Integrative Genomics Viewer
IGV

Tom Carroll

MRC Clinical Sciences Centre

- Introduction to IGV.
  - What is IGV.
  - How to run IGV.
- Navigating IGV.
  - The IGV user interface.
  - Moving around genomes.
- Loading and visualising data.
  - Genome information and annotation.
  - User supplied data.
  - Sample information.
  - External data.
- Displaying genomics data
  - Basic visualisation.
  - Data dependent visualisation.

What is IGV?

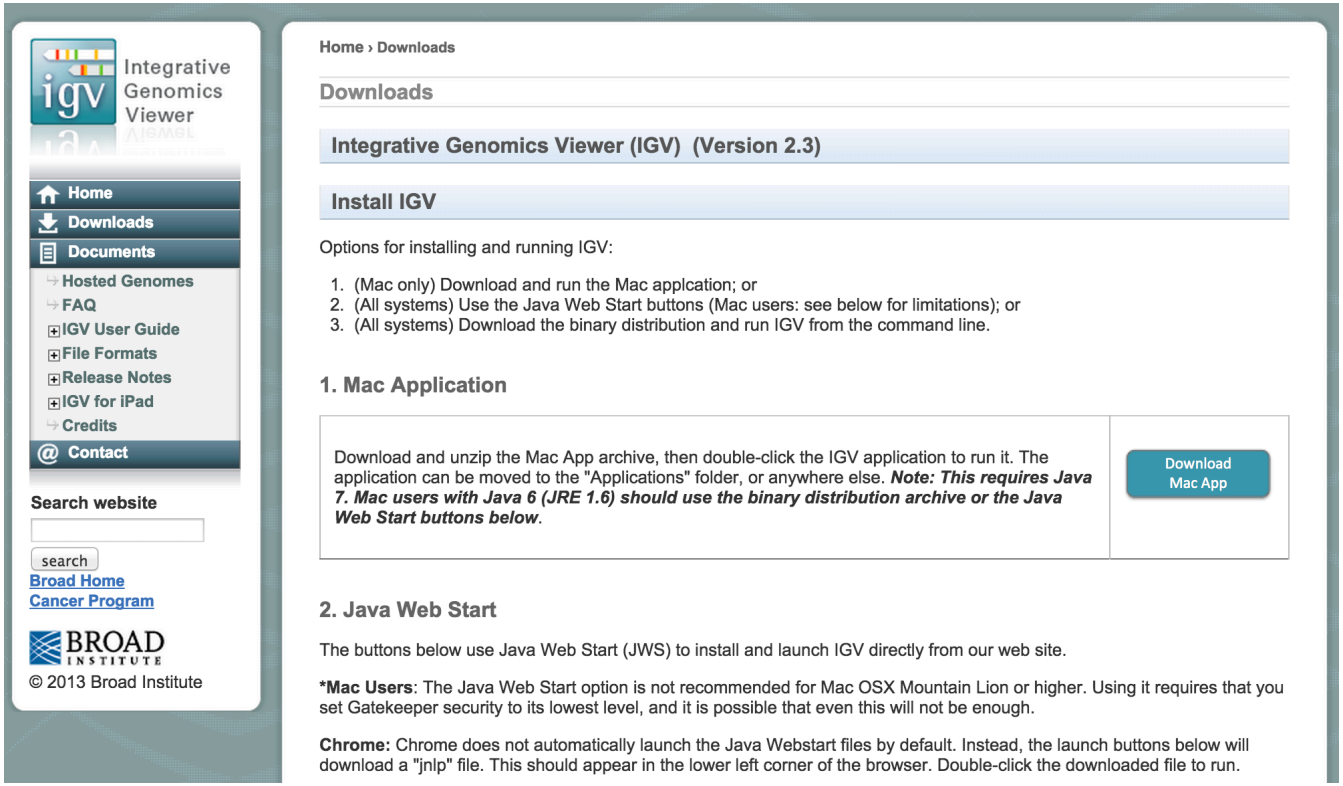
- Created by the Broad institute.
- Genome browser.
  - Visualises genomic data (expression, ChIP, resequencing, multiple alignment, shRNA)
  - Handles most common genomic data types.
- Java Desktop application
  - No dependence on server
  - Loads data locally or from URL, consumes memory and CPU.

How to run IGV?

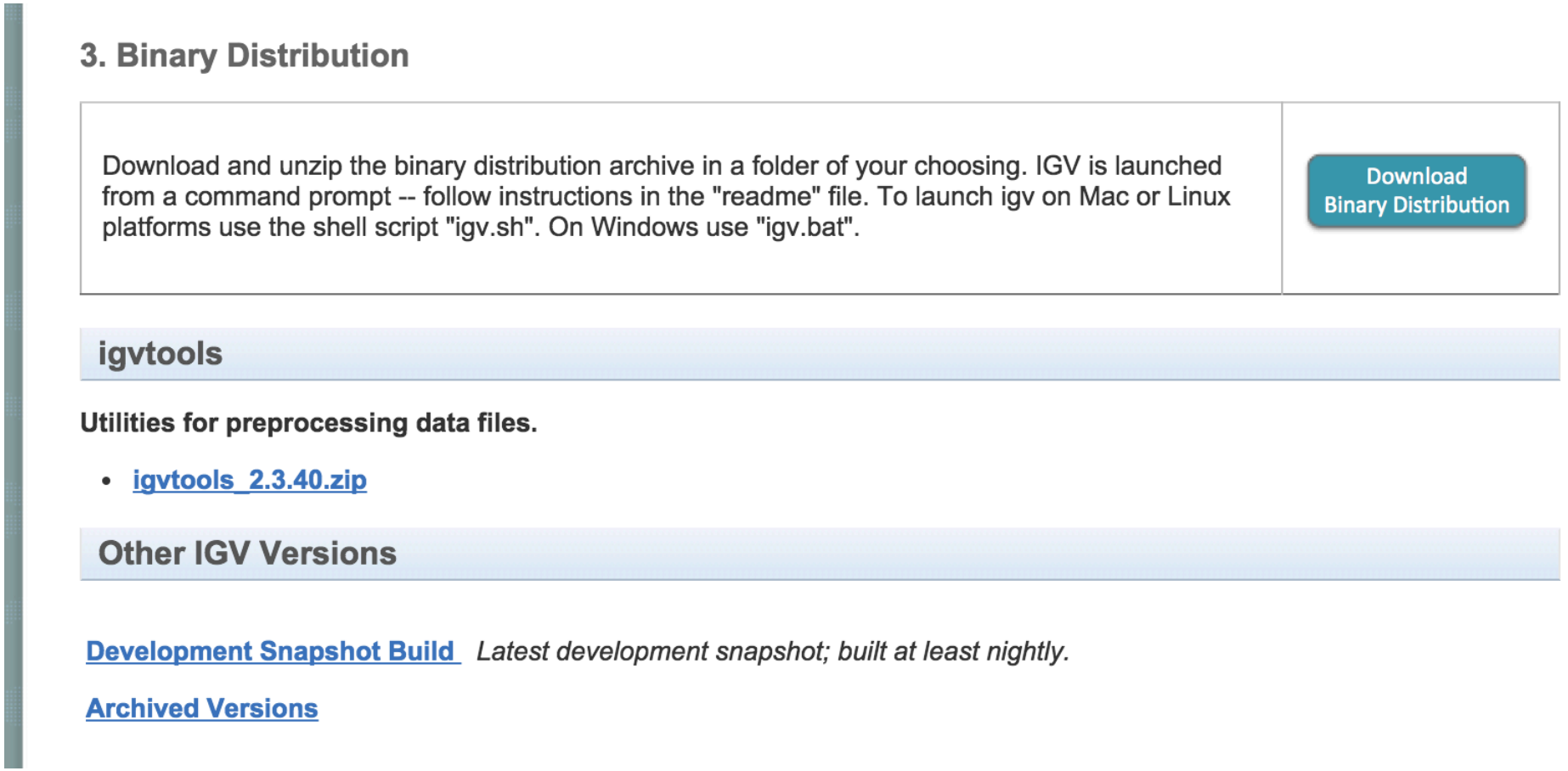
- Requires Java

- IGV available from Broad

http://www.broadinstitute.org/software/igv/download



- Download to computer.
- Runs locally.
- Archived versions available



3. Binary Distribution

Download and unzip the binary distribution archive in a folder of your choosing. IGV is launched from a command prompt -- follow instructions in the "readme" file. To launch igv on Mac or Linux platforms use the shell script "igv.sh". On Windows use "igv.bat".	Download Binary Distribution

igvtools

Utilities for preprocessing data files.

- igvtools_2.3.40.zip

Other IGV Versions

Development Snapshot Build *Latest development snapshot; built at least nightly.*

Archived Versions

- Runs from webstart.
- Always runs latest version of IGV.

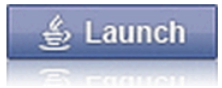
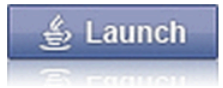
2. Java Web Start

The buttons below use Java Web Start (JWS) to install and launch IGV directly from our web site.

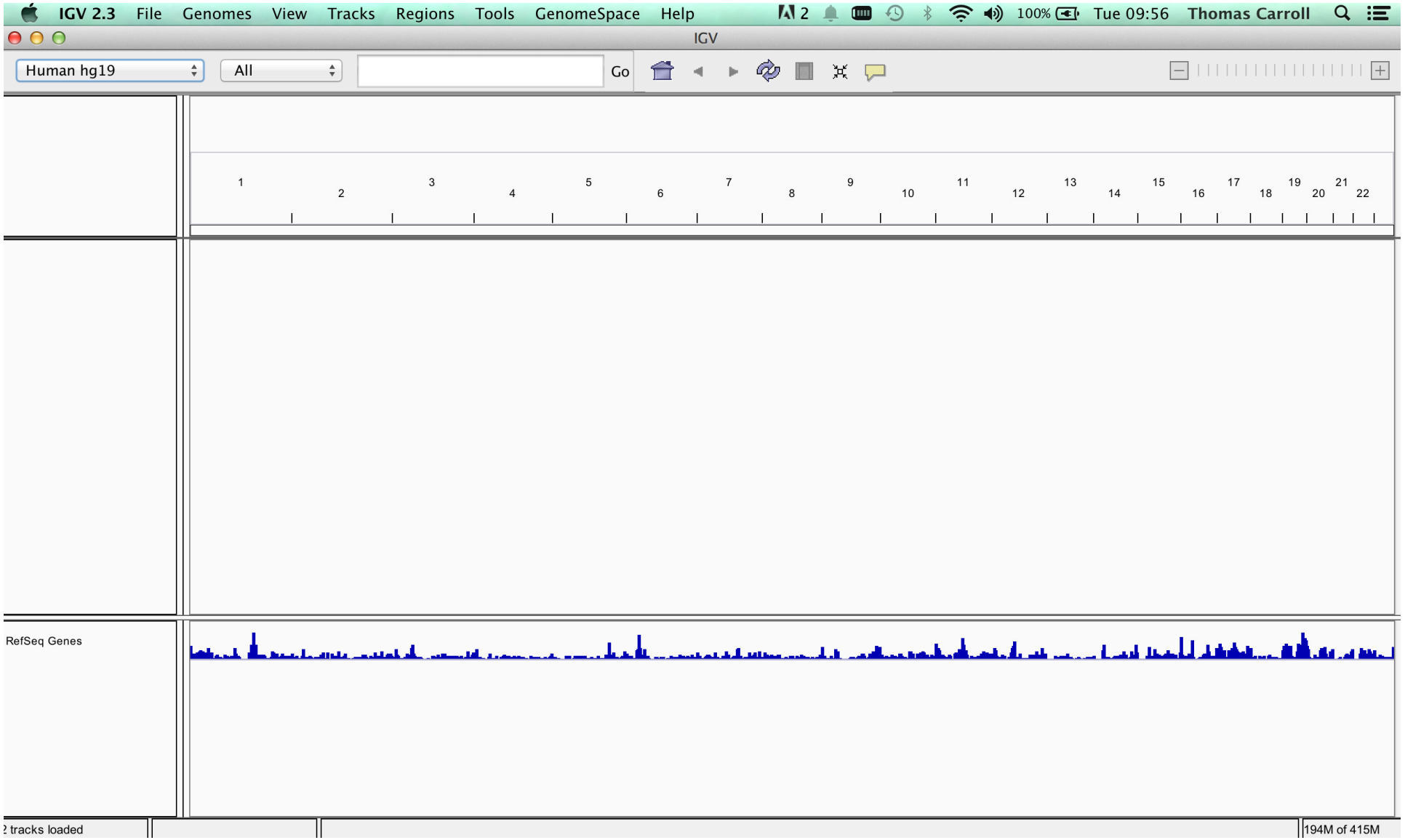
***Mac Users**: The Java Web Start option is not recommended for Mac OSX Mountain Lion or higher. Using it requires that you set Gatekeeper security to its lowest level, and it is possible that even this will not be enough.

Chrome: Chrome does not automatically launch the Java Webstart files by default. Instead, the launch buttons below will download a "jnlp" file. This should appear in the lower left corner of the browser. Double-click the downloaded file to run.

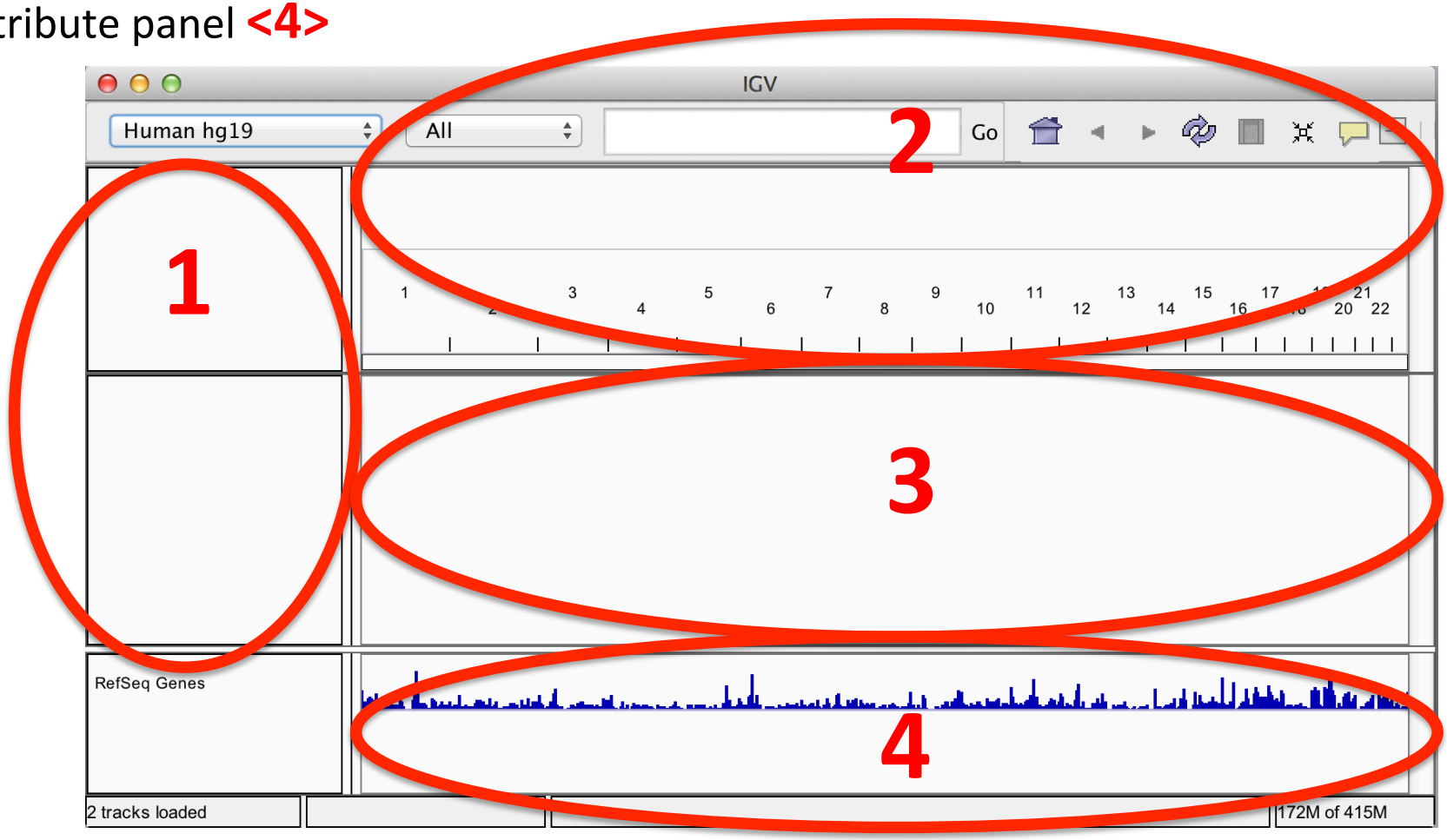
Windows users: To run with more than 1.2 GB of memory you must install 64-bit Java. ***Most Windows installs do not include 64-bit Java by default, even if the operating system is 64-bit***. Attempting to use the 2GB or greater launch options with 32-bit Java will result in the error *"could not create virtual machine"*.

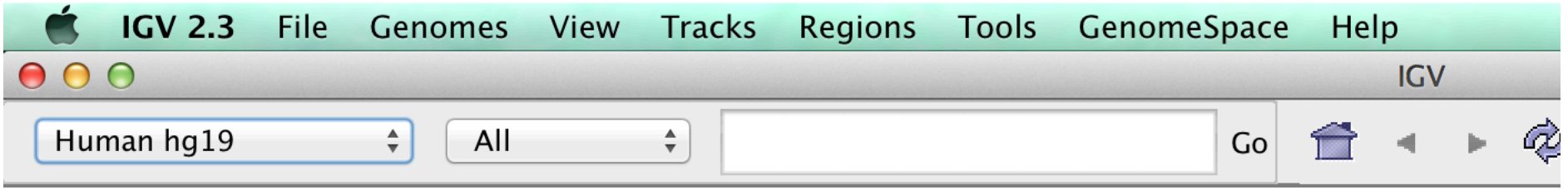
 Launch with 750 MB	 Launch with 1.2 GB Maximum usable memory for Windows OS with 32-bit Java.	 Launch with 2 GB Maximum usable memory for 32-bit MacOS.	 Launch with 10 GB For large memory machines with 64-bit Java.

IGV GUI


IGV 2.3 File Genomes View Tracks Regions Tools GenomeSpace Help
100% Tue 09:56 Thomas Carroll
IGV
Human hg19
All
Go
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
RefSeq Genes
tracks loaded
194M of 415M

- IGV GUI:-
  - Sample information panel **<1>**
  - Genome Navigation panel **<2>**
  - Data panel **<3>**
  - Attribute panel **<4>**

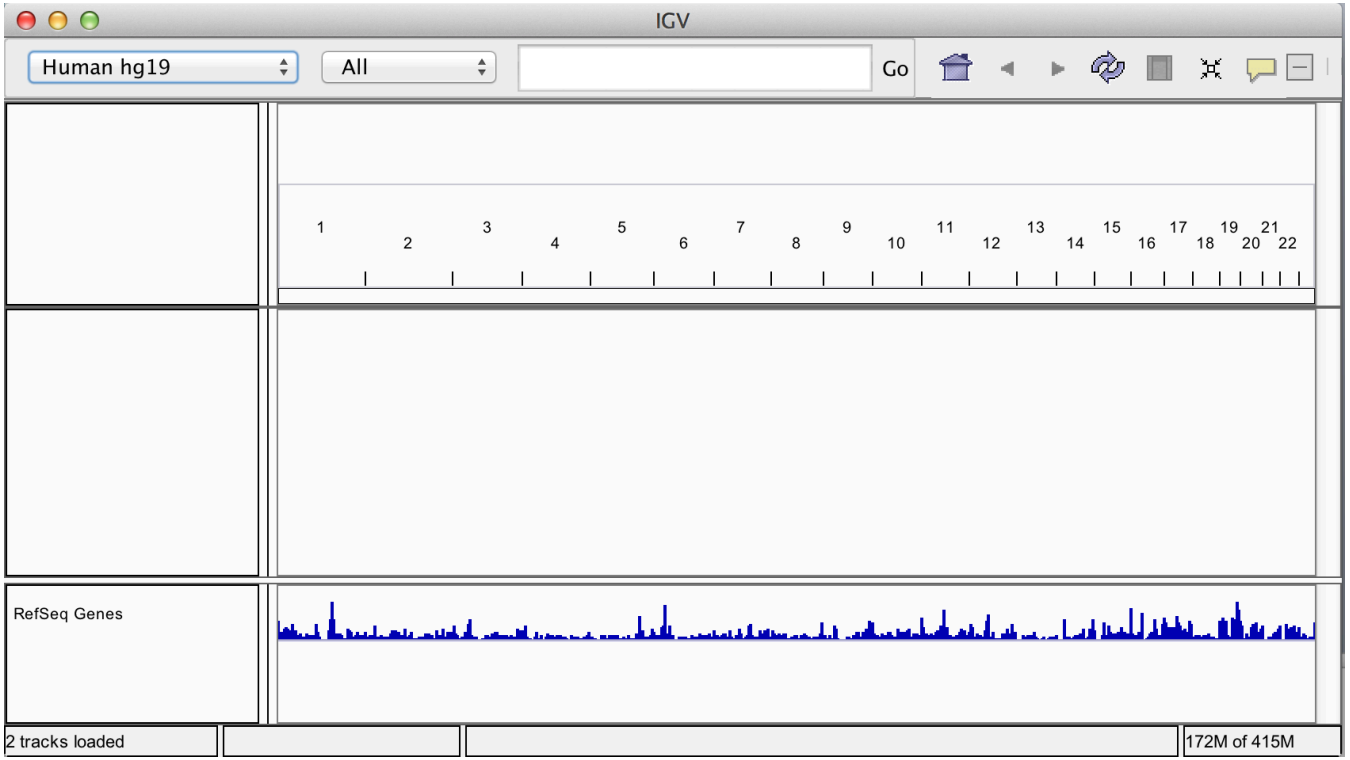




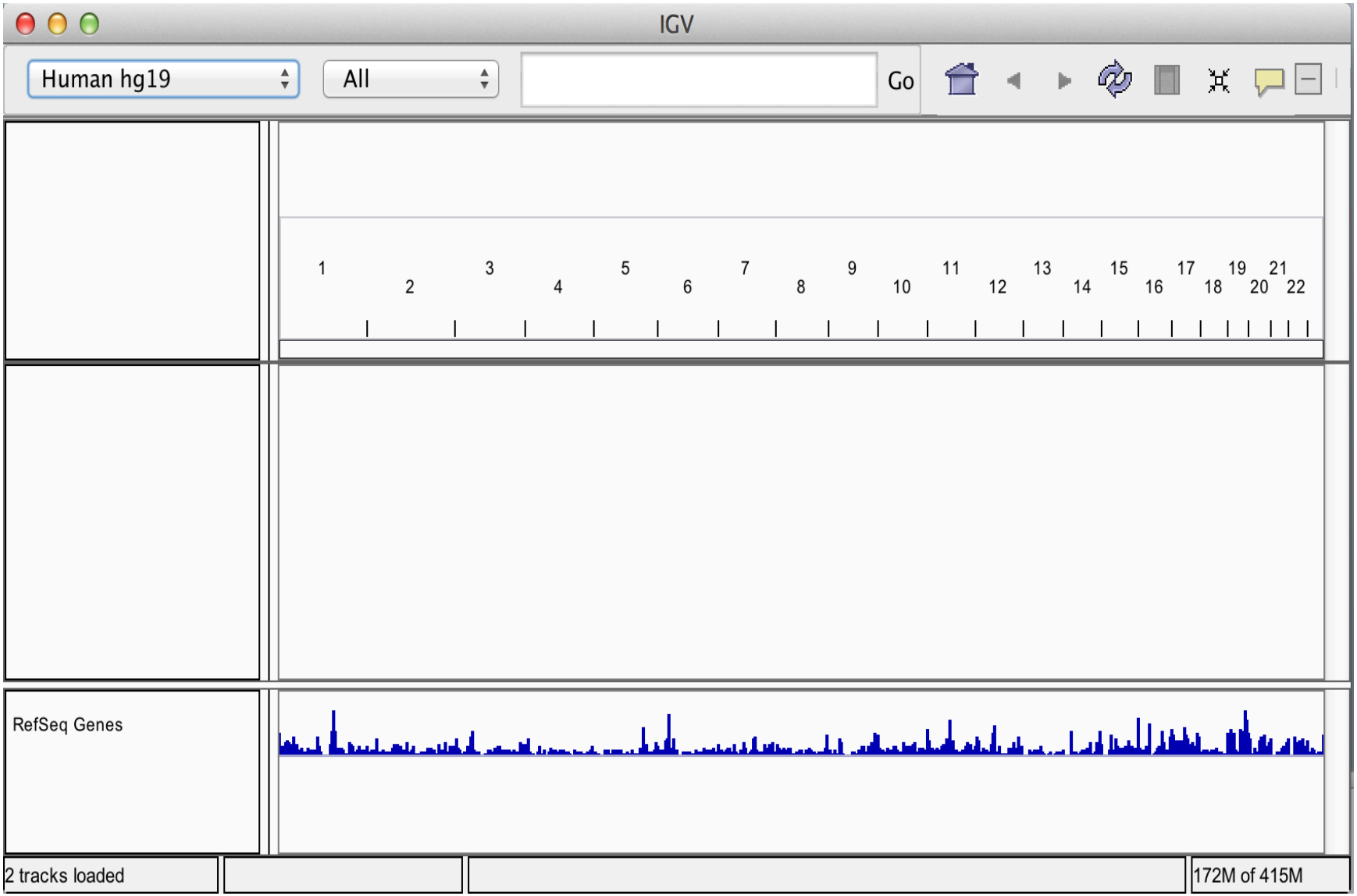
- Menu bar
  - **File,** Load data/sample information.
  - **Genome,** Load and manage genomes.
  - **View,** Display preferences.
  - **Tracks,** Group/sort/filter data tracks.
  - **Regions,** Create region/gene lists.
  - **Tools,** Access to Integrated tools (IGVtools/ Bedtools).
  - **GenomeSpace,** Export/import from Genomespace

Moving around genomes

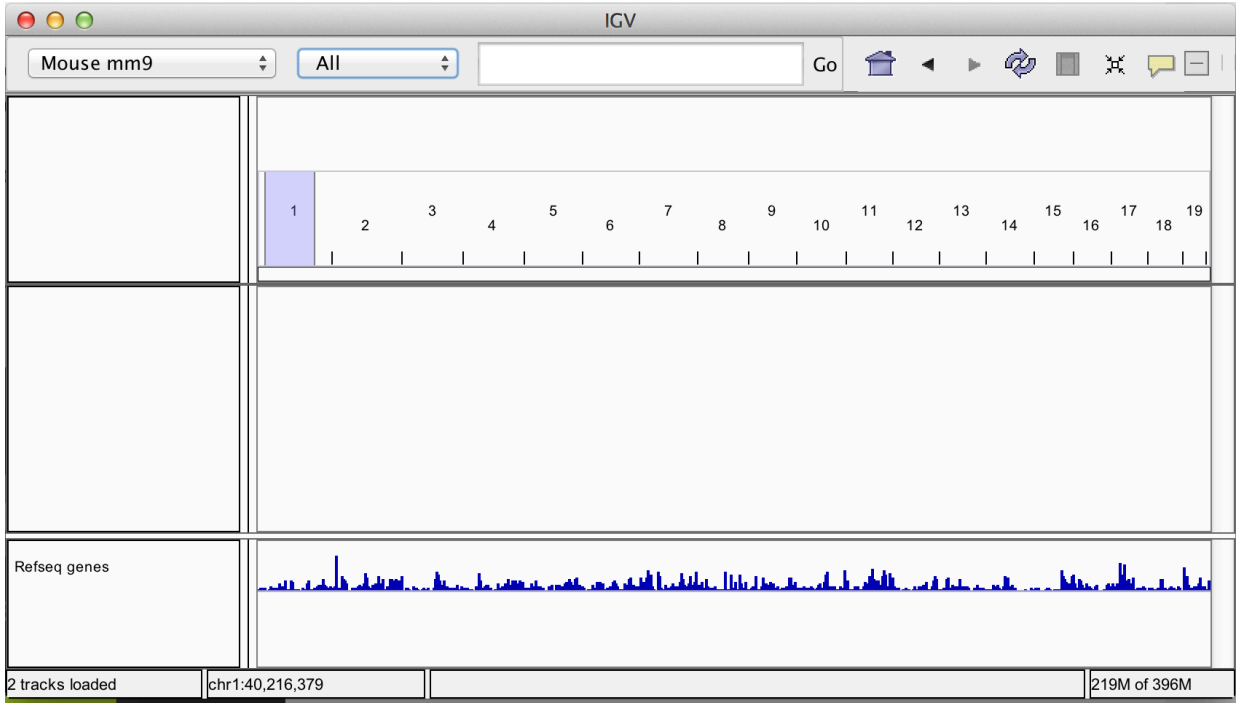
- Cytoband selection and zooming
- Scrolling
- Selection of region of interest

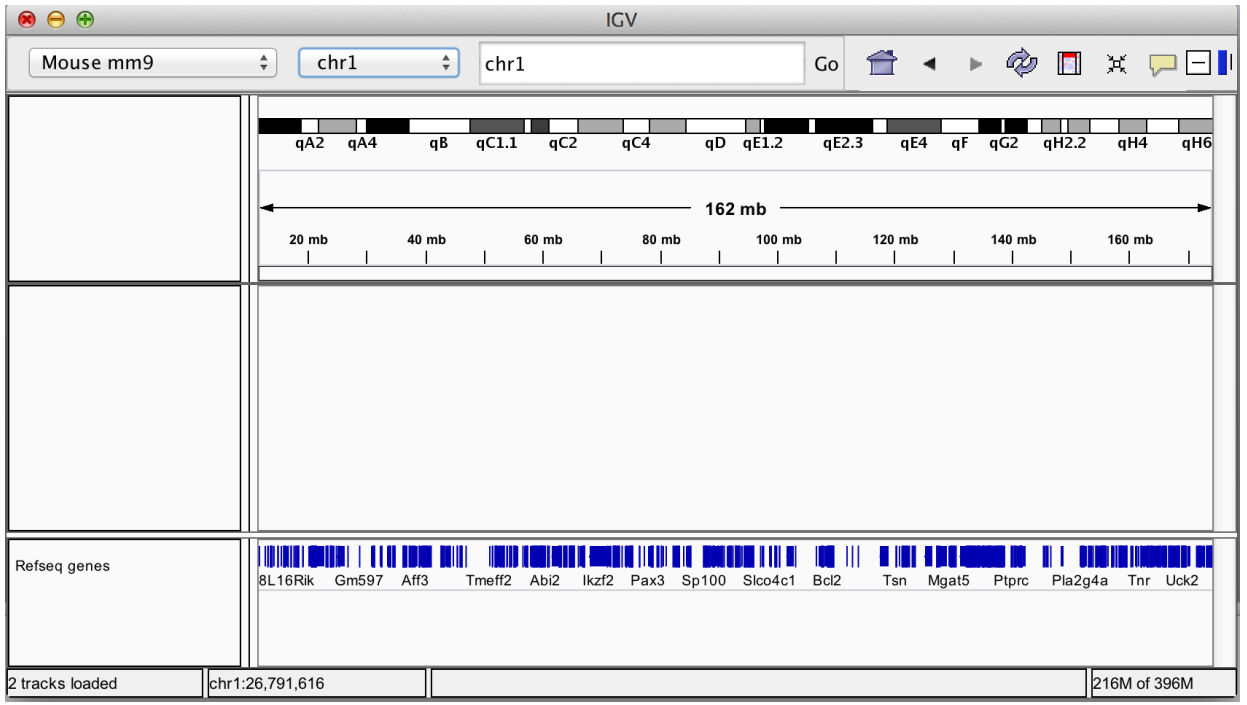


Whole genome view

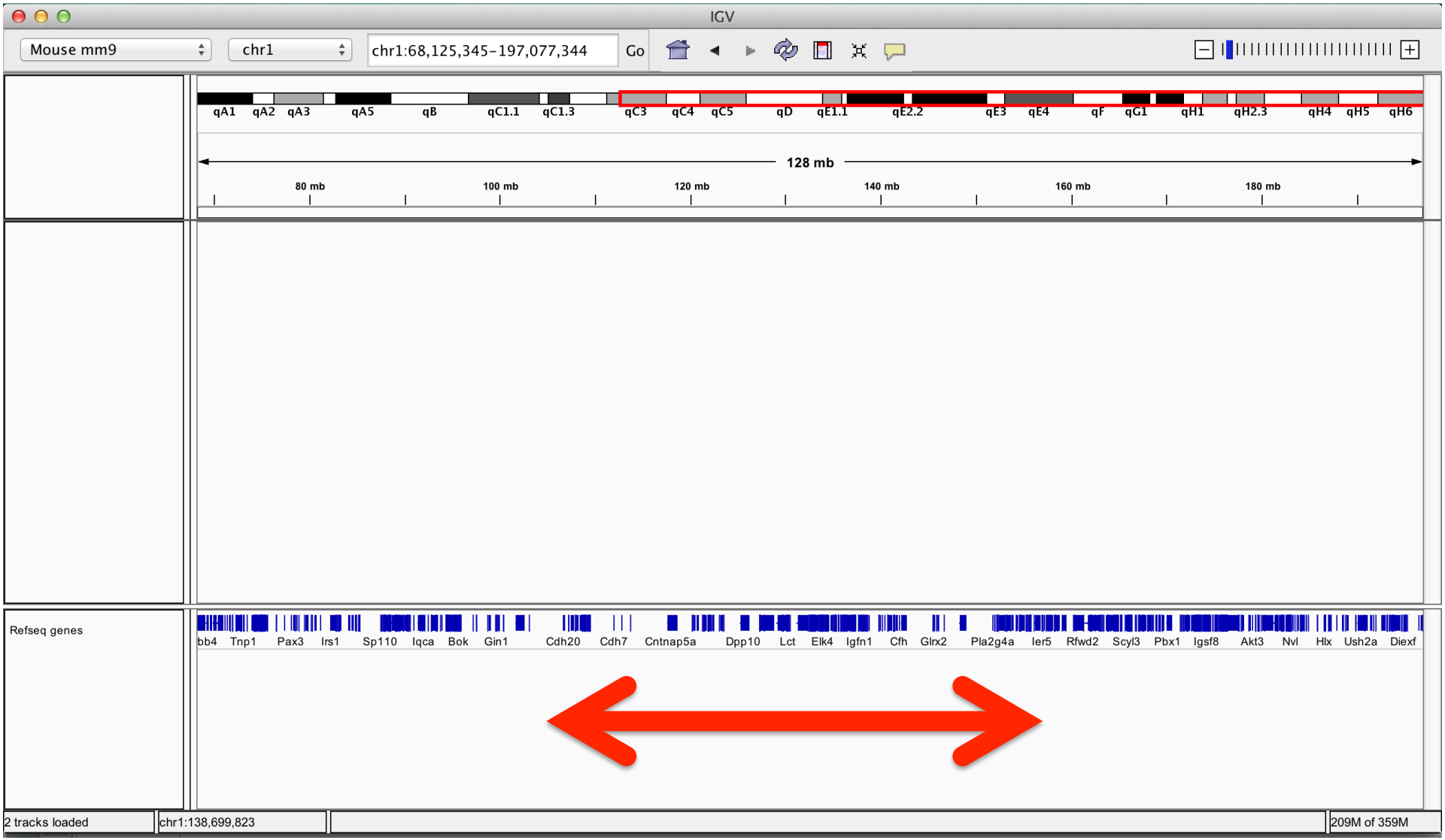


Zooming

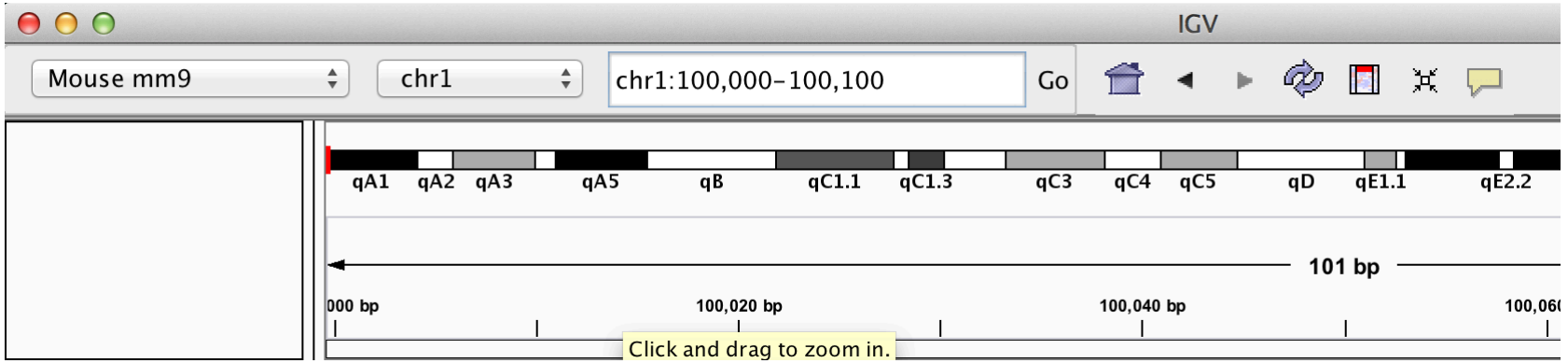

IGV
Mouse mm9
All
Go
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
Refseq genes
2 tracks loaded
chr1:40,216,379
219M of 396M

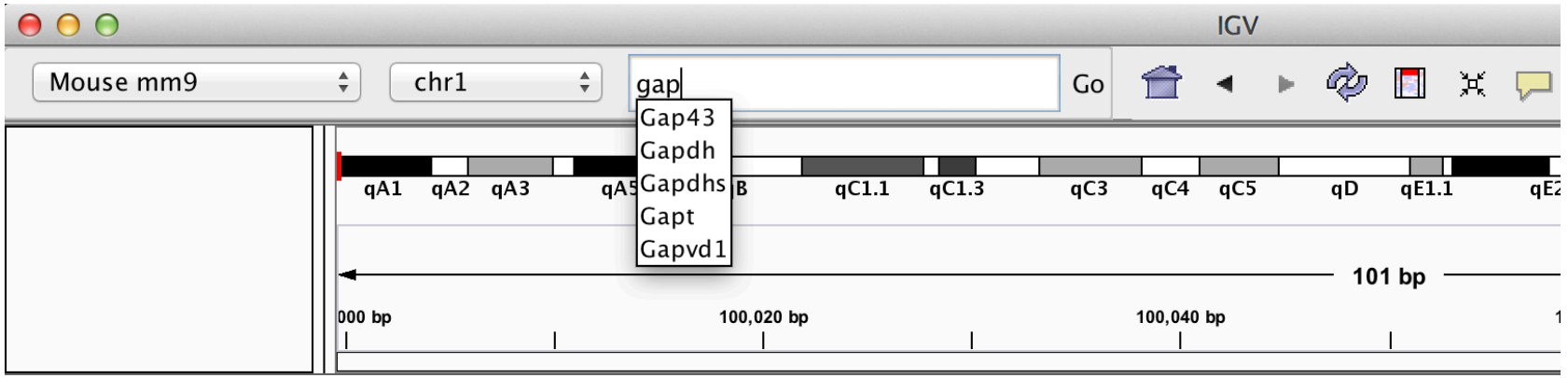

IGV
Mouse mm9
chr1
chr1
Go
qA2
qA4
qB
qC1.1
qC2
qC4
qD
qE1.2
qE2.3
qE4
qF
qG2
qH2.2
qH4
qH6
162 mb
20 mb
40 mb
60 mb
80 mb
100 mb
120 mb
140 mb
160 mb
Refseq genes
8L16Rik
Gm597
Aff3
Tmeff2
Abi2
Ikzf2
Pax3
Sp100
Slco4c1
Bcl2
Tsn
Mgat5
Ptprc
Pla2g4a
Tnr
Uck2
2 tracks loaded
chr1:26,791,616
216M of 396M

Scrolling


IGV
Mouse mm9
chr1
chr1:68,125,345-197,077,344
Go
qA1 qA2 qA3 qA5 qB qC1.1 qC1.3 qC3 qC4 qC5 qD qE1.1 qE2.2 qE3 qE4 qF qG1 qH1 qH2.3 qH4 qH5 qH6
128 mb
80 mb
100 mb
120 mb
140 mb
160 mb
180 mb
Refseq genes
bb4 Tnp1 Pax3 Irs1 Sp110 Iqca Bok Gin1 Cdh20 Cdh7 Cntnap5a Dpp10 Lct Elk4 Igfn1 Cfh Glrx2 Pla2g4a Ier5 Rfwd2 Scyl3 Pbx1 Igsf8 Akt3 Nvl Hlx Ush2a Diexf
2 tracks loaded
chr1:138,699,823
209M of 359M

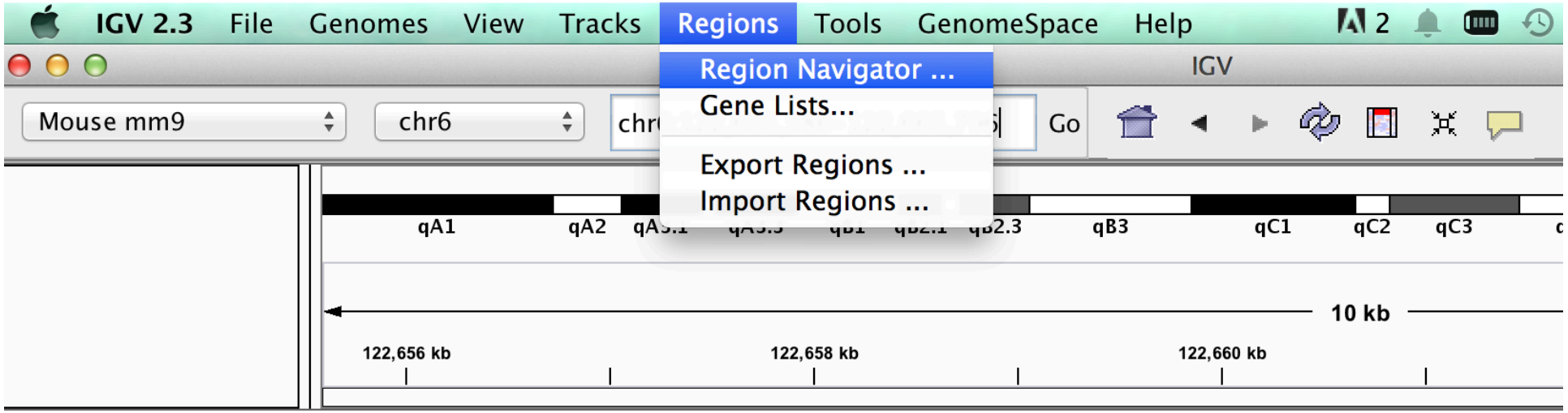
Jump to Region

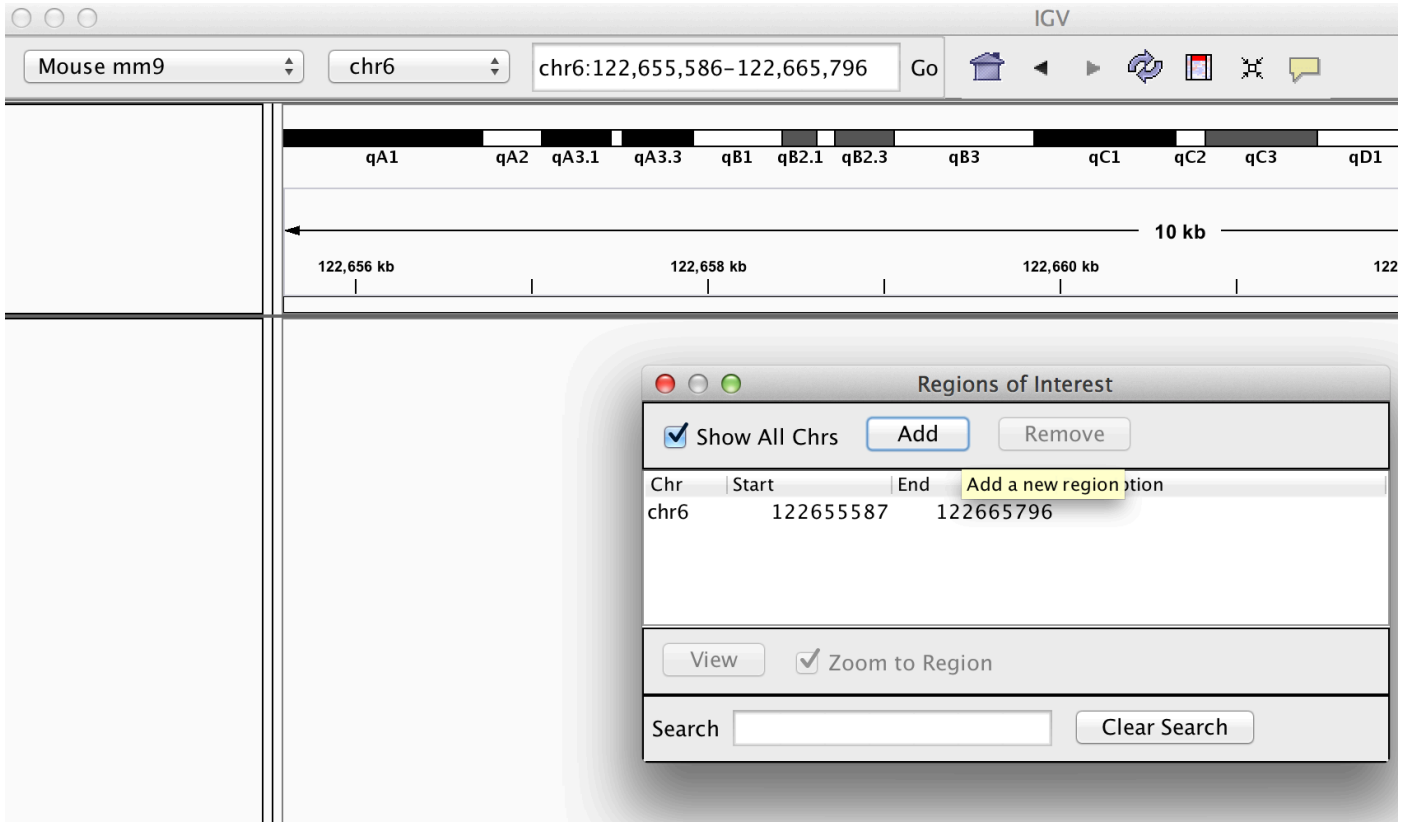




“Bookmarking” regions of interest

- Regions may be added to “Regions of interest”
- These act as bookmarks for areas of particular interest


IGV 2.3
File
Genomes
View
Tracks
Regions
Tools
GenomeSpace
Help
Region Navigator ...
Gene Lists...
Export Regions ...
Import Regions ...
IGV
Mouse mm9
chr6
Go
qA1
qA2
qB3
qC1
qC2
qC3
10 kb
122,656 kb
122,658 kb
122,660 kb

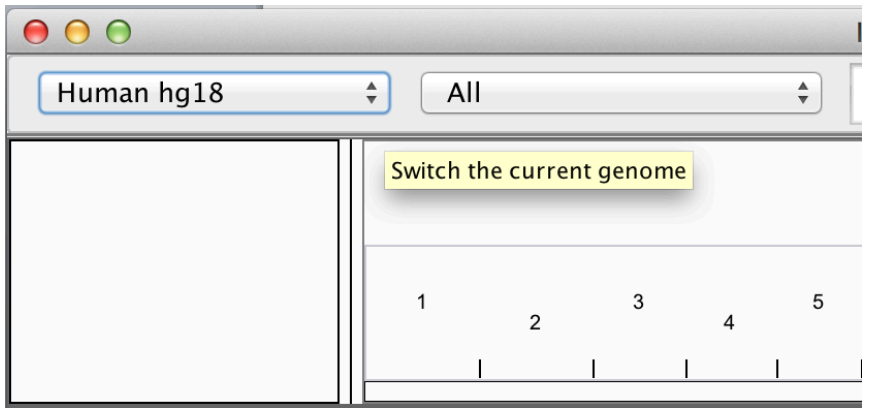

IGV
Mouse mm9
chr6
chr6:122,655,586-122,665,796
Go
qA1
qA2
qA3.1
qA3.3
qB1
qB2.1
qB2.3
qB3
qC1
qC2
qC3
qD1
10 kb
122,656 kb
122,658 kb
122,660 kb
Regions of Interest
Show All Chrs
Add
Remove
Chr
Start
End
Add a new region
chr6
122655587
122665796
View
Zoom to Region
Search
Clear Search

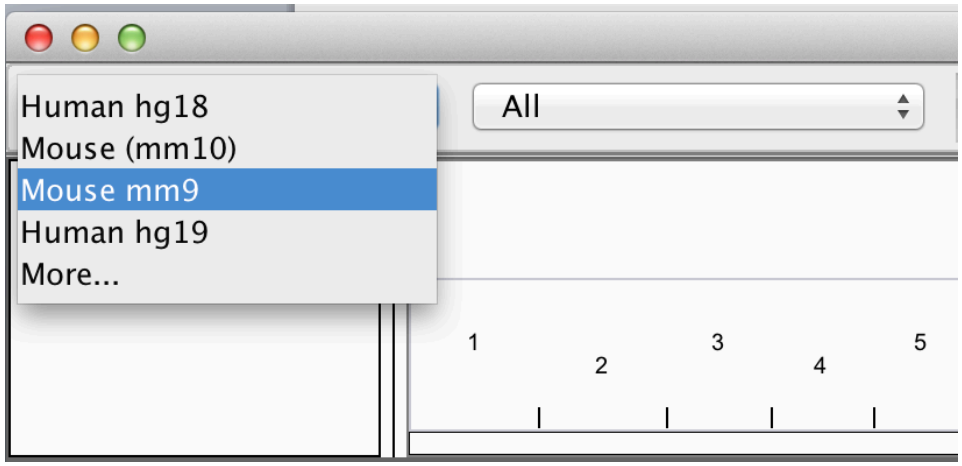
“Book marked” Region Of Interest

Loading data in IGV

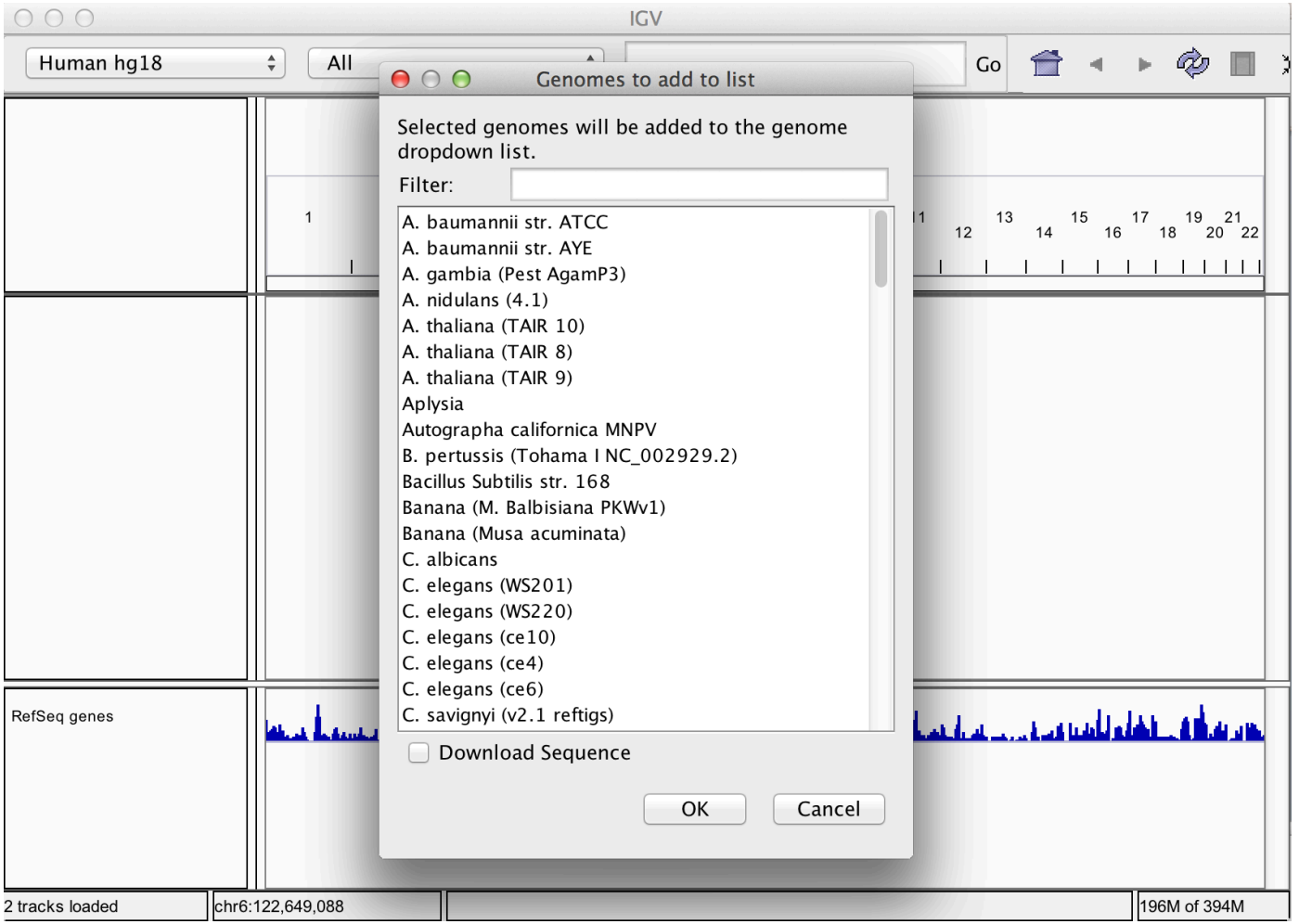
Loading Genome Information

- Most genomes can be selected from dropdown.



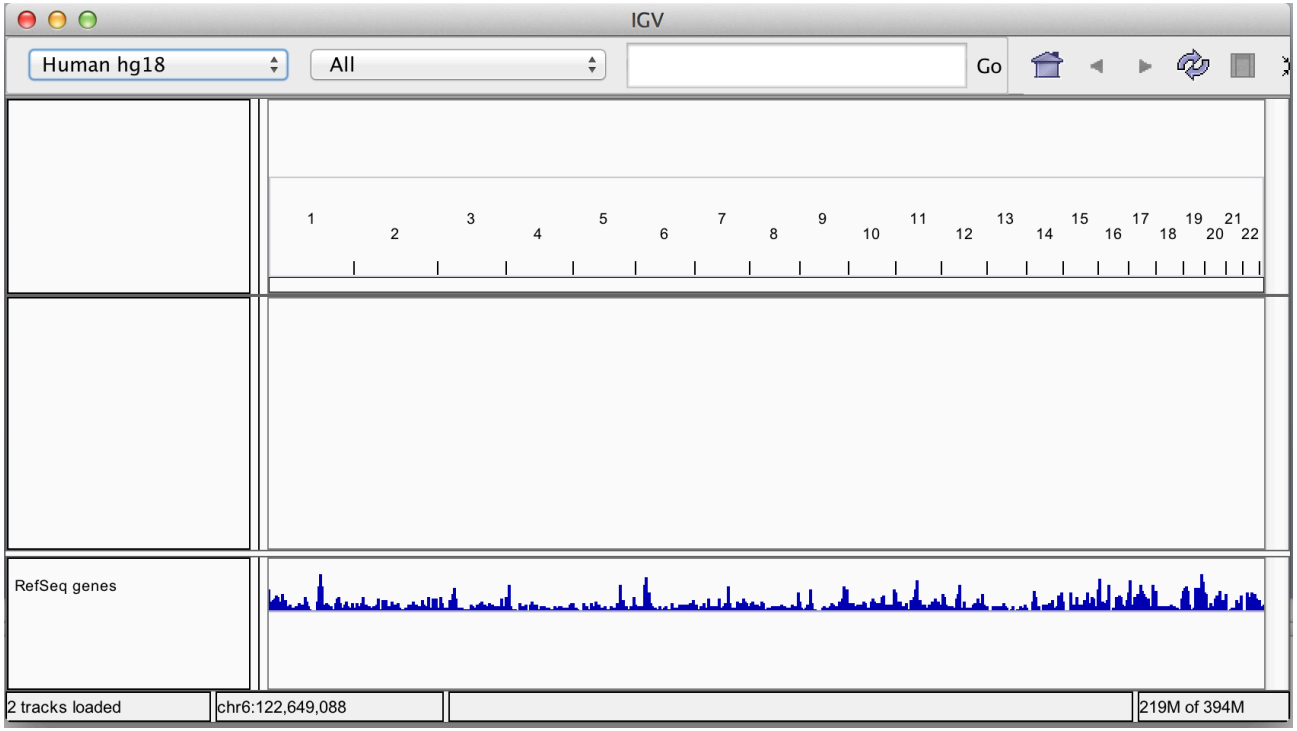


- Genomes not included may be downloaded from repository



Loading genome annotation

- For supported genomes, gene positions are automatically included in “feature” panel.
- Additional gene positions can be loaded into IGV in gff format.



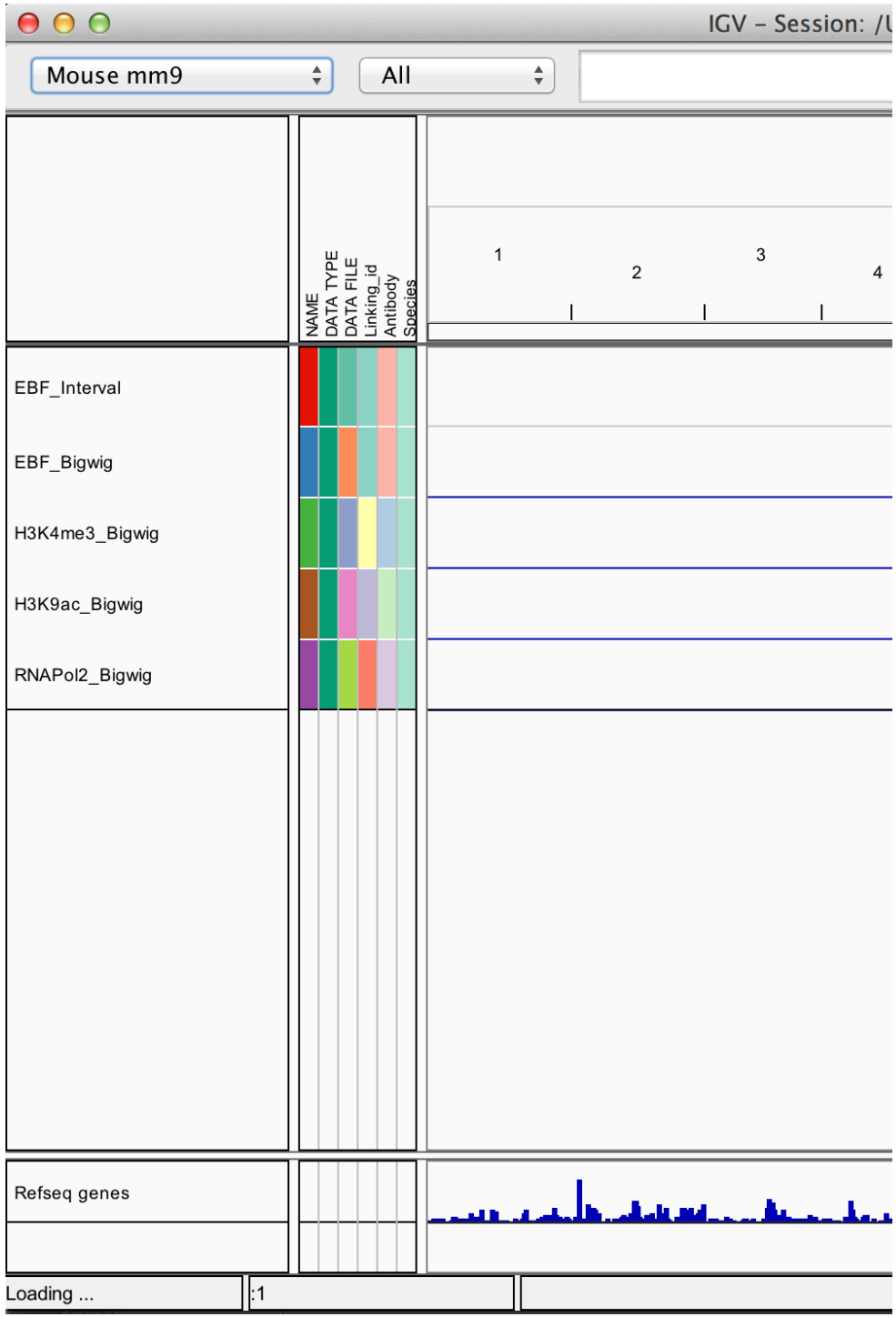
Loading sample data.

- Most common formats can be loaded into IGV through file menu
- Acceptable data formats include:-
    - BED (.bed)
    - BAM and index (.bam with .bai/.bam.bai)
    - BigWig (.bw)
    - BedGraph/Wig (.bedGraph, .wig)

    And many more…

Loading sample metadata

- IGV allows the inclusion of information on samples.
- Sample information is then included in sample information panel.



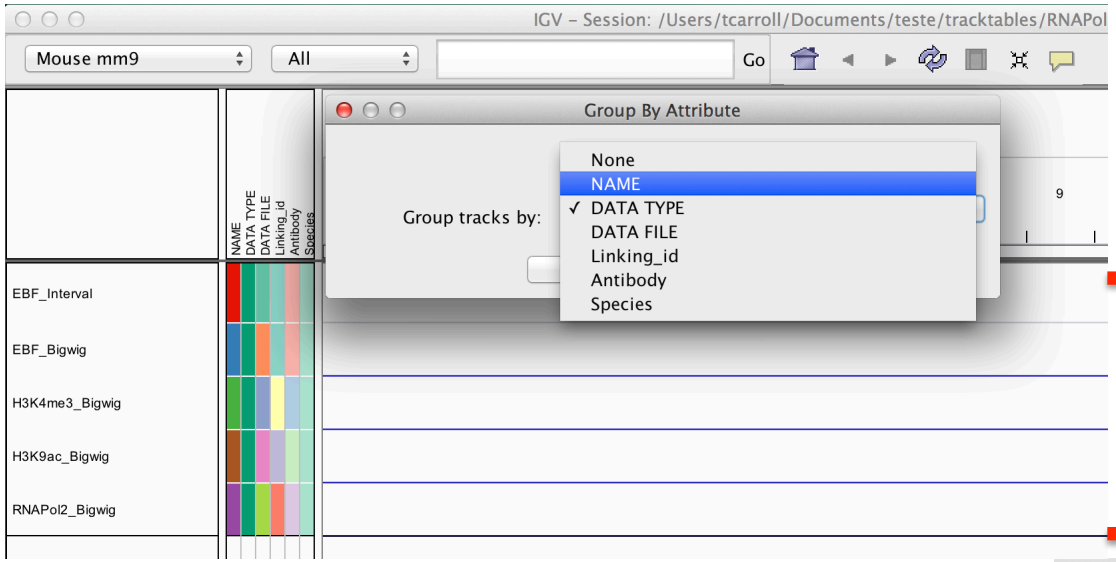
- Example Sample Information file

http://www.broadinstitute.org/igvdata/exampleFiles/exampleSampleInfo.txt

TRACK_ID	Data_Type	PARTICIPANT_ID	SAMPLE_ID	GENDER	T/N	Tumor_type	Treated	Primary/ Secondary	Hypermutated
EX-01-001	Expression	P-01-P001	P-01-S001	M	Tumor	GBM	Y	Primary	Y
CN-01-002	CopyNumber	P-01-P001	P-01-S001	M	Tumor	GBM	Y	Primary	Y
MU-01-003	Mutation	P-01-P001	P-01-S002	M	Tumor	GBM	Y	Primary	Y
EX-01-004	Expression	P-01-P002	P-01-S003	M	Normal	GBM	Y	Secondary	Y
CN-01-005	CopyNumber	P-01-P002	P-01-S004	M	Tumor	GBM	Y	Secondary	N
EX-01-006	Expression	P-01-P002	P-01-S004	M	Tumor	GBM	Y	Secondary	N
ME-01-007	Methylation	P-01-P002	P-01-S004	M	Tumor	GBM	Y	Secondary	N
EX-01-008	Expression	P-01-P003	P-01-S006	F	Tumor	GBM	N	Primary	Y
EX-01-009	Expression	P-01-P004	P-01-S009	F	Tumor	GBM	N	Primary	Y
EX-01-0010	Expression	P-01-P005	P-01-S0011	M	Control				

Using sample information.

- Sample information can include discrete and continuous.
- Can be used to “sort” and “filter” tracks.
- Can split tracks across panels by “group”



Grouped by unique ID

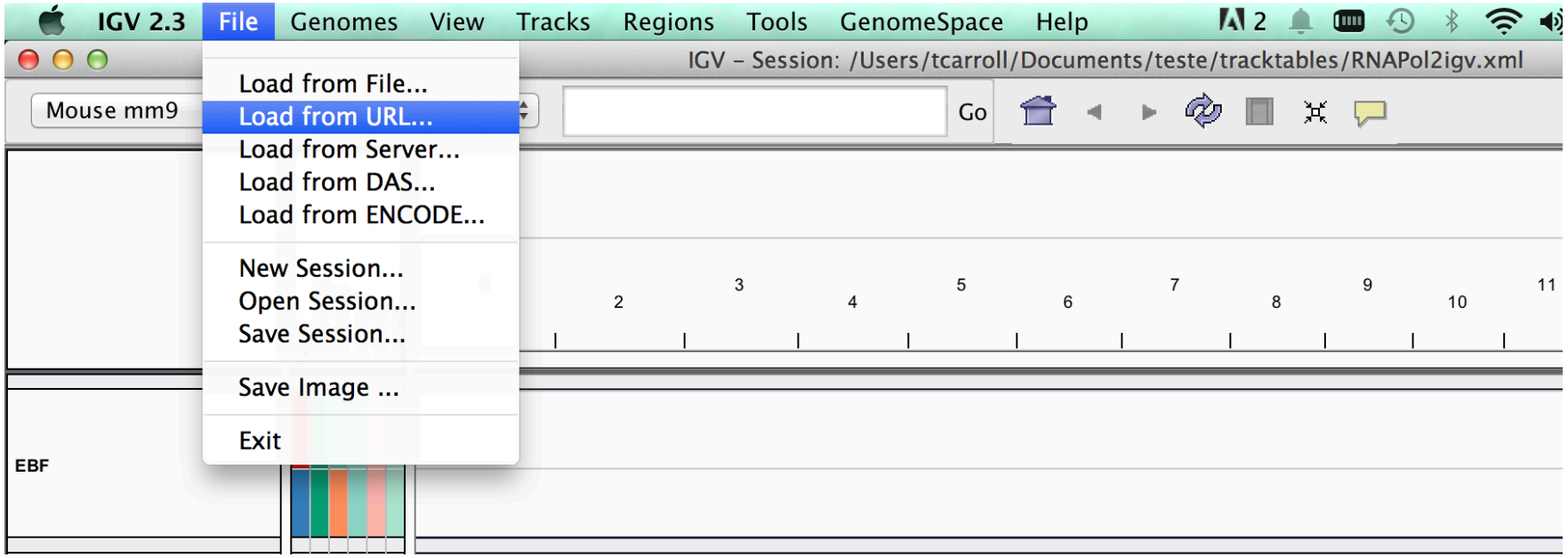
Grouped related files by name

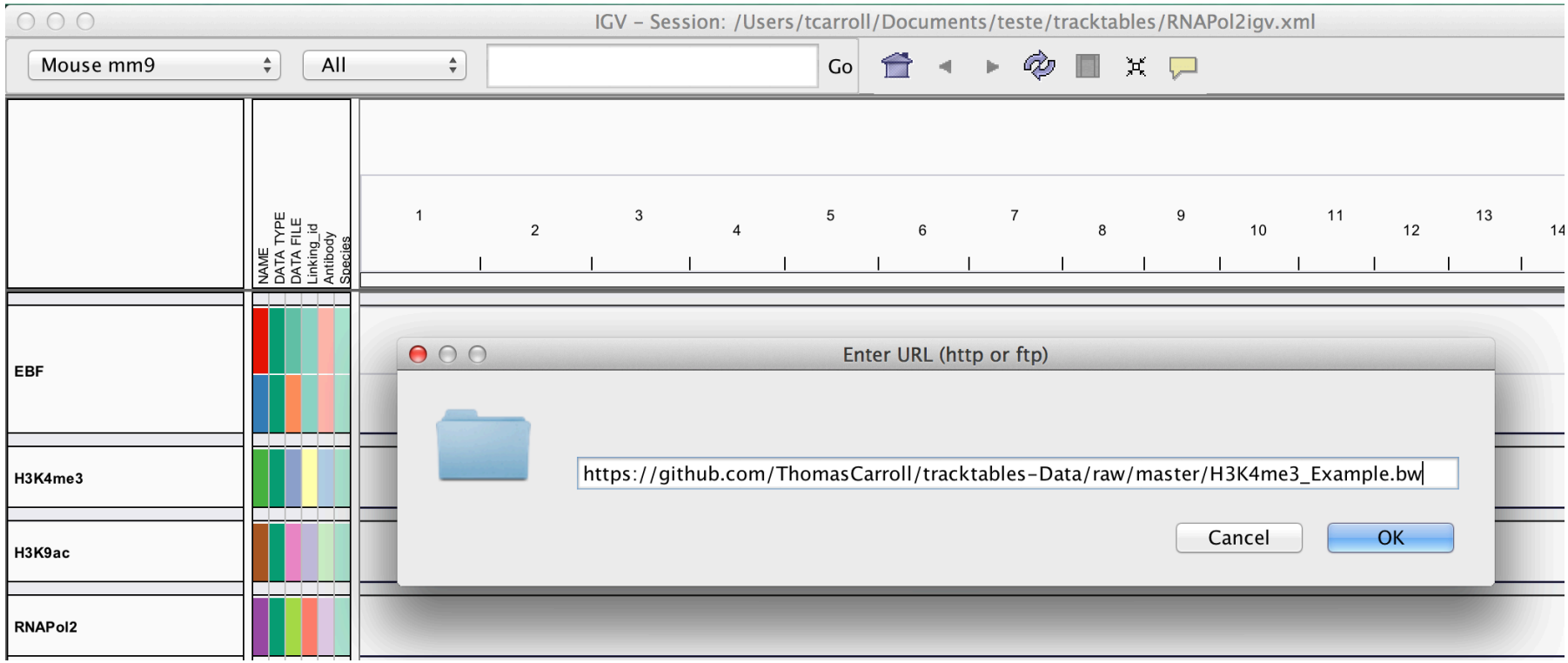
Loading external data and annotation

Load data from a URL.

- As with UCSC, IGV supports data hosted on external servers.

- Data accessible from a URL such as HTTP and FTP can be loaded using the **“Load from URL”.**

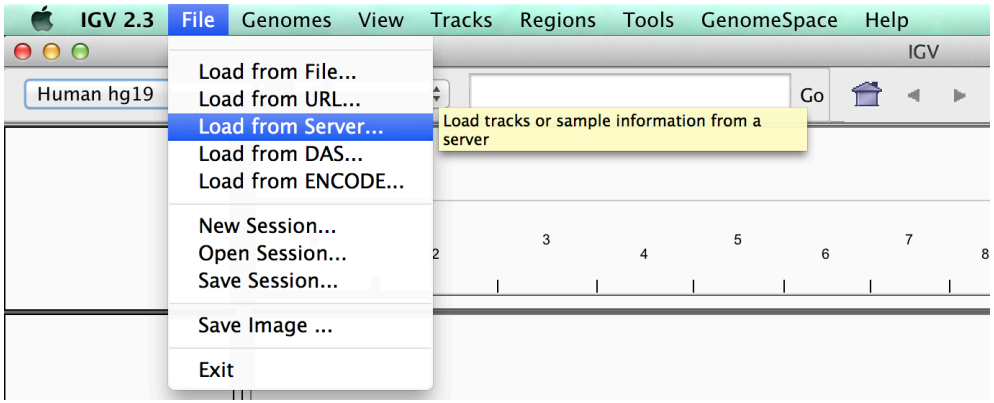

IGV 2.3
File
Genomes
View
Tracks
Regions
Tools
GenomeSpace
Help
IGV – Session: /Users/tcarroll/Documents/teste/tracktables/RNAPol2igv.xml
Load from File...
Load from URL...
Load from Server...
Load from DAS...
Load from ENCODE...
New Session...
Open Session...
Save Session...
Save Image ...
Exit
Mouse mm9
Go
EBF

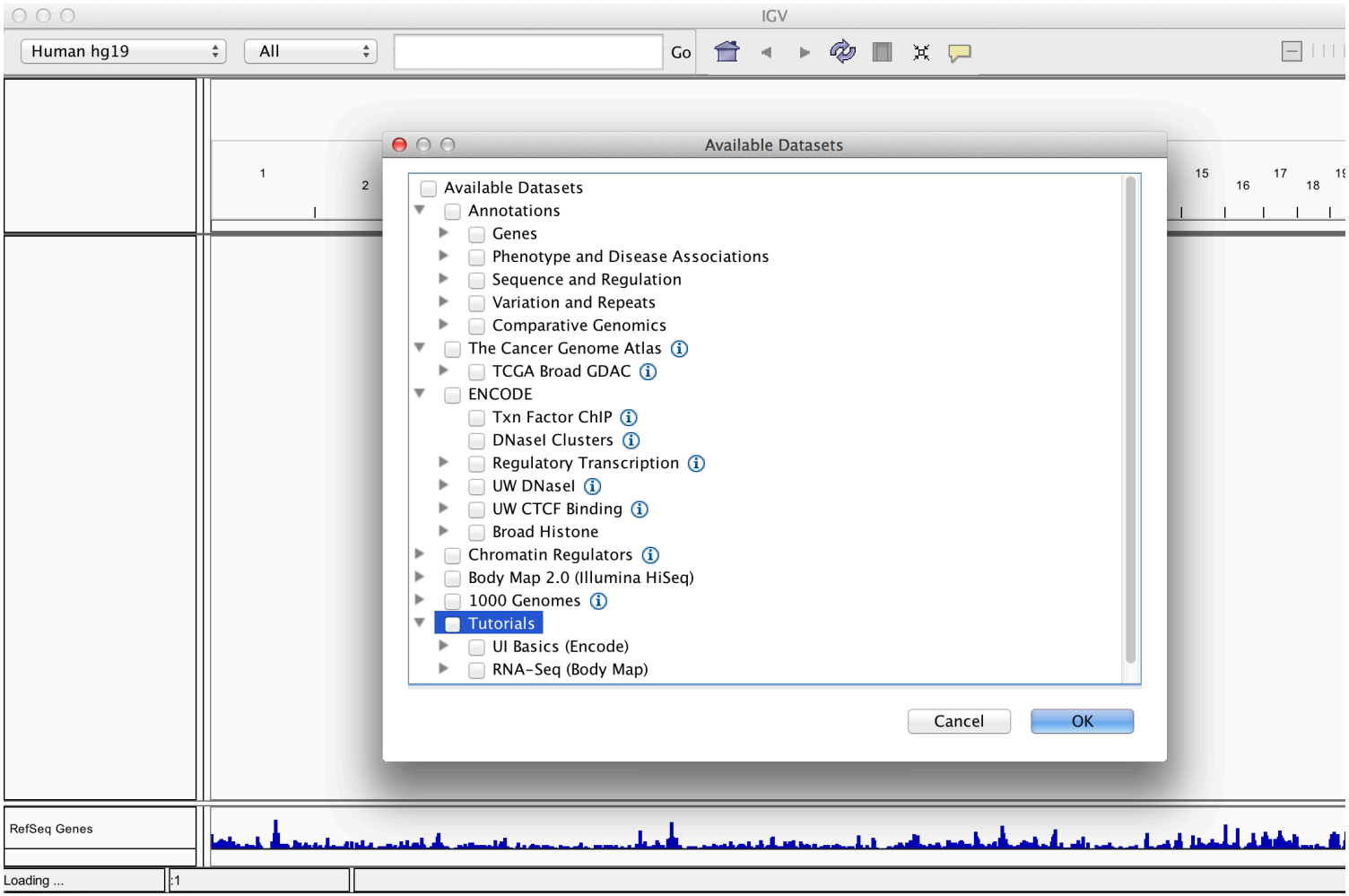

IGV – Session: /Users/tcarroll/Documents/teste/tracktables/RNAPol2igv.xml
Mouse mm9
All
Go
NAME
DATA TYPE
DATA FILE
Linking_id
Antibody
Species
EBF
H3K4me3
H3K9ac
RNAPol2
Enter URL (http or ftp)
https://github.com/ThomasCarroll/tracktables-Data/raw/master/H3K4me3_Example.bw
Cancel
OK

Loading external data and annotation

Load data from a server.

- Unlike UCSC, IGV comes with few external tracks.
- External tracks (relevant to the genome) can be loaded from the IGV server or Encode-IGV server.


IGV 2.3
File
Genomes
View
Tracks
Regions
Tools
GenomeSpace
Help
IGV
Human hg19
Go
Load from File...
Load from URL...
Load from Server...
Load from DAS...
Load from ENCODE...
New Session...
Open Session...
Save Session...
Save Image ...
Exit
Load tracks or sample information from a server


IGV
Human hg19
All
Go
Available Datasets
Available Datasets
Annotations
Genes
Phenotype and Disease Associations
Sequence and Regulation
Variation and Repeats
Comparative Genomics
The Cancer Genome Atlas
TCGA Broad GDAC
ENCODE
Txn Factor ChIP
DNaseI Clusters
Regulatory Transcription
UW DNaseI
UW CTCF Binding
Broad Histone
Chromatin Regulators
Body Map 2.0 (Illumina HiSeq)
1000 Genomes
Tutorials
UI Basics (Encode)
RNA-Seq (Body Map)
Cancel
OK
RefSeq Genes
Loading ...

Viewing data

- IGV associates common file formats with default display methods.

- Most of the time IGV will make a sensible choice how we wish to display data.

Accepted formats and default display.

Information on accepted file formats and default display can be found at
http://www.broadinstitute.org/software/igv/RecommendedFileFormats

Recommended File Formats

The following table lists recommended file formats for data types. **Tip**: Using igvtools to convert a source data file to a binary TDF file can reduce loading time and improve performance.

Source Data	Recommended File Formats
ChIP-Seq, RNA-Seq	TDF format. Use the igvtools package (count command) to generate a binary read count density file in TDF format. Load the resulting TDF file into IGV.
Copy number	CN format, SNP format **Log values**: If copy number data contains negative values, IGV assumes they are scaled log values and displays them without modification. If copy number data contains all positive values, IGV assumes they are unscaled. It centers the unscaled values around 2 (1 for allele specific files), log transforms them (logValue = log2(copyNumber / 2)), and displays the log values.
Gene expression data	GCT format, RES format

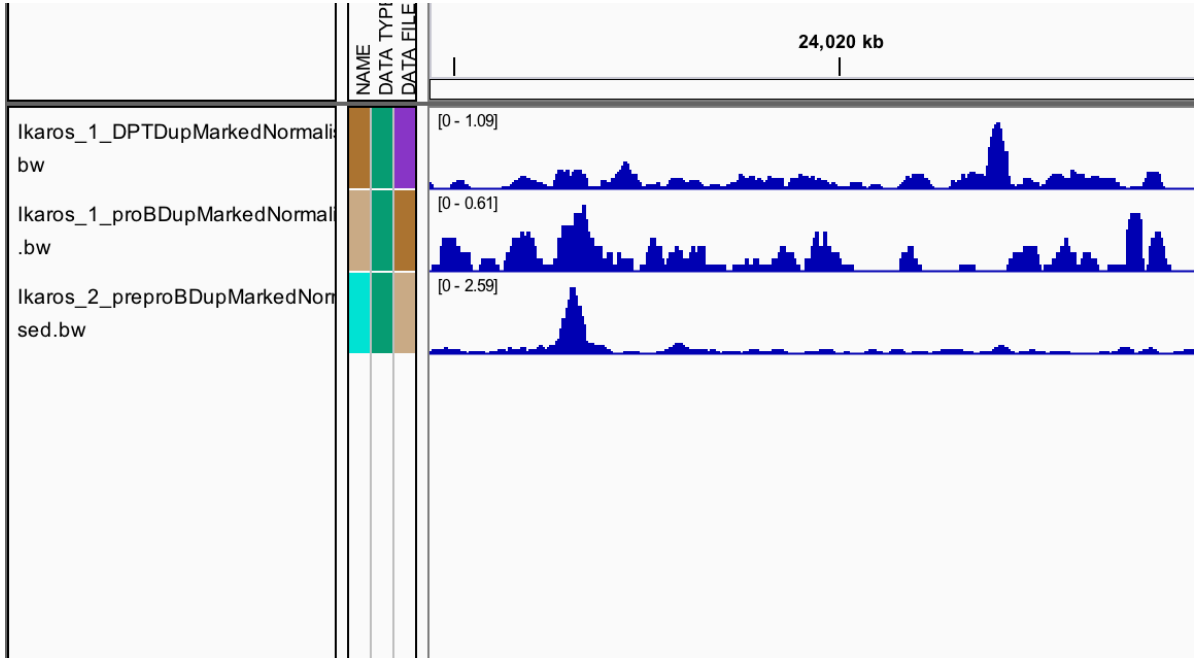
Bed/bigBed

- Basic
  - Tab-delimited

    Chrom,Start,End
- Bed6
- bigBed (recommended)

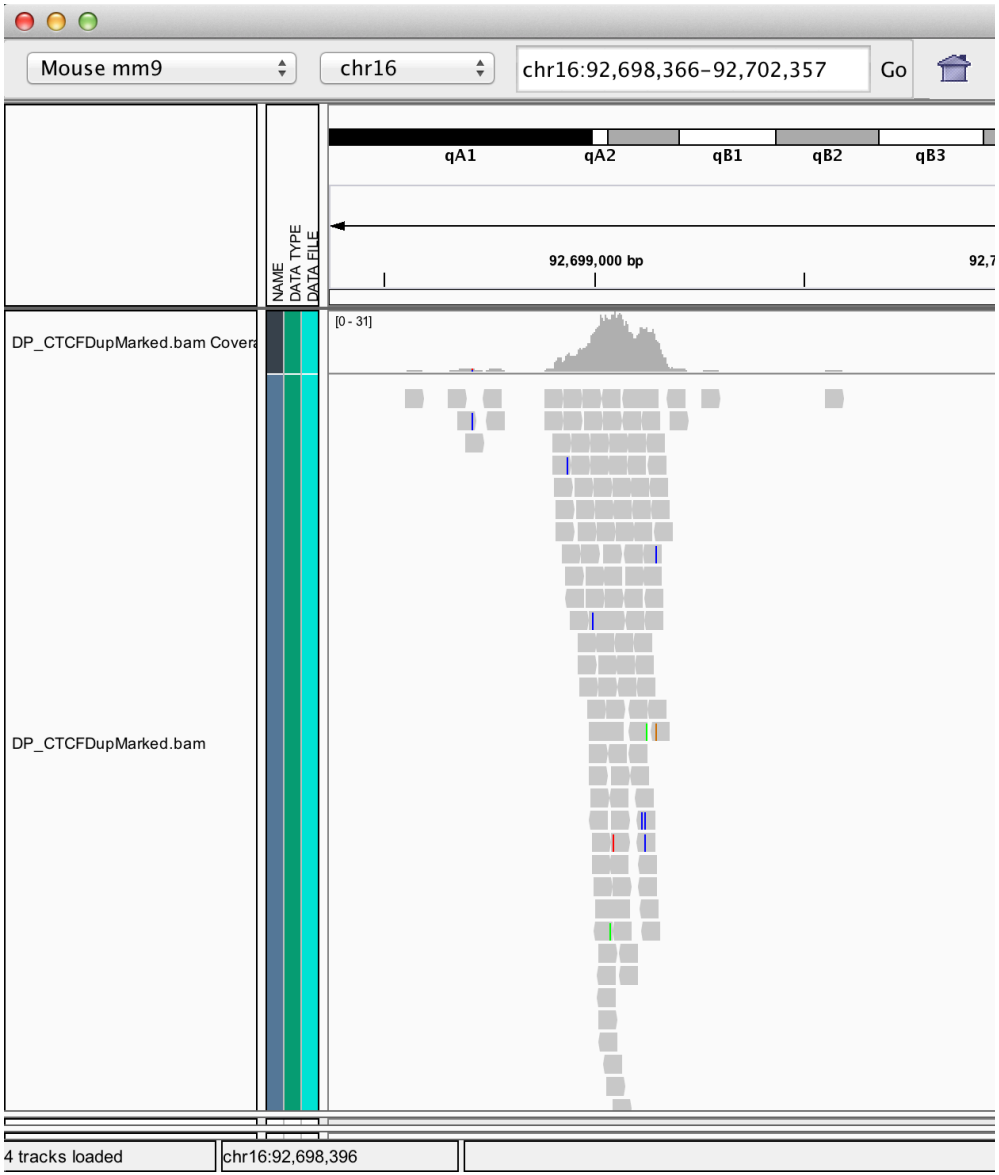
Wig, BedGraph and BigWig

- Wig/bedGraph require high memory load
- Recommended format is bigWig



BAM alignment files

- BAM files contain alignment information.
- Require an accompanying .bai index file for display

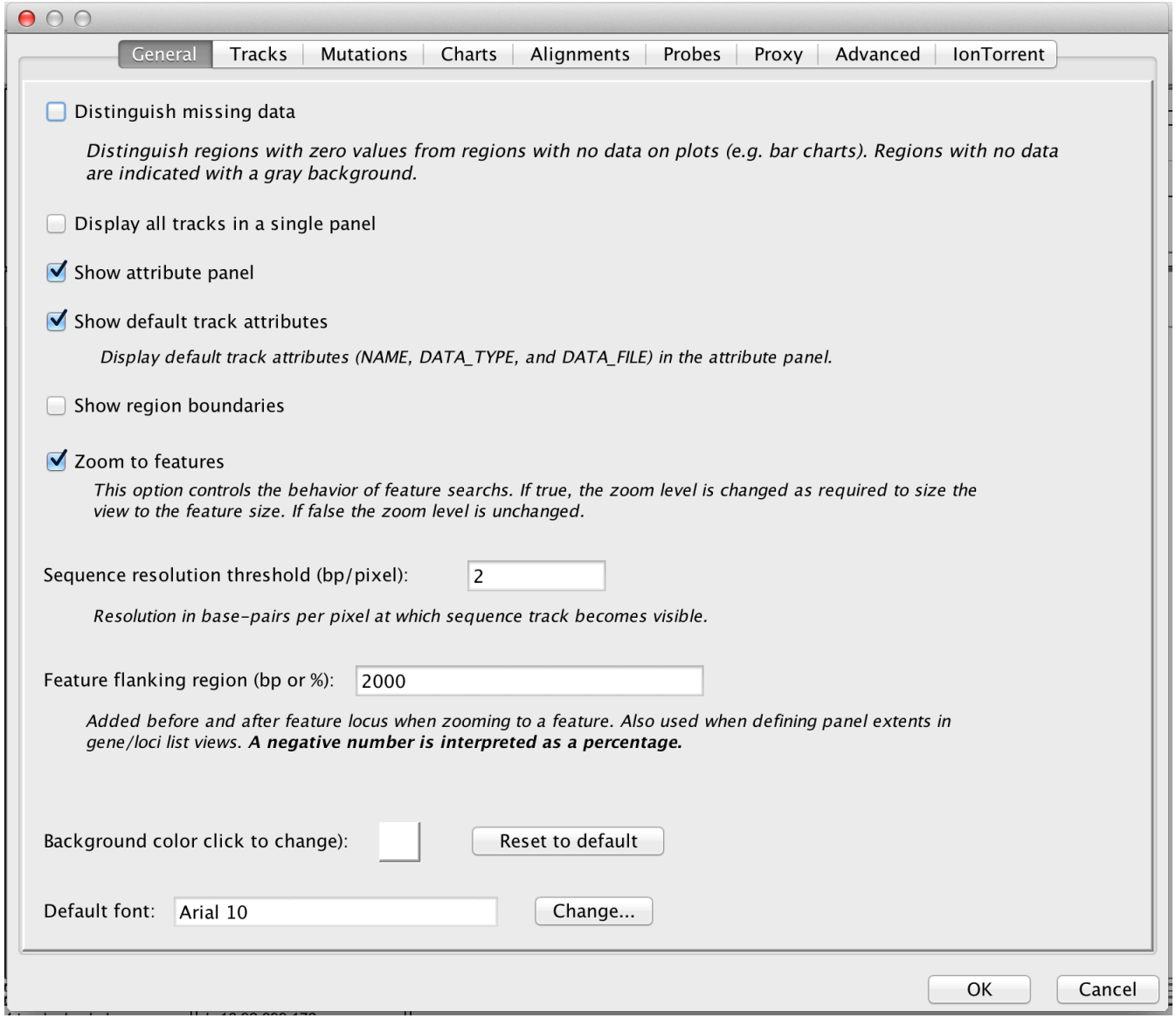


Finer control of display

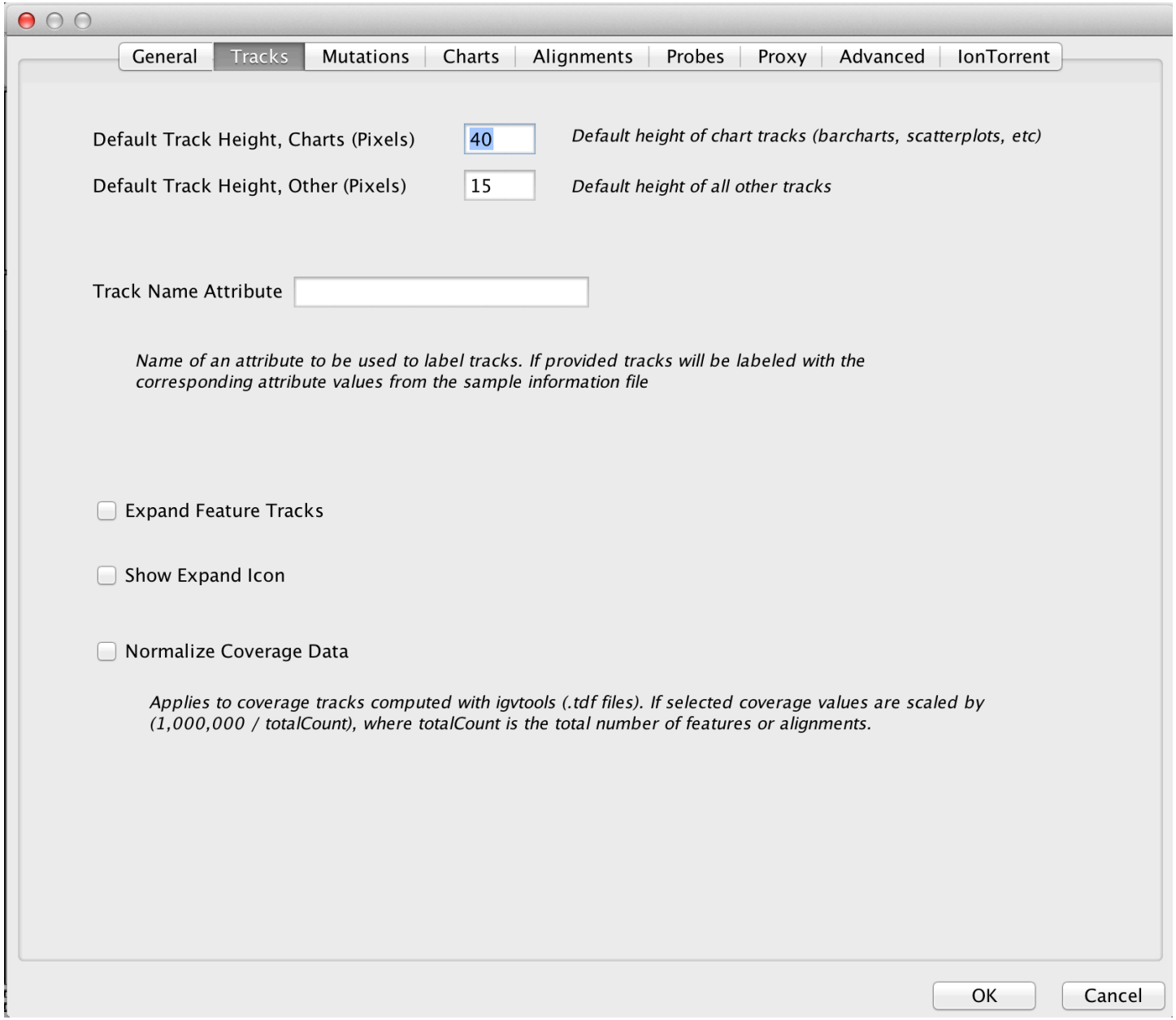
- IGV allows for customization of track display.
  - Menu bar -> View -> Preferences
  - Select track (right click)

Display preferences

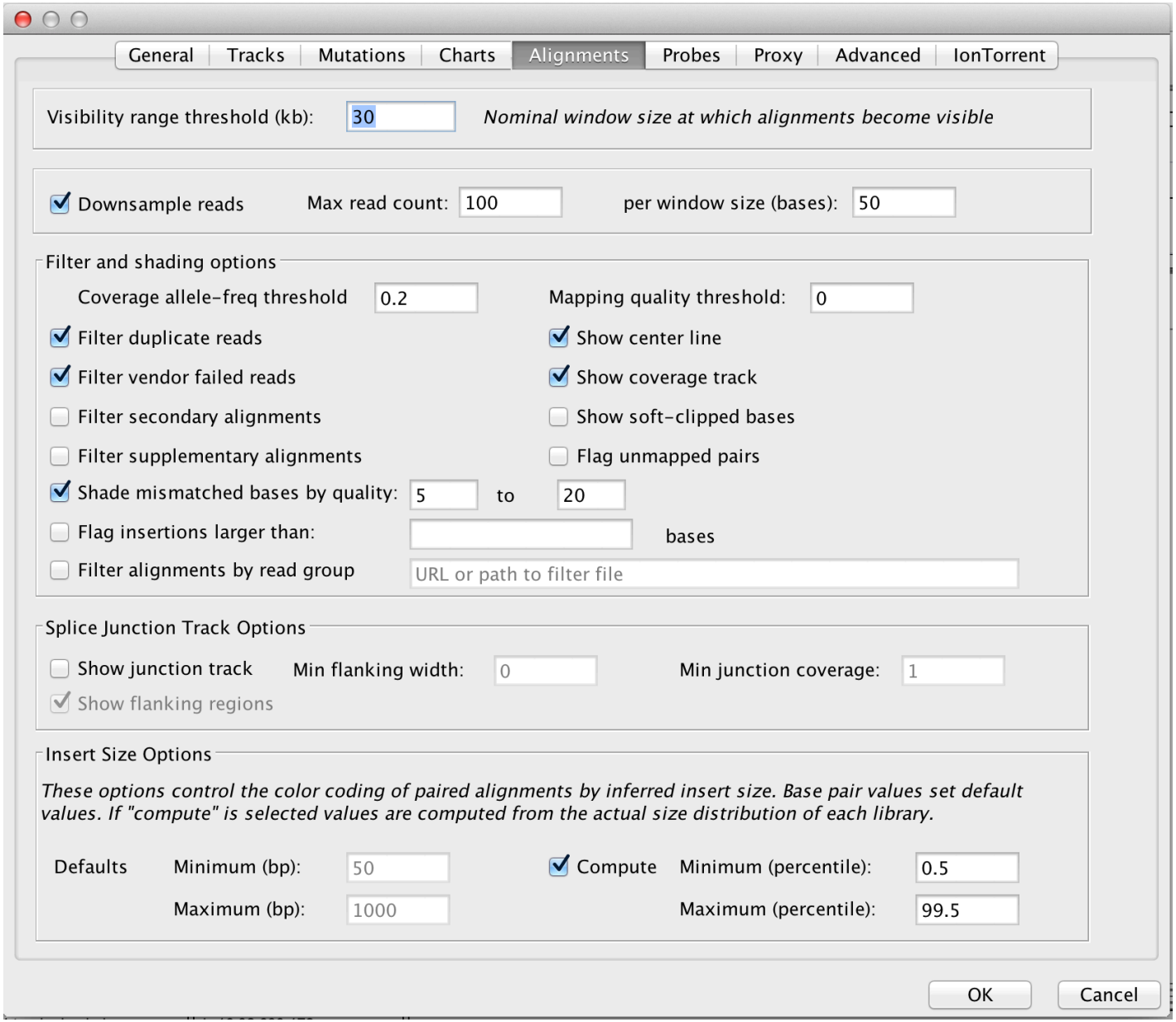
General



Tracks

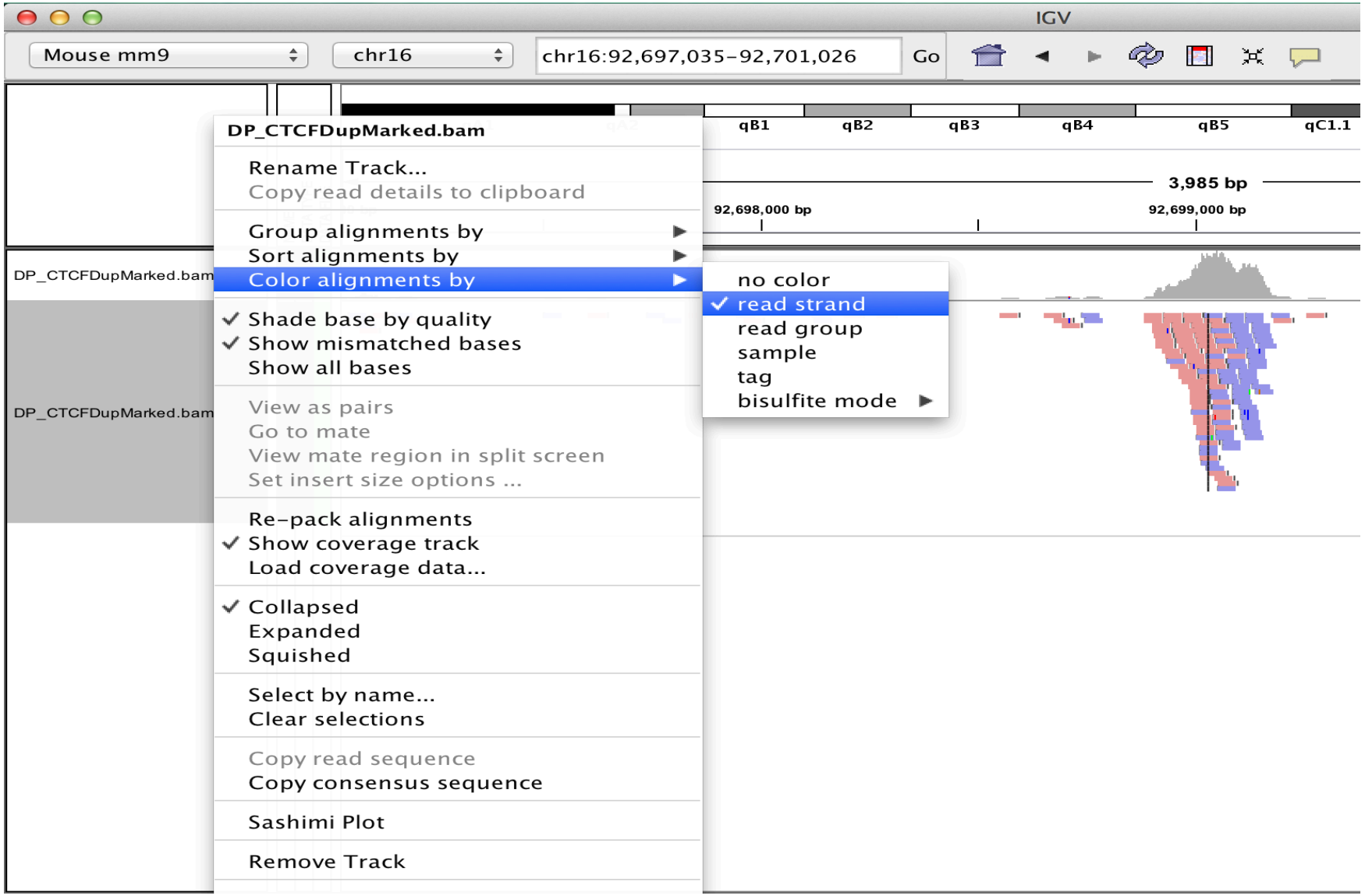

General
Tracks
Mutations
Charts
Alignments
Probes
Proxy
Advanced
IonTorrent
Default Track Height, Charts (Pixels)
40
Default height of chart tracks (barcharts, scatterplots, etc)
Default Track Height, Other (Pixels)
15
Default height of all other tracks
Track Name Attribute
Name of an attribute to be used to label tracks. If provided tracks will be labeled with the corresponding attribute values from the sample information file
Expand Feature Tracks
Show Expand Icon
Normalize Coverage Data
Applies to coverage tracks computed with igvtools (.tdf files). If selected coverage values are scaled by (1,000,000 / totalCount), where totalCount is the total number of features or alignments.
OK
Cancel

Alignments


General
Tracks
Mutations
Charts
Alignments
Probes
Proxy
Advanced
IonTorrent
Visibility range threshold (kb): 30
Nominal window size at which alignments become visible
Downsample reads
Max read count: 100
per window size (bases): 50
Filter and shading options
Coverage allele-freq threshold 0.2
Mapping quality threshold: 0
Filter duplicate reads
Show center line
Filter vendor failed reads
Show coverage track
Filter secondary alignments
Show soft-clipped bases
Filter supplementary alignments
Flag unmapped pairs
Shade mismatched bases by quality: 5 to 20
Flag insertions larger than:
bases
Filter alignments by read group
URL or path to filter file
Splice Junction Track Options
Show junction track
Min flanking width: 0
Min junction coverage: 1
Show flanking regions
Insert Size Options
These options control the color coding of paired alignments by inferred insert size. Base pair values set default values. If "compute" is selected values are computed from the actual size distribution of each library.
Defaults
Minimum (bp): 50
Compute
Minimum (percentile): 0.5
Maximum (bp): 1000
Maximum (percentile): 99.5
OK
Cancel

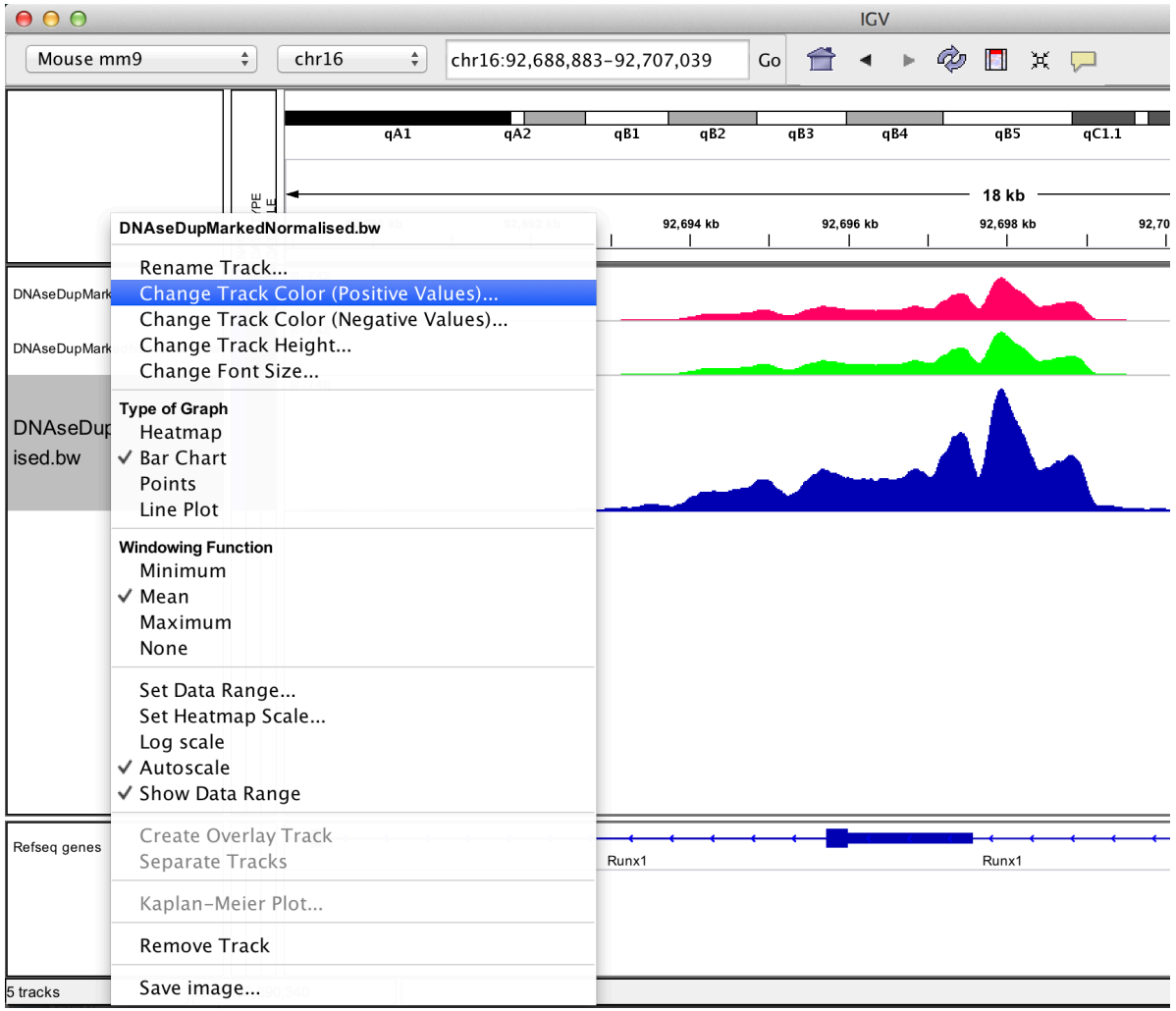
Track display options

- Read Packing, grouping, sorting, colouring options.

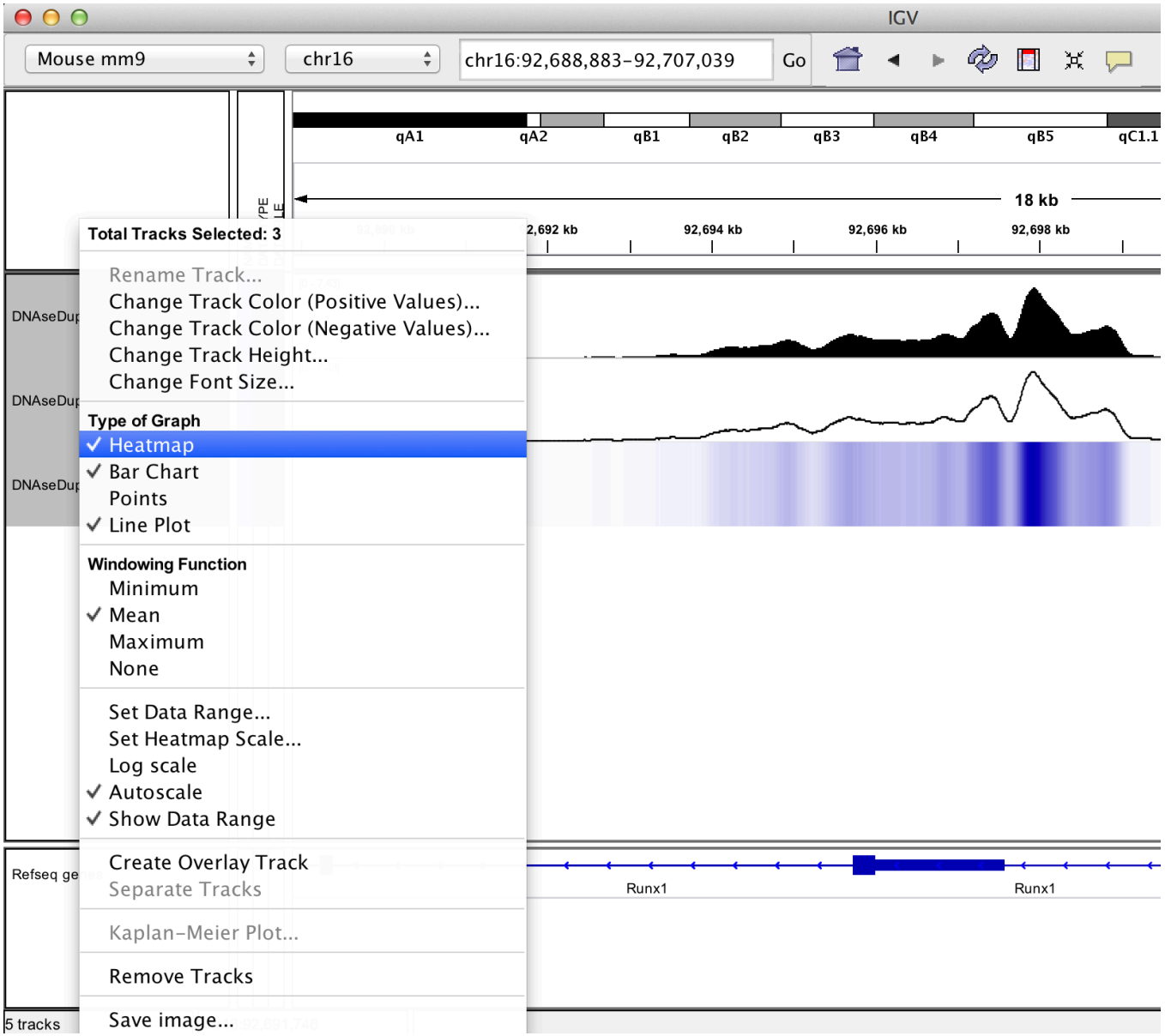


Track display options
Graph/interval files

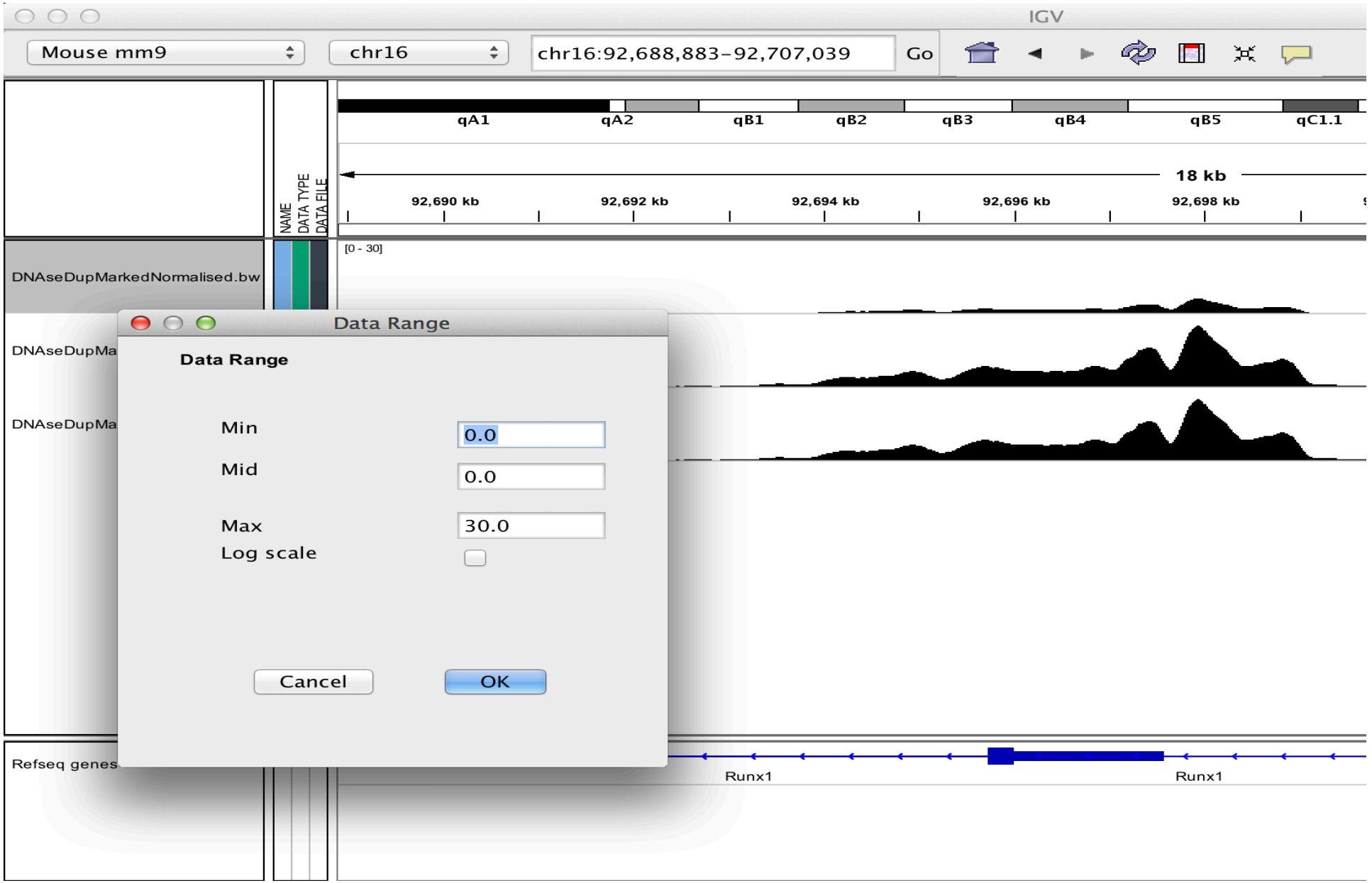
- Track colour/appearance



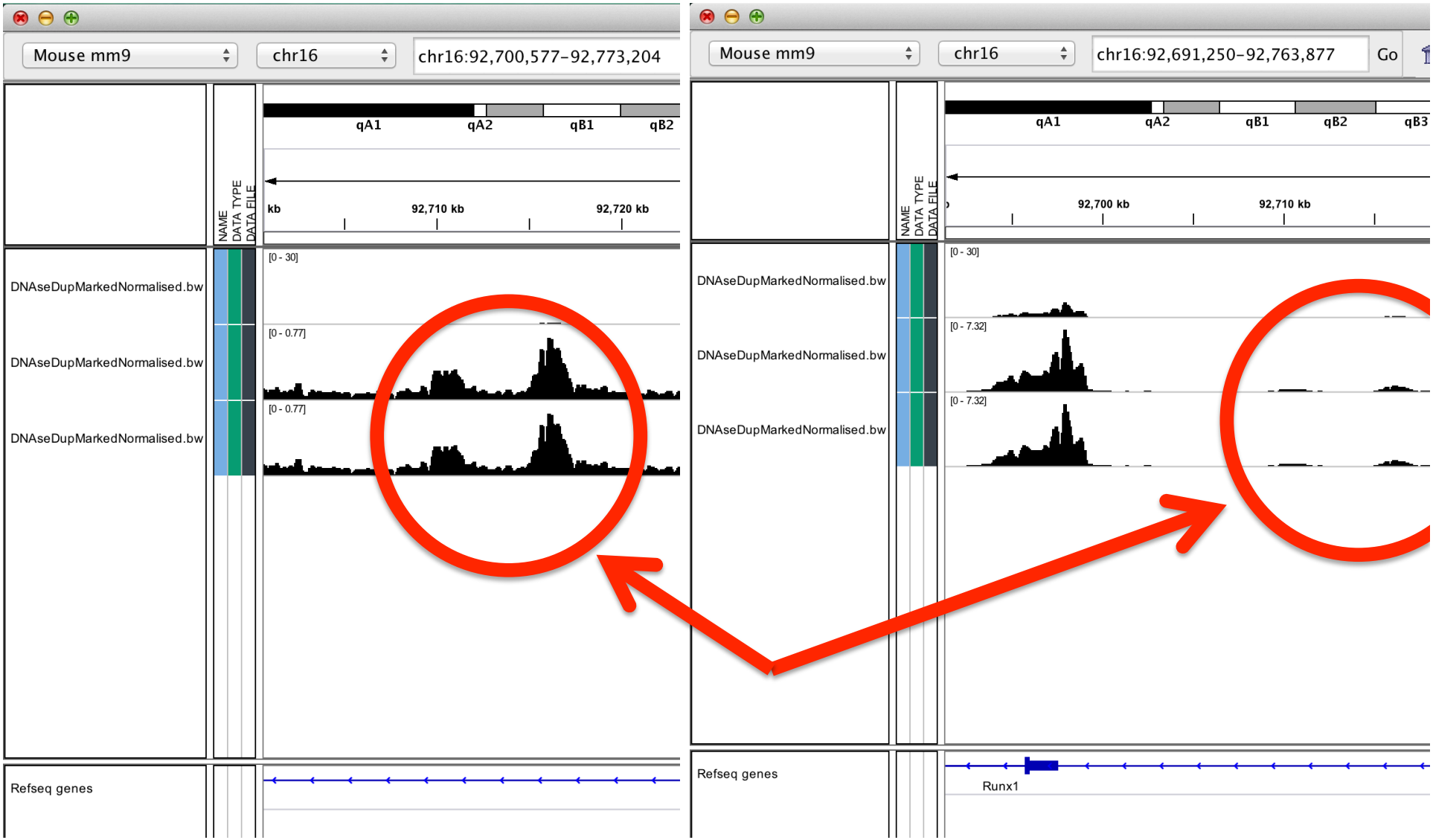
- Graph type


IGV
Mouse mm9
chr16
chr16:92,688,883-92,707,039
Go
qA1
qA2
qB1
qB2
qB3
qB4
qB5
qC1.1
18 kb
92,692 kb
92,694 kb
92,696 kb
92,698 kb
DNAseDup
DNAseDup
DNAseDup
Refseq ge
Runx1
Runx1
5 tracks
Total Tracks Selected: 3
Rename Track...
Change Track Color (Positive Values)...
Change Track Color (Negative Values)...
Change Track Height...
Change Font Size...
Type of Graph
Heatmap
Bar Chart
Points
Line Plot
Windowing Function
Minimum
Mean
Maximum
None
Set Data Range...
Set Heatmap Scale...
Log scale
Autoscale
Show Data Range
Create Overlay Track
Separate Tracks
Kaplan-Meier Plot...
Remove Tracks
Save image...

Data Scaling.

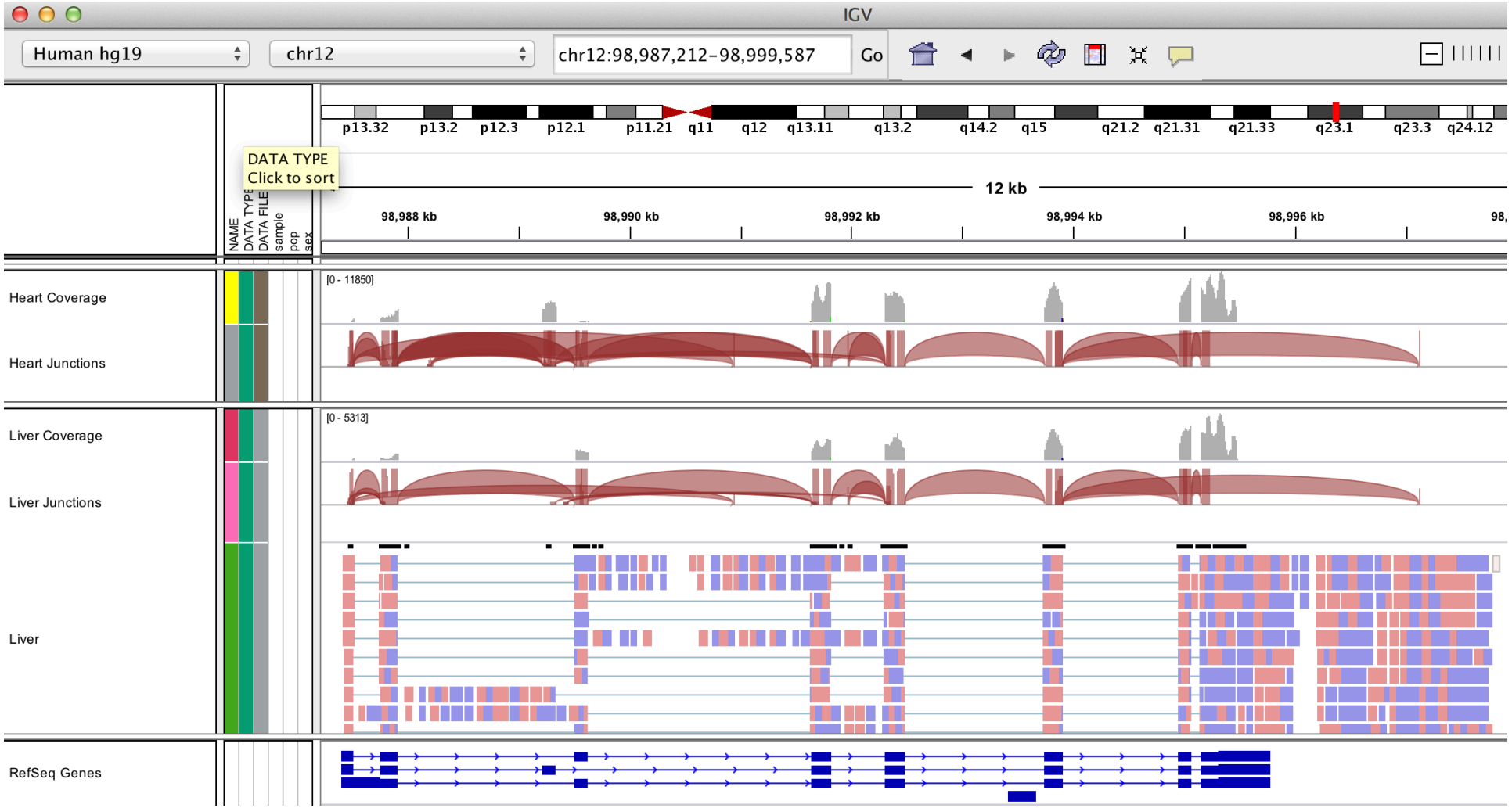


- Autoscaling adjusts to track's visible signal maximum



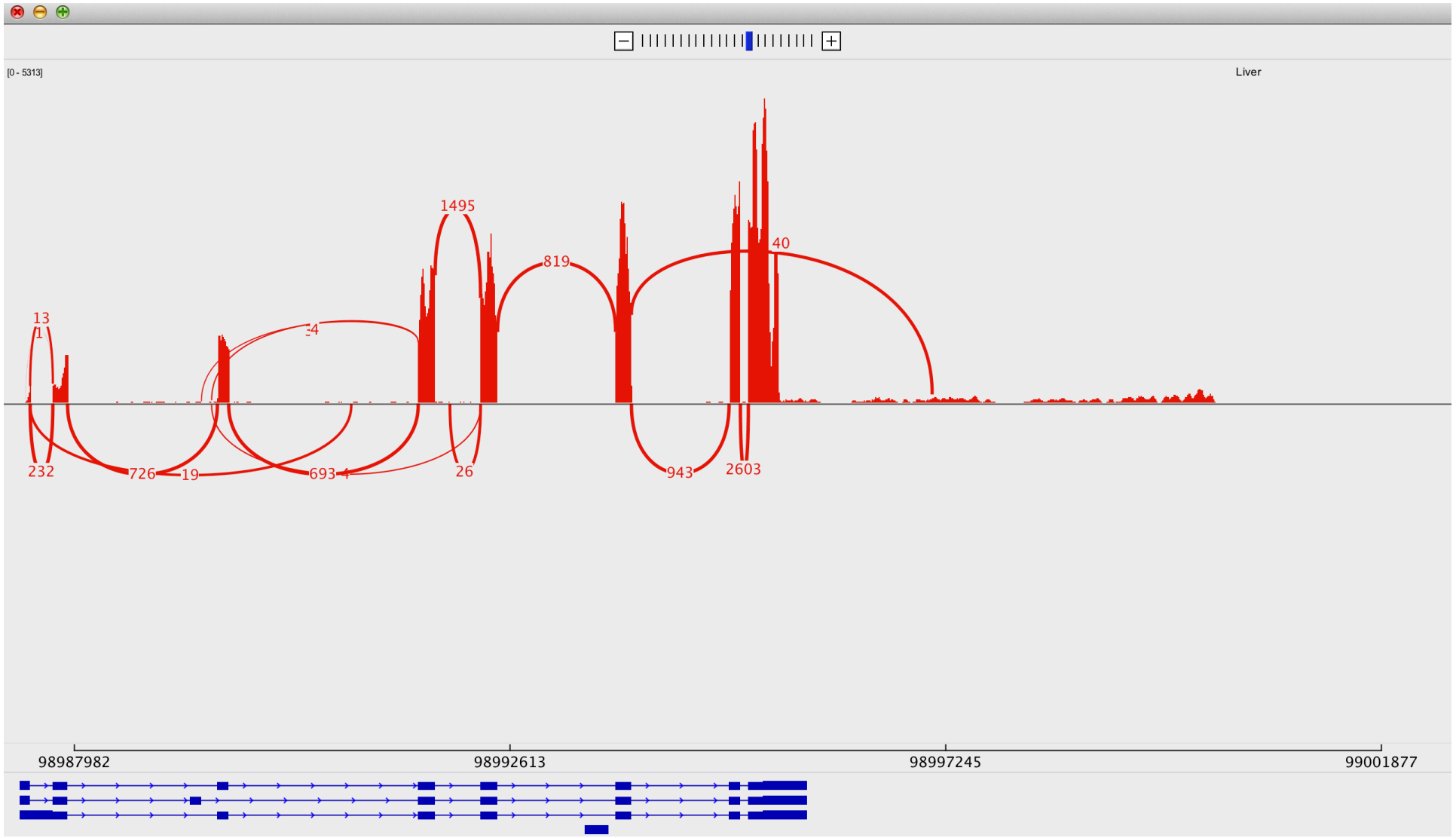
Some cool features.

IGV can display splicing information

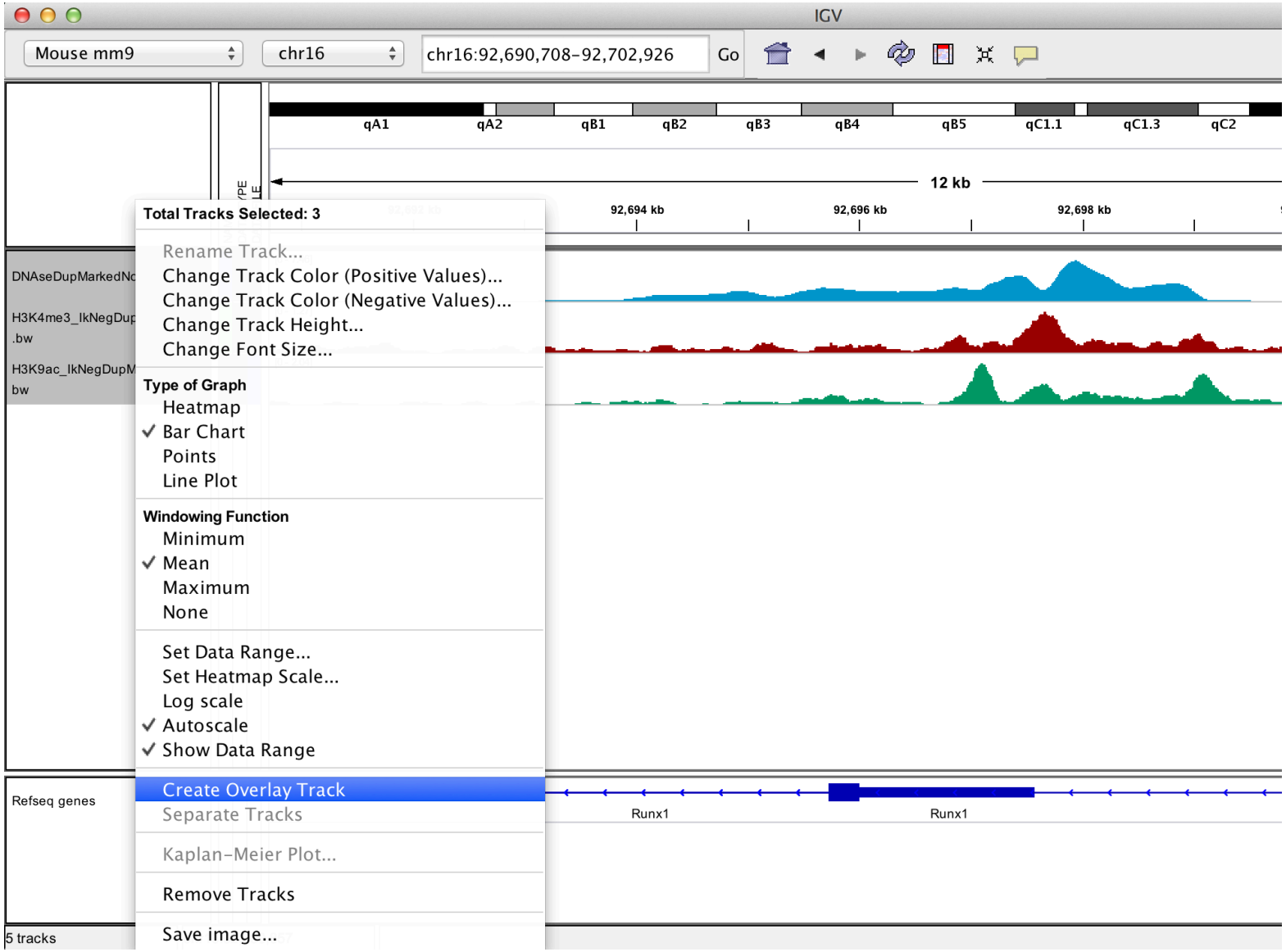


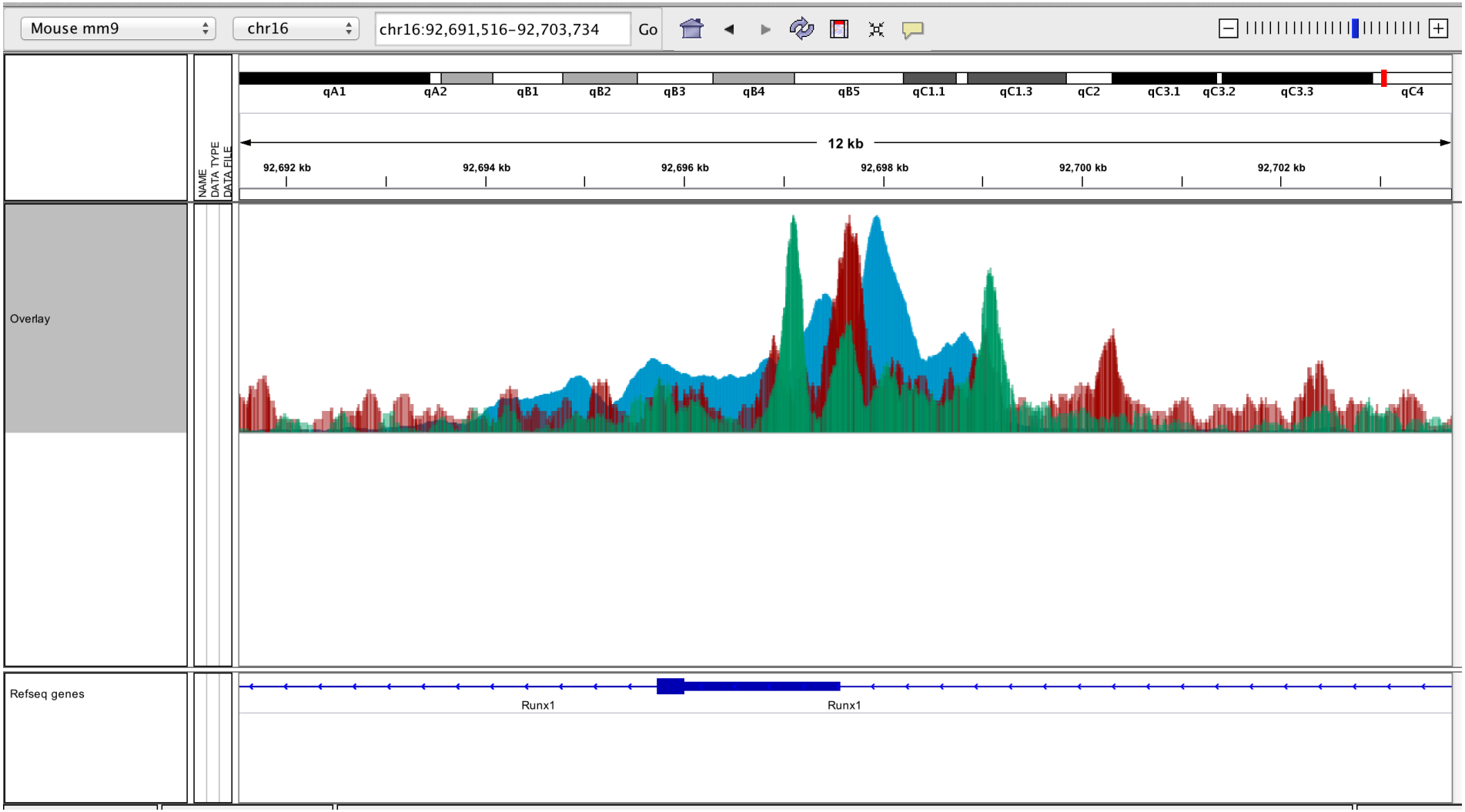
..and Sashimi plots

(http://www.broadinstitute.org/igv/Sashimi)



IGV can overlay tracks too.

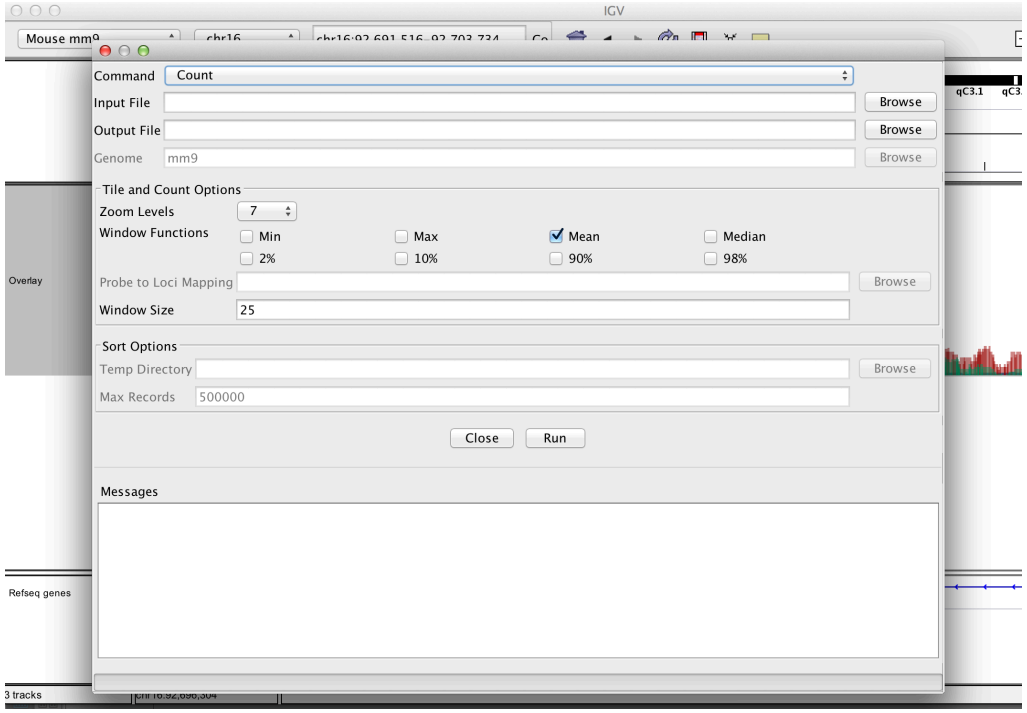



Mouse mm9
chr16
chr16:92,691,516-92,703,734
Go
qA1
qA2
qB1
qB2
qB3
qB4
qB5
qC1.1
qC1.3
qC2
qC3.1
qC3.2
qC3.3
qC4
12 kb
92,692 kb
92,694 kb
92,696 kb
92,698 kb
92,700 kb
92,702 kb
NAME
DATA TYPE
DATA FILE
Overlay
Refseq genes
Runx1
Runx1

IGV includes a simple toolset

IGVTools

- IGVTools can be used to post-process genomics data.
- Includes indexing, sorting and genome graph creation.



Where to get help?

- http://www.broadinstitute.org/igv/UserGuide
- https://groups.google.com/forum/#!forum/igv-help