

Dr. Lightfoot is Phd in Genetics and Professor for Biotechnology and Genomics at SIUC. He has been a member of the Department of Plant, Soil and Agricultural Systems at Southern Illinois University since 1991. He is cross-appointed in the Department of Plant Biology and the Department of Biochemistry and Molecular Biology. He co-authored the RAMP request for the Center for Excellence in Soybean Research, Teaching and Outreach. He is a Chair of the Research Committee in the Center for Excellence in Soybean Research, Teaching and Outreach and Head of the Genomic Science Facility at SIUC. He works on novel gene discovery by genomics, particularly methods to decode and use the DNA sequences of chromosomes. He targets commercially valuable disease resistances, phyto-pharmaceuticals and crop yield boosting elements. His ten research projects are, or have been, funded by the NSF, USDA, USB, ISPOB, CFAR and industrial sponsors. He is involved in collaborations with twelve SIUC faculty in eight additional projects. He collaborates with 85 scientist nationally and 157 scientist worldwide. He has published 75 articles, five book chapters, thousands of sequences. He has trained 5 PhD and 23 MS students.

Using The Soybean Genome Database (SoyGD) To Display and Analyze a 1 Gbp genome sequence.

Abstract:

Genomes like *Glycine max* (soybean) that have been highly conserved following increases in ploidy (by duplication or hybridization) present challenges for bioinformatics and genome analysis. At <http://soybeangenome.siu.edu> the Soybean Genome Database (SoyGD) genome browser has, since 2002, integrated and served the publicly available soybean physical map, BAC fingerprint database and genetic map associated genomic data (1). Duplicated regions have been identified and catalogued with a-d suffix to marker anchor names and contig names that communicate ploidy (ctg>8000 are tetraploid, ctg>9000 are octoploid). DNA sequence data has been used to separate DNA marker anchors from homologs of DNA marker anchors in BAC pools. About 200 gene families were mapped by EST hybridization. About 23,000 minimum tiling path (MTP) BIBAC clones provided BAC end sequences (BES) to decorate the physical map and were added to the database as separate tracks. Predicted gene models were developed for about 15% of the BES. From these models candidate genes underlying disease resistance, seed yield and seed protein, oil or isoflavone content were detected and fine-mapped. In recent additions 1 Gbp of genome sequence was made available in about 1500 scaffolds by DOE. Methods for display were improved by cross-referencing the BES and WGS with Arabidopsis (3). In genome evolution analyses more than a thousand additional microsatellite marker anchors were developed for contigs, 353 on the map and about 700 still in Queue (awaiting placement). Linkage analysis placed one hundred of the 1,053 new microsatellite markers on the genetic map with contigs and associated features. About half of the markers mapped to regions of the genome that formed gaps in earlier maps suggesting marker clustering biases. SoyGD represents the new build 5 for the physical map with 800 contigs from the 76,749 fingerprinted clones publicly available. New QTL data has been incorporated from the newly release 'Essex' by 'Forrest' and 'Flyer' by 'Hartwig' RIL populations. Gene expression data has been added to the gene models represented in SoyGD. This work was supported by NSF project #9878635 and USB 2218-6218.

(1) Shultz et al., 2006. Nucleic Acids Res. Database Issue D758-D765.

(2) Shopinski K, et al 2006. Plant Methods 2:20-28

(3) Shultz JL, Jeffery D. Ray and David A. Lightfoot. 2007. BMC Genomics 8:8

(4) Shultz J, Kazi S., Bashir, R., Afzal J, Lightfoot D.A. 2007. Theor Appl Genet 114:1081–1090