

Customized Gbrowse - visualization that assists research in Genome Projects

Jerzy W. Jaromczyk¹, Ch. L. Schardl², E. Arnaudova¹, U. Hesse², N. Moore¹, M. Ryavec¹, S. Vessels¹.
Departments of Computer Science¹ and Plant Pathology², University of Kentucky, Lexington, Kentucky

Background and Objective

Gbrowse is a popular web-based application for visualizing genomic sequences, used to assist analysis and interpretations of discoveries. Implemented in Perl, HTML, and Javascript, Gbrowse provides researchers with a convenient tool, albeit a difficult one to fully customize for specialized projects. This project addresses three aspects of Gbrowse: (1) providing an annotation system supporting GenBank formats, (2) providing a convenient, modern, and easily-customizable Human-Computer Interface, and (3) improving the performance of time-consuming data imports, while maintaining query speed.

Methods

Annotations are implemented by new data backends and abstract data-access libraries, with an AJAX interface for a fast and fluid user experience. Customization of browser views is achieved by an in-house templating system, separating presentation from code. Efficiency is maintained by using a pair of relational databases: one storing complete genome and analysis information, and another storing a subset, optimized for easy visualization, with links to the complete data. The use of indexed relations allows the databases to be updated and queried efficiently.

Results

The new annotation system ties the Gbrowse data to Genbank annotations. The new HCI makes Gbrowse more intuitive to use for biologists and bioinformaticists, and much simpler to customize. Database optimization techniques allow the database to scale to gigabytes of data without negatively impacting query and update times.

We acknowledge funding from the KBRIN-INBRE NIH 2 P20 RR-16481, the National Center for Research Resources Institutional Development Award (IDeA) Program (sub-award from the University of Louisville)