

JBrowse – 1.3.0 and beyond

Robert Buels
GMOD Community Meeting
April 5, 2012



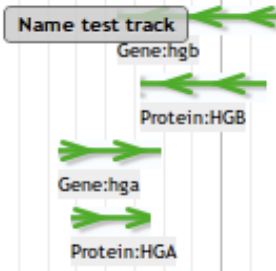
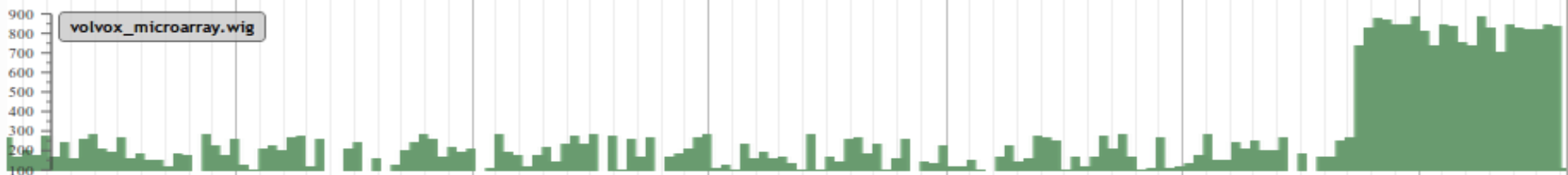
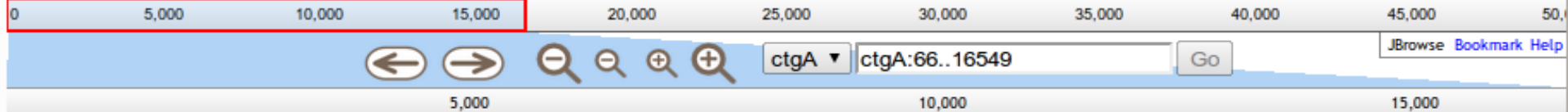
University of California, Berkeley



(JBrowse is currently at 1.2.2)

Available Tracks

- Example motifs
- Frame usage
- Predicted genes
- Exonerate predictions
- Fingerprinted BACs



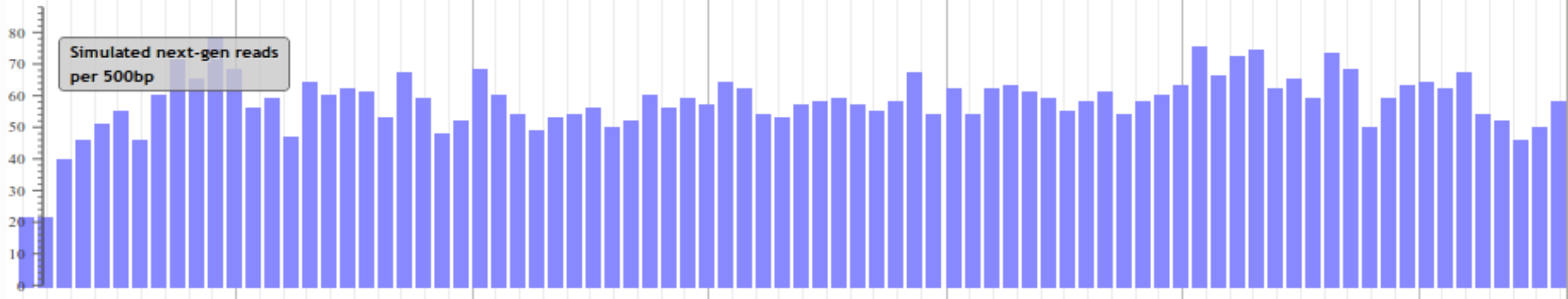
Example alignments



ESTs

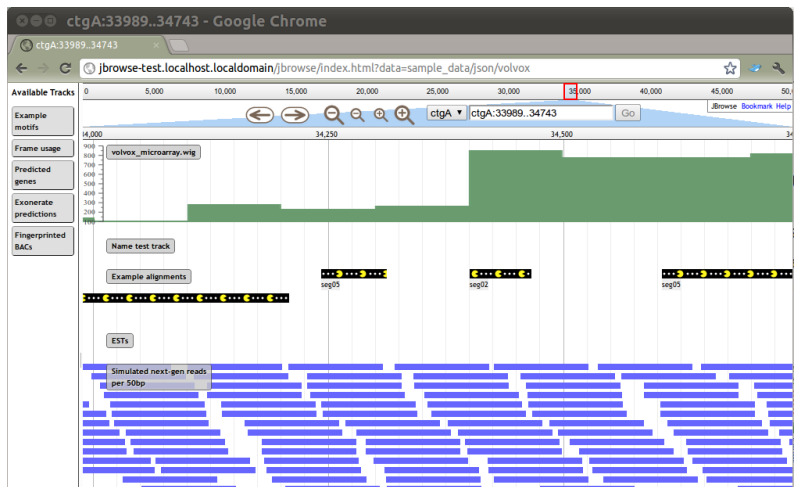


Simulated next-gen reads per 500bp



JBrowse...

- Fast, smooth navigation (*think Google Maps for genomes*)
- Supports BED, GFF, Bio::DB::*, Chado, WIG, BAM, UCSC (*intron/exon structure, name lookups, quantitative plots*).
- Relies on pre-indexed static files, which is highly secure and performant. (*future versions will have an optional server-side component*)
- Is stably funded by NHGRI, many interesting innovations implemented and/or pending integration.
- Is open source, of course.
- Did I mention it's fast?



The JBrowse Project

- free and open source (license: GPL / Artistic)
- a **GMOD** project
<http://gmod.org>
- developed using `git`, hosted on GitHub
<http://github.com/GMOD/jbrowse>
- PIs most involved: Ian Holmes, Lincoln Stein, Suzi Lewis
- just got 3 more years of funding from NHGRI!



A Brief History of JBrowse

- 2006 First prototype (pre-rendered GBrowse).
NHGRI funds 3yr development of project.
Mitch Skinner joins the team.
- 2007 CSHL Biology of Genomes: scaling.
- 2008 JavaScript-only rendering.
- 2009 Genome Research paper.
- 2010 WebApollo funded to develop JBrowse.
NHGRI renews for 6mo.
JBrowse = “official” GBrowse successor.
- 2011 Mitch Skinner joins genomics startup.
US politicians play chicken with economy.
NHGRI renews for 3yr.
- 2012 (January) **Robert Buels joins the team.**

Release Roadmap

1.3.0 – Very Imminent™

1.3.N bugfix releases as needed

1.4.0 – next month

1.4.N bugfix releases as needed

2.0.0 – end of 2012?

Version numbers according to
Semantic Versioning (semver.org) informal
standard.

Course Set for 2.0

- JBrowse needs a lot more **Big Features**:
 - need to upload and share tracks
 - need data export
 - need large numbers of tracks with rich metadata; searchable
 - need powerful, modular extensions
 - data backends
 - client-side plugins

Yes, but ...

JBrowse 1.2.x Urgent Problems

- Flat-file formatting is horribly inefficient.
- Lack of Y-axis labels make quantitative tracks nearly meaningless.
- Lack of dynamic (rubber-band) zooming is very annoying.
- Unusable for large numbers of ref seqs (e.g. rough assemblies) - a major use case.
- No tests for anything: **major** problem for developers.

JBrowse Architecture

- Still has its baby teeth.
- Non-backward-compatible changes needed to for JBrowse to sustainably achieve the **Big New Features**.
- Separation of configuration from data.
- Modular data backends.
- Modular frontend (event publish/subscribe).

JBrowse 1.3.0 Objectives

- Fix the urgent problems that are not so hard.
- Fix minor problems that are easy.
- Add more polish.
- Preserve backward compatibility.

JBrowse 1.3.0 Highlights

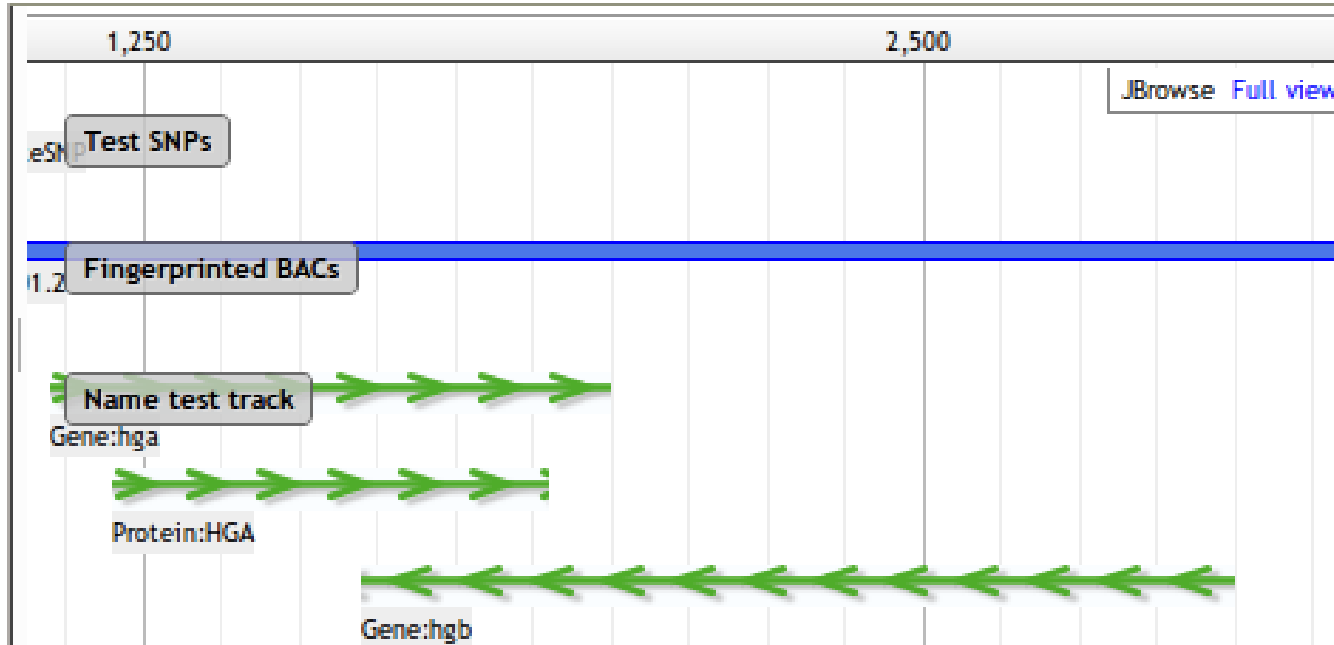
- **“rubber-band” dynamic zooming**
- much better flat file formatting performance
- **support for JavaScript “hooks” to modify FeatureTracks**
- **Y-axis labels**
- sequence track revcom (thanks to Ian Davis)
- stock index.html supports a data URL param to use different data dirs
- **inline help dialog**
- more grid lines
- data format supports more than 2 levels of features
- **NCL and sequence data can be compressed on disk**
- compatibility layer for 1.2.x data
- **1-based coordinates in the user interface**
- Julie Moon - “embedded” mode (like `gbrowse_img`)
- improved documentation - POD, JSdoc, and tutorial
- **(and lots of improvements behind the scenes)**
 - Perl unit and integration tests, JS integration tests with Selenium, internal refactoring

FeatureTrack Hooks

```
"hooks": {
  "modify": "function(track, feat, elem) {
    var fType = feat.get('Type');
    if (fType) {
      elem.className = 'basic';
      switch (fType) {
        case 'CDS':
        case 'thick':
          elem.style.height = "10px";
          elem.style.marginTop = '-3px';
          break;
        case "UTR":
        case "thin":
          elem.style.height = '6px';
          elem.style.marginTop = '-1px';
          break;
      }
      elem.style.backgroundColor = 'red';
    }
  }
}
```

have to hack the data/trackList.json to use these >:-{

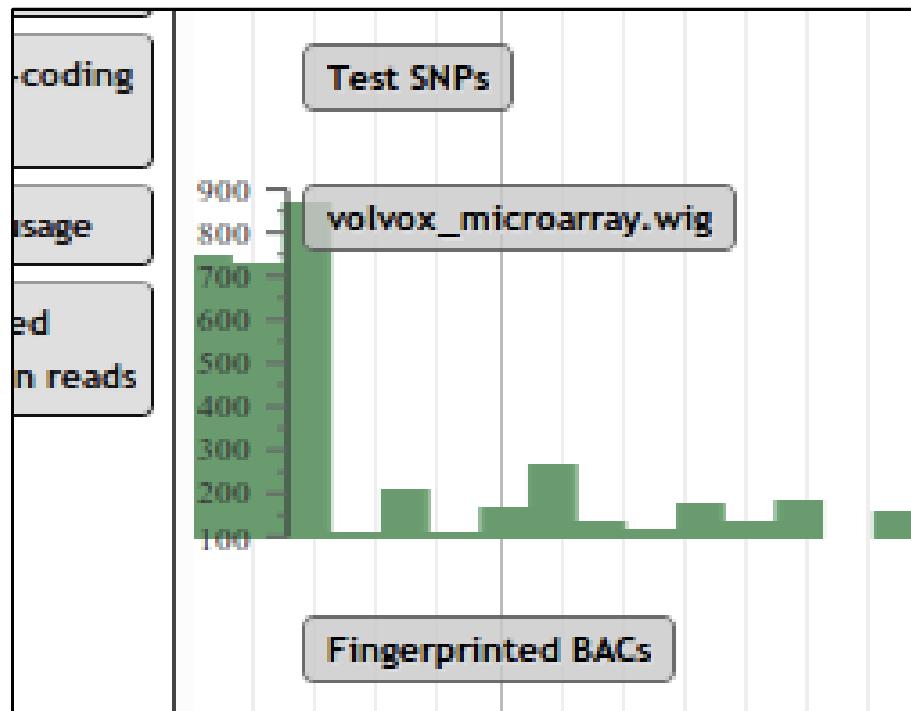
Embedded Mode



Julie Moon, OICR co-op student

```
http://localhost/jbrowse/index.html
?loc=ctgA:1386..6057
&tracks=<list of tracks>
&data=sample_data/json/volvox
&nav=0
&overview=0
&tracklist=0
```

Y-Axis Labels

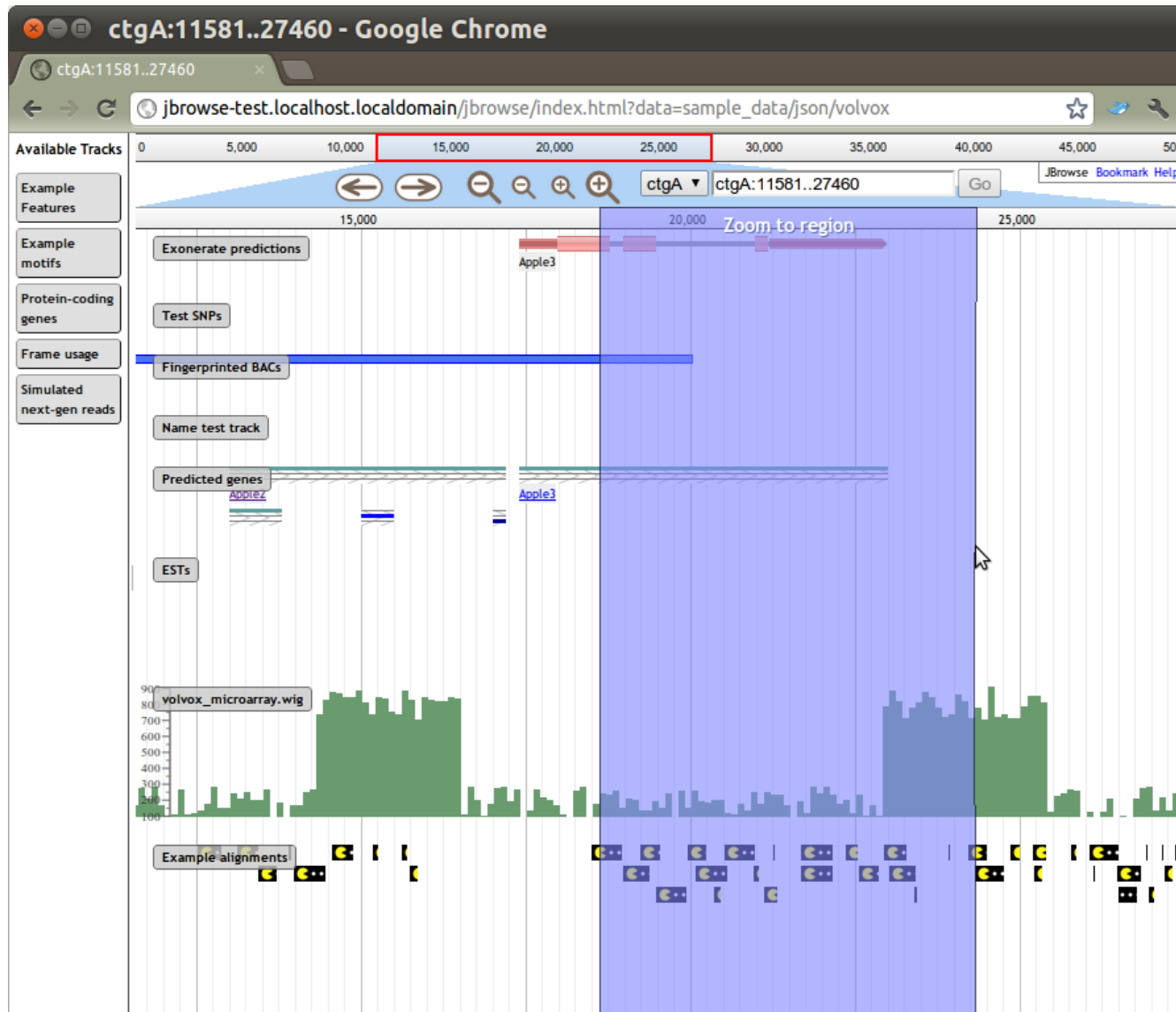


- dojox.charting (SVG/VML)
- fully dynamic, of course

Pre-Compression

- gzips static data files on disk
- shrinks data directories by **about 60%**
 - takes ~~hg19~~GRCh37 from 19GB to 7.6GB.
- needs some webserver configuration for headers
 - Apache: AllowOverride FileInfo (.htaccess files)
 - nginx: add a smallish config snippet

Rubber-band Zooming



Queued for 1.4.0

- performance for large numbers of reference sequences
- autocompletion for refseq and feature names
- probably a few more things
- maybe can hoist faceted track search and rich track metadata here

Low-hanging fruit, really-super-urgent stuff.

Aims for 2.0

- User-editable, direct config file.
- Load and search rich track metadata for thousands of tracks. (what formats?)
- Datasource plugin system, support direct byte-range access to BAM, BigWig, and BigBed files.
- New, documented pub/sub event API ties together both JBrowse components and **any other JS code.**

Config File

- what format?
- leaning toward GBrowse-native format
- will also support injection of configuration information dynamically

Publish / Subscribe Example

- JBrowse publishes “/jbrowse/evt/region/highlighted chr1:12345..67890”
- Some other code has subscribed to this, and
 - calls a web service to launch an analysis on that region
 - pops up a notification saying it has done so
- When the analysis is done, it then publishes: “/jbrowse/command/track/add <track config JSON>”
- Track with the analysis results pops up in the browser

Track Metadata

- What formats do people use to manage this?
 - Excel?
 - RDBMS schemas?
 - XML?
- What should JBrowse do for loading or directly reading it?

Big Thanks To:

Suzi Lewis (LBNL)

Gregg Helt

Ed Lee



Nomi Harris

Ian Holmes (Berkeley)

Mitch Skinner

Stephen Cummings

Lincoln Stein (OICR)

Julie Moon

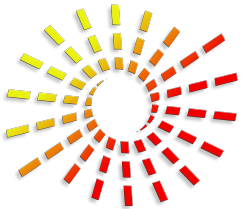
Natalie Fox

Adam Wright



Lukas Mueller (BTI, SGN)

NHGRI



<http://jbrowse.org>



<http://gmod.org/wiki/JBrowse>

github

<http://github.com/GMOD/jbrowse>

(yes, jbrowse needs a logo)

The End

Robert Buels

rbuels@gmail.com

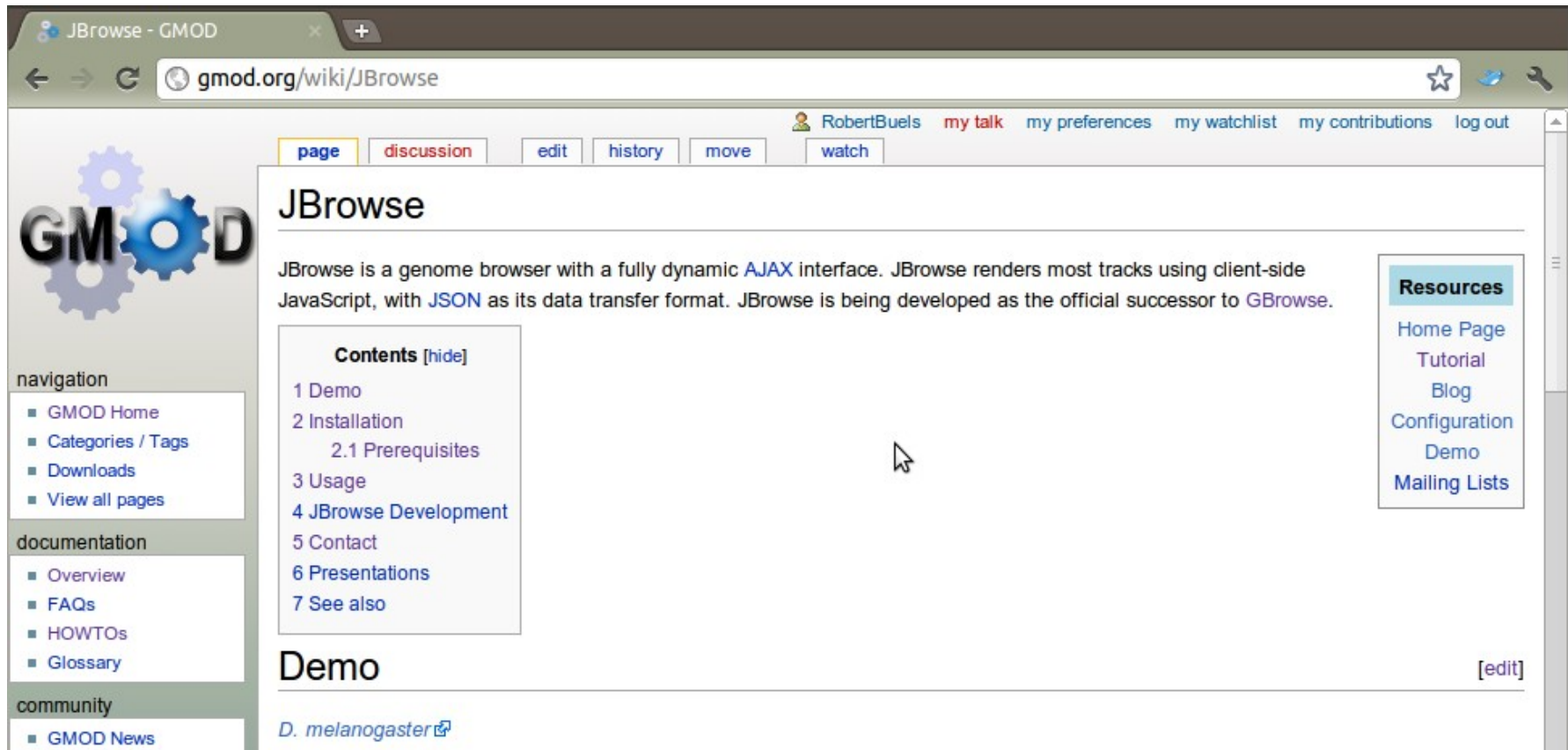
rbuels on freenode IRC
(#bioperl, #gmod)



- Was Engineering Lead at Sol Genomics Network <http://solgenomics.net>, at BTI with Lukas Mueller
- Became new JBrowse Lead Developer as of January 2012 at UC Berkeley with Ian Holmes

JBrowse Documentation

The GMOD wiki is the documentation hub
<http://gmod.org/wiki/JBrowse>



The screenshot shows a web browser window with the address bar displaying `gmod.org/wiki/JBrowse`. The page title is "JBrowse" and the user is logged in as "RobertBuels". The page content includes a description of JBrowse as a genome browser with a fully dynamic AJAX interface, rendered using client-side JavaScript and JSON. A table of contents is visible, listing sections from 1 Demo to 7 See also. The page also features a navigation sidebar on the left and a resources sidebar on the right.

JBrowse - GMOD

gmod.org/wiki/JBrowse

RobertBuels my talk my preferences my watchlist my contributions log out

page discussion edit history move watch

JBrowse

JBrowse is a genome browser with a fully dynamic [AJAX](#) interface. JBrowse renders most tracks using client-side JavaScript, with [JSON](#) as its data transfer format. JBrowse is being developed as the official successor to [GBrowse](#).

Contents [hide]

- 1 Demo
- 2 Installation
 - 2.1 Prerequisites
- 3 Usage
- 4 JBrowse Development
- 5 Contact
- 6 Presentations
- 7 See also

Demo [edit]

[D. melanogaster](#)

Resources

- Home Page
- Tutorial
- Blog
- Configuration
- Demo
- Mailing Lists

navigation

- GMOD Home
- Categories / Tags
- Downloads
- View all pages

documentation

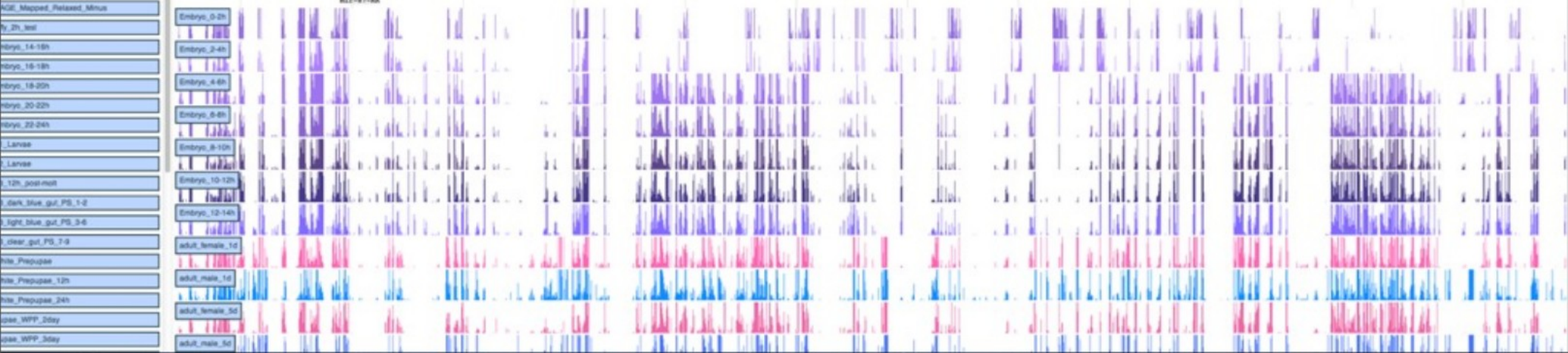
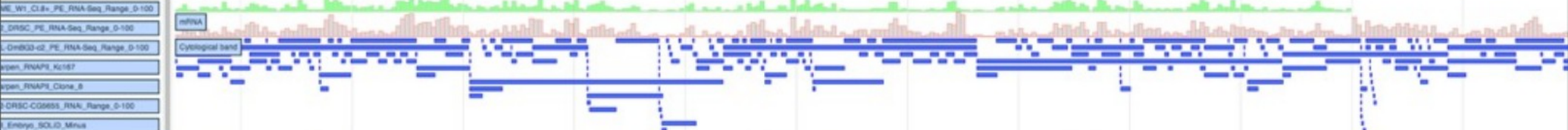
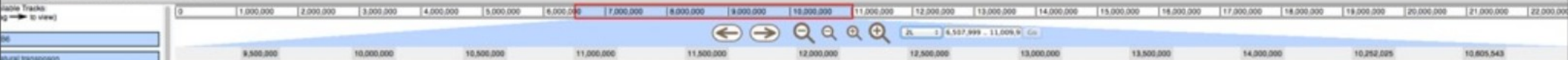
- Overview
- FAQs
- HOWTOs
- Glossary

community

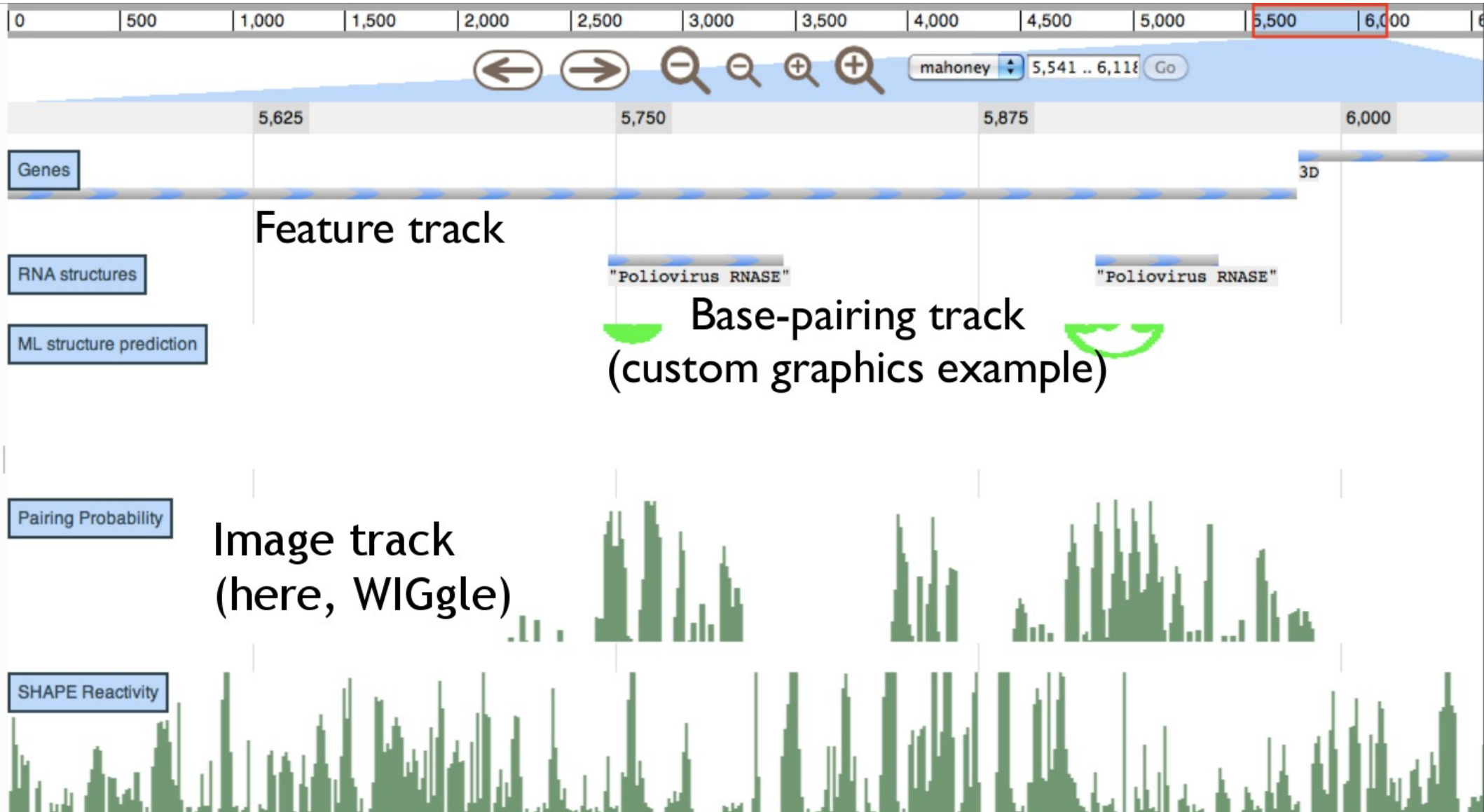
- GMOD News

First look: Live Demo

- a couple of JBrowsers around the web
 - <http://intron.ccam.uchc.edu/JBrowse/Dmel/>
 - <http://jbrowse.org/ucsc/hg19/>
- maybe install a demo JBrowse locally with the tutorial

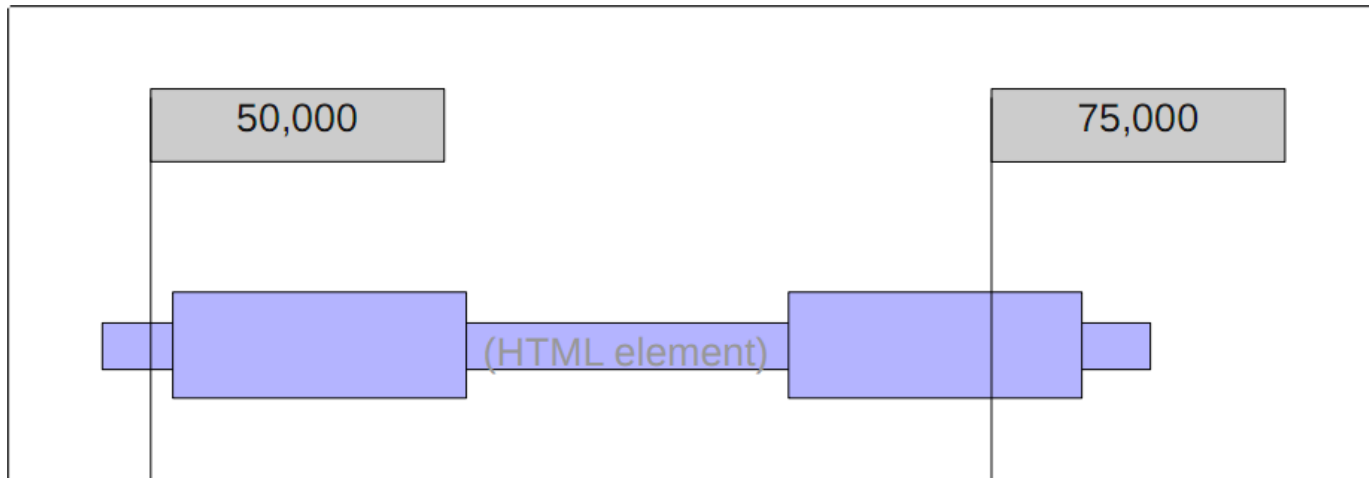


Types of Tracks



Feature Tracks

- rectangles (`<div>s`) with various fills and dimensions to represent the features
- can do a surprising amount, but still limited



Feature Tracks: Density Plots

Zoomed out



Zoomed in

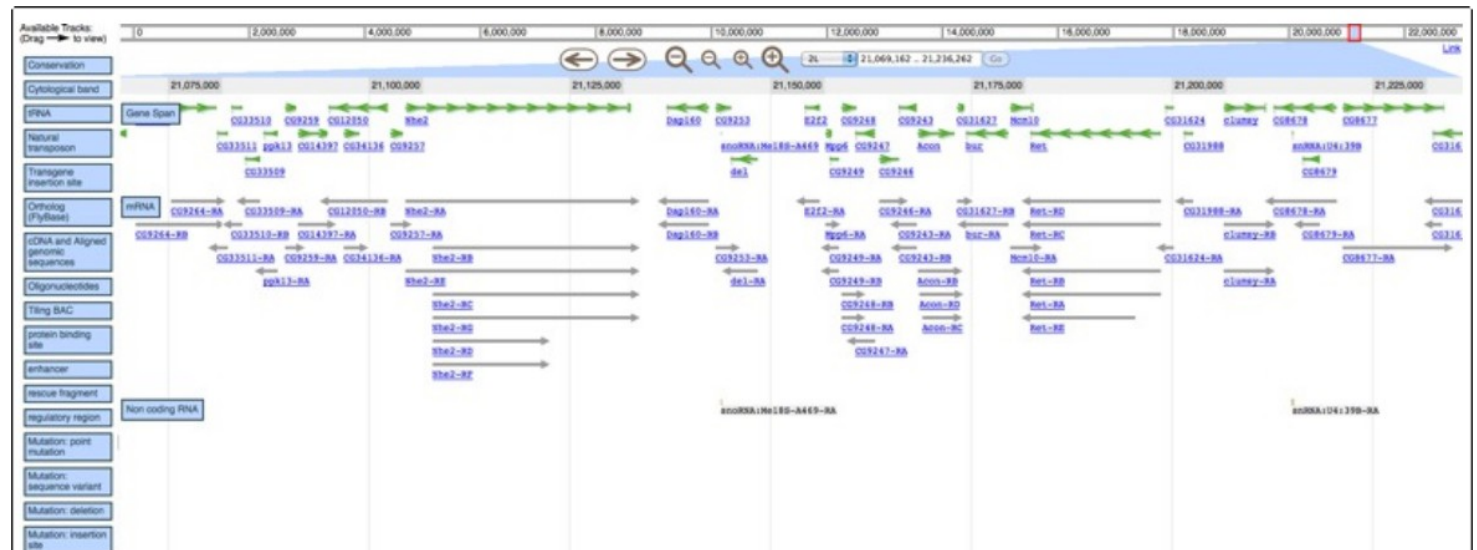
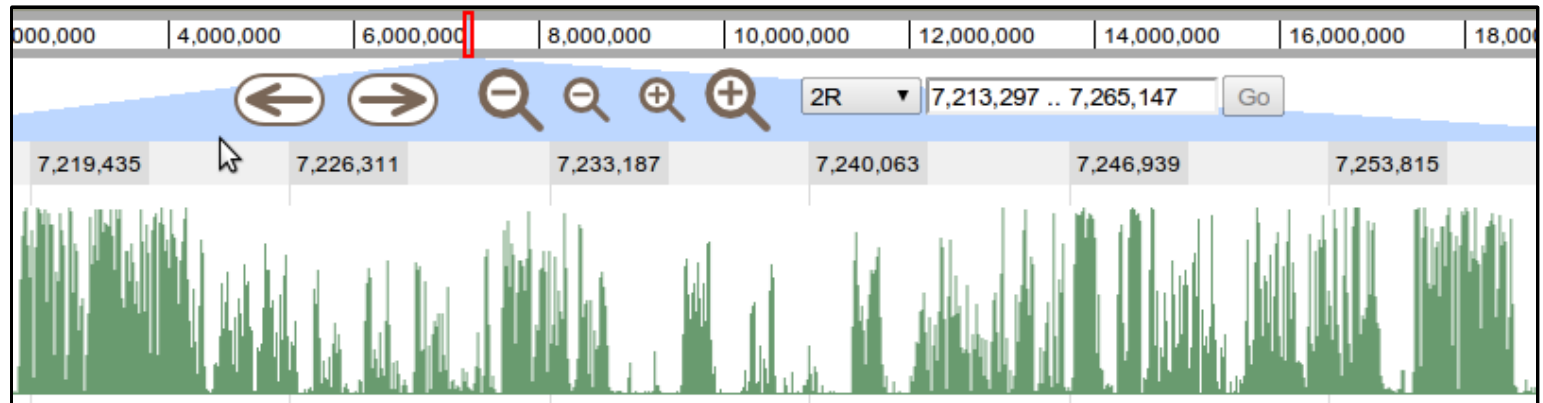


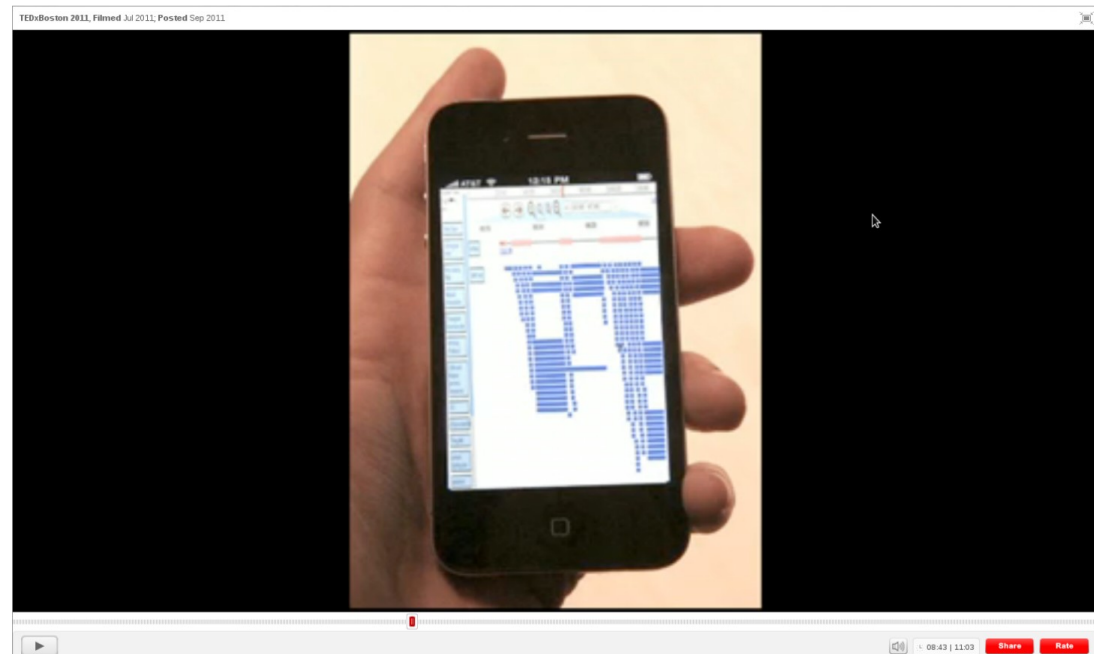
Image Tracks

- any pre-generated images that cover the genome
- quantitative data (wiggle/BigWig)
 - wig-to-json.pl
- RNA base-pairing



± Particular Strengths

- **very** fast and smooth!
- easy things are easy!
- works nicely on an iPad/iPhone



JBrowse on an iPhone
video frame from TEDx Boston talk by Richard Resnick

± Particular Weaknesses

- no user-uploaded data
- no data export
- slow for big numbers of reference seqs (e.g. 5,000 annotated contigs)
- few glyph options, feature tracks are limited by the facts of <div>

These will all be improved in the 2.x series.

Coming Soon

- new track selection
- overview images
- context menus
- track customization
- drag to zoom!

Coming: Hierarchical Track Sel.

Enter text to search track list:

clear search

Available Tracks:
(Drag → to view)

reset list order

- Favourites
- + SequenceTrack
- General
 - Conservation
- Gene Model features
 - Gene Span
 - mRNA
- + Genomic reagents and data
- Genomic features
 - Cytological band
 - Natural transposon
 - Transgene insertion site
- + Mapped features
- + Prediction features
- + Similarity: Synteny features

Select Tracks

0 5,000,000 10,000,000 15,000,000 20,000,000

← → 🔍 🔍 🔍 🔍 2L 13,222,878 .. 13,299,598 Go [Link](#)

13,225,000 13,250,000 13,275,000

Conservation

Cytological band

band - 34A11 band - 34B2 band - 34B3 band - 34B4 band

band - 34B band - 34B1

Tiling BAC

Transgene insertion site

Gene Span

1Bb CG16813 CG15479 CG5945 snRNA:U2:34ABc CG31728 CG6043 CG6116 CG16824

i6 CG15480 CG16815 CG5867 snRNA:U5:34A CG31851 CG31848 CG6108 Rep4 CG31729

CG16812 Prosalpha6T CG16820 CG31730

15

work by OICR intern Natalie Fox

Coming: Faceted Track Selection

clear selected tracks

Click on the track name below to display/hide the track.

22 Items

sorted by: labels; then by... • grouped as sorted

- cDNA and Aligned genomic sequences
- Conservation
- Cytological band
- DNA
- enhancer
- Gene Span
- mRNA
- Mutation: deletion
- Mutation: insertion site
- Mutation: point mutation
- Mutation: sequence variant
- Natural transposon
- Non coding RNA
- Oligonucleotides
- Ortholog (FlyBase)
- protein binding site
- regulatory region
- rescue fragment
- Tiling BAC
- Transgene insertion site
- tRNA
- XRate predictions

Show only the first 10 results

0 5,000,000 10,000,000 15,000,000 20,000,000

← → 🔍 🔍 🔍 🔍 2L

13,222,878 .. 13,274,458 Go

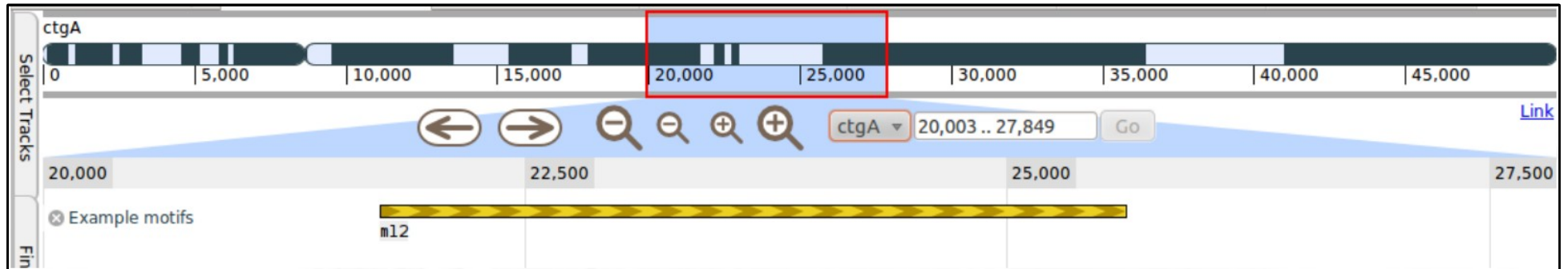
13,225,000 13,250,000

Find Tracks

- Conservation
- Cytological band
 - band - 34A11
 - band - 34B1
 - band - 34B2
 - band - 34B3
- Tiling BAC
- Transgene insertion site
- Gene Span
 - lBb
 - CG16813
 - CG15479
 - CG5945
 - snRNA:U2:34ABc
 - CG31728
 - CG6043
 - l6
 - CG15480
 - CG16815
 - CG5867
 - snRNA:U5:34A
 - CG31831
 - CG31848

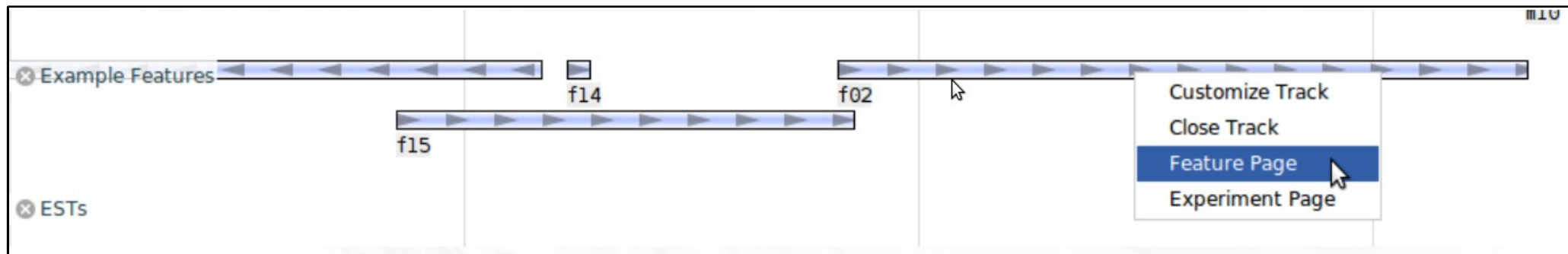
work by OICR intern Natalie Fox

Coming: Overview Images



Can have a karyotype, or any other image.

Coming: Context Menus



Coming: Track Customization

The image displays a genomic browser interface with two main sections: a customization panel on the left and a track view on the right.

Customization Panel (Left):

- Select Tracks:** A section titled "Customizing Example motifs track" showing a yellow dashed track. Below it are "apply changes" and "cancel" buttons.
- Find Tracks:** A section with several dropdown menus for customization:
 - height: select new height
 - fill color: select new fill color
 - track image: select new track image
 - border color: select new border color
 - border color: select new border width

Track View (Right):

- ctgA:** A genomic map at the top with a scale from 0 to 45,000. A red box highlights a region from 20,000 to 25,000.
- Navigation:** Below the map are navigation icons (left, right, zoom in, zoom out) and a search bar containing "ctgA" and "20,000.. 27,314" with a "Go" button and a "Link" text.
- Zoomed View:** The view is zoomed into the 20,000 to 25,000 region, with vertical grid lines at 20,000, 22,500, and 25,000.
- Tracks:** Several tracks are visible:
 - Example motifs:** A yellow dashed track with a feature labeled "m12" at approximately 22,500.
 - Example Features:** A blue track with arrowheads pointing right, containing features "f15" (at ~22,500), "f14" (at ~23,500), and "f02" (at ~25,000).
 - ESTs:** A track with a light blue background.
 - Frame usage:** A green track with a solid bar.
 - Example alignments:** A blue track with arrowheads pointing right, containing a feature labeled "seq02" at the far right.

Migrating: GBrowse to JBrowse

- JBrowse does not yet have all the features of GBrowse
- Configuration and data formats are different
- Adam Wright, intern at OICR, has worked on automated migration tools to help
- Another thing I need to merge into master ;-)



WebApollo

- next generation of the popular Apollo annotation editor
- genome editing in your browser
- based on JBrowse



WebApollo

- clients receive annotation updates **in real time** (like Google Docs)
- saves edits to a central database
- has produced several extensions that will be coming to core JBrowse
 - highlighting and selecting features
 - HTML5 Canvas tracks for wiggle data (opening the door for many more canvas-based techniques)



WebApollo Live Demo!

Step
1

<http://goo.gl/SHByG>

User name: **write_1**
Password: **write_1**

Step
2

<http://goo.gl/mSh3w>

