

# Share and visualize genome-scale data sets using Integrated Genome Browser (IGB)

David Norris<sup>1</sup>, Hiral Vora<sup>2</sup>, Alyssa Gullledge<sup>2</sup>, Michael Lawrence<sup>3</sup>, Lance Frohman<sup>3</sup>, Ann Loraine<sup>2</sup>

<sup>1</sup>Dept. of Computer Science, UNC Charlotte <sup>2</sup>Dept. of Bioinformatics and Genomics, UNC Charlotte <sup>3</sup>Genentech Research and Early Development



## Abstract

Emerging high-throughput sequencing technologies produce vast data sets that can be analyzed in diverse ways. Making the most of these data requires tools that are flexible enough to handle a diverse range of questions. Integrated Genome Browser (IGB) is a fast, flexible, and free genome visualization tool designed to help scientists analyze large-scale data sets. IGB supports a variety of visualization techniques adapted from information visualization fields in order to foster productive exploration of data. More than 8,000 scientists from diverse communities worldwide use IGB to visualize data from next-generation sequencing experiments, tiling arrays, and EST/genome sequencing projects. IGB can run on any computer with Java 1.6 installed and is freely available from [bioviz.org/igb](http://bioviz.org/igb).

## History

The IGB project started at Affymetrix, Inc to support the company's tiling array and genotyping products. IGB was always free to use, and in 2004 Affymetrix released IGB as open source software. Since 2008, an NSF Arabidopsis 2010 grant has supported IGB development for the Arabidopsis and plant genomics communities. The Loraine lab at UNC Charlotte leads IGB development with collaborators from the Huntsman Cancer Institute (Utah), Genentech, and the Molecular Pathology Institute (Vienna). With their help, we've added many new features, including new ways to explore, analyze, and share data from high-throughput sequencing experiments.

Visualization of short read (75 bp, Illumina) alignments loaded from .bam (binary alignment/map) files local files or a remote server

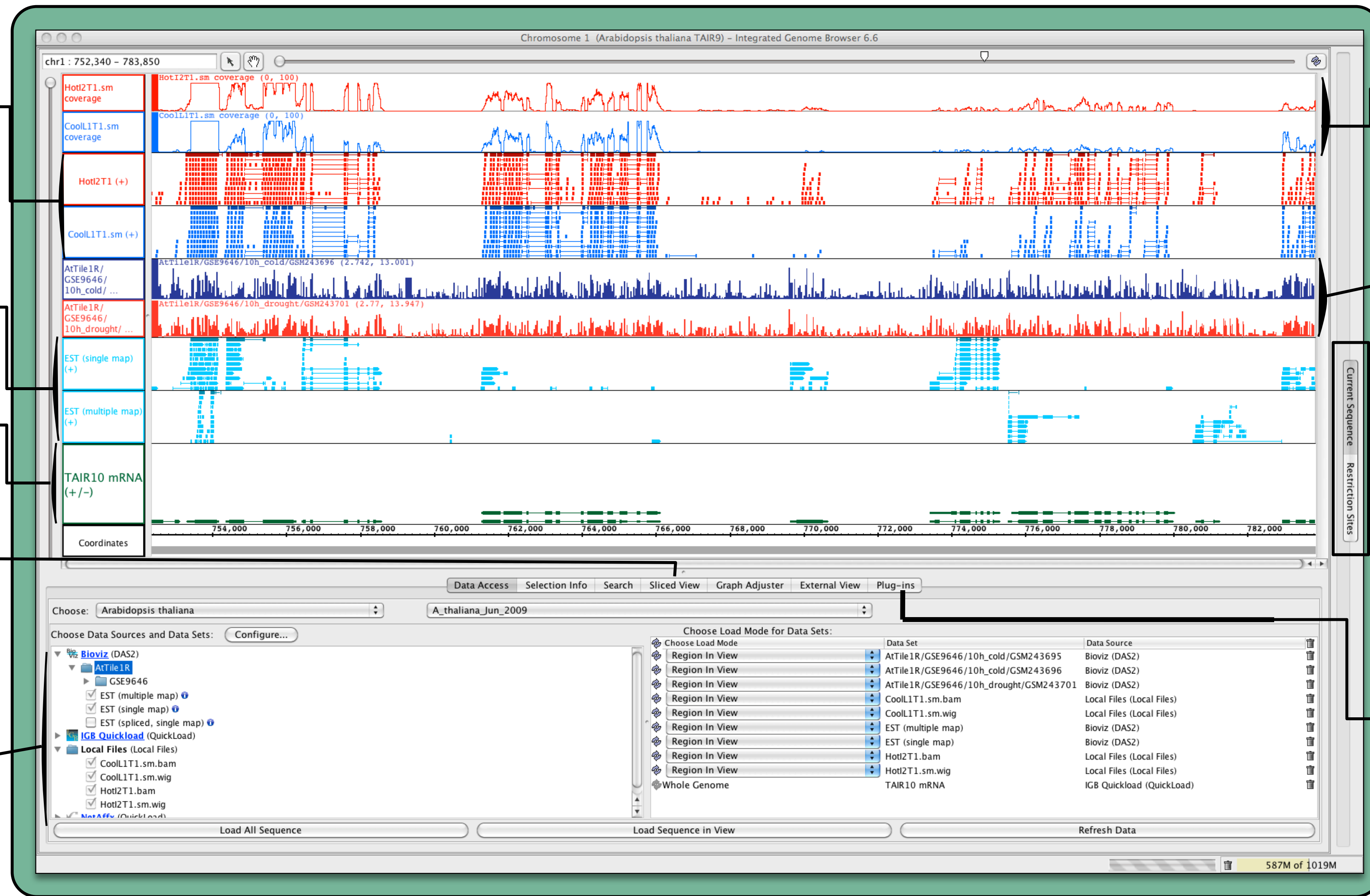
Visualization of EST alignments loaded from the BioViz DAS2 server

TAIR10 protein-coding (mRNA) gene models

Different tabs offer different views of selected items

IGB can load data from diverse data sources and servers, including:

- BioViz QuickLoad site
- BioViz DAS site
- Local files
- NetAffx (Affymetrix)
- Custom sites you set up.

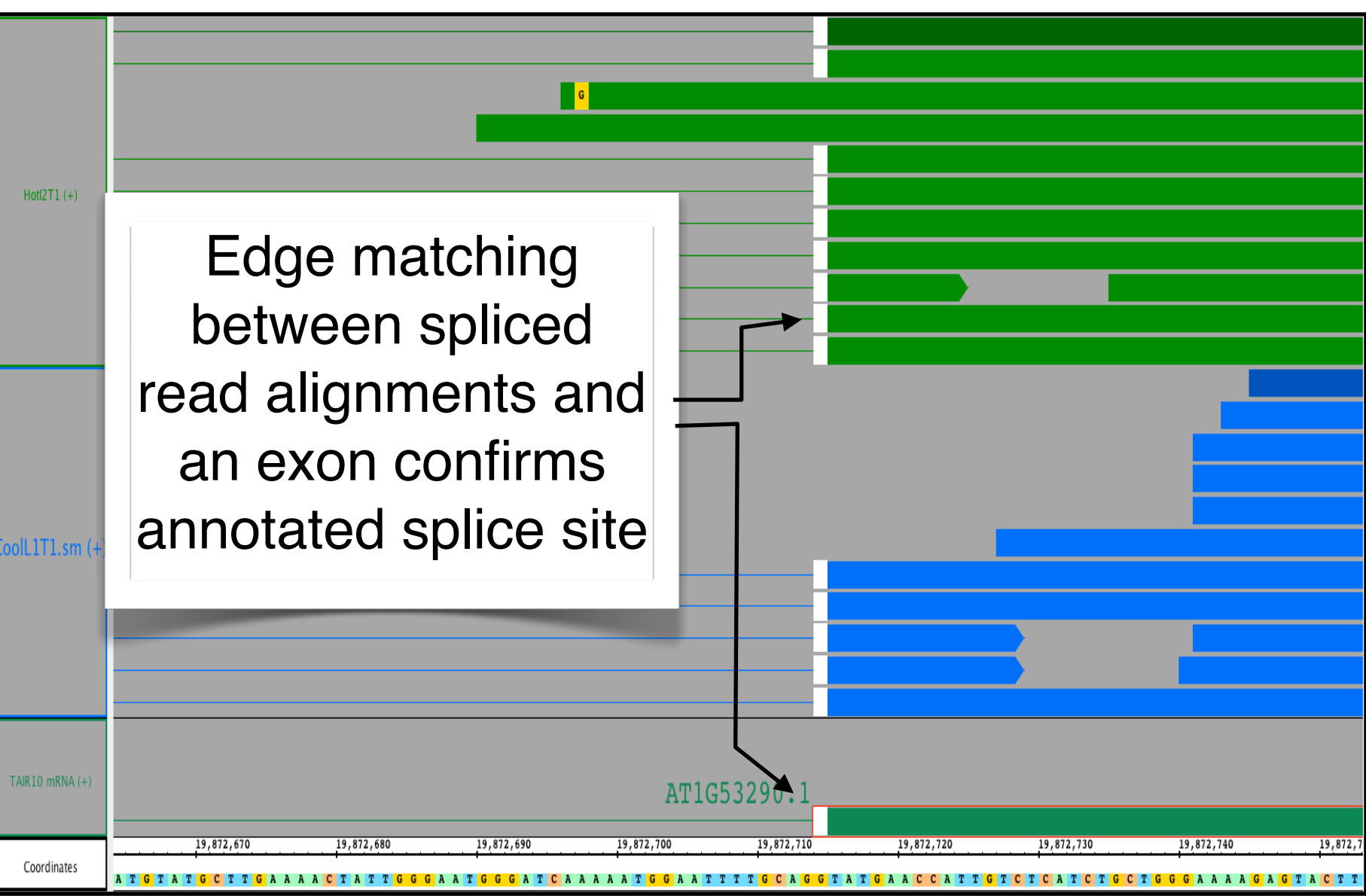


IGB generates coverage graphs from RNA-Seq data. Coverage graphs help to compare expression across samples.

Expression data from tiling arrays harvested from Gene Expression Omnibus

Collapsible tabbed panels maximize your workspace and allow convenient access to IGB functions, such as restriction site mapping.

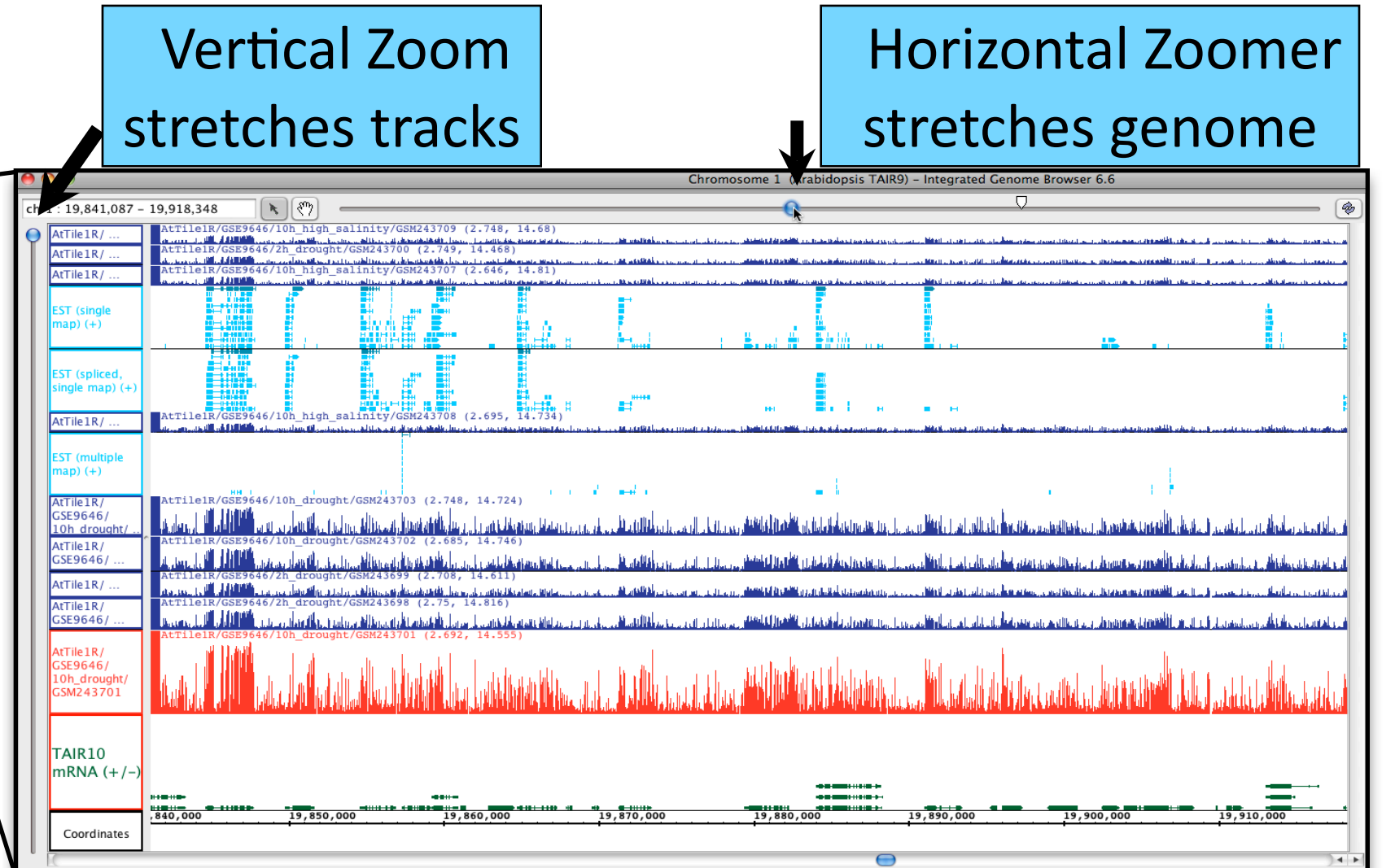
For IGB power users: IGB's all new plug-in system (inspired by Cytoscape) lets developers add new functions and algorithms for statistical analysis, splice variant quantitation, SNP discovery, and more.



Edge matching between spliced read alignments and an exon confirms annotated splice site

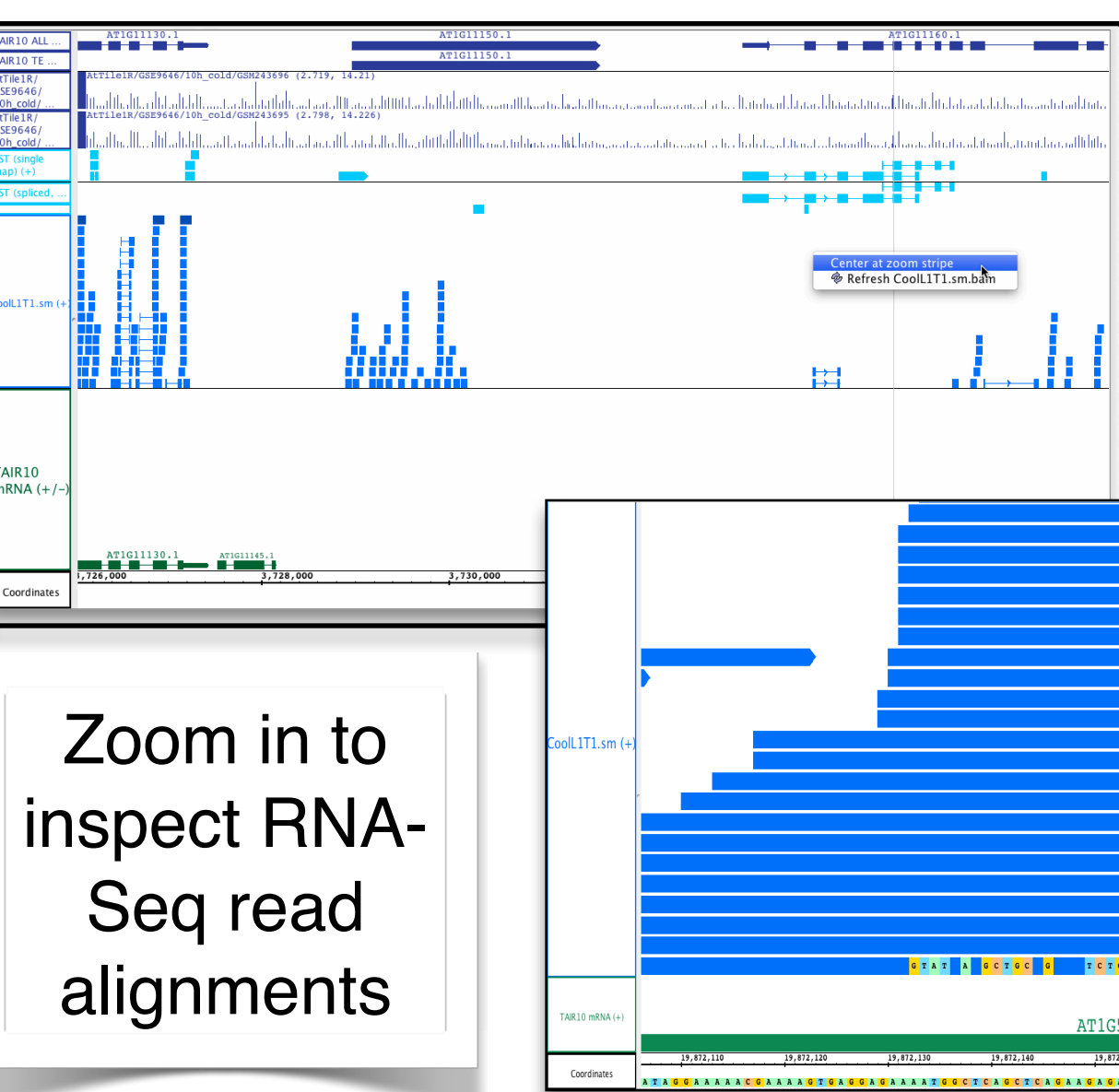
**Animated Zooming**  
In IGB, zooming is animated and dynamic, allowing you to zoom in and out without losing track of where you are.

**Edge Matching**  
IGB edge matching makes comparing alignments and gene models easier. When you select a feature, such as an exon or gene model, the edges of all 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.



Vertical Zoom stretches tracks

Horizontal Zoomer stretches genome

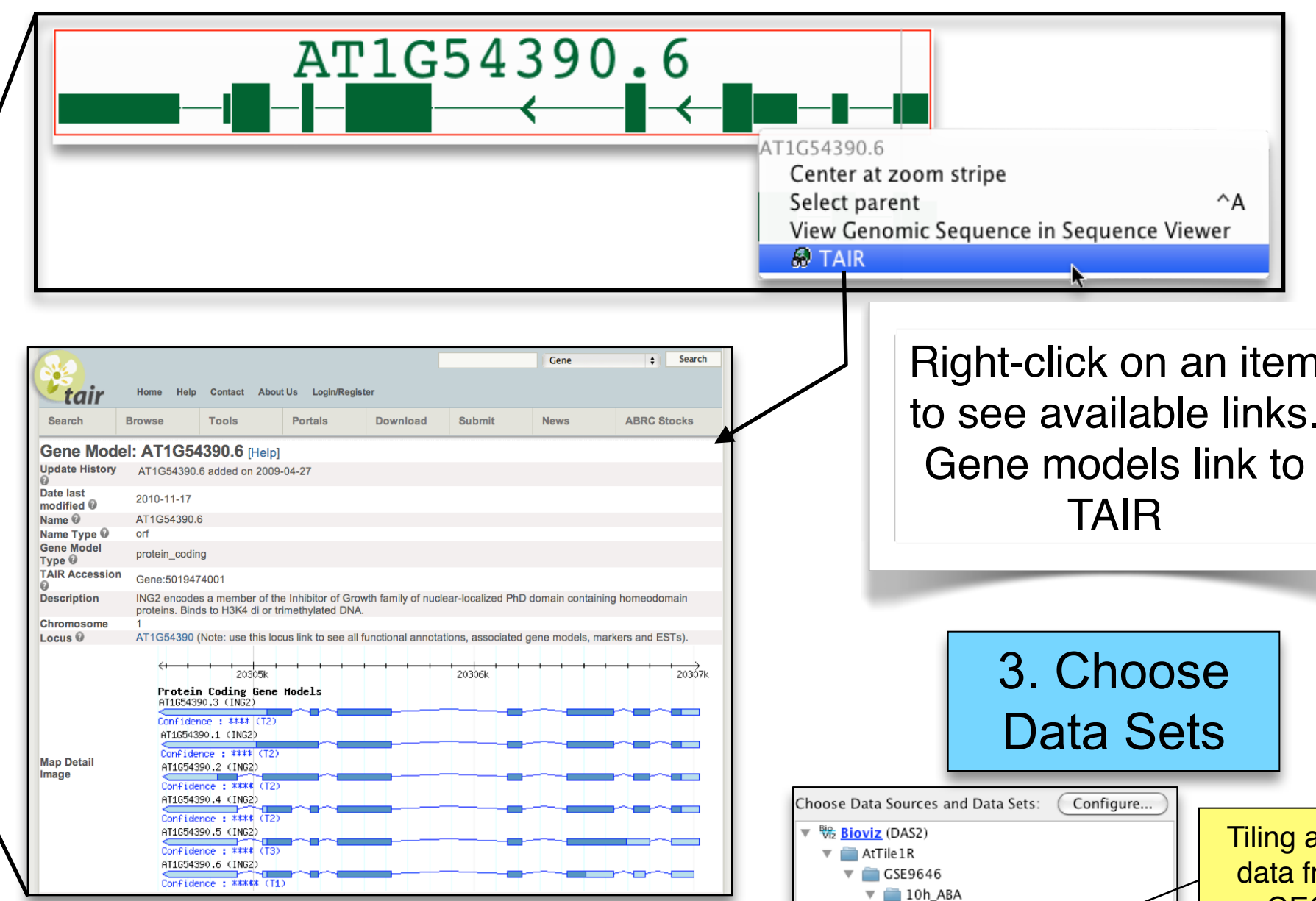


Zoom in to inspect RNA-Seq read alignments

**IGB Tip:** Mismatches reveal alignment problems or SNPs

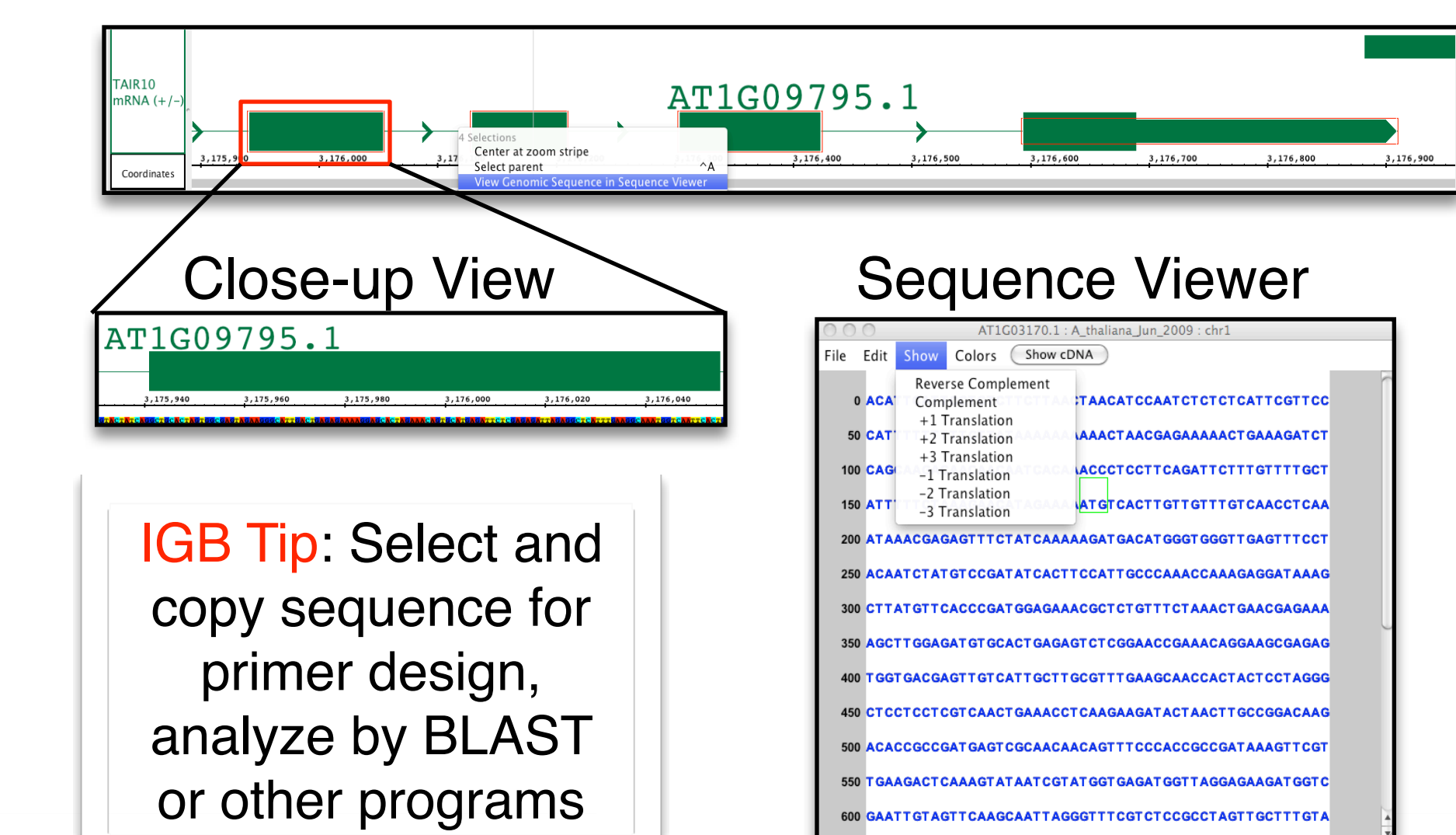
**Linkouts**  
Right-clicking a gene model opens a menu with a link to TAIR. IGB provides convenient access to external Web sources by providing linkouts from track features.

**Zoom Stripe**  
IGB's zoom stripe guide line helps you navigate the genome and also provides a visual aid for matching features across tracks or to sequence.



Right-click on an item to see available links. Gene models link to TAIR

3. Choose Data Sets



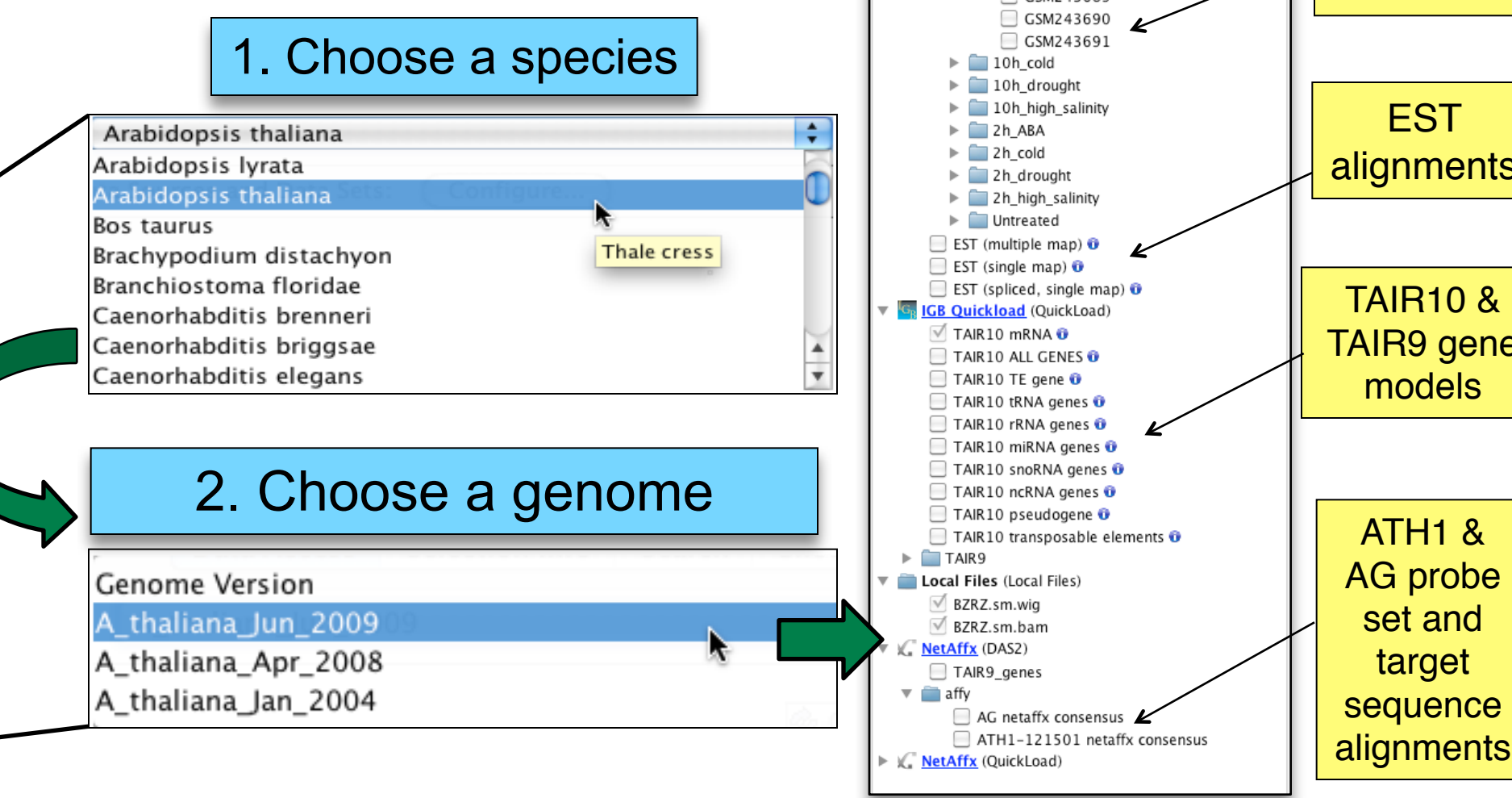
Close-up View

Sequence Viewer

**IGB Tip:** Select and copy sequence for primer design, analyze by BLAST or other programs

**Data Access**  
IGB can load data from QuickLoad, distributed annotation server (DAS) sites or files on your desktop. IGB supports more than 90 species, including 14 plants. You can also set up your own sites for sharing data on the web. All you need is a publicly-accessible Web directory. IGB also supports password protection.

**Viewing Sequence**  
IGB display genomic sequence at high zoom. You can right-click items on display in the viewer to launch the Sequence Viewer, which shows genomic DNA, cDNA and translations in all six frames.



1. Choose a species

2. Choose a genome

Tiling array data from GEO

EST alignments

TAIR10 & TAIR9 gene models

ATH1 & AG probe set and target sequence alignments

IGB is freely available from [bioviz.org/igb](http://bioviz.org/igb) and runs on any computer with Java 1.6 installed

