

# The Contribution of the Solanaceae Coordinated Agricultural Project to Potato Breeding

D. Douches · C. N. Hirsch ·  
N. C. Manrique-Carpintero · A. N. Massa ·  
J. Coombs · M. Hardigan · D. Bisognin ·  
W. De Jong · C. R. Buell



© European Association for Potato Research 2014

**Abstract** Potato variety development in the USA involves extensive interaction between public sector scientists and state and national grower organizations to test and commercialize new varieties. Historically, breeding and germplasm development efforts have relied upon phenotypic evaluations to select and advance germplasm. The goal of the US Department of Agriculture (USDA)-funded Solanaceae Coordinated Agricultural Project (SolCAP) was to translate genomic resources into tools that can be used by breeders and geneticists. A major outcome of this project was the development of a genome-wide single nucleotide polymorphism (SNP) array that can be used to evaluate elite potato breeding germplasm. This array was utilized to genotype numerous biparental tetraploid and diploid populations, a diversity panel, and a core collection of *Solanum* species. The SNP array provides a marker density sufficient to generate genetic maps to identify numerous quantitative trait loci (QTLs) for agronomic, quality, and disease resistance traits. Many new diploid and tetraploid genetic maps have been constructed and used to identify numerous QTLs for important traits. We used SNP markers to assess relationships among germplasm, fingerprint varieties, and identify candidate genes. The Infinium 8303 Potato Array provides a common set of SNP markers that can reliably be used for mapping, germplasm assessment, and

---

D. Douches (✉) · N. C. Manrique-Carpintero · A. N. Massa · J. Coombs · D. Bisognin  
Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI 48824,  
USA  
e-mail: douchesd@msu.edu

C. N. Hirsch  
Department of Agronomy and Plant Genetics, University of Minnesota, Saint Paul, MN 55108, USA

M. Hardigan · C. R. Buell  
Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA

W. De Jong  
School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

fingerprinting. This array has also been a useful tool to advance our understanding of the potato genome. Furthermore, breeders are mapping QTLs across numerous populations that will expand our understanding of economically important traits and lead to marker-assisted selection and breeding, and ultimately improved varieties.

**Keywords** Breeding · Genetic maps · Quantitative trait loci · Single nucleotide polymorphism · Translational genomics

## Introduction

The Solanaceae family is diverse, containing economically important species that collectively provide substantial amounts of starch, sugars, vitamins, and antioxidants to our diet. One member of this family, potato (*Solanum tuberosum* subsp. *tuberosum*) accounts for \$3.7 billion in farm value and is grown on 1.13 M acres in the USA (USDA 2013). The cultivated potato is an autotetraploid ( $2n=4x=48$ ) and is the world's most important dicot food crop ranking fourth in overall production, after rice, wheat, and maize. Unlike many other crops, the majority of US potato breeding is still conducted in the public sector. There are 11 public sector breeding programs along with four US Department of Agriculture–Agricultural Research Services (USDA-ARS) geneticists focused on varietal development, germplasm enhancement, taxonomy, and curation of germplasm. The public programs have a strong history of partnership with regional industries to develop and release varieties for the processing and fresh markets. Breeding activities are directed toward improved germplasm with biotic resistances combined with improved quality and nutritional components. Breeders in the USA also have a history of accessing germplasm from the National Plant Germplasm System (NPGS, NRSP-6) for these quality and resistance traits.

The US potato breeding challenge has been to combine market-driven quality with agronomic performance and host plant resistances needed by growers. Potato variety development in the USA involves extensive interaction between public sector scientists, state potato grower organizations, and the US Potato Board (USPB) to test and commercialize new varieties (e.g., USPB Snack Food Trials, National Chip Processing Trials, and National Fry Processing Trials). The needs of the processing market drive the national breeding focus while specific disease and insect resistances tend to have regional importance. Moreover, the processing and host resistance target traits for breeders have been expanding (e.g., acrylamide reduction, as well as psyllid, zebra chip, *potato virus Y* (PVY), potato mop top virus, powdery scab, and tobacco rattle virus resistance). For many decades, variety development, informed by stakeholder priorities, has been a key component in creating slow but positive change for the overall potato industry in the USA.

One of the challenges of breeding potato, a heterozygous, clonally propagated autotetraploid crop, has been limited marker development. Until recently, breeding and germplasm development efforts have relied upon phenotypic evaluations to select and advance germplasm. A critical need in the Solanaceae breeding and genetics community was access to a sufficient number of markers polymorphic in elite breeding germplasm (Van Deynze et al. 2007). In potato, mapping studies (at the diploid and occasionally tetraploid levels) have been conducted since the late 1980s, but little

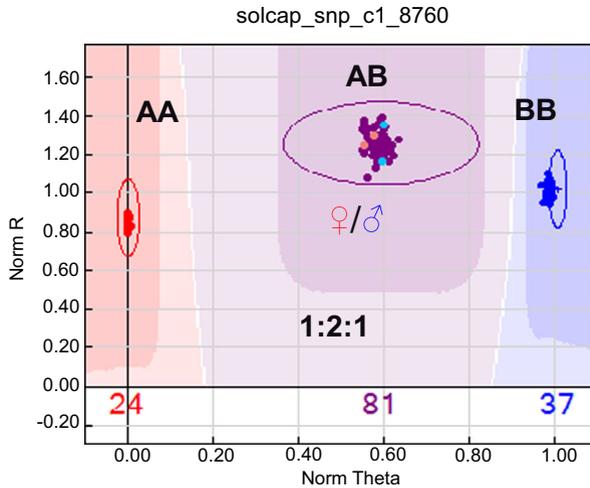
marker-assisted selection (MAS) is yet practiced in varietal breeding. Markers are currently being used to select for PVY resistance (Kasai et al. 2000), golden nematode resistance (own observation), verticillium wilt resistance (Bae et al. 2008), and late blight resistance (Colton et al. 2006). However, the market-limiting traits that are of most interest to growers and processors remain genetically uncharacterized. The potential for markers to analyze whole genomes, to improve estimation of genetic variance, recover elite backgrounds, and select superior varieties has not been fully realized in potato.

The Solanaceae Coordinated Agricultural Project (SolCAP) project (2009–2014) was a USDA-funded project that linked together people from public institutions, private institutions, and industries dedicated to the improvement of potato and tomato (*Solanum lycopersicum*). Through the SolCAP project, we envisioned exploiting homology between genomes to positively impact applied breeding in potato and tomato, the two most important vegetable crops in the Solanaceae. The vision of SolCAP was to move translational genomics beyond commodity boundaries toward an emphasis on taxonomic groups and DNA sequence homology so that knowledge and resources could be leveraged across species. SolCAP integrated research, education, and extension and devoted efforts to workshops, Web sites (<http://solcap.msu.edu> and <http://solanaceae.plantbiology.msu.edu>), and webinars ([https://www.extension.org/plant\\_breeding\\_genomics](https://www.extension.org/plant_breeding_genomics)), as well as hands-on training for the breeding and genetics community. The goal of SolCAP was to reduce the gap between genomics and breeding and provide infrastructure to link variation of SNPs in genes to valuable traits and to translate genomic resources into tools that can be used by breeders and geneticists.

## Outcomes of SolCAP

A major outcome of this project was the development of a genome-wide SNP array that can be used to evaluate elite potato breeding germplasm. The Infinium 8303 Potato Array contains 3018 SNPs from candidate genes, 536 from previously identified genetic markers, and 4749 selected from 69,000 high confidence SNPs for maximum genome-wide coverage (Hamilton et al. 2011). All of the SNPs were identified from the transcriptome of six cultivars and can be found on the Potato Genomics Resource, SpudDB, Web site (<http://solanaceae.plantbiology.msu.edu>). For selecting the candidate genes, community discussions lead to a consensus that consumer and market quality traits would provide strong and mutually beneficial cross-commodity interaction. Sugar, carbohydrate-related phenotypes, and vitamin content represented high impact traits common to both tomato and potato.

A set of over 800 genotypes representing the SolCAP diversity panel (Hirsch et al. 2013), two tetraploid, and two diploid mapping populations (Felcher et al. 2012) were used to define SNP three cluster calls (AA, AB, BB) (Fig. 1). Samples were loaded in a single project in Illumina GenomeStudio software, and used to identify patterns of theta value boundaries for cluster genotype calls of these codominant SNP markers. Of the 8303 SNPs, 7157 SNPs have consistently high quality for three-cluster calling. Additionally, out of the 8303 SNPs, 4860 SNPs can be used for tetraploid genotyping in a dosage-sensitive manner (five-cluster calling: AAAA, AAAB, AABB, ABBB, BBBB) (Fig. 2).

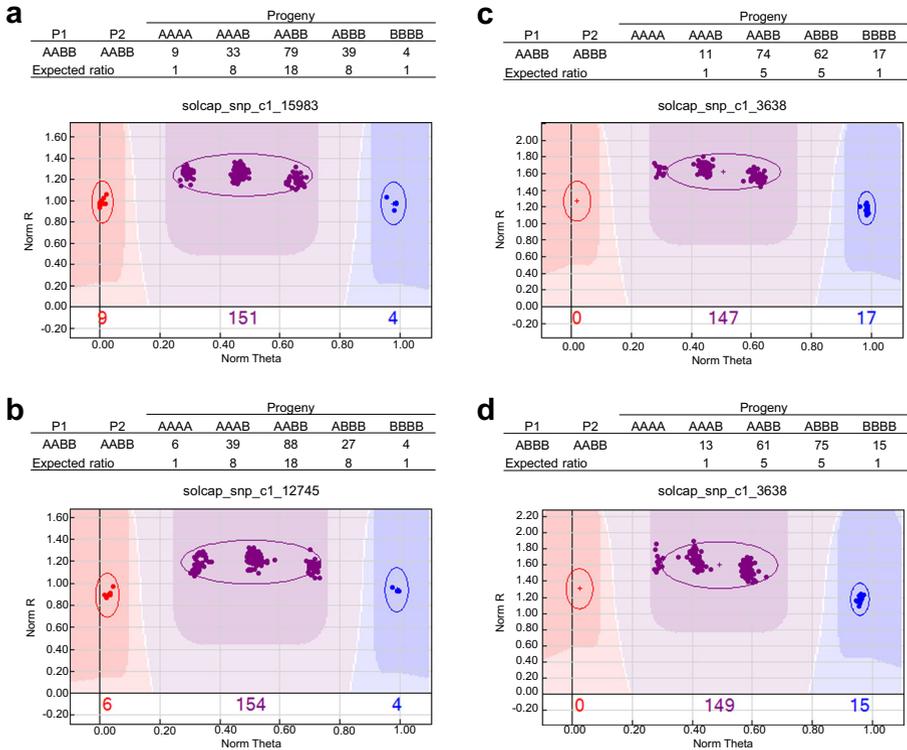


**Fig. 1** Three-cluster single nucleotide polymorphism (SNP) genotype calling. GenomeStudio SNP graph showing a SolCAP SNP name on the top, three genotype clusters (homozygous AA and BB and heterozygous AB) for a SNP that is heterozygous (AB) in both parental lines (♀/♂) of a segregating diploid population, and shaded color boundaries for each genotypic class based on values of theta scores predefined by cluster pattern of more than 800 samples

The Infinium 8303 Potato Array provides a marker density sufficient to generate genetic maps to identify numerous quantitative trait loci (QTLs) for agronomic, quality, and disease resistance traits (Fig. 3). This potato SNP array has been utilized to genotype numerous biparental tetraploid and diploid populations (Felcher et al. 2012), a diversity panel (Hirsch et al. 2013), and a core collection of tuber-bearing *Solanum* species (Hardigan et al. 2014). In addition, multienvironment replicated phenotypic data has been collected on the diversity panel and a russet mapping population (Premier Russet × Rio Grande) for genome-wide association studies (GWAS) and QTL analysis, respectively.

A second major outcome of SolCAP was SNP genotyping of a potato diversity panel that provided a retrospective view of North American potato breeding in the twentieth and twenty-first centuries (Hirsch et al. 2013). Potato germplasm for the diversity panel was focused on specific market classes (russet frozen processing, round white chip processing, etc.), and as such encompasses a broad genetic base. All active US breeding programs contributed varieties and advanced lines of value for the SolCAP diversity panel. A small selection of international germplasm was included for comparative analyses. The collection included the leading US varieties, clones of historical significance, and introgressions from other *Solanum* species. The collection exhibits substantial variation in processing ability, starch content, appearance, and resistance to major diseases. Also included was a small core set of *Solanum* species and accessions that have been introgressed into tetraploid germplasm to provide a taxonomic perspective. Currently, GWAS are being conducted using the cultivated germplasm in the diversity panel. This collection of germplasm from the diversity panel is maintained in tissue culture at Michigan State University for future phenotyping studies.

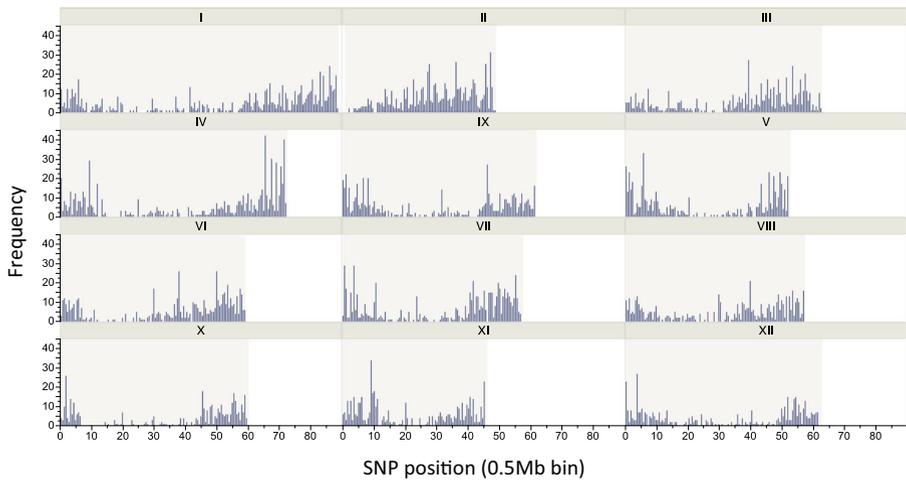
Based upon the analysis of Hirsch et al. (2013), a century of potato breeding has resulted in clear genetic differentiation of germplasm within market classes in the North



**Fig. 2** Dosage cluster single nucleotide polymorphism (SNP) genotype calling. Five-cluster (**a, b**) and four-cluster (**c, d**) single nucleotide polymorphism (SNP) genotype calling. GenomeStudio SNP graphs showing the segregation of SNPs in a tetraploid mapping population Premier Russet × Rio Grande. SolCAP SNP identification is above the graph and four or five genotype clusters are shown in the graph (homozygous AAAA and BBBB, and heterozygous AAAB, AABB, and AAAB). The shaded color boundaries in the graph and cluster counts above the x axis are based upon Genomestudio software three cluster calling

American germplasm. Interestingly, this study also showed that SNP heterozygosity levels in cultivated potato clones has not changed over this period (approximately 51%). However, there has been enrichment for some genes critical to the yellow-fleshed and chip-processing markets. The next steps are to use the SolCAP diversity panel as a reference to compare to other germplasm pools.

NRSP-6, US Potato Genebank used the Infinium 8303 Potato Array to genotype a 74 accession panel representing the broad diversity of *Solanum* section *Petota* from 25 species of the NRSP-6 US Potato Genebank, Sturgeon Bay, WI. This collection included South and North American diploid and tetraploid species, as well as diploid and tetraploid landrace germplasm. Despite the bias of identifying SNPs in the cultivated tetraploid germplasm for the SNP array, there was a high call rate on most species. The Mexican diploid species (*Solanum jamesii*, *Solanum bulbocastanum*, and *Solanum pinnatisectum*), members of the tertiary gene pool, had the highest no-call rate. Using 3846 SNPs, clear differentiation of the species accessions was observed and offered a new taxonomic view of potato germplasm (Fig. 4). Furthermore, when compared to the SolCAP diversity panel, the *Solanum* sect. *Petota* diversity panel was genetically distinct. Analysis of allele frequencies at some SNP loci between the

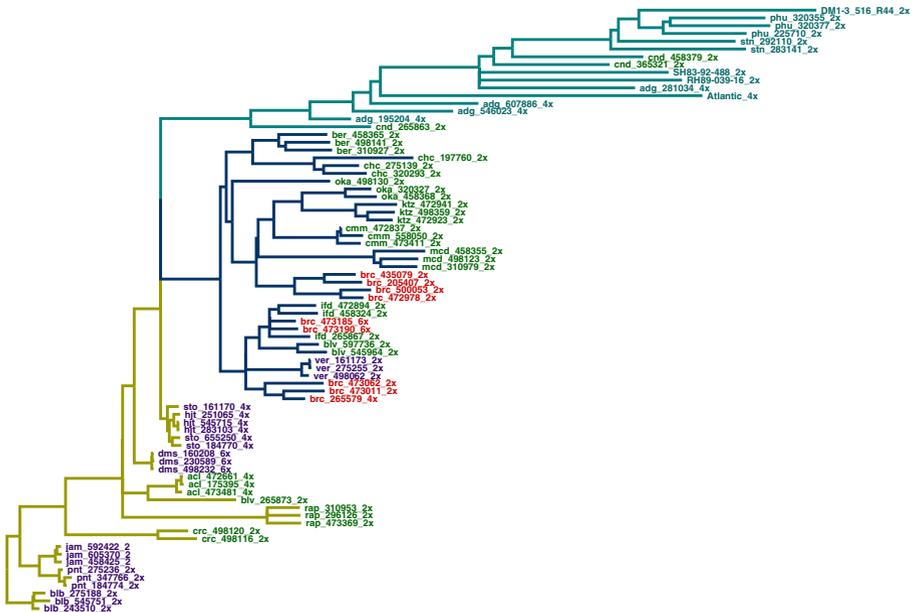


**Fig. 3** Single nucleotide polymorphism (SNP) frequency distribution across potato chromosomes. Frequency is expressed in number of occurrences/500 kbp. A total number of 7157 SNPs from the Infinium 8303 Potato Array were plotted against physical position of the Potato (*Solanum tuberosum* group Phureja DM1-3) PGSC v4.03 Pseudomolecules (<http://potato.plantbiology.msu.edu/cgi-bin/gbrowse/potato/>) to assess SNP coverage on chromosomes 1 through 12 (shaded area). The remaining SNPs (1146) were not graphed because they were either not mapped to the pseudomolecules or mapped to more than one position on the pseudomolecules

*Solanum* sect. *Petota* diversity panel and cultivated tetraploid germplasm from the SolCAP diversity panel identified SNPs with divergence between cultivated potato and its tuber-bearing wild relatives. These SNPs are associated with carbohydrate metabolism and tuber development genes, both of which are related to potato domestication.

NRSP-6, US Potato Genebank the Infinium 8303 Potato Array to assess this set of genome-wide markers for predicting heterogeneity and screening efficiency in wild potato species by using a set of four species breeding system (*Solanum verrucosum*, diploid self-compatible; *S. jamesii*, diploid outcrossing; *Solanum sucrense*, facultative outcrossing autotetraploid; and *Solanum fendleri*, selfing allotetraploid). Samples were constructed from population DNA bulks of 25 plants. When considering among-population versus within-population partitioning of diversity, SNPs usually distinguished the species as expected according to their known breeding system concluding that SNP markers can be used to help breeders and Genebank managers better understand patterns in potato germplasm diversity.

SolCAP has SNP genotyped numerous diploid and tetraploid populations for map construction, QTL analysis, and germplasm comparisons (Table 1). Felcher et al. (2012) used two diploid populations to validate the Infinium 8303 Potato Array. The linkage maps from the two diploid populations were compared with the assembled potato genome sequence. Both populations used the doubled monoploid reference genotype *Solanum tuberosum* Group Phureja Clone DM1-3 516R44 (DM1-3) as the female parent but had different heterozygous diploid male parents (RH89-039-16 (RH) and 84SD22). Use of DM1-3 as the female parent allowed one-way pseudotestcross populations to be used for the first time in generating genetic maps of potato. Between the two populations, over 3500 markers were mapped covering the majority of genome sequence length. This genome coverage by the SNP-based maps was greater than the ultra-high density AFLP maps previously generated (Van Os et al. 2006). Markers with



**Fig. 4** Phylogenetic tree of *Solanum* sect. *Petota* diversity panel genotypes based on Nei's (1972) genetic distance. Groups: cultivated group (I), primarily North American species group (II), and primarily South American group (III). Taxon colors: South American species (green), North American species (purple), *S. brevicaulis* complex species (South American) (red), and cultivated gemplasm (turquoise) (from Hardigan et al. 2014)

distorted segregation ratios occurred in blocks in both linkage maps, accounting for up to 9% of mapped markers. Furthermore, the markers with distorted segregation ratios were unique to each population. The high degree of concordance between the two linkage maps and the 650 Mb of the 12 pseudomolecules genome assembly version 2.1.10 demonstrated both the quality of the potato genome sequence and the functionality of the Infinium 8303 Potato Array.

The broad genome coverage of the Infinium 8303 Potato Array will enable numerous downstream applications. Manrique-Carpintero et al. (submitted) conducted QTL analysis with the diploid pseudotestcross population between DM 1-3 and RH. The population was evaluated for yield, food quality, and plant development traits in 2012 and 2013. Sixteen different QTLs were identified of which a major QTL at 16.8 cM and 3.7 Mb on chromosome V explained between 21.2 and 75.7% of variance for tuber number/plant, specific gravity, vine vigor, maturity, and tuber end rot. The QTL on chromosome V for maturity was 0.7 Mb from PGSC0003DMG400018408, the major regulator gene of plant maturity, tuber initiation, and development in a photoperiod-regulated pathway (Kloosterman et al. 2013). The results confirmed previous QTLs identified for yield, specific gravity, and maturity in diploid and tetraploid populations. This research demonstrated that the Infinium 8303 Potato Array is efficient for constructing high-resolution genetic maps and facilitating the identification of genomic regions closely associated with agronomic traits of interest.

We have constructed many new diploid and tetraploid genetic maps and identified numerous major QTLs linked to important traits that are now candidates for marker-assisted breeding (Table 1). Almost 5000 of the SNPs can be scored for dosage

**Table 1** Populations SNP genotyped by SolCAP

Population	Ploidy	Individuals	Traits evaluated
SolCAP diversity panel	4x	213	Reducing sugars, chip color, yield, tuber shape, etc.
Atlantic × Superior	4x	185	Tuber calcium, reducing sugars, internal defects, specific gravity
B1829-5 × Atlantic	4x	156	Chip color, internal heat necrosis, specific gravity, maturity
Tundra × Kalkaska	4x	206	Scab resistance, chip color, reducing sugars, specific gravity, asparagine, acrylamide
Jacqueline Lee × MSG227-2	4x	158	Specific gravity, late blight resistance, vine maturity
Premier Russet × Rio Grande	4x	162	Reducing sugars, FF, chip color, yield, tuber shape, etc.
CIP core population	2x, 4x	48	CIP breeding lines
DM 1-3 × 84SD22	2x	133	SNP mapping, pigmentation
DM 1-3 × RH	2x	95	SNP mapping, yield, specific gravity, maturity, tuber number, average tuber size
W4 × 524-8	2x	98	Specific gravity, chip color, disease resistance
DM 1-3 × M6 F2	2x	112	Self compatibility
84SD22 × BER83	2x	129	Late blight resistance, tuber dormancy, apical dominance

(AAAA, AAAB, AABB, ABBB, BBBB). This is providing tremendous opportunities for QTL mapping at the tetraploid level. For example, we have genotyped and phenotyped three tetraploid mapping populations to study economically important traits. Using TetraploidMap v2 (BIOSS), adapted from (Hackett et al. 2013), Massa et al. (personal communication) used 3867 SNPs to create genetic maps in the tetraploid mapping population Jacqueline Lee × MSG227-2 to study late blight resistance (Table 2). TetraploidMap v2 expands the number of genetic markers that

**Table 2** Distribution of SNPs in the tetraploid mapping population Jacqueline Lee × MSG227-2

		Jacqueline Lee				
		AAAA	AAAB	AABB	ABBB	BBBB
MSG227-2	AAAA		296	121	35	
	AAAB	335	206	150	49	65
	AABB	96	141	297	223	146
	ABBB	63	57	279	291	416
	BBBB		49	153	399	
				Total	J Lee	G227-2
	Nulliplex X Simplex			1446	751	695
	Nulliplex X Duplex			516	242	274
	Nulliplex X Triplex			212	128	84
	Other configurations			1693	818	875
	Total			3867	1939	1928

can be mapped per chromosome (250 vs 50 markers) and allows for SNP markers segregating with more complex segregations than simplex  $\times$  nulliplex (1:1) or duplex  $\times$  nulliplex (5:1) ratios to be included in the mapping. The SolCAP SNP set in conjunction with advanced analysis tools allowed for the identification of a major QTL for late blight resistance that was mapped to chromosome IX. Two other mapping populations (Premier Russet  $\times$  Rio Grande and Tundra  $\times$  Kalkaska) are being used to map scab resistance, processing quality traits, specific gravity, and tuber shape. Mapping traits and better understanding the cultivated tetraploid genome provides opportunities for breeders to more efficiently develop new varieties. We envision using these parents of the mapping populations to transfer QTLs in future breeding populations.

Furthermore, we are using SNP markers to examine the occurrence and frequency of double reduction along the chromosome arms (a genetic test for autopolyploidy), assess relationships among cultivated germplasm, and fingerprint varieties. The codominant SNP markers were also able to separate conventionally bred potato varieties by market class and were used to evaluate population structure (Hirsch et al. 2013). Using SNP markers that were physically mapped allowed us to compare regions of distorted segregation and marker gaps between populations at the diploid level as well as recombination rates throughout the genome (Felcher et al. 2012).

## Conclusion

The primary research objective of SolCAP was to provide the infrastructure to link allelic variation in genes to valuable traits in cultivated germplasm of potato and tomato. Focusing on elite cultivated potato germplasm increases the probability that potato breeding will benefit from SNP genotype-based selection. The Infinium 8303 Potato Array provides a common set of markers that breeders and geneticists can reliably use and reference for mapping, germplasm assessment, and fingerprinting. This array has been a useful tool to advance our understanding of the potato genome. Furthermore, breeders are mapping QTLs across numerous populations that will expand our understanding of economically important traits and lead to marker-assisted selection and breeding. The extension and education components of SolCAP have integrated training in genomics and plant breeding so that existing breeders can make better use of sequence and genome-wide SNP data in the context of crop improvement.

SolCAP will allow the international Solanaceae community to pursue marker-based breeding by delivering integrated marker tools and breeder-friendly databases. These improved resources and tools have been tied to directed education and extension programs for current and future breeders. SolCAP has provided the tools and training for breeders to link traits with markers, so that potato varieties can be developed more rapidly, precisely, and efficiently.

**Acknowledgments** The SolCAP project was supported by funds from the US Department of Agriculture, National Institute of Food and Agriculture, Agriculture and Food Research Initiative Plant Breeding, Genetics, and Genome grant 2009-85606-05673 to D.S.D.

## References

- Bae J, Halterman D, Jansky S (2008) Development of a molecular marker associated with *Verticillium* wilt resistance in diploid interspecific potato hybrids. *Mol Breed* 22:61–69. doi:[10.1007/S11032-008-9156-8](https://doi.org/10.1007/S11032-008-9156-8)
- Colton LM, Groza HI, Wielgus SM, Jiang JM (2006) Marker-assisted selection for the broad-spectrum potato late blight resistance conferred by gene RB derived from a wild potato species. *Crop Sci* 46:589–594. doi:[10.2135/Cropsci2005.0112](https://doi.org/10.2135/Cropsci2005.0112)
- Felcher KJ et al (2012) Integration of two diploid potato linkage maps with the potato genome sequence. *PLoS One* 7:e36347. doi:[10.1371/journal.pone.0036347](https://doi.org/10.1371/journal.pone.0036347)
- Hackett CA, McLean K, Bryan GJ (2013) Linkage analysis and QTL Mapping using SNP dosage data in a tetraploid potato mapping population. *PLoS One* 8:e63939. doi:[10.1371/journal.pone.0063939](https://doi.org/10.1371/journal.pone.0063939)
- Hamilton JP et al (2011) Single nucleotide polymorphism discovery in elite north american potato germplasm. *BMC Genomics* 12:302. doi:[10.1186/1471-2164-12-302](https://doi.org/10.1186/1471-2164-12-302)
- Hardigan M, Bamberg J, Buell CR, Douches D (2014) Taxonomy and genetic differentiation among wild and cultivated germplasm of *Solanum* sect. *Petota*. *Plant Genome*. Accepted for publication
- Hirsch CN et al (2013) Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. *G3* 3:1003–1013. doi:[10.1534/G3.113.005595](https://doi.org/10.1534/G3.113.005595)
- Kasai K, Morikawa Y, Sorri VA, Valkonen JP, Gebhardt C, Watanabe KN (2000) Development of SCAR markers to the PVY resistance gene *Ryadg* based on a common feature of plant disease resistance genes. *Genome* 43:1–8
- Kloosterman B et al (2013) Naturally occurring allele diversity allows potato cultivation in northern latitudes. *Nature* 495:246–250. doi:[10.1038/Nature11912](https://doi.org/10.1038/Nature11912)
- USDA (2013) Potatoes 2012 summary. United States Department of Agriculture-USDA, National Agricultural Statistics Service-NASS
- Van Deynze A, Douches D, De Jong W, Francis D (2007) Summary of Solanaceae coordinating meetings. *Acta Hort* 745:533–536
- Van Os H et al (2006) Construction of a 10,000-marker ultradense genetic recombination map of potato: providing a framework for accelerated gene isolation and a genomewide physical map. *Genetics* 173:1075–1087. doi:[10.1534/genetics.106.055871](https://doi.org/10.1534/genetics.106.055871)