

CURRICULUM VITA

Name: Seung Yon (Sue) Rhee
Website: <https://dpb.carnegiescience.edu/labs/rhee-lab/>
Address: Carnegie Institution, Department of Plant Biology, 260 Panama Street,
Stanford, CA 94305 USA.
E-mail: srhee@carnegiescience.edu

Education:

Ph.D. (biology), Stanford University, Molecular genetic analysis of cell separation during *Arabidopsis thaliana* pollen development, 1998
B.A. (biology), Swarthmore College, 1992

Employment:

Senior Staff Scientist, Department of Plant Biology, Carnegie Institution for Science, 2018-Present
Acting Director, Department of Plant Biology, Carnegie Institution for Science, 2016-2018
Staff Scientist, Department of Plant Biology, Carnegie Institution for Science, 2005-2016
Staff Associate, Department of Plant Biology, Carnegie Institution for Science, 1999-2005
Founding Director, the *Arabidopsis* Information Resource (TAIR), Department of Plant Biology, Carnegie Institution for Science, 1999-2005
Database Curator, *Arabidopsis thaliana* Database (AtDB), Dept of Genetics, Stanford University, 1998-1999
Independent Science Education Film Producer, Celadon Films, 1998-1999
Research Associate, Department of Plant Biology, Carnegie Institution for Science, 1994-1998
Research Associate, Department of Biochemistry, Stanford University, 1993-1994
Research Assistant, Biology Department, Swarthmore College, 1990-1992
Research Intern, Johns Hopkins University School of Medicine, 1990 (summer)

Professional Services:

Scientific Advisory Boards:

Gene Ontology Consortium (2019-present); Phylos, Inc. (2018-present); ASPB Award Nominations Committee (2018-present); IMPB conference (2018); VIB Department of Plant Systems Biology, Belgium (2016-present); US Dept of Energy's Joint Genome Institute's Plant Group (2015-present); Protein Data Bank (2009-present); Program for International Consortia and Collaboration on AgriBioinformatics in National Agricultural Genome Program (PICCAN) in Korea (2016-2017); NSF C3-C4 Photosynthesis Project (2012-2013); Member, Nominating Committee for the International Society of Biocuration's Executive Committee (2009-2010); Member, Nominating Committee for Plant Cyberinfrastructure Board of Directors (2007); Value-directed Evolutionary Genomics Initiative (VEGI) (2010-2014); CropLink Global Database (2006-2009); Steering Committee Member, International Solanaceae Genome Initiative (2004-2008); Saccharomyces Genome Database (SGD) (2003-2006); GrainGenes (2003-2006); Cornell Genomics (2002-2006); ChromDB (2001-2004)

Grant Review Boards:

DOE (2018); NSF (2018, 2016, 2015, 2014, 2012, 2011, 2008, 2006); USDA-ARS (2002); NHGRI (2002)

International Conference Organization Boards:

Lead organizer, First Plant Cell Atlas Workshop (2020); Co-organizer, 2nd Plant Systems Biology Conference (2020); Co-organizer, Plant Genomes, Systems Biology, and Engineering Conference at Cold Spring Harbor Laboratory (2019); Co-organizer, Plant Genomes and Biotechnology Conference at Cold Spring Harbor Laboratory (2017); Co-organizer, Forth Conference of International Society for Biocuration (2010); Lead organizer, Second International Biocurators meeting (2007); Co-organizer, Solanaceae Genomics meeting (2007); Lead organizer, First International Biocurators Conference (2005); Co-organizer, NSF sponsored workshop on 'National Plant Synthesis Center' (2005)

Scientific Journal Editorial Boards:

In silico Plants Editorial Board (2018-present); Associate Editor, Molecular Plant (2014-2019); Monitoring Editor, Plant Physiology (2002-2008, 2013-2016)

Carnegie Institution for Science Services:

Deputy organizer for Carnegie Workshop on Genomes to Ecosystems (2019); Faculty Advisor for Intrinsically Disordered Proteins Scientific Interest Group (2018-present); Faculty Advisor for DPB Website (2019-present); Carnegie's Center for Scientific Computation Committee (2014-2016); Carnegie Summer Internship Program Coordination (2013-2016); DPB IT Committee (2012-2016); Departmental Website Design (2010); Carnegie Seminar Organization (2008-2010); Departmental Website Design (2002); Internal Seminar Series Initiation and Organization (2000-2001)

Teaching:

Fundamentals and Frontiers in Plant Biology (BIO229) (2020); Career Exploration & Planning (BIO380) (2019); Carnegie Writing Workshop (2018, 2019); Networks in Biology (Freshmen Seminar) (2010, 2012)

Awards:

NSF Predoctoral Fellowship (1993-1996); NSF/DOE/USDA Plant Training Grant Fellowship (1992-1993); Sigma Xi National Society (1991-1992); Howard Hughes Undergraduate Research Fellowship (1990-1991); National Honors Society (1988)

Scientific Society Memberships:

American Society of Plant Biologists (2010-present); International Society of Biocuration (2010-present); American Chemical Society (2014-present); Society for the Study of Evolution (2014-present); Society of Molecular Biology and Evolution (2014-present); Genetics Society of America (2014-present); International Society for Computational Biology (2015-present); California Native Plant Society (2015-present); Northern California Science Writers Association (2016-present); American Society of Cell Biologists (2016-present); American Geophysical Union (2019-present)

Contribution to Science:

1. My early work in graduate school focused on understanding how plant cells separate, which is very rare because plant cells are connected by cell walls. However, cell separation occurs during male gametogenesis in many plants to create individual pollen grains. Through molecular genetics, cell biological, and biochemical approaches, I identified a class of mutants called *quartet*, which are required for cell separation, and subsequently determined the molecular nature of the defects through gene cloning and molecular and biochemical characterization. Using immunolocalization and biochemical analyses, I showed that the phenotype resulted from defects in degrading the temporary cell wall before the secondary cell wall is deposited from the maternal tissue. I then cloned one of the genes, which encoded a pectin methylesterase, the first cell wall degrading enzyme with a demonstrated function *in vivo*. The *quartet* strains are still the *de facto* lines for plant scientists to study a variety of topics including gametophytic function, meiotic drive, genome stability, and centromere mapping. The strains have been used to map Arabidopsis centromeres, which was instrumental in refining the physical map and completing the genome sequencing. In the future, these strains could enable the creation of artificial plant chromosomes.

Preuss D, Rhee SY, and Davis RW. (1994) Tetrad analysis possible in Arabidopsis with mutation of the QUARTET (QRT) genes. **Science** 264(5164):1458-60.

Rhee SY and Somerville CR. (1998) Tetrad pollen formation in quartet mutants of *Arabidopsis thaliana* is associated with persistence of pectic polysaccharides of the pollen mother cell wall. **Plant Journal** 15(1):79-88.

Rhee SY, Osborne E, Poindexter P, and Somerville, CR (2003) Microspore separation in the *quartet 3* mutants of Arabidopsis is impaired by a defect in a developmentally regulated pectinase required for pollen mother cell degradation. **Plant Physiology** 133(3):1170-80.

2. As genome sequencing became feasible towards the end of my graduate work, I became interested in the possibility of genome-enabled biology to understand the functions of all genes and pathways encoded in a genome and elucidate how organisms are hard- and soft-wired. As an early career investigator at Carnegie, I led a team of biologists and software engineers to create a computational infrastructure called the Arabidopsis Information Resource (TAIR) to collect and encode all available genomic and literature data to be computable by algorithms and easily accessible by researchers. TAIR has been a primer for revolutionizing plant research by enabling systematic and quantitative analyses of biological functions and pathways. Some 20,000 scientists around the world are still actively using it. In addition, my group was one of the early developers of the Gene Ontology (GO) system where we contributed to making the system to work for plant genomes. GO is a shared, controlled and structured vocabulary for describing gene attributes. GO has been instrumental in analyzing and interpreting genomic and post-genomic data across many organisms and has been used to analyze data in thousands of research articles, including many studies of various human diseases.

The Gene Ontology Consortium (2001) Creating the Gene Ontology Resource: Design and Implementation. **Genome Research** 11(8):1425-1433.

Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, Garcia-Hernandez M, Huala E, Lander G, Montoya M, Miller N, Mueller LA, Mundodi S, Reiser L, Tacklind J, Weems DC, Wu Y, Xu I, Yoo D, Yoon J, Zhang P. (2003) The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. **Nucleic Acids Research** 31(1):224-228.

Howe D, Costanzo M, Fey P, Gojobori T, Hannick L, Hide W, Hill DP, Kania R, Schaeffer M, St. Pierre S, Twigger S, White O, and Rhee SY (2008) The future of biocuration. **Nature** 455:47-50.

3. One of the biggest problems facing biology in the post-genome era is that we still do not know the functions of many genes (25%-75% of protein-encoding genes are not even predictable for their function based on sequence similarity), even for intensively studied organisms such as *E. coli*, yeast, and human. To systematically infer functions of genes and group them into pathways, my group collaborated with Dr. Ed Marcotte's group to create the first plant genome-wide co-function network called AraNet. It can be used to systematically identify new genes in pathways and infer functions of uncharacterized genes based on the functions of their network neighbors. In addition to contributing to the design and analysis of the network, my group demonstrated that AraNet could be used successfully to guide the functional identification of novel genes. Using molecular genetic approaches, we discovered novel regulators of drought resistance and lateral root development, traits that are essential in engineering drought resistance in plants.

Membrane proteins are perhaps the darkest matter in the pool of uncharacterized proteins because of the difficulty of working with them biochemically and expressing them heterologously. To better understand how proteins function across and within membranes, my group collaborated with Dr. Wolf Frommer's group to develop high-throughput experimental and computational pipelines to systematically identify interactions between membrane proteins and signaling proteins, testing over 6 million binary interactions between 3000 proteins. To date, this is still the largest eukaryotic membrane protein interaction network (such a network previously existed only for yeast, at ~10% of the scale). I led the bioinformatics component of the project where we created a computational pipeline to enable the large-scale experimental pipeline (primer design, sequence validation, and image and statistical analyses of the interactions) and analyzed the resulting protein interaction network. This is a foundational resource for generating many new hypotheses. The vast majority of the membrane protein interactions we found had never before been identified. In addition, the methods we developed for generating high-throughput membrane protein interactions are applicable to any species and the datasets will be useful in identifying patterns of signaling and regulation in plants.

Lee I, Ambaru B, Thakkar P, Marcotte E, and Rhee SY (2010) Rational association of genes with traits using a genome-scale gene network for Arabidopsis thaliana. **Nature Biotechnology** 2(28):149-156.

Jones AM, Xuan Y, Xu M, Wang R-S, Ho C-H, Lalonde S, You CH, Sardi MI, Parsa SA, Smith-Valle E, Su T, Frazer KA, Pilot G, Pratelli R, Grossmann G, Acharya BR, Hu HC, Engineer C, Villiers F, Ju C, Takeda K,

Su Z, Dong Q, Assmann SM, Chen J, Kwak JM, Schroeder JI, Albert R, Rhee SY, and Frommer WB (2014) Border control – a membrane-linked interactome of Arabidopsis. (2014) **Science** 344:711-716.
Bossi F, Fan J, Xiao J, Chandra L, Shen M, Dorone Y, Wagner D, Rhee SY* (2017) Systematic discovery of novel eukaryotic transcriptional regulators using sequence homology independent prediction. **BMC Genomics** 18(1):480

4. Plant metabolism plays a vital role in the health and well-being of our society. Despite our dependence on plants for energy, nutrition, and medicine, plant metabolism remains a surprisingly understudied field. For example, more than 30% of all pharmaceuticals are based on plant natural products, yet our knowledge of plant metabolic pathways accounts for less than 0.1% of the metabolites thought to exist in flowering plants. Understanding how plants evolved this prodigious chemical vocabulary has been a longstanding goal in plant biology. My group developed computational pipelines that systematically annotate enzyme function on the genome-scale. Using this system, we created a unique, unified resource of plant metabolic networks and discovered several properties that illustrate the differential evolution of secondary metabolism, permitting elucidation of novel secondary metabolic pathways. This opportunity is particularly relevant because secondary metabolites often confer upon plants the ability to survive major biotic and abiotic threats, and are the major sources of medicine, fragrance, and flavor. Thus, the molecular components involved in the production of secondary metabolites are a source of great interest across many fields of research, including agricultural biotechnology, synthetic biology, and biomedical and pharmaceutical research.

Mueller LA, Zhang P, and Rhee SY (2003) AraCyc. A Biochemical Pathway Database for Arabidopsis. **Plant Physiology** 132(2):453-60.

Zhang P, Dreher K, Karthikeyan A, Chi A, Pujar A, Caspi R, Karp P, Kirkup V, Latendresse M, Lee C, Mueller LA, Muller R, and Rhee SY (2010) Creation of a Genome-Wide Metabolic Pathway Database for *Populus trichocarpa* Using a New Approach for Reconstruction and Curation of Metabolic Pathways for Plants. **Plant Physiology** 153(4):1479-91.

Chae L, Kim T, Dreher K, and Rhee SY (2014) Genomic signatures of specialized metabolism in plants. **Science** 344:510-513

5. Transcriptional regulation is a fundamental process in biology and has been the subject of an intensive study. However, molecular, genetic, and evolutionary studies suggest that there must be additional layers of control that have not been discovered. To investigate into one of such layers, we used an integrated approach (applying concepts, data, and tools from computer science, genetics, genomics, proteomics, molecular evolution, development, and stress physiology) to uncover a new layer of transcriptional regulation across many domains of life. There are a handful of anecdotal examples of transcription factor-like proteins without a DNA binding domain, coined microProteins (miPs), which regulate evolutionarily related transcription factors. To test the prevalence of this mechanism, my group developed a genome-scale platform to discover, classify, and validate microProteins in Arabidopsis. We found over 400 putative miPs in Arabidopsis along with their putative target transcription factors and their respective biological pathways. In collaboration with experimental biologists at Carnegie and Stanford, we experimentally validated two novel miPs and their predicted target transcription factors using genetic, molecular, and biochemical experiments as a proof-of-concept. Given the prevalence of miPs in Arabidopsis, we applied the same strategy to predict miPs from 19 species, ranging from microbes to plants and metazoans. We detected putative miPs in all organisms examined and paired them with potential targets in almost all known transcription factor families. Our analysis suggests a potential ubiquitous layer of transcriptional regulation by miPs and provides a systematic framework for their future study. The potential universality of miP function may offer new tools to modulate transcription factor function in practical applications ranging from gene therapy to bioengineering.

Magnani E, De Klein N, Nam H-I, Kim J-G, Pham KL, Fiume E, Mudgett MB, and Rhee SY (2014) A comprehensive analysis of microProteins reveals their potentially widespread mechanism of transcriptional regulation. **Plant Physiology** 165(1):149-15.

de Klein N, Magnani E, and Rhee SY (2015) microProtein Prediction Program (miP3): a software for predicting microProteins and their target transcription factors. **International Journal of Genomics** Article ID 734147. 1-4.

Zhao K and Rhee SY* (2019) Epigenomic Landscape of *Arabidopsis thaliana* Metabolism Reveals Bivalent Chromatin on Specialized Metabolic Genes. **bioRxiv** 589036; doi: <https://doi.org/10.1101/589036>

Invited Seminars:

1. York University, York, UK (2000); 2. The Institute for Genome Research (2000); 3. Oxford University, Oxford, UK (2000); 4. Lorne Genome Conference, Melbourne, Australia (2001); 5. University of Arizona, Tucson, AZ (2001); 6. Plant Gene Expression Center, Albany, CA (2001); 7. Entigen, Sydney, Australia (2001); 8. Cornell University, Ithaca, NY (2001); 9. Seoul National University, Seoul, South Korea (2001); 10. University of California at Riverside, Riverside, CA (2002); 11. University of Guelph, Guelph, Canada (2002); 12. VIB, University of Gent, Gent, Belgium (2003); 13. Seoul National University, Seoul, South Korea (2004); 14. University of Missouri, Columbia (2004); 15. Plant Gene Expression Center, Albany, CA (2004); 16. Iowa State University (2007); 17. U.C. Riverside (2007); 18. University of Calgary, Canada (2007); 19. National Research Center-Plant Biotechnology Institute, Saskatoon, Canada (2007); 20. Danforth Center, MO (2007); 21. Dow Agrosciences (2010); 22. U.C. Riverside (2010); 23. U. Florida (2011); 24. Michigan State U. (2011); 25. Danforth Center (2012); 26. National Cheng Kung U., Taiwan (2012); 27. UC Davis (May 2013); 28. Washington State U. (April 2013); 29. University of Missouri-Columbia (Oct 22, 2013); 30. San Francisco State University (April 23, 2015); 31. Second Genome (Nov 9, 2015); 32. Louisiana State U. (Feb 22, 2016); 33. Plant Gene Expression Center, Albany CA (March 10, 2016); 34. Langebio, the National Laboratory of Genomics for Biodiversity, Mexico (May 10, 2016); 35. VIB, Ghent, Belgium (Sept 17, 2018); 36. U. Maryland at College Park (Nov 15, 2018); 37. Danforth Center (May 2019); 38. U. Kentucky (Nov 2019)

Invited Symposia:

1. Agricultural Genomics Conference, San Diego, CA (1999); 2. Advances in Genomic Research, Potentials and Applications, San Francisco State University, SF, CA (1999); 3. Genomic *Arabidopsis* Resource Network Workshop, York, UK (2000); 4. NSF Workshop on U.S.-Australia Interactions, Washington D.C. (2000); 5. AAAS Annual Meeting, San Francisco, CA (2001); 6. Plant & Animal Genome IX Conference, San Diego, CA (2001); 7. Mini-symposium on Plant Bioinformatics, KRIBB, Daejeon, South Korea (2001); 8. Plant Genome Awardees Meeting, San Diego, CA (2002); 9. Future of the National Plant Genome Initiative, National Academy of Sciences, Washington DC (2002); 10. National Science Foundation Managing Plant Genomic Resources Workshop, Asilomar, CA (2002); 11. International Horticultural Congress, Toronto, Canada (2002); 12. 2nd International Conference on Plant Metabolomics, Potsdam, Germany (2003); 13. Digital Archives for Science & Engineering Resources (DASER) Symposium (2003); 14. Crop Functional Genomics (2004); 15. 7th International Plant Cold Hardiness Symposium (2004); 16. Solanaceae Genomics Meeting, Jeju Island, S. Korea (2007); 17. Korea Genome Organization Meeting, Seoul, S. Korea (2007); 18. iPlant Kick-off Conference, Cold Spring Harbor Labs, NY (2008); 19. American Society of Plant Biologists Conference, Merida, Mexico (2008); 20. Banff Plant Metabolism Conference, Banff, Canada (2008); 21. International Conference on *Arabidopsis* Research (2010); 22. A Current Opinion Conference on Plant Genome Research, Amsterdam, the Netherlands (2011); 23. Alliance of Independent Plant Institutes Meeting, St. Louis, MO (2011); 24. ASMS Asilomar Conference on Mass Spectrometry, Asilomar, CA (2011); 25. Phenotype Ontology RCN, NASCent, Durham, NC (2012); 26. Plant Genomics in China XIII, Tai'an, China (2012); 27. Microbial and Plant Genomics Institute (MPGI) Symposium on Systems Biology of Genetic Regulation, St. Paul, MN (2012); 28. International Symposium on Root Systems Biology, Taipei, Taiwan (2012); 29. Evolution of Metabolic Diversity, Banbury, NY (March 2013); 30. Alliance of Independent Plant Institutes Meeting, Ithaca, NY (May 2013); 31. Plants and People Conference, Berlin, Germany (June 2013); 32. Meeting on the Genetic Basis of Unintended Effects in Modified Plants, Ottawa, Canada (Jan 14-15, 2014); 33. JGI DOE Conference: "Genomics of Energy and the Environment" (March 24-25, 2015); 34. Penn State Plant Biology Symposium: "Plant Stress-omics in a Changing Climate" (May 13-16, 2015); 35. Molecular Plant Symposium, Beijing, China (Aug 5-8, 2015); 36. Plant Genomes & Biotechnology: From Genes to Networks, CSHL, NY (Dec 2-5, 2015);

37. NAASC RCN: Arabidopsis Research and Training for the 21st century (May 13-15, 2016); 38. ASPB Conference, Plant Cell Symposium: New Biological Insights from Large-Scale Biology, Austin TX (July 9-13, 2016); 39. MSU - Plant Biotechnology for Health and Sustainability Annual Symposium (Oct 26-27, 2017), East Lansing, MI; 40. First International Plant Systems Biology Conference (Sept 10-14, 2018), Roscoff, France; 41. UC Davis Plant Symposium (April 8, 2019); 42. Plants of the Future, NYU NY (June 13-14, 2019)

PEER REVIEWED PUBLICATIONS

Peer reviewed publications (* corresponding author)

- Zhao K and Rhee SY* (2019) Epigenomic Landscape of *Arabidopsis thaliana* Metabolism Reveals Bivalent Chromatin on Specialized Metabolic Genes. **bioRxiv** 589036; doi: <https://doi.org/10.1101/589036>
- Bouain N, Korte A, Satbhai SB, Rhee SY, Busch W, Rouached H* (2019) Systems approaches provide new insights into *Arabidopsis thaliana* root growth under mineral nutrient limitation. **bioRxiv** 460360; doi: <https://doi.org/10.1101/460360> **PLOS Genetics Accepted**
- Rhee SY*, Birnbaum KD, Ehrhardt DW* (2019) Towards Building a Plant Cell Atlas. **Trends in Plant Science** 24(4):303-310
- Banf M*, Zhao K, Rhee SY* (2019) METACLUSTER - an R package for context-specific functionality analysis of metabolic gene clusters. **Bioinformatics** btz021, <https://doi.org/10.1093/bioinformatics/btz021>
- Lin F, Fan J, Rhee SY* (2018) QTG-Finder: a machine-learning algorithm to prioritize causal genes of quantitative trait loci in plants **bioRxiv** 484204; doi: <https://doi.org/10.1101/484204>
Genes|Genomes|Genetics Accepted
- Friesner J, Assmann SM, Bastow R, Bailey-Serres J, Beynon J, Brendel V, Buell CR, Bucksch A, Busch W, Demura T, Dinneny JR, Doherty CJ, Eveland AL, Falter-Braun P, Gehan MA, Gonzales M, Grotewold E, Gutierrez R, Kramer U, Krouk G, Ma S, Markelz RJC, Megraw M, Meyers BC, Murray JAH, Provart NJ, Rhee SY, Smith R, Spalding EP, Taylor C, Teal TK, Torii KU, Town C, Vaughn M, Vierstra R, Ware D, Wilkins O, Williams C, Brady SM (2017) The Next Generation of Training for Arabidopsis Researchers: Bioinformatics and Quantitative Biology. **Plant Physiol.** 175(4):1499-1509. doi: 10.1104/pp.17.01490. PubMed PMID: 29208732; PubMed Central PMCID: PMC5717721.
- Bossi F, Fan J, Xiao J, Chandra L, Shen M, Dorone Y, Wagner D, Rhee SY* (2017) Systematic discovery of novel eukaryotic transcriptional regulators using sequence homology independent prediction. **BMC Genomics** 18(1):480
- Rouchard H* and Rhee SY* (2017) System-level understanding of plant mineral nutrition in the big data era. **Current Opinion in Systems Biology** 4:71-77
- Schalpfer P, Zhang P, Chuan W, Kim T, Banf M, Chae L, Dreher K, Arvind C, Nilo-Poyanco R, Bernard T, Kahn D, and Rhee SY* (2017) Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants. **Plant Physiology** 173(4):2041-2059
- Chavali A* and Rhee SY* (2017) Bioinformatics tools for the identification of gene clusters that biosynthesize specialized metabolites. **Briefings in Bioinformatics**. bbx020. doi: 10.1093/bib/bbx020
- Banf M and Rhee SY* (2017) Enhancing gene regulatory network inference through data integration with markov random fields. **Nature Scientific Reports** 7:41174.
- Banf M* and Rhee SY* (2016) Computational inference of gene regulatory networks: approaches, limitations and opportunities. **BBA Gene Regulatory Mechanisms** S1874-9399(16)30188-2.
- Walsh JR, Schaeffer ML, Zhang, Rhee SY, Dickerson JA, Sen TZ* (2016) The quality of metabolic pathway resources depends on initial enzymatic function assignments: a case for maize. **BMC Systems Biology** 10:129.
- Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, Banf M, Dai X, Martin GB, Giovannoni JJ, Zhao PX, Rhee SY, Fei Z* (2016) iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. **Molecular Plant** S1674-2052(16)30223-4.
- Rhee SY*, Parker J, and Mockler T (2016) A glimpse into the future of genome-enabled plant biology from the shores of Cold Spring Harbor. **Genome Biology** 17(1):3.

- Fiume E, de Klein N, Rhee SY, and Magnani E* (2015) A framework for discovering, designing, and testing microProteins to regulate synthetic transcriptional modules. **Methods in Molecular Biology** 1482:175-88.
- Guo J, Fan J, Hauser B, and Rhee SY* (2015) Target enrichment improves mapping of complex traits by deep sequencing. **Genes | Genomes | Genetics** 6(1):67-77.
- Ladics G*, Bartholomaeus A, Bregitzer P, Doerrner N, Gray A, Holzhauser T, Jordan M, Keese P, Kok E, Macdonald P, Parrott W, Privalle L, Raybould A, Rhee SY, Rice E, Romeis J, Vaughn J, Wal J-M, and Glenn K (2015) Genetic basis and detection of unintended effects in genetically modified crop plants. **Transgenic Research** 24(4):587-603.
- Kim T, He K, Dreher K, Lee I, Moon S, Bais P, Dickerson J, Dixon P, Fiehn O, Lange BM, Sumner LW, Welti R, Wurtele ES, Nikolau BJ, and Rhee SY* (2015) Patterns of metabolite changes from large-scale gene perturbations in *Arabidopsis thaliana* using genome-scale metabolic networks. **Plant Physiology** 167(4):1685-98.
- de Klein N, Magnani E, and Rhee SY* microProtein Prediction Program (miP3): a software for predicting microProteins and their target transcription factors. (2015) **International Journal of Genomics** Article ID 734147. 1-4.
- Peng J, Uygun S, Kim T, Wang Y*, Rhee SY*, and Chen J* (2015) Measuring genome-specific semantic similarities using Gene Ontology and Gene Co-Function networks. **BMC Bioinformatics** 16(1):44.
- Xu M and Rhee SY* (2014) Becoming data-savvy in a big-data world. **Trends in Plant Science** 19(10):619–622.
- Jones AM, Xuan Y, Xu M, Wang R-S, Ho C-H, Lalonde S, You CH, Sardi MI, Parsa SA, Smith-Valle E, Su T, Frazer KA, Pilot G, Pratelli R, Grossmann G, Acharya BR, Hu HC, Engineer C, Villiers F, Ju C, Takeda K, Su Z, Dong Q, Assmann SM, Chen J, Kwak JM, Schroeder JI, Albert R, Rhee SY*, and Frommer WB* (2014) Border control – a membrane-linked interactome of *Arabidopsis*. **Science** 344:711-716.
[highlighted in F1000]
- Chae L, Kim T, Nilo-Poyanco R, and Rhee SY* Genomic signatures of specialized metabolism in plants. (2014) **Science** 344:510-513. **[highlighted in F1000]**
- Magnani E, De Klein N, Nam H-I, Kim J-G, Pham KL, Fiume E, Mudgett MB, and Rhee SY* (2014) A comprehensive analysis of microProteins reveals their potentially widespread mechanism of transcriptional regulation. **Plant Physiology** 165(1):149-159.
- Rhee SY* and Mutwil M* (2014) Towards revealing the functions of all genes in plants. **Trends in Plant Science** 19(4):212-221.
- Bassel GW, Gaudinier A, Brady SM, Hennig L, Rhee SY, and Smet ID* (2012) Systems analysis of plant functional, transcriptional, physical interaction, and metabolic networks. **Plant Cell** 24(10):3859-75.
- Chen J, Lalonde S, Obrdlik P, Noorani Vatani A, Parsa SA, Vilariño C, Revuelta JL, Frommer WB, and Rhee SY* (2012) Uncovering *Arabidopsis* membrane protein interactome enriched in transporters using mating-based split ubiquitin assays and classification models. **Frontiers in Plant Science** 3(124):1-14.
- Moon S, He Kun, Bais P, Dickerson J, Dixon P, Rhee SY, Wohlgemuth G, Fiehn O, Barkan L, Lange I, Lange B, Cortes D, Shuman J, Shulaev V, Huhman D, Sumner L, Roth M, Welti R, Ilarlan H, Wurtele E, Brachova L, Campbell A, Perera A, and Nikolau B* (2012) Metabolomics as a hypothesis-generating functional genomics tool for the annotation of *Arabidopsis thaliana* genes of “unknown function”. **Frontiers in Plant Science** 3(15):1-12.
- Chae L, Lee I, Shin J, and Rhee SY* (2012) Towards an understanding of how molecular networks evolve in plants. **Current Opinion in Plant Biology** 15(2):177-184.
- Hwang S, Rhee SY*, Marcotte EM*, and Lee I* (2011) Systematic prediction of gene function using a probabilistic functional gene network for *Arabidopsis thaliana*. **Nature Protocols** 6(9):1429-1442.
- Sun Y, Fan X-Y, Cao D-M, He K, Tang W, Zhu J-Y, He J-X, Bai M-Y, Zhu S, Oh E, Patil S, Kim TW, Ji H, Wong WH, Rhee SY, and Wang J-Y* (2010) Integration of Brassinosteroid Signal Transduction with the Transcription Network for Plant Growth Regulation in *Arabidopsis*. **Developmental Cell** 19(5):765-77.
- Lalonde S, Sero A, Pratelli R, Pilot G, Chen J, Sardi MA, Parsa SA, Kim D-Y, Acharya BR, Stein EV, Hu H-C, Villiers F, Takeda K, Yang Y, Han YS, Schwacke R, Chiang W, Kato N, Loqué D, Assmann SM, Kwak JM,

- Schroeder J, Rhee SY, and Frommer WB* (2010) A membrane protein / signaling protein interaction network for *Arabidopsis* version AMPv2. **Frontiers in Plant Science** 1(24):1-14.
- Zhang P, Dreher K, Karthikeyan A, Chi A, Pujar A, Caspi R, Karp P, Kirkup V, Latendresse M, Lee C, Mueller LA, Muller R, and Rhee SY* (2010) Creation of a Genome-Wide Metabolic Pathway Database for *Populus trichocarpa* Using a New Approach for Reconstruction and Curation of Metabolic Pathways for Plants. **Plant Physiology** 153(4):1479-91.
- Bais P, Moon S, He K, Leitao R, Dreher K, Walk T, Sucaet Y, Barkan L, Wohlgemuth G, Wurtele ES, Dixon P, Fiehn O, Lange BM, Shulaev V, Sumner LW, Welti R, Nikolau B, Rhee SY, and Dickerson JA* (2010) PlantMetabolomics.org: A web portal for Plant Metabolomics Experiments. **Plant Physiology** 152(4):1807-16.
- Lee I, Ambaru B, Thakkar P, Marcotte E*, and Rhee SY* (2010) Rational association of genes with traits using a genome-scale gene network for *Arabidopsis thaliana*. **Nature Biotechnology** 2(28):149-156.
[highlighted in F1000]
- Reference Genome Group of the Gene Ontology Consortium (2009) The Gene Ontology's Reference Genome Project: a unified framework for functional annotation across species. **PLOS Computational Biology** 5(7): e1000431.
- Chen J, Ji L, Hsu W, Tan K-L, and Rhee SY* (2009) Exploiting Domain Knowledge to Improve Biological Significance of Biclusters with Key Missing Genes. **IEEE Technical Committee on Data Engineering Conference ICED.2009.205**: 1219-1222.
- Aceituno FF, Moseyko N, Rhee SY, and Gutierrez RA* (2008) The rules of gene expression in plants: Organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. **BMC Genomics** 9:438.
- Howe D*, Costanzo M, Fey P, Gojobori T, Hannick L, Hide W, Hill DP, Kania R, Schaeffer M, St. Pierre S, Twigger S, White O, and Rhee SY* (2008) The future of biocuration. **Nature** 455:47-50.
- Rhee SY*, Wood V, Dolinski K, and Draghici S* (2008) Use and Misuse of the Gene Ontology (GO) Annotations. **Nature Review Genetics** 9(7):509-15.
- Pennycooke JC, Cheng H, Roberts SM, Yang Q, Rhee SY, and Stockinger E* (2008) The low temperature-responsive, *Solanum* CBF1 genes maintain high identity in their upstream regions in a genomic environment undergoing gene duplications, deletions, and rearrangements. **Plant Molecular Biology** 67(5):483-97.
- Lalonde S*, Ehrhardt D, Loqué D, Chen J, Rhee SY, and Frommer WB (2008) Molecular and cellular approaches for the detection of protein-protein interactions and generation of protein interaction maps. **Plant Journal** 53(4):610-35.
- Avraham S, Tung C-W, Ilic K, Jaiswal P, Kellogg EA, McCouch S, Pujar A, Reiser L, Rhee SY, Sachs MM, Schaeffer M, Stein L, Stevens P, Vincent L, Zapata F, and Ware D* (2008) The Plant Ontology Database: a community resource for plant structure and developmental stages controlled vocabulary and annotations. **Nucleic Acids Research** 36:D449-D454.
- Fiehn O*, Sumner LW, Rhee SY, Ward J, Dickerson J, Lange BM, Lane G, Roessner U, Last R, and Nikolau B (2007) Minimum reporting standards for plant biology context information in metabolomic studies. **Metabolomics** 3(3):195-201.
- Caspi R, Foerster H, Fulcher CA, Kaipa P, Krummenacker M, Latendresse M, Paley S, Rhee SY, Shearer AG, Tissier C, Walk TC, Zhang P, and Karp PD* (2008) The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. **Nucleic Acids Research** 38(Database issue):D473-9.
- The Gene Ontology Consortium (2008) The Gene Ontology project in 2008. **Nucleic Acids Research** 36(Database issue):D440-4.
- Ilic K, Kellogg E, Jaiswal P, Zapata F, Stevens P, Vincent L, Pujar A, Avraham S, Reiser L, McCouch SR, Sachs S, Schaeffer M, Ware D, Stein L, and Rhee SY* (2006) Plant Structure Ontology: A Unified Vocabulary for Flowering Plants. **Plant Physiology** 143(2):587-99.
- Stein LD, Beavis WD, Gessler DD, Huala E, Lawrence CJ, Main D, Mueller LA, Rhee SY, and Rokhsar DS* (2006) Save our data! **Scientist** 20(4):24-25.

- Pujar A, Jaiswal P, Kellogg EA, Ilic K, Vincent L, Avraham S, Stevens P, Zapata F, Reiser R, Rhee SY, Sachs MM, Schaeffer M, Stein L, Ware D, and McCouch S* (2006) Whole Plant Growth Stage Ontology: History, Development and Application. **Plant Physiology** 142(2):414-28.
- Leebens-Mack J*, Vision T, Brenner E, Bowers JE, Cannon S, Clement MJ, Cunningham CW, dePamphilis C, deSalle R, Doyle JJ, Eisen JA, Gu X, Harshman J, Kellogg EA, Koonin EV, Philippe H, Pires JC, Qiu YL, Rhee SY, Sjölander K, Soltis DE, Soltis PS, Stevens P, Stevenson DW, Warnow T, and Zmasek C. (2006) Taking the First Steps Towards a Standard for Reporting on Phylogenies: Minimal Information About a Phylogenetic Analysis (MIAPA). **OMICS** 10(2):231-237.
- Rhee SY*, Dickerson J*, and Xu D* (2006) Bioinformatics and its Applications in Plant Biology. **Annual Review of Plant Biology** 57: 335-360.
- Zimmermann P*, Schildknecht B, Craigon D, Garcia-Hernandez M, Grisse W, May S, Mukherjee G, Parkinson H, Rhee SY, Wagner U, and Hennig L. (2006) MIAME/Plant – adding value to plant microarray experiments. **Plant Methods** 2:1-3.
- Gene Ontology Consortium (2006) The Gene Ontology Project in 2006. **Nucleic Acids Research** 34(Database issue):D322-6.
- Caspi R, Foerster H, Fulcher C, Hopkinson R, Ingraham J, Kaipa P, Krummenacker M, Paley S, Pick J, Rhee SY, Tissier C, Zhang P, and Karp P* (2006) MetaCyc: A multiorganism database of metabolic pathways and enzymes. **Nucleic Acids Research** 34(Database issue):D511-6.
- Li S, Ehrhardt D, and Rhee SY* (2006) Systematic Analysis of *Arabidopsis* Protein Localization and Software Tools for Fluorescent Tagging of Full-Length *Arabidopsis* Proteins. **Plant Physiology** 141(2):527-39. **[highlighted in F1000]**
- Jaiswal P*, Avraham S, Ilic K, Kellogg EA, McCouch S, Pujar A, Reiser L, Rhee SY, Sachs MM, Schaeffer M, Stein L, Stevens P, Vincent L, Ware D, and Zapata F. (2005) Plant Ontology (PO): A controlled vocabulary of plant structures and growth stages. **Functional and Integrated Genomics** 6:388-397.
- Rhee SY* (2005) Bioinformatics: Current Limitations and Insights for the Future. **Plant Physiology** 138(2):569-70.
- Yan T, Yoo D, Berardini T, Mueller L, Weems D, Weng S, Cherry JM, and Rhee SY* (2005) PatMatch: a program for finding patterns in peptide and nucleotide sequences. **Nucleic Acids Research** 33(Web Server issue):W262-6.
- Zhang P, Foerster H, Tissier CP, Mueller L, Paley S, Karp P, and Rhee SY* (2005) MetaCyc and AraCyc: metabolic pathway databases for plant research. **Plant Physiology** 138(1):27-37.
- Bard J, Rhee SY, and Ashburner M* (2005) An ontology for cell types. **Genome Biology** 6:R21.
- Schlueter SD, Wilkerson MD, Huala E, Rhee SY, and Brendel V* (2005) Community-based gene structure annotation. **Trends in Plant Science** 10(1):9-14.
- Jenkins H*, Hardy N, Beckmann M, Draper J, Smith AR, Taylor J, Fiehn O, Goodacre R, Bino RJ, Hall R, Kopka K, Lange BM, Liu JR, Mendes P, Nikolau BJ, Oliver SG, Paton NW, Rhee SY, Roessner-Tunali U, Saito K, Smedsgaard J, Sumner LW, Wurtele ES, and Kell DB (2004) A proposed framework for the description of plant metabolomics experiments and their results. **Nature Biotechnology** 22(12):1601-6.
- Zhang X, Fowler S, Cheng H, Lou Y, Rhee SY, Stockinger EJ, and Thomashow MF* (2004) Freezing Sensitive Tomato has a Functional CBF Cold Response Pathway, but a CBF Regulon that Differs from that of Freezing Tolerant *Arabidopsis*. **Plant Journal** 39(6):905-19.
- Berardini TZ, Mundodi S, Reiser R, Huala E, Garcia-Hernandez M, Zhang P, Mueller LM, Yoon J, Doyle A, Lander G, Moseyko N, Yoo D, Xu I, Zoeckler B, Montoya M, Miller N, Weems D, and Rhee SY* (2004) Functional annotation of the *Arabidopsis* genome using controlled vocabularies. **Plant Physiology** 135(2):1-11.
- Tian GW, Mohanty A, Chary SN, Li S, Paap B, Drakakis G, Kopec C, Li J, Ehrhardt E, Jackson D, Rhee SY, Raikhel N, and Citovsky V* (2004) High-Throughput Fluorescent Tagging of Full-Length *Arabidopsis* Gene Products in *Planta*. **Plant Physiology** 135(1):25-38. **[highlighted in F1000]**
- Dolan EL, Soots BE, Lemaux PG, Rhee SY, and Reiser L* (2004) Strategies to Avoid Reinventing the Pre-college Education and Outreach Wheel. **Genetics** 166:1601-1609.

- Weems D, Miller N, Garcia-Hernandez M, Huala E, and Rhee SY* (2004) Design, implementation, and maintenance of a model organism database for *Arabidopsis thaliana*. **Comparative and Functional Genomics** 5(4):362-369.
- Thimm O, Bläsing YG, Nagel, A, Meyer, S, Kruger, P, Selbig, J, Müller, L, Rhee, SY, and Stitt M* (2004) MapMan: A User-Driven Tool to Display Genomics Data Sets onto Diagrams of Metabolic Pathways and other Biological Processes. **Plant Journal** 37(6):914-39. **[highlighted in F1000]**
- Bard JL* and Rhee SY* (2004) Ontologies in biology: design, applications and future challenges. **Nature Review Genetics** 5(3):213-22.
- Rhee SY* (2004) Carpe Diem. Retooling the Publish or Perish Model into the Share and Survive Model. **Plant Physiology** 134(2):543-7.
- Krieger CJ, Zhang P, Mueller L, Wang A, Paley S, Arnaud M, Pick J, Rhee SY, and Karp P* (2004) MetaCyc: Recent enhancements to a database of metabolic pathways and enzymes in microorganisms and plants. **Nucleic Acids Research** 32 Database issue:D438-42.
- Harris MA*, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, Sethuraman A, Theesfeld CL, Botstein D, Dolinski K, Feierbach B, Berardini T, Mundodi S, Rhee SY, Apweiler R, Barrell D, Camon E, Dimmer E, Lee V, Chisholm R, Gaudet P, Kibbe W, Kishore R, Schwarz EM, Sternberg P, Gwinn M, Hannick L, Wortman J, Berriman M, Wood V, de la Cruz N, Tonellato P, Jaiswal P, Seigfried T, and White R; Gene Ontology Consortium. (2004) The Gene Ontology (GO) database and informatics resource. **Nucleic Acids Research** 32 Database issue:D258-61.
- Rhee SY, Osborne E, Poindexter P, and Somerville CR* (2003) Microspore separation in the quartet 3 mutants of *Arabidopsis* is impaired by a defect in a developmentally regulated pectinase required for pollen mother cell degradation. **Plant Physiology** 133(3):1170-80.
- Mueller LA, Zhang P, and Rhee SY* (2003) AraCyc. A Biochemical Pathway Database for *Arabidopsis*. **Plant Physiology** 132(2):453-60.
- Rhee SY*, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, Garcia-Hernandez M, Huala E, Lander G, Montoya M, Miller N, Mueller LA, Mundodi S, Reiser L, Tacklind J, Weems DC, Wu Y, Xu I, Yoo D, Yoon J, and Zhang P (2003) The *Arabidopsis* Information Resource (TAIR): a model organism database providing a centralized, curated gateway to *Arabidopsis* biology, research materials and community. **Nucleic Acids Research** 31(1):224-228.
- Clarke BC, Lambrecht M, and Rhee SY* (2003) *Arabidopsis* genomic information for interpreting wheat EST sequences. **Functional and Integrated Genomics** 3(1):33-38.
- Garcia-Hernandez M*, Berardini TZ, Chen C, Crist D, Doyle A, Huala E, Knee E, Miller N, Mueller L, Mundodi S, Reiser L, Rhee SY, Scholl R, Tacklind J, Weems D, Wu Y, Xu I, Yoo D, Yoon J, and Zhang P (2002) **Functional and Integrated Genomics** 2(6):239-253.
- Reiser L, Mueller LA, and Rhee SY* (2002) Surviving in a sea of data: a survey of plant genome data resources and issues in building data management systems. **Plant Molecular Biology** 48(1):59-74.
- The Gene Ontology Consortium (2001) Creating the Gene Ontology Resource: Design and Implementation. **Genome Research** 11(8):1425-1433.
- Huala E*, Dickerman A, Garcia-Hernandez M, Weems D, Reiser L, LaFond F, Hanley D, Kiphart D, Zhuang J, Huang W, Mueller L, Bhattacharyya D, Bhaya D, Sobral B, Beavis B, Somerville C, and Rhee SY* (2001) The *Arabidopsis* Information Resource (TAIR): A comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. **Nucleic Acids Research** 29(1):102-5.
- Rhee SY* (2000) Bioinformatic resources, challenges, and opportunities using *Arabidopsis thaliana* as a model organism in post-genomic era. **Plant Physiology** 2000 124(4):1460-4.
- Rhee SY, Weng S, Bongard-Pierce DK, Garcia-Hernandez M, Malekian A, Flanders, DJ, and Cherry JM* (1999) Unified display of *Arabidopsis thaliana* physical maps from AtDB, the *A. thaliana* database. **Nucleic Acids Research** 27(1):79-84.

- Rhee SY, Weng S, Flanders D, Cherry JM, Dean C, Lister C, Anderson M, Koornneef M, Meinke DW, Nickle T, Smith K, and Rounsley SD (1998) Genome maps 9. *Arabidopsis thaliana*. Wall chart. **Science** 282(5389):663-7.
- Rhee SY and Somerville CR* (1998) Tetrad pollen formation in quartet mutants of *Arabidopsis thaliana* is associated with persistence of pectic polysaccharides of the pollen mother cell wall. **Plant Journal** 15:79-88.
- Rhee SY and Somerville CR* (1994) Flat-Surface Grafting in *Arabidopsis thaliana*. **Plant Molecular Biology Reporter** 13:118-123.
- Preuss D*, Rhee SY, Davis RW (1994) Tetrad analysis possible in *Arabidopsis* with mutation of the QUARTET (QRT) genes. **Science** 264:1458-60.
- Ruan ZS, Anantharam V, Crawford IT, Ambudkar SV, Rhee SY, Allison MJ, and Maloney PC* (1992) Identification, purification, and reconstitution of OxIT, the oxalate: formate antiport protein of *Oxalobacter formigenes*. **Journal of Biological Chemistry** 267:10537-43.

Book chapters and other publications

- Rhee SY (2014) An interview with Seung Yon (Sue) Rhee. **Trends in Plant Science** 19(4):198-199.
- Ilic K, Stevens PF, Kellogg EA, and Rhee SY (2008) Plant Structure Ontology –anatomical ontology of flowering plants. In: Anatomy Ontologies for Bioinformatics: Principles and Practice. A. Burger, D. Davidson and R. Baldock (Eds). Springer. 27-42. ISBN: 184628884.
- Rhee SY and Crosby W. (2005) Biological Databases for Plant Research. **Plant Physiology** 138(1):1-3.
- Yoo D, Xu I, Berardini T, Rhee SY, Narayanasami V, and Twigger S (2005) PubSearch and PubFetch, a simple management system for semi-automated retrieval and annotation of biological information from the literature. In Current Protocols in Bioinformatics. John Wiley & Sons. Chapter 9. Unit 9.7
- Reiser L and Rhee SY (2005) Using The *Arabidopsis* Information Resource (TAIR) to Find Information About *Arabidopsis* Genes. In Current Protocols in Bioinformatics. John Wiley & Sons. Chapter 1.11.
- Berardini TA and Rhee SY (2004) *Arabidopsis thaliana*: Characteristics and Annotation of a Model Genome. In Encyclopedia of Plant & Crop Science. Marcel Dekker, Inc. 47-50.
- Rhee SY, Zhang P, and Foerster H, and Tissier C (2005) AraCyc: Overview of an *Arabidopsis* Metabolism Database and Its Applications for Plant Research. In Biotechnology in Agriculture and Forestry: Plant Metabolomics. K. Saito, R. Dixon and L. Willmitzer ed., Springer. Volume 57. pp. 141-153.
- Rhee SY (2001) Extending the Frontiers of Plant Biology: Large scale biology, bioinformatics, and information management. **Plant Biotechnology Institute Bulletin** May Issue:10-12.
- Rhee SY and Flanders DJ (2000) Web-based bioinformatic tools for *Arabidopsis* researchers. In *Arabidopsis: A Practical Approach*. pp. 225-265. Zoe Wilson ed., Oxford University Press, UK.

POPULAR SCIENCE ESSAYS:

- Rhee SY (1997) Gregor Mendel (1822-1884). Commissioned and originally published on Genetech. Reprinted at Access Excellence, a national educational program that provides high school biology teacher's access to their colleagues, scientists, and critical sources of new scientific information via the World Wide Web. Currently available at: http://www.cccbotechnology.com/RC/AB/BC/Gregor_Mendel.php, http://www.math.usu.edu/rheal/stat1040/lecture_notes/Chapter_26.pdf, <https://www.verslo.is/home/Raungreinar/lif/ltarefni/Erfdir/Gregor%20Mendel%28e%29.htm>
- Rhee SY (1997) Louis Pasteur (1822-1895). Commissioned and originally published on Genetech. Reprinted at Access Excellence, a national educational program that provides high school biology teacher's access to their colleagues, scientists, and critical sources of new scientific information via the World Wide Web. Currently available at: http://www.cccbotechnology.com/RC/AB/BC/Louis_Pasteur.php
- Rhee SY (1997) Linus Pauling (1901-1994). Commissioned and originally published on Genetech. Reprinted at Access Excellence, a national educational program that provides high school biology teacher's access to their colleagues, scientists, and critical sources of new scientific information via the World Wide Web. Currently available at: http://www.cccbotechnology.com/RC/AB/BC/Linus_Pauling.php

Rhee SY (1997) Kary B. Mullis (1944 -). Commissioned and originally published on Genetech. Reprinted at Access Excellence, a national educational program that provides high school biology teacher's access to their colleagues, scientists, and critical sources of new scientific information via the World Wide Web. Currently available at: http://www.cccbotechnology.com/RC/AB/BC/Kary_B_Mullis.php

People Trained:

	First Name	Last Name	Title	Start Date	End Date	Present Position
135	Elena	Lazarus	RA	7/19	-	
134	David	Huang	Intern	6/19	8/19	UCSD
133	Suzie	Lee	Intern	6/19	8/19	Cal Poly
132	Emily	Fryer	RA	3/19	-	
131	Navadeep	Boruah	Postdoc	10/18	-	
130	Karine	Prado	Postdoc	10/18	-	
129	Sury	Jha	Postdoc	9/18	-	
128	Cheng	Zhao	Postdoc	8/18	-	
127	Charles	Hawkins	Curator	8/18	-	
126	Nienke	Besbrugge	Postdoc	8/18	4/19	Postdoc in Ken Caldeira's lab, Carnegie DGE
125	Angela	Xu	CA	6/18	-	
124	Kevin	Radja	RA	6/18	-	
123	Ankush	Bharadwaj	Intern	6/18	8/18	UCLA
122	Bharti	Parihar	Intern	6/18	8/18	CSU, East Bay
121	Lyn	Vakulenko	Intern	5/18	8/18	Carlton U, Canada
120	Benjamin	Jin	RA	11/17	-	
119	Jiun	Yen	Postdoc	10/17	-	
118	Fan	Lin	Postdoc	9/17	-	
117	Kangmei	Zhao	Postdoc	7/16	-	
116	Hatem	Rouached	Visiting faculty	2016	2018	INRA, France
115	Sam	Craig	Intern	6/16	8/16	PA High
114	Elena	Estrada	Intern	6/16	8/16	UCSD
113	Nikhil	Kaimal	Intern	6/16 6/17	8/16 8/17	UC Davis
112	JeaneAe	Kim	RA	2/16	7/17	PhD candidate, UC Riverside
111	Arvind	Chavali	Postdoc	1/16	3/18	Strategy consultant, ZS Associates
110	Bo	Xu	Intern RA	9/15 2/16	1/15 -	
109	Dylan	Koh	Intern	6/15	8/15	Stanford
108	Vivek	Sriram	Intern	6/15	8/15	Duke U.
107	JeanAe	Kim	Intern	6/15	1/16	PhD candidate at UC Riverside
106	Luong	Mai	Intern	6/15	12/15	Cornell U.
105	Yanniv	Dorone	Ph.D. student	9/15	-	
104	Pascal	Schläpfer	Postdoc	8/14	12/17	Group leader at ETH
103	In-Seob	Han	Visiting faculty	2014	2015	U Ulsan, Korea
102	Jenny	Guarino	Intern	5/14	8/14	Cornell
101	Phong	Nguyen	Intern	5/14	8/14	

			RA	9/14	5/15	
100	Michael	Banf	Postdoc	1/14	2/17	Consultant, Max Planck Institute for Plant Breeding Research
99	Catherine	Doyle	Intern	6/13	8/13	PhD candidate at NC State
98	Jue	Fan	Postdoc	5/13	3/15	Bioinformatics Director, IsoTex
97	Lessley	Peterson	Intern	1/13	10/13	U. Maryland
96	Bernie	Hauser	Visiting faculty	2013	2014	U Floria at Gainesville, USA
95	Chuan	Wang	Postdoc	10/12	2/16	Senior Bioinformatics Scientist, Brightseed
94	Lan	Jiang	Intern	10/12	2/13	
93	Lilyana	Chandra	Intern	6/12	12/12	UC Berkeley
92	Varun	Dwaraka	Intern	7/12	9/12	PhD candidate, UC Santa Cruz
91	Mohan	Avula	Intern	6/12	8/12	Los Altos HS
90	Jim	Guo	Postdoc	7/12	3/15	Staff Scientist, Bioinformatics at the Clinical NGS group at Thermo Fisher
89	Ricardo	Nilo Poyanco	Postdoc	3/12	5/14	Lecturer at School of Biotechnology, Faculty of Science, U. Mayor, Chile
88	Meng	Xu	Postdoc	11/11	8/14	Director of Bioinformatics, Brightseed
87	Taehyong	Kim	Postdoc	3/11	7/15	Principal bioinformatician, Institute for Biomedical Informatics University of Pennsylvania
86	Tam	Tran	Intern	11/11	8/12	Research associate, Mayo Clinic, Rochester, NY
85	Damian	Priamurskiy	Intern	6/11	3/12	
84	Caryn	Johansen	Intern Intern RA	6/11 7/12 10/12	9/11 9/12 7/14	PhD candidate at UC Davis
83	Flavia	Bossi	Postdoc Senior RA	2/10 3/18	2/18 -	
82	Hye-In	Nam	RA	1/10	5/19	Unknown
81	Niek	deKlein	Intern	9/10	1/11	PhD candidate, U. Groningen
80	Kris	Sankaran	Intern	9/10	4/11	PhD candidate, Stanford U.
79	Julian	Huang	Intern	6/10	12/10	Med student, Yale
78	Nathaniel	Leu	Intern	6/10	9/10	Medical student, Eastern Virginia Medical School
77	Kim	Pham	Intern	6/10	12/10	
76	Chang	You	Postdoc	8/09	5/11	Data Science Engineer at MZ
75	Rupa	Paduchuri	Intern	10/09	12/11	
74	Jon	Illoreta	Intern	6/09	12/10	
73	Purva	Karia	Intern	6/09 1/11	7/09 6/11	PhD candidate, McGill U
72	Cherise	Lau	Intern	6/09 7/10	8/09 9/10	UCLA
71	Vibhu	Bakshi	Intern	3/09	6/09	U. North Texas, Ph.D.
70	Pranjali	Karia	Intern	3/09	11/09	
69	Azam	Noorani	RA	12/08	8/09	Unknown
68	Sagaya	Arokiasam	RA	10/08	12/08	Unknown

		y				
67	Lee	Chae	Postdoc	7/08	5/14	Cofounder and CTO, Brightseed
66	Ricardo	Leitão	Intern	7/08	1/09	PhD candidate, UCSC
65	Michael	Ahn	Intern	7/08 8/09	8/08 8/09	Masters, Harvard Grad School of Design
64	Anjo	Chi	Programmer	1/08	1/10	
63	AS	Karthikyan	Curator	1/08	11/09	Unknown
62	Kun	He	Postdoc	9/07	12/09	Decision Sciences Platform Lead, Associate Fellow at Bayer Crop Science
61	Ozgur	Ozturk	Postdoc	8/07	3/08	Software Architect / Data Scientist at Philips Wellcentive
60	Bindu	Ambaru	RA	7/07	4/10	PhD student, Manipal University, India
59	Joy	Zhang	Intern	6/07	8/07	
58	Liping	Ji	Postdoc	5/07	5/08	Associate prof, Harbin Inst. Tech.
57	Adeline	Wong	Intern	4/07	9/07	
56	Suzanne	Fleshman	CA	11/07	2/08	Unknown
55	Kate	Dreher	Curator	11/07	11/13	Germplasm coordinator, CYMMIT
54	Cindy	Lee	Programmer	10/07	7/11	Software Developer at CardioDx
53	Raymond	Chetty	Programmer	10/07	5/10	
52	Larry	Ploetz	SA	3/07	5/13	Operations Engineer at Pinger
51	Jin	Chen	Postdoc	12/06	8/09	Associate prof, U. Kentucky
50	Phillipe	Lamesch	Curator	12/06	3/12	Communication adviser, Luxembourg Centre for Systems Biomedicine, U. Luxembourg
49	Vanessa	Swing	CA Webmaster	10/06 3/07	3/07 11/09	
48	Donghui	Li	Curator	7/06	6/14	Curator, Phoenix Informatics
47	Shanker	Singh	DBA	7/06	3/13	
46	John	McGee	Intern	6/06	8/06	Scientific Founder, FogPharma
45	Tom	Meyer	Programmer	6/06	5/10	Software Consultant, Tegular Software
44	Bob	Muller	IT	5/06	6/14	
43	Tom	Walk	Postdoc	1/06	9/07	Large Plant Breeding Database Manager, North Dakota State U.
42	Noah	Whitman	RA	1/06	4/07	Unknown
41	Natasha	Raikhel	Visiting faculty	2006	2006	UC Riverside, USA
40	Mohammed	Shaikh	Programmer	12/05	1/06	Unknown
39	David	Swarbreck	Curator	10/05	9/10	Group leader, Genome Analysis Team, The Genome Analysis Centre, UK
38	Ryan	Pham	Intern	6/05	8/05	
37	Dan	MacLean	Postdoc	5/05	6/06	Director of bioinformatics, Sainsbury labs, UK
36	Joe	Filla	SA	2/05	3/07	Director of Operations, QuinStreet
35	Christophe	Tissier	Curator	1/05	9/07	Project manager, Learning in Motion, Inc.
34	Hartmut	Foerster	Curator	8/04	8/07	Database curator, SGN
33	Chris	Wilks	Intern Programmer	6/04 6/05	5/05 1/11	Graduate student, UCSC
32	Jon	Slenk	Programmer	4/05	7/06	Software Engineer, Apple

31	Aleksey	Kleytman	CA	2/5	6/6	Tech support
30	Renee	Halbrook	Intern	6/04	8/04	
29	Doug	Becker	Programmer	5/04	4/06	Technical Accountant Manager, NetApp
28	Katica	Ilic	Curator	1/04	10/06	Senior scientist, Fluidigm Corporation
27	Jessie	Zhang	Programmer	9/03	5/04	Unknown
26	Brandon	Zoeckler	CA	8/03	3/6	Research technician, UC Berkeley
25	Thomas	Yan	Intern	7/03	6/06	Senior staff engineer, SanDisk
24	Behzad	Mahini	Programmer	3/03	1/04	Unknown
23	Shijun	Li	Postdoc	10/02	10/05	QC analyst, ThermoFisher
22	Yigong	Lou	Postdoc	9/02	10/04	Bioinformatics analyst, LBL
21	Nick	Moseyko	Curator	9/02	10/05	Bioinformatics analyst, UC Berkeley
20	Gabriel	Lander	CA	8/02	3/03	Assistant professor, Scripps Institute
19	Peifen	Zhang	Curator	4/02	7/18	Curator, TAIR, Pheonix Informatics Inc.
18	Suparna	Mundodi	Curator	3/02	5/06	Bioinformatics analyst, Biorad Inc.
17	Tanya	Berardini	Curator	1/02	6/14	Curator, Phoenix Informatics
16	Julie	Tacklind	Webmaster	12/01	2/07	
15	Chunxia	Xu	Programmer	10/01	1/06	Unknown
14	Danny	Yoo	Programmer	6/01	5/06	Software engineer, Google
13	Mark	Lambrecht	Postdoc	4/01	2/02	Director of the Health and Life Sciences Global Practice at SAS
12	Bryan	Murtha	Programmer	2/01	8/01	Unknown
11	Holly	Nottage	Intern	2/01	7/01	
10	Jill	Larimore	Intern	1/01	4/02	Graduate student, Gladstone Institute, UCSF
9	Aisling	Doyle	CA	11/00	10/02	Biosafety inspector, University College Cork
8	Jungwon	Yoon	CA	9/00	6/03	Unknown
7	Anell	Bengt	Visiting student	9/00	12/01	Software developer, AstraZeneca Inc.
6	Smita	Mitra	Visiting student	8/00	10/00	Software developer, IBM Life Sciences
5	Debika	Bhattacharyya	Intern	7/00	8/00	Enterprise architect, Oracle
4	Lukas	Mueller	Curator	6/00	7/03	Associate professor, Cornell University, BTI
3	Leonore	Reiser	Curator	12/99	3/06	Program manager at Breakout Labs, Thiel Foundation
2	Margarita	Garcia-Hernandez	Curator	9/99	12/06	Research scientist, California Department of Health Care Services
1	Eva	Huala	Curator	9/99	8/05	Executive director, Phoenix Informatics

Research Funding: (Cumulative \$ 43,481,302)

Current Research Grants:

Title: High-throughput determination of a subcellular metabolic network map of plants

Funding Organization: DOE BER

Dates of Project: 09/15/2019 -- 09/14/2022

Award Amount: \$2,311,477

PI: Sue Rhee

Title: Thermo-adaptation of photosynthesis in extremophilic desert plants

Funding Organization: Carnegie

Dates of Project: 06/01/2019 -- 5/31/2022

Award Amount: \$150,000

PI: Sue Rhee

Title: 1st Workshop on the Plant Cell Atlas Initiative

Funding Organization: NSF

Dates of Project: 06/01/2019 -- 05/31/2020

Award Amount: \$82,861

PI: Sue Rhee

Title: An Integrated Pipeline for Accelerated Plant Natural Product Discovery (10643)

Funding Organization: Stanford (NIH): 3U01GM110699-02S1. Subaward #: 61066255-112225

Dates of Project: 08/15/15 – 04/30/20

Award Amount: \$924,810

PI: Christina Smolke

Title: TOOLS-PGR: Computational Infrastructure to Enable High-throughput, High-quality Annotations of Compartmentalized Metabolic Networks for Plant Genomes

Funding Organization: NSF: IOS-1546838 (10708)

Dates of Project: 08/15/16 – 07/31/20

Award Amount: \$2,193,335

PI: Sue Rhee

Title: Prediction and Discovery of Host Metabolites and Metabolic Pathways Required for Proliferation of an Obligate Fungal Biotroph

Funding Organization: UC Berkeley (NSF): MCB-1617020. Subaward #: 00009338 (10722)

Dates of Project: 09/15/16 – 08/31/20

Award Amount: \$417,893

PI: Mary Wildermuth

Title: Using Systems Approaches to Improve Photosynthesis and Water Use Efficiency in Sorghum

Funding Organization: Danforth (DOE DE-SC0018277). Subaward #: 23021-C (10797)

Dates of Project: 09/15/17 – 09/14/22

Award Amount: \$1,866,443

PI: Ivan Baxter

Title: NeuroPlant: Leveraging a botanical armamentarium to manipulate the brain (10897)

Funding Organization: Stanford (NSF): # 62019381 - 140753

Dates of Project: 01/01/19 – 12/31/20

Award Amount: \$70,000

PI: Miriam Goodman

Title: Deciphering Life Functions in Extreme Environment

Funding Organization: Venture Grant / Broccoli Charitable Fund. (10877)

Dates of Project: 10/01/18 – 09/30/20

Award Amount: \$40,000

PI: Sue Rhee

Past Research Grants:

Title: A systems-level analysis of drought and density response in the model c4 grass *Setaria viridis*

Funding Organization: Danforth (DOE): 23009-CI (10412)

Dates of Project: 09/01/12 – 08/31/18

Award Amount: \$2,212,640

PI: Thomas Brutnell

Title: The membrane-based protein Interactome

Funding Organization: NSF: MCB-1052348. (10274)

Dates of Project: 04/15/11 – 03/31/16

Award Amount: \$1,834,556

PI: Wolf Frommer

Title: *Arabidopsis* 2010: Towards a Comprehensive *Arabidopsis* Protein Interactome Map: Systems Biology of the Membrane Proteins and Signalosome

Funding Organization: NSF: MCB-0618402 (2066)

Dates of Project: 09/01/06 – 08/31/2011

Award Amount: \$4,799,186

PI: Wolf Frommer

Title: Building a Network of Plant Metabolic Pathway Databases and Communities

Funding Organization: NSF: DBI-0640769 (2098)

Dates of Project: 03/15/08 – 02/28/2013

Award Amount: \$1,477,869

PI: Sue Rhee

Title: Metabolomics: A Functional Genomics Tool for Deciphering Functions of *Arabidopsis* Genes in the Context of Metabolic and Regulatory Networks

Funding Organization: Iowa State University (NSF): 420-40-71A (10079)

Dates of Project: 03/01/09 – 02/28/2013

Award Amount: \$269,862

PI: Basil Nikolau

Title: TAIR: The *Arabidopsis* Information Resource

Funding Organization: NSF: DBI-0850219 (10107)

Dates of Project: 09/01/09 – 08/31/2014

Award Amount: \$4,170,595

PI: Eva Huala

Title: TRPGR: Building a Highly Automated Metabolic Pathway Reconstruction Infrastructure for Plants

Funding Organization: NSF: IOS-1026003 (10204)

Dates of Project: 09/01/10 – 08/31/2016

Award Amount: \$1,825,569

PI: Sue Rhee

Title: The First International Biocurator Meeting

Funding Organization: Genetics Society of America (5133)

Dates of Project: 5/22/05 – 12/01/2008

Award Amount: \$25,000

PI: Sue Rhee

Title: Enhancing the Quality and Quantity of *Arabidopsis* Metabolism Data in AraCyc and MetaCyc

Funding Organization: Pioneer Hi-Breed International (5119)

Dates of Project: 02/18/05 – 08/06/2006

Award Amount: \$40,000

PI: Sue Rhee

Title: 2nd International Biocurator Meeting

Funding Organization: Villa Bosch (5134)

Dates of Project: 02/13/07 – 02/12/08

Award Amount: \$24,968

PI: Sue Rhee

Title: TAIR: The *Arabidopsis* Information Resource

Funding Organization: NSF: DBI-0417062 (2462)

Dates of Project: 09/01/04 – 08/31/2009

Award Amount: \$7,988,952

PI: Sue Rhee

Title: *Arabidopsis* 2010: Metabolomics: A Functional Genomics Tool for Deciphering Functions of *Arabidopsis* Genes in the Context of Metabolic and Regulatory Networks

Funding Organization: Iowa State Univ. (NSF): 420-40-17 (4306)

Dates of Project: 09/01/05 – 08/31/2008

Award Amount: \$64,875

PI: Basil Nikolau

Title: Low Temperature Regulatory Circuits and Gene Regulation in Higher Plants

Funding Organization: MSU (NSF): DBI-0110124 (4433)

Dates of Project: 09/01/01 – 08/31/2008

Award Amount: \$539,021

PI: Michael F. Thomashow

Title: First International Biocurator Meeting

Funding Organization: NSF: DBI-0551286 (2282)

Dates of Project: 01/01/06 – 12/31/2007

Award Amount: \$29,810

PI: Sue Rhee

Title: The Plant Ontology Consortium

Funding Organization: CSHL (NSF): 22130313 (4417)

Dates of Project: 09/01/03 – 08/31/2007

Award Amount: \$530,928

PI: Lincoln Stein

Title: The MetaCyc Metabolic Pathway Database

Funding Organization: SRI (NIH): 55-000650 (4441)

Dates of Project: 03/01/04 – 02/28/2007

Award Amount: \$839,512

PI: Peter Karp

Title: Gene Ontology Consortium

Funding Organization: The Jackson Laboratory (NIH): 2P41HG002273-04 (4442)

Dates of Project: 03/09/04 – 02/28/2007

Award Amount: \$634,431

PI: Judith Blake

Title: First International Biocurator Meeting

Funding Organization: NIH: 1R13HG004030-01 (1403)

Dates of Project: 12/07/05 – 11/30/2006

Award Amount: \$12,000

PI: Sue Rhee

Title: The Arabidopsis Biological Resource Center at The Ohio State University

Funding Organization: OSURF (NSF): DBI-0091471 (4428)

Dates of Project: 02/01/01 – 02/28/2006

Award Amount: \$402,983

PI: Randy Scholl

Title: AtIR: An *Arabidopsis* Thaliana Information Resource

Funding Organization: NSF: DBI-9978564 (2665)

Dates of Project: 10/01/99 – 09/30/2005

Award Amount: \$5,728,633

PI: Chris Somerville

Title: A Literature Curation Tool for Organism Databases

Funding Organization: NIH: 1R01HG02728-01 (1092)

Dates of Project: 09/30/02 – 08/31/2005

Award Amount: \$896,386

PI: Sue Rhee

Title: Large-scale Fluorescent Tagging of Full-length Genes to Characterize Native Expression Patterns and Subcellular Targeting of *Arabidopsis* Proteins of Unknown Function

Funding Organization: SUNY at Stony Brook (NSF): 1027553 (4413)

Dates of Project: 09/01/02 – 12/31/2004

Award Amount: \$168,497

PI: Vitaly Citovsky

Title: Partnership for Research & Education in Plants

Funding Organization: Virginia Tech (NIH): CR-19501-431399 (4402)

Dates of Project: 09/30/03 – 08/31/2004

Award Amount: \$27,522

PI: Erin L. Dolan

Title: Gene Ontology Consortium

Funding Organization: The Jackson Laboratory (NIH) (4435)

Dates of Project: 01/01/02 – 03/08/2004

Award Amount: \$536,045

PI: Judith Blake

Title: Development of the MetaCyc Metabolic Pathway Database

Funding Organization: SRI (NIH): P11833 (4437)

Dates of Project: 03/01/02 – 02/28/2004

Award Amount: \$238,999

PI: Peter Karp

Title: Sequencing of *Arabidopsis* Chromosome II and Beyond, and Development of Resources for *Arabidopsis* Genome Analysis

Funding Organization: TIGR (NSF): TIGR-00-007 (4432)

Dates of Project: 05/01/01 – 12/31/2001

Award Amount: \$105,824

PI: Claire Fraser