

1. Home	3
1.1 About the User's Guide	3
1.1.1 Archived User's Guides	4
1.2 Quick start	4
1.3 Get IGB	7
1.4 IGB terminology	8
1.5 IGB main window	10
1.6 Loading data	14
1.7 File Formats	16
1.8 Zooming and panning the display	20
1.8.1 Panning the Display	20
1.8.2 Zooming	22
1.9 Selecting Items	26
1.9.1 Using the mouse to select or move	27
1.10 Right-click menus	27
1.11 Tabbed panels	31
1.11.1 Data Access	32
1.11.2 Selection Info	33
1.11.3 Advanced Search	35
1.11.4 Sliced View	37
1.11.5 Annotation	41
1.11.6 Graph	43
1.11.7 External View	47
1.11.8 Plug-ins	49
1.11.8.1 Geometric Mean	49
1.11.8.2 Primer Design	49
1.11.8.3 Tutorial Helper	49
1.11.9 Current Genome	49
1.11.10 Restriction Sites	51
1.11.11 Bookmarks	52
1.12 Visualizing read alignments	56
1.12.1 FindJunctions	60
1.12.2 Paired-end reads	62
1.12.3 Visualizing insertions, deletions, and mismatches	65
1.13 Visualizing probe sets	66
1.14 Sequence Viewer	68
1.15 Graph tracks	71
1.15.1 Graph operations and joining and splitting graphs	72
1.15.2 Applying thresholds to graphs	73
1.15.3 Converting GFF to bedgraph files	75
1.15.4 Creating graph tracks from annotations tracks	78
1.16 Annotation tracks	81
1.16.1 Customize track appearance	81
1.16.2 Annotation track operations	86
1.16.3 Customize annotation colors using Color By	90
1.17 Deleting tracks	93
1.18 Reference sequence	93
1.19 Custom Genomes (Genomes not in IGB)	95
1.20 Sharing data using QuickLoad sites	96
1.20.1 Create a satellite IGB QuickLoad site for genomes that are already supported in the main IGB QuickLoad site	97
1.20.2 Sharing data for a custom genome not already part of IGB QuickLoad	99
1.20.3 Set up a Quickload in Dropbox	103
1.20.4 About annots.xml	106
1.20.5 Password-protect and style your QuickLoad site using .htaccess and Apache web server	107
1.21 Saving images and printing	107
1.21.1 Editing IGB-generated SVG files	108
1.22 Saving and Loading Sessions (Draft)	109
1.23 IGB Gallery	110
1.23.1 Color-coding read alignments by strand	110
1.23.2 A new gene induced by cold in <i>Arabidopsis thaliana</i> ?	111
1.23.3 Edge matching ESTs to assess alternative splicing	112
1.23.4 Counting reads with selection tool	113
1.23.5 Alternative splicing in GRP7 (AT2G21660)	114
1.23.6 Using IGB to view effects of kmer on assembly quality	115
1.23.7 Comparing new and old annotation releases	116
1.23.8 Viewing RNA-Seq and expression tiling array data in the same view	117
1.23.9 Visualizing output from TopHat and BowTie	118
1.23.10 Visualizing RNA-Seq reads aligned onto a genome	119
1.23.11 IGB 6.7.1 running in iPlant Atmosphere	121
1.24 Advanced Features	121
1.24.1 Scripting and the IGB command language	121
1.24.1.1 IGB Script List	124

1.24.2 Scripting with R (Draft)	126
1.24.3 Controlling IGB using IGB Links	131
1.24.3.1 IGB link parameters	132
1.24.4 Command Line Arguments	132
1.24.5 Track lines for configuring track appearance	132
1.24.6 Matching species and genome versions	133
1.24.6.1 Use species.txt to link species names to IGB genome names	133
1.24.6.2 Use synonyms.txt to link genome version names to each other	134
1.24.7 Interactive tutorials	135
1.24.8 Configuring Web links (LinkOuts) to external sites	136
1.24.9 Controlling IGB using IGB Links with Genome Synonyms	136
1.24.10 Personal Synonyms	137
1.25 Preferences	138
1.25.1 Annotation Label Font	138
1.25.2 Other Options	139
1.25.3 Toolbar	141
1.25.4 Data Sources Tab	142
1.25.4.1 Adding and Managing Data Source Servers	145
1.25.5 Tracks	147
1.25.6 Track Defaults	149
1.25.7 Plugin Repositories Tab	150
1.25.8 Sharing Preferences	151
1.26 Troubleshooting	152
1.26.1 Unable to launch the application	152
1.26.2 IGB runs, but acts strangely.	153
1.26.3 I can't see my data!	153
1.26.4 Unable to detect proxy setting	154
1.26.5 My data isn't there	155
1.26.6 The IGB menus have disappeared	155
1.26.7 Where is my IGB download? (Safari)	155
1.26.8 Installing and Updating Java	156
1.27 Tutorials	157
1.27.1 Focus on a Feature Live On-line Tutorials	157
1.27.2 Assessing splicing - Cold-induced alternative splicing of Arabidopsis SR45	157
1.27.3 RNA-Seq Coverage graphs - Visualizing data from maize reproductive transcriptomes	165
1.27.4 View a bacterial genome in IGB	175
1.28 Video Tutorials	179
1.28.1 IGB Quickies	180
1.28.2 Focus on a Feature	180
1.28.3 In-Depth Analysis	181
1.29 Older Versions of IGB	181
1.29.1 Getting Older Versions of IGB	181
1.29.2 System Requirements for Older Versions	181
1.29.3 Troubleshooting Older Versions	181
1.29.3.1 IGB can't be opened because it is from an unidentified developer (Mac)	181
1.29.3.2 Java Web Start starts, but fails to launch IGB	183
1.29.4 Release notes for older versions	183
1.29.4.1 New in IGB 6.5	183
1.29.4.2 New in IGB 6.6	188
1.29.4.3 New in IGB 6.7	194
1.29.4.4 NEW in IGB 7.0	197
1.30 IGB License	205

# Home

## About IGB

The Integrated Genome Browser (IGB, pronounced Ig-Bee) is a fast, flexible, and free desktop genome browser. First developed at Affymetrix in 2001 to support visual analytics of genome tiling arrays, IGB provides an advanced, highly customizable environment for exploring and analyzing large-scale genomic data sets.

Using IGB, you can:

- View your RNA-Seq, ChIP-chip or ChIP-seq data alongside genome annotations and sequence.
- Investigate alternative splicing, regulation of gene expression, epigenetic modifications of DNA, and other genome-scale questions.
- View results from aligning short-read sequences onto a target genome, identify SNPs, and check alignment quality.
- Copy and paste genomic sequences for further analysis into other tools, such as primer design and promoter analysis tools.
- Create high-quality images for publication in a variety of formats.

If you use IGB in your research, please cite:

Nicol JW, Helt GA, Blanchard SG Jr, Raja A, Loraine AE.

**The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets.**

Bioinformatics. 2009 Oct 15;25(20):2730-1. Epub 2009 Aug 4. PubMed PMID: 19654113; PubMed Central PMCID: PMC2759552.

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2759552>

## IGB features

IGB lets you view results from your own experiments or computational analyses alongside public domain gene annotations, sequences, and genomic data sets, thus making it easier for you to determine how your experiments agree or disagree with current thinking and models of genomic structure.

Some features IGB offers include:

- **Animated zooming.** Most genome browsers implement "jump zooming" only, in which you click a zoom button (or other type of control) and then wait for the display to re-draw. In IGB, zooming is animated, allowing you to easily and quickly adjust the zoom level as needed without losing track of your location.
- **Simple Data Sharing System - QuickLoad.** IGB implements a very simple, easy-to-use system for sharing data called QuickLoad. You can use the QuickLoad system to set up a Web site you can use to share your data with colleagues, reviewers, and the public.
- **Draggable graphs.** You can display genome graphs data (e.g., "bar" and "wiggle" files) alongside and even on top of reference genome annotations, thus making it easier to see how your experimental results match up to the published reference genome annotations. You can reset your graphs to "floating" and click-drag them over annotations to compare your results with annotations and others' experiments.
- **Edge-matching across tracks.** When you click an item in the display, the edges of other items in the same or different tracks with identical boundaries light up, highlighting interesting similarities or differences across gene models, sequence reads, or other features.
- **Integration with local and remote external data sources.** IGB can load data from a variety of sources, including Distributed Annotation Servers, QuickLoad servers, ordinary Web sites, and local files.
- **Intron-trimming sliced view.** In many species, introns are huge when compared to the exonic (coding) regions of genes. IGB provides a Sliced View tab that trims uninformative regions from introns.
- **Web-controls.** IGB can be controlled from a web browser or any other program capable of sending HTTP requests. Via IGB links, you can create Web pages that direct IGB to scroll to a specific region and load data sets from local files or servers.
- **Scripting.** IGB understands a simple command language that allows users to write simple scripts directing IGB to show a genome, zoom and scroll to specific regions, and other functions.
- **Open source.** All development on IGB proceeds via a 100% open source model. The license allows developers to incorporate IGB (and its components) into new applications.

## About the User's Guide

The first IGB User's Guide appeared in 2003 when Affymetrix released IGB for the first time. For many years, the User's Guide was published in PDF format. In 2010, we converted the User's Guide to wiki format to allow more people to contribute to the Guide and update it as the IGB application changes and grows.

## Credits

The IGB User's Guide uses Confluence wiki software from [Atlassian](#) Software, which gave us a complimentary license to use their products.

Many people have edited and added content to the User's Guide. Contributors include Ann Loraine, John Nicol, Archana Raja, Hiral Vora, Shira Stav, Vikram Bishnoi, Alyssa Gullledge, Mason Meyer, David Norris, Richard Linchangco and many others.

## Document Conventions

On-screen text (i.e., menus, button labels, folder names, field values, etc.) is indicated by **bold** font.

An item followed by a "greater than" symbol followed by another item means "click the first item; the second item will become visible; then click the second item."

Examples:

- **File > Open File...** means "click on the **File** menu and choose **Open File...**"
- **View > Set AutoLoad Threshold to Current View** means "click on the **View** menu and then click the **Set AutoLoad Threshold to Current View** button"

Note that some images are from earlier versions of IGB and may differ slightly from the latest, released version.

Unless otherwise noted, the manual assumes you are using a two button mouse (right and left buttons). If using a single-button mouse (e.g., on an Apple computer), press the CNTRL (control) key while clicking to perform operations that require right-clicking.

## Archived User's Guides

### Click on link to download pdf of previous versions of the User's Guide

[IGB 8.2](#)

## Quick start

If you are new to IGB, use this **Quick Start Guide** get started.

- [Step 1: Get and start IGB](#)
- [Step 2: Choose species and genome version](#)
- [Step 3: Open data sets](#)
- [Step 4: Zoom in](#)
  - [Other ways to zoom](#)
- [Step 5: Load data](#)
- [Step 6: Configure tracks](#)

## Step 1: Get and start IGB

- Go to the [IGB Download Page](#)
- Select and download the installer for your platform

If you have trouble starting IGB, see [Troubleshooting](#) or ask for help on [BioStars](#).

## Step 2: Choose species and genome version

To choose a species and genome version

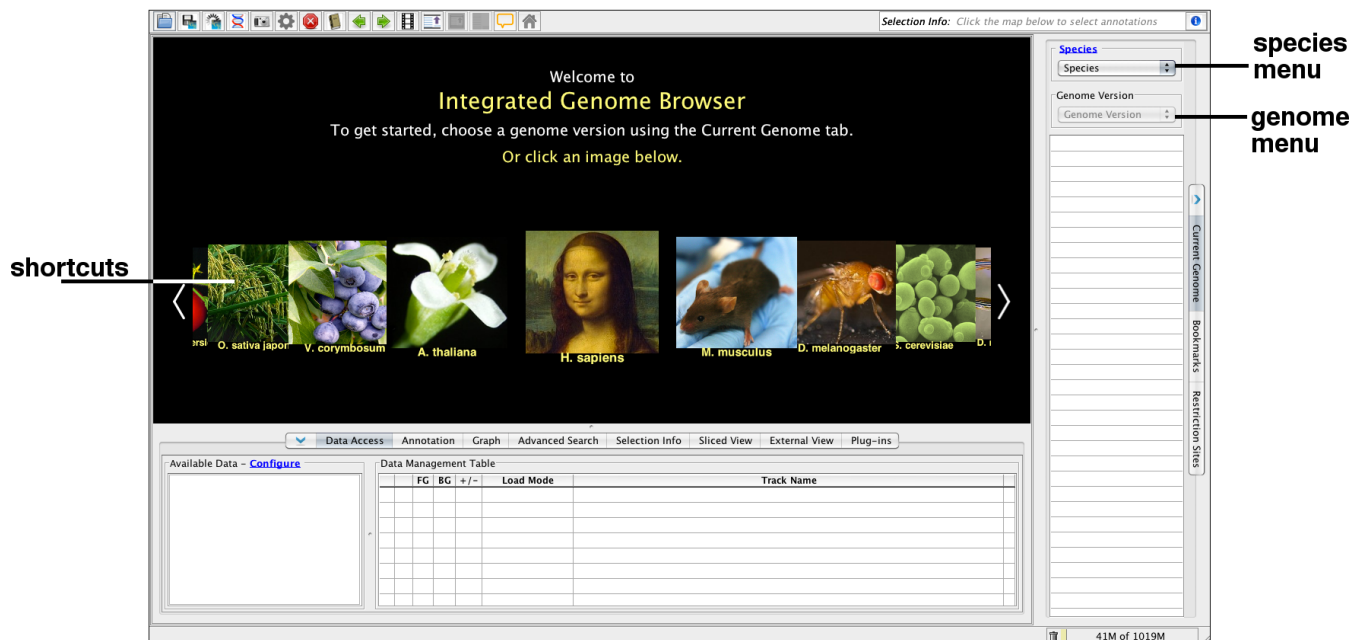
- Click a shortcut image (loads most recent genome)

or

- Choose **Species** and **Genome Version** using the **Current Genome** tab.

**IGB start screen**





If your species or genome version is not listed, you can view it if you have a fasta or 2bit file with sequence data. See [Custom Genomes](#)

### Step 3. Open data sets

Open data sets from remote data sources (**Data Access** tab) or by opening local files.

To open a data set from a remote data source

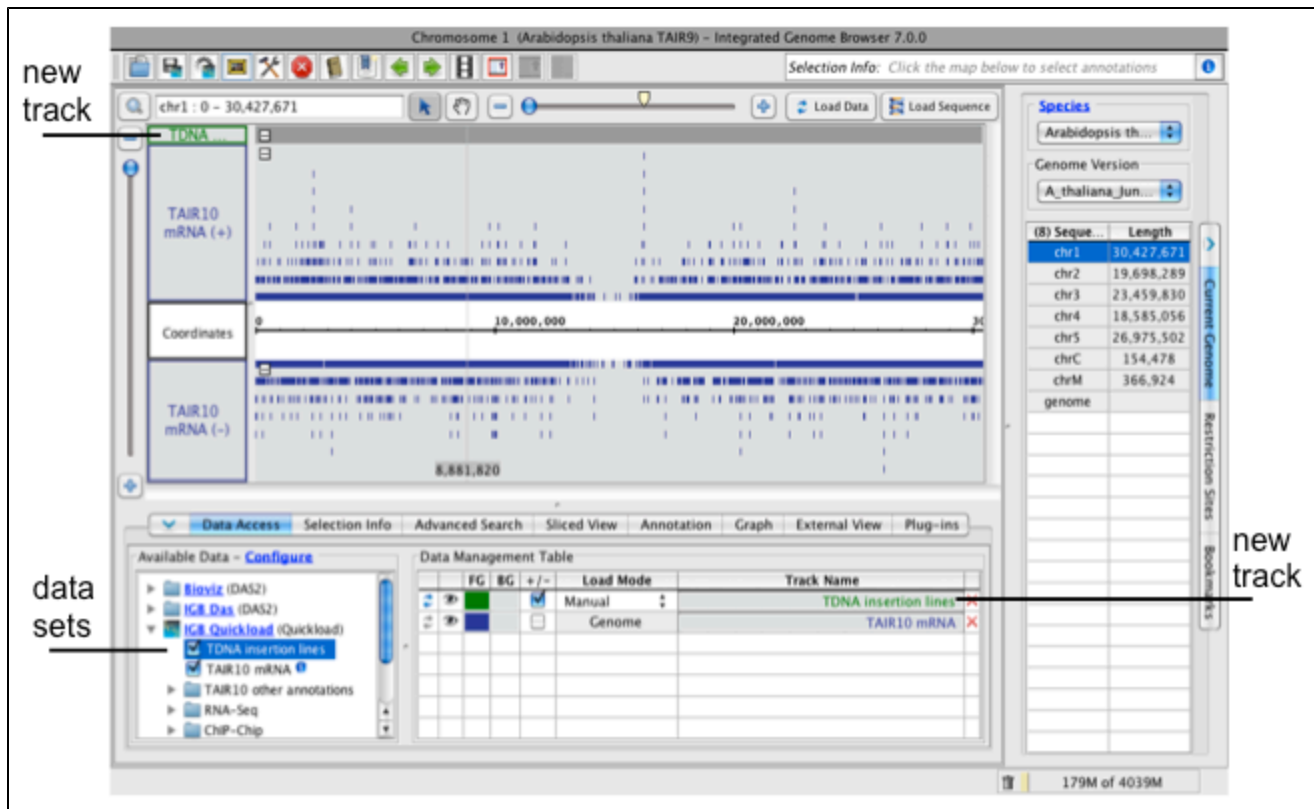
1. Select **Data Access**
2. Select data sets in the **Available Data** section

To open local files on your computer

1. Select **File > Open File...** or **File > Open URL...**
2. Enter file name or URL

When you select a data sets or a file, IGB will add a new empty track to the main view and list it in the **Data Management** table. Regions that do not have data loaded have a darker background color.

**IGB window after opening a data set**



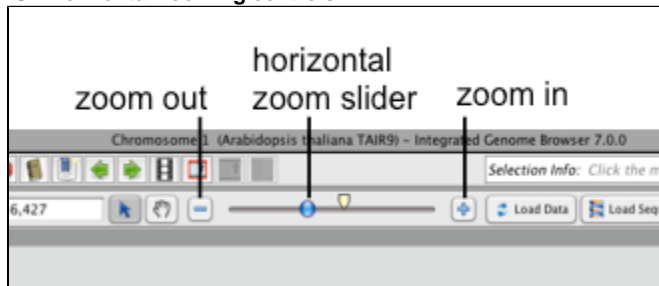
## Step 4: Zoom in

Because many data files contain too much data to view all at once, IGB does not load data into the viewer until you click the **Load Data** button.

Before loading data, zoom in to a region.

1. Click a location in the main view
2. Drag the horizontal zoom slider or use plus and minus buttons

### IGB horizontal zooming controls



## Other ways to zoom

Other ways to zoom include

- Search for a gene by name
- Double-click a feature
- Click-drag within the coordinate axis

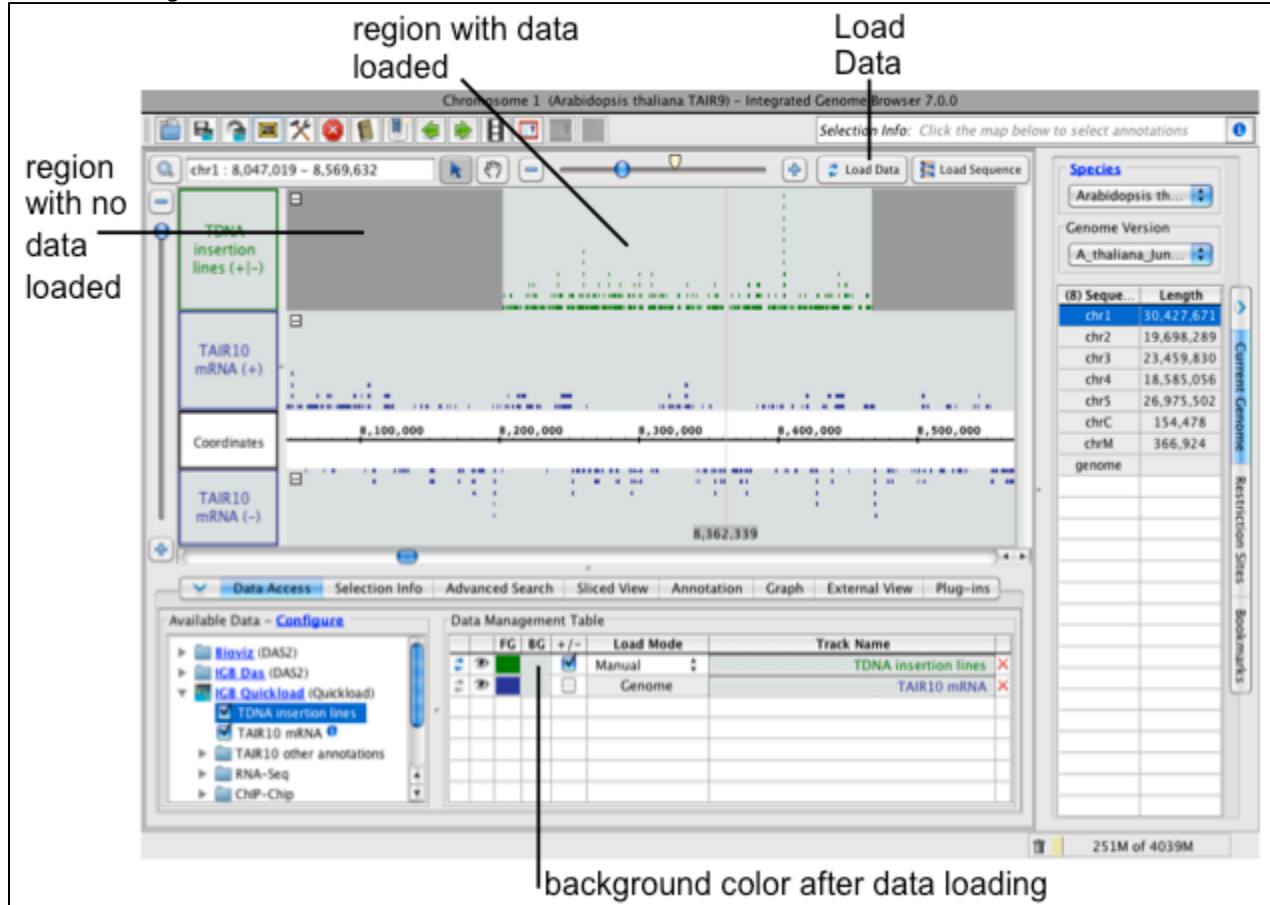
See also:

- [Zooming](#)
- [Panning the display](#)

## Step 5: Load data

To load data, click **Load Data** button. Regions with loaded data show the selected background color; areas without loaded data appear darker.

#### IGB after loading data



See also:

- Main View
- Loading Data

## Step 6: Configure tracks

You can reorder the tracks by dragging the **Track Label** (the **Data Management Table** reflect changes).

To change style elements of a track, click the track label and use the **Annotation** or **Graph** tab to change to change color, track height, annotation label, amount of data shown (stack height), and other options.

See also:

- Customizing tracks
- Preferences
- Annotation tab
- Graph Tab

## Get IGB

- Get IGB using native installer for your platform
- Get compiled IGB (not using native installer)
- Installing on Linux
- How to get older versions of IGB
- How to change IGB memory usage

## Get IGB using native installer for your platform

Get the latest version of IGB from [Bioviz.org](http://Bioviz.org). Follow the downloads link to get an IGB installer for your platform.

Note that the Windows and Mac installers include IGB and also the version of Java that works best for IGB. The Linux installer does not include Java.

## Get compiled IGB (not using native installer)

You can get a "zip" file with compiled IGB from this link: <http://www.bioviz.org/igb/releases/current/igb.zip>

Unpack the "zip" file and launch IGB using the IGB run script for your platform: run\_igb.bat for Windows, run\_igb.command for Mac, and run\_igb.sh for Linux.

## Installing on Linux

Download the Linux installer and run it using:

```
sh IGB_unix_current.sh
```

## How to get older versions of IGB

Installers with older versions of IGB are available from: <http://sourceforge.net/projects/genoviz/files/>

## How to change IGB memory usage

If you used the IGB installer, then when IGB starts, it will use 25% of your system's memory. This means that when IGB runs on a machine with 4 GB of memory (RAM), it will use 1 GB. Or if IGB starts on a machine with 16 GB of RAM, it will use around 4 GB of memory. You can see how much memory IGB is using by looking at the lower right corner of the IGB window.

The only exception is when you run IGB on a 32-bit Windows computer. In that case, IGB will use 1GB of RAM.

You can change how much memory IGB can use.

There is a "vmoptions" file in your IGB installation directory which contains a commented-out jvm argument that sets the max heap size, which is the amount of memory IGB uses.

To change the heap size, un-comment the jvm argument line and modify to suit your needs.

The vmoptions file on Mac is located in /Applications/Integrated Genome Browser/IntegratedGenomeBrowser.app/Contents/vmoptions.txt

On Linux/Windows, the vmoptions file is in \${IGB\_INSTALLATION\_DIR}/IntegratedGenomeBrowser.vmoptions, where \$IGB\_INSTALLATION\_DIR is the folder where IGB resides.

## IGB terminology

- [Genomes and Chromosomes](#)
- [Tracks](#)
  - [Annotations](#)
  - [Graphs](#)
  - [Sequences](#)
  - [Alignments](#)
  - [Probe set alignments](#)

## Genomes and Chromosomes

**chromosome:** In IGB, chromosome refers to any single sequence. Typically this corresponds to the sequence of a physical chromosome. Alternatively, it may represent an assembled contig corresponding to part of a physical chromosome. All chromosomes in IGB are assumed to be DNA, rather than RNA sequences. No two chromosomes within the same genome version can have the same name.

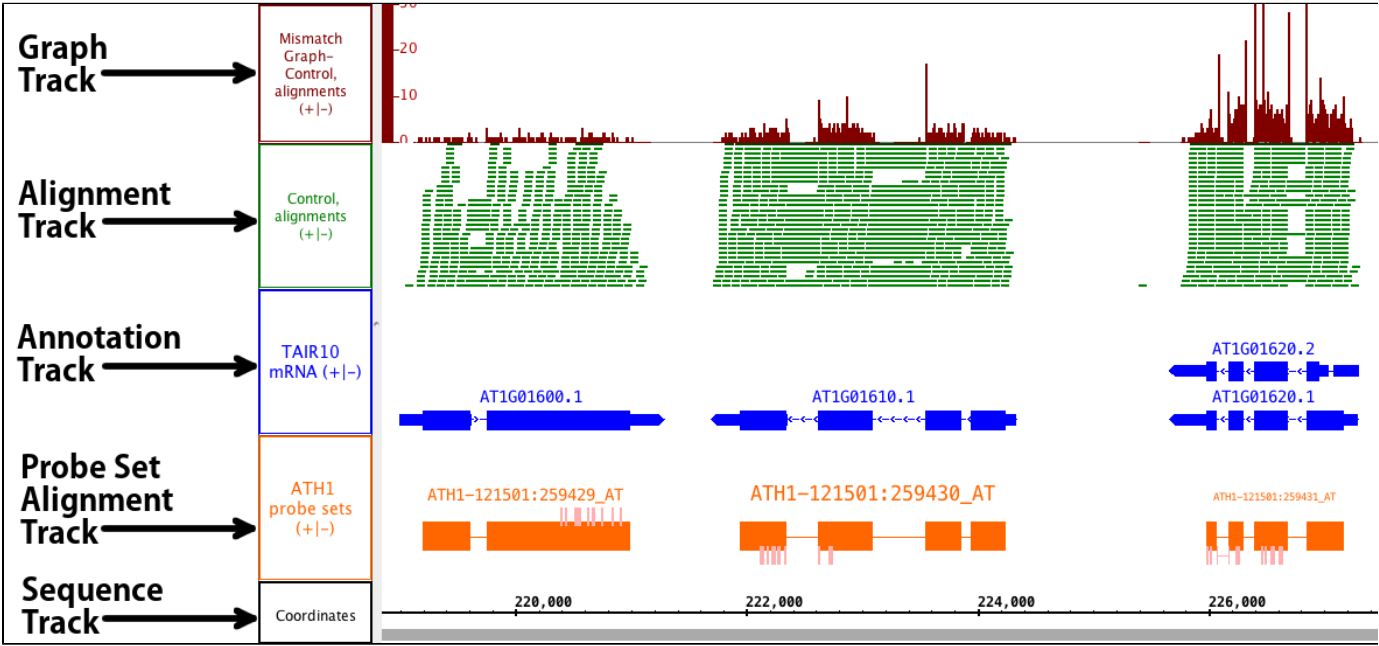
**genome, genome version, assembly:** These terms refer to a group of chromosome sequences that you or another group assembled and made available. IGB designates these using the month and year they were published or made publicly available. For example, NCBI versions 35 and 36

of the human genome are considered to be two separate genome assemblies. Each one contains multiple chromosome sequences, including the expected chromosomes 1 to 22, X and Y. Other sequences, such as "chr22\_random" are also considered distinct chromosomes for the purposes of display in IGB.

Every sequence in IGB is identified by its genome version and chromosome names, which must therefore be distinct.

## Tracks

**Tracks** are rows of data that appear in the IGB interface. When you open a file, the data within the file will appear in one or more tracks. Older versions of IGB referred to tracks as "tiers" and so you may see this term used elsewhere in the User's guide.

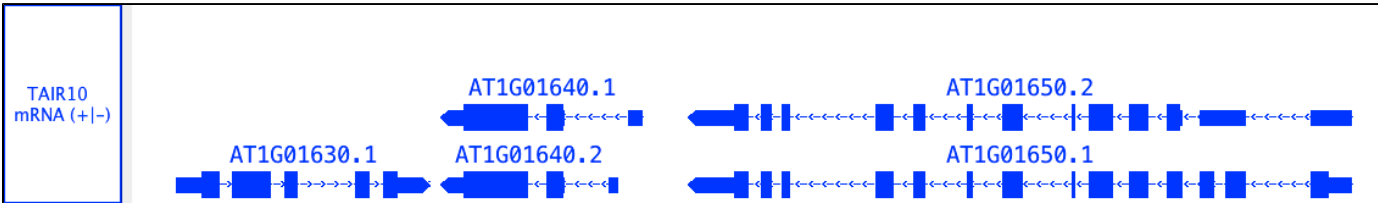


## Annotations

**Annotations** indicate the known or suspected locations of genomic features such as genes, exons, promoter regions, pseudogenes, and so forth. Annotations may consist of a single coordinate, a single span with a start and end positions, or a collection of spans. Most annotations reside on either the plus or minus strand of a chromosome, but some do not.

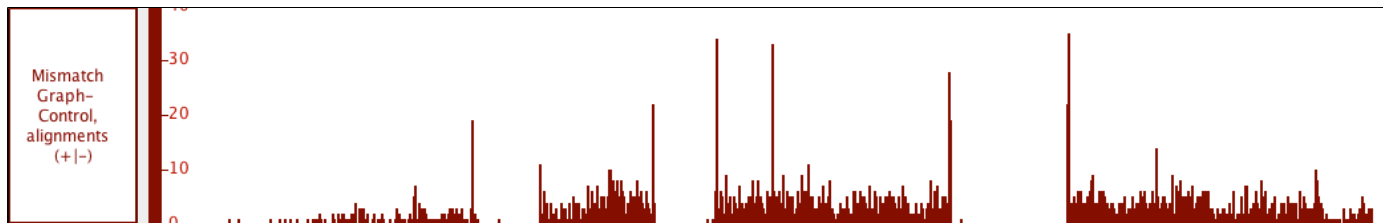
Examples of annotations include:

- **single-coordinate feature**: splice sites
- **single-span feature, with strand**: an exon
- **single-span feature, no strand**: a restriction site
- **multi-span feature, with strand**: a gene model



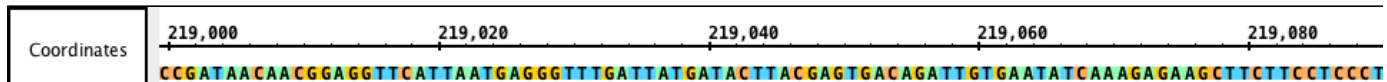
## Graphs

**Graphs** consists of numeric data associated with regions or single-base positions in a chromosome.



## Sequences

**Sequences** are sets of DNA residues comprising a chromosome. Sequences can be fully or partially loaded from local files, QuickLoad sites, or DAS servers. IGB allows you to examine sequence data in a separate window called the Sequence Viewer.



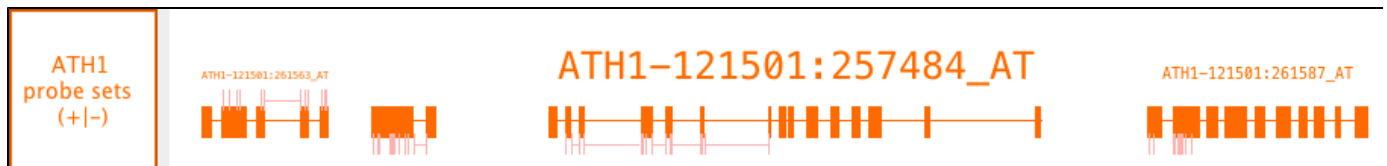
## Alignments

**Alignments** represent how sequences obtained from an experiment (such as reads from an RNA-Seq experiment) align onto the reference genomic sequence. At low zoom they look like annotations, but with marks representing mismatches between the read and the reference genome whenever these data are available. At higher zoom, the aligned sequence bases become visible. File formats for alignments include BAM and PSLX.



## Probe set alignments

**Probe set alignments** consist of a target sequence aligned onto a chromosome and the location of probes relative to the aligned target sequence. This data type is Affymetrix-specific.



## IGB main window

- Introduction
- IGB start screen
- IGB window
- Navigation tools
- Data display area
  - Annotation tracks
  - Zoom stripe

## Introduction

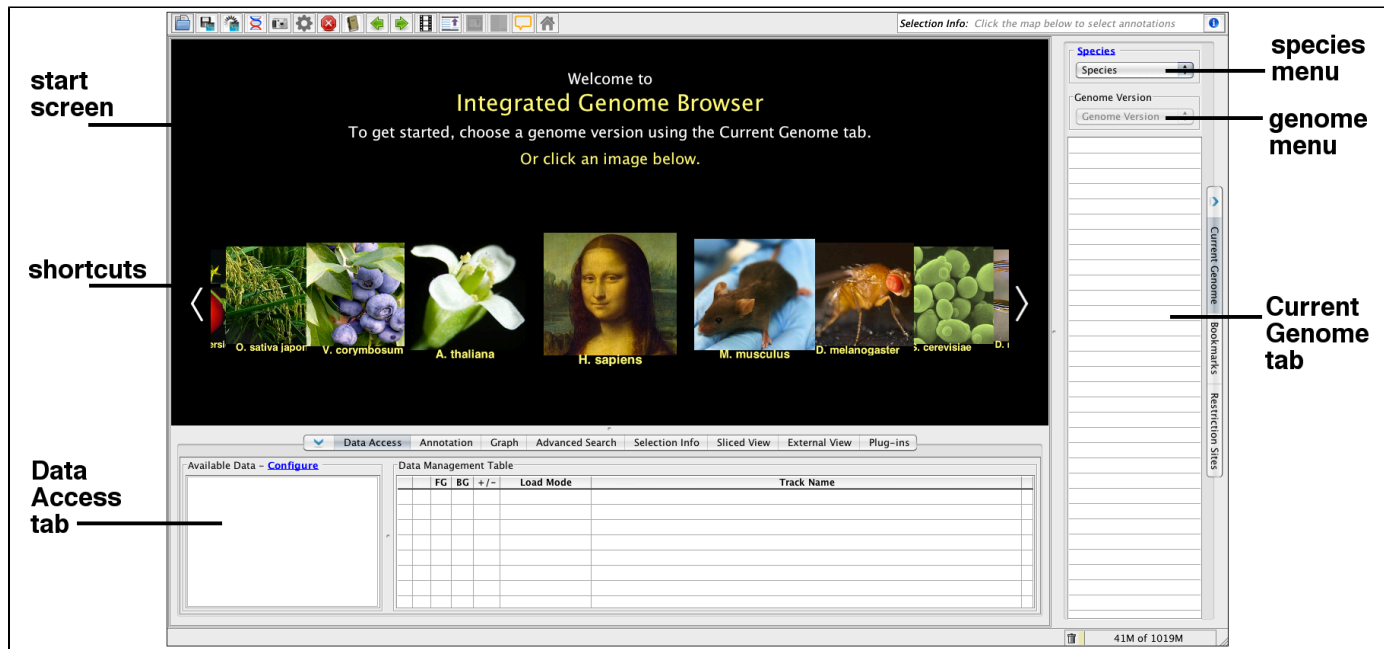
The IGB main window contains tabbed panels, the main data display, toolbar, and navigation tools.

## IGB start screen

When IGB first launches, the IGB window contains

- Start screen with shortcuts to the latest genome versions for popular species
- **Current Genome** tab on the right
- **Data Access** tab on the lower part of the IGB window.

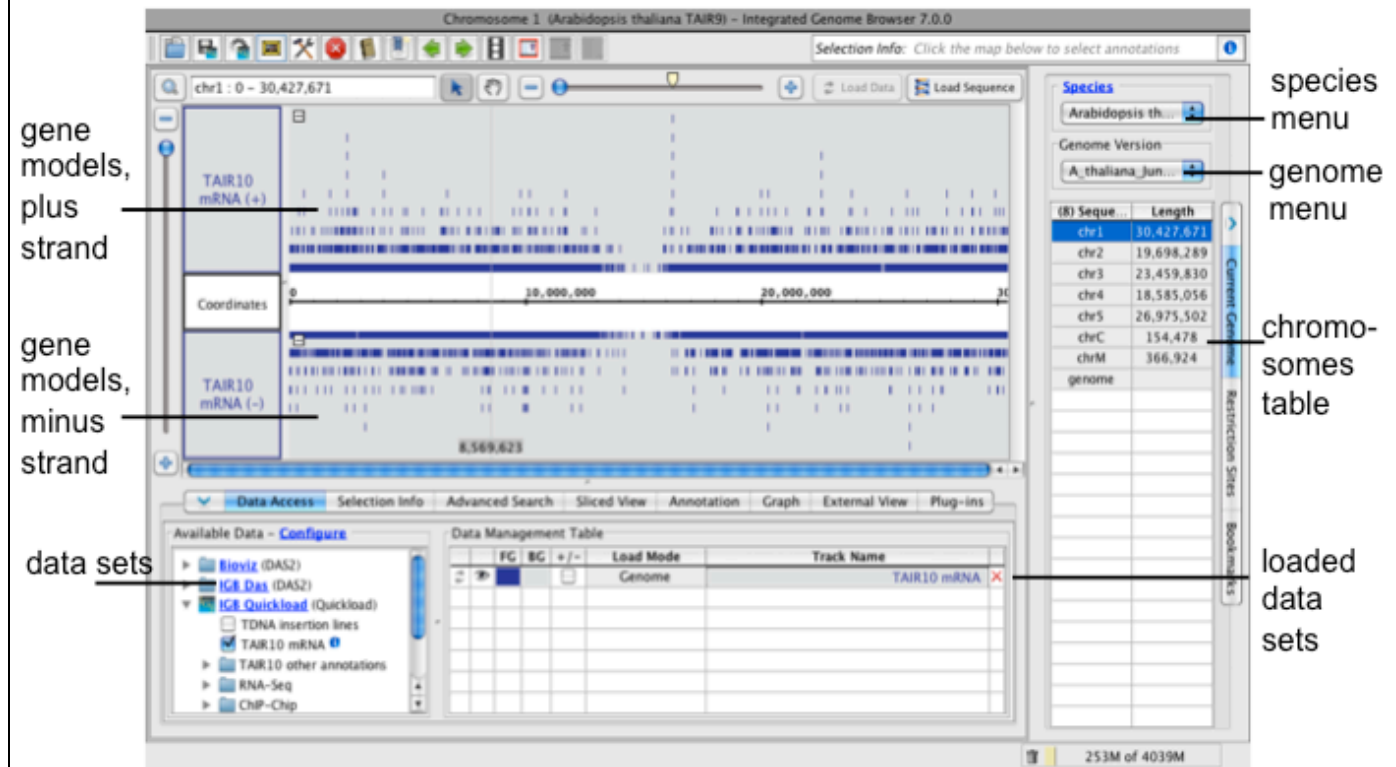
## IGB start screen



When you choose a genome, a view of the first chromosome listed in the **Current Genome** tab replaces the IGB start screen.

If available, reference gene models annotations for the selected genome will load from a remote IGB data source.

## IGB after selecting a genome



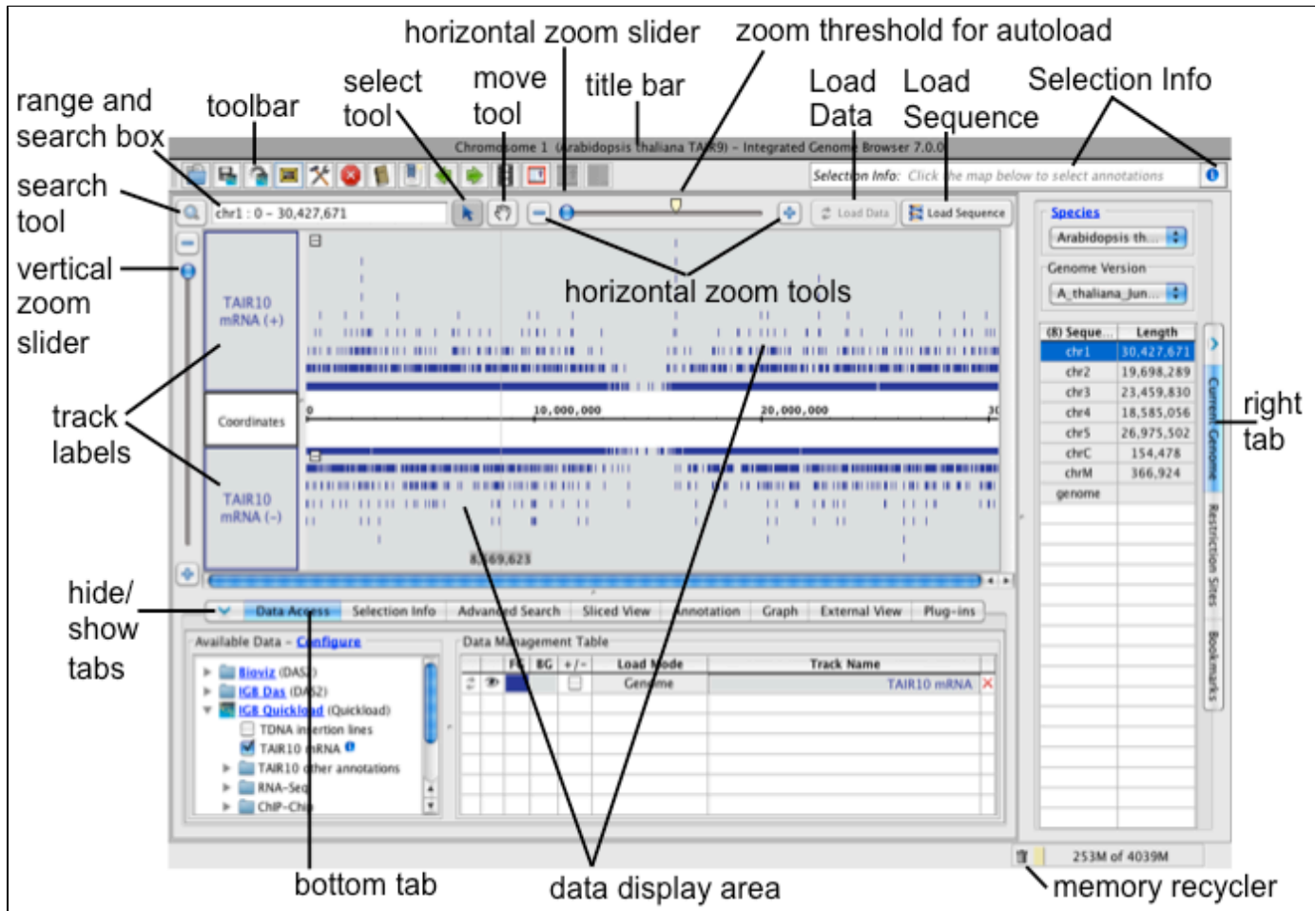
## IGB window

The IGB window contains

- Data display area with central coordinates axis and tracks
- Title bar indicating the chromosome and genome version on display



- Toolbar (new in IGB 7.0) with shortcuts offering one-click access to commonly used functions
- **Selection Info** box which shows information about selected annotations or data sets
- **Load Data** button which triggers loading of data into the viewer. See [Loading data](#) for details
- **Load Sequence** button which loads genomic sequence bases for the region in view
- Navigation tools including zoom sliders, scrollbars



## Navigation tools

The **range box and search box** shows the chromosome and genomic coordinate range currently in view. Use the **range and search box** to search loaded annotations by name or keyword.

Horizontal and vertical zoom sliders control the zoom level. Use the horizontal and vertical zoom sliders to control zoom level. Zooming in shows less data in greater detail and zooming out shows more data in less detail. Use the + and - buttons to zoom incrementally.

The horizontal zoomer also has an **AutoLoad Threshold** mark showing the zoom level at which IGB will automatically load data from data sets whose **Load Mode** setting is to **Auto**. (Details are in [Loading data](#))

Use the **move tool** (hand icon) move the data display field in any direction.

Use the **select tool** select items in the data display field. Click or click-drag over items in the display to select them or click-drag selection tool over the coordinate axis to zoom in. See [Selecting Items](#).

## Data display area

The data display area tracks and the coordinate axis. It is split vertically, with the track labels on the left and data on the right.

Track labels show track names. Right-clicking a track label opens a menu with options to change track appearance, save tracks, and other functions. You can also click and drag track labels to change the vertical ordering of tracks.

The coordinates track displays genomic coordinates and separates minus and plus strand tracks. The coordinates track contains

- A cytobands track with chromosome ideogram, where available

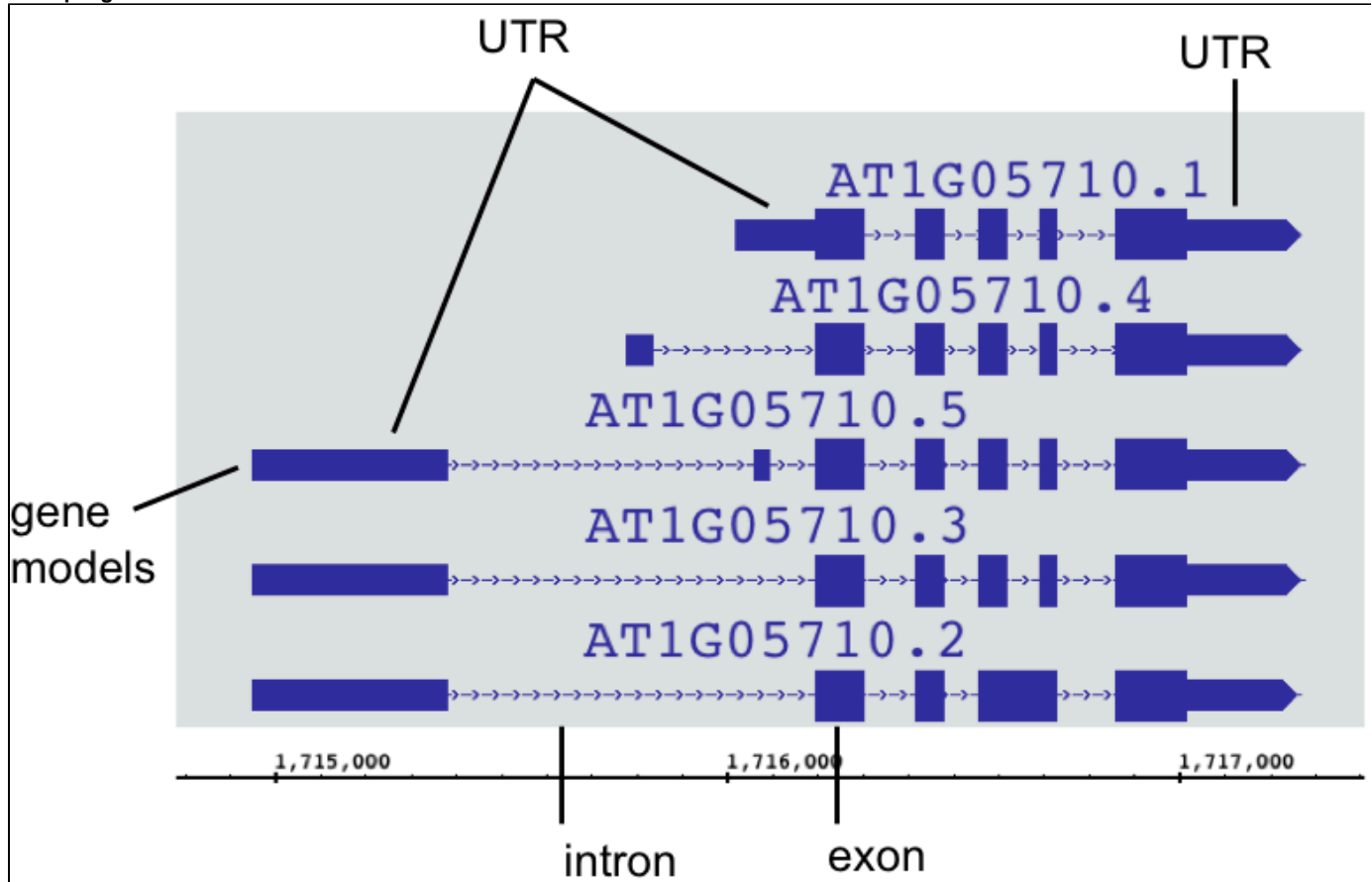


- A number line representing the genomic sequence axis
- The reference genomic sequence, which when loaded appears below the number line.

## Annotation tracks

IGB depicts gene annotations as blocks representing exons linked by lines representing introns. Translated regions (CDSs) appear as taller rectangles, and untranslated regions (UTRs) are indicated by thinner exons, as in the following example.

### Example gene model



When space permits, IGB displays labels next to annotations, depending on settings for the track. To see labels, use the vertical or horizontal zoom sliders to zoom in. Once annotations are large enough, labels appear.

To change the appearance of gene models in a track

1. Select the track by clicking its track label
2. Select the **Annotation** tab
3. Select options under **Style** and **Labels and Strand**

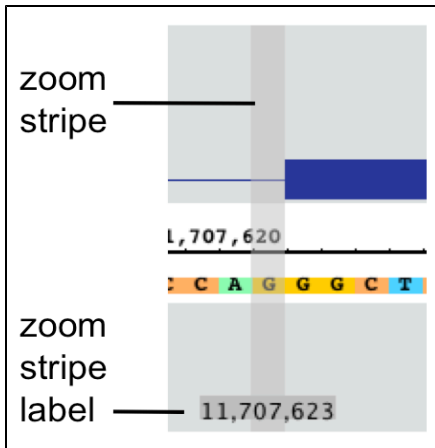


Not every gene model has annotated UTRs. When UTRs are not annotated, all exons are the same height.

## Zoom stripe

The zoom stripe indicates the location of the most recent click. It indicates the focus for zooming, the position in the display that remains stationary when you zoom horizontally.

The zoom strip label indicates the location of the zoom stripe.



## Loading data

IGB can integrate and display data merged from multiple sources in the same view. These includes data sets loaded from your computer, from URL sites, or from various public (and private) DAS, DAS2 and Quickload servers. IGB can also display data from many file types, including:

- Short read alignments data, aka "next generation sequence data" (BAM format)
- Standard annotations such as RefSeq annotations, provided by public repositories
- Alignments of Affymetrix probe sets to the genome, provided by the NetAffx group at Affymetrix (link.psl format).
- Alignments of ESTs or mRNAs produced by blat (psl formats)
- Tiling array graphs, from the TAS program from Affymetrix
- Copy number graphs from the CNAT program Affymetrix
- Data generated from other Affymetrix software tools, such as GCOS, Expression Console and ExACT.
- Annotation and graph files prepared by any method in any of the supported formats.

A list of supported file formats is [here](#).

- [Choose Species and Genome Version](#)

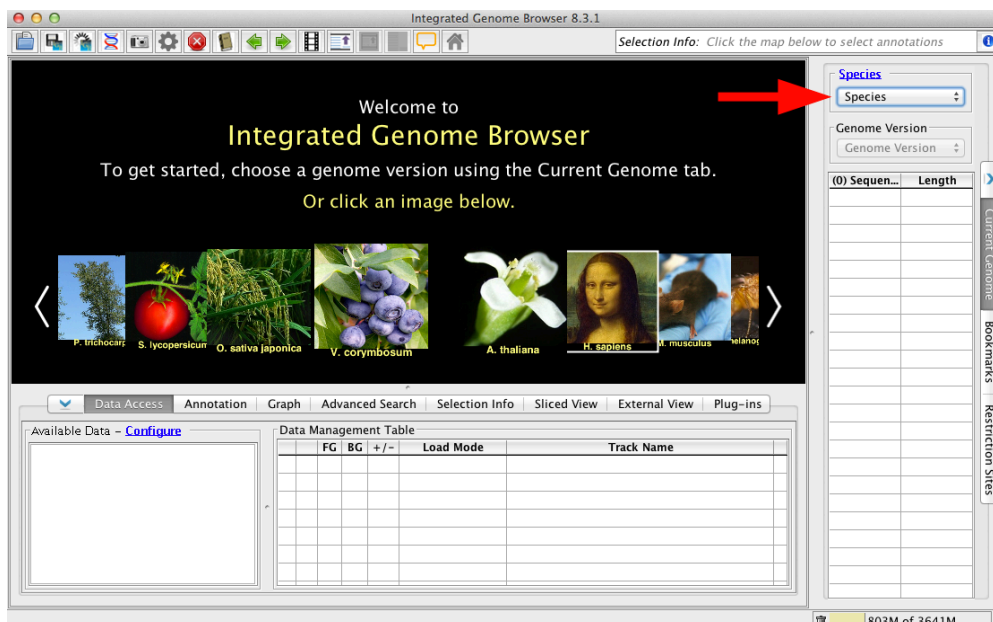
Loading data sets

Data Management Table

Load data to visualize

## Choose Species and Genome Version

The first step to loading data is to choose **Species** and **Genome Version**. IGB uses this information to offer data sources with relevant data sets. To set the species and version, select them in the **Current Genome** panel.



## Loading data sets

There are several ways to get data sets into IGB, from servers/sources, from URLs and from the local computer. To load data from a server, locate the data set in the folders of the **Data Sources** panel. Put a check in the box next to the data you are interested in. This file will be entered into the **Data Management Table**.

For files loaded from URL or from the local computer, just drag and drop into the IGB interface; the file will immediately appear in the **Data Management Table**. Alternatively, use the **File > Open File..** or **Open URL...** to find and load the file(s) you want.



IGB does NOT immediately display loaded files. Many of today's next-gen sequencing files are too big to display all at once. IGB handles this issue by waiting to visualize data sets until you ask it to load data.

While you can immediately refresh to visualize most files, many larger file types, such as BAM and WIG should first have a defined, smaller region selected prior to refreshing the image.

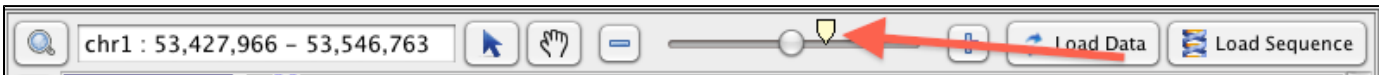
## Data Management Table

IGB is capable of loading and displaying whole genomes, whole chromosomes or just portions of a data file, depending on the file type. Therefore, although there are many load options, not every file type can be loaded the same way. There is always the option of **Don't Load**, which will simply not load the individual file but will maintain it in the list.

- **Don't Load** is the setting for not loading any data, but remaining in the **Data Management Table**
- **Genome** is primarily used with sequence files; this is the default load that IGB uses for the reference sequences for most model organisms (e.g. TAIR10 mRNA)
- **Manual** is the default **Load Mode** for IGB. This setting tells IGB to load only the region in view (e.g. Cool1T1.sm.wig). For large files, such as .bam, .wig or other short read associated files, we strongly recommend that you zoom into a small region of interest, usually about <100Kb depending on your read density, and then **Load Data** to see this region. This is the default setting for most files.
- **Auto** is an option for BAM and SAM files (e.g. Cool1T1.sm.bam), which will allow IGB to automatically load the data in view, if the view is zoomed below the **Auto** threshold, marked with a yellow arrow head (red arrow). The **Auto** zoom level can be adjusted to your needs. First, set the main view to the zoom level you want. Then, **View > Set AutoLoad Threshold to Current View**. This will change the needed zoom level AND will move the yellow arrow to indicate the new threshold setting.

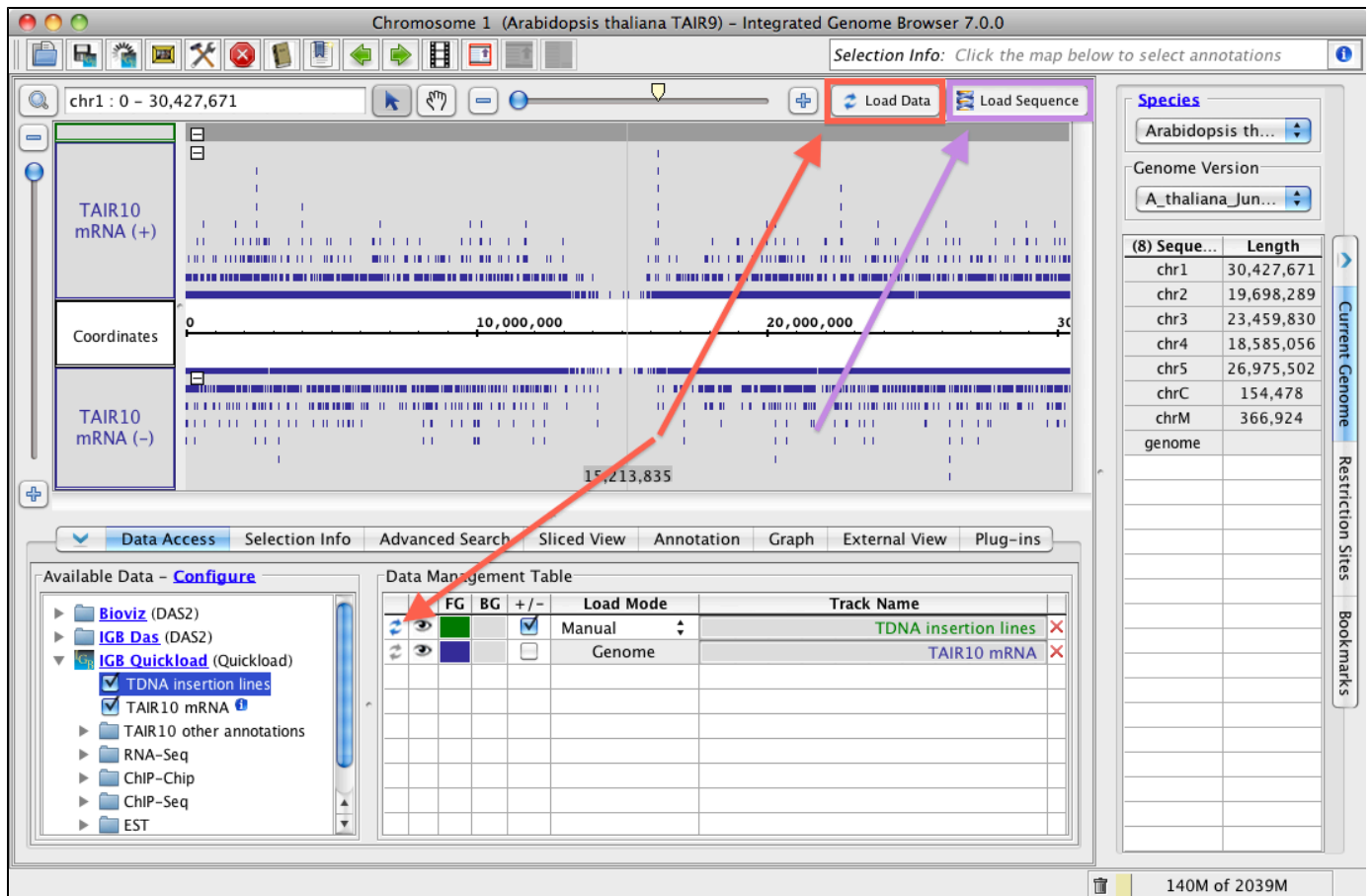


If you set the **Autoload** threshold too high, IGB will try to load all of the indicated data, which could exceed the memory limit.

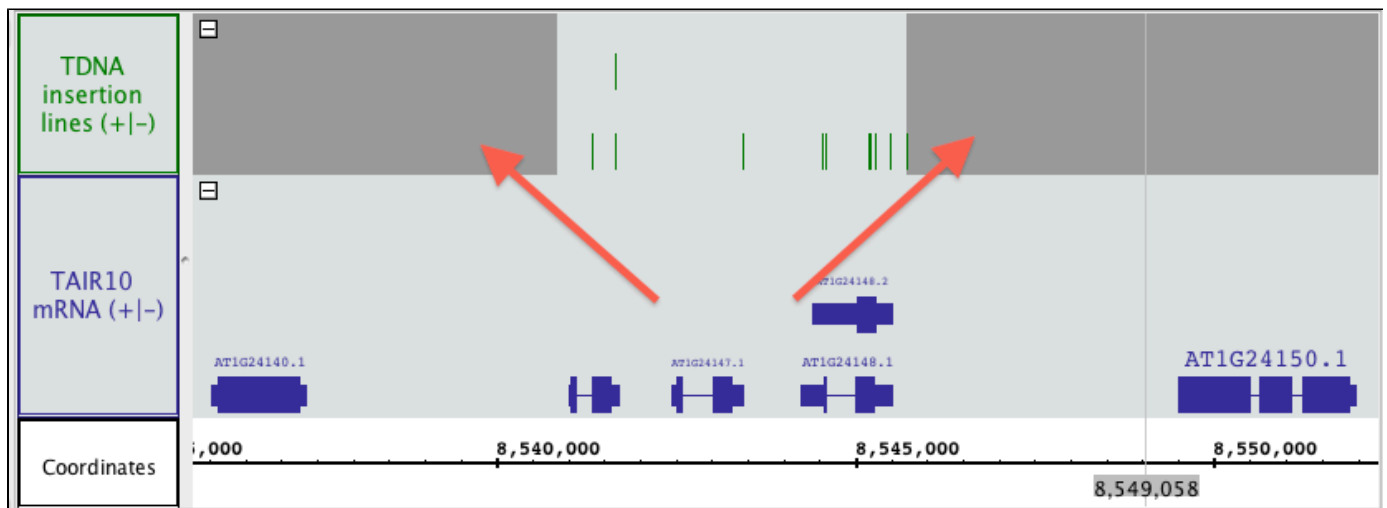


## Load data to visualize

If the data is set to **Load Mode > Manual**, simply click **Load Data**. If you want all the tracks to load data at once, use the **Load Data** button (red box, arrow); to load a single track use the **Load Data** icon in the **Data Management Table** (red arrow). Sequence can be loaded from the **Load Sequence** button (purple box, arrow).



You might load a track and then zoom out to show a greater region of the data sets. IGB will show gray where the track is not loaded yet (red arrows).



- Choose Species and Genome Version


Loading data sets

Data Management Table

Load data to visualize

## File Formats

IGB supports multiple file formats in both compressed and uncompressed formats. See the table below for details and links (when available) to resources describing each format. IGB uses file extensions to recognize file formats, listed below.

 Third-party IGB Apps may support additional formats.

- [Supported file formats](#)
- [About GFF and its variants](#)
- [About bedGraph](#)
- [Partial data loading using tabix indexed files](#)
- [Sequence File Formats](#)
- [.egr and .sin Formats](#)

## Supported file formats

Type	Extension	Description
Affymetrix XML	.axml	A mostly-obsolete XML format used internally at Affymetrix.
BAM	.bam	A binary indexed version of the SAM format used for displaying alignment data. See <a href="#">SAMtools</a> for more details. <b>Note: Be sure you've indexed your BAM file</b> (you should have a .bai file as well). The index file must reside in the same folder as the BAM File unless the location is indicated using the "index" attribute in an IGB Quickload annots.xml configuration file. See
SAM	.sam	Plain text version of BAM format. Index not required. Recommended for smaller files only.
BAR	.bar	Binary graph format developed by Affymetrix. Generated from tiling arrays by TAS (Tiling Analysis Software) from Affymetrix, <a href="#">cisGenome</a> from <a href="#">Hongkai Ji's</a> research group, and others.
BED	.bed	A tabular format developed for the UCSC genome browser. IGB supports four, twelve, and fourteen column BED format. In IGB, the thirteenth and fourteenth columns of fourteen-column BED format (also called BED detail format) are interpreted as title and description, respectively.
BEDGRAPH	.bedgraph	Same as the wiggle format. See below for details.
BigBED	.bigbed	The bigBed format stores annotation items that can either be simple, or a linked collection of exons, much as <a href="#">bed files</a> do. BigBed files are created initially from bed type files, using the program <a href="#">bedToBigBed</a> . The resulting bigBed files are in an indexed binary format. The main advantage of the bigBed files is that only the portions of the files needed to display a particular region are loaded into IGB, so for large data sets bigBed is considerably faster than regular bed files. See <a href="http://genome.ucsc.edu/goldenPath/help/bigBed.html">http://genome.ucsc.edu/goldenPath/help/bigBed.html</a> .
BigWIG	.bigwig	Like the bigBED format, this is an indexed form of a WIG file that facilitates incremental data loading and faster loading than the non-indexed, plain text version of the format. See <a href="http://genome.ucsc.edu/goldenPath/help/bigWig.html">http://genome.ucsc.edu/goldenPath/help/bigWig.html</a> .
BGR	.bgr	Binary graph format developed by Affymetrix.
BNIB	.bnib	Binary format for sequence data originally developed for IGB by Affymetrix to speed up loading sequence data over the network. Replaced by 2bit as of IGB 7.
Cytoband	.cyt	Text format for representing chromosome band (ideogram) data. Examples are available from the <a href="#">IGBQuickLoad.org</a> site under human genome directories.
DAS XML files	.das, .dasxml, .das2xml	XML formats returned from DAS servers. See <a href="http://www.biodas.org">http://www.biodas.org</a> . See <a href="#">DAS/1 specification</a> and <a href="#">DAS/2 specification</a>
Expression Graphs	.egr, .egr.txt, .sin	<a href="#">EGR</a> is a tabular format representing scored genomic intervals. Files generated from Affymetrix GeneChip Operating Software (GCOS) or ExACT (Exon Array Computational Tool) software.
FASTA	.fa, .fasta, .fna, .fsa, .mpfa, .fas	Sequence data in a simple ASCII format. For larger sequence files (e.g., the human genome) use 2Bit. See <a href="#">Sharing data for a custom genome not already part of IGB QuickLoad</a>
GenBank	.gb, .gen	NCBI's file format. IGB has limited support for GenBank files.
GFF (General Feature Format)	.gff, gtf, .gff3	General Feature Format. There are several types of GFF file that use incompatible syntax. The original GFF format is <a href="#">GFF1</a> . A variant called <a href="#">GTF</a> is also used. <a href="#">GFF3</a> has been proposed to extend on GFF and to constrain the specification more tightly to avoid mutually-incompatible versions of GFF. If IGB has difficulty reading your GFF file, make sure the header includes the GFF version, as indicated in the GFF specification documents.

GR	.gr	Tab-delimited graph format. A simple text format containing two columns of numbers separated by a single space or tab. The first number is the base position; the second number is the score. Because this format does not include chromosome names, we recommend you use .sgr or .wig formats instead.
PSL	.psl, .psl3,	PSL is a tabular format used for representing alignments in UCSC's <a href="#">BLAT</a> tool.
Link.psl	.link.psl	Link.psl represents alignments of Affymetrix target sequences and the location of probe set probes within those sequences. Used to display genomic alignments of Affymetrix probe sets. Ann Loraine wrote some python code for creating link.psl files. See <a href="https://bitbucket.org/lorainelab/affyprobesetsforigb">https://bitbucket.org/lorainelab/affyprobesetsforigb</a> and <a href="#">Visualizing probe sets</a>
PSLX	.pslx	PSLX is an extension to the PSL format that includes the aligned sequence. Aligned sequences are displayed similar to BAM files.
Scored Intervals	.sin, .egr, .egr.txt	See <a href="#">EGR</a> .
Scored Map	.map	An outdated format, replaced now by <a href="#">EGR</a> files.
SGR	.sgr	Tab-delimited graph format. Sequence graph files that show base coordinate scores. These files are generated by CNAT (the Affymetrix Chromosome Copy Number Analysis Tool software). The format of .sgr text files is: chromosome identification, then two columns of numbers separated by a single space or a tab. The first number is the base position; the second number is the score
TALLY	.tally	Tally files are created by the bam_tally program (using options -P -B 0), and contain mismatch pileup information. The display is identical to the MisMatch Pileup view mode. The Tally files contain the sequence reference. The plugin will use a tabix index if available.
USeq	.useq	USeq is a binary indexed format used to display graph and annotation data. Supported in IGB 6.2. For more information about it, see: <a href="http://useq.sourceforge.net">http://useq.sourceforge.net</a> .
VCF	.vcf	Variant Call Format (VCF) is a flexible and extendable format for variation data such as single nucleotide variants, insertions/deletions, copy number variants and structural variants. More information on the VCF file format can be found here: <a href="https://github.com/samtools/hts-specs">https://github.com/samtools/hts-specs</a>
Wiggle	.wig	This is a text format for graphical data designed for the UCSC genome browser. IGB supports all 3 subtypes: BED, variableStep, fixedStep. For more information, see the UCSC Web page describing the format: <a href="http://genome.ucsc.edu/goldenPath/help/wiggle.html">http://genome.ucsc.edu/goldenPath/help/wiggle.html</a> . Files in wiggle format can use UCSC track lines to specify colors and other properties.
2Bit	.2bit	2Bit is a compact format for DNA sequences developed by UCSC. See <a href="http://genome.ucsc.edu/FAQ/FAQformat.html#format7">http://genome.ucsc.edu/FAQ/FAQformat.html#format7</a> for more information about it.

## About GFF and its variants

GFF stands for 'general feature format' or 'gene finding format'; it is a tab-delimited file with 9 columns. There are several types of GFF files that use incompatible syntax. The original GFF format is GFF1. A variant called GTF is also used. GFF3 has been proposed to extend on GFF and to constrain the specification more tightly to avoid mutually-incompatible versions of GFF. Some GFF files created by Affymetrix make use of extensions to GFF that are specific to IGB. These are indicated in the file headers by lines beginning with "##IGB-".

IGB can handle most versions of GFF/GTF, but may have difficulty with some rarely-used advanced features. IGB does not read any FASTA data that is included in some GFF3 files. If IGB has difficulty reading your GFF file, make sure that there is a line in the header similar to **##gff-version 2** that identifies the correct format number 1, 2 or 3.

The GFF format is described at [http://www.sanger.ac.uk/Software/formats/GFF/GFF\\_Spec.shtml](http://www.sanger.ac.uk/Software/formats/GFF/GFF_Spec.shtml)

The GTF format is described here <http://genes.cs.wustl.edu/GTF2.html>

The GFF3 format is described here <http://song.sourceforge.net/gff3-jan04.shtml>

## About bedGraph

A bedGraph file associates numerical values (e.g., read coverage) with regions of a genome assembly. Note that the data for an entire genome can reside in a single file. If using this format, you should sort it and index it using tabix for faster and more memory-efficient data loading.

```
track type=bedGraph name="Pollen RNA-Seq coverage"
chr1 0 3722 0
chr1 3722 3797 1
chr1 3797 5890 0
chr1 5890 5893 4
chr1 5893 5897 5
chr1 5897 5902 6
chr1 5902 5910 7
chr1 5910 5939 8
chr1 5939 5944 9
chr1 5944 5957 10
```

Note that the top line of the file contains information (meta-data) about the data set, including its name. When you open the file in IGB, the name will appear as the track label.

## Partial data loading using tabix indexed files

IGB supports partial data loading of several file types using tabix indexing. Supported file types include SAM, BED, BEDGRAPH, PSL, and PSLX.

Indexed files allow for faster searching and loading. The indexed file and its index (.tbi file) must reside in the same folder either on your local computer or on a server. More about tabix can be found [here](#).

## Sequence File Formats

IGB supports fasta and 2bit formats. Older versions of IGB also support an IGB-specific format called bnib. Newer versions of IGB will probably still open bnib files, but as of IGB 7.0, we are no longer including the bnib format in our testing process.

FASTA files contain sequence data in a simple ASCII format. For details, Google search fasta.



We recommend using FASTA for short sequences only. For loading data into IGB and setting up a QuickLoad site, we use the 2bit sequence format.

BNIB is an older format developed for IGB that makes it possible to represent sequence data in a very compact format. 2bit, developed at UCSC is also a compact, binary format for representing sequence data, but a number of open source tools are available for working with this format and so for this reason, IGB now uses 2bit instead of bnib.

## .egr and .sin Formats

EGR (also known as Scored Interval, .sin, format.files) are TAB-delimited files with a header. They can contain one or more scores associated with named annotations or with named or unnamed genomic regions. They have an optional header section which is a list of tag-value pairs, one per line, in the form: **# tag = value** Currently the only tags used by the parser are of the form **score\$*i*** (score name tags are optional). If score name tags are present, then score number **\$*i*** will be named according to the value of the **score\$*i*** tag. If any score name tags are missing, default names will be created.

It is recommended that a tag value pair with the genome version, such as **#genome\_version = H\_sapiens\_May\_2004**, be included to indicate which genome assembly the sequence coordinates are based on. This will ensure that the file is being compared to other annotations from the same assembly.

There are three versions of this format. They can all be described this way, where the parentheses indicate optional elements: (annot\_id) ((seqid) min\_coord max\_coord strand) [score]\*

1. seqid is word string [a-zA-Z\_0-9]+
2. min\_coord is int
3. max\_coord is int
4. strand can be '+', '-', or '.' for "unknown"
5. score is float
6. annot\_id is word string [a-zA-Z\_0-9]+

All lines must have same number of columns. Format 1 has tab-delimited lines with 4 required columns, any additional columns are scores: seqid min\_coord max\_coord strand [score]\* Format 2 has tab-delimited lines with 5 required columns, any additional columns are scores: annot\_id seqid

min\_coord max\_coord strand [score]\* Format 3 has tab-delimited lines with 1 required column, any additional columns are scores:annot\_id [score]\* The IGB parser should be able to distinguish between these, based on combination of number of fields, and presence and position of the strand field. For use in IGB, EGR version 3 is dependent on prior loading of annotations with matching ids.

### Examples

- Format 1:# genome\_version = H\_sapiens\_Apr\_2003

```
1. score0 = A375
2. score1 = FHS
gene1 chr22 14433291 14433388 + 140.642 175.816
gene2 chr22 14433586 14433682 + 52.3838 58.1253
gene3 chr22 14434054 14434140 + 36.2883 40.7145
```

- Format 2:# genome\_version = H\_sapiens\_Apr\_2003

```
1. score0 = A375
2. score1 = FHS
chr22 14433291 14433388 + 140.642 175.816
chr22 14433586 14433682 + 52.3838 58.1253
chr22 14434054 14434140 + 36.2883 40.7145
```

- Format 3:(assumes annotations with the names gene1, gene2, and gene3 are already loaded.)# genome\_version = H\_sapiens\_Apr\_2003

```
1. score0 = A375
2. score1 = FHS
gene1 140.642 175.816
gene2 52.3838 58.1253
gene3 36.2883 40.7145
```

## Zooming and panning the display

### Introduction

To make it easy for you to move through a genome, IGB implements a visualization technique called animated one-dimensional semantic zooming.

This means we provide separate controls for vertical and horizontal zooming - the "one-dimensional" part. "Semantic zooming" means that when you zoom in or out, objects on display change their appearance depending on the zoom level. Also, when you zoom in or out, the position of your last click (indicated by the zoom stripe graphic) stays put while the display stretches or contracts around it.

As a result, zooming in IGB evokes the feeling of effortless flight through a genome; and just as importantly, you can easily adjust the zoom level to match the genes, sequence reads, or other items you're investigating.

Once you've set the zoom level, you can use IGB horizontal and vertical scrollbars to move the display right and left or up and down. IGB also features a move tool that lets you move in any direction.

### Panning the Display

- [Introduction](#)
- [Grab tool](#)
- [Scrollbars](#)
- [Click-drag left or right](#)
- [Arrow toolbar shortcuts](#)
- [Auto Scroll](#)

### Introduction

Moving the IGB display without changing zoom level is called "panning," a term borrowed from film. Panning is also called "scrolling."

When you are zoomed in, only part of your data are visible and much of it is off-screen. The data display area is like the viewfinder of a camera that only shows you part of the surrounding scene. Use IGB's panning controls to move the data display area and view a different part of the genomic data scene.

Because moving through the data scene is an key aspect of data exploration and analysis, IGB offers many ways to pan the display.

To pan the display in IGB, you can

- Click-drag using the **Move** tool
- Use the horizontal or vertical scrollbars



- Click-drag the selection tool right or left
- Click the toolbar panning shortcuts
- Turn on Auto scroll

## Grab tool

The **Move** tool allows you to move in any direction by click-dragging the display.

To use the **Move** tool



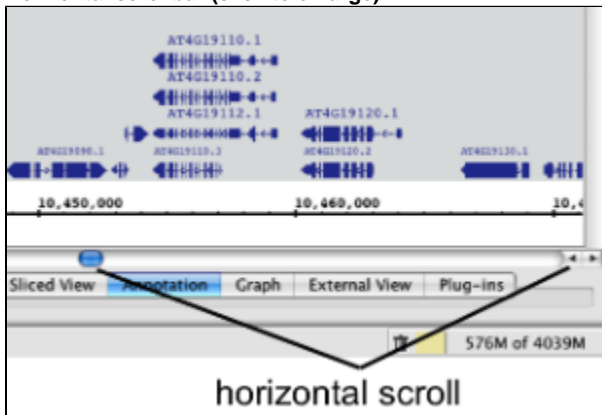
1. Select the **Move** tool
2. Click and drag the display to move in any direction

## Scrollbars

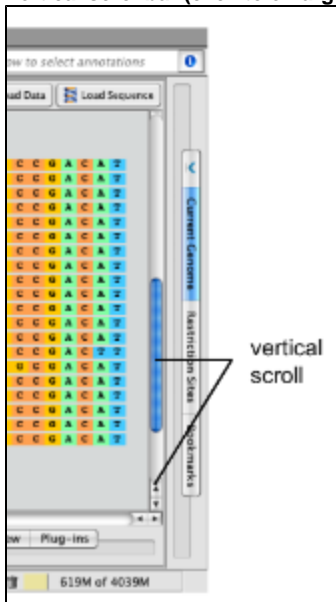
The horizontal and vertical scrollbars move the display up and down, similar to how scrollbars function in other applications.

To pan the display using the scrollbars, click the gutter or click-drag the thumb from left to right. You can also click the ends of the scrollbar.

### Horizontal scrollbar (click to enlarge)



### Vertical scrollbar (click to enlarge)




Horizontal scroll works best when you're zoomed out.

## Click-drag left or right

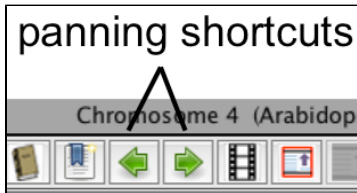
Click-drag the **Selection** tool into the left or right side of the data display area to move left or right.

To move left or right using the **Selection** tool

1. Select the **Selection** tool 
2. Click and hold the mouse in an empty area of the data display area
3. Drag the mouse into the left or right edge of the display
4. Release the mouse to stop scrolling

## Arrow toolbar shortcuts

You can use the right and left arrow shortcuts to pan left or right.



## Auto Scroll

Use Auto Scroll to scan through large regions at medium zoom levels. IGB will move through the image, from left to right, at a speed you define.

To activate autoscroll

- Click the **Autoscroll** icon in the toolbar

or

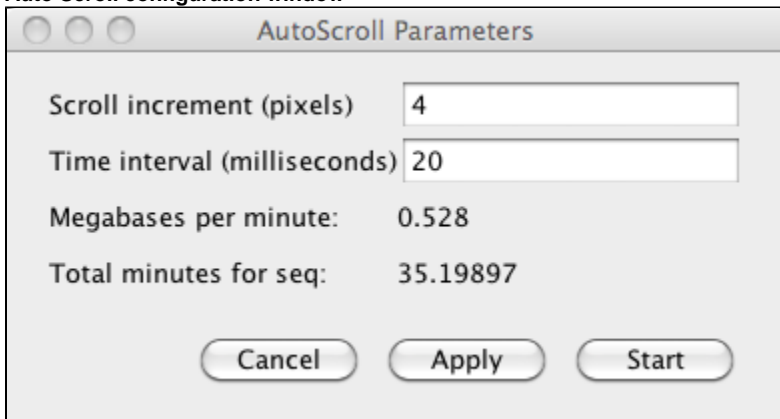
- Select **Tools > Start Auto Scroll**.



To set auto scroll speed

1. Select **Tools > Configure Auto Scroll...**
2. Enter values
3. Click Apply to apply the new speed settings
4. Click Start to start auto scroll

### Auto Scroll configuration window



Scroll increment (pixels)	4
Time interval (milliseconds)	20
Megabases per minute:	0.528
Total minutes for seq:	35.19897

Cancel Apply Start

## Zooming

- [Introduction](#)

- Zoom Stripe
- Horizontal zooming
- Vertical zooming
  - Vertical zooming examples
- Jump zooming

## Introduction

Zooming means: changing the scale of the display. It is essential to be able to change scale quickly when viewing and analyzing genomic data because the scale of what you need to view can vary by many orders of magnitude. You will find that as you work with large genomic data sets, you often need to "zoom" from chromosomes to exons and back again.

To make exploring data easier, IGB implements a visualization technique called animated, one-dimensional semantic zooming. Animated zooming means that during zooming, the scale of the display changes in real time as you operate IGB's vertical and horizontal zoom sliders. Semantic zooming means that the appearance of data objects changes depending on how much area they occupy and the scale of the display. One-dimensional zooming means that you can change horizontal and vertical scaling independently from each other.

When you use the zooming sliders in IGB, the display appears to stretch or contract in real time about a fixed location, called the "zoom focus," which you set by clicking the display. This helps you keep track of your location during zooming. The zoom stripe, which is unique to IGB, marks the horizontal position of the zoom focus and helps you keep track of your location during zooming. It also serves as a pointer at very high zoom levels, making it easy to point out individual bases or compare edges of overlapping objects.

Because changing scale is such a critical element of exploring a genomic data set, IGB provides many tools for zooming, including:

These include:

- Zoom sliders and buttons
- Keyboard and mouse wheel shortcuts
- Double-clicking to zoom in on a feature
- Click and drag the coordinate axis to zoom in on feature
- Searching to zoom in on a gene

### IGB zooming tools

The screenshot shows the IGB interface with several labels pointing to specific features:

- range and search box**: Points to the top search bar.
- horizontal zoom**: Points to the horizontal zoom slider at the top.
- vertical zoom**: Points to the vertical zoom slider on the left.
- autoload threshold**: Points to the icon on the horizontal zoom slider.
- double-click to zoom in**: Points to a gene feature in the main display.
- vertical zoom**: Points to the vertical zoom slider on the right.
- Advanced Search tab**: Points to the 'Advanced Search' tab in the bottom panel.

The bottom panel shows the 'Advanced Search' results for the query 'late lyase family protein' in the 'genome' database. The results table is as follows:

ID	Tier	Gene Name	Description	Start	End	Chromosome	Strand
AT1G09890.1	http://igbquick...	AT1G09890	Rhamnogalacturonate lyase fa...	3217538	3214236	chr1	-
AT1G09910.1	http://igbquick...	AT1G09910	Rhamnogalacturonate lyase fa...	3224541	3220034	chr1	-
AT2G22620.1	http://igbquick...	AT2G22620	Rhamnogalacturonate lyase fa...	9610595	9604771	chr2	-
AT2G22620.2	http://igbquick...	AT2G22620	Rhamnogalacturonate lyase fa...	9610595	9604771	chr2	-
AT4G24430.1	http://igbquick...	AT4G24430	Rhamnogalacturonate lyase fa...	12630017	12633075	chr4	+
AT4G37950.1	http://igbquick...	AT4G37950	Rhamnogalacturonate lyase fa...	17844429	17849274	chr4	+
AT4G38030.1	http://igbquick...	AT4G38030	Rhamnogalacturonate lyase fa...	17866730	17863072	chr4	-

found 8 spans



The autoload threshold icon shows the zoom level at which data sets with **Auto** load mode automatically load into IGB.

## Zoom Stripe

The **Zoom Stripe**, a semi-transparent vertical bar in the image, marks the current focus of zooming, typically the last place you clicked in the display. The zoom stripe label at the bottom of the display indicates its genomic location.

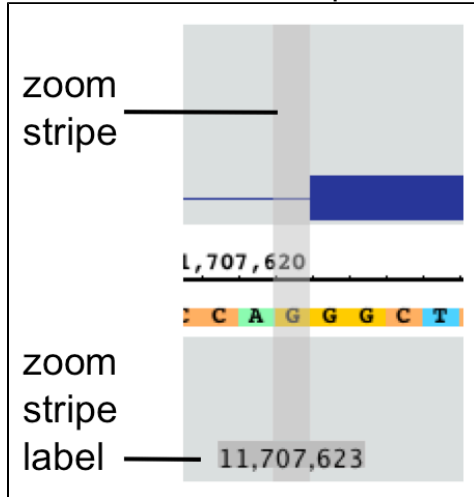
When you zoom horizontally, the zoom stripe remains in place as the display stretches or contracts around it. To set the zoom stripe position

- Click a location in the IGB data display area

To hide or show the zoom stripe or the zoom stripe label

- Select **View > Show Zoom Stripe** or **View > Show Zoom Stripe Label**

Zoomed-in view of the zoom stripe



## Horizontal zooming

Use the horizontal zoom slider to change the scale of the display in real time, or click the "+" and "-" buttons to the right and left of the zoom slider to zoom in small increments.

You can also use keyboard shortcuts for horizontal zooming. Default keyboard shortcuts for zooming include:

- META SHIFT RIGHT zoom in horizontally
- META SHIFT LEFT zoom out horizontally

## Vertical zooming

Use the vertical zoom slider to stretch or compress the display in the vertical dimension. Zooming vertically will expand or compress the vertical dimension of the entire display, with one exception: If you have locked the size of a track, its height will remain the same.

Default keyboard shortcuts for vertical zooming include:

- META SHIFT DOWN zoom in vertically
- META SHIFT RIGHT zoom out vertically

To change keyboard shortcuts, select **File > Preferences > Toolbar**.

You can also click the buttons flanking the vertical zoom slider to zoom in or out.

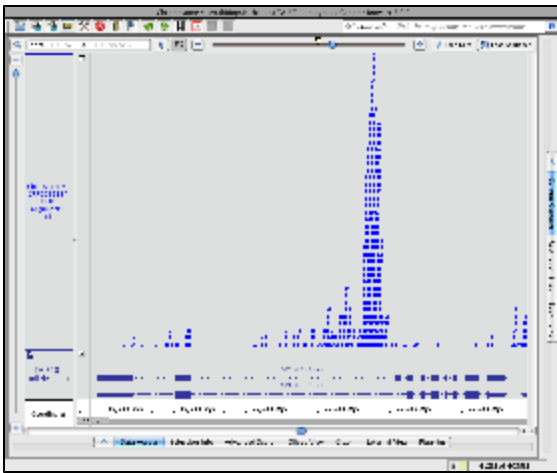
## Vertical zooming examples

The following examples demonstrate how you can use vertical zooming to stretch or compress a track.

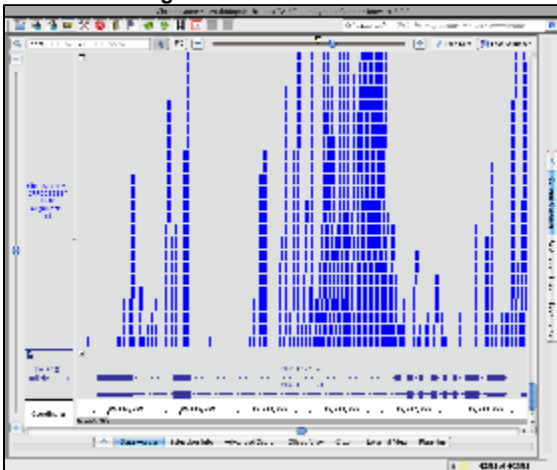
These images show the AG locus of *Arabidopsis thaliana* and a track containing read alignments from a ChIP-Seq experiment investigating binding sites for LFY, a transcription factor that regulates this locus.

In each example, the height of the track containing reference gene models (TAIR10 mRNA) is locked and the vertical zoom settings are different. Note how in the second example, the vertical zoom stretches the LFY IP track but does not affect the locked gene models track. (To hide or show the lock indicator, clear **View > Show Locked Track Icon**.)

### No vertical zoom



### Vertical zooming



### Annotation tab settings for gene models track

Annotation			
Select All	Clear Track	Save Track	Define Track
		Restore Default	Other Options...
<b>Style</b> <input checked="" type="checkbox"/> Foreground <input type="checkbox"/> Background <input checked="" type="checkbox"/> Track Label Track Label Font: 14	Stack Height: 2 <input type="button" value="Apply"/> <input type="button" value="Optimize"/> <input checked="" type="checkbox"/> Lock Track Height (Pixels) Set Height (Pixels): 100 <input type="button" value="Apply"/>	<b>Labels and Strand</b> Label Field: ID Strand: <input checked="" type="checkbox"/> +/- <input checked="" type="checkbox"/> Arrow <input type="checkbox"/> Color by Strand	<b>Operations</b> Single Track: <input type="button" value="Copy"/> <input type="button" value="Apply"/> Multi-Track: <input type="button" value="Apply"/>

## Jump zooming

Often, you'll want to jump to a region or feature.

To jump-zoom to a feature:

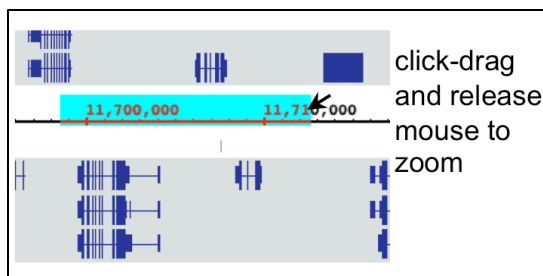
- Double click on it

To jump-zoom to a region:

1. Select the selection tool
2. Click-drag the select tool along the coordinates axis; you should see a blue box indicating the selected region
3. Release the mouse button

The display will "jump" to the region you selected.

### Jump zooming



## Selecting Items

- [Introduction](#)
- [How to select an item \(or items\) in IGB](#)
- [Selecting items activates edge matching](#)

## Introduction

Most objects in the IGB display can be selected by clicking on them. These include annotations, graphs, results from search sequence, and entire tracks.

When an item is selected, IGB draws a red box around it.

IGB also supports selecting and deselecting more than one item at once.

## How to select an item (or items) in IGB



To select items, click the pointer icon to activate the selection tool:

The following are ways to select items in the **Main View** window using the selection tool; selections are outlined in red.

To do this:	Do this:
Select a single annotation or feature.	Click it.
Select a multi-exon gene model	Click a line connecting exons or on the label (if active).
Select a single-exon gene model	Click an exon, right-click and choose <b>Select parent</b> .
Add to the current selection	SHIFT-click to add individual items or shift and drag to add multiple items.
Select multiple items in the display	Click-drag the mouse over all items you wish to select. Start the drag in an empty area. SHIFT-click or SHIFT-drag to add to current selection. <div> <i>Click-dragging selects entire annotations rather than individual parts.</i> </div> <div> <i>IGB reports the number of selected items in the <b>Selection Info</b> box. Use this to count aligned sequences in an alignments track.</i> </div>
Deselect 1 item from a group selection.	Type SHIFT+CNTRL and click the item you want to de-select
Select a sequence region.	Click-drag over the sequence bases you want to select. To select residues, begin dragging in the axis tier.
Deselect everything.	Click an empty region anywhere in the main view.
Select a floating graph.	Click its handle - the rectangle at the left inside the data display area
Select an entire annotation track.	Click its track label in the left-hand panel.

## Selecting items activates edge matching

When you select items in the display, the edges of the selected item and all other items with identical edges are highlighted.

This feature helps in comparing splice variation, exon length, read location, etc.

You can change the edge match color or turn it off. See [Other Options](#).

## Using the mouse to select or move

### Selection and Move modes

You can use the mouse to select items or move the display, depending on which mouse mode you've selected. In **Move** mode, click-dragging on the main data display moves it from side to side, or up and down. In **Selection** mode, clicking or click-dragging on the display selects items, allowing you to find out more information about them, run Google searches, and more.

To set the mouse mode, click the Selection Tool or the Move tool buttons:

- The **Selection** tool 
- The **Move** tool 

To change to **Selection** mode, click the **Selection** tool. You can now select items in IGB by clicking or click-dragging. You can also jump-zoom to a region by click-dragging a region of the sequence axis.

To change to **Move** mode, click the **Move** tool.

### Shortcut

To temporarily switch between mouse modes, click and hold the **ALT** button. When you release the mouse button, you'll return to the currently selected mouse mode.

## Right-click menus

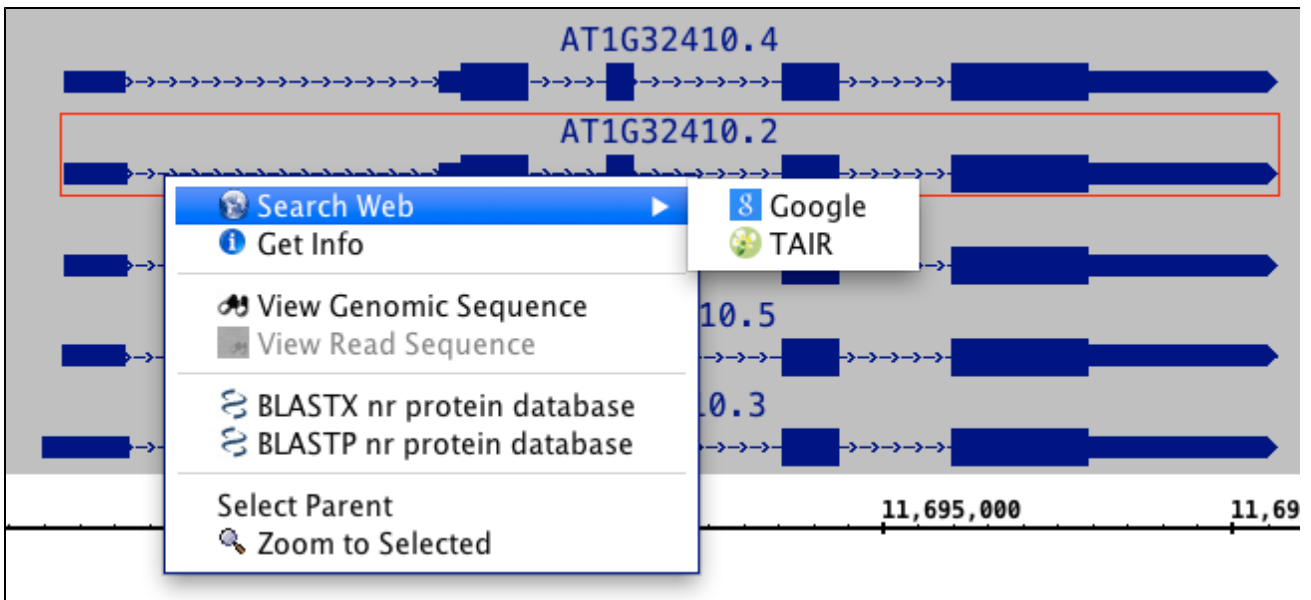
IGB displays context menus when you right-click track labels or items in the data display area. These menus are called "context menus" because the menu options you'll see depend on what you clicked.

- [Right-clicking annotations](#)
- [Right-clicking alignments](#)
- [Track labels](#)
  - [Annotation track right-click menu](#)
  - [Short read \(alignment\) track right-click menu](#)
  - [Graph track right-click menu](#)

## Right-clicking annotations

When you right-click an annotation, you'll see a context menu with the following options:

- **Select Parent** selects the feature of which the right-clicked item is a part. For example, if you right-click an exon and choose **Select Parent**, the entire gene model is selected.
- **Zoom to Selected** zooms to the right-clicked item. Choosing this option is the same as double-clicking.
- **View Genomic Sequence** loads genomic sequence (if not already loaded) and displays it in a new **Sequence Viewer** window.
- **Search Web** opens your web browser to Web pages that have more information about the selected model.
- **Get Info** displays the **Selection Info** panel containing information relevant to your selection.
- **BLASTX nr protein database** searches a non-redundant **protein** database using a **translated nucleotide** query.
- **BLASTP nr protein database** searches a non-redundant **protein** database using a **protein** query.



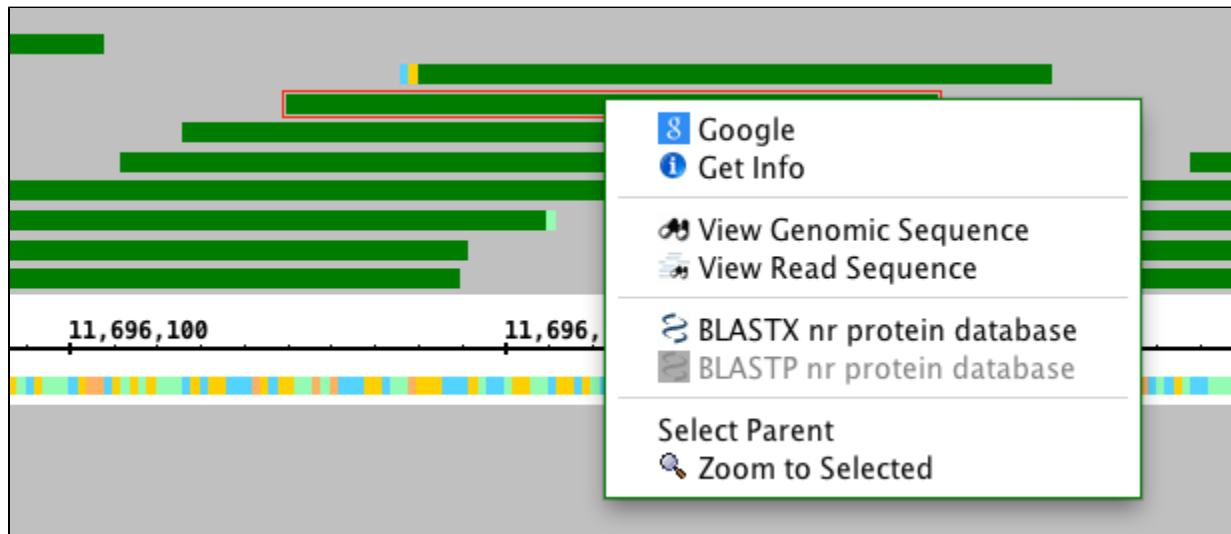
## Right-clicking alignments

When you right-click an alignment, you'll see the same options as in the preceding section plus new options specific to alignments. Context menu items for alignments include:

**View Read Sequence** opens the aligned sequence in the **Sequence Viewer**.

**Restore Read sequence** superimposes the read sequence on the alignment blocks. Ordinarily, if the genomic sequence has been loaded, only mismatched bases are drawn.

For more information about alignments, see [Visualizing read alignments](#).



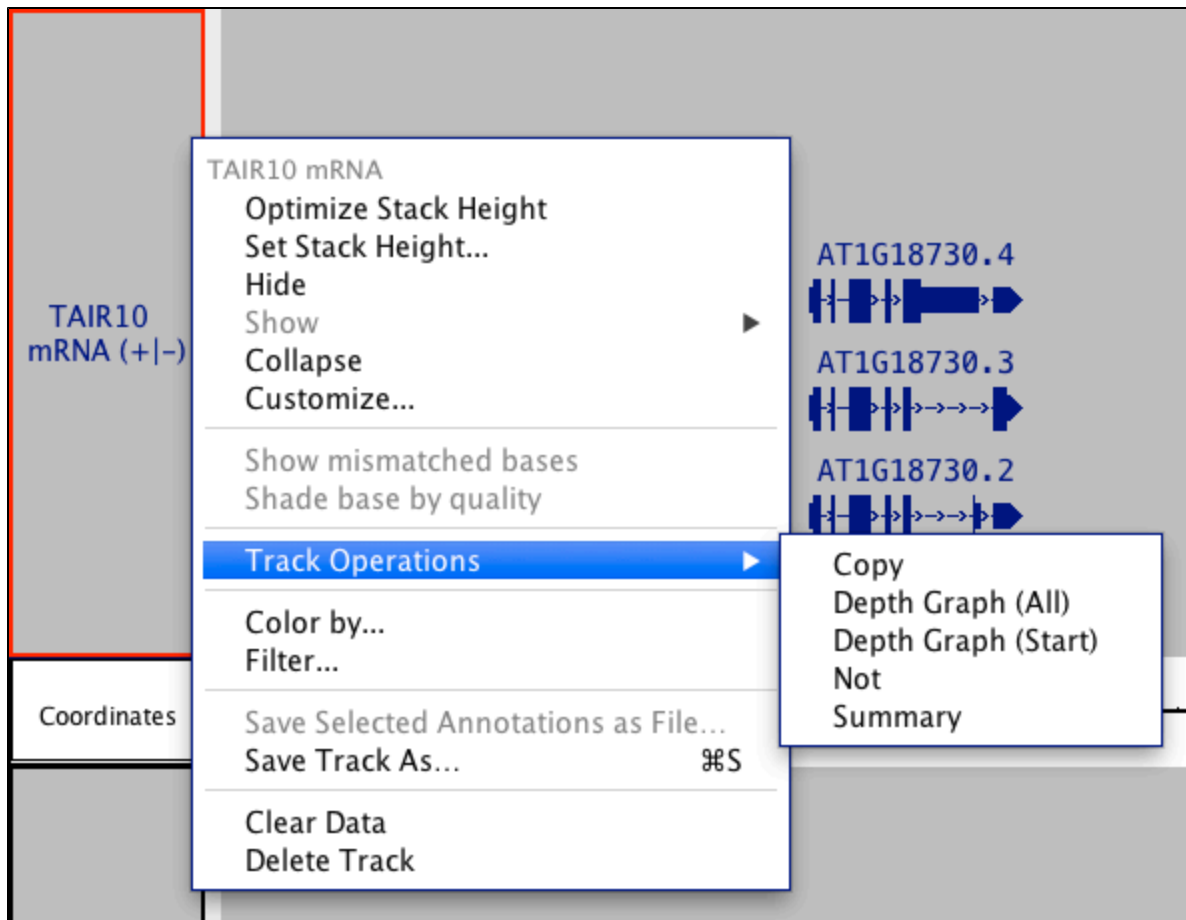
## Track labels

Right click a track label to view options and operations.

The menu you see will vary depending on the track type.

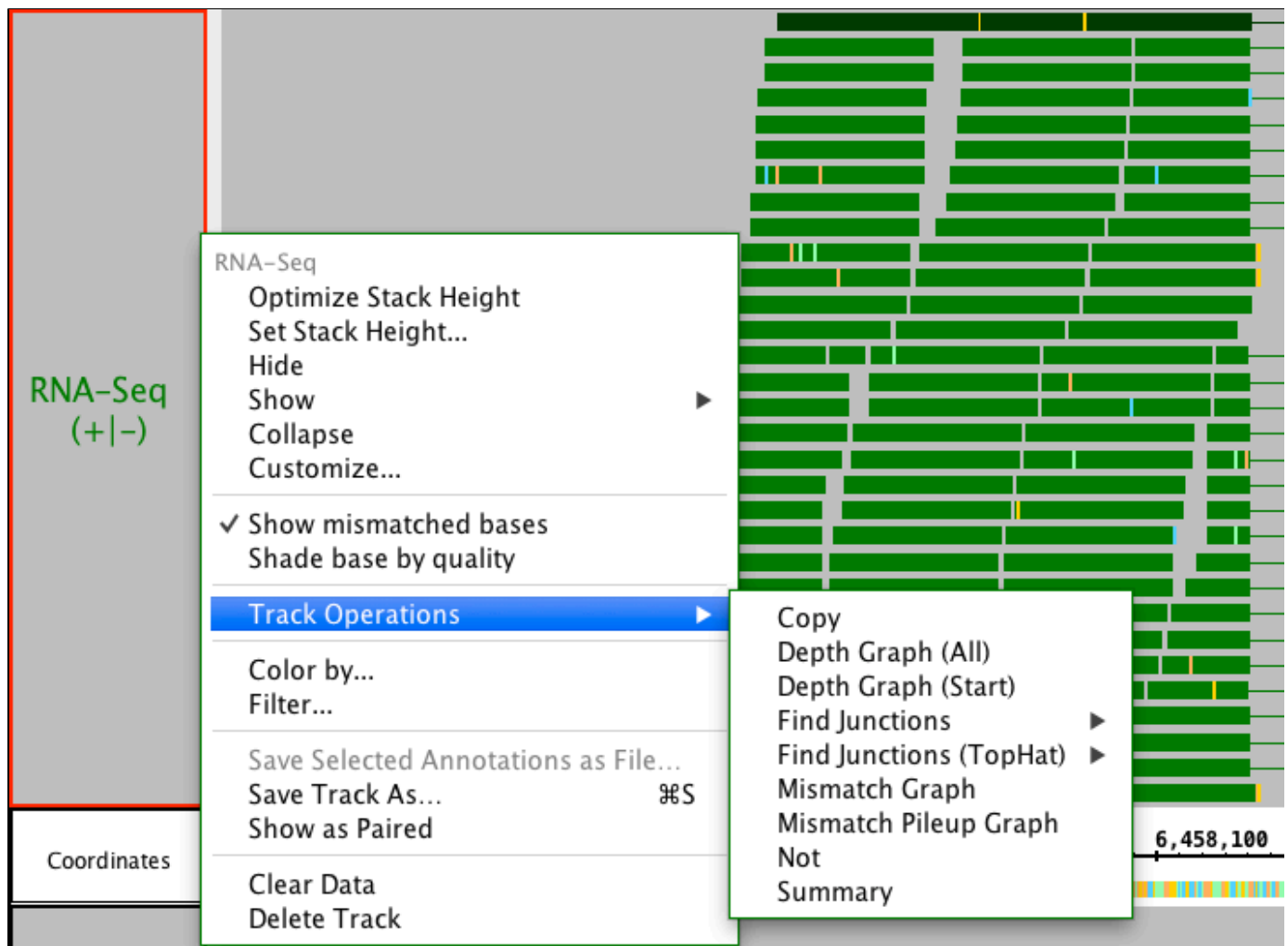
## Annotation track right-click menu



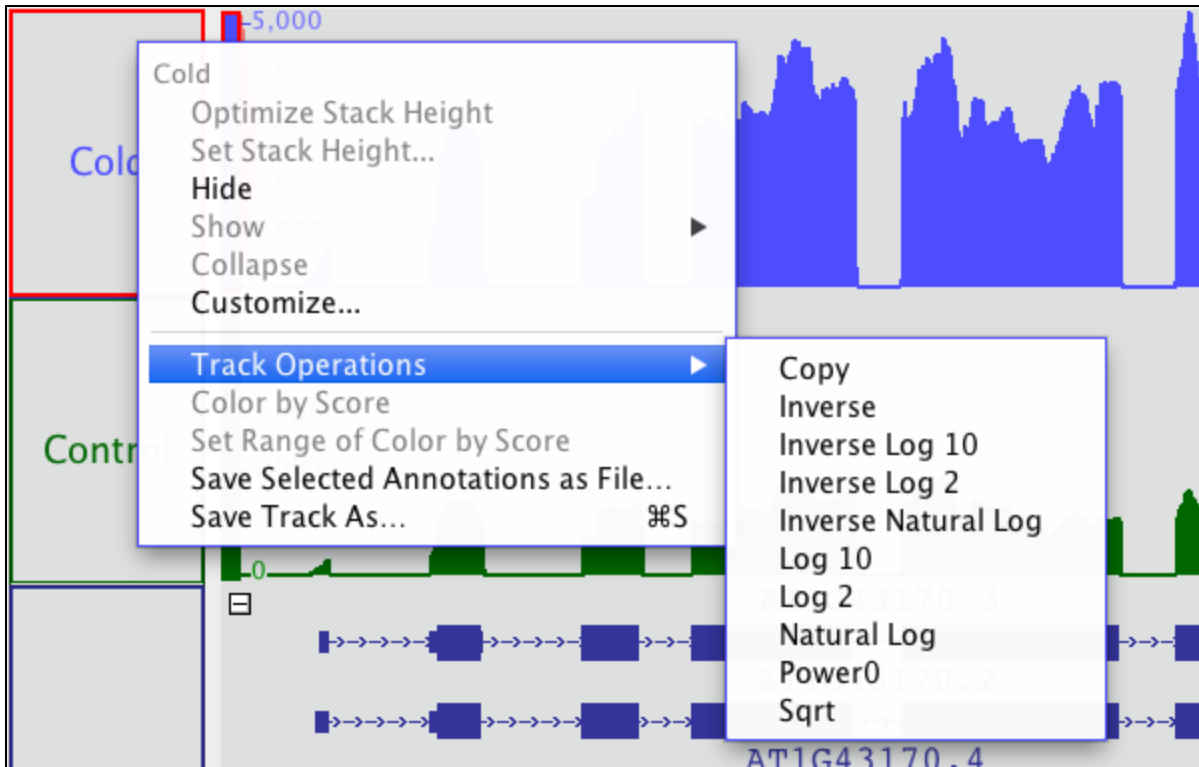


- **Optimize Stack Height** Set stack height so that all available annotations in the currently visible region are shown.
- **Set Stack Height** Open a window that allows you to set the stack height.
- **Hide** Hide the track.
- **Show** Show hidden tracks.
- **Collapse** Compress the track so that all annotations occupy a single row.
- **Customize** Open the **Preferences > Tracks** tab window.
- **Track Operations** Perform operations on a track. See [Annotation track operations](#)
- **Color by Score** Change annotation color depending on score
- **Set Range of Color by Score** Configure color options.
- **Save Selected Annotations as File...** Saves selected annotations to a file
- **Save Track As...** Save all the data in the track as a file.

## Short read (alignment) track right-click menu



Graph track right-click menu



## Tabbed panels

- [Introduction](#)
- [Hide or show tabs](#)
- [Changing tabbed panel locations](#)

### Introduction

IGB tabbed panels provide an easy-to-use way to access multiple functions. Each tab contains information and/or functions involving loading data, exploring data, working with graphs, etc.

By default, the tabbed panels are part of the IGB main window, but you can also display them in separate windows. You can also change their orientation or location within the main IGB display.

### Hide or show tabs

To hide or show tabbed panels

- Click arrow at left end of bottom tabbed panels
- Click arrow at top end of side tabbed panels

To show a tabbed panel

- Click its tab

When you first launch IGB, the **Data Access** and **Current Genome** tabbed panels are open by default.

### Changing tabbed panel locations

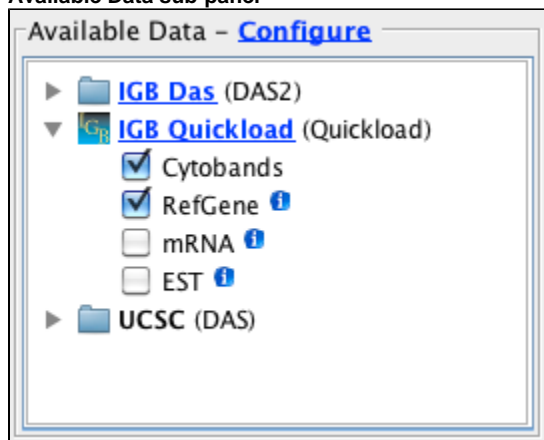
To open all or some tabbed panels in a new window or change their location, select the **Tabs** menu.

**Tabs menu (click to enlarge)**



- Click Data Source links (e.g., IGB QuickLoad) to visit the data source in your Web browser
- Click **Configure** to add or remove data servers (opens the **Data Sources** tab in the **Preferences** window)

#### Available Data sub-panel



### Data Management table

Because many data sets are very large, IGB allows you to control how and when data from a file or data set loads. The Data Management table allows you to set the IGB Load Mode, hide or show tracks, delete tracks, and configure track annotation and background colors, and combined or separate plus and minus tracks. Click in the column labeled **Track Name** to edit a track name.

Possible load modes include:

- **Manual** Data sets load into the viewer only when you click the **Load Data** button
- **Genome** Data sets load all at once
- **Auto** Data sets load only when the zoom level exceeds the threshold indicated by the autoloading threshold marker on the horizontal zoomer

		FG	BG	+ / -	Load Mode	Track Name	
				<input type="checkbox"/>	Genome	RefGene	
				<input type="checkbox"/>	Genome	Cytobands	



You can halt data loading by clicking the stop loading icon on the bottom left of the IGB window

## Selection Info

- [Introduction](#)
- [View information about selected items](#)
- [View or copy table contents](#)

## Introduction

The **Selection Info** tab shows information about selected items in the display, including tracks and features within tracks. Information is shown in a table where each column in the table represents a selected item.

## View information about selected items

To use the tab

- Select one or more items in the main window.
- Click the Selection Info tab

#### Selection Info tab with four items selected

The screenshot shows the Integrated Genome Browser 7.0.0 interface. The top bar indicates 'Chromosome 1 (Homo sapiens (Feb 2009) human being (GRCh37/hg19)) - Integrated Genome Browser 7.0.0'. The main window displays a genomic track for chr1: 155,147,212 - 155,173,793. Four items are selected, highlighted with red boxes: NR\_046328, NM\_001044392, NR\_030281, and NR\_045553. The bottom panel shows the 'Selection Info' tab with a table of properties for these four items.

property	NR_046328	NM_001044392	NR_030281	NR_045553
gene name	TRIM46	MUC1	MIR92B	THBS3
description	tripartite motif containing...	mucin 1, cell surface asso...	microRNA 92b	thrombospondin 3
name	NR_046328	NM_001044392	NR_030281	NR_045553
id	NR_046328	NM_001044392	NR_030281	NR_045553
chromosome	chr1	chr1	chr1	chr1
start	155,146,262	155,162,706	155,164,967	155,177,772
end	155,157,447	155,158,299	155,165,063	155,165,378
length	11,185	4,407	96	12,394
strand	+	-	+	-
cds min		155158610		
cds max		155162634		
score	0.0	0.0	0.0	0.0
forward	true	false	true	false

## View or copy table contents

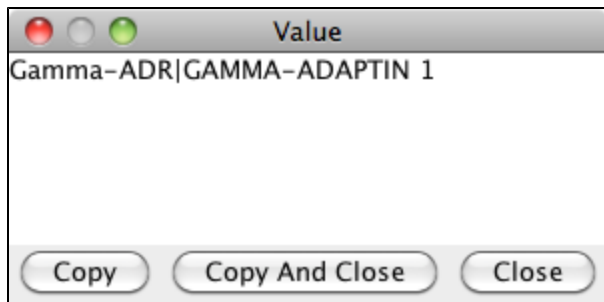
To view and copy the contents of individual cells

- Click the '...' icon

This will open a new window showing the cell contents.

To copy data from the cell

- Click button labeled **Copy** or **Copy and Close**



## Advanced Search

- [Introduction](#)
- [Search types](#)
  - [Search by ID or Name or Keyword](#)
  - [Searching sequence](#)
  - [Searching for multiple residues simultaneously](#)
- [Regular expression and wild cards](#)

## Introduction

Use the Advanced Search tab to search sequence residues or annotations. Both search types support regular expressions and wild card characters.

Using advanced search you can

- look up genes or other annotations by keyword, name, or id
- find for instances of a transcription factor binding sites
- display locations of PCR primers

Search results will appear in the **Advanced Search** results table. Double-click a row in the table to view the result in the main IGB window.

If you search for sequence, IGB will also display color-coded bars in the coordinates track indicating the matched sequence.

## Search types

The Advanced Search supports:

- **ID or name** find features by name
- **Keyword** find features by keyword
- **Residues** find patterns in genomic sequence

### Search by ID or Name or Keyword

To find a feature or annotation by ID or name

1. Select **IDs or names** from the **Search** menu.
2. Choose "genome" or a chromosome to search from the **in** menu.
3. Enter ID or gene name you want to find (**for** textbox).
4. Press <Enter> key or click **Search** button

Only data already loaded into the IGB viewer will be searched unless **also search remotely** is checked.

**IDs or Names** search will search IDs and names of annotations.

**Keyword** search, similar to ID or name search, will search annotation names, but it will also search other information associated with annotations, such as descriptions or other attributes. Choose **Keyword** to search descriptions, names, ids, and other attributes.

Results will appear in the Search Results table. Double-click a row in the table to zoom to that feature.

**Advanced Search tabbed panel after searching the human genome**

Search for annotations or sequence in H\_sapiens\_Feb\_2009

Search  in  for

☒ also search remotely (1 server)

ID	Tier	Gene Name	Description	Start	End	Chromosome	Strand
NR_033387	http://igbquickload...	ADARB2-AS1	ADARB2 antisense RNA 1	1568824	1599179	chr10	+
NR_036460	http://igbquickload...	ADAT1	adenosine deaminase, tRNA-specific 1	75657221	75632246	chr16	-
NR_036636	http://igbquickload...	ADAM1A	ADAM metalloproteinase domain 1A, pseud...	112336866	112339706	chr12	+
NR_037933	http://igbquickload...	ADAM20P1	ADAM metalloproteinase domain 20 pseud...	70950473	70935595	chr14	-
NR_038264	http://igbquickload...	ADAMTS9-AS2	ADAMTS9 antisense RNA 2	64670545	64997143	chr3	+
NR_048577	http://igbquickload...	ADAM15	ADAM metalloproteinase domain 15	155023747	155035252	chr1	+
NR_048578	http://igbquickload...	ADAM15	ADAM metalloproteinase domain 15	155023747	155035252	chr1	+
NR_048579	http://igbquickload...	ADAM15	ADAM metalloproteinase domain 15	155023747	155035252	chr1	+

Search for \*ada\* on genome: 141 local matches, 1936 remote matches

## Searching sequence

To find all instances of a sequence or regular expression

1. Select **Residues** from the **Search** menu
2. Choose chromosome to search from **in** menu.
3. Enter sequence you want to find (**for** textbox).
4. Check **overlay results** to preserve results from previous searches
5. Press <Enter> key or click **Search** button

IGB displays matches in the results table and as colored bars underneath the coordinates axis. Matches on the minus strand appear in a slightly lower position.

### Residue search results.

Chromosome 1 (Arabidopsis thaliana TAIR9) - Integrated Genome Browser 7.0.0

Selection Info: Click the map below to select annotations

chr1 : 9,414,695 - 9,415,016

TAIR10 mRNA (+/-)

Coordinates

9,414,700 9,414,750 9,414,800 9,414,850 9,414,900 9,414,950 9,415,000

AT1G27110.2

AT1G27110.1

AT1G27110.3

Search for annotations or sequence in A\_thaliana\_Jun\_2009

Search  in  for

☒ Overlay results

Pattern	Color	Start	End	Strand	Chromosome	Matched sequence
AGCTTCTCTCAAATGTCCAA		9414719	9414739	+	chr1	AGCTTCTCTCAAATGTCCAA
GGCCAAGTTTATGTTAATG		9414766	9414785	-	chr1	GGCCAAGTTTATGTTAATG
AAACATGATCTTTGTTAT		9414854	9414872	-	chr1	AAACATGATCTTTGTTAT
ATGAGGTATCCATCA		9414884	9414899	+	chr1	ATGAGGTATCCATCA
TGTCGACCCAT		9414951	9414962	+	chr1	TGTCGACCCAT
TGCTATCTTGAAGAGAG		9414982	9415000	-	chr1	TGCTATCTTGAAGAGAG

Found : 0 forward and 1 reverse strand hits. Click row to view hit.

213M of 2039M

## Searching for multiple residues simultaneously

To search for multiple residues simultaneously enter residues in advanced search box separated with the pipe (|) symbol.

Example: atgttc|atggc

This will return a search for atgttc and atggc separately.

## Regular expression and wild cards

IGB searching supports regular expressions and wild cards. This is especially useful when searching for sequence motifs, such as transcription



factor binding sites.

Example queries:

Pattern	Represents	Example	Finds
.	any single nucleotide	ACCT.T	ACCTTT, ACCTAT, ACCTGT, and ACCTCT (4 possibilities)
..	any two nucleotides	ACCT..T	ACCTAAT, ACCTATT, ACTAGT, Etc. (4 x 4 possibilities)
[CG]	a C or a G	ACCT[CG]TC	ACCTCTC and ACCTGTC
T{1,n}	1 to n T's	ACGGT{1,3}C	ACGGTC, ACGGTTC, ACGGTTTC
T*	Zero or more T's	ACGGT*C	ACGGC, ACGGTTTTTC, Etc.
.*?	a string of any length containing any nucleotides	TCGGGGTTAA.*?CTGGACTC	Many possibilities.
.*	the longest possible string of any length containing any nucleotides	TCGGGGTTAA.*CTGGACTC	Differs from the search above in that the longest possible result(s) will be found
(?i)	use case-insensitive matching	(?i)A[CG]T	ACT, agt, acT, AGt, etc.

More information about regular expressions is available from <http://docs.oracle.com/javase/7/docs/index.html>.

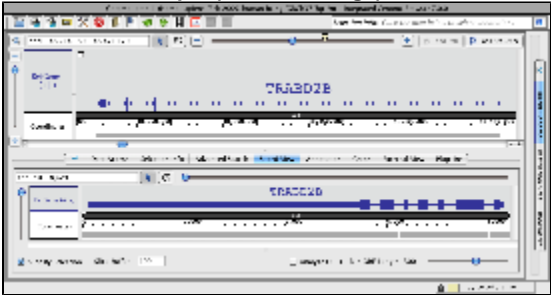
## Sliced View

- [Introduction](#)
- [Display sliced view of a gene](#)
- [Sliced View options](#)
  - [Slice by Selection](#)
  - [Slice Buffer](#)
  - [Analyze ORFs](#)
  - [Min ORF Length](#)
- [Viewing deletions and insertions](#)
  - [Insertion example](#)
  - [Deletion example](#)
- [Showing ORFs and stop codons](#)

## Introduction

The **Sliced View** tab displays an alternative view of items from in the Main View. In the **Sliced View**, introns are shown as having the same size, set in the **Slice Buffer** box). The **Sliced View** also can show open reading frames (ORFs).

**Sliced view example (click to enlarge)**



## Display sliced view of a gene

To view a sliced view of a gene

- Click the Sliced View tab to make it the active bottom tab
- Open the Sliced View tab in a separate window (this is optional) **Tab**s > **Open Current Bottom Tab in New Window**
- Click a gene model in the Main View
- Observe that the same gene model appears in the **Sliced View** but intronic regions are removed

## Sliced View options

The bottom row of the **Sliced View** panel shows options for controlling **Sliced View** appearance.

**Sliced View options (click to enlarge)**



## Slice by Selection

This option controls when the **Sliced View** updates. Check this option to ensure that whatever is selected in the Main View is shown in the **Sliced View** display.

## Slice Buffer

The **Slice Buffer** setting controls the size of introns drawn in the **Sliced View** display. The "buffer" consists of intronic sequence on the five and three prime ends of exons.

## Analyze ORFs

Check this option to activate display of ORFs (open reading frames) underneath the **Sliced View** coordinate axis track.

## Min ORF Length

This option controls the size of ORFs shown when Analyze ORFs is checked.

## Viewing deletions and insertions

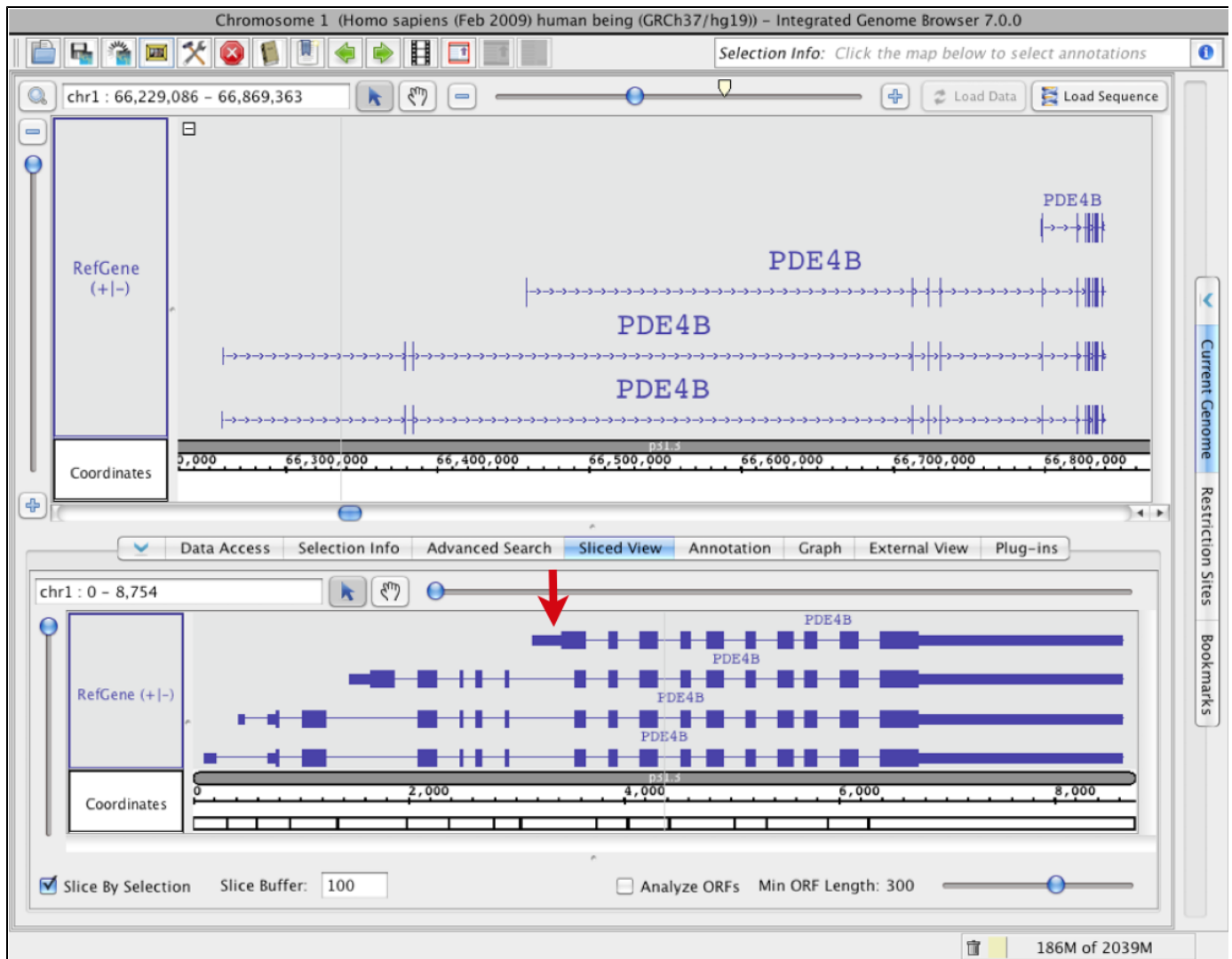
The **Sliced View** can help you see where sequences present in one transcript are absent from the others due to alternative splicing.

To examine alternate splicing using the **Sliced View**

- select one or more transcripts in the Main View
- deletion icons (X characters) IGB draws on the other transcripts
- insertions are shown as larger introns

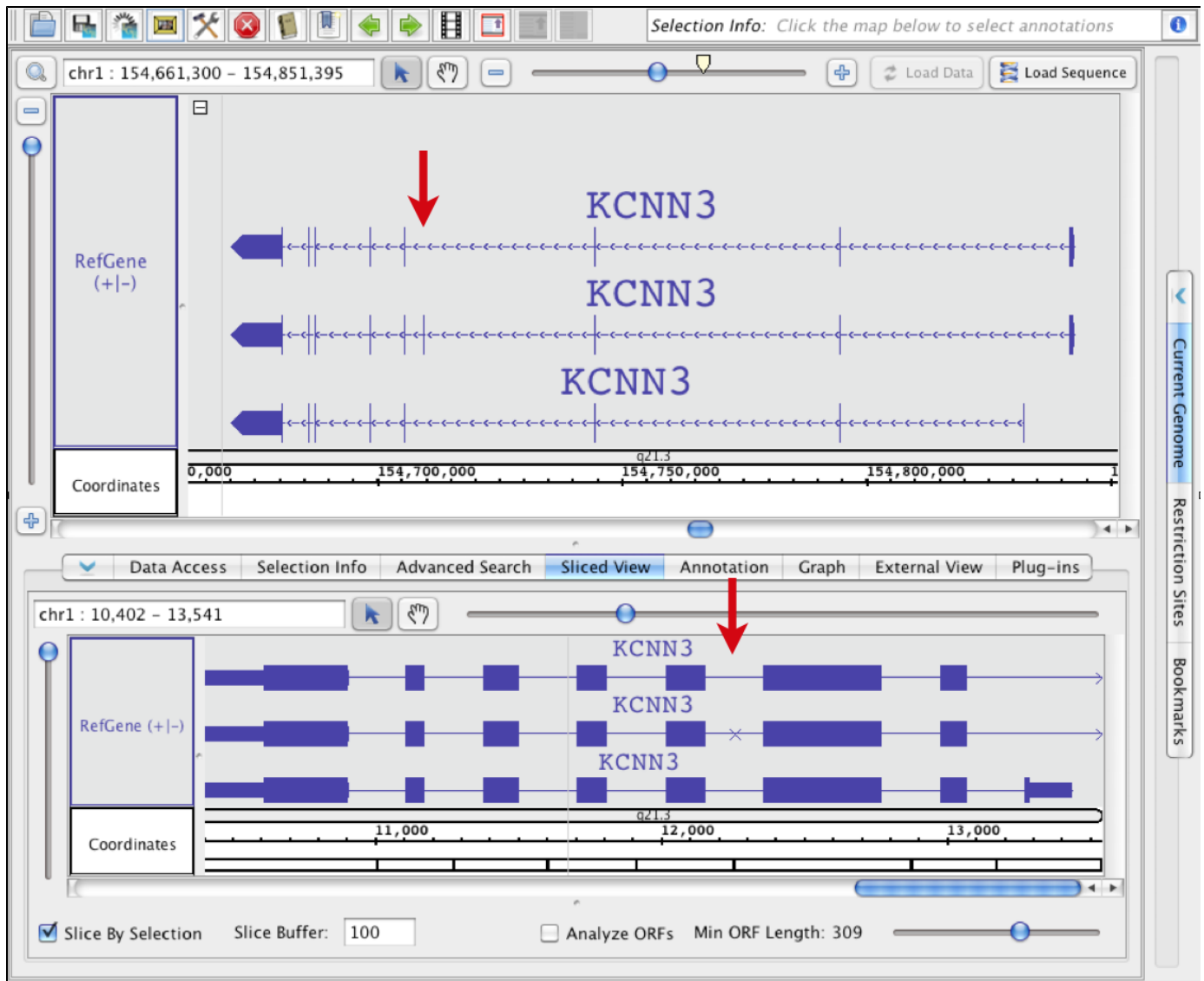
## Insertion example

In this example, all four gene models were selected in the main view when the sliced view was created. In the sliced view, an arrow points to an exon that is present in all four models but is larger in the first one. Introns overlapping this exon in the other three models are larger.



## Deletion example

In this example, the bottom gene was selected in the main view when the sliced view was created. An arrow points to an exon in the main view that is present in the middle gene model but absent in the other two. Since the third model which doesn't include this exon was used to create the sliced view, the skipped exon is indicated using an X character in the **Sliced View** tab.



## Showing ORFs and stop codons

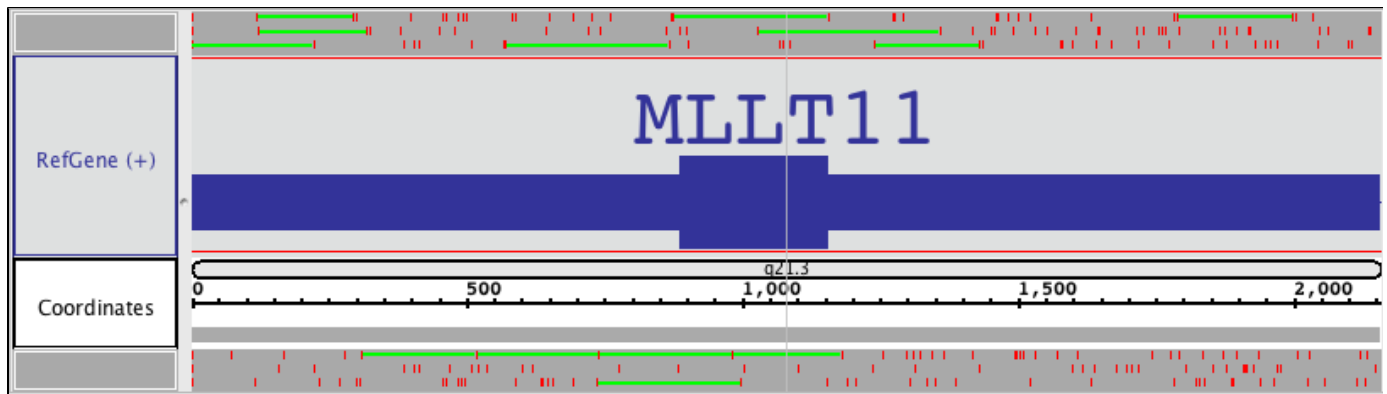
The **Sliced View** panel can show open reading frames and stop codons.

To visualize open reading frames and stop codons

- Select an annotation in the Main View
- Click the **Sliced View** tab
- Check the **Analyze ORFs** checkbox
- Use the Min ORF slider to adjust the minimum length of ORFs to show (in base pairs)

ORFs will appear for all six translation frames. Stop codons are shown as red marks. ORFs that are as long as or longer than the minimum ORF setting will be shown as green lines. Stop codon and ORF colors can be set using **Preferences > Other Options**.

**Sliced View ORF display.**



## Annotation

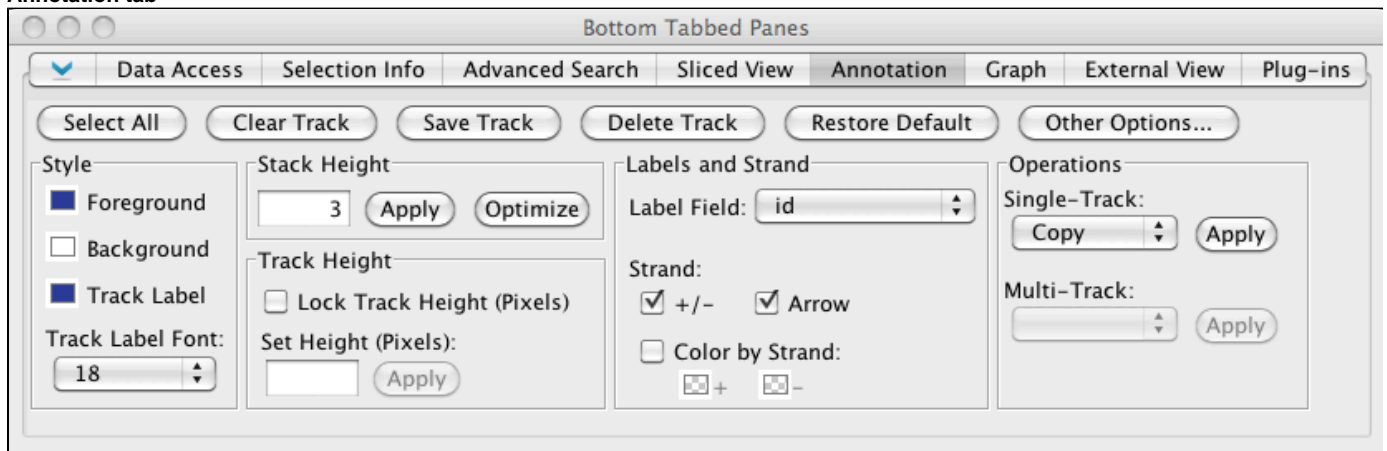
- [Introduction](#)
- [Changing track appearance](#)
  - [Style sub-panel](#)
  - [Stack Height sub-panel](#)
  - [Track Height sub-panel](#)
  - [Labels and Strand](#)
- [Operations](#)

## Introduction

The Annotation tab provides convenient access to annotation track configuration options and operations. To use **Annotation** tab,

1. Select one or more tracks by clicking track labels
2. Select **Annotation** tab
3. Change settings in **Annotation tab** sub-panels

### Annotation tab



## Changing track appearance

**Annotation** tab sub-panels on the left of the **Annotation** tab offer options for changing track appearance. The following sections describe track appearance configuration options and how to set them.

Click **Other Options...** as a shortcut to opening the **Tracks** tab in the **Preferences** window. Clicking **Other Options** is the same thing as selecting **File > Preferences > Tracks**.

Clicking **Restore Defaults** restores how a track appeared when first loaded into IGB.

Also see [Tracks](#)

### Style sub-panel

Use the **Style** sub-panel settings to change track colors and track label font. Changing the background color changes background color for both track and track label. Changing the foreground color changes the color of annotations within a track and the font color of the track label. Changing the track label color changes the color of the track label only.

To change color and font settings

1. Select one or more tracks
2. Select options in the **Style** sub-panel

## Stack Height sub-panel

Stack height is the number of annotations that can appear stacked in individual rows within a track. If a location in the genome contains more annotations than can be stacked in their own rows, IGB places them in a summary row. Annotations in the summary row are darker than annotations stacked in individual rows.

Change the stack height setting to change the number of annotations that can be stacked into individual rows.

To change stack height so that every annotation in the current view appears in its own row and no annotations are shown in the summary row

1. Select one or more tracks
2. Click **Optimize**

To change stack Height to a specific value

1. Select one or more tracks
2. Enter a new value in the **Stack Height** box
3. Click **Apply** or type ENTER

## Track Height sub-panel

Beginning in IGB 7.0, you can specify the absolute height (in pixels) for one or more tracks. To specify the height of a track

1. Select one or more tracks
2. Select **Lock Track Height**
3. Enter a value in the **Set Height** box
4. Click **Apply** or type ENTER

To change the track height, **Lock Track Height** option must be selected.

## Labels and Strand

Use the **Label Field** menu to specify which annotation property IGB display next to each annotation in a track.

To find out what properties are available to use as annotation labels

1. Select an annotation
2. Click the Selection Info button in the top right of the IGB window

A window will open listing annotation properties in the left column and their values in the right column.

To use a property's value as the annotation label in an annotation track

1. Select a track by clicking the track label
2. Select **Annotation** tab
3. Select the property from the **Label Field** menu

Use the **Strand** checkboxes to determine how strand will be indicated in a track. You can choose any combination of options in the **Strand** sub-panel.

Available options for indicating strand include the following:

- **+/-** Selecting this option combines plus and minus strand annotations into the same tier. De-selecting this option separates plus and minus strand annotations into separate tiers.
- **Arrow** Selecting this option causes annotations to be drawn with arrows pointing in the direction of transcription.
- **Color by Strand** Selecting this option causes annotations to be colored according to their strand. This option over-rides the foreground color set in the **Style** sub-panel.

## Operations

You can perform operations on tracks by

- Click operations buttons in the top row of the **Annotation** tab
- Using the **Operations** sub-panel

The **Operations** sub-panel contains operations you can perform on one or more annotations tracks.

For details, see [Annotation track operations](#).

The operations buttons in the top row of the **Annotation** tab include

- **Save** track contents to a file
- **Delete** a track, which removes the track from the IGB data display area
- **Clear** a track, which removes data from a track without deleting the track from the IGB data display area

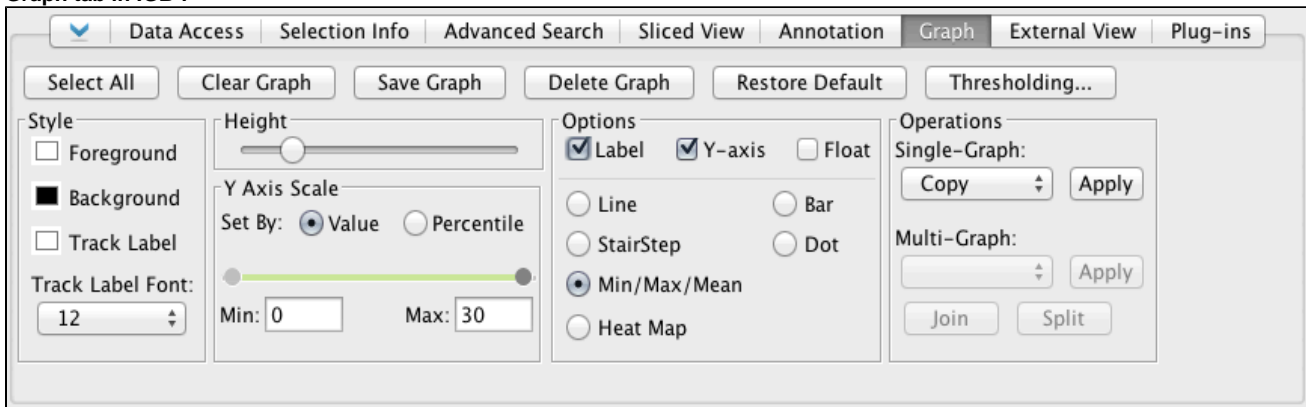
## Graph

- [Introduction](#)
- [Changing the appearance of graphs](#)
  - [Graph Labels](#)
  - [Y Axis and Scaling](#)
  - [Floating Graphs](#)
  - [Change the height of a graph](#)
  - [Change graph styles](#)
  - [Change graph color](#)

## Introduction

Use the **Graph** tab to change graph appearance and perform operations on graphs.

### Graph tab in IGB 7



## Changing the appearance of graphs

Most graphs are similar to the annotation tracks; you can modify many properties of graph tracks in the same way you modify properties of annotation tracks.

To change Right-click on the track handle and choose from the menu. See [Customize track appearance](#). Floating graphs do not appear in such a track. Settings and adjustments that are specific for graphs can be made using the **Graph** tab; to enable most of these options, a graph track must first be selected.

To change the settings for one or more graphs that are displayed in IGB:

1. Click the Graph Adjuster tab.
2. Select the graph(s) to change by doing one of the following:
  - a. Click the track label to select a graph.
  - b. Shift-click to select additional graphs.
  - c. To change the settings for all graphs that are currently displayed, click the **Select All** button in the **Graph Adjuster** panel.
3. Make changes to the graph settings by typing in new values or by operating sliders in the **Graph Adjuster** panel.

Any changes you make to the values in the Graph Adjuster panel will apply to all currently selected graphs.

## Graph Labels

You can show the track name, and y-axis min, max on graphs by checking the **Label** option in the **Manipulations** panel.

## Y Axis and Scaling

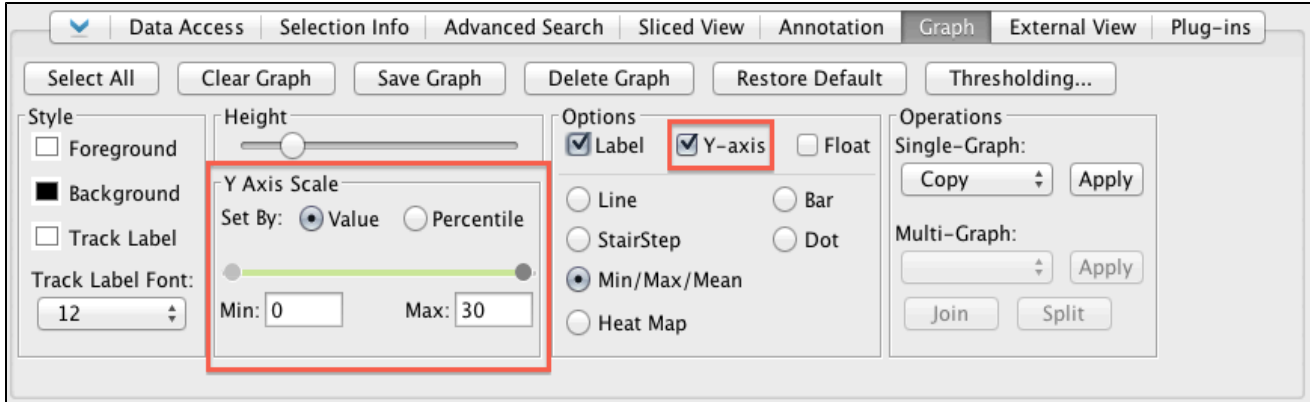
You can set the **Y axis** to be displayed by checking the **Y Axis** box. The axis markings will appear at the left side of the track.

The **Y-axis scale** can be altered using the **Y-Axis Scale** sliders. Scaling is an important issue to consider when viewing expression values because the distribution of these values across the range defined by their minimum and maximum may be highly asymmetrical. For example, the majority of values could range between 100 and 1,000, with just one or two percent exceeding 100,000. These extremely high values (relative to the rest) would tend to dwarf the others if displayed to scale because the vertical axis would be forced to show a range of 1 to 100,000. Although

technically accurate, such a graph would not be very informative because most of the information – the relative differences in value between different base pair positions – would be impossible to discern visually.

Setting minimum and maximum bounds **By Value** allow the user to set the cut-off values based on the actual **Y-axis** number. This does not take into consideration where most of the actual data lies. **By Percentile** offers thresholds defined by the actual percentage of data. For instance, using our theoretical 1 to 100,00, if you set **By Percentile** with **Max**: 90, and 90% of your data is below 1000, the **Y-axis** range would run from 0 to 1000. Any extreme values that are above or below the chosen thresholds are truncated from the display, and the values in between are displayed across a range that allows their differences to be noted.

To set these visible bounds, use the **Y-axis scale** section of the **Graph** tab:



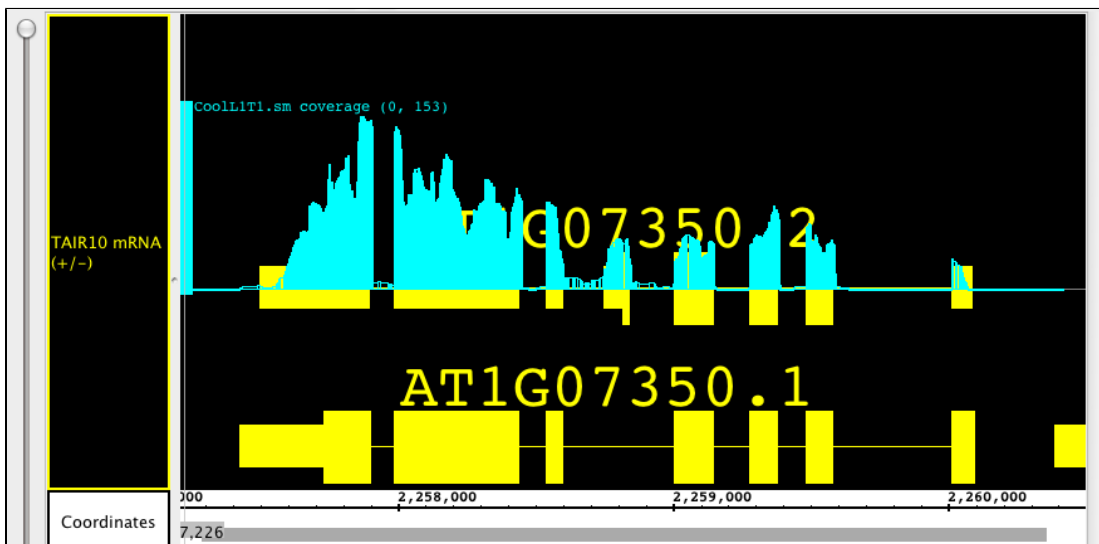
- To set specific values, select **By Value**. The chosen minimum and maximum value will be applied to each selected graph. You are free to set maximum and minimum values that cover a range smaller or larger than the actual range of your data. Either enter the number in the entry box or use the sliders.
- To set values by percentile, select **By Percentile**. Percentile ranges will be computed separately for each graph, which means that the absolute minimum and maximum range values for each selected graph may be different. Either enter the number in the entry box or use the sliders.

To display the full range of the data, choose **By Percentile** and set the **Min** to 0 and **Max** to 100.

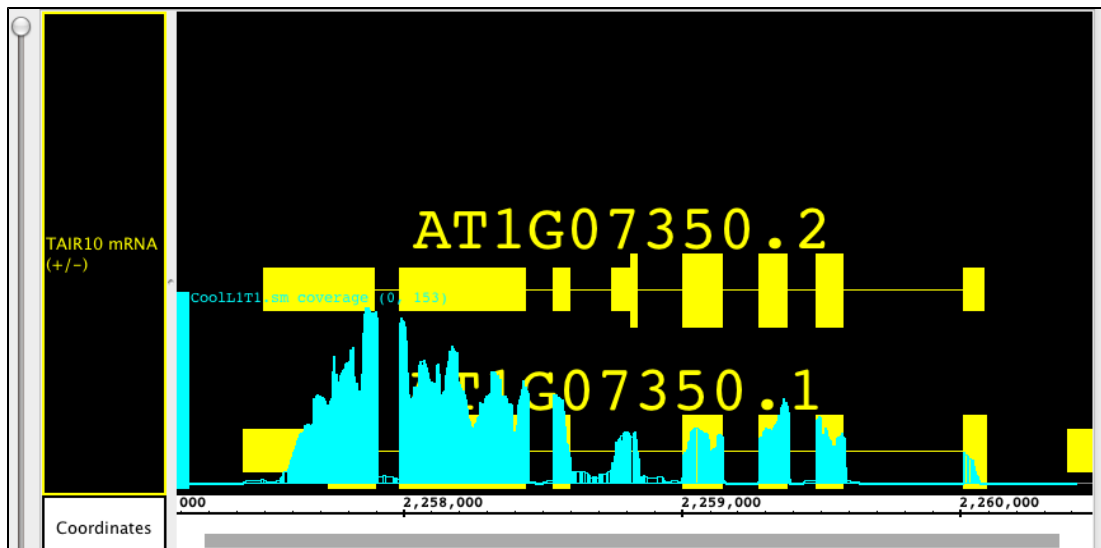
## Floating Graphs

In the **Options** panel, check the option **Floating**. By selecting this option, the graph track becomes 'disengaged' from the track framework. This allows a graph track to be moved any where (vertically) within the main image, AND overlay whatever it is placed over. By placing this quantitative track (graph) over or nearer to annotation tracks, you might find it easier to understand and analyze the data. If the graph is floating, you can change it's height by holding down the shift key and dragging your mouse up or down on the colored bar at the left end of the graph (this block is the 'graph handle').

To move or select a floating graph, click the graph-handle. In the pictures below you can see how the floating graph can be moved to overlay two different gene models, showing the alignment of exons and introns with the depth graph (.wig) from a next-generation sequencing experiment.







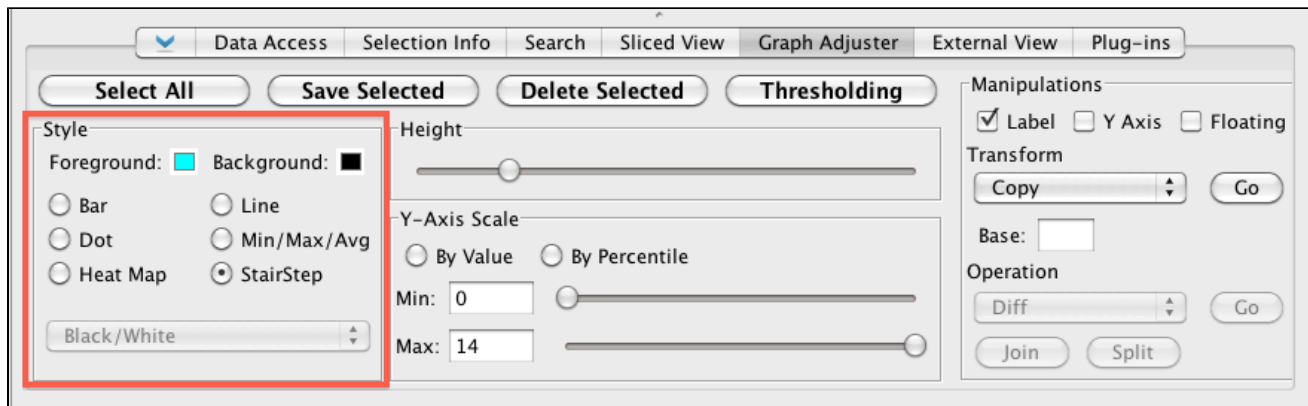
## Change the height of a graph

There are several ways to change the vertical height of a graph. Select the graphs you want to change by clicking on their track label, then drag the **Height** slide bar in the **Graph** tab. Alternatively, move your cursor to the boundary between track labels, and shift the edge of the track label up or down.

## Change graph styles

Graphs can be shown in various representational styles; the type of graph that is most appropriate depends on the type of question being asked about the data. For example, when comparing trends and patterns, it is very useful to use the line graph display method. The number of expression intervals being shown also can affect the graph display choice. The user is encouraged to experiment with the different display types to find out which method works best for specific purposes.

To change the graph style, select the graph track, then select the type of graph in the **Style** panel.

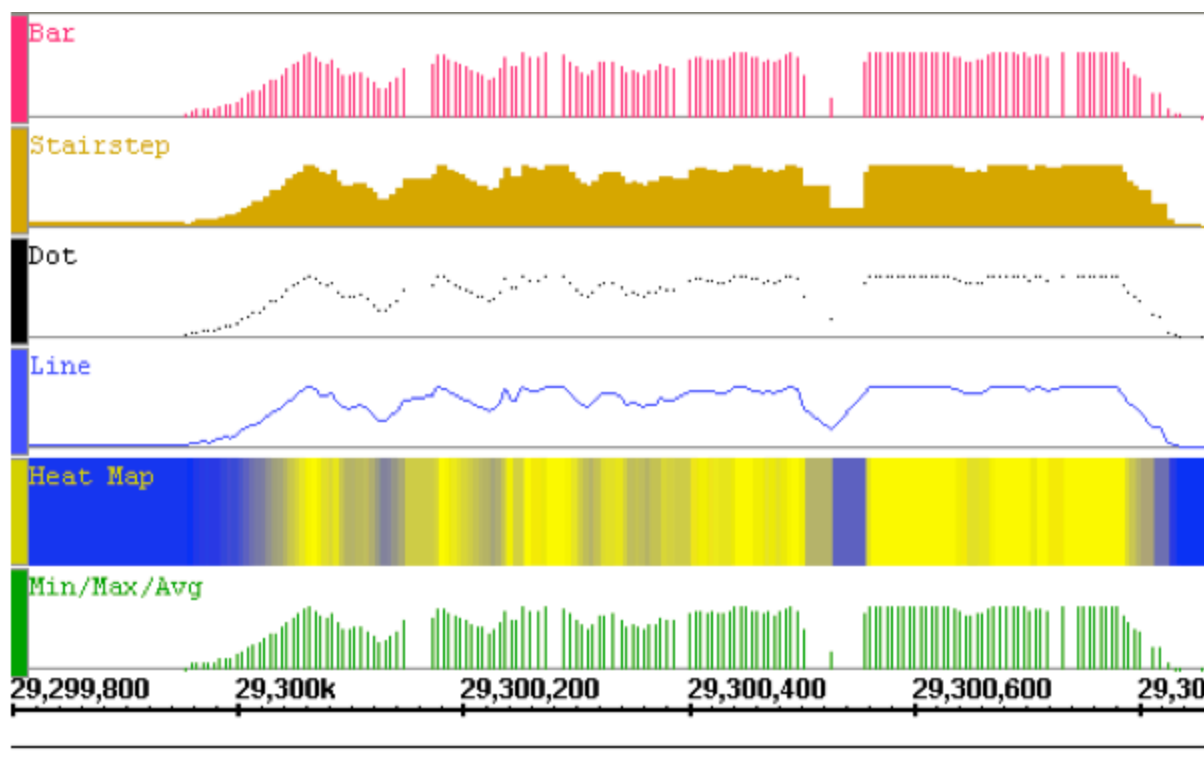


The following graph styles are available in IGB:

- **Bar** -- Individual values are shown as vertical bars that are one base wide for position graphs and of variable width for interval graphs.
- **Line** -- Subsequent values are linked with a line. Even if the input file was not sorted, the values will be connected in order along the genomic coordinate axis.
- **Dot** -- Shows a single dot for each data value. For interval graphs, horizontal lines will be connect the start and end points.
- **Min/Max/Avg** -- This is the default style for most graph data types and is usually fine for most purposes. It is especially useful for showing very densely populated graphs with data points for large numbers of positions. When IGB is zoomed all the way in, the display is equivalent to the Bar style. When zooming out, IGB starts to summarize values. When the scale of the display reaches the point where individual x-values are associated with multiple score values, IGB picks the maximum and minimum values and draws a vertical bar between them. In addition, IGB draws lines through the average of all the data points represented at each x value.
- **Stairstep** -- Similar to the bar graph style, except that bar widths along the horizontal axis are stair-stepped. For example, if position 100 has a value of 50 and position 200 has a value of 75 and there are no values in between, then IGB will draw a bar of height 50 that starts at position 100 and stops at position 200. Then, at position 200, IGB will draw a new bar of height 75 that terminates at the next location with a value. This style is particularly useful for viewing .egr and .sgr files, ESTs, or other high-density data.
- **Heat map** -- Instead of showing relative intensity via the height of the line at each pixel or coordinate as in most other graph styles, a heat map shows expression levels via color or brightness of the line at each pixel or coordinate. This graph style is useful if you want

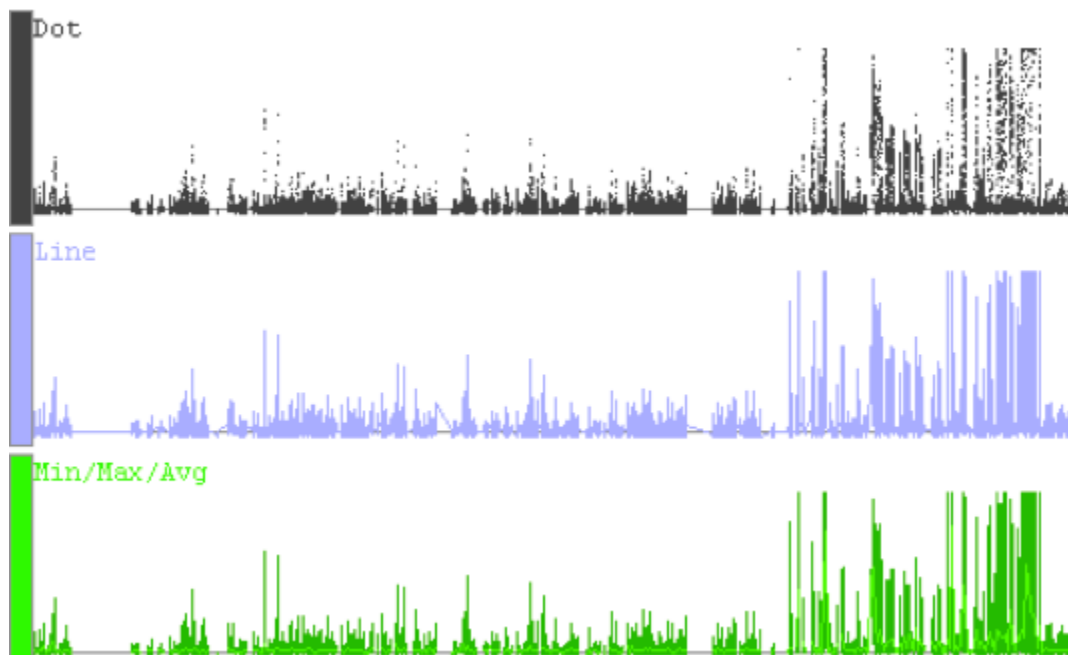
areas of greater expression to jump out at you. If a graph does not render or is hard to see, adjust the visible bounds of the graph until features are readily visible. Several heat map color maps are available to choose from. Transparent color maps are available and may be useful for floating graphs.

The following image illustrates how a single data set would look in each of the graph styles. The graph used for this image is a simple position graph. An example of an interval graph was given in the introduction to this chapter.



Note that with the heat map style for position graphs, the background (here in white) does not show through in regions where there are no data points. This is because the heat map first applies the stair-step function on the data before choosing colors to display. Compare the widths of the color bands in the heat map graph to those in the stair-step graph. For interval graphs displayed in the heat map style, the background will show through in regions with no data.

Due to the high zoom level in the above image, the Min/Max/Avg graph style is displayed exactly the same as the Bar style. When zoomed out, it reverts to showing the minimum, maximum and average values of all the data in the whole range of coordinates under each horizontal pixel width.



## Change graph color

To change graph color, like changing graph style, select graph, then use the color boxes in the **Style** panel.

For graphs in the heat map style, the chosen color will only be used on graph labels and handles. The heatmap colors come from the chosen heat map.

## External View

- [Introduction](#)
- [View a region in an external genome browser](#)
- [Import UCSC genome browser settings](#)
- [Advice on region sizes](#)

## Introduction

Use this tab to view the currently visible region in the Ensembl or UCSC Genome Browsers.

Use this tab to

- See an alternative view on the same data
- Find tracks to download and import into the IGB environment

Both Ensembl and UCSC Genome Bioinformatics sites are accessible via DAS1 data sources in IGB. And the UCSC Genome site has a Table Browser that allows you to download Genome Browser data tracks in formats IGB supports.

So if you see something worth further investigation, you can import the data directly into IGB.

## View a region in an external genome browser

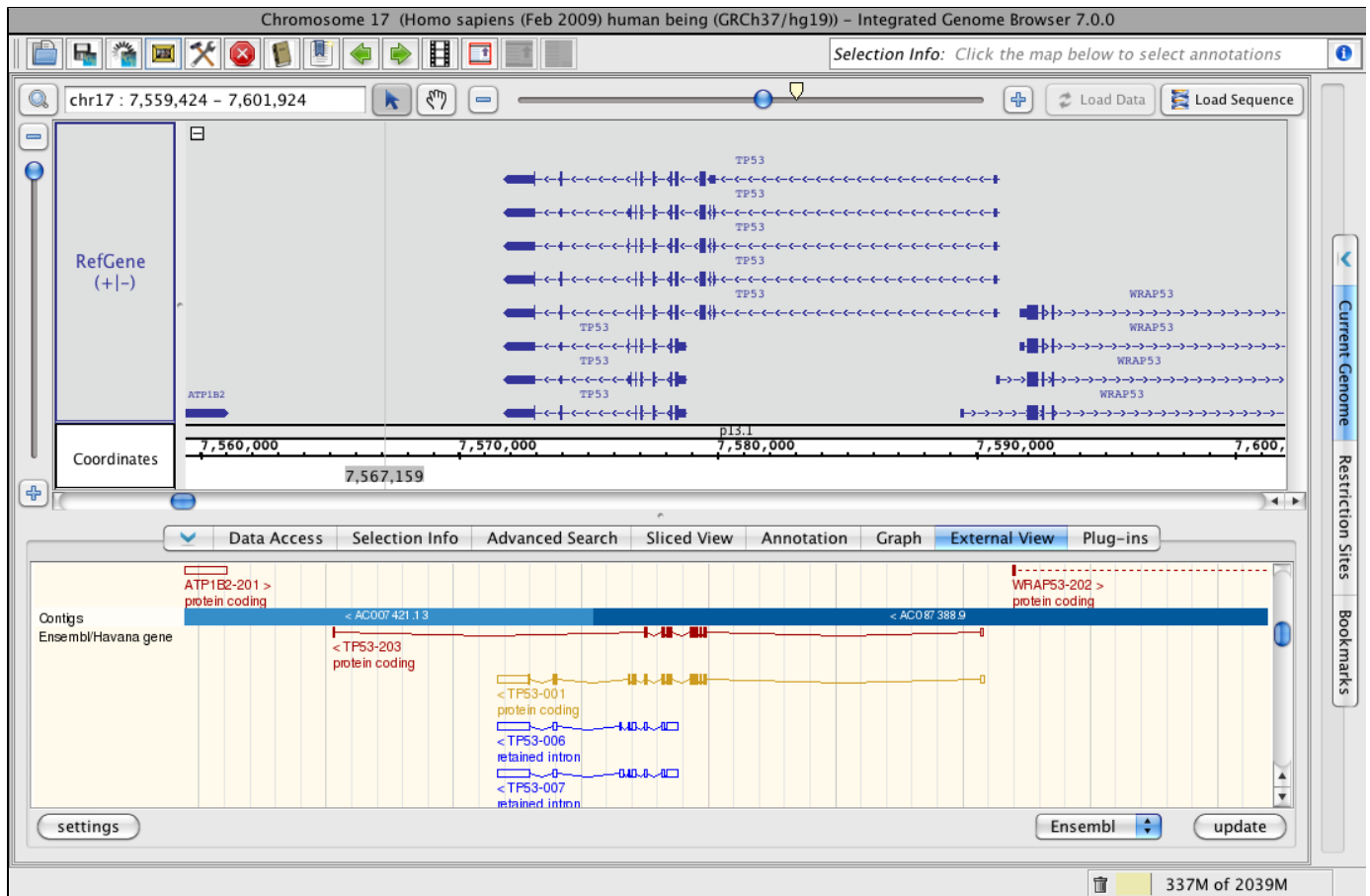
Depending on the genome, you can use **External View** tab to view the currently visible region in the UCSC or Ensembl genome browsers.

To view a region

- Click **External View** tab to open it
- Choose external browser from the menu
- Click **Update**

If you navigate to a new region, click the **Update** button.

**External View showing TP53 in IGB and the Ensembl Genome Browser**



## Import UCSC genoe browser settings

If you have used a web browser to access the UCSC server and have changed the UCSC view through the browser, you can import those settings into IGB, permitting IGB to display the view in the same way as the web browser. Click the **settings** button (blue box in previous image) and follow the directions in the window.

### Settings window

Customize UCSC settings

### Setting the UCSC user id

If you have already customized the USCS view in a web browser, IGB can use those settings. Follow the instructions below to link your settings to IGB.

1. Obtain your user id by clicking on the "UCSC info" button. Or open <http://genome.ucsc.edu/cgi-bin/cartDump> in your browser
2. Then scroll down in the opened window and copy the value of hguid into the "UCSC user id" field.
3. Click the submit button.
4. Your IGB UCSC View is now synchronized with your browser configuration. The settings in your browser now change the view.

UCSC info

UCSC user id (hguid):

submit

## Advice on region sizes

While you can access information on very large regions (spanning 10-15 genes), a smaller region or specific gene may return more data due to more details being available. To use ENSEMBL you must have a fairly zoomed in picture (<100,000 bp) and it must be a species that ENSEMBL

curates. The same is true for UCSC, although it can handle slightly larger regions of the chromosome. A smaller region or specific gene may return more data due to more details being available.



Larger regions take longer to load.

## Plug-ins

- [Introduction](#)
- [How to install or un-install a plug-in](#)
  - [Adding new plug-in repositories](#)
- [Updating plug-ins](#)

## Introduction

Plug-ins are small programs you can install while IGB is running to add new functionality. Use the Plug-ins tab to find out what plug-ins are available, install or un-install them, update installed plug-ins, and add new plug-in servers.

## How to install or un-install a plug-in

- Click the **Plug-ins** tab
- Click checkbox in **Installed** column

### Plug-ins tab

<div><div></div>Data AccessSelection InfoAdvanced SearchSliced ViewAnnotationGraphExternal View<b>Plug-ins</b></div>				
Installed	Name	Description	Version	Repository
<input type="checkbox"/>	TallyHandler <a href="#">i</a>	plugin to allow the user to load tally files	1.0.0	Bioviz
<input checked="" type="checkbox"/>	TranscriptIsoform <a href="#">i</a>	BETA release of transcript isoform evidence visualization	0.0.9	
<input type="checkbox"/>	TutorialHelper <a href="#">i</a>	plugin to allow the user to see widget ids to help develop tutorials	0.0.9	Bioviz
<input type="checkbox"/>	SampleSelection <a href="#">i</a>	tab to allow the user to select sample data	1.0.0	Bioviz
<div><input checked="" type="checkbox"/> Show Installed<input checked="" type="checkbox"/> Show Uninstalled<div>Update all</div><div>Update selected</div><div>Repositories...</div></div>				

## Adding new plug-in repositories

To add a new plug-in repository

- Click **Repositories**
- Enter URL of the repository

## Updating plug-ins

IGB shows you the version number of a plug-in in the **Version** column. The **Update All** button will become available (no longer be grayed out) if newer versions of a currently installed plug-in are available.

Click **Update Selected** or **Update All** to update selected or all plug-ins.

## Geometric Mean

## Primer Design

## Tutorial Helper

## Current Genome

- [Introduction](#)
- [Species link](#)
- [Choosing species and genome version](#)
- [Choose a sequence to view](#)
- [View the entire genome](#)

## Introduction

Use the Current Genome tab to select the species, genome version, and chromosome to view in IGB.

## Species link

Clicking the Species link ([blue arrow](#)) returns you to the IGB start screen.

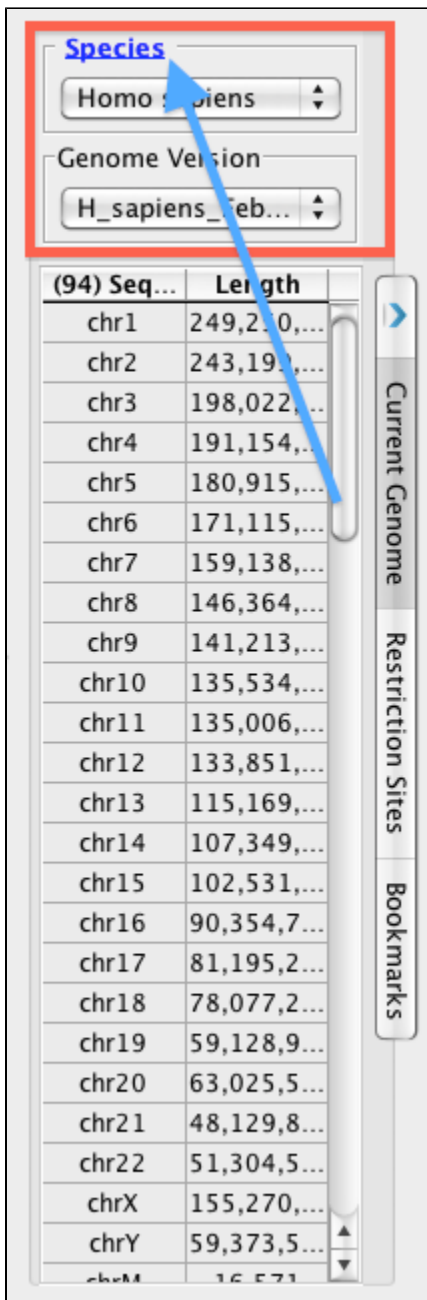
## Choosing species and genome version

All species and genome versions supported by IGB data sources appear in the Species and Genome Version menus.

If you have added your own custom genome, it will also be listed here.

To view a new genome

- Select a species from the Species menu ([red box](#))
- Select a genome version from the Genome Version menu



## Choose a sequence to view

The **Current Genome** tab contains a table listing chromosomes and chromosome sizes for the current genome version.

Use this table to sort chromosomes by name or size or to choose a new chromosome to view.

To change to a new sequence

- Click its row in the table

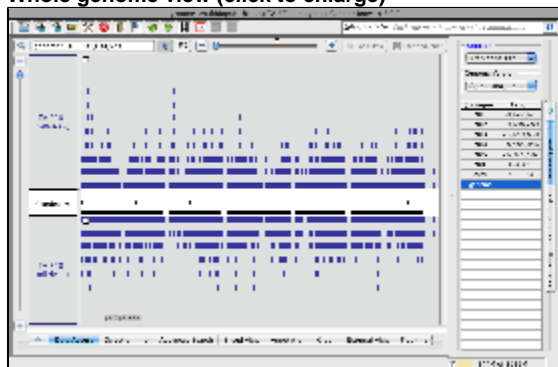
To sort by size or name

- Click the column title

## View the entire genome

The chromosome table contains a row named **genome**. Clicking this row activates a whole-genome view in which all chromosomes are shown in IGB.

### Whole genome view (click to enlarge)



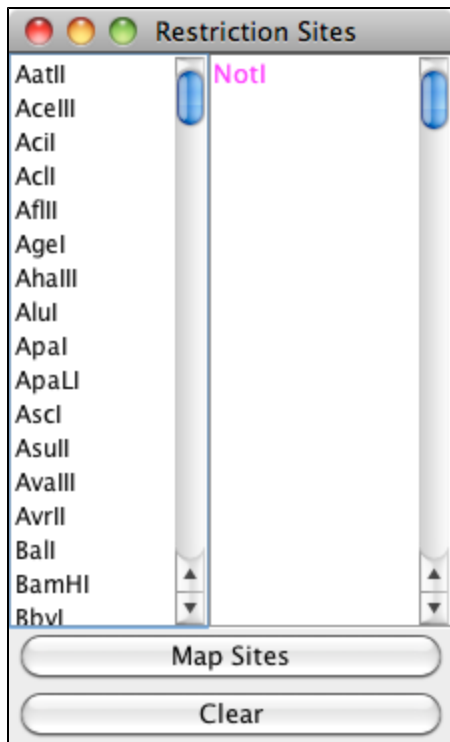
## Restriction Sites

- [Introduction](#)
- [How to map sites](#)
- [Viewing results](#)

## Introduction

The **Restriction Sites** tab maps restriction enzyme sites within the reference genomic sequence.

**Restriction Sites** tabbed panel



## How to map sites

To display restriction sites for an enzyme (or group of enzymes)

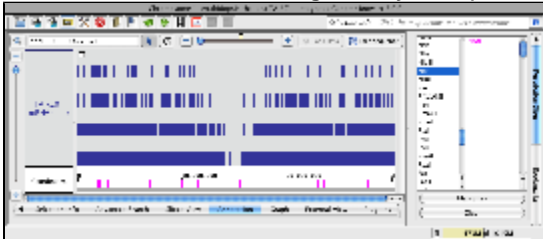
- Click the enzyme name
- Click **Map Sites**

Multiple selections are possible by SHIFT- or CONTROL-clicking (platform dependent)

## Viewing results

Sites are shown underneath the coordinate axis in the main view. To zoom in on a site, double-click it.

**Zoomed-out view of sites recognized by Not1 (a rare cutter) mapped onto Arabidopsis chr1 (Click to enlarge)**



**Zoomed-in view of a single Not1 site (click to enlarge)**



Click the **Clear** button to remove sites.

## Bookmarks

- [Introduction](#)



- [Bookmarks tab overview](#)
- [Returning to a bookmarked region](#)
- [Creating new bookmarks](#)
  - [Bookmark types](#)
- [Managing bookmarks](#)
- [Exporting and Importing bookmarks](#)
- [Viewing bookmark properties \(for advanced users\)](#)
- [IGB bookmarks file](#)

## Introduction

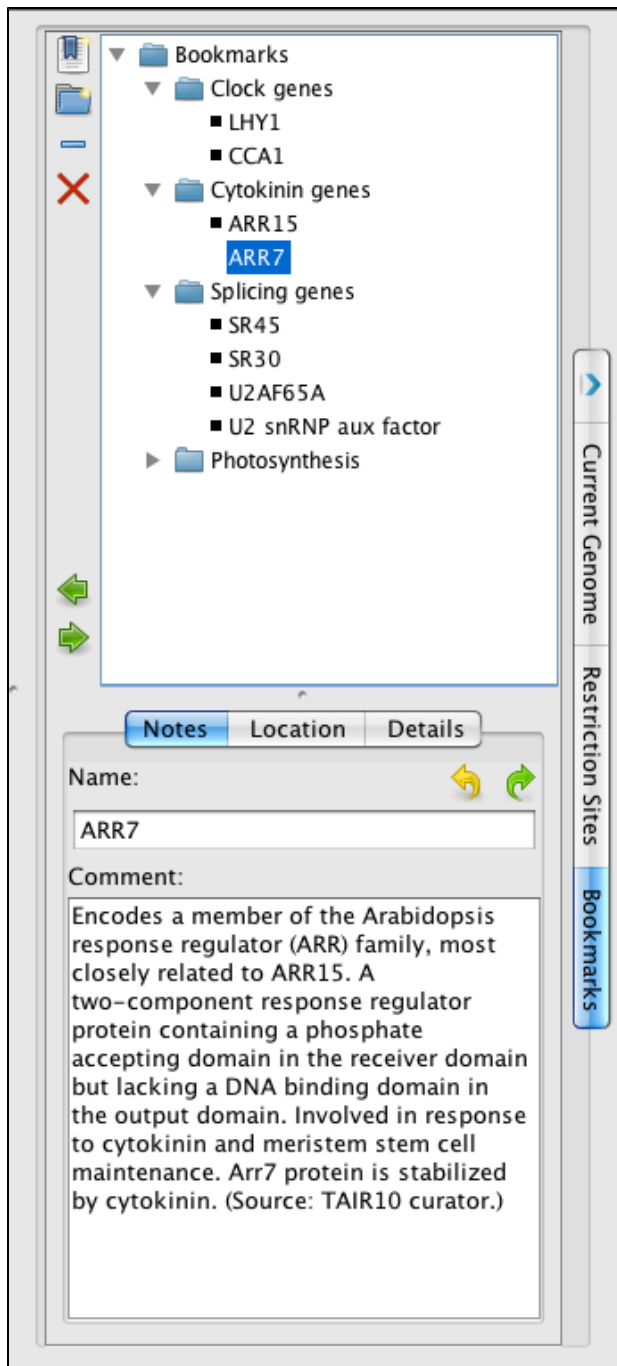
Use IGB bookmarks to:

- Jump to a bookmarked location and re-load data.
- Make notes on a genomic scene.
- Share analysis results with other users.

## Bookmarks tab overview

The **Bookmark** tab allows you easy access to the bookmark list and provides tools for managing and sharing bookmarks.

**IGB bookmarks tab with example bookmarks**



## Returning to a bookmarked region

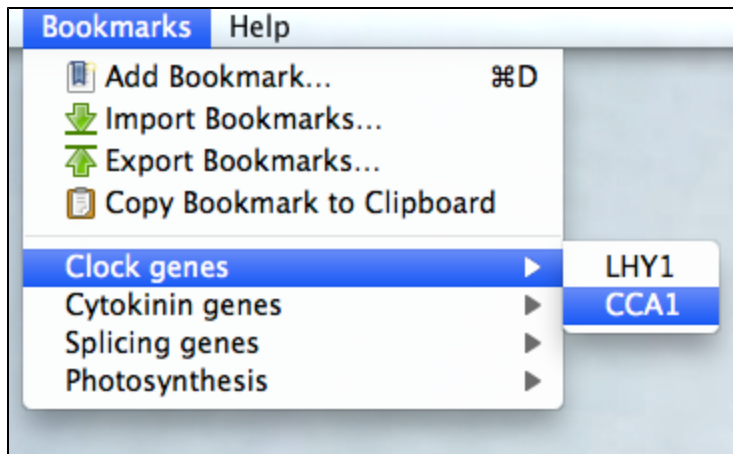
To return to a bookmarked region,

- Double-click it

or

- Choose the bookmark from the **Bookmarks** menu

**Bookmarks menu**




## Creating new bookmarks

To bookmark a location and data scene

- Choose **Bookmarks > Add Bookmark**

or

- Click **Bookmarks** tab
- Click add the **New Bookmark** icon 

A window will appear with places to enter a name, notes, and bookmark type.

### New bookmark window

## Bookmark types

**Position Only** bookmarks record genome, current position and zoom level.

**Position and Data** bookmarks record genome, current position, zoom level, and data sets.

## Managing bookmarks

You can sort your bookmarks into folders and associated comments with comments on folders.

To edit the name or comments associated with a bookmark or folder

- Click **Bookmarks** tab.
- Click to select a bookmark or folder.
- Edit name or comments in the Notes sub-tab.

To delete a bookmark or folder

- Click to select it
- Click the delete bookmark icon (red X) or press Delete

## Exporting and Importing bookmarks

To save a copy of the bookmarks to a file

- Choose **Bookmarks > Export Bookmarks...**

To import another user's bookmarks

- Choose **Bookmarks > Import Bookmarks...**

Bookmarks are saved as a single file in an HTML-based format. Use caution if you choose to edit the file.

You can import multiple sets of bookmarks during the same IGB session. Imported bookmarks will be added to the end of the list of existing bookmarks.

## Viewing bookmark properties (for advanced users)

To see when a bookmark was created and the region it refers to

- Click **Bookmarks** tab
- Click the bookmark to select it
- Click **Location** sub-tab

To see other properties of a bookmark

- Click **Bookmarks** tab
- Click the bookmark to select it
- Click **Details** sub-tab

or

- Click **Bookmarks** tab
- Right-click a bookmark
- Select **Properties..**

A new window will open that reports information about the bookmark and how IGB represents it internally.

To see how IGB combines these properties into an IGB bookmark, export your bookmarks and open the bookmarks file in a text editor to inspect bookmark syntax.

## IGB bookmarks file

IGB saves your bookmarks at the end of each session in a file named bookmarks.html in a folder named .igb in your home folder.

## Visualizing read alignments

- [Introduction](#)
- [Open short-read alignment \(BAM or SAM\) file](#)
  - [Set the load mode \(optional\)](#)
    - [Manual \(the default\)](#)
    - [Auto](#)
- [Zoom in on a region and load data](#)
- [Load sequence data to view mismatches](#)
- [Adjust stack height](#)

## Introduction

Starting with IGB version 6.3, you can use IGB to visualize sequence alignment files alongside annotations and genome sequence.

IGB supports visualization of alignment files in

- BAM format (binary alignment format)
- SAM format (plain text version of BAM)
- PSLX format (output from the blat alignment tool)
- VCF (with the appropriate plug-in)

As with other file formats, IGB can open files from your local hard drive, from a Web site, or from a QuickLoad data source.

If you are loading alignments from a Web site, IGB will first download your index (BAI) file and then use it to determine which part of the much larger BAM file to download whenever you click the **Load Data** button.



Each BAM file requires an index file (BAI) located in the same location (same folder, same website, etc.) as the BAM file. SAM files do not require indexes.

## Open short-read alignment (BAM or SAM) file

If your species and genome version of interest are available, first select it using menus in the Current Genome tab or click a species shortcut image on the IGB start screen.

Next, open the file. You can open a file in IGB by

- using **File > Open File** menu (allows you to enter a custom species and genome)
- using **File > Open URL...** menu
- click-dragging from your desktop into the IGB window
- click-dragging a URL from a Web page
- or loading from a QuickLoad server

When you first open a BAM file, a new empty track will appear in the main IGB window, labeled with the name of the file or the name of the data set if loaded from an IGBQuickLoad site.

A new row representing the newly opened file or data set will also appear in the **Data Management Table** in the **Data Access** tab.



If your genome of interest is not available, you can enter it in the file chooser window.

## Set the load mode (optional)

### Manual (the default)

When you first open a BAM file, its **Load Mode** will first be set to **Manual**, meaning that to load the data into IGB, you click the **Load Data** button. Also, only data for the currently shown region will load into IGB.

Before clicking the **Load Data** button, first zoom in to a region or search for a gene of interest.

Attempts to load data for large regions will prompt a warning from IGB.

Since alignment data sets are typically very large, it is a good idea to only load a few genes' worth of data, depending on the file and the experiment.

### Auto

The Auto load mode setting triggers automatic loading of data when the display is zoomed in past the **Autoload** threshold, indicated by the yellow marker on the horizontal zoom control.

Drag the horizontal zoom control to the right of the Autoload marker to activate automatic loading of data sets.

To change the Autoload threshold to the current zoom level, choose **View > Set Autoload Threshold to Current View**

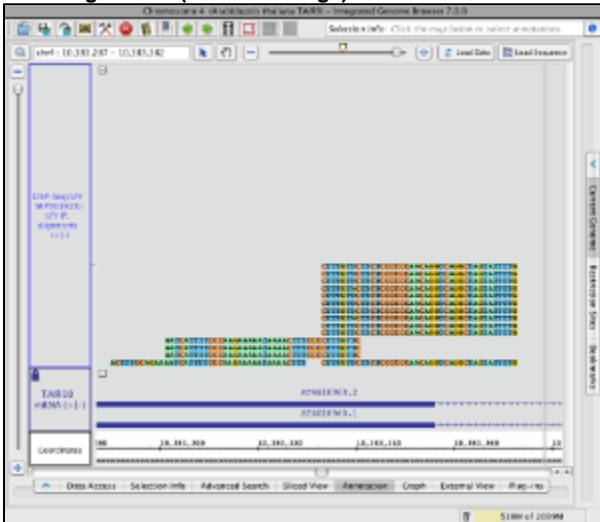
## Zoom in on a region and load data

- Use the slider, double-click a gene model, run a search, or click-drag the sequence axis to zoom in.
- Click the **Load Data** button to load data into the display.

IGB will then load data for the currently visible region.

If genomic sequence data have not already been loaded for the current region, IGB will show the sequence bases of reads when zoomed in, as in the image below.

### Read alignments (click to enlarge)



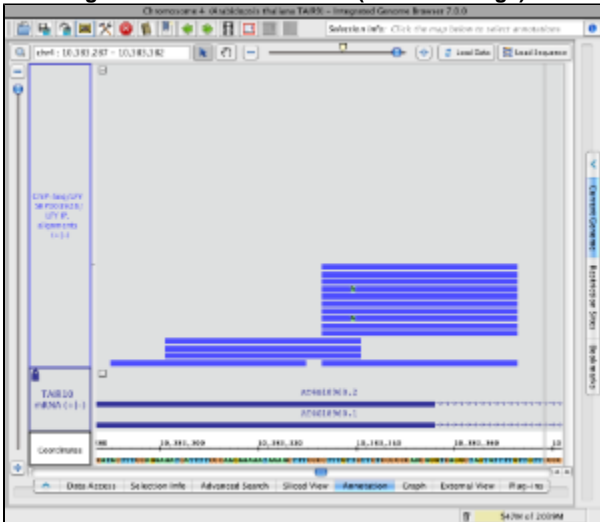
## Load sequence data to view mismatches

To view just mismatches between reads and the reference genome

- Click **Load Sequence in View** in the upper right corner of the IGB display

This will trigger loading of genomic sequence and allow IGB to show just the mismatches rather than the full sequence of each aligned read.

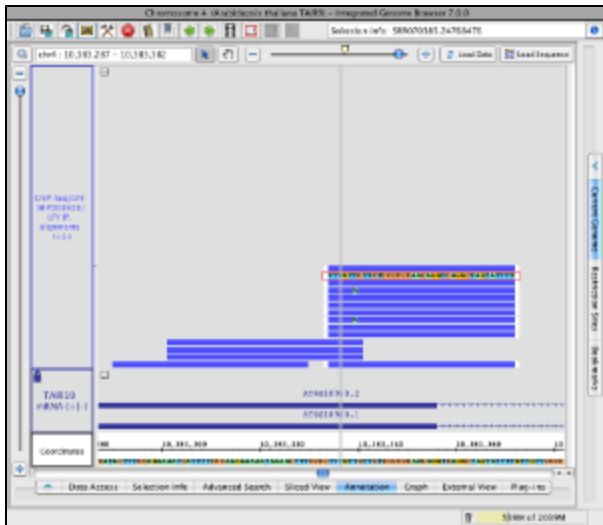
### Read alignments with mismatches (click to enlarge)



To restore read sequence

- Right click a read and select **Restore Read Sequence**.

### Read alignment with read sequence (click to enlarge)



## Adjust stack height

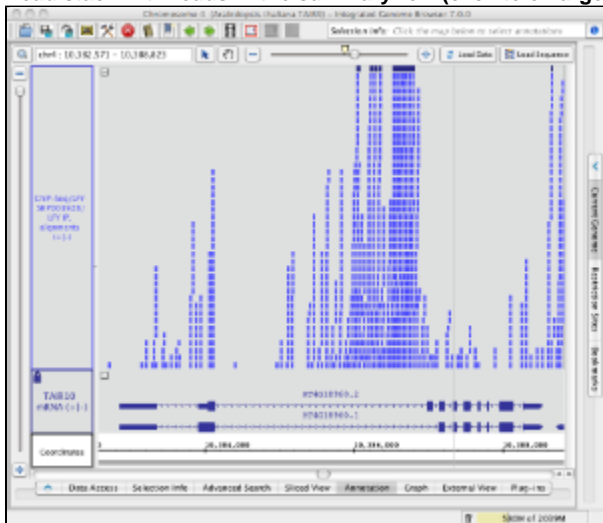
Note that in the image below showing read alignments from a ChIP-Seq experiment, there is a row of reads at the top of the track which appear darker than the rest. The darker row indicates there is some number of reads that are not being shown because they exceed the current stack height setting for the track.

The darker row of reads in the top row of the track are called the "summary row." To count them, click-drag the arrow cursor over the summary row. In this case there are more than 1,000 reads in the summary row.

The number of rows shown below the summary row is determined by the stack height setting. To change the stack height setting

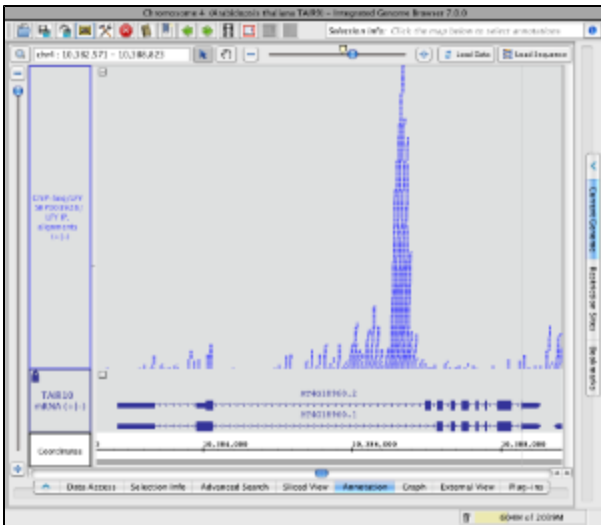
- Right-click a track label and choose **Set Stack Height** or **Optimize Stack Height**

### Read stack with reads in the summary row (click to enlarge)



The image below shows the same scene as in the image above after using the **Optimize Stack Height** feature. Note that the height of the annotation track (TAIR10 mRNA) is locked to ensure that increasing the read track's stack height would not affect the annotation track's height, which is set to 100 pixels.

### Read stack with all reads showing (click to enlarge)



## FindJunctions

- Introduction
- FindJunctions algorithm
- How to use FindJunctions within IGB
- Using FindJunctions from the command line
- To run FindJunctions as a stand-alone program, visit <https://bitbucket.org/lorainelab/findjunctions>

## Introduction

FindJunctions uses gapped RNA-Seq read to genome alignments to identify exon-exon junctions, i.e., introns. FindJunctions is implemented both as a visual analytics function within IGB and as a stand-alone, command-line program.

Within IGB, FindJunctions produces a new track showing exon-exon junction features labeled with the number of alignments that supported the junction. The command-line program produces a BED file in which the score field contains the number of reads supporting the junction.

## FindJunctions algorithm

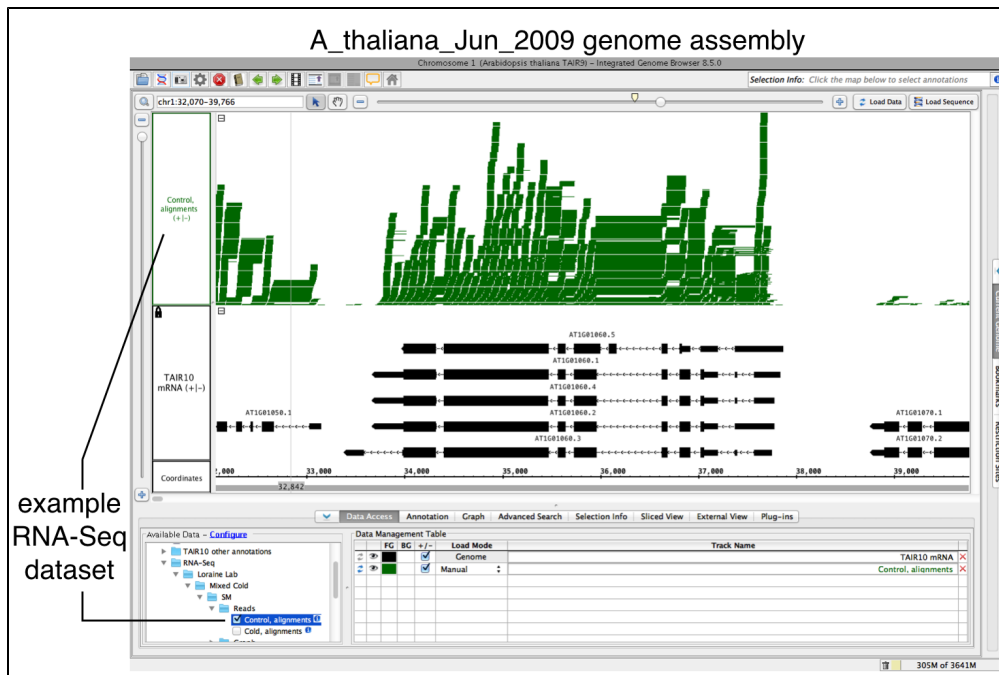
FindJunctions operates on RNA-Seq read alignments, using alignments that contain gaps in the read sequence relative to genomic sequence. These gaps in the read sequence correspond to introns and typically start and end with the so-called canonical splice site consensus sequences "GT" (5' end) and "AG" (3' end) for genes transcribed from the plus (forward) strand. For genes transcribed from the minus (reverse) strand, the consensus sequences relative to the plus strand are the reverse complement of the consensus splice site sequences, i.e., "CT" on the 5' end of the gap and "CA" on the 3' prime end. For each alignment containing a gap, FindJunctions inspects the start and end coordinates of the gap and uses the genomic sequence to infer the strand, if available. FindJunctions creates a list of all such gaps, recording the strand and the genomic coordinates of the start and end coordinates. For each unique triplet of start, end, and strand, FindJunctions creates a scored junction feature and increments the score each time a gap supporting that feature is encountered in a dataset. Options are available to limit scoring to read alignments that have a given minimum number of bases flanking a gap and/or which having one unique mapping onto the genomic sequence.

## How to use FindJunctions within IGB

To use FindJunctions within IGB, first load a BAM file:

1. Open BAM file.
2. Zoom in on the gene or region of interest.
3. Click **Load Data**.

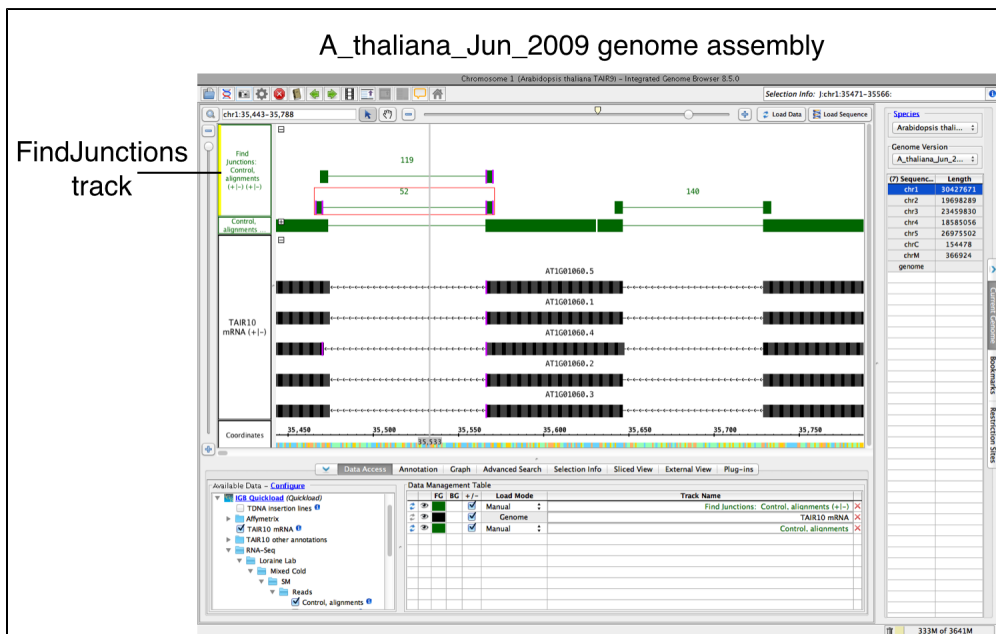




Run FindJunctions:

1. Right-click the RNA-Seq track label
2. Select **Track Operations > FindJunctions**
3. Enter a value or use the default. At least this many bases must align across a putative intron for a read to be counted as support for a junction.
4. Select OK to run FindJunctions.

A new track will then appear containing junction features bracketing introns. Labels report the number of spliced alignments that supported the junction.



## Using FindJunctions from the command line

To run FindJunctions as a stand-alone program, visit <https://bitbucket.org/lorainelab/findjunctions>

Follow the instructions to compile FindJunctions and create a "jar" file.

FindJunctions takes one argument - the name of a BAM file, and multiple options:

- -u option (unique) ensures that only uniquely mapping spliced reads (with NH tag equal to 1) will be used to construct junctions. Not required, default is to use all reads regardless of mapping quality or number of mappings obtained.
- -n option is the number of bases that must map to either side of a putative intron for a spliced read to be used to create or support a junction feature. Default is 5.
- -b is the absolute path to a .2bit format genomic sequence file that will be used to identify junction strand. This is required.
- -o (output) is the name of the junctions file that will be written. Default is print to stdout.

Output is tab-delimited BED12 format. The name field contains a name constructed from the location of the junction and the score field contains the number of spliced alignments supporting each junction.

Example)

```
java -Xmx1g -jar FindJunction_exe.jar -u -n 5 -b Genome.2bit -o FJ.bed
sample1.bam,sample2.bam
```

Here, -Xmx1g option specifies that the program can run with up to 1 Gb of computer memory (RAM) using the code in jar file (-jar) FindJunction\_exe.jar.

## Paired-end reads

- [Introduction](#)
- [Viewing paired-end data](#)

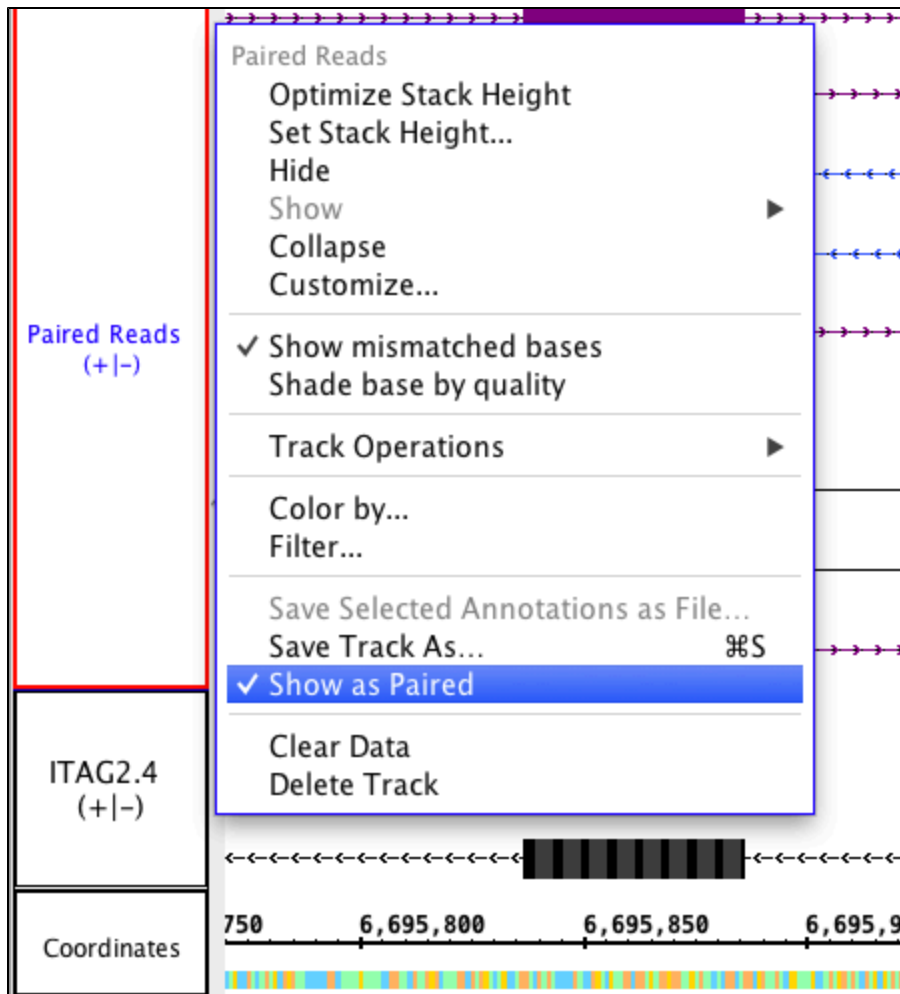
### Introduction

Paired-end sequencing allows you to sequence both ends of a fragment. Many aligners can then use the paired-end data to better align reads, especially in cases of repetitive regions.

### Viewing paired-end data

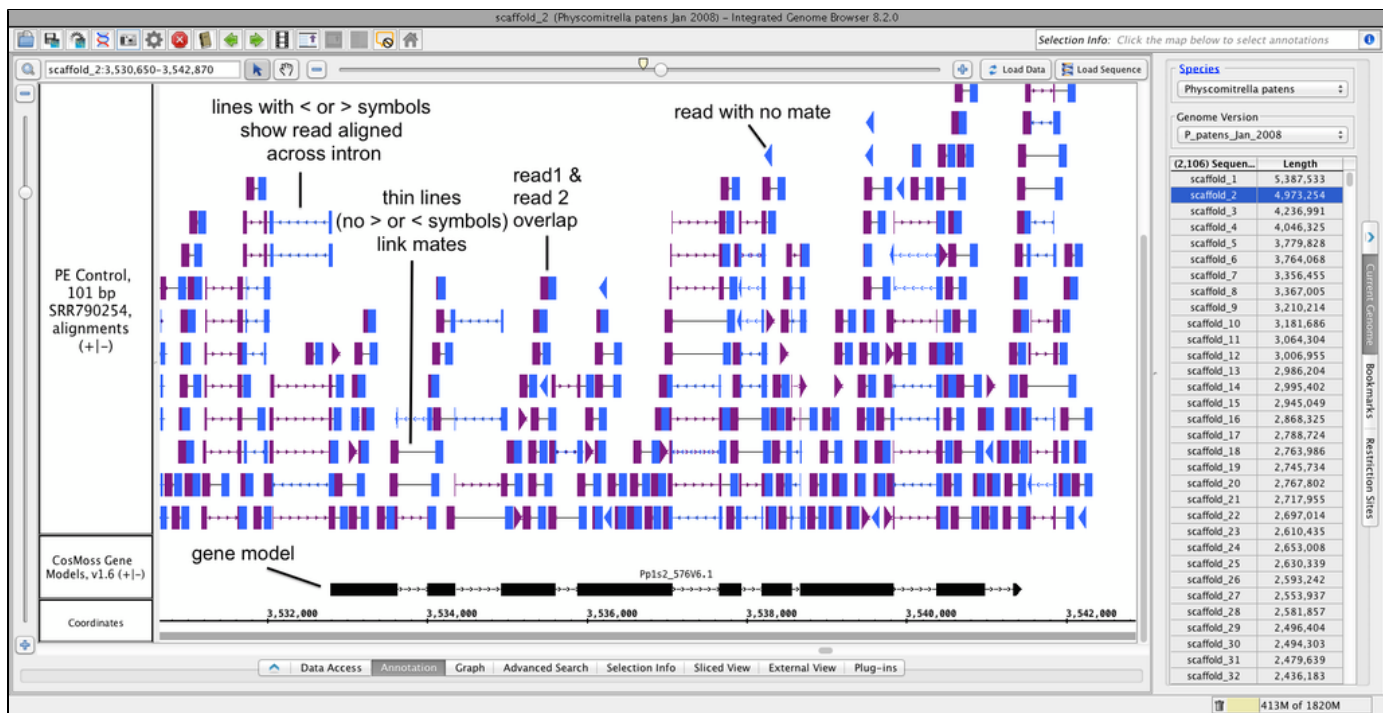
You can load paired-end files, such as aligned reads in BAM format, into IGB just like any other data file (see - [Loading data](#)).

To enable paired-end visualization, right click on the track or track label and select **Show as Paired**.

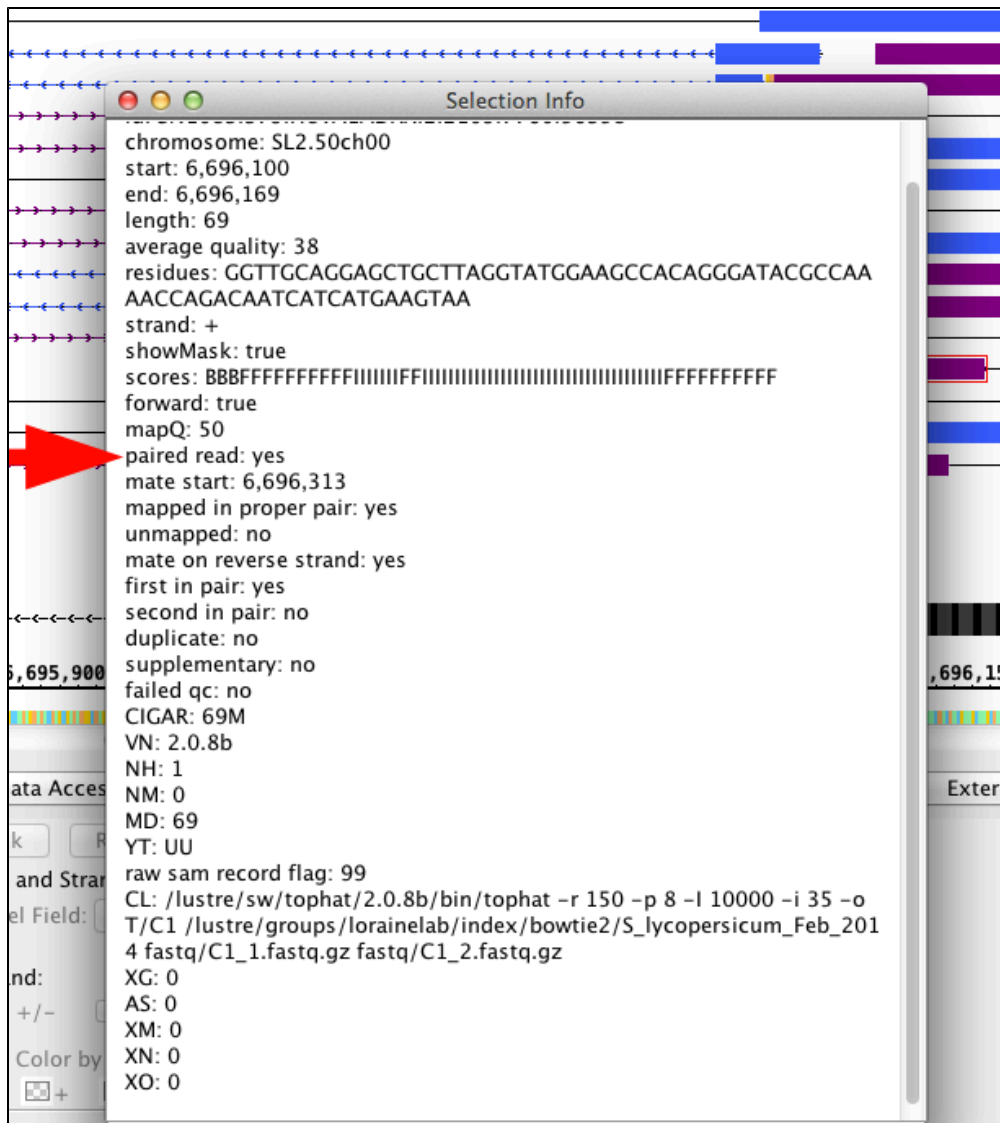


Selecting **Show as Paired** will change the visualization in several ways:

- Paired reads will be connected by a thin black line.
- Data that were separated by strand will be placed on the same strand. So long as **Show as Paired** is active, the strands cannot be separated.
- **Color by Strand** will be activated. This option can be disabled within the **Annotation** tab.
- Arrows depicting the direction of each strand will be enabled. The **Arrow** checkbox can be disabled within the **Annotation** tab.



If you want to confirm the status of a read, hover over it to bring up the **tooltips**, or select the read and click on **Selection Info**. The paired status of the read can be found at **paired read**.

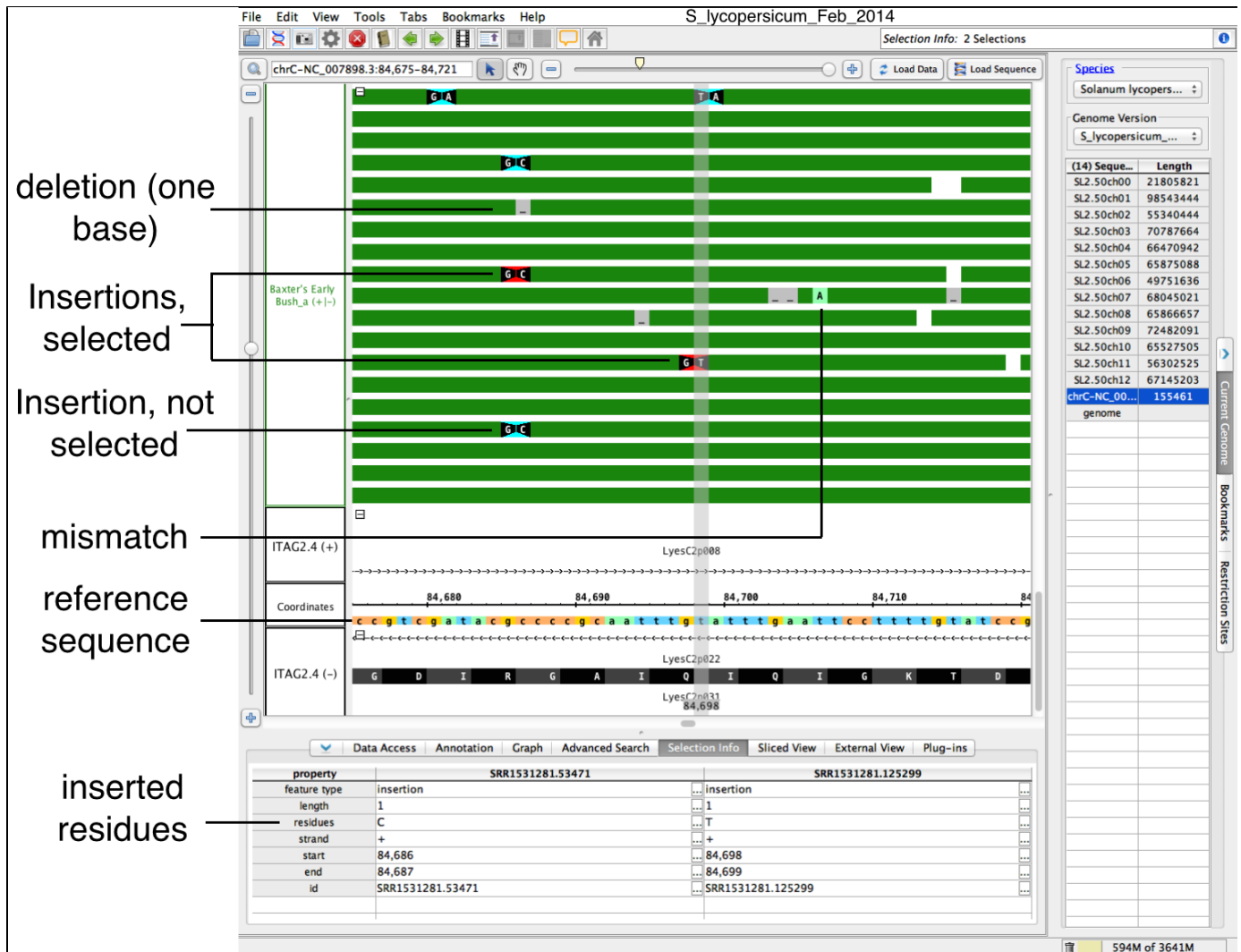


## Visualizing insertions, deletions, and mismatches

IGB displays insertions, deletions and mismatches as follows:

- **Insertions** - residues flanking insertion site appear are shown in white with a black background; triangular shapes between them point to the location of the inserted base(s) present in the aligned sequence but not in the reference. Select an insertion to display the sequence of inserted bases.
- **Deletions** - residues deleted in the aligned sequenced relative to the reference appear in gray with underscore characters
- **Mismatches** - residues in the aligned sequence that do not match the corresponding base in the reference appear as highlighted bases. To highlight mismatches, click **Load Sequence**

Use the zoom stripe (shown in gray) as a guide to compare mismatches, deletions, and insertions between different read alignments and the reference.



To re-create this scene in IGB:

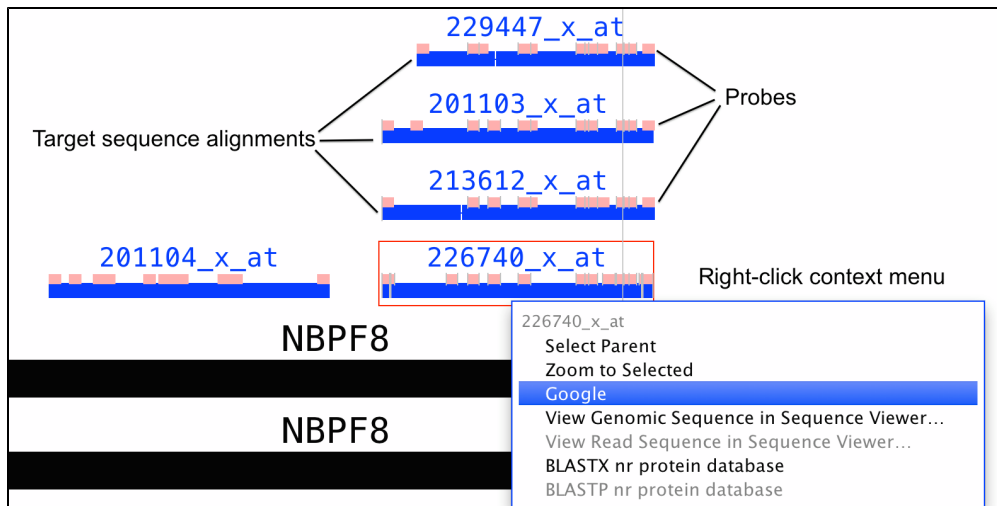
- Open tomato genome S\_lycopersicum\_Feb\_2014 (ITAG2.4)
- Select **Current Genome** > **chrC-NC\_007898.3** (tomato chloroplast genome)
- Click **Load Sequence** to load chloroplast genome sequence
- Zoom to chrC-NC\_007898.3:72,724-102,016 (enter in **Quick Search** box, top left)
- Select **Data Access** > **Available Data** > **Plastid Resequencing** > **Baxter's Early Bush\_a**
- Click **Load Data** to load alignments for Baxter's Early Bush tomato cultivar
- Expand stack height to 150 reads per track (right-click track and select **Set Stack Height**)
- Use vertical zoomer to stretch read alignments vertically
- Enter chrC-NC\_007898.3:84,675-84,721 in **Quick Search** box to zoom to region

## Visualizing probe sets

- [Probe set alignments - link.psl files](#)
- [About probe set alignment visualizations](#)
- [Why this is useful](#)

## Probe set alignments - link.psl files

IGB can display Affymetrix probe sets aligned onto a reference genome - it can show probe set design sequences aligned onto a genome with the locations of the probes indicated as blocks.



When IGB was first developed at Affymetrix, the company distributed probe set alignment files for its catalog 3' IVT arrays. In recent years, however, they've stopped updating these files. So for some genomes, the alignment files you can find on the Affymetrix Web site reference obsolete reference genomes. If you need to work with more up-to-date genomes, we recommend you create your own alignment files or request them from the IGB team. For some genomes, we've added probe set alignments to the main [IGBQuickLoad.org](http://IGBQuickLoad.org) site. Mainly we've done this at the request of individual researchers, and so if you would like to request an array, let us know. If probe set alignments are available from our site, you'll typically find them in a folder named **Affymetrix** under the **Data Sources** section of the **Data Access** panel.

You can also make your own probe set alignment files using `blat`, `tabix`, and a python script we wrote. For more information, see [this Bitbucket repository](#).

## About probe set alignment visualizations

Depending on when the arrays were designed, Affymetrix typically used expressed sequences from GenBank to select probes for probe sets - these expressed sequences were sometimes called "exemplar" or "consensus" sequences. They then selected individual probes from regions near the 3' end of the expressed sequence. Affymetrix (as of 2014) distributes probe and target sequences on their Web site, where "target sequences" contain the 3' end regions from which the probes were selected.

Probe set visualizations in IGB show the alignments of target, exemplar, or consensus sequences onto the genome. They also show the locations of probes that were selected from the design sequence. See the preceding figure for an example.

Because probes were selected from the expressed sequences, sometimes a probe will be shown as split across an intron. Also, sometimes probes overlap. And sometimes probes may be missing. If the target sequence contains a region that can't align onto the reference, and if this unaligned sequence contains a probe, then that probe will not be shown.

If you have questions about what you see in a probe set alignment, let us know.

## Why this is useful

Often multiple, seemingly redundant probe sets interrogate one gene. This situation mainly arises when a gene has multiple, alternative three-prime ends due to alternative splicing or alternative termination sites. If an experiment identifies genes where redundant probe sets are differentially expressed with different fold-changes or in opposite directions, this can indicate that the treatment affects splicing as well as the overall abundance of RNAs arising from the gene.

Thus if you observe redundant probe sets that give different or contradictory results, it's a good idea to view them in IGB and compare their alignment to annotated genes and transcripts.

# Sequence Viewer

The IGB **Sequence Viewer** displays sequence data in a text-based window. When you right-click a selected item in the main display window or select a region of genomic sequence, IGB displays a menu offering **View Genomic Sequence**. You can access the same functionality using the **View** menu.


The viewer offers options to display protein translations and reverse complements. You can also copy the sequence to the system clipboard.

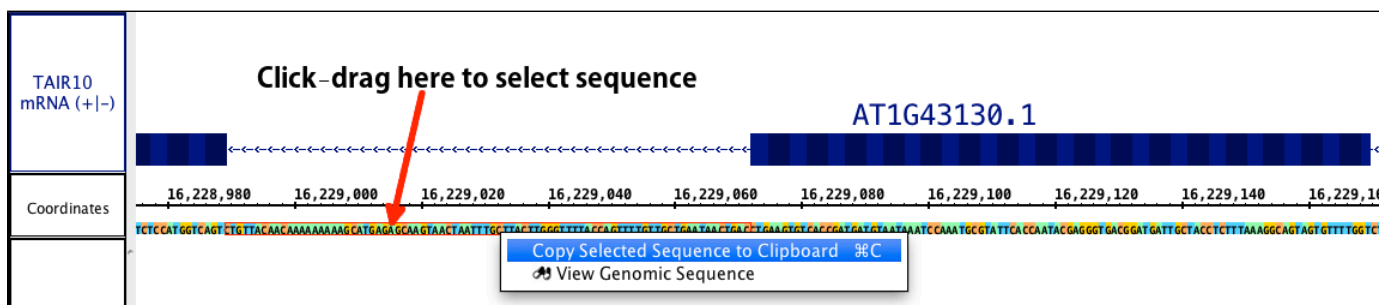
## View Sequence

By default, the viewer depicts bases in yellow on a black background. Use the **Colors** menu to change the color scheme to **Blue on white** (exons in blue and introns in black)

### Selected Sequence or Cytobands

You can look at the selected sequence highlighted in the refseq, or (if present) the sequence of DNA contained in a cytoband. Select the sequence you wish to view by click-dragging the cursor over the sequence bar at any zoom level. A red outline will appear around the selected sequence.

 This method of genomic DNA selection also allows a direct copy to the clipboard, if you wish to skip the Sequence viewer. Simply right click within the selected sequence, and chose **Copy**.




Right click within the selected area or open the **View** menu and chose **View Genomic Sequence**. The sequence viewer will open, and the title will show the genomic range of the sequence, the genome version, and the chromosome number. The **Show cDNA** option will be disabled, but all other options will remain. The DNA sequence will always be presented in the + (plus) direction. However, if your sequence of interest is the other strand, **Show > Complement** will provide the complement of the sequence (this is usable with all sequences).

The sequence viewer also allows you to save the reverse complement of the sequence you are viewing if you wish. Go to **File > Save As FASTA (Reverse Complement)**.

### Gene Models or Other Annotation

For gene models/ annotations, right click on the annotation or open **View** menu and choose option **View Genomic Sequence**.

 To choose a single exon (part of) a gene model, right-click the exon. To choose the entire gene model, right-click any of the introns or the label, if available.



TAIR10  
mRNA (+/-)

## AT1G54330.1

- Search Web
- Get Info
- View Genomic Sequence**
- View Read Sequence
- BLASTX nr protein database
- BLASTP nr protein database
- Select Parent
- Zoom to Selected

Coordinates: 20,279,600   20,279,800   20,280,000   20,280,200   20,280,400   20,280,600

**Right-click to View Genomic Sequence**

Choosing **View Genomic Sequence** opens the Sequence Viewer as shown below. By default, the viewer depicts exonic bases in yellow and intronic bases in white against a black background. The start codon (ATG) is shown with a green box around it, and the stop codon has a red box around it.

AT1G54330.1 : A\_thaliana\_Jun\_2009 : chr1

File Edit Show Colors Show cDNA

0	CTAATCACTTTAATTAATCTCTCTCATCTCTTTCTCTCGTTTTCTAAC
50	GCATACACTTTGAATGGCGCCCATGAGTTTGCCTCCAGGATT CAGGTTTC
100	ATCCCACTGATGAAGAACTGGTGGCTTACTATCTTGATAGGAAGGTCAAC
150	GGCCAAGCCATTGAGCTCGAGATCATCCAGAAGTTGATCTTTATAAATG
200	CGAGCCATGGGACTTGCCTGGTACGATCCAACCTTTTATATCTCTCTCTA
250	TATATGGTT CAGCTTGTTTATGTCAATCTGACGTTATAGATTACATATAT
300	TGGTCATGATGCATTCATACATATGATATTGGAGGTTTTATGAGTTATTT
350	ATTTGGATAATACGCGGAACAGAAAAGTCATTTTTGCCGGGAAACGACAT
400	GGAATGGTACTTTTACAGCACAAAGGAT AAGAAGTATCCAAATGGCTCTA
450	GGACGAACCGTGCGACCCGAGCGGGTTACTGGAAGGCCACGGGGAAAGAT
500	CGTACAGTAGAATCAAAGAAGATGAAGATGGGAATGAAGAAGACACTGGT
550	TTATTATAGAGGAAGGGCTCCTCATGGCCTTCGTACTAATTGGGT CATGC
600	ATGAATATCGTCTCAGCACGCTCCTTCCTCCTCCTGAAGGTAATTAT

The title bar shows the name of the selected item. For instance, the image below shows genomic sequence for gene model AT1G54330.1, from the A\_thaliana\_Jun\_2009 genome version, on chromosome 1.

AT1G54330.1 : A\_thaliana\_Jun\_2009 : chr1

File Edit Show Colors Show cDNA

0	CTAATCACTTTAATTAATCTCTCTCATCTCTTTCTCTCGTTTTCTAAC
50	GCATACACTTTGAATGGCGCCCATGAGTTTGCCTCCAGGATT CAGGTTTC

## Short Reads/ Alignments

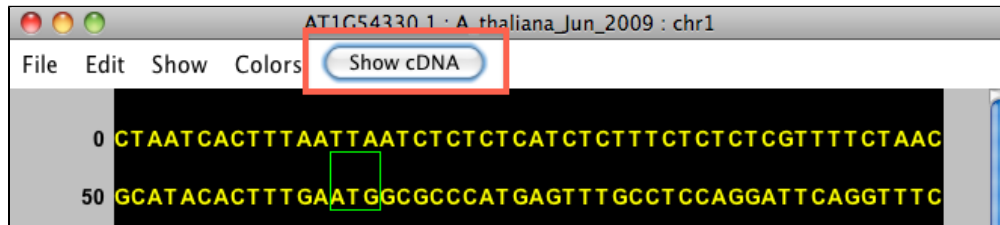
The Sequence Viewer has a special option for short reads. If you select a short read and select **View Genomic Sequence** you will see the

genomic sequence that the read is aligning to. The option **View Read Sequence** will show the sequence of the read itself, including any inserted sequence. If a deletion is present, the deleted nucleotide(s) will appear as a dash (-).

## Additional Features for Annotated Gene Models

### Show cDNA

After selecting a gene model, the Sequence Viewer allows you to select **Show cDNA**, which will eliminate introns from the displayed sequence. We recommend selecting this option before using the translation feature. If an individual exon is selected, this button will be available, but will not change the sequence.



After selecting this feature, you'll see just the exonic (block) sequence, and the button will read **Show Genomic**.



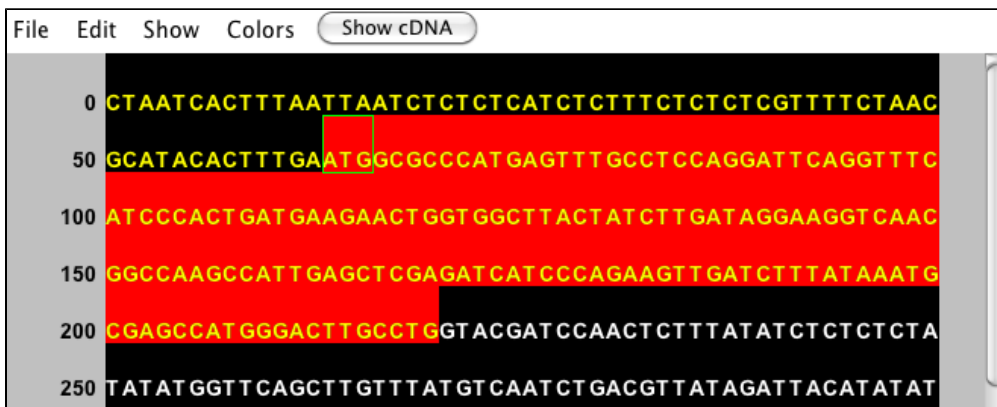
### Show translation

Translation options are accessed from the **Show** menu. **Reverse Complement** provides the reverse complement in place (changes what is displayed). The **Complement** provides the complement under the main sequence. The + (forward) translations appears over the sequence; the - (reverse) translations below. To copy these translations, see **Copy** below

Reverse Complement Complement
+1 Translation
+2 Translation
+3 Translation
-1 Translation
-2 Translation
-3 Translation
All + Translations
All - Translations

## Copying and Saving Sequence

As mentioned, you can copy selected sequence to the clipboard directly from the main window of IGB. Additionally, you can copy a selection or all of your sequence from the sequence viewer. Highlight the sequence of interest (as shown below) and use **Copy > Copy Selected DNA Sequence**. Then open a text editor of choice and paste the sequence.



You can also **Copy DNA with Lowercase Introns and Start/Stop Codons** (NEW!) which copies the DNA sequence with the annotations. When you paste this sequence, you will be able to determine intron and exon boundaries, and find the annotated start and stop codons within the nucleotide sequence. This can be very helpful for PCR primer design or alternative splicing studies.

**Copy Annotated Translation** (NEW!) will copy the annotated translation for the selected gene model - no highlighting needed.

**Copy Selected Translation(s)** (NEW!) requires that you **Show** 1-6 translations. The user then highlights the region of interest (possibly the whole sequence). This function then copies the translation(s). When pasted each translation is concatenated so that all of +1 shows together, then all of +2, then all of +3, etc. with a label indicating which translation is which.

You also have the option to save the sequence directly from the viewer into a fasta file. Fasta files can be shared and used for bioinformatics in this file format, but fasta files can also be opened and viewed from any text editor to access the sequence directly. The file will contain the sequence as well as a line with all of the relevant information from the title of the sequence viewer window. We allow you to save the sequence directly or in the reverse complement, as you need. Open **File > Save As Fasta** or **Save As Fasta (reverse complement)**.

Finally, you can export an image of the viewer directly using the **Save Image** function.

## Graph tracks

Genome graphs associate numerical data associated with base pair positions. You can open and view genome graph data files in IGB and manipulate them using functions in the **Graph** tab.

Using the **Graph** tab settings, you can display graphs as line graphs or heat graphs, change their scale, convert them to floating (draggable) graphs, and perform mathematical or other types types of transformations. You can also compute averages or sums of multiple graphs.

You can create graphs from annotation tracks, such as an RNA-Seq coverage graphs showing the number of read alignments starting at each position.

Examples of genome graphs include:

- Probe intensity values from genome tiling arrays.

- Percentage methylation across chromosomes
- GC content along chromosomes
- Measures of conservation between two genomes

There are two basic types of genomes graphs:

- Position graphs associate scores with single genomic positions.
- Interval graphs associate scores with ranges of genomic positions

## Graph operations and joining and splitting graphs

### Operations - making new graphs from old ones

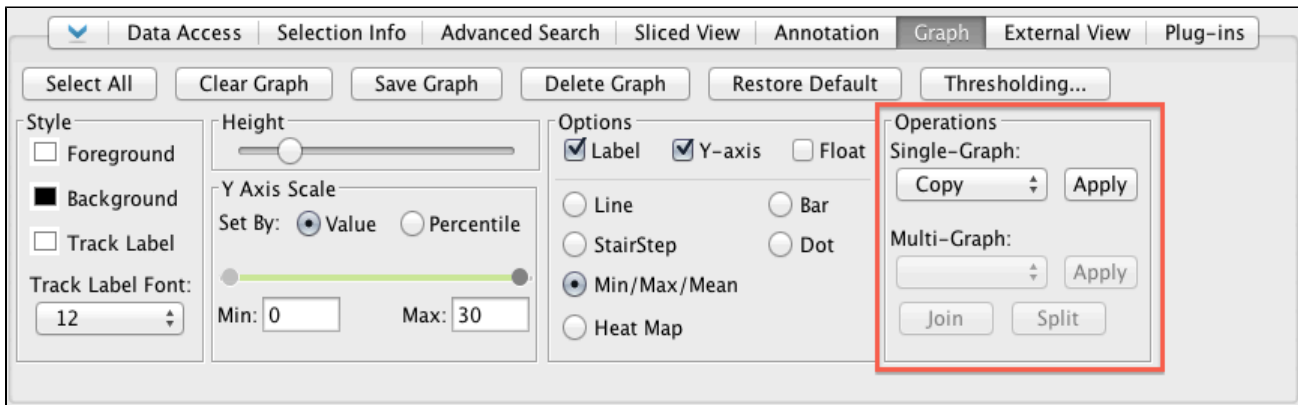
IGB supports multiple types of graph manipulations via the **Operations** section of the Graph tab. To perform a graph operation:

- Select the graph (or graphs) by clicking the corresponding track label. SHIFT-click to select more than one graph.
- Select the **Graph** tab
- Select the operation you wish to perform

Note that single-graph operations operate on single graph tracks. Multi-graph operations operate on two or more tracks.

The result will appear in a new track, labeled with the operation used to create it. When you load new data, the track containing the result of the operation will update.

**Tip:** Save the result to a file by right-clicking the track and select Save As. Only data for the current chromosome will be saved. (This may change in future releases of IGB.)



### Joining and splitting graphs

It is possible to join two or more graphs into a single track. Graphs that are joined into one track share a track label. Operations performed on the joined graph apply to each of the joined graphs.

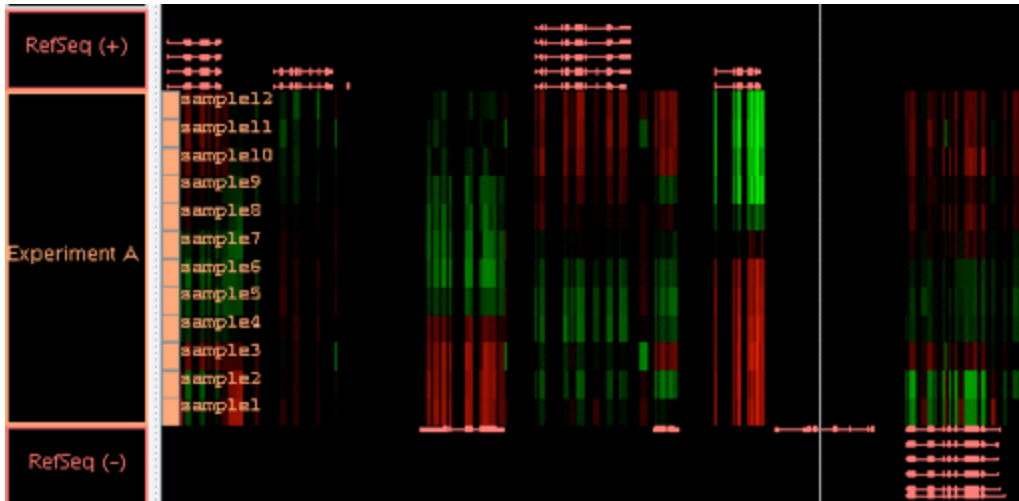
To join graphs:

- Select two or more graphs by SHIFT-clicking their track labels
- Select **Graph** tab
- Click the **Join** button (under **Operations**)

To separate joined graphs:

- Select the joined graph by clicking its track label
- Select **Graph** tab
- Click the **Split** button (under **Operations**)

**Tip:** Use the Y Axis Scale settings to put joined graphs onto the same scale. Click the "collapse" tool (upper left of joined graph track) to overlay and compare the joined graphs. Or set their style to Heat Map to view multiple experimental samples in the same track. For example, the following image shows data from a developmental time course highlighting expression changes over time.



## Applying thresholds to graphs

- [Introduction](#)
- [The Thresholding Window](#)
  - [Direction](#)
  - [Changing the threshold](#)
  - [Specify offsets](#)
  - [Control gaps](#)
- [Make an Annotation Track of the Threshold Results](#)

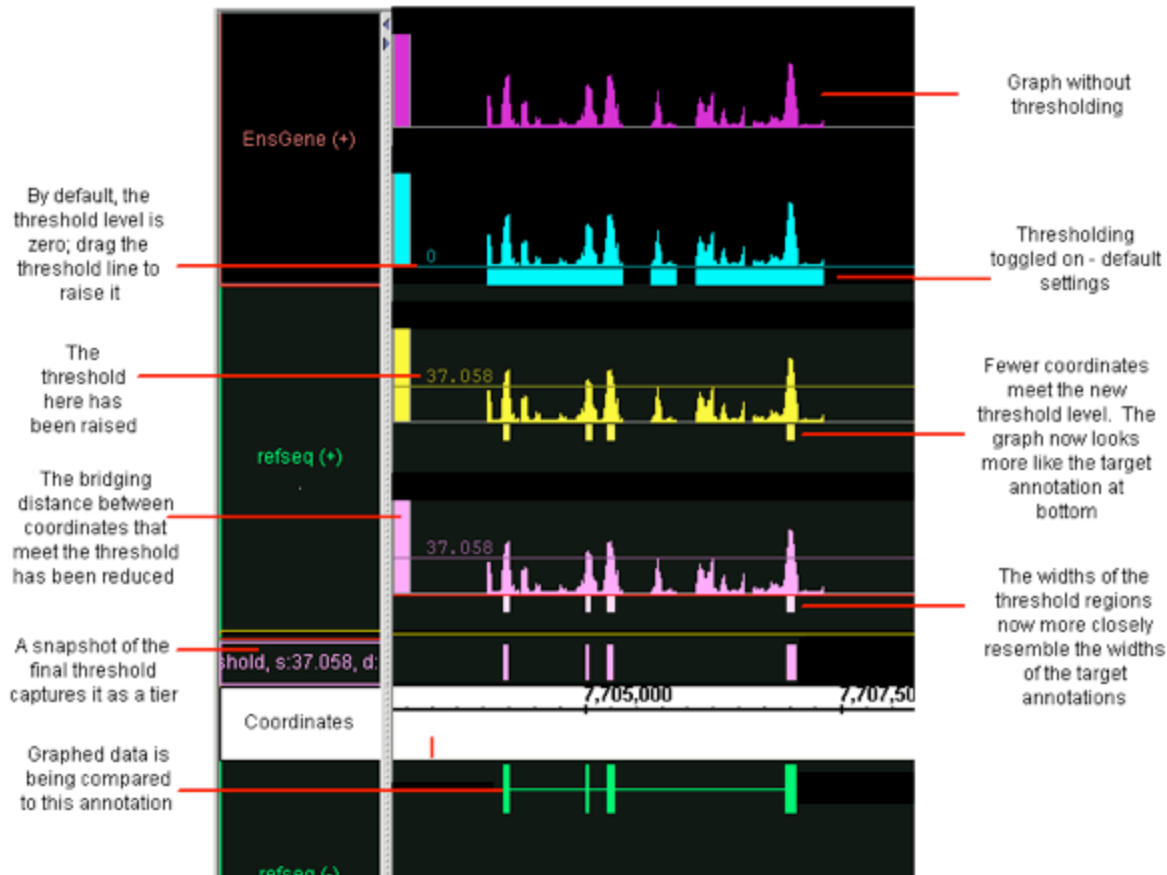
### Introduction

A common analysis step is to determine meaningful regions on a sequence based on graph values being above or below a certain threshold. To screen out non-meaningful values and to view meaningful features of a graph as annotations, use the thresholding feature in IGB. Thresholding displays graph features as annotation-like bars in locations where the graph value at the coordinate meets your defined threshold.

You can convert this data into a standard annotation track in order to analyze it using all of the features in IGB for working with and comparing annotations.

In some cases, you may want to adjust your graph thresholds so the bars correlate to existing annotations; any threshold bars visible at that point that don't have known annotation counterparts (or vice-versa) may indicate areas for further investigation.

Use thresholding to compare graphed data with known annotations



## The Thresholding Window

To use thresholding, select one (or more) graphs. Click the **Thresholding** button in the **Graph** tab panel. This will open the **Graph Thresholds** window.

The first thing you will need to do is set the visibility to 'On'. By default, the threshold is set midway between the graph minimum and maximum values. To obtain more meaningful thresholds, adjust the threshold settings using the methods described in the following sections.

### Direction

You can tell IGB to show you either the data above the threshold (**> thresh**) or below the threshold (**<= thresh**). For example, for expression results graph score represents the level of hybridization activity. For a hybridization activity level to be considered significant, the score must be

above a certain threshold. So you would choose the **> thresh** option and set your desired threshold value. If you need help to pick a threshold level, toggle the Y-axis on to see the graph's scale on the Y-axis. See [Graph](#).

## Changing the threshold

To change the thresholds of selected graphs so that more or fewer points meet the threshold limit, modify these settings in the Graph Thresholds window:

- **By Value** specifies the exact threshold value.
- **By Percentile** sets the threshold based on the percentage of graph values greater or less than your threshold.

## Specify offsets

The **Offsets for Thresholded Regions** feature is designed specifically for working with tiling array data. Captured data from tiling array probes should usually be shifted for display. Since the probes are 25 base pairs long, but the x-coordinates given represent the starting coordinate, you should shift the threshold data so that it starts at 12 base pairs past the given beginning, and ends at 13 base pairs past the beginning. This is the default placement of all graph threshold bars; if you are viewing typical tiling array data, IGB default is set to the correct parameters. If these offsets are not correct for the data you are analyzing, you should change them.

When you view non-tiling array data in IGB, you should adjust the offset to 0 (set **Start** to 0 and **End** to 1) to correctly align the threshold bars with their actual coordinates.

## Control gaps

Experimental data can be noisy. If you are looking for general trends and want to ignore small local variations, you can use the Max Gap and Min run settings. Slide the sliders to adjust, or enter values into the boxes for each parameter:

**Max Gap:** Groups of coordinates that meet the threshold may be separated by gaps where the threshold is not met. By default, thresholding bridges such gaps if they are less than 100 bps long. For example, if there are 300 bps that meet the threshold, then 99 bps that don't, followed by 100 that do, the threshold bar would bridge the 99 bps to produce a single threshold annotation 499 bps long. To change the length of this bridge, enter a different value into the Max Gap text-entry field. Often times, if you have short introns in your species/ gene models, you might need to adjust this lower to have potential introns drawn correctly.

**Min run:** minimum number of bases in a row that must meet the threshold before an annotation bar will appear. For example, by default the Min run is 30 base pairs, so a sequence of 29 base pairs that meet the threshold will not be marked by an annotation bar. Again if you species/ gene models typically have very short exons, you may need to reduce this number. On the other hand, if your species has a minimum size for exon length to be meaningful, you can adjust this higher.

## Make an Annotation Track of the Threshold Results

As shown above, while you are working with threshold values, the 'bars' appear in the same track as the graph itself. However, you may wish to record the threshold results before changing values again; you may wish to work with the thresholding results as a separate track; you may wish to save the results to share with others. For all of these reasons, you can create an annotation (NOT a graph) track of the thresholding results. After you have captured threshold "annotations" from a graph or other set of data, you can examine these "annotations" using all of the tools in IGB for working with and comparing annotations.

After you have set the threshold levels to yield results, you click the **Make Track** button. After an annotation track has been created with **Make Track**, it is frozen and is no longer linked to changes in the **Graph Thresholds** window. However, you can continue to modify the criteria in the **Graph Thresholds** window, altering the threshold results for the same graph track and then create another annotation track.

This track will be given a default label containing all of the relevant parameters, which you can later change if desired. An example name is

"threshold, [60 to infinity] offsets: (+12, +13), max\_gap=20, min\_run=30, graph: depth: HotI2T1"

The name of the new track will always start with 'threshold' and end with the name of the graph track that was used (in this case, the graph was a depth graph made from a .bam track called HotI2T1). The value '60 to infinity' shows that the threshold level was set to 60 and looked at **> thresh**. The offsets were NOT corrected as they should have been, and are set to the tiling array values. Finally, the max gap and min run values are recorded.

## Converting GFF to bedgraph files

- [Introduction](#)
- [Color by score](#)
- [Convert to BEDGRAPH to view the data as a graph in IGB](#)
  - [Step-by-step instructions](#)
    - [Step One: Launch Terminal](#)
    - [Step Two: Change into folder with your files](#)
    - [Step Three: Download python script, tabix, and bgzip.](#)
    - [Step Four: Make the python script, bgzip, and tabix programs executable.](#)
    - [Step Five: Convert your file](#)
    - [Step Six: Sort, compress and index your file OPTIONAL](#)
- [Open the file in IGB](#)

## Introduction

Some analysis programs output scored GFF files, which are GFF (general feature format) files that specify numeric scores for genomic intervals. An example is the Nimblegen/Roche DEVA program for processing genome tiling array files, but there are many other programs that output this file format.

Here is an example:

```
# Roche NimbleGen, Inc. DEVA 1.2. For Life Science Research only. Not for use in
diagnostic procedures.
chr1    DEVA    S1:BLOCK1:Cy5/Cy3    20600676    20600725    -0.3410    .    .
```

The preceding example has two lines. The first line starts with a '#' character indicating it's a comment and contains no data. The second line is contains tab-delimited data for one annotation block from chromosome 1 (chr1). The block starts at position 20600676 and ends at position 20600725. The block has score -0.3410. The remaining fields are not relevant.

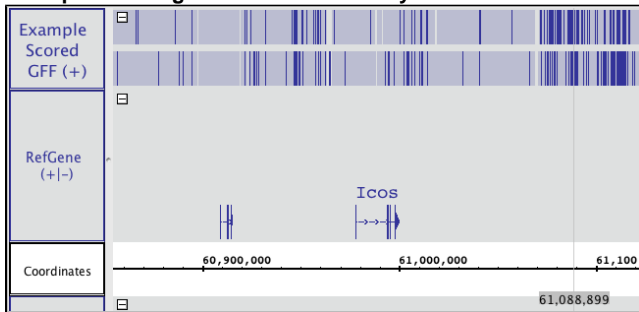
Note that the second field (DEVA) is the name of the program that produced the file.

## Color by score

If you open these files in IGB, the data will appear as individual blocks within a single track. Because the blocks have scores (column six in the file), you can use IGB's **color by score** feature to modify their appearance. The result will be that the highest-scoring blocks will look darker.

To color by score, open the file, load the data, right-click the track label, and choose **Color by Score**.

### Example showing GFF file with Color by Score



## Convert to BEDGRAPH to view the data as a graph in IGB

However, depending on your data, you may wish to view your data as a graph rather than as block annotations. To view your data as a graph, you must convert the data from GFF to a graph format such as bedgraph. The advantage of converting to a bedgraph file is that when you load the data into IGB, the data will be shown as a genome graph instead of as annotation blocks and you will be able to use **Graph** tab functions to manipulate its appearance and perform graph operations on the file, such as calculating averages across multiple files.

## Step-by-step instructions

The following steps describe how you can use UNIX utilities (and a simple python program) to convert a GFF file to BEDGRAPH.

The instructions assume you are using a Mac computer, but they will work on any UNIX-based computer.

### Step One: Launch Terminal

Go to **Applications > Utilities** and launch the Terminal program.

### Step Two: Change into folder with your files

When Terminal starts, by default it's current working directory will be your home directory. This is fine, but in order to convert your files, you'll need to change into the same folder where the file is saved. Use the **cd** (for "change directory") command to change into the folder where your file is saved.

As an example, assume your file is saved in a folder on your Desktop called **my gff files**



To change into that directory inside your Terminal, you would type the following command after the command prompt. Press ENTER to execute the command.

```
cd "Desktop/my gff files"
```

Note that you have to type the quote marks if the folder name contains any spaces.

Now, if you type **ls** (for list files) the Terminal should print the names of the files that are saved in the folder.

For example, assume your folder contains two files: File1.gff and File2.gff. When you type **ls**, the Terminal will print the names of the files it finds.

```
ls
```

This will cause the Terminal to list the files in the directory:

```
File1.gff  
File2.gff
```

### Step Three: Download python script, tabix, and bgzip.

On this page, go to **Tools > Attachments** and download:

- DEVAGff2Bedgraph.py.

If you are using a Mac computer, also download:

- tabix
- bgzip

Save the files to the same folder as the files you want to convert.

### Step Four: Make the python script, bgzip, and tabix programs executable.

Inside the Terminal window, type:

```
chmod a+x DEVAGff2Bedgraph.py  
chmod a+x bgzip  
chmod a+x tabix
```

Press ENTER after each line to execute it.

### Step Five: Convert your file

Run the python script from the command line, giving it the name of the file you want to convert and save the output to a new file.

Type:

```
DEVAGff2Bedgraph.py FILE1.gff > FILE1.bedgraph
```

Where FILE1.gff is the name of the file you want to convert and FILE1.bedgraph is the name of the file you want to create.

At this point, you can open your file in IGB.

However, if your file is very large, you may want to compress it and index it. This will make opening and loading the data into IGB much faster.

### Step Six: Sort, compress and index your file OPTIONAL

Use UNIX sort and your newly downloaded tabix and bgzip programs to compress and index your new bedgraph file.

Type the following lines, pressing ENTER after each line.

```
sort -k1,1 -k3,3n FILE1.bedgraph | bgzip > FILE1.bedgraph.gz
tabix -p bed FILE1.bedgraph.gz
```

The first command runs UNIX sort command and "pipes" the output to bgzip, which outputs compressed data which is then saved to the file named FILE1.bedgraph.gz. The second command creates a .tbi index named FILE1.bedgraph.gz.tbi. Note that the name of the index file is always the same as the file it's indexing. Always store the index ('tbi') file IN THE SAME FOLDER as the file it targets.

**Note:** The tabix and bgzip programs you downloaded were compiled on a Mac and so the above commands will only work on Mac. You can obtain versions for your platform from <http://samtools.sourceforge.net/>.

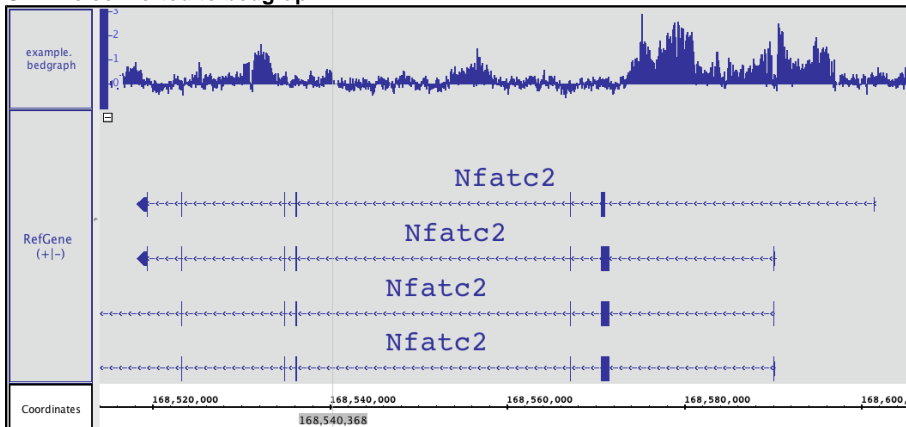
## Open the file in IGB

You can now open FILE1.bedgraph.gz in IGB. Because it's sorted, compressed, and indexed, IGB will be able to open it more quickly than the non-sorted, non-compressed version.

To open the file, choose **File > Open File** and then choose the file. Or click-drag the file into the IGB window.

Press the **Load Data** button to load the data into the IGB window. Once the data are loaded, click the track label to select the track and then choose the **Graph** tab to change the graph appearance.

### GFF file converted to bedgraph



## Creating graph tracks from annotations tracks

- Depth Graph (All)
- Depth Graph (Start)
- Mismatch Graph
- Mismatch Pileup Graph

Graphs can be created by right-clicking on a track's label and highlighting **Track Operations**. A list of graphs and other track operations compatible with the track are displayed and may be selected to create a child track. The track operations that create graphs are:

- Depth Graph (All)
- Depth Graph (Start)
- Mismatch Graph
- Mismatch Pileup Graph

After these graphs are created, click **Load Data** to load graphs for regions that have not yet been loaded. Graph tracks may be saved as .wig formatted files.

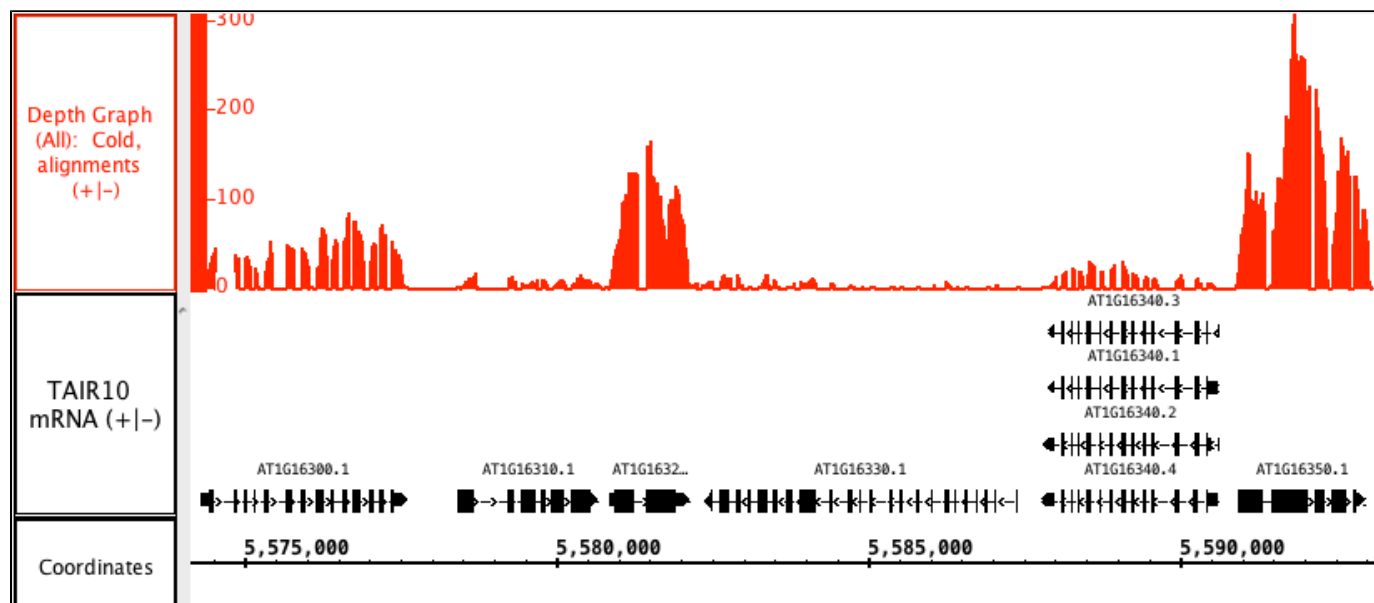
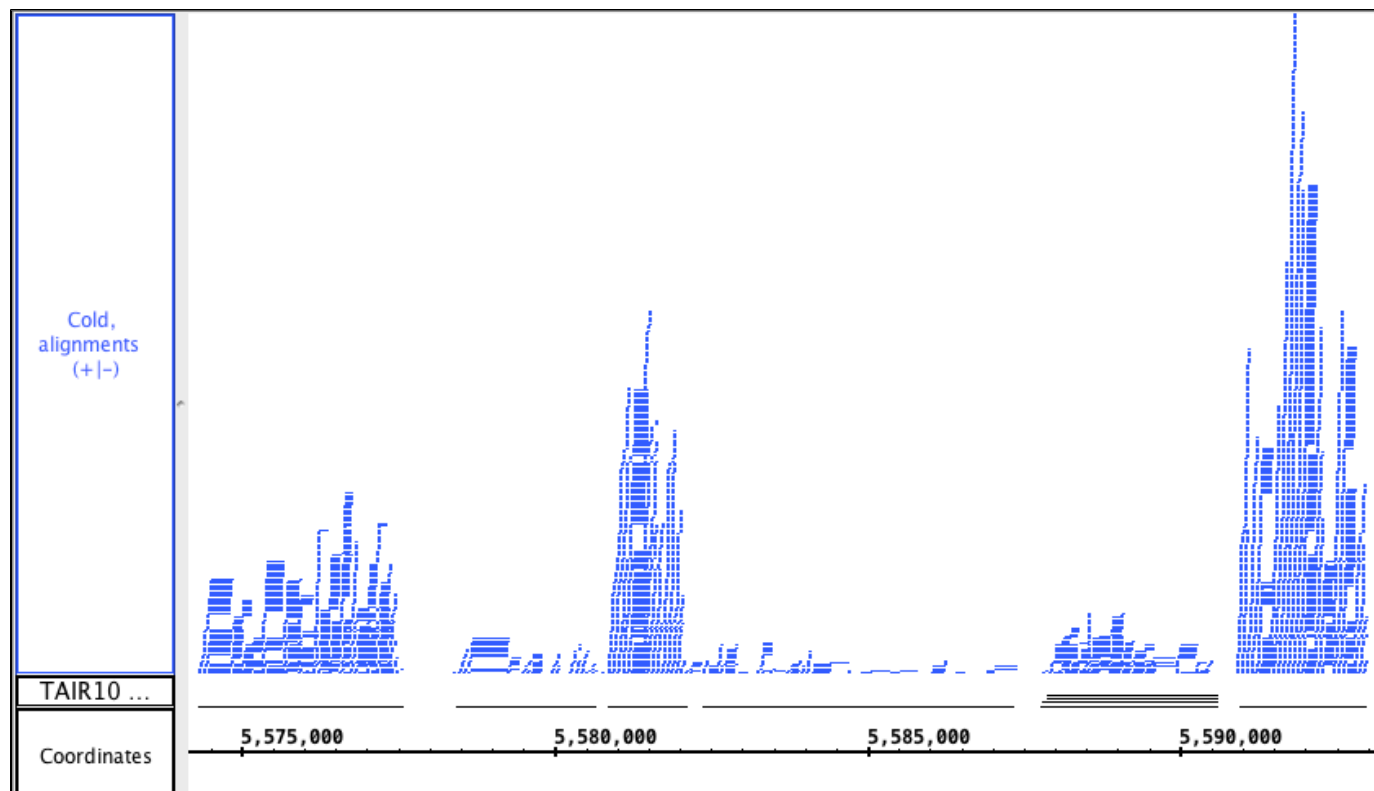
The **Depth Graph** feature works for all annotations but is most useful in combination with a **Mismatch Graph**. This allows you to see how many reads are present at each nucleotide position (**Depth Graph**) and where mismatches occur as compared to the genomic sequence file (**Mismatch Graph**).

## Depth Graph (All)

**Depth Graph (All)** allows you to generate an overview of the relative density of annotation coverage across an entire chromosome (this is also known as a coverage graph). This feature dynamically graphs the number of annotations in a track that are present at each nucleotide across the genomic sequence. When you are completely zoomed in, and each pixel represents a single nucleotide base, the graph shows actual values for the number of reads aligned at that genomic location. Using the **Select** tool to hover over a particular point in the graph will open a tooltip with the actual number of reads at that point.

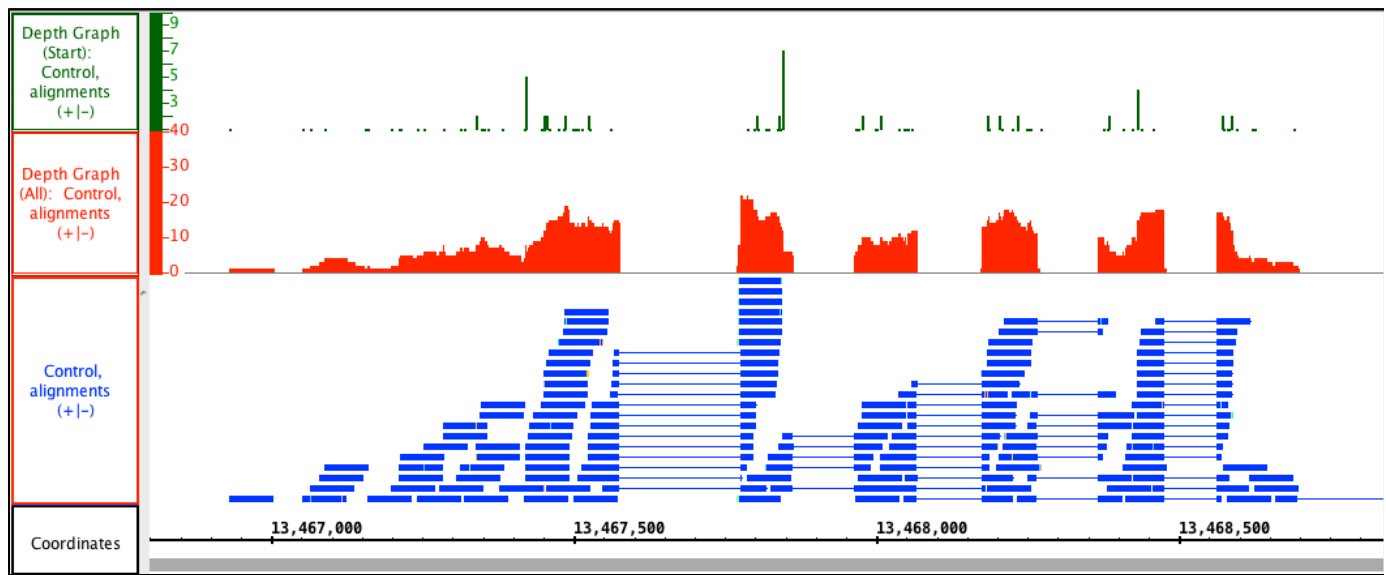
Right-click on the track's label and choose **Track Operations > Depth Graph (All)**.

The picture below shows a fully expanded **BAM** file, showing the difficulty of seeing any details. The second picture shows just the depth graph (\*BAM \*track is 'hidden').



## Depth Graph (Start)

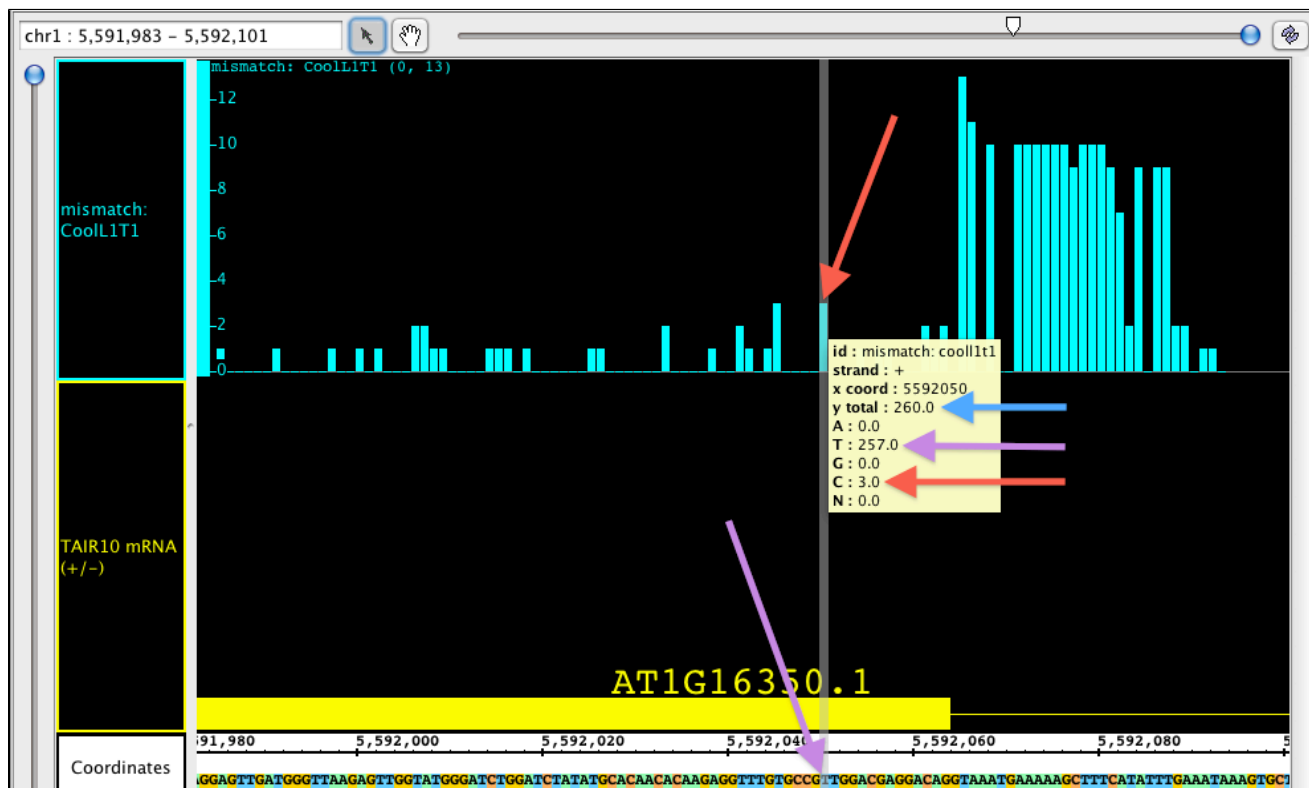
**Depth Graph (Start)** only graphs the first nucleotide of each read. This is most useful when assessing coverage bias in RNA-Seq or other high-throughput sequencing experiments. Often some regions are better represented than others, thanks to PCR amplification bias or other artifacts.



## Mismatch Graph

The **Mismatch Graph** is specific to short read alignment files, such as BAM. The graph shows only the number of mismatched nucleotides across all reads at a specific genomic location, which can be very helpful in the detection of allelic variation, SNP identification and also error checking. Right click in the label to access the menu; choose **Track Operations > Mismatch Graph**. This function requires that the genomic sequence be loaded; if it is not loaded, IGB will load the sequence for you.

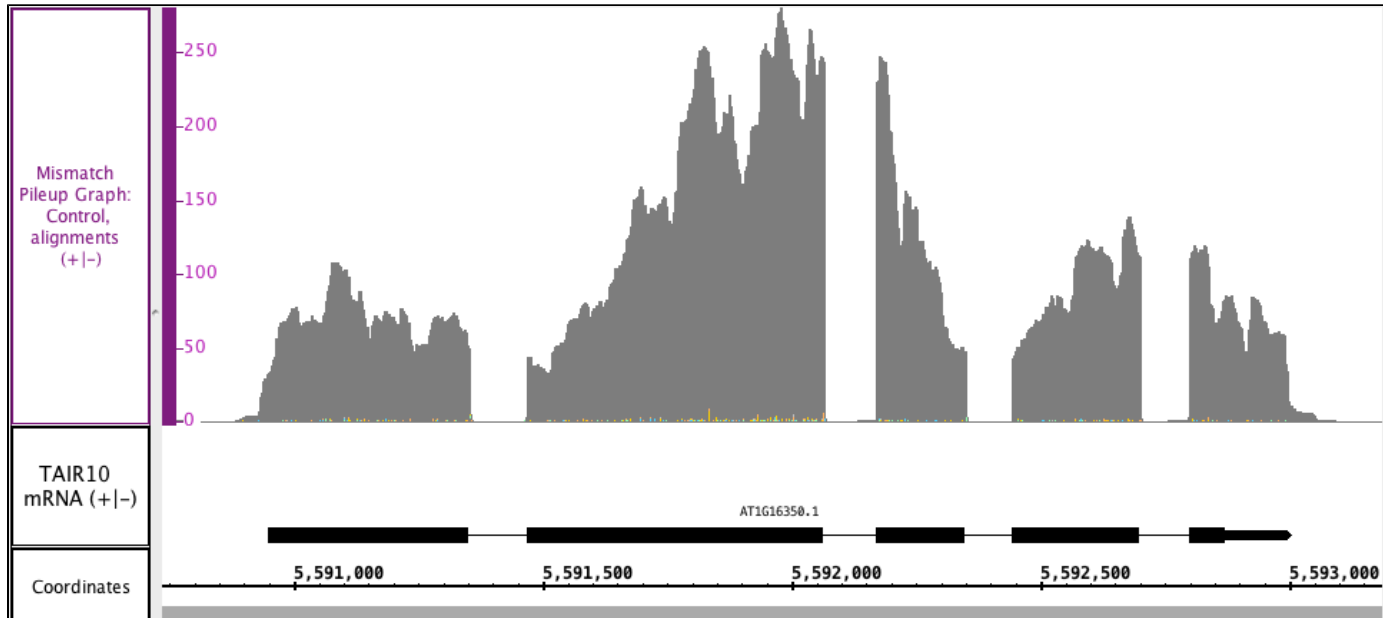
If you hover the **Select** tool (arrow cursor) over an individual bar of the graph, you will get a tool tip showing the total number of reads, and then the number of reads broken down by nucleotide. In the picture below, you can see that the total number of reads at the specified position is 260 (blue arrow). Of those reads, 257 at 'T', which you can see is the matching nucleotide (purple arrows). 3 of those reads contain a 'G' instead; since G is a mismatch, the graph shows a bar with height of 3 (red arrows).



## Mismatch Pileup Graph

In the **Mismatch Pileup**, you have the **Depth Graph** portion represented as a dark gray graph, with the **Mismatch Graph** overlayed in the matching nucleotide colors. The colors show which and how many of each nucleotide(s) are present at each mismatch position. The **Mismatch Pileup** graph, like the **Mismatch Graph**, is also specific to short read alignment files, such as BAM. Right click in the label to access the menu; choose **Track Operations > Mismatch Pileup Graph**. This function requires that the genomic sequence be loaded; if it is not loaded, IGB will load the sequence for you.

In the image below, you can see the gray portion is a **Depth Graph**, but you can also see the mismatched nucleotides (very small numbers in this section), which are color coded by G, A, T and C.



## Annotation tracks

Annotation tracks contain gene models, short read alignments, EST alignments, or other genomic features. IGB allows you to change how annotations appear - you can change colors, change the Track Label names, and so on.

IGB also allows you to create new tracks by combining existing tracks. For example, you can create a new track that represents the intersection of two existing tracks. This is useful when looking for new novel genes or splicing variants, for example.

## Customize track appearance

- Introduction
- Track Colors
- Track label text
- Track label width
- Edge Match color
- Annotation labels
- Strand controls
  - Combine or separate plus and minus strand annotations
  - Use Arrows to indicate strand
  - Use color to indicate strand
- Collapsing and expanding tracks
- Stack height
- Using color to represent scores

## Introduction

IGB offers many options to customize the appearance of annotation tracks.

IGB remembers your customizations from session to session.

Some of the available customizations include:

- Annotation color

- Track background color
- Track label text, color
- Track height (pixels)
- Annotation stack height
- Annotations labels (e.g., using score or gene name)
- Changing track height and position in the display
- Changing strandedness of the tracks
- Changing the amount of data shown within the track (Stack Height)

You can change appearance of annotation tracks

- By clicking Annotation tab
- By right-clicking a track label
- By opening **File > Preferences > Tracks**

You can also set the default values for newly opened tracks based on file type using the **Track Defaults** tab in the **Preferences** window.

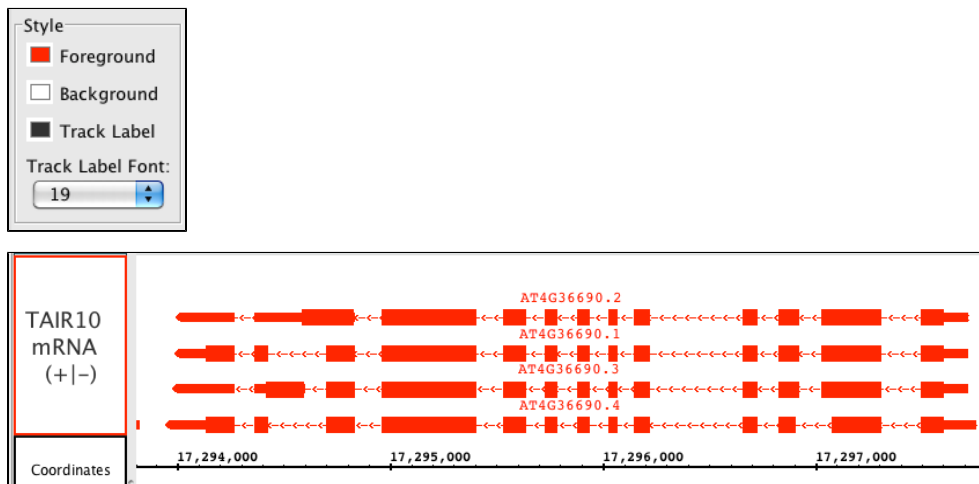
## Track Colors

In IGB you can modify the color schemes for annotation and graph tracks. Using controls in the **Annotation** and **Track Preferences** tabs, you can

- change annotation and graph colors (foreground, FG)
- change background color (BG)
- change the color of the track labels

Unless you change it, the track label font color always matches the foreground color.

The following images shows the **Style** settings from the **Annotation** tab for a selected track. Note that the track label outline is red, indicated the track is selected.



## Track label text

All annotation tracks have a track label in the panel on the left side of the display. You can change several aspects of how the label appears. You can

- change the text (called the "display name") displayed on the label
- change the track label text font size and color

To change the text in a track label, select the track by clicking its track label and

- choose **Preference > Tracks** tab and enter new text under the Track Name column
- or choose the **Data Access** panel and enter new text under the Track Name column in the **Data Management Table**

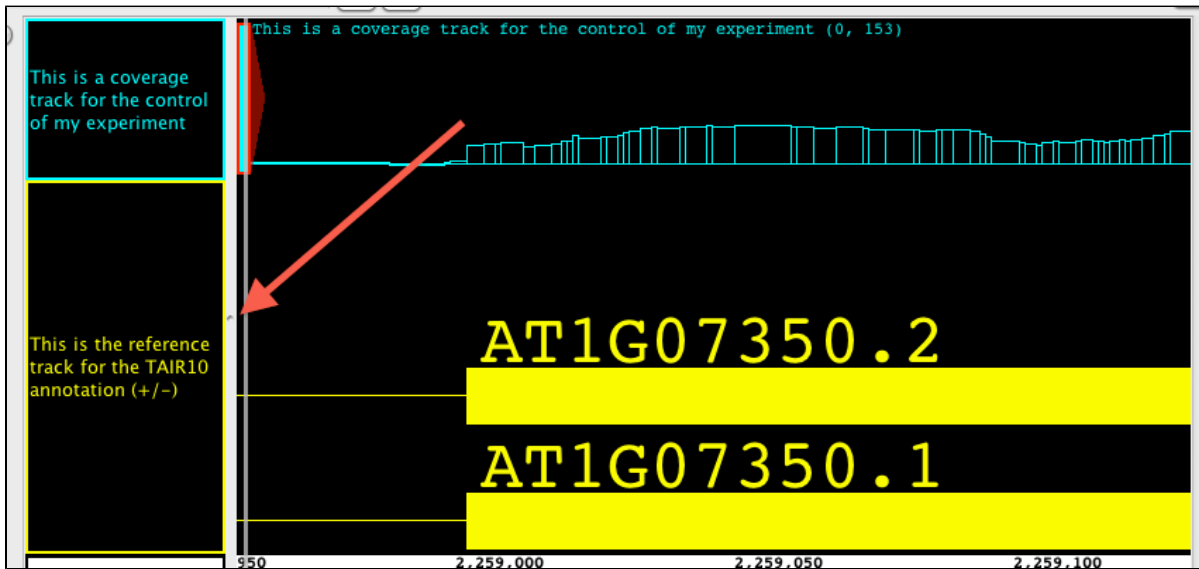
To change the track label font size, select the track by clicking its track label and

- choose **Preferences > Tracks** and enter a new value using the Track Label Font menu
- or choose the **Annotation** tab and enter a new value in the **Style** box

## Track label width

You can also change the width of a track label. To increase or decrease track label width, drag the vertical bar that separates the labels from the

data. The red arrow in the image below shows the 'grab point' which can be used to change the label column width.



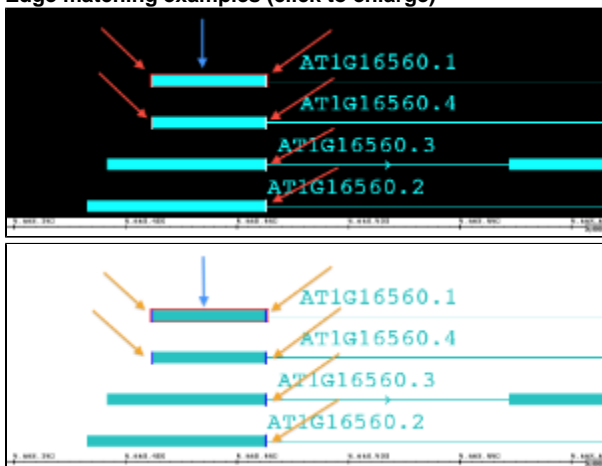
## Edge Match color

If you use white or light gray shades for backgrounds, you may want to change the color of the **Edge Match** graphics so that it is easier to see.

To change the edge match color

- Open **File > Preferences > Other Options**
- Click on the swatch labeled **Edge Match Color**

Edge matching examples (click to enlarge)



## Annotation labels

Every annotation has a set of associated attributes, such as its ID, score, name, gene name, description, start and end positions, and so on.

You can configure IGB to display any one of these attributes next to an annotation.

To change what values are shown as the annotation label

- Select the track by clicking its track label
- Click **Annotation tab**
- Choose an option in the **Label Field** menu



Choose score field for junction BED files created by TopHat

## Strand controls

## Combine or separate plus and minus strand annotations

IGB can show plus and minus strand features in separate tracks or combined into one track. To combine plus and minus strand annotation into a single track

- Select the track by clicking its track label
- Click **Annotation tab**
- Click +/- checkbox in the **Labels and Strand** sub-panel

## Use Arrows to indicate strand

IGB can draw arrows within introns that indicate strand. To use arrows to indicate strand

- Select the track by clicking its track label
- Click **Annotation tab**
- Click **Arrow** checkbox in the **Labels and Strand** sub-panel

## Use color to indicate strand

IGB can indicate annotation strand using color. To turn this option on or off or choose colors to use for minus and plus strand feature

- Select the track by clicking its track label
- Click **Annotation tab**
- Click **Color by Strand** checkbox in the **Labels and Strand** sub-panel
- Select colors by clicking the **Color by Strand** swatches

## Collapsing and expanding tracks

Collapsing a track merges all annotations in the track into a single row.

To collapse a track

- right-click the track label and choose **Collapse**
- click the collapse track icon in the top left corner of the track

## Stack height

The stack height setting controls the number of annotation rows that appear within a track. If there are more annotations than can be shown, IGB draws these in a single summary row in the top or bottom row of the track, depending on whether plus or minus strand features are being shown.

To change stack height, right-click the track label and

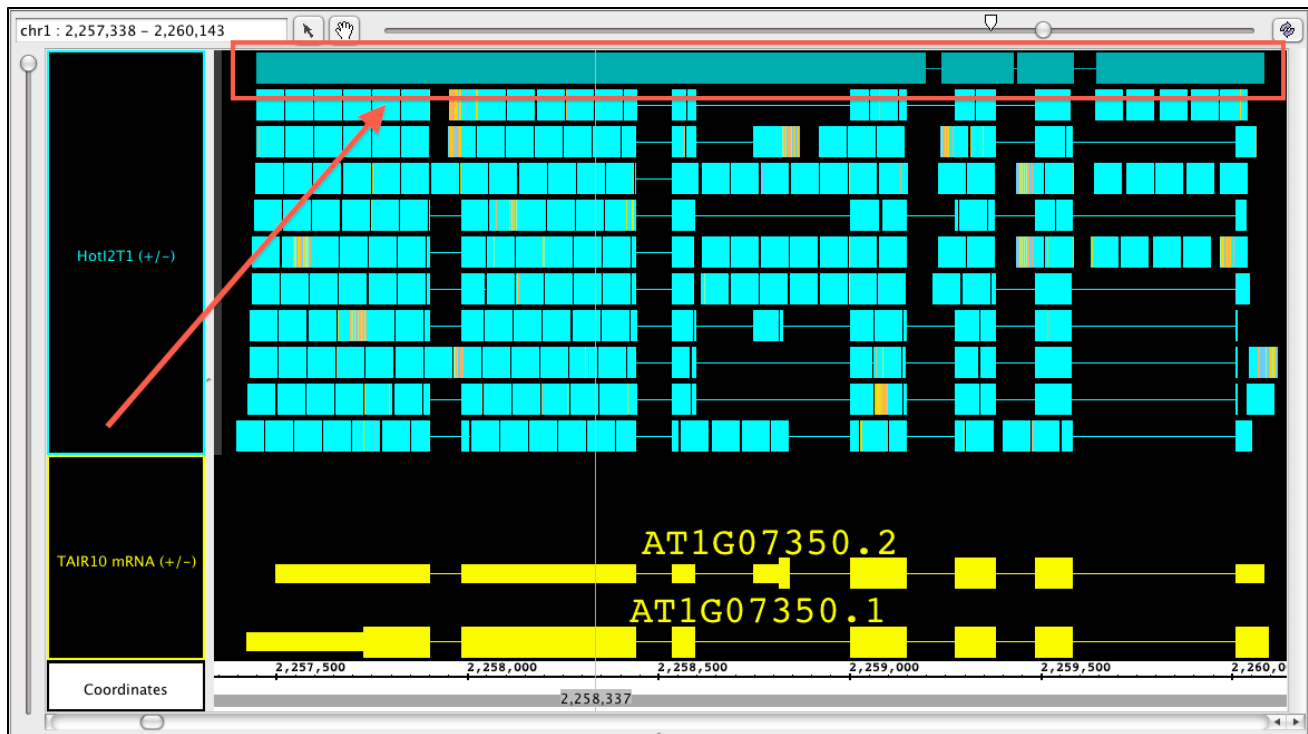
- choose **Set Stack Height**
- choose **Optimize Stack Height** to set a value that is optimal for the current view, e.g., ensure that all annotations will be drawn individually as part of a stack

or

- click the track label (to select the track) and click the corresponding stack height icons in the toolbar

In the image below, the **Stack Height** is set to 10. The red arrow points to the summary row.





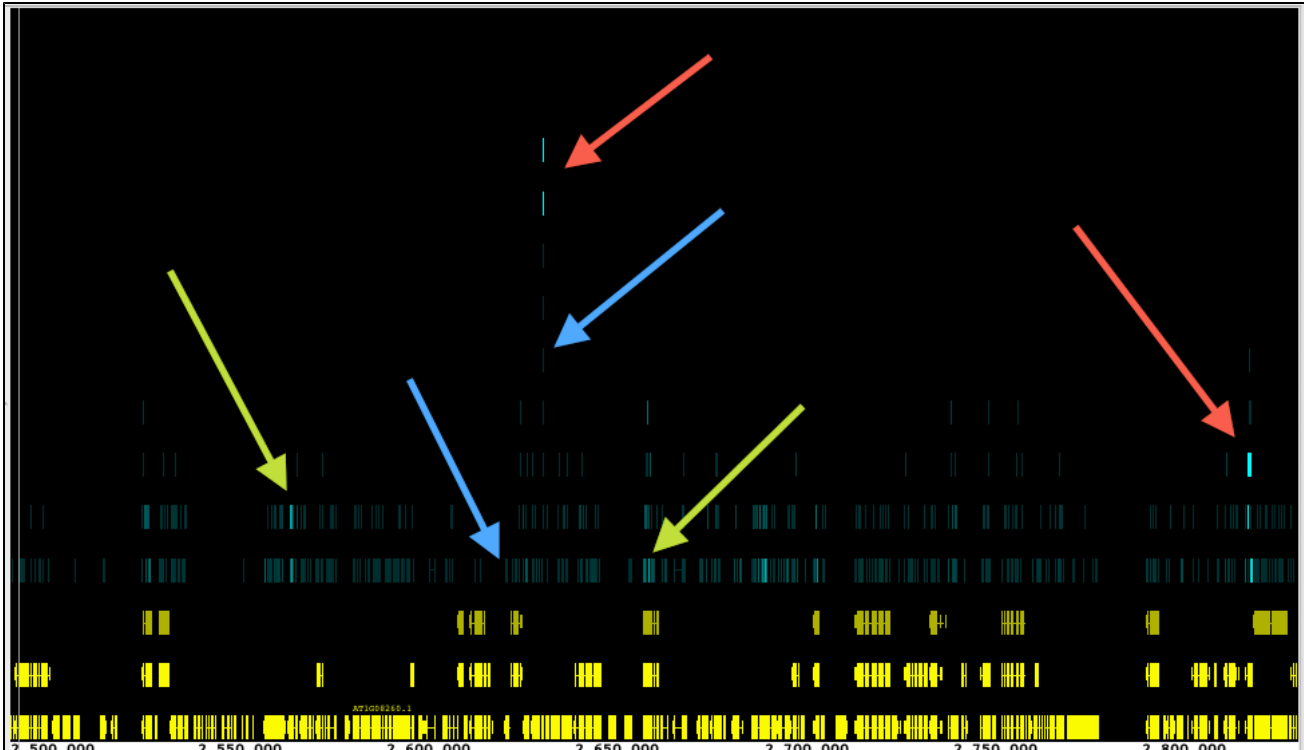
## Using color to represent scores

For annotation data that contains scores, IGB can display the data with color intensity reflective of the scores. The scores must have been included in the data when it was loaded. Files in GFF and BED format allow the specification of scores. To turn color-by-score on or off, right-click the track name and choose **Color by Score**.

The default range that IGB uses is 0 to 1000+. New for IGB 7.0, we have added the ability to Set Range of Color by Score. This allows the user to set both the top, as well as bottom scores to best display their data. You can set the lower value so that any data with a score below your threshold will be barely visible; you can set the upper score to the highest score in your data, or any other value. After setting the lower and upper values, IGB will scale the color intensity across the new range so that the highest values are most intense, and the lowest values are the least intense (see image below).

This function has no effect on data that doesn't contain scores.

In this image, exons are the scored value. Exons with the highest scores will display in the chosen foreground color (**red arrows**). Exons with the lowest scores will display in a color intensity that blends with the background color (**blue arrows**). All other scores will be displayed at color intensities associated with their numerical score, where mid range numbers will show with mid range color intensity (**green arrows**).



## Annotation track operations

- Introduction
- Single track functions
  - Not
  - Copy
- Multi-track functions
  - Intersect
  - Union
  - A not B, B not A
  - XOR

## Introduction

IGV contains several functions that allow you to create all new annotation and graph tracks from existing tracks.

These functions can help identify new genes, new splicing variants, and help expose biologically meaningful aspects of your data. You can save the new tracks as BED format files, which you can re-open later in IGV and share with others.

There are two types of track functions

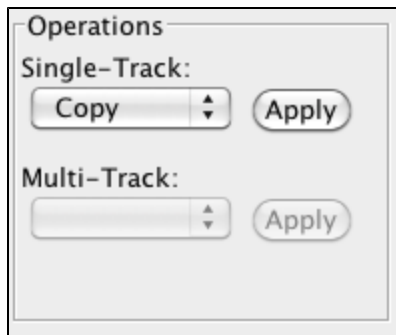
- Single-track functions that apply to one track only
- Multi-track functions that apply to two or more tracks

## Single track functions

To access single-track functions

- Select a track
- Click Annotations sub-panel
- Choose a function from the **Operations** sub-panel

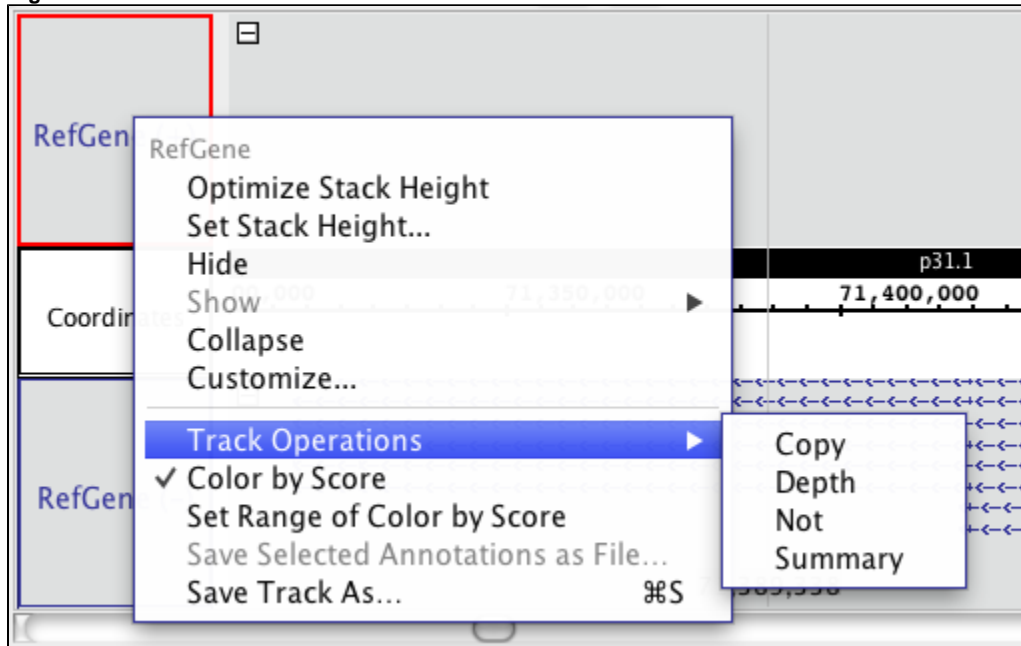
### Annotation tab



OR

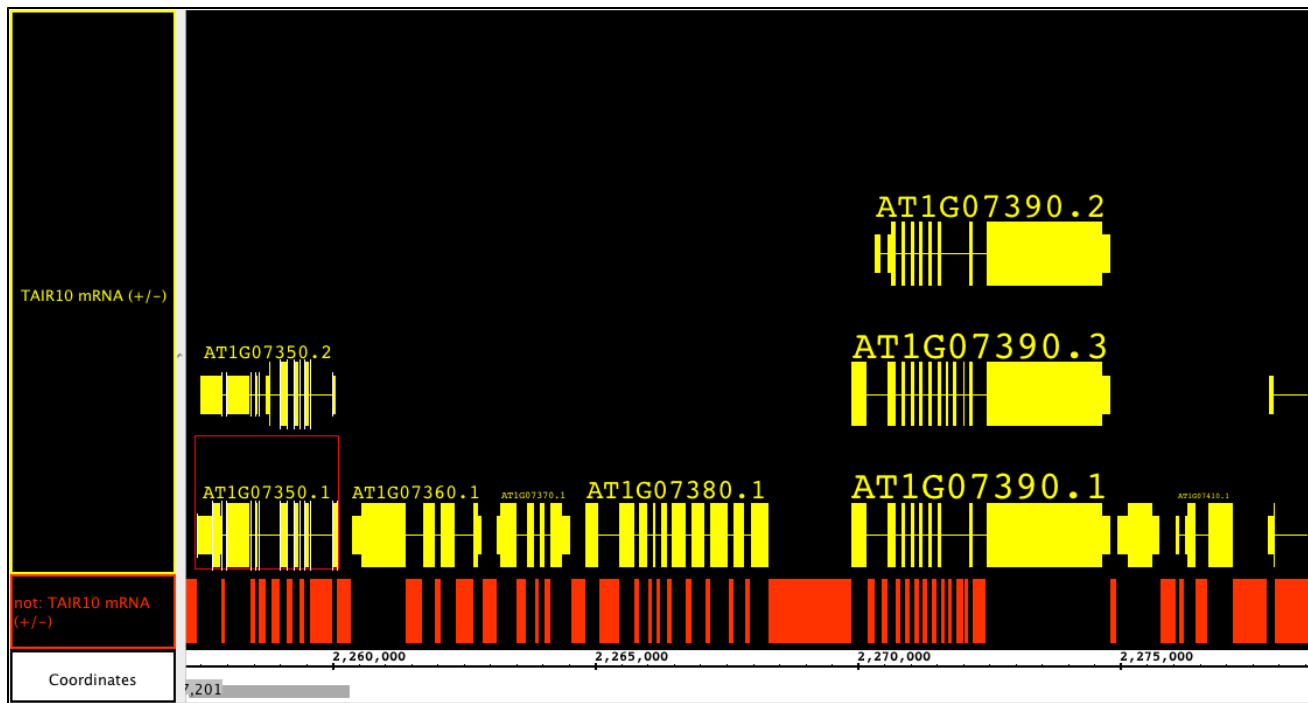
- Right-click a track label
- Choose **Track Operations**

Right-click menu



Not

**Not** creates a new track with annotations representing empty regions from the selected track.



## Copy

This function simply produces a copy of the original, with no changes to the annotations.

## Multi-track functions

To access multi-track functions

By performing the functions below, IGB can provide new annotation tracks showing a uniquely specified/analyzed data set. This can be merged annotations, or annotations of areas where only one or the other track have data. This powerful set of functions can help discover new features of all types within the data. For instance, selecting a RNA-seq experimental track followed by a reference sequence track and choosing **A not B** will show you all of the places where there are RNA reads that do not align to the annotated genome; all of the annotations in the newly created track could represent retained introns, new exons, new exon-intron boundaries, and even whole new genes.

These actions are specific to the tracks and strands that you selected, i.e. you can select just the + strand or just the - strand of a data set. This can be used to your advantage; for instance, if you want to know where on the chromosome there are annotations that overlap on the forward and reverse strands, then choose the + strand and the - strand and use the **Intersect** operation.

Another place where performing track comparisons is useful is when you are comparing treatments of samples versus controls. Depending on the function chosen, the new tracks will show differences or similarities, or even regions left completely unaffected. This can even be used with thresholding tracks created from graphs. By comparing 2 different tracks with different threshold settings, these operations can show you what is different between the two cases. To capture the graph thresholds in a track, see [Applying thresholds to graphs](#).

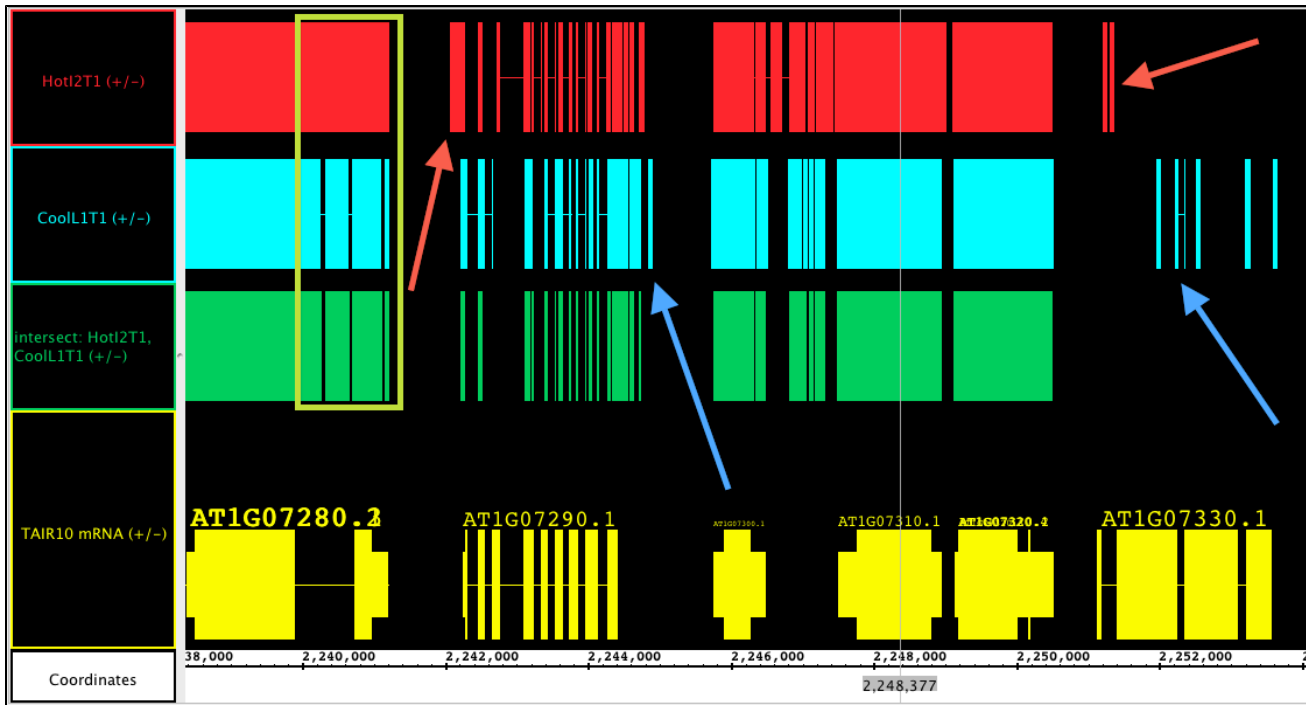
To combine tracks:

- Select two or more tracks by clicking and SHIFT-clicking the track labels. The first track selected is designated 'A' and the second track is 'B'.
- Right-click a selected track label and choose **Track Operations**.

Choose one of the options listed. The effects of each option is described below.

## Intersect

Yields only the overlapping portions of A and B. This is good for looking at similarities between tracks. Although not quantitative, this is a good way to find out what similar elements are present or not. The picture below shows the intersection track, in green, of two .bam files (shown collapsed) and their reference sequence (also collapsed). Here you can see that they generally have the same genes expressed, even though they may be expressed at different levels. The red arrows are pointing out where **ONLY** the treatment track has reads; blue arrows show only control reads. The green box highlights the fact that if either track has gaps (such as introns) they will show in the intersection track.



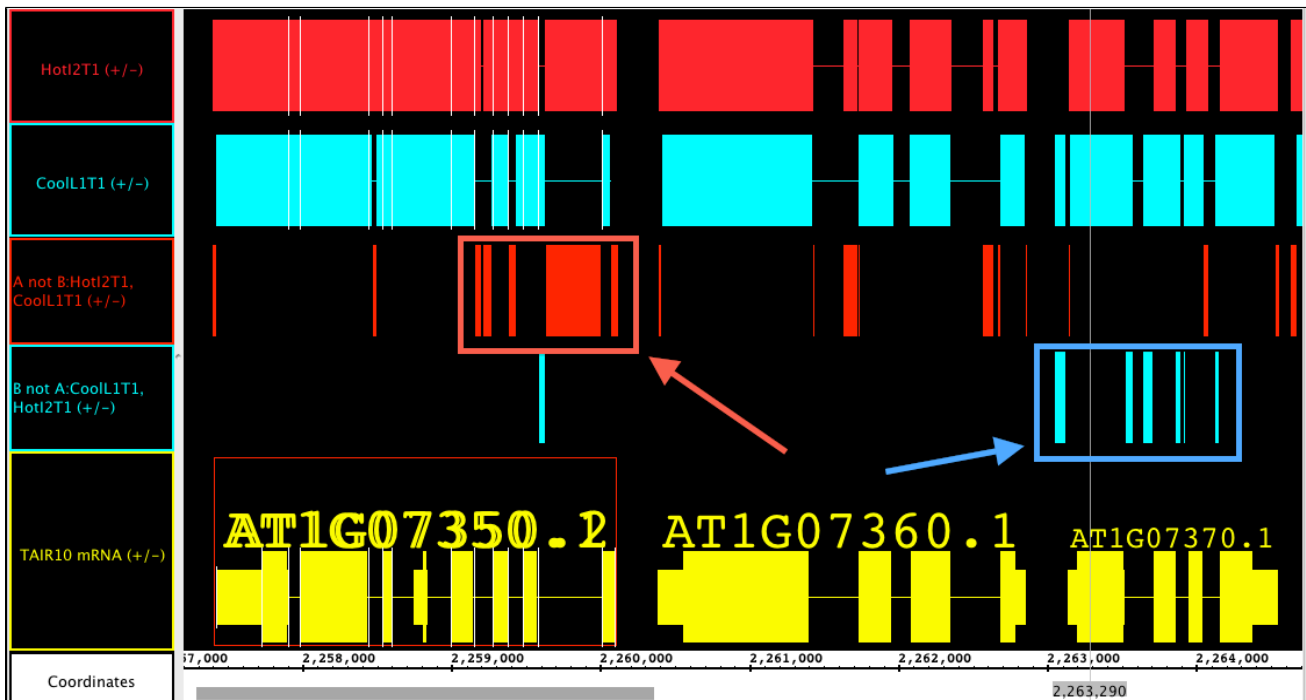
## Union

Yields all bases that are in either A **OR** B

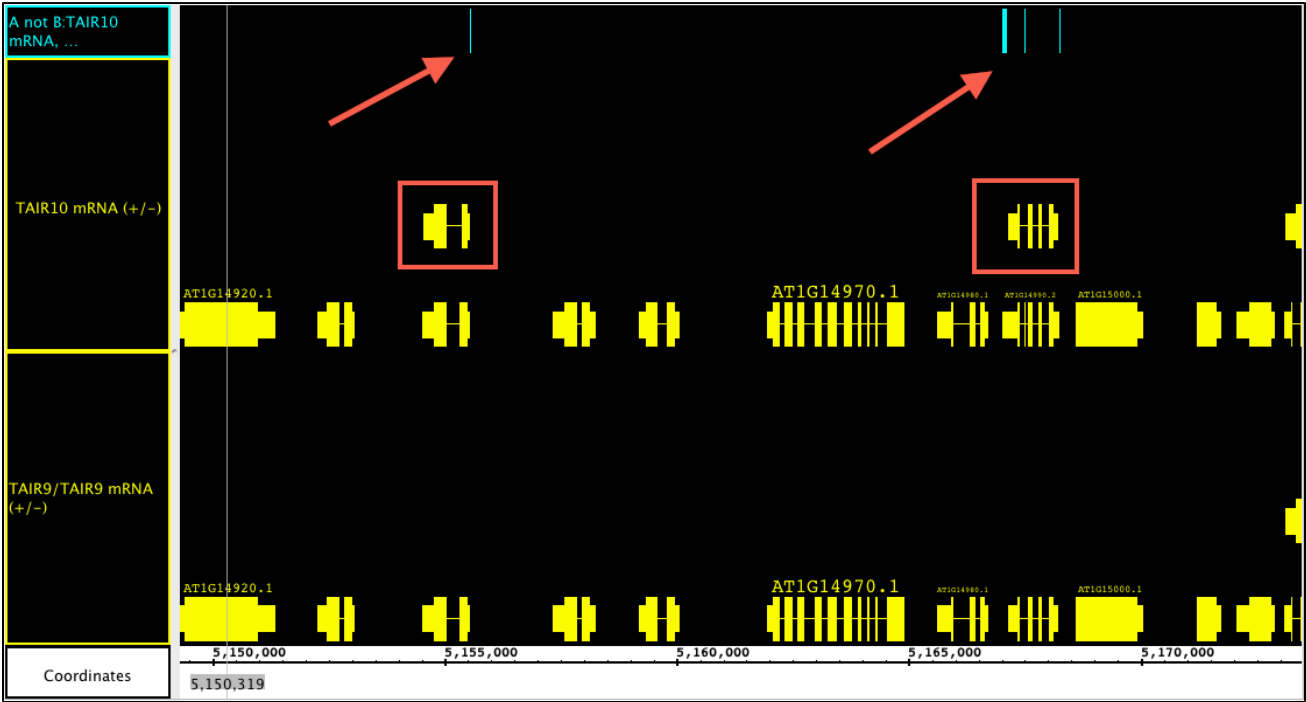
## A not B, B not A

**A not B** and **B not A** tracks show the part of track 'A' that does not overlap with track 'B' and vice versa. These tools are incredibly powerful for finding differences between two tracks. You can find differences due to treatments, between expressed sequences (RNA-seq) and annotated tracks (reference sequences) or between different versions of data sets. There are many other applications as well.

To use these functions in IGB, select track 'A' first, and then track 'B'. The new track label indicates the function and the order of the track selection, i.e. in the image below the track label "A not B:Hotl2T1, Cool L1T1(+/-)" indicates **A not B** was used to generate the track and that Hotl2 is 'A' and CoolL1 is 'B'. In this image, the **A not B** track shows all of the instances where there is data present in Hotl2 that are not matched in CoolL1 (red box, red arrow). Meanwhile, the **B not A** track shows CoolL1 data that is not present in Hotl2 (blue box, blue arrow).

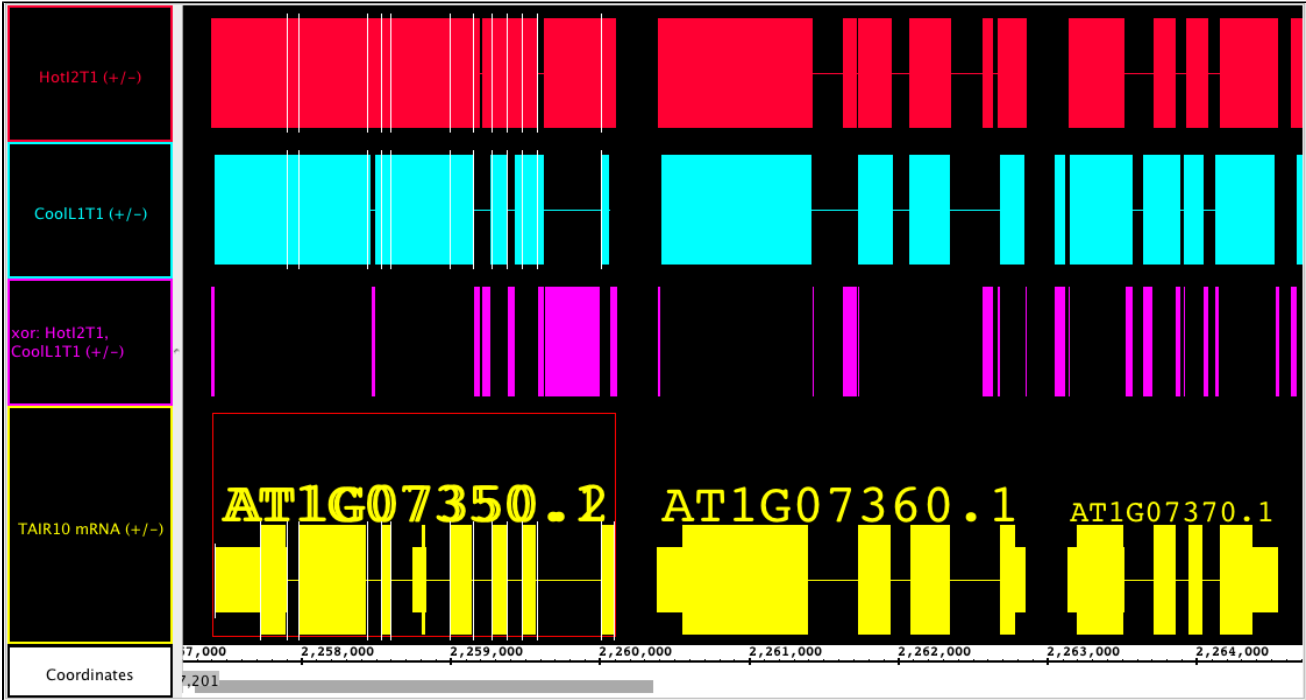


In this track, we compare TAIR10 annotations to older TAIR9 annotations to find differences. You can see that there are 2 new splice variants included in this region, but that they do not look very different from the original by a quick glance. the **A not B** track identifies exactly where they are different.



XOR

**XOR** tracks shows the union of the material that is in A or B but not both; this is like a summary track of both **A not B** and **B not A**. Order of track selection is irrelevant here. While this track summarizes the differences into one track, you cannot determine which of the original tracks contains the difference, but it makes an excellent way to skim across data looking for general regions of difference.



Customize annotation colors using Color By

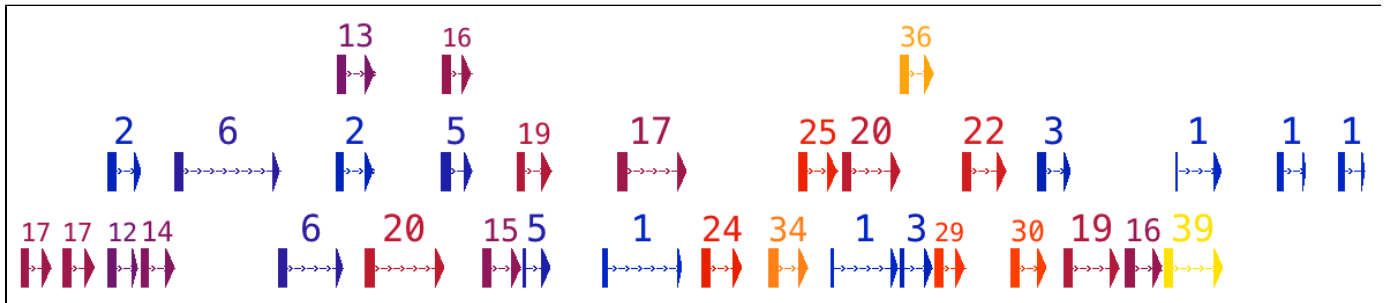
- Introduction
- How to customize annotation colors using Color By
- Attributes and Options for setting annotation color

- Color by Property
- Coloring by numeric values: Using the heatmap editor
  - Heatmap Handle Settings panel

## Introduction

The **Color By** feature in IGB lets you customize annotation color based on annotation attributes like **length**, **property**, **RGB**, and **score**.

For example, here is an image from IGB showing splice junction features colored by their score. Redder features have higher scores and blue features have smaller scores.

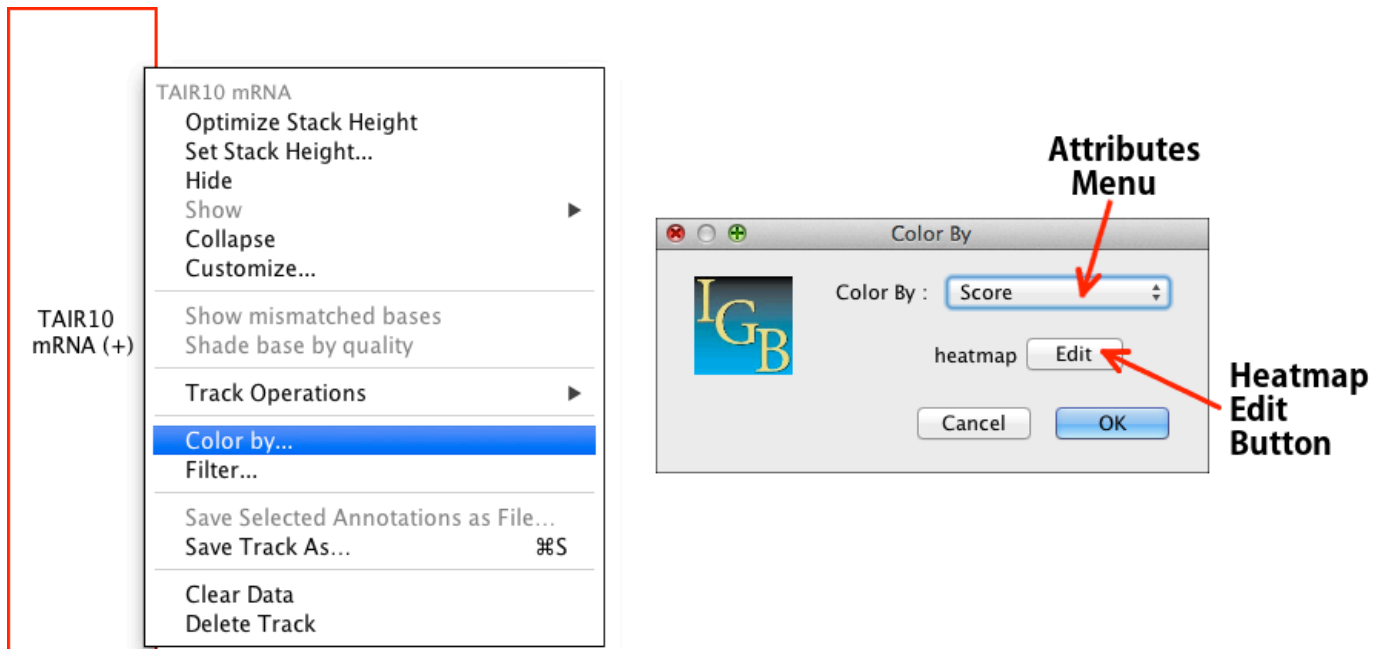


## How to customize annotation colors using Color By

To customize annotation colors in a track

- Right-click the track label
- Select Color by to open the **Color by** dialog
- Choose an attribute from the menu
- Select options and colors. Note that available options depend on attribute type.

Figure 1. Right-click annotation track label and select Color by... to open Color by dialog



## Attributes and Options for setting annotation color

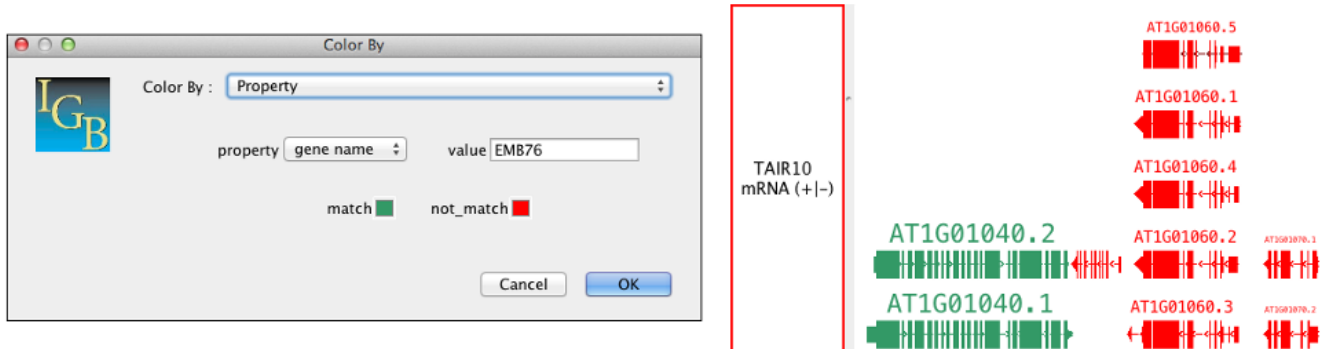
The menu in the **Color By** dialog shows annotation **attributes** you can use to configure annotation colors. The attribute names you see will depend on the type of data set (or file) you loaded into the track. When you select an attribute, the **Color by** dialog will then show you options for choosing colors. The options you see after selecting an attribute will depend on the **attribute** type - whether it's a number (like score) or a

category (like name).

## Color by Property

Coloring by **Property** is unique in that it lets you select a property from a drop down menu, enter a value for the property, and then select a color that indicates a match or a mismatch.

To select a color, click a color swatch and then choose a color from the color palette that appears. Click the **OK** button to confirm your selection.

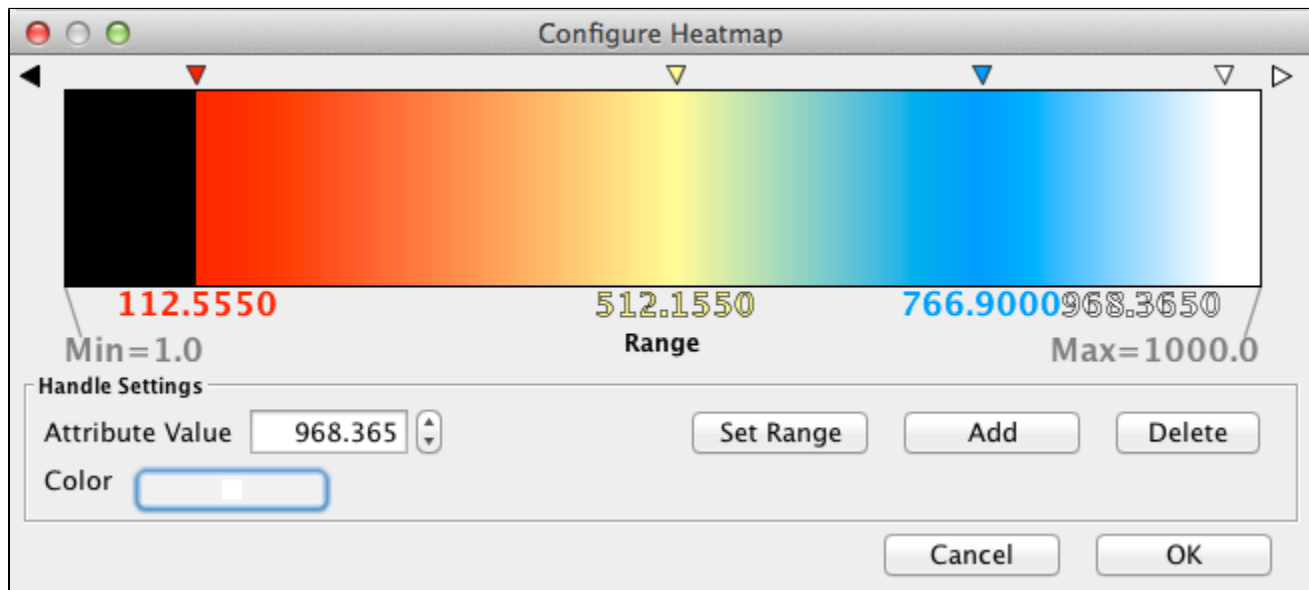


## Coloring by numeric values: Using the heatmap editor

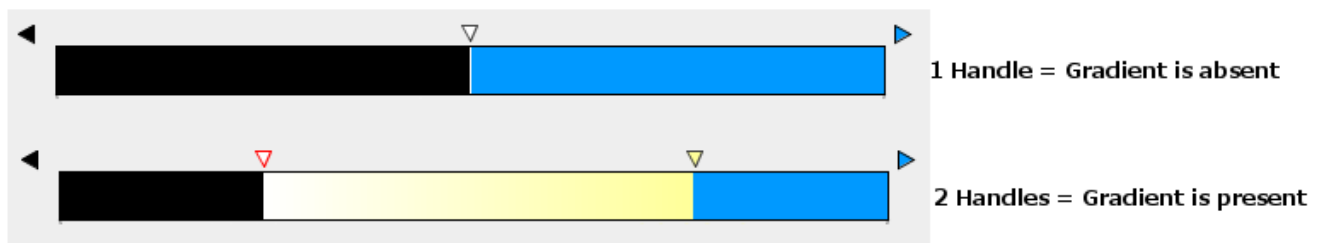
You select a numeric attribute (like score) to color by, then the **Color by** dialog will show a heatmap **Edit** button that lets you color annotations using customizable heatmap.

To use the **Heat Map Editor**:

- Click **Edit** in the **Color by** dialog to open the **Heatmap Editor**
- To set **Min** and **Max** values, click **Set Range**
- Use the color handles (triangles above the heatmap) to configure the heat map.
- To change a color range, click-drag a color handle triangle to a new location or edit its settings in the **Handle Settings** panel.
- Add new handles by clicking the **Add** button.
- To remove a handle, select the handle and click the **Delete** button.
- When the editor has at least two handles a gradient appears as in the example below.







Heatmap Handle Settings panel

1. **Attribute Value** displays the value for the selected handle. Enter a new value to change the setting.
2. **Set Range** sets the *Min* and *Max* values of the heatmap.
3. **Add Handle** adds a new handle to the editor.
4. **Delete Handle** deletes the selected handle.
5. **Color swatch** lets you choose a new color for the selected handle.

\*Note: Population of the Attributes Menu is currently *file format* driven rather than *data* driven. Therefore, if the attribute is not present in your file, the Color By feature will not actually be applied when the OK button is clicked. Make sure the attribute you are trying to color by exists for your file!

## Deleting tracks

There are three ways to remove (delete) a track.

To delete either a graph or annotation track:

- Right-click the track label and select **Delete track**
- Select the **Data Access** tab and click the red "x" next to the track name in the **Data Management Table**
- Select the track, select **Annotation** or **Graph** tab, and click the **Delete** button

## Reference sequence

- [Introduction](#)
- [Loading reference genomic sequence data](#)
  - [What to do if your genome is not supported](#)
- [Loading sequence into a new track](#)
- [Viewing the reference sequence](#)
  - [Examples](#)
  - [Copy and paste sequence bases in IGB](#)

## Introduction

IGB display reference genomic sequence in the Coordinates track. IGB supports many interactions and views on genomic sequence, including copying sequence, viewing sequence in the **Sequence Viewer**, and running BLAST searches. Other features may be available via IGB plug-ins.

## Loading reference genomic sequence data

IGB does not load genomic sequence until you request it, giving you total control over how much data is loaded and when.

To load genomic sequence

- Click **Load Sequence** button

Sequence data will then load from a remote data source or local file, depending on whether you're viewing a genome version supported by an external data source or if you're viewing a genome you loaded using the Load Custom Genome option.



If you are working with a genome that is not supported by the IGBQuickLoad data source, [let us know](#). It is easy to add new genome versions to IGBQuickLoad.org.



Some external sequence data servers will reject requests for residues from a region that is too large. Others will return the data, but only after a very long wait. If you have problems, let us know.

## What to do if your genome is not supported

If you have a sequence file in one of the formats that IGB can read ([File Formats](#)), you can use that sequence as a reference sequence.

To use your own file as the reference genome sequence

- Choose **File > Open Custom Genome**
- Enter a species name into the Species text box
- Enter a genome version name into the Genome Version text box
- Use the "... " button to browse and open your reference sequence file
- Click **OK**

IGB will then open the file and it will appear in the **Data Management** table.

To load data, click **Load Sequence** button (NOT **Load Data**) to view the sequence in the coordinate axis.



If the chromosome names do not follow standard synonym conventions you may need to create a personal synonyms file. See: [Personal Synonyms](#).



If the sequence length contradicts your annotation files, IGB will not be able to display them together. Please [contact us](#) if you need help.

## Loading sequence into a new track

If you want to load an additional sequence as a track rather than as a reference sequence (perhaps a substrain or an individual), open it as if it were an ordinary file.

- You can drag and drop the file into IGB from either a desktop folder or from a URL.

or

- You can also choose **File > Open File...** or **Open URL....**
- Click **Load Data** to load the sequence into the track

As with any other track, a new empty track will appear labeled with the name of the file you opened. The file name will appear in the **Data Management** table in the **Data Access** tab.

## Viewing the reference sequence

At low zoom (many bases are shown) loaded sequence is shown as a gray bar beneath the coordinates axis.

When you are zoomed in, the sequence bases become visible. If sequence has not yet been loaded, '-' (hyphen) characters are shown. If the sequence is loaded, the nucleotide base letters are shown. By default, IGB uses cool colors to represent A and T bases and warmer colors to represent G and C bases.

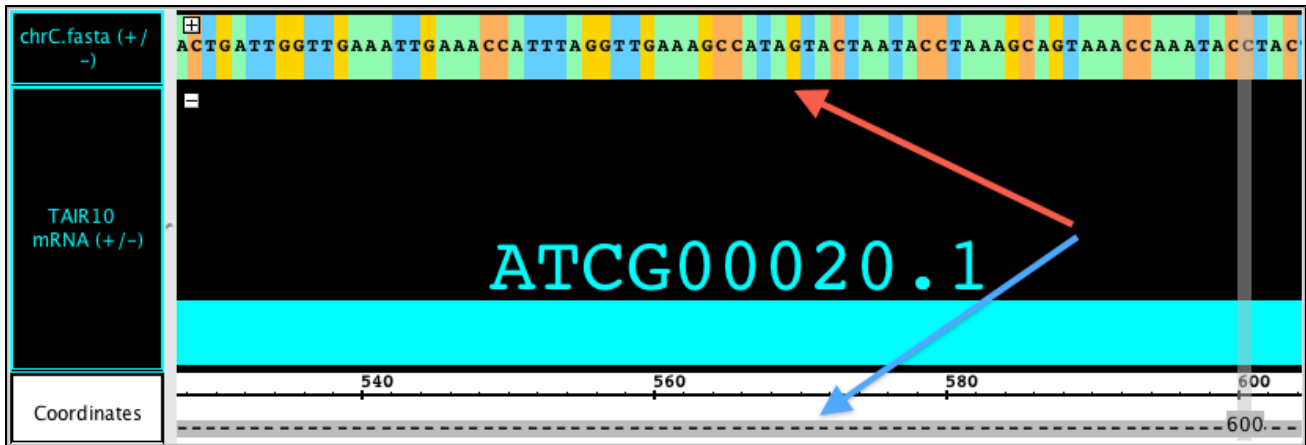
Change base colors in the [Preferences > Other Options](#) tab.



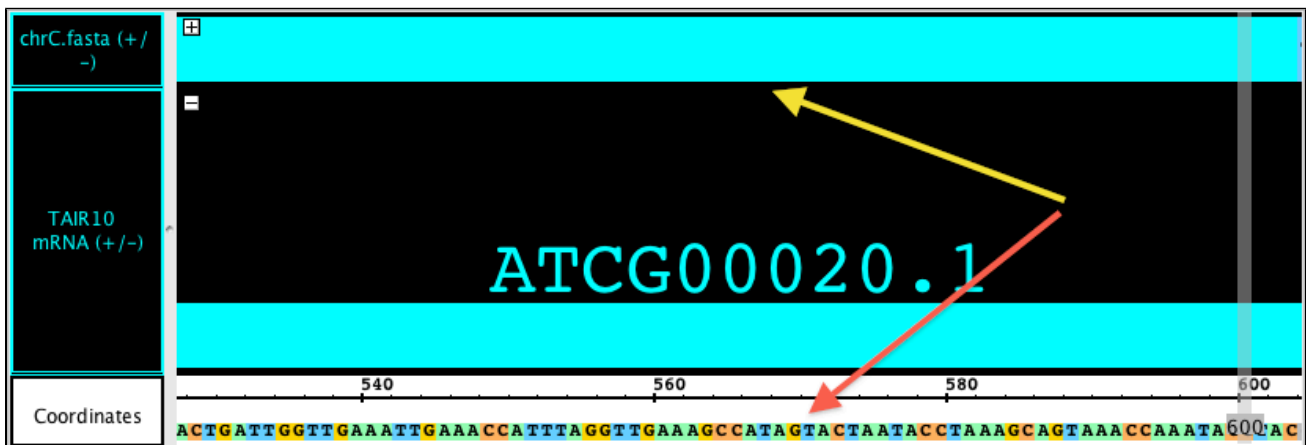
Note that depending on the sequence data server, file, or URL that supplied the sequence data, bases may appear in upper or lower case.

## Examples

In the first image, you can see a sequence loaded as a track ([red arrow](#)). Notice that the refseq has not yet been loaded ([blue arrow](#)) so it appears gray with '-' marks; when you are zoomed out, the sequence track shows a gray bar (default color), and the coordinate axis is empty.



After loading the refseq (red arrow), the sequence in the track (yellow arrow) changes to the FG color for that track (cyan in this case) where ever it matches the refseq. In this state, it will only highlight differences/mismatches with a letter and a color change. Zoomed out, the track appears as a solid bar in the FG color for the track. The refseq appears as a gray bar in the coordinate axis.



## Copy and paste sequence bases in IGB

You can copy a sequence of bases into another application, such as Notepad.

To capture a sequence of bases:

- Make sure that the sequence residues are already loaded.
- Either select a single annotation or click and drag in the **Coordinates** track to select the desired coordinate range.
- Choose **Edit** menu > **Copy Selected Residues** to Clipboard.
- Paste the text string into the target application or location.



If you have selected an annotation that has introns, *the selected residues will only include the exon regions.*

If you have selected a negative-strand annotation, *the selected residues will be reverse-complemented.*

If you want to capture all the residues including the intron regions, select the region by dragging in the **Coordinates** track

## Custom Genomes (Genomes not in IGB)

IGB supports many species and genome versions, not just the species shown on the start screen. To check whether a genome is available in IGB, click the **Current Sequence** tab and use the **Species** and **Genome Version** menus to look for your genome of interest.

However, if your genome is not available, you can still use IGB. Here's how:

## How to open a custom genome

1. Select **File > Open Genome from File...** (or click the DNA icon in the **Toolbar**.)
2. Select a sequence file to use as the reference genome (fasta or 2bit format).
3. **Optional:** Enter an IGB-friendly genome version name (optional)

4. **Optional:** Enter the species name, including genus, species and variety as appropriate (optional)
5. Click **OK** and wait for the genome to load.
6. Open data files as usual. And to view sequence, zoom in and click **Load Sequence**.

1. choose reference sequence (2bit is best; fasta is fine. )

2. enter IGB-friendly genome version name (optional)

3. enter genus, species, variety (optional)

4. click OK

IGB-friendly genome version names look like:

**G\_species\_MMM\_YYYY**  
or  
**G\_species\_variety\*\_MMM\_YYYY**

- **G** - first letter of genus
- **species** - species name
- **variety\*** - zero or more strain, cultivar, or accession names (optional)
- **MMM** - 3-letter abbreviation of month when the genome version was released
- **YYYY** - year of release

For examples, see the **Genome Version** menu in the **Current Genome** tab.

## Sharing data using QuickLoad sites

### About IGB QuickLoad

Use IGB QuickLoad (QL) to share data files on-line using a Web site or cloud storage service like Dropbox. Using IGB QuickLoad, you can set up and configure data sets so that they'll be listed in the IGB **Data Access** panel, making it easy for lab members, collaborators, and members of the public to browse and load your data as new tracks in IGB.

QuickLoad sites can be local or remote. Typically, you'll create a QuickLoad site on your local computer, test it there, and once you're happy with how the data look when loaded into IGB, you'll transfer it to a Web site or Dropbox folder.

See also: [Set up a Quickload in Dropbox](#)

### How to set up an IGB QuickLoad site

First, find out if your genome is already supported by the [main IGB QuickLoad site](#).

If your genome is already there, then you should make a satellite QuickLoad site. A satellite site should not contain sequence data or gene models annotations as these are already available from the main IGB QuickLoad site.

To find out if your genome is already supported, look for it in the the main IGB QuickLoad site's [contents.txt file](#).

IGB genome version names

**contents.txt**

H_sapiens_Dec_2013	Homo sapiens (Dec 2013)	human being (GRCh38/hg38)
H_sapiens_Feb_2009	Homo sapiens (Feb 2009)	human being (GRCh37/hg19)
H_sapiens_Mar_2006	Homo sapiens (Mar 2006)	human being (NCBI36/hg18)
H_sapiens_May_2004	Homo sapiens (May 2004)	human being (NCBI35/hg17)
G_gallus_Nov_2011	Gallus gallus (Nov 2011)	chicken (galGal4/ICGC Gallus-gallus-4.0)
G_max_Aug_2010	Glycine max Glyma1	
G_max_Jan_2014	Glycine max Gmax_275_v2.0	

These appear in the "genome version" menu in the IGB **Current Sequence** tab

descriptions, including commonly used synonyms  
IGB displays this text in the title bar of the main IGB window.

If your genome is listed, go to: [Create a satellite IGB QuickLoad site for genomes already supported in the main IGB QuickLoad.](#)

If your genome is **not** listed, go to: [Sharing data for a new or custom genome not already part of IGB QuickLoad.](#)

## Create a satellite IGB QuickLoad site for genomes that are already supported in the main IGB QuickLoad site

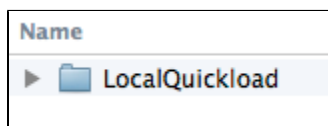
Follow these instructions if your genome is already supported in the main IGB QuickLoad site. To find out if your genome is already supported, see the instructions in [Sharing data using QuickLoad sites.](#)

- [Create a QuickLoad root directory.](#)
- [Create a directory for your genome version.](#)
- [Create contents.txt file listing your genome.](#)
- [Add genome.txt file from public QuickLoad site.](#)
- [Create annots.xml meta-data file for each genome version directory.](#)
- [Add data files.](#)
- [Test your new QuickLoad site in IGB.](#)

### Create a QuickLoad root directory.

Create a QuickLoad root directory (folder) on your local computer or on a Web server. This folder will contain your genome directories and a "meta-data" file called **contents.txt**.

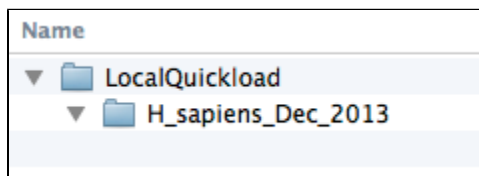
Here is an example:



### Create a directory for your genome version.

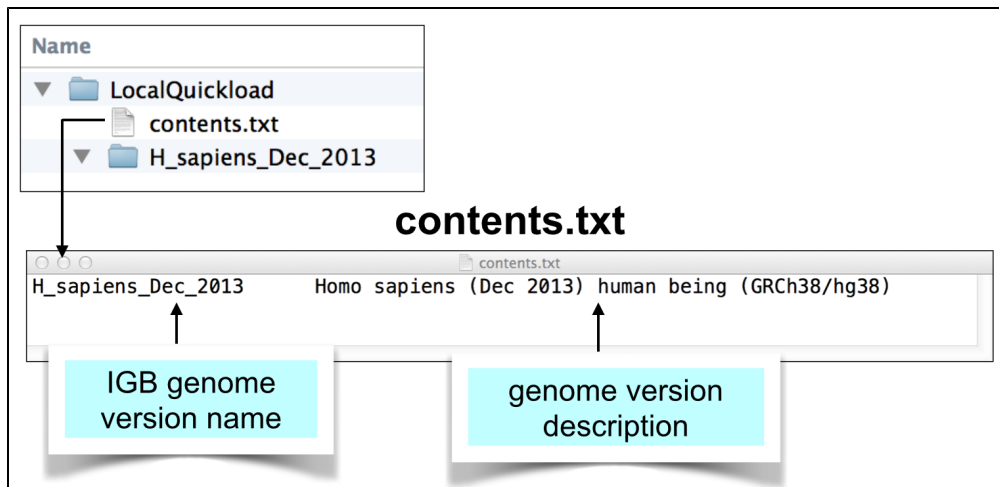
Create a genome directory corresponding to the genome assembly for which you want to share data via your QuickLoad. Like the main IGB QuickLoad site, you can support multiple genome versions from different species. Use the same names as in the [main IGB QuickLoad site](#).

Here is an example:



### Create contents.txt file listing your genome.

Create a plain text file called **contents.txt** and save it your QuickLoad root directory. Copy the text for your genome from the main IGB QuickLoad [contents.txt](#) file into your file and save it.



## Add genome.txt file from public QuickLoad site.

Visit the [public QuickLoad](#) site in your Web browser and navigate to the directory for your genome version. Download the genome.txt from that directory and save it to your local genome version directory.

Note that IGB uses this file to create the sequence selection table under the **Data Access Panel**. The Data Access panel reports both sequence names and sizes, and so that's why the genome.txt file needs to contain both.

**IGB Current Genome tab**

Species: Homo sapiens

Genome Version: H\_sapiens\_Dec\_2013

(455) Sequenc...	Length
chr1	248,956,422
chr2	242,193,529
chr3	198,295,559
chr4	190,214,555
chr5	181,538,259
chr6	170,805,979
chr7	159,345,973
chrX	156,040,895
chr8	145,138,636
chr9	138,394,717
chr11	135,086,622
chr10	133,797,422
chr12	133,275,309
chr13	114,364,328
chr14	107,043,718

genome.txt

chr1	248956422
chr2	242193529
chr3	198295559
chr4	190214555
chr5	181538259
chr6	170805979
chr7	159345973
chrX	156040895
chr8	145138636
chr9	138394717
chr11	135086622
chr10	133797422
chr12	133275309

names      sizes

## Create annots.xml meta-data file for each genome version directory.

Each genome version subdirectory needs an **annots.xml** file that is either empty or contains a listing of annotations you want to appear the **Available Data Sets** section of IGB's **Data Access** tabbed panel.

The annots.xml file is a simple "markup" file that specifies all the files that a user can load into IGB. Each file corresponds to a one track.

Here is a simple example:

```
<files>
  <file name="bamfiles/Treatment1.bam"
    title="RNA-Seq/Treatment Sample 1"
    description="RNA-Seq alignments from Treatment Sample 1"
    background="FFFFFF"
    foreground="0000FF"
    max_depth="15"
    name_size="12" />
</files>
```

The **name** attribute of the **file** tag indicates the physical location of the file **relative to the directory where the `annots.xml` file resides**. However, you can also use a URL to indicate the location of the file. This means that if the file is located on another server, you can leave it there; you don't have to copy it into your QuickLoad site.

For example, in this case, you could enter "<http://www.example.com/bamfiles/Treatment1.bam>" instead of "`bamfiles/Treatment1.bam`" for file.

Note also that if you are hosting BAM files, the BAM file index ".bai" file needs to reside in the same directory as the BAM file. Otherwise, IGB will display an error message when users try to access the data.

For details on supported file attributes in an `annots.xml` file, see: [About `annots.xml`](#)

## Add data files.

Place your annotation, BAM, or graph (bedgraph or bigwig) files to the locations specified in your **`annots.xml`** file. You can use any format IGB supports.

IGB QuickLoad can support BED, bedgraph, and GFF files that have been sorted, compressed, and indexed using the `bgzip` and `tabix` utilities. Doing this helps speed up data loading in IGB and allows loading data by region. For an example of how we used `tabix` to distribute coverage and junction files from an RNA-Seq data set, see [Creating a new genome release for IGB QuickLoad](#) from the IGB Developer's Guide.

## Test your new QuickLoad site in IGB.

Tell IGB to use your new QuickLoad site.

Follow the directions in [Adding and Managing Data Source Servers](#). To add the local server, click the "..." button and select the folder which contains the **`contents.txt`** file, your QuickLoad root directory.

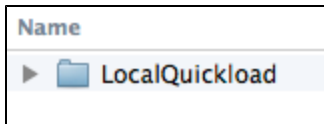
## Sharing data for a custom genome not already part of IGB QuickLoad

Follow these instructions if your genome is **not** already in the main IGB QuickLoad site. To find out if your genome is already supported, see the instructions in [Sharing data using QuickLoad sites](#).

- [Create a QuickLoad root directory](#).
- [Create a genome directory](#).
- [Create `contents.txt` file](#)
- [Save sequence data in twoBit format in the genome directory](#)
- [Make a `genome.txt` file using `twoBitInfo`](#)
- [Create `annots.xml` file listing available annotations and data sets](#).
- [Add annotation files](#).
- [Create `species.txt` file](#)
- [Create `synonyms.txt`](#)
- [Add your new QuickLoad site to IGB](#).

## Create a QuickLoad root directory.

Create a QuickLoad root directory (folder) on your local computer or on a Web server.



## Create a genome directory.

Create a genome directory for the new, custom genome version you want to share via QuickLoad. To do this, create an IGB friendly name for your genome version and use this as the name of the genome directory.

Doing this will ensure that your genome version will be displayed correctly in the **Current Genome** tabbed panel menus. The IGB-friendly genome version name look like:

```
G_species_variety MMM_YYYY
```

where

- "G" is the first letter of the genus
- "species" is the species name
- "variety" is the sub-species, cultivar or strain (this is optional)
- MMM is the three-letter English abbreviation in English for the month the genome version was released
- YYYY is the year the genome version was released

For example, **H\_sapiens\_Dec\_2013** is the IGB name for the human genome assembly released in Dec 2013.

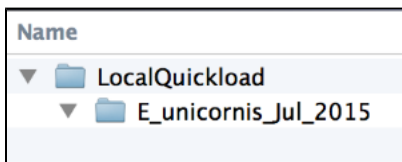
You can also indicate subspecies, varieties, or individuals by including additional suffixes.

For example, there are two *Oryza sativa* (rice) subspecies in wide cultivation: japonica and indica. The first published rice genome sequence was from the japonica subspecies. Because an indica genome sequence is now available, we need names that distinguish them. We distinguish the japonica assemblies using prefix O\_sativa\_japonica and the month and year of release: **O\_sativa\_japonica\_Oct\_2011**. Annotations and sequence data reside in genome version folder O\_sativa\_japonica\_Oct\_2011 on the main IGB QuickLoad site - view its contents by visiting [http://igbquickload.org/quickload/O\\_sativa\\_japonica\\_Oct\\_2011](http://igbquickload.org/quickload/O_sativa_japonica_Oct_2011)

There is no limit on the number of "\_variety" suffixes you can include, provided the genome version name always ends with MMM\_YYYY. By including the month and year in the genome version names, you ensure that IGB will display the latest genome version releases first in the **Genome Version** menu of the **Current Sequence** tab.

Before you pick a name, also look at the [species.txt](#) file that is packaged with every version of IGB. If your species is listed there, then use the genus and species prefix listed in column three of the file. I

After picking an IGB-friendly name, create a sub-folder named for the genome version within the QuickLoad root directory:



## Create contents.txt file

Create a simple, plain text file called **contents.txt** and add it to the top-level directory.

This file should be plain text, tab-delimited, with two columns separated by a tab character.

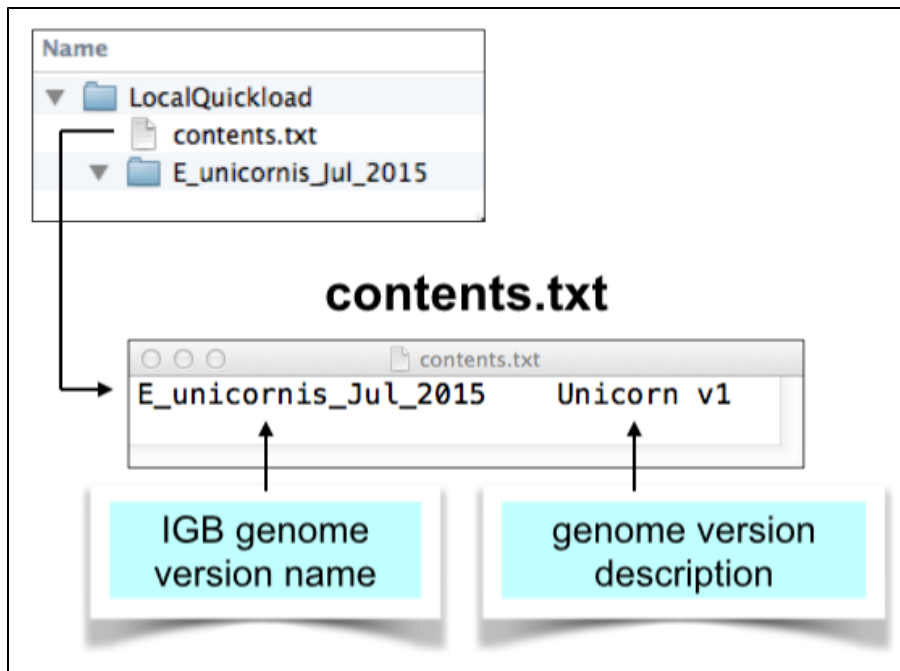
The first column should list names of the genome directories you have created and want to share.

The second column is optional - it contains text IGB will display on the title bar when you open IGB. View an example contents.txt on IGB QuickLoad: <http://igbquickload.org/quickload/contents.txt>

To create a **contents.txt** file for your QuickLoad site:

1. Open a text editor (e.g., Notepad or TextEdit).
2. Type the genome version name.
3. Type TAB.
4. Type a brief description of the genome; this will appear in the IGB title bar when users open your genome version.
5. Repeat 2 through 4 for each genome version you want to support in your QuickLoad site.
6. Save the file as a plain text (".txt") file in your root QuickLoad directory - call it "contents.txt"





## Save sequence data in twoBit format in the genome directory

Obtain sequence files and convert them to 2bit format, a binary format IGB can use to obtain all or part of the sequence data when you press the **Get Sequence** button, run a blast search, or open the **Sequence Viewer** window.

The 2bit format was first developed for the blat cDNA-to-genome alignment tool. IGB 2bit uses to support efficient retrieval of sequence data. To convert a fasta format file to twobit, use the `faToTwoBit` command-line tool from <http://hgdownload.cse.ucsc.edu/admin/exe/>.

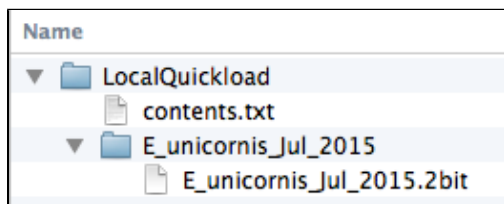
To convert a fasta file to 2bit, open a terminal window and enter:

```
Make a 2bit file

faToTwoBit FASTA.fa G_species_MMM_YYYY.2bit
```

The first argument (FASTA.fa) should be the fasta file with your sequence. The second argument (G\_species\_MMM\_YYYY.2bit) should have the same name as your genome version and use the .2bit file extension.

Please this file inside your genome version directory:



## Make a genome.txt file using twoBitInfo

For each genome directory, create a **genome.txt** file that lists all the chromosomes in your genome, together with their sizes.

As with the **contents.txt** file, it should be tab delimited.

The first column lists the chromosome names and the second column lists their sizes. IGB uses this file to create the sequence selection table under the **Data Access Panel**.

Here is an example: [http://igbquickload.org/quickload/A\\_thaliana\\_Jun\\_2009/genome.txt](http://igbquickload.org/quickload/A_thaliana_Jun_2009/genome.txt)

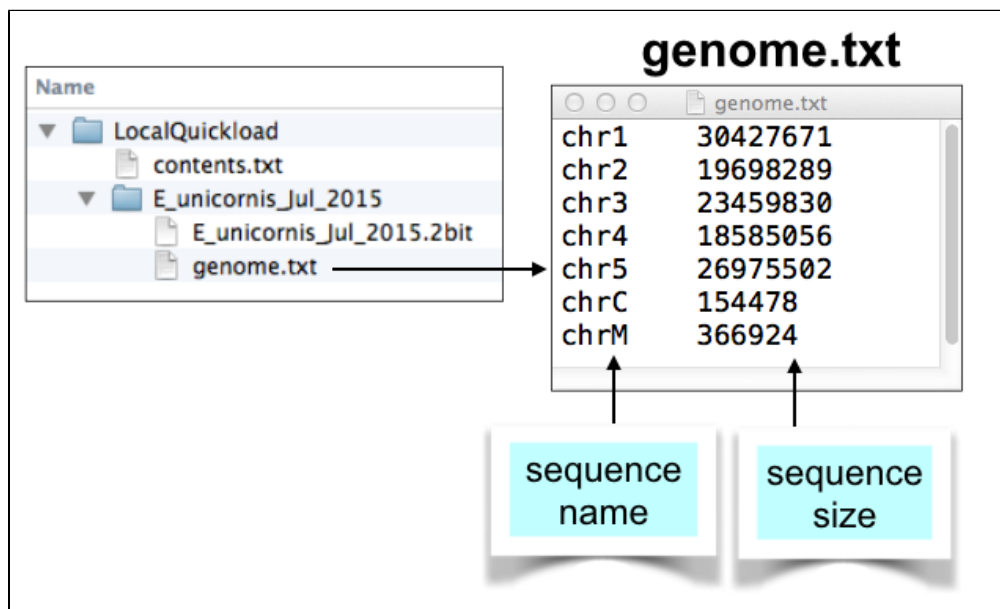
Create the **genome.txt** file from a sequence 2bit file (see below) using the `twoBitInfo` command-line tool, which is available from <http://hgdownload.cse.ucsc.edu/admin/exe/>.

To run twoBitInfo, open a terminal, change into the genome directory, and enter:

### Making a genome.txt file

```
twoBitInfo G_species_MMM_YYYY.2bit genome.txt
```

The first argument is the name of the 2bit file you made in the previous step. The second argument is always the same: "genome.txt," the name of the genome information file twoBitInfo program creates.



## Create annots.xml file listing available annotations and data sets.

Each genome version subdirectory needs an **annots.xml** file that is either empty or contains a listing of annotations you want to appear the **Available Data Sets** section of IGB's **Data Access** tabbed panel.

Here is a simple example:

```
<files>
  <file name="E_unicornis_Jul_2015.bed.gz"
    title="Gene models/Maker annotations"
    description="Results from Maker annotation pipeline"
    load_hint="Whole Sequence"
    label_field="title"
    background="FFFFFF"
    foreground="000000"
    max_depth="10"
    name_size="12" />
</files>
```

When loaded into IGB, each file will create one track.

Note that the **name** attribute of the **file** tag indicates the physical location of the file relative to the directory where the annots.xml file resides. However, you can also use a URL to indicate the location of the file. This means that you if you want to reference other resources hosted on other machines, you can do it.

Also note that if you are hosting BAM files, the BAM file index ".bai" file needs to reside in the same directory as the BAM file. Otherwise, IGB will display an error message when users try to access the data.

For more information about annots.xml format, see: [About annots.xml](#)

## Add annotation files.

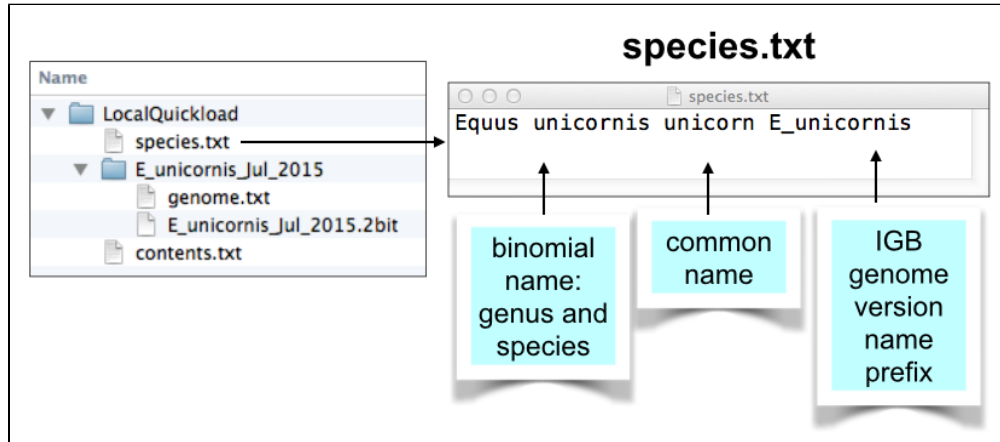
Place your annotation, BAM, or graph (bedgraph or bigwig) files into the appropriate genome sub-directories. You can use any format IGB supports.

IGB QuickLoad can support BED, bedgraph, and GFF files that have been sorted, compressed, and indexed using the bgzip and tabix utilities. Doing this helps speed up data loading in IGB and allows loading data by region.

## Create species.txt file

This file tells IGB how to populate the species and genome version selection menus in the **Current Sequence** tabbed panel. IGB contains a "built-in" list of species it supports, many of which are not yet sequenced or supported in the main IGB QuickLoad site.

To create your own species.txt file, [follow these instructions](#). Be sure to include the G\_species prefix in column three of the file.



## Create synonyms.txt

This is a list of synonyms for genomes. This list allows you to match names across different QuickLoad sites and also the Galaxy workflow site. For example, if a QuickLoad site at another location uses different names to refer to the same genome version, you can specify these in the synonyms.txt. Each line contains any number of synonyms for a genome, separated by tabs.

To create your own synonyms.txt file, [follow these instructions](#). Be sure to include the IGB genome version name in column 1 of the file.

## Add your new QuickLoad site to IGB.

Tell IGB to use your new QuickLoad site.

Follow the directions in [Adding and Managing Data Source Servers](#). To add the local server, click the "..." button and select the folder which contains the **contents.txt** file, e.g. "Quickload servers".

## Set up a Quickload in Dropbox



You can use file hosting service [Dropbox](#) to share IGB Quickload sites with colleagues. To use Dropbox as a QuickLoad site:

1) Create QuickLoad folders and files as described in [Sharing data using QuickLoad sites](#). Before moving your new site to your Dropbox, ensure that your Quickload works correctly by adding it as a data source on your [local machine](#).

2) Copy your QuickLoad folder into your **Public** Dropbox folder.



3) Open the QuickLoad folder and select the **contents.txt** file. (You should then see an option to **Copy public link...**)

contents.txt	Copy public link...	Download	Delete...	Rename	Move...	More ▾
 contents.txt	document	12/12/2014 9:29 AM				
 C_mellificae_Apr_2014	folder	--				

4) Select **Copy public link...** and then select **Copy to Clipboard**.

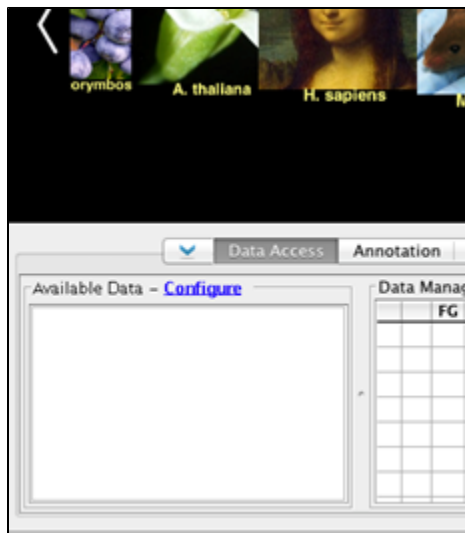
Copy public link...

To share this file with others, copy and paste this link:
Shorten link

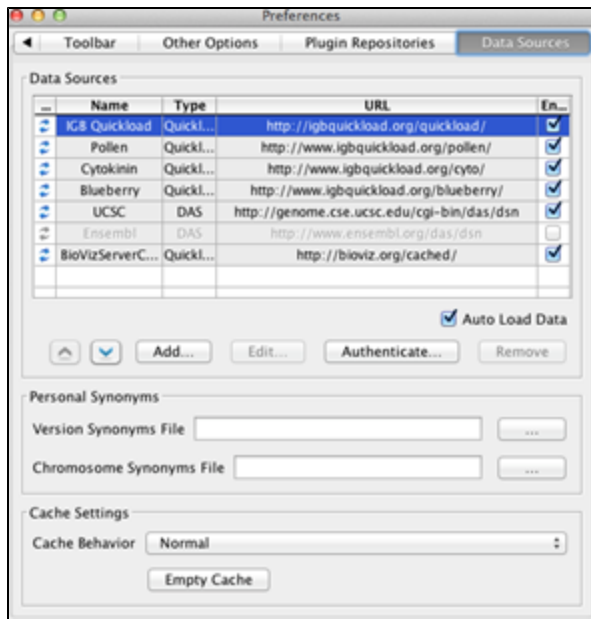
<https://dl.dropboxusercontent.com/u/33652246/Crithidia/contents.txt>

Copy to clipboard
Cancel

5) Now open IGB and click **Configure** under the **Data Access** tab in the bottom left corner of IGB,

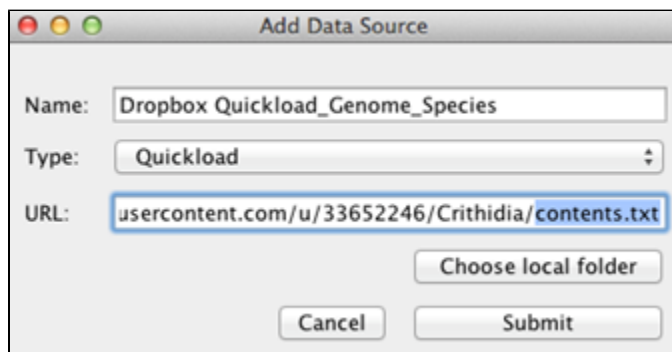


This will open the **Data Sources** tab within IGB Preferences.



6) Click the **Add...** button and enter:

- Under name, enter whatever you want this server to be called within your IGB
- Under type, leave as **Quickload**.
- Under URL, paste the Public Link you copied from Dropbox. Then, delete "contents.txt" from the link.



**IMPORTANT:** remove the "contents.txt" from the end of the public link.

7) Enter **Submit**.

You should see a message stating your data are now available.



8) If you added a new species and genome, then these should now be available in the **Species Menu**. If not, then select your species from the IGB Start screen or from the **Species** and **Genome Version** menus. Your new QuickLoad data source should be listed in the **Available Data** section of the **Data Access** tab.

## About annots.xml

### Use annots.xml to specify track color, annotation style, and more

An **annots.xml** file resides in an IGB QuickLoad genome directory. It lists all data files available for that genome.

The file contains one or more **file** tags, enclosed in a **files** element. The file tag has attributes that control how the data looks when loaded as new track into IGB. Many options correspond to options you can configure using the **Annotation** or **Graph** tabbed panels in IGB.

attribute	optional	Description
name	Required	<p>The name of the file <b>or</b> a URL. The URL can point to any file, anywhere, provided it's accessible via the internet.</p> <p>You can specify the location of the file using an absolute URL or a relative URL. Relative means: relative to the genome version directory where the annots.xml file resides.</p> <p>ex)</p> <p>name="bamfiles/treatment.bam" (relative)</p> <p>name="http://www.example.com/bamfiles/treatment.bam" (absolute)</p>
title	Optional	<p>User-friendly text IGB will show in the Data Access Panel as the title of the data set. If you don't provide a title, IGB will display the name attribute instead.</p> <p>If you would like IGB to display data set titles in nested folders in the Available Data Sets section of the Data Access panel, insert the folder separator character "/" used under the Unix operating system (a forward slash) into the title.</p> <p>For example, if a file title is RNA-Seq/Treatment/Sample1, then IGB displays the data set named Sample1 inside a named Treatment inside another folder named RNA-Seq.</p>
description	Optional	<p>IGB displays this text in the tooltip that appears when the mouse hovers over the data set title in the <b>Data Access</b> tab.</p>
url	Optional	<p>Use this tag to specify the location of an external Web page describing the data set. If provided, IGB will display an "info" icon next to the data set title. When users click the icon, their Web browser will open showing the contents of the URL you provide.</p> <p>You can specify the location of the file using an <b>absolute</b> URL or a <b>relative</b> URL.</p> <p>An absolute URL looks like this: "http://www.example.com"</p> <p>Relative means: relative to the Quickload root directory.</p> <p>For example, to show "About.html" residing inside a genome version directory named E_unicornis_Jun_2009, use relative url "E_unicornis_Jun_2009/About.html"</p>
load_hint	Optional	<p>If used, its value should be "Whole Sequence". Using this tag will force IGB to load the entire file when users select the genome version, which is usually appropriate only for reference gene model annotations or other equally small data sets.</p>
label_field	Optional	<p>Use this field to indicate the annotation property (e.g., "score" or "id") used to label individual annotations within a track. The default is <b>none</b>.</p> <p>To find out what annotation properties are available, view the <b>Selection Info</b> for an annotation loaded from the file. To view annotation properties:</p> <ul style="list-style-type: none"><li>• Open the file and load some data.</li><li>• Click an annotation to select it.</li><li>• View <b>Selection Info</b> by clicking the "i" button (top left) or opening the <b>Selection Info</b> tab.</li></ul> <p>Annotation properties are listed in the first column.</p>

background	Optional	Use this field to set a track's background color.  Set colors using <a href="#">six-digit, hexadecimal triplets</a> , but do not include a leading "#" character.  For example, to set a white background, use background="FFFFFF".
foreground	Optional	Use this field to define annotation or graph colors (e.g. foreground="00FFFF"). See above.
max_depth	Optional	Use this field to define the default maximum stack height for an annotations track. This is the number of annotations shown individually in a stack (e.g., max_depth="10" or max_depth="0" for unlimited)
name_size	Optional	Use this field to define the default track label name font size (e.g. name_size="12")
connected	Optional	Use this field to define the default boolean value for the connected field (e.g. connected="true" or connected="false"). The default is "true". This tells IGB to draw lines linking the blocks belonging to compound features assembled from smaller ones.
show2tracks	Optional	Use this field to define the default boolean value for the show2Tracks field (e.g. show2tracks="true" or show2tracks="false"). If you would like minus and plus strand features to be displayed in the same track, set this to "false." The default is "true."
direction_type	Optional	Use this field to dictate whether annotations or alignments will be shown using arrows and/or color to indicate direction. (e.g. direction_type="arrow", direction_type="color", direction_type="both", direction_type="none"). The default is "none."
positive_strand_color	Optional	Use this field to define the default positive strand color (e.g. positive_strand_color="CCFFFF") for when the direction_type setting is "color."
negative_strand_color	Optional	Use this field to define the default negative strand color (e.g. negative_strand_color="33FFFF") for when the direction_type setting is "color."

## Password-protect and style your QuickLoad site using .htaccess and Apache web server

If you set up your QuickLoad site in the Web directories of an Apache server, you can modify and fine-tune how files are displayed when users visit your site in their Web browser by adding directives to the **.htaccess** file in that directory.

You can configure Apache to allow users to view the contents of your QuickLoad directory and include a short description of each file or data set.

Here is an example:

```
Options Indexes
IndexOptions FancyIndexing IgnoreCase FoldersFirst DescriptionWidth=*
AddDescription "Arabidopsis thaliana TAIR9 genome release" A_thaliana_Jun_2009
AddDescription "Annotation file list" annots.[txt|xml]
AddDescription "Chromosome lengths and assembly information" genome.txt
AddDescription "Available genome assemblies" contents.txt
```

You can also use the .htaccess file to password-protect entire folders with data or an entire QuickLoad site. This can be very useful when you're working on a data set and are not yet ready to make it available to the public. When IGB requests a resource that is password-protected in this way, it opens a dialog for users to enter a user name and password.

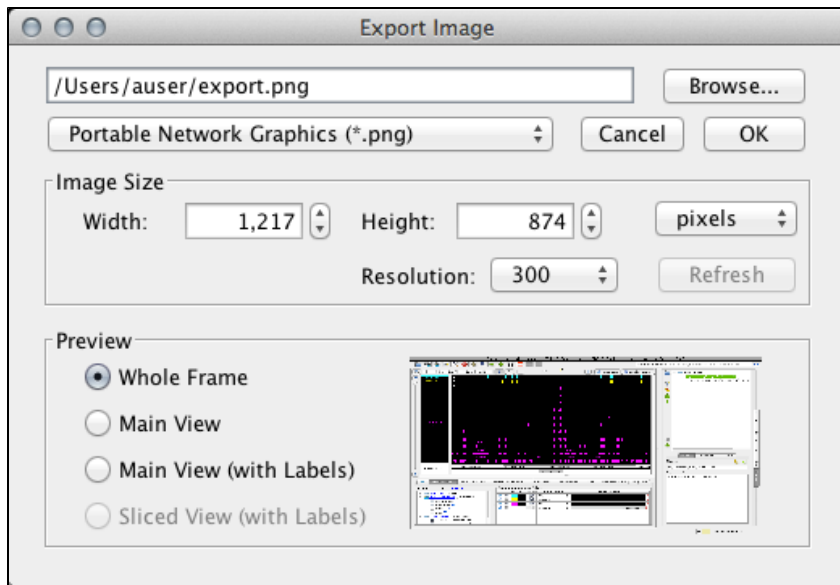
For more information, see: <http://httpd.apache.org/docs/2.0/howto/auth.html>.

## Saving images and printing

- [Saving images](#)
  - [Image export options](#)
- [Printing](#)

### Saving images

To save an image, select **File > Save Image....** A window similar to the image below will open.



## Image export options

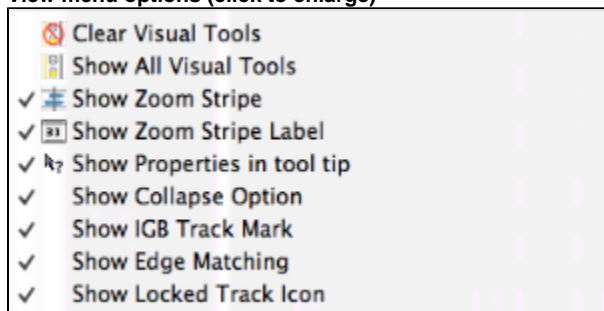
- **Image format** IGB supports PNG, JPEG, and SVG.
- **Image size** Select image size. Default values are the current dimensions of the IGB window. Aspect ratio is locked. To change the aspect ratio, resize the IGB window and click the **Refresh** button.
- **Export options** (Preview)
  - **Whole Frame** includes the entire IGB window
  - **Main View (with Labels)** includes tracks and track labels only.
  - **Sliced View (with Labels)** is available only when the Sliced View tab is open. It will export the image from the **Sliced View** tab.



Due to limitations of the SVG file format, SVG files exported using the **Whole Frame** option may contain some undesired artifacts.

Use the **View** menu to show or hide parts track decorations and other visual tools, such as the zoom stripe, edge matching marks, or the zoom stripe label.

### View menu options (click to enlarge)



## Printing

- **File > Print Main View with Labels...** prints tracks and track labels only
- **File > Print Whole Frame...** prints the entire IGB window.

## Editing IGB-generated SVG files

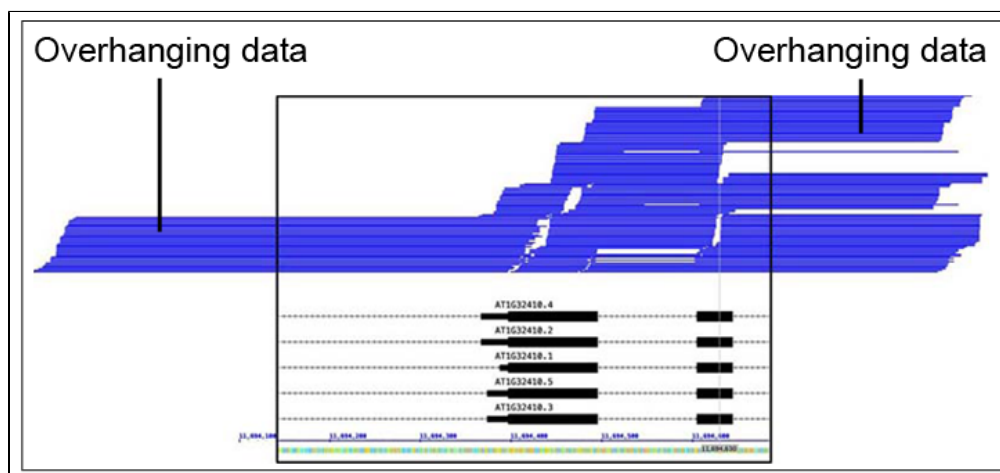
- [Introduction](#)
- [How to crop in Adobe Illustrator.](#)
- [How to crop in Inkscape \(free\)](#)

## Introduction



SVG files generated by IGB sometimes include vector graphics that extend outside image area. This is due to problems with the SVG export library IGB uses, and we aim to fix this in future versions of IGB.

Until then, you'll need to use a vector graphics editor to remove the overhanging data (see below) if necessary.



## How to crop in Adobe Illustrator.

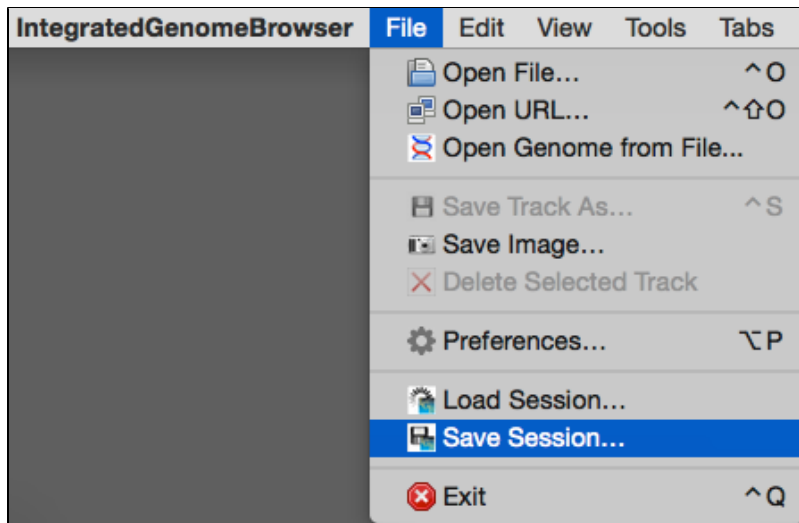
1. Open the SVG file.
2. Adobe may display the message: "Clipping will be lost on roundtrip to Tiny." If yes, click **OK**.
3. Select the Rectangle Tool.
4. Draw a rectangle over the part of the image you wish to use.
5. Change the Opacity to 0%.
6. Select the Selection Tool.
7. Highlight both the rectangle and the image.
8. Select **Object > Clipping Mask > Make**.

## How to crop in Inkscape (free)

1. Open the SVG file.
2. Select the Create rectangles and squares tool.
3. Place a rectangle over the part of the image you wish to use.
4. Under the Fill and Stroke tab, change Opacity to 0%.
5. Select the Select and transform objects tool.
6. Highlight both the rectangle and the image.
7. Select **Object > Clip > Set**

## Saving and Loading Sessions (Draft)

In IGB you can save sessions that preserve your loaded data and track styling. The IGB session, saved as an **XML** file, can be loaded later so you can resume your work.



The **Save Session...** feature allows you to save the following:

- Genomic Location
- Loaded Data (from a remote data provider) *Need more accurate wording here.*
- Track Names
- Track Foreground and Background Color
- Track Label Color
- Track Label Font Size
- Annotation Label Field

## Saving a Session

To save a session, select **File > Save Session....** You will then be able to save your IGB session (as an **XML** file) to a desired location on your computer.

## Loading a Session

To load a session, select **File > Load Session....** Select your saved IGB session and choose **Open** to load your session into IGB.

## IGB Gallery

### Introduction

See links below for images from the IGB software showing a variety of data types and data sets.

Most of these images are from data sets created in the [Lorraine Lab](#).

Some examples (e.g., the tiling array examples) come from data harvested from the [Gene Expression Omnibus](#).

### Color-coding read alignments by strand

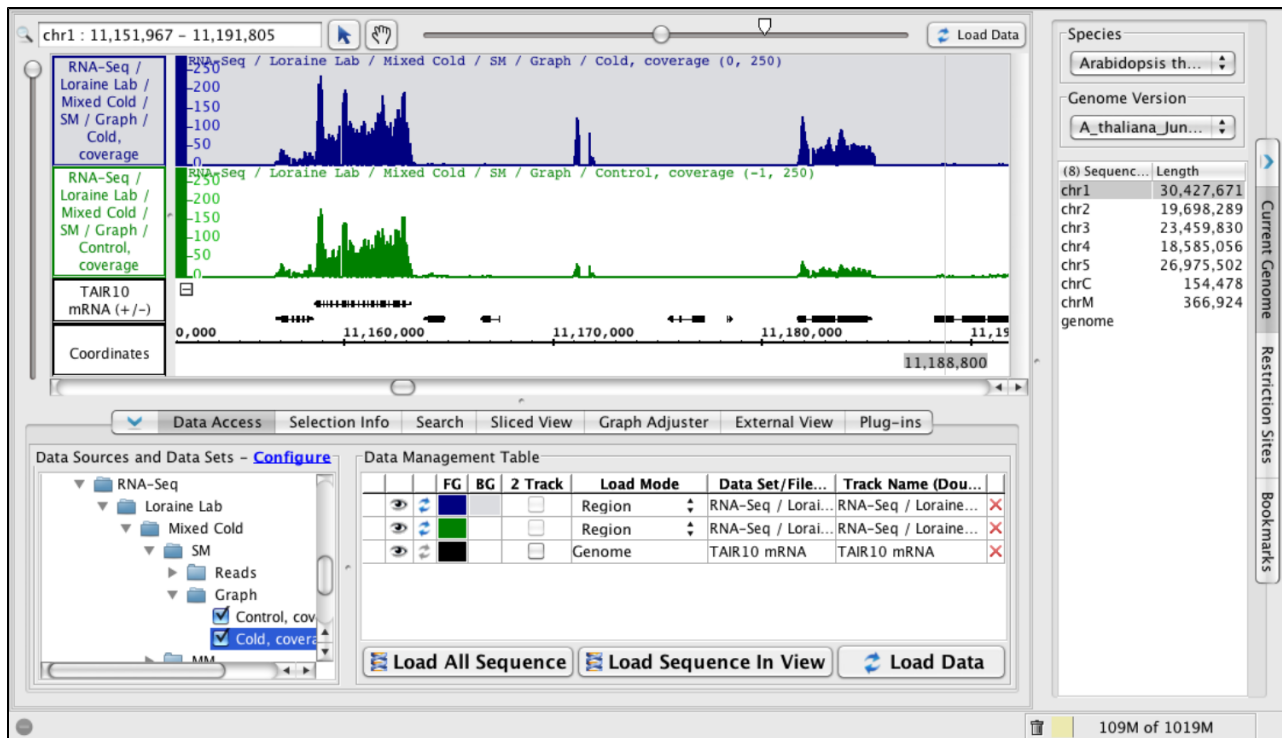
**IGB Preferences allows configuring display of read alignments.**

Sometimes it is useful to use color to indicate strand in RNA-Seq reads. The image below shows IGB configured to display plus and minus strand reads using different colors. Users can choose the colors on a track by track basis. These data were loaded from a custom QuickLoad site.



## A new gene induced by cold in Arabidopsis thaliana?

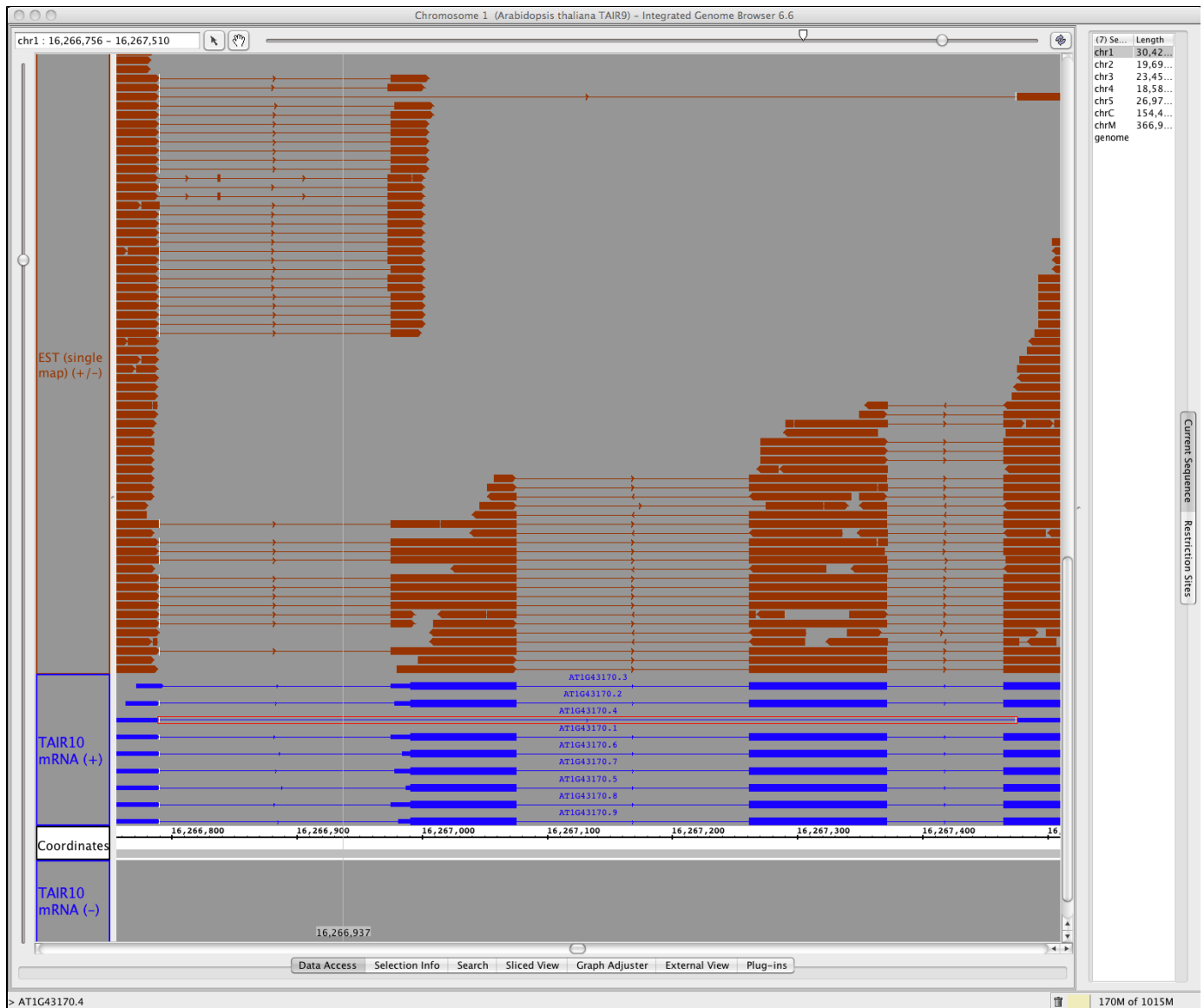
This image from IGB 6.7 shows two coverage graphs from an RNA-Seq experiment. The dark blue graph shows RNA from seedlings undergoing cold stress and collected at multiple time points over several days. The green graph shows RNA from control seedlings collected at the same time points. The gene models data set contains protein-coding genes only. The middle region with a taller peak in the blue (cold-treated) sample corresponds to a gene annotated as "other\_na" in the TAIR annotations, available from IGB QuickLoad.



## Edge matching ESTs to assess alternative splicing

### Edge matching ESTs to assess alternative splicing

Edge matching is a specialized visualization technique for genomics data that makes it easy to compare boundaries. Edge matching is especially useful for genes that undergo alternative splicing. We can assess the degree to which individual splicing choices are made by looking at overlapping data from EST alignments or RNA-Seq data. The example below shows a visualization in IBG 6.6, which introduces a new iconography for gene models that uses arrows to indicate strand. Arrows pointing right mean the annotation is on the plus (forward) strand and arrows pointing left indicate minus strand annotations. Note how in this case that one of the gene models suggests an exon skipping event but only one overlapping EST supports it.

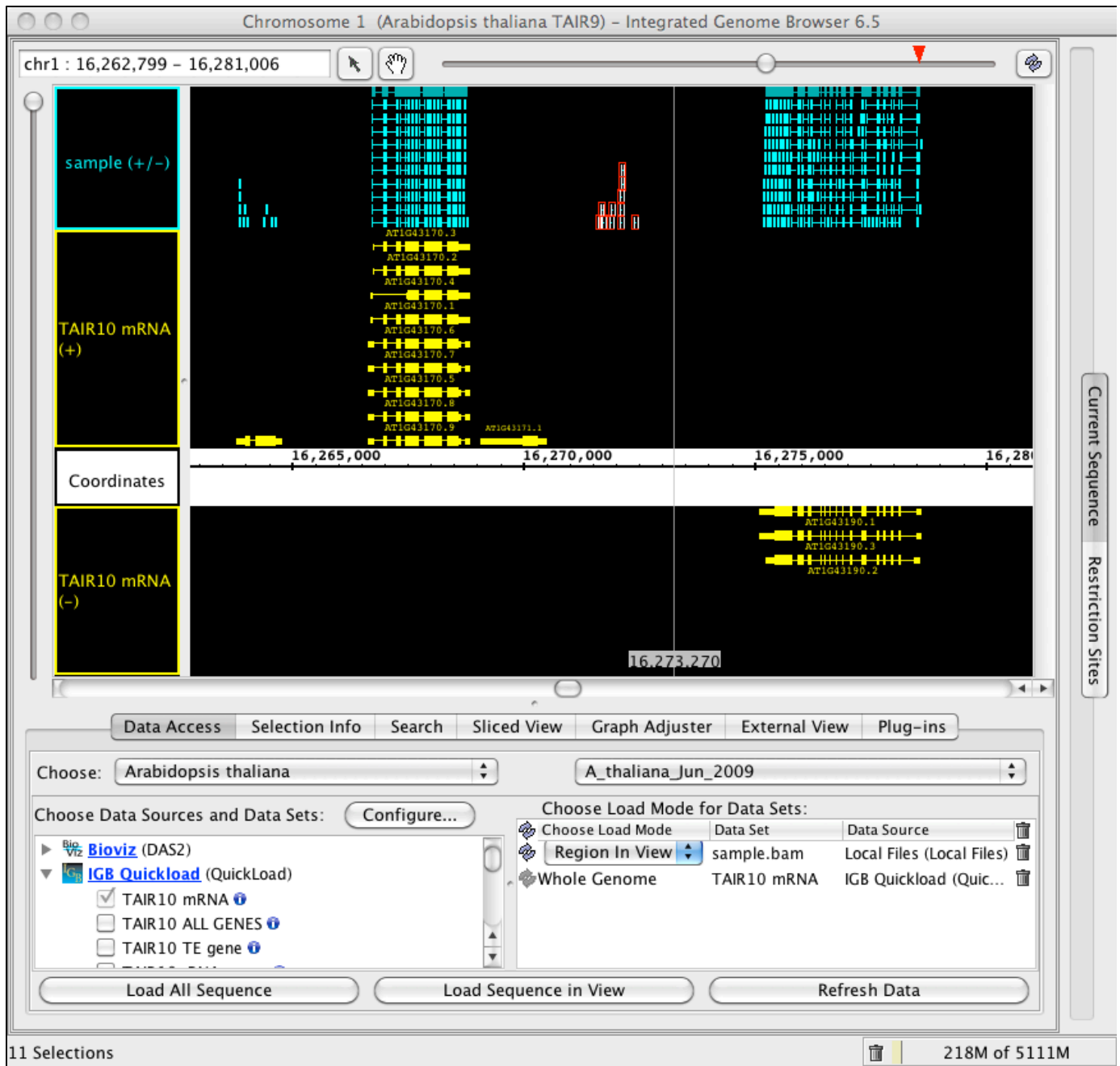


## Counting reads with selection tool

### Counting reads in a region with no annotated protein-coding gene

This image illustrates using the selection tool (the arrow icon) to select a set of reads over a region of the *Arabidopsis* genome with no annotated protein-coding gene. This region might contain a gene of another type, e.g., an "other RNA" gene. To check this, load the TAIR10 ALL data set.

Note that the number of selected items appears in the lower left corner of the IGB window. The version of IGB is 6.5., released in April, 2011.

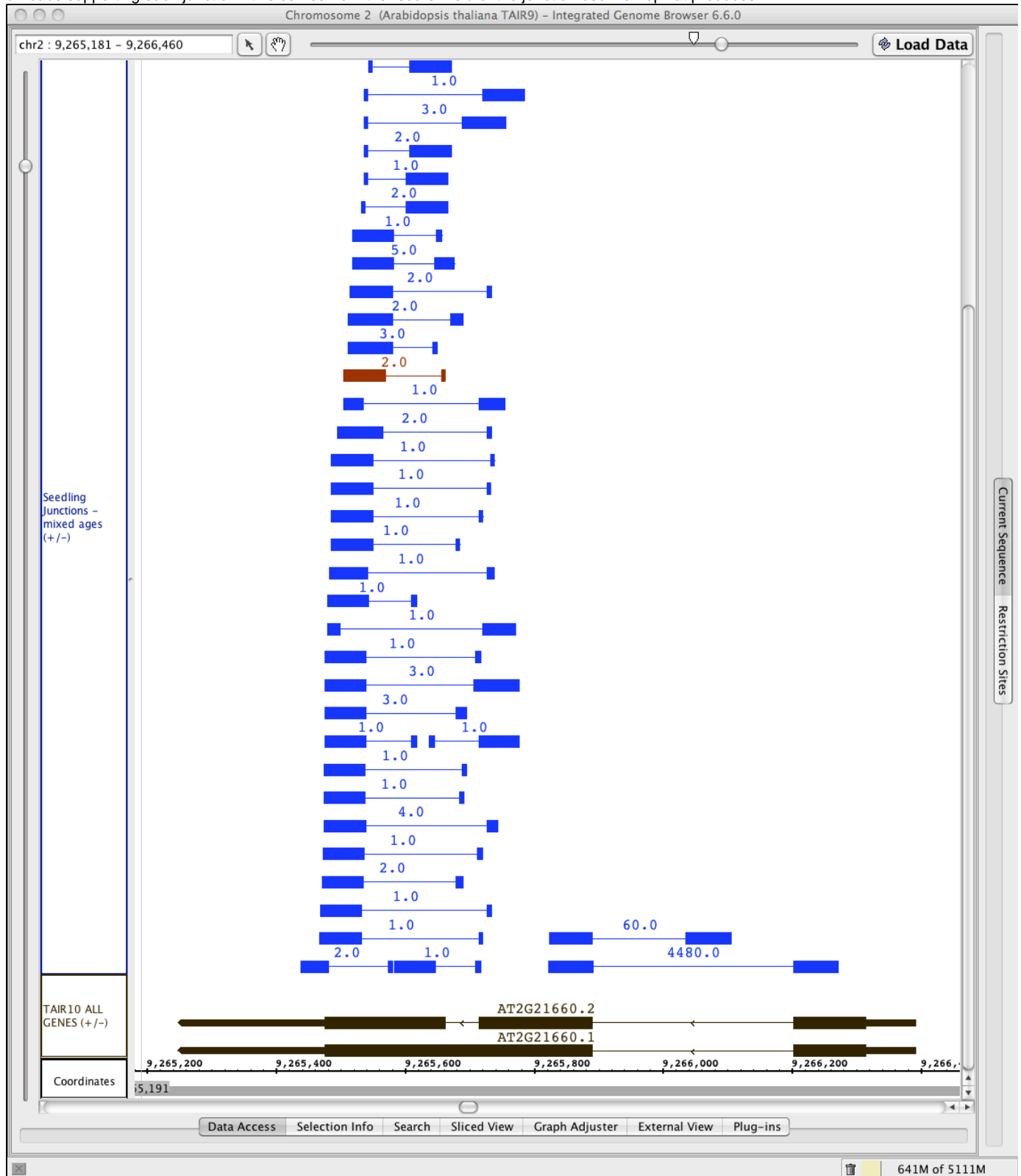


## Alternative splicing in GRP7 (AT2G21660)

### TopHat analysis of RNA-Seq data reveals alternative splicing in Arabidopsis thaliana GRP7 (AT2G21660)

This image from IGB 6.6 shows how IGB can capture a rich data scene and convey details about alternative splicing. The image depicts Arabidopsis thaliana GRP7 (AT2G21660) encoding "a small glycine-rich RNA binding protein" that "regulates..circadian oscillations of its own transcript." (Functional annotation comes from The Arabidopsis Information Resource [locus page](#) for AT2G21660.) Splicing pattern of GRP7 is

complex, as depicted in the image below. Here, junctions predicted by the TopHat spliced alignment program are shown together with the number of reads supporting each junction - this comes from the "score" field of the junction.bed file TopHat produces.



## Using IGB to view effects of kmer on assembly quality

This image from IGB 6.6 shows a close-up view of two nearby *A. thaliana* Jun 2009 genes on opposite strands. Tracks shown in yellow display alignments of transcriptome assemblies of RNA-Seq data from Arabidopsis seedlings undergoing a severe drought stress and unstressed controls. Transcriptome assemblies were performed using Oases run with the same parameters and data, but with different kmer values. Note how different kmer values yield very different results. Note also how the assembly performed using kmer 21 includes a transcript model that appears to merge adjacent genes from opposite strands.

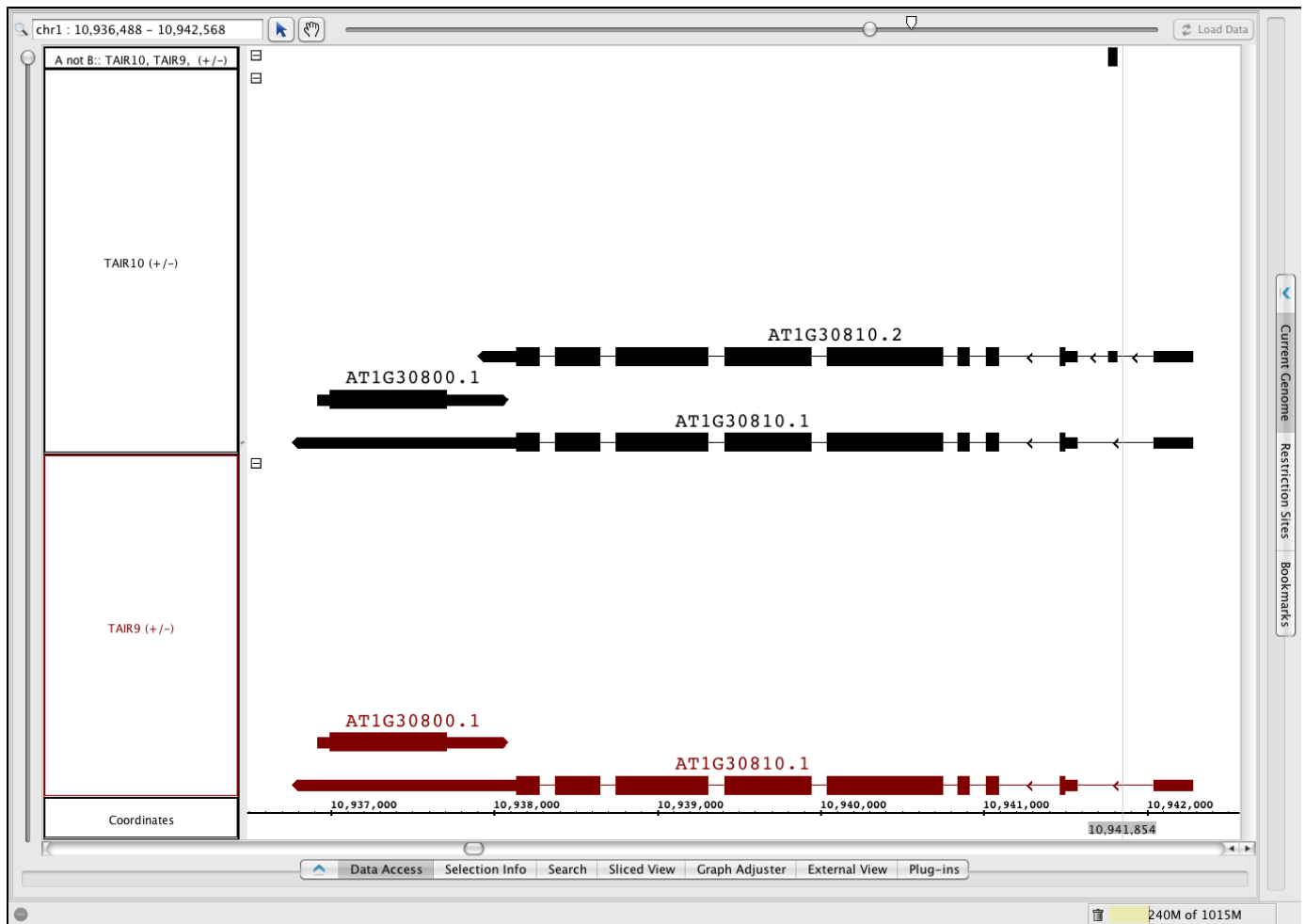


## Comparing new and old annotation releases

This image from IGB 6.7 demonstrates track operation "A not B" that creates a new track containing annotated regions present in A but absent from B. An "A not B" track can highlight differences between runs of the same analysis software using different parameters, new additions to an annotated gene set, and even differential expression across RNA-Seq data sets.

Here, the "A not B" track occupies the top row of the IGB display.

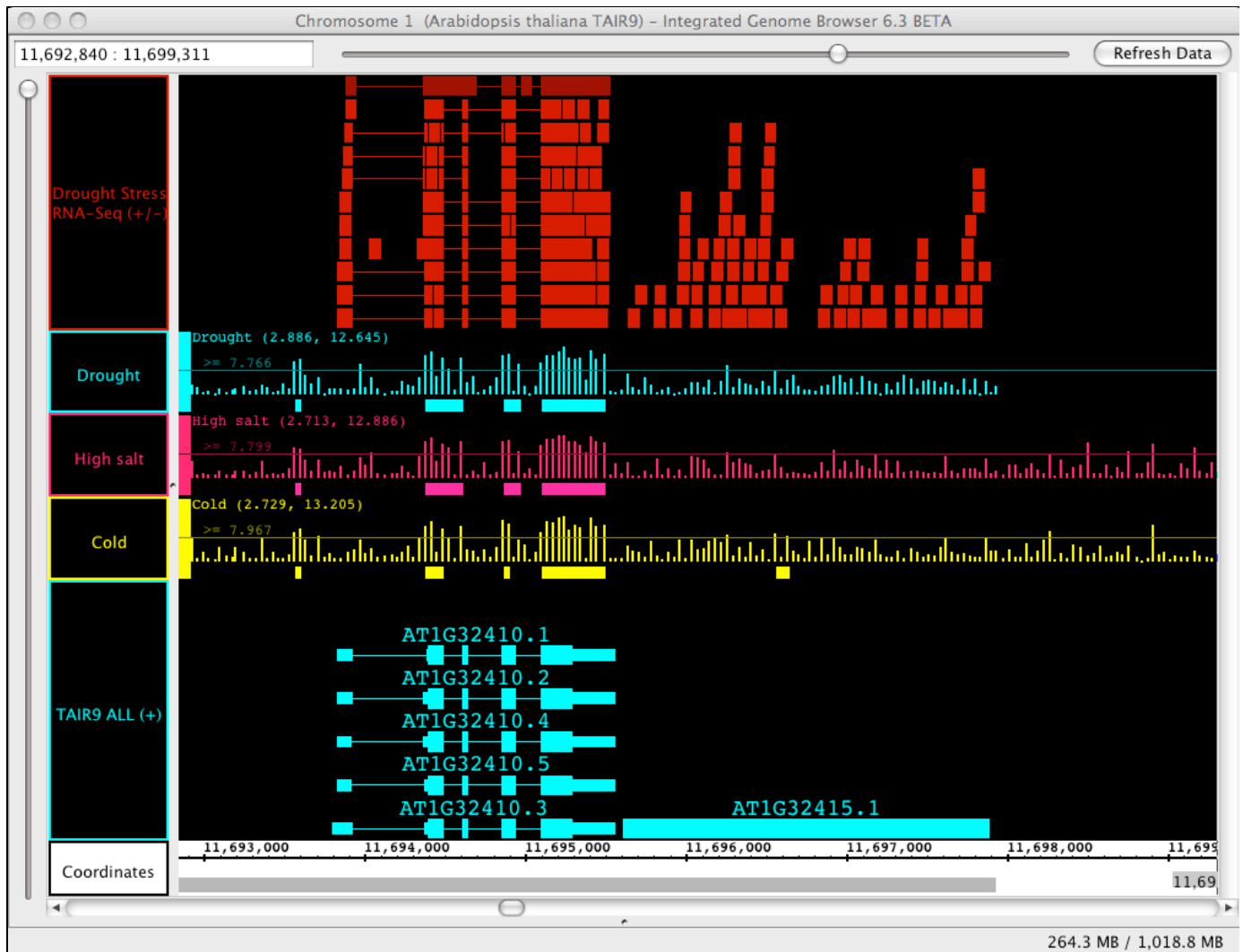




## Viewing RNA-Seq and expression tiling array data in the same view

### Viewing RNA-Seq and expression tiling array data in the same view in IGB 6.3

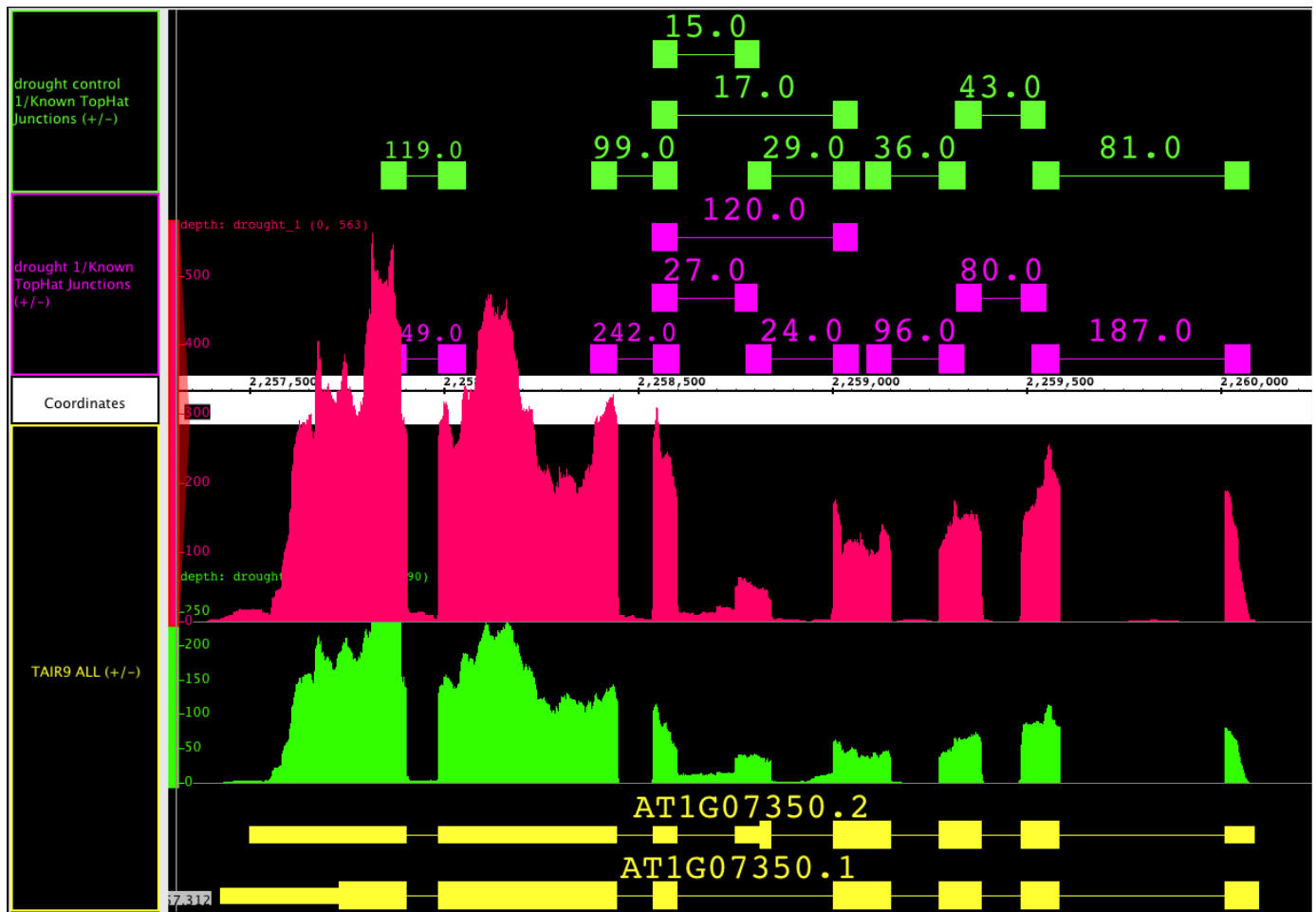
In this image, you see three different *Arabidopsis* expression tiling array data sets corresponding to cold (GSM243694), high salt (GSM243703), and drought (GSM243707) treatments assayed using the Affymetrix AtTile1R tiling array platform. The data were loaded in simple graph format, where probe intensities are shown as vertical bars. The graphs were then configured via the Graph Adjuster tab's Graph Thresholding option to display a bar underneath groups of consecutive probes with intensity values above a certain threshold. Note how the bars seem to correspond to known exons in the gene models displayed in the TAIR9 track. At the top of the display are short read Illumina RNA-Seq data (75 bases per read) from plants undergoing severe drought stress. Note that graph thresholding seems to suggest that this gene contains a previously undiscovered five prime exon. However, the RNA-Seq data contain no reads that support this idea. This example illustrates some of the ambiguities that can arise from using data from high-throughput methods - like tiling arrays and Illumina sequencing. Consider that the data sets are enormous and noisy! Thus, purely through chance we will observe at least some genes adjacent to what tiling arrays seem to suggest are unannotated exons. If the RNA-Seq data supported the idea that this region of high probe intensity is indeed an exon, then we would be much more likely to believe the conclusion, because the odds of two entirely different expression measurement technologies giving the same spurious result are very small.



## Visualizing output from TopHat and BowTie

### Visualizing output from TopHat and BowTie

The image below shows output from TopHat and BowTie, programs that align short reads from Illumina sequencing experiments onto a genome. The top two tracks are from a BED format file that TopHat creates in which each line of the BED file represents a splicing choice and the score field indicates the number of reads that supported that choice. The two floating graphs (red and green) are from WIG files the programs produce that summarize the number of reads covering individual regions. In this particular experiment, the control sample (green) produced many more reads than the treatment sample (red), and so the fact that there are more reads overall in the treatment (red) sample tells us this gene is up-regulated under the treatment.



## Visualizing RNA-Seq reads aligned onto a genome

### Visualizing RNA-Seq reads aligned onto a genome

This image from IGB shows short read sequences aligned onto the Arabidopsis A\_thaliana\_Jun\_2009 (TAIR9) genome. The track above the Coordinates track presents part of two Arabidopsis gene models, with sequence data loaded. Note how the reads seem to support two alternative splicing variations in the overlapping gene. Also note the nucleotide differences between the aligned reads (upper track) and the reference sequence.

Refresh Data

TAIR9 mRNA  
(+)

AT4G10760.1  
 AT4G10760.2  
 14,982,920 14,982,940 14,982,960 14,982,980 14,983,000 14,983,020  
 2,915 TCTTTTACAGAGTGTGTGTGTGTGTGATATATATTAACCTTAACCTTTCAGCAGGGGATAAATATGTCAGACATTGTTCCTTGTATGGTGACATTATGCTGTGGTCACTG

A\_thaliana\_Jun\_2009

Choose Load Mode for Data Sets:

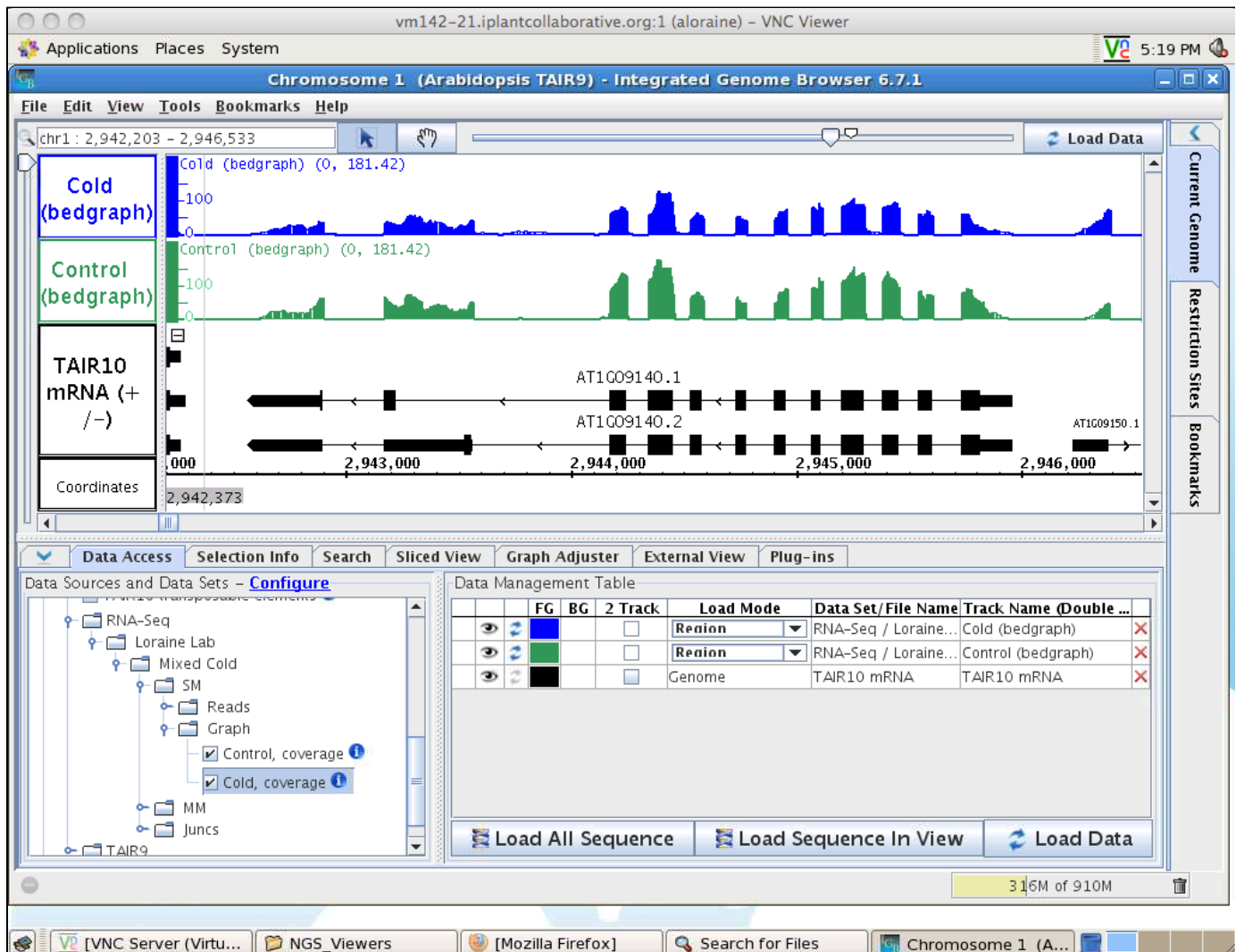
Choose Load M...	Data Set	Data Source
Region...	sample.b...	Local Files (Loc...
Whole Genome	TAIR9 mR...	Bioviz (QuickLo...

Sequence	Length
chr1	30427671
chr2	19698289
chr3	23459830
chr4	18585056
chr5	26975502
chrC	154478



## IGB 6.7.1 running in iPlant Atmosphere

iPlant Atmosphere offers a nifty way you can run IGB on an iPlant hosted virtual machine. To make this image, I logged into Atmosphere, requested a new virtual machine, and then used a free software tool called VNC Viewer to log in. From there, I downloaded and launched IGB from <http://bioviz.org/igb> just as I would do on my local desktop or laptop computer. What's great about this is that now I can use IGB to access files I've made in the iPlant Discovery Environment or files I've made by running software running on the VM itself. I can also load data from the IGBQuickLoad site in the usual way, letting me compare my results to publicly available data sets, like the one shown below. The data shown here come from an RNASeq experiment from the Loraine lab that we deployed the the publicly-accessible [IGBQuickLoad site](#) at UNC Charlotte.



## Advanced Features

### Scripting and the IGB command language

- Introduction
- Syntax
  - Commands
    - Open a genome version
    - Go back to the IGB start screen.
    - Go to a region
    - Add a new data set to IGB's list of active data sets
    - Remove a data set from IGB's list of active data sets
    - Change the load mode for an active (opened) data set
    - Load data from currently active data sets
    - Make images

- [Select genes](#)
  - [Wait \(sleep\) for commands to finish](#)
- [Comments](#)
- [Examples](#)
  - [Loading a BAM file](#)
  - [Exporting images](#)
  - <https://gist.github.com/Hillrunner2008/25976458488b65a3a25b>

## Introduction

You can control IGB from scripts, text files containing simple commands for navigation, loading data, and making snapshots.

To run an IGB script:

- Open a script file using **File > Open**. Script files should end with .igb.
- Click an IGB script link. See: [Controlling IGB using IGB Links](#)
- Drag and drop a script file into an open IGB window

**Note:** A prototype plug-in enabling python scripts (.py extension) was developed but is not yet released. If you would like to know more, please get in touch.

## Syntax

The script file has a simple, space-delimited syntax. Each command should occupy a separate line.

### Commands

#### Open a genome version

```
genome <version>
```

Open genome version (e.g., H\_sapiens\_Mar\_2006). Any genome version or synonym that's accepted by IGB is accepted here.

#### Go back to the IGB start screen.

```
homescreen
```

Return to the IGB start screen.

#### Go to a region

```
goto <region>
```

Go to a region of a chromosome (e.g., "chr1:40000-60000" or "chr1:40,000-60,000").

#### Add a new data set to IGB's list of active data sets

```
load <file or URL>
```

Load the specified file (identified by full path) or URL into IGB. The data set will appear in the user's **Data Management** table upon completion.

#### Remove a data set from IGB's list of active data sets

```
unload <file or URL>
```

### Change the load mode for an active (opened) data set

```
loadmode <NO_LOAD | REGION_IN_VIEW | CHROMOSOME | GENOME> <file or URL>
```

Change the load mode for the specified data set or file, identified by full path.



The **loadmode** command will work for any loaded track. Make sure to load your desired track before calling the **loadmode** command.

### Load data from currently active data sets

```
refresh
```

This is equivalent to clicking the **Load Data** button in IGB. This triggers loading of data from the active data sets available in IGB, as specified in the load command.

### Make images

```
snapshot <filename>
```

```
snapshotwholeFrame <filename>
```

```
snapshotmainView <filename>
```

```
snapshotmainViewWithLabels <filename>
```

```
snapshotsllicedViewWithLabels <absolue path to file>
```

Export the entire frame, main view, main view with labels, or sliced view, respectively, to a PNG format file. You must include the full path to the file or IGB will write the file to the location from where you launched IGB.

The filename defaults to "<genome\_version><timestamp>.png" if not specified.



The **snapshot** and **snapshotwholeFrame** commands only work on Mac and Linux if attempting to export an SVG file.

### Select genes

```
select <geneids>
```

Selects the given gene IDs, which are passed in as a comma-separated list.

### Wait (sleep) for commands to finish

```
sleep <milliseconds>
```

Sleep the application for x milliseconds.



IGB is multi-threaded, which means that the user may be able to operate IGB while data are being loaded. If your IGB script includes loading data from remote sites (e.g, RNASeq data from the IGBQuickLoad.org site) then you should include a sleep command to give IGB time to load the data.

## Comments

Lines starting with the IGB script comment character '#' are ignored.

## Examples

### Loading a BAM file

```
# select genome
genome A_thaliana_Jun_2009
# go to a region in chr4
# commas are optional
goto chr4:8,860,484-8,872,096
# load a file
load
http://igbquickload.org/quickload/A_thaliana_Jun_2009/cold_stress/cold_control.sm.bam
# load the data
refresh
```

## Exporting images

<https://gist.github.com/Hillrunner2008/25976458488b65a3a25b>

## IGB Script List

Commands	Arguments	Description	Examples
bringToFront	none	Make IGB the top window on the Desktop.	bringToFront
genome	<version>	Load a genome version. Accepts IGB-friendly genome version names (e.g., H_sapiens_Feb_2014) and synonyms.	genome H_sapiens_Mar_2006
goto	<region>	Go to a region. Any format that's accepted by IGB is accepted here. Commas are optional.	goto chr1:13,640,211-13,643,760 goto chr1:13640211-13643760



hidetrack	<file or URL>	Hide the track. Accepts one or more comma separated files identified by full path or URL.	hidetrack <a href="http://igbquickload.org/H_sapiens_Dec_2013/H_sapiens_Dec_2013_al.sl.gz">http://igbquickload.org/H_sapiens_Dec_2013/H_sapiens_Dec_2013_al.sl.gz</a>
homescreen	none	Return to the IGB Start Screen.	homescreen
load	<file or URL>	Load the specified file (identified by full path) or URL into IGB. The data set will appear in the <b>Data Management</b> table upon completion.	load <a href="http://igbquickload.org/D_melanogaster_Jul_2014/D_melanogaster_Jul_2014.sl.gz">http://igbquickload.org/D_melanogaster_Jul_2014/D_melanogaster_Jul_2014.sl.gz</a>
loadmode	<NO_LOAD   REGION_IN_VIEW   GENOME> <file or URL>	Change the load mode for a data set. <b>Note:</b> The <b>loadmode</b> command will work for any loaded track. Make sure to load your desired track before calling the <b>loadmode</b> command.	loadmode GENOME <a href="http://igbquickload.org/M_musculus_Dec_2011/M_musculus_011_refGene.bed.gz">http://igbquickload.org/M_musculus_Dec_2011/M_musculus_011_refGene.bed.gz</a>
loadsequence	none	This is equivalent to clicking the <b>Load Sequence</b> button in IGB. This triggers loading of the reference sequence for the region in view.	loadsequence
refresh	none	This is equivalent to clicking the <b>Load Data</b> button in IGB. This triggers loading of data from the active data sets available in IGB, as specified in the load command.	refresh
select	<geneids>	Selects the given gene IDs, which are passed in as a comma-separated list. If the command contains a single gene ID then IGB will select and jump-zoom to the given gene.	select AT1G36280.2,AT1G36240.1,AT1G36340.1
selectfeature	<geneids>	Selects the given gene IDs, which are passed in as a comma-separated list.	selectfeature
showtrack	<file or URL>	Unhide the specified track. Accepts one or more comma separated files identified by full path or URL.	showtrack <a href="http://igbquickload.org/O_sativa_japonica_Oct_2011/O_sativa_japonica_011.bed.gz">http://igbquickload.org/O_sativa_japonica_Oct_2011/O_sativa_japonica_011.bed.gz</a>
sleep	<milliseconds>	Sleep the application for x milliseconds.	sleep 6000

snapshot	<filename>  Note: specify the full path to file.	Create an image showing the entire IGB window. the entire frame an image file. The filename extension determines the format. Supported formats are .png, .jpeg, and .svg. <b>Note:</b> This command only works on Mac and Linux if attempting to export an SVG file.	snapshot alternativesplicing1.png
snapshotmainView	<filename>	Same as snapshot but only makes an image showing the main IGB data display.	snapshotmainView testimage.jpeg
snapshotmainViewWithLabels	<filename>	Export the main view with labels an image file. The filename extension determines the format. Supported formats are .png, .jpeg, and .svg.	snapshotmainViewWithLabels imageexport1.svg
snapshotslicedViewWithLabels	<filename>	Export the sliced view an image file. The filename extension determines the format. Supported formats are .png, .jpeg, and .svg.	snapshotslicedViewWithLabels April_1_2015.png
snapshotwholeFrame	<filename>	Export the entire frame to an image file. The filename extension determines the format. Supported formats are .png, .jpeg, and .svg. <b>Note:</b> This command only works on Mac and Linux if attempting to export an SVG file.	snapshotwholeFrame GeneAT1G36280wholeframe.jpeg
unload	<file or URL>	If a matching track is found for the specified URL/path it will be deleted. Accepts one or more comma separated files identified by full path or URL.	unload <a href="http://igbquickload.org/A_thaliana_Jun_2009/cold_stress/cold_control.sm">http://igbquickload.org/A_thaliana_Jun_2009/cold_stress/cold_control.sm</a>

## Scripting with R (Draft)

- [Introduction](#)
- [Getting Started](#)
  - [Start IGB](#)
  - [Time](#)
  - [Data set](#)
  - [Source file](#)
- [Bulk collection of images](#)
- [Images of a modified file](#)

## Introduction

Starting with IGB version 8.2, R users can launch and control Integrated Genome Browser from an R session.

This functionality depends on the [SRAdb](#) library from [BioConductor](#).

The scripts and the primary documentation for controlling IGB through R scripts can be downloaded from our Bitbucket repository: [R to IGB](#)

The current R functions are wrappers for the existing [IGB script syntax](#).

This document discusses two major uses of IGB in an automated setting. (The file `TestImages.Rmd` is an R markdown version of this page.)

In the first example, [a bulk collection of images](#), all of the genes that meet some criteria are listed and an IGB image for each one is generated and collected in a folder. You can come back to this folder later to quickly flip through this bulk image collection. This allows you to dedicate time to just a few images to make them into nice figures after looking at a huge number of images as candidate figures.

In the second example, [an image of a modified file](#), a new bed file is made and an up-to-date image of that file is generated and incorporated into the mark down that made the file. This makes it easy to see how the output of a process looks, without manually updating the image, so the image and the code are always in sync.

## Getting Started

The examples shown here use function from the `IGBfunctions.R` source file and a small test file, both of which are available in the [Bitbucket repository](#).

### Start IGB

Before running your R script, open IGB and use the user-interface to make tracks look the way you want. Set up the main view to have the aspect ratio you would like. For tracks that do not need to be modified, you can set up the stack height. If you are going to generate a file to view, you can make a test file with the same name and path and use that to set up the colors, name fields, etc. IGB will remember these preferences after you have closed the file, overwritten it with real file, and reloaded it. It is possible to start IGB through the R interface, but it requires java web start and on many machines that requires changing permissions. Opening IGB ahead at the start is recommended. Be aware, it may be necessary to close the file, and perhaps IGB as well, between runs of the script.

### Time

Some of these functions take time but right now they don't return any indication of having completed. To ensure that the images shown in this markdown are the ones generated by this run through (rather than the previous run), we will make the program wait for images generated after the start time by using the `waitForFile()` function. This wait function allows you to use these less-than-instantaneous functions in an automated way.

```
bedfile="data/ModifiedMiniBed.bed"
write.table(x=NULL, file=bedfile) #write the real file later, for now use this to get
a start time.
startTime = file.info(bedfile)$ctime
```

### Data set

The `Ath-test.bed` file includes a subset of lines from the `TAIR10.bed.gz` file: [http://www.igbquickload.org/A\\_thaliana\\_Jun\\_2009/TAIR10.bed.gz](http://www.igbquickload.org/A_thaliana_Jun_2009/TAIR10.bed.gz). The subset includes several genes from chromosomes 1, 2 and 5.

```
T10 = "data/Ath-test.bed"
T10 = read.table(file=T10, sep="\t", quote="", header=F)
bednames=c("chromosome", "chromStart", "chromEnd", "name", "score", "strand",
"thickStart", "thickEnd", "itemRGB", "blockCount", "blockSizes", "blockStarts", "ID",
"description")
names(T10)=bednames
# all of the output images will saved in a folder called images
system("mkdir images")
```

### Source file

`IGBfunctions.R` is a sourceable version of `IGBFuctions.Rmd` with a few added functions.

Set up the IGB interface.

```
source("IGBfunctions.R")
# system('rm -rf ~/.igb/bundles')
# startIGBdev()
socket = IGBsocket()
```

```
## [1] "Success! 5"          "Success! localhost" "Success! 7084"
```

```
IGBbringToFront()
IGBgenome(socket, 'A_thaliana_Jun_2009')
```



Note: startIGBdev() and startIGB() use Java Web Start, so (at least on some machines) it requires permission, and can't be part of an automated process unless you change your security settings.

## Bulk collection of images

Which genes in the test set have more than 3 transcripts?

```
set = names(table(T10$ID))[table(T10$ID) > 3]
```

This set includes 6 genes.

Take a picture of each one.

```
for (gn in set){
  # define image range
  w = which(T10$ID == gn)
  edges=c(c(T10[w,"chromStart"]),c(T10[w,"chromEnd"]))
  low = min(edges)
  high = max(edges)
  length = high - low
  side = round(length*.2, 0)
  imageStart = low - side
  imageEnd = high + side
  goto = paste0(T10$chromosome[w][1], ":", imageStart, "-", imageEnd)

  # save IGB image
  id = as.character(T10$name[w][1])
  IGBgoto(socket, goto)
  IGBselect(gn) # selects the first gene model, to show edge matching
  fname=paste0("images/", id, ".png")
  IGBsnapshotMainViewWithLabels(socket, fname)
  waitForFile(startTime, fname)
}
```

Now you should have a folder full of IGB images showing the edge matching of the first transcript compared to all other transcripts for each gene that has more than three models.

```
dir("images")
```

```
## [1] "AT1G01060.3.png"      "AT1G01100.2.png"
## [3] "AT2G43160.1.png"      "AT2G43180.2.png"
## [5] "AT2G43190.4.png"      "AT5G24630.1.png"
## [7] "test-newView-main.png" "test-unaltered-main.png"
## [9] "test-unaltered.png"
```

## Images of a modified file

Select a random gene from the taир10 set.

```
geneRow = sample(1:nrow(T10), size=1)
gene = T10[geneRow,]
```

We have selected the AT2G42540.3 transcript of the gene COR15|COR15A.

Take a picture of that gene, as it appears in the IGB quickload file.

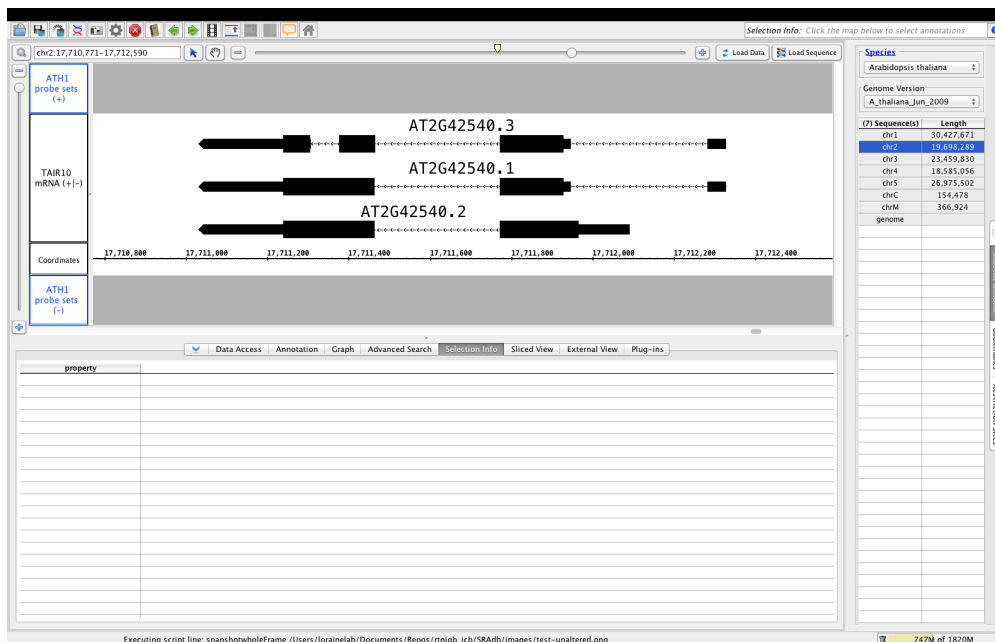
```
length = gene$chromEnd - gene$chromStart
side = round(length *.2, 0)
imageStart = gene$chromStart - side
imageEnd = gene$chromEnd + side
goto = paste0(gene$chromosome, ":", imageStart, "-", imageEnd)

IGBgenome(socket, 'A_thaliana_Jun_2009')
IGBgoto(socket, goto)

fname = "images/test-unaltered.png"
IGBsnapshotWholeFrame(socket, fname)
waitForFile(startTime, fname, 20)
```

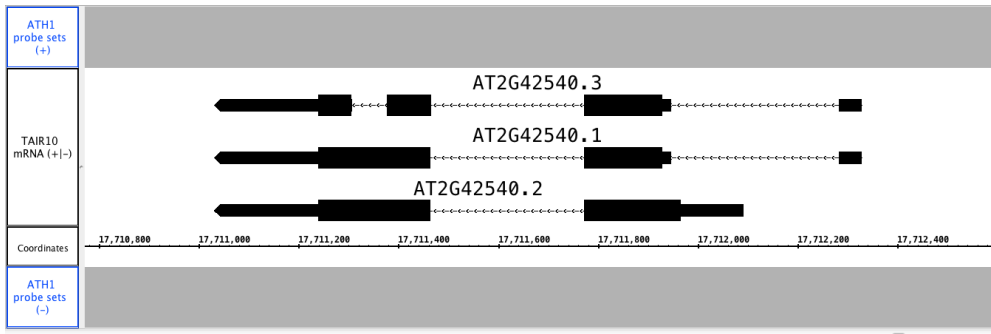
Use html to embed the image into the html output of knitr.

```
<img src=`r fname` alt="oops, was the file `r fname` created?">
```



Get just the main view with labels.

```
fnameMain = "images/test-unaltered-main.png"
IGBsnapshotMainViewWithLabels(socket, fnameMain)
waitForFile(startTime, fnameMain)
```



Copy that record and make two duplicates that are slightly upstream and downstream of that gene.

```
gene[2,]=gene[1,]
gene[3,]=gene[1,]
shift = length - 100
cols=c("chromStart", "chromEnd", "thickStart", "thickEnd")
gene[2,cols] = gene[2,cols] - shift
gene[3,c("chromStart", "chromEnd")] = gene[3,c("chromStart", "chromEnd")] + shift
gene[3,c("thickStart", "thickEnd")] = gene[3,c("chromStart", "chromEnd")] + shift
```

Put the copied gene record and its two duplicates into a bed file.

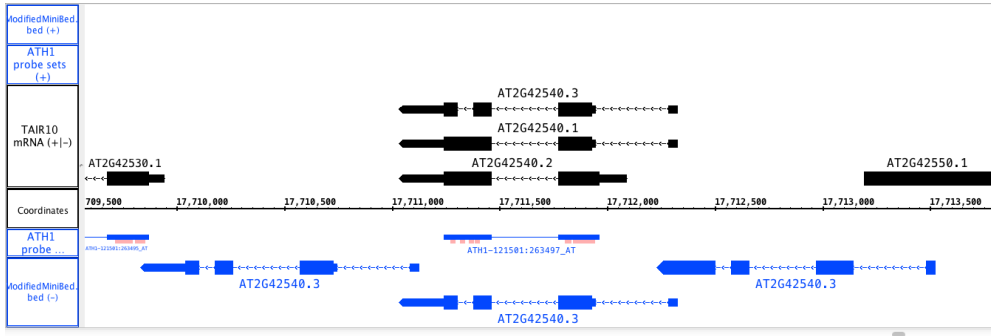
```
write.table(x=gene, file=bedfile, quote=F, sep='\t', row.names=F, col.names=F)
```

Create the parameters for the new view window.

```
imageStart = gene$chromStart[2] - side
imageEnd = gene$chromEnd[3] + side
goto = paste0(gene$chromosome, ":", imageStart, "-", imageEnd)
```

View this file in IGB next to the original.

```
IGBload(socket, paste0(getwd(), "/", bedfile))
IGBgoto(socket, goto)
IGBrefresh()
fnameNew = "images/test-newView-main.png"
IGBsnapshotMainViewWithLabels(socket, fnameNew)
waitForFile(startTime, fnameNew)
```



## Controlling IGB using IGB Links

- Introduction
- IGB links syntax
  - Show a region
- Other options
  - Scriptfile
  - Scriptfile examples
    - Load a species and navigate to a region
    - Open a data set

### Introduction

You can control IGB by embedding IGB links in Web pages. This is useful when you need to manually inspect a large number of regions. By clicking links in an IGB links page, you can quickly move to a new region.

### IGB links syntax

When IGB launches, it starts a "bookmark server" at localhost, port 7085.



If you launch a second instance of IGB while one is already running, the new instance will listen to a different port. For this reason, bookmark links will only work for the first IGB instance only.

IGB uses this internal server to handle HTTP requests.

An IGB link base URL is structured as

```
http://localhost:7085/UnibrowControl?[parameter]
```

### Show a region

IGB link syntax to show a region is:

```
http://localhost:7085/UnibrowControl?version=[version]&seqid=[seqid]
start=[start]&end=[end]
```

Parameters include:

- version - genome version
- start - start position of the region to show
- end - end position of the region to show
- seqid - chromosome sequence

Example:

[http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=A\\_thaliana\\_Jun\\_2009](http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=A_thaliana_Jun_2009)



Other parameters are also available. If you're developing IGB links, it helps to open up an exported IGB bookmarks file in a text editor and take a look at how the links are structured.

## Other options

### Scriptfile

As of IGB 6.3, you can use the scriptfile parameter to send commands to IGB, where scriptfile designates a local file or URL

```
http://localhost:7085/UnibrowControl?scriptfile=[file or URL]
```

- For IGB command line syntax, see [Scripting and the IGB command language](#).

### Scriptfile examples

#### Load a species and navigate to a region

[http://127.0.0.1:7085/UnibrowControl?scriptfile=http://wiki.transvar.org/confluence/download/attachments/17269487/load\\_and\\_navigate.igb](http://127.0.0.1:7085/UnibrowControl?scriptfile=http://wiki.transvar.org/confluence/download/attachments/17269487/load_and_navigate.igb)

#### Open a data set

[http://127.0.0.1:7085/UnibrowControl?scriptfile=http://wiki.transvar.org/confluence/download/attachments/17269487/add\\_new\\_data.igb](http://127.0.0.1:7085/UnibrowControl?scriptfile=http://wiki.transvar.org/confluence/download/attachments/17269487/add_new_data.igb)

## IGB link parameters

## Command Line Arguments

- [Run IGB from the command line](#)
- [IGB command line options and arguments](#)

### Run IGB from the command line

To run IGB at the command line, download the latest release (igb.zip) from <http://sourceforge.net/projects/genoviz/>.

Uncompress igb.zip to make a new directory igb which contains the igb\_exe.jar, which you can run at the command line.

The following command runs IGB with 1 Gb of RAM.

```
$ java -Xmx1g -jar igb_exe.jar [IGB options]
```

### IGB command line options and arguments

-href [URL]

- Open an IGB control bookmark.

## Track lines for configuring track appearance

- [Introduction](#)
- [Track line syntax](#)
- [Supported track line properties](#)

### Introduction

For many text-based file formats, it is possible to configure the track name, color, and display properties through the use of track lines in the file header or comments.



Track lines were invented for use in the UCSC genome browser, and the full format is described on the UCSC Genome Bioinformatics web site.

IGB understands a subset of the full format, and will ignore property settings that it doesn't use.

## Track line syntax

In ".wig", ".bed", ".psl", or ".gff" format (version 1 or 2), the track line begins with "track".

In ".gff" format version 3, the track line must begin with "##track" to indicate it as a meta-directive, since track lines would otherwise disobey the GFF3 specification.

GFF3 has its own rich methods of specifying annotation styles. A future version of IGB may make use of this.

Here is a sample track line that can be understood by IGB:

track name="Example" description="Example Description" visibility=full color=200,100,0 useScore=1 url="http://sample.com/search?id=\$\$"

## Supported track line properties

IGB understands these track line properties:

name	Short track name.
color	Track color, in the format r,g,b; where r, g, and b are numbers from 0 to 255.
description	Human-readable short description of the track.
visibility	In IGB, visibility=1 and visibility=dense will cause the track to be displayed in "collapsed" mode. Otherwise, the track displays in "expanded" mode.
url	A URL for linking to an outside website. Any instances of \$\$ in the URL will be replaced with the item id. See <a href="#">Configuring Web links (LinkOuts) to external sites</a> .
useScore	Whether to convert score data to color values. Valid values are useScore=0 and useScore=1 Files in ".bed" and ".gff" format can contain scores.
itemRgb	If itemRgb=On, use the data in column 9 of ".bed" files to determine the colors of individual annotations. Using this function requires additional memory.
graphType	Sets the graph type for graph files in the "wiggle" format. Can be graphType=bar or graphType=points.

## Matching species and genome versions

### How IGB recognizes species and genome versions

IGB is distributed with two tab-delimited files called "species.txt" and "synonyms.txt" that allow it to match genome version names across QuickLoad sites, Galaxy, and the UCSC Genome Bioinformatics system. Also, whenever IGB accesses a QuickLoad site, tries to download species.txt and synonyms.txt from the top-level, root directory of the site. If the files are not present, no harm is done, but if they are, IGB can use them to add new species and genome versions to the **Species** and **Genome Version** menus in the **Current Genome** tab. IGB does this so that if you have set up your own [custom IGB QuickLoad site](#) with new genome, IGB will display your data in the same way as other genomes.

See:

- [Use species.txt to link species names to IGB genome names](#)
- [Use synonyms.txt to link genome versions names to each other](#)

## Use species.txt to link species names to IGB genome names

The IGB file **species.txt** is a tab-delimited file that lists

- Column 1: binomial (Latin) names for species
- Column 2: common name for the species
- Column 3: IGB-friendly genome version name prefix (e.g., H\_sapiens or A\_gambiae)
- Column 4, 5, etc. (optional): Genome version name prefixes used by Galaxy, UCSC Genome Bioinformatics, or other data providers

IGB uses these data to populate the Species menu in the **Current Genome** tab. It also uses these data to associate species with genome assemblies. This is why the species.txt file also lists the IGB-friendly genome version prefix in the file in column 3. Because many data providers also use common prefixes to indicate genome versions for a species, we also include those whenever possible.

When IGB accesses a QuickLoad site, it tries to retrieve a species.txt file from the QuickLoad root directory. If the file is available, IGB reads the file and adds any new species to the Species menu.

The species.txt file that is distributed with IGB is version-controlled as part of the IGB source code repository. To see the latest version of this file, go to <https://bitbucket.org/lorainelab/integrated-genome-browser>, select the source button, and navigate to the resources folder as shown below.

For developers: This is the "master branch" which we use to create the released versions of IGB you download at BioViz. See also [Developing IGB from the IGB Developers Guide](#).

**2. go to species.txt**

**1. select source**

1	Ailuropoda melanoleuca	Giant panda	A_melanoleuca	ailMel	Ailuropoda_melanoleuca
2	Alligator mississippiensis	American alligator	A_mississippiensis	allMis	
3	Anolis carolinensis	Anole Lizard	A_carolinensis	anoCar	Anolis_carolinensis
4	Anopheles gambiae	Mosquito	A_gambiae	anoGam	

binomial name      common name      IGB genome name prefix      Other data provider prefixes

Also, [here is a version](#) that was current as of December, 2014.

### How to make a species.txt file for a QuickLoad site

First check to see if your species and genome version prefixes are already present in the species.txt file. If they are, then you do not need to distribute a species.txt file with your QuickLoad site. If not, then to make a species.txt file for a QuickLoad site:

- Open an editor (Microsoft Word is fine, but you'll need to save the file as a **plain text** (.txt) file.)
- Type the Latin name for the species, including a subspecies (e.g., Zea mays B73).
- Type TAB
- Type the common name for the species (e.g., maize)
- Type TAB
- Type the **IGB genome version prefix** you'll use for every genome assembly from this species (e.g., Z\_mays\_B73)
- Type TAB
- Type the **Galaxy/UCSC genome version prefix**, if available. You only need to do this if the genome versions is supported by Galaxy and UCSC and you would like to use open files and data from Galaxy in IGB.
- Repeat the previous steps for all new species you'd like to include in your QuickLoad site.
- Save the file as plain text.
- Place the file in the root directory of your QuickLoad site.

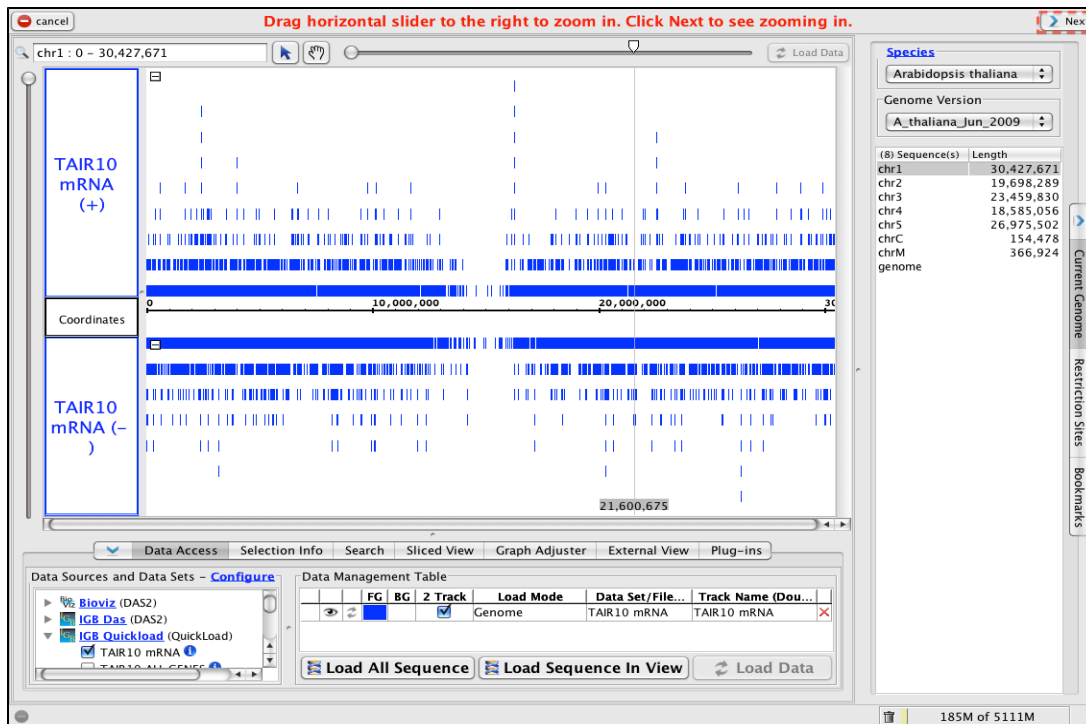
### Use synonyms.txt to link genome version names to each other

The IGB file **synonyms.txt** is a tab-delimited file that lists alternative names for the same genome version. Each line contains one or more synonyms.

Before making your own synonyms.txt file, first find out if there is entry for it in the synonyms.txt file that is packaged with the latest IGB release. To see the latest version of this file, go to <https://bitbucket.org/lorainelab/integrated-genome-browser>, select the source button, and navigate to the resources folder as shown below.

For developers: This is the "master branch" which we use to create the released versions of IGB you download at BioViz. See also [Developing IGB from the IGB Developers Guide](#).





## Configuring Web links (LinkOuts) to external sites

- Introduction
  - LinkOut contents
- How to create a new LinkOut
- Links to documentation on Java regular expressions

### Introduction

When users right-click a gene model or other item in the display, IGB displays a menu with links to external Web sites. These links to external Web sites are called "LinkOuts."

You can create new LinkOuts using the Web links tool.

### LinkOut contents

A LinkOut consists of

- **URL** the Web page IGB will open when users right-click an item
- **Name** the LinkOut name displayed in the right-click menu
- **Pattern** the pattern IGB matches against track id or item id to create the LinkOut

Every item in the IGB display has an **id** or identifier.

IGB uses this id to create the URL.

### How to create a new LinkOut

Click **Tools > Web Links** to open the **Web Links** tool.

Use the Web Links tool to create new LinkOuts. Use the existing LinkOuts as a guide to making new ones.

To find out an item's id

- Click the item to selected it
- Click the **Selection Info** tab or Info button in the **Selection Info** box (top right)

### Links to documentation on Java regular expressions

Java Regular expression: <http://docs.oracle.com/javase/1.4.2/docs/api/java/util/regex/Pattern.html>

Regular Expression Online Generator: <http://txt2re.com/>

Regular Expression Online tester: <http://www.regexplanet.com/advanced/java/index.html>

# Controlling IGB using IGB Links with Genome Synonyms

IGB recognizes commonly-used synonyms for a variety of genomes.

If you are creating IGB bookmarks, you can use these synonyms to refer to genome versions.

## Examples:

### Example 1. To load *Arabidopsis thaliana* TAIR9 for chr1: 20,633,626 - 20,636,669

[http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=A\\_thaliana\\_Jun\\_2009](http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=A_thaliana_Jun_2009)

<http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=TAIR9>

<http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=Tair9>

### Example 2. To load *Homo sapiens* Feb 2009 for chr1 : 52,065,515 - 52,267,144

[http://localhost:7085/UnibrowControl?seqid=chr1&end=52267144&start=52065515&version=H\\_sapiens\\_Feb\\_2009](http://localhost:7085/UnibrowControl?seqid=chr1&end=52267144&start=52065515&version=H_sapiens_Feb_2009)

<http://localhost:7085/UnibrowControl?seqid=chr1&end=52267144&start=52065515&version=hg19>

[http://localhost:7085/UnibrowControl?seqid=chr1&end=52267144&start=52065515&version=human\\_ncbi37](http://localhost:7085/UnibrowControl?seqid=chr1&end=52267144&start=52065515&version=human_ncbi37)

### Example 3. To load *Mus musculus* July 2007 for chr1 : 120,520,690 - 120,583,464

[http://localhost:7085/UnibrowControl?seqid=chr1&end=120583464&start=120520690&version=M\\_musculus\\_Jul\\_2007](http://localhost:7085/UnibrowControl?seqid=chr1&end=120583464&start=120520690&version=M_musculus_Jul_2007)

<http://localhost:7085/UnibrowControl?seqid=chr1&end=120583464&start=120520690&version=mm9>

<http://localhost:7085/UnibrowControl?seqid=chr1&end=120583464&start=120520690&version=Mm:NCBI37>

### Example 4. To load *Caenorhabditis elegans* May 2008 for chrII : 4,082,693 - 4,089,972

[http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=C\\_elegans\\_May\\_2008](http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=C_elegans_May_2008)

<http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=ce6>

<http://localhost:7085/UnibrowControl?seqid=chr1&end=20636664&start=20633631&version=WS190>

### Example 5. To load *Bos taurus* Oct 2007 for chr1 : 50,368,385 - 50,708,462

[http://localhost:7085/UnibrowControl?seqid=chr1&end=18659361&start=18080205&version=B\\_taurus\\_Oct\\_2007](http://localhost:7085/UnibrowControl?seqid=chr1&end=18659361&start=18080205&version=B_taurus_Oct_2007)

<http://localhost:7085/UnibrowControl?seqid=chr1&end=18659361&start=18080205&version=bosTau4>

[http://localhost:7085/UnibrowControl?seqid=chr1&end=18659361&start=18080205&version=Cow\\_Oct\\_2007](http://localhost:7085/UnibrowControl?seqid=chr1&end=18659361&start=18080205&version=Cow_Oct_2007)

## Personal Synonyms

- [Introduction](#)
- [How to make a chromosomes.txt file](#)
- [How to make a synonyms.txt file](#)
- [Loading synonym files](#)

## Introduction

Different data sources and files sometimes use different names to mean the same thing. To ensure that data reference to the same thing can be displayed together, IGB supports the concept of genome version and chromosome name synonyms.

IGB supports two synonym file types:

- `synonyms.txt` - lists synonyms for genome version names
- `chromosomes.txt` - lists synonyms for sequence names

Synonym files are tab-delimited. Each row of the file reports a list of synonyms for the same thing, with synonyms separated by tabs. A row can contain two or more columns of data. Synonyms can contain spaces or any other character except a tab.

The value listed first in each row is typically what IGB will display.

**Note:** If you create your own custom QuickLoad site, you can add `chromosome.txt` and `synonyms.txt` synonym files to your QuickLoad root directory (same level as `contents.txt`) and IGB will read them and adds the synonyms to its internal list of synonyms. Including one or both files is optional.

For more information about creating a QuickLoad site, see: [Sharing data using QuickLoad sites](#).

## How to make a chromosomes.txt file

Do this if you have an alignment file or other data file that made with a reference genome sequence that used different names for chromosomes than what you see when you load that genome into IGB.

Create a tab-delimited file that uses the IGB preferred name in the first column and the synonym in the second column. You can add more synonyms in subsequent columns if you like.

To create a chromosomes.txt synonym file:

1. Open a new file in a plain text editor (such as Notepad or TextEdit).
2. Enter the preferred IGB name (the display name) in column 1.
3. Press TAB to insert a tab character.
4. Enter the next synonym.
5. Repeat step 3 and 4 until all synonyms are entered.
6. Save file as **chromosome.txt**

## How to make a synonyms.txt file

A synonyms.txt file lists synonyms for genome versions. IGB uses this file to integrate data from multiple sources when those sources use different names for the same genome. For information about how to create a synonyms.txt file, see: [Use synonyms.txt to link genome version names to each other](#).

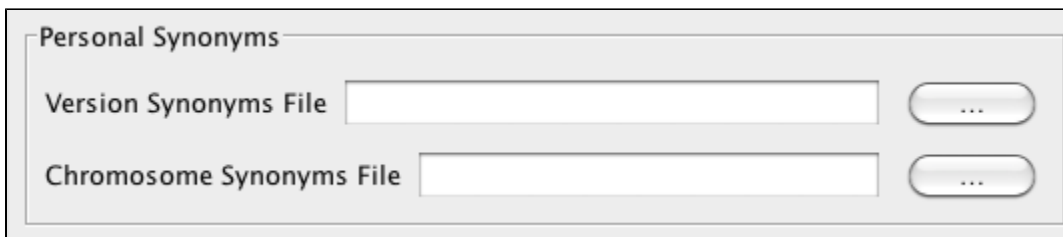
## Loading synonym files

After you create your synonym file, you need to load it into IGB.

To load a synonym file into IGB:

1. Select **File > Preferences > Data Sources**.
2. Under Personal Synonyms, select the file you want to add.
  - a. Select **Version Synonyms** to add synonyms.txt
  - b. Select **Chromosome Synonyms** to add chromosomes.txt
3. Choose and open file.
4. Restart IGB.

**Note:** If you clear your preferences or install a new version of IGB, you may need to reload your synonym files.



## Preferences

### Preferences

IGB offers many options to customize track appearance, add new data servers, and change data loading behavior.

Most customization options are available via the **Preferences** window which contains several tabbed panels devoted to different types of preferences.

To open the **Preferences** window

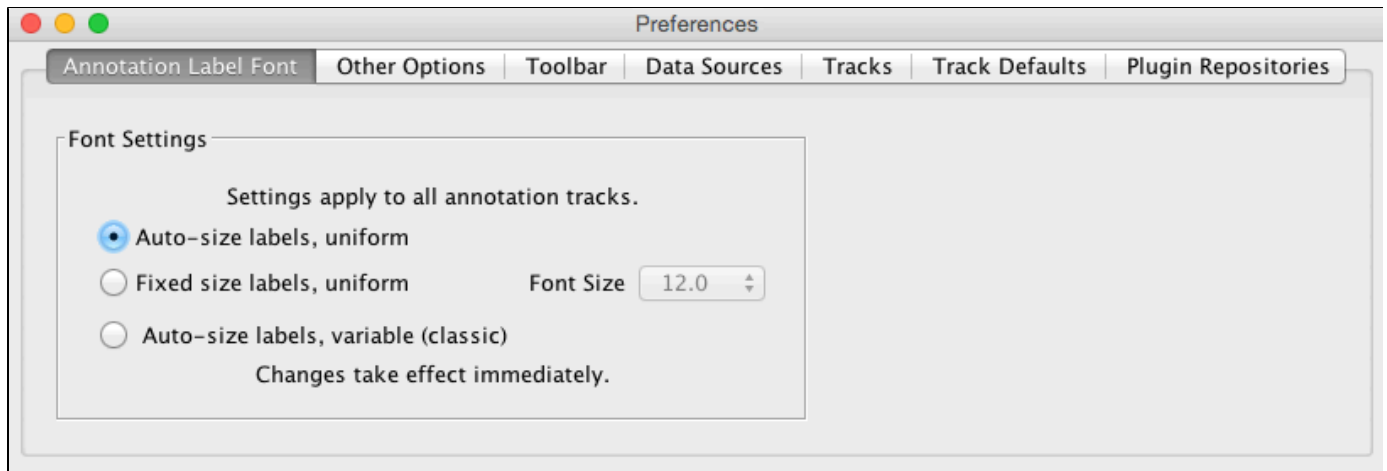
- Choose **File > Preferences** or
- Click the **Preferences** shortcut icon in the Toolbar

The following links contain descriptions of each tab in the Preferences window.

## Annotation Label Font

### Introduction

The **Annotation Label Font** tab lets you select **Font Settings** for all annotation labels in IGB.



 Settings apply to all annotation tracks and changes take effect immediately.

## Auto-size labels, uniform

With the **Auto-size labels, uniform** setting track height sets the label font size. All annotations in a track have the same font size. Partial labels with ... sometimes appear.

## Fixed size labels, uniform

With the **Fixed size labels, uniform** setting you set the label font size. All annotations in a track have the same font size. Partial labels with ... sometimes appear.

## Auto-size labels, variable (classic)

With the **Auto-size labels, variable (classic)** setting annotation size and track height set the label font size. Different annotations can have different font sizes.

## Other Options

- [Introduction](#)
- [Coordinates](#)
- [Residue Colors](#)
  - [ORF \(Open Reading Frame\) Analyzer](#)
  - [Edge Match](#)
- [Codon Display](#)
- [Track Resize Behavior](#)
- [Data loading and exit behaviors](#)
  - [Confirmations](#)
- [Reset all preferences to defaults](#)

## Introduction

Select **File > Preferences > Other Options** to configure the following:

- [Coordinates track number format](#)

- Sequence base colors
- ORF analyzer
- Edge matching
- Gene model translations
- Track resizing
- IGB exit and data set loading behaviors

#### Preferences Other options

The screenshot shows the 'Preferences' dialog box with the 'Other Options' tab selected. The dialog has a title bar with standard macOS window controls. Below the title bar are four tabs: 'Tracks', 'Track Defaults', 'Toolbar', and 'Other Options'. The 'Other Options' tab contains several sections: 'Coordinates' with a 'Number format' dropdown set to 'COMMA'; 'Residue Colors' with color swatches for A (green), T (blue), G (yellow), C (orange), and Other (gray); 'ORF Analyzer (Sliced View tab)' with 'Stop Codon' (red), 'Dynamic ORF' (green), and 'Background' (gray); 'Edge Match' with a 'Color' dropdown and a checked 'Show Edge Matching' checkbox; 'Codon Display' with radio buttons for 'Hide', 'One Letter' (selected), and 'Three Letter'; 'Track Resize Behavior' with radio buttons for 'Adjust All Tracks' and 'Adjust Adjacent Tracks' (selected); and a section with four checkboxes: 'Ask before exit' (checked), 'Confirm before loading large data set' (checked), 'Confirm before delete' (checked), and 'Display Errors on Status Bar' (unchecked). At the bottom is a 'Reset preference to defaults' button.

## Coordinates

These are options for changing number appearance. **FULL** shows the number with no thousands separators and the unit 'kb'. **COMMA** shows the numbers with commas at every  $10^3$  (U.S. style) and with no unit designations. This is the default. **ABBREV** will show the numbers as abbreviated as is reasonable, with unit designations. For example, when zoomed out, numbers appear as 10M, 20M etc. Changes made here to **Number Format** take effect immediately.

## Residue Colors

IGB displays letters representing DNA bases (residues) when zoomed-in. Residues are shown in alignment tracks, sequence tracks, and in underneath the coordinate axis.

By default, IGB uses cool colors for A and T residues and warmer colors for G and C residues. **Other** refers to residues that are either 'N', 'X' or '-' (missing) and can be set to a bright color if you are looking for deletions in short read sequences; otherwise you may wish to leave it as the neutral gray.

Click a color swatch to change base colors.

## ORF (Open Reading Frame) Analyzer

These preferences affect colors used to indicate stop codons, the ORFs and the background of the ORF tracks that are shown in the **Sliced View** tab.

## Edge Match



Edge matching highlights features that have identical boundaries. Edge matching becomes active when you click items in the display.

Click the swatch to change the edge match color. The selected color applies all annotation tracks.



Choose a color that has good contrast between foreground and background colors

## Codon Display

The **Codon Display** is activated when the sequence of a gene model is loaded and the user zooms in close enough to see the sequence. The codon translation appears within the gene model. User can set the amino acids to appear in one letter format (default), three letter format or be hidden. The colors is based on the chosen foreground color for the track.



Codon display is active only for tracks loaded from 14-column BED detail files.

## Track Resize Behavior

**Track Resize Behavior** is how IGB adjusts other tracks when a specific track height is changed. **Adjust All Tracks** means that as a single track is increased in height, all other tracks are reduced by a commensurate amount. **Adjust Adjacent Tracks** (default) means that all added or subtracted height is removed/added to the bordering track, leaving the rest of the display unaffected.

## Data loading and exit behaviors

### Confirmations

IGB warns you when you are about to exit the program and before you delete a file or files. Each of these warnings has the option to not be shown again. To turn off these warnings or reset your previous choice, check (or uncheck) these options.

## Reset all preferences to defaults

Click **Reset preferences to defaults** to reset all the preferences to the IGB defaults. This is sometimes useful if you need to troubleshoot IGB.



**Be careful**, as this action will reset ALL preferences and will also cause IGB to forget the location you were last viewing.

## Toolbar

- [Introduction](#)
- [Adding Shortcuts to the Toolbar](#)

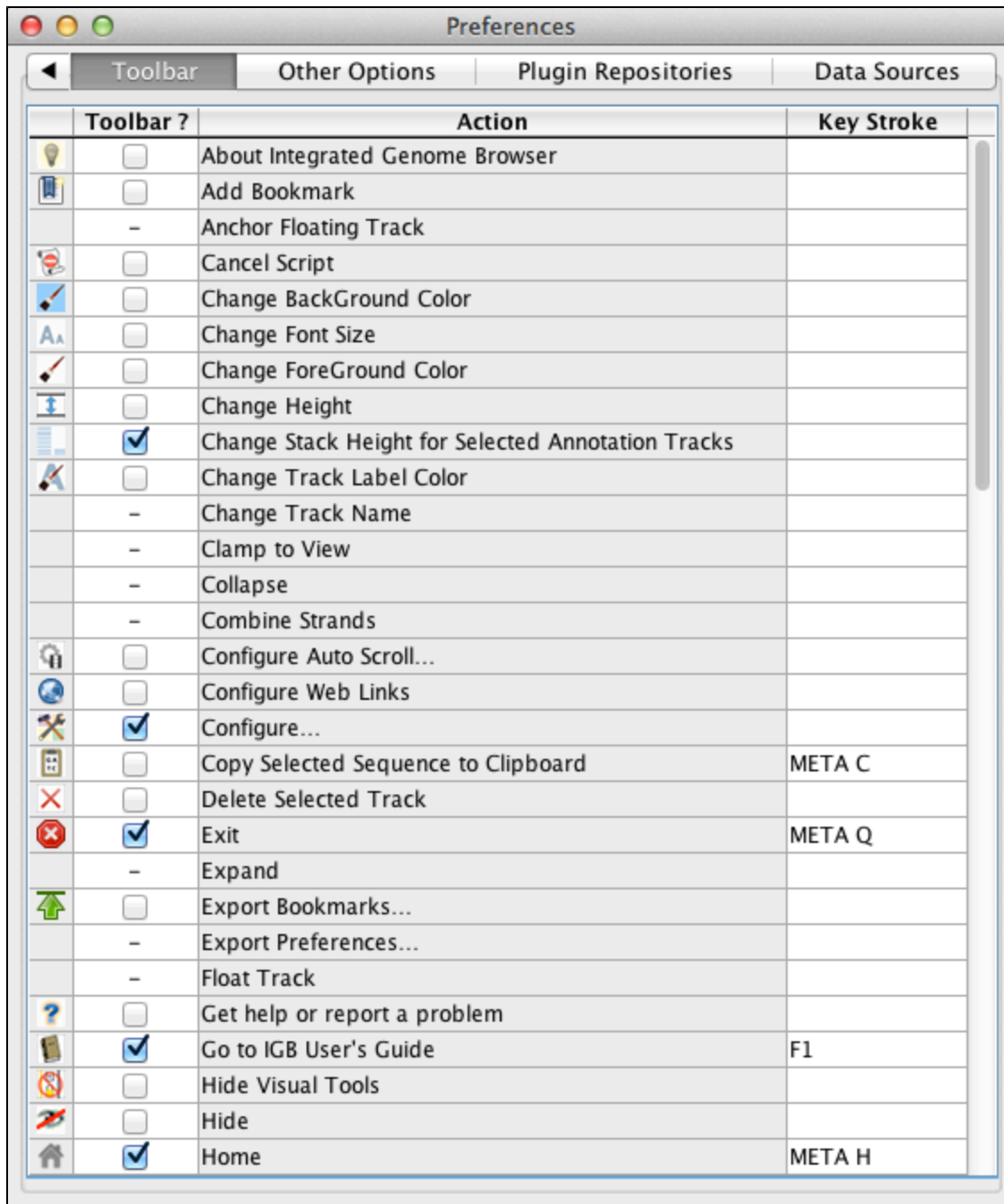
### Introduction

The IGB 7.0 introduced the **Toolbar**, which provides one-click access to most IGB functions. To see a list of IGB functions you can add to the **Toolbar**, select **File > Preferences > Toolbar**. Use the Toolbar tab in the IGB to view or change keystroke shortcuts for these and other IGB functions.

### Adding Shortcuts to the Toolbar

To select a shortcut function to add to the **Toolbar**, check its box in the **Toolbar?** column. To the left you can see the icon as it will appear in the **Toolbar**, and to the right is the name of the function. If you would like to add a function to the **Toolbar** that does not have an icon, let us know! If you would like a function that does not appear in the list at all, let us know! We are always happy to help out with new features and options.

If you would prefer to set and use **Key Stroke** short cuts, click in the **Key Stroke** field for your selected shortcut, and the cursor will appear in an open text box. Type the command key(s) you want to use (CONTROL/META, ALT or SHIFT, or combinations) + a letter or number key. Type ENTER after you set the key stroke combination.



Key Stroke shortcuts for zooming and scrolling the view will apply both in the main view and in the sliced view. The effect can be confusing when both views are visible in the same window at the same time. Opening the **Sliced View** tab in a separate window to disconnect the two views.

## Data Sources Tab

- Introduction
- Data sources
- Data sources order
- Adding new data sources
  - Enter user names and passwords
- Data loading settings
- Personal synonyms files
  - Genome synonyms File
  - Chromosome Synonyms File
- Cache Settings and clear the cache

## Introduction

Use the **Data Sources** Preferences tab to

- view, add, or edit data source sites from which IGB can retrieve data.
- load personal genome version and chromosome synonyms files
- change how IGB caches data from remote and local sites

When you choose a species and genome version, IGB downloads meta-data files about available genomes from each site listed in the Data Sources list. If a site contains data for the currently displayed genome, then IGB lists those sites in the **Available Data Sources/Data Sets** section of the **Data Access** tab. And when you load sequence data into IGB by clicking the Load Sequence button, IGB attempts to obtain the sequence data from the data sources in the same order they appear in the Data Sources list.

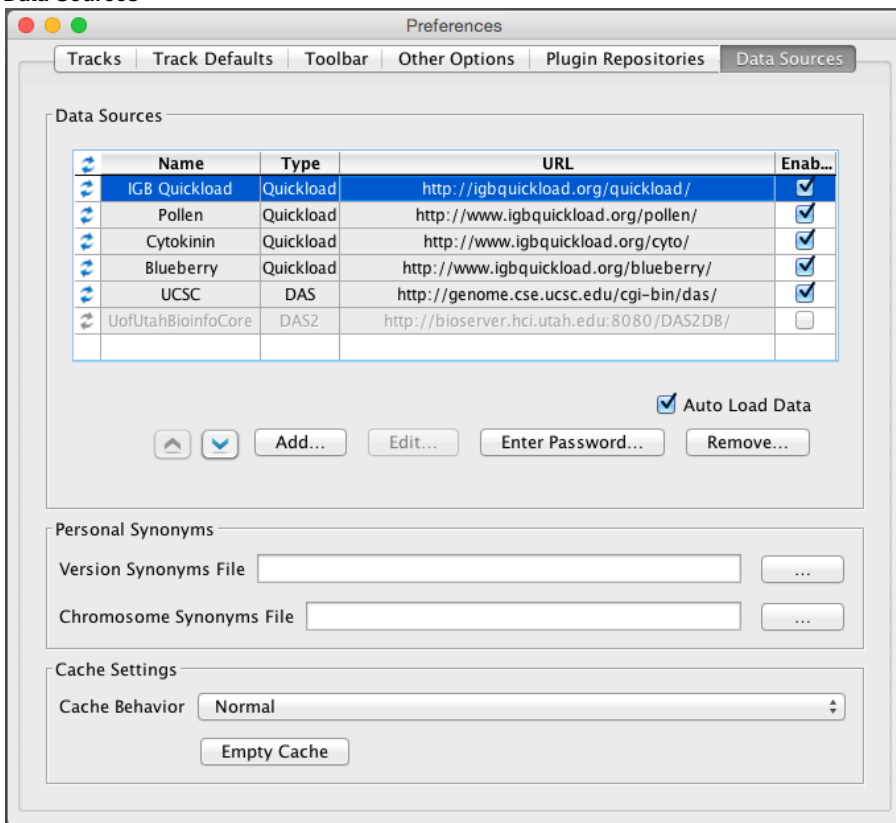
To open the Data Sources Preferences tab

- Select **Data Access > Configure**

or

- Select **File > Preferences > Data Sources**

### Data Sources



Name	Type	URL	Enab...
IGB Quickload	Quickload	http://igbquickload.org/quickload/	<input checked="" type="checkbox"/>
Pollen	Quickload	http://www.igbquickload.org/pollen/	<input checked="" type="checkbox"/>
Cytokinin	Quickload	http://www.igbquickload.org/cyto/	<input checked="" type="checkbox"/>
Blueberry	Quickload	http://www.igbquickload.org/blueberry/	<input checked="" type="checkbox"/>
UCSC	DAS	http://genome.cse.ucsc.edu/cgi-bin/das/	<input checked="" type="checkbox"/>
UofUtahBioinfoCore	DAS2	http://bioserver.hci.utah.edu:8080/DAS2DB/	<input type="checkbox"/>

☒ Auto Load Data

Up Down Add... Edit... Enter Password... Remove...

Personal Synonyms

Version Synonyms File  ...

Chromosome Synonyms File  ...

Cache Settings

Cache Behavior: Normal

Empty Cache

## Data sources

The **Data sources** table lists all active and inactive data sources that have been added to IGB. Active data sources have a check in the enabled column. When you select a genome version, IGB contacts each active server to determine if the server has any data sets for the selected genome. If a server has data sets for the selected genome, the available data sets will appear as options in the **Available Data Sources/Data Sets** section of the **Data Access** tab.

## Data sources order

For some genome versions, more than one data source may provide sequence data. When you click the **Load Sequence** button (top right), IGB loads sequence directly from the first server on the list that supports the currently selected genome. Starting with IGB 6.6, you can change the order of data sources listed in the **Data Sources** preferences tab.

To change sequence order, click **Up** and **Down** buttons (red box).

**Data Sources**

	Name	Type	URL	Enabled
	IGB Quickload	Quickload	http://igbquickload.org/quickload/	<input checked="" type="checkbox"/>
	Pollen	Quickload	http://www.igbquickload.org/pollen/	<input checked="" type="checkbox"/>
	Cytokinin	Quickload	http://www.igbquickload.org/cyto/	<input checked="" type="checkbox"/>
	Blueberry	Quickload	http://www.igbquickload.org/blueberry/	<input checked="" type="checkbox"/>
	UCSC	DAS	http://genome.cse.ucsc.edu/cgi-bin/das/	<input checked="" type="checkbox"/>
	UofUtahBioinfoCore	DAS2	http://bioserver.hci.utah.edu:8080/DAS2DB/	<input checked="" type="checkbox"/>

☒ Auto Load Data

## Adding new data sources

Use **Add...** to add a new data source. A new window will open with boxes and a menu for you to specify a new data sources.

## Enter user names and passwords

Click **Enter Password...** to enter or change the user name and password IGB uses to access private data sources.



Most data sources don't require authentication. If you select a public (open) data source and click **Enter Password...**, IGB will allow you to enter a user name and password but won't use it when contacting the data source.

## Data loading settings

Some servers are configured to automatically load certain foundation data sets, such as the RefSeq gene models, into IGB as soon as the corresponding genome is selected. To block this behavior, uncheck **Auto Load Data**.

## Personal synonyms files

Different data sources often refer to the same genome version or chromosome by different names.

For example, hg17, ensembl1834 and H\_sapiens\_May\_2004 all refer to the same genome assembly and sequence. Similarly, some data providers use **chr1** to refer to chromosome one while others may use **1** to mean the same thing.

IGB provides a system for designating synonyms for sequences and genome versions.

## Genome synonyms File

A version synonyms file matches genome versions and allows you to display data from multiple sources in the same window.

Each time IGB restarts, synonyms will be loaded and merged from all the servers in use as well as from any personal synonyms file that you have specified.

A genome version synonyms file follows a simple tab-delimited format in which synonyms for the same genome appear in the same line, separated by tab characters.

## Chromosome Synonyms File

A chromosome synonyms file is tab-delimited. Each line lists possible names for the same sequence.

Each time IGB starts, synonyms will be loaded and merged from all the servers in use as well as from any personal chromosome synonyms file entered here.

Personal Synonyms

Version Synonyms File  ...

Chromosome Synonyms File  ...

## Cache Settings and clear the cache

IGB caches data that you have loaded over the web in order to speed the loading process the next time you use it. Files loaded from your local file system will not be cached.

Clicking the **Empty Cache** button forces IGB to clear its local cache of downloaded data sets.

**Tip.** If you suspect that the cache files have somehow been corrupted, which may happen if a download was halted unexpectedly, unexpected behavior may result. Click **Empty Cache** if data sets you've opened from a QuickLoad site don't appear to have any data or cause IGB to freeze.

Cache Settings

Cache Behavior

click to empty cache →  Cache Cleared

## Adding and Managing Data Source Servers

- [Data Sources](#)
- [Add...](#)
- [Edit...](#)
- [Refreshing a Data Source](#)
- [Data Source Status Indicators](#)

### Data Sources

- The **Data Sources** panel shows a table with the columns **Name**, **Type**, **URL**, and **Enabled**. To turn a source off or on, click the **Enabled** checkbox. When data sources are disabled, IGB shows them as grayed out.
- The **Auto Load Data** feature (when checked) allows IGB to automatically load data sets with `load_hint="Whole Sequence"` configured in its `annots.xml` file.

Data Sources

	Name	Type	URL	Enab...
	IGB Quickload	Quickload	<a href="http://igbquickload.org/quickload/">http://igbquickload.org/quickload/</a>	<input checked="" type="checkbox"/>
	Pollen	Quickload	<a href="http://www.igbquickload.org/pollen/">http://www.igbquickload.org/pollen/</a>	<input checked="" type="checkbox"/>
	Cytokinin	Quickload	<a href="http://www.igbquickload.org/cyto/">http://www.igbquickload.org/cyto/</a>	<input checked="" type="checkbox"/>
	Blueberry	Quickload	<a href="http://www.igbquickload.org/blueberry/">http://www.igbquickload.org/blueberry/</a>	<input checked="" type="checkbox"/>
	UCSC	DAS	<a href="http://genome.cse.ucsc.edu/cgi-bin/das/">http://genome.cse.ucsc.edu/cgi-bin/das/</a>	<input checked="" type="checkbox"/>
	UofUtahBioinfoCore	DAS2	<a href="http://bioserver.hci.utah.edu:8080/DAS2DB/">http://bioserver.hci.utah.edu:8080/DAS2DB/</a>	<input type="checkbox"/>

☒ Auto Load Data

Server order affects the order in which IGB retrieves genomic sequence data. When you click the Load Sequence button, IGB requests sequence data from servers in the order in which they are listed. Once it finds a server that can deliver the data, it uses that server's data and ignores the rest of the data sources on the list.

Re-order the list using the up and down arrows. Select the server, and click the arrows to move it up and down in the priority list.

**Data Sources**

	Name	Type	URL	Enabled
↻	IGB Quickload	Quickload	http://igbquickload.org/quickload/	<input checked="" type="checkbox"/>
↻	Pollen	Quickload	http://www.igbquickload.org/pollen/	<input checked="" type="checkbox"/>
↻	Cytokinin	Quickload	http://www.igbquickload.org/cyto/	<input checked="" type="checkbox"/>
↻	Blueberry	Quickload	http://www.igbquickload.org/blueberry/	<input checked="" type="checkbox"/>
↻	UCSC	DAS	http://genome.cse.ucsc.edu/cgi-bin/das/	<input checked="" type="checkbox"/>
↻	UofUtahBioinfoCore	DAS2	http://bioserver.hci.utah.edu:8080/DAS2DB/	<input checked="" type="checkbox"/>

☒ Auto Load Data

## Add...

Use the **Add** button to add a new Quickload, DAS, or DAS2 server.

To add a data source:

- Click the **Add...** button to open the **Add Data Source** window.

**Add Data Source**

Name:

Type:

URL:

- Choose a unique name for this new data source. This is the name that appears in the **Data Access** panel.
- Choose the server type (Quickload, DAS, or DAS2).
- The URL name should start with protocol strings "file:/", "http://", or "https://".
  - If this is a local QuickLoad server, you can use the **Choose local folder** button to select a folder on your local file system.
  - A URL for a local directory on a Windows machine might look something like this: "file:/C:/Users/Username/Documents and Settings/QuickLoad/"
- Click the **Submit** button.
- If the server is private and requires authentication, IGB will ask you to enter a user name and password.

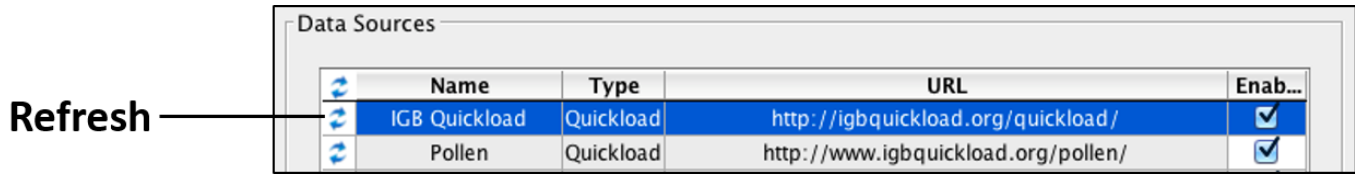
## Edit...

To edit the name or URL associated with a data source, select the source you want to edit and click **Edit...** A new window will open identical to the **Add...** window. Use it to enter new values. (You will not be able to edit any of the default IGB data sources.)

After completing the edit, click the load icon to the left the server name to refresh server settings. This forces IGB to re-read the server's configuration files.

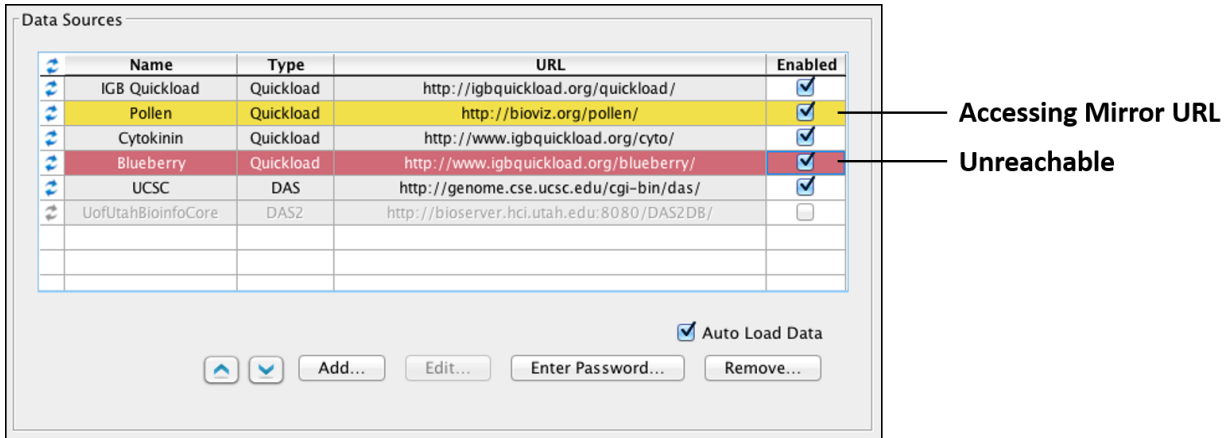
## Refreshing a Data Source

If you are developing a new data source for IGB, you may need to force IGB to re-read its configuration files. Click the Refresh button to the left of a data source to force IGB to re-read the data source server's configuration files.



## Data Source Status Indicators

The status of data sources are indicated by the color of the data source's row in the Data Sources table.



- Rows in yellow indicate the data source is accessing its mirror URL. The mirror URL is accessed when the data source's primary URL is not reachable. IGB is able to load files from the data source in this case.
- Rows in red indicate the data source is unreachable. This means that the data source's primary URL and its mirror URL are not reachable. IGB is not able to load files from the data source in this case.
- Rows with a gray background are accessing the data source's primary URL. IGB is able to load files from the data source in this case.

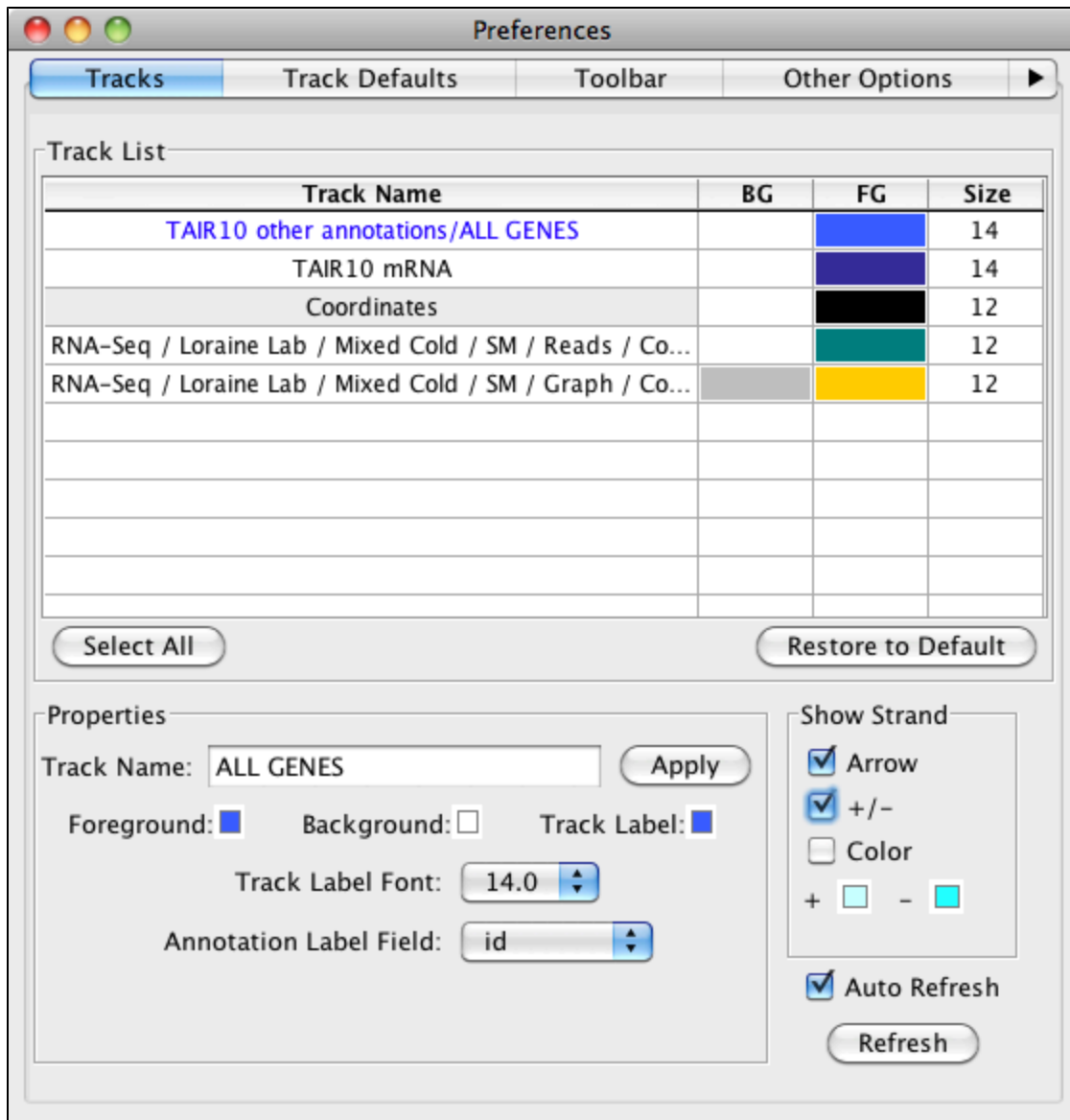
## Tracks

- [Introduction](#)
- [Track List](#)
- [Properties](#)
- [Show Strand](#)
- [Restore to Default](#)

## Introduction

Use **Preferences > Tracks** to configure the appearance of loaded tracks. Changes take effect immediately if **Auto Refresh** is selected or when you click the **Refresh** button. You can also change many aspects of track appearance using the **Annotation** tab or by right-clicking a track label.

For more about changing the appearance of tracks, read [Customizing track appearance](#)



## Track List

The **Tracks** tab has a list showing all loaded tracks (all tracks listed in the **Data Access > Data Management** table). You can select a track by clicking on it in this list. This table allows you to select any number of tracks, so that you can apply changes to one, all or a subset.

To select multiple tracks

- Click a track while pressing CNTRL (on Mac, press CMD)

## Properties

**Track Name** can be changed by typing in the entry box. This can also be changed from the **Data Access > Data Management Table**.

To change the **Foreground** color (FG), the **Background** color (BG) or the **Track Label** color, the color square and choose a color. To change the **Track Label Font** size, use the font size dropdown box. Color and font size can also be changed from the **Annotation** or **Graph** tabs; color can be changed from the **Data Access** panel.

The **Annotation Label Field** box allows you to have a specific label displayed over your annotations. The default option is gene name, however you can use any property available within the file including, name, ID, gene descriptions or score. The **Annotation Label Field** box dynamically adjusts to offer you all options for a selected track. You can turn off the labels by setting **Label Field** to \*none\*.

## Show Strand



IGB allows you to control the strand display. Checking **Arrow** will show directional arrow heads within introns, and will put a pointed end onto the last exon of each gene model. **+/-** allows the user to combine both plus and minus strands into a single track, as opposed to the default which draws the + strand above the coordinate axis, and the - strand below the axis. Checking **Color** will change the color of all annotations from the **Foreground** color to the + color or the - color. The colors can be used as default or changed to your preference by clicking on the color box(es). You can all options at the same time, and these options can be used whether or not the strands are combined.

## Restore to Default

Sometimes after changing the appearance of a track in a variety of ways, colors and font sizes, etc., you want to return it to the original settings. **Restore to Default** will restore the track to the **Preferences > Track Default** settings.



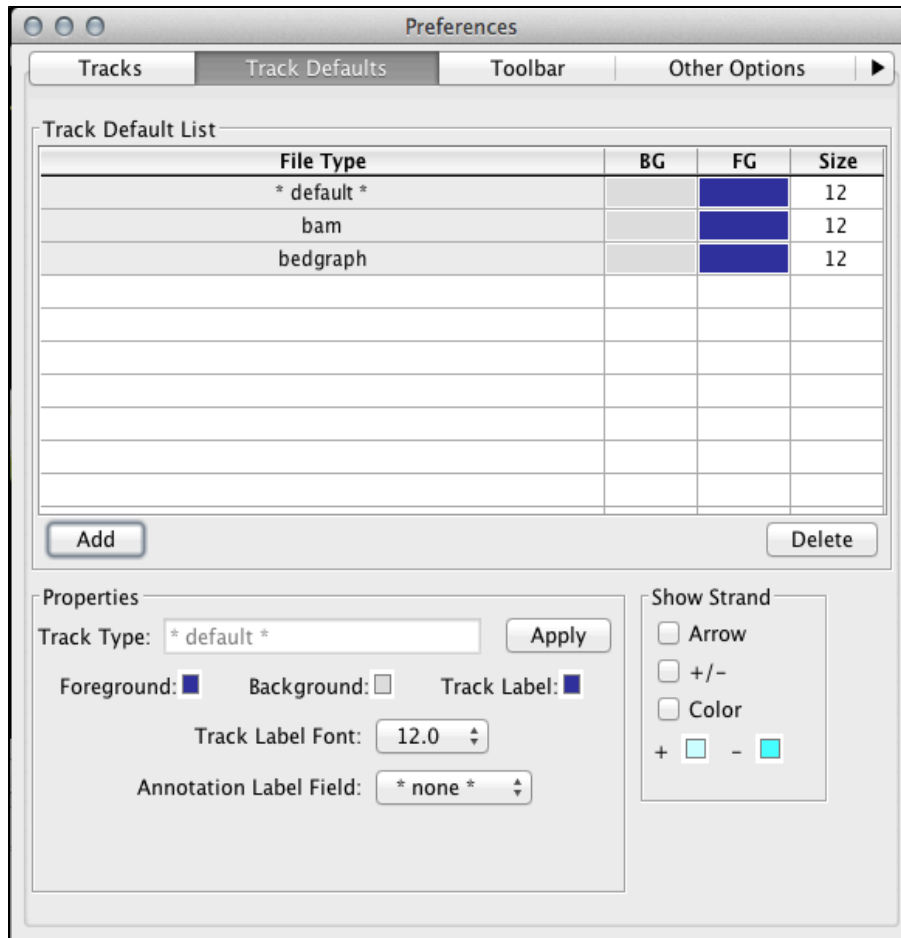
If the default has been changed or updated *after* loading the selected track, using the **Restore to Default** button will change all the setting to be in line with the *current Track Default*.

## Track Defaults

- [Introduction](#)
- [Setting defaults by file type](#)

### Introduction

Use **Track Defaults** to set defaults for how data from different file types will appear once you open them.



### Setting defaults by file type

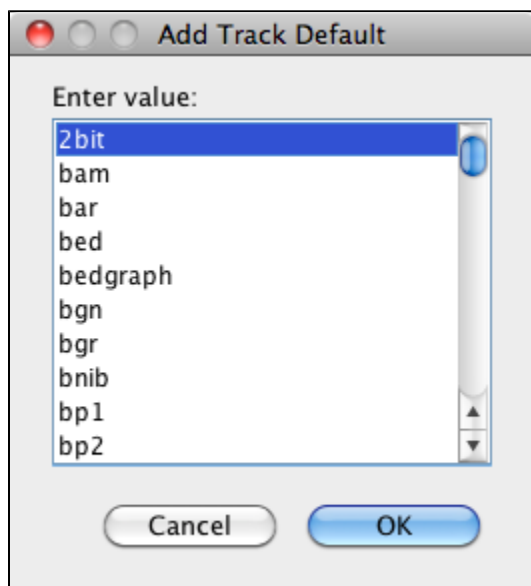
To pre-configure how data from a file type will appear once opened and loaded into IGB


1. Select a file type from the table or use the **Add** button to add new file type to the table
2. After selecting, make the style changes such as color, label, max expand and strand options.

This will not affect any files already opened and loaded into IGB.

Also, if you change settings for a track after loading it, your settings will over-ride these defaults.

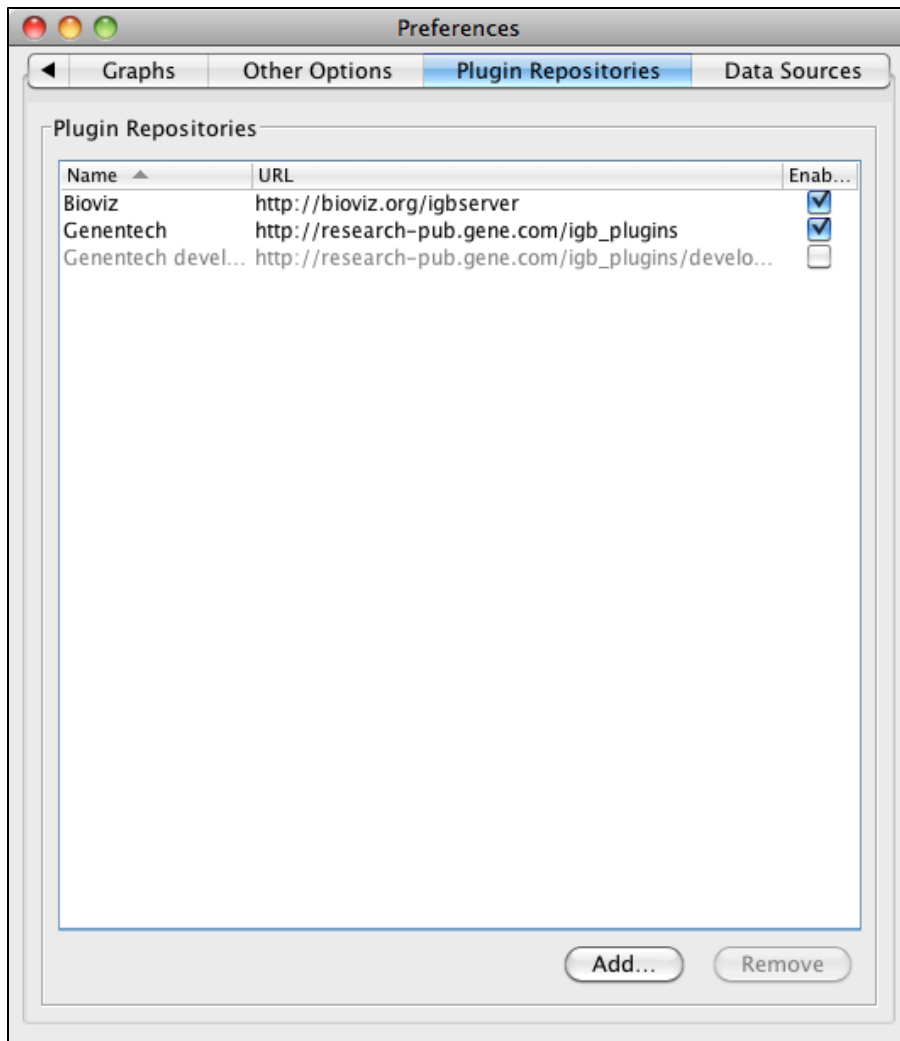
The **Add** button lists file formats that IGB supports.



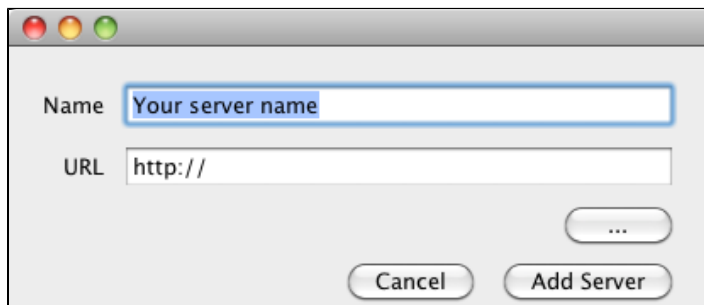
 Files in QuickLoad directories may already have styling information set by an **Annots.xml** file. If this is the case, changes made in **Track Defaults** will not affect the files.

## Plugin Repositories Tab

With the inclusion of Plugins in IGB, we provide an interface for including new Plugin repository sites. IGB supported repositories are shown.



Repositories can be accessed by checking the **Enabled** box; they can be added by clicking the **Add** button, and the following window opens:



Provide the information, and be sure the new repository has **Enabled** checked.

Repositories can be made inaccessible by unchecking the Enabled box; they can be completely removed by selecting the repository and clicking the **Remove** button.

## Sharing Preferences

### Importing and Exporting Preferences

You can set each preference manually, or you can export and import preferences to an XML file. This can be done to save them or share them among a group of people who are working on similar or related projects. This is also useful if you want to move IGB from one computer to another, or if you would like to share preferences with other users.



In earlier versions of IGB, preferences were set by the user by editing a file called "igb\_prefs.xml". That file is deprecated, though it is still used by some users for certain advanced functions.

## Exporting Preference Settings

You can export preference settings for use in different instances of IGB or to share them with other IGB users.

- Choose **File > Preferences**.
- In the Preferences dialog box, choose **Preferences** menu > **Export Preferences**.
- Choose the destination and filename to which you want to save your preference settings.
- Click **Save**.

## Importing Preference Settings

You can import preference settings that you or another IGB user has exported. You can freely exchange preferences between Windows, Linux, and Macintosh systems. Preferences in the file that you import are merged with your existing preferences.



Be sure you trust the provider of the file.

- Choose **File** menu > **Preferences**.
- In the Preferences dialog box, choose **Preferences** menu > **Import Preferences**.
- Choose the saved preferences file and click **Open**.

## Troubleshooting

See the links below for common problems and their solutions.

IGB is available for the Mac, Windows, and Linux operating systems with the following requirements:

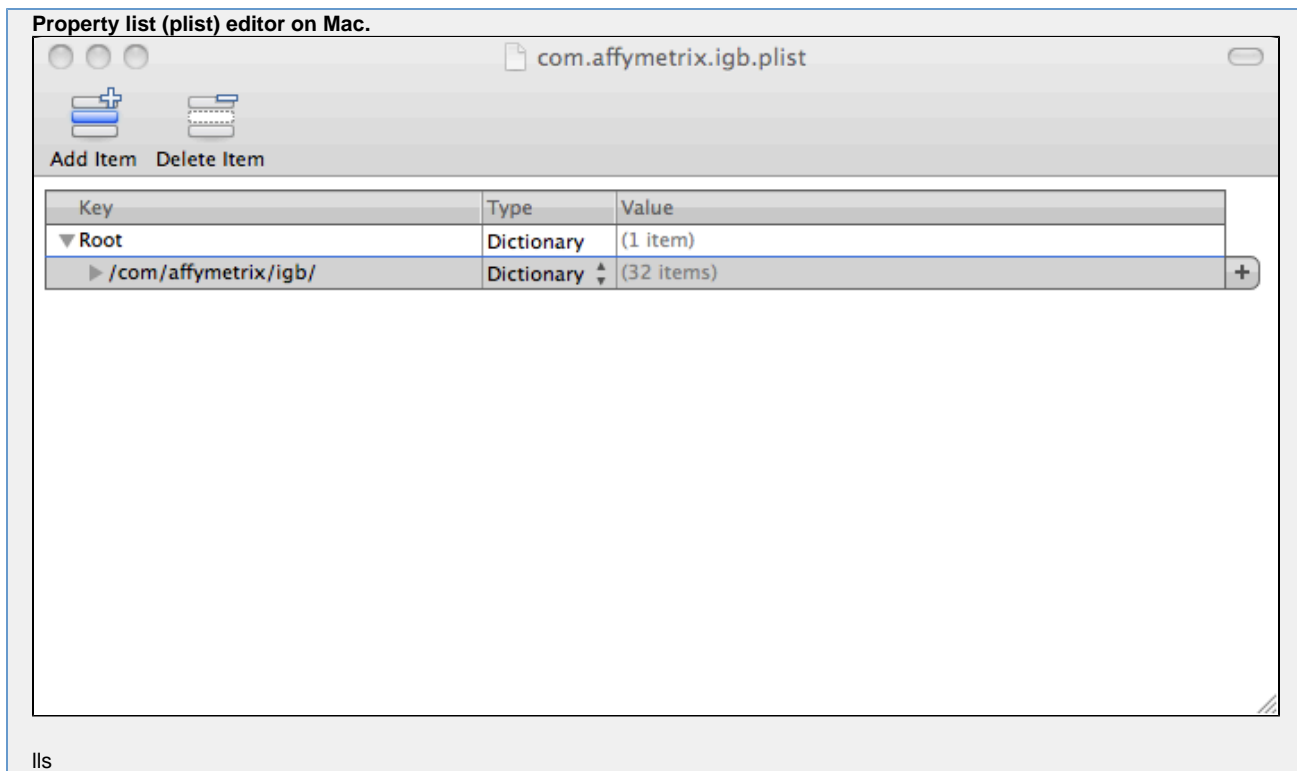
- Mac requires OS X Version 10.7.3 or later.
- Windows requires Vista or Later (32-bit and 64-bit).
- Linux requires Java 1.8 installed

## Unable to launch the application

If you typically download and run IGB using the Bioviz Java Web Start launch pages (the [Download IGB page](#)), you may be experiencing problems with old-of-date files or preference settings. Alternatively, your organization may be blocking certain types of internet-based applications (like IGB) from downloading and running on your desktop.

If you are unable to download and launch IGB from the Bioviz site, try the following:

- **Clear your Web browser cache** and re-open your browser. Sometimes your browser will attempt to run IGB using an out-of-date, cached version, which affects various elements of IGB including memory requirements. For example, IGB uses a deployment descriptor file that can get out of date. Deleting your browser cache gets rid of out-of-date files, allowing you to get the latest versions, which then will allow you to run IGB.  
For instructions see [How do I clear my web browser's cache](#) from Indiana University IT Knowledge Base.
- **Clear your Java Cache.** The Java cache may contain outdated or corrupted copies of the deployment descriptor or libraries. This affects application start-up and arguments passed to the application. See <https://elearn.usu.edu/java/clear-java-cache.html> for instructions on how to clear the Java cache on Windows or Mac OS X. For Linux, simply delete the directory `~/.java`
- **Delete the IGB Cache folder.** This will clear all cached data, as well as remove any corrupted data that could be blocking IGB from running. These directories are hidden; you will need to enable viewing of hidden files, or delete them from the command line. **Warning:** Do not delete **ENTIRE** .igb directory because if you do, you will lose your bookmarks along with cached data. **Only** remove the IGB cache folder.
  - MacOSX: Delete folder `UserHome/.igb/cache`
  - Windows: Delete folder `UserHome/Application Data/IGB/cache`
  - Linux: Delete folder `~/.igb/cache`
- **Clear the plist (For Mac users only).** IGB stores the users' Preferences information in different locations on different platforms. On Windows, Preferences are stored in the windows registry. On Linux, preferences are stored in ".xml" files. In both of these cases, we recommend against editing these files directly. In Mac OS, Preferences are stored in a file called a "plist," which you can edit using a Mac utility called the Property List Editor. To edit your plist using the Property List Editor, open a new Finder window and open `Library > Preferences` in your home folder. Double-click the file named `com.affymetrix.igb.plist`, which should launch the Property List Editor. Select the item called `/com/affymetrix/igb` and click "Delete Item." (See image below.) This will re-set all your preferences to their defaults. **Warning:** This action will also delete all data sources and passwords you've added; you will need to add them back manually once IGB is able to run.



- **Download IGB executable.** If none of the above solutions help, it's possible that your organization is preventing Java applications (like IGB) that are downloaded from the internet via Java Web Start from running on yours and other users' computers. If this is the case, then you should download IGB directly from the Sourceforge Web site as a "zip" file, unpack it, and then use the "run\_igb" scripts we've developed for users who can't run IGB from Java Web Start. To do this:
  - Visit <http://sourceforge.net/projects/genoviz/files/>.
  - Choose the latest release of IGB (e.g., IGB 6.4.1 as of this writing.)
  - Download the file named **igb.zip**
  - Once it downloads, unpack it. (On most systems, all you have to do is double-click the file, and on some systems, it will unpack automatically.)
  - Once unpacked, you should see a folder called "igb"
  - Open the folder and double-click run\_igb.sh (on Mac or Linux) or run\_igb.bat (Windows.)
  - IGB should then launch and run.

## IGB runs, but acts strangely.

If you are able to run IGB but are observing strange behaviors, try the following:

- **Clear your browser cache** and re-open your browser. Sometimes your browser will attempt to run IGB using an out-of-date, cached version, which affects various elements of IGB including memory requirements. For example, IGB uses a deployment descriptor file that can get out of date. Deleting your browser cache gets rid of out-of-date files, allowing you to get the latest versions, which then will allow you to run IGB. For instructions see [How do I clear my web browser's cache](#) from Indiana University IT Knowledge Base.
- **Clear your Java Cache.** The Java cache may contain outdated or corrupted copies of the deployment descriptor or libraries. This affects application start-up and arguments passed to the application. See <https://elearn.usu.edu/java/clear-java-cache.html> for instructions on how to clear the Java cache on Windows or Mac OS X. On Linux, simply delete the directory ~/.java
- **Clear your IGB cache.** This will clear any cached data, but will also remove any corrupted data that could be causing crashes or odd behavior while the application is running. Go to **File > Preferences > Data Sources** and click the 'Empty Cache' button at the bottom of the screen.
- **Reset Preferences to default.** You should use this as a last resort, because **this will remove all saved preferences, including servers and passwords you've entered.** **\*To do this, go to \*File > Preferences > Other Options** and click the 'Reset all preferences to defaults' button at the bottom of the screen.

## I can't see my data!

If you are having trouble viewing your data, see the following:

- [Ensure that your file is in a format we support](#)
- [Click Load Data button](#)
- [.bam files need a .bai file](#)
- [I opened my sequence file \(".fasta" or ".bnib" or "twobit"\) but all I see is a gray bar and dashes.](#)

## Ensure that your file is in a format we support

IGB can open and read nearly all of the common data and graph files used with genomics data. For more information about supported file types, please see [File Formats](#). We also take requests - if you would like us to support a new type of file format, please let us know.

## Click Load Data button

IGB allows you to open files from your desktop, local servers, URLs or other data sources.

However, IGB will not instantly start loading the data into the display.

When you open files, IGB adds them to the list of opened files in the **Data Management** table in the **Data Access** tab. Note that the Data Management table lists a Load Mode for each opened file or data set. By default, IGB adds most files and data sets using the **Region in View** load mode. Data sets and files with load mode **Region in View** will load into the main data display when you click the **Load Data** button. To display data from these files, zoom in to a region you want to see, and click **Load Data**. Whatever data in the file that overlaps with the current region in view will appear in the display.

**Tip:** If your data file is not large, change **Load Mode** to **Genome**. If you do that, IGB will then read and display all the data in the file. Examples of data sets that work best when loaded in the **Genome** mode include: gene models for the entire human genome or significant regions identified in a ChIP-Seq experiment.

For more about loading files, please see [Loading data](#).

## .bam files need a .bai file

IGB can display alignments data from .bam files. However, it requires a .bam index file (called a .bai) to be present in the same folder as the .bam file.

## I opened my sequence file (".fasta" or ".bnib" or "twobit") but all I see is a gray bar and dashes.

To save memory, IGB doesn't load sequence data until you click the **Load Sequence** button. To view the bases, click the **Load Sequence** button and then zoom in.

## Unable to detect proxy setting

### *If you have downloaded IGB as igb.zip*

Open 'run\_igb.sh' in a text editor. Look up for following lines.

```
# Use 32MiB as initial heap size
echo $VMARGS[*] | grep -q '\-Xms\|\-ms' || VMARGS[${#VMARGS[*]}]="-Xms32m"
```

Add following lines after it.

```
# Use proxy server
echo $VMARGS[*] | grep -q '\-DproxySet' || VMARGS[${#VMARGS[*]}]="-DproxySet=true"

# Replace the IP address of the server to the name/IP of your proxy-server
echo $VMARGS[*] | grep -q '\-Dhttp.proxyHost' ||
VMARGS[${#VMARGS[*]}]="-Dhttp.proxyHost=wwwproxyhost"

# Replace the port to the port that is used at your proxy-server
echo $VMARGS[*] | grep -q '\-Dhttp.proxyPort' ||
VMARGS[${#VMARGS[*]}]="-Dhttp.proxyPort=wwwproxyport"
```

### *If you are running IGB using a jnlp file.*

Open jnlp file in a text editor. Look up for following lines.

```
<argument>-prefs</argument>
<argument>http://www.bioviz.org/igb/releases/igb-6.4/igb_prefs.xml</argument>
```

Add following lines after it.

```
<argument>-DproxySet</argument>
<argument>true</argument>

<argument>-Dhttp.proxyHost</argument>
<argument>wwwproxyhost</argument>

<argument>-Dhttp.proxyPort</argument>
<argument>wwwproxyport</argument>
```

Replace wwwproxyhost and wwwproxyport with the IP/name of your working www proxy and proxy port respectively.

e.g wwwproxyhost can be replaced by 192.168.0.1 and wwwproxyport can be replaced by 3128

(Contributed by [Youri V. Kravatsky](#))

Note: This solution is not yet verified by IGB team.

## My data isn't there

When you open or drag a file into IGB, the file will appear in the **Data Management Table** on the **Data Access** panel. Be sure that your file was properly opened, and that it appears in the list. IGB will also open an empty track for this file with a grayer background, waiting for you to load the data.

There can be several reasons that your data has not appeared:

- Data not yet loaded

IGB does not immediately load your data when you open a file or a URL. Most of today's data sets are too large to visualize all at once. After you zoom in to a more refined area, click **Load Data**. Your data will be loaded and appear in the **Data Display Area**.

- The file names for genome version or chromosomes do not match

Often times, different files will use different names for either the genome version (e.g. "human genome build 35" is also named "hg17", "ensembl1834", or "H\_sapiens\_May\_2004") or chromosome name (e.g. "chromosome 1" is also named "1", "chr1", "chr 1", "chr1" or "chrom1"). If the files have mismatched names, IGB is not able to align them. This can be easily corrected by the instructions on the [Personal Synonyms](#) page to adjust the names.

- You are in an area where there is no data

Sometimes, while zoomed into low data areas, it will seem like your data is not there. However, if the background color in your data track is the selected **BG** color (and not a darker/grayer color) IGB has loaded all information in the region. Zoom out or scroll to a new region and click **Load Data**.

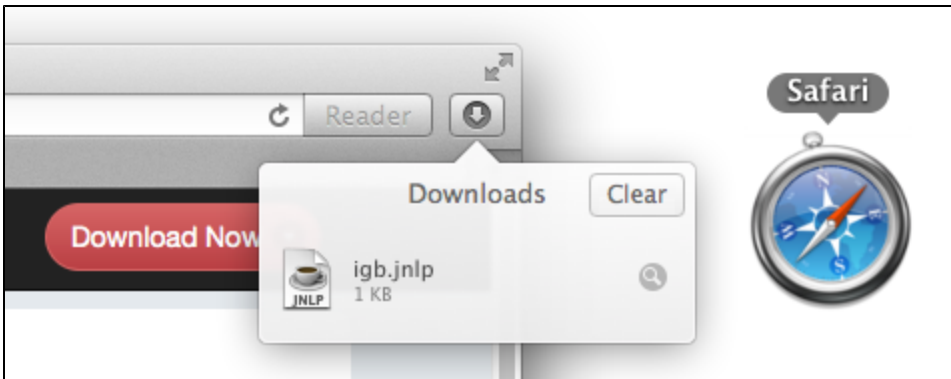
## The IGB menus have disappeared

There is a Java bug (in version 1.7, update 13) in which opening a pop-up window from within IGB causes the IGB menus to disappear. This issue has been seen with the survey, Preferences and even sequence viewer. Until the next Java update, this will likely remain an issue.

The simple fix is to minimize and reopen the IGB window. All of the menus will be restored.

## Where is my IGB download? (Safari)

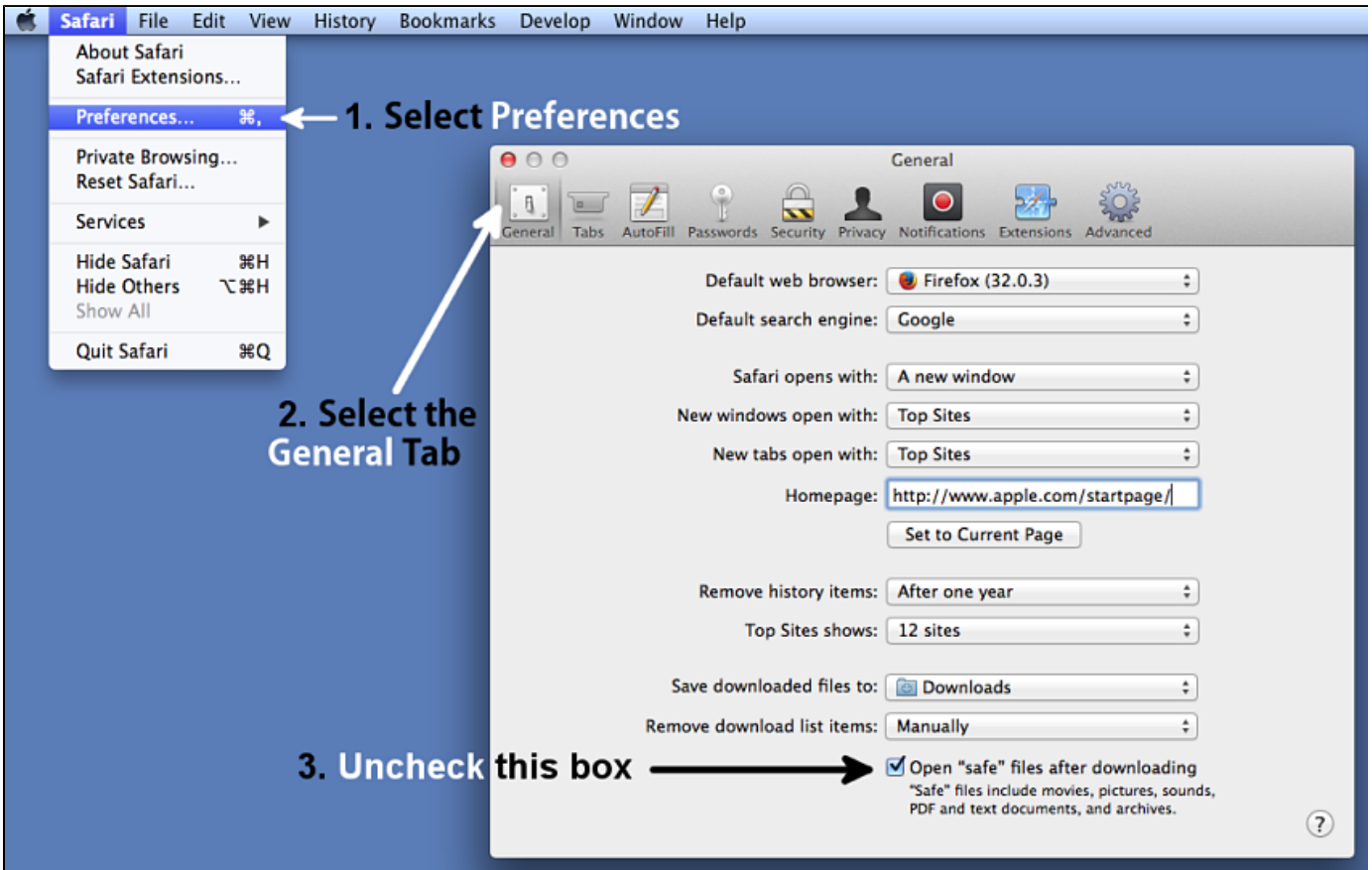
If you are using Safari as your web browser, you may be wondering where your IGB JNLP file is located after downloading IGB. This is because Safari downloads JNLP files into an uncommon directory (/private/var/.../T).



## Make Safari download JNLP files into the Downloads directory

1. Open a Safari window and select **Preferences** from the **Safari** menu located at the top of your screen.
2. Select the **General** Tab.
3. Uncheck the box next to **Open "safe" files after downloading**.
4. Quit and then reopen Safari.

JNLP files will now download into the **Downloads** folder.



## Installing and Updating Java

Running IGB on a Linux machine requires Java 1.8. Your system already may have Java installed. If not, follow the link below for instructions on how to install or update Java on your computer.





[Click here for instructions](#) on installing Java (version 8) on your computer.

## Tutorials

### Welcome

To help new and veteran users get acquainted with IGB features, we've developed the following tutorials.

If you have ideas for improvement or new tutorials, please let us know.

Email your suggestions to:

Ann Loraine  
aloraine – at – uncc.edu

Thank you and best of luck in your research,

The IGB Team

### Focus on a Feature Live On-line Tutorials

To help new and experienced IGB users get more out of using IGB to analyze data, we are hosting a series of live, on-line tutorials called "Focus on a Feature."

Announcements for new Focus on a Feature tutorials are made on [Biostars](#) and [Twitter](#).

To view recordings from previous tutorials, view them under [Video Tutorials](#) or visit the [IGB Channel](#) at YouTube.

### Assessing splicing - Cold-induced alternative splicing of Arabidopsis SR45

- [Introduction](#)
- [Setting up](#)
  - [Open read alignment files](#)
  - [Customize the tracks](#)
- [Analyze Splicing](#)
  - [Background on SR45, a plant-specific, putative splicing regulator](#)
    - [A note on nomenclature - AGI locus codes](#)
    - [SR45 mutant phenotype](#)
  - [View SR45 in IGB](#)
    - [Use the search tab to search for AT1G16610 \(SR45 AGI locus code\)](#)
    - [View genomic sequence and protein translations for SR45 splice variants](#)
    - [Load RNA-Seq reads from cold stress experiment](#)
    - [Assess relative abundance of SR45 splicing isoforms](#)
- [Finishing up - export an image showing what you found](#)

### Introduction

The *Arabidopsis* and other plant genomes include many genes predicted to encode multiple splice variants.

However, it isn't always clear which of these splicing variants are the most highly expressed in specific cell types, organ structures, or conditions.

One way to assess which isoform is most abundant is to visualize aligned reads from an RNA-Seq experiment.

If the read lengths are sufficiently long (75 or more bases) and the gene of interest has sufficient coverage, you can usually tell which variant is the most abundant.

And if you have data from experiments that investigate different treatments, you can sometimes also find out if the ratio of isoforms has changed as a result of the treatment. This is called condition-dependent alternative splicing, which just means: a treatment has changed the pattern of splicing at a locus.

In this tutorial, you'll learn how to use IGB to investigate splicing using RNA-Seq data. Using data from a cold-stress experiment in *Arabidopsis thaliana*, you'll learn how to:

- Identify the dominant isoform for a gene
- Observe condition-dependent alternative splicing

This tutorial will feature SR45/AT1G16610 from *Arabidopsis thaliana*, a putative splicing regulator that is itself alternatively spliced.

## Setting up

Follow these instructions to configure IGB to get started looking at a cold stress RNA-Seq data from *Arabidopsis thaliana*:

- Start IGB
  - Download and launch IGB using Java Web Start from <http://www.bioviz.org/igb/download.shtml>
  - Or download igb.zip from <http://sourceforge.net/projects/genoviz/files/>.
    - To launch IGB from igb.zip, double-click igb.zip to unpack it, open folder igb, and double-click run\_igb.bat (Windows) or run\_igb.command (Mac)
- Choose species **Arabidopsis thaliana**, Genome Version **A\_thaliana\_Jun\_2009**

**Note** that the gene models should automatically load.

**Tip:** Color the gene track black (foreground) and white (background) to make attractive images for slides and printing. Just click the color swatches (squares) in the Data Management table to change colors.

**IGB 6.7 showing protein-coding (TAIR10 mRNA) chr1 gene models, with gene models changed to black and background changed to white.**

The screenshot shows the IGB (Genome Browser) interface. The main window displays the TAIR10 mRNA gene models for chromosome 1. The coordinates range from 0 to 30,427,671. The Data Management Table at the bottom shows the loaded data sets and their configurations.

	FG	BG	2 Track	Load Mode	Data Set/File N...	Track Name (Doub...	
			<input type="checkbox"/>	Region	RNA-Seq / Lorai...	Cold, alignments	
			<input type="checkbox"/>	Region	RNA-Seq / Lorai...	Control, alignments	
			<input checked="" type="checkbox"/>	Genome	TAIR10 mRNA	TAIR10	

Buttons: Load All Sequence, Load Sequence In View, Load Data

## Open read alignment files

Now start opening read alignment files from chilled and unchilled samples.

- Under the **Data Access** tabbed panel, open
  - Data Source **IGB Quickload > RNaseq > Loraine Lab > Mixed Cold > SM > Reads**
- Click the checkboxes next to data sets named:
  - Control, alignments
  - Cold, alignments

**Please Note:** When you click the checkboxes, IGB will access the files to make sure they are readable, but it won't immediately start loading the data. The reason for this is that many genomic data sets, especially next-generation sequencing data, are very large. To save memory and time, IGB gives you a chance to zoom and pan to a region you are most interested in viewing before loading the data into the viewer.

If it was able to access the files, IGB will add the data sets to the **Data Management Table** under the **Data Access Tab** and will also create new (empty) tracks for the data sets. Once the data are loaded, it will appear in these new tracks.

## Customize the tracks

If you don't like the default color scheme, take a moment to customize your track settings colors.



Foreground and background color schemes are set by the IGBQuickLoad server, but you can change them in IGB if you like. IGB will remember your color choices between sessions.

Click the color squares in the **Data Management Table** (FG is foreground, BG is background) or choose **File > Preferences > Tracks** to open a new Window that let's you customize different aspects of the display at several at once. Click each row in the table and give each row a white background. Choose colors for individual rows by clicking on them and then selecting options.

**Tip:** Customize multiple tracks at once by choosing **File > Preferences > Tracks**

**Tip:** Simplify the track name and increase track name label size to make the Track Labels easier to read.

## Analyze Splicing

The remainder of the tutorial will focus on one gene - SR45, also known as AT1G16610.

### Background on SR45, a plant-specific, putative splicing regulator

SR45 encodes a putative splicing regulator resembling members of the so-called SR protein family, a group of proteins that contain one or more RNA-recognition motifs (RRM) and regions enriched with serine-arginine (SR) or arginine-serine (RS) amino acid repeats.

The function of the RS/SR repeats is not well-understand, but one idea is that they help stabilize interactions between pre-mRNA transcripts and proteins during splicing or spliceosome assembly.

- Read about SR45 at [the Arabidopsis Information Resource](#)
- Read about SR45 in [Entrez Gene at the National Center for Biotechnology Information \(USA\)](#)

### A note on nomenclature - AGI locus codes

Like most Arabidopsis genes, SR45 has a second name (AT1G16610) called an AGI locus code that was first assigned by the Arabidopsis Genome Initiative, which sequenced and annotated early versions of the Arabidopsis genome. Use the AGI code to locate the gene in Integrated Genome Browser. (More on this later.)

### SR45 mutant phenotype

One mutant line (called sr45-1) carrying a disruption in the SR45 gene has been characterized. These plants exhibit multiple defects, including misshapen petals and reduced root growth. SR45a expresses at least three known splicing variants.

An article titled **Two alternatively spliced isoforms of the Arabidopsis SR45 protein have distinct roles during normal plant development** investigated how Variants 1 and 2 differ with respect to SR45 function.

The authors found that transforming the sr45-1 plants with splice variant 1 corrected the petal phenotype, while transforming with variant 2 corrected the root phenotype. (Variant 3 was not tested.)

The difference between the variants was not large, but clearly the two different splice forms affect development in different ways.

Another article titled **The plant-specific SR45 protein negatively regulates glucose and ABA signaling during early seedling development in Arabidopsis** reported that SR45 plays a role in sugar sensing and ABA pathways during early seedling development. (ABA is a plant hormone that helps keep seeds from sprouting prematurely.) This article represents one of the first discoveries of a link between a splicing regulator and a hormone response pathway in plants.

- Read [Two alternatively spliced isoforms of the Arabidopsis SR45 protein have distinct roles during normal plant development](#)
- Read [The plant-specific SR45 protein negatively regulates glucose and ABA signaling during early seedling development in Arabidopsis](#)

## View SR45 in IGB

### Use the search tab to search for AT1G16610 (SR45 AGI locus code)

To view the SR45 gene in IGB, click the **Search** tab and enter the AGI locus code for SR45: AT1G16610.

Three rows of results will appear. Double-click one to navigate to the SR45 region.

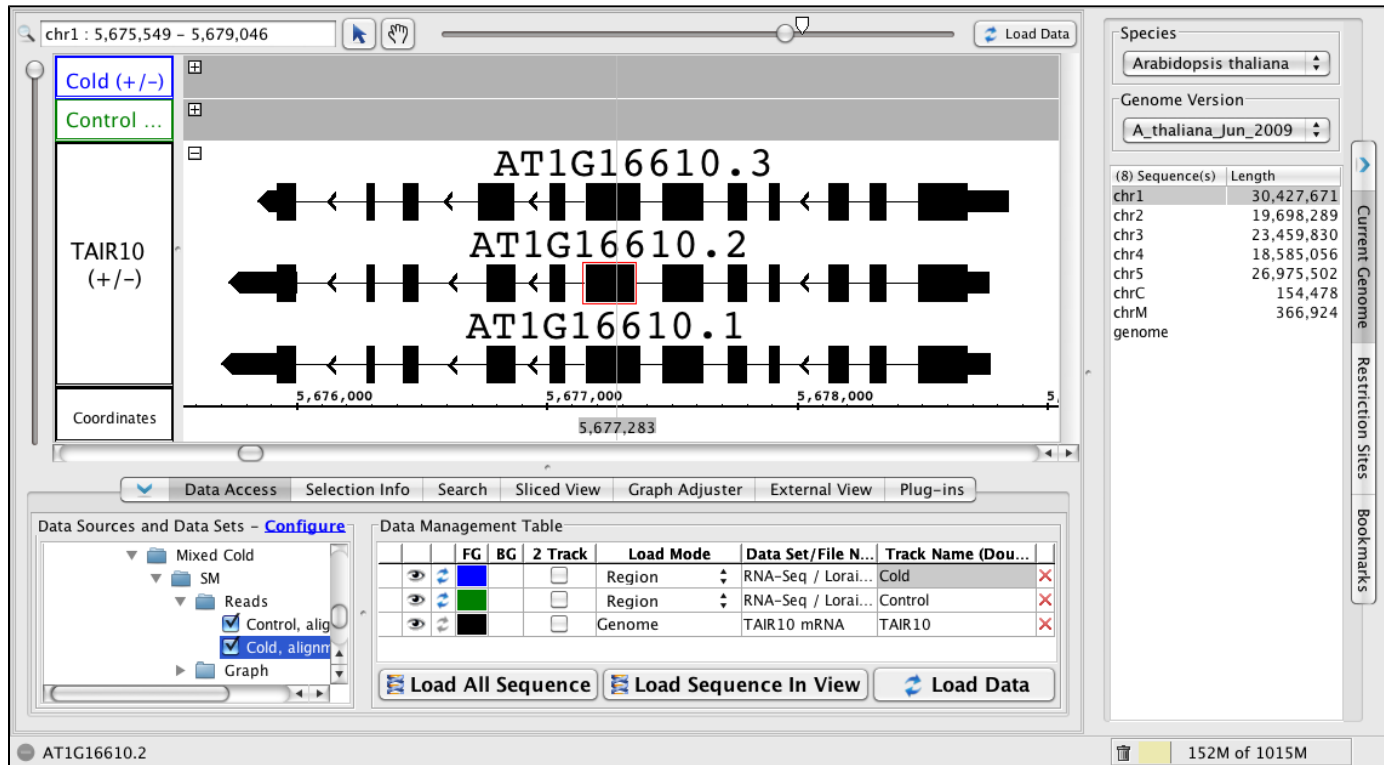
You'll see three variants. Note that transcription proceeds from right to left - the promoter is on the right in this case because the gene resides on the minus strand of the chromosome.

Let's customize the display even further - change the "max stack depth" and show all gene models in one track:

- Right-click (or control-click on Mac) the TAIR10 track label
  - Choose **Change > Adjust Max Stack Depth** and enter 3 to remove extra white space above the gene models. (Note that if you zoom out, gene models from genes with more than three gene models may be drawn on top each other.)
  - Choose **Change > Show 1 track (+/-)** to combine plus and minus strand gene models into the same track

When you're done, you'll see something like the image below.

Note that one exon is selected in variant AT1G16610.2. This is the differentially spliced exon that differs between variants 1 and 2. In variant 2, the exon has a different five prime end. It encodes a slightly shorter protein than Variant 1.



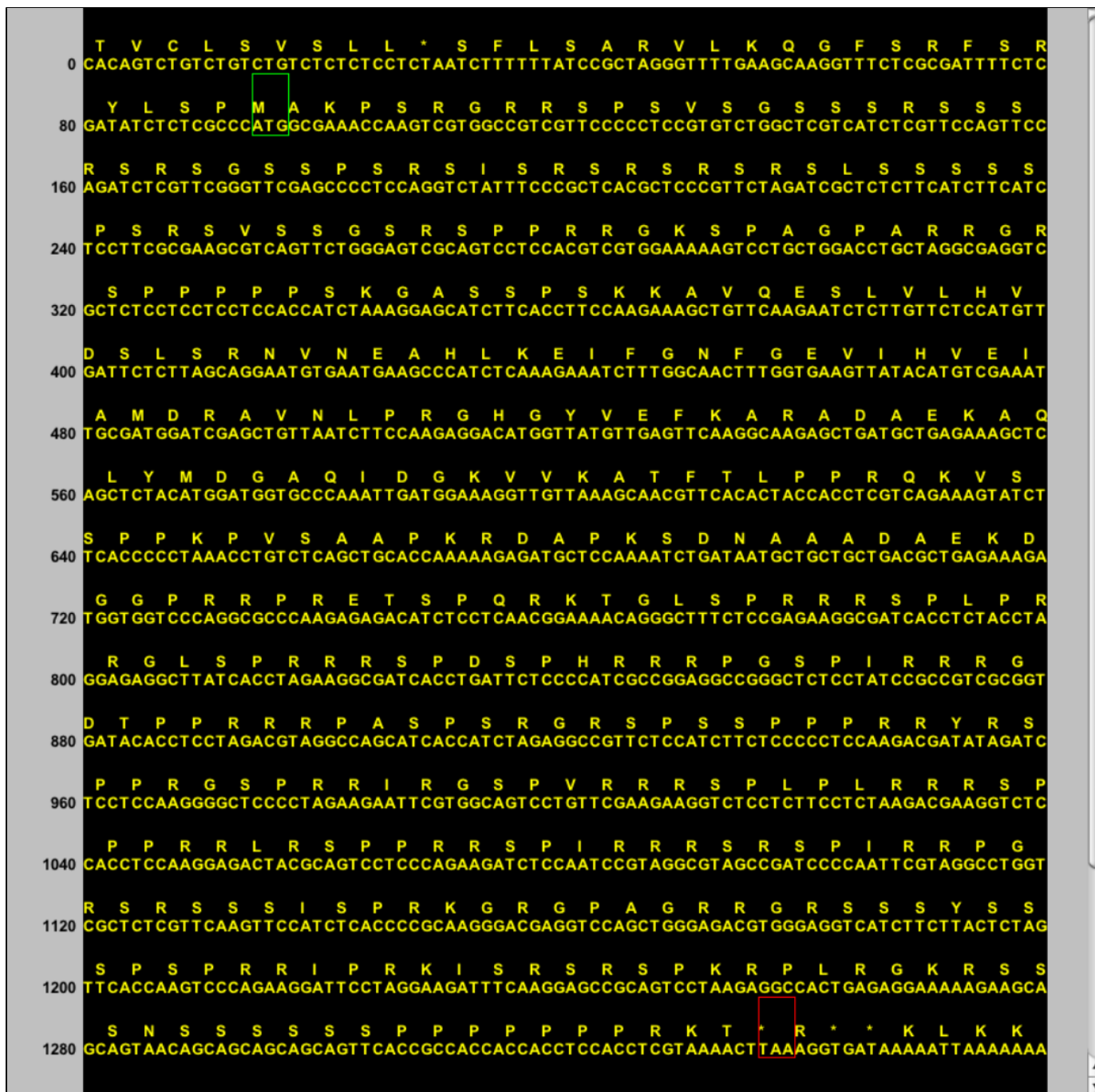
### View genomic sequence and protein translations for SR45 splice variants

- Right-click (or control-click on Mac) variant 1 - AT1G1660.1.
- Choose **View Genomic Sequence in Sequence Viewer**
  - When the viewer opens, click **Show cDNA**
  - Click **Show > +2 Translation** to view the amino acids encoded by variant 1.

Note how the SR45 protein contains C-terminal and N-terminal RS/SR motifs. This is unusual for an SR protein and some researchers believe that this unusual structure makes SR45 a plant-specific SR protein, and others argue that it should not properly be considered an SR protein.

- Read this review article on SR proteins: [Implementing a rational and consistent nomenclature for serine/arginine-rich protein splicing factors \(SR proteins\) in plants.](#)

### Genomic sequence viewer with amino acid translation of splice variant AT1G16601.1:



## Load RNA-Seq reads from cold stress experiment

If you have already accessed the read alignment files, you should see two empty tracks in the IGB main view (as in the images above.) To load the reads mapping to the SR45 region, click the **Load Data** button.

- Click Load Data to load reads for the region
- Right-click (or control-click) the Control and Cold Treatment track labels
  - Choose **Change > Adjust Max Stack Depth** and enter "0" to ensure that all reads will be shown
  - Use the vertical slider to stretch the display vertically to stretch the reads and gene models

**About Max Stack Depth:** The **Max Stack Depth** setting takes some getting used to, but as you learn to use IGB, you'll find it gives you extra power to configure the IGB display and fine-tune images you create for publication, slides, and reports.

The **Max Stack Depth** setting allows you to control the number items (reads, gene models, or other annotations) that can be stacked on top of each other in a given track.

This is a useful setting when you're working with BAM files (read alignments) because for some genes, there will be tens, hundreds, or even thousands of aligned reads that overlap the same region of the genome.

If you set the **Max Stack Depth** to a number other than 0, then IGB will never build "stacks" of overlapping reads taller than that number. Any extra reads that would make the stack too tall (and exceed the Max Stack Depth setting) will be drawn on top of each other in a special overlap row at the very top of the track (for plus strand reads) or at the bottom of the track (for minus strand reads.)

This special "slop" row of extra reads or annotations will be drawn in a slightly darker color than the individually stacked items in the rest of the track.

You can interact with both the individually stacked and "slop" reads or annotations; you can click on them, click-drag across them to select them, and so on, which is useful when you need to count the number of reads overlapping a gene or an intron. (More on this later.)

### Assess relative abundance of SR45 splicing isoforms

The SR45 gene is annotated as producing three possible variants. Which of these is the most highly expressed in the cold treatment and control samples?

To answer this question, use the aligned reads from the two RNA-Seq samples.

Focus on the differentially spliced region:

- Double-click the alternative exon in variant 1 (sixth from the left) to zoom in on it. Note it has two possible boundaries and that the intron to its immediate right (5-prime side relative to transcription) is an alternatively spliced intron.
- Scroll up through the treatment and control read alignments. Observe how many reads support the V1 versus the V2 version of the alternative intron. Which form is more common? Is the same form more common in the treatment or in the control sample?

To answer this question, we have to look at reads that align across the intron immediately to the right of the alternative exon.

The intron has two alternative 3-prime boundaries, also called "acceptor sites." Remember that the gene is transcribed from right to left (see arrows in IGB) because it is located on the minus strand of the chromosome.

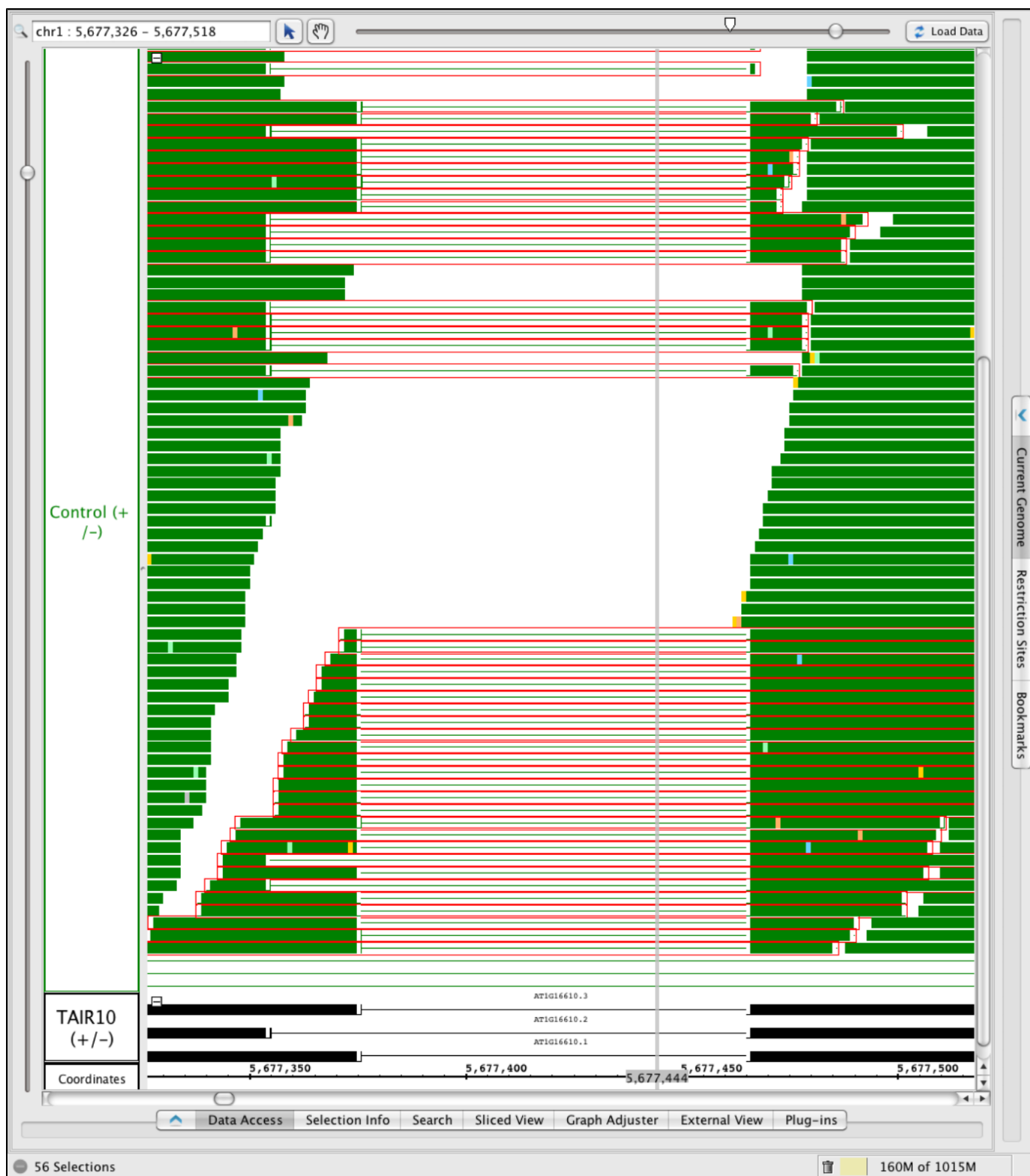
With respect to splicing, the most informative reads are spliced reads, reads that align across the alternative intron. These spliced reads can support one or the other choice, but not both.

You can count all the reads that overlap the intron simply by selecting them. When you click-drag over a region using the arrow (selection cursor), then anything you capture in the drag will be added to the list of currently selected items.

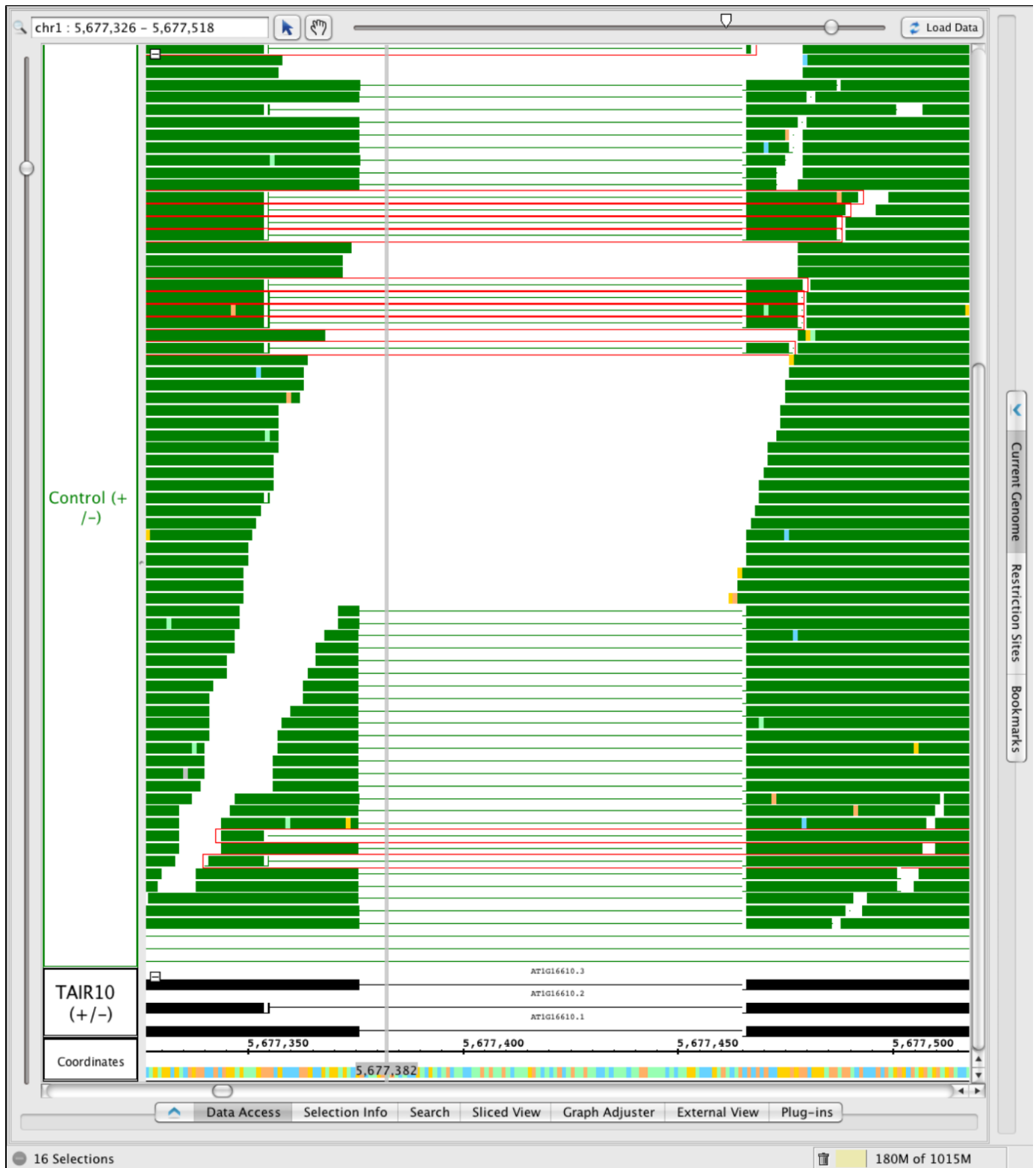
- **Select** reads by click-dragging over a region
- To **add** an item to the selection list, **SHIFT-CLICK** it
- To **remove** an item to the selection list, **CONTROL-SHIFT-CLICK** it.

Use selection to count the number of reads that support V1 or V2.

**IGB with 56 reads selected.** Selected reads include just the reads that were spliced across the alternative intron - selections include reads supporting either V1 or V2.



IGB with just the V2 supporting reads selected:



Based on this analysis, it appears that the V1 form (AT1G16610.1) is more abundantly expressed in the control sample.

That is, out of 56 spliced reads, only 16 (around 25%) supported V2 and the rest supported V1.

- Repeat the process for the Cold-treated sample.
  - Were there fewer or more reads overlapping the alternative intron in the cold-treated sample?
  - What percentage of reads supported V2?

What preliminary conclusions can you draw about cold stress and splicing in SR45?



## Finishing up - export an image showing what you found

If you would like to make images like the ones you see above, choose File > Export image. A new window will appear giving you the option to create image files in various formats, set image size and resolution, and save either the whole IGB window, the main view, or the main view with labels.

- Use **File > Export Image** to create a "souvenir" image from your session with IGB.

## RNA-Seq Coverage graphs - Visualizing data from maize reproductive transcriptomes

- [Introduction](#)
  - [About the data](#)
  - [What's a coverage graph?](#)
- [Setting up](#)
  - [Setting IGB Preferences - The preferences window from IGB 6.7.0](#)
  - [Rescale the graphs, add y-axis](#)
- [View coverage graphs from four samples](#)
  - [Explore the data.](#)
    - [Use IGB's zoom controls to zoom around in the data](#)
    - [Stretch the display vertically to show more detail, including gene model labels](#)
    - [Jump-zoom to a gene model by double-clicking on it.](#)
  - [Examples of new genes and/or differential expression across sample types](#)
    - [A multi-kilobase view of chromosome 1](#)
    - [Gene expression high in ovule and lower in leaves and embryo samples.](#)
    - [A single-exon gene expressed in pollen but not the other sample types.](#)
    - [A new gene?](#)
- [Get more info from Phytozome, maizegenome.org, or google.](#)
- [A complex locus with multiple predicted splicing variants.](#)
- [Data loaded for all of chromosome one](#)

### Introduction

A paper published in 2011 titled **Utility of RNA Sequencing for Analysis of Maize Reproductive Transcriptomes** describes a study that used RNA-Seq to characterize genes expressed in reproductive structures from maize.

### About the data

The study sequenced fourteen libraries prepared from fourteen sample types, including pollen, ovules, tassels, anthers, etc. One replicate per sample type was included.

To find out more about the study,

- [Read the paper \(PDF\)](#)
- [View sequencing records from the Short Read Archive](#)

In this tutorial, you'll learn how to load these data into IGB, view genome coverage graphs for a subset of samples, and identify genes that are expressed primarily in ovules, pollen, or other sample types.

To make these data available for visualization in IGB, we downloaded the sequence data from the Short Read Archive and then aligned the sequences onto a maize reference genome assembly using the spliced alignment tool TopHat. We used a program called wiggles (from the TopHat distribute) to generate coverage graph files for each sample. We then set up the files on IGB's QuickLoad site at <http://www.igbquickload.org/quickload>.

For a description of how we set up the data in the IGB QuickLoad system, please see this page from the IGB Developer's Guide:

- [Creating a new genome release for IGB - an example from Zea mays](#)

### What's a coverage graph?

A coverage graph in an RNA-Seq experiment is a type of genome graph that reports the number of reads overlapping regions in the genome. In a coverage graph, the x-coordinates are the genomic base pairs and the y coordinates represent the number of reads overlapping each genomic base.

In this experiment, graphs represent read counts; no normalization has been done. Thus simply comparing graph heights "by eye" can be misleading; you will need to take into account the fact that different samples produced different numbers of mapped reads. However, the number of mapped reads per sample type was similar across the different samples and so the relative heights of coverage graphs can give a reasonably good approximation of gene expression levels per sample.

Here is an example of the first few lines from a bedgraph file:

1	0	6	1
1	6	9	2
1	9	10	4
1	10	35	5
1	35	41	4

The first column reports the name of the sequence (chromosome 1), the next two columns give the start and end coordinates of the first region, and the third number reports the number of reads that overlap that region. Note that the coordinates are in interbase.

- Information about interbase: [Introduction to Chado - Interbase Coordinates](#).

When we processed these data to make them available via IGB QuickLoad, we also computed read alignment distributions for each library. You can view the read alignment distribution files in the same Web directories where the coverage graphs and read alignment files reside. Click the IGB Quickload Data Source to view the data repository and navigate to this folder:

[http://igbquickload.org/quickload/Z\\_mays\\_B73\\_Mar\\_2010/SRP006463/](http://igbquickload.org/quickload/Z_mays_B73_Mar_2010/SRP006463/)

## Setting up

Follow these instructions to configure IGB to get started looking at gene expression in maize reproductive tissues.

- Start IGB
  - Download and launch IGB using Java Web Start from <http://www.bioviz.org/igb/download.shtml>
  - Or download igb.zip from <http://sourceforge.net/projects/genoviz/files/>.
    - To launch IGB from igb.zip, double-click igb.zip to unpack it, open folder igb, and double-click run\_igb.bat (Windows) or run\_igb.command (Mac)
- Choose species **Zea mays B73**, Genome Version **Z\_mays\_B73\_Mar\_2010**

**Note** that the gene models for maize should automatically load.

- Customize gene models appearance
  - Right-click (or control-click on Mac) the Zea mays gene models track and choose
    - Choose **Change > Show track (+/-)** to combined plus and minus gene models into the same track
      - Note that when you zoom in far enough, you'll see arrows appear that indicate the direction of transcription
  - Change background color to white and foreground color to black to maximize readability for printouts and slides

Now start opening coverage graphs from different samples from the RNA-Seq experiment.

- Under the **Data Access** tabbed panel, open
  - Data Source **IGB Quickload > RNASeq > SRP006463 Reproductive > Single-map reads**
- Click the checkboxes next to data sets or click-drag them over to the **Data Management Table**.
  - coverage SRR189770 Ovule
  - coverage SRR189771 Pollen
  - coverage SRR189766 Embryo
  - coverage SRR189761 Leaves

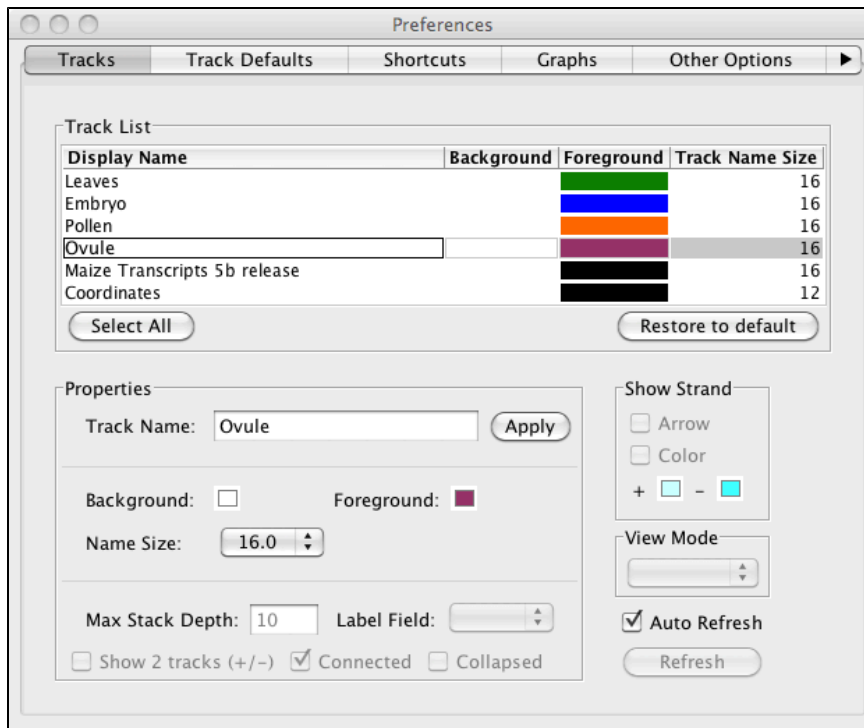
The four data sets will automatically be added to the **Data Management Table**. Don't worry if you accidentally add a different data set. You can remove it by unchecking the box or clicking the red x.

Take a moment to customize the colors. Click the color squares in the **Data Management Table** (FG is foreground, BG is background) or choose **File->Preferences->Tracks** to open a new Window that let's you customize different aspects of the display at several at once. Click each row in the table and give each row a white background. Choose colors for individual rows by clicking on them and then selecting options.

**Tip:** Simplify the track name to make the Track Labels easier to read.

**Tip:** Color the gene track black (foreground) and white (background) to make attractive images for slides and printing.

## Setting IGB Preferences - The preferences window from IGB 6.7.0



Once you've set up the tracks, zoom in on a region.

- Enter 1:51631058-61254530 into the Search/Range Box (upper left corner) and type **Return**.

IGB should then scroll and zoom to the region you entered. Next, click **Load Data** and proceed from the first example below, after adjusting Y-axis scales as described in the next section.

**TIP:** You can also use IGB links from a Web site to navigate a genome in IGB. An IGB link looks like this: [http://localhost:7085/UnibrowControl?seqid=1&end=61254530&start=51631058&version=Z\\_mays\\_B73\\_Mar\\_2010](http://localhost:7085/UnibrowControl?seqid=1&end=61254530&start=51631058&version=Z_mays_B73_Mar_2010). When you click the link, IGB should scroll and zoom to the location indicated in the link, provided your computer allows access to non-standard ports. (This may not work on every computer.)

Once the data load, you'll see a view that looks a bit like the image below. To match the image below, you'll need need to rescale the graphs, add a y-axis, and so.

**Note to advanced users:** If you are running the high-memory version of IGB and have a fast internet connection, you can probably load all the data for chromosome 1 for all four graphs. However, if you try this and out of memory, you can always delete a graph (using the red X) and click the trashcan icon at the bottom right of IGB to force what is called "garbage collection," a programming term that means recovering memory.

## Rescale the graphs, add y-axis

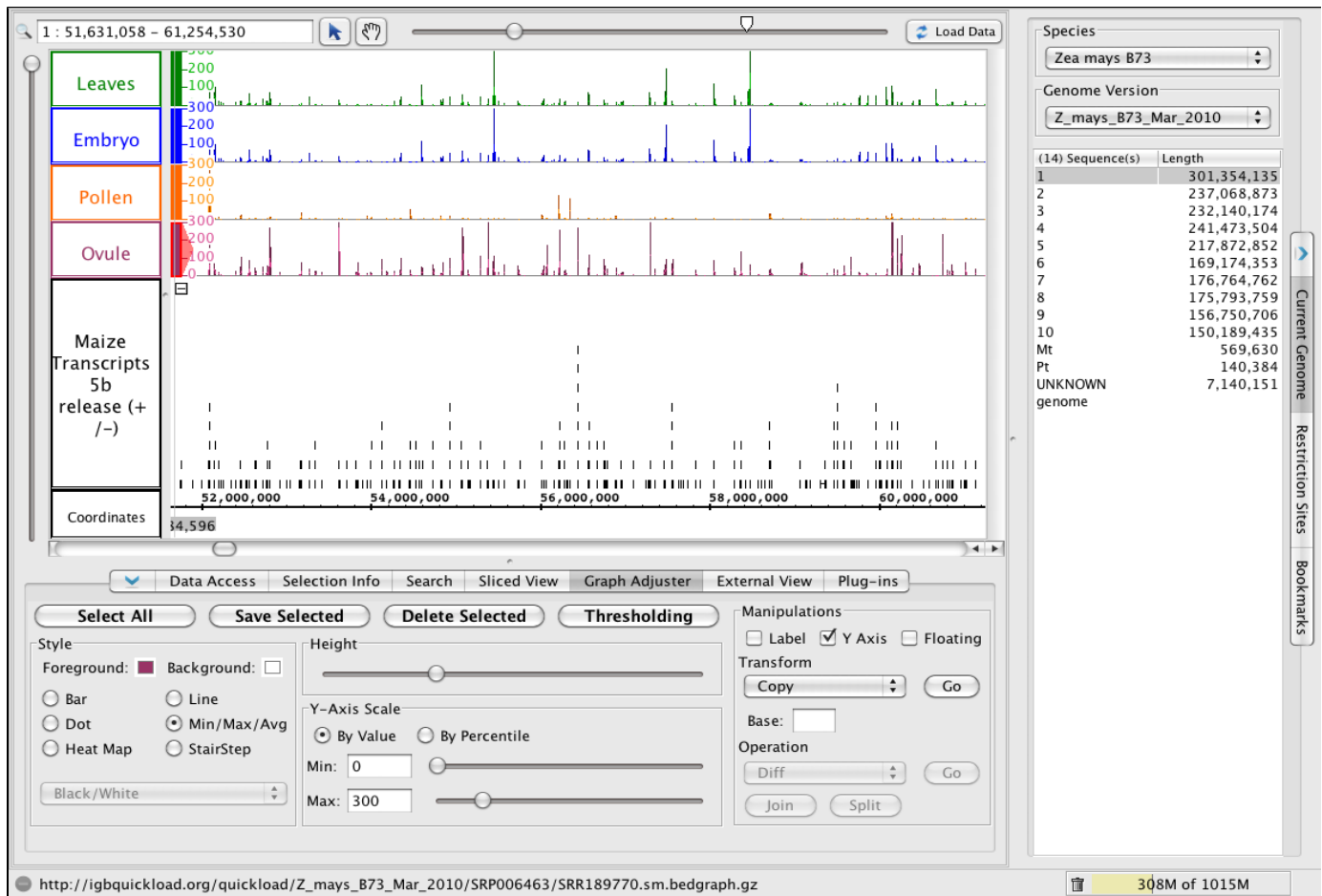
Use the **Graph Adjuster** tab to put all the graphs into the same scale, thus make it possible to compare them. Click the **Graph Adjuster** tab and then click the **Select All** button to select all the graphs. While all are selected to the following:

- Click **Y Axis** to add y axis scale to every graph.
- Unclick **Label** to hide the label for each graph.
- Under **Y-Axis Scale**, click **By Value** and then enter a new maximum. 300 is a good place to start for these data.
- Use the Height slider to increase the height for each graph.

After setting the colors and configuring the Graphs, you should see something like the following image. (The Ovule graph is selected here to show its **Graph Adjuster** settings.)

## View coverage graphs from four samples

- To jump to this region, enter 1:51631058-61254530 (same as above) into the Search/Range Box (upper left corner) and type **Return**.
- Click **Load Data** button (top right) to load the data into IGB.



## Explore the data.

Now that the graphs use the same scale, you can zoom and pan around the chromosome and observe how different genes are expressed in different samples.

## Use IGB's zoom controls to zoom around in the data

To zoom in, click the place you want to zoom - the vertical stripe (called a zoom stripe) will hop to the location of your most recent click and also will focus zooming so that when you zoom in or out, the zoom stripe will remain stationary while the scene around it expands or contracts. To zoom in, drag the horizontal slider (top right) to the right. To zoom out, drag the horizontal slider to the left.

## Stretch the display vertically to show more detail, including gene model labels

To stretch the display in the vertical dimension, drag the the vertical slider (left side of the IGB window) up. To compress the display (and make it fit inside the IGB main data display), drag the vertical slider down.

To zoom in a region, click-drag your mouse along the number line. During the drag, you'll see a blue highlight. When you release the mouse, you'll jump zoom to the region you selected.

## Jump-zoom to a gene model by double-clicking on it.

To zoom in on a gene model, just double-click it. This shortcut will help you quickly inspect interesting genes and their structure.

Scroll and zoom around the chromosome, making note of places where one or more samples appear to have more reads covering a particular region and double-click on gene models to zoom in for a closer look.

Also, keep your eyes open for regions with reads but no annotated gene. When you find one, click the corresponding read alignments file for the sample whose graph contained expression data. Load the reads to get an idea of the intron/exon structure for the unannotated gene.

Visit the following regions (see below) to see examples of new genes and differential expression across sample types.

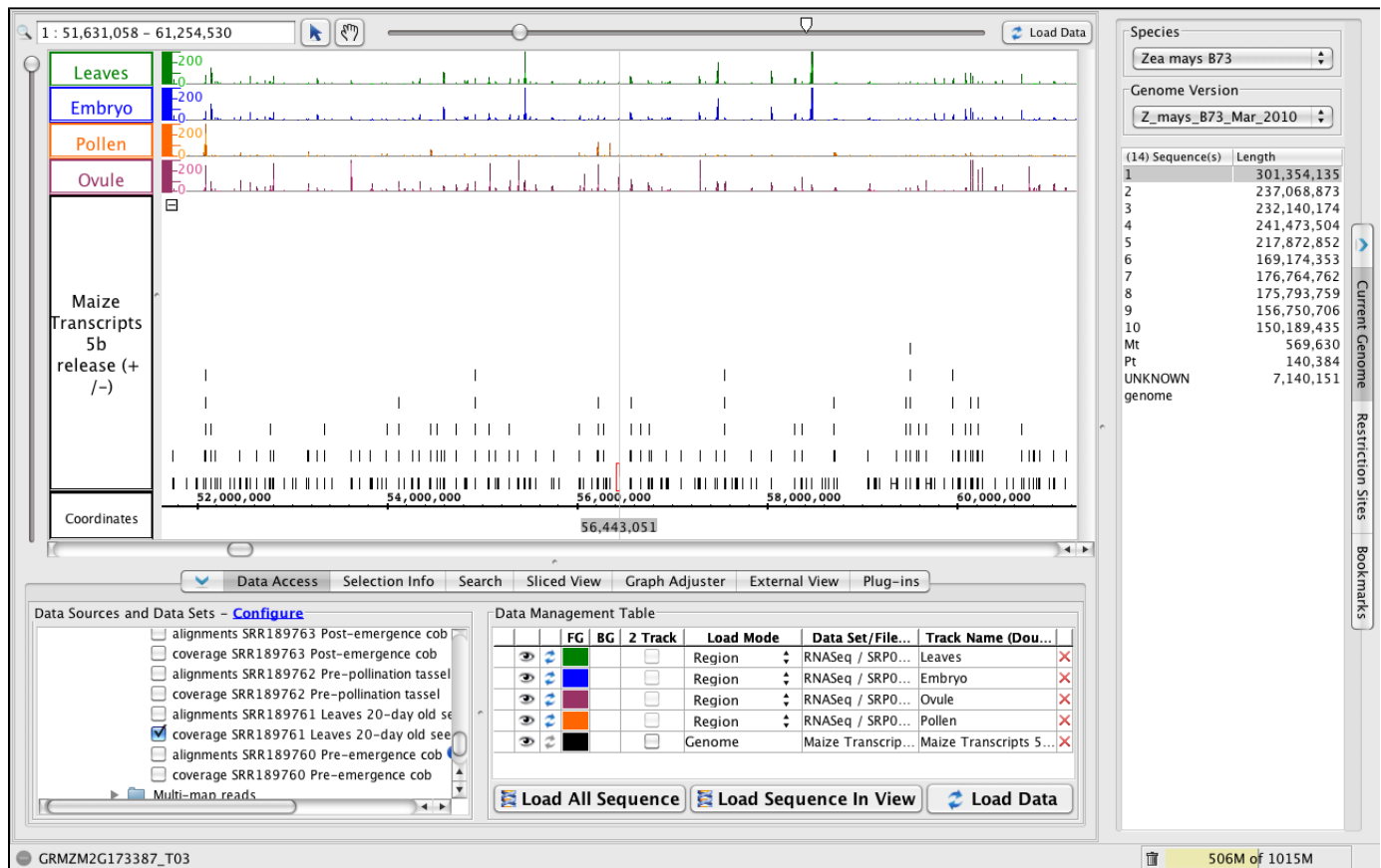
**Tip:** To navigate quickly to the regions shown below, enter the regions (as shown) in the upper left text box and press the Return key on your keyboard. IGB will then jump to the region you entered. You can also search for gene models by name using the **Search** tab.

## Examples of new genes and/or differential expression across sample types

### A multi-kilobase view of chromosome 1

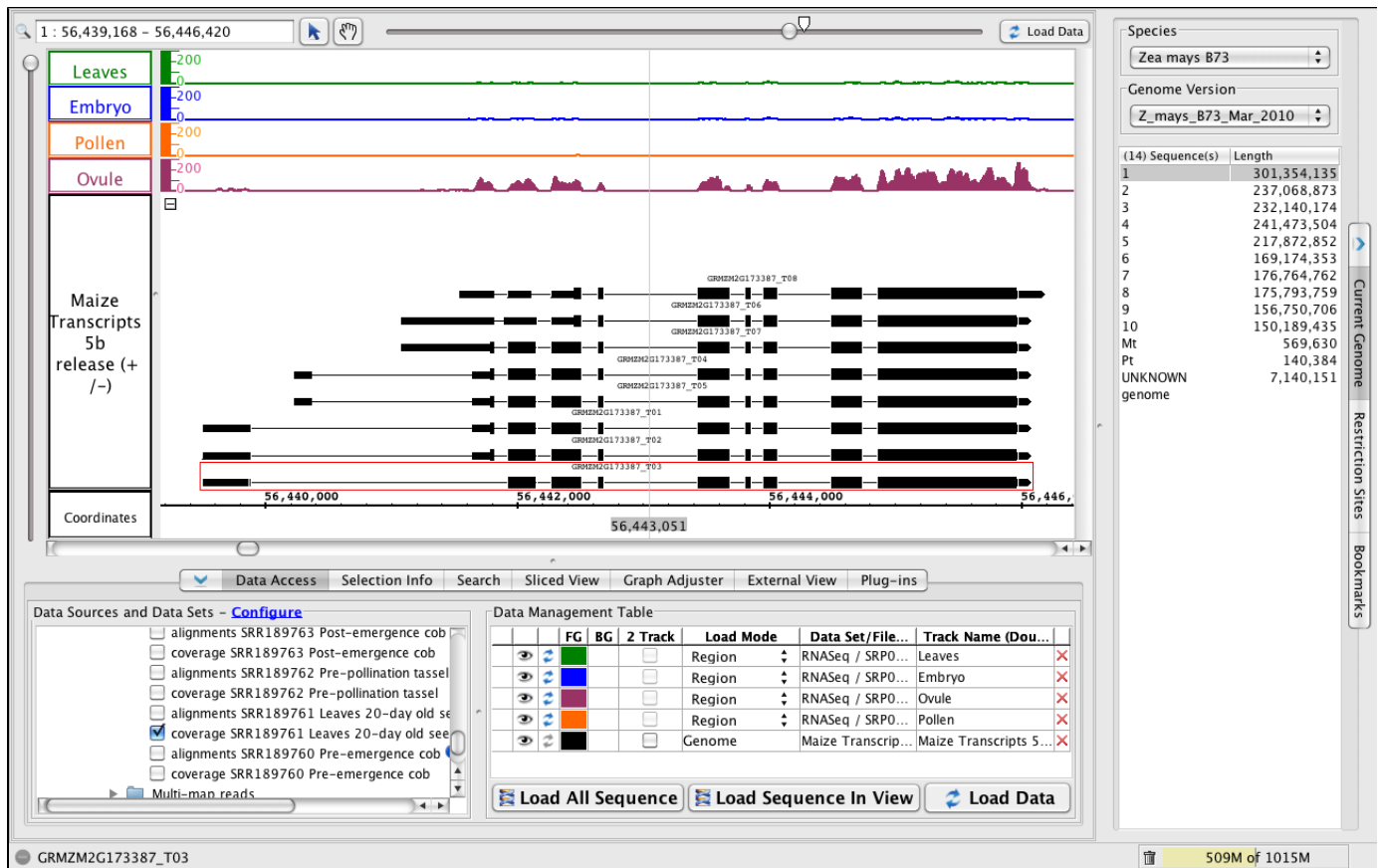
- To jump to this region, paste 1:51,631,058-61,254,530 into the Search/Range box and hit **Enter**.

Note how different graphs have peaks at different regions, but all graphs use the same scale.



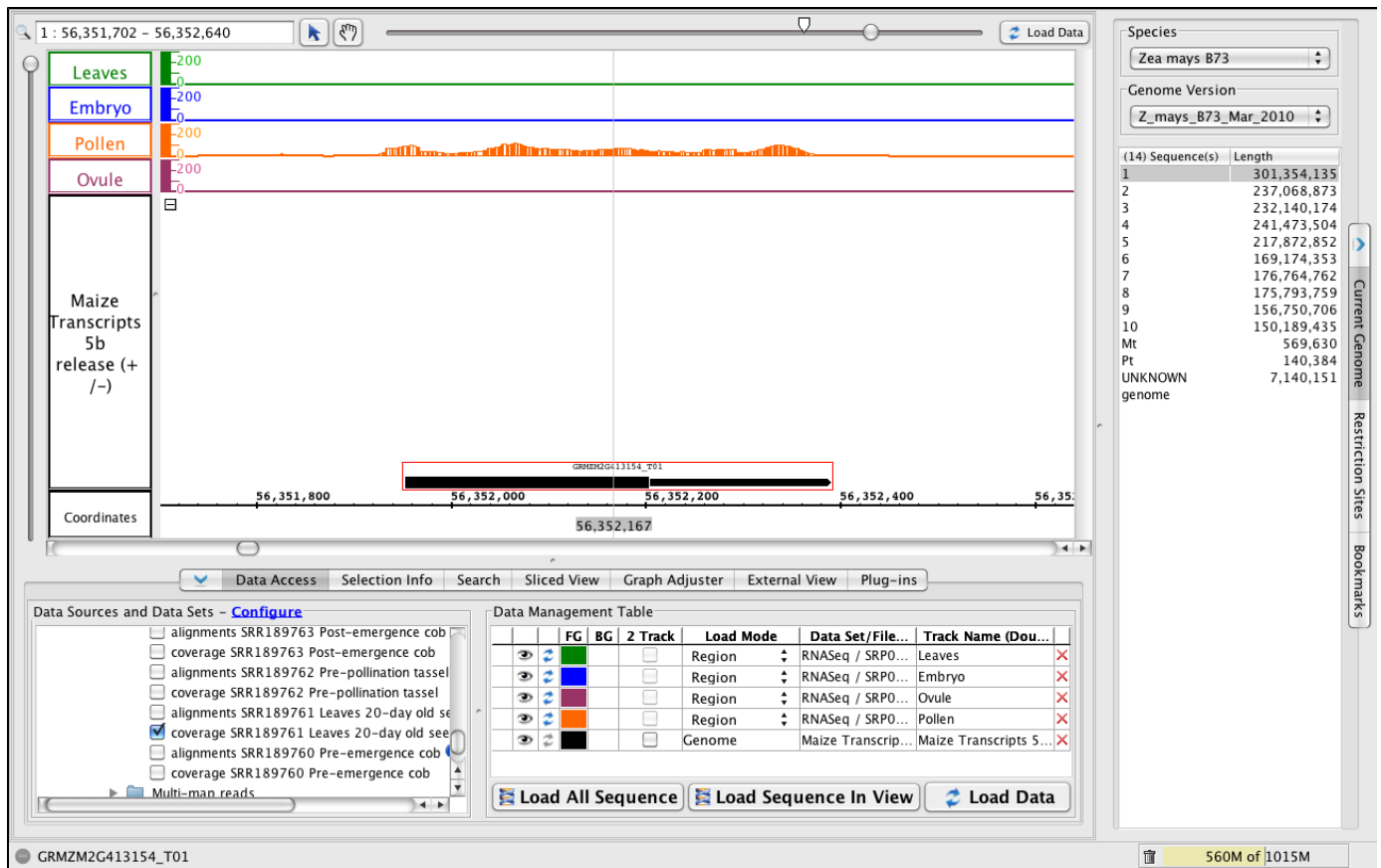
### Gene expression high in ovule and lower in leaves and embryo samples.

- To jump to this region, paste 1:56,439,168-56,446,420 into the Search/Range box and hit **Enter**.



A single-exon gene expressed in pollen but not the other sample types.

- To jump to the region, paste 1:56,351,702-56,352,640 into the Search/Range box and hit **Enter**.

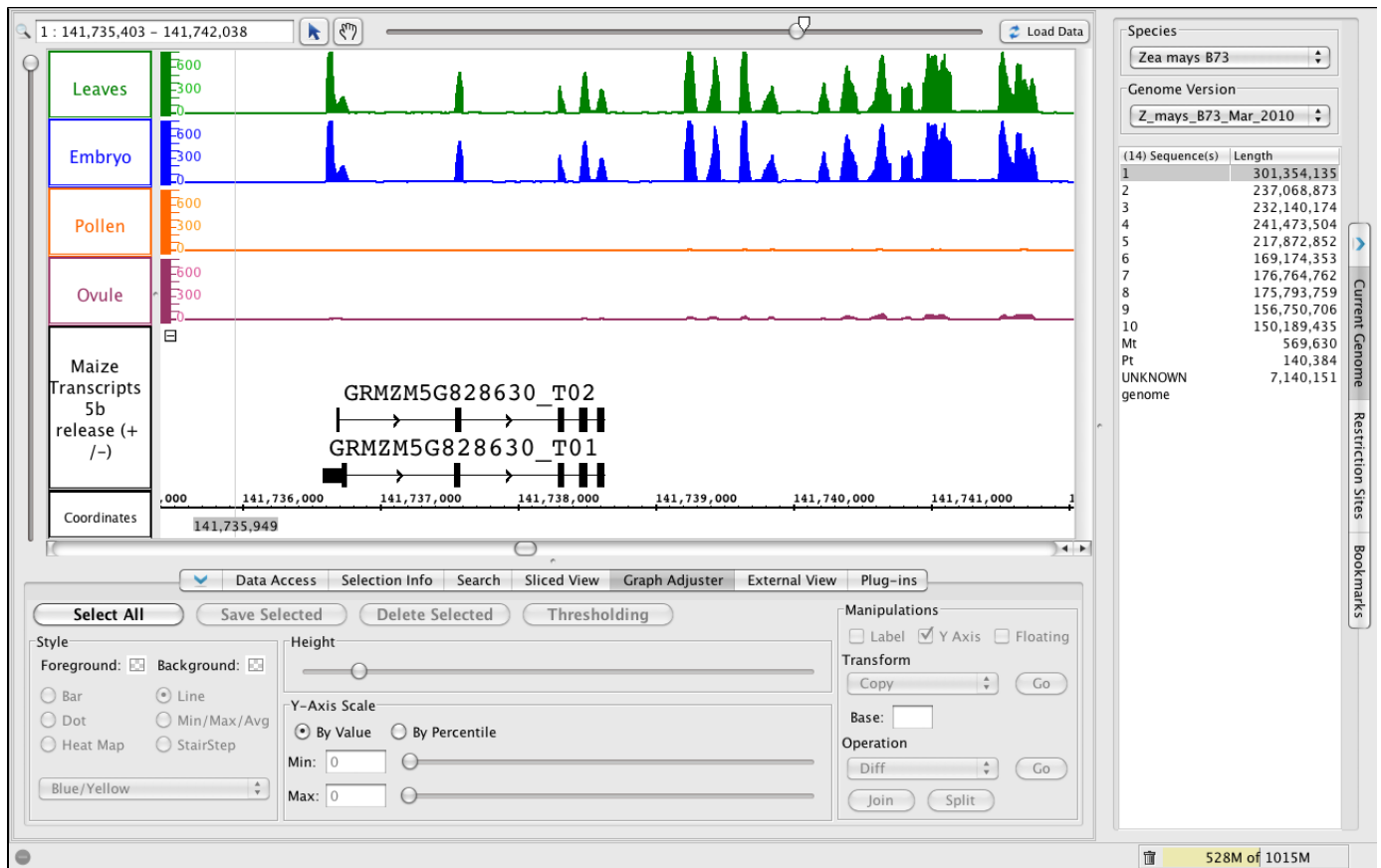


## A new gene?

- To jump to the region, paste 1:141,735,403-141,742,038 into the Search/Range box and hit **Enter**.
- Click **Load Data** to load the data for this region, which is outside the range of data you loaded earlier in the tutorial.

To replicate this image, modify the **Max Stack Depth** setting for this track, which controls the amount of vertical space above the two gene models.

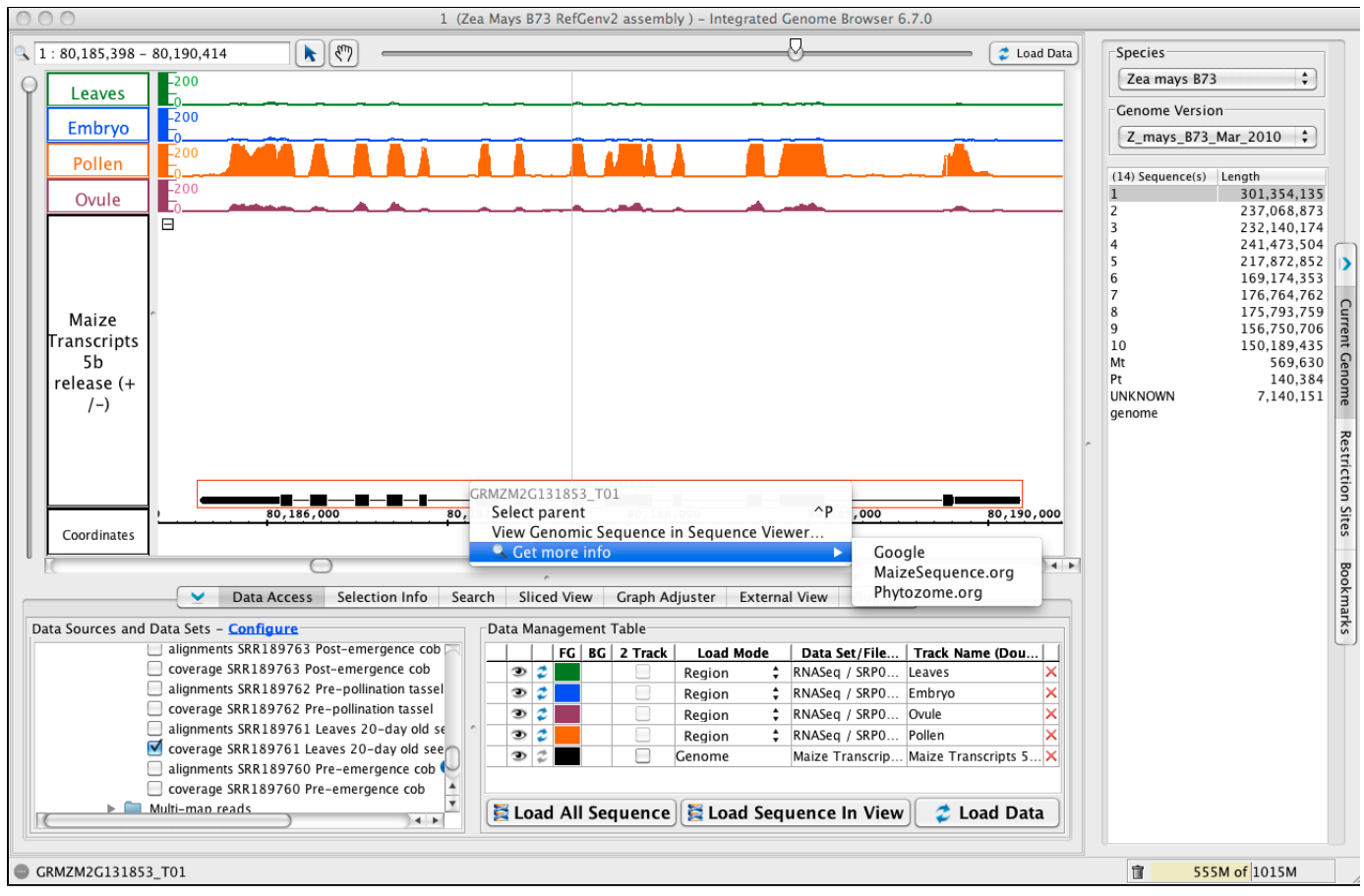
- Right-click the Maize Transcripts 5b track label and choose **Change > Adjust Max Stack Depth**. Enter 2. This will tell IGB to never make a "stack" of individually drawn gene models that is larger than 2. Any other gene models will be drawn on top of each other in the very top row (the "slop" row) at the top of the track.
- When you look at a gene with more than two gene models, use **Change > Adjust Max Stack Depth** to change this setting to a larger number to see all the gene models individually, or enter 0 to allow unlimited stacking.



## Get more info from Phytozome, maizegenome.org, or google.

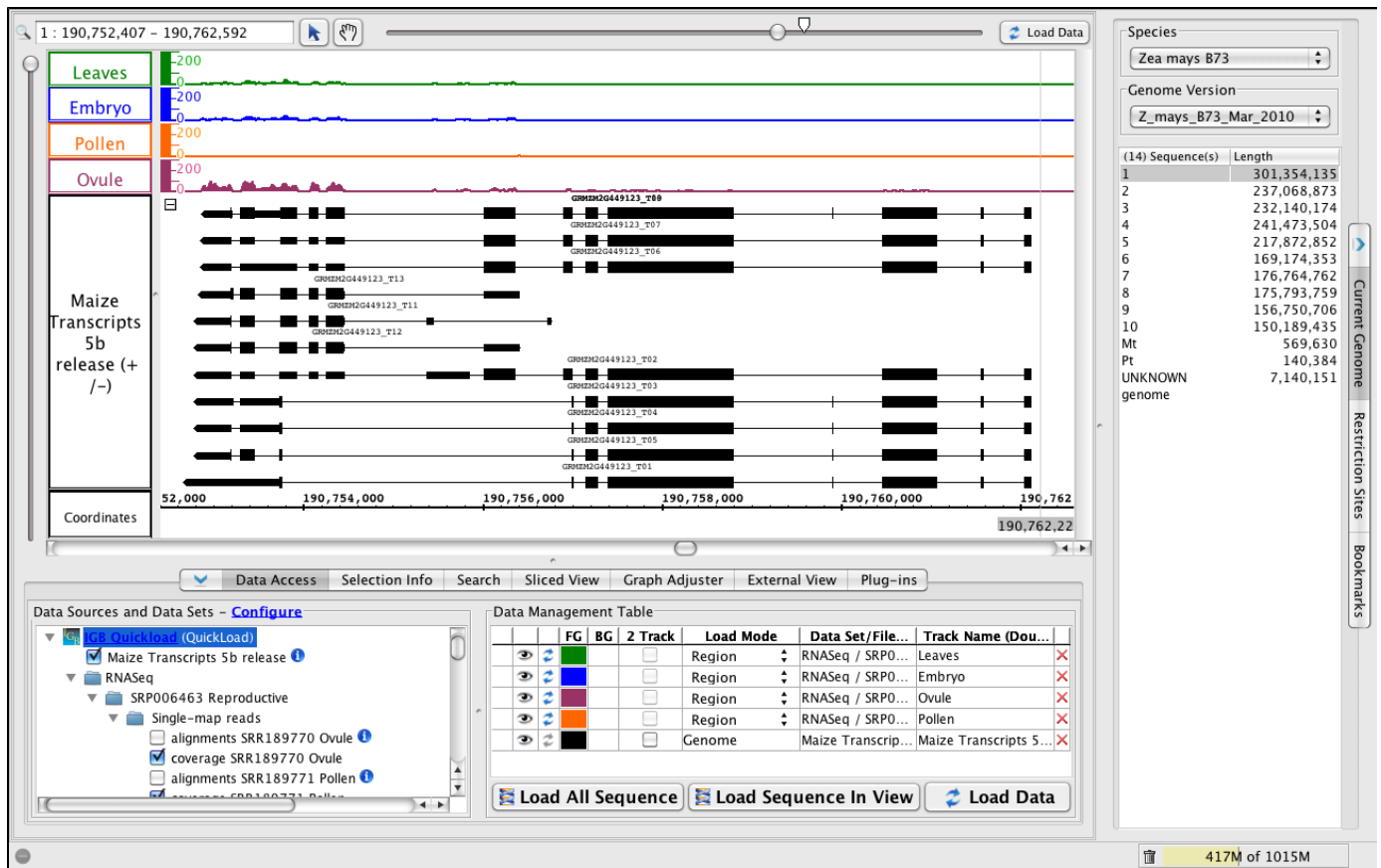
- To jump to the region, paste 1:80,185,398-80,190,414 into the Search/Range box and hit **Enter**.
- Click **Load Data** to load graphs data for this region, which is outside the first region you loaded.
- Right-click the gene model (as shown below) and choose **Get More Info** to search Phytozome, MaizeGenome.org, or Google for information about the selected gene.





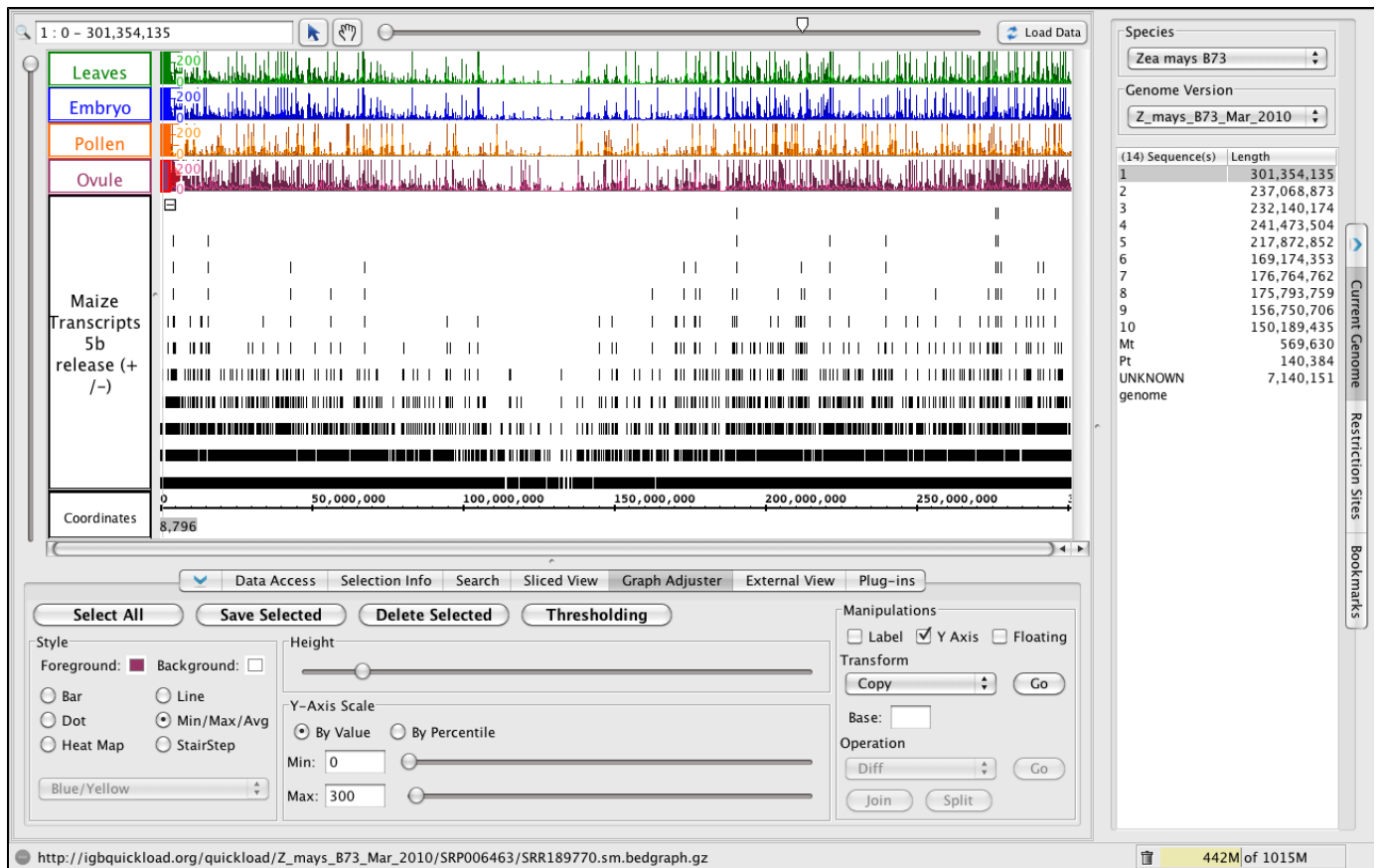
## A complex locus with multiple predicted splicing variants.

- To jump to the region, paste 1:190,752,407-190,762,592 into the Search/Range box and hit **Enter**.
- Click **Load Data** to load graphs data for the region.
- To view all the variants, change the **Max Stack Depth** for the track.
  - Right-click the Maize transcripts 5b track label and choose **Change > Adjust Max Stack Depth**. Enter 13 or 0 for unlimited.



## Data loaded for all of chromosome one

- Warning: Loading all of the data from the remote server (igbquickload.org) may take some time! Also, your computer may not have sufficient memory to display all four graphs. But if you would like to try it and view all the data at once, use the horizontal slider (top right) to zoom all the way out. Then, click **Load Data**. If the loaded data take up too much memory (e.g., close to your maximum memory limit shown on the lower right corner of IGB), then you can delete a data sets by clicking the red X or unchecking it in the **Data Sources and Data Sets** folder tree.
- To recover memory, click the trash can icon in the bottom right of the display.
- To halt data loading, click the minus symbol that appears in the bottom left whenever data are being accessed from a server or local file.



## View a bacterial genome in IGB

This tutorial describes how to import sequence and annotations into IGB when your genome of interest is not already available from an IGB QuickLoad or Distributed Annotation site. In this tutorial, we'll demonstrate importing a bacterial genome (E. coli) using files downloaded from NCBI.

### Get genome data from NCBI

To view a bacterial genome in IGB, first visit NCBI to retrieve the sequence data and annotations.

- Go to the NCBI GenBank record for E. coli K-12 substr. MG1655: [http://www.ncbi.nlm.nih.gov/nuccore/NC\\_000913.2](http://www.ncbi.nlm.nih.gov/nuccore/NC_000913.2)

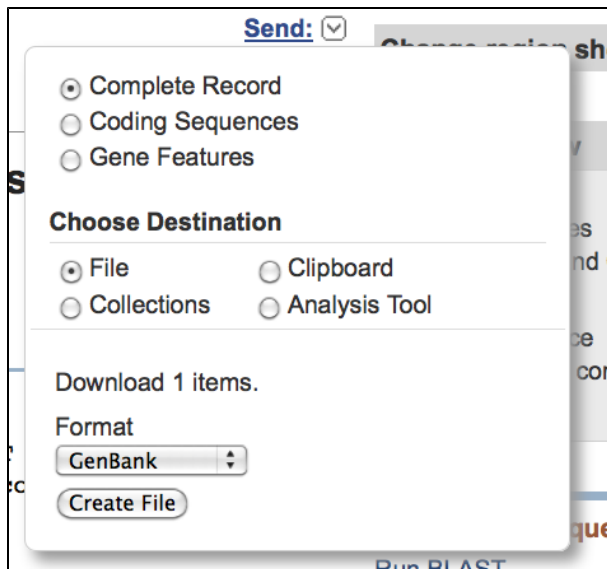
### Download annotations file

Click the **Send** menu (on the right side of the record) and choose **Complete Record**, **File**, **GenBank**, and **Create File**.

Save the file to your computer; by default, it will be named "sequence.gb".



If you change the name, be sure to use the ".gb" file extension so that IGB can recognize the file and the file format.



Next, click on the FASTA link just under the name and reference of the record. This will load up the sequence for this genome. In the picture below, after FASTA has been selected, the Display Settings indicate that the record shown is the FASTA; the options below the name and the NCBI Reference Sequence number offer GenBank and Graphics.

Again, using the **Send** function select **Complete record**, **File**, **FASTA** (default), **Create file**.

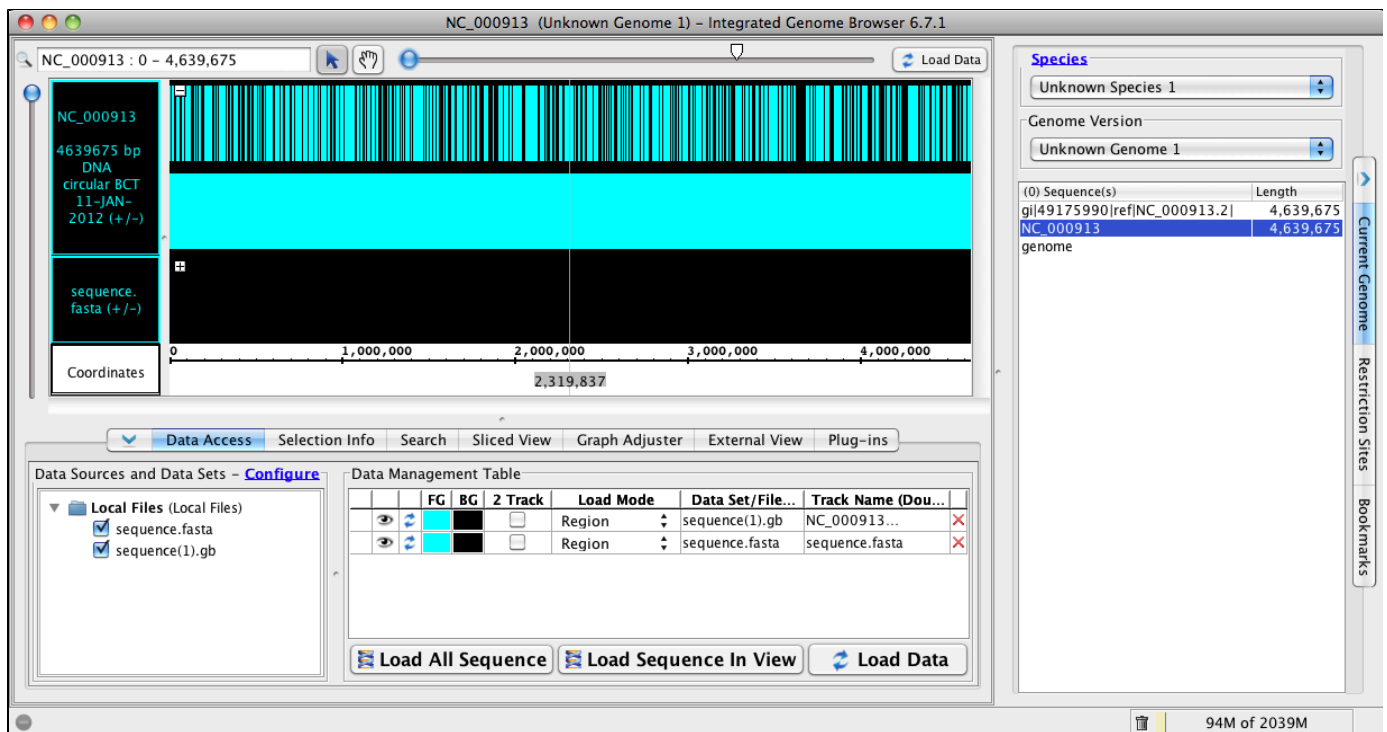
This will create a new file called 'sequence.fasta'. This file contains the reference sequence. Make sure that the extensions for the fasta file is either '.fasta' or '.fa'. If the extension is '.txt', you can safely change it to '.fa'.

## Step 2: Synonyms

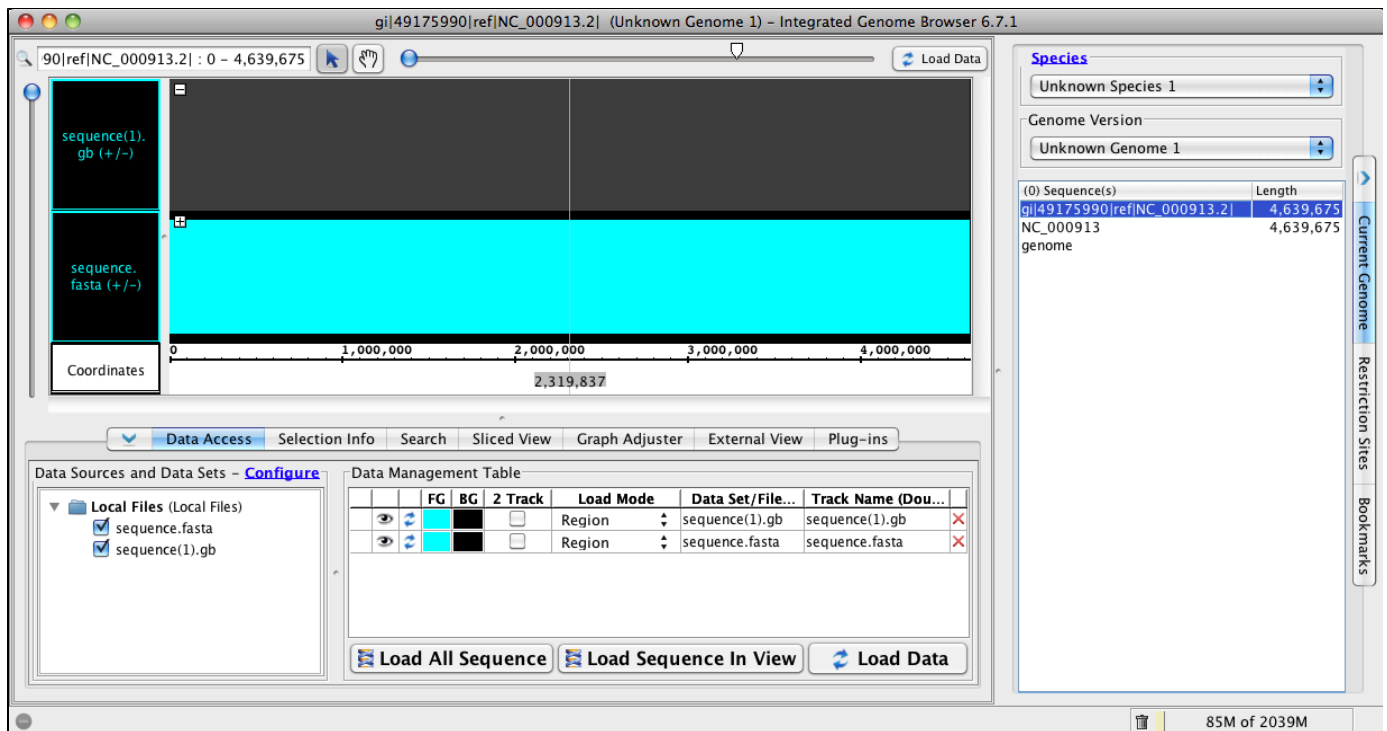
When IGB shows these files, it uses the header of each file to identify the 'chromosome' that is being displayed. Unfortunately (especially with NCBI files), the headers for each file type (FASTA sequence files, annotation files, .bam/alignment files, etc.) can ALL be different for the exact same chromosome/sequence! IGB will see different headers and treat each one as a separate chromosome; they will not be visualized together, even though they are all the same length.

To overcome this, we have the ability to enter 'synonyms'; IGB already has some internal synonyms so that "1" and "chr1" are all equivalent. However, the NCBI reference numbers are specific to each chromosome, for each organism. You will need to know the header of each of these files; if you are not sure, a quick way is to drag all of the files into IGB at the same time.

In the picture below, we have dragged (in order) the gene models (.gb) and the sequence file (.fasta). In the **Current Genome** panel we have "NC\_000913" and "gi|49175990|ref|NC\_000913.2|". Notice that both **Sequences** have the same **Length** in the Current Genome panel. If we select "NC\_000913", and then **Load Data**, we can show the annotations but we can NOT show the sequence.

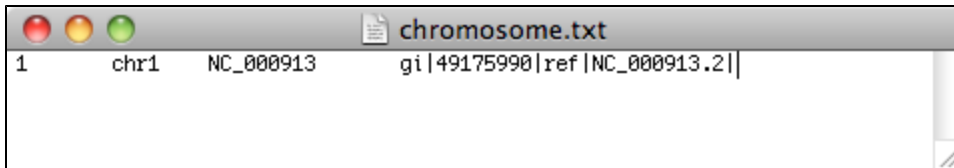


Alternatively, if we select the 'gi' number and hit load, the sequence file loads (as a track) but the models cannot be seen.



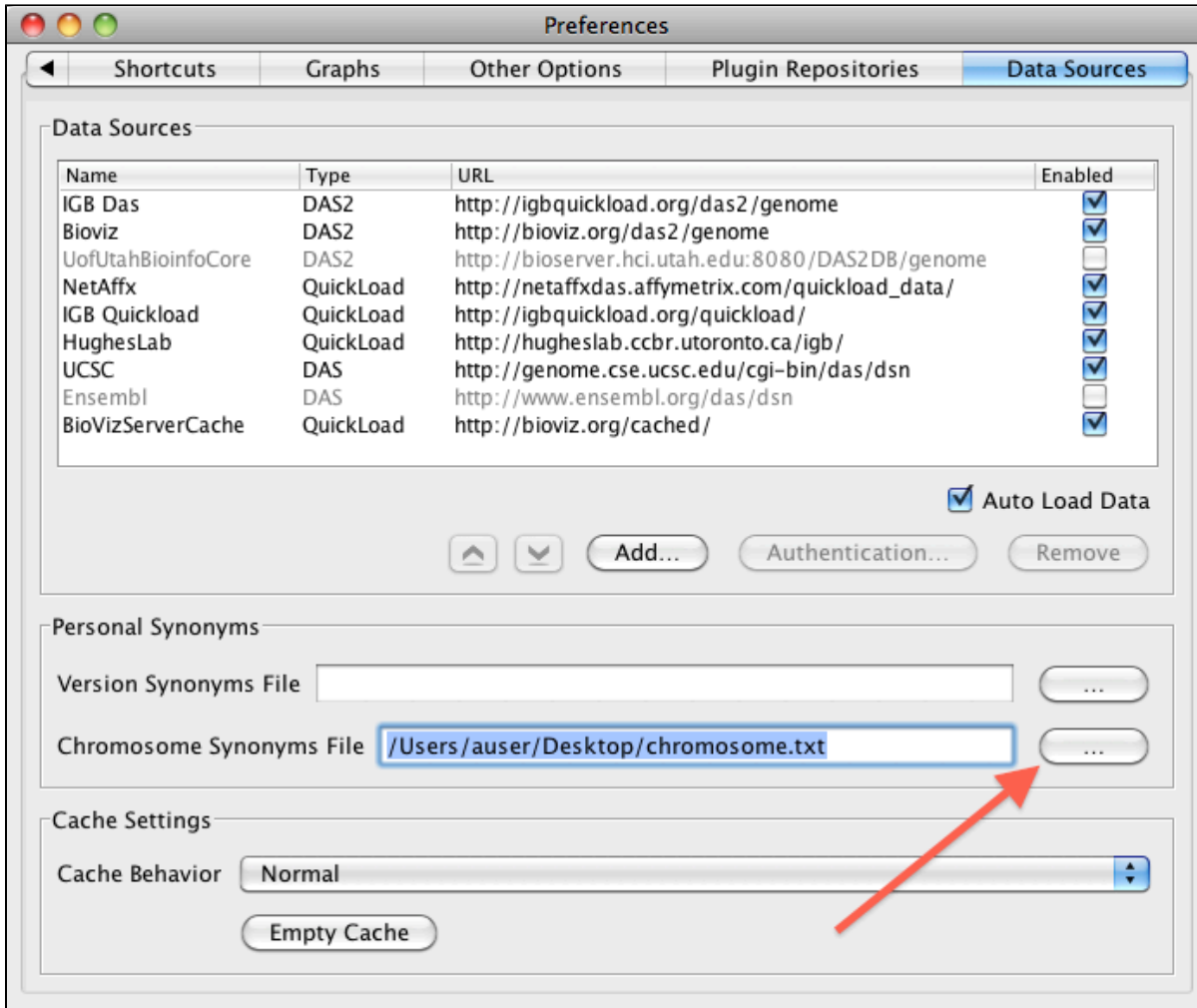
To fix this issue, we will create a 'personal synonym' file. This is a simple text file that is tab separated. All terms included in a single line will be 'equivalent' and IGB will know that they are to be shown together. To create and load the synonym file is very simple:

- Open a text editor; this must be simple text (.txt), and not rich text format (.rtf)
- Type in 1 and then a tab. Type in chr1 and then a tab. (1 and chr1 are standard names for the first chromosome, and many files use these as their headers. Typically, we always include these two options in a synonym file)
- Type in the name of the first file (i.e. NC\_000913) and then hit the TAB key.
- Type the second name (i.e. gi|49175990|ref|NC\_000913.2|)
- Save this file with the name chromosome.txt



Just that easy! If you are making a synonym file for a multi-chromosomal organism, then make a new line in the file for each chromosome, and just add all of the 'names' associated with it (make sure that there is a 'tab' between each name!). If you include a file that has a new name, open the chromosome.txt, and add the name to the proper line (be sure to include the tab between the last name and the new name!).

- Now, while IGB is open, choose **Configure** from the **Data Access** panel or open **File > Preferences > Data Sources**. Click the '...' button next to **Chromosome Synonyms File**, select your chromosome.txt file and click open. The file path will appear in the open box, showing that it is loaded. You can close the preferences window at this point, and IGB will remember to load your chromosome.txt file for you every time it opens. This means if you make a change in your chromosome.txt while IGB is open, you will need to re-open IGB to see those changes take place.

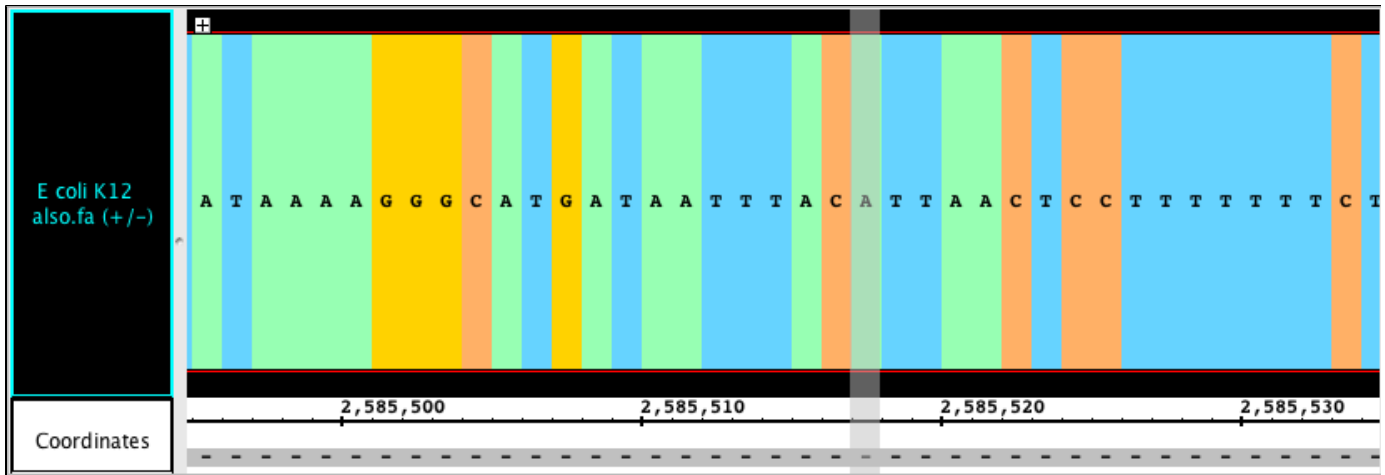


- IGB will need to be restarted to properly load and recognize your chromosome.txt; close this instance of IGB and then reopen IGB.

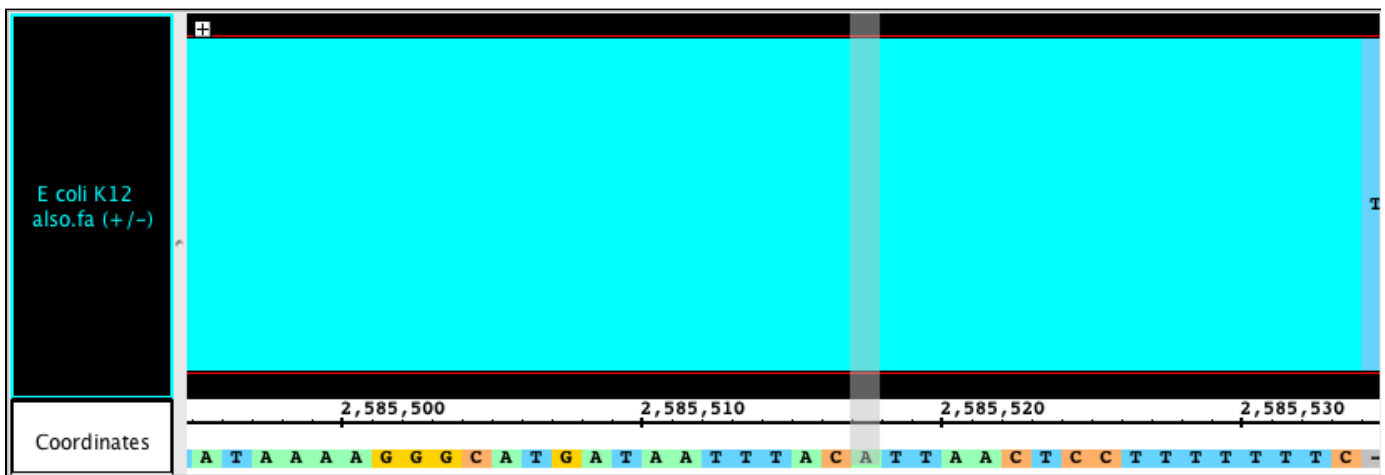
### Step 3: Visualizing the Reference Sequence and Models

At this point, we have the sequence file and we have the gene models. When you open the new instance of IGB, your synonyms will be loaded for you.

We will now open these two files so that you can begin analysis of your own data. The first step will be to open the sequence file as a reference sequence rather than as a track. This will allow the sequence to load into the 'sequence' area of the **Coordinate Axis**, and IGB will use it for sequence comparison with your data. If a sequence file is loaded as a track (and there is NO reference sequence) it shows the actual nucleotide sequence (the reference sequence line shows dashes and it greyed out) but it cannot be used for comparison to your .bam files or other data:



If the reference sequence is loaded, however, the sequence in the track will ONLY show mismatches, but otherwise appear as a solid bar. The reference sequence line in the Coordinate Axis area will now have the nucleotides show (when zoomed in).



To open up the sequence and annotation files:

- **File > Open Reference Sequence...**
- Chose the 'sequence.fasta' file.
- Drag the sequence.gb (gene models) into IGB.
- Click the **Load Data** button. **NOTE:** you need to click the **Load Data** button and NOT the **Load Sequence** button to visualize your sequence file!

Both the sequence and the models will load. You will be zoomed out, so the sequence will appear as a grey bar; as you zoom in the colors and nucleotides will become visible. At this point, you can drag in your files of interest and begin viewing them with the models and sequence. Keep in mind that if the name of the chromosome(s) is different in **your** files, you will have to add the name to the chromosome.txt and then reopen IGB so it can load in the new information.

## Video Tutorials

- IGB Quickies
- Focus on a Feature
- In-Depth Analysis

## IGB Quickies

Short how-to videos demonstrating specific features within IGB.

## Focus on a Feature

Previously recorded on-line tutorials demonstrating in depth use of IGB.

## In-Depth Analysis

Longer videos describing more in-depth analysis of data and features in IGB.

### IGB Quickies

- IGB Quickies
  - Quick Start
  - Zooming
  - Semantic Zooming
  - Panning the Display
  - Adjusting Vertical Data Display
  - Adjusting Track Height and Track Order
  - Adding a New or Custom Genome
  - Exporting Images
  - Creating Bookmarks

### IGB Quickies

#### Quick Start

#### Zooming

#### Semantic Zooming

#### Panning the Display

#### Adjusting Vertical Data Display

#### Adjusting Track Height and Track Order

#### Adding a New or Custom Genome

#### Exporting Images

#### Creating Bookmarks

### Focus on a Feature

- Focus on a Feature
  - Visualization of Whole Genome DNA Methylation
  - ChIP-Seq Alignment and Visualization
  - All New Features in IGB 8.2
  - RNA-Seq Alignment and Visualization
  - Alternative Splicing and Primer Design
  - Working with Sequence Data in IGB
  - Working with Graphs in IGB

### Focus on a Feature

#### Visualization of Whole Genome DNA Methylation

#### ChIP-Seq Alignment and Visualization

#### All New Features in IGB 8.2



## RNA-Seq Alignment and Visualization

## Alternative Splicing and Primer Design

## Working with Sequence Data in IGB

## Working with Graphs in IGB

## In-Depth Analysis

- [In-Depth Analysis](#)
  - [Viewing RNA-Seq Data](#)

## In-Depth Analysis

### Viewing RNA-Seq Data

## Older Versions of IGB

## Getting Older Versions of IGB

### Get IGB from [bioviz.org](http://bioviz.org) Java Web Start

If you prefer to continue using IGB 8.2.3 with Java Web Start, you can access this version from [this page](#). Please be aware we advise against using this version.

### Get older versions of IGB from [SourceForge.net](http://SourceForge.net)

#### Get igb.zip

The IGB distribution available from [SourceForge](http://SourceForge.net) ([igb.zip](#)) includes a compiled "jar" (java archive) file called **igb\_exe.jar**, which you can use to run IGB on almost any system.

- Download igb.zip file and unpack it.

#### Open the igb folder

Inside this folder you will find three files: run\_igb.command (for Mac), run\_igb.bat (for Windows) and run\_igb.sh (for other UNIX systems).

On Windows or Mac systems, you can double-click the file to launch IGB.

You can also run IGB from the command line by typing

```
$ run_igb.command
```

## System Requirements for Older Versions

### IGB System Requirements

IGB version 8.2.3 and earlier can run on any computer that has the Java Runtime Environment (JRE) version 1.6 or higher. (This is also called Java 6.) IGB is developed and tested mostly on Macintosh and Windows, but should work on any computer with Java enabled.

The amount of RAM memory and disk space required depends on the type of data you intend to view. We recommend running IGB with at least 2 Gb RAM.

## Troubleshooting Older Versions

See the links below for common problems and their solutions.

**Note:** To run older versions of IGB, you will need Java installed on your computer. Most computer types already have Java, but if you need to install it yourself, visit the [Java download page](#). To find if you already have Java, click the link labeled [Do I have Java?](#)

### IGB can't be opened because it is from an unidentified developer (Mac)

If you are a Mac user, you may see the following message when you attempt to download and run IGB from the [IGB download page](#):



The IGB download page uses Java Web Start to distribute IGB, and so you may see this message if your Mac is configured to block Java Web Start applications developed by non-Apple verified developers.

There are at least three (maybe more!) workarounds. Each one has benefits and drawbacks. If you have questions, let us know.

#### Workarounds

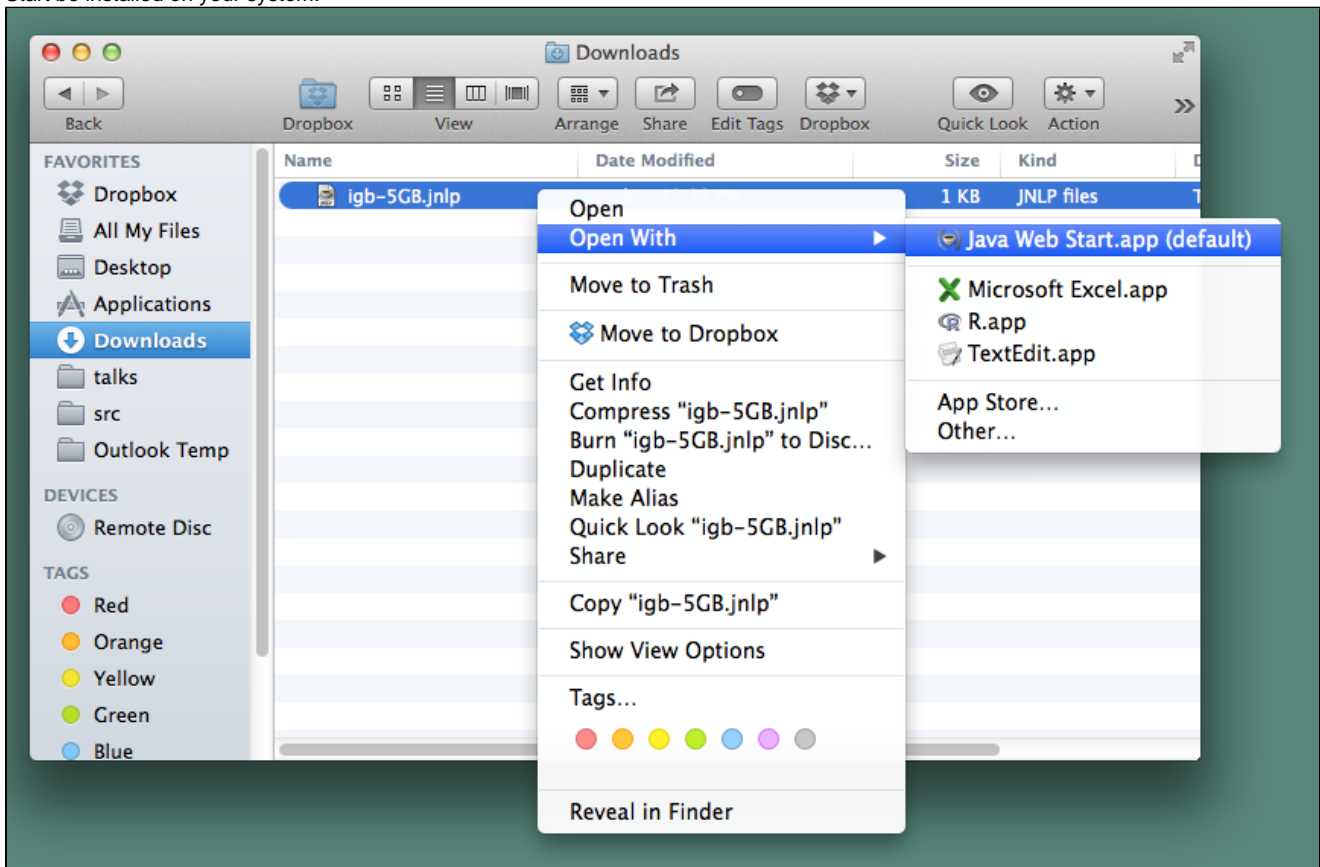
- [Use Java Web Start app to open IGB JNLP file](#)
- [Download igb.zip and run IGB using an IGB script](#)
- [Modify the security settings on your Mac](#)

#### Use Java Web Start app to open IGB JNLP file

When you download or launch IGB from the BioViz.org download page, your browser downloads a special file called a "JNLP" file. If you open this file using Java Web Start.app on your computer, then IGB will download and run.

1. Open your Downloads folder and locate the JNLP file downloaded from BioViz.
2. CNTRL-click (or right-click) the file.
3. Select **Open With > Java Web Start.app**

This is probably the safest option but may be less convenient than the other options listed in the following section. Also it requires that Java Web Start be installed on your system.



#### Download igb.zip and run IGB using an IGB script

One big disadvantage of this workaround is that you won't automatically get newer versions (including bug fix releases) when they become available on the BioViz site. The major benefit is that it nearly always works no matter how your system is configured.

1. Go to the [IGB downloads page](#) and download igb.zip (See the section titled **Other ways to get IGB**)
2. Unzip igb.zip by double-clicking or using "unzip" command in Terminal. This creates a new folder named igb.
3. Open the igb folder and double-click a run\_igb.command script.
4. A new Terminal will open and IGB will launch. Don't close the Terminal window or IGB will shut down without saving any bookmarks you've made or giving the opportunity to save your session.

## Modify the security settings on your Mac

This is least safe option, but the most convenient. Don't do this unless you're confident you can avoid mistakenly running programs downloaded from the Internet.

See the section titled **Disable GateKeeper's Unidentified Application** in [this documentation](#) for instructions on how to modify the security settings on your Mac.

## Java Web Start starts, but fails to launch IGB

The IGB download page uses Java Web Start to download and launch IGB on your computer. The way this works is that your browser downloads what's called a "JNLP" (java network launch protocol) file, which is also sometimes called a "deployment descriptor." When your browser downloads the file, it then launches a second program called Java Web Start on your computer that initiates the IGB download and starts IGB for you. (It may also ask if you would like to create an IGB shortcut on your desktop.)

If Java Web start launches and then fails to start IGB, try the following steps to trouble-shoot:

- **Java Web Start fails with an error like "Cannot grant permissions to unsigned jars."**  
Likely this is caused by a recent bug introduced in OpenJDK. Easiest workaround is to use the Sun JDK for Java Webstart. See [http://sourceforge.net/tracker/?func=detail&aid=3063085&group\\_id=129420&atid=714744](http://sourceforge.net/tracker/?func=detail&aid=3063085&group_id=129420&atid=714744) for more details about this bug in OpenJDK.
- **Under Windows, the Java Web Start version of IGB fails without presenting a window or small loading window flashes on the screen.**  
Reset your browser cache. For details, see: [How do I clear my web browser's cache](#) from Indiana University IT Knowledge Base. This is caused by a bug in a stale copy of the deployment descriptor. One possible cause is that the deployment descriptor requested more memory than 32-bit Windows could allocate and causes Java to crash.
- **Java Web Start fails with an error like "Java virtual machine unable to start".**  
You may have requested more memory for IGB than your machine can handle. Clear your browser cache, close and reopen your browser, and try the "small memory" version of IGB. For details on how to clear the cache, see [How do I clear my web browser's cache](#) from Indiana University IT Knowledge Base.
- **Under Macintosh OS X with Firefox, the Java Web Start version of IGB fails. Error message says "java.lang.UnsupportedClassVersionError: Bad version number in .class file".**  
This is a Firefox bug. Firefox will not correctly launch IGB once the Java Plug-In has been started. Restart Firefox and try again or use a different browser to launch IGB.
- **Under Macintosh OS X 10.8 (Mountain Lion), the Java Web Start version of IGB fails. Error message says "The digital signature is not trusted. Java will not allow any access to this application."**  
With 10.8, Apple has introduced new security feature i.e. Gatekeeper. By default Gatekeeper is enabled. This will not allow to run any apps that has not been download from Mac app store. Here's the link on how to disable it. <http://answers.uchicago.edu/page.php?id=2548>

1

## Release notes for older versions

### New in IGB 6.5

#### Hello and welcome to IGB 6.5!

The following new features debuted in IGB 6.5, released in April 2011.

Most of these new features come from you - the IGB user and developer community.

As always, the IGB team relies on you, the user, to suggest ways we can improve the IGB software. [Please keep sharing your ideas!](#)

#### Sequence Viewer

For the release of IGB 6.5, we have created a **Sequence Viewer**. When you right-click an item in the main IGB data display, you'll see a new option offering to display genomic sequence. Choosing this option opens the **Sequence Viewer**, shown below.

The **Sequence Viewer** can display all six frames of translation (**Show** menu), a spliced cDNA (**Show cDNA** button), or change the color scheme. Exons and introns in gene models appear as different colors, and the viewer also marks start and stop codons, if available. You also have the option to view complementary strand sequence. You can copy and paste sequence or save it to a file using the **File** menu.

IGB 6.5 also makes it easier to work with genomic sequence bases. Just load the sequence in the usual way and then click-drag across the sequence of interest to select it. Right-click to open the **Sequence Viewer** or just press Control-C to copy.

For more information, read [Genome sequence viewer](#).

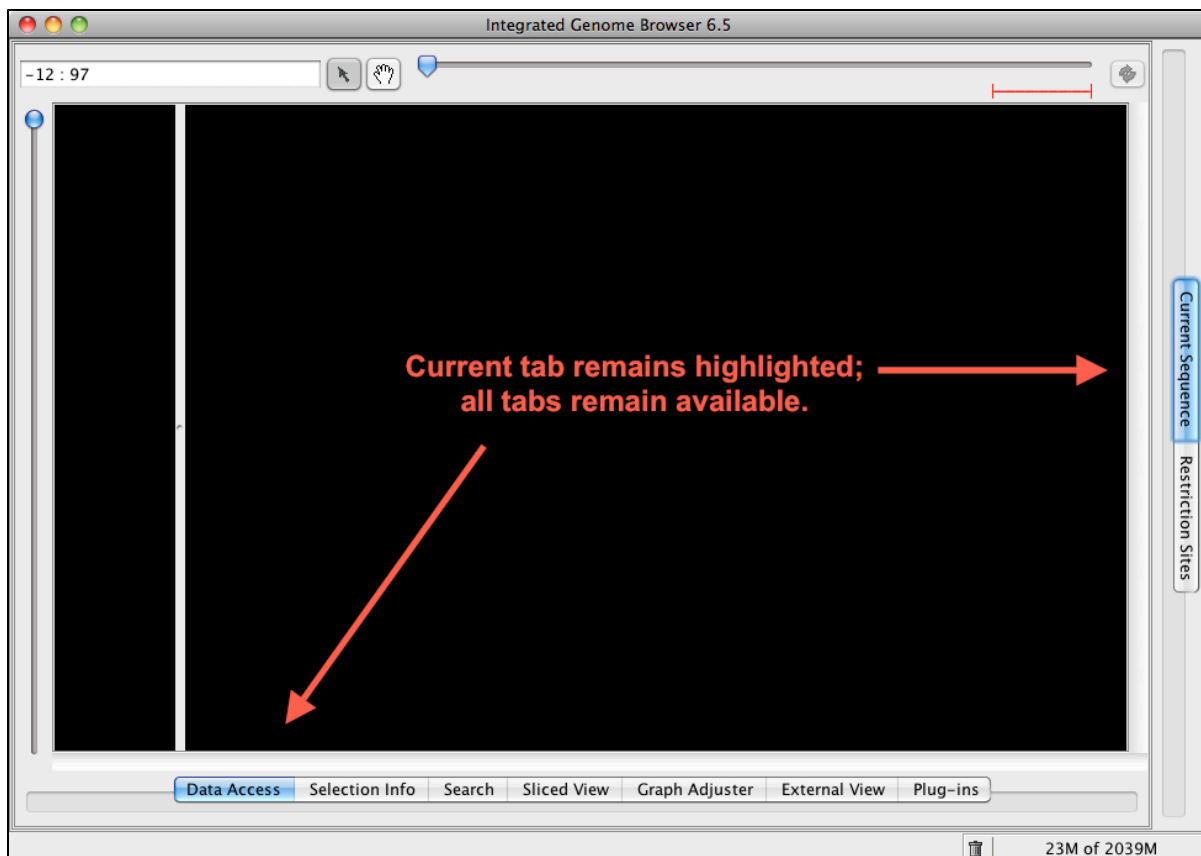
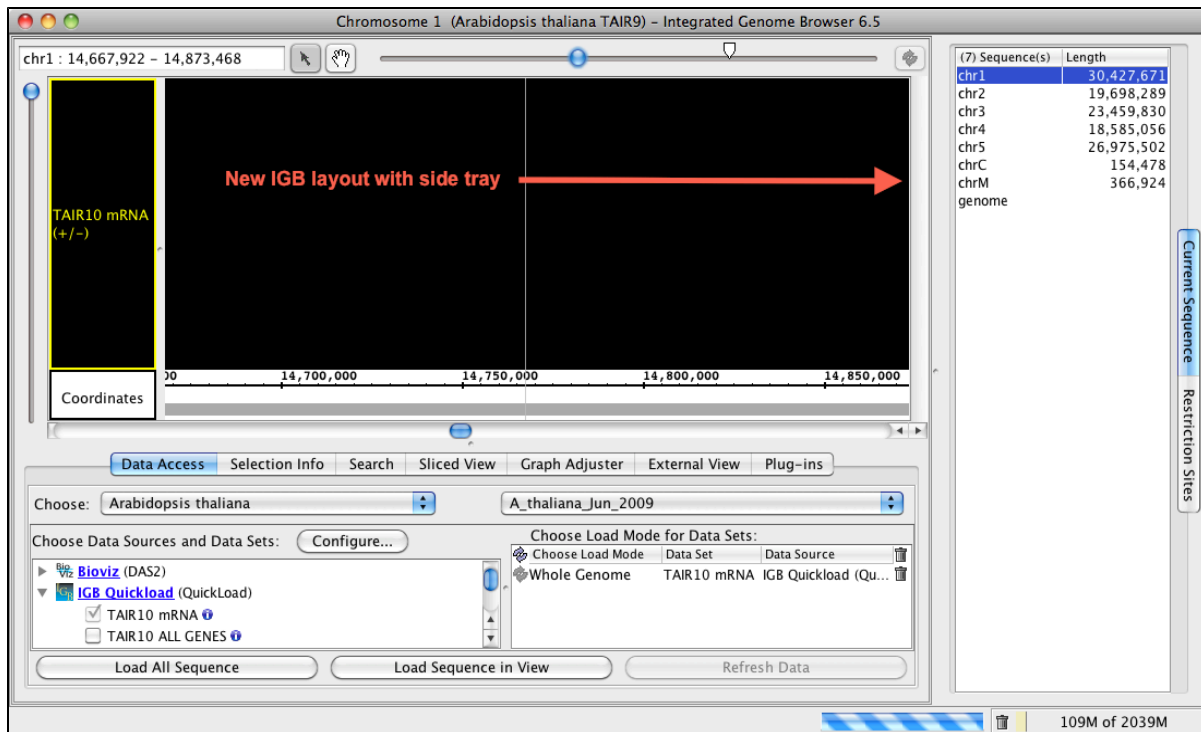


### Tabbed Panels and Tray Layouts

To give you greater flexibility in setting up your personal IGB environment, we have made some changes in the main IGB window. Upon start up, you will now see panels on the right as well as the bottom of the main window. These panels (also called trays) contain all of the usual tabs. The main difference is that you can move them out of the way by clicking the active tab, freeing more space for data exploration.

As always, clicking any tab will open it in the usual way. As with previous versions, you can open tabs in a new window. Now, however, you can move them to left, right, or bottom trays, depending on the tab.

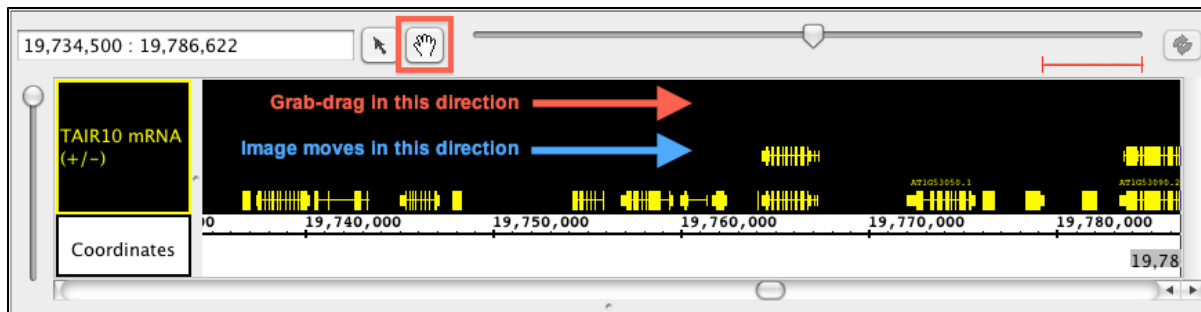
For more information, read [Tabbed panels](#).



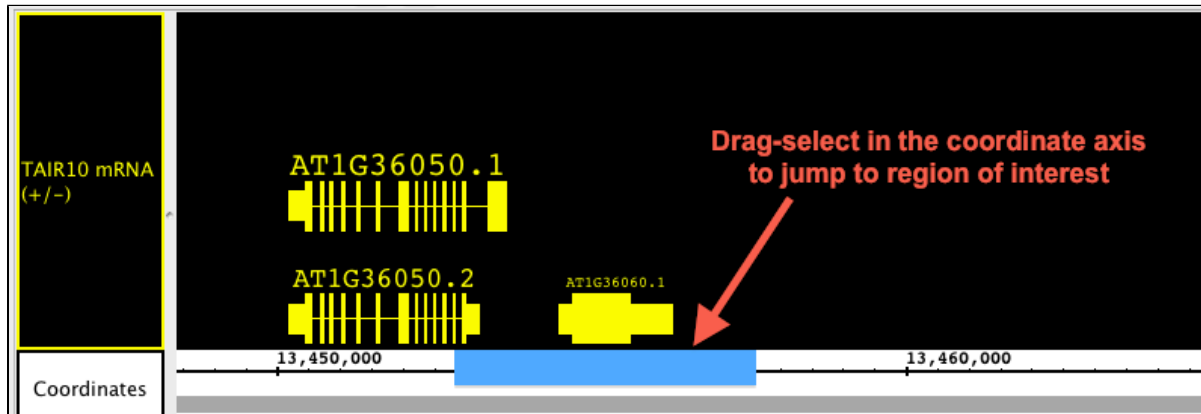
## More Zooming and Panning Options

IGB 6.5 introduces several new ways to pan (scroll) and zoom.

IGB now includes two tool buttons, **Select** and **Grab**, next to the horizontal zoomer. Click the 'hand' button (outlined in red below) to choose the **Grab** tool. Then click and drag the tool on the display to move left or right, and up or down. To select items, choose the **Select** tool (arrow cursor).



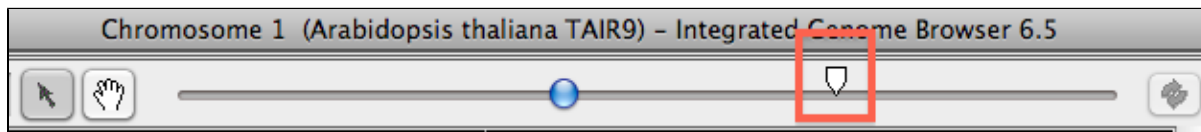
A new zoom feature allows you to "jump zoom" to a region by click-dragging (using the **Select** tool) over the central axis - the number line in the center of the display. Just click-drag over the region of interest, and release. IGB will jump and zoom to the highlighted region.



For more information, read [Panning the Display](#) or [Zooming](#).

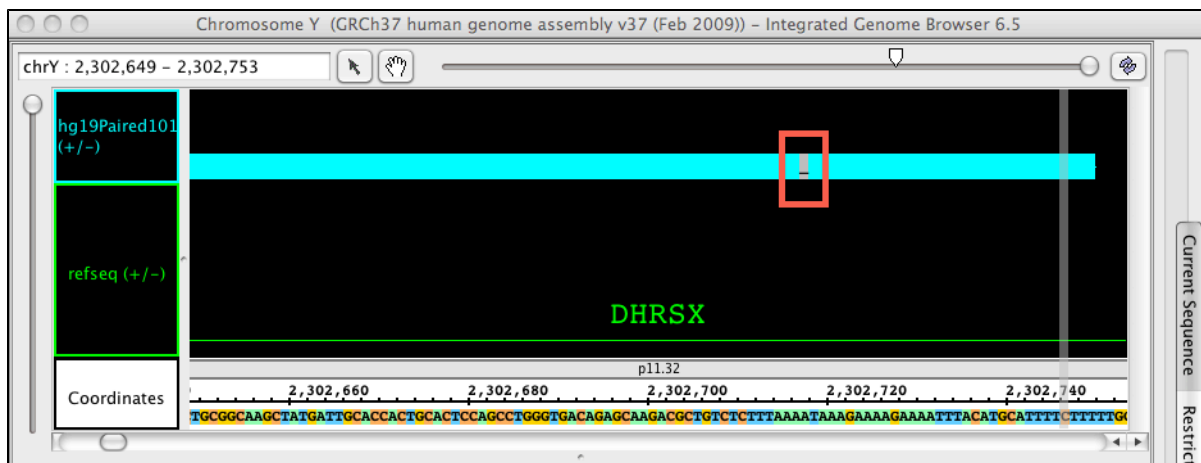
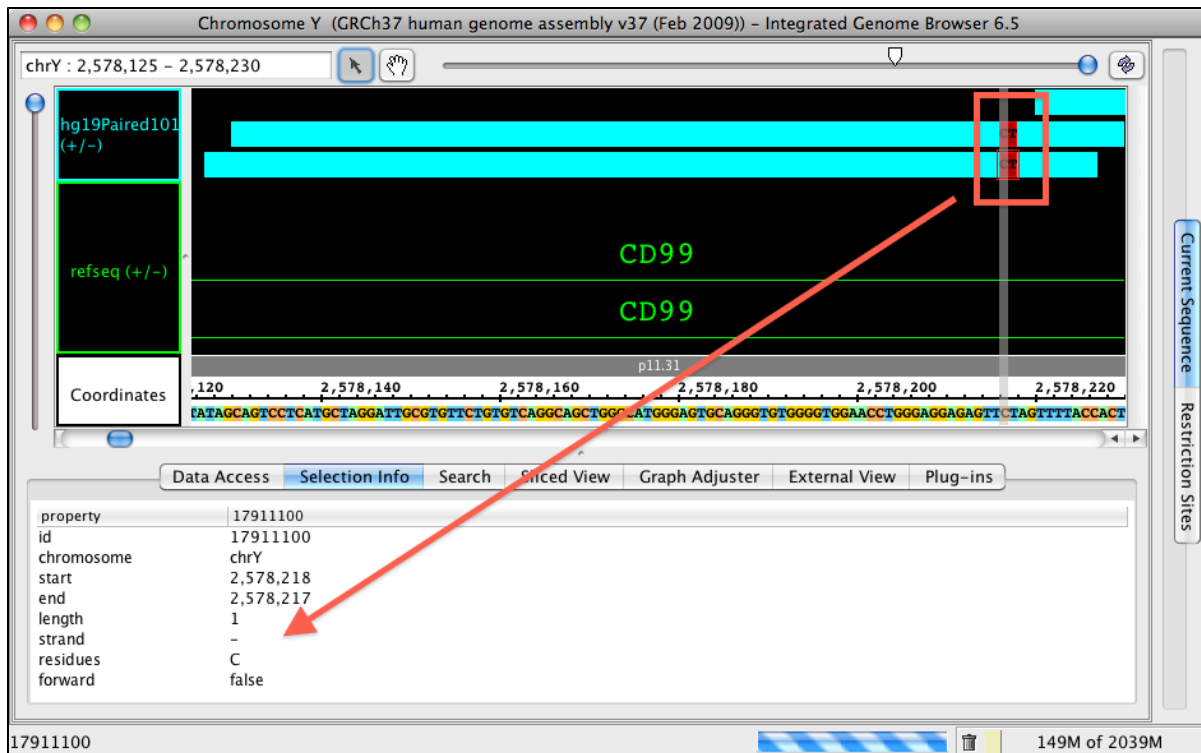
#### Autoload for BAM files

A new optional **Autoload Mode** (for BAM files) tells IGB to automatically load data from a file or URL once you've zoomed in far enough. A marker appears above the horizontal zoomer indicating the zoom level that triggers automatic loading of data.



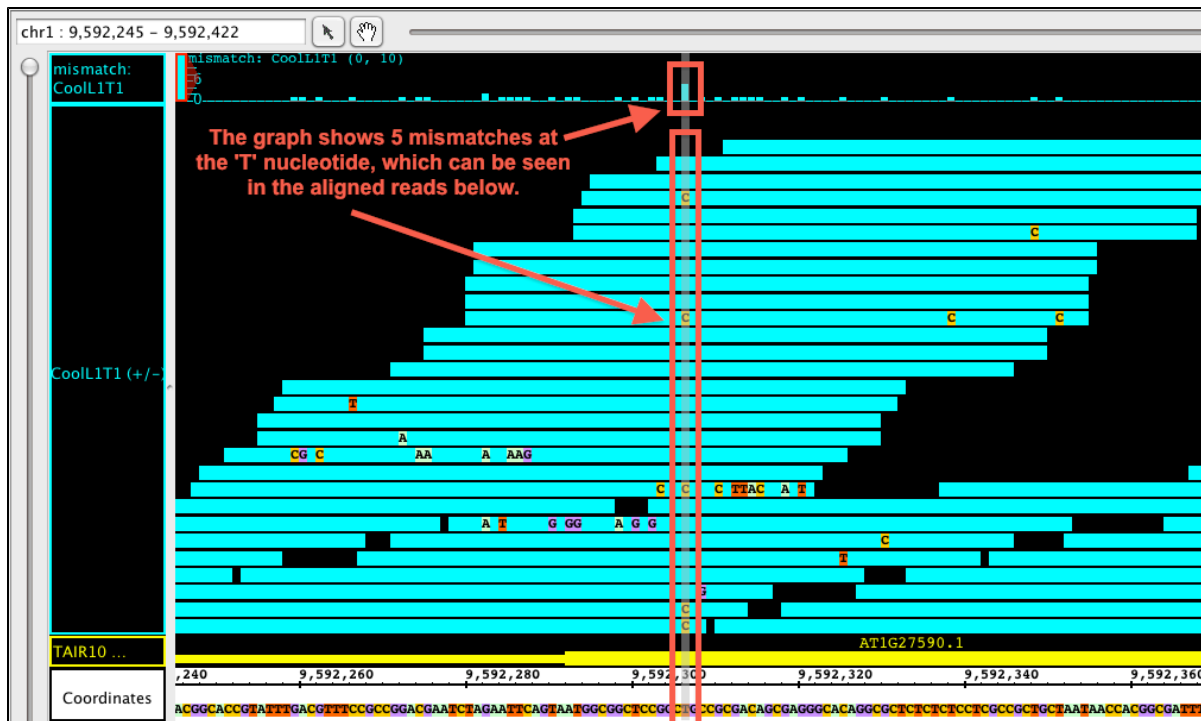
#### Visualize insertions and deletions

Another feature is that IGB can now show both insertions and deletions in short reads. Insertions appear as a pair of red residues, indicating the point of insertion. Details about the insertion can be gotten from the **Selection Info** panel or by hovering the mouse over the red residues to activate the tooltip; this information contains the length and sequence of the insertion. Deletions appear as a 'dash', highlighted by default as gray; by opening **File > Preferences > Other Options**, you can set the 'other nucleotides' to appear as any color you like. This will highlight the deletions, as well as unloaded sequence, or 'N' reads in your selected color.



## Create a mismatch graph

IGB now provides the opportunity to create a mismatch graph from BAM files. This bar graph shows how many mismatched reads are present at each nucleotide. You simply select the track, right click in the label and open the menu; you are now offered the option of **Make Mismatch Graph**, along with the regular option of **Make Annotation Depth Graph**. By hovering over a position in the mismatch graph, a tooltip with information about nucleotide distribution will appear. This allows a user to quickly and easily find out how many mismatched nucleotides in total, as well as how many of which type, are represented in the reads at any single genomic location. This can be very useful in SNP identification and allelic determination. For more information on making either graph type, read [Creating graph tracks from annotations tracks](#).



## Plug-Ins

We have instituted the ability for IGB to accept plug-ins, allowing developers and users to create small 'add in' programs that customize the functionality of IGB. Each plug-in can perform a function or series of functions for a specific purpose. By selecting the right plug-ins to add, the user expands the usefulness of IGB and tailors it to their needs, without cluttering IGB with unnecessary functions. Some plug-ins, such as **Invers** **eTransformer** and **PowerTransformer** are already available in the IGB program. These can be activated from the **Plug-ins** tab. The IGB developers will continue to author and make available some plug-ins; users can also write and use their own plug-ins, suited to their needs.

Data Access   Selection Info   Search   Sliced View   Graph Adjuster   External View   Plug-ins			
Name	Description	Version	Installed
InverseTransformer	Graph Transformer that takes the inverse of the input - $1/x$	1.0.0	<input type="checkbox"/>
PowerTransformer	Graph Transformer that takes the input to a given power - $x^a$	1.0.0	<input type="checkbox"/>

☒ Show Installed   
 ☒ Show Uninstalled   
 Update all   
 Update selected   
 Repositories...

## Expanded Support for Plant Genomics

IGB continues to allow access to publicly available DAS servers that support some plant genomes, but with version 6.5, the IGB team debuted a new IGB QuickLoad site with several sequenced plant or algal genomes not previously supported. We also moved the site to a faster server with improved internet connectivity, which should make your experience with IGB feel faster and more responsive.

To view the new IGB QuickLoad site, and see the full list of genomes we directly host, visit <http://www.igbquickload.org/quickload>.

## New in IGB 6.6

### New Features for 6.6

Welcome to IGB 6.6! (Did you see our new welcome screen?) We have made many changes and improvements to make IGB even easier for users. Many changes, such as the Data Panel redesign and the Preference tab Changes are immediately obvious, but there are many more new and exciting functions you should explore.

### Redesign of the Data Access Panel

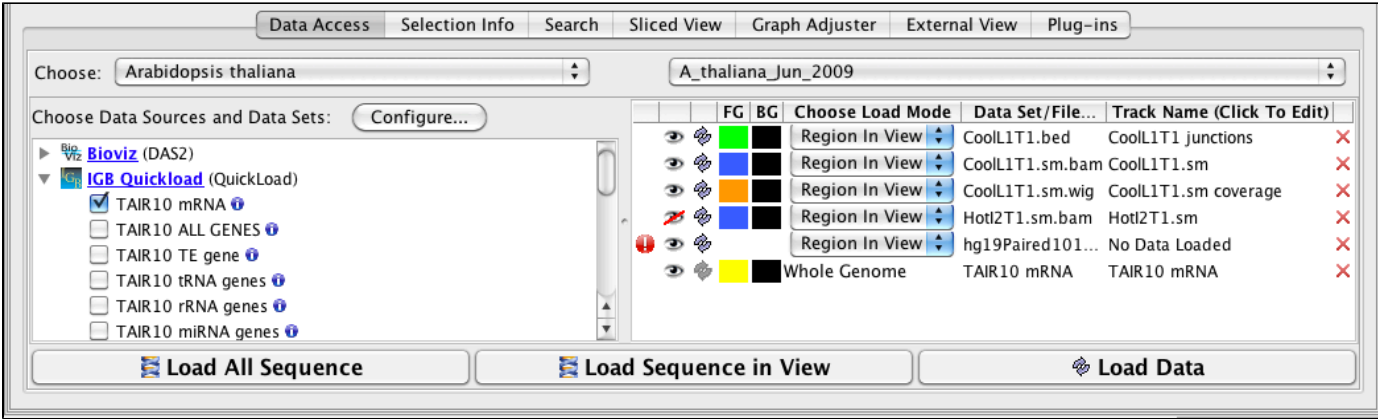


The **Data Access** panel has been redesigned to allow users to easily change the foreground and background colors, plus change the name of the track. The **Delete** icon (trashcan) and the **Reload** icon (arrows) are joined by the **Hide** icon (open eye); **Hide** can be used to hide tracks from view in the main window without deleting the track. A new warning icon indicates when your more recent **Load Data** request failed to add anything to the display.

More subtle is a new feature affecting the Data Sources tree; you can now remove a track in the display by clicking its checkbox in the **Data sources and Data Sets** panel.

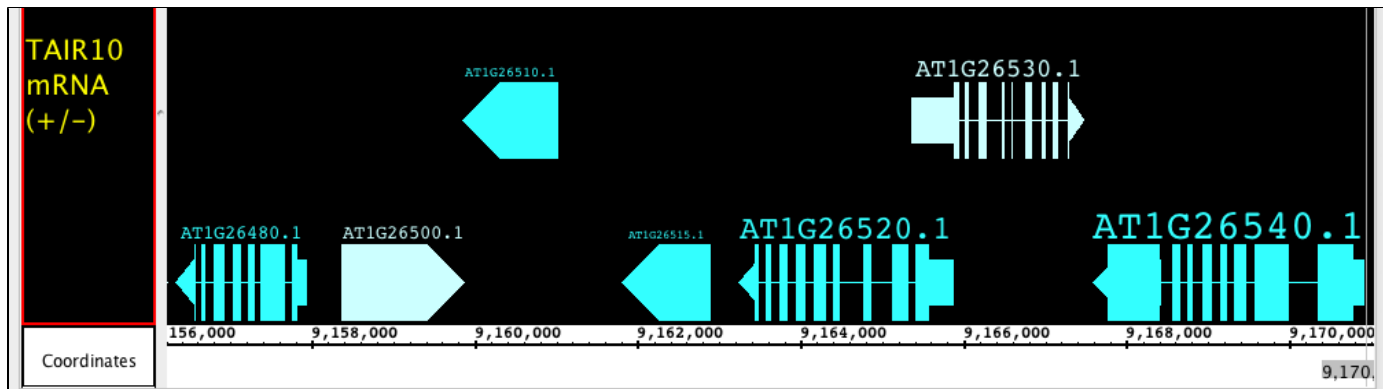
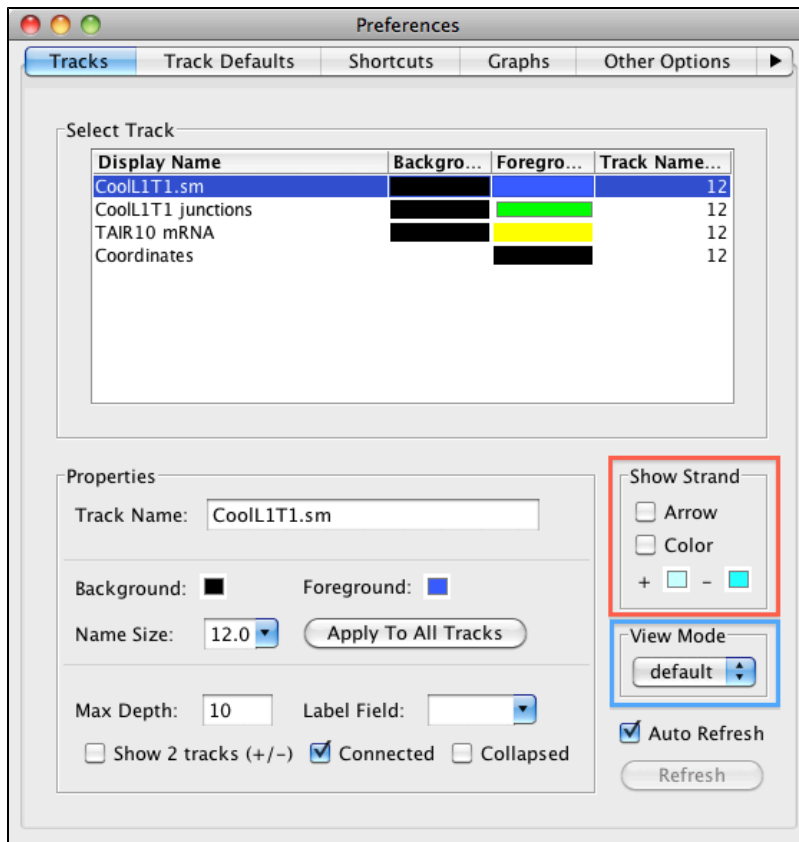
if loading a file is taking too long (or it might fill the memory) a user can stop the loading process by using the **X** box at the bottom of the main window.

Here is a close-up of the redesigned **Data Access** panel showing the 'No Data!' alert, the **Hide** function activated (red line through eye), the color change options, and the now-editable **Track Name** column.



### New Options for Changing Track Appearance

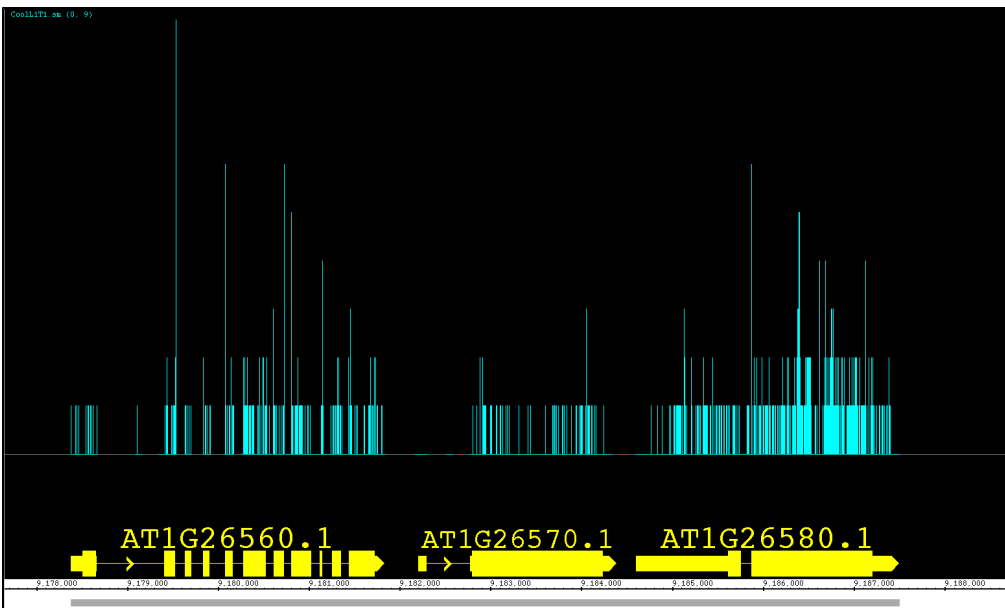
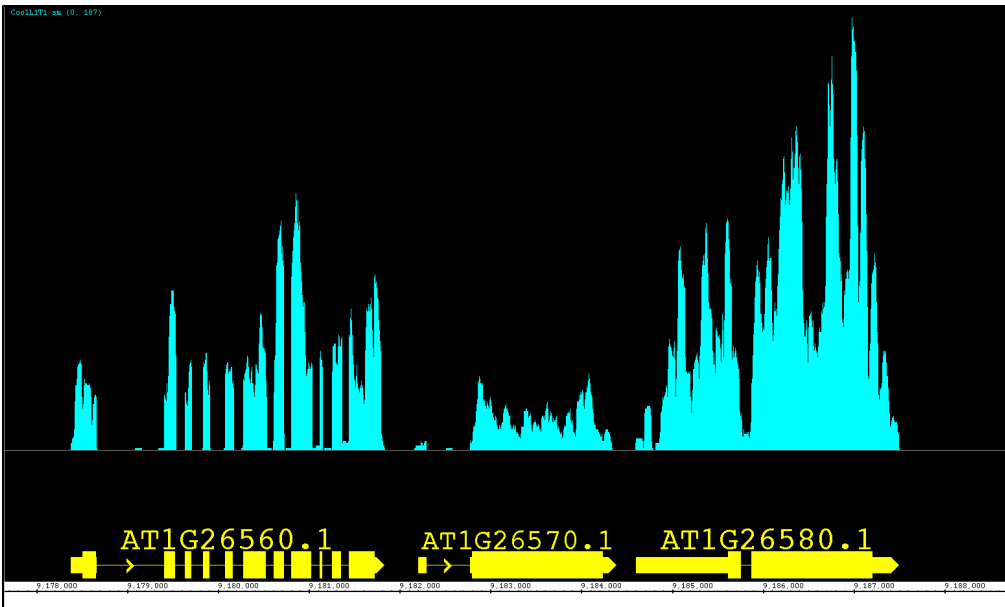
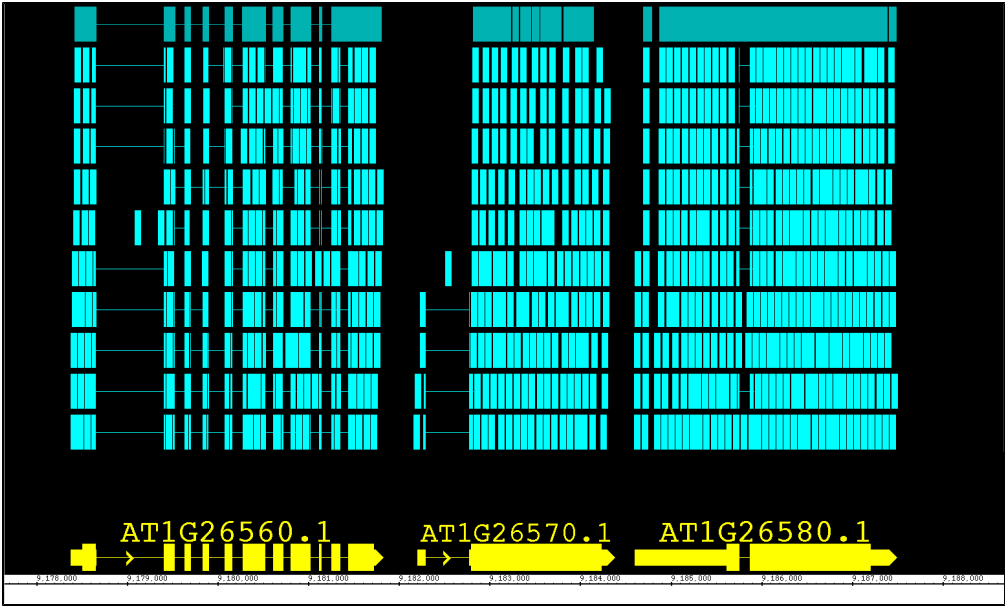
Including the redesign of the Data Access panel, we have made changing the style of a track much simpler and we have expanded the track appearance options. Track parameters can be changed through the right-click on the label menu (current method), **Data Access** panel (new) or the newly enhanced **Tracks** tab in the **Preferences** panel. The **Tracks** tab now has many great features to help you adjust the presentation of the tracks. As always, colors, name, track depth and label field can be accessed, but in addition, users can now change the font size of labels, or highlight strand differences (see red box in the image below). This is a new set of feature which allows the user to show direction of gene models. The options include **Arrow**, which uses an arrow head to indicate direction of transcription, and **Color**, which uses 2 colors (can be changed) to indicate + or - strand. This feature is an excellent way to show strandedness when + and - strands are shown as a single track (see second image).

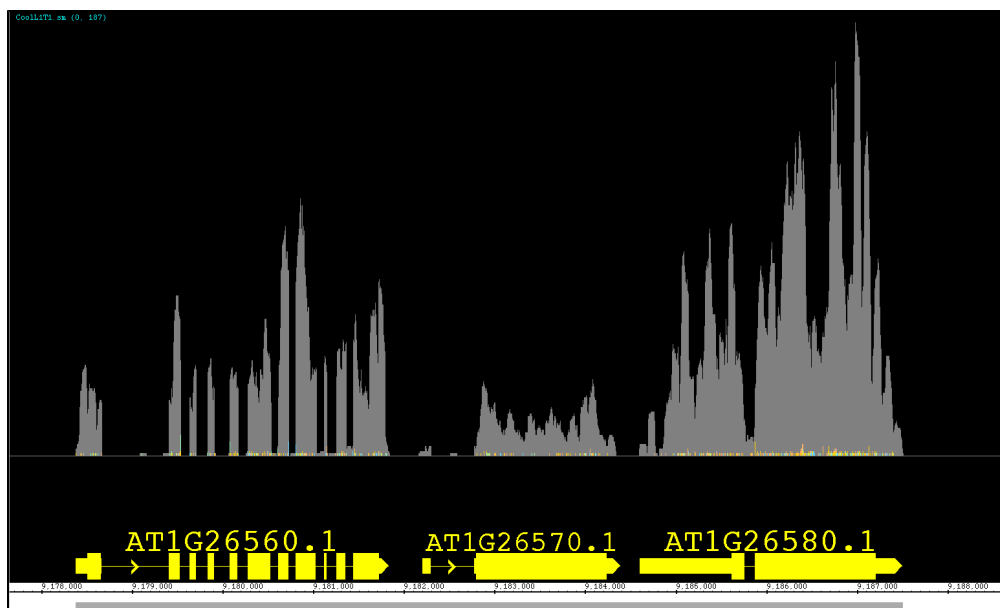


### Prototype: View Mode

We have been developing a fantastic new feature called **View Mode**. We have been so pleased with this feature we are releasing it as a prototype with version 6.6. This feature allows you to get more information out of your annotation tracks without having to create several independent graphs! Accessible from the right-click menu or the **Preferences > Tracks** tab, you can set an annotation track to show **depth**, as if you created a depth graph. For BAM files you can change the **View Mode** to show **mismatches** as if it were a mismatch graph. We have also introduced a new view (also available as a **Create Graph** option) called **Mismatch Pileup**; this new function highlights mismatches (color coordinated to each nucleotide) overlayed on a dark gray depth graph, giving a lot of information compacted into a single track. Unlike graphs, you can readily change the view options at will, getting graphical information and returning to a default representation if you choose. As we continue to expand this feature, it will gain more options for adjusting the graphical display and possibly even additional views! Feel free to [contact us with your ideas](#).

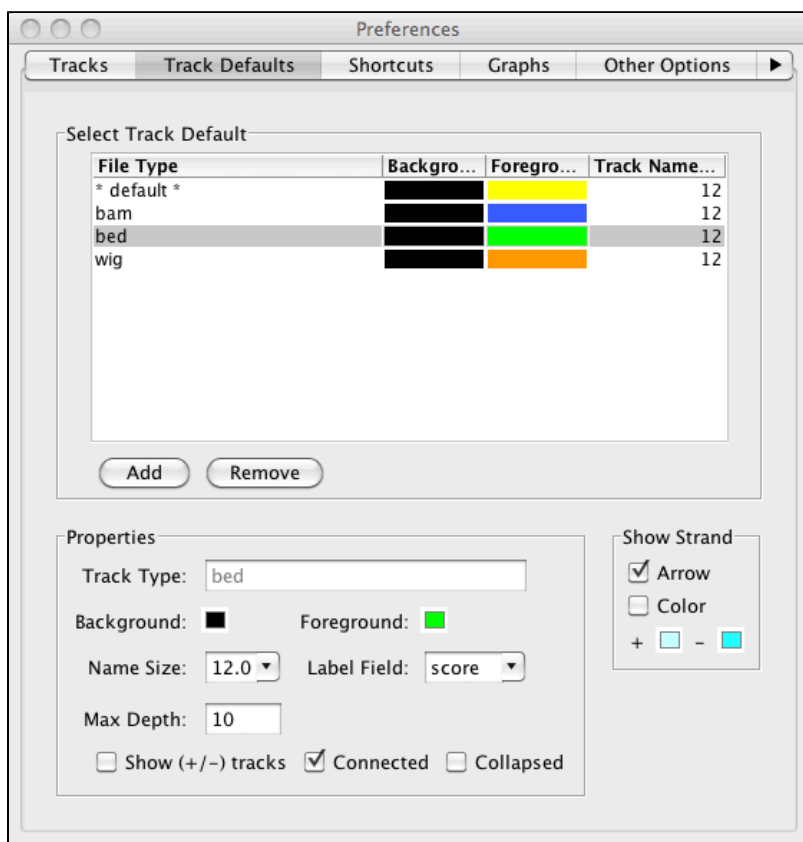
The pictures below show Default view, Depth view, Mismatch view and Mismatch Pileup view.





## Improved Preference Tabs

Almost all of our **Preferences** tabs have been redesigned to make easier to use. This includes the incredibly expanded functions of the **Tracks** tab (see *Track Appearance* below for details), the simplification of the **Other Options** tab, and the addition of source ordering buttons in the **Data Sources** (see *Faster Performance* below for details). We have also created the entirely new **Track Defaults** tab where you can now set the default colors, font size, label, max depth and strand determination for different annotation file formats. You can also set the colors and label size for different graph file types.



## New Features for the Main View

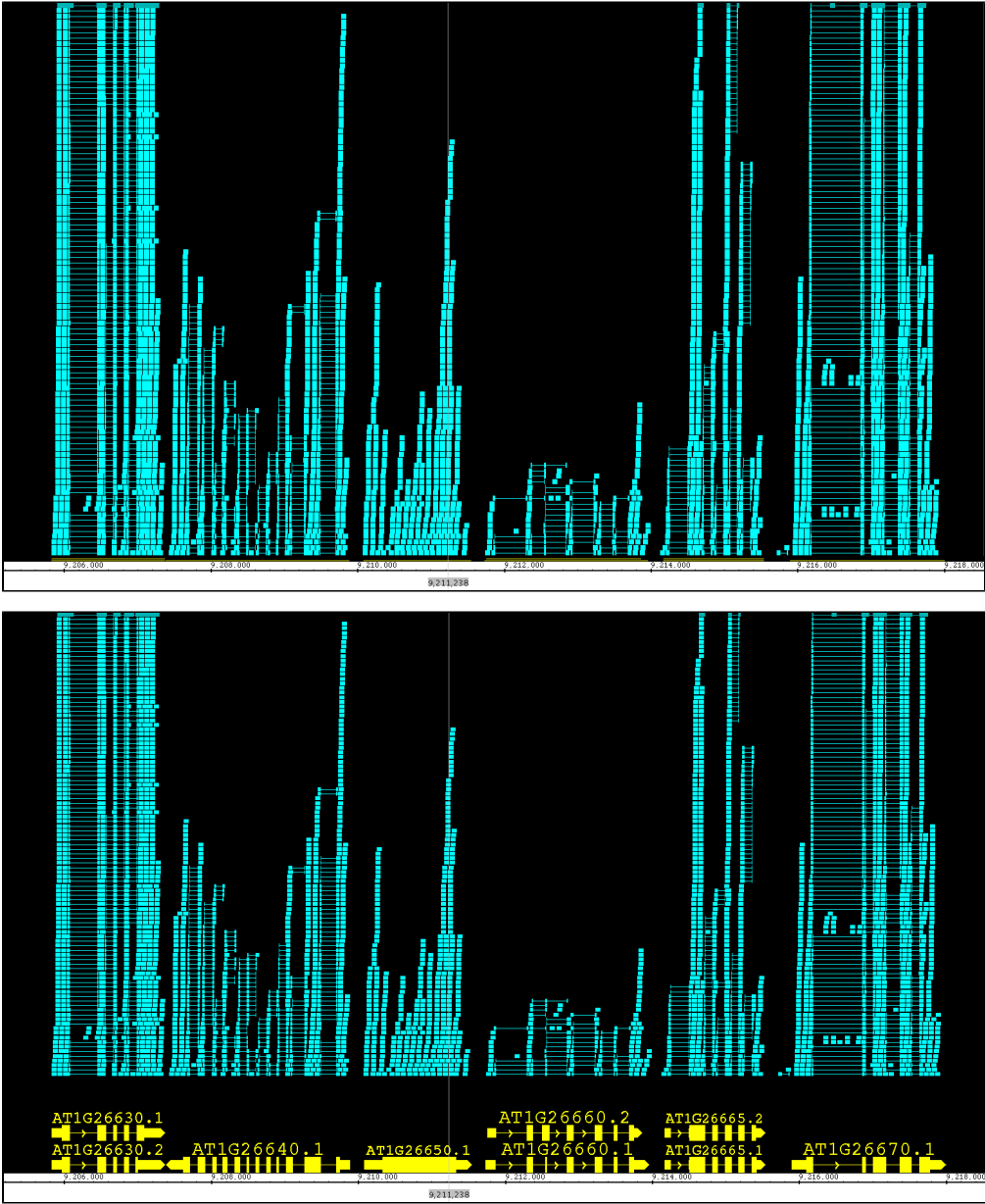
There are many new and improved features to enhance the main view in IGB. We have made the **Zoom Stripe** translucent so that you can easily see your data at the same time you can see point of focus, but both the **Zoom Stripe Label** and the **Zoom Stripe** itself can be turned off from the

**View** menu. Another fantastic feature accessible from the right-click menu is the **Focus On Selected Track** option, which will allow the user to 'zoom in' and fill the viewable space with the track of interest. We have also changed the sequence to a dynamic font which both allows you to zoom in on the track just a little more and increases visibility of each nucleotide. By installing the **MergeAnnotationOperator Plugin**, you can now join annotations into a single track, just as the **Join** function works for graphs.

**Prototype: Track Height Adjustment**

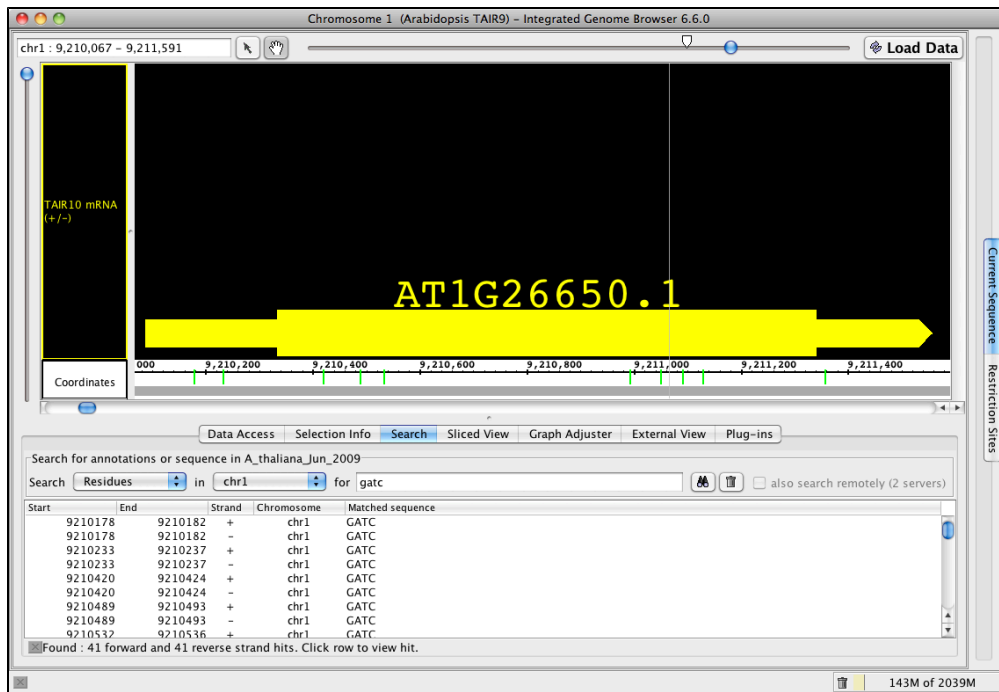
Here is another experimental function that was just too useful to leave out! You can now adjust the track height for annotation and graph tracks. Move the cursor over the boundary between two track labels until the height adjustment cursor icon appears. Click and drag to the new height. As this feature is still under development, it may exhibit some quirky behavior, but we found it to be so useful overall that we included it with this release. Look for this feature to be even better in future releases!

In the pictures below you can see a common situation, where a short reads file (BAM) overwhelms the reference sequence models. In the second picture, you can see the image readjusted by simply increasing the height of the ref seq track.



**Improvements to Search**

Now when using the **Search** tab for sequences, IGB provides a tabular output in the **Search** panel as well as more visible (green) hash marks on the sequence bar. A search that might be running too long or was incorrectly entered can be stopped using the **X** box below table.



## New Features for Graphs

Our new **Mismatch Pile-up** (described in **View Mode**) graph shows both the **Depth AND Mismatch** graphs as one over-layered graph, and is readily adjusted by using the **Graph Adjuster** tab. We also have redesigned the Manipulations portion of our Graph Adjuster tab to accommodate the functions provided by Plugins. Three new Plugin functions include **Geometric Mean**, **Power Transformation** and **Inverse Transformation**; after installing from the **Plugins** tab, these functions can be found in the **Transform** drop down box.

## New File Types, Faster Performance

With more and more data being present in more and more formats, we have expanded the types of files IGB can accept; users can load SAM files (with or without matching index files), fasta files that have been indexed and traditional BED, PSL, SAM and WIG that have been indexed using tabix. After installing the appropriate **Plugins**, IGB can additionally use BigBed, BigWig and VCF formats. We have made many behind the scenes changes to speed up file handling, but we now also permit the user to prioritize the order in which IGB queries servers for data; this can allow IGB to preferentially go to faster servers first OR to get sequences from a preferred server (**Preference > Data Sources**).

## Improved Bookmarks

We have added two new features to **Bookmarks**. We have activated tooltips for bookmarks, so that you can find the chromosome and nucleotide range for each bookmark without needing to open it. We have also introduced a **History** function; using the 'fast forward' and the 'rewind' buttons in the **Bookmarks** tab, you can scan back and forth through your bookmark browsing history, letting you go back for a second look, without losing your place.

## New in IGB 6.7

### New Features in IGB 6.7

Welcome to IGB 6.7. We have made many changes and improvements to make IGB even easier to use.

Many changes, such as moving the **Species** and **Genome Version** to **Current Sequences** tab and the redesigned **Bookmarks** tab will be immediately obvious to IGB veterans, but there are many more new and exciting functions you should explore in IGB 6.7.

Major thanks are due to IGB software developer Hiral Vora, who led the effort and kept us all on track for the current release. Thank you Hiral!

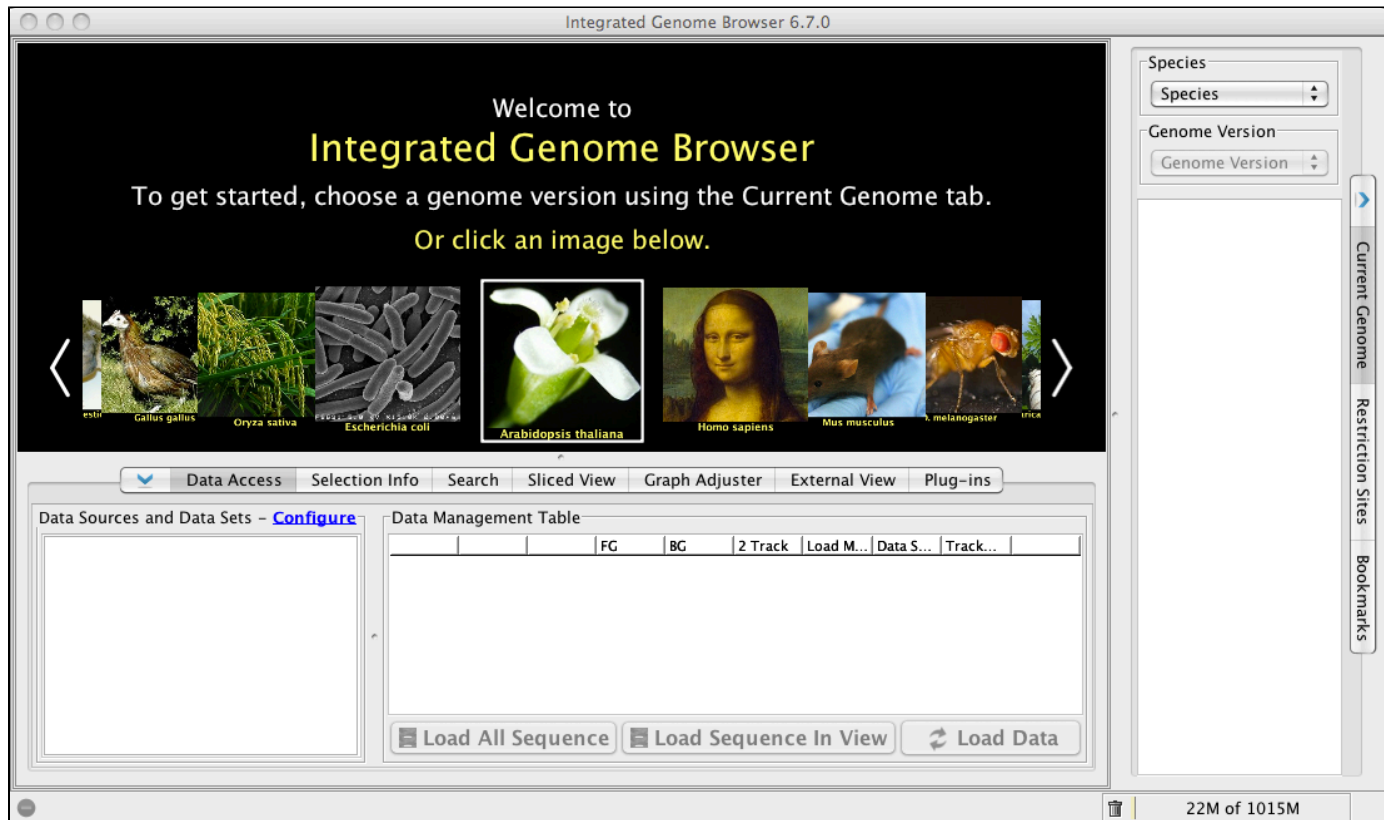
### Panels and tabs redesigned

To make IGB easier to use, we re-configured the user interface for several IGB tabs and panels including

- **Bookmarks** tab
- the **Export File** window
- the **Web Links Tools** interface.

We also re-designed the IGB start screen. Now, when you start IGB for the first time, you'll see a new start screen that provides quick and easy access to a variety of model organisms, helping new users jump right in.

Credit for this new interface go to Dr. Jeremy Villalobos and Graduate Research Assistant David Norris.



### Better Bookmarks tab - and support for comments

The **Bookmarks** tab now supports adding comments to bookmarks and bookmark folders.

When you create a new bookmark, IGB will give you the option to add a free text comment. (As before, you can name bookmarks whatever you like or just use the IGB-supplied default.)

The **Add Bookmark** function offers the option of making one of two types of bookmarks:

- a **Position Only** bookmark which records genomic location
- or a **Position and Data** bookmark, which also records data sets currently being displayed

Returning to a Position Only bookmark returns you to the bookmarked genomic position but won't load data.

Returning to a **Position and Data** bookmark returns you to the bookmarked genomic position and also re-loads data sets that were in the original scene (provided they are available).

Right clicking a bookmark will open a new window displaying bookmark attributes.

Thank you to Graduate Research Assistants Zhong Ren and David Norris for programming and also to Dr. Alyssa Gullledge for design and usability evaluation.

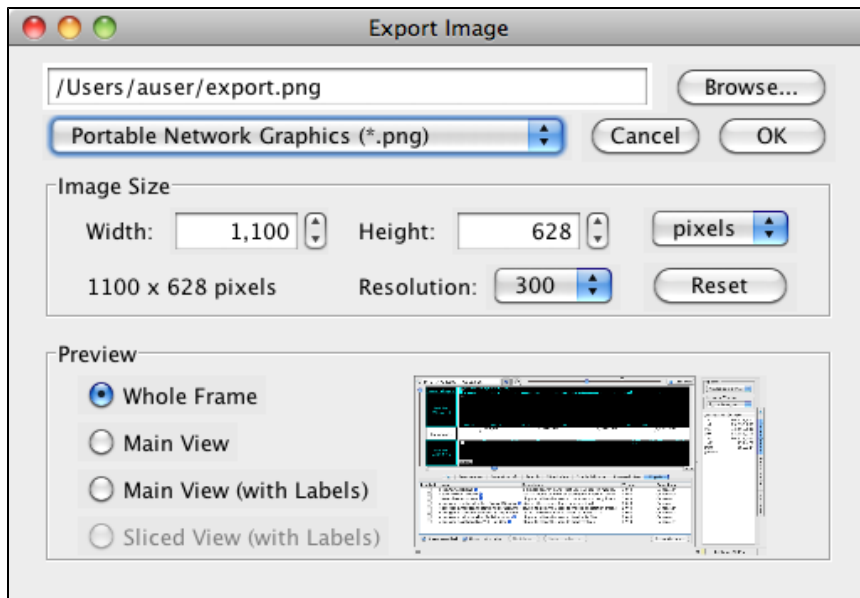
### New options for exporting image files

Along with a redesign to improve ease of use, the **Export** interface also features new options not previously available. Users can now specify image size and file type for export. You can still specify which part of the IGB display you'd like to export, e.g. the main display image, the main image with the track labels, or the entire window with tabbed panels (such as **Selection Info**) showing. IGB can currently export images in PNG or JPEG formats.

We are also added a new feature that allows you to set image resolution/DPI. IGB now makes it easier to create high-quality images of 300 dpi or better (the minimum required for images destined for publication in printed journals).

**Note to current users:** Unfortunately, we had to remove export to PDF and other vector graphics formats. There were a number of show-stopping bugs in the external libraries we were using to support export in those formats. We are looking for new libraries that will be able to support those formats and hope to aim this functionality in the next release.

Thank you to Zhong Ren (CS Masters student) for programming the Export Image window.



### Autoload for features that use sequence data

A number of IGB features that require access to sequence data, such as **Restriction Site** mapping, the **Sequence Viewer** for gene models and **Search Residues** functions. Previously, IGB would ask you before loading sequence data. Starting with IGB 6.7, IGB will automatically load sequence data from available IGB QuickLoad or DAS sites, speeding the process of retrieving information you need for analysis and visualization. However, you can still **Load Sequence in View** or **Load All Sequence** information from the **Data Access** panel.

**Tip:** You can control which server IGB will use for accessing sequence data by changing the ordering of **Data Sources** in the **Data Sources** Preferences tab. Data providers differ on whether it is better to deliver masked versus unmasked genomic sequence to users, and so we decided to give users the ability to control which data source provides the reference genome sequence. Thanks are due to developer Lance Frohman (Genentech) who added this feature.

### Expanded residue Search function displays multiple sequences

With IGB 6.7, we have introduced the ability to overlay multiple sequence **Search** results (up to 7 different searches).

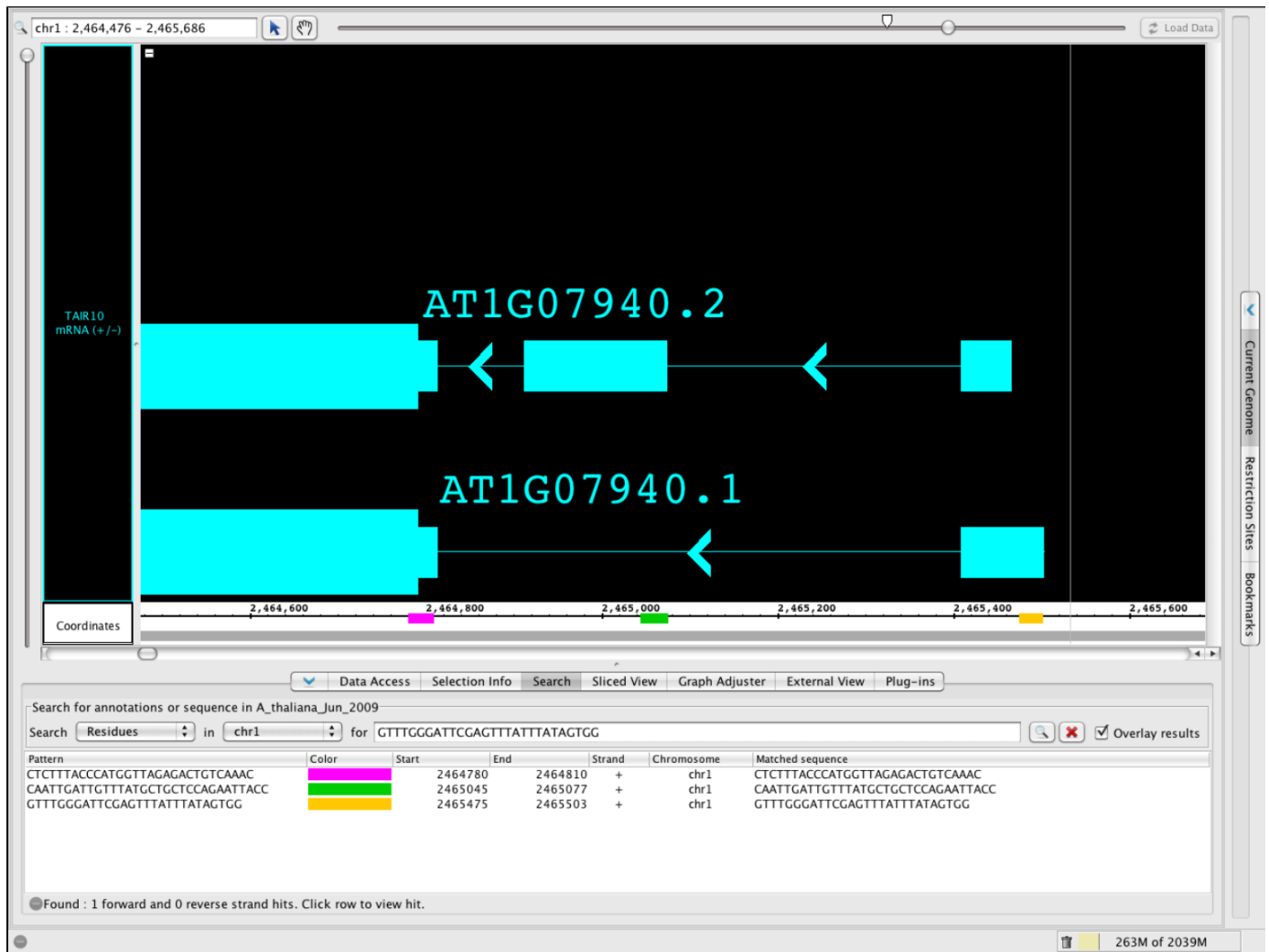
The basic **Search Residue** function hasn't changed; as before, you type in the sequence you are searching for and IGB returns a list of 'hits' and adds colored bars to the sequence indicating the location of matches.

However, if you check the new **Overlay** option, subsequent residue searches will be added to the list and highlighted in new colors without removing previous search results.

Use this new feature when you need to display multiple sequence features in the same view, such as forward and reverse primers for PCR primers, transcription factor binding sites, and so on.

Thanks are due to Dr. Alyssa Gullledge for designing and testing this new feature.





This image shows the 3 primers needed for determining splice variant choice for the AT1G07940 gene of Arabidopsis thaliana, all shown in the same view.

## Expanded file types and increased save options

We are always expanding our ability to support new file types. With IGB 6.7, we can now support [BedDetail files](#) and [tabix indexed GFF files](#). If there is a file type we don't support yet, let us know! IGB has allowed users to save annotation tracks (including tracks created through **Track Operations**) as .bed files. Now, IGB can now save graph tracks (including tracks generated with IGB tools in the **Graph Adjuster** panel!) as .bedgraph files.

## NEW in IGB 7.0

- [Welcome to IGB 7.0](#)
  - [New features and improvements - a short list](#)
- IGB 7.0 highlights
  - [Intron arrows in gene models](#)
  - [Protein translations for BED Detail files](#)
  - [Search tips](#)
  - [Selection Info textbox](#)
  - [Saving and loading sessions](#)
  - [Saving tracks and annotations](#)
  - [More control over Color by Score](#)
  - [Features for creating publication-quality images](#)
  - [New and improved tabbed panels](#)
    - [Data Access panel](#)
    - [Annotation panel](#)
    - [Graph panel](#)
    - [Selection Info panel](#)
  - [New Tab Menu](#)
  - [Toolbar shortcuts](#)
  - [Locking and setting track height \(in pixels\)](#)

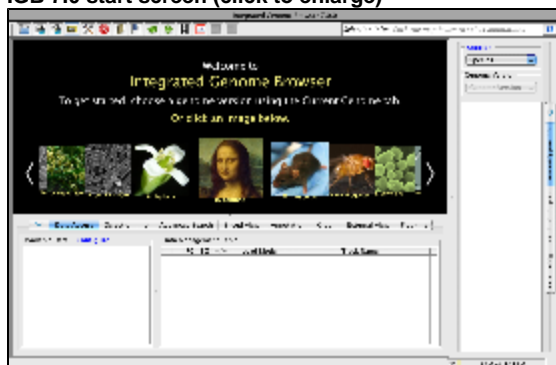
- [Stack Height shortcuts](#)
- [Auto Scroll Updates](#)
- [New genomes supported in IGBQuickLoad](#)
  - [Retiring Affymetrix DAS2 data source in progress](#)
- [New features for QuickLoad sites](#)
- [Web Link link-out builder tool improved](#)

## Welcome to IGB 7.0

We are excited to bring you many new and improved features with the release of IGB 7.0.

As always, we strive to improve all aspects of IGB, including functionality, usability, and connectivity.

### IGB 7.0 start screen (click to enlarge)



## New features and improvements - a short list

IGB 7.0 has includes new features for data analysis and exploration, including:

- Better-looking, more informative gene models
  - Protein translation display when zoomed in
  - Better indicators for strands (intron arrows)
- Easier, faster searching to help you find genes by gene name or keyword
- New **Sequence Viewer** features, including the ability to export conceptual translations of gene models or other features
- A new **Selection Info** textbox that reports the name or number of selected features
- A way to save and load IGB sessions, which helps you save your place in a long-term analysis or share your analysis with others
- Expanded **Save Track** features, including the ability to save selected annotations and/or tracks as BED detail file format
  - Support for BED detail format (replaces Affymetrix "BRS" and "BGN" formats)
- More control over the numerical range of the **Color by Score** feature
- Several improvements to help you make publication quality images

We redesigned the UI to make IGB easier to use for advanced and beginning IGB users. New features include:

- A new **Annotation** tab for configuring and styling annotation tracks (reads, gene models, ESTs, mRNAs)
- A new **Toolbar** allows users to customize their IGB environment with one click access to favorite shortcut functions
- Enhanced **Stack Height** and **Track Height** options give more control over the amount and size of data to be shown
- Improved **Auto Scroll** feature makes scanning through large amounts of data smoother and easier

IGB was designed from the start to allow sharing and accessing data sets from diverse providers. IGB 7.0 has new features to promote and enable data integration, including:

- Many reference genomes data sets updated to Nov 2012 versions (replacing many out-of-date Affymetrix NetAffx DAS data sets)
  - New plant species add to the IGB data source IGBQuickLoad.org
- Continued support and improvements to QuickLoad sites
- Improved **Configure Web Links** tool for building customized IGB linkout patterns, useful for data providers setting up their own QuickLoad sites

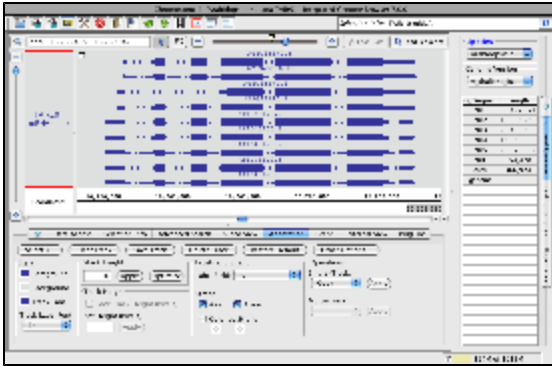
See below for highlights and images showing the new IGB 7.0 interface.

## IGB 7.0 highlights

### Intron arrows in gene models

To make it easier to identify strand of origin for gene models, we augmented the Arrow strand style to include arrow marks along the entire length of introns. (This change was inspired in part by UCSC Genome Browser, which uses greater-than symbols to indicate strand.)

### New arrow strand style (click to enlarge)



## Protein translations for BED Detail files

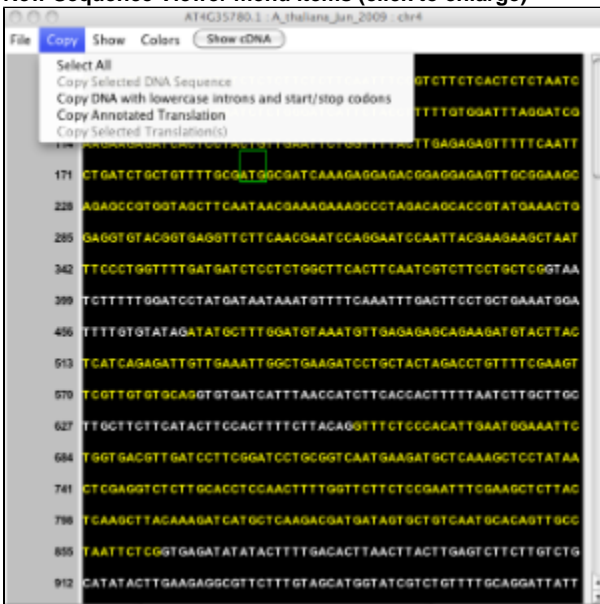
We have expanded the ability to include protein translation during analysis in two ways. In the main view, users see translations overlaid on gene models, where amino acids are either displayed as single letter or three letters (or hidden, if you prefer). Simply click **Load Sequence** and zoom in on the gene; the amino acid sequence will appear according to the annotated start and stop of the gene model.

Close-up on a gene model (click to enlarge)



The **Sequence Viewer** features new options for copying gene model conceptual translations and intron/exon sequences.

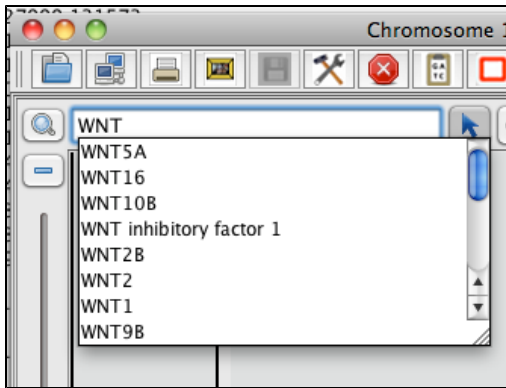
New Sequence Viewer menu items (click to enlarge)



## Search tips

IGB 7.0 introduces search tips. Start typing in either the Quick Search/Coordinates display (top left) or in the **Advanced Search** tab and IGB will suggest search terms.

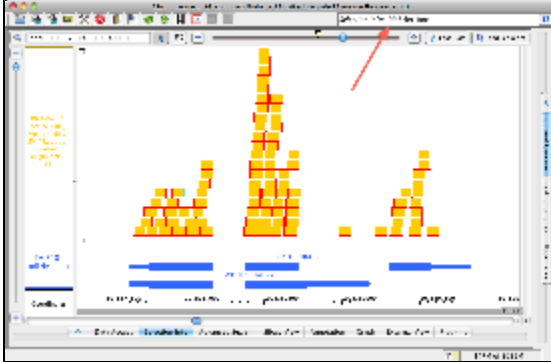
Search tips



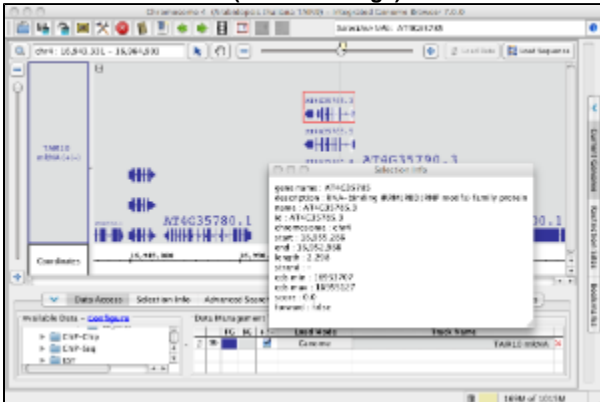
### Selection Info textbox

In addition to the added copy functions in the **Selection Info** tab, we have also added a new **Selection Info** textbox (top right), which reports the name of a currently selected item or the number of items if more than one is selected. If you have selected a region of genomic sequence, the number of selected bases pairs is shown. Clicking the "i" icon opens a new window showing information about the selected item.

### Counting reads and selection info box (click to enlarge)



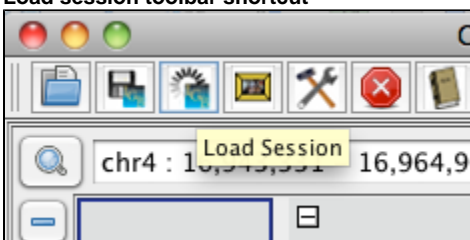
### Selection info window (click to enlarge)



### Saving and loading sessions

Now you can easily "save your place" in a lengthy IGB session and also share your sessions with collaborators. A new **Save Session** feature creates a simple .xml file that remembers your location, files you've opened, data you've loaded, track preferences, and more.

### Load session toolbar shortcut



## Saving tracks and annotations

IGB now supports saving selected annotations as well as entire tracks. IGB also supports more formats for saving annotations to local files, including BED Detail, bedgraph (wig), and more.

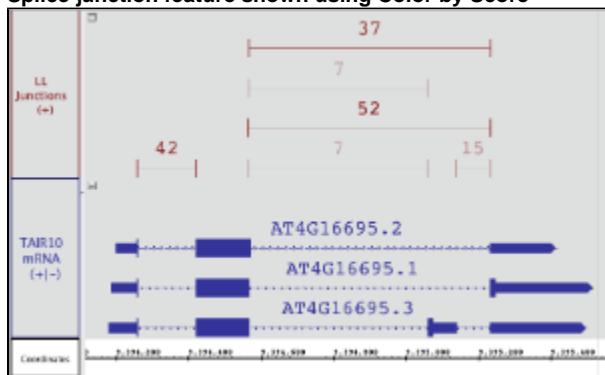
## More control over Color by Score

**Color by Score** allows you to make annotations appear darker or lighter based on their score. This feature aids in identifying outliers and assessing the distribution of scores in a large data set. However, prior to this release, the scale of values was pre-set as 1 to 1000.

Starting with IGB 7.0, you can use the **Set Range for Color by Score** feature to designate your own range of values.

In the example below, higher-scoring splice junction features with greater support (more supporting reads) appear darker than less well-supported junctions. You can also tell IGB to display the score above each feature by choosing "score" as the Label Field.

### Splice junction feature shown using Color by Score



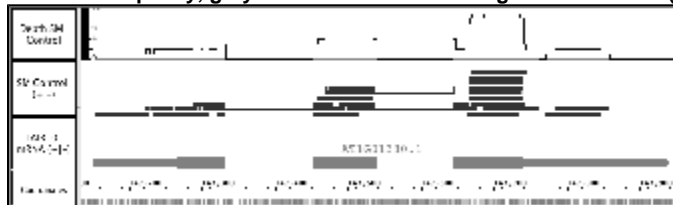
## Features for creating publication-quality images

- You can now set the track label font color and font size independently from track data.
- IGB now supports larger image resolution settings. You can export an image with resolution up to 1000 DPI.
- A new feature **Clear Visual Tools** and **Show All Visual tools** (View menu) hides graphical elements such as the zoom stripe, the IGB Track Mark, Locked Track Mark, etc.

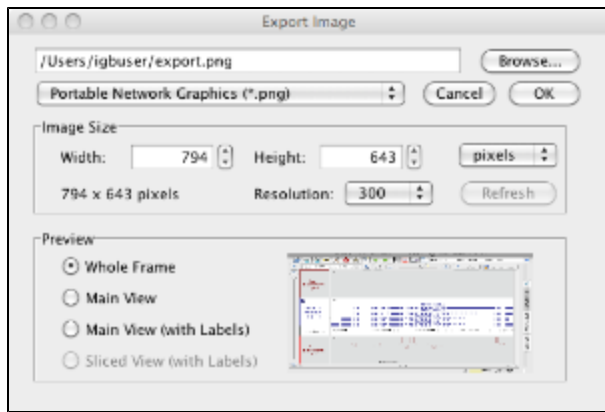
### Image showing visual tools you can hide using the new Clear Visual Tools feature (click to enlarge)



### Publication-quality, grayscale version of the image shown above (click to enlarge)



### IGB Export Image window (click to enlarge)



## New and improved tabbed panels

### Data Access panel

The order, background color, and font color of data sets in the **Data Management Table** now reflects their appearance and ordering in the IGB main window display. We right-justified the text in the **Data Management Table**, making long file paths easier to read. File information is still available via the **Selection Info** tab.

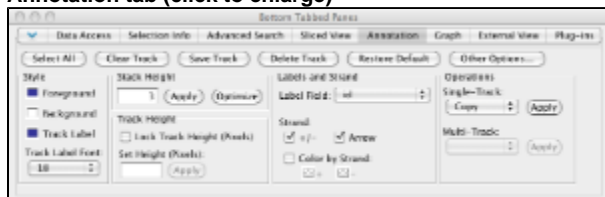
We also updated and simplified data set **Load Mode** settings, which now include **Auto**, **Manual** and **Genome** load modes.

- **Manual** means data only load when you request it by pressing the **Load Data** button
- **Auto** means that data load automatically whenever the zoom level passes the **AutoLoad** threshold.
  - Set the zoom level **AutoLoad** threshold using the **View** menu
- **Genome** means data for the entire genome load at once

### Annotation panel

The new **Annotation** tabbed panel provides easy access to annotation track settings and operations.

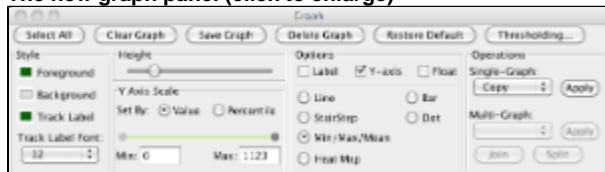
#### Annotation tab (click to enlarge)



### Graph panel

The **Graph** panel now includes a new double-headed slider for setting the lower and upper bounds on the visible range of graphs.

#### The new graph panel (click to enlarge)



### Selection Info panel

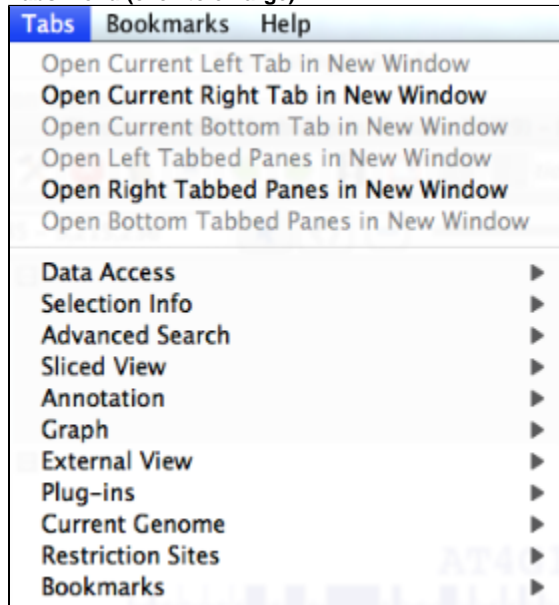
Each cell in the **Selection Info** table now contains a button you can click to view and copy cell contents. This is useful when the cells contain a large amount of text or text you would like to copy into a Google search or other tool.

### New Tab Menu

The new **Tab** menu makes it easier to access tab control features. Tabs, as always, can be shown or hidden, can be displayed as part of the IGB main window (as IGB "trays") or as separate windows.

Portrait (side) tabs can be shown in on the left or right side of the main window.

#### Tabs menu (click to enlarge)



#### Toolbar shortcuts

We added a new toolbar that allows users to access commonly-used functions with one click of a button. This is personalized by the user: open the **Preferences > Toolbar** tab, and select the functions you want in the **Toolbar**! This saves the need to continually flip between tabs for a commonly used function. This has been added in addition to **Key Stroke** shortcuts, which can still be personalized as well.

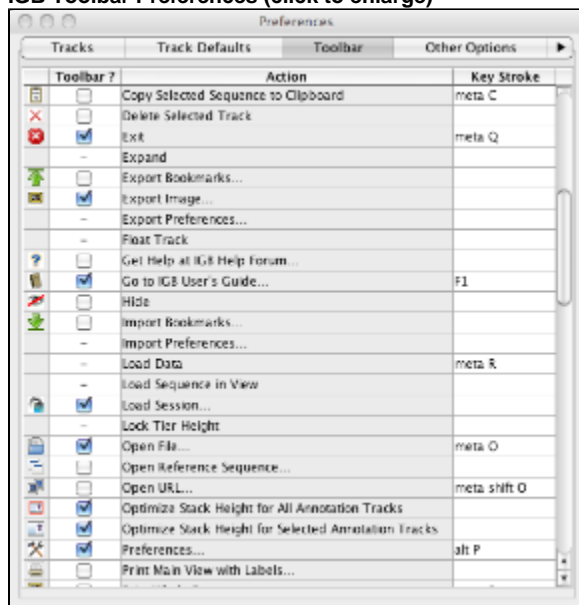
As part of this new feature, **Preferences > Toolbar** has been redesigned. Any function that can be put in the **Toolbar** has an icon and a checkbox; simply check the box on and the button will appear. After filling your **Toolbar** with functions you like, they can be rearranged into an order that suits you by simply clicking on the icon and dragging it to the new position. If you decide you do not want a function in the **Toolbar**, uncheck it from the **Preferences > Toolbar** menu. All of the functions in **Preferences > Toolbar** remain available for Key Stroke shortcuts as well.

Did we miss your favorite function? Let us know! Use **Request a Feature or a Shortcut...** to get to our request forum!

#### IGB 7.0 toolbar

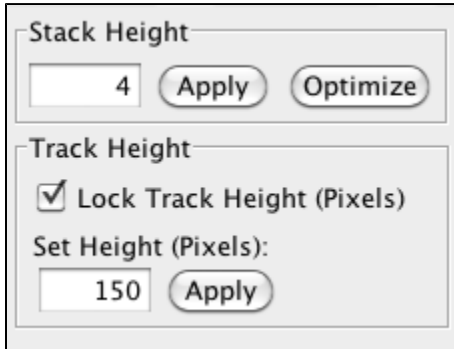


#### IGB Toolbar Preferences (click to enlarge)



#### Locking and setting track height (in pixels)

**Lock Track Height** is a new feature that allows locking the vertical dimension (in pixels) of a track. No matter how tall the other tracks become, a locked track's height never varies. **Lock Track Height** and **Set Height** for an annotation track using the new **Annotation** tab.



The screenshot shows the 'Annotation' tab in the IGB interface. It has two sections: 'Stack Height' and 'Track Height'. The 'Stack Height' section has a text input field containing the number '4' and two buttons: 'Apply' and 'Optimize'. The 'Track Height' section has a checked checkbox labeled 'Lock Track Height (Pixels)', followed by the text 'Set Height (Pixels):', a text input field containing the number '150', and an 'Apply' button.

### Stack Height shortcuts

IGB allows users to set the number of overlapping annotations that appear in "stacks" along the genomic sequence axis. If the number of annotations overlapping a region exceeds the track's stack height setting, the extra annotations are drawn in a single summary row at the top of the track. (IGB draws them in a darker color to signal they are part of the summary row.)

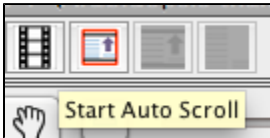
Previously, you could change the stack depth by entering new values in the Adjust Max Stack Depth window.

Starting with IGB 7.0, we added new **Optimize Stack Height** shortcuts to the **Toolbar** and **Annotation** tabs.

These shortcuts re-set the stack height for a track so that all annotations appear in a stack, with no "left overs" in the summary row.

### Auto Scroll Updates

We simplified configuring and running the animated IGB **Auto Scroll** feature, which scrolls tracks scroll from left to right automatically.



### New genomes supported in IGBQuickLoad

We added many new genomes and species to the IGBQuickLoad site, which is version controlled via publicly accessible subversion repository located at <https://svn.transvar.org/repos/genomes/trunk/pub/quickload>.

The IGBQuickLoad data repository allows us to document data sets using subversion log messages and allows everyone to retrieve older versions of data sets when needed, a key feature for ensuring reproducible research in complex data analyses.

### Retiring Affymetrix DAS2 data source in progress

We are in the process of retiring the Affymetrix DAS2 data source, which the company has not updated in more than a year. For each data set available via the Affymetrix NetAffx DAS2 source, we are setting up sequence data and reference genome annotations in the IGBQuickLoad.org site. To support Affymetrix probe set alignments, we created a variant of Affymetrix' link.psl format that allows indexing by the tabix utility, which allows partial, region-based manual loading of these data possible.

### New features for QuickLoad sites

Users can easily make Quickload sites to hold their data or make it easy to share their data with others. We have improved their ability to control the Quickload environment and to control how data loaded from the site appears.

- **File attribute in annots.xml now supports URLs**  
Annotation files listed in annots.xml files can now point to remote URLs instead of local files. To use this feature, replace the "name" attribute with a URL. IGB will then attempt to load the data from the remote site. This feature allows QuickLoad data providers to integrate data sets stored on diverse sites into a single Data Source under the IGB Data Access panel.
- **Continued support for foreground, background, and other style attributes**  
QuickLoad data providers can continue to specify how data sets will appear once loaded into IGB using file element attributes. Once loaded, of course, users can change all of these elements to personalize their data analysis experience.
- **QuickLoad site-specified Web links**  
Data providers can now specify link-out patterns by including a file named preferences.xml at the top level of the QuickLoad root directory.



## Web Link link-out builder tool improved

The **Web Links Tool** is an easy way for you to create and test new link-out patterns. Use the **Web Links Tool** to specify regular expressions for data sets or annotation ids; web links are used by IGB to access external Web sites when users choose the **Get More Info** option (right-click menu). IGB built-in links (e.g., Google searches, links to NCBI) provide examples of how to build your own link-out patterns.

### Web Links link-out builder (click to enlarge)

The screenshot shows the 'Web Links Tool' interface. It features a table titled 'Server Provided Web Link' with columns: Name, Regular Expression, URL Pattern, and Matches. The table lists several built-in links like Dictybase Gene, Dictybase, Entrez cross-data, etc. Below this is a section for 'Custom Web Links - Click To Edit' with a similar table. At the bottom, there are buttons for 'Create New', 'Delete', 'Import', and 'Export', followed by a 'Web Link Builder' section with input fields for Name, Regular Expression, and URL, and radio buttons for 'Track Name' and 'Annotation ID'.

Name	Regular Expression	URL Pattern	Matches
Dictybase Gene	(?<ODD>Gid+\$	http://dictybase.org/gene/\$\$	Annotation ID
Dictybase	(?<ODD>id+\$	http://www.dictybase.org/db/cgi-bin/sea...	Annotation ID
Dictybase	(?<dictybase>.*\$	http://dictybase.org/db/cgi-bin/search/s...	Annotation ID
Entrez cross-data...	(?<NCBI>.*\$	http://www.ncbi.nlm.nih.gov/query/term...	Annotation ID
Entrez	(?<GeneID>.*\$	http://www.ncbi.nlm.nih.gov/gene/\$\$	Annotation ID
Hybase	(?<HYBASE>.*\$	http://hybase.org/bin/fbldp.html/\$\$	Annotation ID
GeneDB	(?<^sanger_pumbe\$	http://www.genedb.org/genedb/search/h...	Track Name
Google	.*	http://www.google.com/search?q=\$\$	Annotation ID

Name	Regular Expression	URL Pattern	Matches
NCBI Example	(?<EST>((mRNA) R...	http://ncbi.nlm.nih.gov/nuccore/\$\$	Track Name

Create New Delete Import Export

Web Link Builder

Name: NCBI Example Regular Expression: (?<EST>((mRNA)|RefGene) URL: http://ncbi.nlm.nih.gov/nuccore/\$\$

Regular Expression Matches: ☒ Track Name ☐ Annotation ID

## IGB License

Source code for IGB is released under the Common Public License, v1.0, an OSI approved open source license. For your convenience, the license appears below:

### Common Public License Version 1.0

THE ACCOMPANYING PROGRAM IS PROVIDED UNDER THE TERMS OF THIS COMMON PUBLIC LICENSE ("AGREEMENT"). ANY USE, REPRODUCTION OR DISTRIBUTION OF THE PROGRAM CONSTITUTES RECIPIENT'S ACCEPTANCE OF THIS AGREEMENT.

#### 1. DEFINITIONS

"Contribution" means:

- a) in the case of the initial Contributor, the initial code and documentation distributed under this Agreement, and
- b) in the case of each subsequent Contributor:

- i) changes to the Program, and
- ii) additions to the Program;

where such changes and/or additions to the Program originate from and are distributed by that particular Contributor. A Contribution 'originates' from a Contributor if it was added to the Program by such Contributor itself or anyone acting on such Contributor's behalf. Contributions do not include additions to the Program which: (i) are separate modules of software distributed in conjunction with the Program under their own license agreement, and (ii) are not derivative works of the Program.

"Contributor" means any person or entity that distributes the Program.

"Licensed Patents " mean patent claims licensable by a Contributor which are necessarily infringed by the use or sale of its Contribution alone or when combined with the Program.

"Program" means the Contributions distributed in accordance with this Agreement.

"Recipient" means anyone who receives the Program under this Agreement, including all Contributors.

#### 2. GRANT OF RIGHTS

- a) Subject to the terms of this Agreement, each Contributor hereby grants Recipient a non-exclusive, worldwide, royalty-free copyright license to reproduce, prepare derivative works of, publicly display, publicly perform, distribute and sublicense the Contribution of such Contributor, if any, and such derivative works, in source code and object code form.

b) Subject to the terms of this Agreement, each Contributor hereby grants Recipient a non-exclusive, worldwide, royalty-free patent license under Licensed Patents to make, use, sell, offer to sell, import and otherwise transfer the Contribution of such Contributor, if any, in source code and object code form. This patent license shall apply to the combination of the Contribution and the Program if, at the time the Contribution is added by the Contributor, such addition of the Contribution causes such combination to be covered by the Licensed Patents. The patent license shall not apply to any other combinations which include the Contribution. No hardware per se is licensed hereunder.

c) Recipient understands that although each Contributor grants the licenses to its Contributions set forth herein, no assurances are provided by any Contributor that the Program does not infringe the patent or other intellectual property rights of any other entity. Each Contributor disclaims any liability to Recipient for claims brought by any other entity based on infringement of intellectual property rights or otherwise. As a condition to exercising the rights and licenses granted hereunder, each Recipient hereby assumes sole responsibility to secure any other intellectual property rights needed, if any. For example, if a third party patent license is required to allow Recipient to distribute the Program, it is Recipient's responsibility to acquire that license before distributing the Program.

d) Each Contributor represents that to its knowledge it has sufficient copyright rights in its Contribution, if any, to grant the copyright license set forth in this Agreement.

### 3. REQUIREMENTS

A Contributor may choose to distribute the Program in object code form under its own license agreement, provided that:

a) it complies with the terms and conditions of this Agreement; and

b) its license agreement:

i) effectively disclaims on behalf of all Contributors all warranties and conditions, express and implied, including warranties or conditions of title and non-infringement, and implied warranties or conditions of merchantability and fitness for a particular purpose;

ii) effectively excludes on behalf of all Contributors all liability for damages, including direct, indirect, special, incidental and consequential damages, such as lost profits;

iii) states that any provisions which differ from this Agreement are offered by that Contributor alone and not by any other party; and

iv) states that source code for the Program is available from such Contributor, and informs licensees how to obtain it in a reasonable manner on or through a medium customarily used for software exchange.

When the Program is made available in source code form:

a) it must be made available under this Agreement; and

b) a copy of this Agreement must be included with each copy of the Program.

Contributors may not remove or alter any copyright notices contained within the Program.

Each Contributor must identify itself as the originator of its Contribution, if any, in a manner that reasonably allows subsequent Recipients to identify the originator of the Contribution.

### 4. COMMERCIAL DISTRIBUTION

Commercial distributors of software may accept certain responsibilities with respect to end users, business partners and the like. While this license is intended to facilitate the commercial use of the Program, the Contributor who includes the Program in a commercial product offering should do so in a manner which does not create potential liability for other Contributors. Therefore, if a Contributor includes the Program in a commercial product offering, such Contributor ("Commercial Contributor") hereby agrees to defend and indemnify every other Contributor ("Indemnified Contributor") against any losses, damages and costs (collectively "Losses") arising from claims, lawsuits and other legal actions brought by a third party against the Indemnified Contributor to the extent caused by the acts or omissions of such Commercial Contributor in connection with its distribution of the Program in a commercial product offering. The obligations in this section do not apply to any claims or Losses relating to any actual or alleged intellectual property infringement. In order to qualify, an Indemnified Contributor must: a) promptly notify the Commercial Contributor in writing of such claim, and b) allow the Commercial Contributor to control, and cooperate with the Commercial Contributor in, the defense and any related settlement negotiations. The Indemnified Contributor may participate in any such claim at its own expense.

For example, a Contributor might include the Program in a commercial product offering, Product X. That Contributor is then a Commercial Contributor. If that Commercial Contributor then makes performance claims, or offers warranties related to Product X, those performance claims and warranties are such Commercial Contributor's responsibility alone. Under this section, the Commercial Contributor would have to defend claims against the other Contributors related to those performance claims and warranties, and if a court requires any other Contributor to pay any damages as a result, the Commercial Contributor must pay those damages.

### 5. NO WARRANTY

EXCEPT AS EXPRESSLY SET FORTH IN THIS AGREEMENT, THE PROGRAM IS PROVIDED ON AN "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, EITHER EXPRESS OR IMPLIED INCLUDING, WITHOUT LIMITATION, ANY WARRANTIES OR CONDITIONS OF TITLE, NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Each Recipient is solely responsible for determining the appropriateness of using and distributing the Program and assumes all risks associated with its exercise of rights under this Agreement, including but not limited to the risks and costs of program errors, compliance with applicable laws, damage to or loss of data, programs or equipment, and unavailability or interruption of operations.

#### 6. DISCLAIMER OF LIABILITY

EXCEPT AS EXPRESSLY SET FORTH IN THIS AGREEMENT, NEITHER RECIPIENT NOR ANY CONTRIBUTORS SHALL HAVE ANY LIABILITY FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING WITHOUT LIMITATION LOST PROFITS), HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OR DISTRIBUTION OF THE PROGRAM OR THE EXERCISE OF ANY RIGHTS GRANTED HEREUNDER, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

#### 7. GENERAL

If any provision of this Agreement is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this Agreement, and without further action by the parties hereto, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

If Recipient institutes patent litigation against a Contributor with respect to a patent applicable to software (including a cross-claim or counterclaim in a lawsuit), then any patent licenses granted by that Contributor to such Recipient under this Agreement shall terminate as of the date such litigation is filed. In addition, if Recipient institutes patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Program itself (excluding combinations of the Program with other software or hardware) infringes such Recipient's patent(s), then such Recipient's rights granted under Section 2(b) shall terminate as of the date such litigation is filed.

All Recipient's rights under this Agreement shall terminate if it fails to comply with any of the material terms or conditions of this Agreement and does not cure such failure in a reasonable period of time after becoming aware of such noncompliance. If all Recipient's rights under this Agreement terminate, Recipient agrees to cease use and distribution of the Program as soon as reasonably practicable. However, Recipient's obligations under this Agreement and any licenses granted by Recipient relating to the Program shall continue and survive.

Everyone is permitted to copy and distribute copies of this Agreement, but in order to avoid inconsistency the Agreement is copyrighted and may only be modified in the following manner. The Agreement Steward reserves the right to publish new versions (including revisions) of this Agreement from time to time. No one other than the Agreement Steward has the right to modify this Agreement. IBM is the initial Agreement Steward. IBM may assign the responsibility to serve as the Agreement Steward to a suitable separate entity. Each new version of the Agreement will be given a distinguishing version number. The Program (including Contributions) may always be distributed subject to the version of the Agreement under which it was received. In addition, after a new version of the Agreement is published, Contributor may elect to distribute the Program (including its Contributions) under the new version. Except as expressly stated in Sections 2(a) and 2(b) above, Recipient receives no rights or licenses to the intellectual property of any Contributor under this Agreement, whether expressly, by implication, estoppel or otherwise. All rights in the Program not expressly granted under this Agreement are reserved.

This Agreement is governed by the laws of the State of New York and the intellectual property laws of the United States of America. No party to this Agreement will bring a legal action under this Agreement more than one year after the cause of action arose. Each party waives its rights to a jury trial in any resulting litigation.