Michigan State University, Undergraduate student trained in DNA extraction and PCR in Douches lab.

Kevin Stoffel: UC Davis, Senior research associate, was trained in normalized cDNA development in Van Deynze lab.

Lee Heekyoung: UC Davis. Lee is a visiting scientist from Korea that was trained in molecular marker validation and analysis in Van Deynze lab.

Sebastian Reyes Cin-Wo, UC Davis. Sebastian is an undergraduate student from Costa Rica who was trained in tomato field testing, bioinformatics analysis of sequence data and primer design in Van Deynze lab.

Shawn Yarnes: UC Davis. Shawn is a postdoc in the Van Deynze lab. She is managing the research trials for tomato in 2010.

Selene Clay: UC Davis. Selene is a high school student in the Van Deynze lab who has been trained in genotyping and analysis of tetraploid segregation in potato using the data generated by SolCAP.

Doug Maxwell, University of Wisconsin, Madison-Mapping of virus resistance in tomato

Jay Scott, University of Florida, Mapping of virus resistance in tomato

Sam Hutton, University of Florida, Mapping of virus resistance in tomato

Martha Mustchler, Cornell University, mapping disease resistance in tomato

Dilip Panthee, North Carolina State University, mapping disease resistance in tomato

Neelima Sinha and Julin Maloof; UC Davis-Deciphering leaf traits in tomato

Leonie Moyle, Indiana State University, Mapping fruit traits in tomato

Collaborations:

As noted above, there are strong collaborative arrangements among the public potato and tomato breeding and genetics community.

Shelly Jansky (USDA/ARS at University of Wisconsin) and her staff and students have been collaborating on the SNP genotyping of a diploid mapping population.

Paul Bethke (USDA/ARS at the University of Wisconsin) has been collaborating on the HPLC sugar analysis and the SNP genotyping of a tetraploid mapping population.

Craig Yenchu (North Carolina State University) has been collaborating on the SNP genotyping of a tetraploid mapping population.

Jiwan Palta (University of Wisconsin) has been collaborating on the SNP genotyping of a diploid mapping population.

Hermann van Eck and Chris Maliepaard, Wageningen University – SNP dosage calling.

Great Lakes Vegetable Working Group (GLVWG) is a network of 149 (with 45 active) extension specialists from 7 states and Ontario, Canada that are participating in the

evaluation of “heirloom” tomato varieties. The collaboration added two full sites to the vintage variety evaluation and 10 partial sites.

Sanwen Huang, Potato Genome Sequencing Consortium, Beijing Genomics Institute, Chinese Academy of Agriculture Sciences.

Liping Jin, Potato Breeding and Genetics, Institute of Flowers and Vegetables, Chinese Academy of Agriculture Sciences.

Glenn Bryan, Scottish Crop Research Institute, Dundee, Scotland.

Dan Ronis/Aliya Momotaz, Frito-Lay Potato Breeding, SNP genotyping.

Merideth Bonierbale, Geneticist at the International Potato Center, Lima, Peru is collaborating germplasm diversity evaluation.

Luis E. Rodriguez-Saona, Ohio State University Food Science and Technology, is collaborating on analysis of metabolites in the tomato collection.

Campbells Breeding. Collaborator evaluated processing tomato panel in 2009 and 2010

Martin Ganal, Trait Genetics, Gatersleben, Germany provided SNP sequences and Design advice for tomato Infinium chip.

Mathilde Causse, INRA, Avignon, France provided SNP sequences for tomato Infinium chip

Cindy Lawley, Illumina co-developed consortiums for Potato and Tomato arrays and Genome Studio software.

APPENDIX

Project Evaluation

At the January 2009, 2010, 2011 and 2012 Plant and Animal Genome (PAG) meetings we held our advisory committee meetings. Verbal feedback and written reports from the advisory group have been used by the executive committee to further refine project activities and 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 has provided 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 SolCAP workshops held at the Potato Association of America 2009, 2010, 2011, and 2012 annual meetings, the 2009 Tomato Breeders Round Table (TBRT), the 2010 and 2011 Tomato Disease Workshops, and the SolCAP colloquium held as part of the 2012 American Society for Horticultural Science conference. Evaluation surveys were used to gather feedback on the specific workshop sessions as well as the workshop as a whole, including the extent to which advertised content was covered, learning objectives were met, presentation of theory and application was balanced, and participants gained new knowledge, as well as whether the technical level of presentations was appropriate for the audience and whether participants would recommend the workshop to others. Participants in these workshops for potato and tomato professionals rated the workshop sessions as being relevant to their work and effectively presented. SolCAP events during 2011 and 2012 were also broadcast in webinar format for participants who were not able to travel to the meetings, and feedback from these online participants was included in the evaluation data.

During 2011 and 2012, an additional series of educational webinars was developed and provided via eXtension.org through a collaboration among NIFA CAP grants led by SolCAP. An online questionnaire was used to obtain feedback from participants in these webinar sessions. Through November 2012, responses from eight webinars indicated that these educational events were viewed by participants as highly effective and aimed at an appropriate technical level for the audience.

Evaluation feedback from SolCAP-related university courses was also analyzed and reported during 2010. After the courses, participants reported significantly increased confidence (self-efficacy) in their ability to apply genetic markers in their work, and in their ability to communicate or teach others about marker applications. The clarity and effectiveness of the course materials and presentations were rated as “very effective” or “OK” by 80 to 100 percent of participants. Student comments in response to open-ended questions were very positive. In addition, participants provided suggestions for course improvements, as well as suggestions for focusing genomic research on particular practical problems in crop improvement.

Detailed reports of these findings as well as participant recommendations for future educational events were prepared by the external evaluator and are being used to plan future education and extension activities.