Integrated Genome Browser
User's Guide

January 10, 2006

Contents

Getting started

Introduction

The Affymetrix Integrated Genome Browser (IGB) is a powerful tool to help genome researchers:

- Compare experimental results to computational results
- Visualize and compare multiple genomic annotation sets from a variety of public and private sources
- Target areas of interest to explore further via other tools and/or experimentation
- Provide guidance for fine-tuning or modification of experiments

Overview of IGB functionality

Generally, use IGB to:

- Select and load genomic annotations to analyze from a variety of sources including your own experimental results
- Compare variations in annotations and expression activity in different data sets
- Search for and drill down in areas of interest

Unique features and benefits of IGB

IGB combines in one viewer your own experimental or computational results, common reference information, and access to public and private data banks.

Other features unique to IGB include:

- Slicing: a tool that facilitates examination and manipulation of alternate splicings, intron and exon boundaries
- Ability to merge graphs of scores at specified points along the genome
- Facility in to handling large amounts of genomic annotations with speed and efficiency

About this document

Getting the most current version

Visit the main page for IGB on the Affymetrix web site: http://www.affymetrix.com/support/developer/tools/download_igb.affx and look for the link to this document or to IGB documentation generally.

Other IGB documentation

Guides for using data from other Affymetrix programs are available on the Affymetrix Tools web page here:

http://www.affymetrix.com/support/developer/tools/affytools.affx

and on the Affymetrix Developer's Network web page here:

http://www.affymetrix.com/support/developer/tools/devnettools.affx.

Those pages contain information for dealing with data from these sources, among others:

- GCOS
- CNAT
- Tiling Arrays

Document conventions

Onscreen text (i.e., menus, button labels, folder names, field values, etc.) is indicated by **Arial** 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** menu > **Open File** means "click on the **File** menu and choose **Open File**."

QuickLoad tab> **Load Sequence Residues** means "click the **QuickLoad** tab and then click the **Load Sequence Residues** button"

Installation and setup

System requirements

Current system requirements are posted at:
http://www.affymetrix.com/support/developer/tools/download_igb.affx

Hardware/ OS

IGB has been tested on Windows 2000 and Linux but should work on any Java environment.

Software

Java Runtime Engine 1.4.2 or higher must be installed before IGB can run. This software is available from http://java.sun.com/downloads.

Installing

Go to the webpage:
http://www.affymetrix.com/support/developer/tools/download_igb.affx

If you do not see one or more buttons labeled “Launch IGB”, or something similar, you need to install Java Web Start, which is included with the Java Runtime Engine version 1.4.2 or higher.

See Appendix 1 Using Java Web Start.

After Java Web Start is installed, return to that web page and press one of the buttons labeled “Launch IGB”.

If IGB has not already been installed, or if an update is available, it will be installed automatically. If IGB has already been installed, pressing these buttons will simply start the program that is already installed on your machine.

The multiple buttons give the choice to run IGB with various amounts of RAM memory. All choices run the same program, but specify different maximum amounts of RAM to allow. Please choose a version that requires no more than the total RAM you have available to dedicate to IGB.

Setting preferences

See Appendix 3 Preferences.

Launching IGB

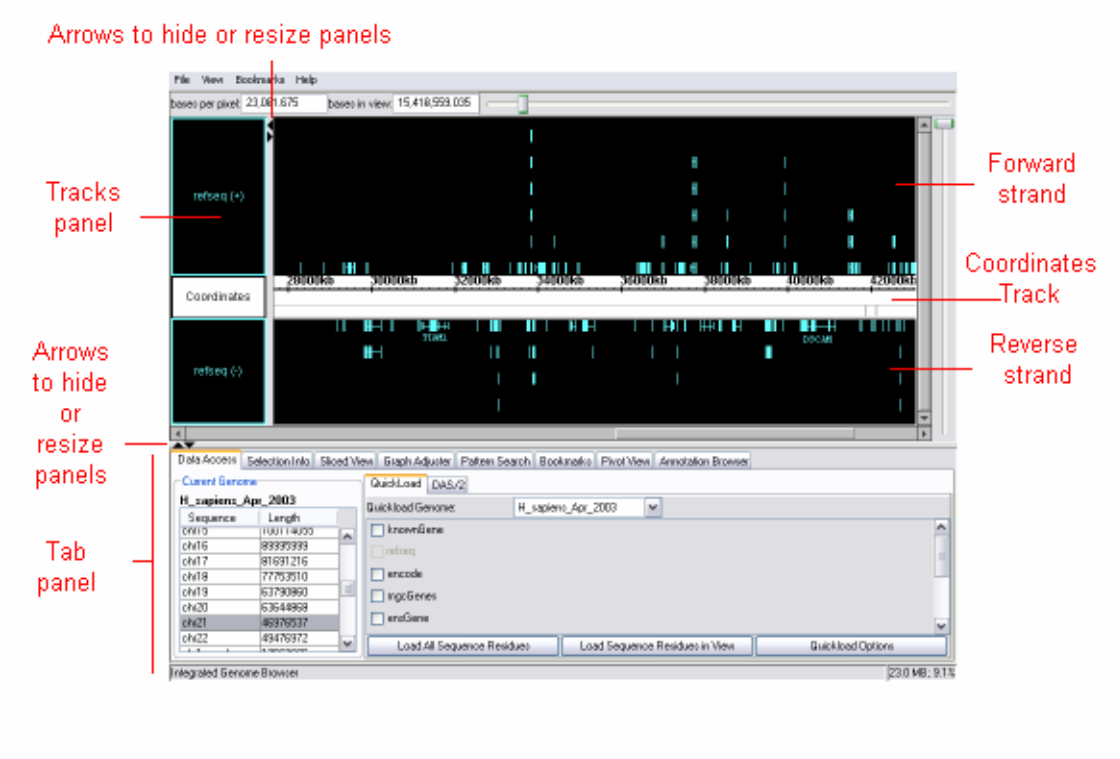
After you install IGB using the procedure in Appendix 1 Using Java Web Start, launch IGB using one of the following methods:

- If you chose to integrate IGB with your desktop environment when you installed IGB, you can launch IGB by clicking the IGB icon on your desktop or choose it from **Start** menu > **Programs**.
- If you did not choose to integrate IGB with your desktop environment when you installed IGB, choose **Start** menu > **Programs** > **Java Web Start** > **Java Web Start**. You will see the Web Start Console. Click IGB in the list of applications in Java Web Start, then click **Start**.
- If you do not want to install the most current version of IGB, make sure that you are not connected to the internet when you launch IGB. By default, when you launch IGB, if you are connected to the internet, the most current version of IGB is downloaded and installed on your computer.

Launching multiple instances of IGB simultaneously is not recommended.

Overview of the IGB window

After you have loaded annotations into IGB, the browser looks like this:



"Tracks" in IGB are comparable to "annotation tracks" in other genome browsers. The main portion of the browser window is the viewer window.

Viewing Tip

Information in each tab panel can be displayed in a separate window. You can then drag the arrows (shown in the illustration above) to hide or maximize the viewer panel in the IGB window, giving you a larger area for viewing annotations and graphs.

To open a tab panel in a new window:

- Click the tab to open, then choose **View** menu > **Open tab in new window**.

To return the window to the tab panel:

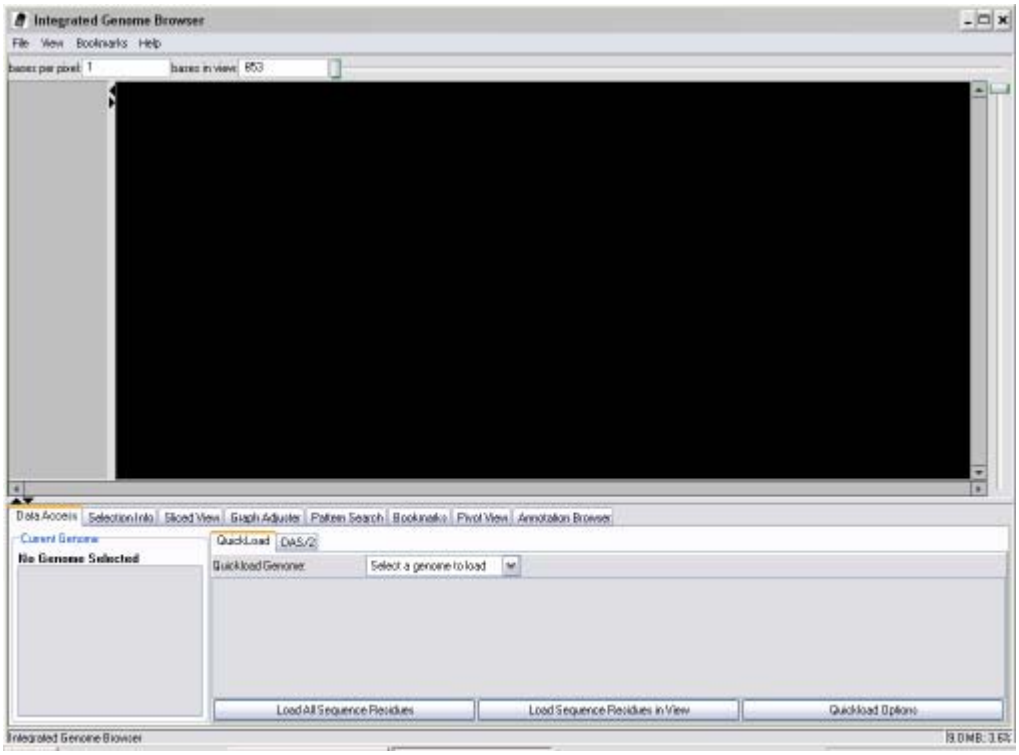
- Click the [X] in the upper right corner of the window to close the window.

Keyboard shortcuts

You can create keyboard commands (keyboard shortcuts) for some operations that you perform frequently. See Changing keyboard commands on page 85.

Note for Macintosh users: Operations in this document that require right-clicking can be performed on Mac by holding down the control key while clicking the mouse.

Choosing and opening data sources



IGB opens with a blank viewer. Load genomic information from any source(s) below to compare annotations.

Data sources that IGB can work with include:

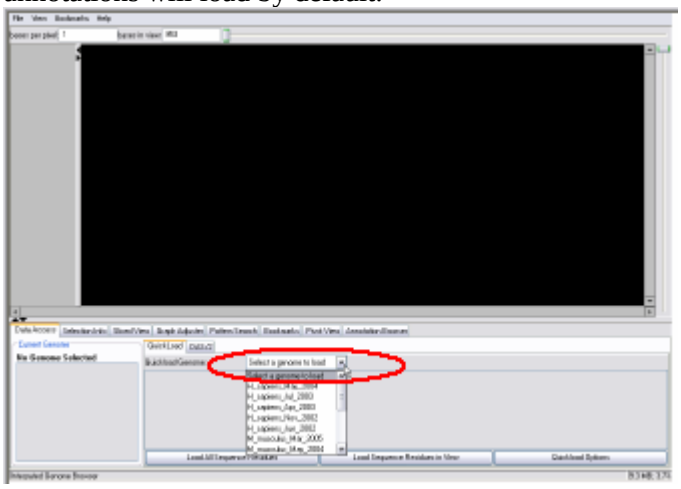
- Files generated from Affymetrix software tools, such as CNAT, GCOS, or ExACT.
- Tiling array data
- Fast-loading standard reference data such as RefSeq annotations
- Publicly available data sets from sources such as UCSC
- Data files of various formats

For more about each of these, see the appropriate sections below.

Fast-loading standard reference data

"QuickLoad" standard annotation types, predictions, and/or contigs for a whole genome version you choose.

1. Make sure you're connected to the internet (this may not work on slow connections)
2. Click the **Data Access** tab. (The Data Access panel is already displayed by default when you launch IGB.)
3. Click the **QuickLoad** sub-tab. (The QuickLoad sub-panel is already displayed by default when you launch IGB.)
4. From the drop-down list, select a genome assembly to load. RefSeq annotations will load by default.



5. Wait a few moments for the annotation type to load.
6. Initially, IGB displays no annotations. To view annotations, click a chromosome in the table at the left side of the tab panel.
7. Contigs are indicated by gray-outlined bars at the bottom of the Coordinates track. If no gray-outlined bar is present, this means that there is no contig data available in the annotation type or for the region that you are viewing.
8. (Optional) Load any of the reference annotation types shown in the panel by clicking to select the associated checkbox. Available annotation types will vary depending on the genome you select. You may need to click the referenced chromosome to view the loaded data.
9. (Optional) Load sequence residues for the selected chromosome. See page 18.

   If the desired sequence does not appear, it may not be available for the assembly you have selected.

Using advanced QuickLoad features

Advanced QuickLoad features, including the ability to create your own QuickLoad server, are described in Advanced Features below.

Loading annotation types from a DAS server

DAS (Distributed Annotation System) servers are sources of annotation types. Loading annotations from a DAS server gives you more flexibility than QuickLoad to specify the amount of data that you load. For example, instead of loading an entire genome, you can load annotations for a range of coordinates.

There are two types of DAS servers, and you access each via a different method in IGB. The original DAS protocol is referred to in this document as DAS/1. DAS/2 is a newer version of the protocol that provides faster access to annotations as well as other improvements. For more information about DAS in general, visit http://biodas.org.

See the following procedures to load annotation types from DAS servers:

- Loading data from a DAS/1 server
- Loading data from a DAS/2 server

Loading data from a DAS/1 server

DAS servers are described in the preceding section.

You can add to or change the list of DAS/1 servers that you can access. See Adding DAS/1 data sources on page 84.

To load data from a DAS/1 server:

1. If you are already viewing other data, make sure the viewer is displaying the genome and chromosome to be loaded from DAS.
2. If you are already viewing other data, and depending on the density of the annotations you expect to be loaded (EST's, for example, are dense) you may want to zoom in so the range displayed in the viewer is about 2 Mb.
3. Choose **File** menu > **Load DAS Features**
4. Select a DAS Server from the first drop-down list; the list will contract when access is established. This may take some time.
5. If not automatically entered, specify genome and chromosome from the drop-down lists. These must match any data you are viewing already.
6. Set the range min and max if you did not zoom in on the region of interest in step 2 above. The recommended maximum range is about 2 Mb.
7. Click the checkboxes for the specific info you seek. If none are available, there may not be data in the data source you have chosen for the sequence you are investigating.
8. For information about the resources listed under the UCSC option, visit

http://genome.ucsc.edu/goldenPath/help/hgTracksHelp.html#IndivTracks

9. Click **OK**.

Loading data from a DAS/2 server

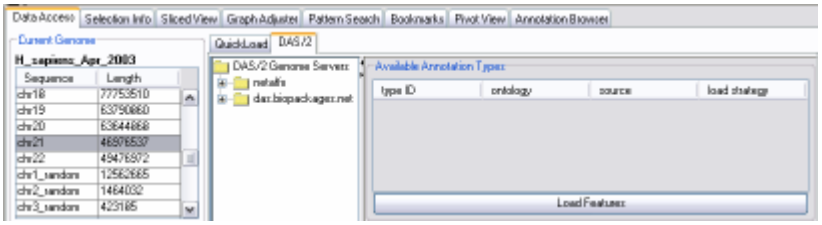
DAS servers are described in Loading annotation types from a DAS server on page 11.

For more information on the DAS/2 specification, visit http://biodas.org/.

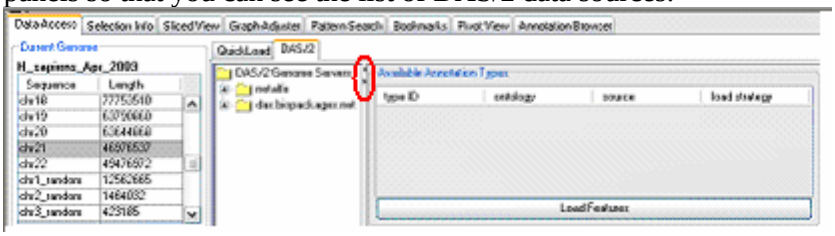
Access DAS/2 servers using the method described in the procedure in this section. Access DAS/1 servers using the method described above11.

To load annotation types from a DAS/2 server:

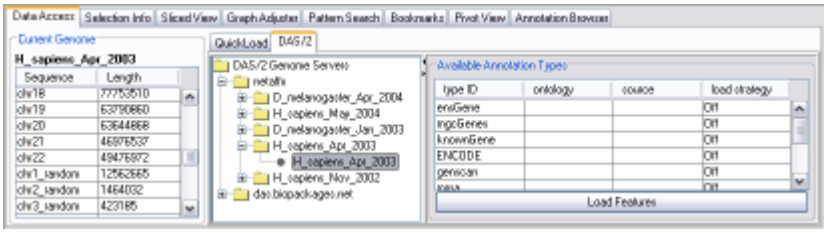
1. Make sure you are connected to the internet. (This may not work on slow connections)
2. Click the **Data Access** tab. (The Data Access panel is already displayed by default when you launch IGB.)
3. Click the **DAS/2** sub-tab.



4. If necessary, drag the arrows indicated in the illustration to resize the panels so that you can see the list of DAS/2 data sources:



5. Click a DAS/2 server folder to view the genome assemblies that are available from that source, and then click the genome assembly folder to display the assembly.
6. To load RefSeq annotations and see a list of the annotation types available from the source you selected, click the genome assembly in the folder:



7. Choose which annotations to load: For each annotation type, click in the **Load strategy** column (by default, **Off** is displayed) for that annotation type and choose the annotations to load:
    - Annotations for the entire genome (Choose **Whole Sequence**). Loading a whole sequence may take a very long time if the server or your connection is slow.
    - Annotations for the range of bases that you currently see in the viewer (Choose **Visible Range**). Loading large ranges may take a long time if the server or your connection is slow.
    - No annotations from the annotation type (Choose **Off**).
8. Click the **Load Annotations** button.

    The annotation types that you loaded appear in separate tracks in the viewer. If you specified **Visible Range** and an annotation extends beyond the range that you specified, the entire annotation is loaded.

    The next time you access this server, IGB will automatically load the same version and sequence that you loaded from this server this time.

The columns that appear in the table in the DAS/2 sub-tab after you select an assembly depend on the information available from the source you selected. If data does not appear in the Ontology and Source columns, the DAS/2 source that you chose does not include this data.

Some columns you may see include:

- Source: The derivation of the feature. It can be used to indicate the algorithm or curatorial procedure used to generate this type of feature.
- Ontology: The sequence ontology category that describes the selected annotation type.

After you have accessed a particular sequence from a particular source, IGB caches that sequence so that it loads quickly the next time you use it. To clear the cache, see Clearing the cache on page 65.

Loading a file of data

IGB can view data from one or more files. A variety of file formats is supported.

All filenames must include the appropriate filename extension, for example ".psl". Compressed files can be read if a recognized filename extension is part of the filename, for example "myfile.psl.gz" or "myfile.psl.zip".

Viewable file formats

IGB can display several types of standard genomics file formats.

There are two types of genomic file formats: those which describe annotations, and those which give sequence residues.

Graph file formats are addressed separately.

Annotation file formats

These files give sequence coordinate locations of exons, genes, transcripts, etc. on the genome. Some accepted formats are described in the table below.

.gff	More info is available at: http://www.sanger.ac.uk/Software/formats/GFF/index.shtml or http://genome.ucsc.edu/goldenPath/help/customTrack.html#GFF
.gtf	For information, visit http://genome.ucsc.edu/goldenPath/help/customTrack.html#GTF For IGB use, the track name is the **source name**. **Score** and **frame** will be ignored. Coordinates will be transformed from *Base 1* to *Interbase 0*.
.psl	Displays annotations with **target name** matching the loaded sequence. For additional information about .psl files and how to use them, see http://www.soe.ucsc.edu/~kent/exe/doc/psLayout.doc or http://genome.ucsc.edu/goldenPath/help/customTrack.html#PSL. Files of interest are available via the DAS server. When you load them, you will be prompted to specify whether to use the Query or Target sequence. See Comparing genome to genome alignments.
.bed	Standard genomic file format. For information, visit http://genome.ucsc.edu/goldenPath/help/customTrack.html#BED

Sequence file formats

Sequence files contain sequence residues only. If an annotation file you use lacks sequence residues, you can merge a sequence file for the same region.

IGB can use the following sequence file formats:

- .fasta
- .fa

Files of either of these types that are compressed. See below.

Compressed data files

IGB can also read compressed files of any of the above file types with any of the following extensions:

- .zip

- .Z (case-sensitive)
- .gz
- .gzip

The filename must include both the filetype extension and the compression extension, such as “myfile.psl.gz”. Zip files must contain only a single annotation file.

Viewing files

You can view a single file, view multiple files or sources of data in a single instance of IGB. Launching separate instances of IGB simultaneously is not recommended.

Viewing a single file

Use this option to view a file that you don’t intend to compare with any other annotation tracks. Anything already displayed in the viewer will be obliterated when you use this option.

To view a single file of data:

- Choose **File** menu > **Open File.** Uncheck the checkbox to **Merge with currently-loaded data**.

Viewing multiple files

To view and compare annotations from multiple files in a single IGB window, or if annotations from other sources are currently loaded:

1. Make sure the genome version and chromosome of files that you want to open match those of the data already displayed.

   IGB recognizes standard terms for version and sequence (For specifics, see the Synonyms file on the QuickLoad server: (http://netaffxdas.affymetrix.com/quickload_data/synonyms.txt).

   If in doubt, check the text at the beginning of the file to be opened. If genome version and chromosome cannot be matched, you may still be able to open a file if the sequence name matches.

2. Choose **File** menu > **Open File.**
3. Make sure that the **Merge with currently loaded data** checkbox is selected.
4. Choose the file and click **Open**.

Annotations from merged files will load as additional tracks alongside tracks already in the viewer.

If merging brings in the file as a narrow band at the top of the viewer panel, use the vertical scroll box at the right side of the viewer to make the annotations easier to see, and to view the name of the annotation type in the tracks panel. For instructions, see Expanding vertically.

Other sources of data

Other sources of loadable data include:

- Bookmarked data from IGB, imported
- Bookmarked data from IGB, via web browser

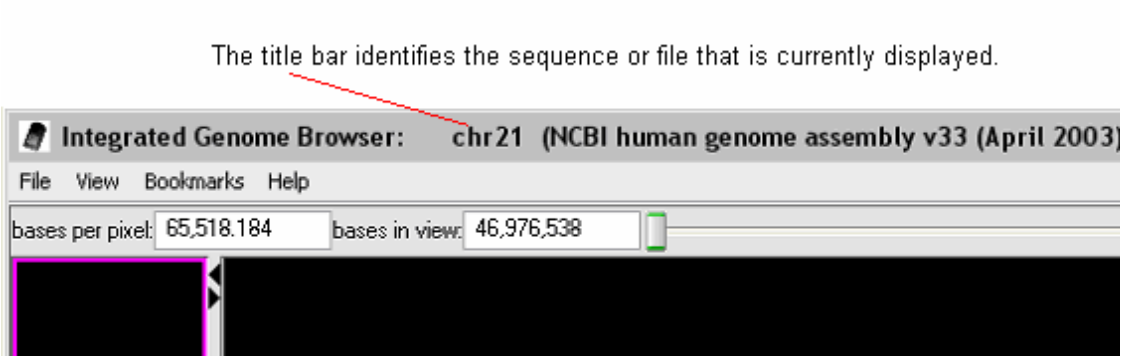
Viewing, navigating, selecting, and finding data

This section describes different ways to:

- determine what you're looking at
- view annotations and information about them
- navigate around a genome assembly or other loaded data set
- zoom in on areas of interest
- find annotations, sequences, or patterns
- select annotations for use with various features of IGB

Identifying the sequence displayed in the viewer

If you are viewing a standard data set, the genome version and chromosome or sequence being viewed are identified in the window title. If you are viewing a file, the filename will be displayed.



Clearing the IGB viewer

To clear the contents in the viewer in order to work with a different data set:

- Choose **File** menu > **Clear All**

QuickLoad annotations and any loaded sequence residues remain loaded; click a chromosome to display it. Bookmarks remain in the Bookmarks menu.

Loading and viewing sequence residues

Many IGB features require sequence residues to be loaded, so you may want to load them initially during each IGB session.

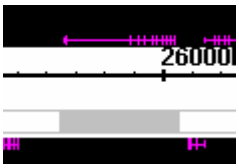
Loading sequence residues

The method of loading sequence residues differs depending on the data source you are viewing.

For QuickLoad data

Load sequence residues separately for each chromosome, or load them for just the region of a chromosome that you are currently viewing.

To load sequences residues for the region that is currently visible in IGB:

1. Zoom in to display the desired region of the chromosome.
2. Click the **Data Access** tab > **QuickLoad** sub-tab > **Load sequence residues in view** button.
3. Wait while the sequences load. Sequence residues contain a large amount of data, so this may take a few minutes.
4. A gray bar at the bottom of the Coordinates track indicates the region(s) for which you have loaded sequence residues:
5. 

To load sequence residues for the entire chromosome selected:

1. **Data Access** tab > **QuickLoad** sub-tab > click a chromosome in the list > **Load all sequence residues** button.
2. Wait while the sequences load. Sequence residues contain a large amount of data, so this may take a few minutes. Your screen may go gray, but this is normal. The sequence data will be cached and will thus load faster the next time, unless you have turned off the caching feature.

By default, sequence residues are loaded from http://genome.cse.ucsc.edu/cgi-bin/das. To change the source, click **Data Access** tab > **QuickLoad** sub-tab > **QuickLoad Options** button. Enter the new source in the **DAS DNA Server URL** field.

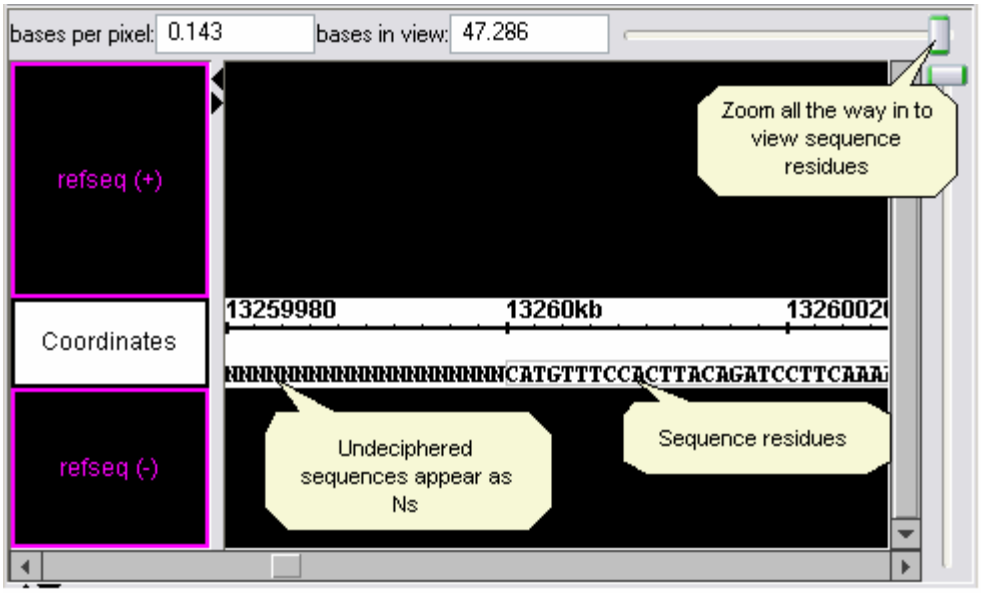
If loading over the internet is too slow and you are an advanced user, consider copying the sequence residues from the QuickLoad server onto your local network. See Clearing the cache on page 65.

For file data

For file data, if sequence residues are not included in your file, merge a .fasta sequence file for the identical sequence.

Viewing sequence residues

Sequence residues, when loaded, are visible in the Coordinates track when zoomed all the way in (see picture).



Unsequenced regions appear as Ns.

Navigating around the genome

In addition to the following navigation methods, see also Finding sequence patterns on page 25.

Going to chromosome

If you have loaded a standard reference annotation type, to go to chromosome:

- **Data Access** tab > **QuickLoad** sub-tab > click a row in the table at the left of the panel.

Going to an annotation

There are several ways to go to a particular annotation:

- Click the **Pattern Search** tab, then enter the (case-sensitive) identifier for the annotation you want to jump to in the second box (**enter id of annotation to find**), then press the <Enter> key. (Ignore the other options in this panel.)
- Click the **Annotation Browser** tab, then click the annotation ID you want to go to. Wait a moment for your annotation to display.

IGB displays the region of the annotation, if it is present in the data set being viewed. If QuickLoad data is displayed, the window switches to the chromosome the annotation is on.

See also Zooming in on an annotation on page 22 and the **Select parent** feature in Selecting annotations or sequences to work with on page 24.

Going to a specified coordinate

To go to the coordinate you specify and focus the zoom on it:

1. Click the **Pattern Search** tab.
2. Enter the coordinate number you want to jump to in the first box (**enter coordinate to center on**) then press the <Enter> key

If you have clamped to a view (see page 23) or are looking at a file that does not include the coordinate you want to jump to (and a standard reference annotation type such as QuickLoad is not loaded), this feature does not work.

Scrolling along the chromosome/ file data

When you are zoomed in, you can browse the view in several ways:

- Automatically scroll along the chromosome. Choose **View** menu > **AutoScroll**.
- Use the horizontal scroll box slider at the bottom of the viewer panel to scroll along the sequence. Slider size and speed of movement will vary according to the current zoom level.
- Click in the scroll bar on either side of the slider to move it in the direction of your click.
- For finest precision when zoomed in, click the arrows at the ends of the scroll bar.

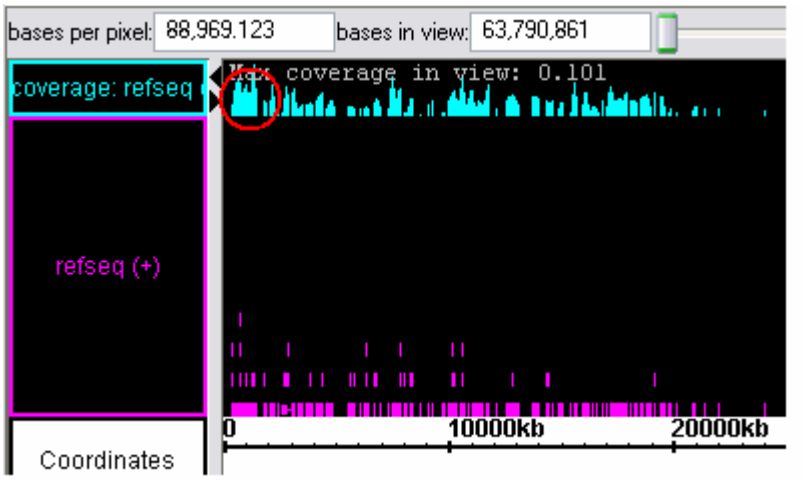
To locate the horizontal scroll bar and its scroll box slider, see the illustration in the Zooming in on a coordinate or region section on page 21.

When you are zoomed all the way out (the horizontal zoom control slider is at the far left), you are already viewing the entire range of the loaded data and the scroll box slider is not available.

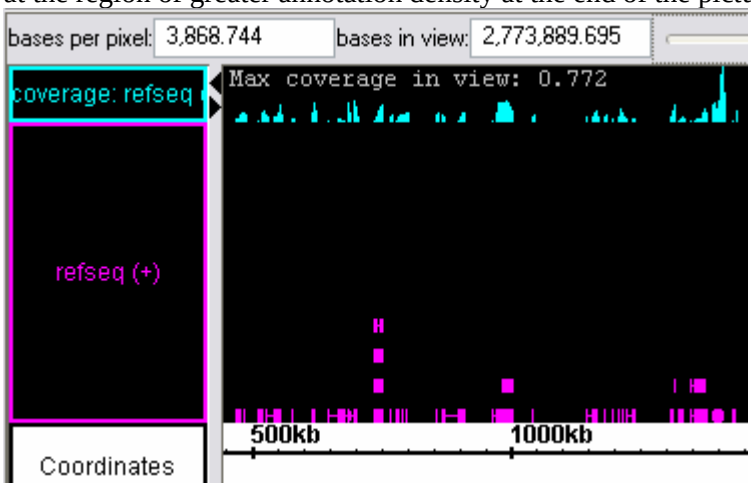
Looking at density of annotation coverage

You can see an overview of the relative density of annotation coverage across an entire chromosome. The Coverage Track feature dynamically graphs the relative number of annotations in a track that are present at each pixel across the main window. As you zoom in and each pixel represents a smaller number of nucleotides, the graph automatically adjusts to indicate the relative number of annotations present in the smaller region covered by that pixel. When you are completely zoomed in, and each pixel represents a single nucleotide base, the graph shows either a 0 (no annotation is present) or a 1 (an annotation is present) for each pixel.

The following illustration shows the relative annotation coverage at each pixel when the view is zoomed all the way out. In this example, suppose you conclude that the density of the circled region is significant and you want to examine it more closely.



After you zoom in to display the relative annotation density within the region that is circled in the illustration above, you might decide to look more closely at the region of greater annotation density at the end of the pictured region.



To show the density of annotation coverage for an annotation type that you are viewing:

- Right-click on the annotation type in the Tracks panel and choose **Make Annotation Coverage Track**.

Zooming and locking

Zoom in on an area of interest, and then lock on it to protect your view.

Slicing (see page 36) is a separate feature that shares some similarities with zooming.

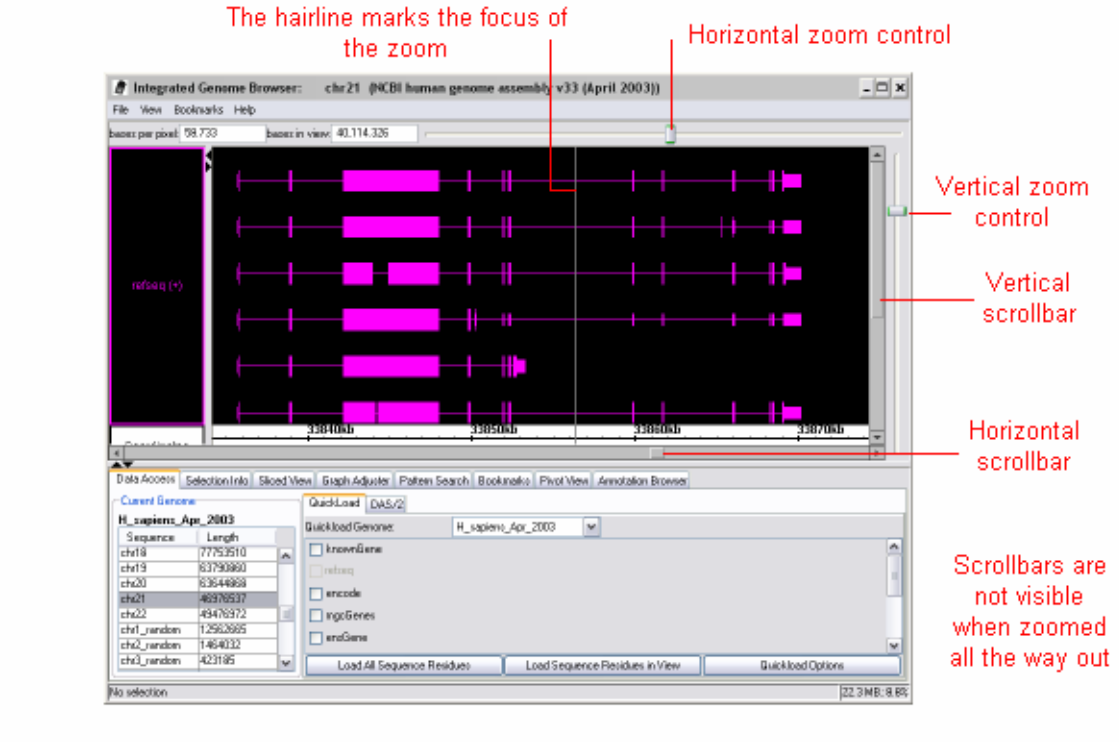
Zooming in on a coordinate or region

Click anywhere in the viewer to set the focus of the zoom. A hairline will mark the zoom focus.

To zoom in, move the slider at the top of the viewer to the right. To zoom out, move the slider at the top of the viewer to the left.

To see the exact number of the coordinate marked by the hairline, choose **View** menu > **Toggle hairline label.** The hairline will display the coordinate number with each subsequent click. To turn off the label, choose **View** menu > **Toggle hairline label** again.

To keep the hairline always in view when you scroll, you can set a preference. See the Setting other options section on page 89 of Appendix 3 Preferences.



See also Going to a specified coordinate on page 20.

Zooming in on an annotation

To center on and fill the viewer with the selected object:

- Right-click on an item > **Zoom to selected**

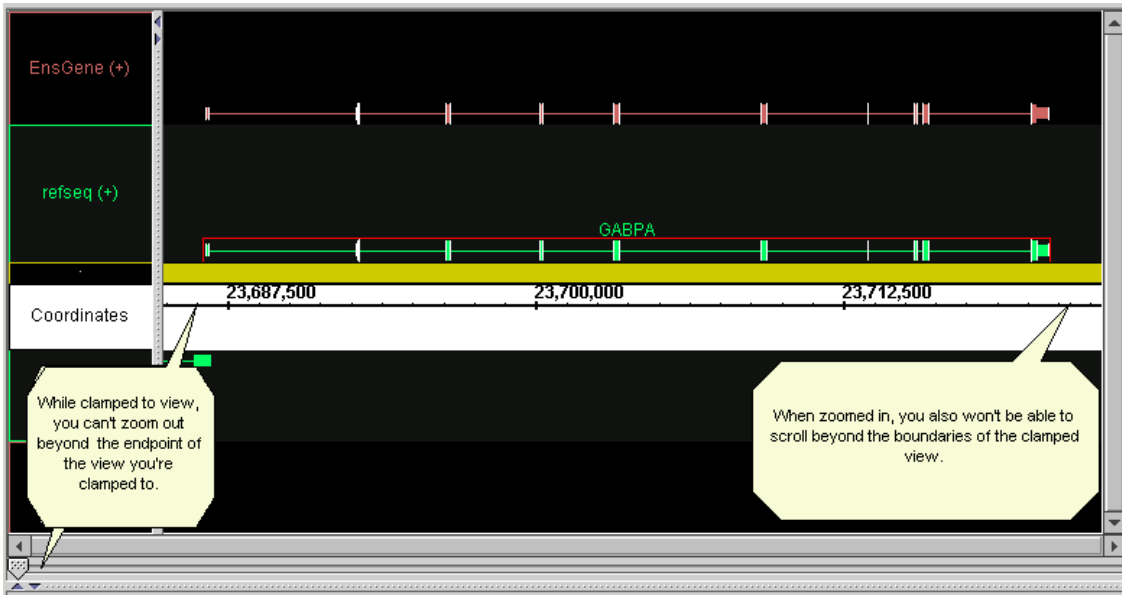
To zoom out:

- Slide the horizontal zoom control back to the far left.

See also Going to an annotation on page 19.

Clamping on a view

Normally, if you are zoomed in on a small region, it is very easy to scroll away from the target region and lose it. Clamp to view so you don't lose what you're examining as you browse the region using the zoom and scroll sliders.



Clamping restricts the max and min coordinates of the view to the boundaries of the clamped view, so that you cannot zoom out or scroll beyond the range you've clamped to.

For example, if you want to examine the gene GABPA more closely, zoom in so the gene is the width of the window, then clamp to view. When you then zoom in further for a closer look, the scrollbar now will not scroll beyond the begin- and end-points of the gene.

Clamping also allows finer control when scrolling by moving the horizontal slider.

To clamp to a view:

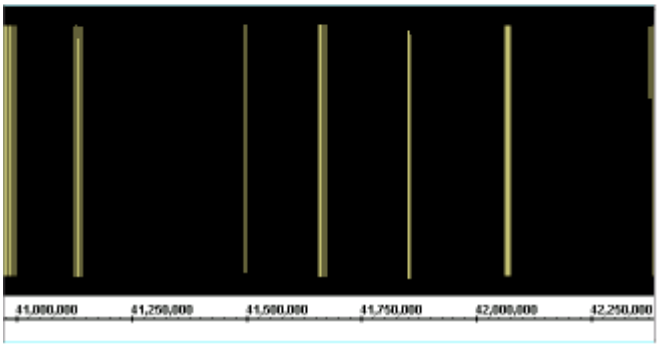
1. Zoom in on a region of interest. If you're zoomed all the way in or out, you can't clamp.
2. Choose **View** menu > **Clamp to View**
3. To unclamp and zoom out to or scroll along the entire chromosome or file content, choose **View** menu > **Unclamp**.

Zooming in and locking on a small region

When working with a file that contains annotations that only cover a small region of the sequence (for example, a set of annotations on a contig rather than on an entire chromosome), you can tell IGB to ignore most of the sequence range and zoom in on the region of interest:

Before shrink wrapping:

After shrink wrapping:



To zoom in on a focused region in a file:

- Choose **View** menu > **Toggle Shrink Wrapping**.

To turn off shrink wrapping:

- Again choose **View** menu > **Toggle Shrink Wrapping**.

Note that this will not work if you have also loaded other annotation types that have no empty black space to eliminate.

Selecting annotations or sequences to work with

Select annotations in order to perform certain operations on them. The following are ways to select items or sequences in the viewer window.

To do this:	**Do this:**
Select a single annotation	Click on it.
Select multiple annotations	Drag the mouse through a rectangular region to select all items touched by or enclosed in the drag boundaries. A drag must begin in the empty area, not on top of any annotation. Dragging may select both the parent genes and the individual exons. To select only the parent genes, right-click in the empty background area

	and choose **Select parent**
Select a sequence region	Drag along the axis of the coordinates track to select a set of residues. To select residues, the drag must begin in the axis tier.
Deselect everything	Click the empty black space in the viewer window.
Select the gene(s) containing a selected exon(s)	Right-click in the blank background or on a selected annotation > **Select parent** You can directly select a whole gene by clicking on the line connecting the exons.
Add to the current selection	Shift-click to add individual items or shift and drag to add multiple items.
Bring-up a pop-up menu	Right click on an item that is already selected or on the empty background.

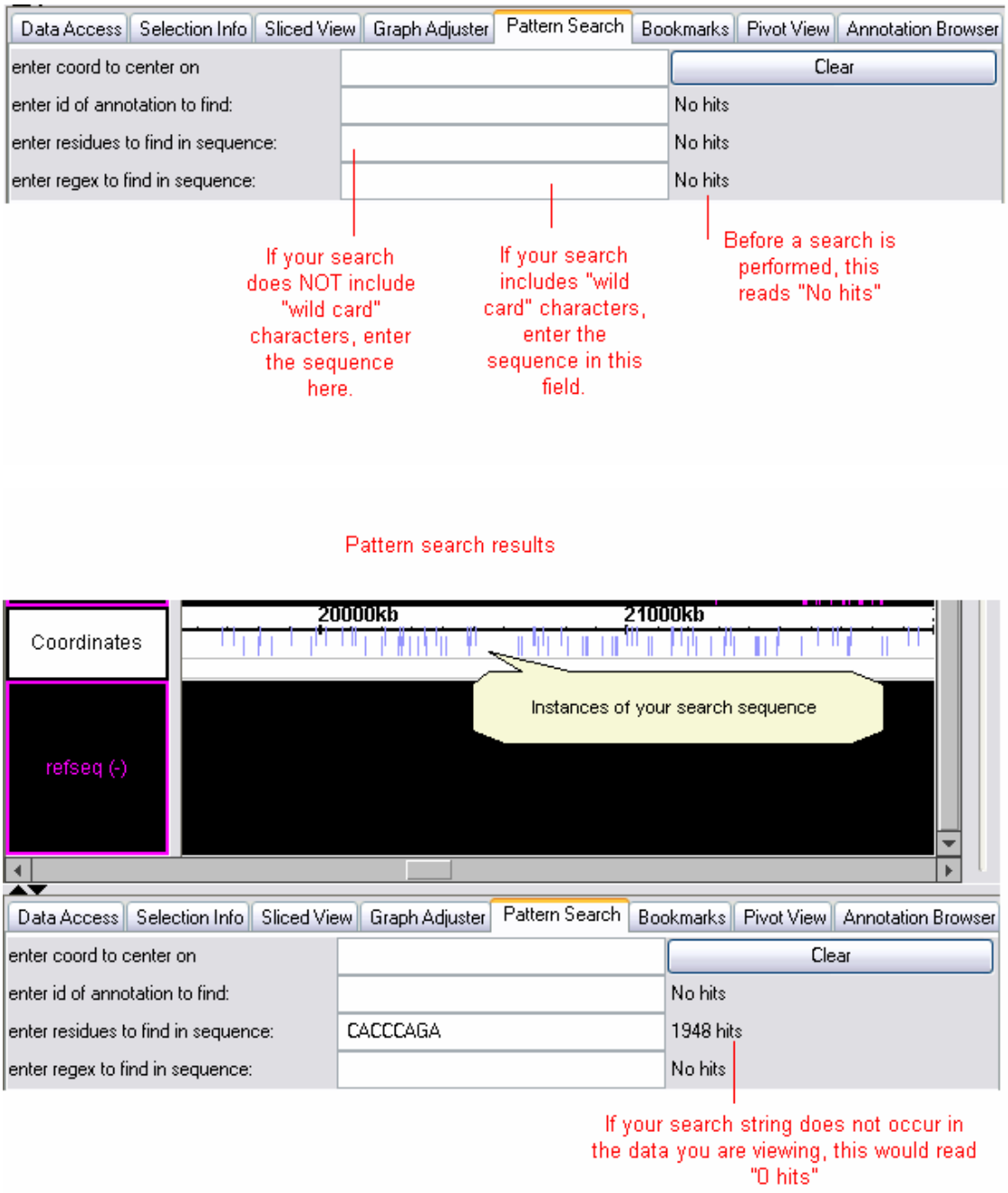
Important: Mistakenly selecting the exon rather than the gene may cause some functions to appear not to work. To select the parent gene, right-click in the blank background or on a selected annotation > **Select parent**.

If selecting seems slow, and you are working with a sliced view (see page 36), do the following:

- Choose **Sliced View** tab > uncheck **Slice based on selection**.

Finding sequence patterns

Use the Pattern Search feature to locate each instance of a short sequence you specify, either exact bases or including placeholders for unknown bases. All instances of that sequence on the chromosome will be marked on the Coordinates track.



Searching by ID works only for data loaded as files or QuickLoaded standard reference annotation types, not data loaded from remote servers (that is, via DAS servers). Sequence residues must be loaded before you search based on residues.

“Enter residues to find in sequence”. This searches for an exact sequence of residues. It is not case-sensitive and it searches the forward and reverse strand.

“Enter regex to find in sequence”. This searches for sequence data based on a regular expression. Only the forward strand is searched. This search is

case-sensitive for efficiency. Fasta files may contain both uppercase and lowercase sequences (lowercase is for repeating sequences); if you are working with data from a .fasta or .fa file, search for upper and lower case separately as desired. It is possible to make regular-expression searches be case-insensitive by using the flag "(?i)" at the beginning of the regular expression.

Found instances of your search sequence will be marked in blue on the Coordinates track (see the illustration above). Instances on the forward strand are positioned slightly above instances on the reverse strand.

The number of instances found appears in the tab window next to the field in which you entered the sequence string.

No hits means no search was performed for that function. **0 hits** means no matches were found. (See the illustration.) Watch this indicator to know when IGB is done processing your search. The cursor will also begin to blink in the text-entry field when processing is done.

Finding instances of a known sequence

In the chromosome being viewed, to identify all instances of an exact base sequence on the forward or reverse strand:

1. See the information in the previous section.
2. Click the **Pattern Search** tab and enter a sequence or select a sequence of bases in the Coordinates track (see Selecting annotations or sequences to work with on page 24), then paste the bases in the third text-entry field (**enter residues to find in sequence**).
3. Press the <Enter> key. Wait a few moments.

Finding instances of a sequence containing unknowns

In the chromosome being viewed, identify all sequences on the forward strand matching a regular expression (a sequence containing "wild card" nucleotides):

1. Read information about Finding sequence patterns on page 25.
2. Click the **Pattern Search** tab and enter a case-sensitive string of bases and wild cards (see table below for common ones) in the text-entry field called (**enter regex to find in sequence**). Press the <Enter> key.
3. Wait a few moments; this may take some time.

Example wildcards:

Wild card	Represents	Example entry	Finds sequences
.	any single nucleotide	ACCT.T	ACCTTT, ACCTAT, ACCTGT, and ACCTCT

...	any three nucleotides	ACCT…T	ACCTGAGT, Etc.
[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, and ACGGTTTC
T*	Zero or more T's	ACGGT*C	ACGGC, ACGGTTTTC, 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.

The full list of regular expression syntax that you can use in IGB is available at http://java.sun.com/j2se/1.4.2/docs/api/index.html.

For more detailed information about wildcard searches in general, see Friedl's book *Mastering Regular Expressions*.

Deleting pattern markers from the Coordinates track

To delete pattern markers from the Coordinates track:

- In the **Pattern Search** panel, press the **Clear** button.

Getting further details about annotations

IGB can provide specific information about annotations in various ways.

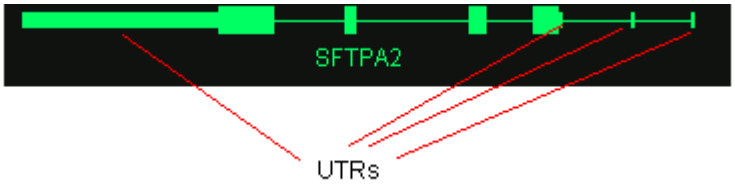
Annotation ID's

To show or hide annotation IDs, see Choosing annotation labels on page 83.

To make annotation names visible if they are not hidden but you cannot see or read them, use the vertical zoom slider at the right side of the viewer window to increase the annotation size until the text is visible.

Untranslated Regions (UTRs)

UTRs are indicated graphically by a "tongue" at either end of a gene, or a thinner-than-usual exon, as in this example:



Only some annotation types and files will indicate UTRs.

Item details as text

View start, end, chromosome, and strand for all annotations in the loaded sequence that have ID's

1. Click the **Annotation Browser** tab.
2. Scroll down to the annotation of interest.

Details for currently selected features

To get details about annotations, graphs, or primers:

1. Click the **Selection Info** tab
2. Select the item(s) of interest in the viewer. These can be an annotations, primers, or graphs.
3. Shift-click to select more than one item.
4. If you click an item and less information than you expect appears in the Selection Info panel, you may have selected an exon rather than a

gene. Right-click on the exon or on the blank background and choose **Select parent** to see information about the gene.

The kinds of information displayed vary by item selected and source.

Selection Info shows one column for each gene that you shift-click to select in the viewer.

Viewing a region in the UCSC web genome browser

The UCSC web genome browser is a rich source of genomic information.

For more information on what's available there, visit http://genome.ucsc.edu/.

To view in the UCSC browser the same region currently displayed in IGB:

1. Select an item in the viewer window
2. Choose **View** menu > **View region in UCSC browser**

Results may be slow to appear. (You may want to set your web browser preferences to show loading progress -- for example, Microsoft Internet Explorer has a bar on the bottom of the web window. See the documentation for your web browser.)

See also Loading data from a DAS/1 server on page 11, plus the following section related to DAS/2 servers, to load annotations from UCSC and other sources into the IGB viewer.

Viewing information on public web sites

For many annotation types, more information about an annotation is available directly from the source. This feature automatically takes you to NCBI (for RefSeq annotations), Ensemble (for EnsGene annotations), UCSC (many DAS annotations), or other web site depending on the source of the annotation you seek information on. Not all annotation sources contain the necessary link information for this feature; if the menu option is unavailable for the annotations you are examining, this means that the annotation type has not been associated with a website.

To view information for an annotation from a public web site:

- Right-click a gene in the viewer > **Get more info**

  Make sure you're selecting the gene, not just an exon.

Comparing data sets

You can view and compare information from different sources, customize the tracks, and use unique IGB tools for analyzing and comparing.

View and compare annotations and graph data from various and multiple sources:

- Graphs (Tools for comparing graphs are discussed in the Working with graphs section)
- QuickLoad data
- Data from DAS servers
- Files (Merged)
- Bookmarks (imported or via Web)
- IGB tools such as pattern searches

Working with tracks

Each annotation type that you load appears in its own track. You can facilitate viewing and comparison by customizing the appearance of each annotation type and by adjusting the position and size of tracks relative to each other.

You can manipulate tracks in the following ways:

- Changing the color of annotations in a track
- Changing the width of the tracks panel
- Re-ordering tracks top-to-bottom
- Deleting tracks
- Hiding and Showing (Unhiding) tracks
- Expanding all tracks vertically
- Collapsing and Uncollapsing tracks
- Limiting the number of annotation rows in a track
- Restoring the view of all annotation rows in a track
- Permanently setting the appearance of tracks
- Working with multiple tracks

Changing the color of annotations in a track

When you change the color of annotations in a track, the changes will remain after you close and re-launch IGB and load the same data set.

To change the color of annotations in a track:

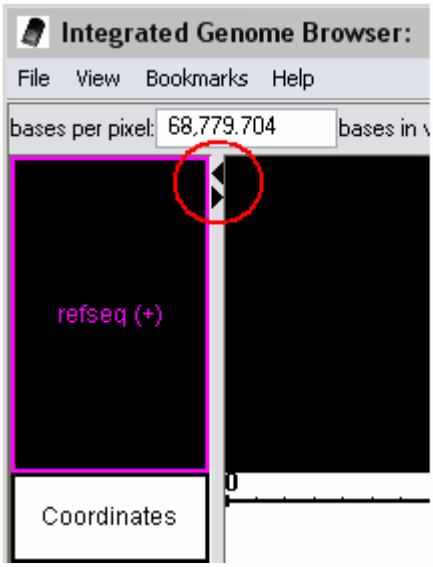
1. Right-click the track name and choose **Change Annotation Color**.

2. Click a color in the palette, then click **OK**.

Changing the width of the tracks panel

You can make the Tracks panel wider or narrower. For example, the name of the annotation type in a track may be too long to fit into the default track size, and you may want to make it easier to read.

To resize the tracks column, drag the separating bar or click the tiny arrows:



Re-ordering tracks top-to-bottom

To re-order the tracks:

- Click a track name (in the panel on the left) and drag the track up or down to the desired location.

Deleting tracks

Tracks can't be deleted; hide them instead. See the next section.

Hiding and Showing (Unhiding) tracks

To hide a track:

- Right-click on the track or tracks you want to hide > **Hide**

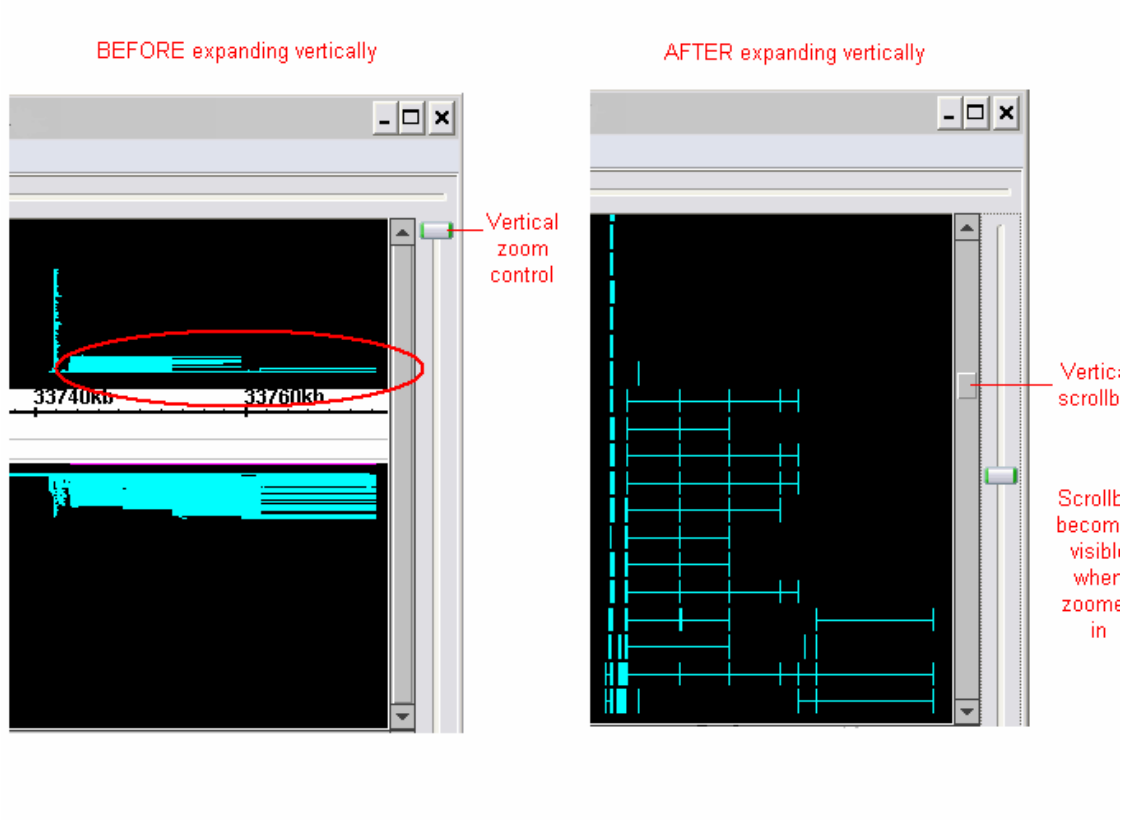
To show hidden tracks:

- Right-click on any track > **Show All**. If you moved the track, it may reappear in its original position.

Expanding all tracks vertically

This feature is especially useful for viewing ESTs or any other annotation types with many rows of annotations.

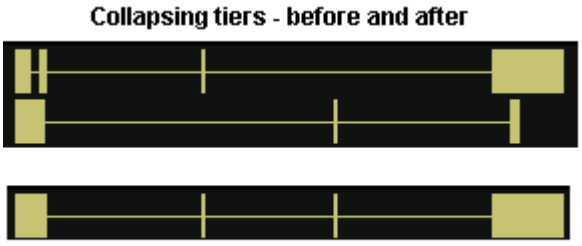
To vertically expand the view, drag the vertical slider:



After expanding vertically, use the vertical scrollbar to view different annotation rows.

Collapsing and Uncollapsing tracks

For tracks containing multiple rows of annotations, collapsing tracks consolidates all rows within a track into a single row. Annotations on any of the rows will be included in the collapsed track. Larger annotations may obscure smaller ones; annotations with introns may be obscured by annotations that don't show the intron.



To collapse or expand one or a few tracks:

- Right-click on track name(s) > **Collapse** or **Expand**

To collapse or expand all tracks at once:

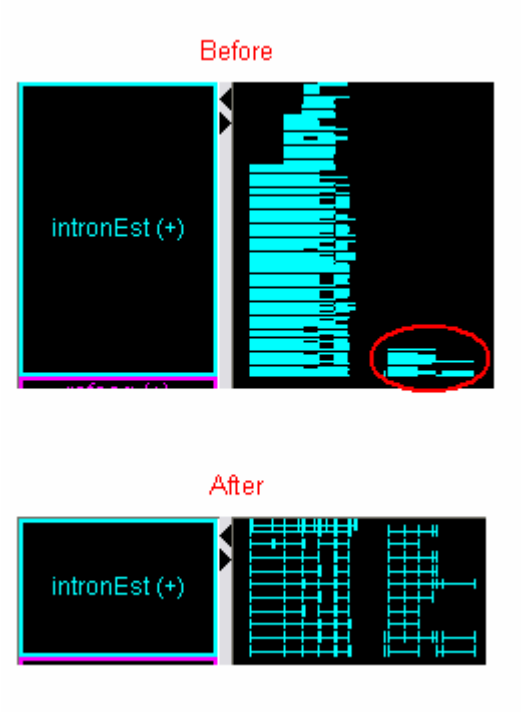
- Right-click on track name > **Collapse All** or **Expand All**

Changes that you make will apply each time you load this annotation type. See also Collapsing annotation rows into a single row in the Preferences section.

Limiting the number of annotation rows in a track

If a single track contains many rows of annotations, some annotations may be represented by many instances while others have few representations. To focus on annotations that have fewer representations, or to reduce the amount of space that a track occupies in the viewer, limit the number of rows that IGB displays.

In the example below, to make the circled annotations more visible, the number of annotation rows in the track was limited to 5:



To limit the number of annotation rows in a track:

1. Right-click on the track > **Adjust max expand**
2. Enter the new maximum track height
3. Click **OK**

To limit the number of annotation rows in *all* tracks in the viewer:

1. Right-click on the track > **Adjust max expand all**
2. Enter the new maximum track height

3. Click **OK**

To set the number of annotation rows to display each time you load a particular data set, see the Preferences section Setting the maximum number of annotation rows to display on page 82.

Restoring the view of all annotation rows in a track

To return to the original view and make all annotation rows in a track (or in all tracks) visible:

1. Right-click on a track > **Adjust max expand** (or **Adjust max expand all**).
2. Enter **0** for the new maximum track height
3. Click **OK**

Permanently setting the appearance of tracks

To set other preferences that apply each time you work with a particular annotation type, see the Preferences section, Customizing the default appearance of tracks on page 80.

Working with multiple tracks

To select more than one track, hold down the Shift key while you click each track. To perform right-click operations on multiple tracks simultaneously, right-click one of the selected tracks to display the menu of track options.

Tools for comparing annotations

IGB offers several tools to highlight patterns to aid in making comparisons:

- Highlighting matching endpoints
- Slicing
- Summarizing tracks
- Combining tracks

Highlighting matching endpoints

To see if different data sources show the same start and endpoints for an annotation:

- Select an item; matching edges of corresponding items in all tracks being viewed will be highlighted in white.

To specify the degree of closeness required for a match (how many base pairs difference is still considered a match):

1. Choose **View** menu > **Adjust Edge Match fuzziness.**
2. Slide the slider to the maximum number of base pairs difference that you want to consider a match.

By default, if edge match fuzziness is adjusted, all matching edges are highlighted in gray rather than in white.

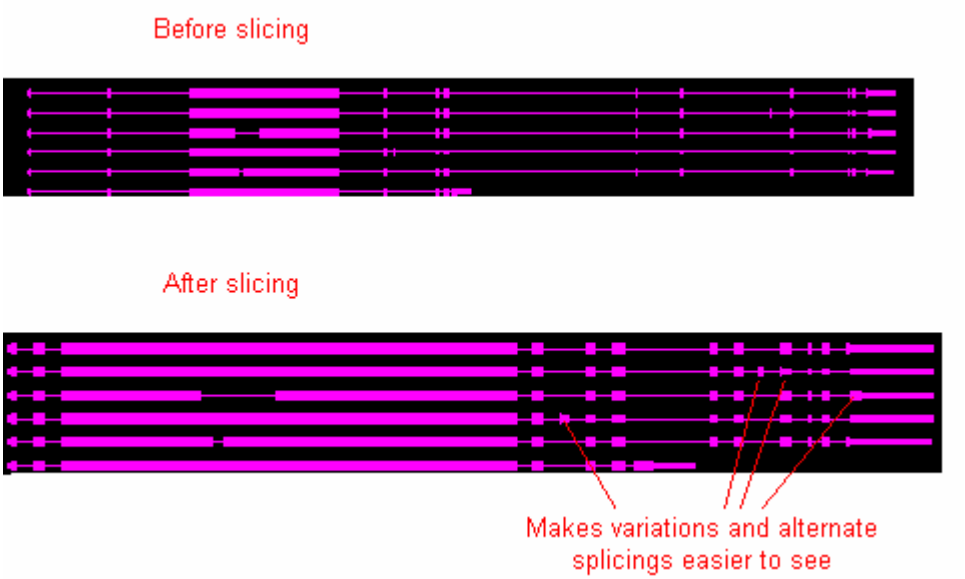
To change the default colors of exact and/ or fuzzy matches:

- Choose **View** menu > **Adjust Edge Match fuzzinesss**

You can also set preferences for the colors that indicate exact and fuzzy matches. See the Setting edge-matching colors section in Appendix 3 Preferences.

Slicing

Slicing makes slight variations more visible when comparing multiple annotations on the same sequence. Slicing cuts out long introns and realigns the exons of a selected annotation or set of annotations, making pattern irregularities and possible alternate splicings more obvious.



To slice a selection:

1. In the main viewing window, select all annotations to include in the slice. In the example above, you would select all 6 “copies” of the gene.
2. Click the **Sliced View** tab.
3. If desired, resize the tab panel or window to enlarge the viewing area of the Sliced View window.
4. Make sure the **Slice By Selection** checkbox is checked.

IGB may take a few moments to display the sliced region. For easy comparison, the sliced view displays all annotations and graphs in the region being viewed.

Note that the numbers in the Coordinates track in the sliced view panel indicate scale; they don't correspond to genomic coordinates. Endpoint matching in the upper and lower windows are independent of each other.

To examine the sliced view:

- Use the zoom and scroll bars in the sliced view panel.

Note: If you use keyboard shortcuts for zooming and scrolling, the shortcuts will apply to both 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 in a separate window can give more predictable shortcut behavior.

You can view the sliced region in a separate window by opening the tab in its own window. See page 7.

To further enhance viewing and comparison, especially of ESTs, summarize the track after slicing. See page 37.

Adjusting the sliced view

Each sliced region includes a buffer, in order to allow examination of the bases in the intron-exon boundary. The buffer is initially 100 bps on each side of each exon; if an exon in the same range that was not selected for slicing extends beyond the buffer, it may be truncated. To avoid this problem, increase the buffer enough to accommodate the entire exon by changing the number of base pairs before and after each exon.

To adjust the buffering number of base pairs before and after each exon:

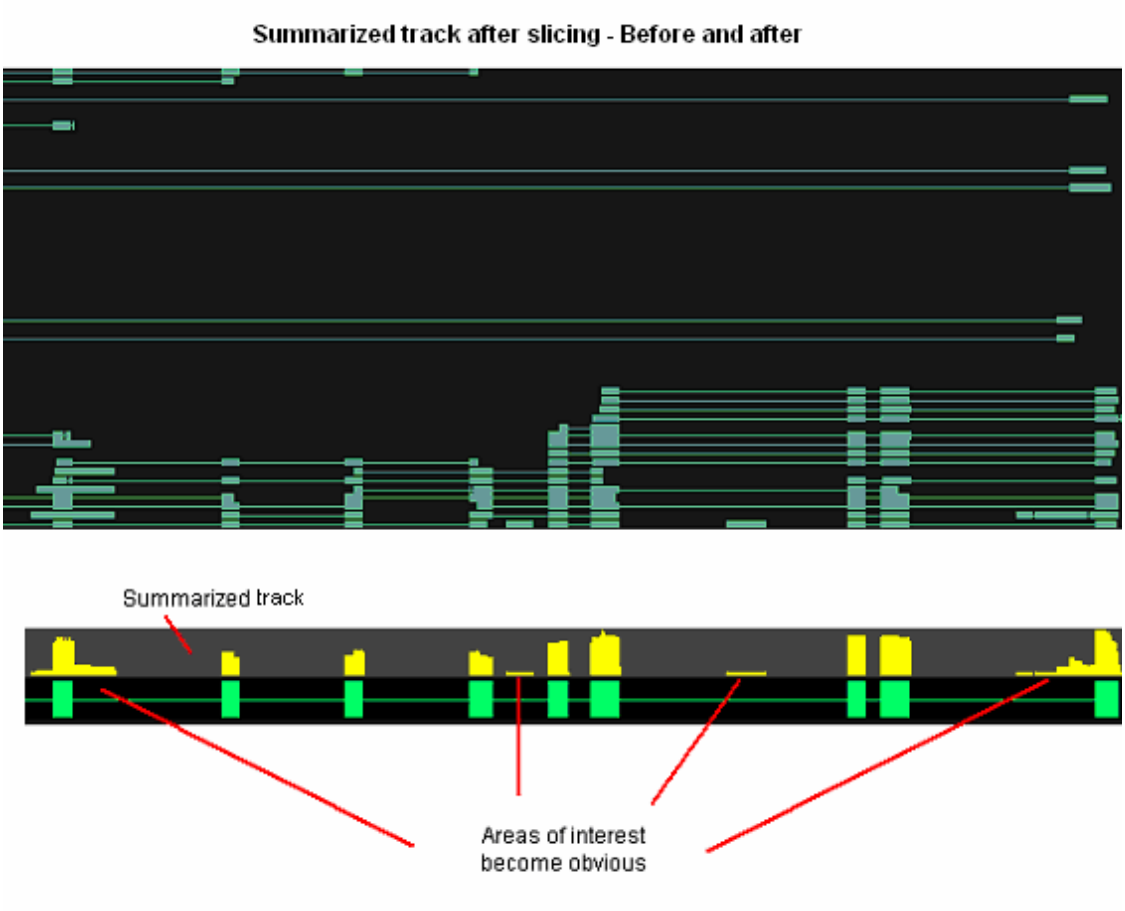
1. Visually estimate the number of base pairs required to achieve the adjustment needed
2. In the **Sliced View** tab, change the value in the **Slice Buffer** field.
3. Press the <Enter> key on your keyboard.

You may also set the slice buffer to 0 to completely eliminate the introns. This can be useful in conjunction with the **Analyze ORFs** function to find open reading frames.

Summarizing tracks

For tracks containing multiple rows of data, summarizing collapses a stack of annotations so that anomalies are readily visible. The height of the results at each position indicates the number of annotation rows that include each annotation. An area where the height is very low compared to the others indicates an anomaly in one or more of the rows.

Summarizing a track is especially useful after slicing (see page 36) and when examining ESTs.



To summarize a track:

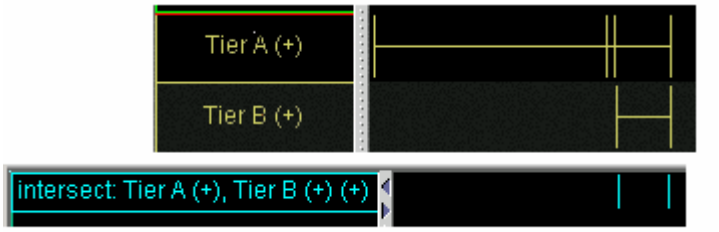
1. If you will use slicing (see page 36) during this analysis, slice now
2. Right-click on the track label > **Make annotation depth track**
3. Wait a few moments for processing.

The summary track appears at the top of the viewer window. The name of the track is "depth: <name of the track you summarized>".

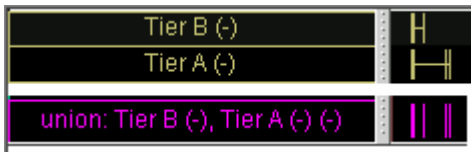
Combining tracks

You can highlight differences or commonalities between different annotations or graph thresholds on two tracks by combining the tracks in any of several ways:

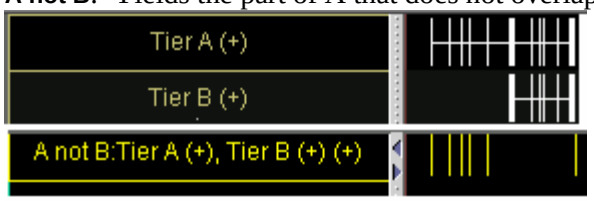
- **Intersect:** Yields only the overlapping portions of A and B



- **Union**: Yields all bases that are in either A or B



- **A not B**: Yields the part of A that does not overlap with B



- **B not A**: Yields the part of B that does not overlap with A

- **XOR**: the union of the yields of (A not B) and (B not A); i.e., all of the material that is in A or B but not both.

- **Not**: (Select only one track) Yields all coordinates that are not in the selected track

For example, if you are comparing graphs, you can capture annotation-like bars resulting from different threshold settings, then compare the captured snapshots, then modify the threshold settings and compare them again.

To compare thresholded data from graphs by combining tracks:

1. If you are working with graphs, first capture the graph thresholds in a track. (See page 52.)
2. Shift-click the names of two tracks to compare. (If you will use **Not**, click only one track.)

3. Right-click a selected track > **Combine Selected Tracks** > choose an expression above for comparing tracks.

Working with graphs

Introduction

Graphs contain numerical data associated with base pair positions along the genomic sequence axis, such as sets of scores in specific positions along the genome. Graphs allow you to visualize numerical data (scores) associated with base positions along the genome.

Examples of suitable data include:

- Files generated from Affymetrix software tools, such as CNAT, GCOS, or ExACT.
- Expression values from whole-genome tiling arrays.
- Density of EST's across a chromosome
- GC content along the genome
- Measures of conservation between two genomes

Overview

Use IGB to make important features in your data easier to see. This chapter explains how to use IGB in a variety of ways to examine your data:

- Adjust graphs to highlight features
- Fine-tune graph scaling to improve visibility
- Compare data from different sources for the same region.
- Use thresholding to transform graph data into annotation-like format, in order to examine it further using other tools in IGB.

Working with GCOS, CNAT, ExACT, and Tiling Array data

For overviews of working with these data types in IGB, see the separate IGB Quick Start guide for each source. See Other IGB documentation on page 5.

Information about working with ExACT data in IGB is located in the documentation for ExACT at:
http://www.affymetrix.com/products/software/specific/exact.affx.

Opening graph files in IGB

Data must be formatted as described by one of the graph file formats below.

Graph file formats

IGB can display graphs in several file formats, all developed at Affymetrix.

All files must include the filename extension, for example ".sgr".

Graph files may be in the following formats:

.bar Generated from tiling arrays by GTAS (GeneChip Tiling Analysis Software.)

For information about working with GTAS data in IGB, see the separate Quick Start guide posted with the most current version of this document on the Affymetrix web site.

.egr Files generated from Affymetrix GeneChip Operating Software (GCOS) or ExACT (Exon Array Computational Tool) software.

For information about generating .egr files from GCOS, see the separate document Visualizing_Affymetrix_Expression_data_in_IGB, posted with the most current version of this document.

For information about generating .egr files from ExACT, see the documentation for ExACT at: http://www.affymetrix.com/products/software/specific/exact.affx

.sgr Sequence graph files that show base coordinate scores. These files are generated by CNAT (the Affymetrix Chromosome Copy Number Analysis Tool software).

For information about working with CNAT data in IGB, see the Quick Start Guide posted with the most current version of this document on the Affymetrix web site.

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.

The name of the file is the name of the graph.

Opening graph files in IGB

To open graph files in any of the supported formats:

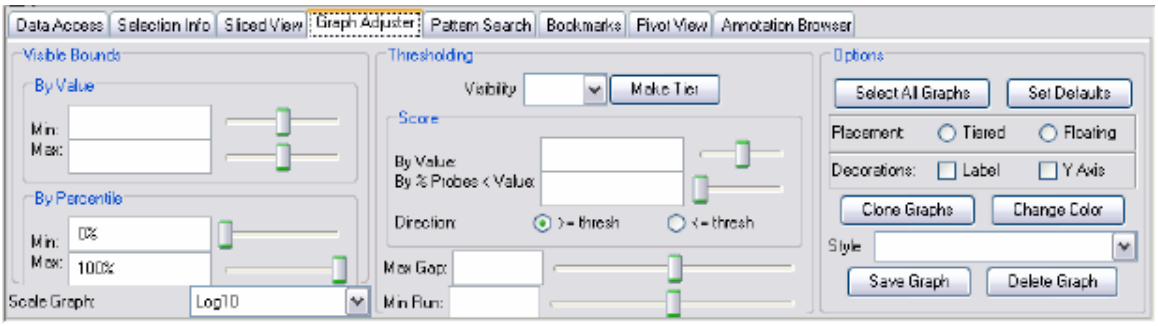
- Follow one of the procedures in Viewing files on page 15.

Important: Graph files may display even if they don't match up with the coordinates currently loaded, but any comparisons will be meaningless.

Changing graph appearance

Settings and adjustments for graphs can be made using the **Graph Adjuster** tab. Note that earlier version of IGB used right-clicking on graphs to change settings; the current version now uses the Graph Adjuster tab instead.

The Graph Adjuster panel will appear in the tabbed panel sections of the IGB display:



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:
   - To change a single graph, click the colored bar at the left side of a graph to select it.
   - To change the settings for multiple graphs all at once, shift-click to select each graph.
   - To change the settings for all graphs that are currently displayed, click the **Select All Graphs** 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. For details, see the sections that follow this procedure.

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

Attach graph to a track or let it float

A graph that is attached appears in its own track, like the other annotation types.

A floating graph can be easily dragged to new locations across other tracks.

To put a graph into its own track:

- Click the colored bar at the left side of the graph to select it, then click the **Attached** radio button for the **Placement** option.

When a graph is attached as a track, the title may be hard to read. To view the full title of the graph in the tracks column, resize the tracks column. See page 32. You can also rename a graph, just as you can rename any other track, by right-clicking on the label and choosing **Customize**.

To detach the graph from the track:

- Click the colored bar at the left side of the graph to select it, then click the **Floating** radio button.

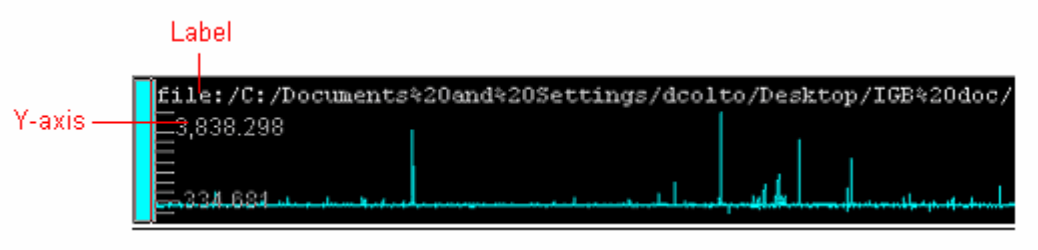
You can set preferences so that all graphs load either floating or attached as a track. See Setting graph preferences on page 86.

Change graph labeling

You can modify the labels for graphs that are currently selected by checking checkbox options in the Graph Adjuster panel.

- Show or hide the **Y-axis** scale
- Show or hide the identifying **label** of the graph source (typically the name of the graph file)

An example graph showing both options turned on is shown below:



Change graph color

To change graph color:

1. Click the colored bar on the left end of the graph to change.
2. Click the **Graph Adjuster** tab. In the **Options** section, click the **Change color** button.
3. Choose the desired color, and then click **OK**.
4. If you choose not to use the selected color, click **Reset**.

Change the height of a graph

There are several ways to stretch out the graph in the vertical direction:

5. If the graph is attached as a track, detach it so it floats. See page 43.
6. Hold down the shift key and drag your mouse up or down on the colored bar at the left end of the graph.

To change preferences so that all graphs initially load with the height that you specify, see Setting graph preferences on page 86.

After you change graph preferences, you can use the procedure described in that section to apply the new height settings to graphs that are currently loaded in IGB.

Change visible graph bounds

Changing the visible bounds involves changing the scale of the graph while at the same time setting the maximum and minimum values to be displayed.

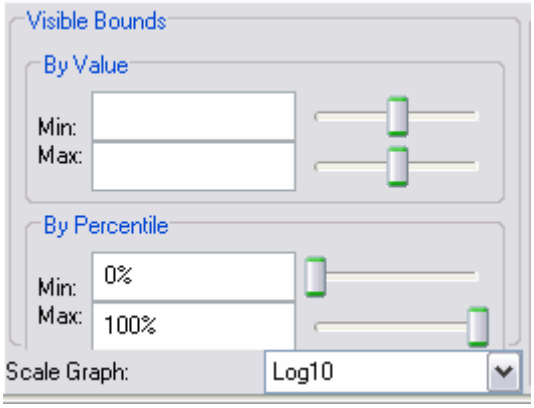
Scaling is an important issue to consider when viewing expression values because the distribution of these values across the range defined by their minimum and maximum may be highly asymmetrical.

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 in a graph because the vertical axis would be forced to show a range of 1 to 100,000. Thus, the majority of the values in the lower range of 100 to 1,000 would appear be dwarfed or "squashed" by the minority values in the higher range. 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. Thus, if you are using graphs to display numerical values, you should find out what the overall distribution pattern of your values looks like. If the distribution is asymmetrical (for example, not following a bell-shaped, normal curve) then you need to come up with good **Visible Bounds** settings.

Min/Max Visible Bounds scaling offers one solution to the asymmetrical distribution problem. In Min/Max scaling, minimum and maximum thresholds are chosen by some criteria, and these thresholds define the visual range of values displayed in the graph. Then, any expression value that is above the maximum threshold is occupies the maximum possible vertical space within the graph. Likewise, any expression value that is below the minimum threshold is shown as having a barely non-zero height.

Thus, any extreme values that are above or below the chosen thresholds appear the same visually. And the values in between are displayed across a range that allows their differences to be noted.

To set these Min/Max Visible Bounds, use the **Visible Bounds** section of the Graph Adjuster tabbed panel:



1. Select the graph(s) to change: Click the colored bar at the left end of each graph to change. (Shift-click to select more than one.).
2. If thresholding is on, toggle it off before you adjust the graph bounds, then on again after you are finished. See page 50.
3. Click the **Graph Adjuster** tab.
4. Adjust the visible graph bounds:
    - To make smaller changes to the graph view, modify the values in the **By Value** section. You can use the sliders or enter values in the **Min** and **Max** fields.
    - To make larger changes to the graph view, modify the values in the **By Percentile** section. You can use the sliders or enter values in the **Min** and **Max** fields.

To undo an adjustment and return to the original graph boundaries:

- In the **By Percentile** section, reset the **Min** to 0 and **Max** to 100

Setting the graph scale

By default, all graphs display in IGB in linear scale.

Note: Change the scale of a graph only if you know that its source was based on a mathematical scale other than linear. Graph files do not carry this information.

IGB duplicates your graph in a new track and applies your change to the new graph. If you make multiple modifications to a single graph, transformations are cumulative.

To change the graph scale:

1. Select the graph(s) to change: Click the colored bar at the left end of each graph to change. (Shift-click to select more than one.).
2. Click the **Graph Adjuster** tab.

3. In the **Visible Bounds** section, choose an option from the **Scale Graph** menu.

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, it is very useful to use the line graph display method because this method makes it easier to compare trends and patterns. 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.

The following graph styles are available in IGB:

- **MinMaxAvg** – This is the default style 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, values at individual bases resolve into bars, one for each data position/numerical value pair. When zooming out, IGB starts to summarize values. When the scale of the display reaches the point where individual pixels are associated with multiple values, IGB picks the maximum and minimum values and draws a bar between them. In addition, IBG draws lines through the centerpoints (the average) of all the bars.

- **Line** – In this style, adjacent values are linked with a line. If the zoom level forces a single pixel to represent multiple values, the average of all them is shown.

- **Bar** – Individual values are shown as vertical bars that are one base wide. Height is determined by the value and also the scaling. Zooming is handled as with line graphs.

- **Dot** – Instead of bars, dots are shown. Zooming is handled as with bar and line graphs.

- **StairStep** – This style is particularly useful for viewing .egr and .sgr files, ESTs, or other density data. .Egr graphs render by default as stairstep graphs, and .sgr graphs export from CNAT as stairstep graphs. The **StairStep** style is 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.

- **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 brightness of the line at each pixel

or coordinate. Thus, a brighter line indicates greater hybridization. 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.

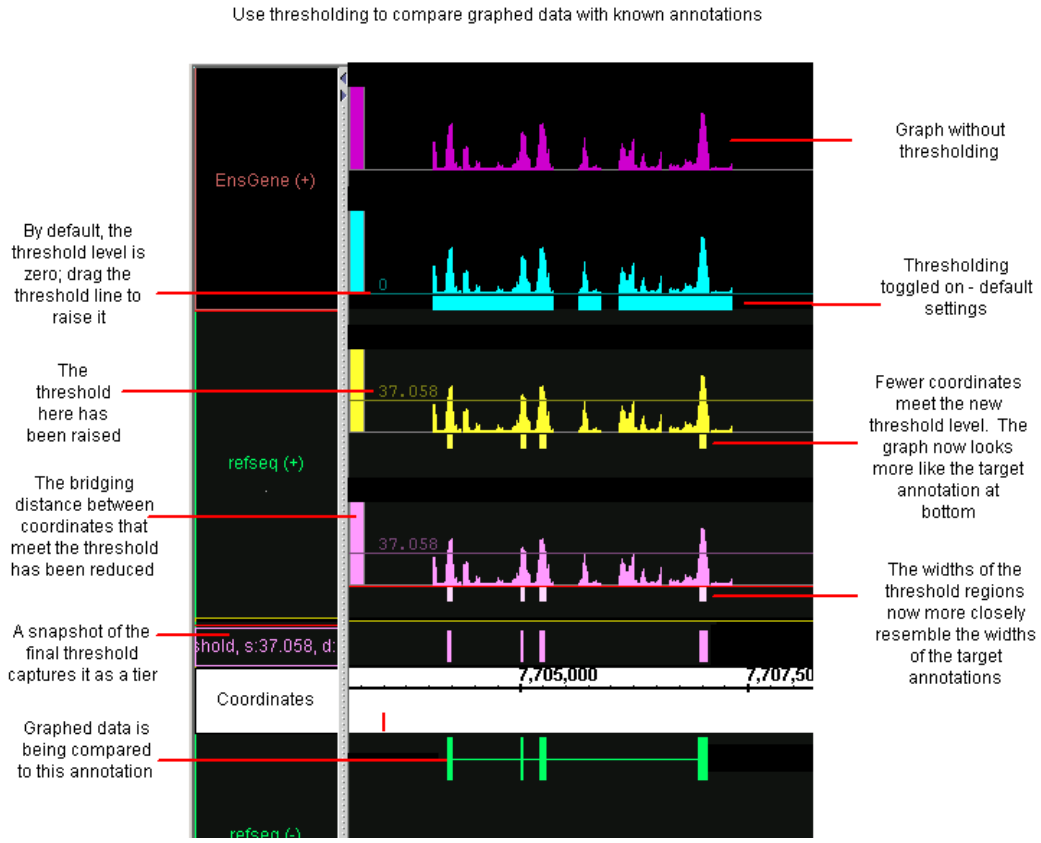
To change the graph style:

1. Click the colored bar at the left end of the graph to threshold.
2. Click the **Graph Adjuster** tab.
3. In the **Options** section, click the **Style** menu and choose one of the options described above.

View graph thresholds

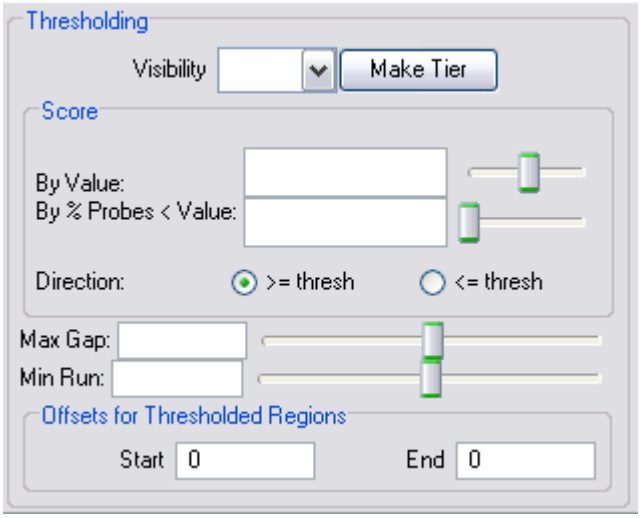
A common analysis step is to determine meaningful regions on a sequence based on graph values being above or below a certain threshold. To screen out non-meaningful values and to view meaningful features of a graph as annotations, use the thresholding feature in IGB.

Thresholding displays graph features as an annotation-like bar when the threshold you define is met (i.e., when the level of activity at the coordinate meets your defined threshold.) The thresholding feature lets you analyze your graph data using all of the features in IGB for working with and comparing annotations.



Adjust your graph's threshold 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.

To work with thresholds, make settings in the Thresholding section in the center of the Graph Adjuster panel:



Turn thresholding on

To turn thresholding on:

1. Click the colored bar at the left end of the graph to threshold.
2. Click the **Graph Adjuster** tab.
3. In the **Thresholding** section, from the **Visibility** menu, choose **On**.
4. (Optional) Make a separate track to hold the threshold bars.

By default, the threshold is set to the average of the min and the max value for a coordinate. To obtain meaningful thresholds, adjust the threshold settings using the methods described in the following sections.

Specify the threshold level

To change the threshold so that more or fewer points meet the threshold limit:

1. Click the colored bar at the left end of the graph to threshold.
2. Click the **Graph Adjuster** tab.
3. In the **Thresholding** section, specify **Score** settings using one of the following methods:
    - **By Value** – For expression results, this score represents the level of hybridization activity and is measured on the Y-axis. For a hybridization activity level to be considered significant, the score must be above a certain threshold. Enter a number greater than zero to raise the threshold to a meaningful level. If needed, toggle the Y-axis on See page 44.
    - **By % Probes < Value** – You can set the threshold based on the percentage of probes that are less than the threshold you set.
4. Slide the sliders to adjust thresholding, or enter values into the boxes for each of the following parameters:
    - **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.
    - **Min run** -- Min run is the 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.

Specify the threshold direction

Specify whether the threshold bars represent values that are above the threshold you specified or below it.

In the **Thresholding** section of the **Graph Adjuster** tab, for the **Direction** option, choose a value:

- To display a bar when graph values are greater than or equal to the threshold you specified, click the **>= thresh** radio button.
- To display a bar when graph values are less than or equal to the threshold you specified, click the **<= thresh** radio button.

Specify offset values

Offset values in IGB must match any offsets in your source data.

Update offsets for Tiling Array data

The Offsets for Thresholded Regions feature is designed specifically for working with tiling array data.

Captured data from tiling array probes is shifted by default; it starts at 12 base pairs past the intended beginning, and ends at 13 base pairs past the intended end. By default, the placement of all graph threshold bars currently reflects this fact.

Important: The Offset fields in the Graph Adjuster tab do not currently reflect the actual offset. If you want to change the offset values in the Graph Adjuster panel to reflect the actual thresholding behavior, use the following procedure.

To adjust the offset values for Tiling Array data:

1. Click the colored bar at the left end of the graph that you are thresholding.
2. Click the **Graph Adjuster** tab.
3. In the **Thresholding** section, look for the **Offsets for Thresholded Regions** options at the bottom of the window. If you do not see these options, make the IGB window or the tabs panel larger until they come into view at the bottom of the panel.
4. Change the **Start** value to 12.
5. Press the <Enter> key on your keyboard.
6. Change the **End** value to 13, then press <Enter>.

Remove offsets for other graphs

By default, the placement of threshold bars in the IGB viewer is offset from the actual coordinates to which they correspond; each threshold bar starts at

12 base pairs past the actual beginning, and ends at 13 base pairs past the actual end. When you view non-tiling array data in IGB, you must adjust the offset to 0 to correctly align the threshold bars with their actual coordinates.

Important: Currently the alignment of the threshold bars is offset by default, but the Offsets fields in the Graph Adjuster panel do not reflect this fact. Although both offset values appear to be zero, in fact they are not.

To adjust the offset values for non-tiling array data:

1. Click the colored bar at the left end of the graph that you are thresholding.
2. Click the **Graph Adjuster** tab.
3. In the **Thresholding** section, look for the **Offsets for Thresholded Regions** options at the bottom of the window. If you do not see these options, make the IGB window or the tabs panel larger until they come into view at the bottom of the panel.
4. Change the **Start** value to 0, then press the <Enter> key.
5. Change the **End** value to 0, then press the <Enter> key.

Capture the threshold bars

Capture the threshold bars (which look similar to annotations) in their own track to more easily compare different threshold settings for a single graphed region. (See the thresholding illustration, above)

To capture the threshold bars for use in this IGB session:

1. Click to select the graph
2. Click the **Make Track** button.

The snapshot does not capture the graph ID; it does, however, give the score (the level at which the threshold is set, or the "s" value in the track name) and distance (the length of the "bridge" joining the gaps where the threshold is not met, or the "d" value in the track name).

To permanently capture the threshold snapshot for later use or sharing:

1. Right-click the track and choose **Save Track as BED file**.
2. Name your file. Be sure to include the .bed filename extension.

What to do next

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.

Compare your "annotation" track with other annotation types using the tools in IGB for working with and comparing annotation tracks.

Useful graph customizations

Some common useful graph customizations are:

- Dampen the distorting effects of outliers – for example, ignore any points that are above the 99.5th percentile or below the .05th percentile. This makes it easier to see the bulk of your data.

  Under **Visible Bounds,** set **By Percentile** to the desired lower and upper percentile bounds.

- Change scales to match when comparing multiple graphs of different scales. For example, if one graph has a minimum value of 12 and a max value of 24, and another has a min value of 24 and a max value of 48, make them easier to compare by changing the visible bounds so they have the same scale. For each graph being compared, toggle the Y-axis on to view the min and max values for each, then adjust the visible bounds by Value to make them compatible
- Display the optimal type of graph for a particular purpose, for example use a stairstep style graph for analyzing ESTs or performing other density analyses. Then slice the selection to see details.
- Use the thresholding feature to find areas of activity that may be significant, for example areas that do not coincide with known exons. Adjust your threshold bars to match the known annotations; any "extra" threshold bars visible in your graph indicate potential areas of interest.

Getting details about a graph

Textual information about some graphs is available in Selection Info.

Click the colored bar at the left of the graph, then click the **Selection Info** tab. Resize the columns if needed.

Comparing two graphs

You may need to compare the relative differences in expression levels between two graphs that represent data that occupy the identical genomic position on the chromosome. For example, a graph of cancerous cells will show what's turned on or off when compared to a graph of normal cells. The following features make the differences between two sets of data easier to see.

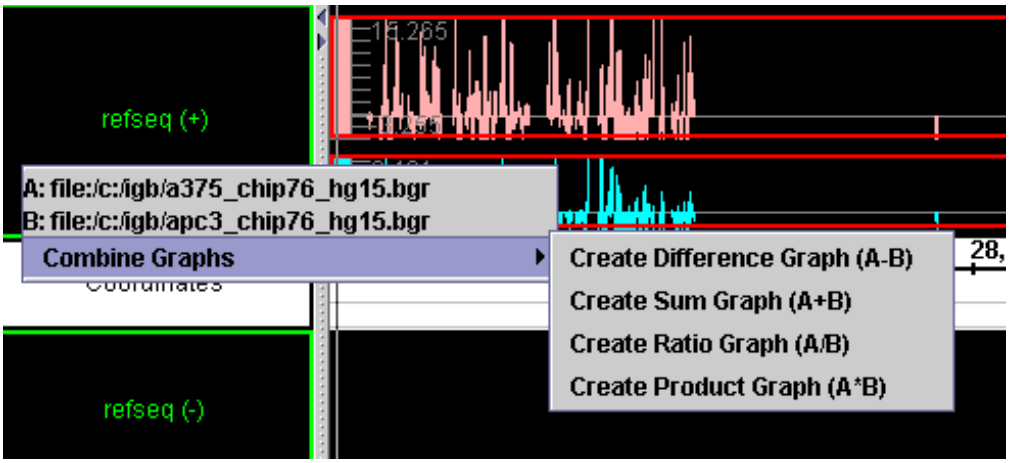
To create a difference or ratio graph:

1. Shift-click the bar at the left end of each graph to select the graph. The graph you click first is A, the graph you click second is B.
2. Right-click the bar at the left end of either graph and choose **Combine Graphs**, then choose one of the following:
   - **Create Ratio Graph (A/B)**
   - **Create Difference Graph (A-B)**

Results that show significantly more or less activity than a control graph for the same region may indicate an area of interest.

To create a sum or product graph:

1. Shift-click the bar at the left end of each graph to select the graph. The graph you click first is A, the graph you click second is B.
2. Right-click the bar at the left end of either graph and choose **Combine Graphs**, then choose one of the following:
    - **Create Sum Graph (A+B)**
    - **Create Product Graph (A*B)**



Saving and bookmarking graph views

Certain aspects of graph manipulation can be captured for future use by saving or bookmarking. You can also export bookmarked views.

Saving a graph as a file

You can save a graph as a separate file for future reference or offline use. Note that this saves the generic graph, without any visualization customizations you may have made to it.

To save a graph to file:

1. Select the graph by clicking the colored bar at the left end of the graph.
2. Click the **Graph Adjuster** tab. In the **Options** section, click the **Save Graph** button.
3. Specify location and graph name, then add the **.gr** filename extension.
4. Click **OK**.

Bookmarking graph views

See Bookmarking on page 62.

You can also export bookmarked views.

Setting graph preferences

To set preferences that apply by default when you load graphs:

- See Setting graph preferences on page 86.

Duplicating graphs

You may want to duplicate a graph to experiment with different graph adjustments without losing adjustments that you have already made.

To duplicate a graph:

1. Click the colored bar at the left end of a graph to select it.
2. Click the **Graph Adjuster** tab. In the **Options** section, click the **Clone graphs** button.

Deleting graphs

Deleting graphs from the viewer window

Delete graphs to remove them from the viewer. Hiding graphs is no longer available.

Deleting a graph

To delete a single graph from the viewer:

1. Select the graph to delete by clicking the colored bar on its left end.
2. Click the **Graph Adjuster** tab. In the **Options** section, click the **Delete graph** button.

Deleting all graphs

To delete all graphs currently displayed:

- Choose **File** menu > **Clear graphs**

Visualizing probe sets

Introduction

Chip design procedures have varied somewhat from chip to chip, but, in general, probes on Affymetrix' standard commercial expression microarrays (such as the Human U133 chip) are designed to match the 3' regions of known or computed mRNA sequences. These probes are typically 18 to 25 bases long depending on the chip.

In standard arrays, probes are grouped conceptually into probe sets, groups of probes that are expected to measure expression for individual known or computationally-deduced mRNA molecules.

Probe sequences are typically selected from the 3' regions of these design sequences in order to maximize the amount of sample mRNA to be measured. (This so-called "3-prime bias" will most likely change as sample preparation protocols improve over time.)

These design sequences may be identical to known mRNA sequences in GenBank, or they may have been produced computationally by merging ESTs or mRNA sequences into a single sequence, sometimes called a "consensus" sequence.

IGB can be used to visualize the location of design sequences and probes within the genomic sequence. Being able to see where these sequences are located within the genome can be extremely useful when several overlapping probe sets recognize diverse mRNAs originating from the same gene. Around 60% of human genes produce multiple variants, and these diverse variants often exhibit very complex configurations of exons and introns. As a result, it often helps to be able to view a diagram of these complex structures together with probes in order to determine which individual mRNAs are being detected by a given probe set.

To make this easier, IGB shows where the design sequences and their probes align to the genome together with alignments between the genomic sequence and known or predicted mRNAs. By showing probes, known mRNAs, and design sequences in the same view, IGB makes it possible to determine quickly which known mRNA a probe set could detect.

However, it is important to note that these genome alignments can only be as good as the genomic sequence on which they are based. Genomes vary in their level of quality – for example, the human and fruit fly genomes are very high quality, having gone through several releases involving many refinements. More recently-sequenced genomes (such as mouse or rat) are not yet as reliable. For the less-refined genomes, there may be many examples of design or mRNA sequences that simply do not align anywhere to the genomic sequence or, if they do align, do so imperfectly. Readers interested in understanding more about these issues can read more about this

topic at the following link:
http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=12149135/.

Invoking IGB via links from NetAffx

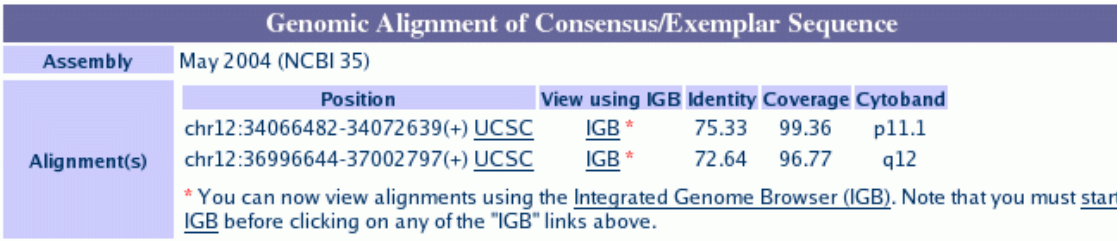
Probe set pages on the NetAffx customer support Web site contain links to the IGB browser under the “Genomic Alignments” section of the page. Clicking these links (labeled “IGB” as shown below) will tell IGB to show the region of the genome containing the probe set and its design sequence.

IGB must currently be running for these links to work. If IGB is not currently running, start it in your usual way or by clicking the “start IGB” link provided for that purpose. If IGB is already running, then the current scene will be replaced with a view of the probe set design sequence alignment region.

Depending on the probe set, there may be more than one reasonably good alignment for the design sequence. Most genomes contain large amounts of duplicated sequence, and so it is to be expected that many design sequences will align to multiple locations.

For example, the figure below shows links to two different alignments for a single probe set and its corresponding design sequence.

IGB links in probe set page.



Genomic Alignment of Consensus/Exemplar Sequence					
Assembly	May 2004 (NCBI 35)				
Alignment(s)	Position	View using IGB	Identity	Coverage	Cytoband
	chr12:34066482-34072639(+) UCSC	IGB *	75.33	99.36	p11.1
	chr12:36996644-37002797(+) UCSC	IGB *	72.64	96.77	q12
	* You can now view alignments using the Integrated Genome Browser (IGB). Note that you must start IGB before clicking on any of the "IGB" links above.				

Visualizing probe sets: loading from data files

Sometimes it is useful to load up an entire chip’s design sequence alignments into IGB. To do this, download the files from the Affymetrix customer support Web site.

These can be obtained on-line at the NetAffx customer support Web site. To obtain these files for individual chips, go to http://www.affymetrix.com/support and choose the array of interest. On the array’s support page, look for alignment data files in the "NetAffx Annotation Files" section. Note that alignment data files will only be available for species with well-characterized genomes.

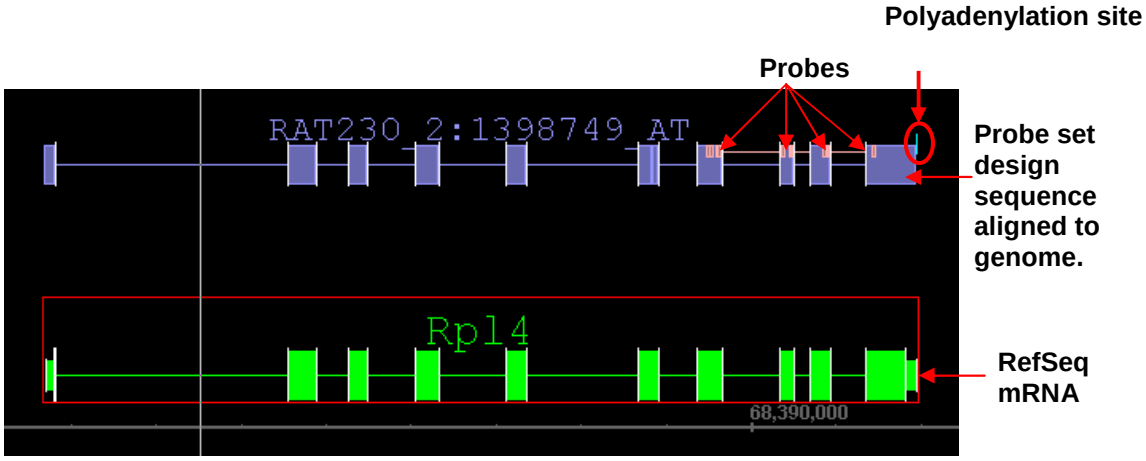
These alignment files are in a special format called “PSL” (pslayout) which was first designed by Jim Kent of U.C. Santa Cruz. For more details on this format, visit the U.C.S.C. Genome Bioinformatics Web site at http://genome.ucsc.edu.

To visualize all the probe sets and associated design sequences for a particular chip, load the genome of interest and merge a probe set data file, see Viewing multiple files in a single instance of IGB.

Probes and probe sets on display

Once the data have been loaded, the probes and design sequence alignments will appear in a separate track labeled by chip. Within the track, probes and their matching design sequences always appear together. The following figure shows an example from the rat 230 chip.

This image shows a probe set that detects mRNA transcripts from the rat Rp14 gene, which encodes a ribosomal protein.

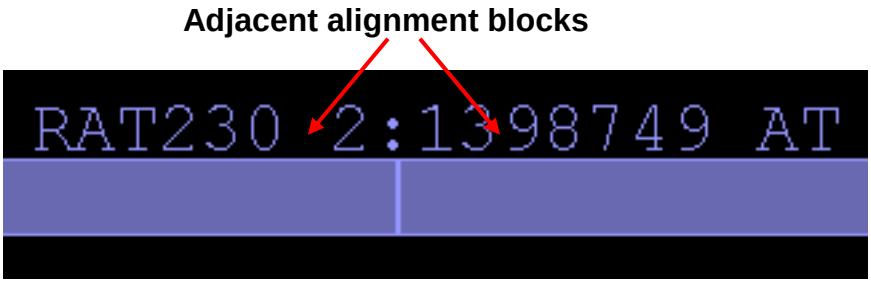


Probe set design sequence

The alignment between the probe set design sequence and the genome is represented at the top of the figure as a series of blocks. Each block represents a block of alignment in which each base in the genome matches a corresponding base in the design sequence. Gaps between the blocks typically represent areas where the genomic sequence contains inserts relative to the aligned design sequence. Usually, these gaps are due to introns.

There are some exceptions to this, however. For example, the sixth and seventh blocks in the figure above are so close together that they almost appear as a single block at this level of zoom.

Zooming in for a closer view reveals that these two alignment blocks are immediately adjacent to each other. This indicates that these two blocks of alignment were separated by an insert in the design sequence relative to the genomic sequence. That is, the design sequence contained some bases that were not present in the genomic sequence. This may present a problem if this missing region (in the genome, that is) contains some probe sequences. In this particular case, however, the alignment irregularity occurs in a 5' region, outside the area covered by the probes. (See the discussion below.)



How can this happen? There are a number of reasons, but gaps in the genomic sequence – regions that simply have not yet been sequenced – are the most common culprit.

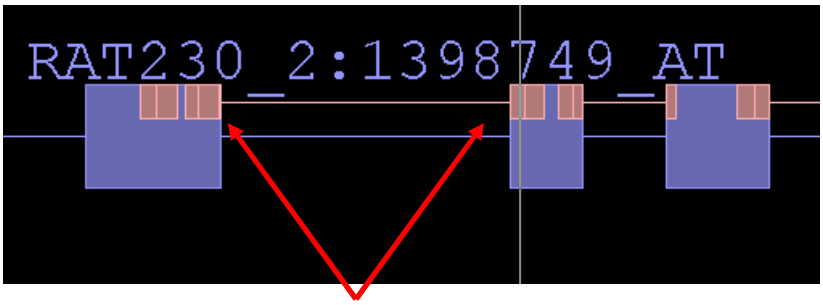
Probes and probe sets.

Each design sequence is shown with its corresponding probe set. Each probe set consists of a group of probes, which are shown superimposed on the alignment blocks of the design sequence.

The figure below shows a close-up view near the 3' end of the design sequence. The 3' end of the design sequence is annotated with blocks which represent individual probes.

Two things are important to notice about this image. First, sometimes individual probes are split across gaps in the alignment, which typically correspond to introns. When this occurs, the two halves of the probe are connected by a line.

Second, sometimes probes overlap with each other. This can be seen by clicking on probes. If a selection outline is visible in the middle of what otherwise looks like a single block, then there is really more than one block.



Probe split across genomic gap (intron)

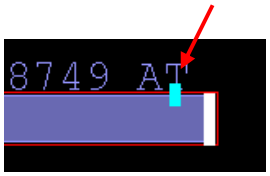
Probe set labels

Each probe set and probe set design sequence is labeled with the name of the chip (Rat 230, in this case) and the probe set identifier.

Polyadenylation sites

For many probe set design sequences, Affymetrix bioinformatics scientists have used computational methods to predict putative polyadenylation sites near the 3' end of the design sequence. These are shown as dark or light blue boxes riding on top of the terminal alignment span.

Predicted
polyA addition site



If the box is light blue, then the site was deduced from the overlap of multiple expressed sequences (usually 3' ESTs) whose genome alignments all terminate at a common location. This is sometimes called a "polyA stack."

If the box is dark blue, then the site has been deduced through sequence analysis of an individual, exemplar sequence, such an mRNA sequence record from GenBank. This is sometimes called a polyA site to distinguish it from a polyA stack.

mRNAs in the region – putative probe set targets

It is usually a good idea to load additional tracks of data besides just the probe set information. For example, in the figure above, the track below the probe set track shows the probe set's likely target, an mRNA from the Rp14 locus. Based on their relative alignments to the genome sequence, it appears that both sequences overlap in the region that contains the probes. Thus, it is very reasonable to assume that this probe set does indeed detect the mRNA shown below it.

Getting more information about a probe set

Right-clicking a probe set or its design sequence and selecting the **Get More Info** option will open a Web browser window showing detailed information about that probe set, including its design sequence.

About alignments

Probe mappings

A program called "blat" was used to map probe set target sequences onto various versions of various genomes. More information can be obtained about this program from the U.C. Santa Cruz Genome Web site, which also uses the blat program to map mRNA sequence onto genomic sequence. The link to this Web site is: http://genome.ucsc.edu/ .

The location of the probes within the probe set design sequence is already known from the design process. This information, in combination with information about the location of the probe set design sequences within the larger genomic sequence, was used to map probes onto the genome.

Many chip probe names use special suffixes to indicate when a probe set may detect multiple distinct target mRNAs. For details, consult the data sheet for your array of interest or contact Affymetrix technical support. (See: http://www.affymetrix.com/site/contact/index.affx for contact information.)

Viewing tips

Use the spliced view display to examine probe sets, probe set design sequences, and mRNA transcript annotations. Introns in mammalian genomes tend to be very large. The "sliced" view takes away the uninformative intronic sequence, allowing one to take in the entire structure of gene without having to scroll or zoom.

Capturing, saving, and sharing IGB data

There are several ways to save data for use in other applications, sharing, or later use in IGB.

Save or share data you from IGB in any of several ways:

- Saving graph files
- Making and exporting bookmarks
- Saving graph threshold snapshots
- Copying and pasting to another location (sequences or data in IGB tables)
- Printing

Saving graph files

See Saving a graph as a file on page 54.

Bookmarking

To return later to a region or annotation and zoom level of interest, you can bookmark it. Items that you bookmark will appear under the Bookmarks menu.

Any bookmarks that you create during a session will be saved when you exit, and will be available the next time you launch IGB.

Creating bookmarks

Before you add bookmarks, make sure you are not running multiple instances of IGB.

There are two types of bookmarks in IGB:

1. Position bookmarks.

   Position bookmarks record the currently shown genome, the location, and the current zoom level.

   - To create, choose **Bookmarks** menu > **Add position Bookmark**

2. Position-and-graphs bookmarks.

   Position-and-graph bookmarks record the same things as position bookmarks, but also keep track of any graphs that you have loaded thus far. In addition to remembering the data files that you have loaded, position-and-graphs bookmarks can also remember many of

your customizations for graph appearance, such as color, graph bound adjustments, thresholding, and labeling.

- To create, choose **Bookmarks** menu > **Add position & graphs bookmark.**

Using bookmarks

To use a bookmark to load a view that you have bookmarked, do one of the following:

- Choose **Bookmarks** menu > [name of your bookmark]
- Click the **Bookmarks** tab, then click the bookmark to use, then click the **Go To** button. You may also need to adjust the vertical zoom and scroll.

The view that you bookmarked will load.

Viewing and changing bookmark details

You can view details about your bookmarks, such as their position, which genome they refer to, or customizations for position-and-graphs bookmarks. You can also modify some details, such as the name of a bookmark or customizations of a position-and-graphs bookmark.

To view or change bookmark details:

1. Click the **Bookmarks** tab.
2. Click a bookmark, then click the **Properties** button.
3. The information will appear in tabular format in a new window.
4. (Optional) Modify the bookmark details.
5. Click **Apply Changes** or **Cancel**.

Exporting bookmarks

To save a copy of the bookmarks to a file:

- Choose **Bookmarks** menu > **Export Bookmarks**

Bookmarks are saved in a single file in the Netscape bookmarks file format. This format resembles XML and HTML, but is not identical to either of those, so use caution if you choose to edit the file.

Importing bookmarks

You can import and use bookmarks, for example bookmarks that other people give you:

- Choose **Bookmarks** menu > **Import Bookmarks** > choose the file containing the bookmark of interest.

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. All imported bookmarks will be persistent between sessions of IGB.

Capturing a sequence of base codes to clipboard

You can copy a sequence of bases and paste the sequence into another application for use there.

To capture a sequence of bases:

1. Make sure sequence residues are loaded. See page 18.

   Click and drag in the Coordinates track to select the desired coordinate range, or click to select a single annotation.

2. Choose **View** menu > **Copy selected residues to clipboard**.
3. Paste the text string into the target application or location.

If you have selected an annotation that has introns, the selected residues will include only the exonic regions. If you want to capture all the residues including the intronic regions, select the region by dragging in the Coordinates track.

Capturing data in table cells

You can copy and paste table cells from the Selection Info panel (or any other tabular format) and paste them into Excel or some other program.

To capture data in table cells:

1. Click and drag to select the cells you want.
2. Copy (using Control-C) and then paste into the target application.

Printing

To print the entire IGB window including the tab section:

- Choose **File** menu > **Print Whole Frame**

To print the contents of the viewer window only (not including the tab section):

- Choose **File** menu > **Print**

Advanced Features

The following advanced features are more complicated or less commonly used than other features in IGB. The advanced features in this section include:

Using advanced QuickLoad features

The features described in this section are intended for advanced users only.

Changing the QuickLoad source

By default, QuickLoad data loads from http://netaffxdas.affymetrix.com/quickload_data/

To add a secondary personal QuickLoad data source:

- Click the **QuickLoad Options** button and specify a QuickLoad source URL.
- You may use the “file://”, “http://”, or “https://” protocols.

Clearing 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 filesystem will not be cached.

You can set preferences for caching, or clear the cache.

Click the **QuickLoad Options** button and specify caching options, or clear the cache.

Setting up your own QuickLoad server

If loading sequence residues over the internet is too slow, consider copying the annotations and sequence residues from the QuickLoad server for your preferred genomes onto your local network. For information about setting up your own QuickLoad server, see the document titled *How to create a QuickLoad Directory for use in IGB* on http://sourceforge.net/docman/?group_id=129420.

Setting up secure data sources

QuickLoad and DAS data can be made available via secure https connections. Set your IGB preferences for that data source to point to the URL of the secure server. For information, see the section above and Appendix 3 – Preferences.

Pivot View – visualizing expression across many experiments

Introduction

Use the Pivot View tabbed panel to examine and compare the results from multiple gene expression experiments.

Using Pivot View is similar to using Graphs in that both show expression values in the context of genomic sequence. The difference is that in the Pivot View, the information on the x- and y-axes in the main view is flipped and individual sequence segments are separated into individual rows.

With Graphs, expression values from a single experiment appear as a graphs laid out across the horizontal sequence axis. In Pivot View, columns represent individual experiments, while rows represent individual regions.

Expressed intervals and .egr files

Pivot View displays data from (file name suffix: .egr) files, tab-delimited plain text files containing summarized expression data for expressed intervals.

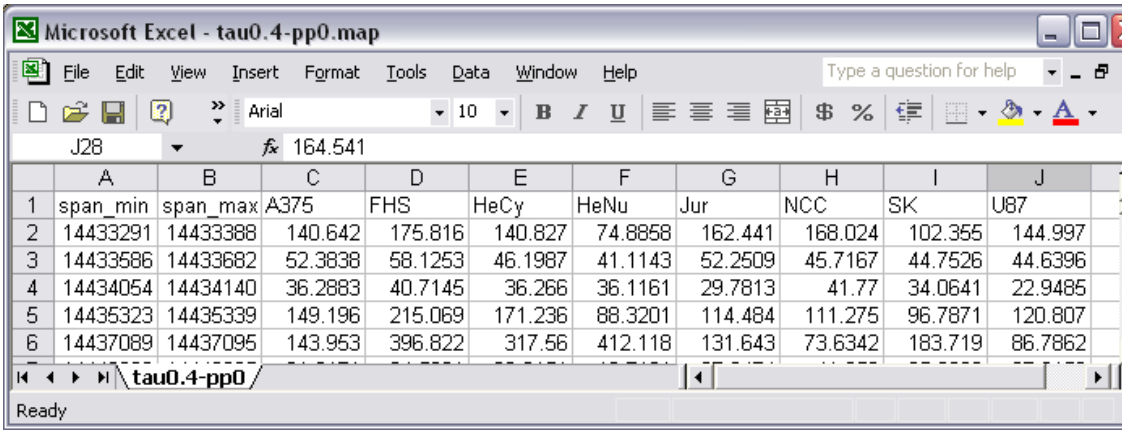
Expressed intervals are segments of genomic sequence that a data analysis has determined are expressed together. For example, a data analysis algorithm might determine that an exon whose boundaries never vary is an expressed interval. Conversely, if an exon coincides with two expressed intervals, then it is possible that the exon may be alternatively spliced, depending on the tissue type or developmental state. Another way to look at expressed intervals is to regard them merely as contiguous runs of bases whose expression values can be intelligently summarized using a single value.

The image below shows an example IGB expression map file that was opened using Excel. Note that the first row of data contains the column headings and that the values are separated by tabs.

The first two columns give start and end positions for the expressed intervals described in the file. The columns after that list expression values for expressed intervals in eight different samples, including cytoplasmic and nuclear fractions from HeLa cells and samples from other cell lines, as well.

Note that as of this writing, map files do not contain information regarding which chromosome or genome release the expressed intervals come from. Future versions may contain this information, however. As with graphs, it is

important to match up the map files with the appropriate segment of the genome being viewed.



	A	B	C	D	E	F	G	H	I	J
1	span_min	span_max	A375	FHS	HeCy	HeNu	Jur	NCC	SK	U87
2	14433291	14433388	140.642	175.816	140.827	74.8858	162.441	168.024	102.355	144.997
3	14433586	14433682	52.3838	58.1253	46.1987	41.1143	52.2509	45.7167	44.7526	44.6396
4	14434054	14434140	36.2883	40.7145	36.266	36.1161	29.7813	41.77	34.0641	22.9485
5	14435323	14435339	149.196	215.069	171.236	88.3201	114.484	111.275	96.7871	120.807
6	14437089	14437095	143.953	396.822	317.56	412.118	131.643	73.6342	183.719	86.7862

Loading .egr files

Loading .egr files is similar to how other annotation files are loaded:

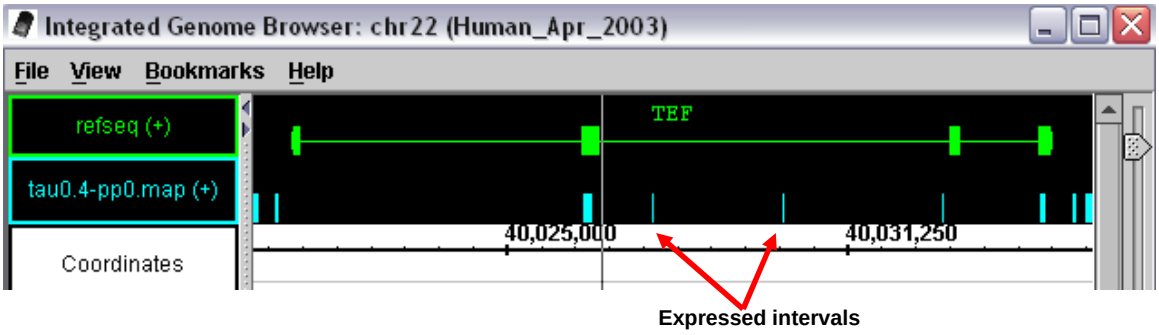
1. Make sure the genome version and chromosome of files you want to open match those of the data already displayed.
2. Choose **File** menu > **Open**
3. Check the **Merge with currently loaded data** checkbox.
4. Select an .egr file.

The data from this merged file will automatically be added to an additional track alongside the tracks already in the viewer.

Map data in the main window

When an .egr file is first loaded, the data appear in a new track named after the data file. The image below shows a new track that appeared after loading a data file called **tau0.4-pp0.egr**.

Expressed intervals in main display window.



In the picture above, the blue blocks above the genomic sequence axis represent expressed intervals occurring at the indicated positions. These purely positional data correspond to the interval information that was stored in the first two columns of the tabular expression map file shown above.

Note that in this case, some of these expressed intervals overlap with exons in the RefSeq mRNA transcript annotation shown in the track directly above the expression interval track. However, there are also some smaller expressed intervals which do not overlap with any RefSeq annotations. In fact, they appear to overlap with intronic regions which should not be detected on a GeneChip because they are removed from the transcript during splicing. However, it is possible that these intervals correspond with genes and RNAs that have not been entered into the RefSeq database. If so, these unreported genes or transcripts could be transcribed from either the plus (top) or minus (bottom) strands. In this case, at least, the experimental protocol used to generate the expression map data cannot distinguish strand information.

Thus the expression map data format does not include strand information. By default, expression intervals are always shown as features on the plus (top) strand of the genomic sequence.

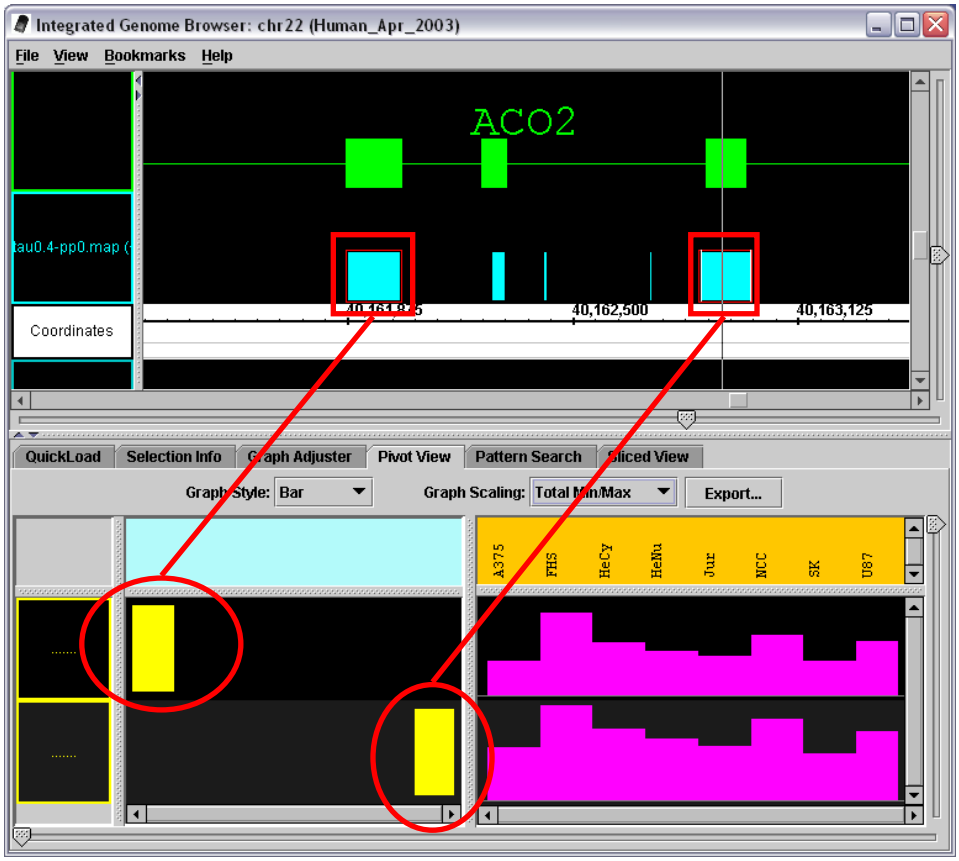
Map data in the Pivot View

To view expression data associated with the expressed intervals, select them in the main window: Anything that is selected (outlined in red) in the main window is automatically shown in the Pivot View.

To select a region of expressed intervals for display in Pivot View, click-drag over the expressed interval blocks. To add more expressed intervals to the selection, click or click-drag over additional blocks while pressing the **SHIFT** key.

The image below shows a sample Pivot View display that was built by selecting two expressed intervals from the Aco2 region. (Aco2 encodes aconitase, a mitochondrial protein that catalyzes the second step of the TCA cycle.)

Pivot View showing expression values for two expressed intervals



In this figure, the two expressed intervals that were selected are outlined in red in the main IGB display. They also appear as yellow blocks in the left-hand portion of the Pivot View tabbed panel. Note that individual expressed interval blocks occupy separate rows, and their positions within the row indicate where they lie with respect to the genomic sequence axis. Thus, the leftmost block represents the 5'-most interval, while the rightmost block represents the 3'-most interval.

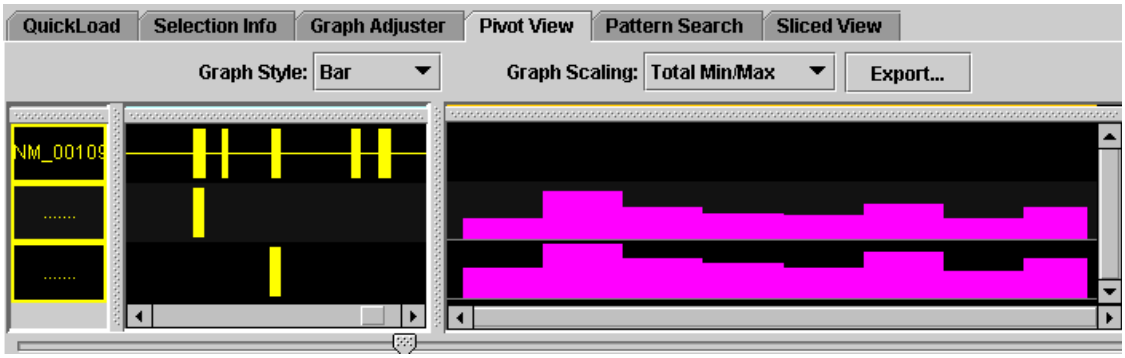
Expression patterns for each interval appear in the expression values panel occupying the right side of the Pivot View. This panel shows the expression values for each expression interval (rows) measured across several experiments (columns).

In this example, the expression panel shows expression values for as a bar graph that represents expression for an individual expression interval across a panel of experiments. Within each graph, the height of individual bars represents expression intensity. Taller bars indicate higher levels, while shorter bars indicate lower levels of expression.

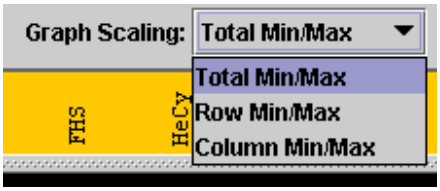
What should be clear from this very simple example is that the overall pattern of expression for both blocks is similar. That is, if you compare the height of bars in adjacent cells for both blocks, you notice that when the expression for one block goes up from experiment to experiment, the

expression for the other block goes up as well. Similarly, if expression goes down from experiment to experiment for one block, the expression of the other blocks also decreases. This makes perfect sense when you consider that the two blocks probably represent nearby exons in the same mRNA transcript. Thus, it is very likely that these blocks measure aconitase expression.

To compare the blocks shown in Pivot View with RefSeq annotations in the main view, just add the annotation of interest to the currently selected items by clicking the item while holding the SHIFT key. The following figure shows how this looks for our aconitase example.



Expression panel graph scaling



The **Graph Scaling** menu at the top of the Pivot View tab contains settings that control the vertical scaling and range of values displayed in the expression panel.

For example, if a row occupying 1 cm of vertical space has an expression value range of 0 to 1000, then a bar that is 0.10 cm high indicates an expression value of 1000 expression units / cm x 0.10 cm = 10 expression units.

The **Graph Scaling** menu contains two options:

- Total Min/Max
- Row Min/Max

The default setting is **Total Min/Max** for newly displayed graphs. You may be able to change the default setting in a future version of IGB.

In **Total Min/Max** scaling, each row shows the same range of values. The top end of this range is set as the largest expression value across all the rows,

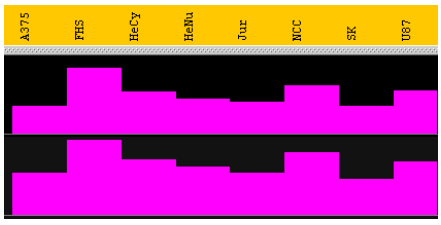
while the bottom end of this range is set by the minimum value across all the rows.

In **Row Min/Max** scaling, scaling is done on a per-row basis. That is, the maximum and minimum values within a row determine the relationship between vertical space and expression values. Use **Row Min/Max** scaling when one row has unusually high values.

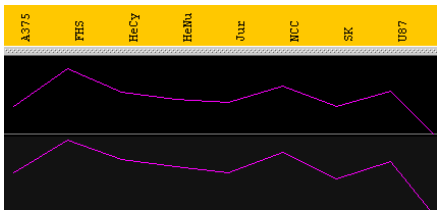
Changing the graph style

Using the **Graph Style** menu, it is possible to switch between three different graph presentation styles in the Pivot View expression panel. The three options currently available include:

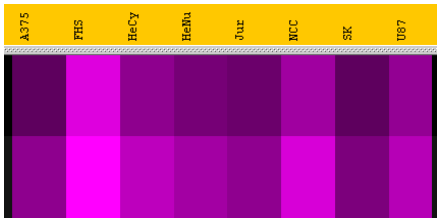
- Bar graph



- Line Graph



- Heat map



The type of graph that is most appropriate depends on the type of question being asked about the data. For example, when comparing trends, it is very useful to use the line graph display method because this method makes it easier to compare trends and patterns. 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.

Display Tips

When viewing data from a large number is of expression experiments and map files, it helps to put the Pivot View panel into its own separate window on the desktop. This is easily done by choosing **View** menu > **Open Tab in New Window**.

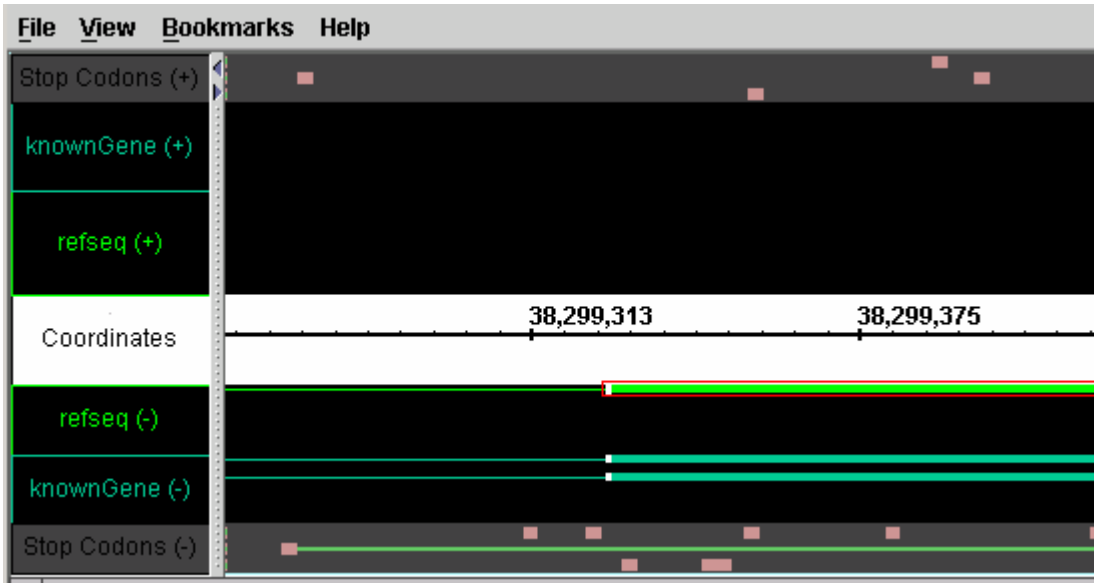
Showing ORFs and stop codons

Use ORFs to find coding regions; the longer the ORF is, the more likely it is to contain live coding regions.

To show ORFs and stop codons:

1. Make sure sequence residues are loaded.
2. Select an item of interest in the main view and click the **Sliced View** tab.
3. ORF settings are at the bottom of the panel.
4. Adjust the slider to set the minimum length of ORFs you want to see, if necessary. By default, only ORFs 300 bps or longer are shown.
5. Click the **Analyze ORFs** checkbox.

Three rows of ORFs in each direction appear, one for each start position. Stop codons are pink. ORFs of the specified length or longer appear as a green line:



If too many green lines (ORFs) appear, increase the minimum ORF length (see step 4 above). If too few appear, decrease the length. If you do not see the result of each change, click the Analyze ORFs checkbox after you change the minimum length.

Note: You can change the colors that indicate ORFs and stop codons. See the Setting ORF and Stop-codon colors section in Appendix 3 Preferences.

Clearing ORFs from the viewer

To clear ORF and stop codon markers from the viewer window:

- In the **Sliced View** tab, deselect the **Analyze ORFs** checkbox.

Identifying a translation frame

To identify a translation frame:

1. Load sequence residues for the region you're viewing. See page 18.
2. Slice the region. See page 36.
3. Adjust the slicing to a buffer of 0 bases to completely remove the introns
4. Analyze the ORFs as above.
5. Look for the long, continuous green line that corresponds to your strung-together exon set. This is your translation frame.

   Currently, if a stop codon is located beyond the region being viewed, the green line is not visible; instead, the evidence of a translation frame is the lack of stop codons over a long sequence.

Controlling IGB through a browser, using bookmarks

If you open a bookmarks file in a Web browser and then click one of the links *while IGB is running*, then IGB will respond by changing the display to show the selected bookmark.

This will only work for the first instance of IGB that you are running. For example, if you are running two copies of IGB, then the browser will only affect the IGB window that you started first.

Technical details on how this works.

If you open a bookmarks file in a text editor, you will find that the URL for each bookmark looks something like **http://localhost:7085/UnibrowControl?seqid=chr21&end=44542686&version=Human_May_2004&start=44534188**.

In this case, "localhost" simply refers to your local machine and "localhost:7085" refers to port 7085 on your local machine. IGB, by default, "listens" to port 7085 on whatever machine it happens to be running. So if a Web browser sends a request to "localhost:7085," IGB will hear it and respond. If the URL contains the right parameters (version, seqid, start, end,

etc.) then IGB responds by showing the genomic “scene” described in the bookmark.

However, if you start up a new instance of IGB while one is already running, the new instance will listen to a different port – the first port above 7085 that happens to be vacant. This is why the bookmark links will only work for the first IGB instance and not for any others you start running after that.

Appendix 1 Using Java Web Start

You can use a free utility from Sun Microsystems called Java Web Start to simplify installation and use of IGB and ensure that you always have the most current version of IGB. When you use this utility to launch IGB, it will automatically check for, download and install updates as they become available. Java Web Start is a component of the Sun J2SE Standard Edition.

In many cases, Java Web Start will already be installed on your machine and you will not need to do anything to install it.

Obtaining Web Start and IGB

Note: the following process is current at the time this document was written, but what you will actually see on Sun's web site is subject to change.

To use IGB with Java Web Start:

1. Go to the Affymetrix site for IGB: http://www.affymetrix.com/support/developer/tools/download_igb.affx. If you see buttons labeled "Launch IGB" or something similar, then you should be able to start IGB by pressing one of those buttons. If you do not see those buttons, continue with the rest of these instructions.
2. Click the Java Web Start link to go to the Sun Microsystems Java downloads website, or go directly to: http://java.sun.com/products/javawebstart/
3. On the Sun Microsystems Java downloads website, look for the link to download **Java 2 Platform, Standard Edition (J2SE JRE)**. Click the link to download J2SE 1.4.2 or higher. (J2SE 5.0 or higher is also acceptable.)
4. Click to download the application for the platform that you are using (Windows, Linux, etc.) You can use either the offline or the regular version.

   Many of the downloads at the Sun site have similar names; choose carefully, since you do **not** need the SDK, or the Enterprise Edition, or the Micro Edition, or Net Beans, or a Cobundled package. You may download those if you wish, but you only need the J2SE JRE.

   If you already have a version of the JRE or the JDK, this download should not interfere with that installation.
5. Read and accept the license.
6. Before you click the final link to perform the download, verify that you are downloading the correct software.

7. Click **Open** to run the installation from the Sun server, or click **Save** to save the installer to your hard disk and run the installation from there. When you run the installer, accept the defaults.
8. After Java Web Start installation is complete, return to the Affymetrix site in Step 1 of this procedure and click the **Launch IGB** button.
9. Click **Start**

   You will see a message asking if you would like to integrate IGB into your desktop environment.
10. If you wish to place a shortcut to IGB on your desktop and in your **Start** menu > **Programs**, click **Yes**. Otherwise, click **No**.

    IGB will launch.

Note: If you change web browsers, then the next time you launch IGB, the Java Web Start process may start from the beginning and download the Web Start application again.

Appendix 2 Troubleshooting

If you have problems, first double-check the instructions for the specific task you're working on, and then try solutions below that are specific to the area of IGB you're working on, and finally try the general troubleshooting items at the end of this appendix.

Troubleshooting specific areas

Below are troubleshooting tips for some areas of IGB.

IGB seems slow or sluggish

Try the following:

- Launch IGB using a larger RAM memory allocation. See Launching IGB.
- Turn off graph thresholding if you are not actively using it.
- Turn off slicing: Choose **View** menu > **Turn off slicing**. Or make sure the tab is not visible.

Keyboard command doesn't work

You may have changed the command in the Preferences. Or some component other than the viewer window has the "keyboard focus". For example, your key strokes may be going to a text entry box or to another application. Click anywhere in the viewer to return the keyboard focus to it.

Out of memory errors

If you are working with many graphs or sequences, you may run out of memory. After you receive an out-of-memory error, most functions will become unusable and you will have to close and re-launch IGB.

You may need to launch IGB using a larger RAM memory allocation. See Launching IGB.

Tab or tab window is missing

If you have opened a tab in a new window, the tab window may have become hidden behind the main IGB window or another window.

General troubleshooting

If anything you're working on doesn't work, try these common fixes, even if they don't seem intuitive:

- Read the section in this document related to the task you are trying to accomplish
- Load Sequence residues for the chromosome or file you're using
- Make sure the genome version and chromosome in the viewer match those of whatever you're trying to load or compare.
- Make sure you've selected the gene, not the exon, or vice-versa. See select parent.
- Unclamp the view
- Try limiting the region in the viewer to less than 2Mb
- Make sure you're connected to the network and the server or data source that you're connecting to is up and running…
- Make sure you haven't hidden any tracks or graphs

Appendix 3 Preferences

Preference settings that you make will remain in effect after you shut down and re-launch IGB. Some apply to data that is already loaded, while others apply only to future activity or require that you close and re-launch IGB.

You can set preferences for the following:

- Appearance of the viewer, tracks, annotations, and graphs
- Locations of data sources
- Keyboard commands for tasks that you perform frequently
- Appearance and function of specific features.

Changing preferences

To set track-related preferences, see the section on Customizing the default appearance of tracks on page 80.

To change preferences related to QuickLoad and caching of standard reference annotations, see the QuickLoad Options section.

To change all other preferences:

1. Choose **File** menu > **Preferences**.
2. In the **Preferences** dialog box, choose an option from the **Preferences** menu or click a tab to see more options.
3. Set options as described in the rest of this appendix.
4. The Help for Preferences may have more current information than this document. To view the Help, choose **File** menu > **Preferences**, then choose an option from the **Help** menu in the Preferences dialog box.

Managing preference settings

You can set your own preferences, or you can export and import preferences to save them or share them among a group of people who are working on similar or related projects. Affymetrix may provide preferences files to optimize IGB for analyzing certain data sets.

Exporting preference settings

You can export preference settings for use in different instances of IGB or to share with other IGB users.

1. Choose **File** menu > **Preferences**.
2. In the **Preferences** dialog box, choose **Preferences** menu > **Export Preferences**.

3. Choose the default destination and filename to which you want to save your preference settings.
4. Click **Save**.

Importing preference settings

You can import preference settings that you, Affymetrix, or another IGB user has exported using the procedure above.

Preferences in the file that you import are merged with your existing preferences.

WARNING: be sure you trust the provider of the file.

1. Choose **File** menu > **Preferences**.
2. In the **Preferences** dialog box, choose **Preferences** menu > **Import Preferences**.
3. Choose the saved preferences file and click **Open**.

Customizing the default appearance of tracks

You can customize the way an annotation type looks in the track.

For most annotation types, IGB remembers your settings so that the annotation type appears the same way each time you load it.

IGB does not remember settings for graphs and for temporary data resulting from arithmetic manipulations, such as intersections and unions of tracks.

You can make the following customizations to the appearance and behavior of tracks and the annotations they contain:

- Changing the name of the track
- Changing the color of annotations
- Changing the background color of the track
- Displaying both strands in one track or two
- Collapsing annotation rows into a single row
- Setting the maximum number of annotation rows to display
- Grouping exons
- Choosing annotation labels

You can set default values for most of these parameters so that all annotation types display by default using settings that you specify.

Your settings apply to both forward and reverse strands.

Some of these settings can also be made by right-clicking the name of the annotation type in the Tracks panel. The effect is the same.

To customize the appearance of a track:

1. Load the annotation type to customize.
2. Right-click on the track name in the Tracks panel and choose **Customize.**
3. In the **Preferences** dialog box, make selections for the annotation type to customize, following instructions in the sections that immediately follow this procedure.
    - If the name of the track is shown in italic font, the settings for that track will apply only to the current session. (This is the case for tracks that were created through manipulations such as intersections or unions of other tracks.)
    - If the track name is shown in normal font, the settings will be used for future sessions.
4. When you are finished setting preferences, close the Preferences dialog box.

Changing the name of the track

You can change the label of the data set as it appears in the tracks panel.

1. Double-click the cell in the **Display Name** column. Your cursor will blink in the field.
2. Type the name you want to use.
3. Press the <Enter> key on your keyboard.
4. If you are finished setting preferences, close the Preferences dialog box.

Changing the color of annotations

For an alternate method, see Changing the color of annotations in a track on page 31.

1. Click the color that is showing in the **Color** column.
2. Click a color in the palette.
3. Click **OK.**
4. If you are finished setting preferences, close the Preferences dialog box.

Changing the background color of the track

1. Click the color that is showing in the **Background** column.
2. Click a color in the palette.
3. Click **OK.**
4. If you are finished setting preferences, close the Preferences dialog box.

Displaying both strands in one track or two

To display forward and reverse strands in two tracks (+) and (-):

1. Click the checkbox in the **2 Tracks** column to add a checkmark.
2. When you are finished setting preferences, close the Preferences dialog box.

To display both strands in one track (+/-):

1. Click the checkbox in the **2 Tracks** column to remove the checkmark.
2. When you are finished setting preferences, close the Preferences dialog box.

Collapsing annotation rows into a single row

For more information about what this feature does, see Collapsing and Uncollapsing tracks on page 33.

To always display the data set with all annotation rows collapsed into the minimum height for the data set:

1. Click the checkbox in the **Collapsed** column to add a checkmark.
2. To display all annotation rows in the data set:
3. Click the checkbox in the **Collapsed** column to remove the checkmark.
4. When you are finished setting preferences, close the Preferences dialog box.

Setting the maximum number of annotation rows to display

If annotation rows are not collapsed, specify the maximum number of annotation rows to display:

1. Double-click the cell in the **Max Depth** column. Your cursor will blink in the field.

2. Type the maximum number of annotation rows to display. A suggested value is 5.
3. Press the <Enter> key on your keyboard.
4. If you are finished setting preferences, close the Preferences dialog box.

To cancel a restricted maximum that you have set:

1. Double-click the cell in the **Max Depth** column. Your cursor will blink in the field.
2. Clear the value in the field.
3. Press the <Enter> key on your keyboard.
4. If you are finished setting preferences, close the Preferences dialog box.

Grouping exons

This preference determines whether to connect groups of exons into transcripts.

For data with no intron-exon structure, such as repeats or contigs, exons should not be connected. For items that do have intron-exon structure, not connecting the exons may affect whether labels are displayed.

To always display annotations in the data set with exons connected or not connected:

1. Click the checkbox in the **Connected** column to add or remove a checkmark.
2. When you are finished setting preferences, close the Preferences dialog box.

Choosing annotation labels

You can hide annotation labels or display any label information that is available for the data set.

To see what label options are available for the data set that you are customizing, click an annotation in the browser window, then click the **Selection Info** tab. Look at the properties listed in the **property** column.

1. Double-click the cell in the **Label field** column. Your cursor will blink in the field.
2. Type the name of the source for annotation labels:

- **id** - Most data sets include **id** values.
- To display no label, leave the field blank. This reduces the amount of memory that IGB requires to run, and also speeds-up drawing of the display.
- Other property names, if they appear in the **Selection Info** tab. Enter them as they appear in Selection Info. For example, "gene name" is a possible choice for RefSeq entries.

3. Press the <Enter> key on your keyboard.
4. When you are finished setting preferences, close the Preferences dialog box.

Note: If you have chosen to display exons without connection lines, labels will not be visible if the individual exons do not have the property that you specify.

Adding DAS/1 data sources

DAS/1 sources for annotation types are introduced in Loading annotation types from a DAS server on page 11.

If you need to use data from a DAS/1 source that is not listed in the **Load DAS Features** dialog box, you can add that data server. There is currently no way in to add a DAS/2 data source to the available servers listed in IGB.

To add to or modify the list of DAS servers that appears in the **Load DAS Features** dialog box:

1. Obtain the URL(s) for the DAS server(s) that you want to use.
2. Choose **File** menu > **Preferences**.
3. Click the **DAS Servers** tab.
4. Perform one of the following tasks.

To add a DAS server to the list:

1. Click the **Add** button.
2. Enter the URL and a convenient name to identify this data source, check the **Enabled** checkbox if you want to enable this data source now, then click the **Apply Changes** button.

Both the *http://* and *https://* protocols are accepted in the URLs.

To change the URL or identifying name of an existing DAS server on the list:

1. Click the DAS server to change.

2. Click the **Edit** button.
3. Make changes.
4. Click the **Apply Changes** button.

To temporarily hide a DAS server from the list in the **Load DAS Features** dialog box:

1. Click the DAS server to hide.
2. Click the **Enabled** checkbox to deselect it.

To again view a hidden DAS server:

1. Click the DAS server to show.
2. Click the **Enabled** checkbox to place a checkmark in it.

To remove a DAS server from the list (you cannot remove UCSC or NetAffx-Align; hide them instead if you do not want to view them):

1. Click the DAS server to remove.
2. Click the **Remove** button.

Changing keyboard commands

You can set keyboard shortcuts for many menu and other functions. The list of functions for which you can set keyboard shortcuts appears in the Preferences described in the following procedure.

Some action names may appear that don't seem to have any function. This could happen if you have installed and then un-installed some plugins, or if some actions have been re-named.

Keyboard Shortcut keys 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 in a separate window can give more predictable shortcut behavior.

To set or change keyboard commands:

1. Choose **File** menu > **Preferences**.
2. Click the **Shortcuts** tab.
3. Click the action for which you want to establish a keyboard command.
4. Click in the field labeled **Type a shortcut for** <the command you selected>.

5. Press the keys that you want to assign to the function you selected, for example, Ctrl-Shift-X. You cannot use keystrokes that your operating system uses, such as Ctrl-Alt-Delete on Windows. If you make a mistake, simply type the keystroke(s) that you want, and the new keystrokes will replace the erroneous keystrokes that you entered.
6. Click the **Apply this shortcut** button.
7. Your new keyboard shortcut(s) will be available the next time you launch IGB.

To clear keyboard commands:

1. Click the shortcut from which you want to remove the command.
2. Click **Clear this shortcut**.
3. Your changes take affect after you close and then re-launch IGB.

Setting graph preferences

You can set preferences for the way graphs display in the viewer window.

These preferences are very useful when the scale of the graphs that you are viewing makes it difficult to view significant features, either because the score values are very small or because they are very large.

Changes affect graphs that you load after you complete the procedure below.

To change default characteristics of graphs:

1. Choose **File** menu > **Preferences**.
2. Click the **Graphs** tab.
3. Choose options for the following:
    - Whether graphs float or are attached to tracks
    - For floating graphs, specify the default pixel height. Choose a larger number for easier visibility, or choose a smaller number to minimize graph height if you are viewing many graphs and/ or tracks simultaneously.
    - For graphs that are attached as a track, specify a y-axis coordinate height. Choose a larger number for easier visibility, or choose a smaller number to minimize the height of the track if you are viewing many graphs and/or tracks simultaneously.
4. If you are attaching floating graphs to a track, choose whether to retain in the attached graph the pixel height changes that you have made to the preferences, or use the default y-axis coordinate height values. If you want to apply the height that you specified for attached

graphs to graphs that are currently displayed in IGB, choose **Default Coord Height** for this option.

5. If you are finished setting preferences, close the Preferences dialog box.
6. If you want to apply the settings that you just made to an attached graph that is currently displayed in IGB, click the colored bar at the left end of the graph, then click the **Graph Adjuster** tab, then click the **Floating** radio button and then the **Attached** radio button. Toggling between Floating and Attached will cause the graph to redraw in the browser with the coordinate height that you specified in the graph preferences.

Setting color and axis display preferences

Setting axis display – option descriptions

You can set the following preferences for the appearance of the viewer window:

Item	Description
Foreground	The color of the background on which annotations appear in the viewer.
Background	The color of the Coordinates track that runs horizontally across the center of the viewer window.
Number format	The way base position numbers are noted in the Coordinates track (X-axis).

Setting viewer background color

This setting determines the color of the background on which annotations appear in IGB. By default this is black.

1. Choose **File** menu > **Preferences**.
2. Click the **Other Options** tab.
3. To change the color of the background on which annotations appear in the viewer, click the button beside **Foreground**.
4. (Optional) To see alternate methods to choose colors, click the HSB and/or RGB tabs.

5. Click a color in the palette.
6. Click **OK**.

Setting Coordinates track color

This setting determines the color of the axis on which coordinates appear in IGB. By default this is white.

1. Choose **File** menu > **Preferences**.
2. Click the **Other Options** tab.
3. To change the color of the Coordinates track, click the button beside **Background**.
4. (Optional) To see alternate methods to choose colors, click the HSB and/or RGB tabs.
5. Click a color in the palette.
6. Click **OK**.

Specifying Coordinates track number format

This setting determines the number format of coordinates as they appear in the Coordinates track.

1. Choose **File** menu > **Preferences**.
2. Click the **Other Options** tab.
3. For **Number format**, choose one of the following options:
    - **Full** - Displays in the format of this example: 20000kb
    - **Comma** - Displays in the format of this example: 20,000,000
    - **Abbrev** - Displays in the format of this example: 20M

Setting edge-matching colors

By default, perfect edge matches are indicated with white, and fuzzy edge matches are indicated with gray. To change these indicator colors:

1. Choose **File** menu > **Preferences**.
2. Click the **Other Options** tab.
    - To change the indicator of perfect edge matches, click the button beside **Standard color**.
    - To change the indicator of fuzzy edge matches, click the button beside **Fuzzy matching color**.
3. Click a color in the palette.

4. Click **OK**.

Setting ORF and Stop-codon colors

By default, ORFs are indicated in the sliced view with green and stop codons are indicated with a pinkish-beige.

To change ORF and stop codon indicator colors:

1. Choose **File** menu > **Preferences**.
2. Click the **Other Options** tab.
    - To change the indicator of stop codons, click the button beside **Stop Codon**.
    - To change the indicator of ORFs, click the button beside **Dynamic ORF**.
3. Click a color in the palette.
4. Click **OK**.

Setting other options

Other options you can set:

Option	Description
Ask before exiting	Whether IGB asks you to confirm that you really want to exit when you choose **File** menu > **Exit** or close the IGB window by clicking the [X] in the title bar.
Keep hairline in view	Whether to automatically prevent the hairline from moving outside the view as you scroll.
Make graphs from scored intervals	Whether to automatically create graphs from data in scored interval (.egr) files. When importing .egr files, the default is to display the data as annotation-style rectangular blocks. Check this checkbox to display the data in graph format instead. If you are planning to use the pivot view feature, deselect this option.
Show DAS query genometry	Intended for advanced users, for debugging of DAS servers. Shows the coordinate regions used in queries to the DAS server.

To access these options:

1. Choose **File** menu > **Preferences**.
2. Click the **Other Options** tab.
3. Check the checkboxes beside the options you want to enable.

License

Source code for IGB is released under the Common Public License, v1.0, an OSI approved open source license. IGB uses other open source software packages that are not distributed with the IGB source code release, including Xerces and Jetspeed from Apache, and Jetty from Mortbay Consulting, which are covered by their own open source licenses. IGB is Copyright (c) 2000-2004 Affymetrix, Inc. Research and development of IGB is supported in part by NIH grant R01HG003040.

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.

Index