

Introduction IGB for Arabidopsis: a resource for Arabidopsis genome data visualization

Ann Loraine (1), Gregg Helt (2), Ed Erwin (2), Sue Rhee (3), Eva Huala (3)

(1) Section on Statistical Genetics, University of Alabama at Birmingham

(2) Bioinformatics Department, Affymetrix, Inc, Emeryville, CA

(3) The Arabidopsis Information Resource, the Carnegie Institute, Stanford, CA

The Integrated Genome Browser (IGB, pronounced 'Ig-Bee') is an experimental, Java-based, desktop software program that allows rapid, interactive visualization and exploration of genes, sequence, microarray and other types of data from multiple genomes. Until now, IGB has only been able to display data from animal genomes, simply because Arabidopsis and other plant data have not been available in compatible formats. To remedy this, we created a prototype Arabidopsis genome 'Quickload' server, which is an on-line data repository containing annotations and sequence data that users can load and view in IGB. Currently, the server delivers foundation gene annotations from TAIR, probe-to-genome alignments from Affymetrix ATH1 and AG GeneChip platforms, and reference sequence data from version 5.0 of the Arabidopsis genome.

IGB is designed to allow and encourage loading annotations for an entire genome at once. To make this possible, IGB uses optimized data models that reduce memory requirements, binary file formats that speed up loading time, and a local data caching system. IGB also implements techniques for exploratory data analysis, such as the ability to display and manipulate numerical data alongside gene structures. This feature is especially useful for viewing probe readings from DNA expression microarrays but has other applications, as well. In addition, IGB can load data from user-generated files so long as they conform to one of several compatible formats. We are working to develop tools and documentation that will make the process of file format conversion easier and more transparent.

Information about IGB, including links to download pages and instructions on how to use the program, is available at www.transvar.org/at_annots. To access Arabidopsis data after launching IGB, select the 'Arabidopsis' option under the 'Quickload' tab. To load individual data sets, click the checkboxes to the right. The 'TAIR' option loads gene models, and the 'ATH1' option loads probe sets from the Affymetrix ATH1 chip. To load sequence for an entire chromosome, click the 'Load Entire Sequence' button. To view Web pages describing individual gene models, right-click (Windows, Unix) or control-click (Mac) the gene and select the 'Get More Info' option.

IGB is part of the Genoviz open source software development project [sourceforge.net/projects/genoviz], a project dedicated to creating and distributing visualization software for biology and bioinformatics.