

Translating Genome Sequence Resources into Applied Outcomes in the Solanaceae: The Vision of SolCap.

David Francis^{1*}, David Douches², Walter De Jong³, Shanna Moore³, Allen Van Deynze⁴, C. Robin Buell², Esther van der Knaap¹, Matthew Robbins¹, and Sung-Chur Sim¹.

¹ The Ohio State University/OARDC, Wooster, OH 44691

² Michigan State University, East Lansing, MI 48824

³ Cornell University, Ithaca, NY 14853

⁴ University of California, Davis, CA, 95616

Data from expressed sequence tags and from whole-genome sequencing has opened opportunities for plant breeding application in the Solanaceae. These sequence resources are also driving agricultural research toward a perspective that emphasizes taxonomic groups and DNA sequence homology rather than traditional commodity boundaries. The long-term objective of the Solanaceae Coordinated Agricultural Project (SolCAP) is to provide infrastructure to link allelic variation in genes for valuable traits in the three most important vegetable crops in the Solanaceae: potato, tomato and pepper. Leveraging resources between commodities has resulted in tools to manipulate phytonutrient biosynthesis pathways affecting potato tuber and tomato fruit nutritional quality. Strategies to develop Single Nucleotide Polymorphisms (SNPs) as molecular markers for whole-genome analysis include analysis of polymorphisms in EST databases, hybridization to oligonucleotide arrays to detect single feature polymorphisms (SFPs), and *de novo* sequencing using ultra-high throughput techniques. In our work with tomato, application of SNPs has emphasized genotyping of a core collection (n = 99) of *S. lycopersicum* varieties representing heirloom (19), fresh market (23), and processing germplasm (39). Greenhouse varieties (3), land races (5) and wild species accessions from *S. lycopersicum* var. *cerasiformae*, *S. pimpinellifolium*, *S. pennellii*, and *S. habrochaites* are also included in the core collection. A variety of indexes and clustering procedures were used to analyze population structure within the cultivated germplasm. These analyses reveal subpopulations consistent with market class differentiation ($P = 0.0001$) and adaptation to arid or humid growing environments ($P = 0.003$). By extending sequence resources horizontally to encompass both accessions of wild relatives and populations of elite varieties, a greater understanding of how variation in DNA sequence affects variation in phenotype is expected.