



VirtualPlant: A software platform to support Next-Generation Systems Biology research

Manpreet S. Katari¹, Rodrigo A. Gutierrez^{1,3}, Gabriel Krouk¹, Rebecca Davidson^{1,2}, Tamara Tereshakovec², Dennis E. Shasha², Gloria M. Coruzzi¹,

(1)Department of Biology, New York University (2) Courant Institute of Mathematical Sciences, New York University. (3) Department of Molecular Genetics & Microbiology, P. Universidad Católica de Chile.

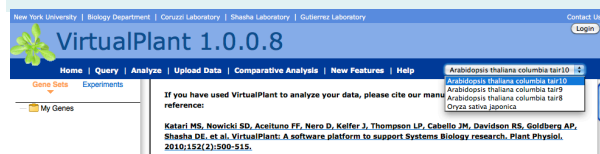
<http://www.virtualplant.org>

VirtualPlant enables biologists to mine lists of genes, microarray experiments, next-generation sequencing data and gene networks to address biology's grand challenge questions. VirtualPlant achieves this by enabling researchers to integrate, analyze, and visualize genomic data in a systems biology context. VirtualPlant simplifies data analysis by integrating the tools into a single platform. The unique "gene cart" functionality enables iterative data analysis and suggest additional rounds of experimentation.

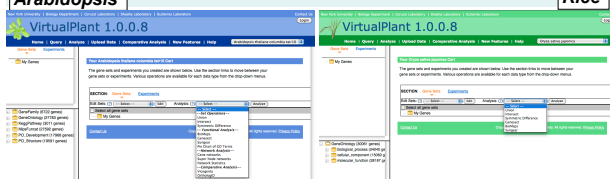
Please see the following manuscript for details and example case studies: Katari et al. "VirtualPlant: a software platform to support systems biology research." *Plant Physiol.* 2010 Feb;152(2):500-15.

Funded by: NSF DBI (0445666) Software development NSF 2010 (IOB-0519985) Data generation

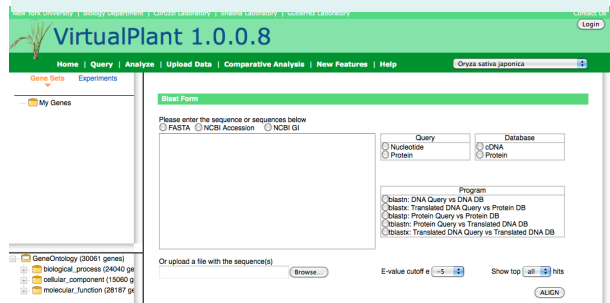
Tair 10 and other versions: The Arabidopsis annotation has been updated to Tair 10. A feature in the gene cart allows users to convert their gene lists from an older version of Tair to the newer version



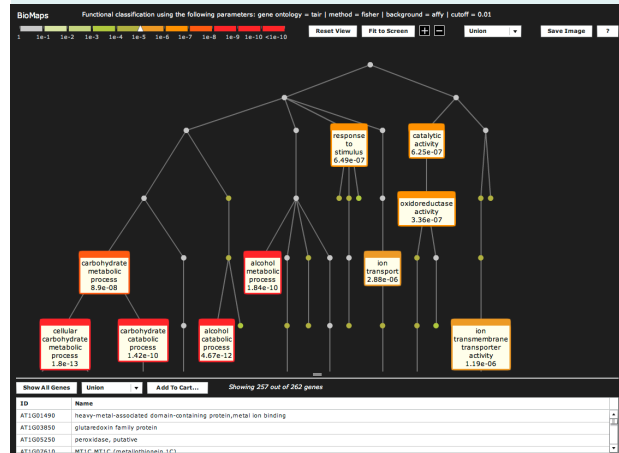
Multi-Species: VirtualPlant can support any number of species or ecotypes. Rice is already available and we are in the process of supporting more sequenced plant genomes.



Comparative Analysis: Currently users can align their sequencing using Blast to identify putative orthologs based on top match. Other methods and databases will be supported in the near future.



Biomaps: The gene set enrichment tool, now has an interactive graph written in Flash with the ability to save the GO-hierarchy image as a high quality figure.



GeneSect: A new tool that uses non-parametric randomization test to determine whether overlap between two gene sets are significant.

GeneSect: P-value (size of intersection)

	A: ATGE7 (288)	B: ATGE8 (289)	C: ATGE8 (178)	D: ATGE8 (99)
A: ATGE7 (288)		<0.001 (82)	<0.001 (32)	0.124 (3)
B: ATGE8 (289)	1 (62)		<0.001 (11)	<0.001 (2)
C: ATGE8 (178)	1 (32)	1 (11)		0.17 (2)
D: ATGE8 (99)	0.984 (3)	1 (28)	0.965 (2)	

Overrepresented
Above the diagonal and in yellow: P-value <0.05 - the size of the intersection is higher than expected

Underrepresented
Below the diagonal and in blue: P-value <0.05 - the size of the intersection is lower than expected

GeneSect: Z-score (number of standard deviations)

	A: ATGE7 (288)	B: ATGE8 (289)	C: ATGE8 (178)	D: ATGE8 (99)
A: ATGE7 (288)		36.1	36.0	1.6
B: ATGE8 (289)			36.0	36.0
C: ATGE8 (178)				1.6
D: ATGE8 (99)				

Overrepresented
Green: Z-score >6

Underrepresented
Red: Z-score <-5

ECOTYPES

COMING SOON

METAGENOMICS



RNA-seq: A new tool is being developed where users can submit the read counts for each gene in the genome and then use tools such as edgeR to determine differentially expressed genes

VirtualPlant Development Team attending the meeting:
Email: virtualplant-discuss@googlegroups.com

