



Gramene: A Platform for Comparative Plant Genomics

DBI-9723510

¹Edward Buckler Co-PI, ²Pankaj Jaiswal Co-PI, ¹Susan McCouch Co-PI, ³Doreen Ware PI, ³Ken Youens-Clark Project Manager
¹Cornell University, ²Oregon State University, ³Cold Spring Harbor Laboratory

Introduction

The Gramene project is developing resources for knowledge discovery by integrating genomic and phenotypic resources. The three research aims of the project are to: (1) Use comparative genomics to identify functional elements and sequence variants that may have phenotypic consequences; (2) Annotate biological pathways in order to provide the infrastructure to understand how those sequence variants lead to phenotypes; (3) Collect and uniformly reanalyze QTL and diversity data in order to connect genetic diversity to phenotypic variation.

Specific Aim #1: Provide an infrastructure of comparative genomic data to allow for the mining and analysis of functional data on the genomes of rice and other plants.

In the last year this objective has focused on data acquisition, analysis and software development and updates. In the last year we have added three new genomes (Sorghum, Arabidopsis Lyrata and Brachypodium) as well as three short arms of rice from the OMAP project and updated the annotations for two genomes (Arabidopsis and rice). In addition to the genome sequences that have become available for Arabidopsis and grape in the last year, we have added diversity data that will be available in the Gramene release 30 (September 2009). We have updated to Ensembl version 55 which includes improved performance and visualizations, a new layout for the browser (Figure 1), improvements to the tree browsers (Figure 2), and new views for variation (Figure 3). We have also updated the CMap module to now include over 200 genetic, QTL, sequence, and physical maps from 29 species. We have extended our existing collaboration with EBI to formally collaborate on the development of Plant Ensembl (<http://www.ensemblgenomes.org>), a subset of the genomes that are being hosted as part of the larger Ensembl Genome project.

Specific Aim #2: Enhance the value of the comparative maps with pathway, phenotypic and other functional information from rice, maize, and Arabidopsis.

Gramene uses the Pathway Tools package from SRI to display information from ten species. In Gramene's 29 and 30 builds, the focus has been on updating our rice pathways with the newest versions of the TIGR version 5 updates and MSU version 6 as well as the curation of more than 10 rice pathways including Chorismate/Shikimate biosynthesis, ent-kaurene biosynthesis, Flavonoid and flavonol biosynthesis, Geranylgeranyldiphosphate biosynthesis, Iron uptake pathways, Leucine biosynthesis, Phenylpropanoid biosynthesis, Phytoalexin and diterpenoid biosynthesis (Momiactone [Figure 4], Oryzalexin S, Phytocassane, and Oryzalexin A-F), Starch biosynthesis and Tryptophan biosynthesis. To improve usability of the "Omics viewer" (Figure 5) a microarray probe ID validator was developed and implemented in the last year. In release 30, the Plant Gene Wiki is being made available to facilitate community annotation of Gramene's genes.

Specific Aim #3: We will acquire genotypic and phenotypic diversity data for each of the sequenced monocot genomes. We will recalculate this data using a standardized methodology that allows us to integrate the QTL values across species and to relate phenotypic diversity to candidate genes via pathway information.

In the last year, efforts have been centered on data acquisition, software development and improved usability. Release 30 will see the addition of Arabidopsis data from 931 ecotypes against 216,509 SNPs. To enhance usability of the website, new data summary pages have been generated (Figure 6) as well as a new section aimed at plant breeders that highlights our germplasm resources (Figure 7).

Specific Aim #4 Support the Plant Ontology (PO) Two Years Only
Gramene's specific aim #4 has been moved to [NSF award Award Number: 0822201](http://www.nsf.gov/awardsearch/showAward.do?awardNumber=0822201).

Specific Aim #5: Education, Outreach and Diversity

The experimental plan consists of 3 specific components 1) a community-supported wiki targeted to pure and translational plant sciences researchers; an example of this is the Plant Gene Wiki (Figure 8) and Pathways; 2) traditional "push" presentations at meetings, staffed by the co-PIs and senior curators; in the last year project members attended six meetings; 3) virtual and physical tutorials sponsored by a commercial partner, OpenHelix, targeted at students in under-served institutions and underrepresented minorities who will be starting in the fall of 2009.

Publications

Youens-Clark, K., Faga, B., Yap, I., Stein, L., Ware, D. 2009. CMap 1.01: A comparative mapping application for the Internet. Bioinformatics, doi: 10.1093/database/bap005.

Ni, J., Pujar, A., Youens-Clark, K., Yap, I., Jaiswal, J., Teale, I., Tung, C.W., Ren, L., Spooner, W., Wei, X., Avraham, S., Ware, D., Stein L., and McCouch, S. 2009. Gramene QTL database: development, content and applications. Database. doi: 10.1093/database/bap005.

Liang, C., Mao, L., Ware, D. and Stein, L. 2009. Evidence-based gene predictions in plant genomes. Genome Research. doi:10.1101/gr.088997.108.

Figure 1

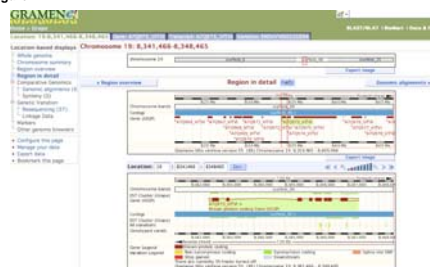


Figure 2

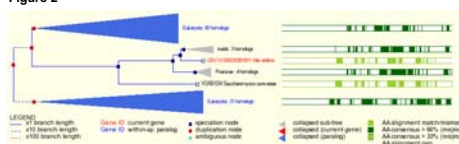


Figure 3

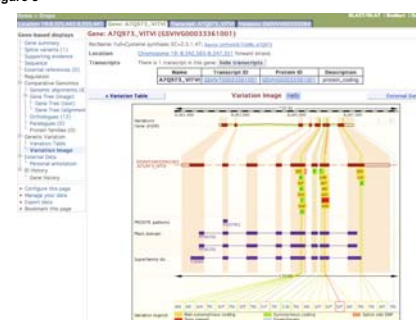


Figure 4



Figure 5

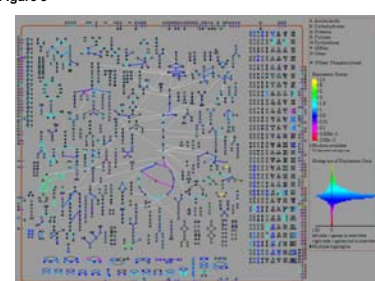


Figure 6



Figure 7



Figure 8

