

WebGBrowse

A Web Server for GBrowse Configuration

Ram Podicheti B.V.Sc. & A.H. (D.V.M.), M.S.

Staff Scientist – Bioinformatics

Center for Genomics and Bioinformatics

Indiana University



01/16/2009

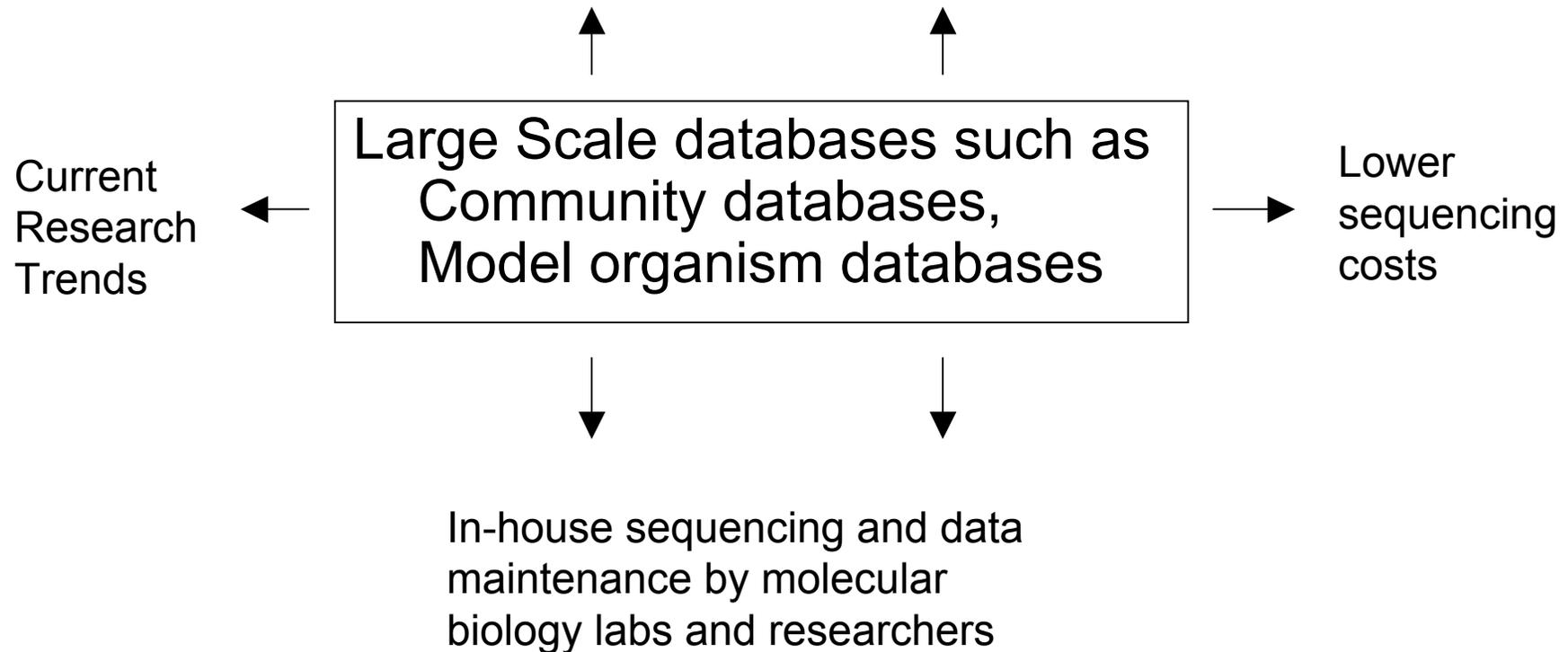
GMOD Conference 2009

San Diego CA

Generic Genome Browser

- Most popular web based genome browser
- Visualize genome features along a reference sequence
- Open Source
- Highly customizable
- Excellent usability
- Rich set of “glyphs”
 - Genome features
 - Quantitative Data
 - Sequence Alignments

Scope of GBrowse Usage



Scope of GBrowse Usage

Large Scale databases such as
Community databases,
Model organism databases

More specific, but smaller databases such as a lab
owned database or Individual Researcher's database

GBrowse Setup

- Software installation and maintenance
- GFF3 dataset preparation
- Writing the configuration file

GBrowse Setup

- Software installation and maintenance
- GFF3 dataset preparation
- Writing the configuration file



Goal

Make GBrowse Available to Biologists
without

- installation hassles
- worries about GBrowse configuration semantics

WebGBrowse

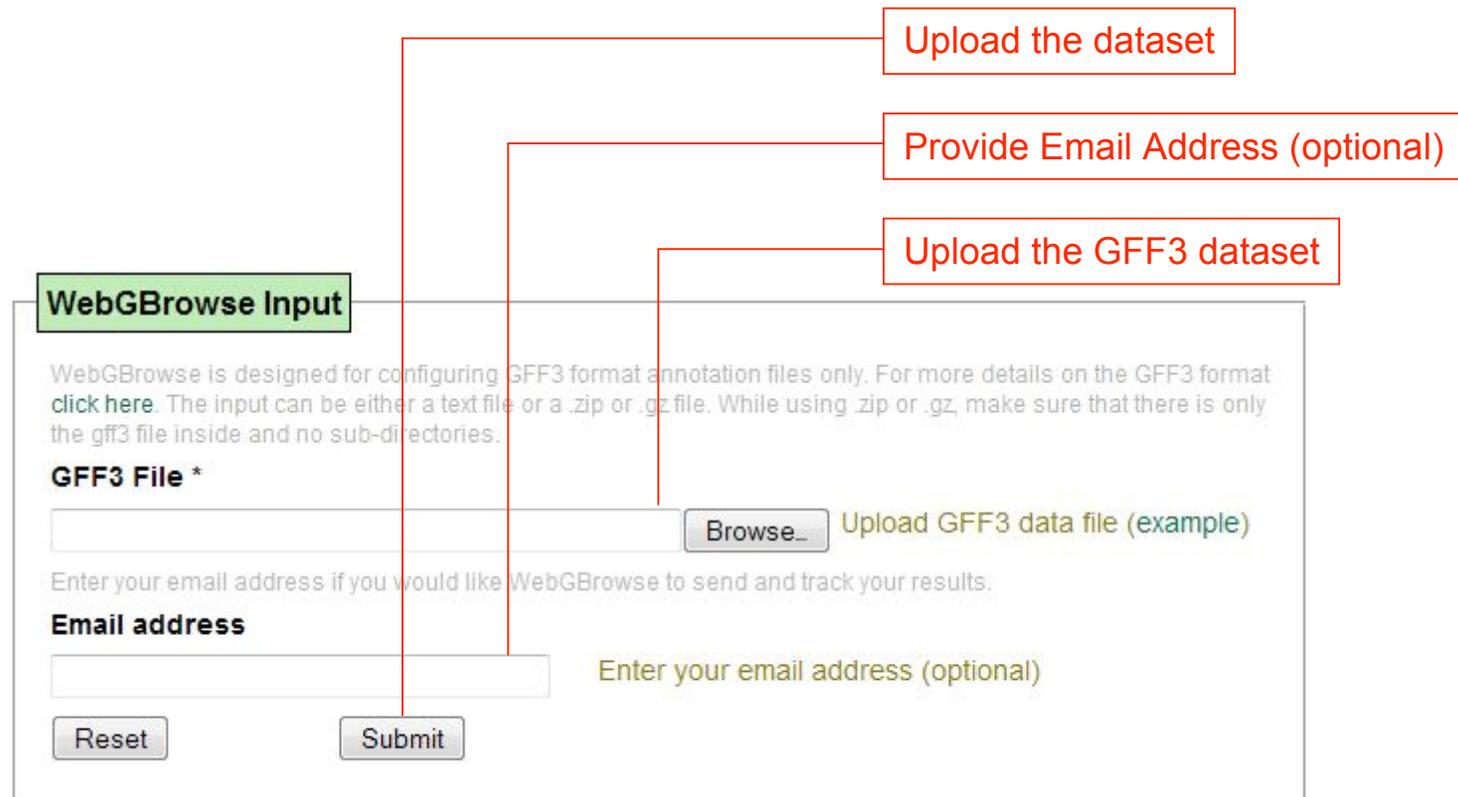
WebGBrowse

- Allows users to upload their GFF3 datasets
- Powered by a Glyph Library
- Configuration information for 40+ glyphs
- Assists in Configuring the display of each genomic feature into individual tracks
- Hosts the datasets with the specified configuration settings on an integrated GBrowse server

WebGBrowse

<http://webgbrowse.cgb.indiana.edu/>

WebGBrowse Input



The image shows a web form titled "WebGBrowse Input" with three red callout boxes pointing to specific fields:

- Upload the dataset**: Points to the "GFF3 File *" input field.
- Provide Email Address (optional)**: Points to the "Email address" input field.
- Upload the GFF3 dataset**: Points to the "Browse..." button next to the "GFF3 File *" input field.

The form itself contains the following elements:

- WebGBrowse Input** (Title)
- Introductory text: "WebGBrowse is designed for configuring GFF3 format annotation files only. For more details on the GFF3 format [click here](#). The input can be either a text file or a .zip or .gz file. While using .zip or .gz, make sure that there is only the gff3 file inside and no sub-directories."
- GFF3 File *** (Label)
- Input field for the GFF3 file.
- Browse...** (Button) and [Upload GFF3 data file \(example\)](#) (Link)
- Text: "Enter your email address if you would like WebGBrowse to send and track your results."
- Email address** (Label)
- Input field for the email address.
- [Enter your email address \(optional\)](#) (Text)
- Reset** (Button) and **Submit** (Button)

Configuration Panel

Add New Track

Please select a feature from the list on the left and a glyph from the drop down menu on the right, then click the button **Add Track** to add a new track for display. Upon selection, a brief description and a sample image for the corresponding glyph will be shown.

Feature *

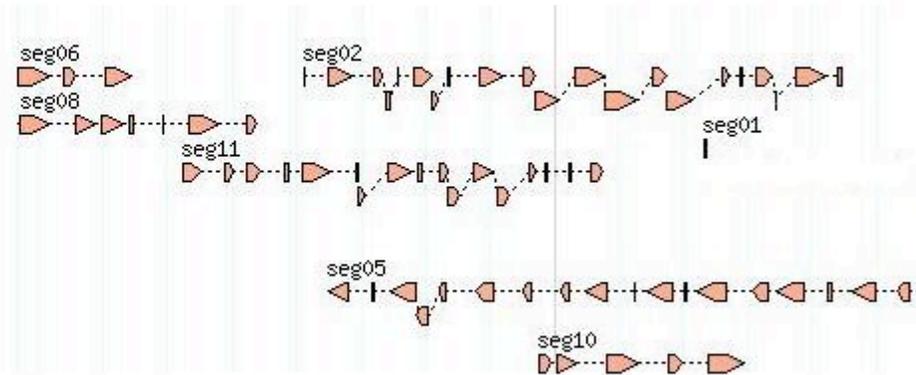
BAC
CDS
CDS:example
CDS:predicted
CDS:exonerate
EST_match
UTR
clone_end
clone_start
contig

Glyph (What is a glyph?)

generic

Generic is identical to the "box" glyph except that it will draw the subparts of features that

Add Track



Unique Feature set Identified from the uploaded dataset

Configuration Panel

```
ctgA example remark 1659 1984 . + . Name=f07;Note=This is an e:
ctgA example remark 3014 6130 . + . Name=f06;Note=This is anot!
ctgA example remark 4715 5968 . - . Name=f05;Note=Ok! Ok! I ge!
ctgA example remark 13280 16394 . + . Name=f08
ctgA example remark 15329 15533 . + . Name=f10
ctgA example remark 19157 22915 . - . Name=f13
ctgA example remark 22832 24633 . + . Name=f15
ctgA example remark 23072 23185 . + . Name=f14
ctgA example remark 24562 28338 . - . Name=f02
ctgA example remark 36034 38167 . + . Name=f09
ctgA example remark 36649 40440 . - . Name=f03
ctgA example remark 37242 38653 . + . Name=f04
ctgA example remark 44705 47713 . - . Name=f01
ctgA example remark 46990 48410 . - . Name=f11
ctgA example remark 49758 50000 . - . Name=f12
```

Configuration Panel

ctgA	example	mRNA	1300	9000	.	+	.	ID=EDEN.3;Parent=EDEN;Name=E
ctgA	example	five_prime_UTR	1300	1500	.	+	.	Parent=EDEN.3
ctgA	example	five_prime_UTR	3000	3300	.	+	.	Parent=EDEN.3
ctgA	example	CDS	3301	3902	.	+	0	Parent=EDEN.3
ctgA	example	CDS	5000	5500	.	+	1	Parent=EDEN.3
ctgA	example	CDS	7000	7600	.	+	1	Parent=EDEN.3
ctgA	example	three_prime_UTR	7601	9000	.	+	.	Parent=EDEN.3
ctgA	predicted	CDS	10000	11500	.	+	0	Name=Apple1;Note=A for Apple
ctgA	predicted	CDS	13000	13800	.	+	0	ID=cds-Apple2;Name=Apple2;No
ctgA	predicted	CDS	15000	15500	.	+	1	ID=cds-Apple2;Name=Apple2
ctgA	predicted	CDS	17000	17200	.	+	2	ID=cds-Apple2;Name=Apple2
ctgA	exonerate	mRNA	17400	23000	.	+	.	ID=rna-Apple3;Name=Apple3;No
ctgA	exonerate	UTR	17400	17999	.	+	.	Parent=rna-Apple3
ctgA	exonerate	CDS	18000	18800	.	+	0	Parent=rna-Apple3
ctgA	exonerate	CDS	19000	19500	.	+	1	Parent=rna-Apple3

Configuration Panel

Add New Track

Please select a feature from the list on the left and a glyph from the drop down menu on the right, then click the button **Add Track** to add a new track for display. Upon selection, a brief description and a sample image for the corresponding glyph will be shown.

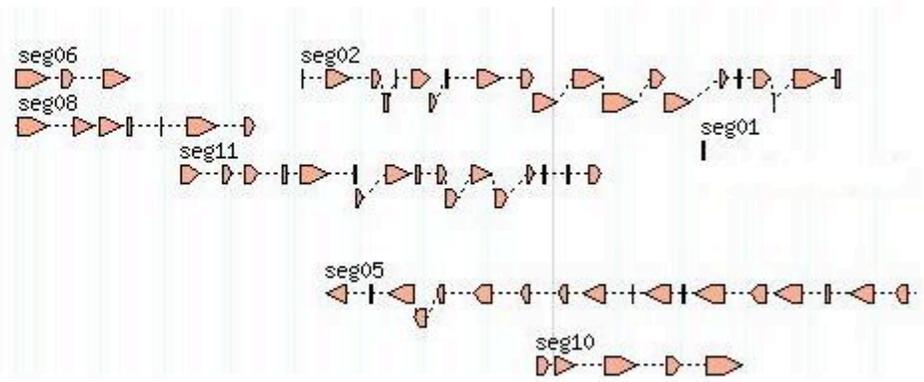
Feature *

- BAC
- CDS
- CDS:example
- CDS:predicted
- CDS:exonerate
- EST_match
- UTR
- clone_end
- clone_start
- contig

Glyph (What is a glyph?)

generic

Generic is identical to the "box" glyph except that it will draw the subparts of features that



Add Track

Unique Feature set Identified from the uploaded dataset

Configuration Panel

List of glyphs supplied by the glyph library

Add New Track

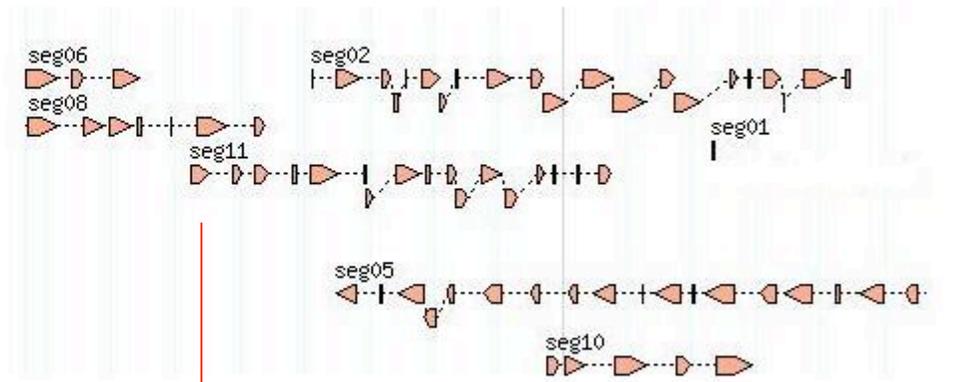
Please select a feature from the list on the left and a glyph from the drop down menu on the right, then click the button **Add Track** to add a new track for display. Upon selection, a brief description and a sample image for the corresponding glyph will be shown.

- Feature *
- BAC
 - CDS
 - CDS:example
 - CDS:predicted
 - CDS:exonerate
 - EST_match
 - UTR
 - clone_end
 - clone_start
 - contig

Glyph (What is a glyph?)

generic

Generic is identical to the "box" glyph except that it will draw the subparts of features that



Add Track

A Sample Image of the selected glyph

Unique Feature set Identified

Brief description of the selected glyph

Add Track Button

Glyph Parameters Form

Cancel Changes Save and Continue

Generic Glyph Parameters

*** Key**
The key controls the descriptive key that is drawn in the key area at the bottom of the image. It also appears in the checkboxes that the end user uses to switch tracks on and off. It defaults to the track name.

The name to identify the track

*** Glyph Background Color**
Glyph Background Color can be specified by name (e.g. yellow), or in HTML #RRGGBB format.

default background color of the glyph

*** Glyph Outline Color**
Glyph Outline Color can be specified by name (e.g. yellow), or in HTML #RRGGBB format.

default outline color of the glyph

*** Glyph Height**
It is the height of the glyph expressed in pixels.

Height of the glyph

Stranded
This option if checked, will indicate the strandedness of the feature, usually by drawing an arrow.

Show Strandedness
Show/Ignore strandedness

Show Label

Parameter
Description

Configured Tracks shown in the configuration panel

Tracks Added

Select a track from the list on the left and you will see the configuration settings for that feature in the box on the right. Click on **Edit Track** to edit the configuration settings of the selected track. In case you want to delete the selected track, click **Delete Track**. Once you are done adding tracks, you can proceed to the actual GBrowse display by clicking on **Display in GBrowse**. You will be allowed to come back and make further changes if desired.

Configured Tracks *

- remark

Configuration Settings

Track Configuration for "remark"

```
Glyph: generic
key: Track1001
bgcolor: lightsalmon
fgcolor: black
height: 10
stranded: 1
label: 1
description: 1
citation: Track with Generic Glyph
```

Edit Track **Delete Track** **Display in GBrowse**

The diagram shows a configuration panel with three main sections: a list of tracks, a configuration settings box, and three buttons. Red lines connect callout boxes to these elements:

- Edit Track Button**: Points to the 'Edit Track' button.
- Delete Track Button**: Points to the 'Delete Track' button.
- List of tracks added so far**: Points to the 'Configured Tracks' list.
- Configuration for the selected track in GBrowse**: Points to the 'Configuration Settings' box.
- Button to display the selected track in GBrowse**: Points to the 'Display in GBrowse' button.

GBrowse Display with WebGBrowse Control Panel

WebGBrowse Control Panel

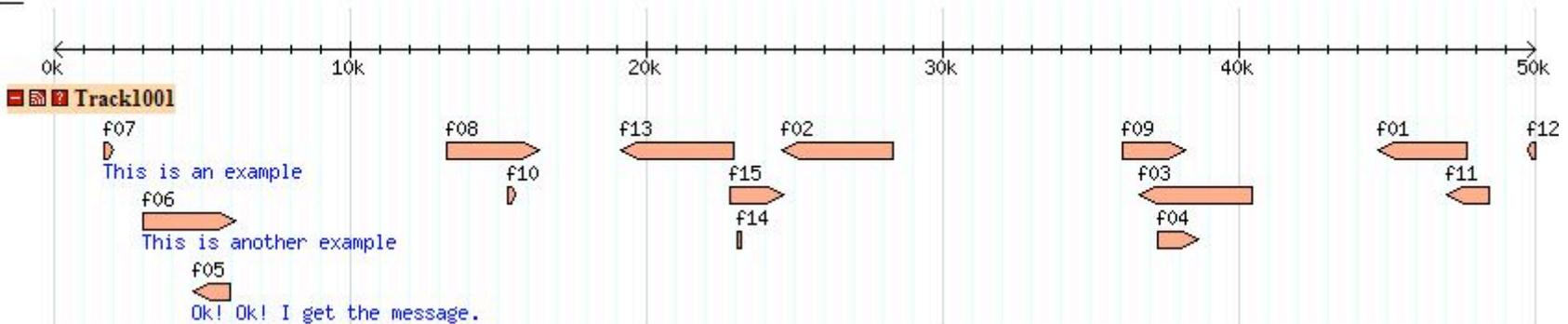
WebGBrowse Control Panel

Edit Configuration	Click Edit Configuration to return to the track configuration panel
Download Configuration	Click Download Configuration to download the configuration file

Showing 50 kbp from ctgA, positions 1 to 50,000

[Bookmark this] [Upload your own data] [Hide banner] [Share these tracks] [Link to Image] [Help] [Reset]

- [Search](#)
- [Overview](#)
- [Details](#)



Architecture

- Data Driven
- Glyph Library
- Configuration information
 - Initialize a data structure compatible with HTML::FormEngine
 - Load into the data structure
 - Serialize into a YAML file (<http://www.yaml.org/>)

WebGBrowse Demo

<http://webgbrowse.cgb.indiana.edu/>

Important resources on the website

- [Glyph Library](#)
- [Tutorial](#)
- [Software](#)
- [FAQ](#)

To Do List

- Expand the glyph library
- Allow uploading of a pre-existing conf file and start from there
- Provide "General Section" configuration (optional)
- Add more features (Balloons, plugins etc.)
- Categorizing the glyphs
- Tutorial on how to add new glyphs
- Callbacks?
- Suggestions from GMOD group

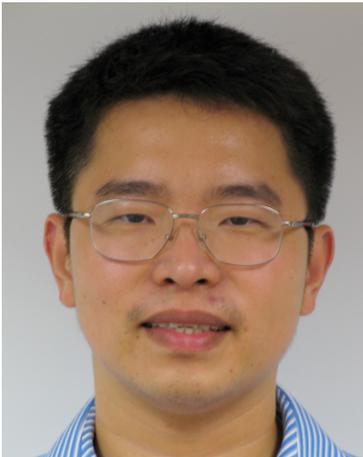
References

- Karolchik, D. *et al.* (2003) The UCSC Genome Browser Database, *Nucleic Acids Res*, 31, 51-54.
- Schlueter, S.D. *et al.* (2006) xGDB: open-source computational infrastructure for the integrated evaluation and analysis of genome features, *Genome Biol*, 7, R111.
- Stalker, J. *et al.* (2004) The Ensembl Web site: mechanics of a genome browser, *Genome Res*, 14, 951-955.
- Stein, L.D. *et al.* (2002) The generic genome browser: a building block for a model organism system database, *Genome Res*, 12, 1599-1610.

Acknowledgements



Rajesh Gollapudi
Graduate Student
School of Informatics
Indiana University



Dr. Qunfeng Dong
Director Bioinformatics
Center for Genomics and Bioinformatics
Indiana University

Acknowledgements



Chris Hemmerich

Staff Scientist & Database Unit Leader
Center for Genomics and Bioinformatics
Indiana University

Acknowledgements



Acknowledgements

This research was supported in part by the Indiana METACyt Initiative of Indiana University, funded in part through a major grant from the Lilly Endowment, Inc.

