

Soybean Genomics Research Program Accomplishments Report (2012-2016)

Overview

The soybean genomics research community has an impressive history of strategic planning and implementation to achieve research and resource milestones. This philosophy was nicely summarized in the Executive Summary from the 2008-2012 Accomplishment Report:

“The soybean research community has engaged a transparent process for developing and implementing a strategic framework for a national program to unlock the secrets of the soybean genome. The success of this process is evidenced by achieving the soybean genome sequence in record time, compared to similar efforts in other major crops, and many useful tools to expedite elite variety development.” – R.F. Wilson and D. Grant (August, 2010)

This strategic process includes periodic meetings and drafting of documents to update accomplishments and devise new strategies to meet community-wide goals. Previous strategic planning efforts were conducted in October, 1999; July, 2001; May, 2003; July, 2005; May, 2007; and July, 2010. The next meeting will be held in July, 2015.

Two key documents are produced with each strategic planning cycle: the “Soybean Genomics Research Program Strategic Plan” and the “Soybean Genomics Research Program Accomplishments Report”. The “Strategic Plan” identifies the emerging challenges and opportunities for the research community; it lays the groundwork for the next 5 years by setting community-wide research and tool development goals. The “Accomplishments Report” summarizes the accomplishments achieved during the past period with respect to each performance measurement area. This informs the community of recent accomplishments and assists the development of a new Strategic Plan that will benefit soybean researchers and stakeholders. Below you will find the current draft of the 2012-2016 Accomplishments Report.

Goal 1: Genome Sequence: Improve the quality and utility of the soybean genome sequence.

Performance Measure 1.1: Ensure the accuracy of reference sequence assembly.

- Re-releases of soybean genome assembly (versions 1.1 and *Glycine max* Wm82.a2.v1) and annotations, available at phytozome.org and at soybase.org.
- Gene atlases to improve gene prediction.
 - <http://www.soybase.org/soyseq>
 - http://soykb.org/data_sources.php (Resources were migrated to SoyBase in 2015).
- Ongoing work to resolve the heterozygous regions of the Williams 82 reference genome.
 - Haun WJ, Hyten DL, Xu WW, et al. (2011) The composition and origins of genomic variation among individuals of the soybean reference cultivar Williams 82. *Plant Physiol* 155: 645-655.

Performance Measure 1.2: Capturing and leveraging existing genetic diversity in soybean germplasm.

- New 50K SNP-chip, applied to USDA germplasm collection (data at soybase.org).
 - Song Q, Hyten DL, Jia G, et al. (2013) Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PLoS ONE* 8: e54985.
- The Soybean NAM Project. Forty NAM populations have been developed with approximately 140 lines/population. Populations were developed by crossing ISU cultivar IA3023 with 40 lines selected by the soybean breeding community for high yield, genetic diversity or as plant introductions. These populations have been phenotyped for a variety of traits including yield, seed composition and other agronomic traits. NAM parents have been sequenced and have been analyzed with the 50K SNP-chip. Comparative genome hybridization has been used to identify copy number variants (CNV) in the NAM population.
 - Anderson JE, Kantar MB, Kono TY, et al. (2014) A roadmap for functional structural variants in the soybean genome. *G3 Genes Genom Genet* 4: 1307-1318.
- The Soybean Milestones Project. DNA sequences from eighty soybean lines reflecting 90 years of soybean have been assembled relative to the Williams 82 reference genome. SNP, CNV, genotype and phenotype data can be viewed at SoyBase.org. The project has also produced small RNA, transcriptome and methylome data. In addition a depository of cloned transcription factors has been generated.
 - <http://shiny.soybase.org/SNP/>
 - <http://shiny.soybase.org/CNV/>
- Other major resequencing efforts in *G. max* and *G. soja* have identified genes involved in soybean domestication and improvement.
 - Zhou Z, Jiang Y, Wang Z, et al. (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat Biotechnol*, in press.

- Chung, WH, Jeong N, Kim J, et al. (2014) Population structure and domestication revealed by high-depth resequencing of Korean cultivated and wild soybean genomes. *DNA Res* 21: 153-167.
- Li YH, Zhao SC, Ma JX, et al. (2013) Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. *BMC Genomics* 14: 579.
- Joshi T, Valliyodan B, Wu JH, Lee SH, Xu D, Nguyen H (2013) Genomic differences between cultivated soybean, *G. max* and its wild relative *G. soja*. *BMC Genomics* 14: S5.
- Lam HM, Xu X, Liu X, et al. (2010) Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat Genet* 42: 1053-1059.
- BAC libraries and BAC end sequences for perennial and annual *Glycines*, aligned to reference genome.
 - <http://www.soybase.org/soymap2>
 - Ha J, Abernathy B, Nelson W, et al. (2012) Integration of the Draft Sequence and Physical Map as a Framework for Genomic Research in Soybean (*Glycine max* (L.) Merr.) and Wild Soybean (*Glycine soja* Sieb. and Zucc.) *G3 Genes Genom Genet* 2: 321-32.
- Preliminary Pan-genome for wild soybean.
 - Li YH, Zhou G, Ma J, et al. (2014) De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nat Biotechnol* 32: 1045–1052

Performance Measure 1.3: Improving bioinformatic resources for genomic analysis and practical applications.

- Soybean gene ontology enrichment tool. This tool provides gene ontology (GO) information (inferred from Arabidopsis) for all genes in the soybean genome. Statistical analyses can be performed to identify statistically overrepresented GO terms within a supplied gene list relative to all genes in the genome.
 - http://www.soybase.org/goslimgraphic_v2/dashboard.php
- Soybean gene annotation tool. Given a supplied gene list, this tool will provide additional annotation information including BLAST homology information to the Uniprot protein database and predicted Arabidopsis proteins. Gene ontology information is also provided.
 - <http://www.soybase.org/genomeannotation>
- Fast neutron mutant data for phenotypes and genotypes are publically available at SoyBase for a subset of the populations.
 - <http://www.soybase.org/mutants>
 - Bolon YT, Stec AO, Michno JM, et al. (2014) Genome resilience and prevalence of segmental duplications following fast neutron irradiation of soybean. *Genetics* 198: 967-981.
- Soybean Virus Induced Gene Silencing (VIGS) Database. Provides a database of soybean genes for which VIGS constructs are publicly available.

- <http://soybase.org/SoyVIGS>
- Soybean 50K chip data. Provides SNP data developed from genotyping the soybean germplasm collection with the SoySNP50K iSelect BeadChip.
 - <http://www.soybase.org/dlpages/#snp50k>
- Improved soybean trait ontology for growth and development.
 - <http://www.soybase.org/ontology.php>
 - The SoyBase soybean trait ontology was improved by adding traits for newly reported QTL, and updating the disease traits with additional information such as disease and taxonomic synonyms. These comprehensive ontologies have been used to provide metadata for most data types in SoyBase. This allows searches to return not only what was requested (i.e. a strict lexical search) but also any other related information. Thus a user can submit a general query such as "oil" and have returned all data in SoyBase related that term. For example a lexical search would be expected to find QTL for seed oil content but the trait-based metadata search would additionally return enzymes involved in fatty acid biosynthesis and other related information.
- Soybean genome categorization and visualization software to facilitate exploration of diversity and sequence variation.
 - <http://soykb.org/SNPviz/>
 - Langewisch T, Zhang H, Vincent R, et al. (2014) Major soybean maturity gene haplotypes revealed by SNPviz analysis of 72 sequenced soybean genomes. PLoS ONE 9(4): e94150.

Performance Measure 1.4: Reveal function of targeted genome sequences to facilitate gene discovery and application.

- Development and use of Virus Induced Gene Silencing (VIGS) for soybean.
 - Kandoth P, Heinz R, Yeckel G, et al. (2013) A virus-induced gene silencing method to study soybean cyst nematode parasitism in *Glycine max*. BMC Research Notes 6: 255.
 - Zhang C, Yang C, Whitham SA, et al. (2009) Development and use of an efficient DNA-based viral gene silencing vector for soybean. Mol Plant Microbe In 22: 123-131.
- VIGS has been used to identify genes and gene pathways conferring resistance to Asian soybean rust and soybean cyst nematode, and tolerance to iron deficiency chlorosis.
 - Liu S, Kandoth PK, Warren SD, et al. (2012) A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. Nature 492: 256-260.
 - Pandey AJ, Yang C, Zhang C, et al. (2010) Functional analysis of the Asian Soybean Rust resistance pathway mediated by *Rpp2*. Mol Plant Microbe In 24: 194-206.
 - Meyer JDF, Silva DCG, Yang C, et al. (2009) Identification and analyses of candidate genes for *Rpp4*-mediated resistance to Asian soybean rust in Soybean. Plant Physiol 150: 295-307.

Performance Measure 1.5: Leveraging genomic information from Phaseoloids and other species.

- Development of Legume Information System (LIS).
 - <http://www.comparative-legumes.org>
- Reference genomes for multiple legumes are now available including *Phaseolus vulgaris*, *Cajanus cajan* and *Cicer arietinum*.
 - Schmutz, J, McClean PE, Mamidi S, et al. (2014) A reference genome for common bean and genome-wide analysis of dual domestications. Nat Genet 46: 707-713.
 - Varshney RK, Song C, Saxena RK, et al. (2013) Draft genome sequence of chickpea (*Cicer arietinum*) provides resource for trait improvement. Nat Biotechnol 31: 240-246.
 - Varshney RK, Chen W, Li Y, et al. (2012) Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume of resource poor farmers. Nat Biotechnol 30: 83-89.

Performance Measure 1.6: Determine the role of epigenetics in soybean improvement.

- Prediction of small RNAs and their targets.
 - Joshi T, Yan Z, Libault M, Jeong DH, Park S, et al. (2010) Prediction of novel miRNAs and associated target genes in *Glycine max*. BMC Bioinformatics 11: S14.
- Database of small RNAs funded by USB.
 - Arikrit S, Xia R, Kakrana A, et al. (2014). An atlas of soybean small RNAs identifies phased siRNAs from hundreds of coding genes. Plant Cell: in press.
 - <http://mpss.udel.edu/#legume>
- Analyses of small RNAs from different tissues.
 - Zabala G, Campos E, Varala KK, et al. (2012) Divergent patterns of endogenous small RNA populations from seed and vegetative tissues of *Glycine max*. BMC Plant Biol 12: 177.
- Role of small RNAs in regulating flavonoid and defense pathways.
 - Wong J, Gao L, Yang Y, et al. (2014) Roles of small RNAs in soybean defense against *Phytophthora sojae* infection. Plant J 79: 928-940.
 - Cho YB, Jones SI, Vodkin L (2013) The transition from primary siRNAs to amplified secondary siRNAs that regulate chalcone synthase during development of *Glycine max* seed coats. PLoS ONE 8: e76954.
 - Zhai J, Jeong DH, De Paoli E, et al. (2011) MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. Gene Dev 25: 2540-2553.
- Degradome analysis of small RNA targets in developing seed.
 - Shamimuzzaman M, Vodkin L (2012) Identification of soybean seed developmental stage-specific and tissue-specific miRNA targets by degradome sequencing. BMC Genomics 13: 310.
- Analyses of DNA methylation patterns.
 - Zabala G, Vodkin LO (2014) Methylation affects transposition and splicing of a large CACTA transposon from a MYB transcription factor regulating anthocyanin synthase genes in soybean seed coats. PLoS ONE 9: e111959.

- Cook DE, Bayless AM, Wang K, et al. (2014) Distinct copy number, coding sequence, and locus methylation patterns underlie *Rhg1*-mediated soybean resistance to soybean cyst nematode. *Plant Physiol* 165: 630-647.
- Song QX, Lu X, Li QT, et al. (2013) Genome-wide analysis of DNA methylation in soybean. *Mol Plant* 6: 1961-1974.
- Schmitz RJ, He Y, Valdés-López O., et al. (2013). Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. *Genome Res* 23: 1663-1674.

Goal 2: Gene Function: Develop functional genomic technologies to optimize utility of genome sequence information in germplasm enhancement.

Performance Measure 2.1: Develop comprehensive gene expression data for soybean.

- Gene atlases to improve gene prediction.
 - Severin AJ, Woody JL, Bolon Y-T, et al. (2010) RNA-seq atlas of *Glycine max*: A guide to the soybean transcriptome. *BMC Plant Biol* 10: 160.
 - Libault M, Farmer A, Brechenmacher L, et al. (2010) Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection. *Plant Physiol* 152: 541-552.
 - Libault M, Farmer A, Joshi T, et al. (2010) An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. *Plant J* 63: 86-99.
- Analysis of RNA-Seq data from developing seeds.
 - Kour A, Boone AM, Vodkin LO (2014) RNA-Seq profiling of a defective seed coat mutation in *Glycine max* reveals differential expression of proline-rich and other cell wall protein transcripts. *PLoS ONE* 9 (5) e96342.
 - Jones SI, Vodkin LO (2013) Using RNA-Seq to profile soybean seed development from fertilization to maturity. *PLoS One* 8L: e59270.
- Analysis of RNA-Seq in seven stages of seedling cotyledons.
 - Shamimuzzaman M, Vodkin L (2014) Transcription factors and glyoxylate cycle genes prominent in the transition of soybean cotyledons to the first functional leaves of the seedling. *Funct Integr Genomic* 14: 683-696.
- Analysis of RNA-Seq in response to pathogen infection.
 - Lin F, Zhao M, Baumann DD, et al. (2014) Molecular response to the pathogen *Phytophthora sojae* among ten soybean near isogenic lines revealed by comparative transcriptomics. *BMC Genomics* 15: 18.
 - Tremblay A, Hosseini P, Li S, et al. (2013) Analysis of *Phakopsora pachyrhizi* transcript abundance in critical pathways at four time-points during infection of a susceptible soybean cultivar using deep sequencing. *BMC Genomics* 14: 614.
 - Kim KH, Kang YJ, Kim DH, et al. 2011. RNA-Seq analysis of a soybean near-isogenic line carrying bacterial leaf pustule-resistant and-susceptible alleles. *DNA Res* 18: 483-497.

Performance Measure 2.2: Develop near isogenic lines (NIL) to help reveal genetic mechanisms that mediate useful traits.

- There has been no organized (funded) effort to develop new NILs. Some NILs have been developed and analyzed on a lab by lab basis.
 - Lin F, Zhao M, Baumann DD, et al. (2014) Molecular response to the pathogen *Phytophthora sojae* among ten soybean near isogenic lines revealed by comparative transcriptomics. *BMC Genomics* 15:18.
 - Diers BW, Kim KS, Frederick RD, et al. (2013) Registration of eight soybean germplasm lines resistant to soybean rust. *J Plant Regist* DOI: 10.3198/jpr2012.11.0052crg.
 - Kim KS, Diers BW (2013) The associated effects of the soybean aphid resistance gene *Rag2* from PI 200538 on agronomic traits in soybean. *Crop Sci* 53: 1326-1334.
 - Kandoth PK, Ithal N, Recknor J, et al. (2011) The Soybean *Rhg1* locus for resistance to the soybean cyst nematode *Heterodera glycines* regulates the expression of a large number of stress- and defense-related genes in degenerating feeding cells. *Plant Physiol* 155: 1960-1975.

Performance Measure 2.3: Develop an improved infrastructure to facilitate genome annotation.

- New genome annotation released via DOE/JGI/HudsonAlpha. Soybean is a priority genome for DOE with continued investment.
- The 2nd genome assembly was released in 2014. It and the annotations are now in SoyBase as the default. The SoySNP50K SNPs have been mapped to the Wm82.a2 assembly. The Severin and Libault RNA-seq expression atlases are being mapped to Wm82.a2. NCBI RefSeq genes, SNP markers for several large genetic mapping projects, SoyMap2, and several expression data sets have been mapped to Wm82.a2. Other annotation tracks will be migrated to the Wm82.a2 assembly as needed.

Performance Measure 2.4: Achieve high-definition genomic characterization of biological mechanisms and regulatory systems in soybean.

- A database of soybean transcription factors.
 - Wang Z, Libault M, Joshi T, et al. (2010) SoyDB: a knowledge database of soybean transcription factors. *BMC Plant Biol* 10: 14.
 - Soybean transcription factor data will be migrated to SoyBase in 2015.

Performance Measure 2.5: Use functional genomic methods to characterize transcription regulated pathways.

- VIGS coupled with microarray or RNA-seq to identify the gene networks regulating defense and abiotic stress responses.
 - Morales A, O'Rourke JA, van de Mortel M, et al. (2013) Transcriptome analyses and virus induced gene silencing identify genes in the *Rpp4*- mediated Asian Soybean Rust resistance pathway. *Funct Plant Biol* 40: 1029-1047.
- Characterization of transcription factors involved in seedling development.

- Jones SI, Tan Y, Shamimuzzaman Md, et al. (2014) Direct detection of transcription factors in cotyledons during seedling development using sensitive silicon-substrate photonic crystal protein arrays. *Plant Physiol* 167, in press.
- Shamimuzzaman M, Vodkin L (2013) Genome-wide identification of binding sites for NAC and YABBY transcription factors and co-regulated genes during soybean seedling development by CHIP-Seq and RNA-Seq. *BMC Genomics* 14: 477.

Performance Measure 2.6: Advance gene modification technologies to help associate candidate genes with a discrete phenotype.

- Site directed mutagenesis using genome engineering methods.
 - Curtin SJ, Zhang F, Sander JD, et al. (2011) Targeted mutagenesis of duplicated genes in soybean with zinc finger nucleases. *Plant Physiol* 156: 466-473.
- Tilling used to identify novel traits.
 - Hoshino T, Watanabe S, Takagi Y, et al. (2014) A novel *GmFAD3-2a* mutant allele developed through TILLING reduces α -linolenic acid content in soybean seed oil. *Breeding Sci* 64: 371.

Performance Measure 2.7: Create a saturated transposon insertion population with defined flanking sequences that can be used to identify mutants by BLAST sequence comparison.

- Efforts towards addressing performance measure 2.7 are ongoing, albeit with current resource allocations to this effort the creation of a “saturated transposon population” will be a challenge. Nonetheless, the community is currently exploiting multiple transposon tagging systems as a functional genomics resources for the crop. Moreover, fast neutron mutagenized soybean populations have been generated with ongoing genotypic characterizations at both the University of Missouri and the University of Minnesota.
 - Gillman JD, Stacey MG, Cui Y, et al. (2014) Deletions of the *SACPD-C* locus elevate seed stearic acid levels but also results in fatty acid and morphological alterations in nitrogen fixing nodules. *BMC Plant Biol* 14:143.
 - Raval J, Baumbach J, Ollhoff AR, et al. (2013) A candidate male-fertility female-fertility gene tagged by the soybean endogenous transposon, Tgm9. *Funct Integr Genomics* 13: 67-73.
 - Cui Y, Barampuram S, Stacey MG, et al. (2013) Tnt1 retrotransposon mutagenesis: a tool for soybean functional genomics. *Plant Physiol* 161: 36-47.
 - Hancock CN, Zhang F, Floyd K, et al. (2011) The rice miniature inverted repeat transposable element mPing is an effective insertional mutagen in soybean. *Plant Physiol* 157: 552-562.
 - Bolon YT, Haun WJ, Xu WW, et al. (2011) Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. *Plant Physiol* 156: 240-253.
 - Mathieu M, Winters EK, Kong F, et al. (2009) Establishment of a soybean (*Glycine max* Merr. L) transposon-based mutagenesis repository. *Planta* 229: 279-289.
 - The Ac/DS, mPing, TGM and T-DNA insertion mutants will be added to SoyBase.

Performance Measure 2.8: Implement outreach opportunities for education and use of genomic databases.

- Tutorials for using SoyBase are available.
 - <http://www.soybase.org/tutorials>
 - SoyBase and SoyKB training available annually at the Soybean Breeders Workshop.

Performance Measure 2.9: Develop an ORFeome library from agronomically important genes and gene families.

- An ORFeome of transcription factors has been developed and is available through SoyKB.
 - Chai C, Wang Y, Joshi T, et al. (2015) Soybean transcription factor ORFeome: valuable resource to accelerate research on abiotic stress resistance. BMC Genomics: submitted.
 - http://soykb.org/TF_ORFeome

Goal 3: Transformation/Transgenics: Optimize and expand transgenic methods and improve understanding of natural genes for modification of trait expression.

Performance Measure: 3.1: Establish a soybean genetic repository and distribution center.

- USDA/ARS allocated \$100K/year to establish and maintain a repository to house the collection of transposon and FN characterized seed lots at the University of Illinois. However, this performance measure is still in the discussion phase.

Performance Measure: 3.2: Develop next-generation transformation and targeting technologies and utilize these transgenic approaches to help elucidate gene function and deploy genes of interest.

- Given the rapid pace of technology development in the area of genome editing tools, performance measure 3.2 is well positioned to complement the ongoing transposon-based/FN platforms for functional genomics programs that focus on soybean. Next-generation technology includes the implementation of transcription activator like effector nucleases (TALENs) and the clustered regulatory interspaced short palindromic repeat Cas-based RNA-guided DNA endonuclease (CRISPR/Cas9).
 - Haun W, Coffman A, Clasen BM, et al. (2014) Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. Plant Biotechnol J 12: 934-940.
 - Gaj T, Gersbach CA, Barbas III CF (2013) ZFN, TALEN, and CRISPR/Cas based methods for genome engineering. Trends Biotechnol 31:397-405.
 - Voytas DF (2013) Plant genome engineering with sequence-specific nucleases. Annu Rev Plant Biol 64:327-350.
 - Shan Q, Wang Y, Li J, et al. (2013) Targeted genome modification of crop plants using a CRISPR-CAS system. Nat Biotechnol 31:686-688.

- Belhaj K, Chaparro-Garcia A, Kamoun S, et al. (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* 9:39:1-10.
- Bogdanove AJ, Voytas DF (2011) TAL Effectors: Customizable proteins for DNA targeting. *Science* 333:1843-1846.
- Improvements in soybean transformation methods have been developed.
 - Benzle KA, Finer KR, Marty DM, et al. (2015) Isolation and characterization of novel *Agrobacterium* strains for soybean and sunflower transformation. *Plant Cell Tiss Organ Cult* 121:71-81.
 - Yamada T, Takagi K, Ishimoto M (2012). Recent advances in soybean transformation and their application to molecular breeding and genomic analysis. *Breeding Sci* 61: 480.
- Targeted mutagenesis/editing platforms such as CRISPR are now available for plant species and are being applied to soybean.
 - Curtin SJ (2014). A Genome Engineering Toolbox for Legume Functional Genomics. In *Plant and Animal Genome XXII Conference*. Abstract.
 - Jacobs TB, LaFayette PR, Ortega MA, et al. (2014) A Comparison of ta-siRNA and CRISPR/Cas9 Mediated Gene Silencing Systems in Soybean. In *Plant and Animal Genome XXII Conference*. Abstract.

Goal 4: Translational Genomics: Optimize breeding efficiency with robust sequence-based resources.

Performance Measure: 4.1: *Develop analytical approaches to characterize soybean germplasm diversity based on the SoyHapMap 1.0 data to identify parental lines for breeding purposes.*

- An expanded array of 50,000 SNPs has been used to create haplotypes for over 18,000 accessions in the USDA soybean germplasm collection. These data have been incorporated into SoyBase and the complete collection of haplotypes or a user-selected subset can be downloaded.

Performance Measure: 4.2: *Discover gene/QTL for qualitative traits and develop tightly linked DNA markers.*

- Identification of salt tolerance genes.
 - Qi X, Li M, Xie M, et al. (2014) Identification of a novel salt tolerance gene in wild soybean by whole-genome sequencing. *Nat Commun* 5: 4340.
 - Guan R, Qu Y, Guo Y, et al. (2014) Salinity tolerance in soybean is modulated by natural variation in *GmSALT3*. *Plant J* 80: 937-950.
- Identification of shattering related genes.
 - Funatsuki H, Suzuki M, Hirose A, et al. (2014) Molecular basis of a shattering resistance boosting global dissemination of soybean. *P Natl Acad Sci USA* 111: 17797-17802.
 - Dong Y, Yang X, Liu J, et al. (2014) Pod shattering resistance associated with domestication is mediated by a NAC gene in soybean. *Nat Commun* 5:3352.

- Identification of phosphorus utilization genes.
 - Zhang D, Song H, Cheng H, et al. (2014) The acid phosphatase-encoding gene *GmACP1* contributes to soybean tolerance to low-phosphorus stress. *Plos Genetics* 10: 1371.
- Identification of genes associated with determinant/indeterminant phenotypes.
 - Ping J, Liu Y, Sun L, et al. (2014) *Dt2* is a gain-of-function MADS-domain factor gene that specifies semideterminancy in soybean. *Plant Cell* 26: 2831-2842.
- Identification of genes associated with leaf shape and seeds per pod.
 - Jeong N, Suh S, Kim M, et al. (2012) *Ln* is a key regulator of leaflet shape and number of seeds per pod in soybean. *Plant Cell* 24: 4807-4818.
- Identification of genes associated with tocopherol balance.
 - Dwiyanti M, Yamada T, Sato M, et al. (2011) Genetic variation of γ -tocopherol methyltransferase gene contributes to elevated α -tocopherol content in soybean seeds. *BMC Plant Biol* 11: 152.
- High oleic soybeans.
 - Pham A, Shannon JG, Bilyeu K (2012) Combinations of mutant *FAD2* and *FAD3* genes produce high oleic acid in soybean. *Theoret Appl Genet* 125:503-515.
- Identification of genes associated with iron deficiency in soybean.
 - Moran Lauter A, Peiffer GA, Yin T, et al. 2014. Identification of candidate genes involved in early iron deficiency chlorosis signaling in soybean (*Glycine max*) roots and leaves. *BMC Genomics*, 15: 702.
 - Atwood SE, O'Rourke JA, Peiffer GA, et al. (2013) Replication Protein A and the iron efficiency response in soybean. *Plant Cell Environ* 37: 213-234.
 - Peiffer G, King K, Severin A, et al. (2012) Identification of candidate genes underlying an iron efficiency quantitative trait locus in soybean. *Plant Physiol* 158: 1745-1754.
- Identification and fine mapping of aphid resistance genes.
 - Jun TH, Mian MR, Michel AP (2012). Genetic mapping revealed two loci for soybean aphid resistance in PI 567301B. *Theor Appl Genet* 124: 13-22.
 - Kim KS, Bellendir S, Hudson KA, et al. (2010) Fine mapping the soybean aphid resistance gene *Rag1* in soybean. *Theor Appl Genet* 120: 1063-1071.
 - Zhang G, Gu C, Wang D (2010). A novel locus for soybean aphid resistance. *Theor Appl Genet* 120: 1183-1191.
 - Kim KS, Hill CB, Hartman GL, et al. (2010) Fine mapping of the soybean aphid-resistance gene *Rag2* in soybean PI 200538. *Theor Appl Genet* 121: 599-610.
- New genes conferring resistance to *Phytophthora sojae*, including *RpsYu25*, *RpsUN1*, and *RpsUN2*, have been mapped.
 - Sun S, Wu XL, Zhao JM, et al. (2011) Characterization and mapping of *RpsYu25*, a novel resistance gene to *Phytophthora sojae*. *Plant Breeding* 130: 139-143.
 - Lin F, Zhao M, Ping J, et al. (2013) Molecular mapping of two genes conferring resistance to *Phytophthora sojae* in a soybean landrace PI 567139B. *Theor Appl Genet* 126: 2177-2185.

Performance Measure: 4.3: Discover gene/QTL for quantitative traits and develop tightly linked DNA markers.

- The soybean genetics community continues to identify loci for quantitative traits. Over 200 publications identifying QTL in soybean have been published since 2010.
- Identification of genes associated with maturity.
 - Zhang D, Cheng H, Hu Z, et al. (2013) Fine mapping of a major flowering time QTL on soybean chromosome 6 combining linkage and association analysis. *Euphytica* 191: 23-33.
 - Xia Z, Watanabe S, Yamada T, et al. (2012) Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *PNAS* 109: E2155-E2164.
 - Watanabe S, Xia Z, Hideshima R, et al. (2011) A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. *Genetics* 188:395-407.
- Identification of genes associated with seed traits.
 - Kato S, Sayama T, Fujii K, et al. (2014) A major and stable QTL associated with seed weight in soybean across multiple environments and genetic backgrounds. *Theor Appl Genet* 127: 1365-1374.
 - Xie FT, Niu Y, Zhang J, et al. (2014). Fine mapping of quantitative trait loci for seed size traits in soybean. *Mol Breeding* 34: 2165-2178.
 - Gillman JD, Tetlow A, Lee J-D, et al. (2011) Loss-of-functions affecting a specific Glycine max R2R3 MYB transcription factor result in brown hilum and brown seed coats. *BMC Plant Biol* 11: 155.
- Identification of genes associated with triterpenoid saponins.
 - Sayama T, Ono E, Takagi K, et al. (2012) The *Sg-1* glycosyltransferase locus regulates structural diversity of triterpenoid saponins of soybean. *Plant Cell* 24: 2123-2138.
- Increase available phosphorus in soybean seeds with reduced-phytate soybeans
 - Vincent JA, Stacey M, Stacey G, et al. (2015) Phytic acid and inorganic phosphate composition in soybean lines with independent IPK1 mutations. *Plant Gen*: in press.
 - Gillman JD, Baxter I, Bilyeu K (2013) Phosphorus partitioning of soybean lines containing different mutant alleles of two soybean seed-specific adenosine triphosphate-binding cassette phytic acid transporter paralogs. *Plant Gen* 6: 1.
- Identification genes associated with aluminum tolerance.
 - Abdel-Haleem H, Carter TE, Rufty TW, et al. (2014) Quantitative trait loci controlling aluminum tolerance in soybean: candidate gene and single nucleotide polymorphism marker discovery. *Mol Breeding* 33: 851-862.
- Identification of genes associated with nematode resistance.
 - Xu X, Zeng L, Tao Y, et al. (2013) Pinpointing genes underlying the quantitative trait loci for root-knot nematode resistance in palaeopolyploid soybean by whole genome resequencing. *PNAS* 110(33):13469-13474.
 - Cook DE, Lee TG, Guo, X, et al. (2012) Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science* 338: 1206-1209.

Performance Measure: 4.4: *Develop and populate a user-friendly database of validated QTL for use in marker assisted breeding applications.*

- This exact database has not been developed as there is so far no agreed on definition of "validated". The SoyBase QTL section contains all of the information needed but it is currently up to users to decide what is meant by validated.

Performance Measure: 4.5: *Define the molecular genetic signatures of selection in 70+ years of U.S. soybean breeding by use of the 50,000 SNP Illumina Infinium Assay.*

- This performance measure is currently underway using whole genome resequencing data (see Performance Measure 1.2).

Performance Measure: 4.6: *Define optimum breeding models for different breeding situations using in silico analysis.*

- The use of genomic selection has been explored for quantitative traits including yield, plant height, days to maturity, and resistance to SCN.
 - Bao Y, Vuong T, Meinhardt C, et al. (2014) Potential of association mapping and genomic selection to explore PI 88788 derived soybean cyst nematode resistance. *The Plant Genome* 7(3).
 - Jarquín D, Kocak K, Posadas L, et al. (2014) Genotyping by sequencing for genomic prediction in a soybean breeding population. *BMC Genomics* 15: 740.