



GMOD Tools for Comparative Genomics

Dave Clements¹, David Riley², Ben Faga³, Hervé Tettelin², Sam Angiuoli², and Jonathan Crabtree⁴

¹National Evolutionary Synthesis Center, Durham, NC ²Institute for Genome Sciences, Baltimore MD

³Cold Spring Harbor Laboratory, Cold Spring Harbor, NY ⁴The J. Craig Venter Institute, Rockville MD

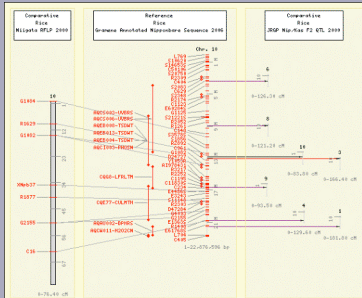


Abstract

GMOD is a collection of interoperable open source software components for managing, annotating and visualizing genomic data. GMOD includes several components for managing and visualizing comparative genomics data. GMOD is used in many smaller research and emerging model organism communities, where informatics budgets are often tight. GMOD is also a community of developers and users that support and enhance it.

CMap: Comparative Map Viewer

The web-based CMap enables users to view comparisons of many types of genomic data. CMap is data type agnostic and shows disparate data types together such as sequence, genetic, physical, QTL and deletion maps.



A rice sequence assembly compared against a genetic map (left side) and several QTL maps (right side). In this case, the genetic map and the sequence map appear to agree. The QTL maps have their correspondences aggregated into one line to reduce clutter but can be expanded if desired.

CMap displays correspondences between features such as markers, HSPs or any other annotation. CMap comes with tools for creating these correspondences based on feature names, or correspondences can be imported directly.

CMap can display correspondences as lines or ribbons. This figure compares optical map data against an in silico digest of assembled maize sequence. Correspondences linking same sized fragments were imported into CMap. This allows for easy determination of where differences between the sequence data and the optical data occur.

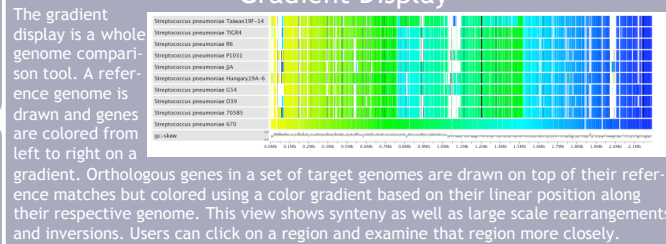
See <http://gmod.org/CMap> for more.

Figures from the Gramene project: <http://www.gramene.org>
Optical map data: Shiguo Zhou, Michael Bechner, Steve Goldstein, Dan Forrest, David C. Schwartz.
CMap is work by Ben Faga and Ken Youens-Clark.

Sybil: Comparative Genomics Visualization

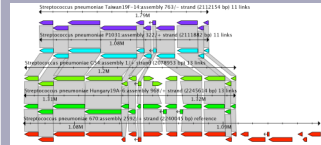
Sybil is a web based tool for visualizing and mining comparative genomic data that relies on a Chado database containing clusters of orthologous genes.

Gradient Display



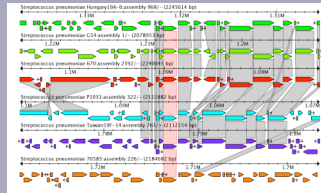
The gradient display is a whole genome comparison tool. A reference genome is drawn and genes are colored from left to right on a gradient. Orthologous genes in a set of target genomes are drawn on top of their reference matches but colored using a color gradient based on their linear position along their respective genome. This view shows synteny as well as large scale rearrangements and inversions. Users can click on a region and examine that region more closely.

Region Comparison



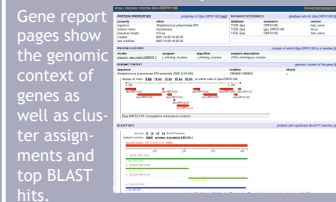
The genomic region comparison display searches for regions that match a reference. Matching regions are drawn above the reference and matching genes are connected via polygons. Users can click on genes and gene links to retrieve additional information and navigate to other views.

Cluster Report



The cluster report page shows the genomic context of a set of orthologous genes. Cluster members are linked with red polygon and are centered in the display. Orthologous genes in flanking regions are connected via gray polygons. Like the region comparison view, users can click on genes for more information and to navigate to another view.

Gene Reports



Gene report pages show the genomic context of genes as well as cluster assignments and top BLAST hits.

Sybil is open source and is freely available with documentation and demo databases at <http://sybil.sourceforge.net>

Sybil is implemented in Perl using a tiered architecture. It has an API for retrieving data from Chado, and utilities for rendering publication quality images in SVG and PDF formats.

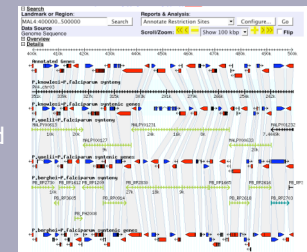
Sybil: Methods and Software for Multiple Genome Comparison and Visualization. Jonathan Crabtree, Samuel V. Angiuoli, Jennifer R. Wortman and Owen R. White; in Gene Function Analysis, edited by Michael F. Ochs (2007)

GBrowse Based Synteny Viewers

GBrowse is a web based viewer for displaying genomes and their annotation. Three extensions to GBrowse for viewing synteny data are available in GMOD.

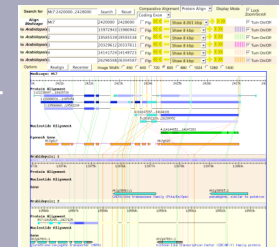
SynView

SynView displays synteny at the region and/or gene level. Users select a reference genome and then synteny with other selected genomes is displayed relative to that genome. SynView can be layered on top of an existing GBrowse instance and uses the full range of GBrowse's display and configuration options.



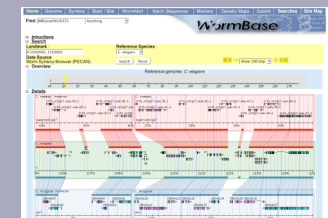
SynBrowse

SynBrowse displays synteny based on nucleotide or protein alignment. It uses color intensity to indicate degree of similarity. SynBrowse is built on several open source packages, including parts of GBrowse, resulting in a similar look and feel.



GBrowse_syn

GBrowse_syn is a synteny specific version of GBrowse. It displays synteny between regions, using grid lines to show where syntenic regions have grown or shrunk across genomes.



SynView: a GBrowse-compatible approach to visualizing comparative genome data. Haiming Wang, Yanqi Su, Aaron J. Mackey, Eileen T. Kraemer, and Jessica C. Kissinger; in Bioinformatics 22 (18)
SynBrowse: a Synteny Browser for Comparative Sequence Analysis. Pan, X., Stein, L. and Brendel, V.; in Bioinformatics 21 (17)
GBrowse_syn work is by Sheldon McKay.

Poster authors are funded by the National Institutes of Health, PATH, Washington University in St. Louis, the National Science Foundation, and the University of Arizona.