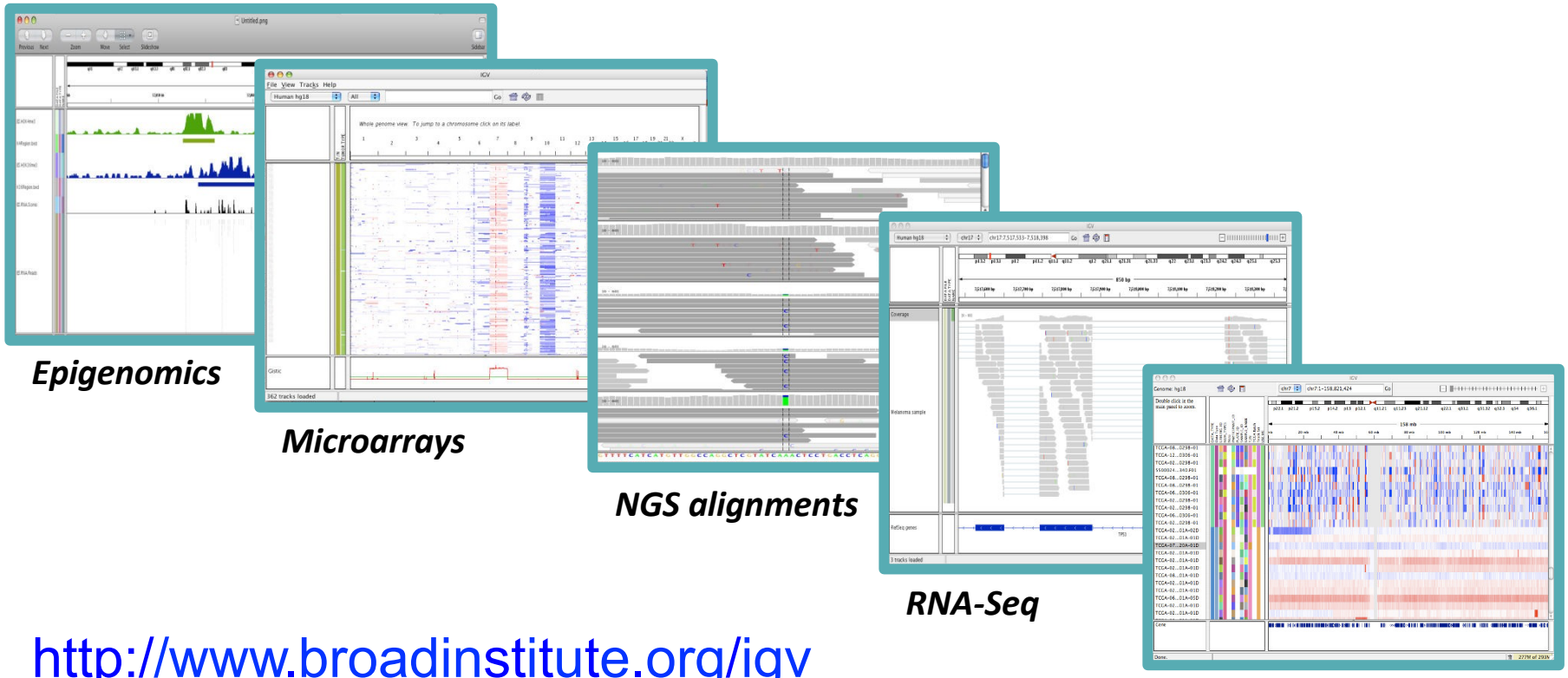
# Introduction to
# NGS Visualization with the
# Integrative Genomics Viewer (IGV)

BROAD INSTITUTE

# Integrative Genomics Viewer (IGV)

*Desktop application for the interactive visual exploration of integrated genomic datasets*



***Epigenomics***

***Microarrays***

***NGS alignments***

***RNA-Seq***

***mRNA, CNV, Seq***

http://www.broadinstitute.org/igv

65,000 registrations

# Features

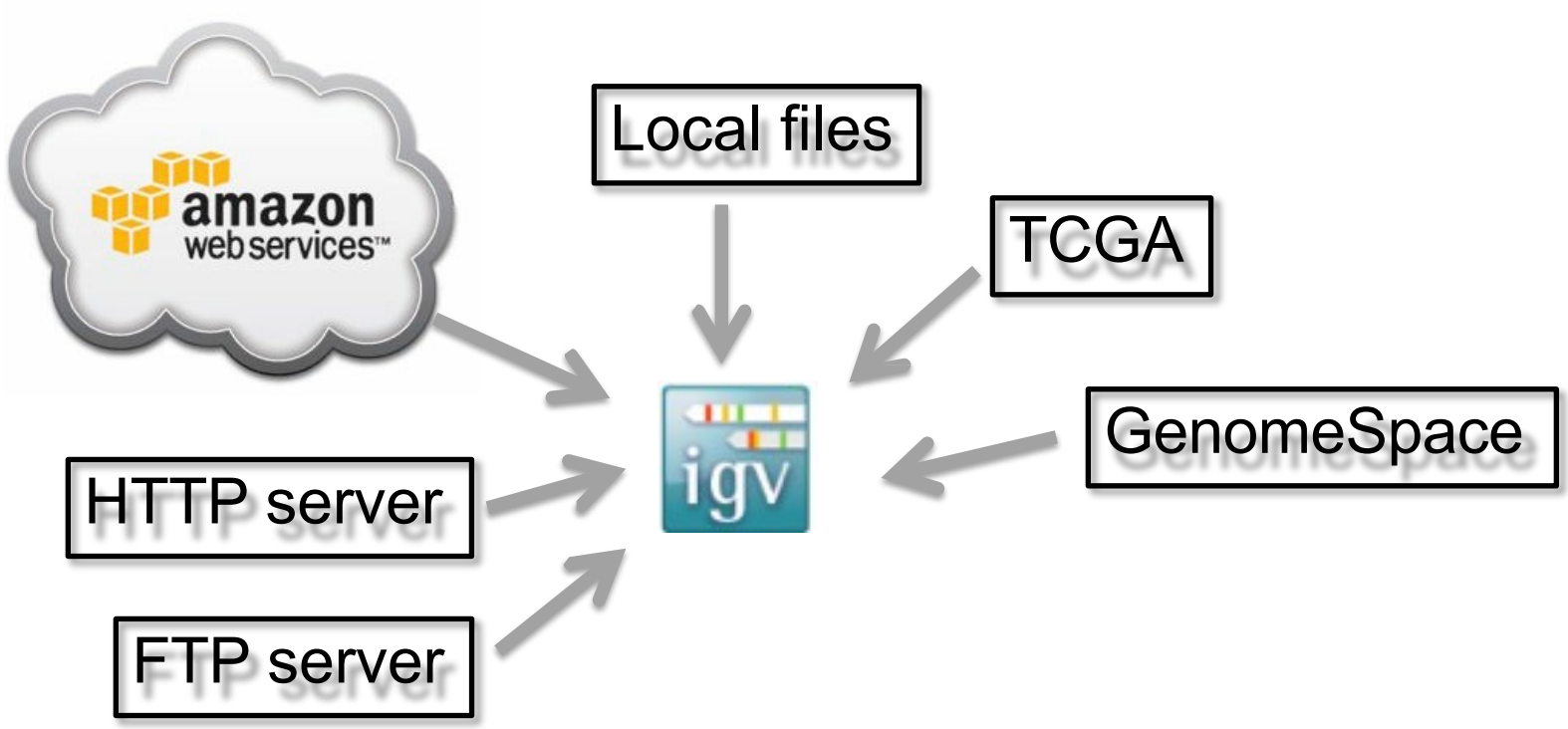
Integrative Genomics Viewer

With IGV you can…

- Explore large genomic datasets with an intuitive, easy-to-use interface.
- Integrate multiple data types with clinical and other sample information.
- View data from multiple sources:
  - local, remote, and “cloud-based”.

BROAD INSTITUTE

# IGV data sources




- View **local** files without uploading.
- View **remote** files without downloading the whole dataset.

# Using IGV: The Basics

BROAD INSTITUTE

# Using IGV: the basics

Hands-on exercise

- Launch IGV
- Select a reference genome
- Load data
- Navigate through the data

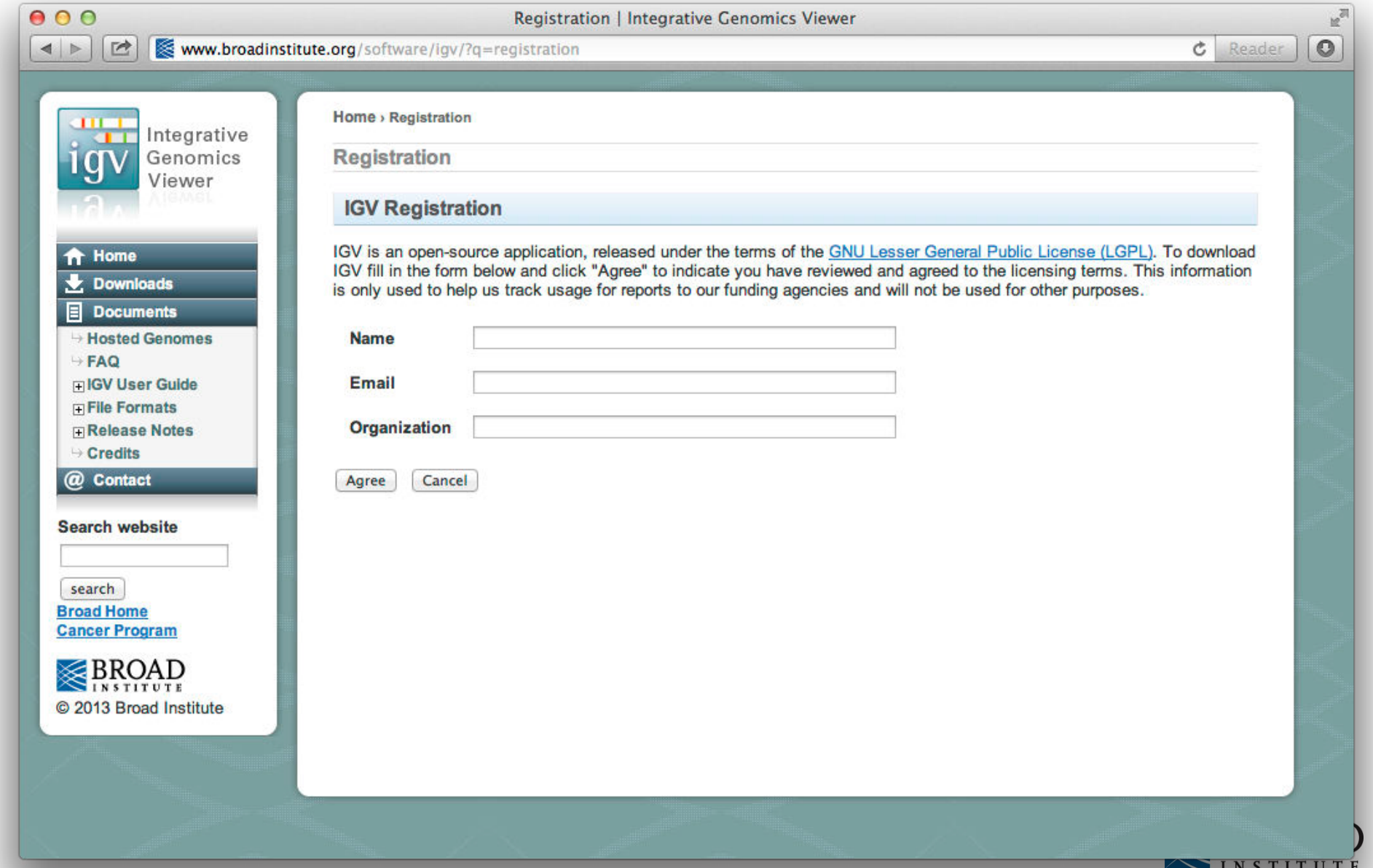
# Launch IGV

Integrative Genomics Viewer

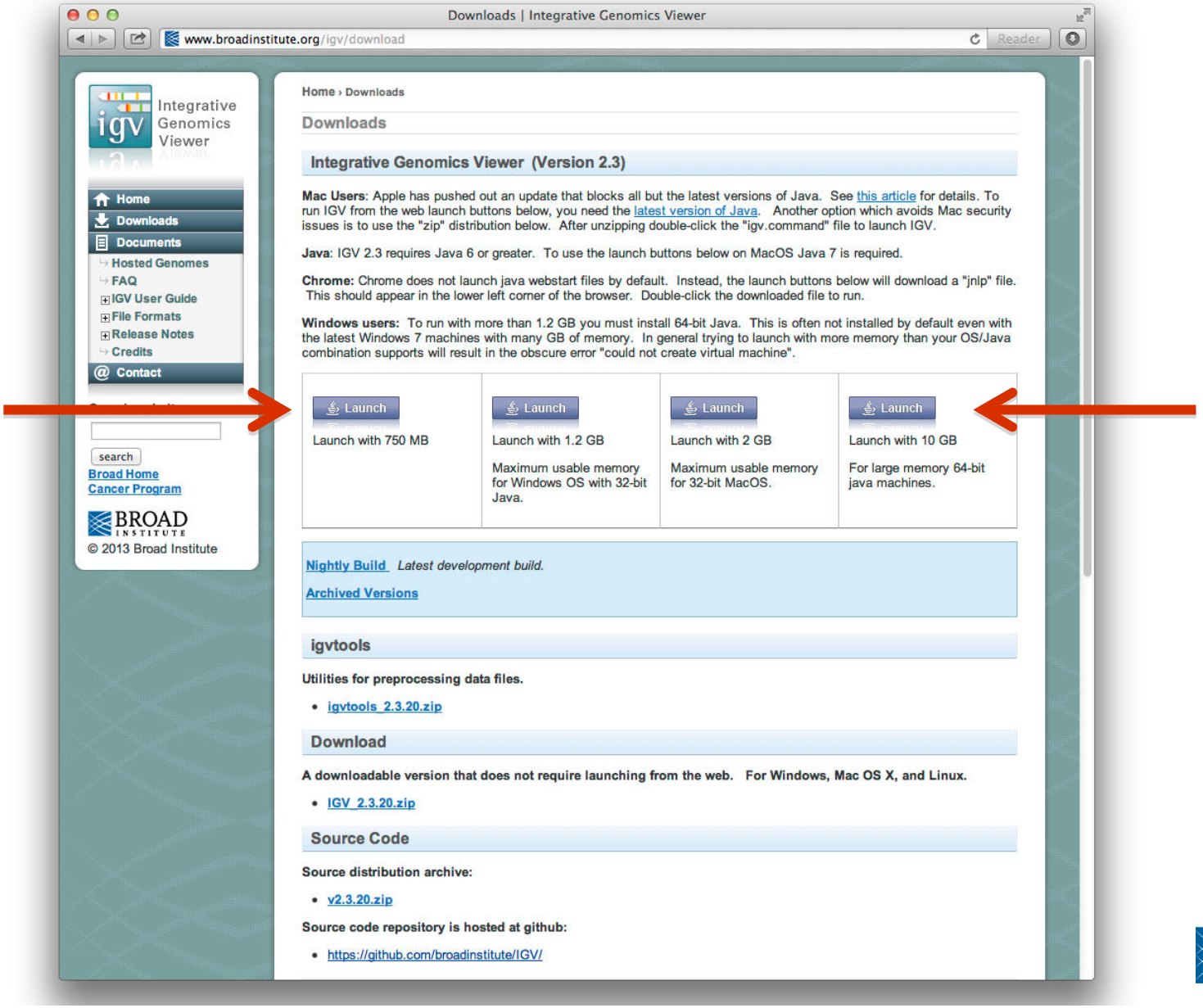
## http://www.broadinstitute.org/igv


BROAD INSTITUTE

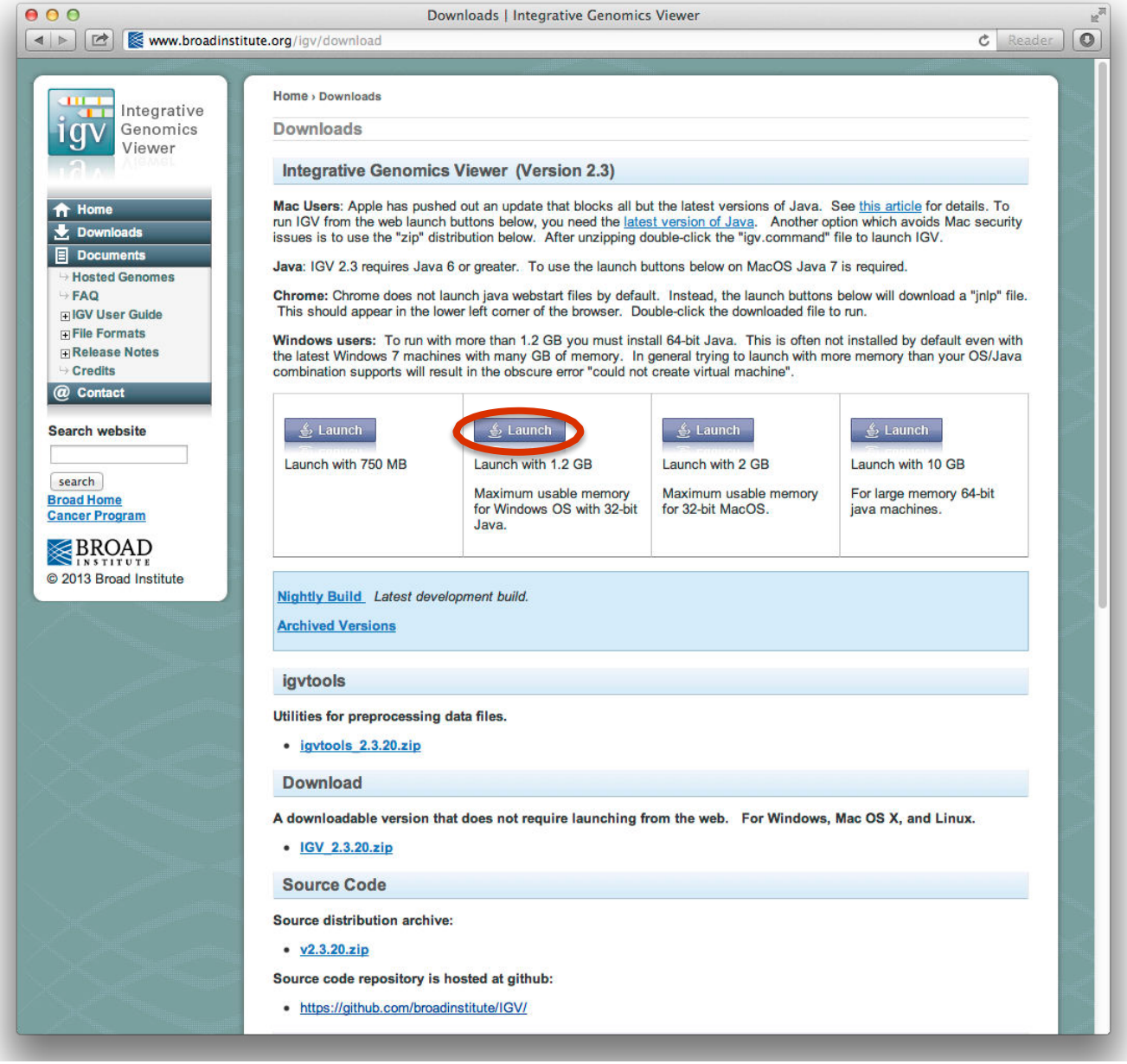
# Launch IGV


Registration | Integrative Genomics Viewer
www.broadinstitute.org/software/igv/?q=registration
Reader
Integrative Genomics Viewer
Home
Downloads
Documents
Hosted Genomes
FAQ
IGV User Guide
File Formats
Release Notes
Credits
Contact
Search website
search
Broad Home
Cancer Program
BROAD INSTITUTE
© 2013 Broad Institute
Home › Registration
Registration
IGV Registration
IGV is an open-source application, released under the terms of the GNU Lesser General Public License (LGPL). To download IGV fill in the form below and click "Agree" to indicate you have reviewed and agreed to the licensing terms. This information is only used to help us track usage for reports to our funding agencies and will not be used for other purposes.
Name
Email
Organization
Agree
Cancel

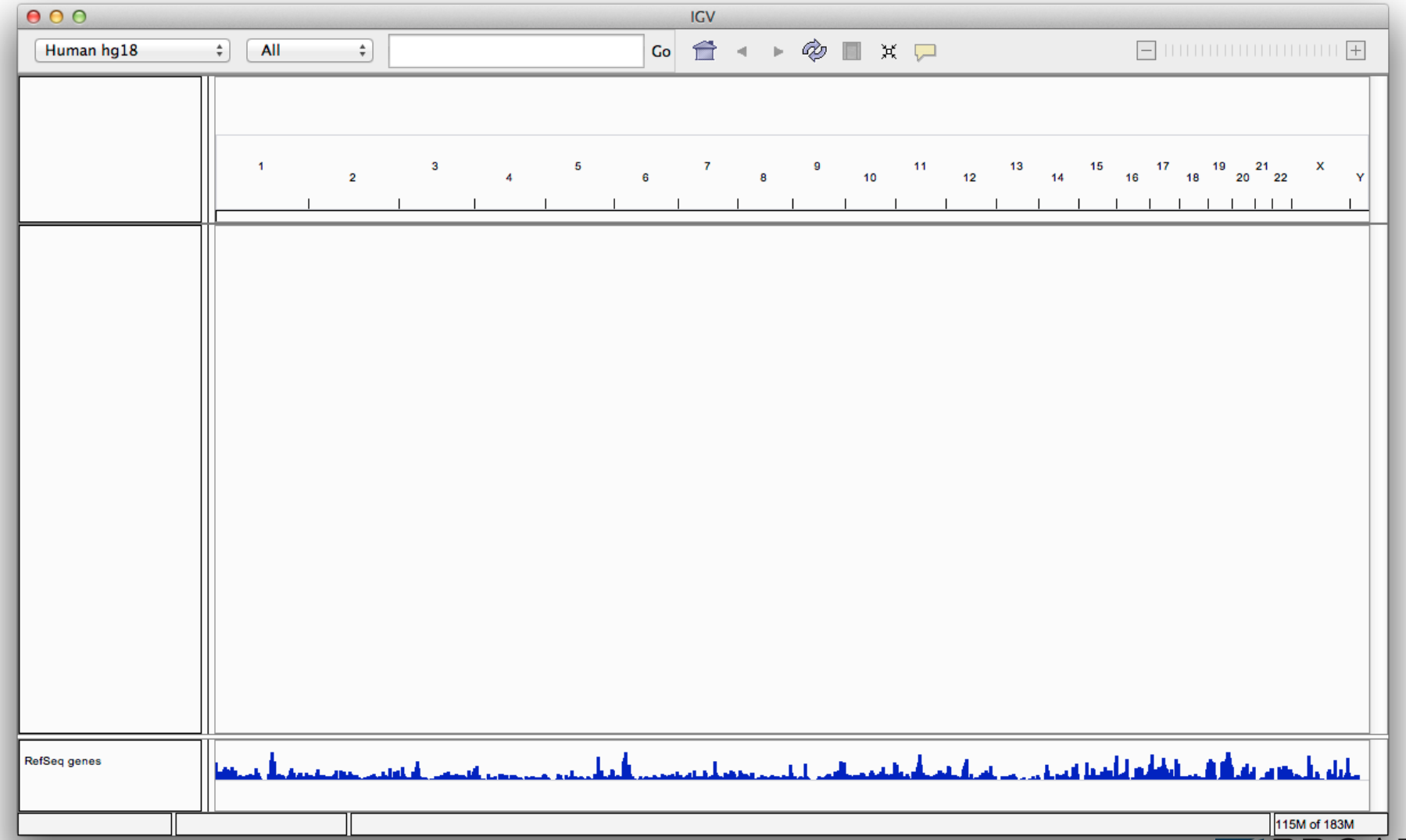
# Launch IGV


Downloads | Integrative Genomics Viewer
www.broadinstitute.org/igv/download
Reader
Integrative Genomics Viewer
Home
Downloads
Documents
Hosted Genomes
FAQ
IGV User Guide
File Formats
Release Notes
Credits
Contact
search
Broad Home
Cancer Program
BROAD INSTITUTE
© 2013 Broad Institute
Home › Downloads
Downloads
Integrative Genomics Viewer (Version 2.3)
Mac Users: Apple has pushed out an update that blocks all but the latest versions of Java. See this article for details. To run IGV from the web launch buttons below, you need the latest version of Java. Another option which avoids Mac security issues is to use the "zip" distribution below. After unzipping double-click the "igv.command" file to launch IGV.
Java: IGV 2.3 requires Java 6 or greater. To use the launch buttons below on MacOS Java 7 is required.
Chrome: Chrome does not launch 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 you must install 64-bit Java. This is often not installed by default even with the latest Windows 7 machines with many GB of memory. In general trying to launch with more memory than your OS/Java combination supports will result in the obscure error "could not create virtual machine".
Launch
Launch with 750 MB
Launch
Launch with 1.2 GB
Maximum usable memory for Windows OS with 32-bit Java.
Launch
Launch with 2 GB
Maximum usable memory for 32-bit MacOS.
Launch
Launch with 10 GB
For large memory 64-bit java machines.
Nightly Build Latest development build.
Archived Versions
igvtools
Utilities for preprocessing data files.
igvtools_2.3.20.zip
Download
A downloadable version that does not require launching from the web. For Windows, Mac OS X, and Linux.
IGV_2.3.20.zip
Source Code
Source distribution archive:
v2.3.20.zip
Source code repository is hosted at github:
https://github.com/broadinstitute/IGV/

# Launch IGV

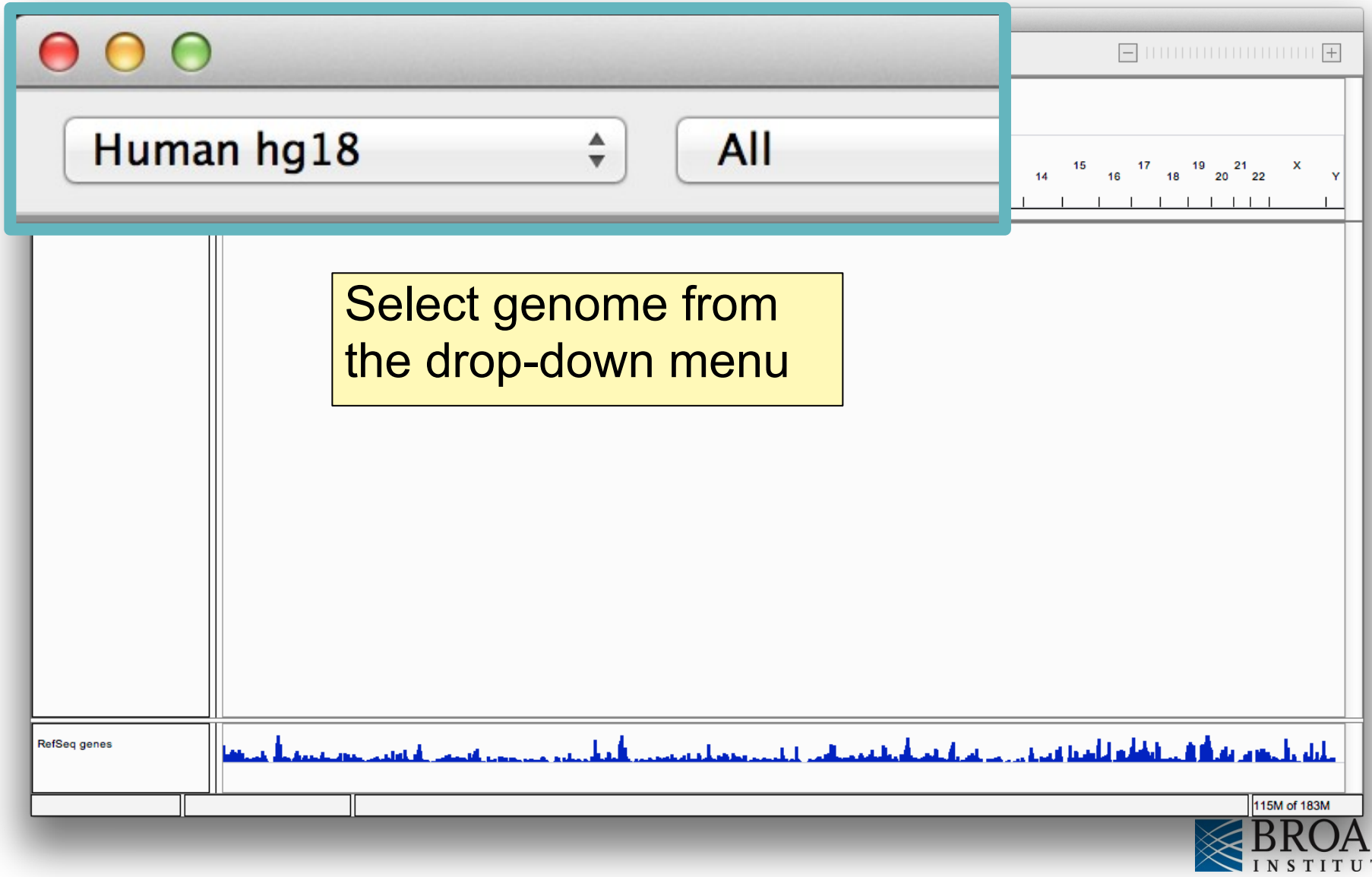

Downloads | Integrative Genomics Viewer
www.broadinstitute.org/igv/download
Reader
Integrative Genomics Viewer
Home
Downloads
Documents
Hosted Genomes
FAQ
IGV User Guide
File Formats
Release Notes
Credits
Contact
Search website
search
Broad Home
Cancer Program
BROAD INSTITUTE
© 2013 Broad Institute
Home › Downloads
Downloads
Integrative Genomics Viewer (Version 2.3)
Mac Users: Apple has pushed out an update that blocks all but the latest versions of Java. See this article for details. To run IGV from the web launch buttons below, you need the latest version of Java. Another option which avoids Mac security issues is to use the "zip" distribution below. After unzipping double-click the "igv.command" file to launch IGV.
Java: IGV 2.3 requires Java 6 or greater. To use the launch buttons below on MacOS Java 7 is required.
Chrome: Chrome does not launch 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 you must install 64-bit Java. This is often not installed by default even with the latest Windows 7 machines with many GB of memory. In general trying to launch with more memory than your OS/Java combination supports will result in the obscure error "could not create virtual machine".
Launch
Launch with 750 MB
Launch
Launch with 1.2 GB
Maximum usable memory for Windows OS with 32-bit Java.
Launch
Launch with 2 GB
Maximum usable memory for 32-bit MacOS.
Launch
Launch with 10 GB
For large memory 64-bit java machines.
Nightly Build Latest development build.
Archived Versions
igvtools
Utilities for preprocessing data files.
igvtools_2.3.20.zip
Download
A downloadable version that does not require launching from the web. For Windows, Mac OS X, and Linux.
IGV_2.3.20.zip
Source Code
Source distribution archive:
v2.3.20.zip
Source code repository is hosted at github:
https://github.com/broadinstitute/IGV/

Integrative Genomics Viewer

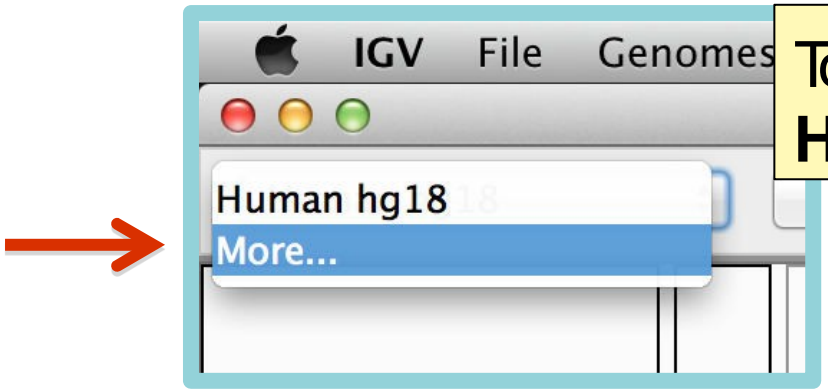



BROAD INSTITUTE

# Select the reference genome



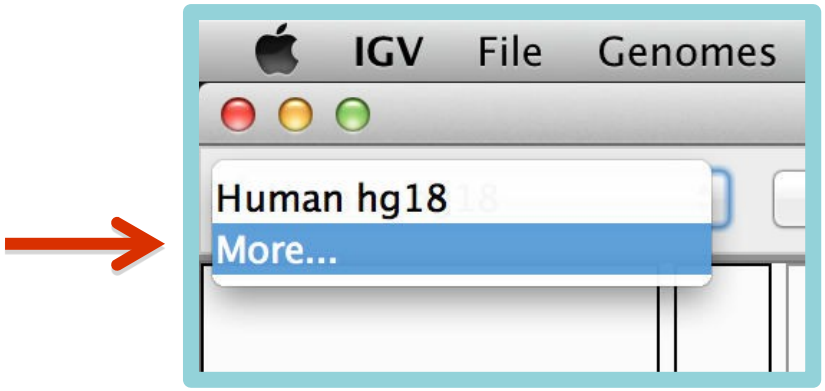
# Select the reference genome

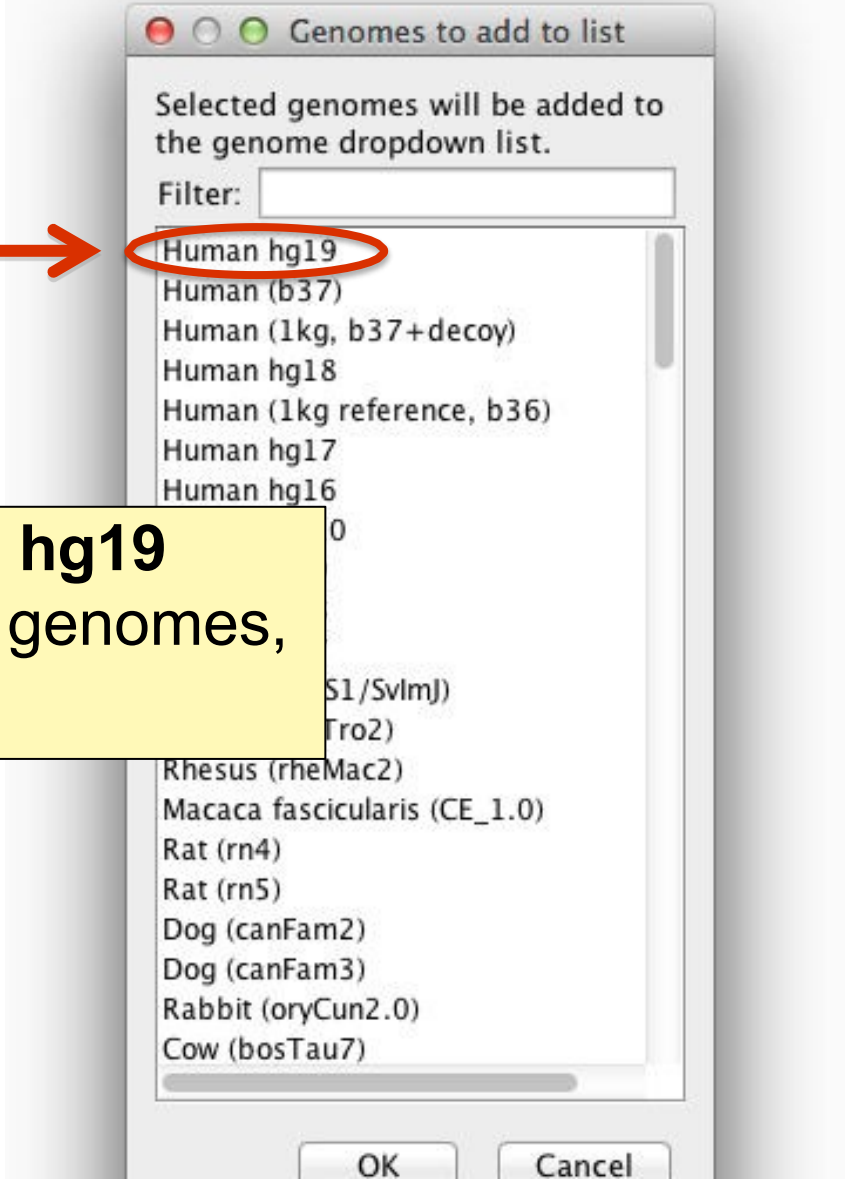



Today, we will use both **Human hg18** and **hg19**

If **Human hg19** is not in the menu, then click on ***More…***

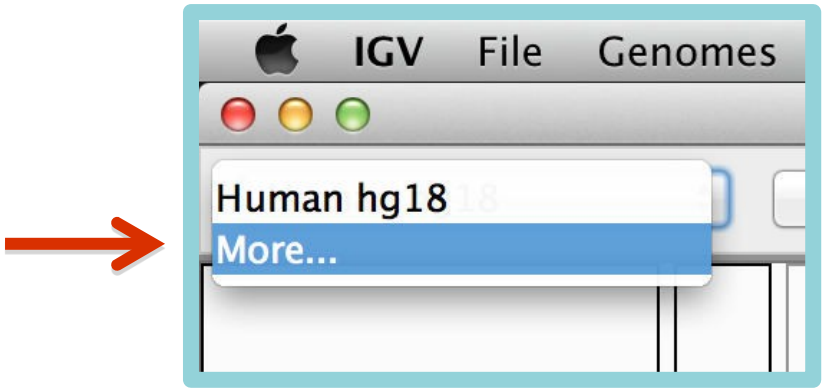
# Select the reference genome

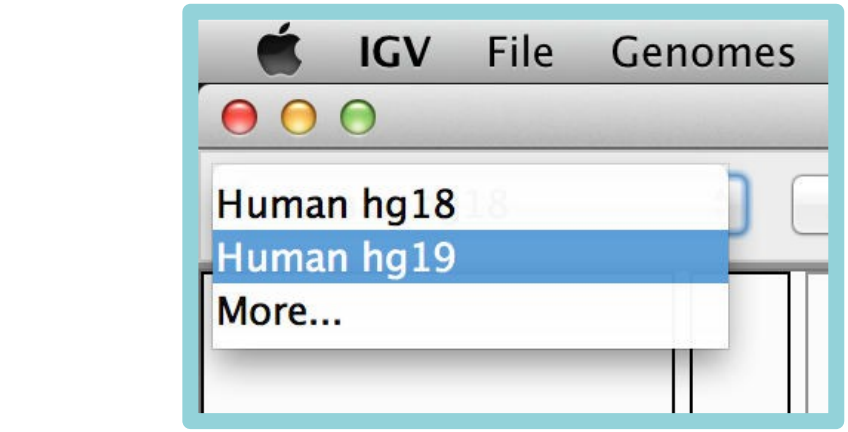


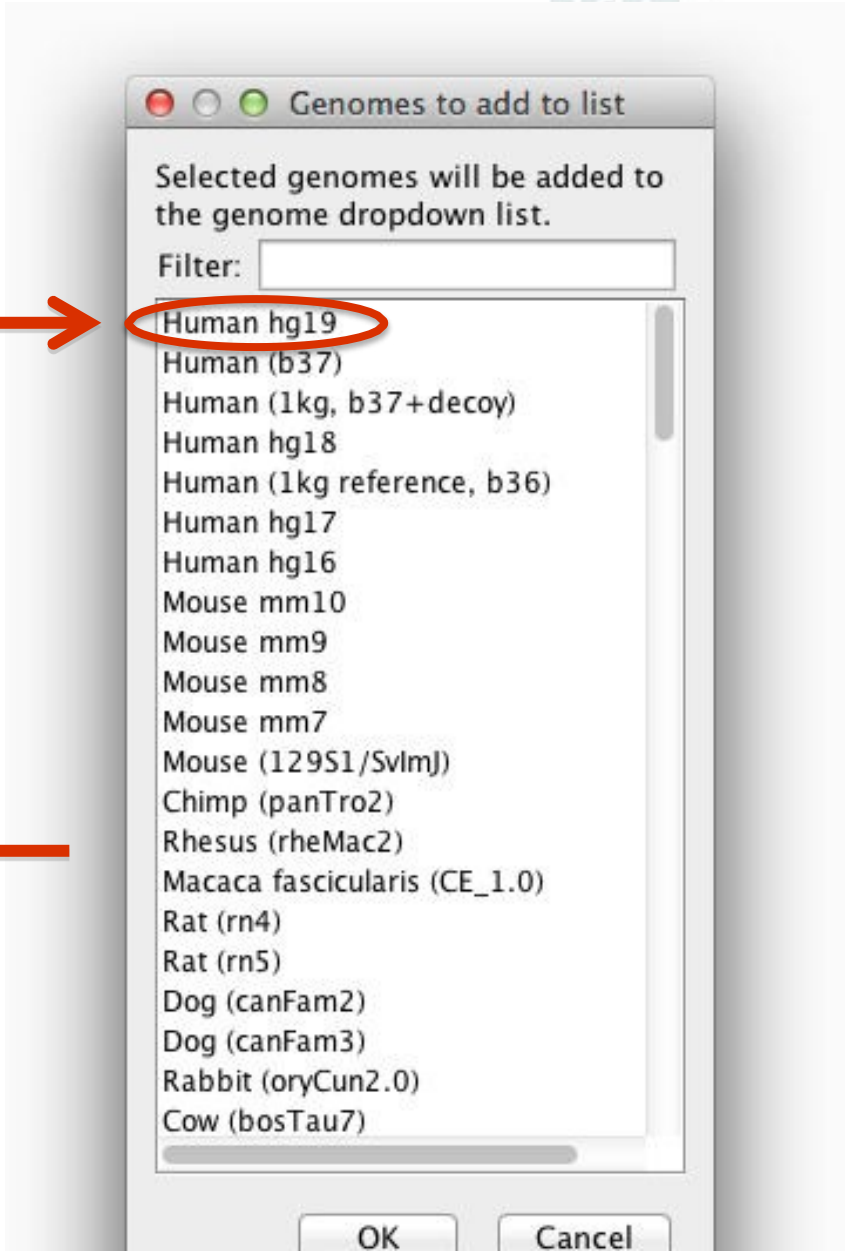




Select **Human hg19** from the list of genomes, and click **OK**

# Select the reference genome



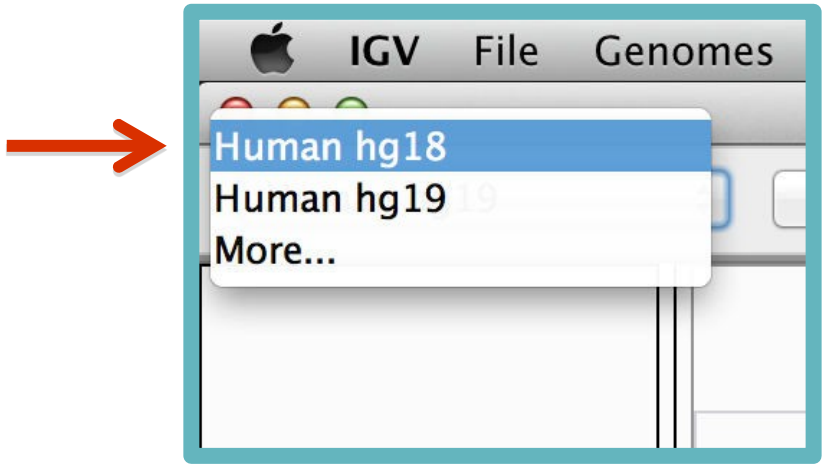



# Select the reference genome

Integrative Genomics Viewer

Select **Human hg18**

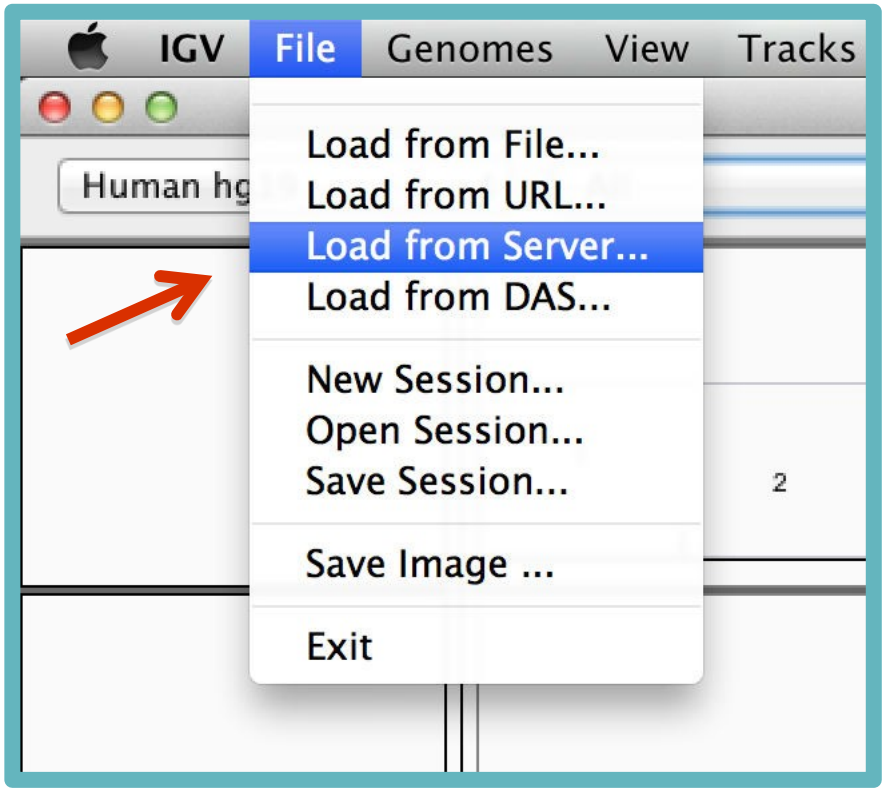



BROAD INSTITUTE

# Load data

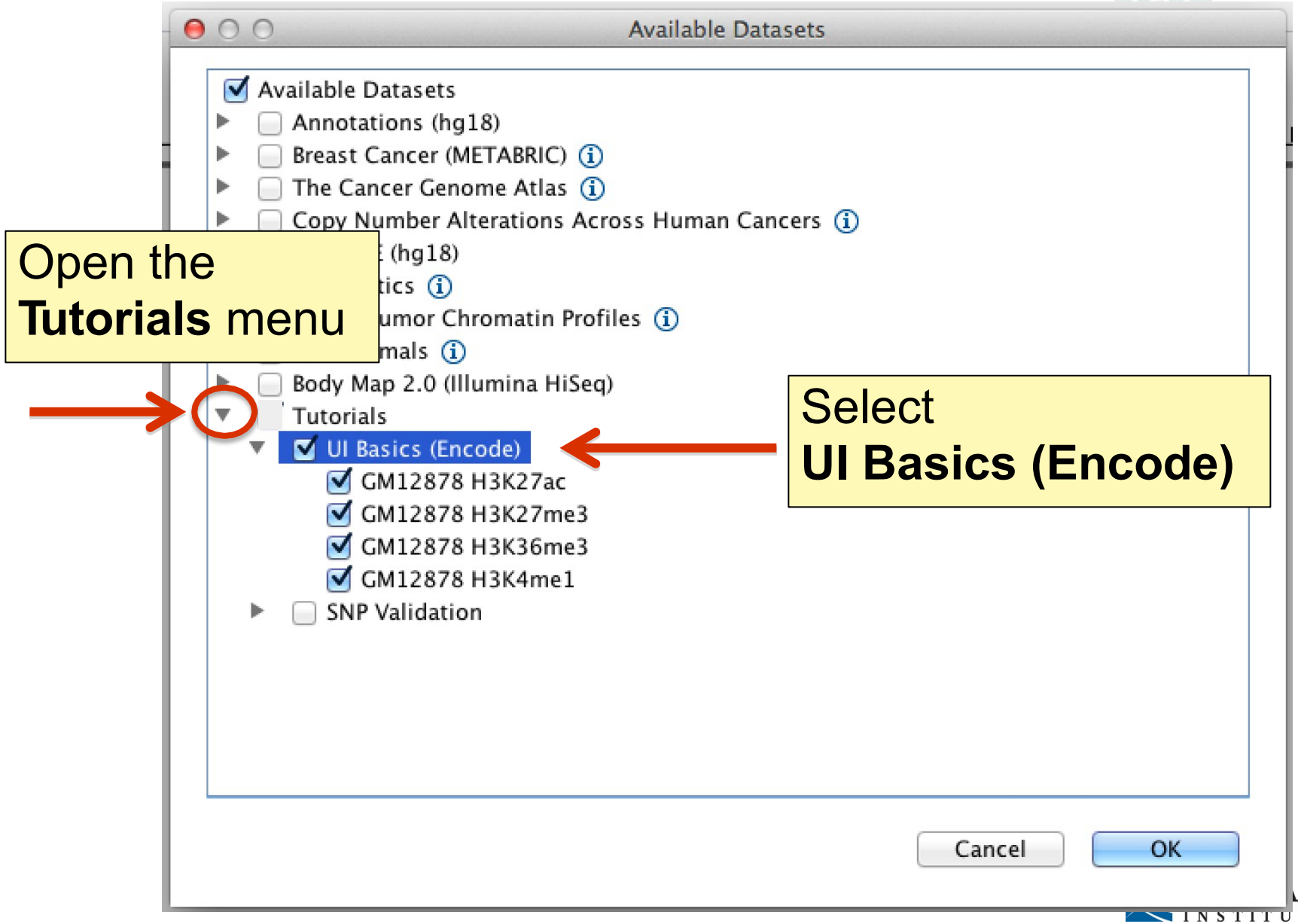
Integrative Genomics Viewer

Select **File > Load from Server…**

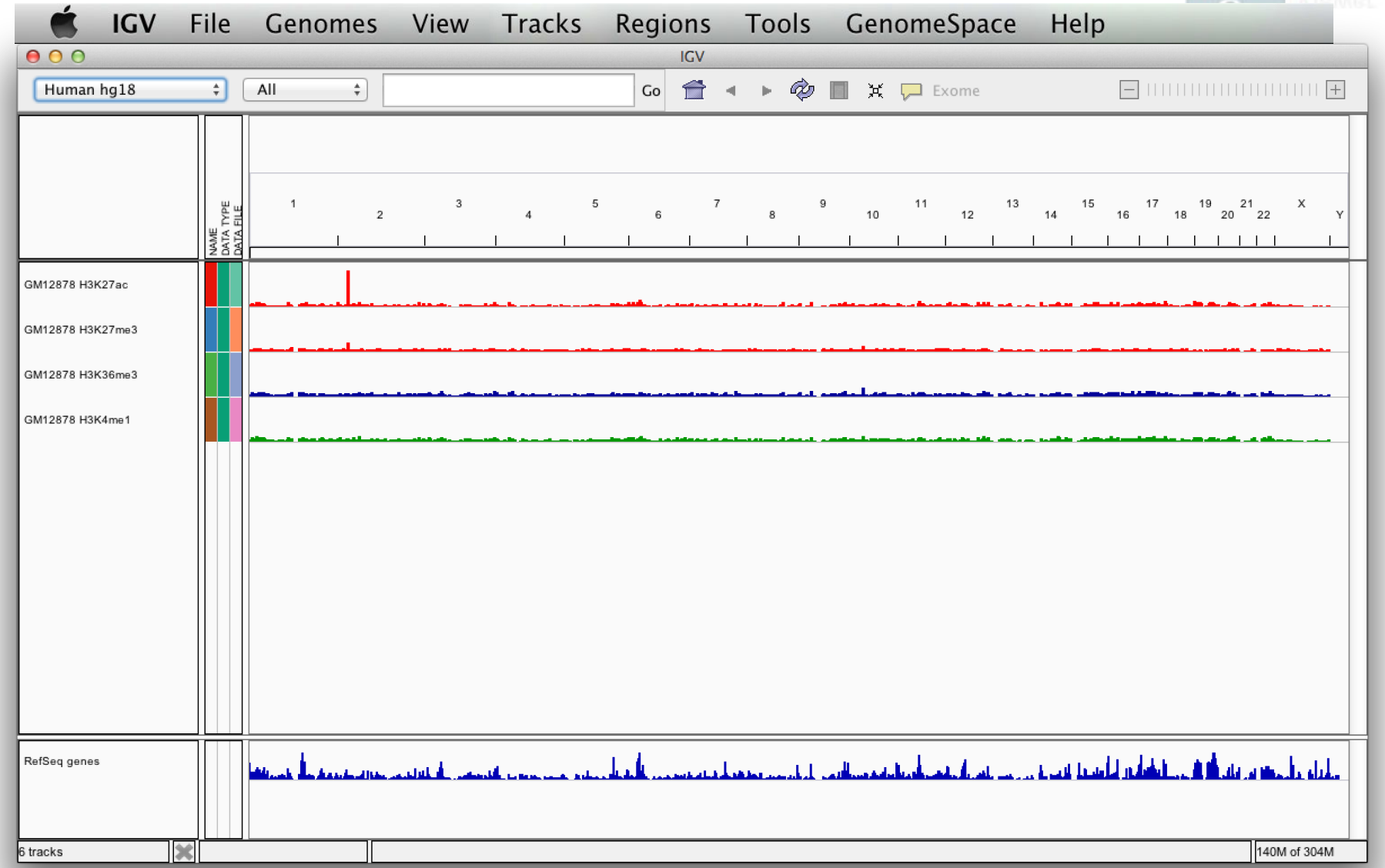



BROAD INSTITUTE

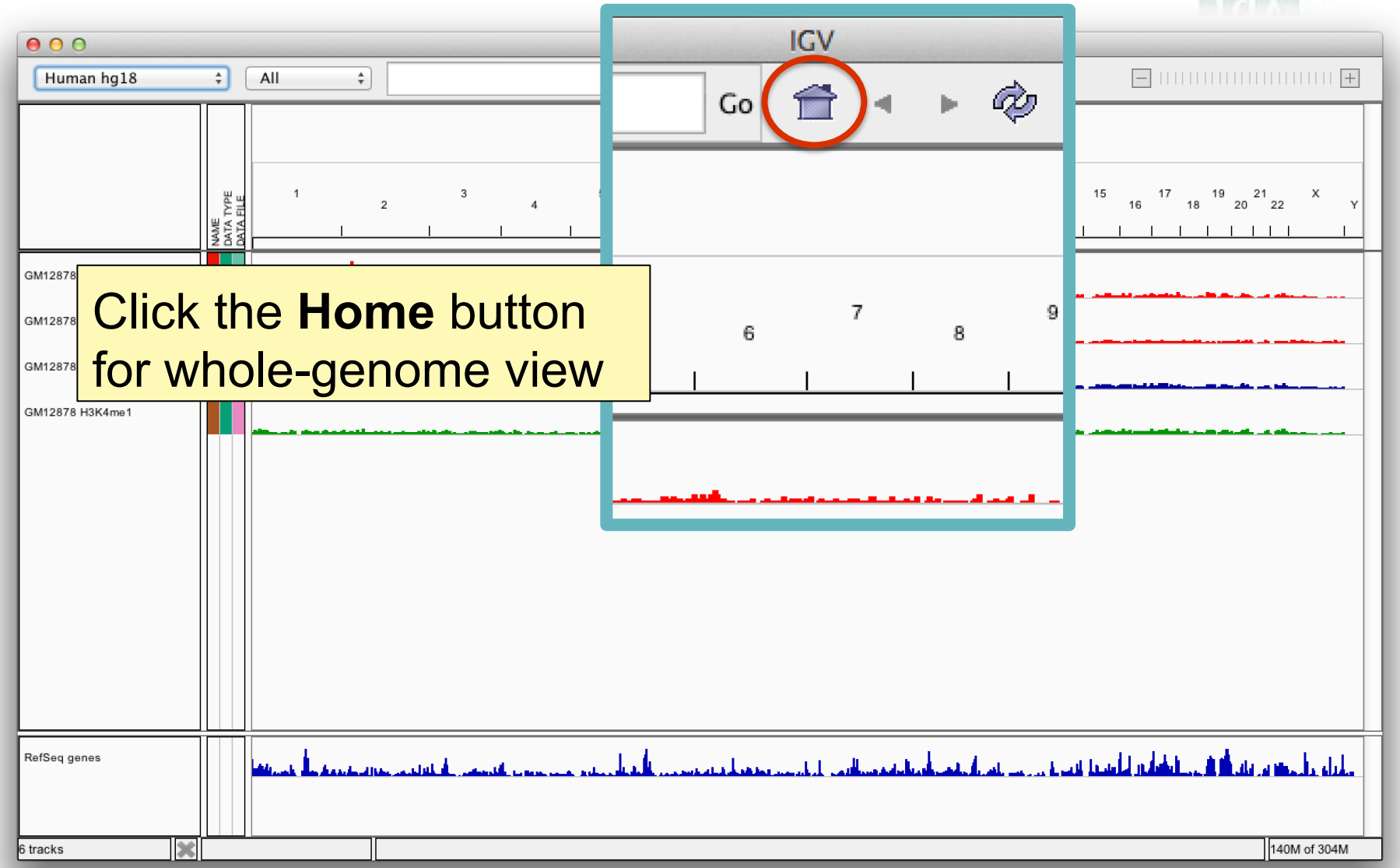
# Load data


Available Datasets
Available Datasets
Annotations (hg18)
Breast Cancer (METABRIC)
The Cancer Genome Atlas
Copy Number Alterations Across Human Cancers
Body Map 2.0 (Illumina HiSeq)
Tutorials
UI Basics (Encode)
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
SNP Validation
Cancel
OK
Open the **Tutorials** menu
Select **UI Basics (Encode)**

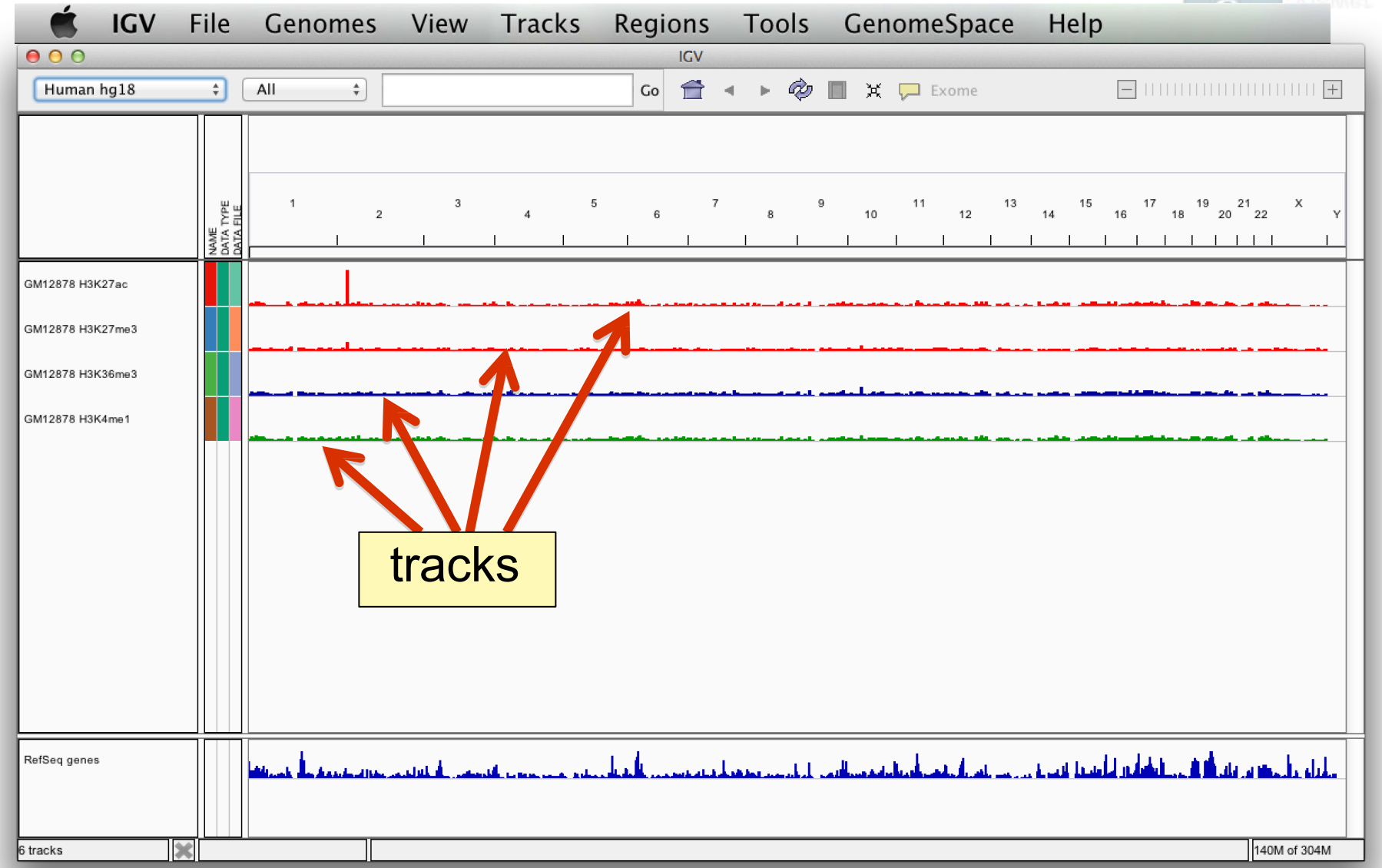
# Screen layout



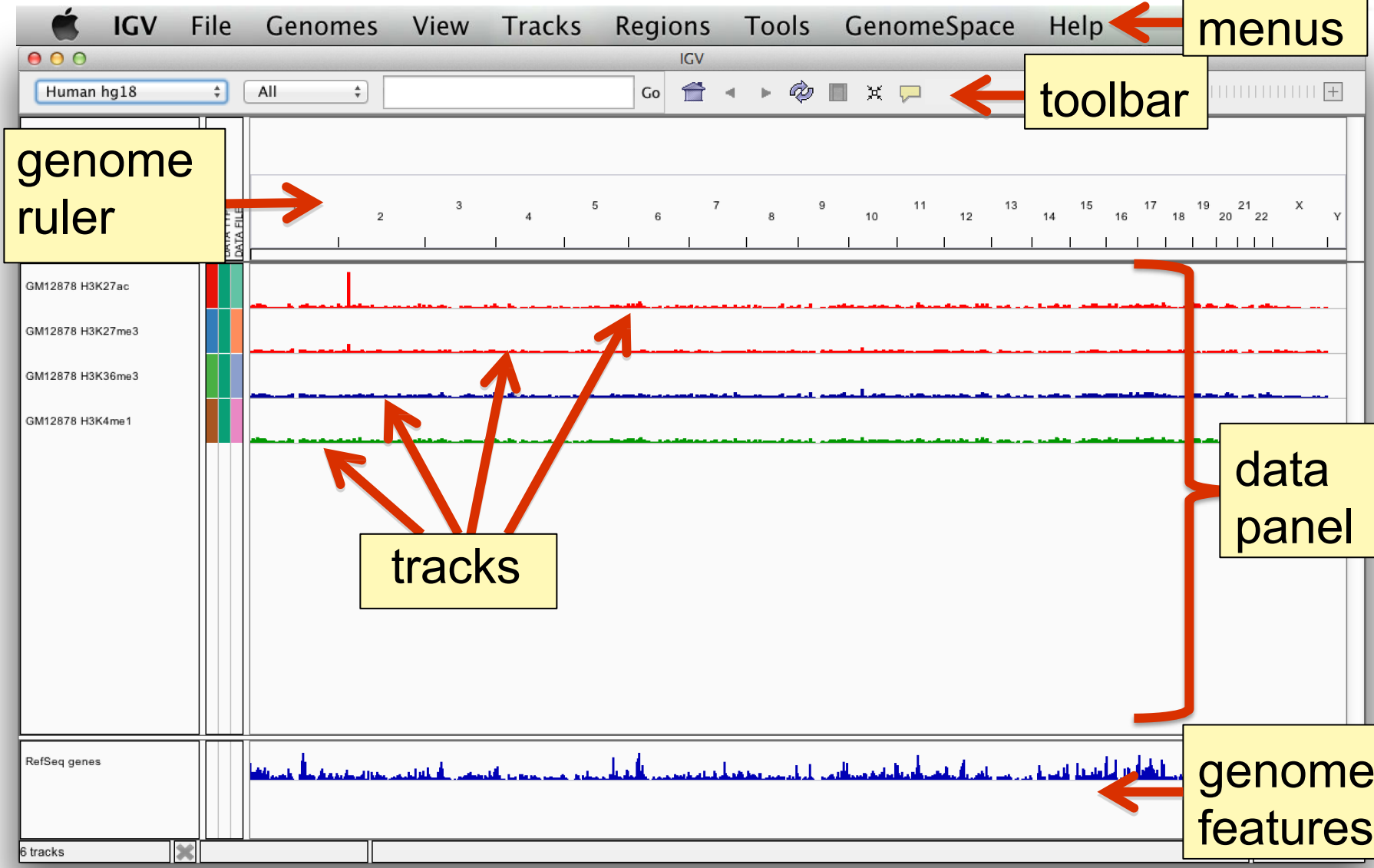
# Screen layout



# Screen layout



# Screen layout




Integrative Genomics Viewer

BROAD INSTITUTE

# File formats and track types

- The **file format** defines the track type.
- The **track type** determines the display options

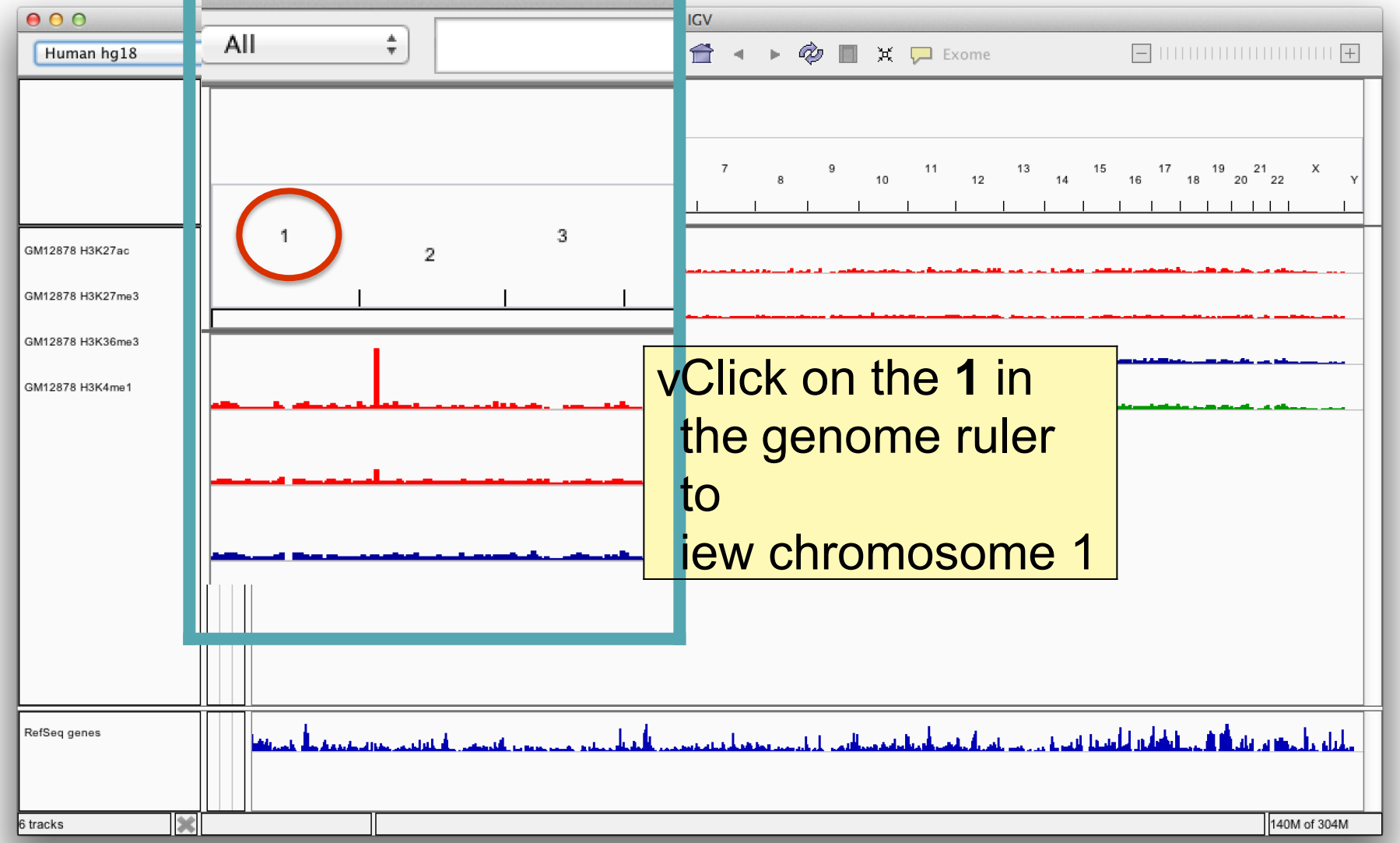
# File formats and track types


- The **file format** defines the track type.
- The **track type** determines the display options
- IGV supports many different file formats.

  - BAM
  - BED
  - BedGraph
  - bigBed
  - bigWig
  - Birdsuite Files
  - broadPeak
  - CBS
  - CN
  - Cufflinks Files
  - Custom File Formats
  - Cytoband
  - FASTA
  - GCT
  - genePred
  - GFF
  - GISTIC
  - Goby
  - GWAS
  - IGV
  - LOH
  - MAF (Multiple Alignment Format)
  - MAF (Mutation Annotation Format)
  - Merged BAM File
  - MUT
  - narrowPeak
  - PSL
  - RES
  - SAM
  - Sample Information
  - SEG
  - SNP
  - TAB
  - TDF
  - Track Line
  - Type Line
  - VCF
  - WIG

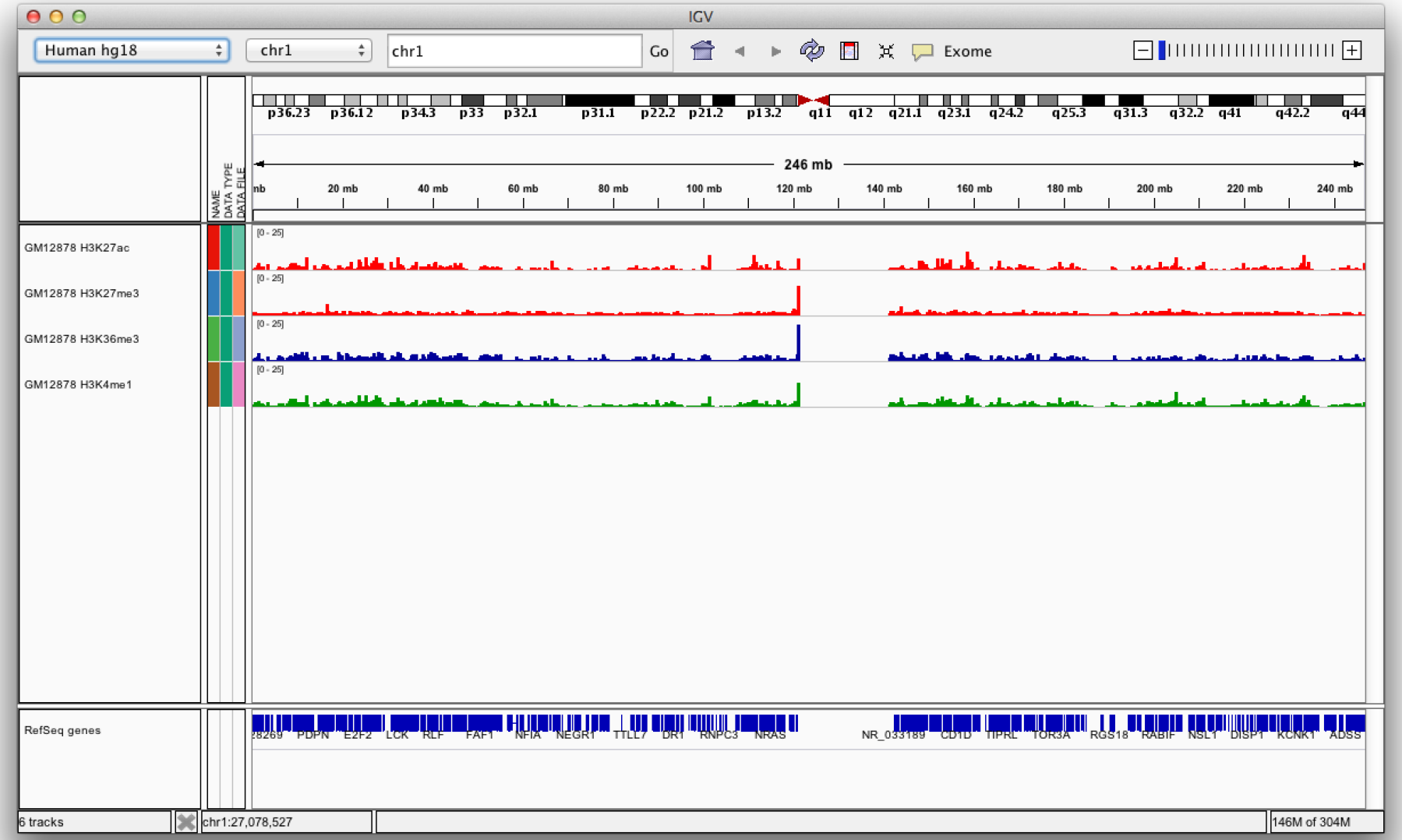
- For current list see: www.broadinstitute.org/igv/FileFormats

# Navigate

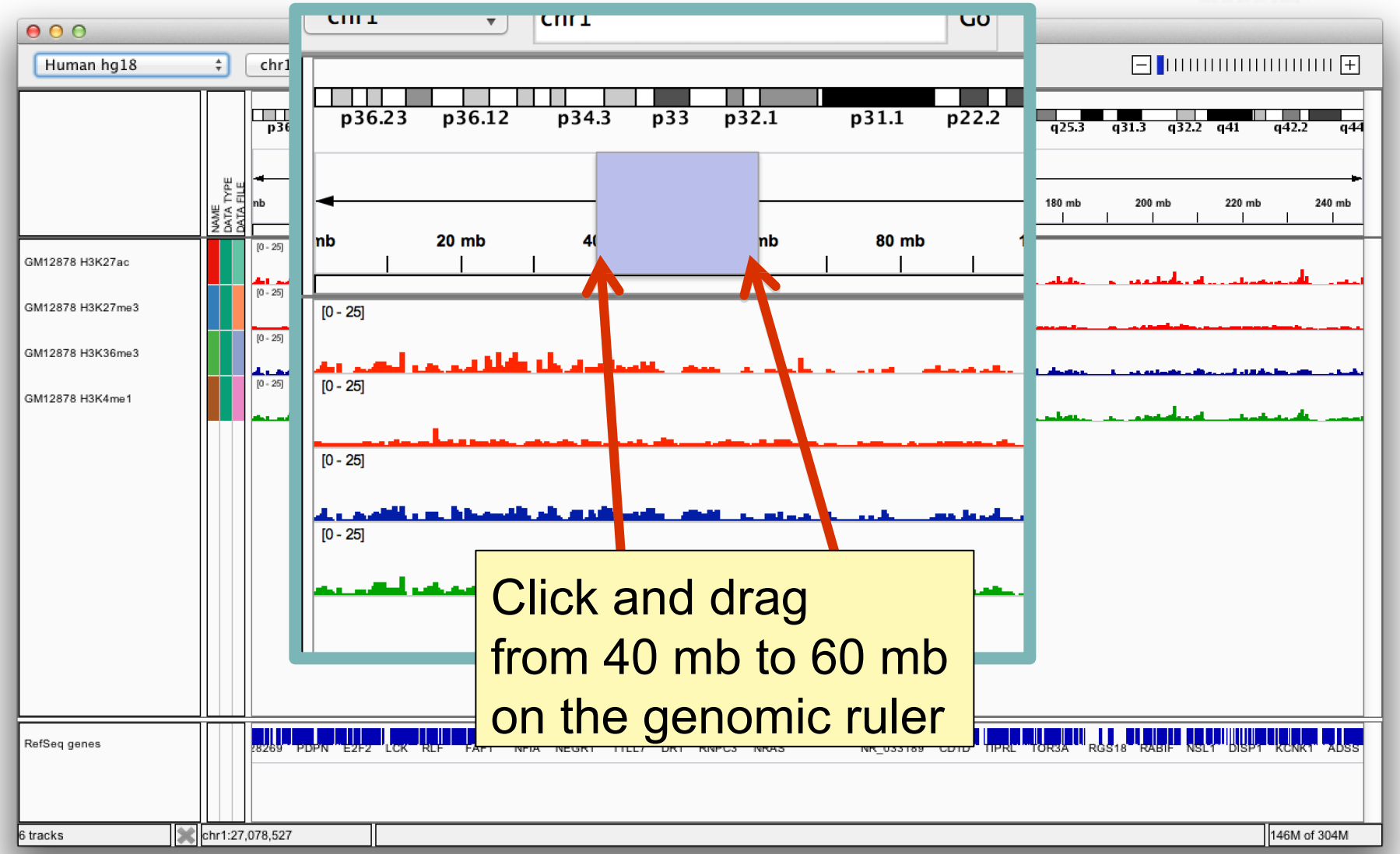



vClick on the **1** in the genome ruler to iew chromosome 1

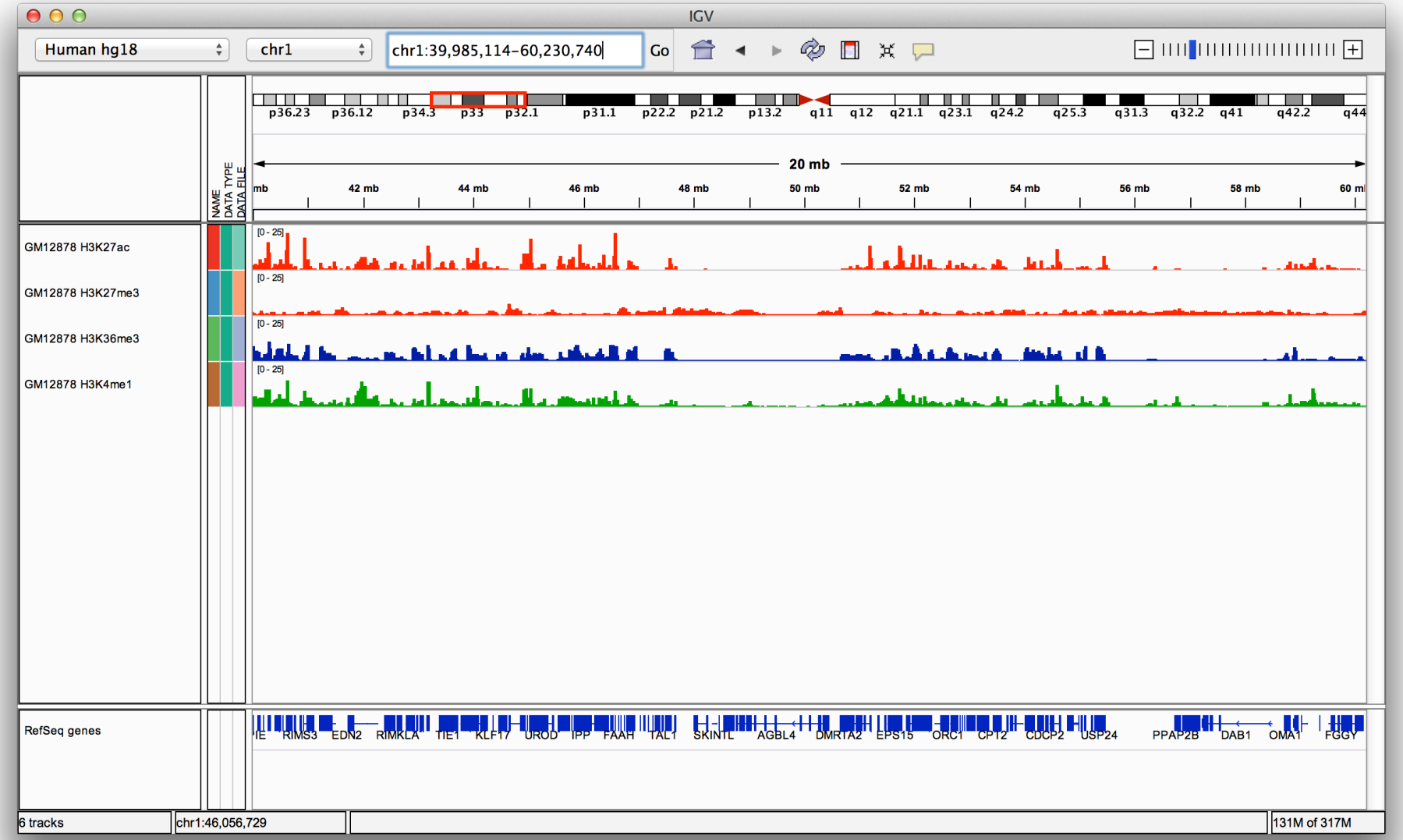
# Navigate


IGV
Human hg18
chr1
chr1
Go
Exome
246 mb
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
RefSeq genes
6 tracks
chr1:27,078,527
146M of 304M

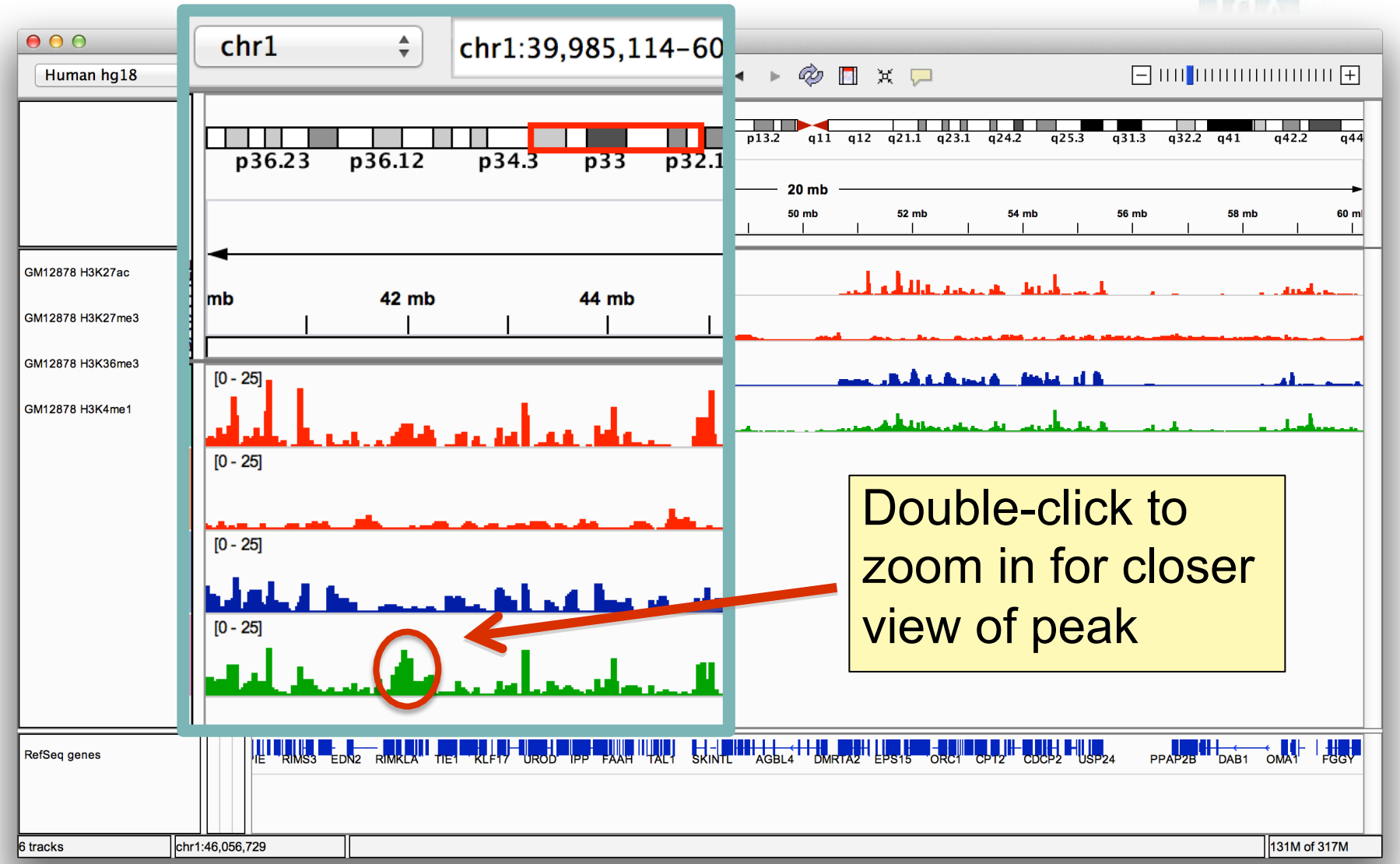
# Navigate


Human hg18
chr1
Go
p36.23
p36.12
p34.3
p33
p32.1
p31.1
p22.2
20 mb
80 mb
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
[0 - 25]
RefSeq genes
6 tracks
chr1:27,078,527
146M of 304M
Click and drag
from 40 mb to 60 mb
on the genomic ruler

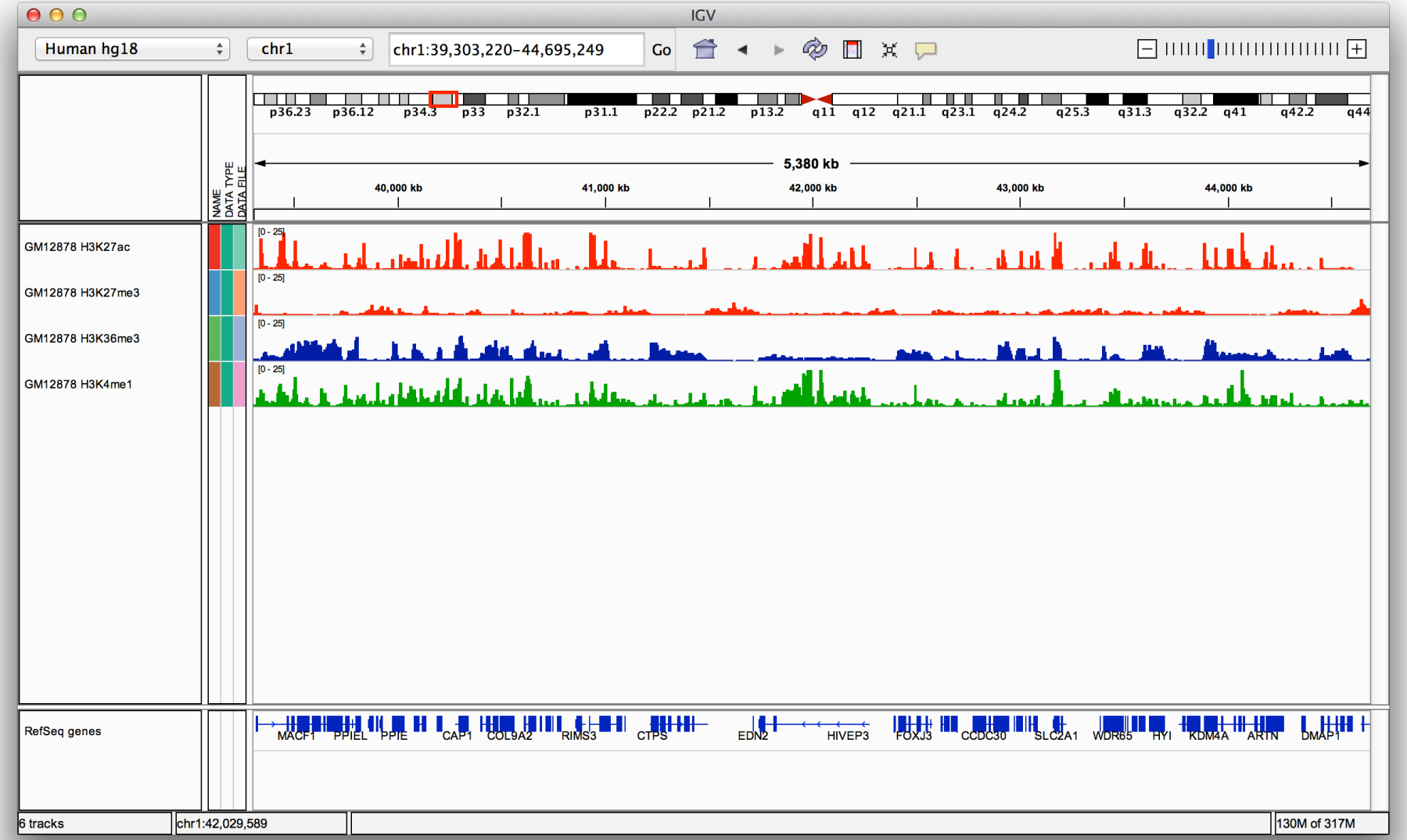
# Navigate


IGV
Human hg18
chr1
chr1:39,985,114-60,230,740
Go
20 mb
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
RefSeq genes
6 tracks
chr1:46,056,729
131M of 317M

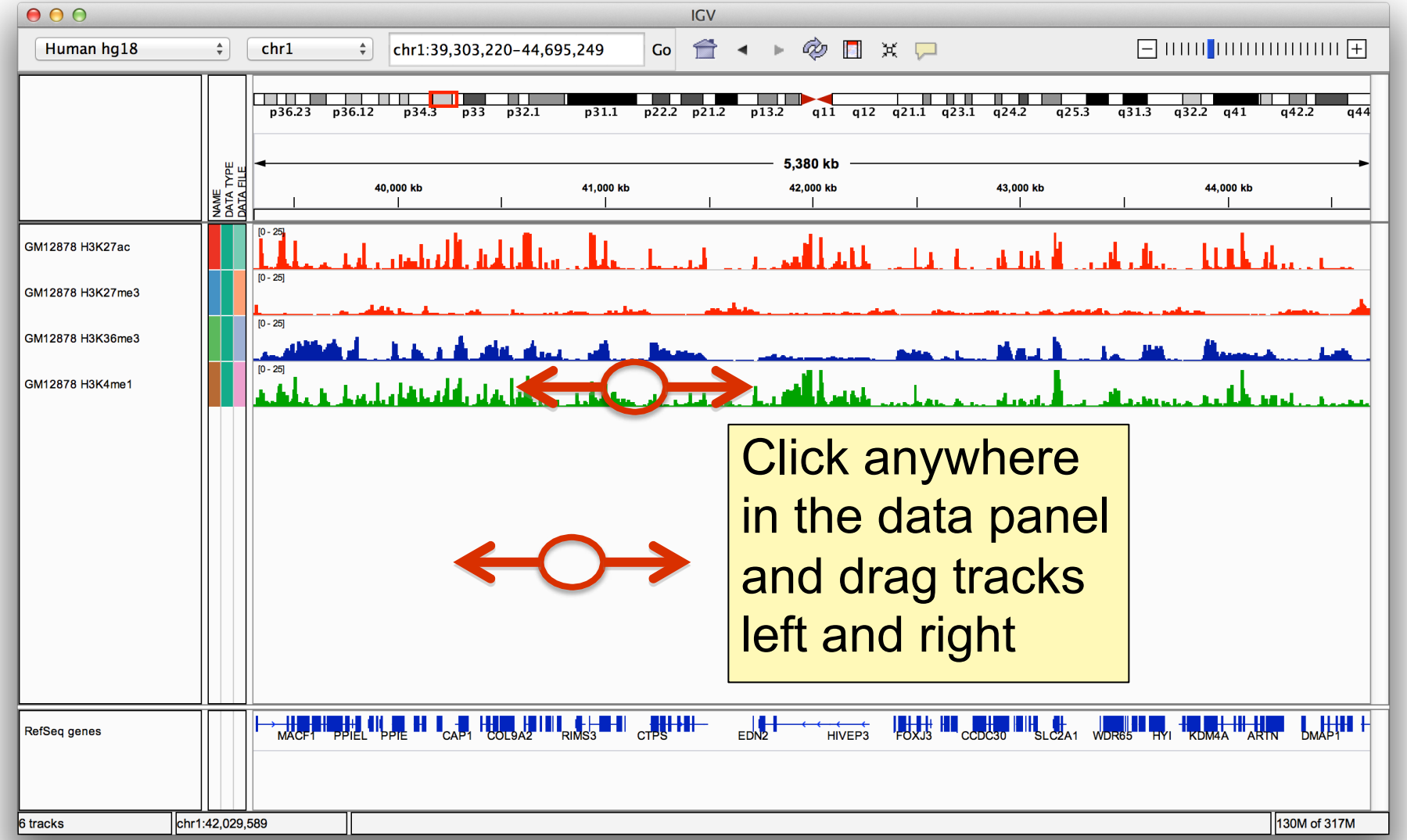
# Navigate


chr1
chr1:39,985,114-60
Human hg18
p36.23
p36.12
p34.3
p33
p32.1
p13.2
q11
q12
q21.1
q23.1
q24.2
q25.3
q31.3
q32.2
q41
q42.2
q44
20 mb
50 mb
52 mb
54 mb
56 mb
58 mb
60 mb
mb
42 mb
44 mb
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
[0 - 25]
[0 - 25]
[0 - 25]
[0 - 25]
Double-click to zoom in for closer view of peak
RefSeq genes
RIMS3
EDN2
RIMKLA
TIE1
KLF17
UROD
IPP
FAAH
TAL1
SKINTL
AGBL4
DMRTA2
EPS15
ORC1
CPT2
CDCP2
USP24
PPAP2B
DAB1
OMA1
FGGY
6 tracks
chr1:46,056,729
131M of 317M

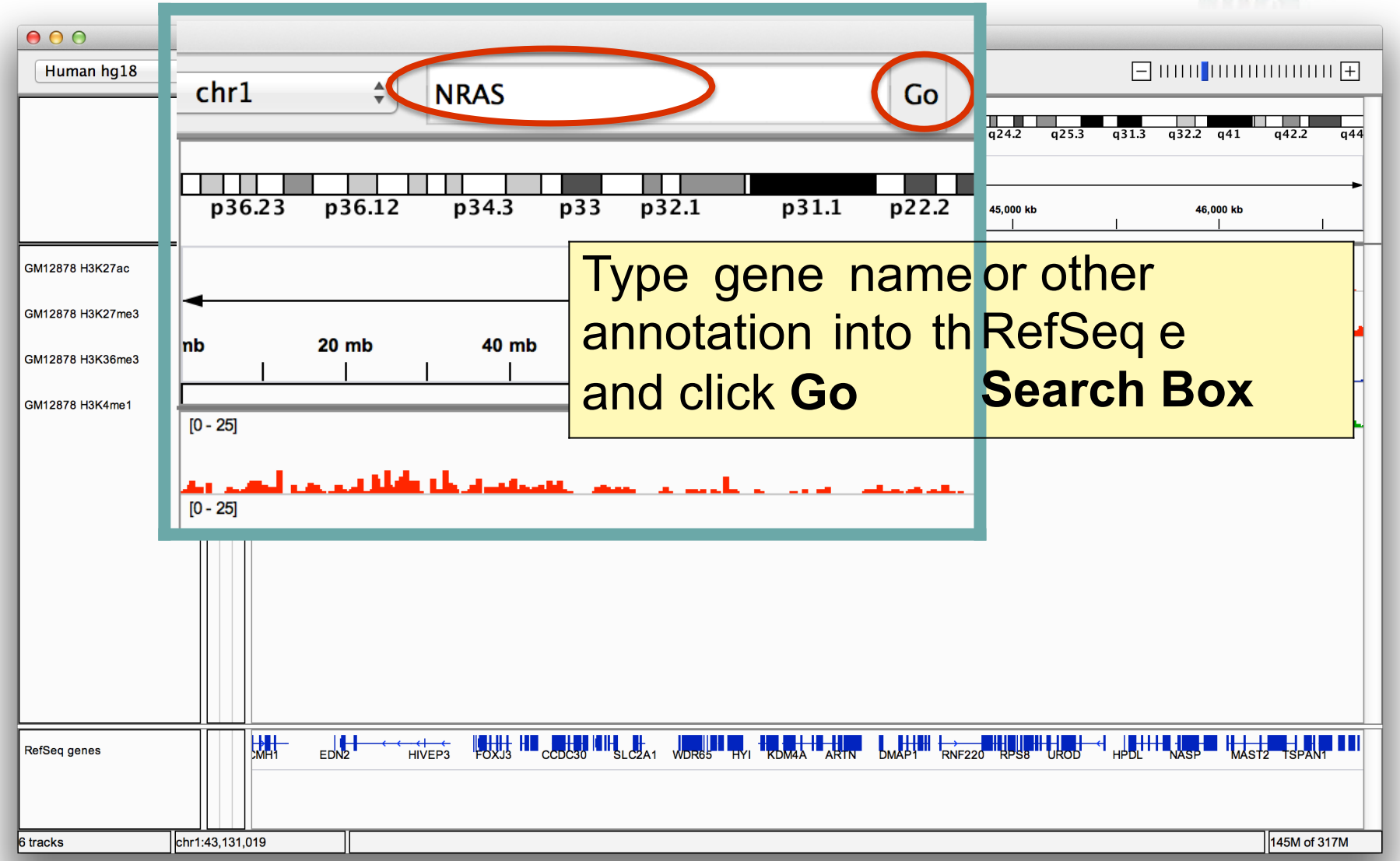
# Navigate


IGV
Human hg18
chr1
chr1:39,303,220-44,695,249
Go
p36.23 p36.12 p34.3 p33 p32.1 p31.1 p22.2 p21.2 p13.2 q11 q12 q21.1 q23.1 q24.2 q25.3 q31.3 q32.2 q41 q42.2 q44
5,380 kb
40,000 kb
41,000 kb
42,000 kb
43,000 kb
44,000 kb
NAME
DATA TYPE
DATA FILE
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
[0 - 25]
RefSeq genes
MACF1 PPIEL PPIE CAP1 COL9A2 RIMS3 CTPS EDN2 HIVEP3 FOXJ3 CCDC30 SLC2A1 WDR65 HYI KDM4A ARTN DMAP1
6 tracks
chr1:42,029,589
130M of 317M

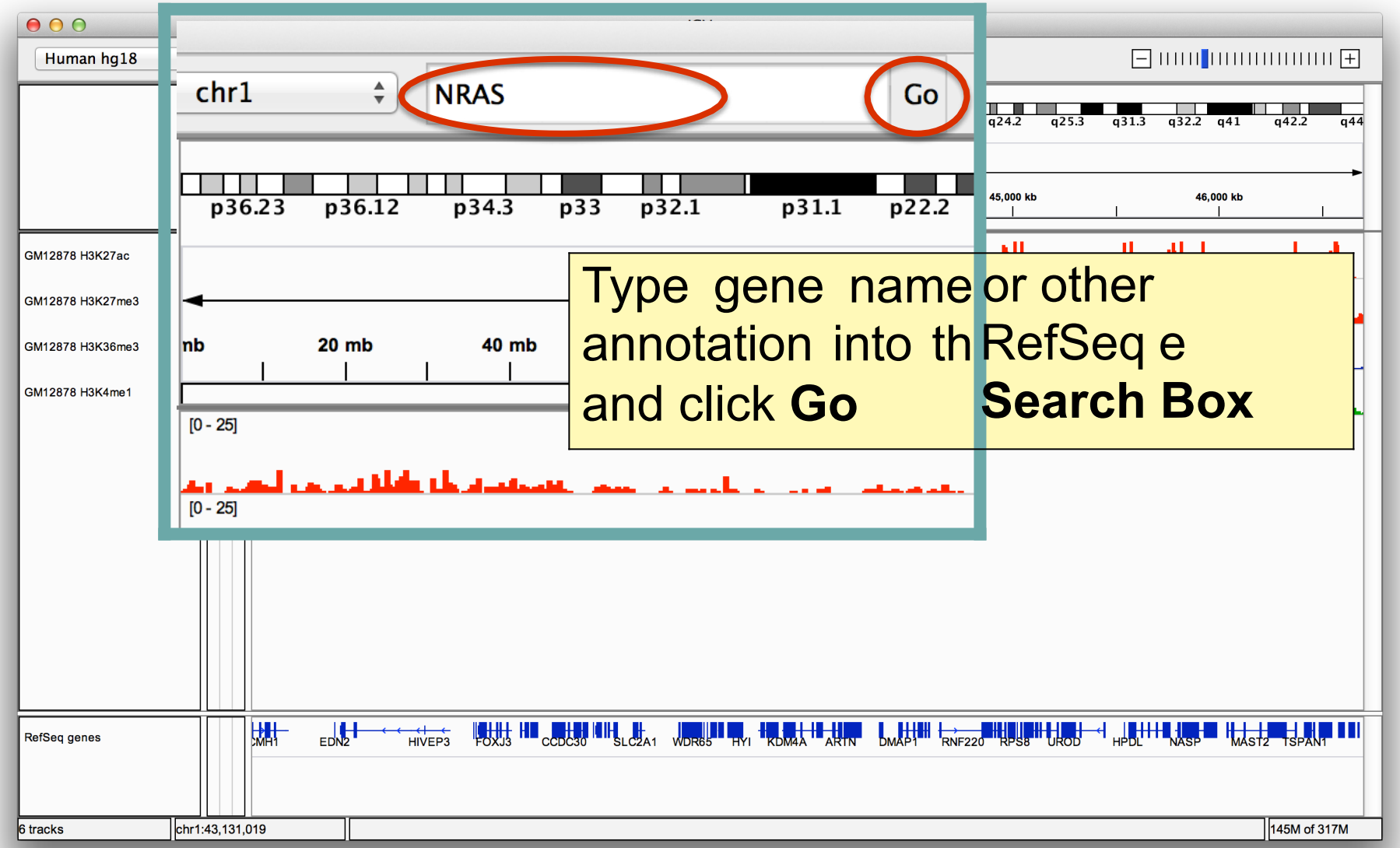
# Navigate


IGV
Human hg18
chr1
chr1:39,303,220-44,695,249
Go
5,380 kb
40,000 kb
41,000 kb
42,000 kb
43,000 kb
44,000 kb
NAME
DATA TYPE
DATA FILE
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
[0 - 25]
RefSeq genes
MACF1
PPIEL
PPIE
CAP1
COL9A2
RIMS3
CTPS
EDN2
HIVEP3
FOXJ3
CCDC30
SLC2A1
WDR65
HYI
KDM4A
ARTN
DMAP1
6 tracks
chr1:42,029,589
130M of 317M
Click anywhere in the data panel and drag tracks left and right

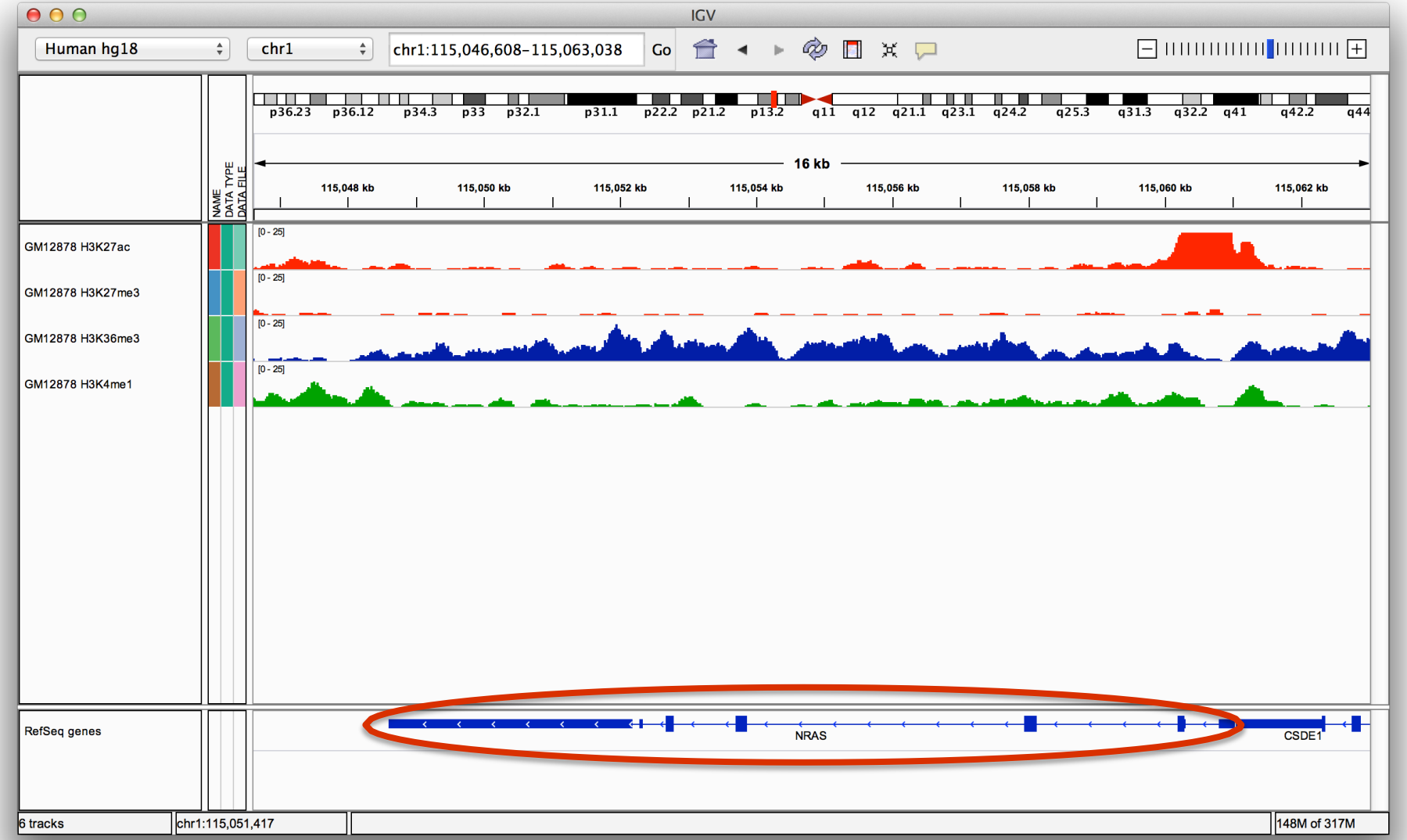
# Navigate


Human hg18
chr1
NRAS
Go
p36.23
p36.12
p34.3
p33
p32.1
p31.1
p22.2
q24.2
q25.3
q31.3
q32.2
q41
q42.2
q44
45,000 kb
46,000 kb
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
nb
20 mb
40 mb
[0 - 25]
[0 - 25]
Type gene name or other annotation into the RefSeq e and click Go
Search Box
RefSeq genes
CMH1
EDN2
HIVEP3
FOXJ3
CCDC30
SLC2A1
WDR65
HYI
KDM4A
ARTN
DMAP1
RNF220
RPS8
UROD
HPDL
NASP
MAST2
TSPAN1
6 tracks
chr1:43,131,019
145M of 317M

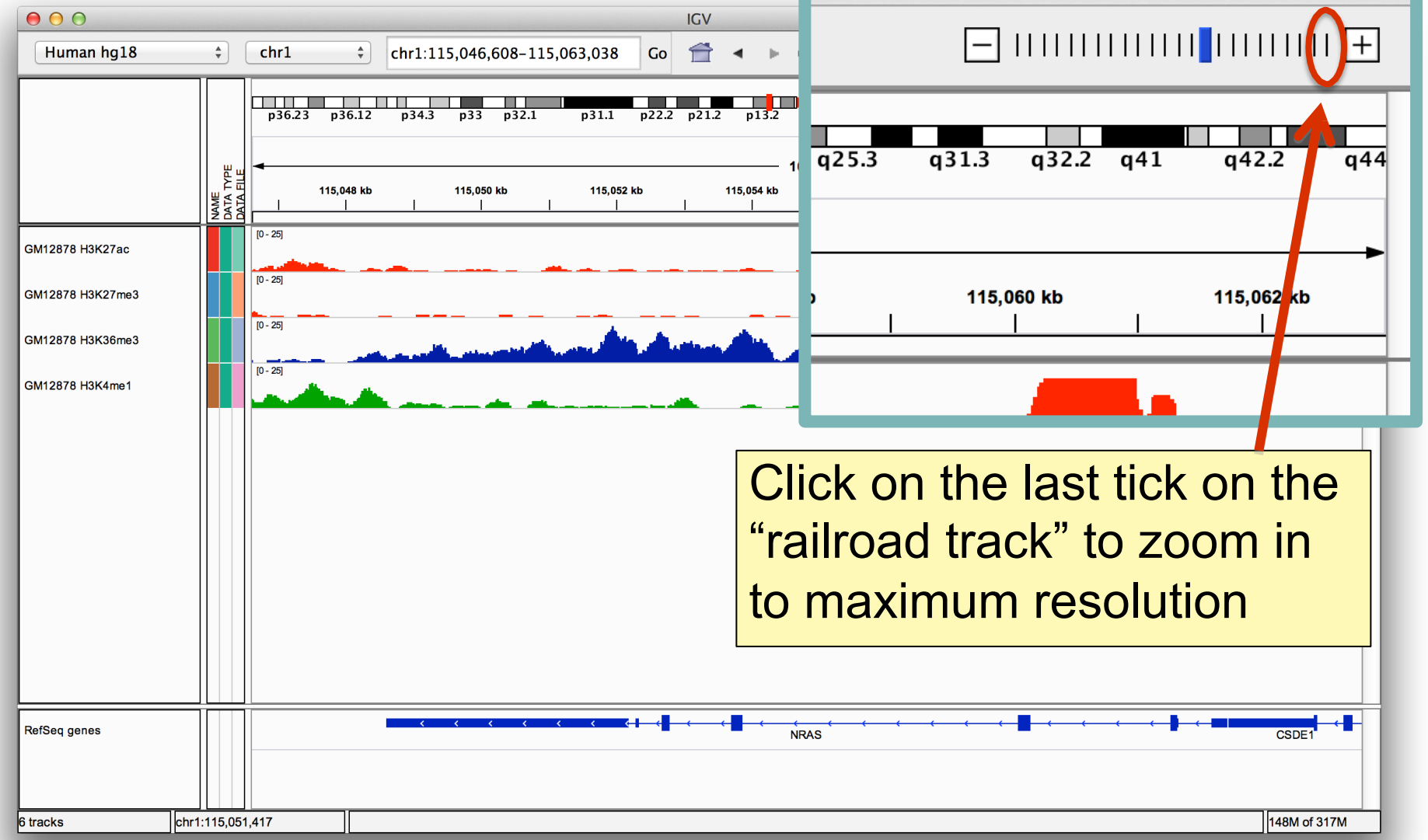
# Navigate


Human hg18
chr1
NRAS
Go
p36.23
p36.12
p34.3
p33
p32.1
p31.1
p22.2
q24.2
q25.3
q31.3
q32.2
q41
q42.2
q44
45,000 kb
46,000 kb
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
nb
20 mb
40 mb
[0 - 25]
[0 - 25]
Type gene name or other annotation into the RefSeq and click Go
Search Box
RefSeq genes
CMH1
EDN2
HIVEP3
FOXJ3
CCDC30
SLC2A1
WDR65
HYI
KDM4A
ARTN
DMAP1
RNF220
RPS8
UROD
HPDL
NASP
MAST2
TSPAN1
6 tracks
chr1:43,131,019
145M of 317M

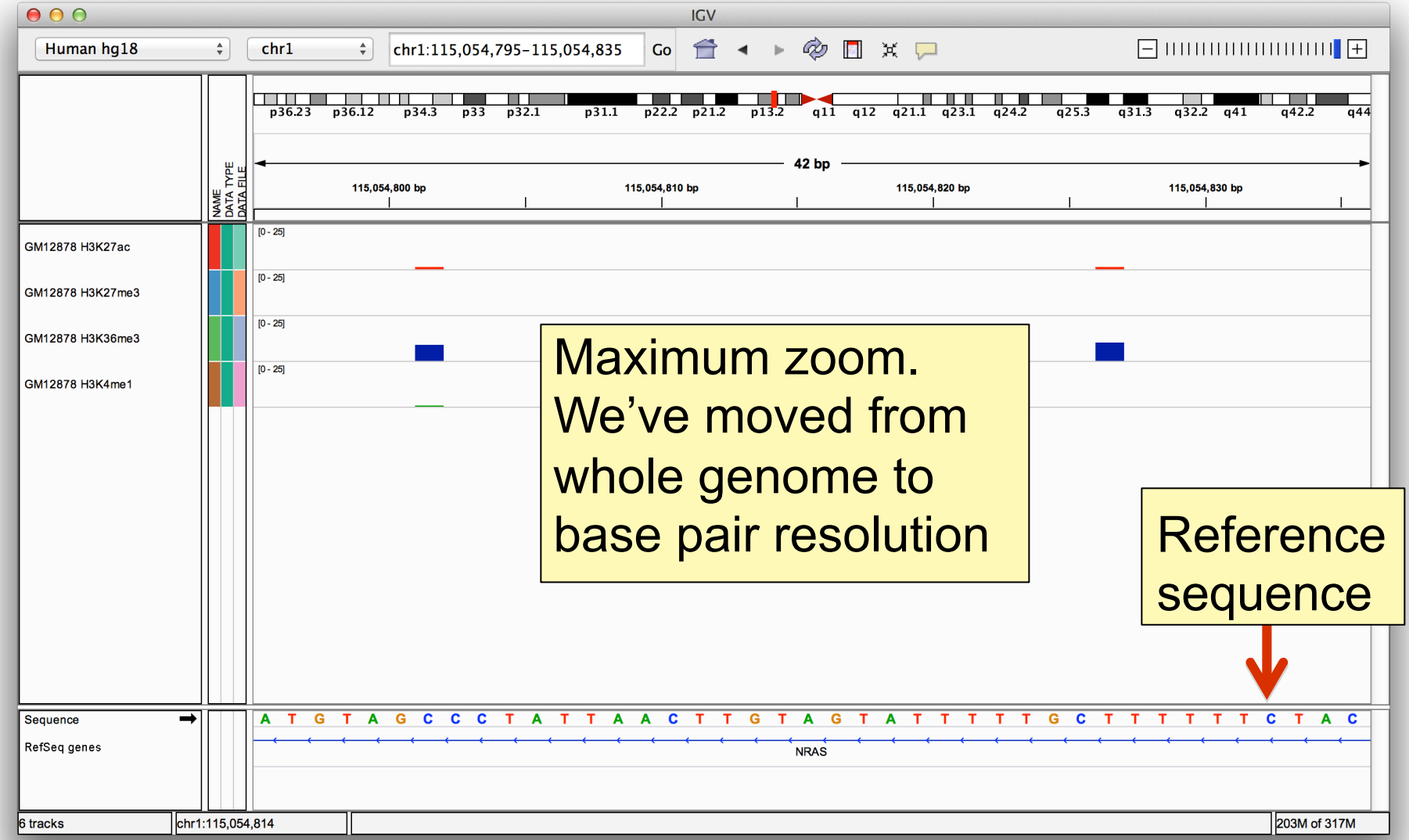
# Navigate


IGV
Human hg18
chr1
chr1:115,046,608-115,063,038
Go
p36.23 p36.12 p34.3 p33 p32.1 p31.1 p22.2 p21.2 p13.2 q11 q12 q21.1 q23.1 q24.2 q25.3 q31.3 q32.2 q41 q42.2 q44
16 kb
115,048 kb 115,050 kb 115,052 kb 115,054 kb 115,056 kb 115,058 kb 115,060 kb 115,062 kb
NAME
DATA TYPE
DATA FILE
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
[0 - 25]
RefSeq genes
NRAS
CSDE1
6 tracks
chr1:115,051,417
148M of 317M

# Navigate


IGV
Human hg18
chr1
chr1:115,046,608-115,063,038
Go
p36.23 p36.12 p34.3 p33 p32.1 p31.1 p22.2 p21.2 p13.2
115,048 kb
115,050 kb
115,052 kb
115,054 kb
NAME
DATA TYPE
DATA FILE
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
[0 - 25]
RefSeq genes
NRAS
CSDE1
6 tracks
chr1:115,051,417
148M of 317M
q25.3 q31.3 q32.2 q41 q42.2 q44
115,060 kb
115,062 kb
Click on the last tick on the "railroad track" to zoom in to maximum resolution

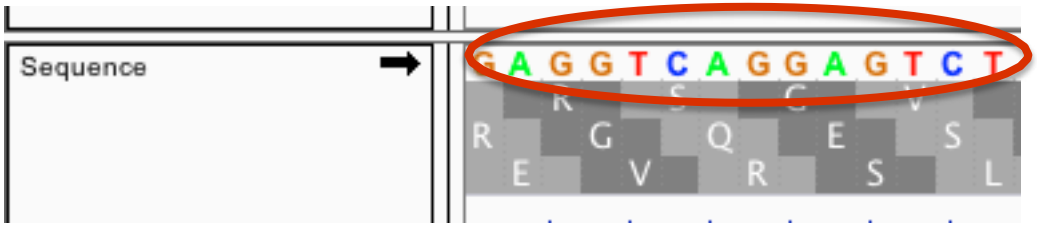
# Navigate


IGV
Human hg18
chr1
chr1:115,054,795-115,054,835
Go
p36.23 p36.12 p34.3 p33 p32.1 p31.1 p22.2 p21.2 p13.2 q11 q12 q21.1 q23.1 q24.2 q25.3 q31.3 q32.2 q41 q42.2 q44
42 bp
115,054,800 bp
115,054,810 bp
115,054,820 bp
115,054,830 bp
NAME
DATA TYPE
DATA FILE
GM12878 H3K27ac
GM12878 H3K27me3
GM12878 H3K36me3
GM12878 H3K4me1
[0 - 25]
Maximum zoom. We've moved from whole genome to base pair resolution
Reference sequence
Sequence
RefSeq genes
A T G T A G C C C T A T T A A C T T G T A G T A T T T T T G C T T T T T T C T A C
NRAS
6 tracks
chr1:115,054,814
203M of 317M

# Reference sequence

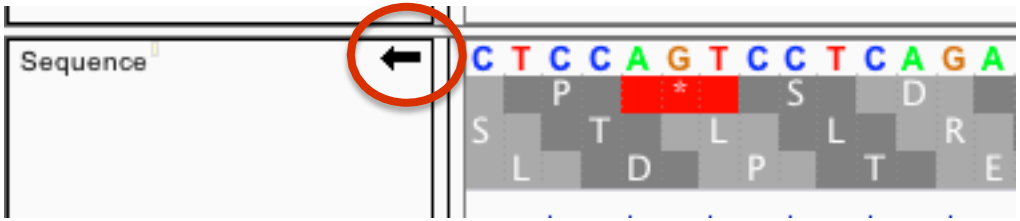
Integrative Genomics Viewer

Click anywhere on the sequence to see a 3 frame translation.




By default the sequence for the forward strand is shown.


Click the arrow on the left to reverse the strand.




BROAD INSTITUTE

# Genome annotation track

Integrative Genomics Viewer

**UCSC style gene representation**

5’ UTR Intron Exons 3’ UTR

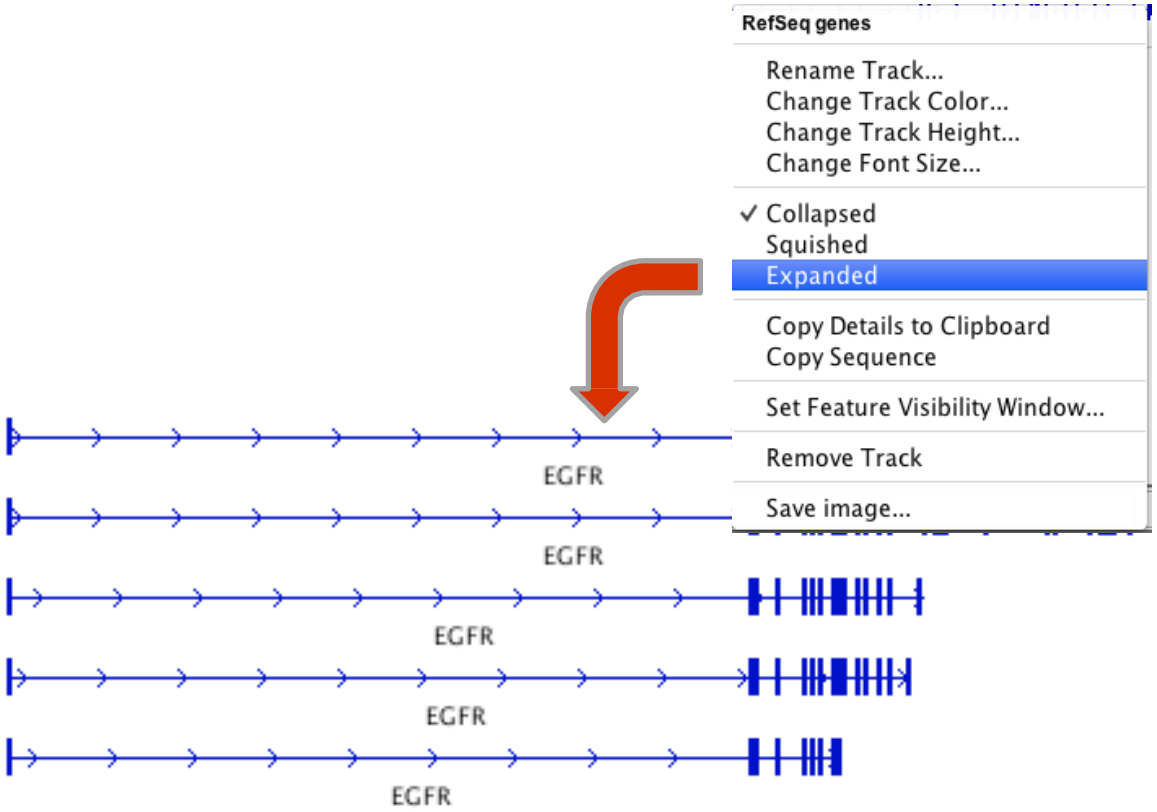
**Zoomed in views**

**Zoomed out views**

# Annotation display mode

1. Features are drawn in a single row, by default

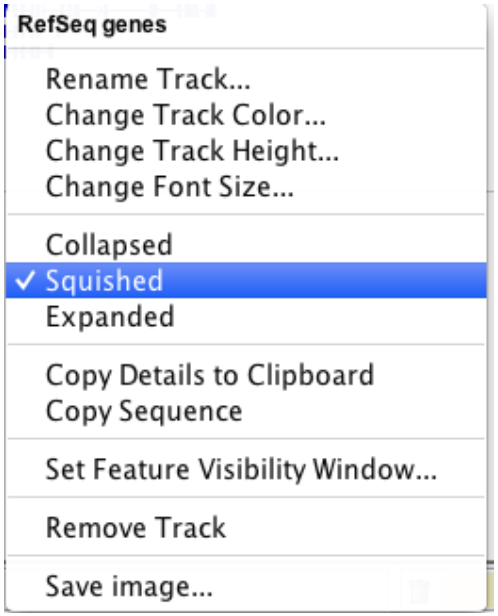

2. Expand the track using the popup menu

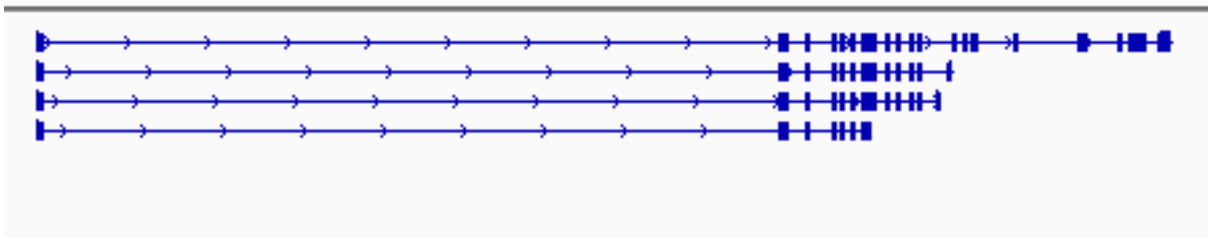


# Annotation display mode

Integrative Genomics Viewer

3. For a compact view of all variants use “Squished”



BROAD INSTITUTE

# Viewing multiple regions

Integrative Genomics Viewer

BROAD INSTITUTE

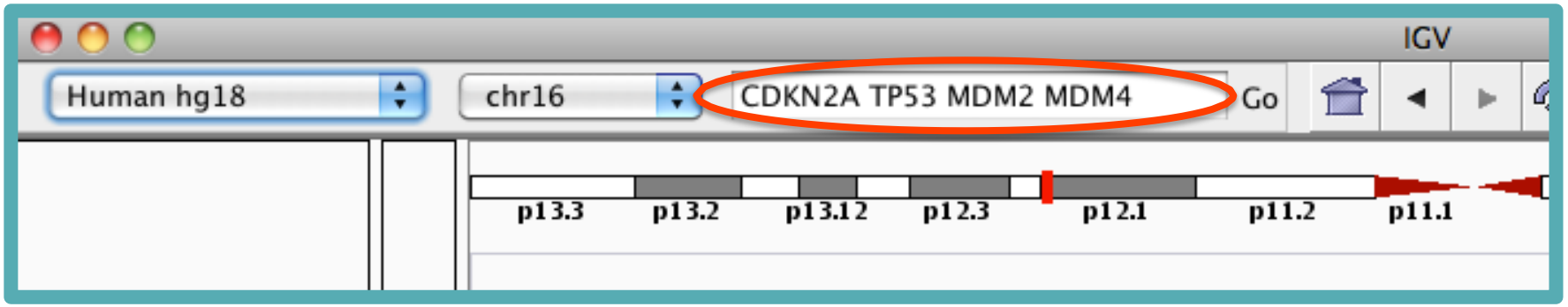
# Viewing multiple regions

# Viewing multiple regions

Integrative Genomics Viewer

- **Search box**

  Enter multiple loci or features in the search box




- **Regions > Gene Lists…**

  Select from a number of pre-defined gene lists, or

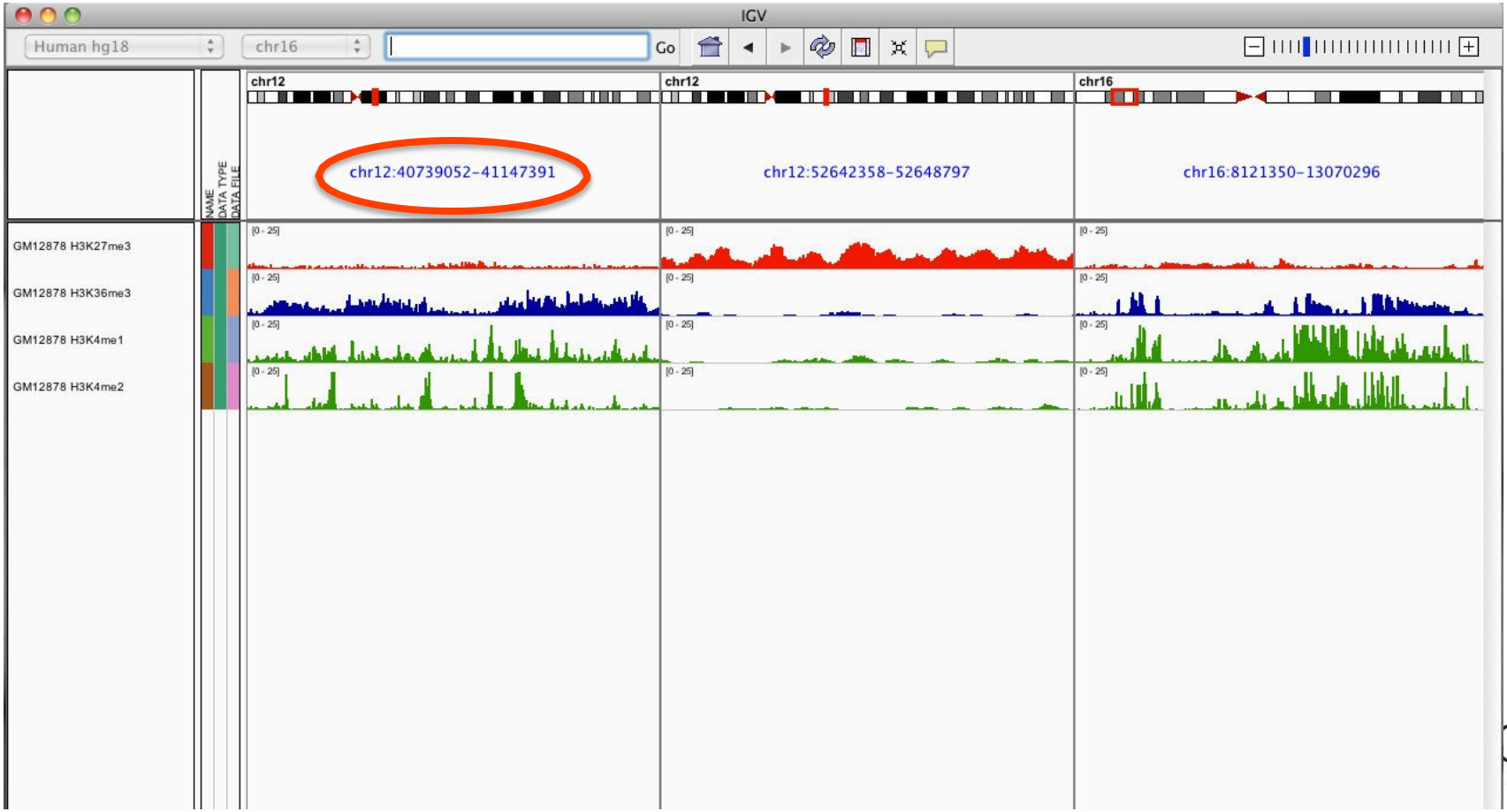
  Create your own persistent list

BROAD INSTITUTE

# Viewing multiple regions

To go back to the standard, single-region view:

- *double-click* on a region label – or –
- *right-click* and select “Switch to standard view”



# Viewing NGS Data

BROAD INSTITUTE

# Viewing alignments

Integrative Genomics Viewer

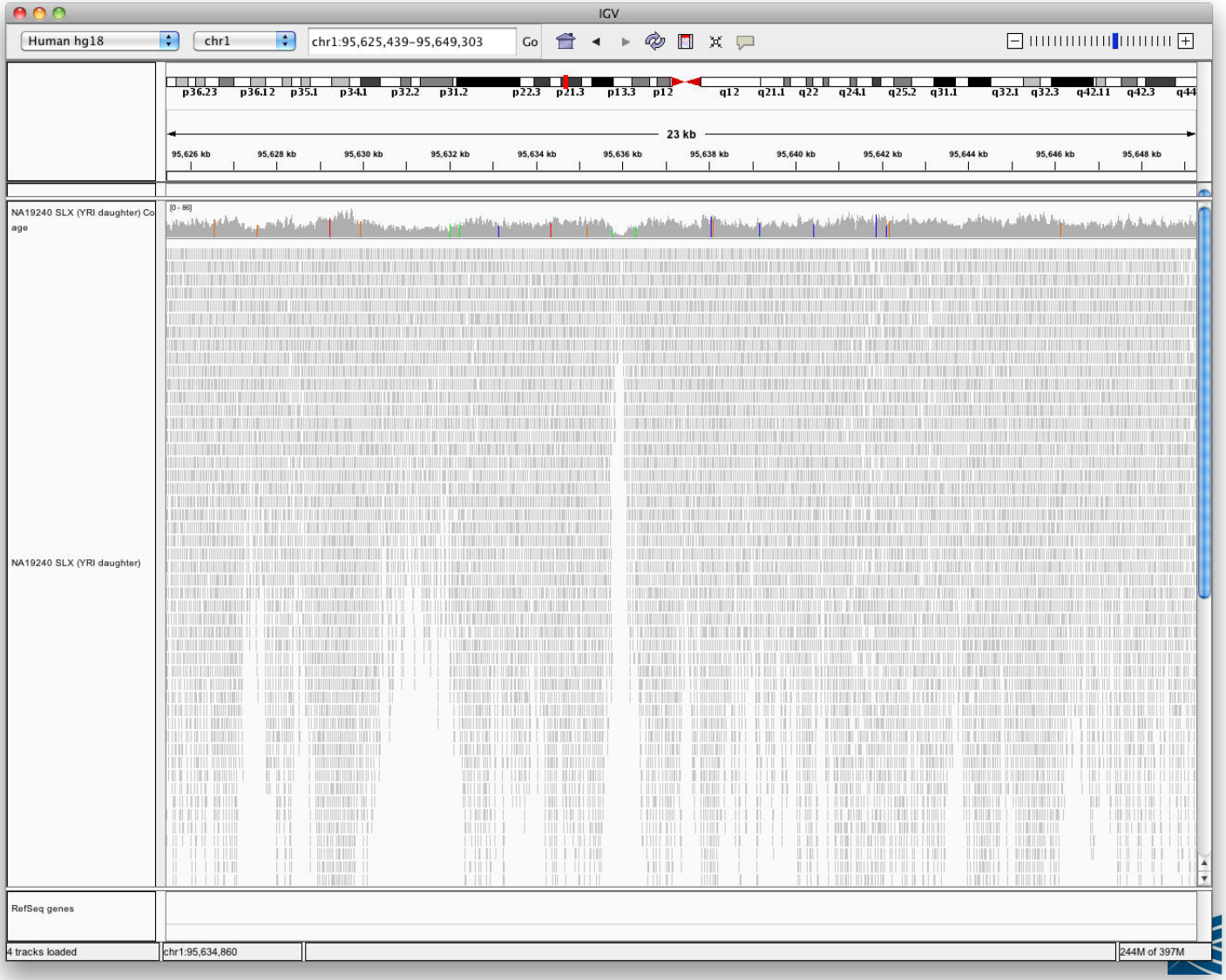
## Whole chromosome view


BROAD INSTITUTE

# Viewing alignments

Integrative Genomics Viewer

## Zoom in to view alignments

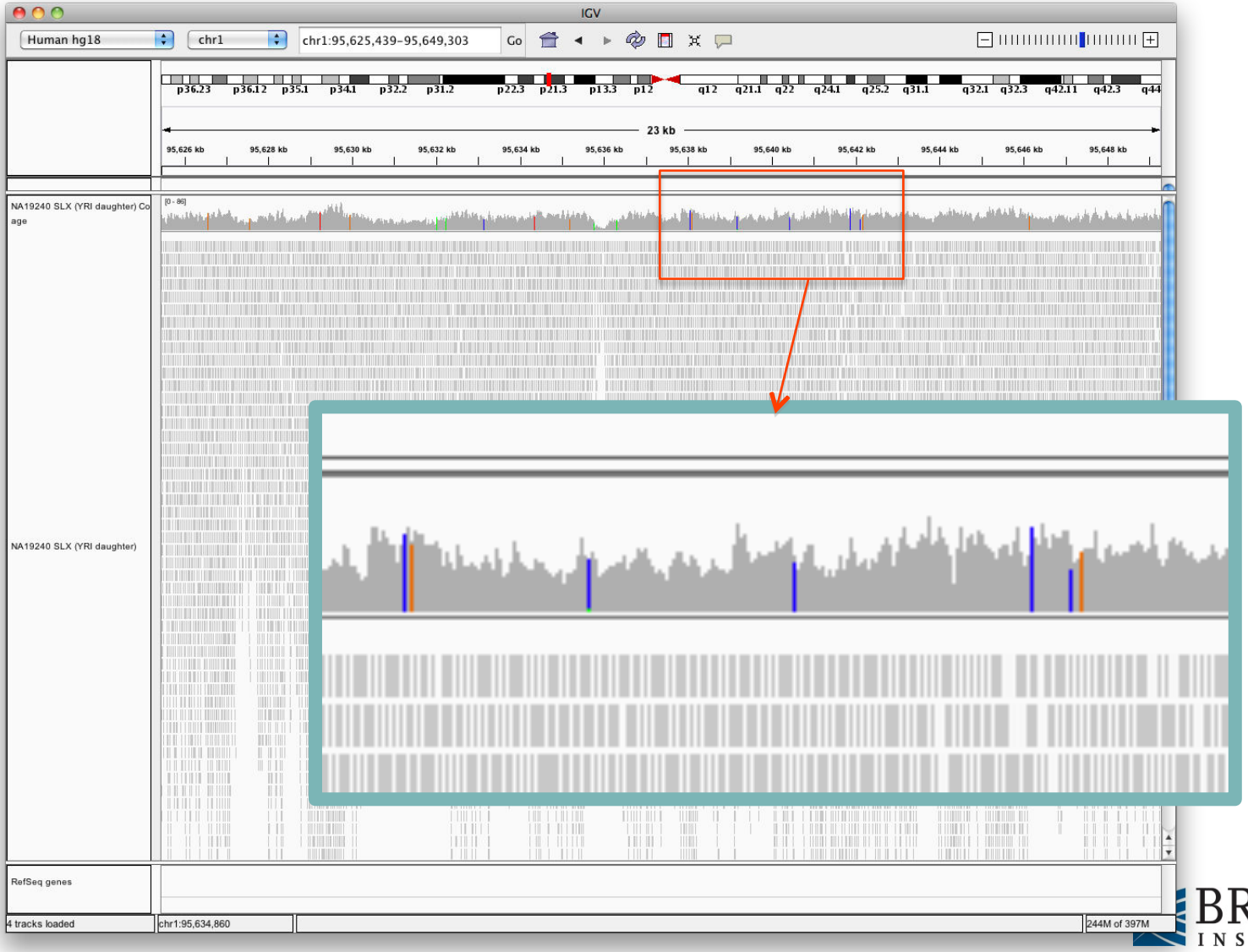



BROAD INSTITUTE

# Viewing alignments

Integrative Genomics Viewer

## Coverage track now has more detail

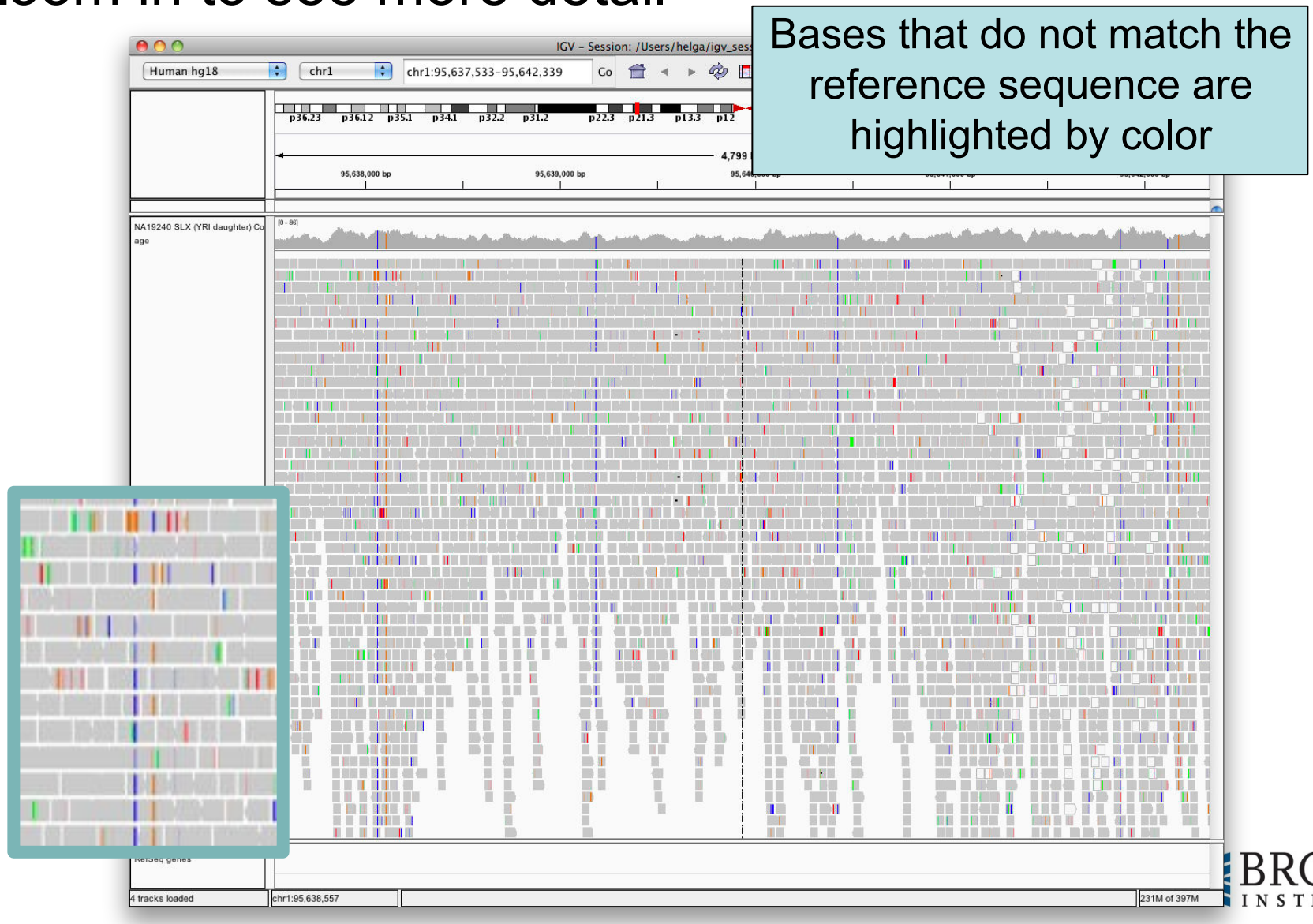



BROAD INSTITUTE

# Viewing alignments

Integrative Genomics Viewer

## Zoom in to see more detail

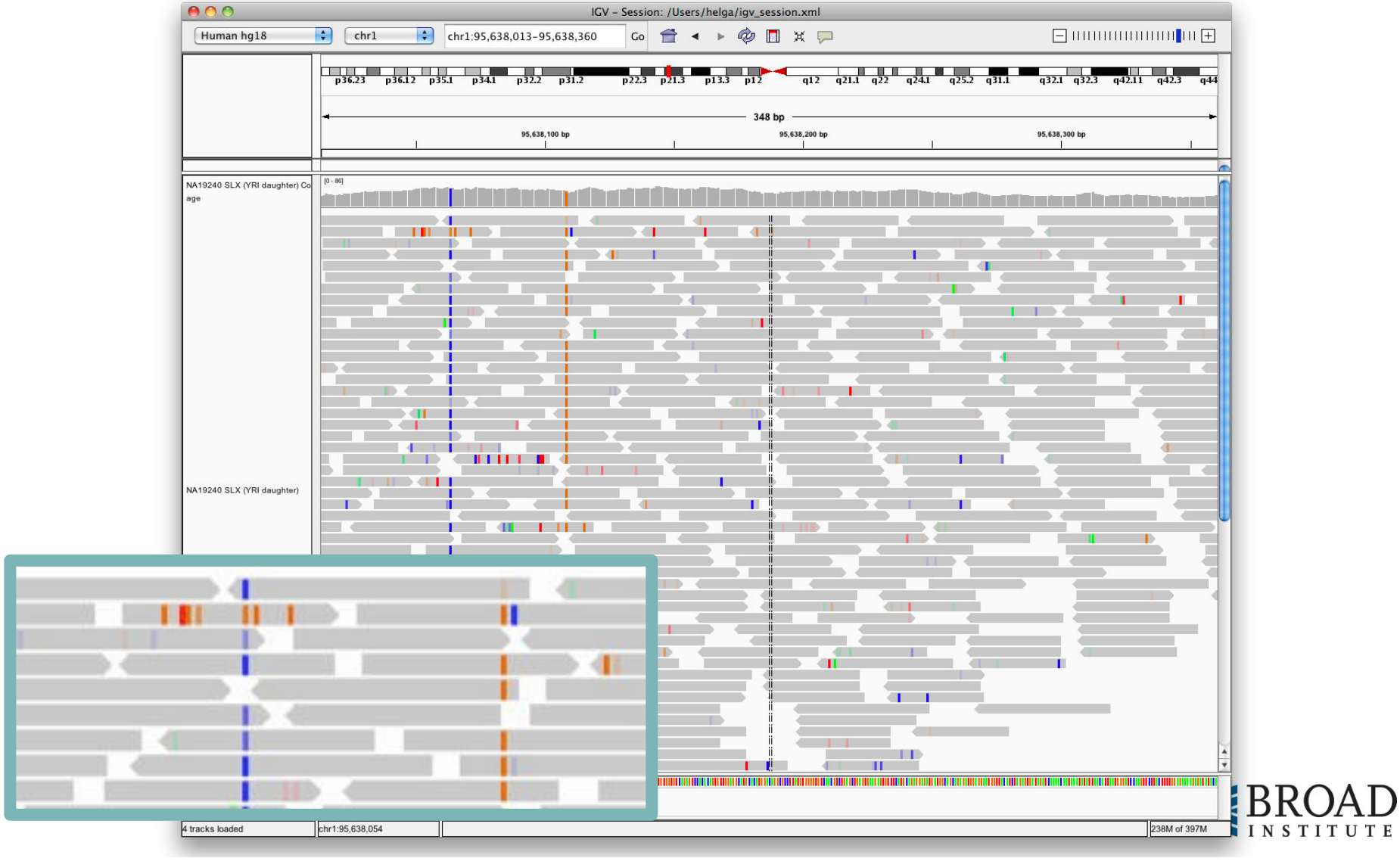



BROAD INSTITUTE

# Viewing alignments

Integrative Genomics Viewer

## Zoom in to see more detail

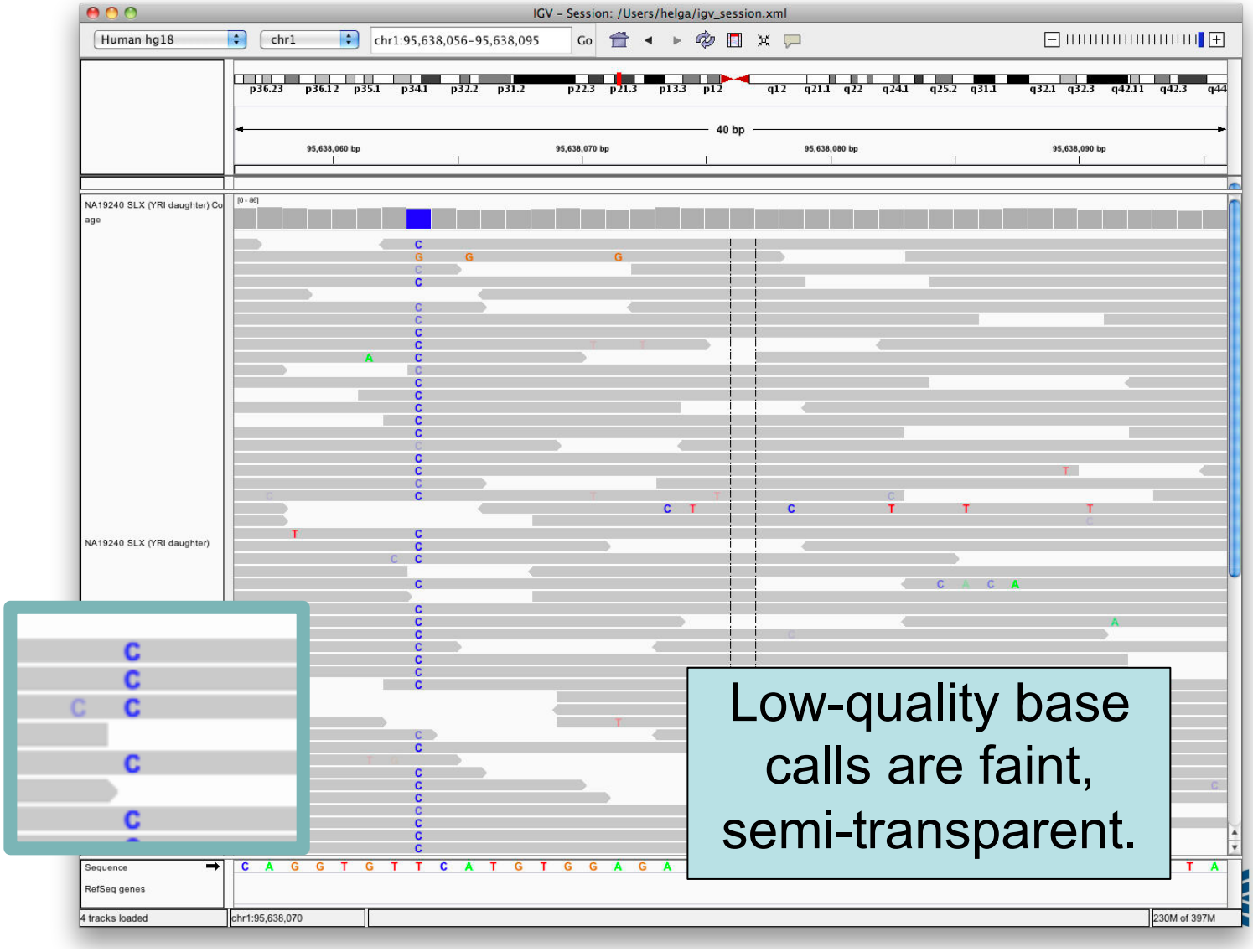



BROAD INSTITUTE

# Viewing alignments

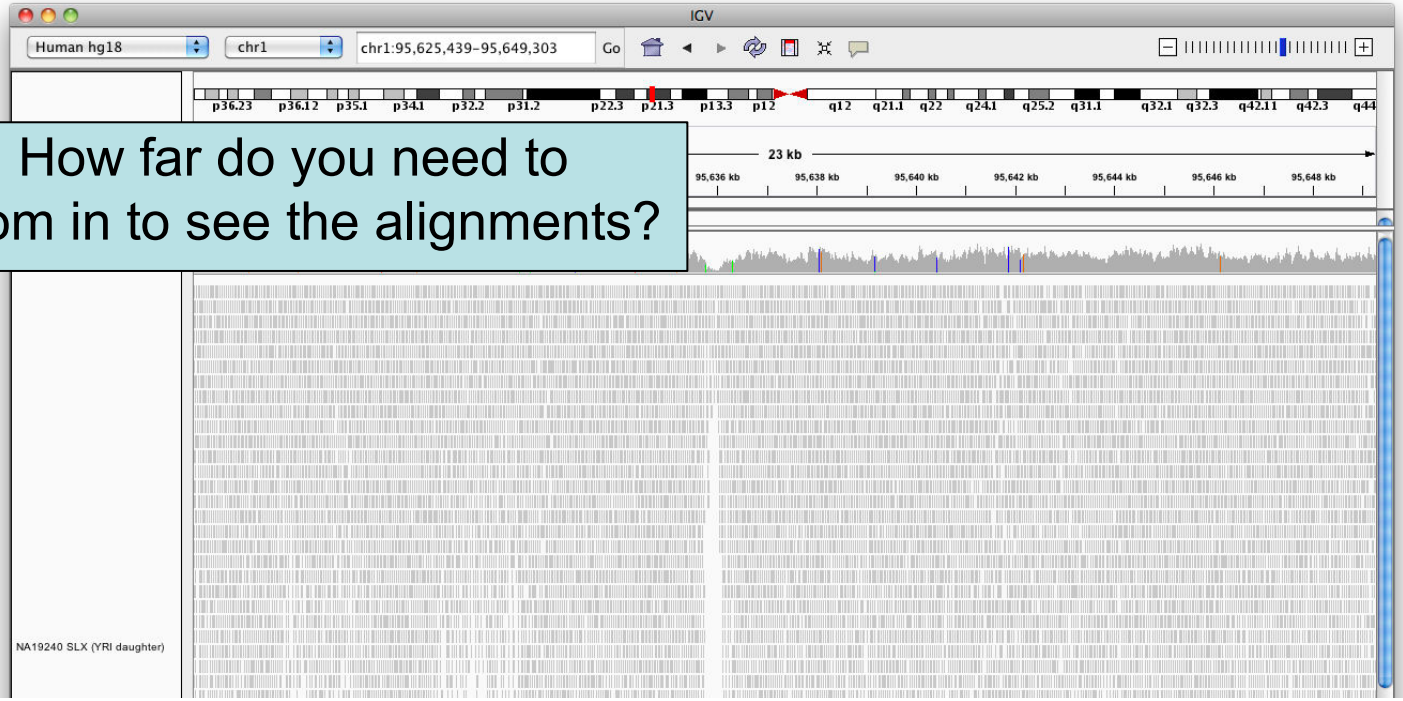
Integrative Genomics Viewer

## Zoom in to see more detail



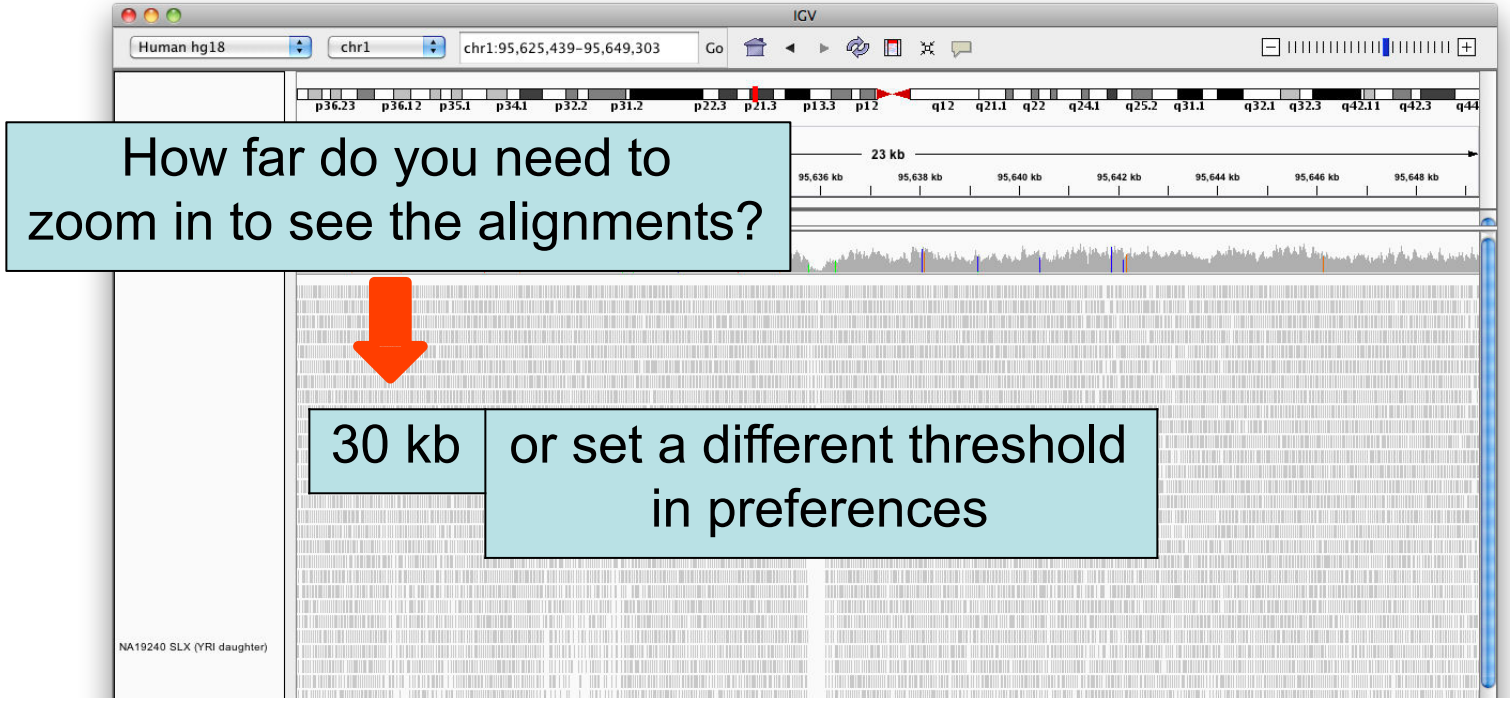

BROAD INSTITUTE

# Viewing alignments



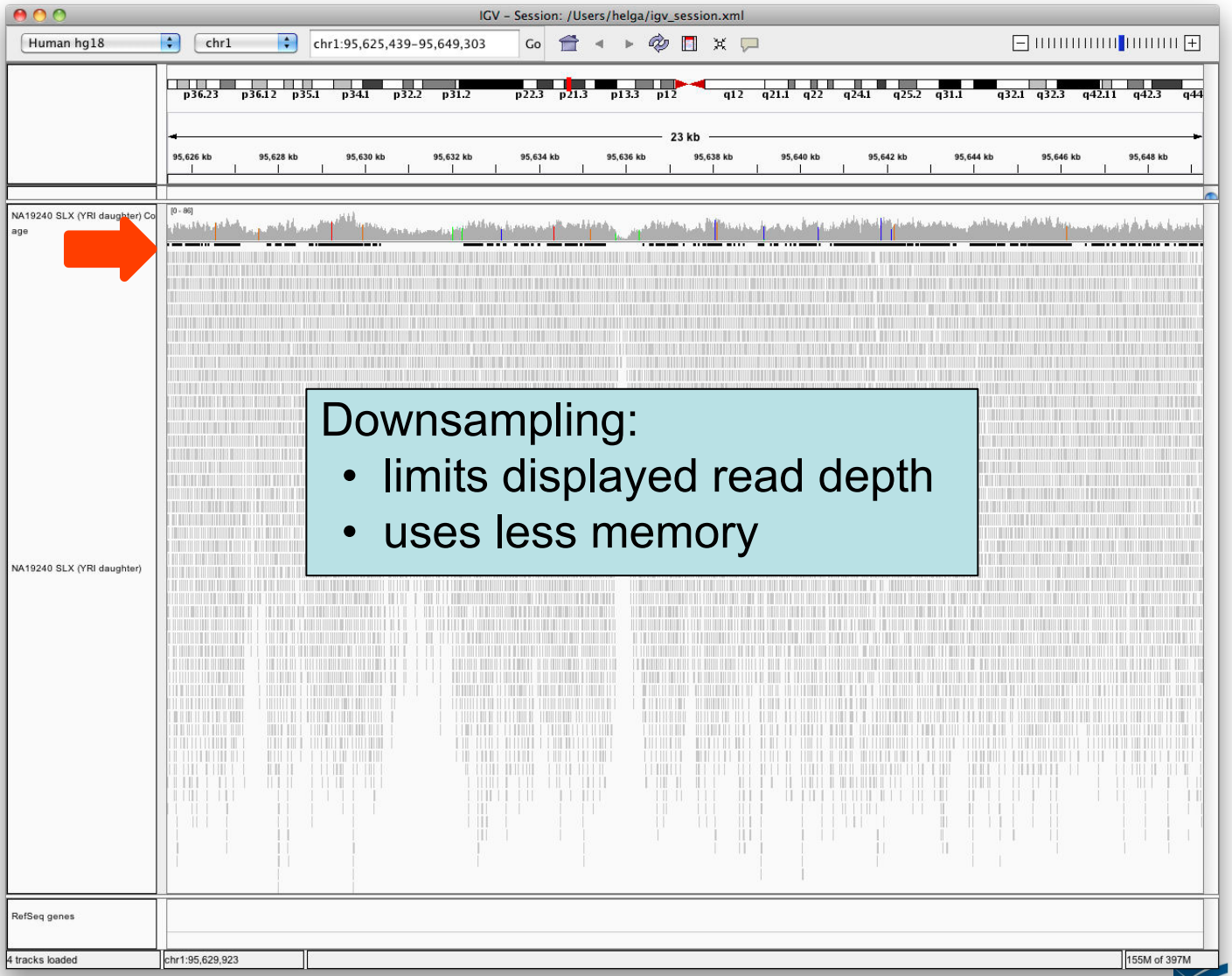

How far do you need to zoom in to see the alignments?

# Viewing alignments




- Higher value (larger region) ➔ requires more memory
- Low coverage files ➔ ok to use higher value
- Very deep coverage files ➔ use lower value

# Viewing alignments

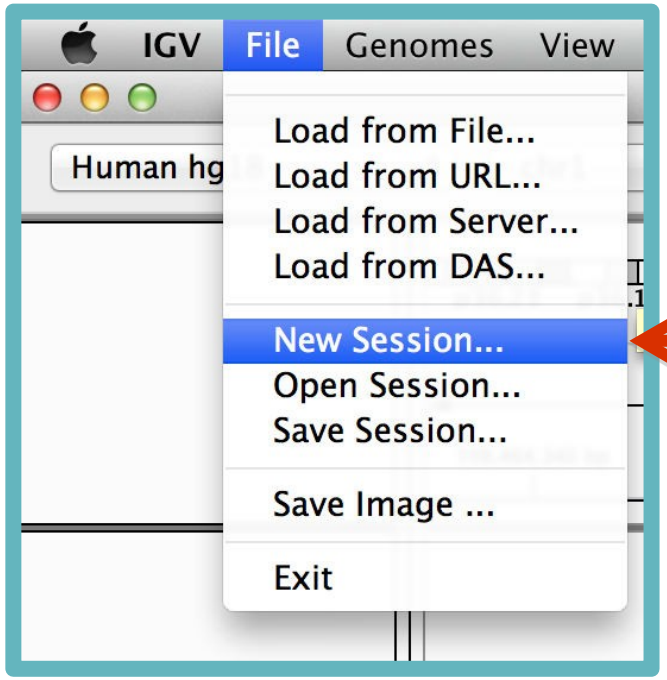


# Viewing SNPs

Hands-on exercise

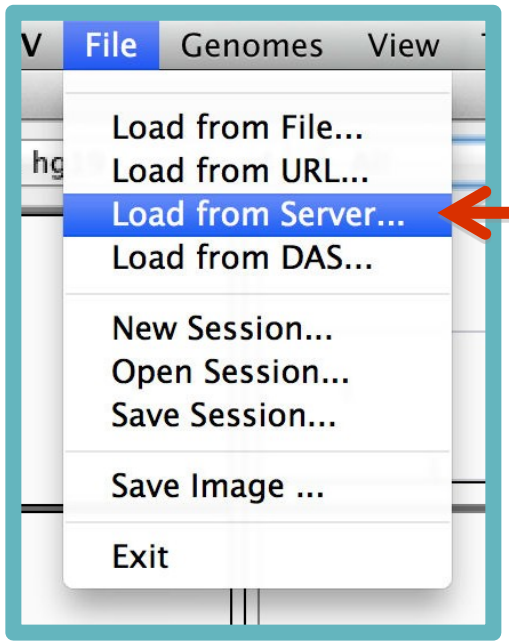
- Load alignments from whole genome sequencing
- View sites where SNPs were called
- Sort and color to highlight patterns

# Viewing SNPs



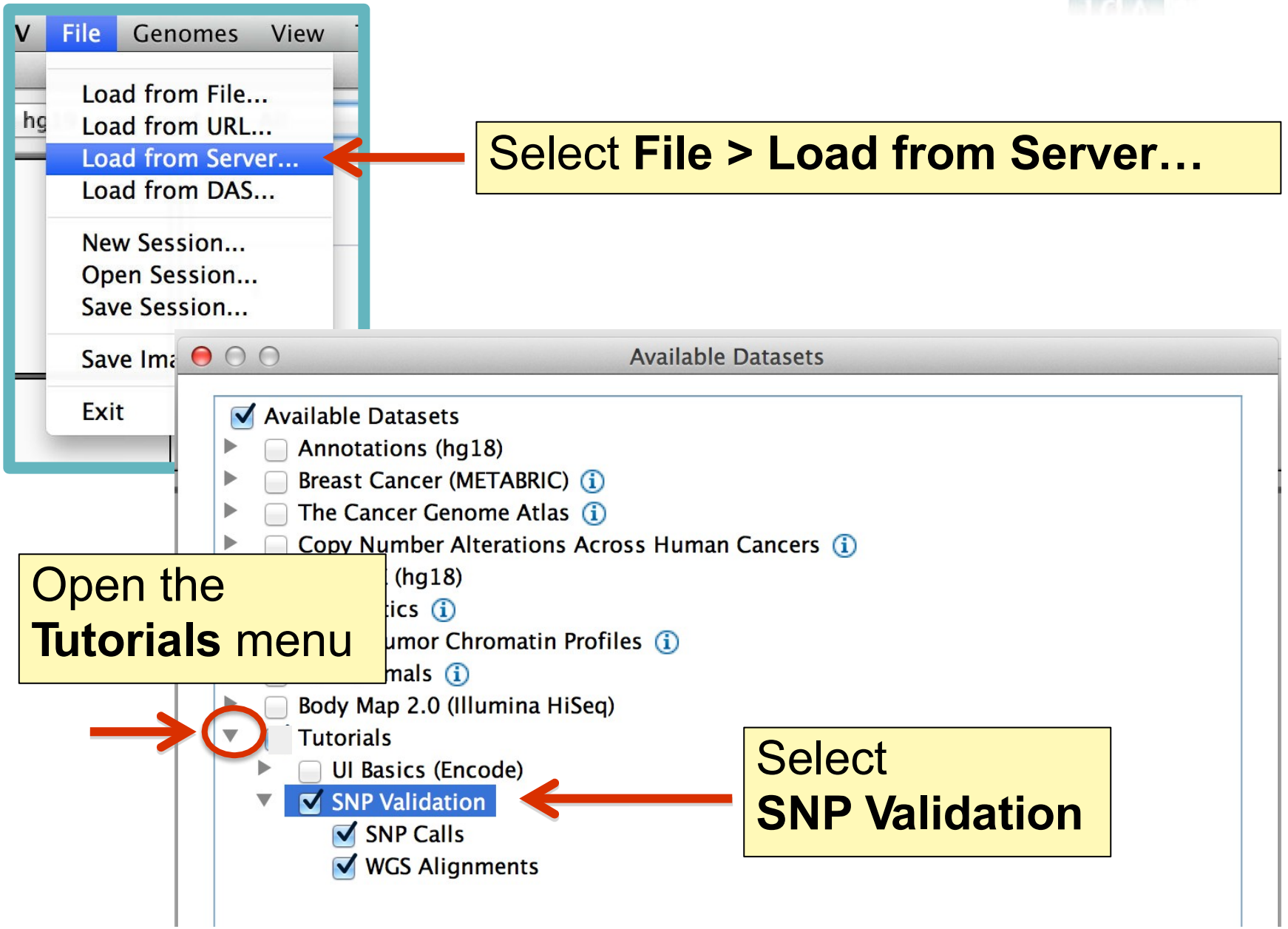

Before we start:
Select **File > New Session**
to clear IGV window

# Viewing SNPs

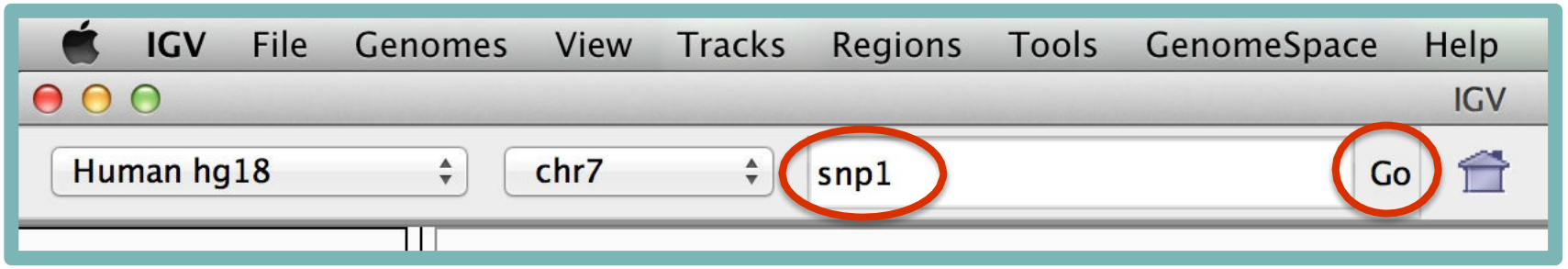



Select **File > Load from Server…**

# Viewing SNPs

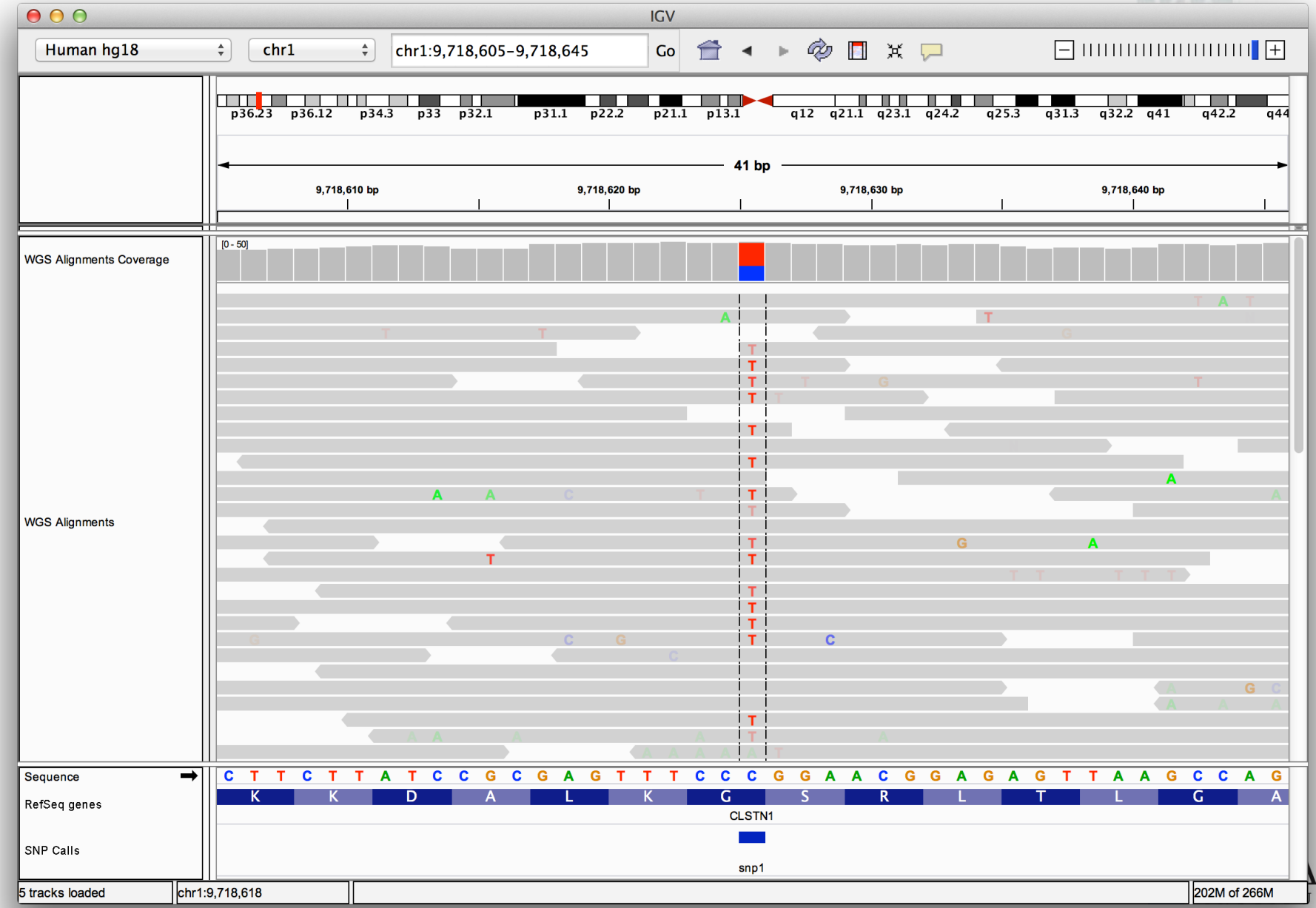


# Viewing SNPs

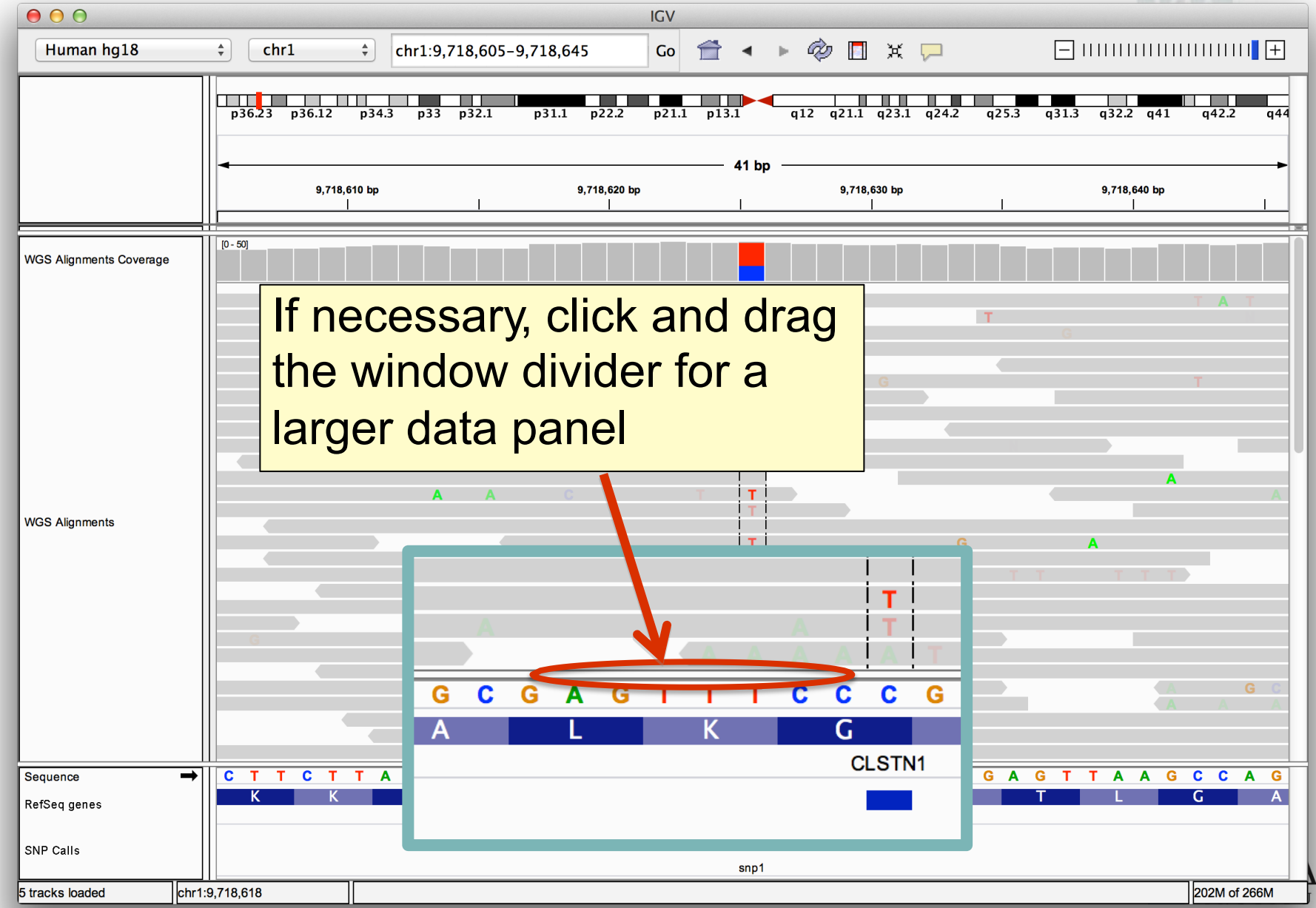



Type “snp1” in the **Search Box** and click **Go**

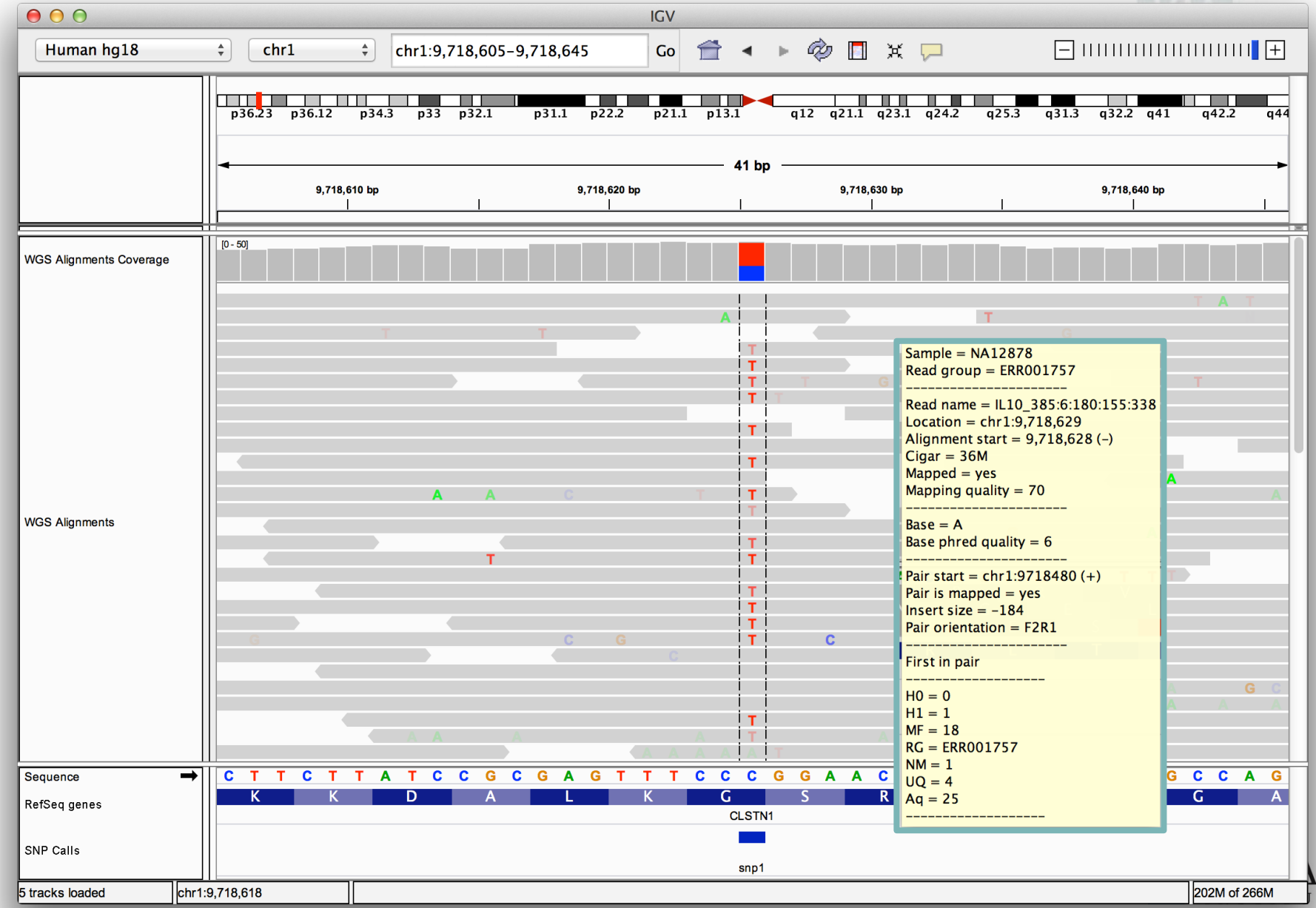
# Viewing SNPs



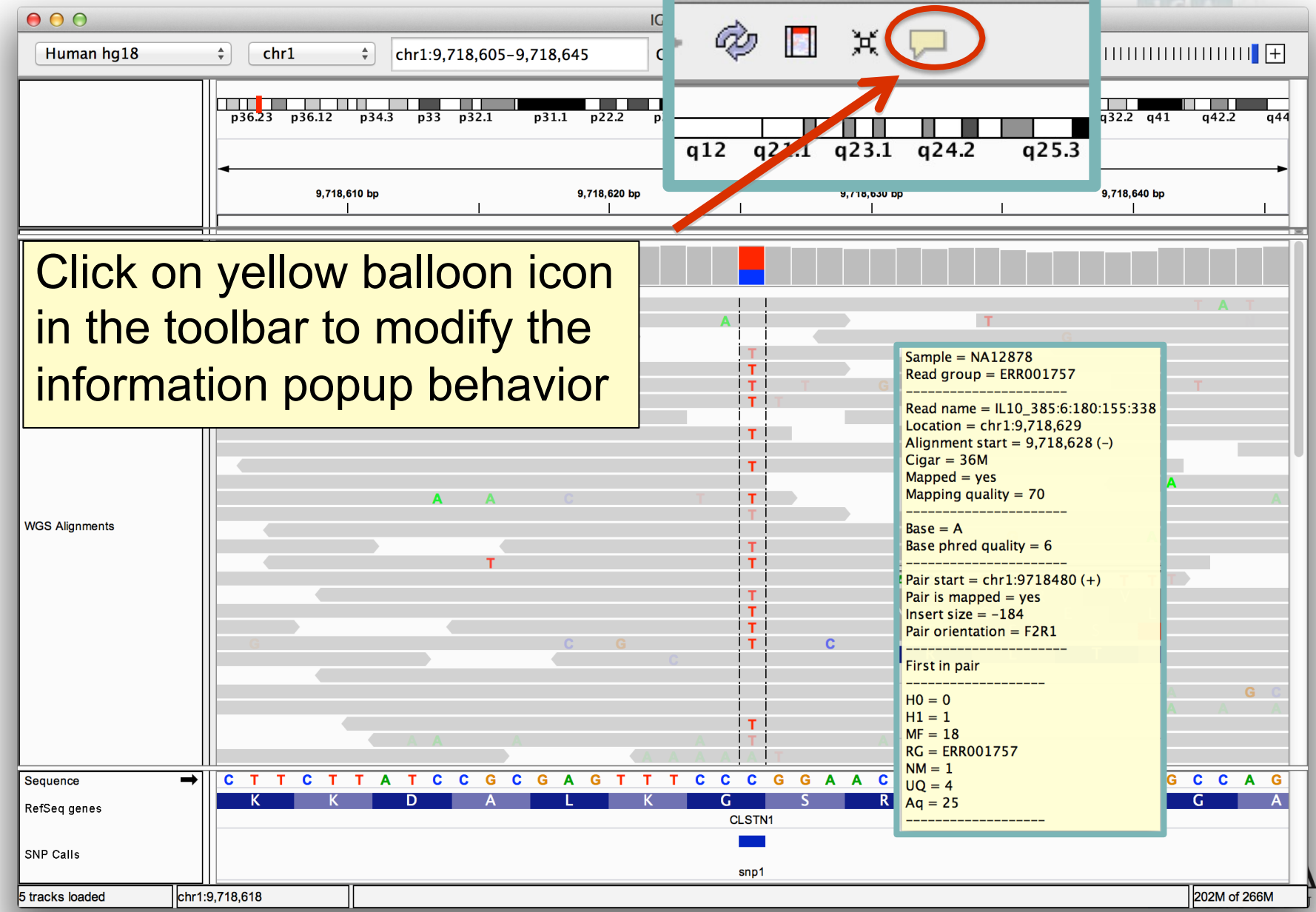
# Viewing SNPs



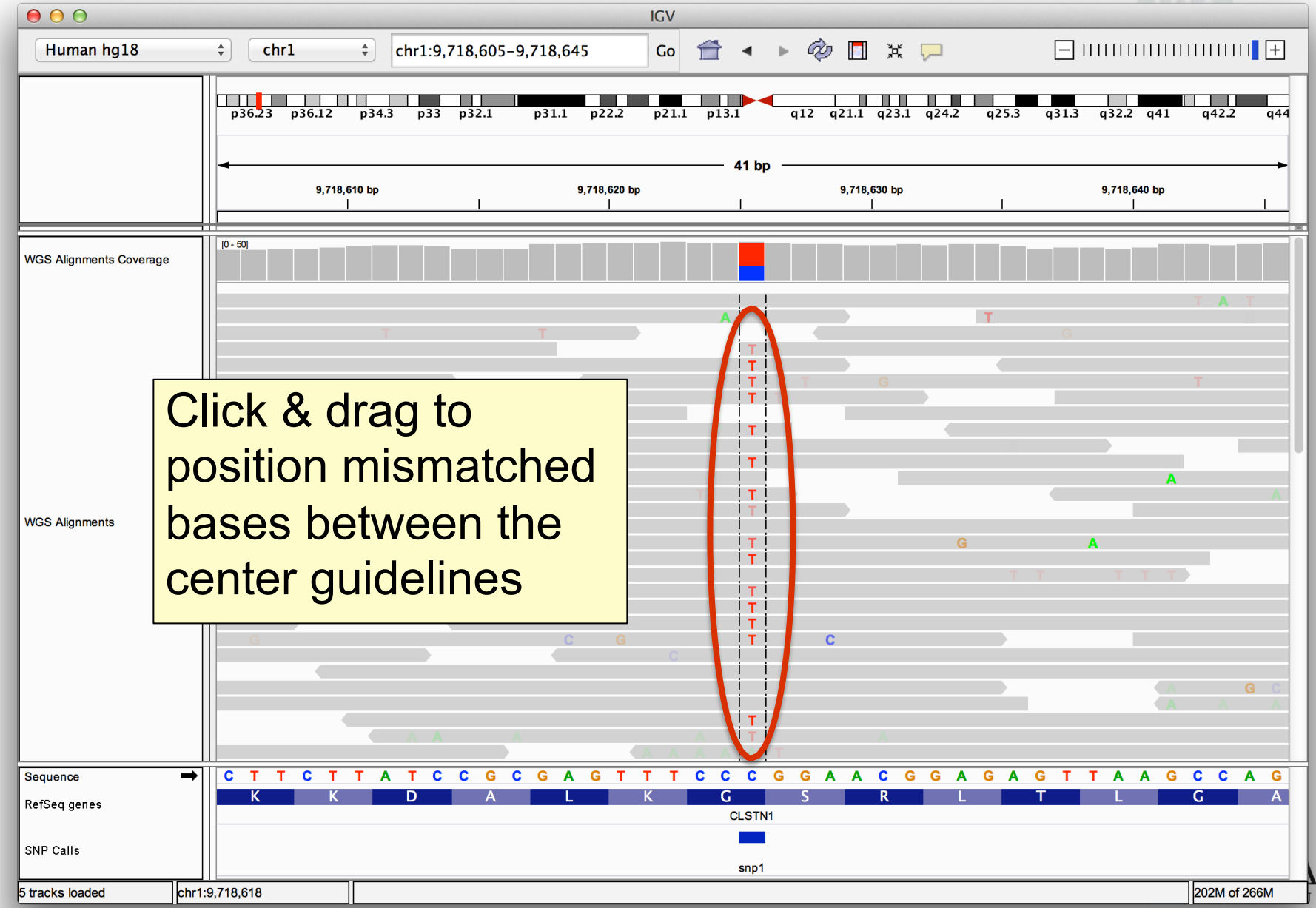
# Viewing SNPs



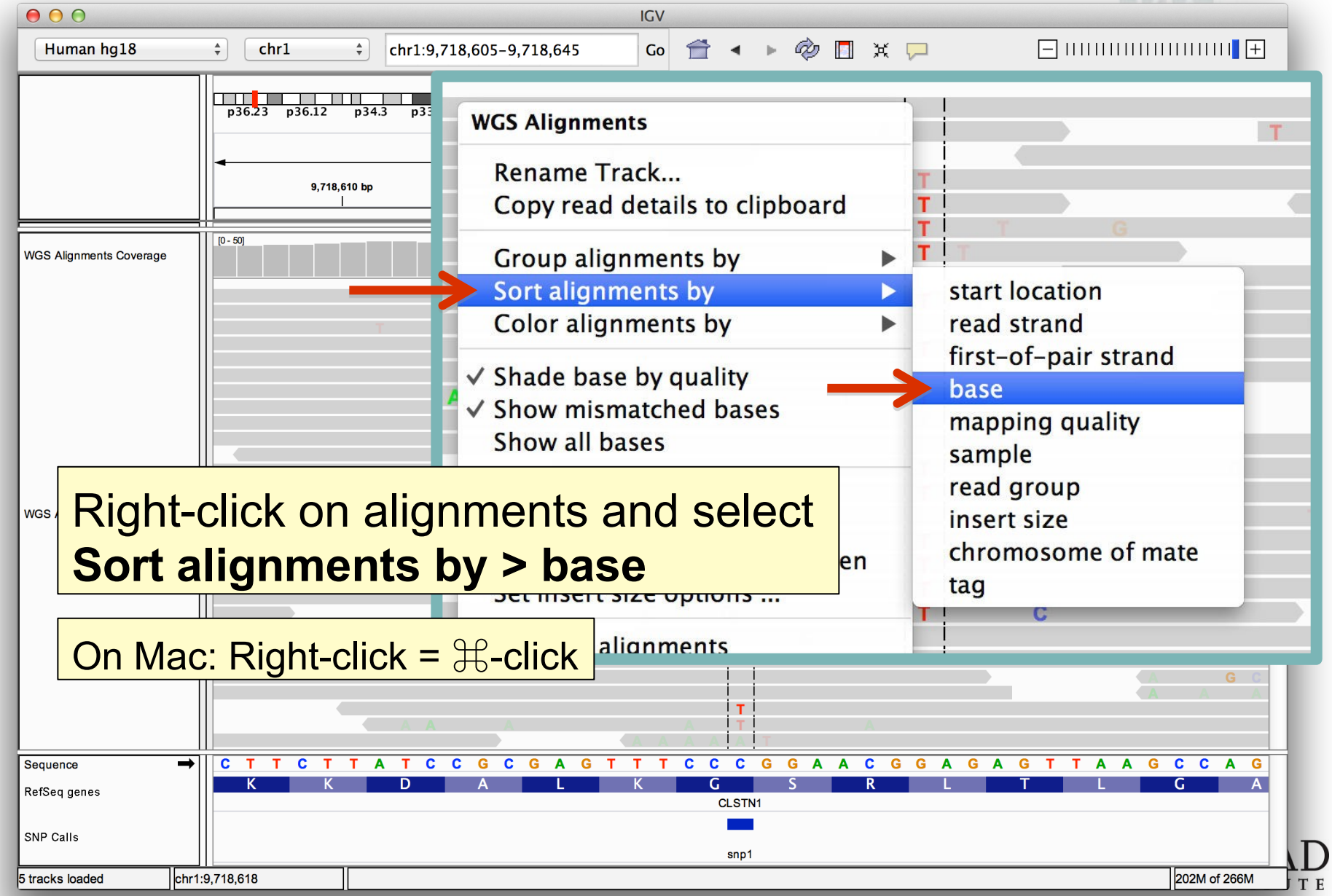
# Viewing SNPs



# Viewing SNPs



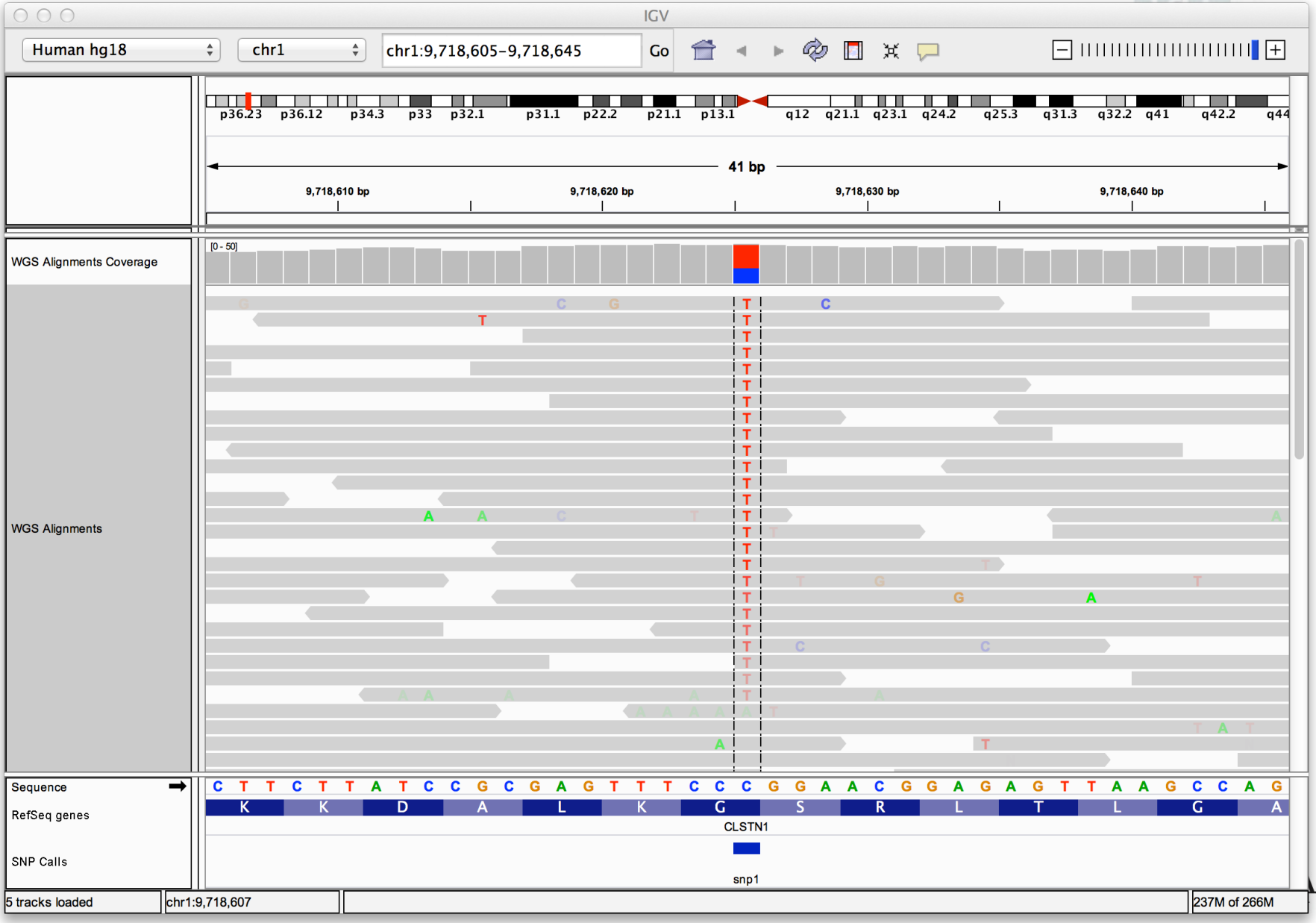
# Viewing SNPs



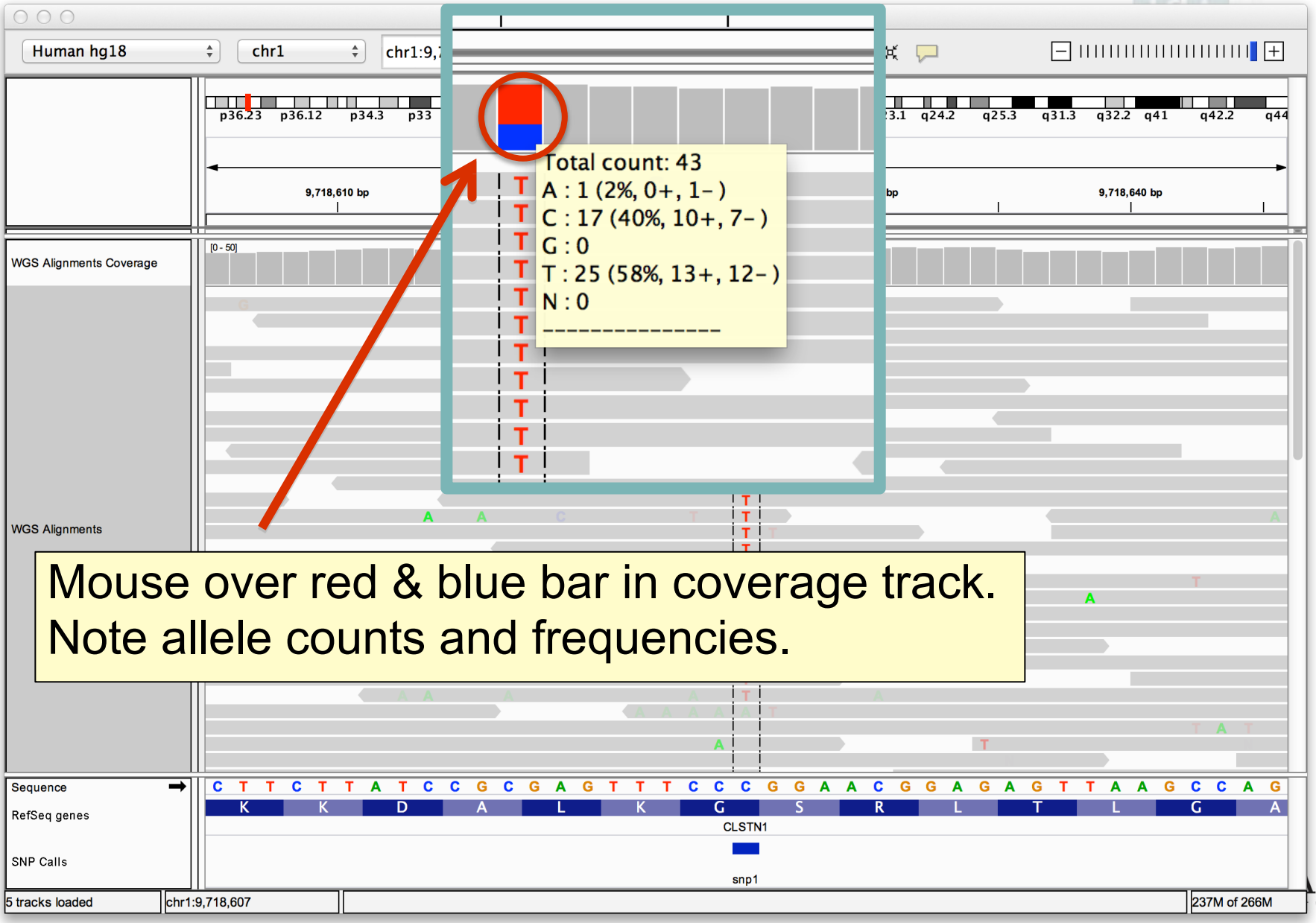

Right-click on alignments and select **Sort alignments by > base**

On Mac: Right-click = ⌘-click

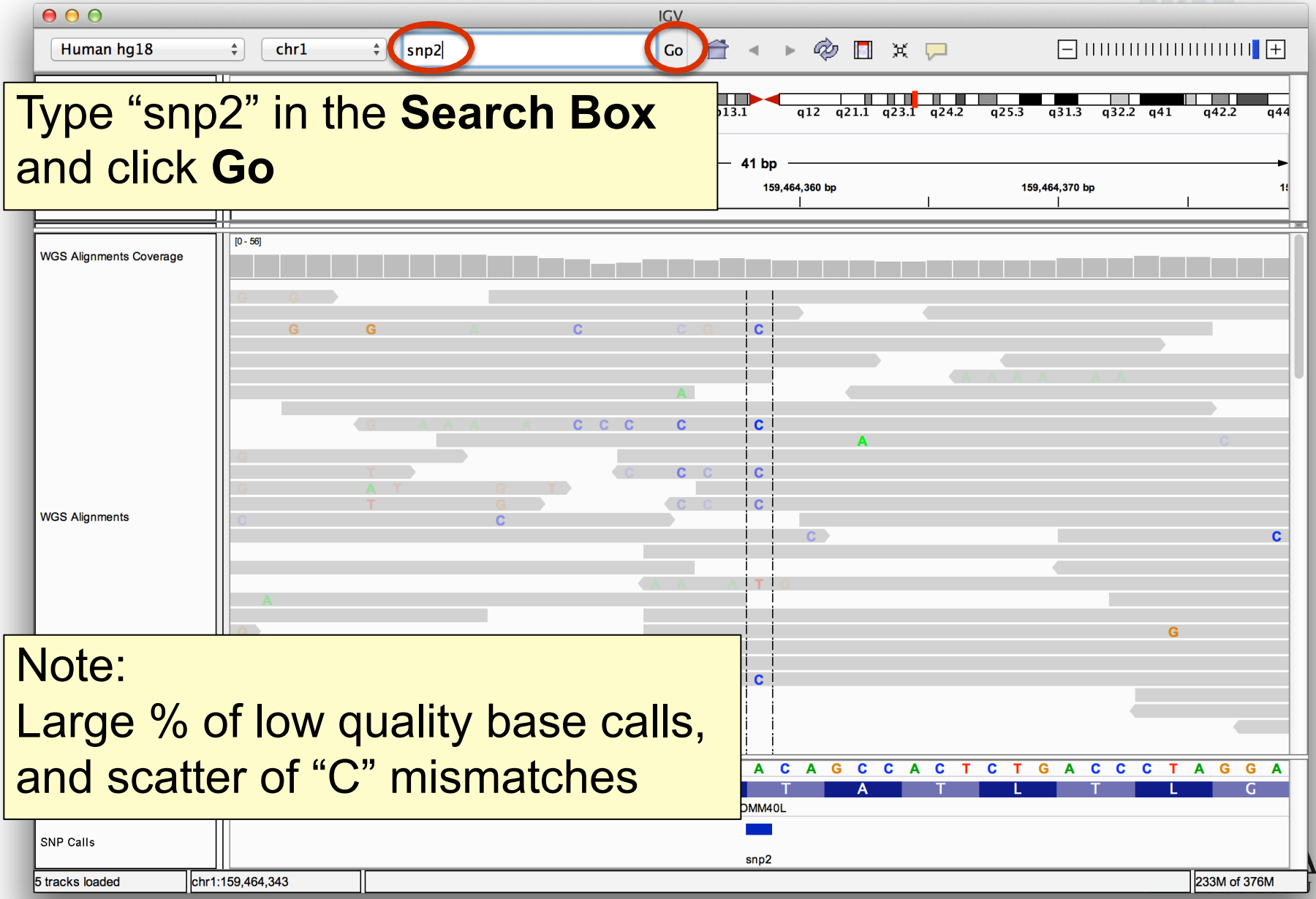
# Viewing SNPs



# Viewing SNPs



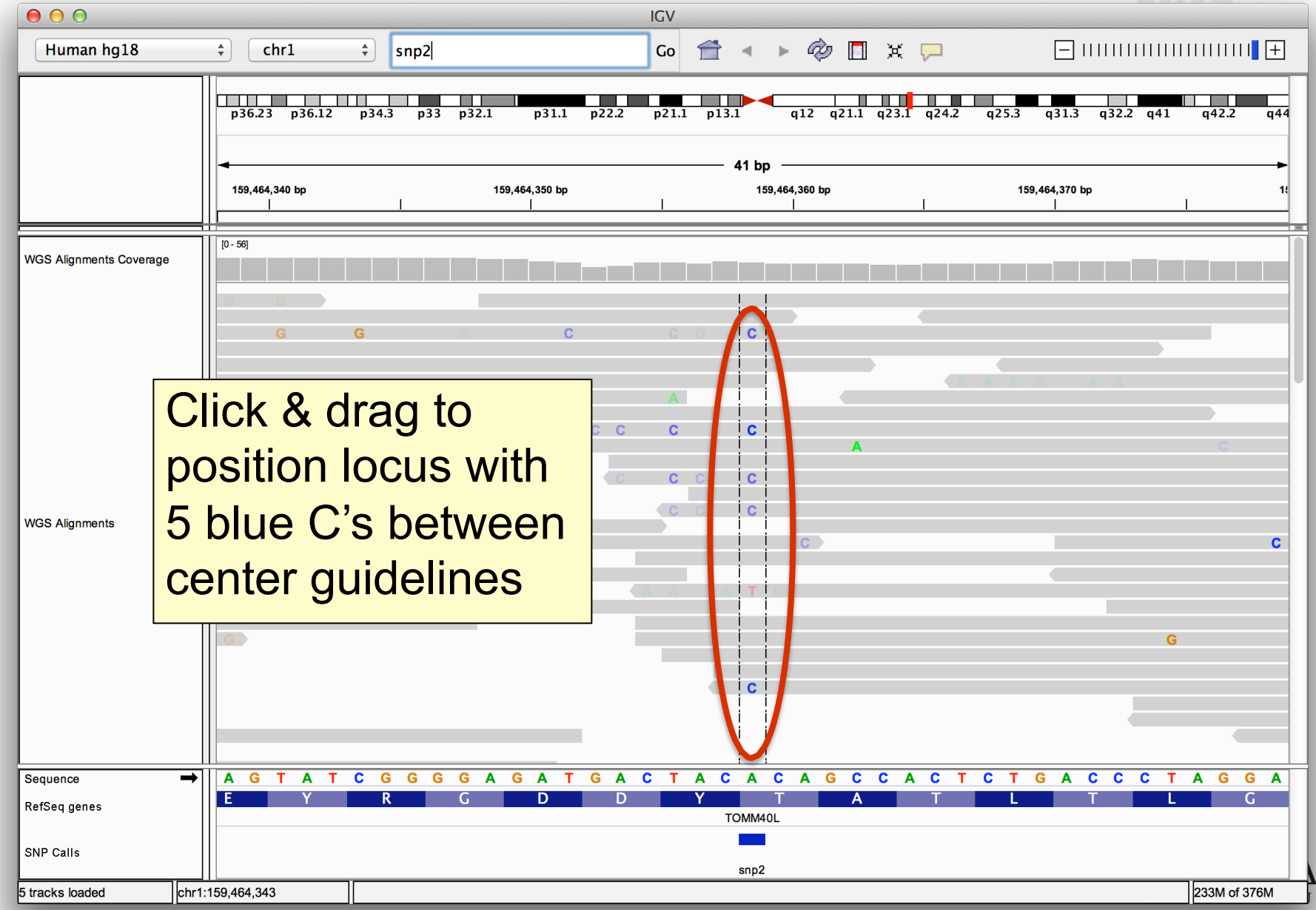
# Viewing SNPs



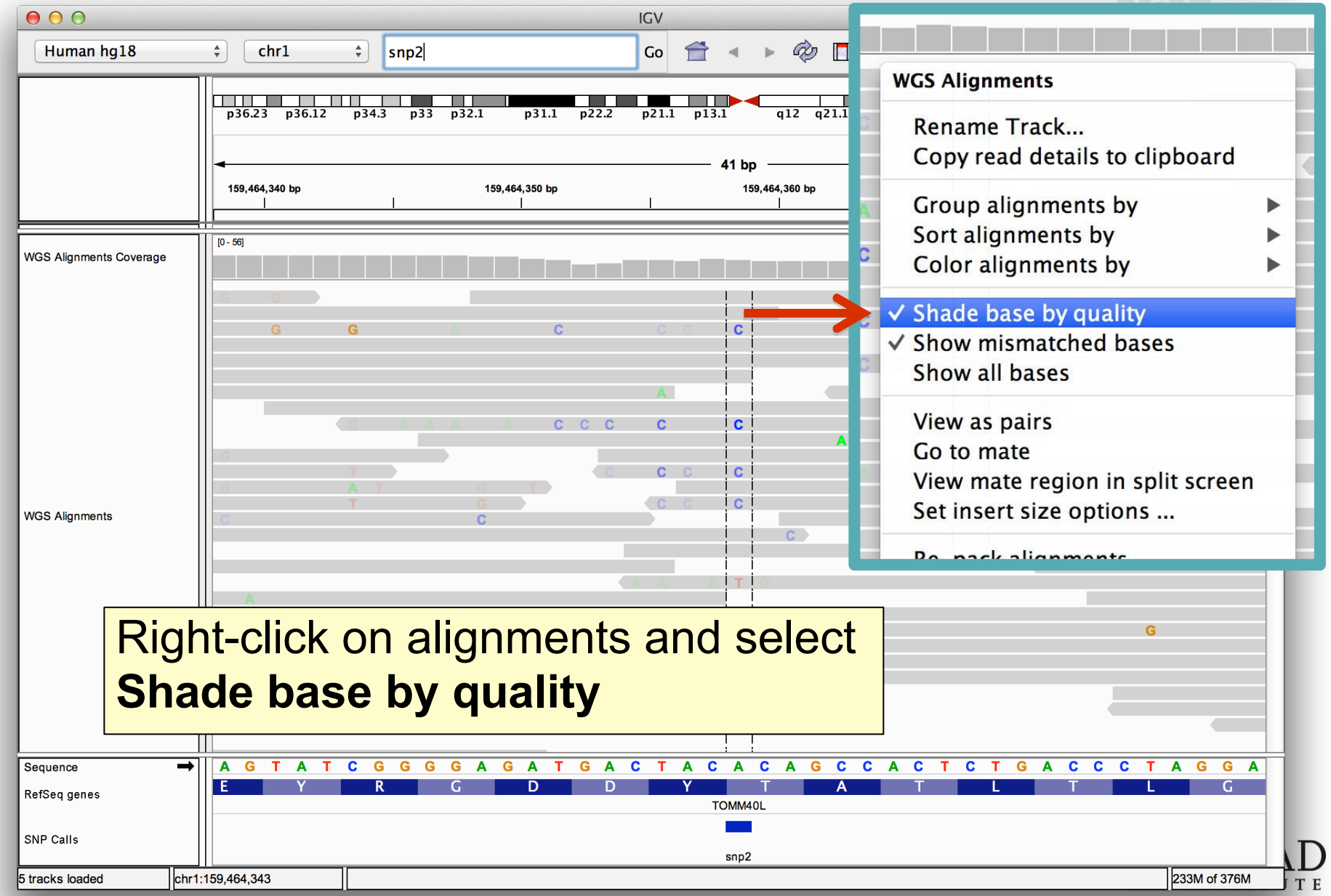

Type “snp2” in the **Search Box** and click **Go**

Note:
Large % of low quality base calls, and scatter of “C” mismatches

# Viewing SNPs

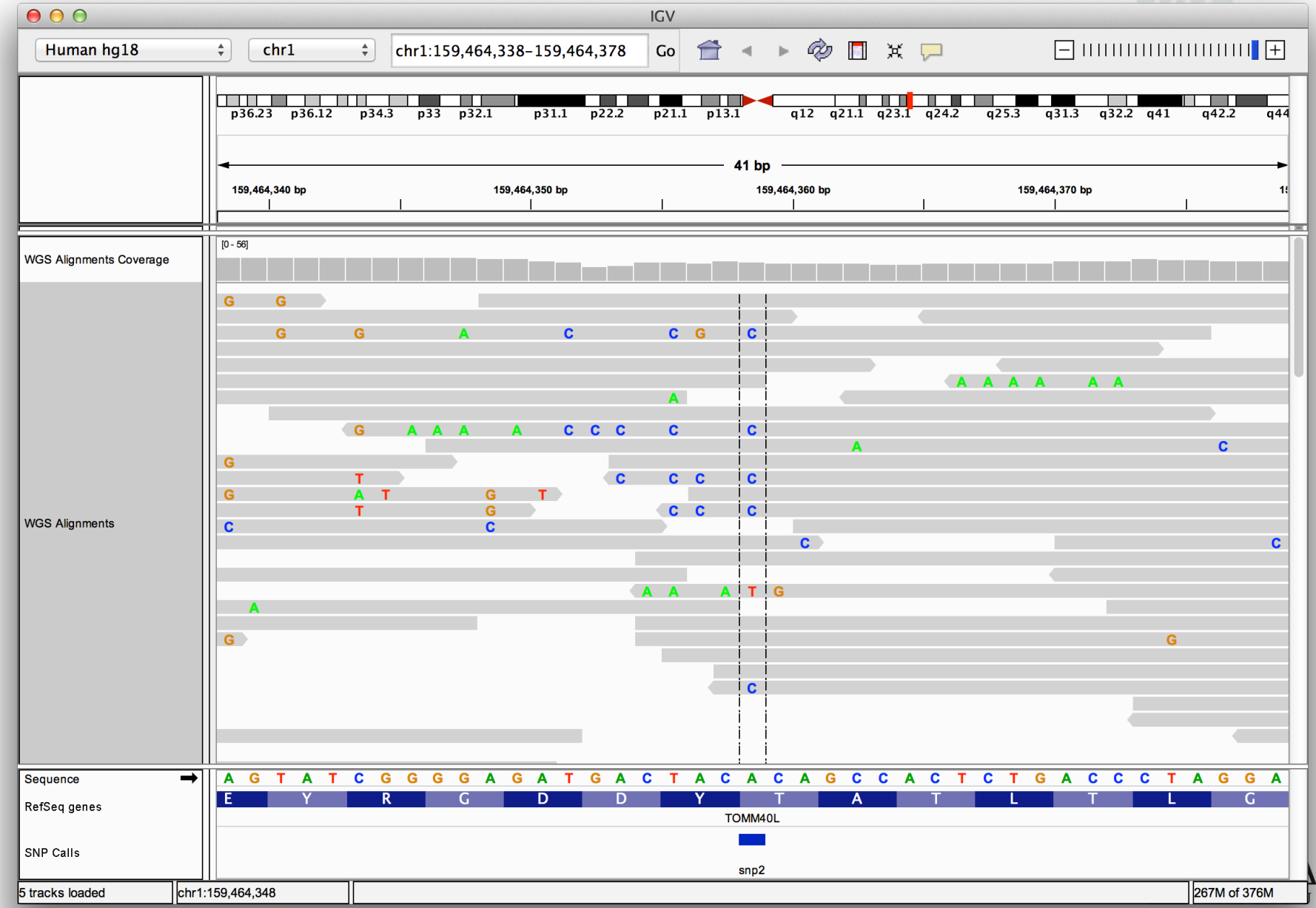


# Viewing SNPs

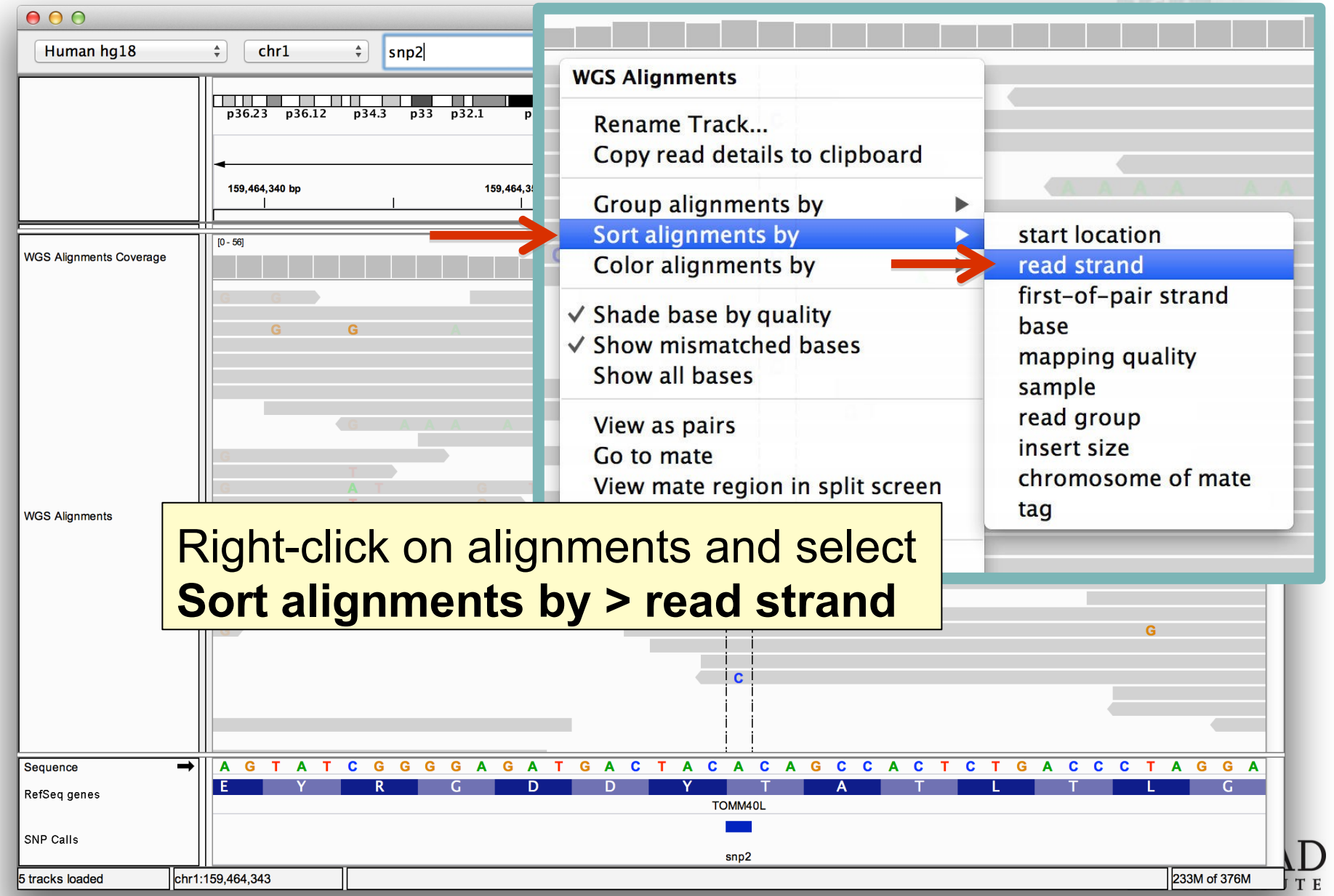



Right-click on alignments and select **Shade base by quality**

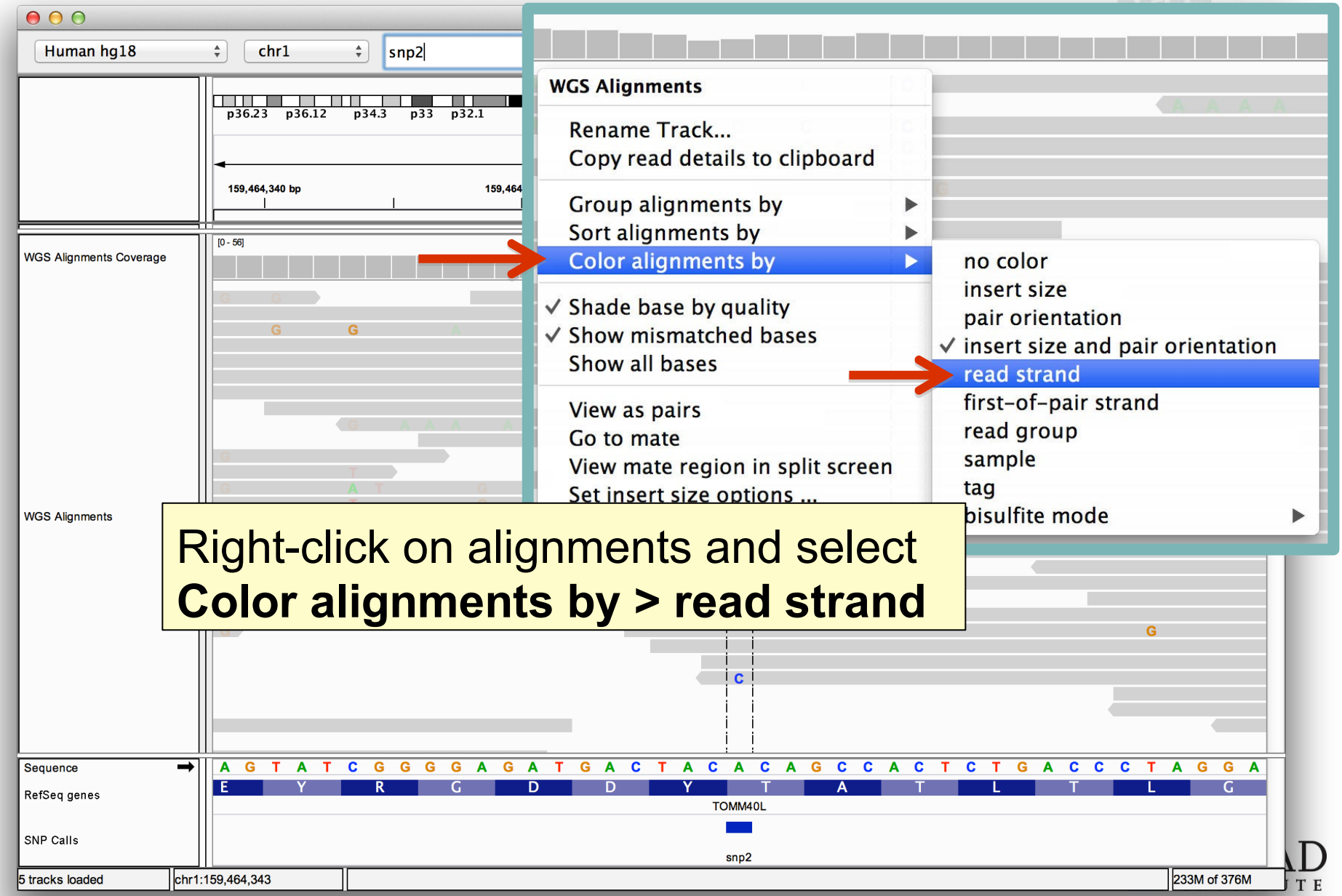
# Viewing SNPs



# Viewing SNPs

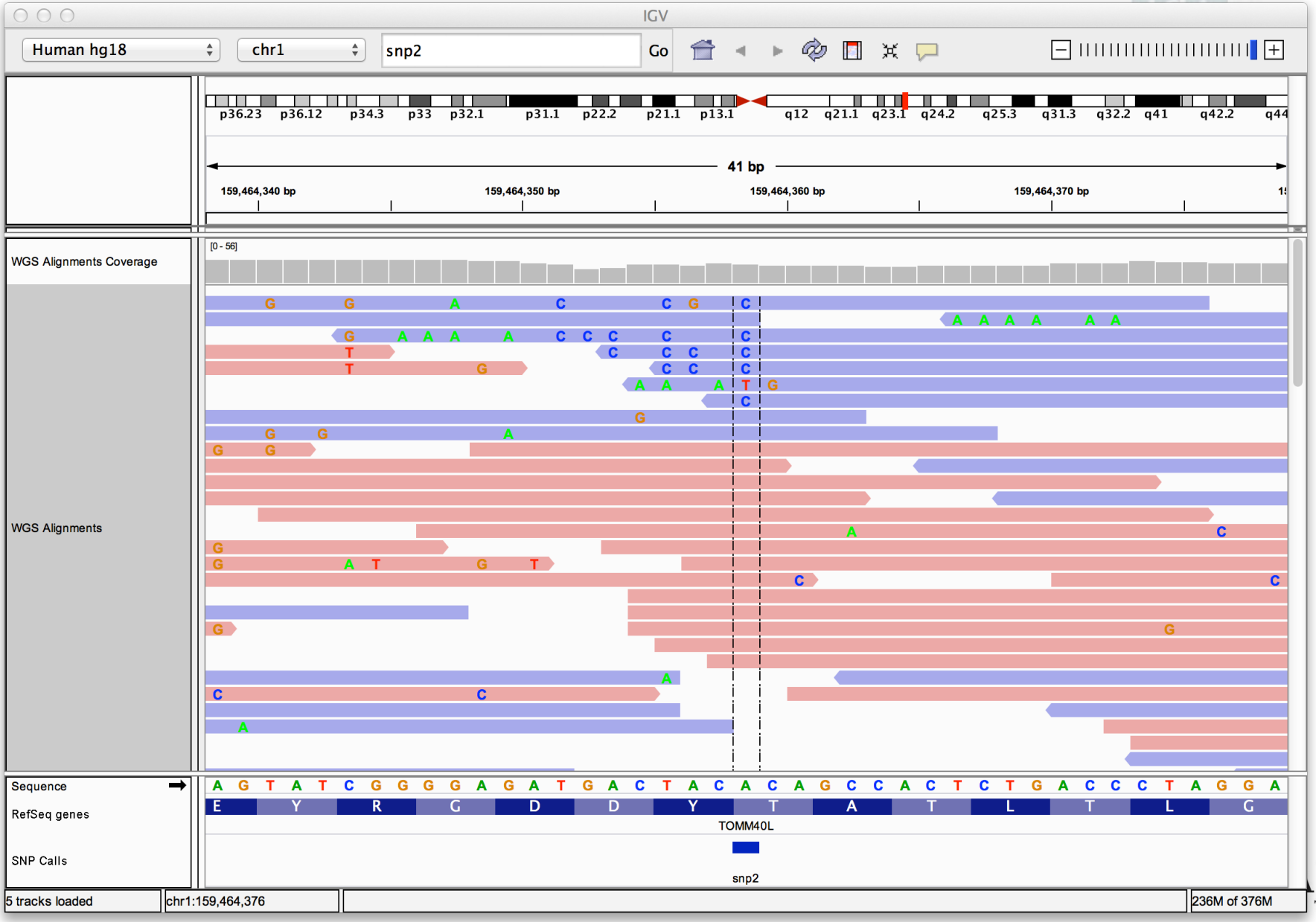


# Viewing SNPs




Right-click on alignments and select
**Color alignments by > read strand**

# Viewing SNPs



# Viewing Structural Events

BROAD INSTITUTE

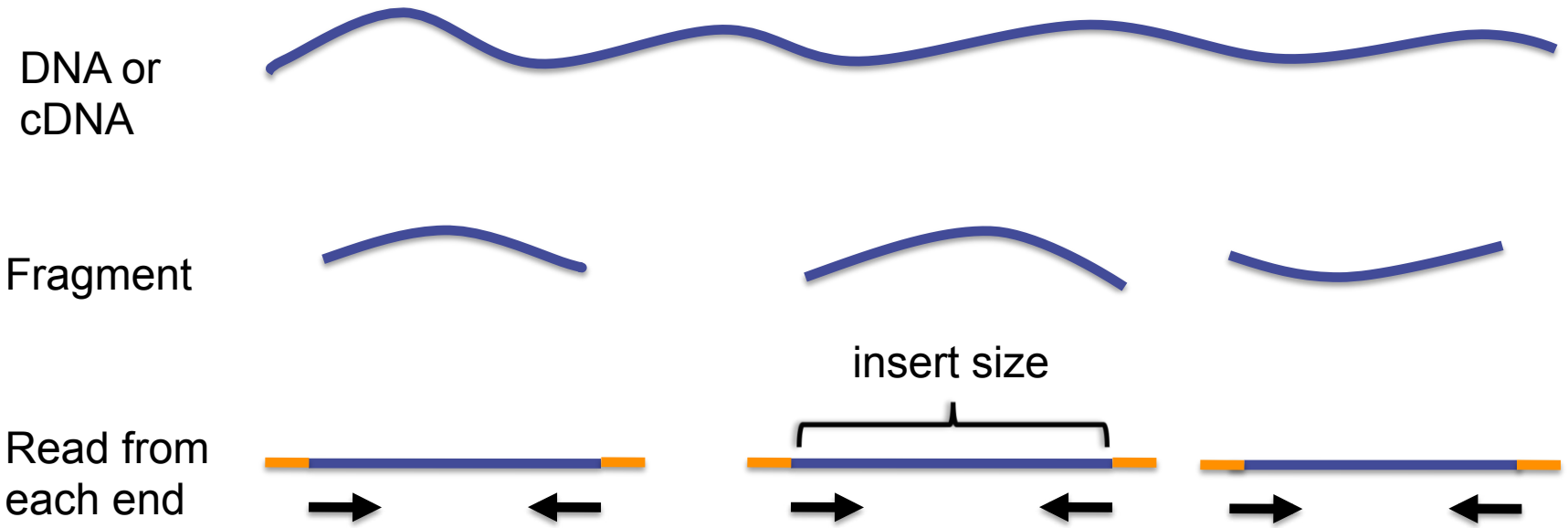
Integrative Genomics Viewer

# Structural events

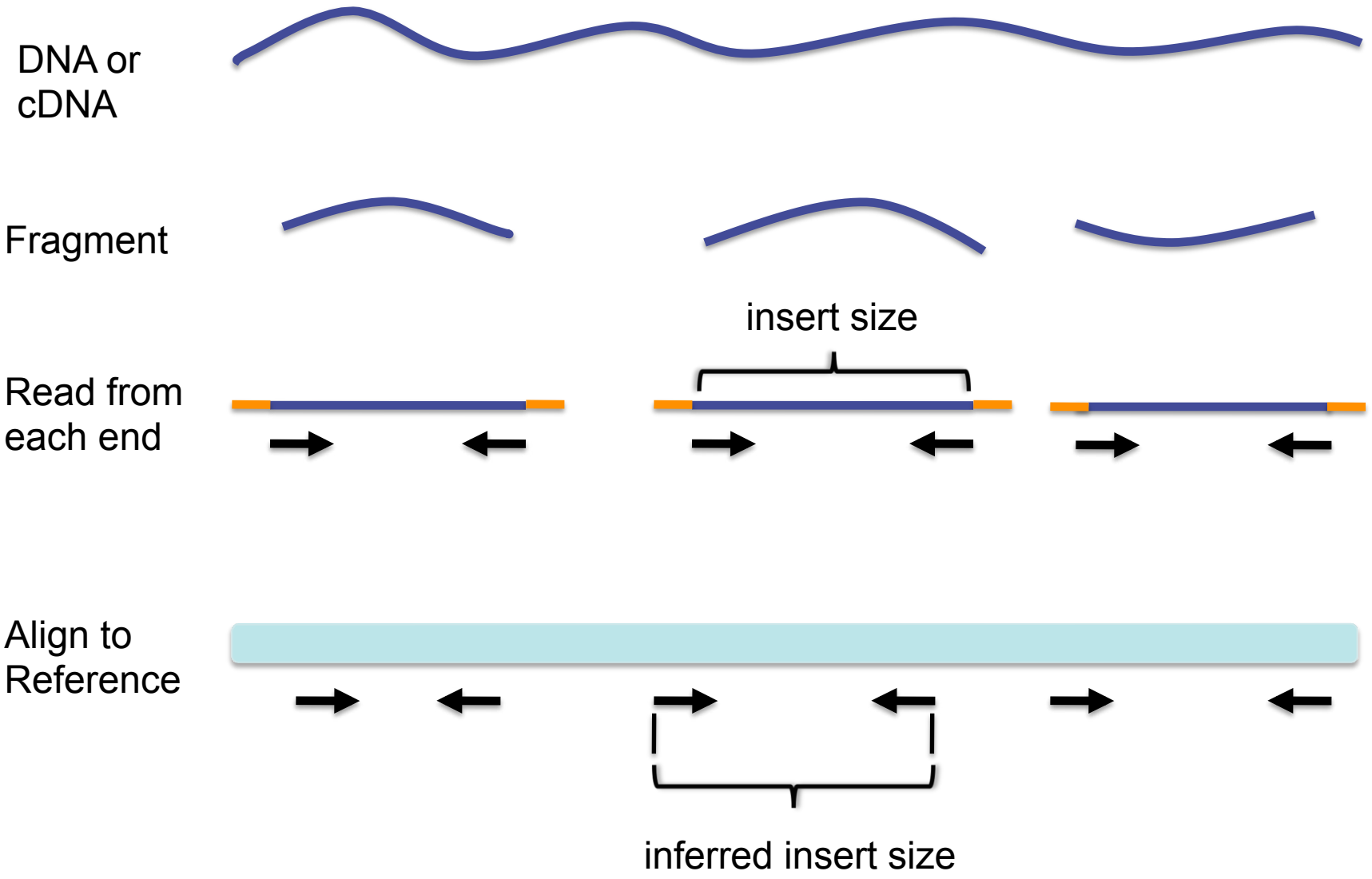

- Paired reads can yield evidence for genomic “structural events”, such as deletions, translocations, and inversions.
- Alignment coloring options help highlight these events based on:
  - Inferred insert size (template length)
  - Pair orientation (relative strand of pair)

BROAD INSTITUTE

# Paired-end sequencing



# Paired-end sequencing



# Interpreting Insert Size

BROAD INSTITUTE

Integrative Genomics Viewer

# Interpreting inferred insert size


The “inferred insert size” can be used to detect structural variants, including:

- Deletions
- Insertions
- Inter-chromosomal rearrangements: (Undefined insert size)

BROAD INSTITUTE

# Deletion

What is the effect of a deletion on inferred insert size?

# Deletion

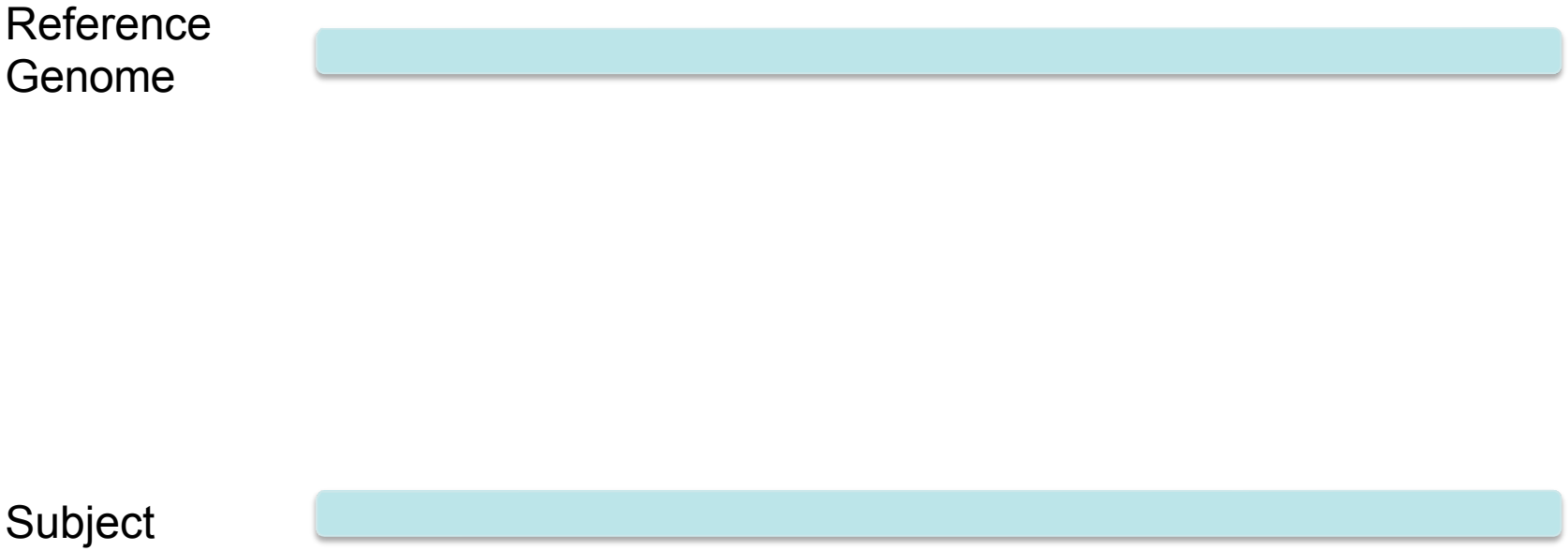
Integrative Genomics Viewer

Reference Genome



BROAD INSTITUTE

# Deletion

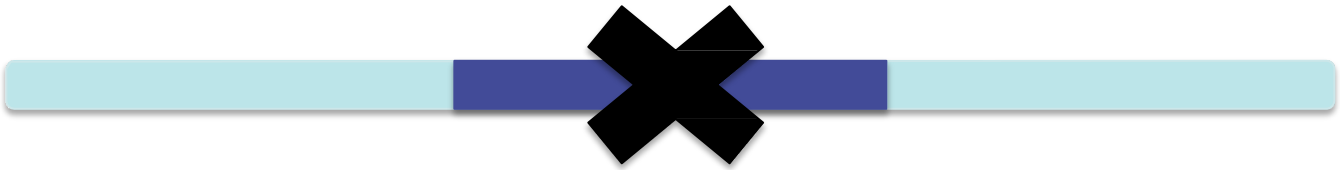


# Deletion

Integrative Genomics Viewer

Reference Genome

Subject



BROAD INSTITUTE

# Deletion

Integrative Genomics Viewer

Reference Genome

Subject

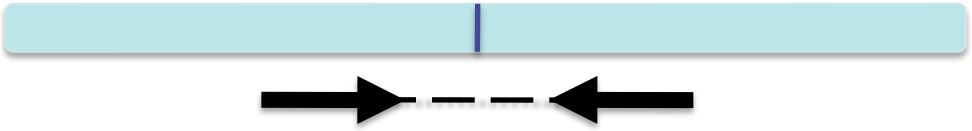
BROAD INSTITUTE

# Deletion

Integrative Genomics Viewer

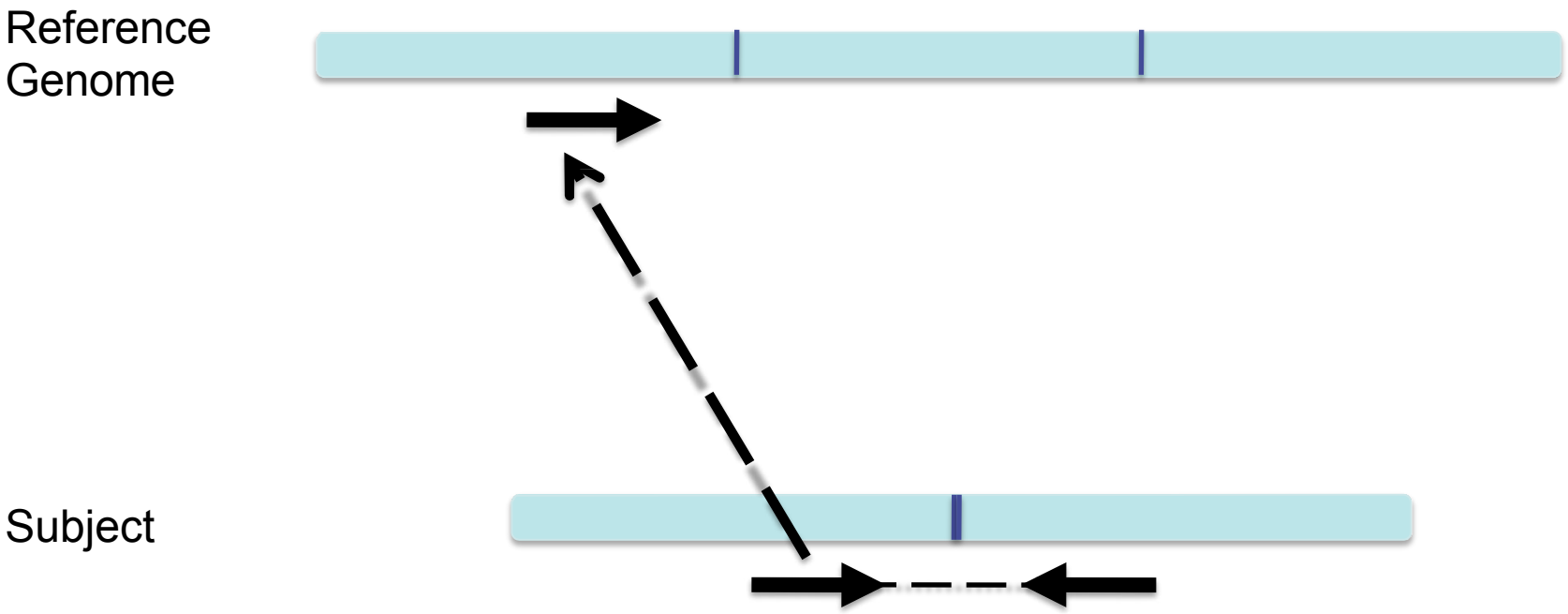
Reference Genome

Subject

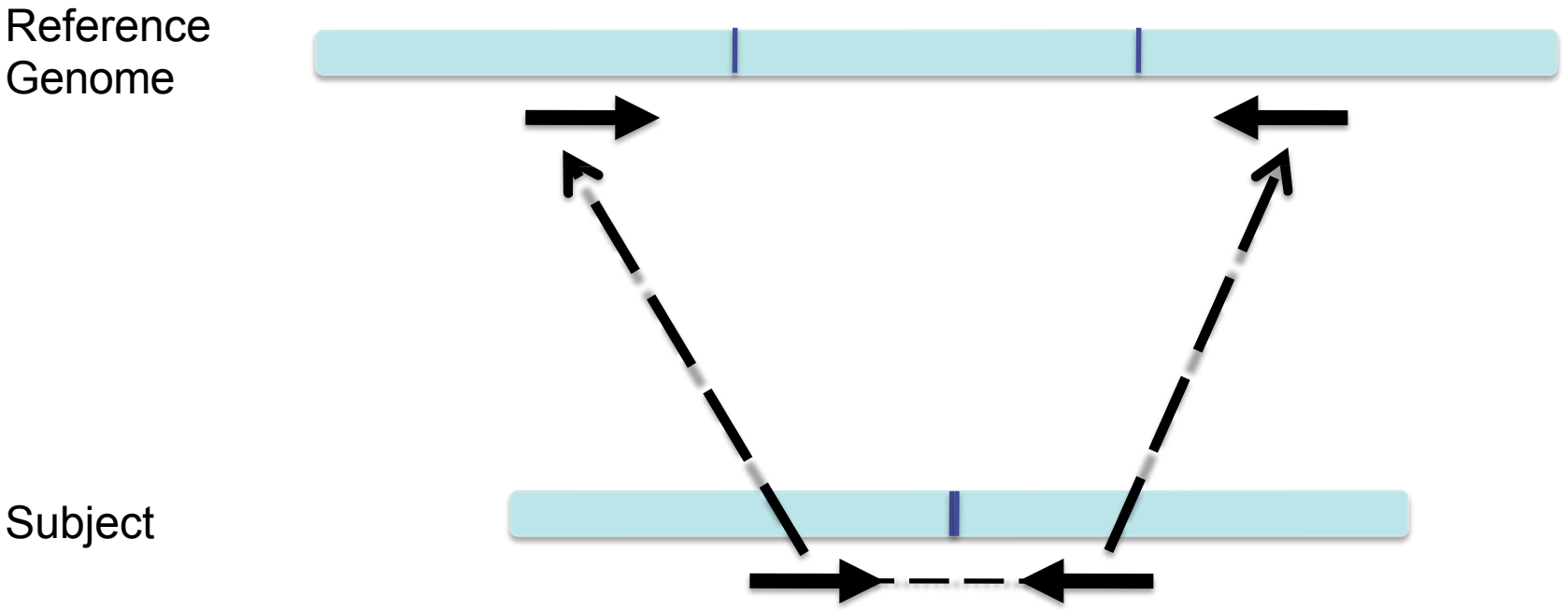


BROAD INSTITUTE

# Deletion

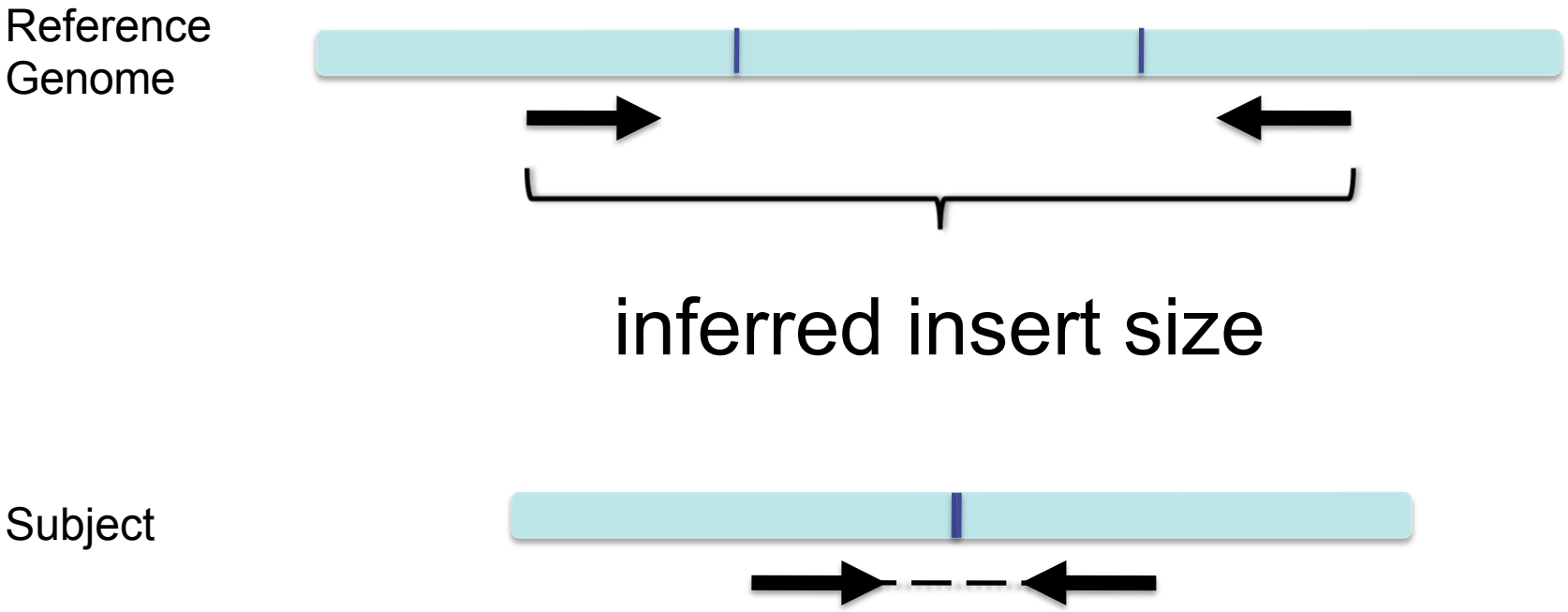


# Deletion



# Deletion

Integrative Genomics Viewer




BROAD INSTITUTE

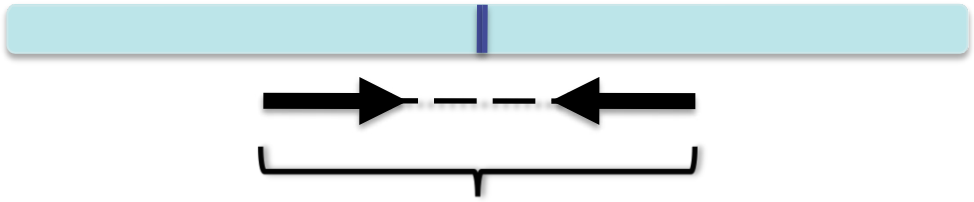
# Deletion

Integrative Genomics Viewer

Reference Genome

inferred insert size

Subject



expected insert size

BROAD INSTITUTE

# Deletion

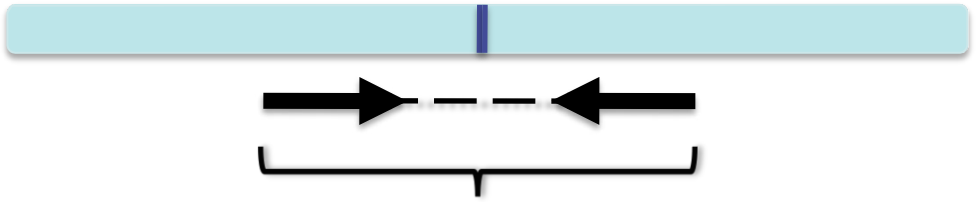
Integrative Genomics Viewer

## Inferred insert size is > expected value

Reference Genome

inferred insert size

Subject



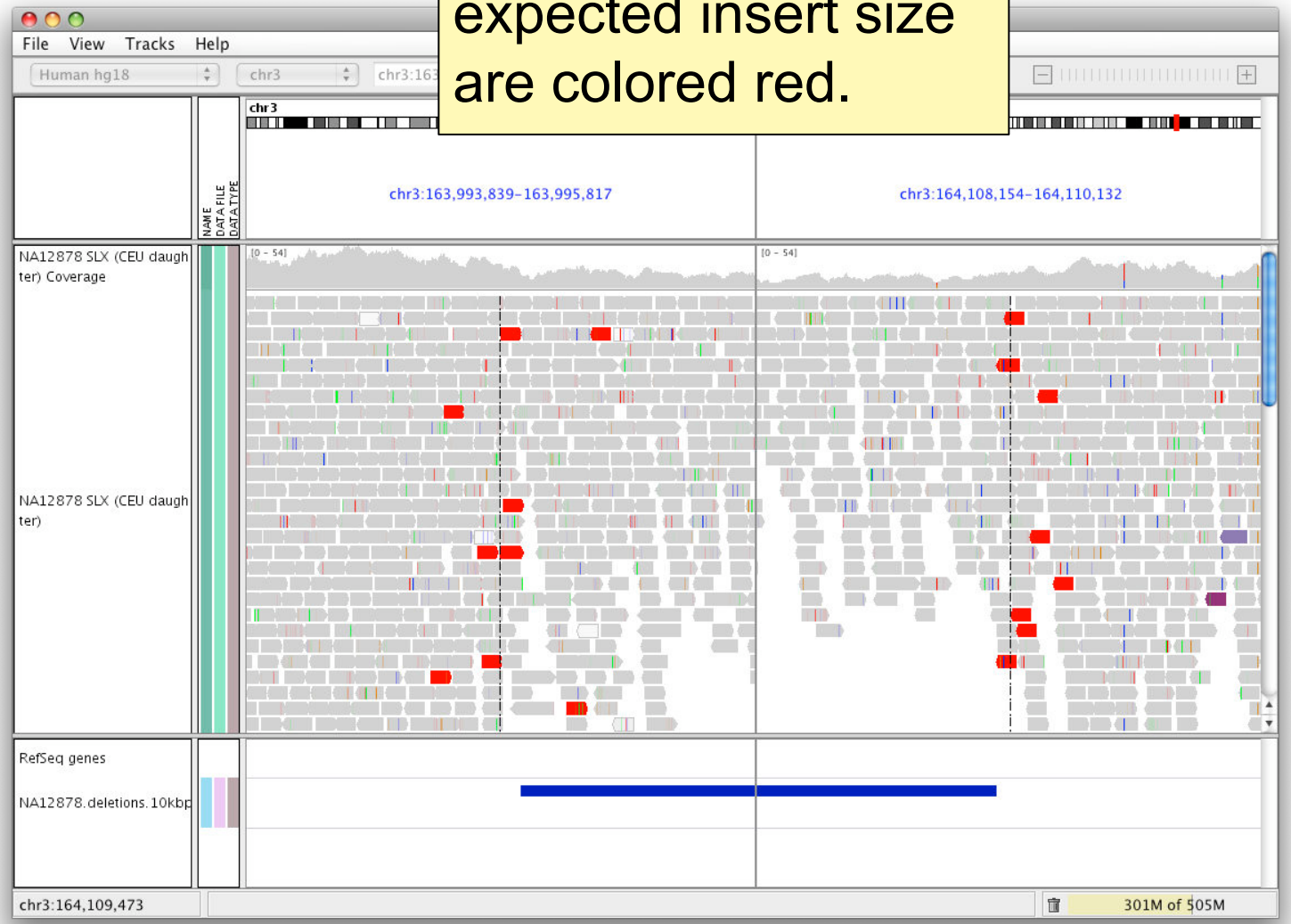
expected insert size

BROAD INSTITUTE

# Deletion

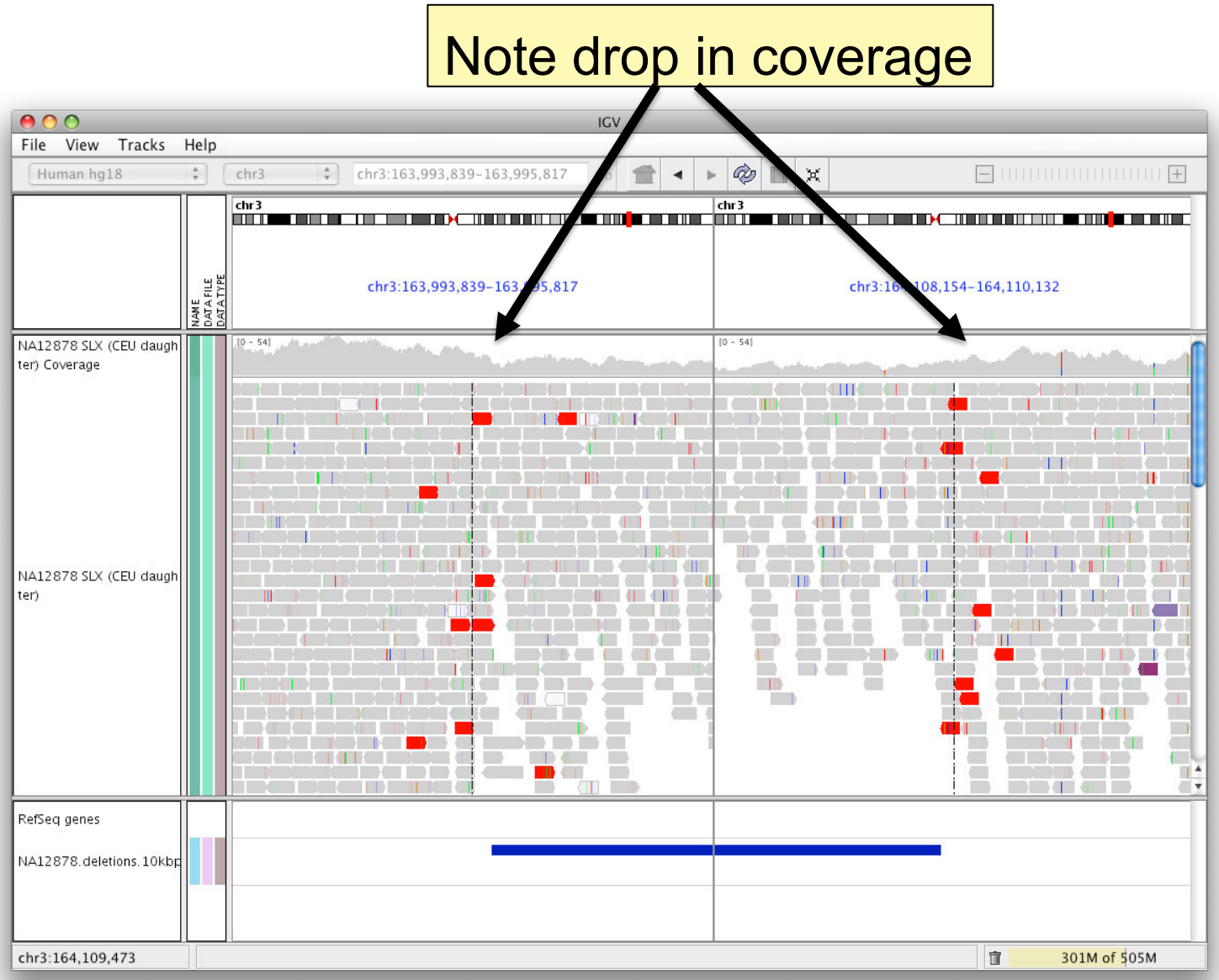
Integrative Genomics Viewer

Pairs with larger than expected insert size are colored red.




BROAD INSTITUTE

# Deletion

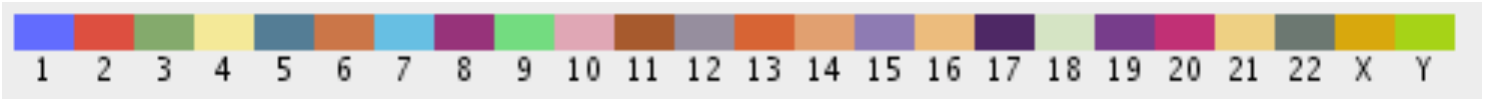

Note drop in coverage
IGV
File View Tracks Help
Human hg18
chr3
chr3:163,993,839-163,995,817
chr3
chr3
chr3:163,993,839-163,995,817
chr3:164,108,154-164,110,132
NA12878 SLX (CEU daughter) Coverage
NA12878 SLX (CEU daughter)
RefSeq genes
NA12878.deletions.10kbp
chr3:164,109,473
301M of 505M

# Insert size color scheme

Integrative Genomics Viewer

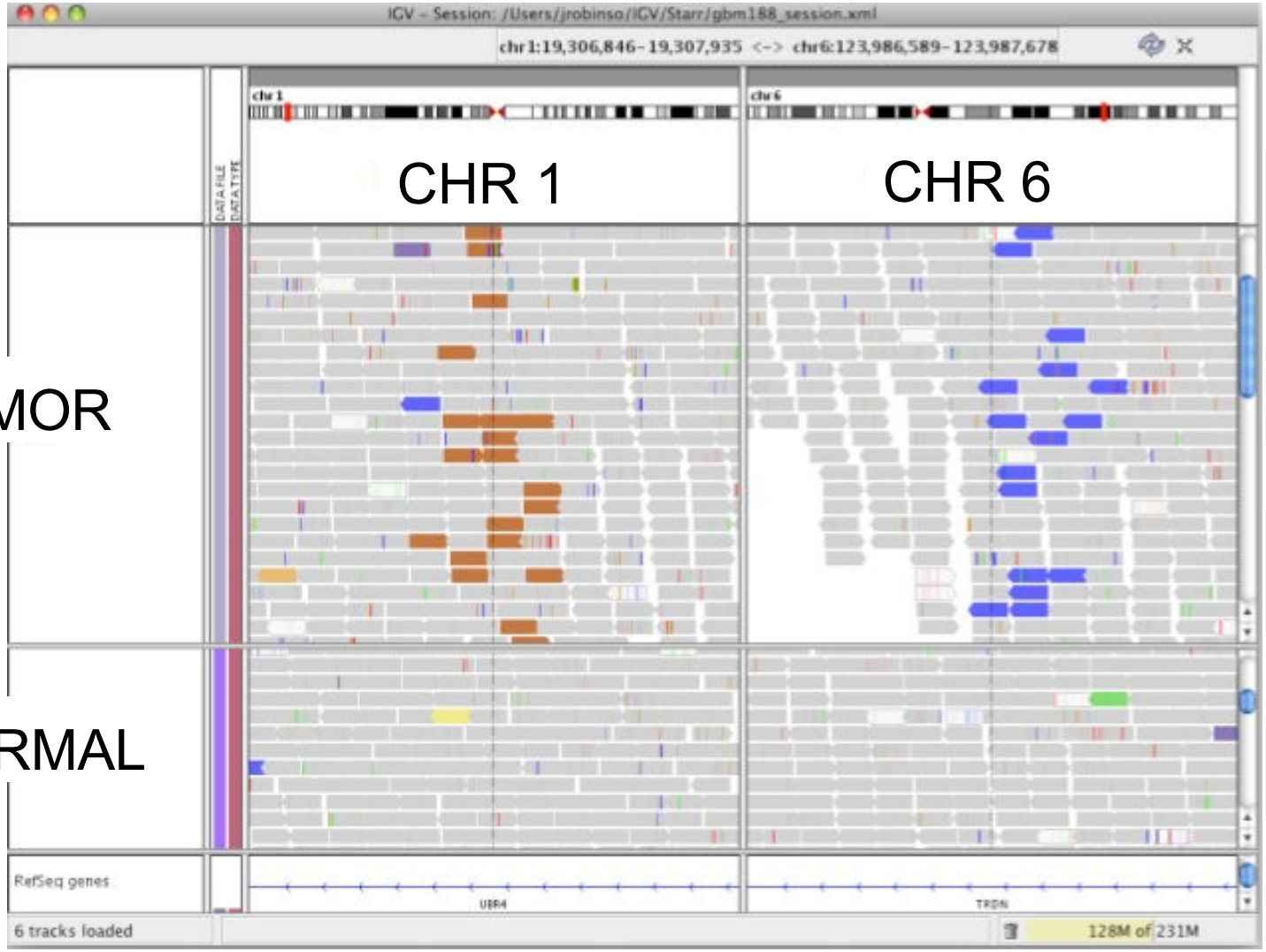
- Smaller than expected insert size: 
- Larger than expected insert size: 
- Pairs on different chromosomes

*Each end colored by chromosome of its mate*

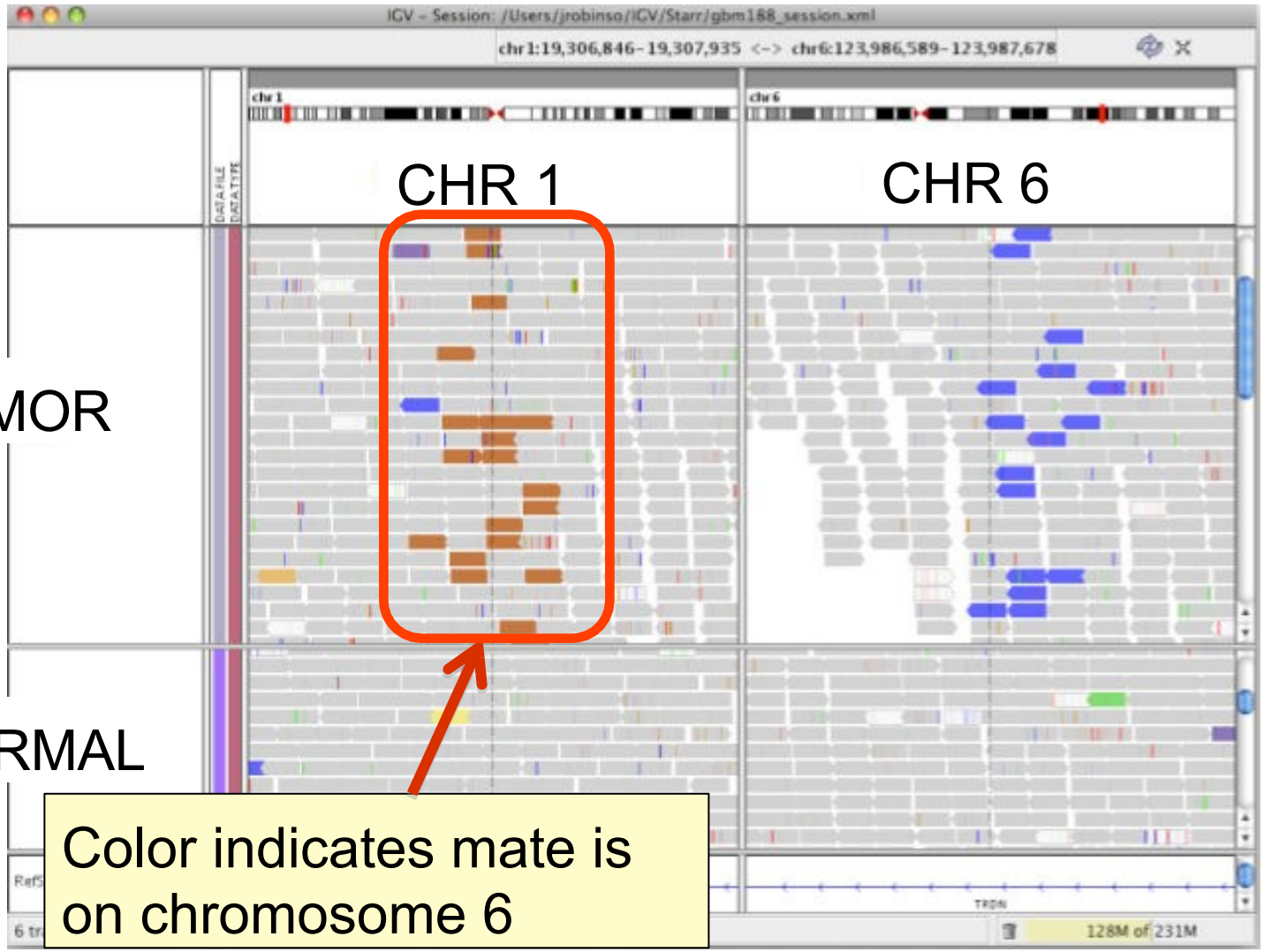



BROAD INSTITUTE

# Rearrangement



# Rearrangement



# Interpreting Pair Orientations

BROAD INSTITUTE

igv Integrative Genomics Viewer

# Interpreting pair orientations

Integrative Genomics Viewer

Orientation of paired reads can reveal structural events, including:

- inversions
- duplications
- translocations

Orientation is defined in terms of

- read strand, left *vs* right, *and*
- read order, first *vs* second

BROAD INSTITUTE

# Inversion

Integrative Genomics Viewer

Reference genome



BROAD INSTITUTE

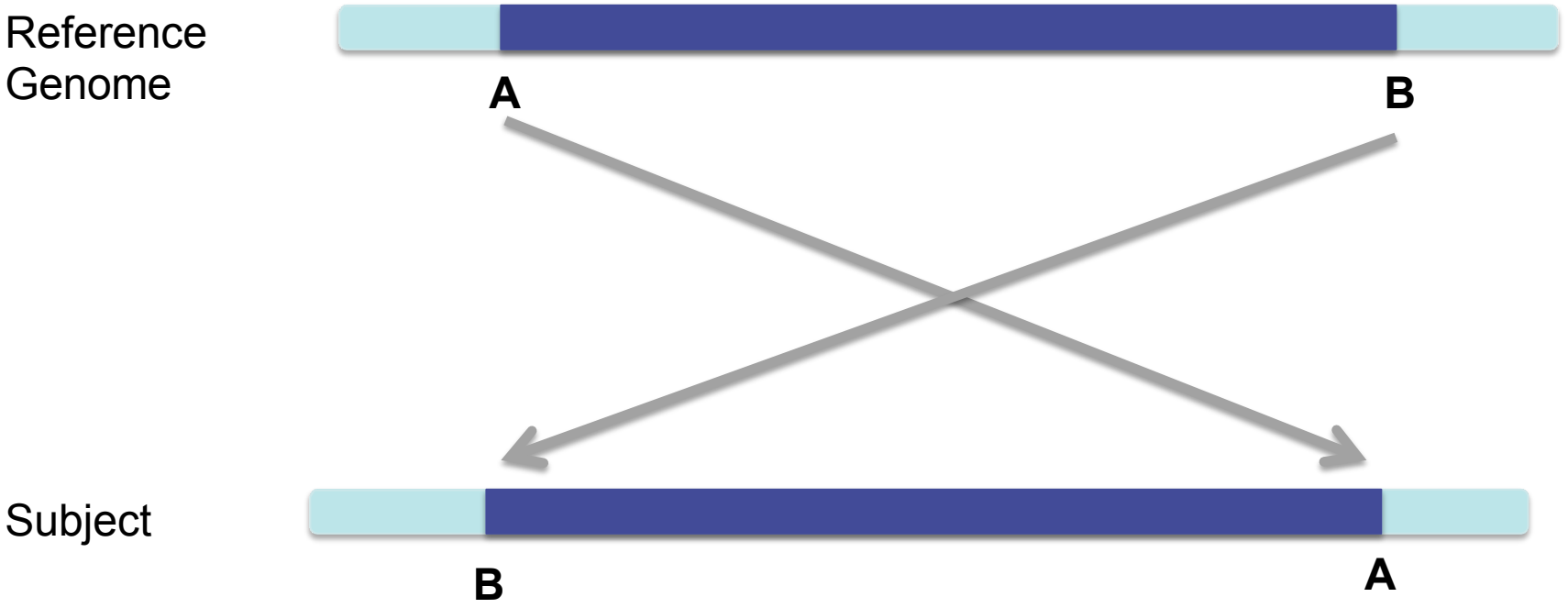
# Inversion

Integrative Genomics Viewer

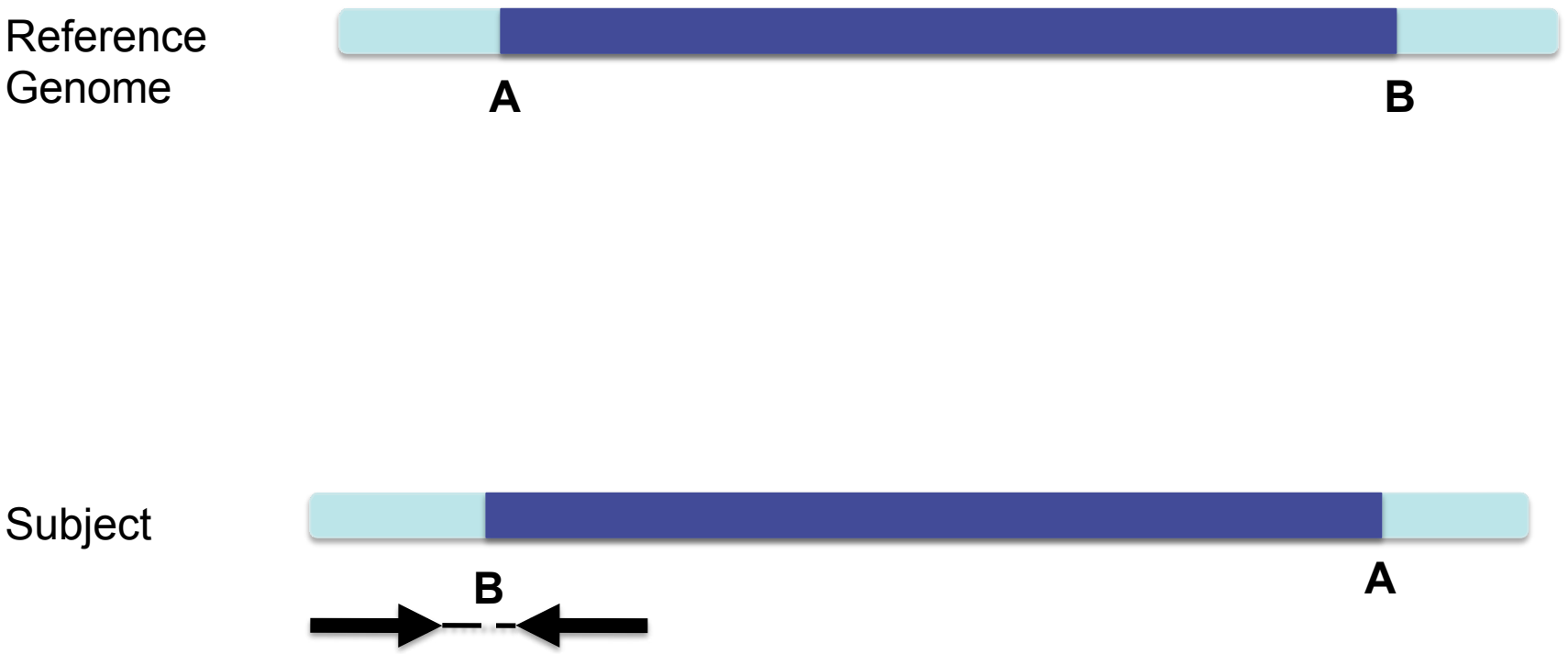
Reference genome


BROAD INSTITUTE

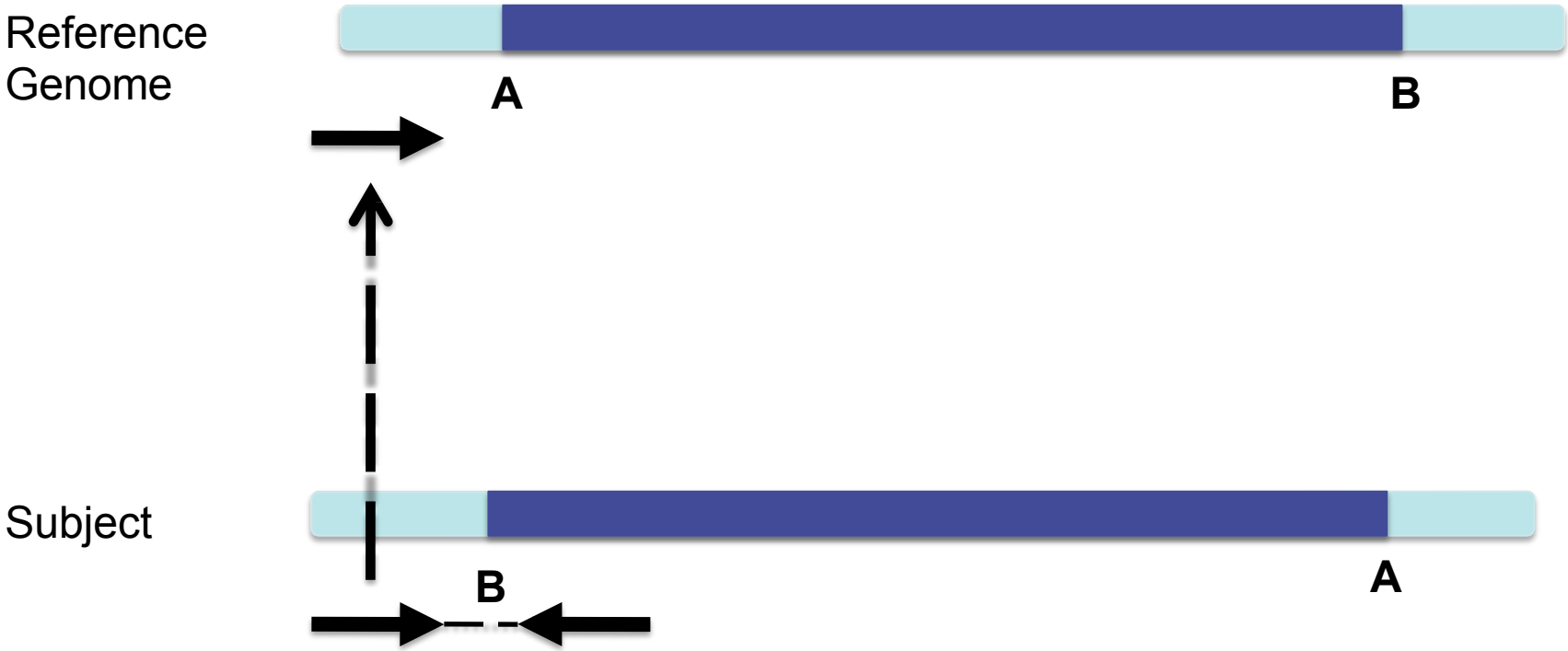
# Inversion



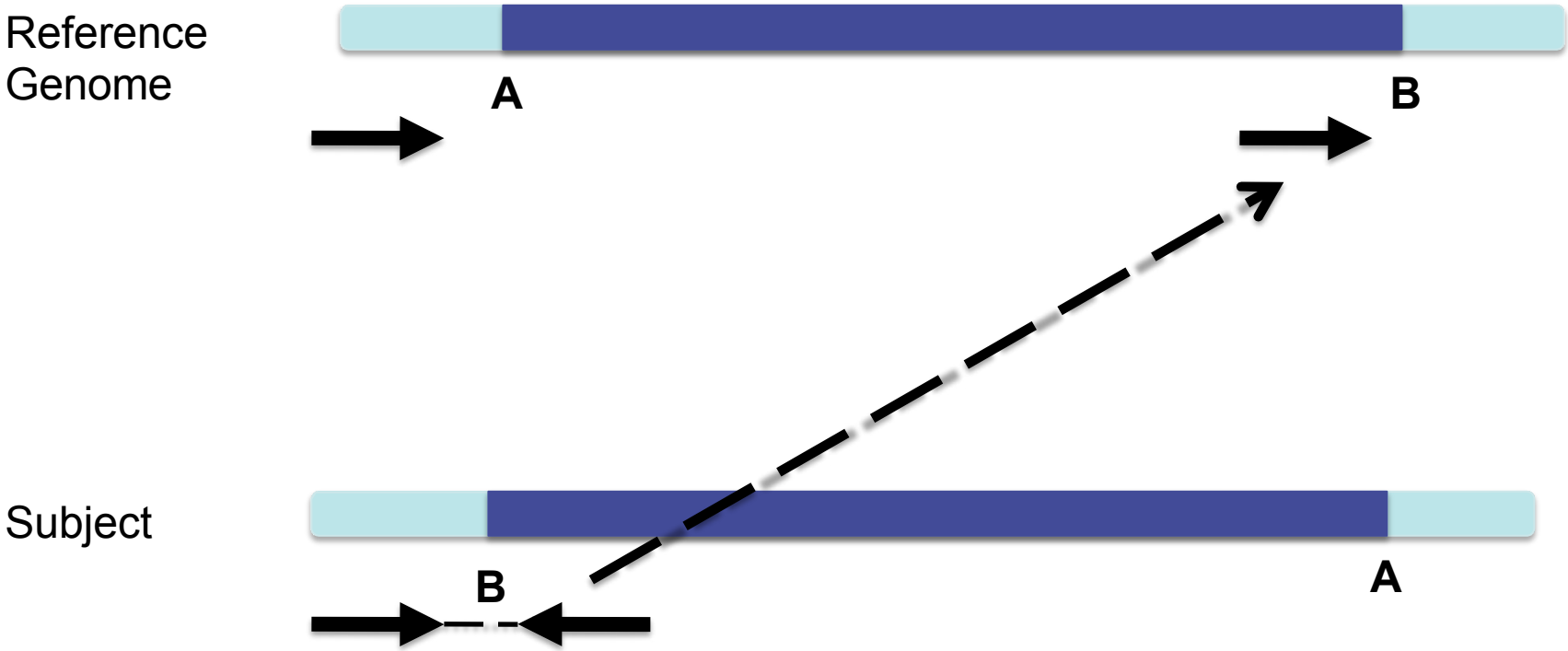
# Inversion



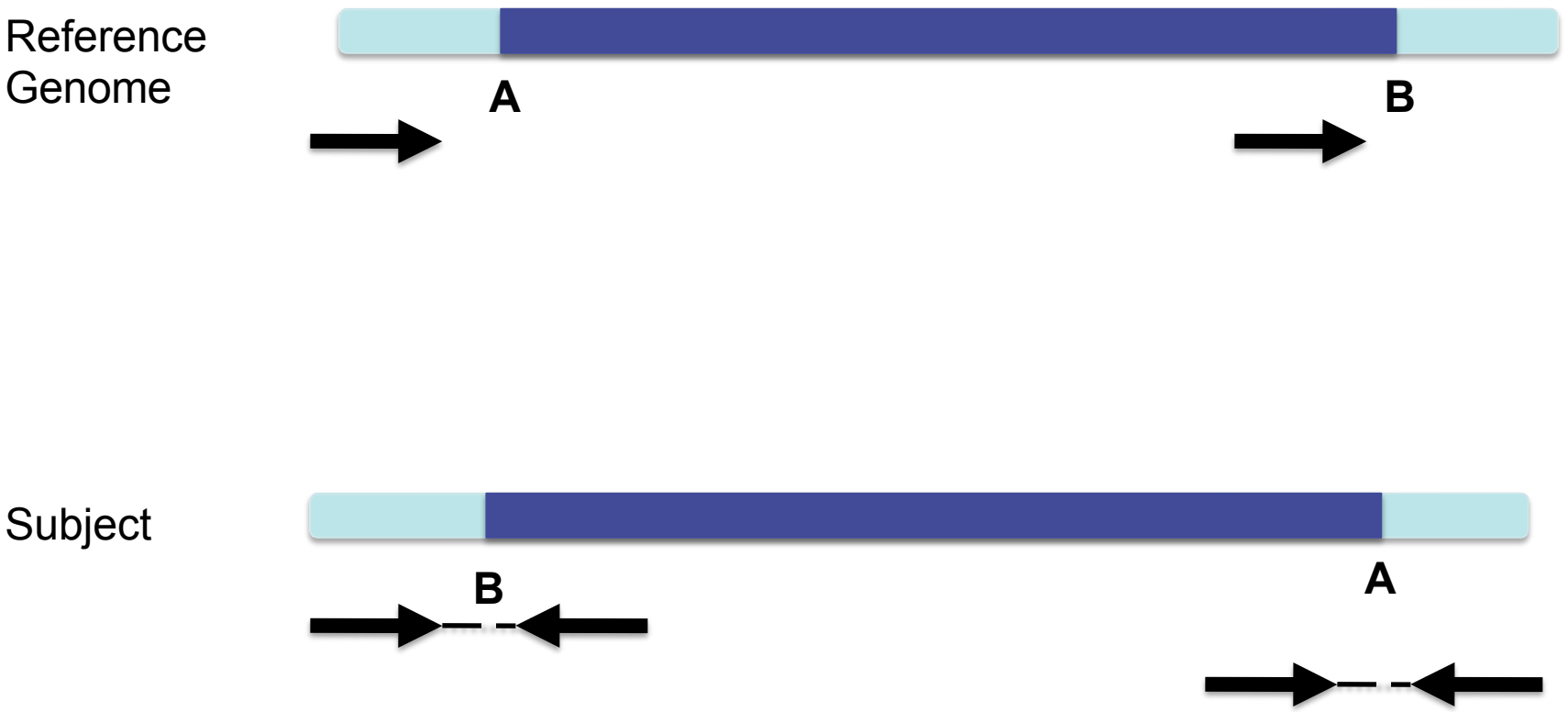
# Inversion



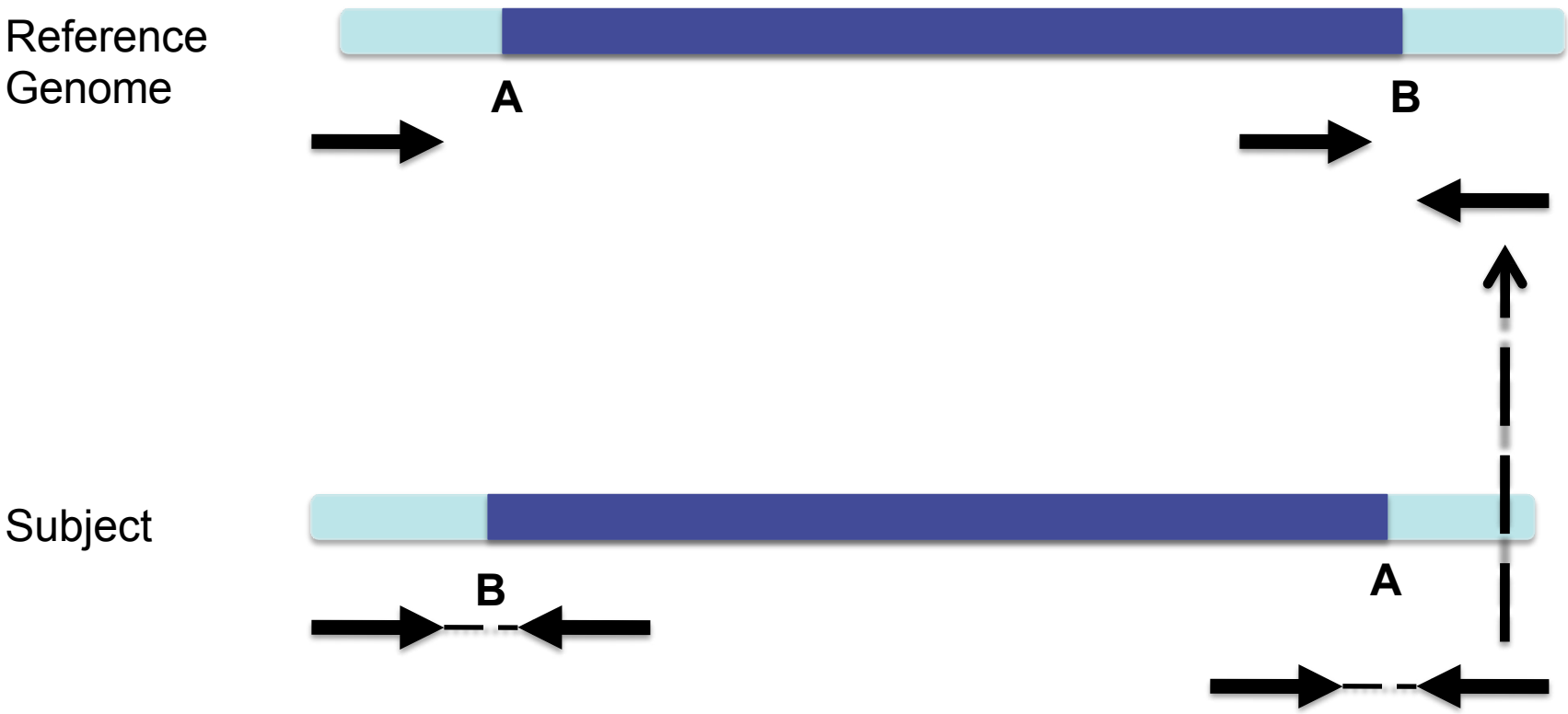
# Inversion



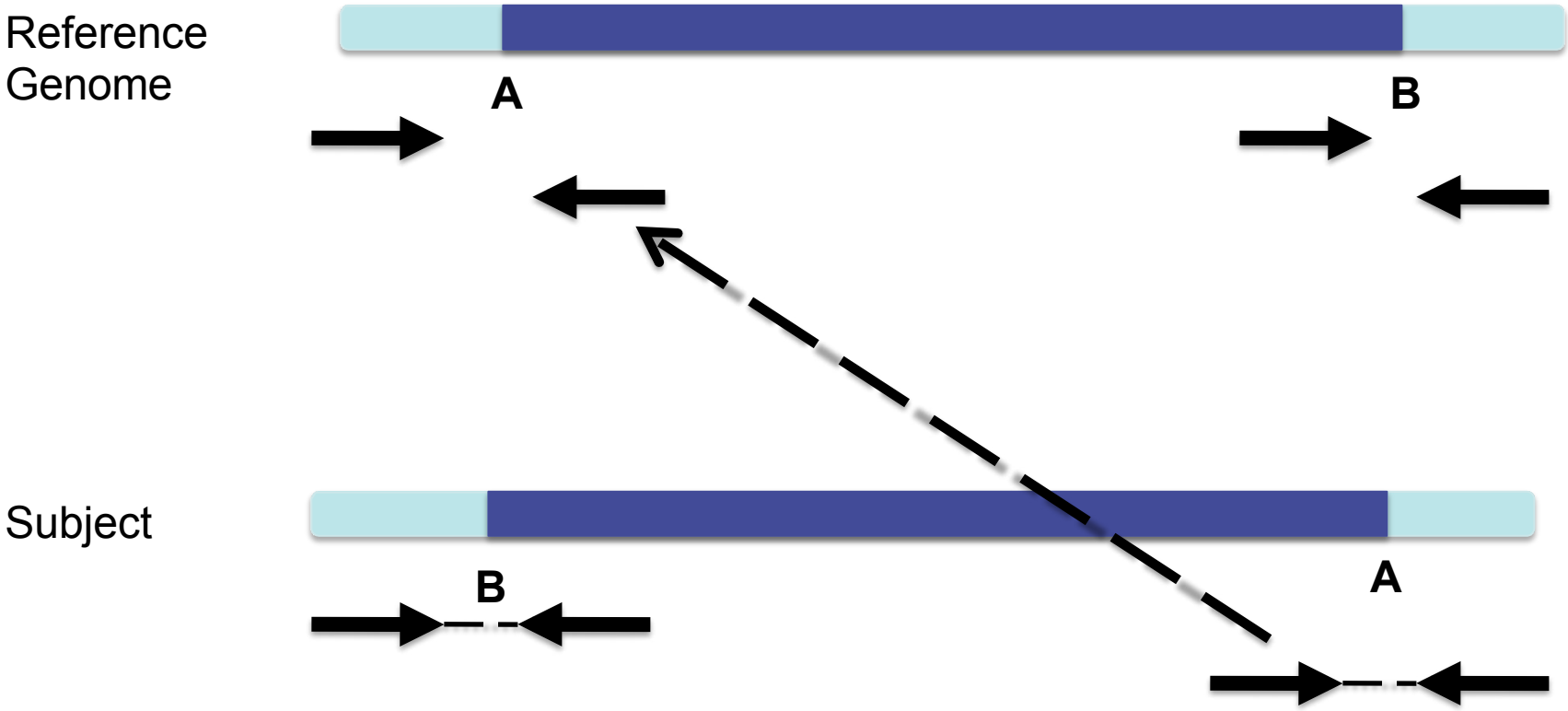
# Inversion



# Inversion

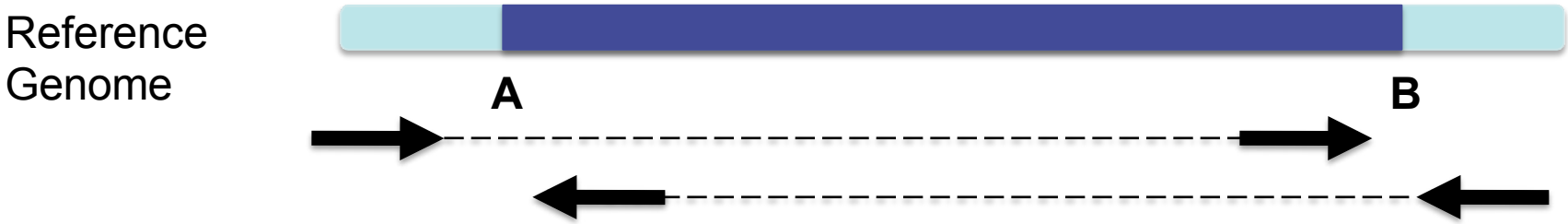


# Inversion



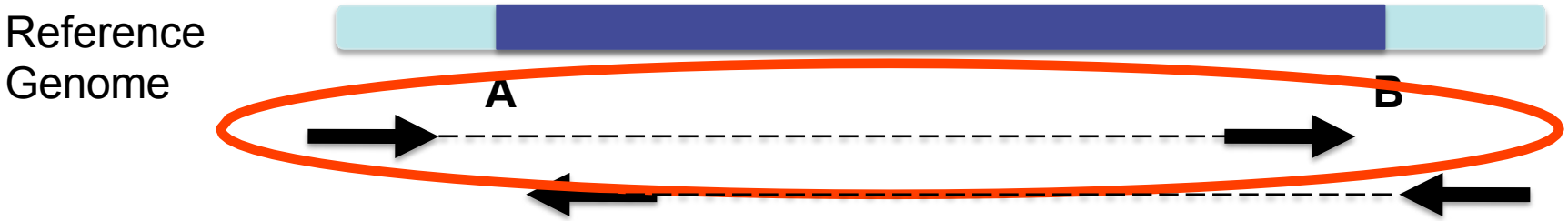
# Inversion

Integrative Genomics Viewer



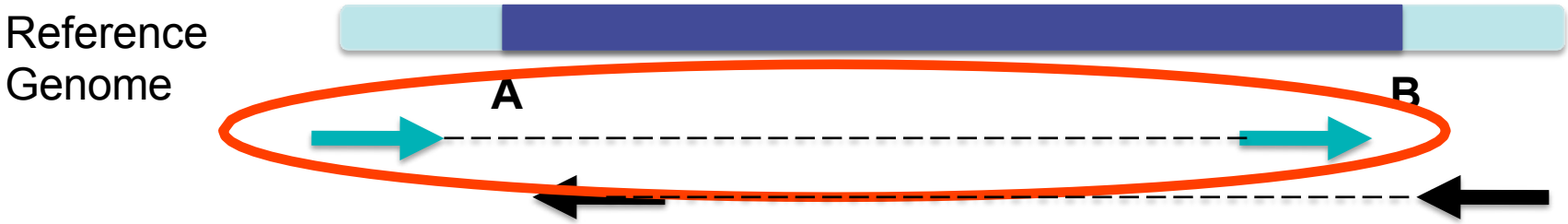

BROAD INSTITUTE

# Inversion




Anomaly –
Expected pair orientation is inward facing ( → ← )

# Inversion

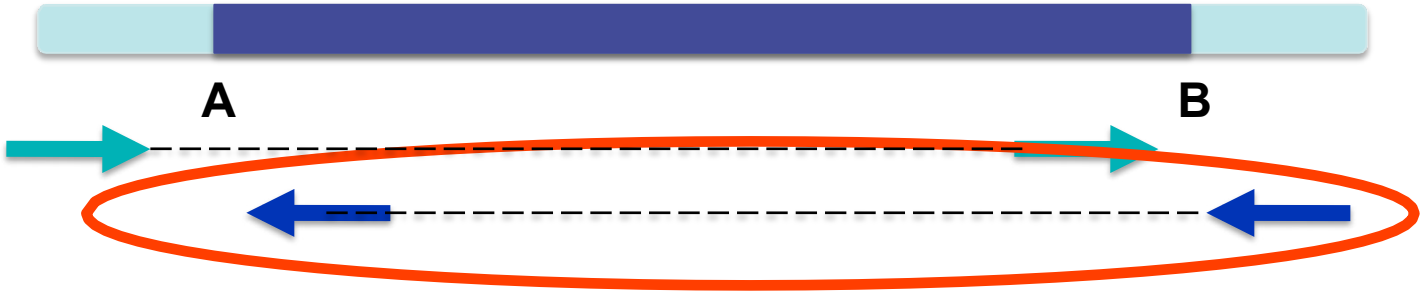



"Left" side pair

# Inversion

Integrative Genomics Viewer

Reference Genome

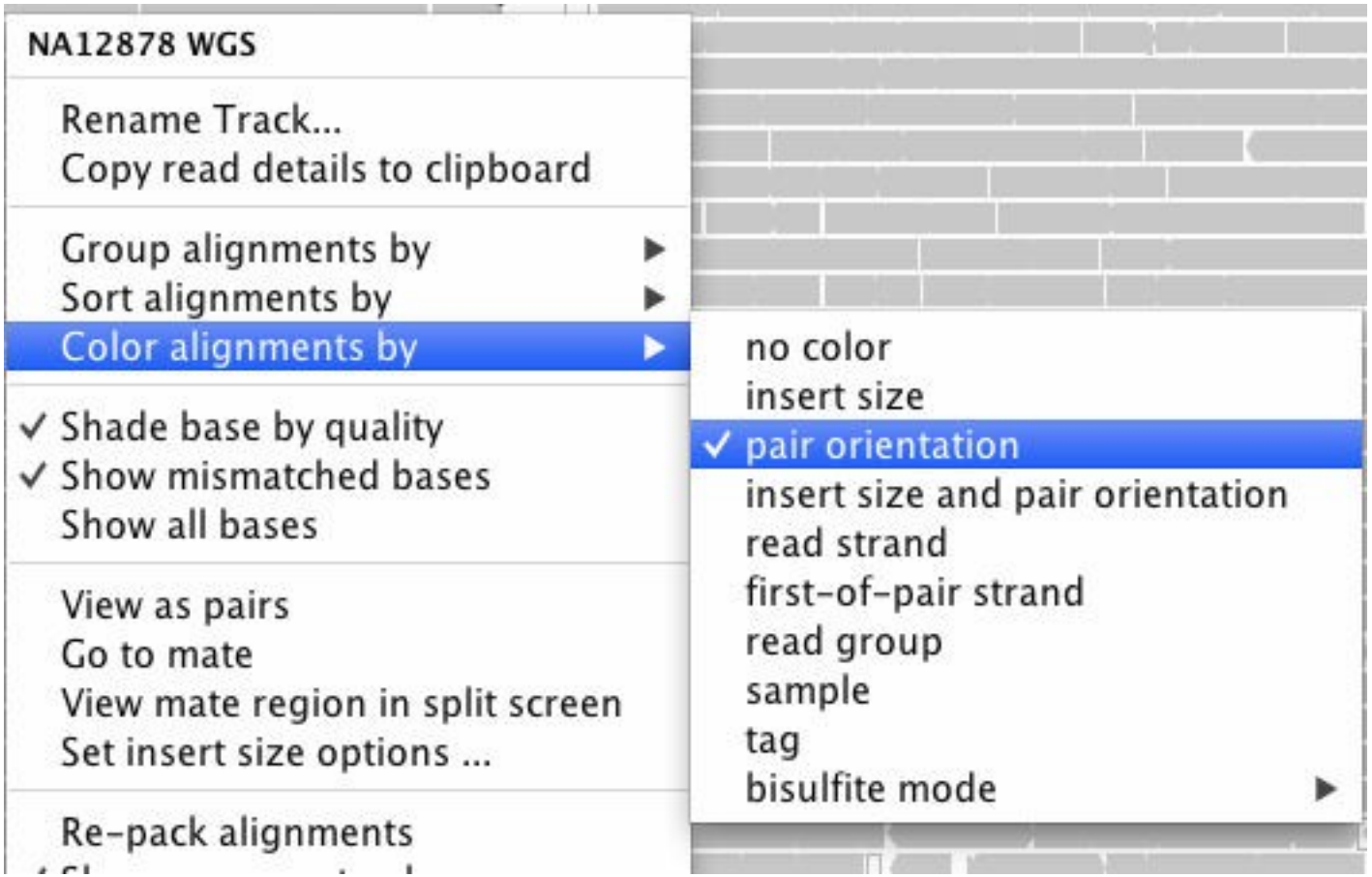



“Right” side pair

BROAD INSTITUTE

# Color by pair orientation

Integrative Genomics Viewer




BROAD INSTITUTE

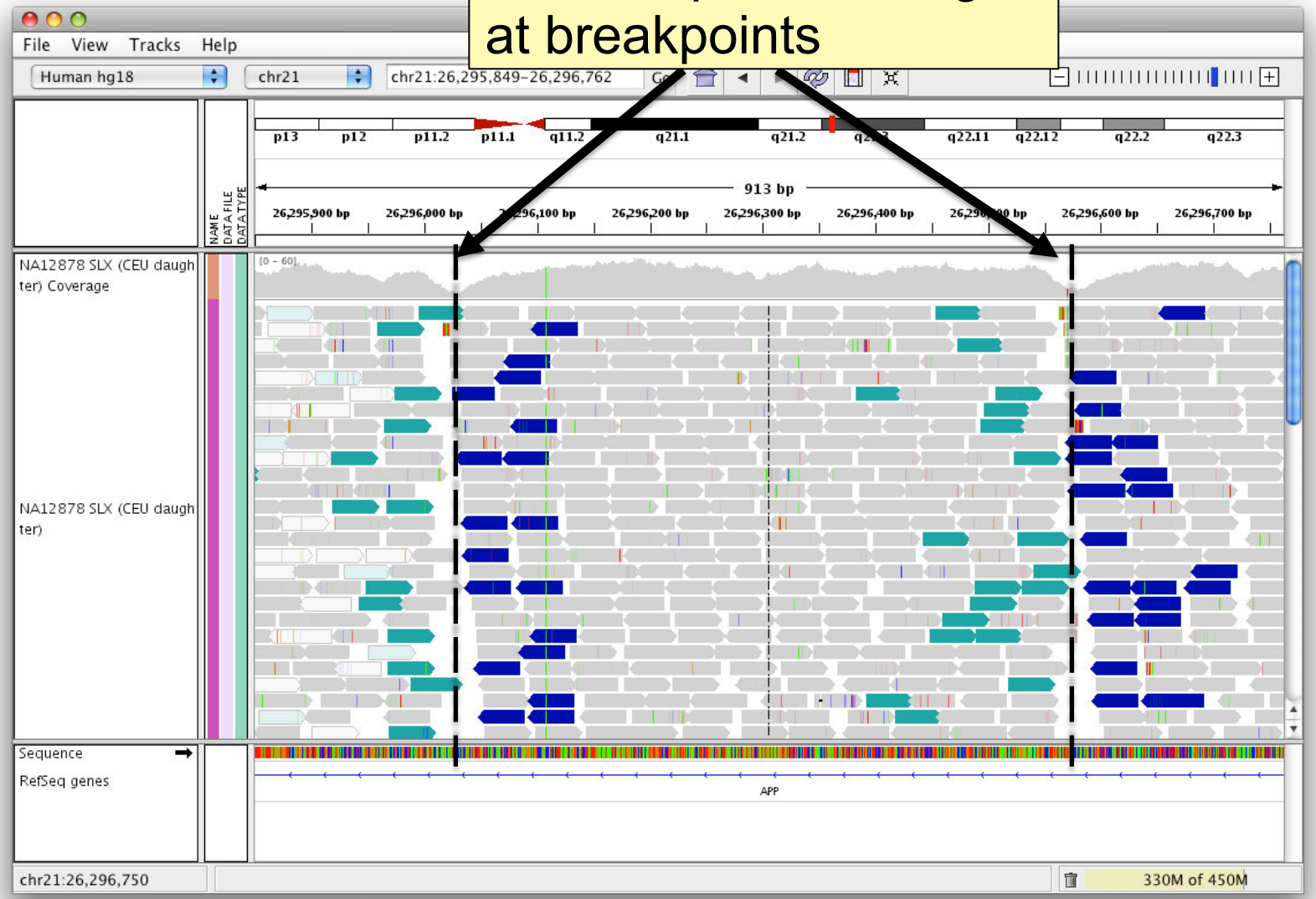
# Inversion

Integrative Genomics Viewer


BROAD INSTITUTE

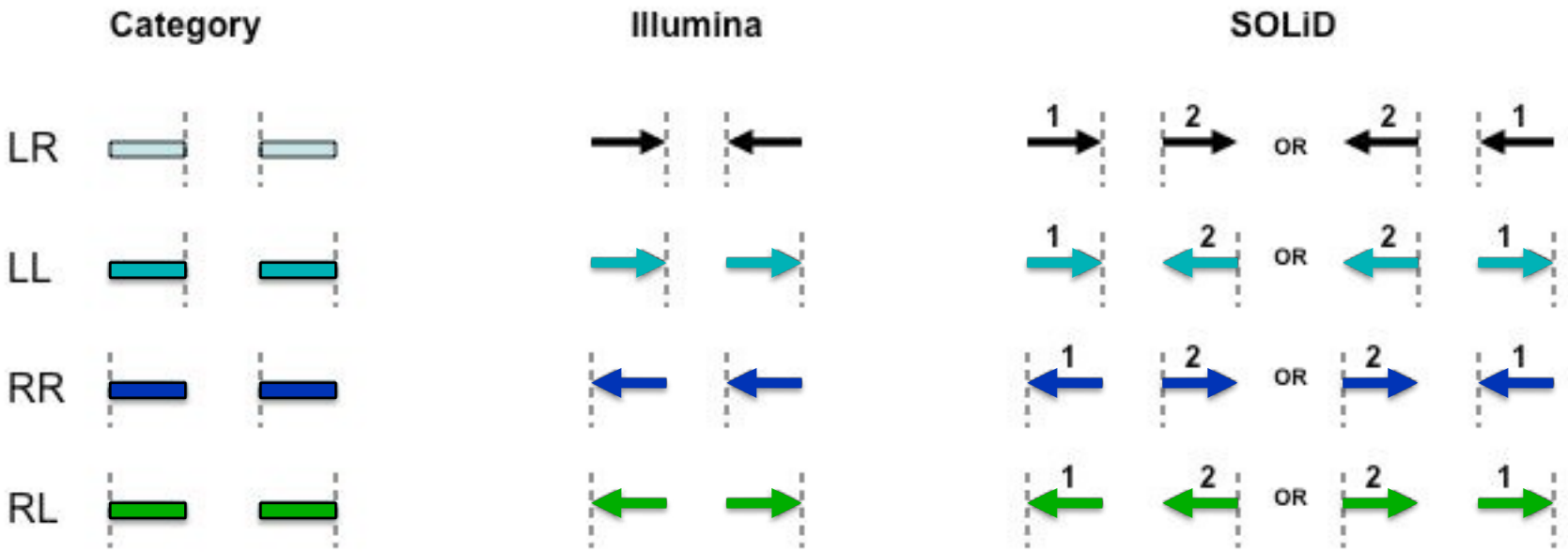
# Inversion

Integrative Genomics Viewer




BROAD INSTITUTE

## Interpretation of read pair orientations




| | |
|---|---|
| LR | Normal reads.<br>The reads are left and right (respectively) of the unsequenced part of the sequenced DNA fragment when aligned back to the reference genome. |
| LL,RR | Implies inversion in sequenced DNA with respect to reference. |
| RL | Implies duplication or translocation with respect to reference. |

These categories only apply to reads where both mates map to the same chromosome.

*Figure courtesy of Bob Handsaker*

# RNA-Seq

Integrative Genomics Viewer

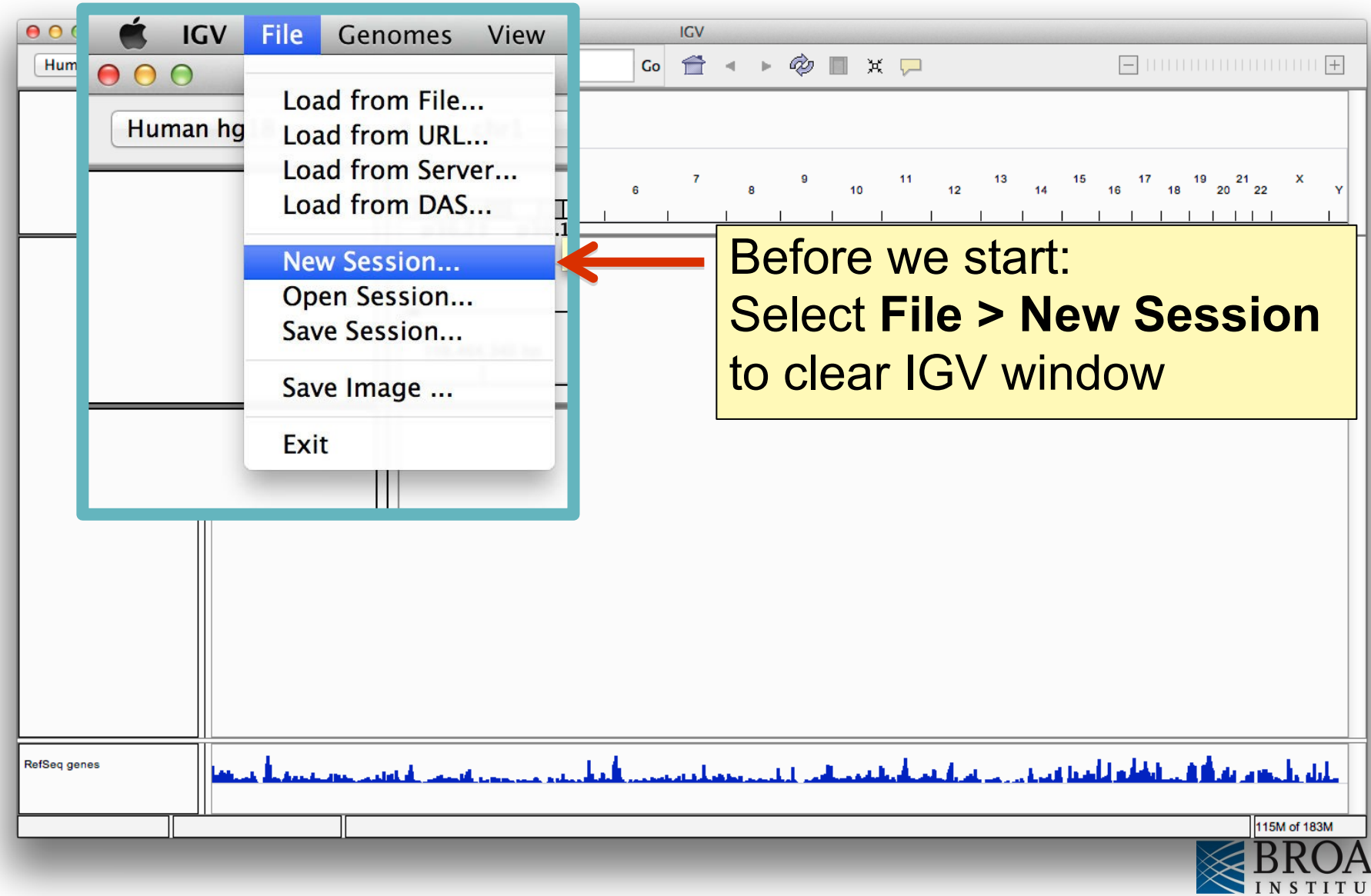
Hands-on exercise

- Examine tissue-specific alternative splicing.
- Data:  Illumina BodyMap 2.0

http://www.illumina.com/science/data_library.ilmn

BROAD INSTITUTE

Integrative Genomics Viewer




BROAD INSTITUTE

# RNA-Seq Setup

Integrative Genomics Viewer

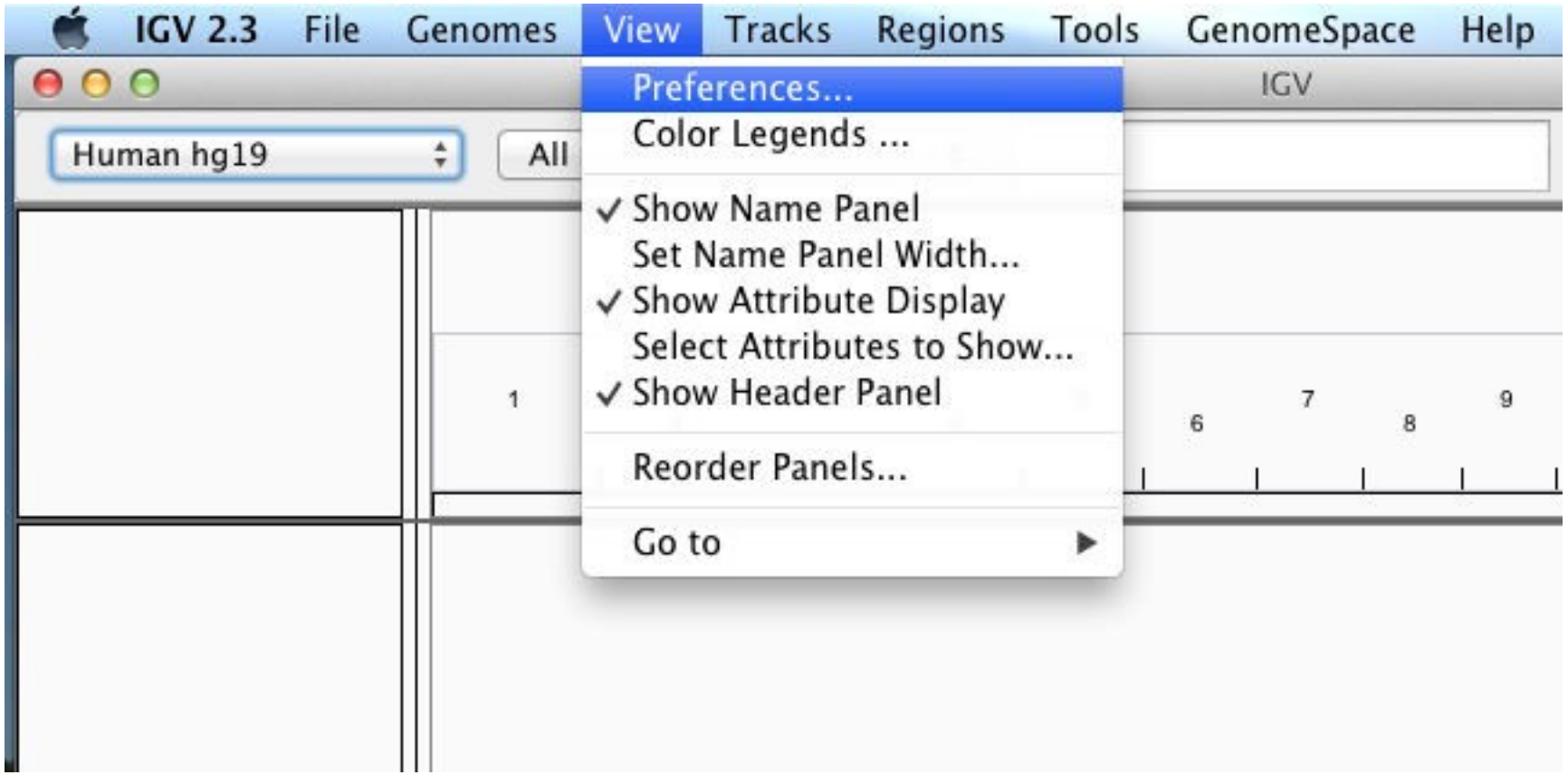
- Step 1:  Tune settings for RNA.

BROAD INSTITUTE

# RNA-seq alignments

Integrative Genomics Viewer

Select **View > Preferences…**

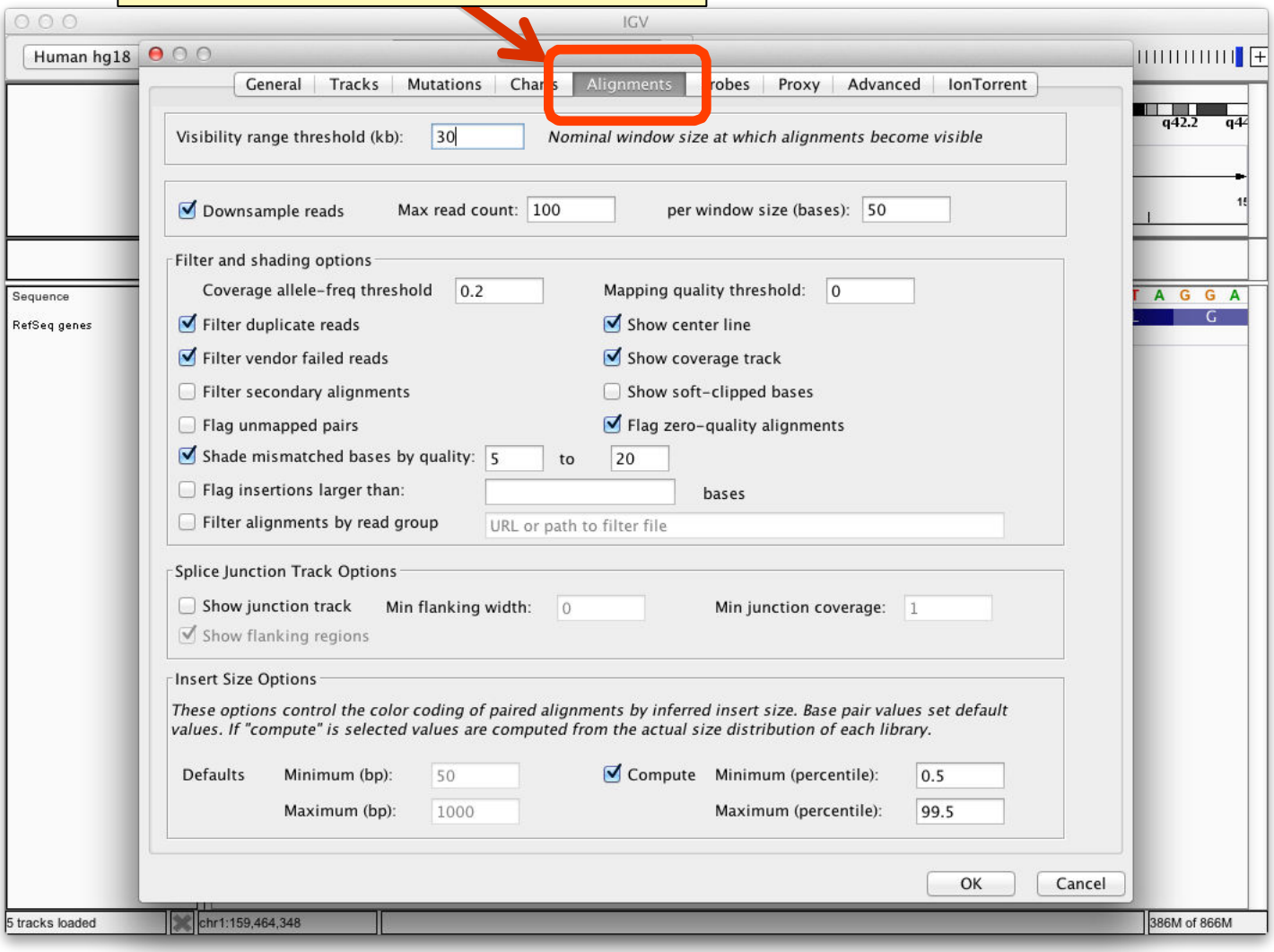



BROAD INSTITUTE

# RNA-seq alignments

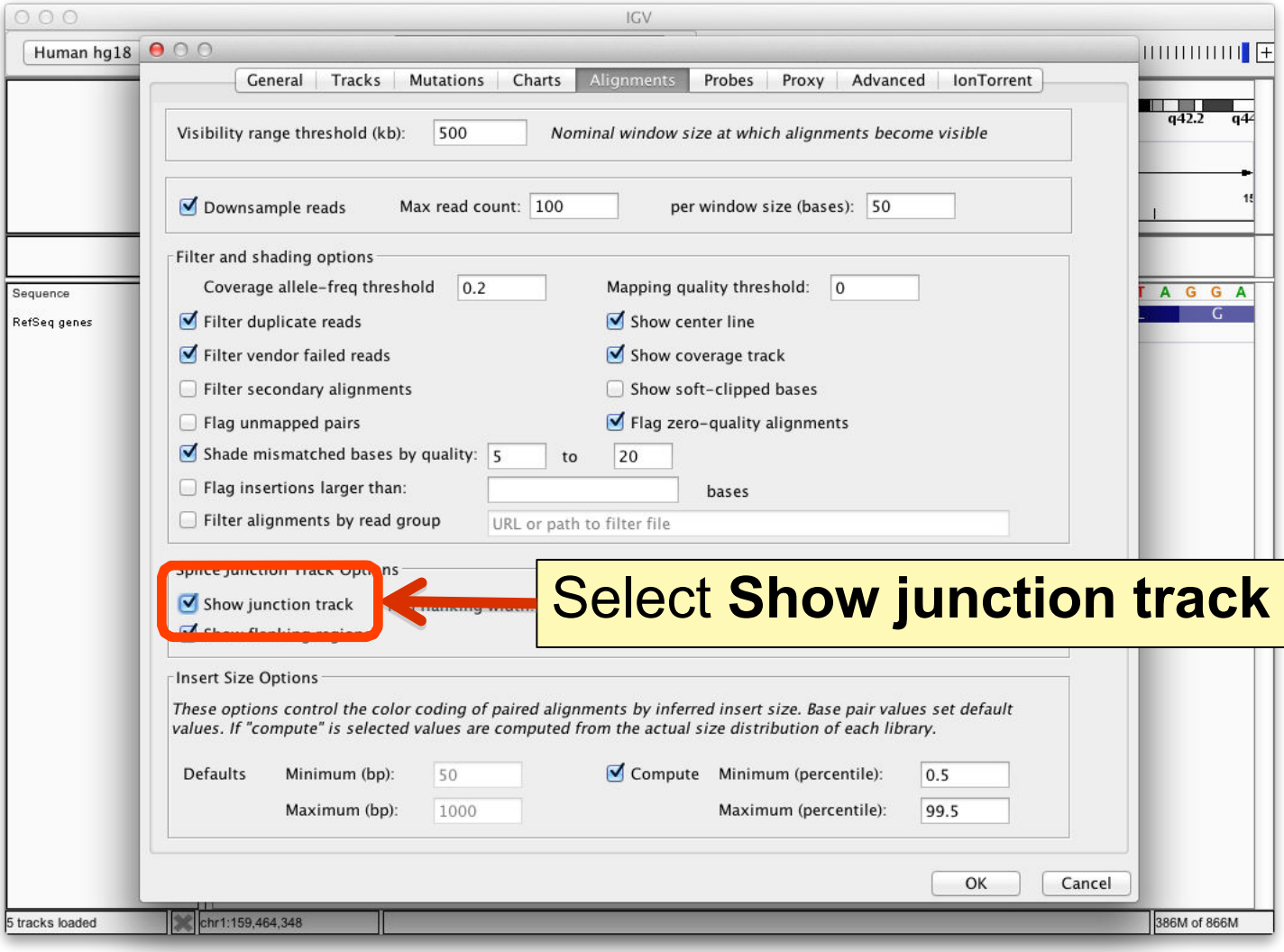
Integrative Genomics Viewer

Click **Alignments** tab

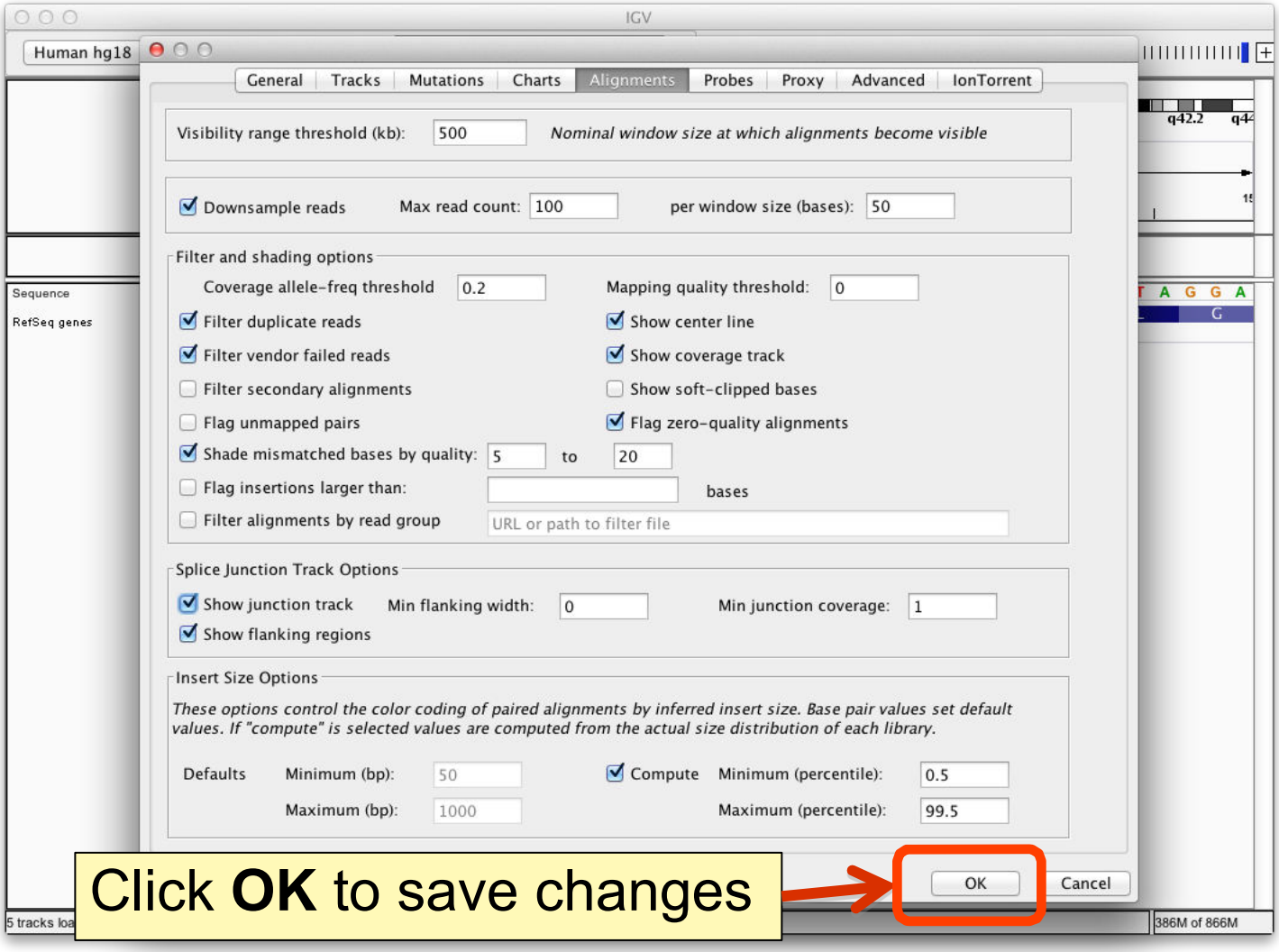



BROAD INSTITUTE

# RNA-seq alignments

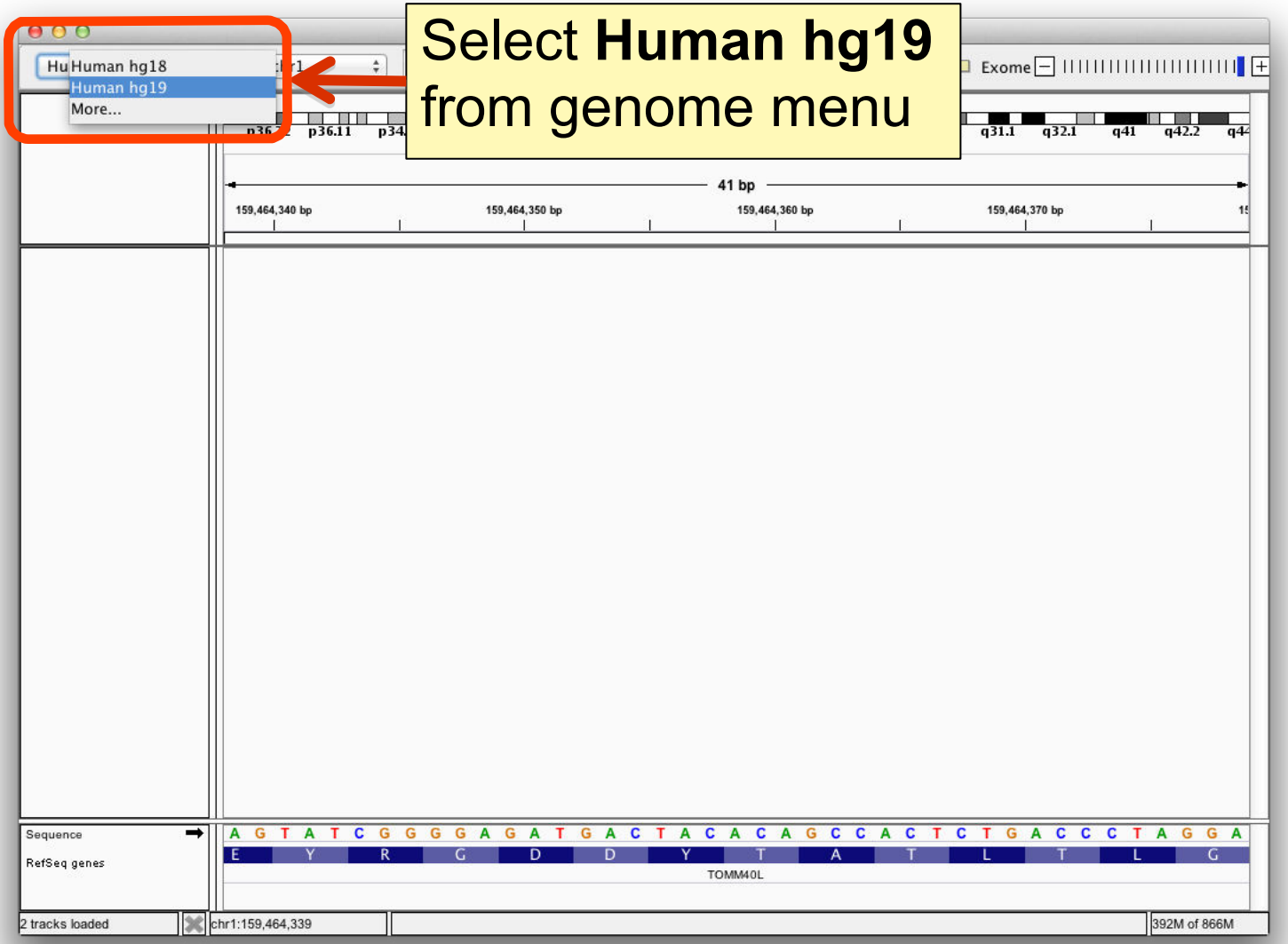


# RNA-seq alignments

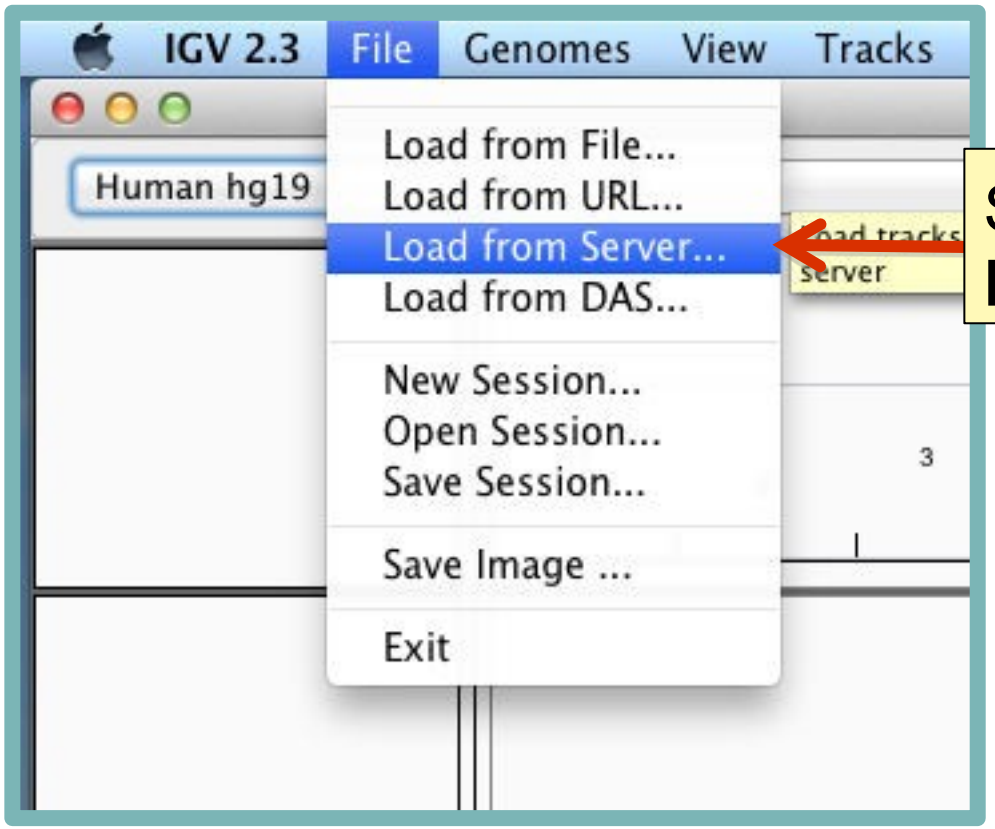



Click **OK** to save changes

# RNA-seq alignments

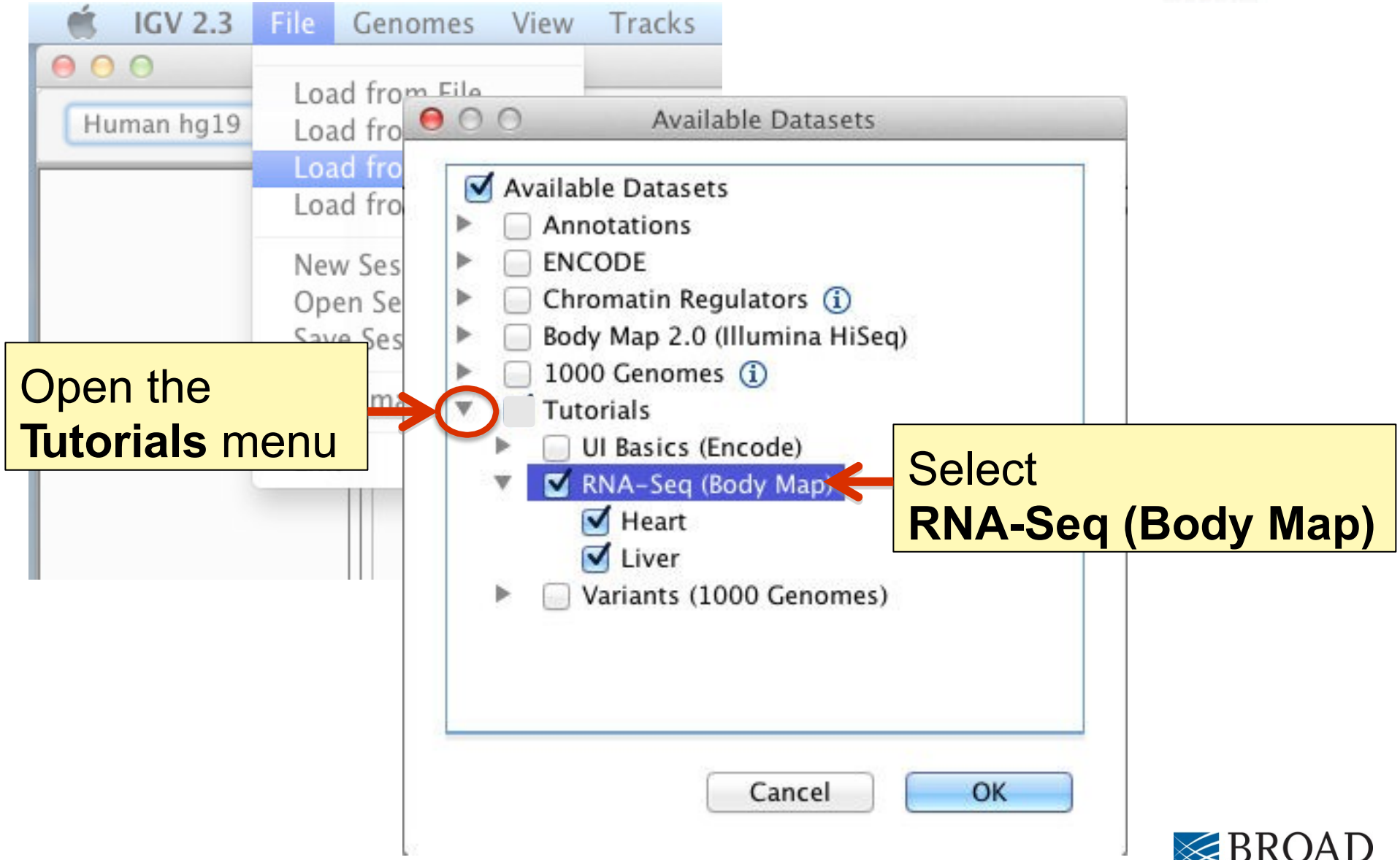


# RNA-seq alignments



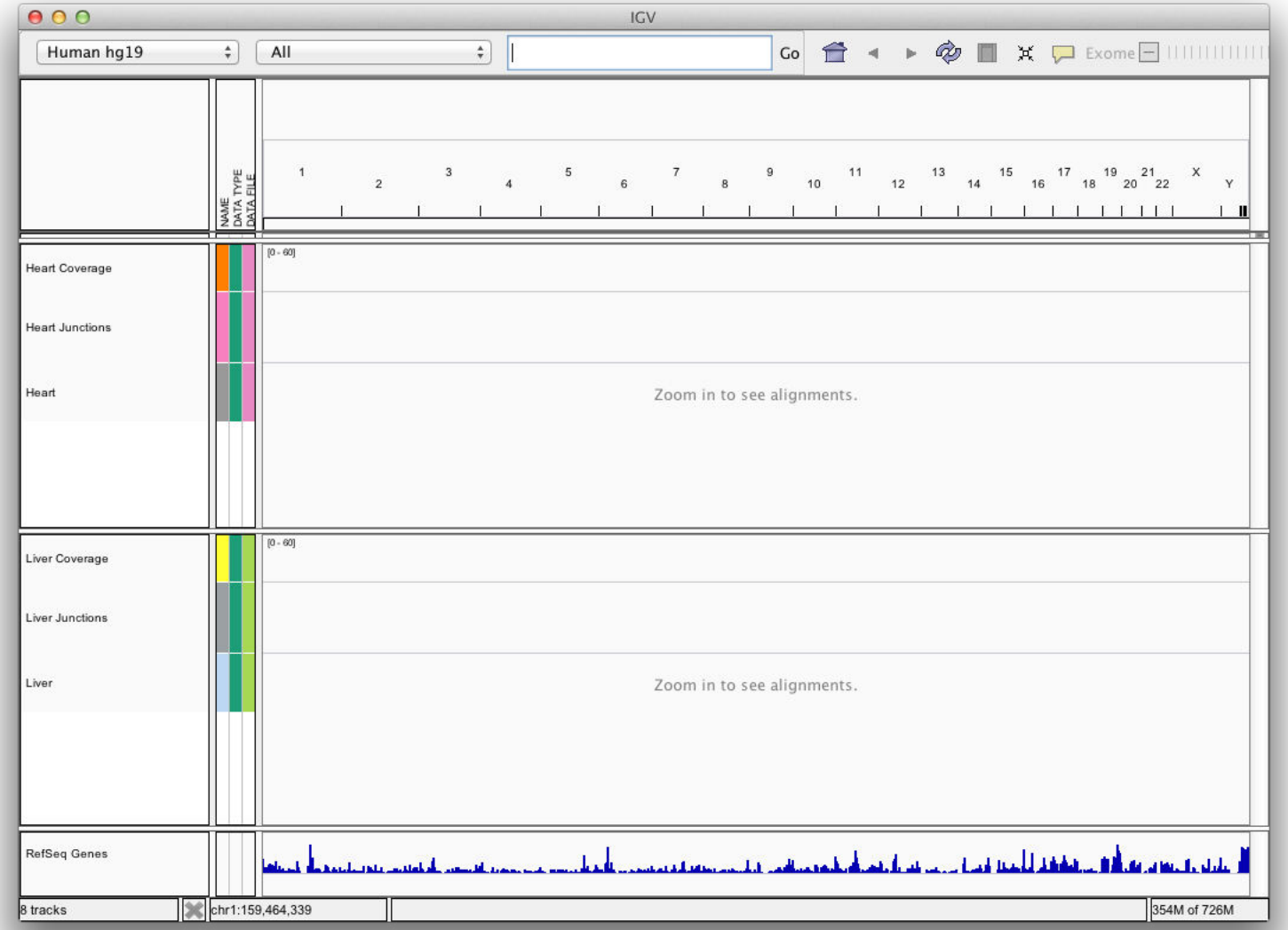

Select:
**File > Load from Server…**

# RNA-seq alignments



# RNA-seq alignments

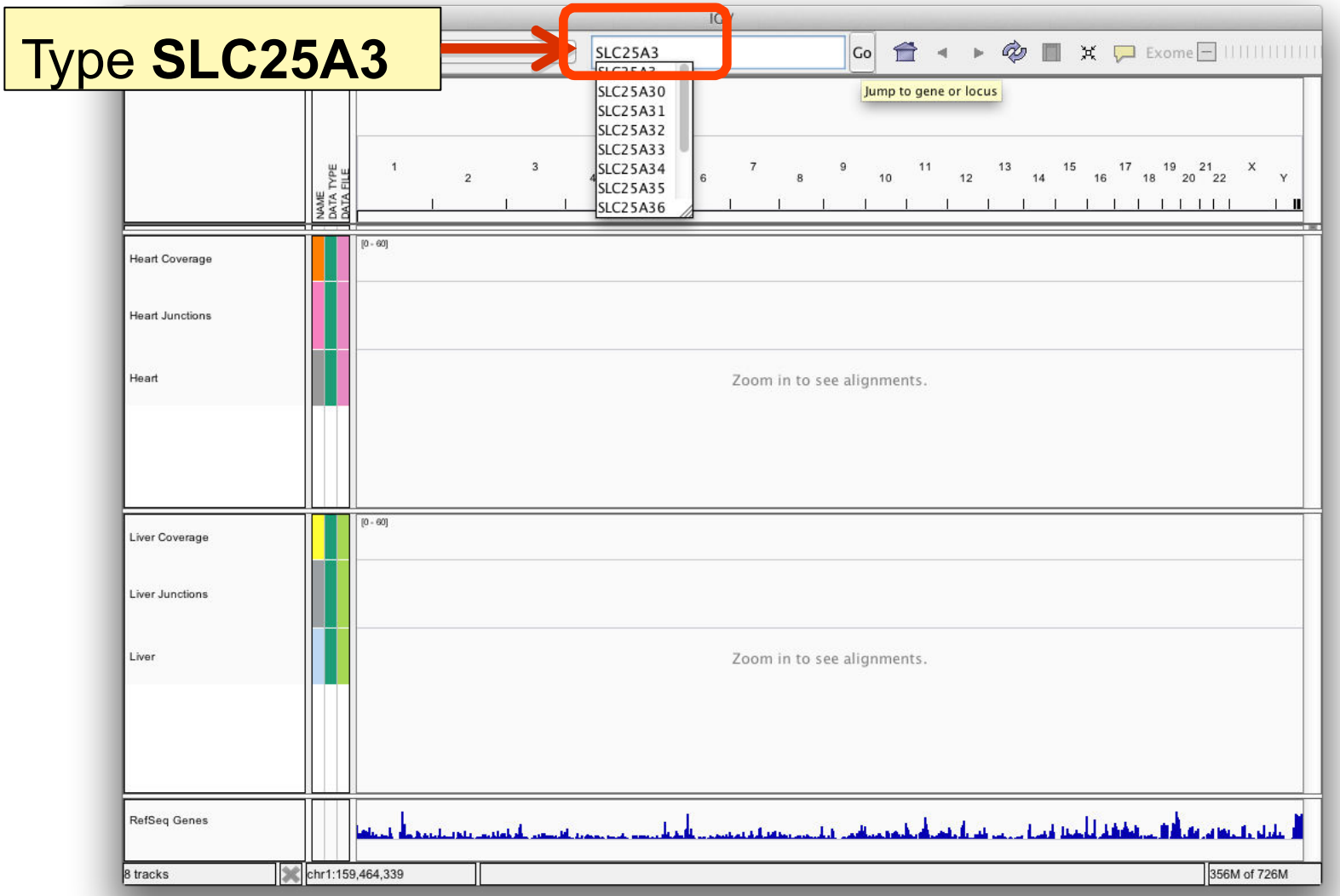
Integrative Genomics Viewer




BROAD INSTITUTE

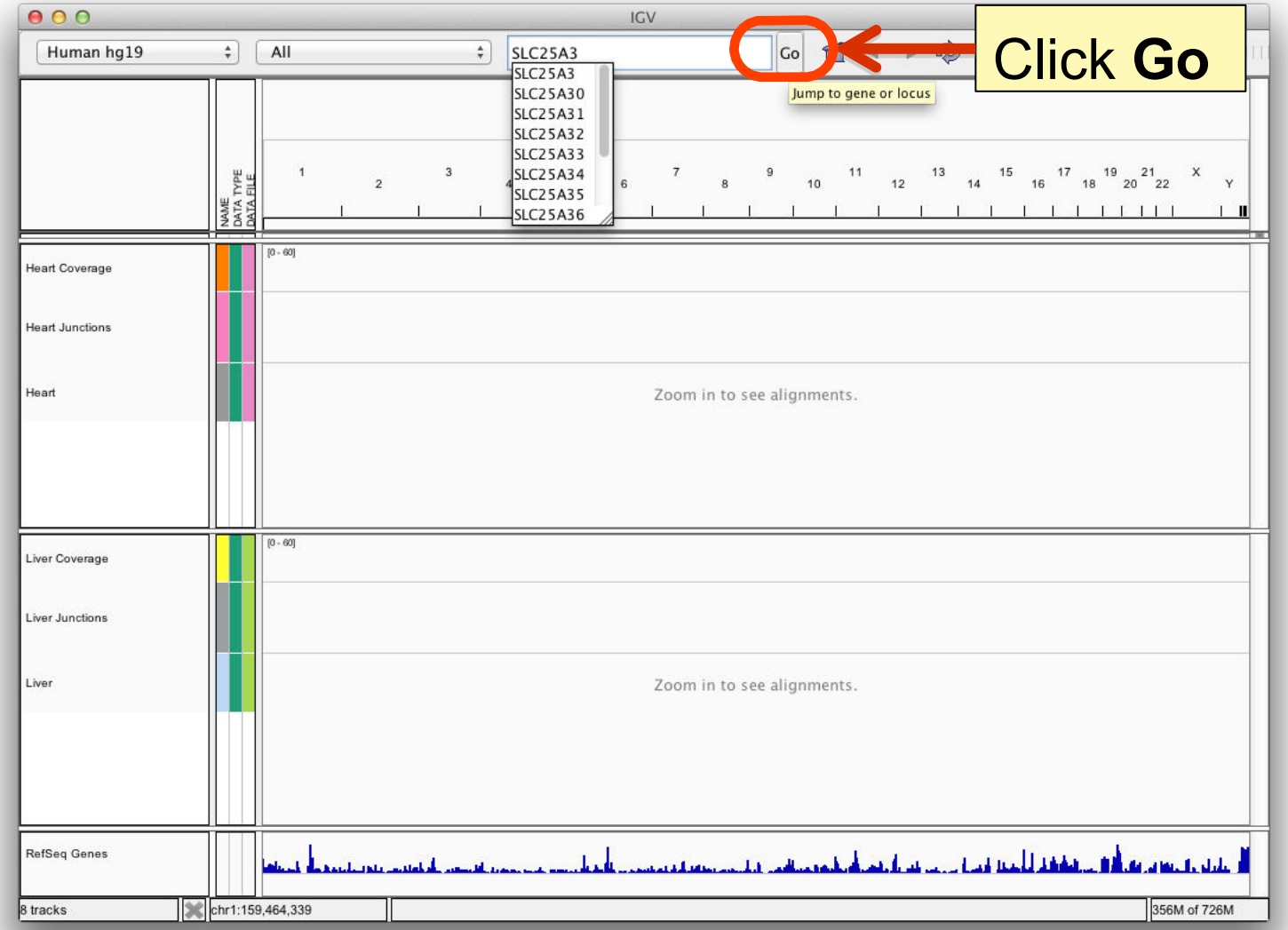
# RNA-seq alignments

Integrative Genomics Viewer



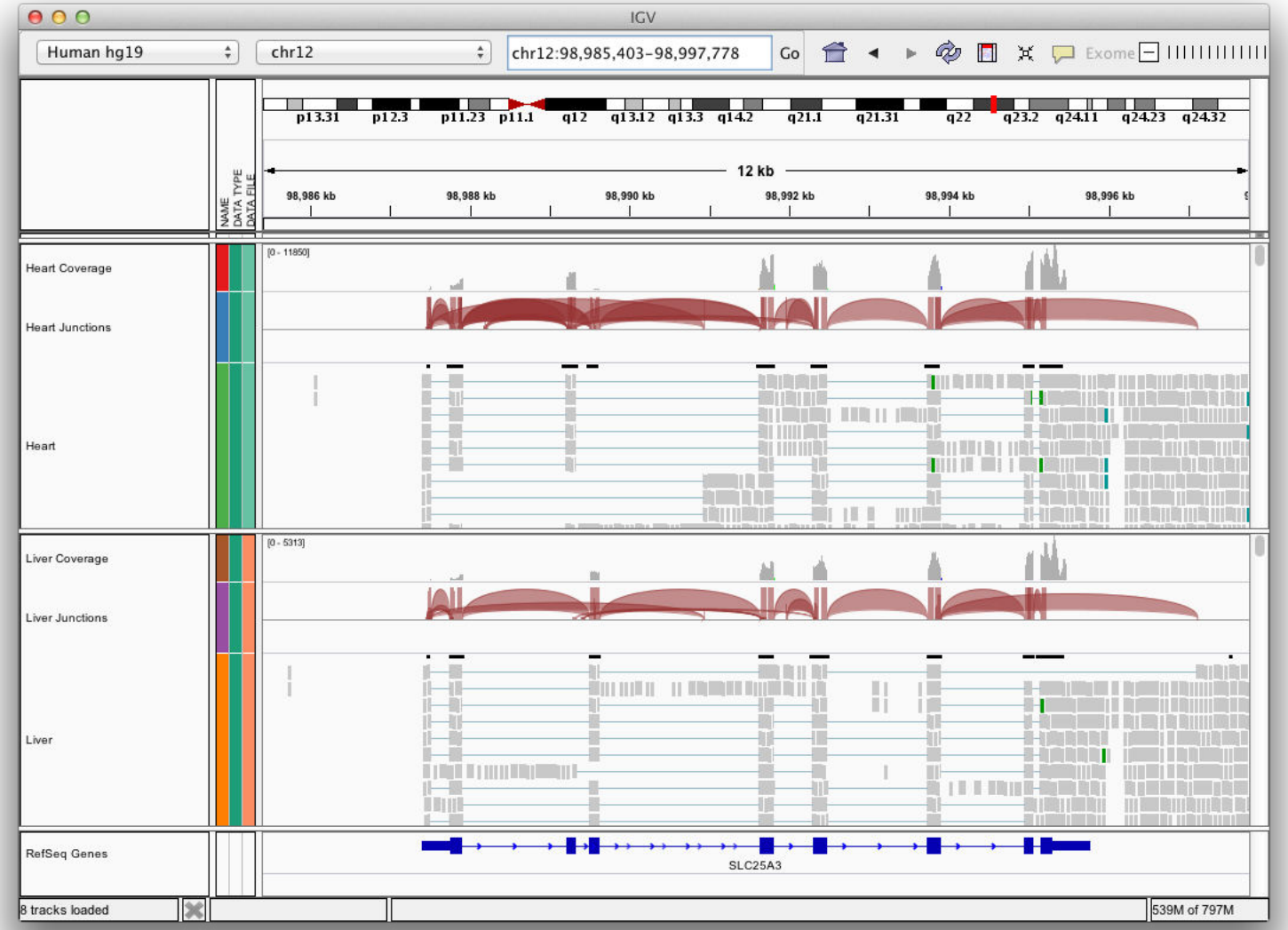

BROAD INSTITUTE

# RNA-seq alignments



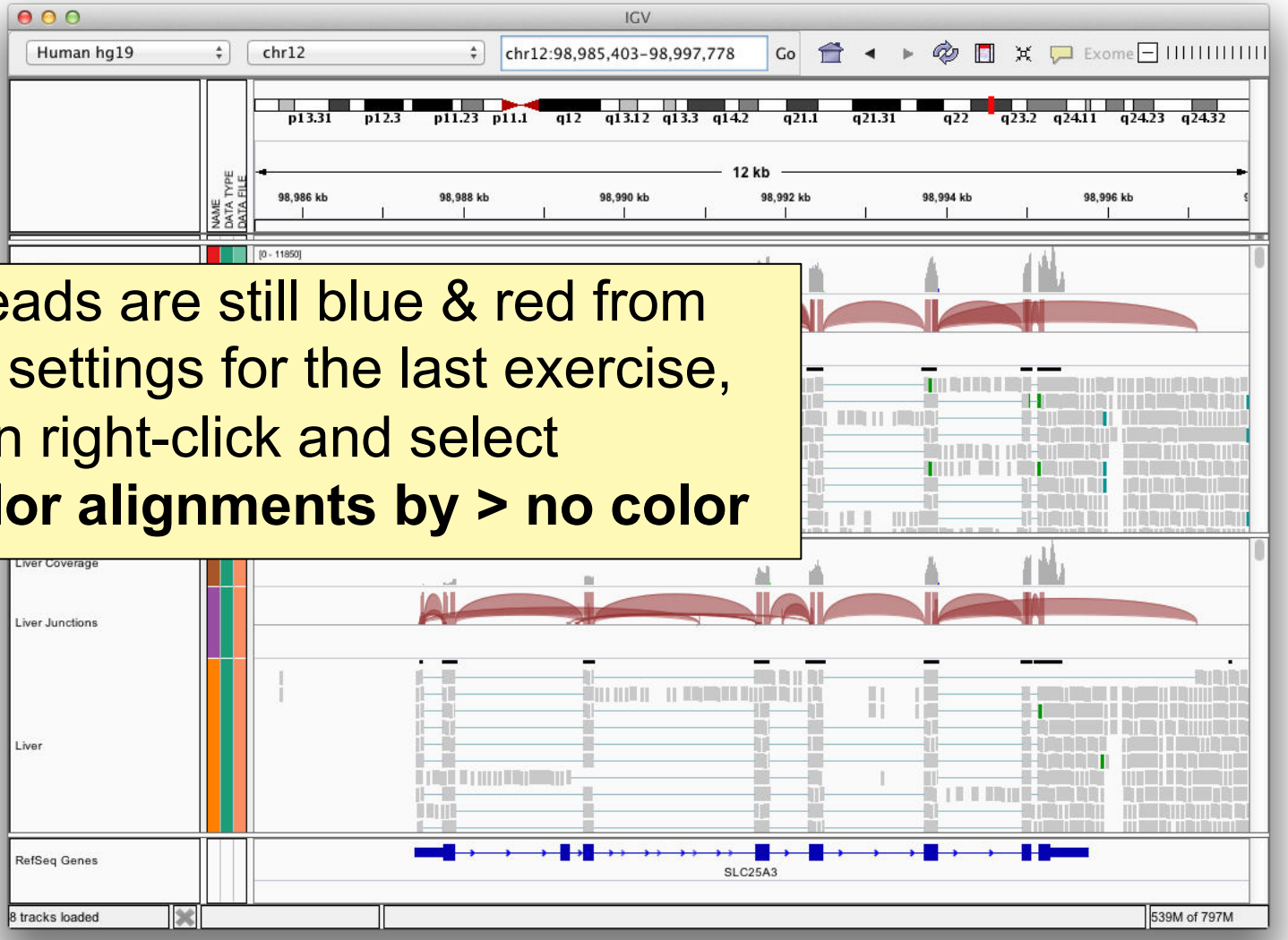
# RNA-seq alignments

Integrative Genomics Viewer




BROAD INSTITUTE

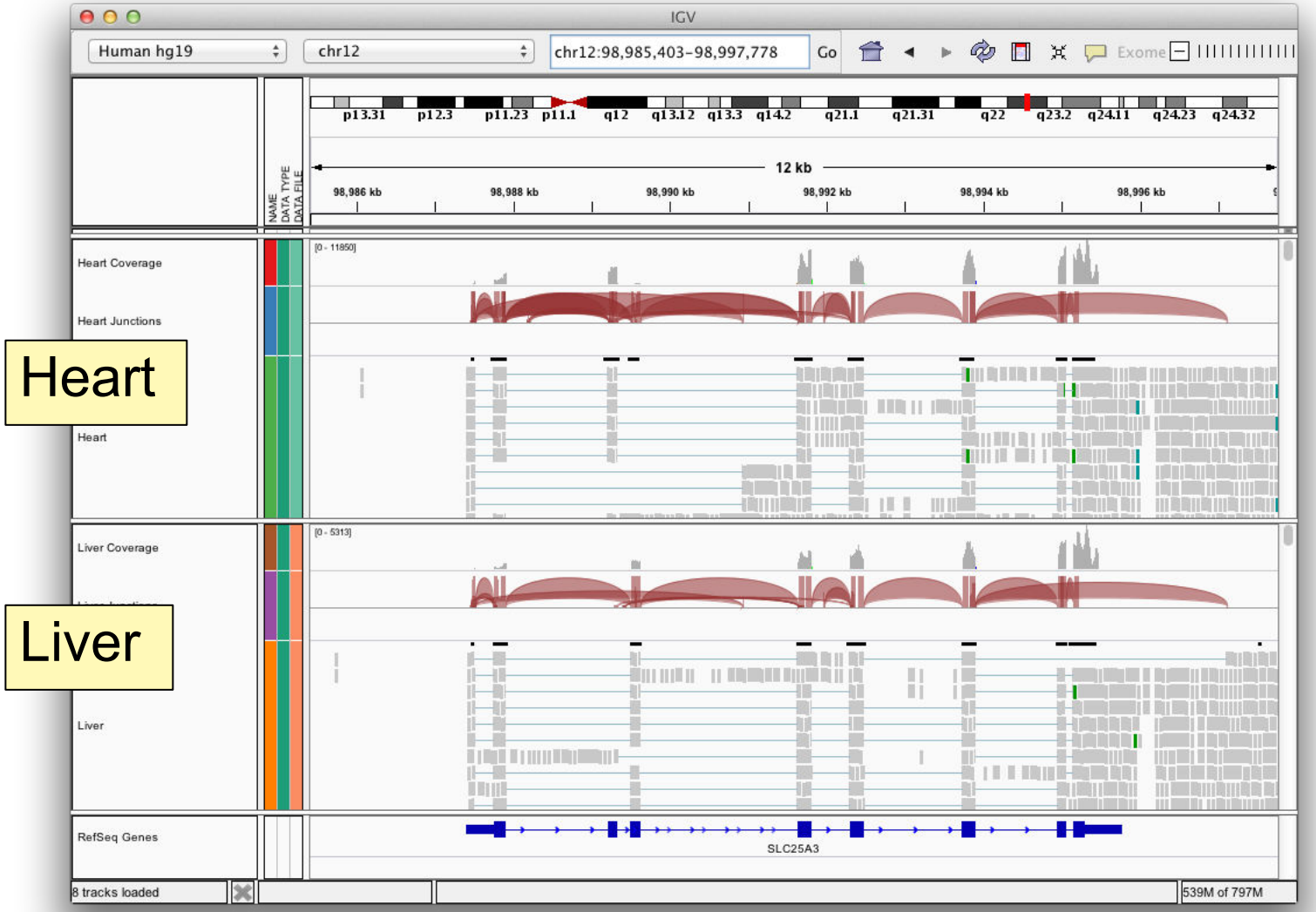
# RNA-seq alignments




If reads are still blue & red from the settings for the last exercise, then right-click and select **Color alignments by > no color**

# RNA-seq alignments

Integrative Genomics Viewer

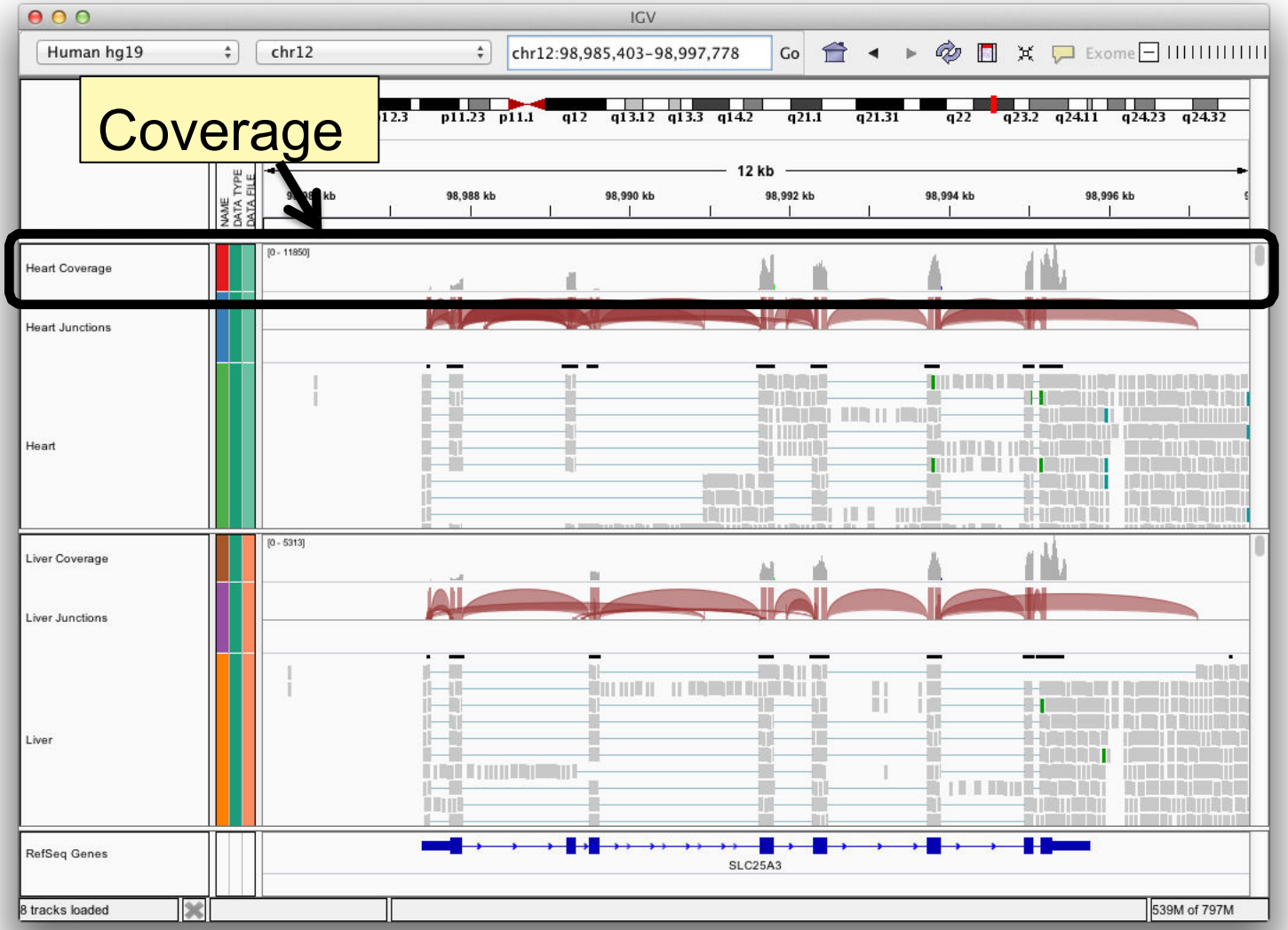



Heart

Liver

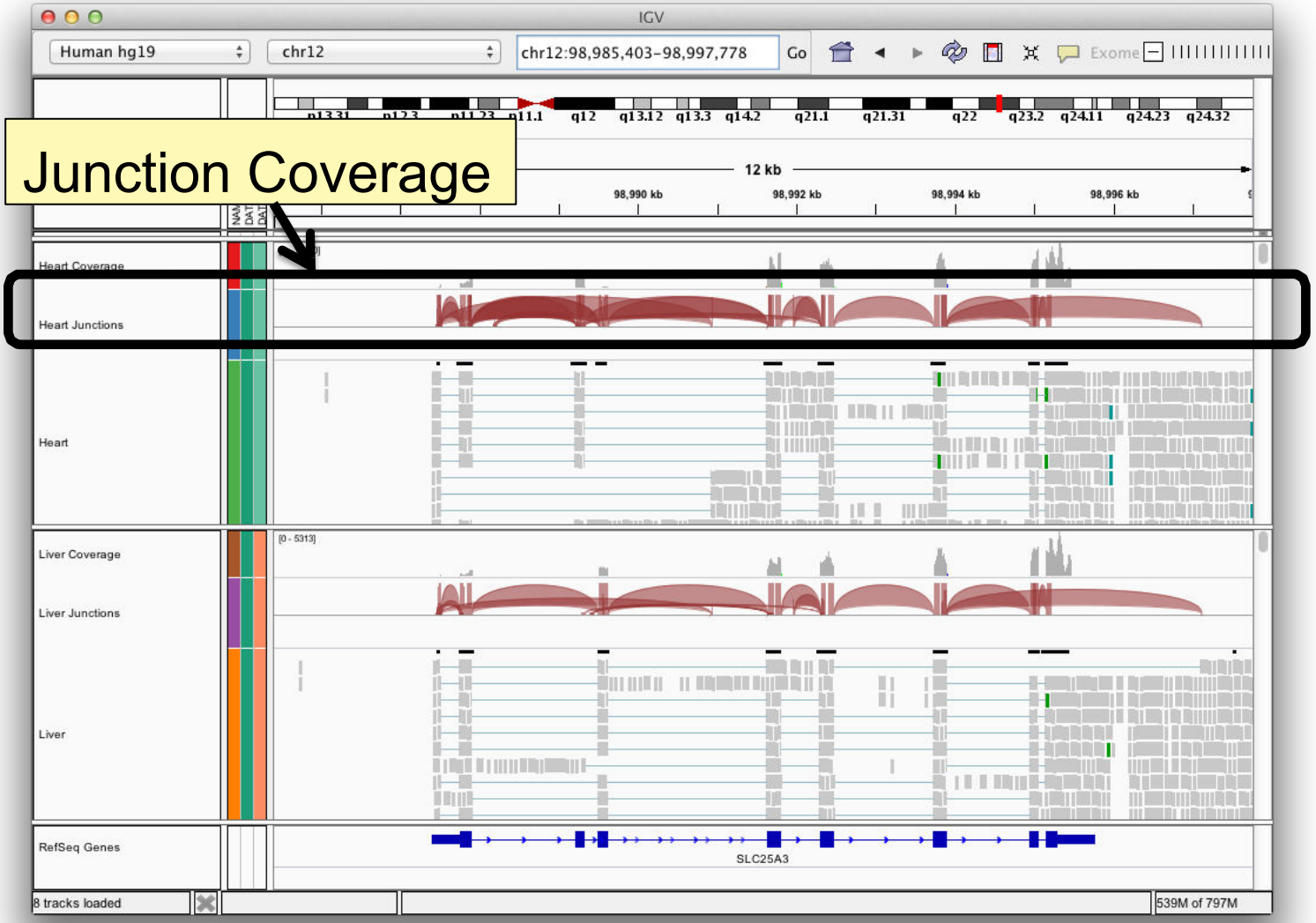
BROAD INSTITUTE

# RNA-seq alignments



# RNA-seq alignments

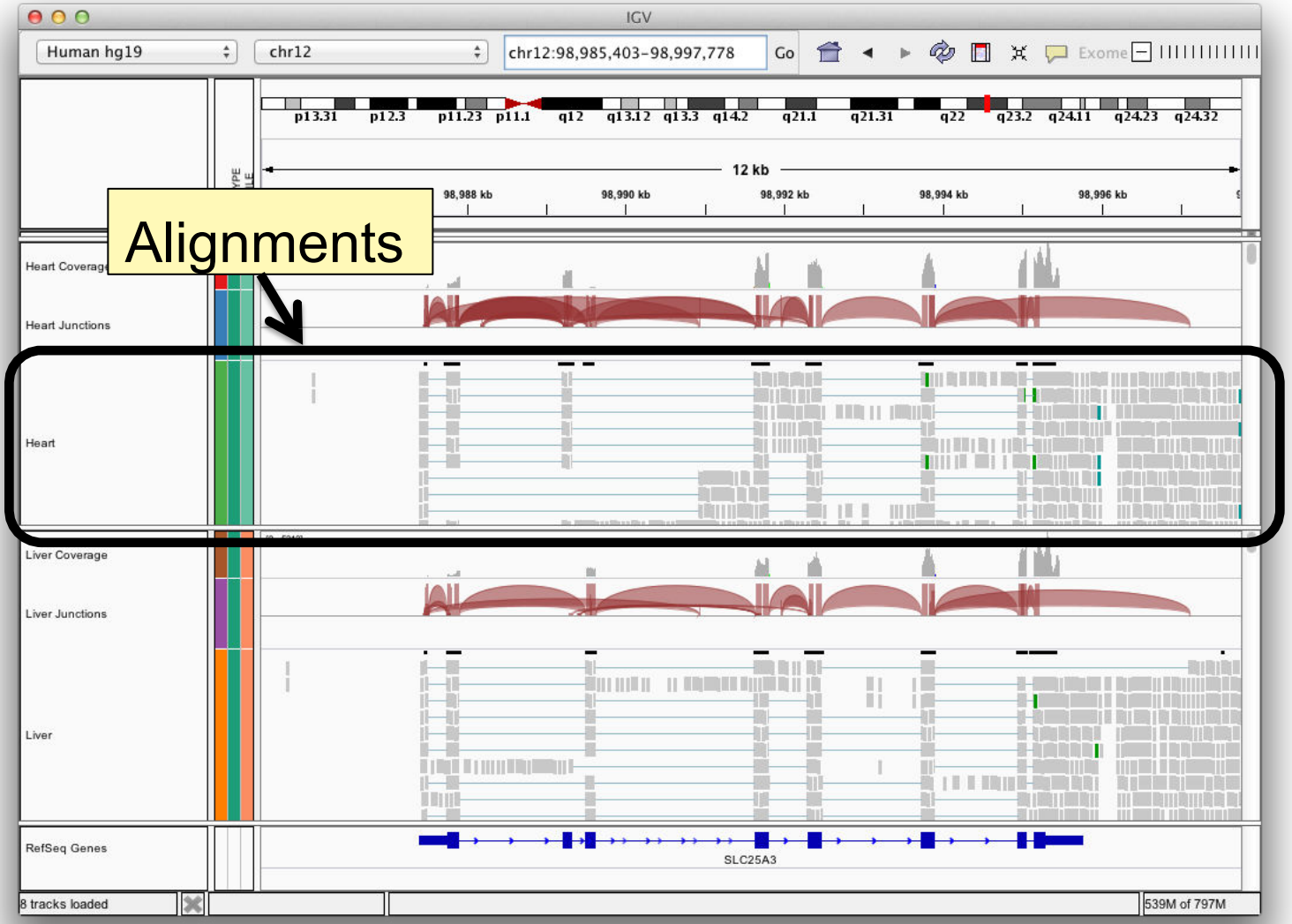
Integrative Genomics Viewer




BROAD INSTITUTE

# RNA-seq alignments

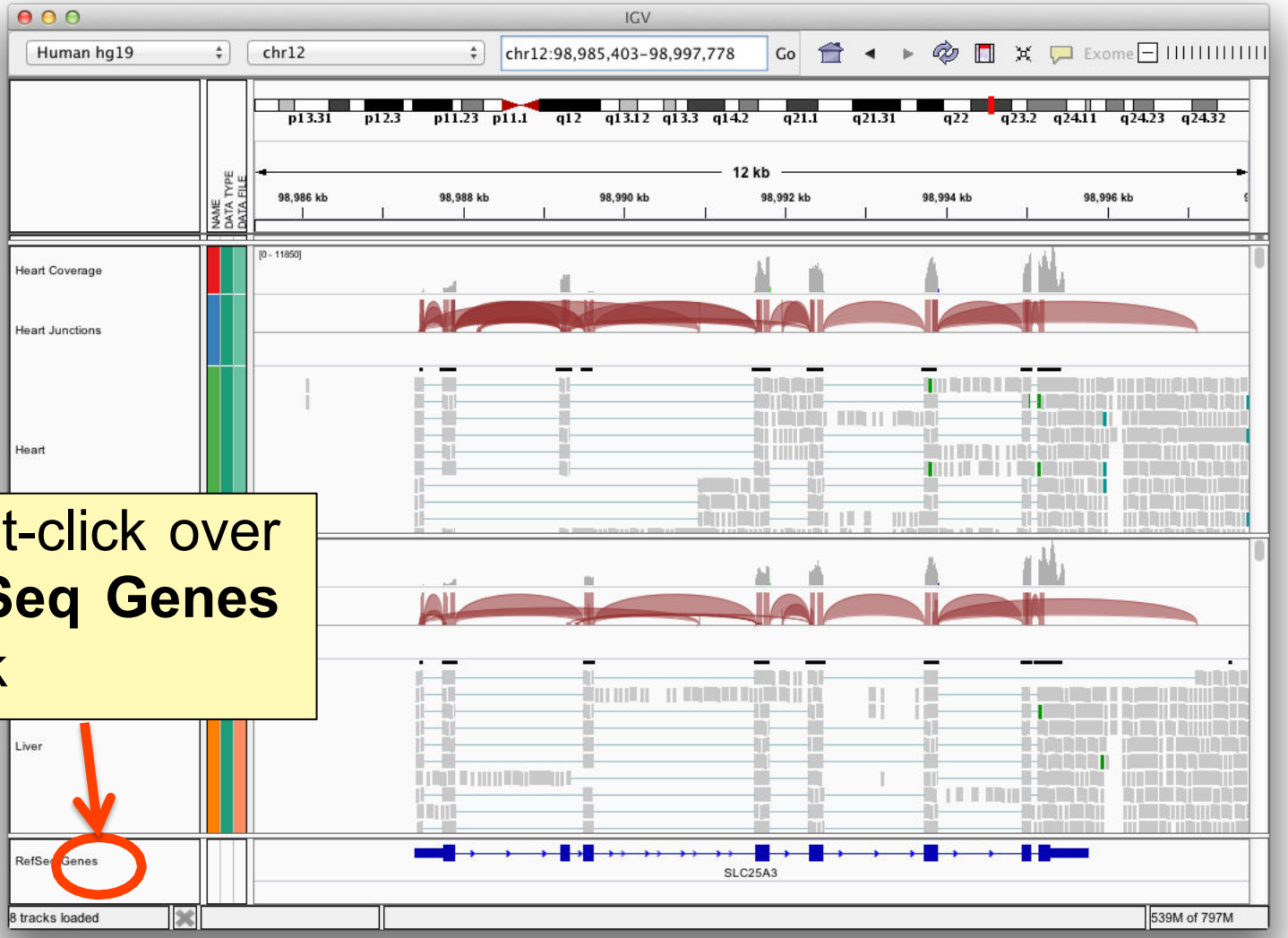
Integrative Genomics Viewer




BROAD INSTITUTE

# RNA-seq alignments

Integrative Genomics Viewer

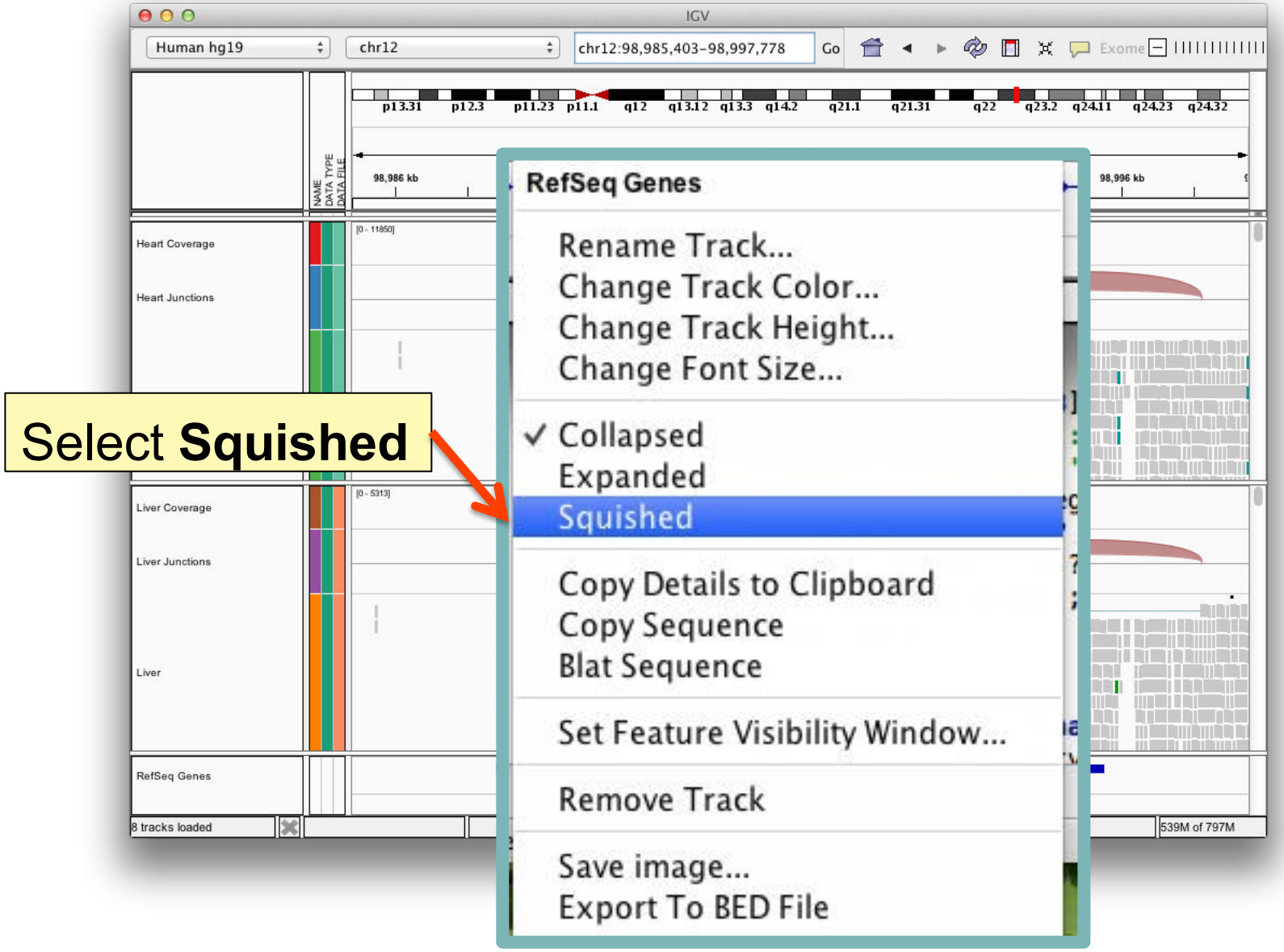



Right-click over **RefSeq Genes** track

BROAD INSTITUTE

# RNA-seq alignments

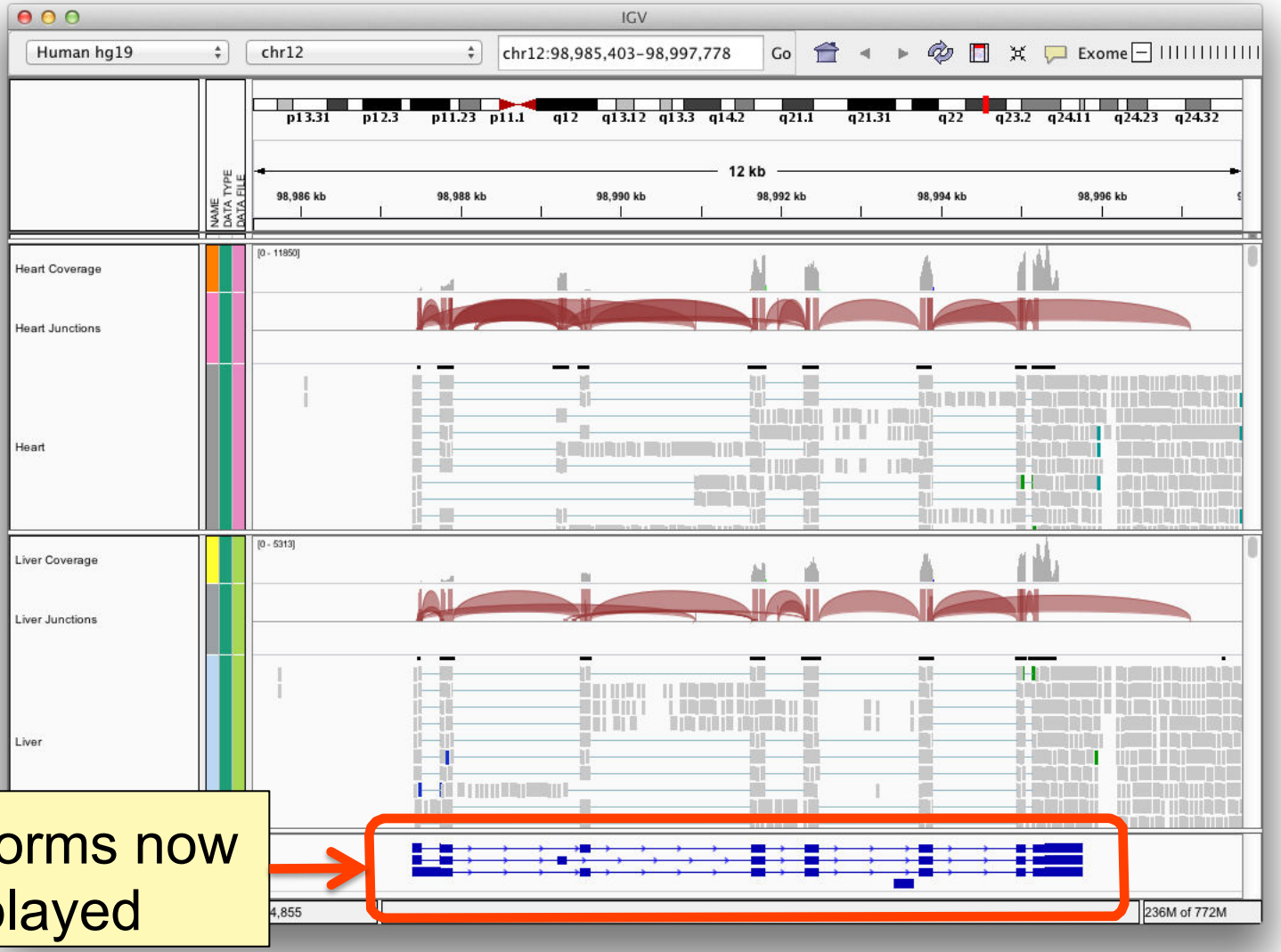
Integrative Genomics Viewer




BROAD INSTITUTE

# RNA-seq alignments

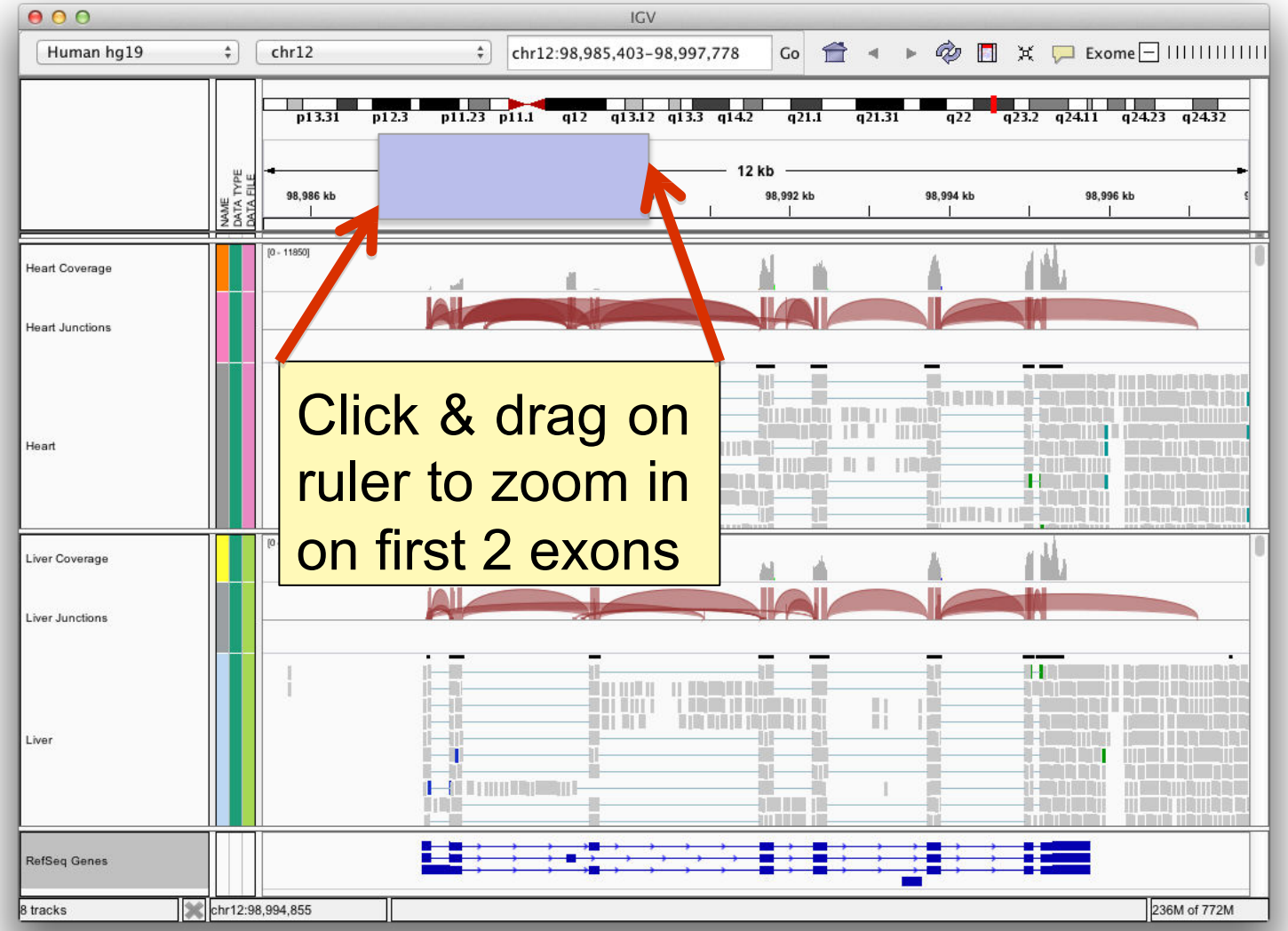
Integrative Genomics Viewer




Isoforms now displayed

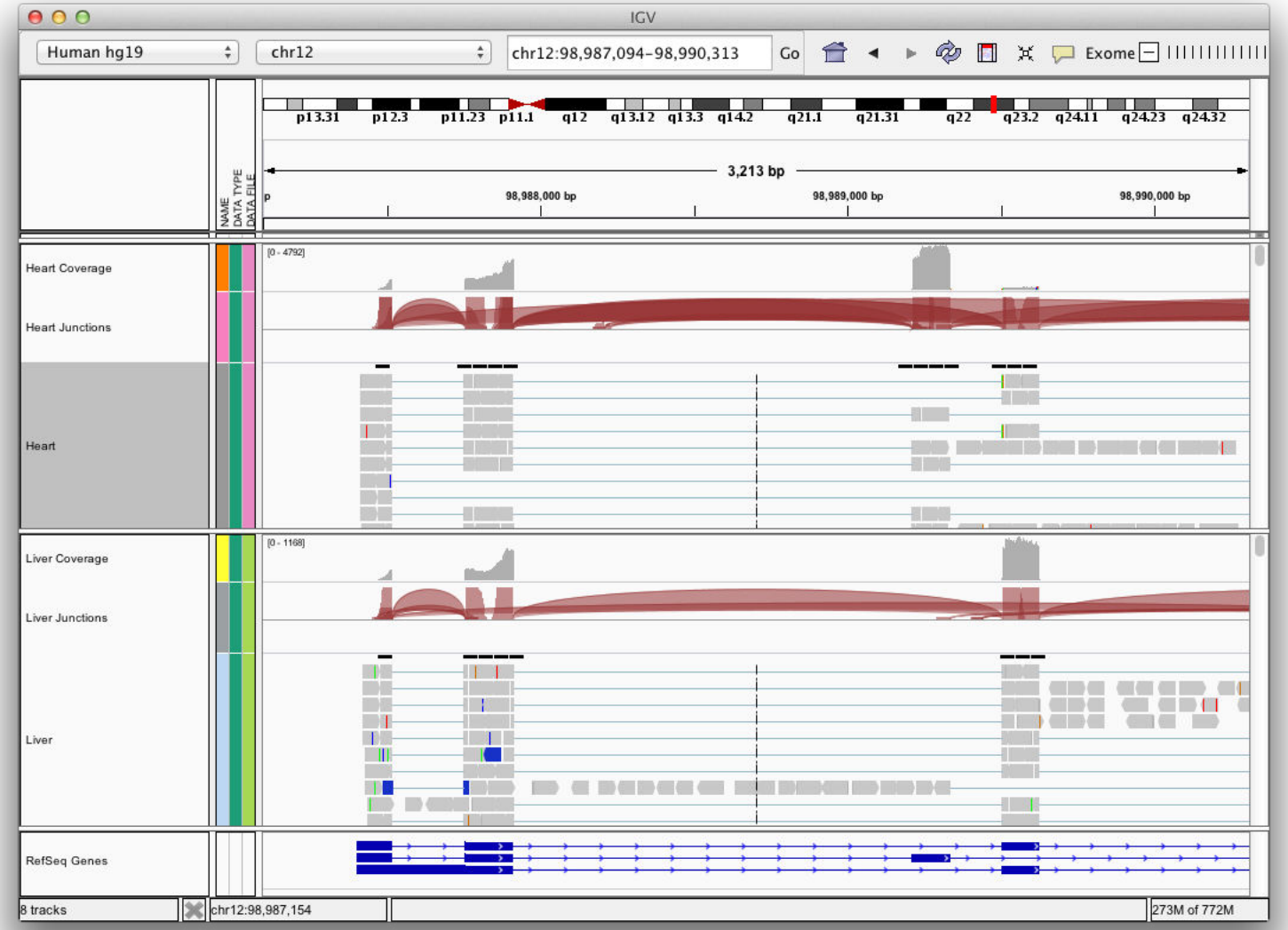
BROAD INSTITUTE

# RNA-seq alignments



# RNA-seq alignments

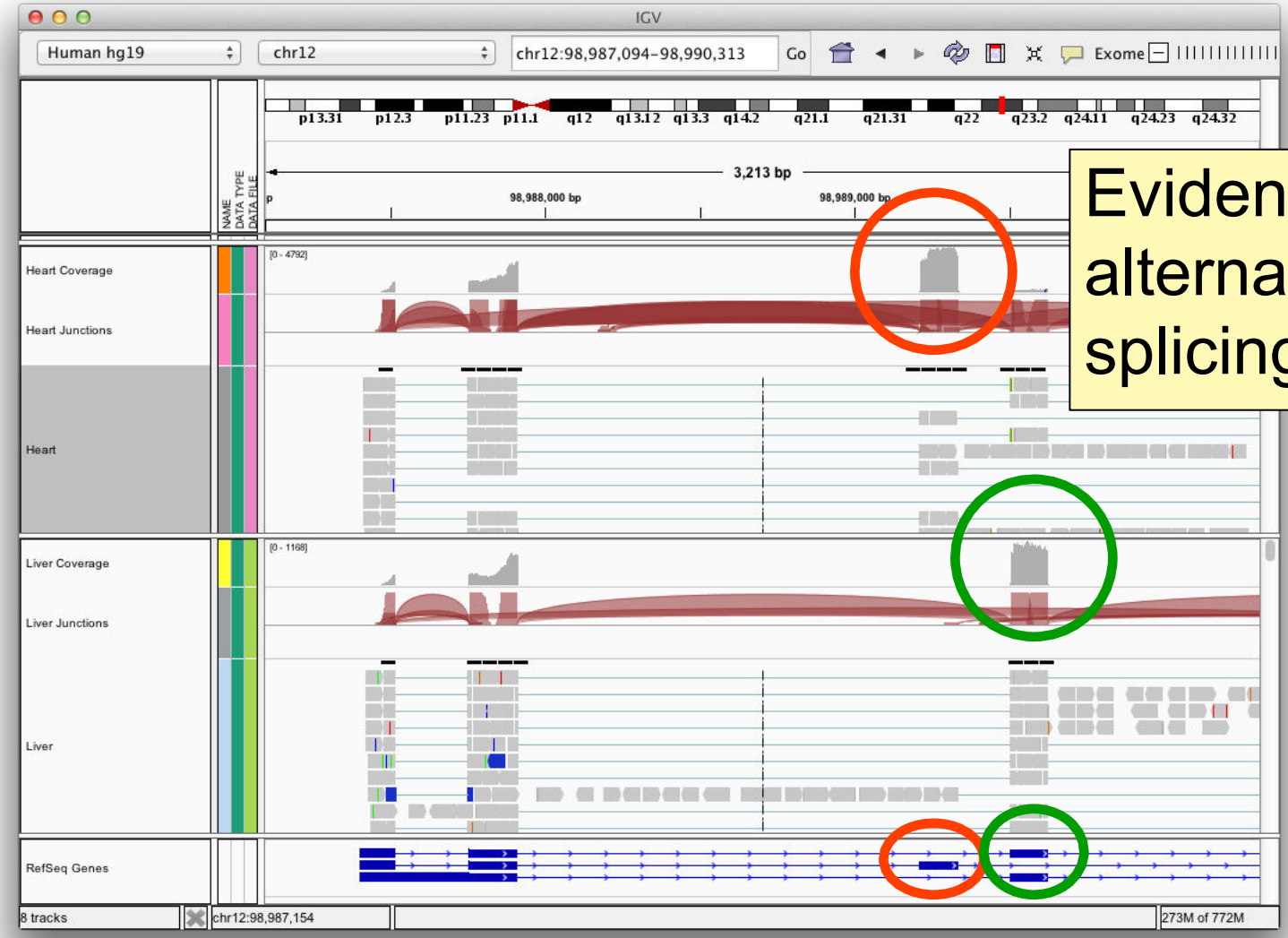
Integrative Genomics Viewer




BROAD INSTITUTE

# RNA-seq alignments

Integrative Genomics Viewer




Evidence of alternative splicing

BROAD INSTITUTE

# Sashimi plot

Integrative Genomics Viewer

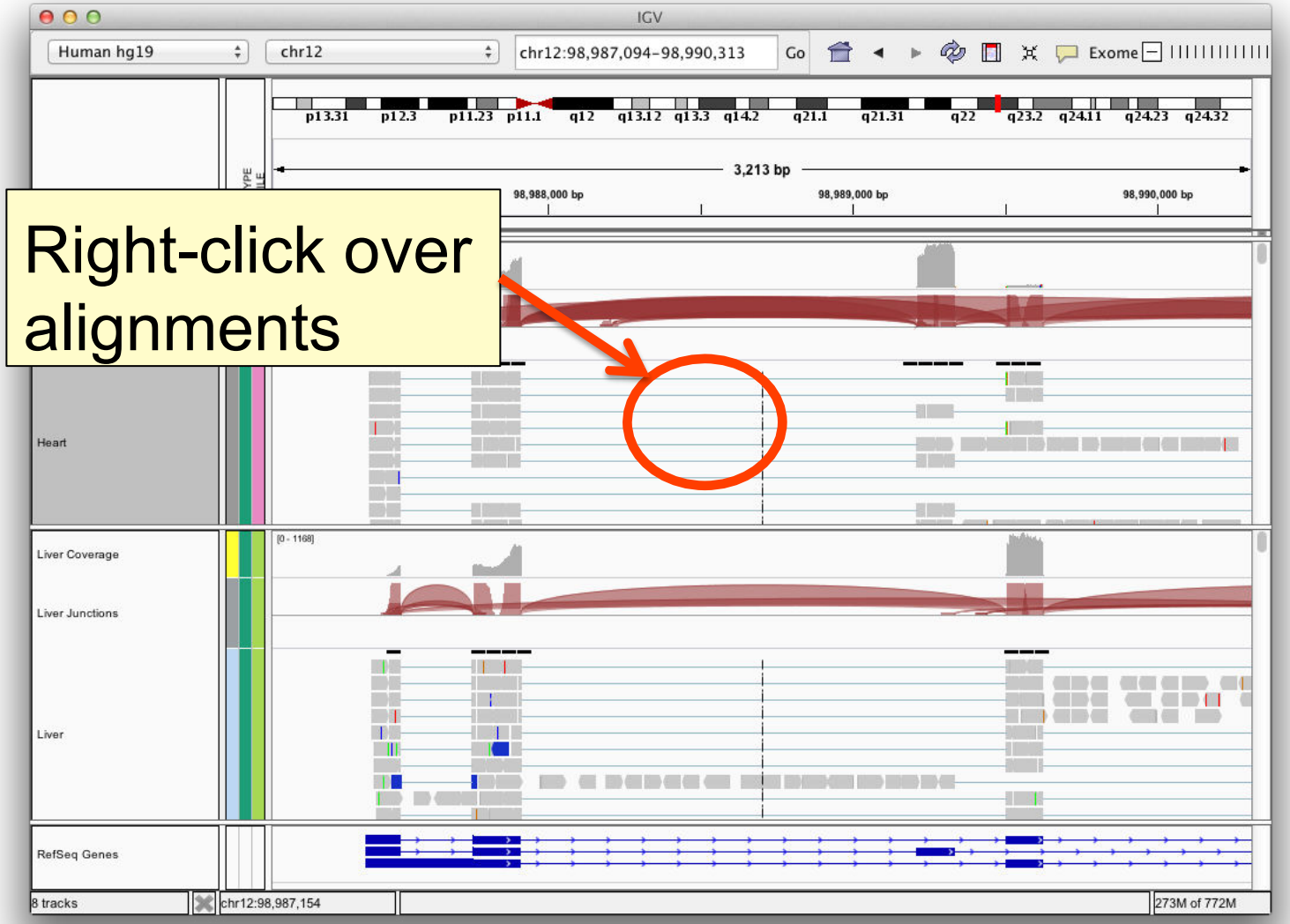
Viewing RNA splicing with Sashimi Plots

Reference: Katz Y, Wang ET, Silterra J, Schwartz S, Wong B, Mesirov JP, Airoldi EM, Burge, CB. ***Sashimi plots: Quantitative visualization of RNA sequencing read alignments.*** arXiv:1306.3466 [q-bio.GN], 2013

BROAD INSTITUTE

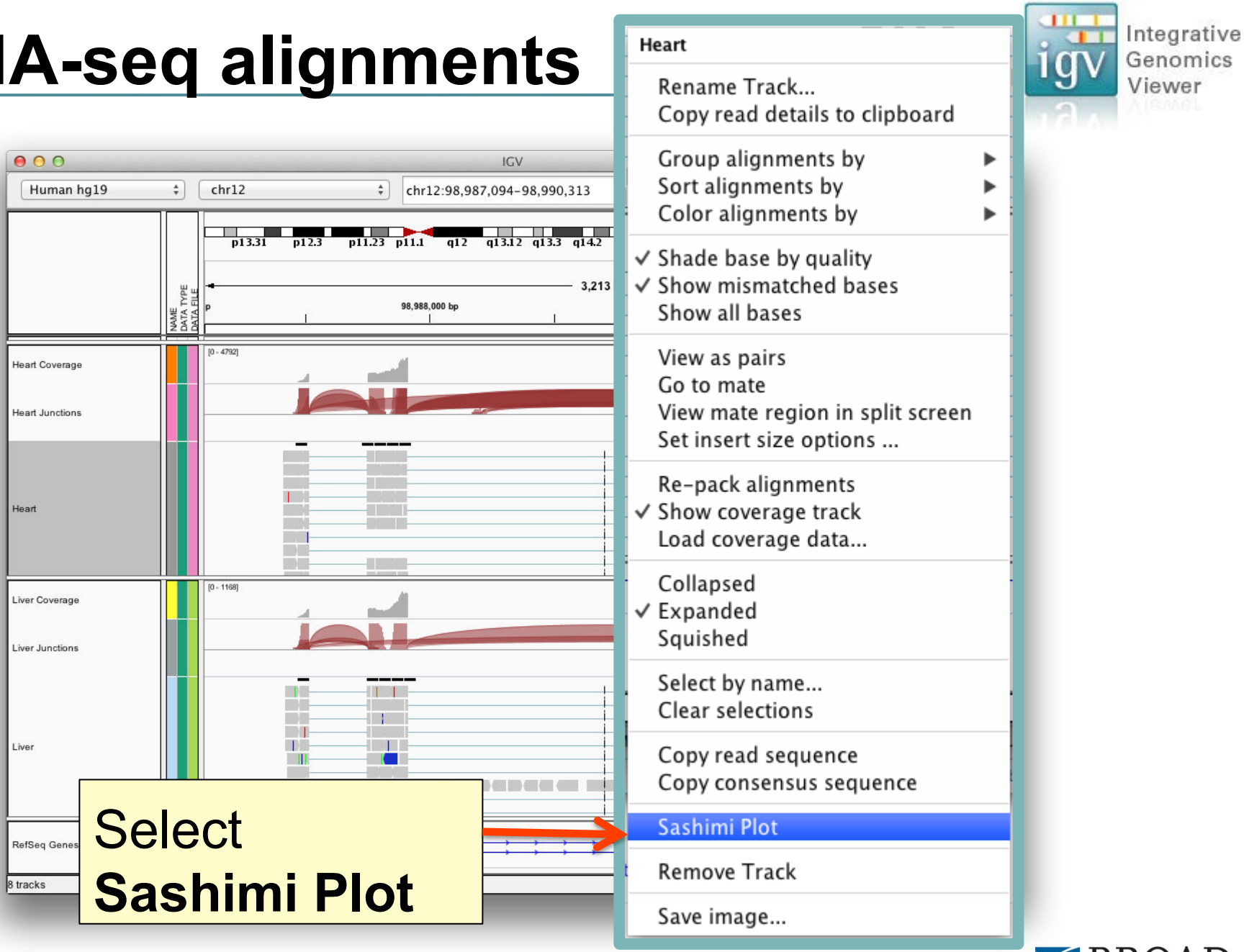
# RNA-seq alignments

Integrative Genomics Viewer




BROAD INSTITUTE

# RNA-seq alignments

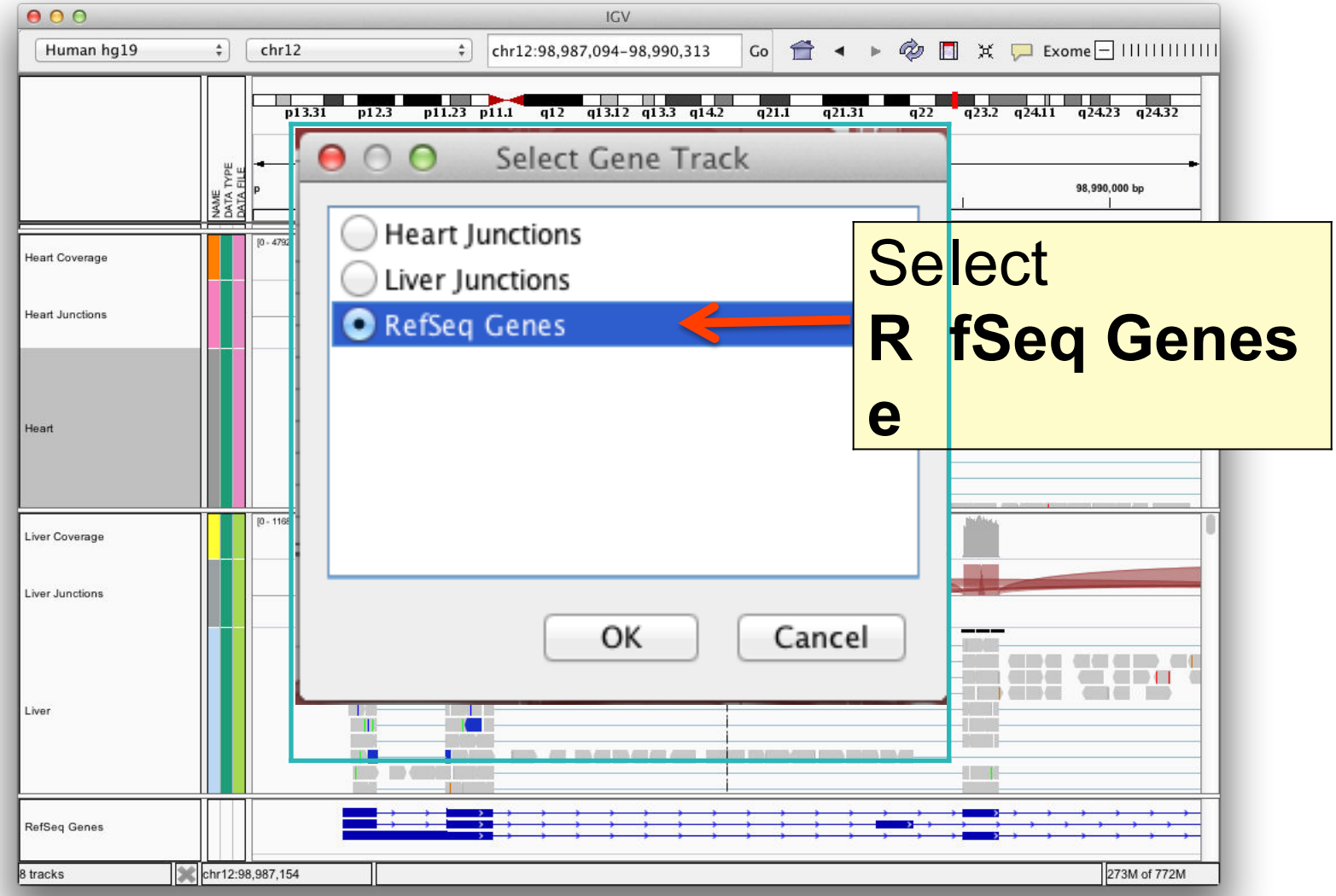



Integrative Genomics Viewer

BROAD INSTITUTE

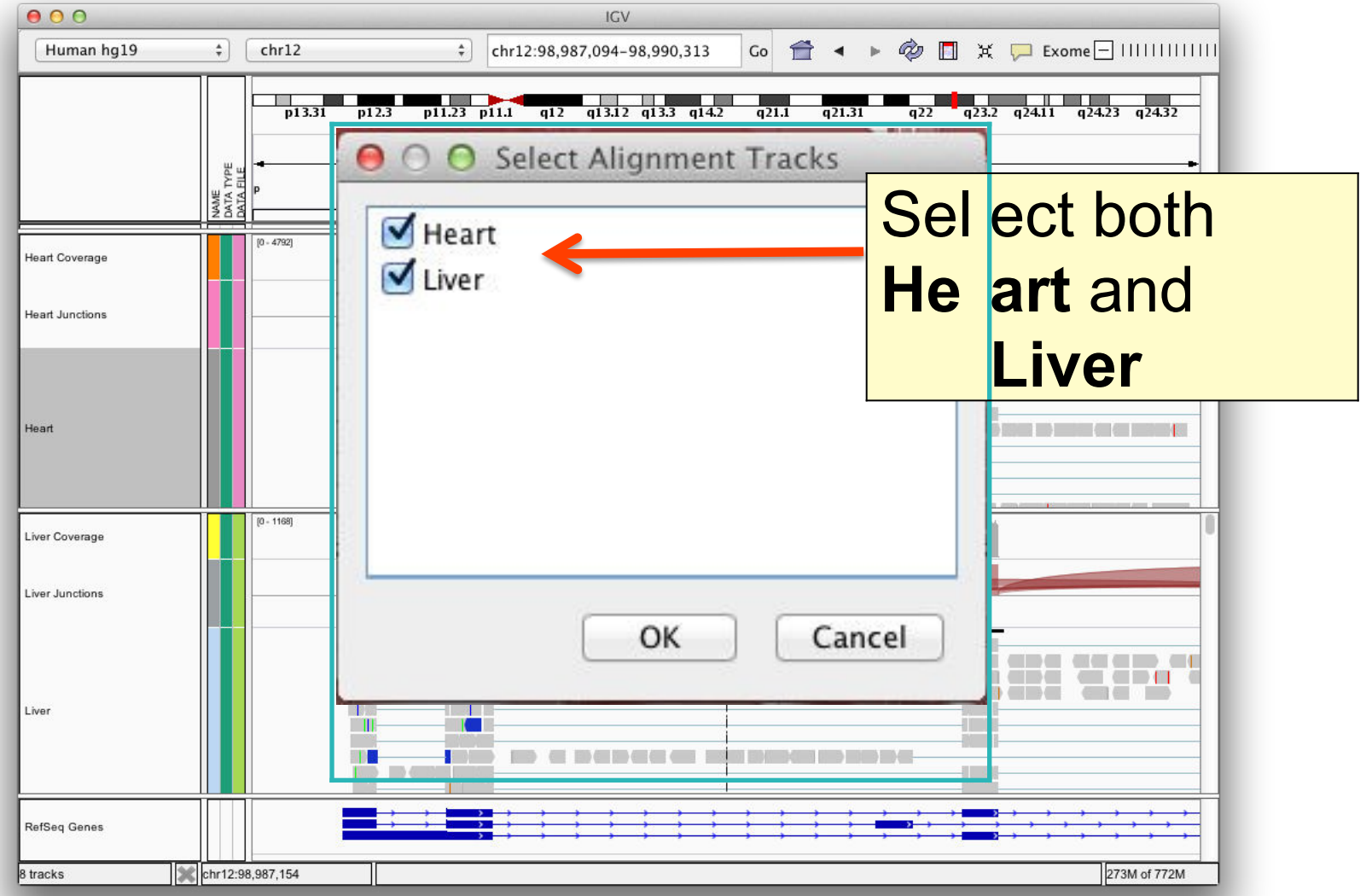
# RNA-seq alignments

Integrative Genomics Viewer




BROAD INSTITUTE

# RNA-seq alignments




Select both **Heart** and **Liver**

# RNA-seq alignments

Integrative Genomics Viewer


BROAD INSTITUTE

# RNA-seq alignments

Integrative Genomics Viewer


A little busy. Let's filter out low-count events

BROAD INSTITUTE

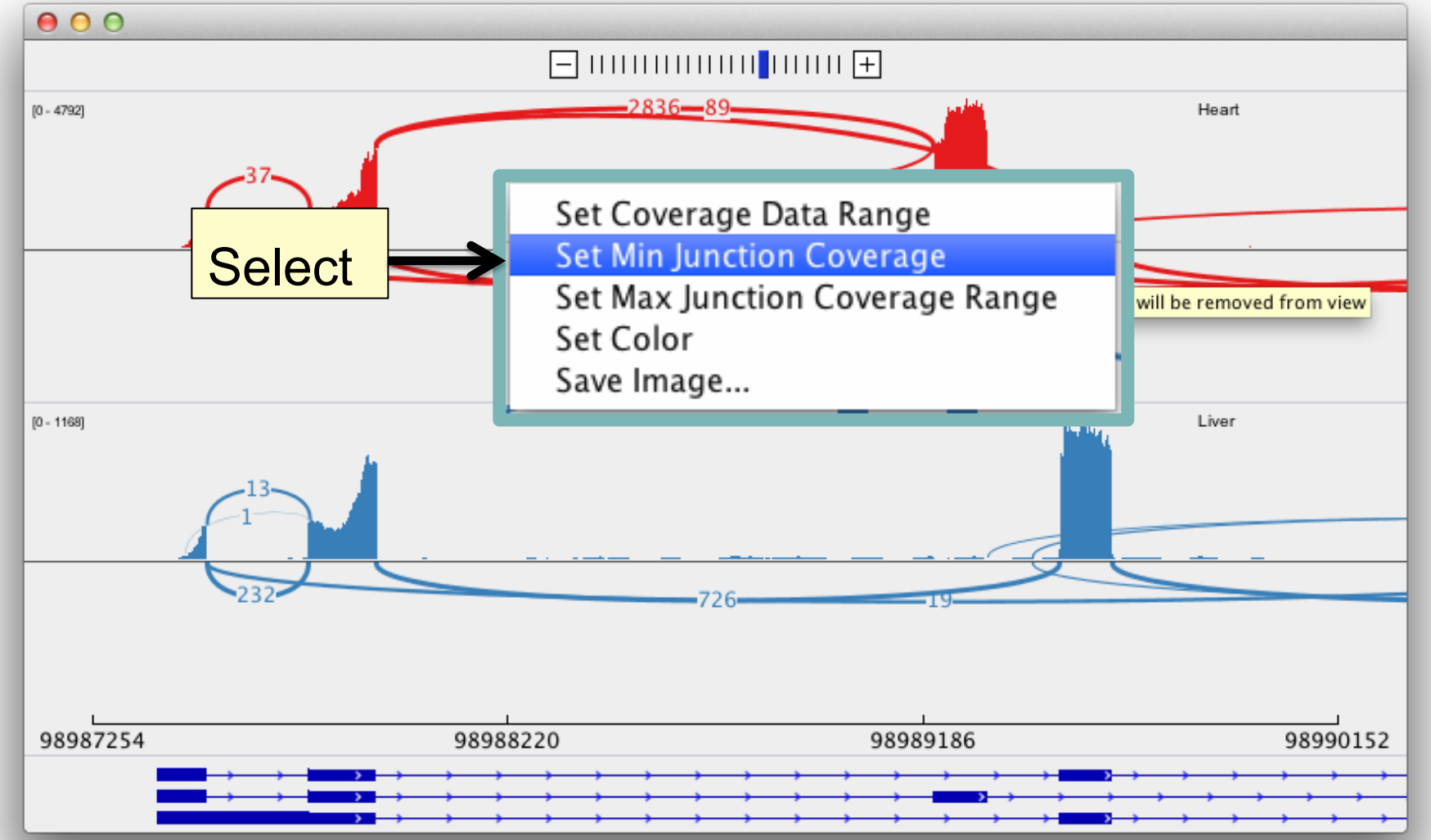
# RNA-seq alignments

Integrative Genomics Viewer


BROAD INSTITUTE

# RNA-seq alignments

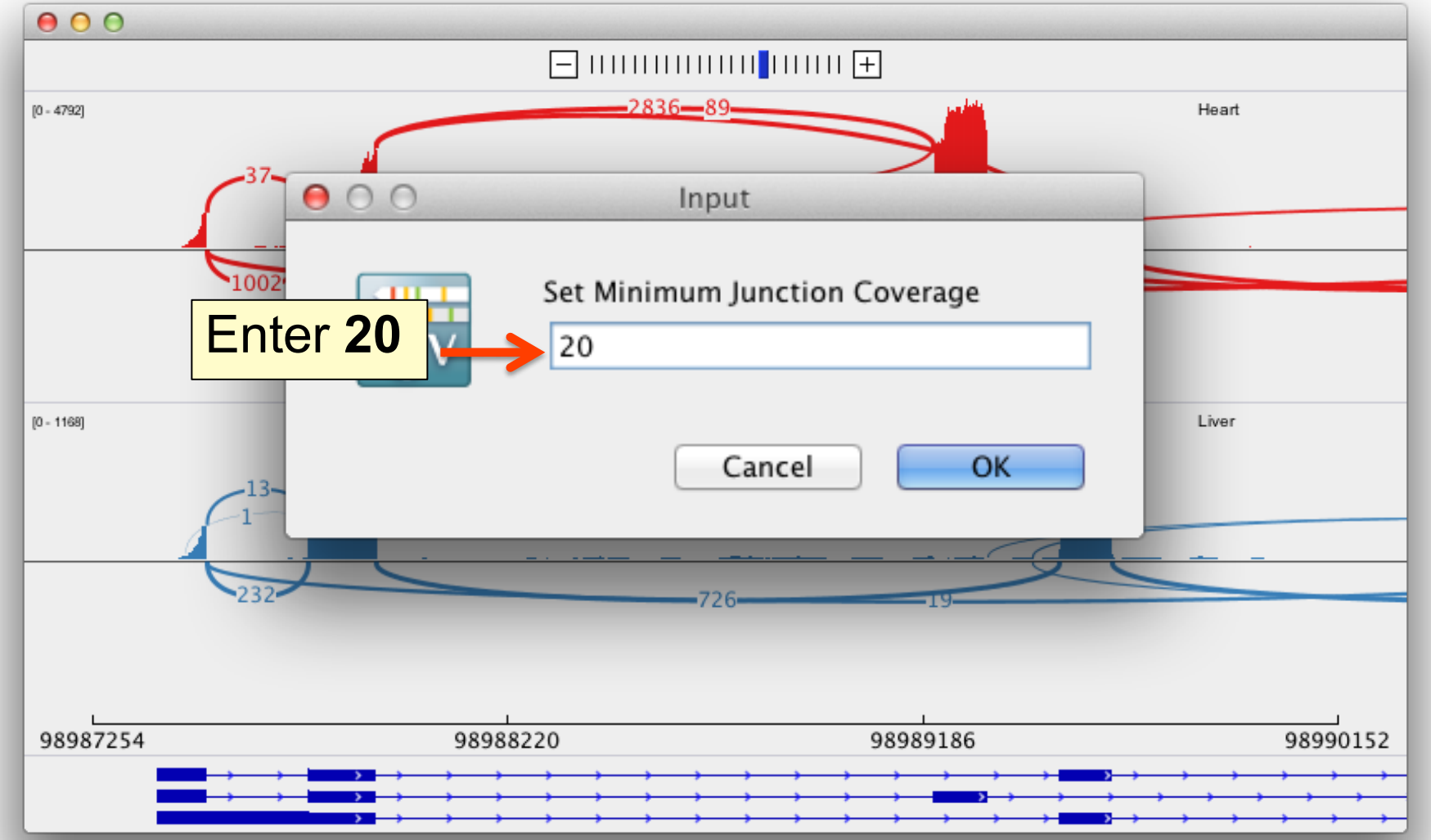
Integrative Genomics Viewer




BROAD INSTITUTE

# RNA-seq alignments

Integrative Genomics Viewer




BROAD INSTITUTE

# RNA-seq alignments

Integrative Genomics Viewer


BROAD INSTITUTE

# igvtools

BROAD INSTITUTE

Integrative Genomics Viewer

# igvtools


A set of utilities for preparing files for efficient display.

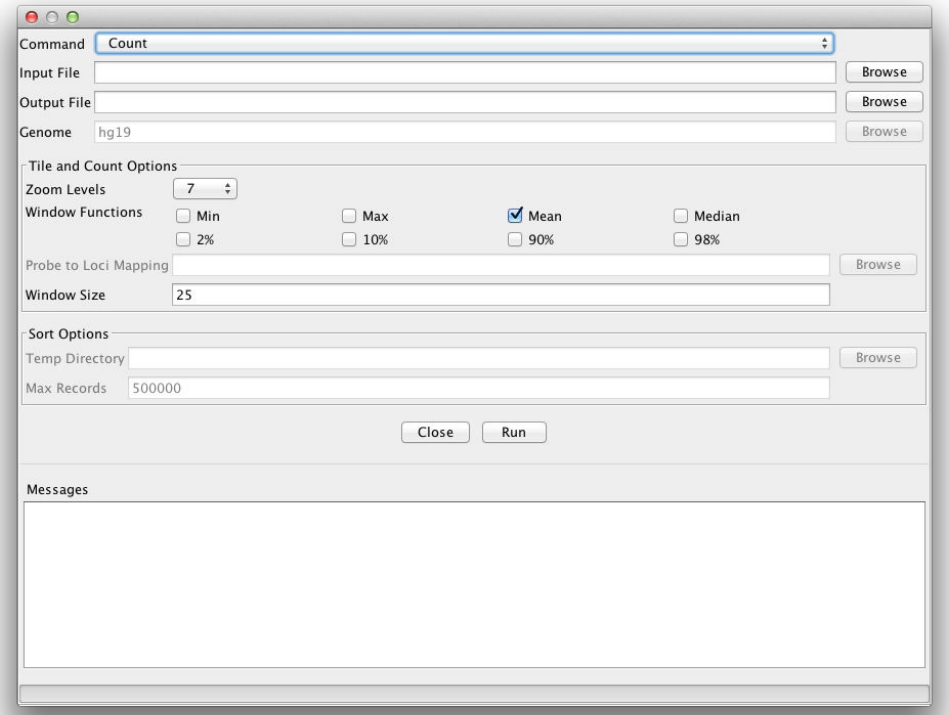
| | |
|---|---|
| **toTDF** | • Converts sorted data file to a binary tiled data file (TDF).<br>• Supported file formats: .wig, .cn, .snp, .igv, .gct |
| **count** | • Computes average alignment or feature density over a specified window size across the genome.<br>• Supported file formats: .sam, .bam, .aligned, .sorted.txt, .bed |
| **sort** | • Sorts file by genomic start position.<br>• Supported file formats: .cn, .igv, .sam, .aligned, .bed. |
| **index** | • Creates an index file for alignment or feature file.<br>• Supported file formats: .sam, .aligned, .sorted.txt, .bed |

BROAD INSTITUTE

# igvtools

Integrative Genomics Viewer

- Can be launched from the IGV user interface *File > Run igvtools…*
- Or run from the command line




BROAD INSTITUTE

# igvtools toTDF

Integrative Genomics Viewer

The **toTDF** utility converts large ASCII data files into tiled data format (.tdf) files.

TDF files have the following advantages:

- Data is indexed for efficient retrieval.
- Data is preprocessed for zoomed out views.
- TDF files are web friendly – large data files can be shared over the web. Only small slices of the file are actually transferred as needed.

BROAD INSTITUTE

# igvtools count

The **count** command is used to transform alignment files to read density TDF files, e.g. for ChIP-Seq, RNA-Seq, and similar alignment counting experiments.


**Alignments**

Alignments in bam/sam, .aligned, or bed format

**Read Density**

TDF format, indexed and optimized for fast retrieval at multiple resolution scales

# igvtools sort


- Sorts IGV-supported genomic formats by start position.
- The index command requires sorted files.

**Example:**
igvtools sort -m 1000000 –t ~/myTmpDir inputFile.sam outputFile.sorted.sam

- Uses combination of memory and disk to handle large files.

  -m = maximum # of lines to hold in memory. When this number is exceeded a temporary file is created.

  -t = directory used to create temporary files during sorting.

BROAD INSTITUTE

# igvtools index

Integrative Genomics Viewer

Creates an index file for viewing large files in bed, gff, or vcf formats. An index is optional for bed or gff files, but required for vcf files.

An alternative indexing tool is “tabix”. Tabix both compresses and indexes genomic files. IGV can read either type of index (igvtools or tabix).

**Example:** igvtools index myFeatures.bed

The index file must remain in the same directory as the input file

BROAD INSTITUTE

# Computing coverage: igvtools


Hands-on exercise

- Compute alignment coverage from a BAM file using igvtools count command.

**Data source**

Illumina BodyMap

BROAD INSTITUTE

Integrative Genomics Viewer

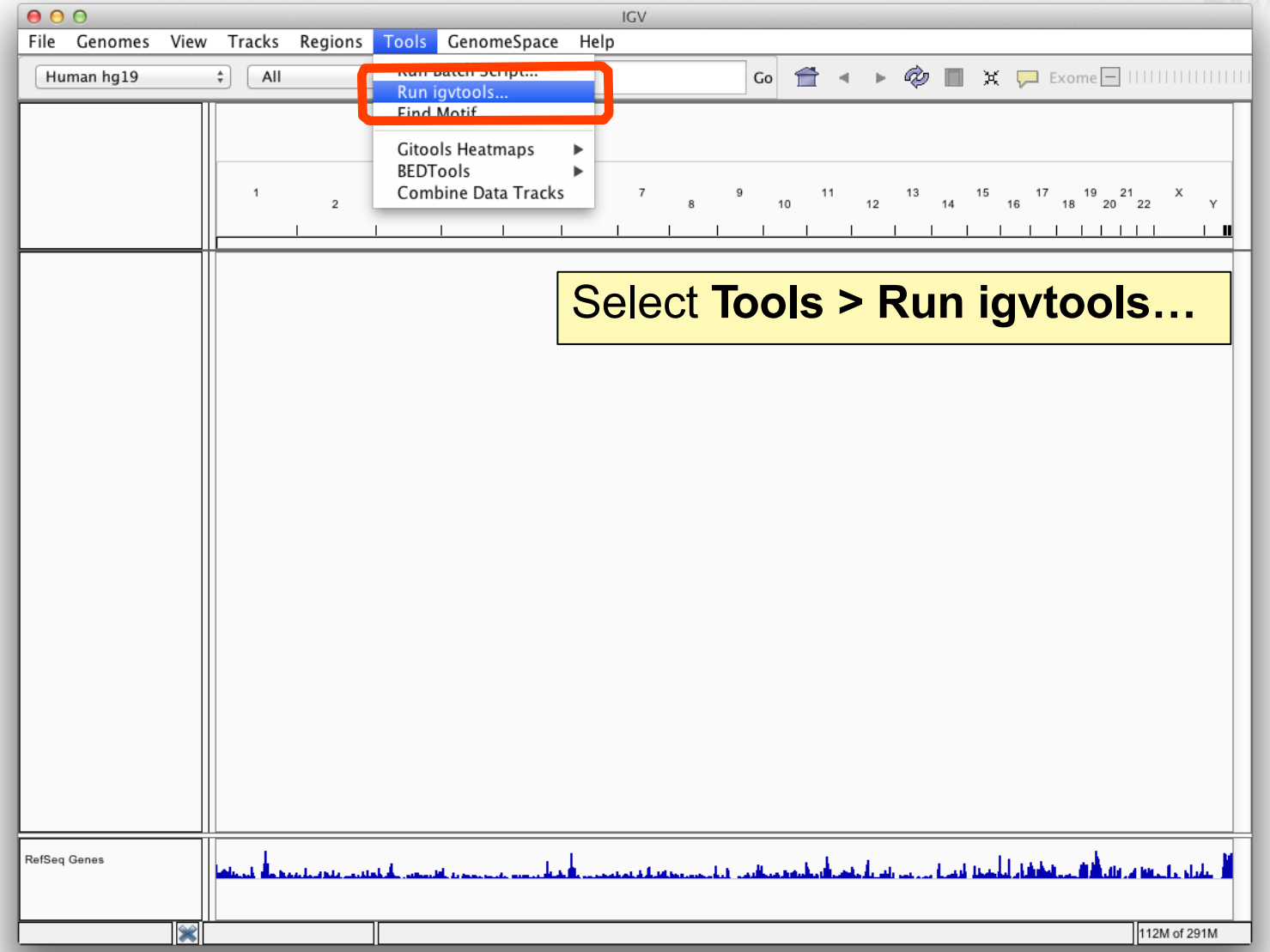
Download data files required for this exercise from:
ftp://ftp.broadinstitute.org/pub/igv/CSH_2013/files.zip

Files included in the zip:

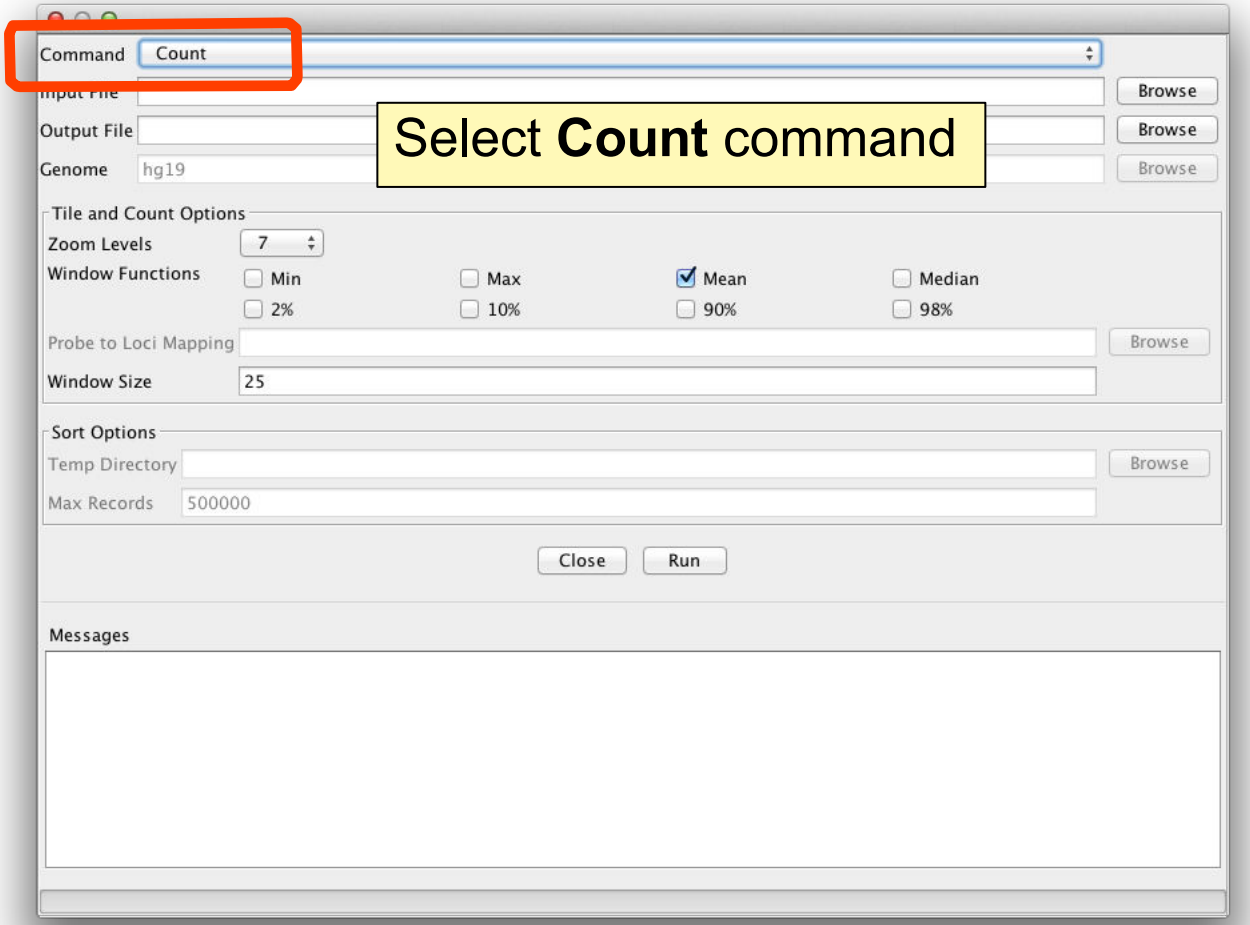
- heart.bodyMap.bam
- heart.bodyMap.bam.bai
- sacCer3.fa (used in next exercise)

BROAD INSTITUTE

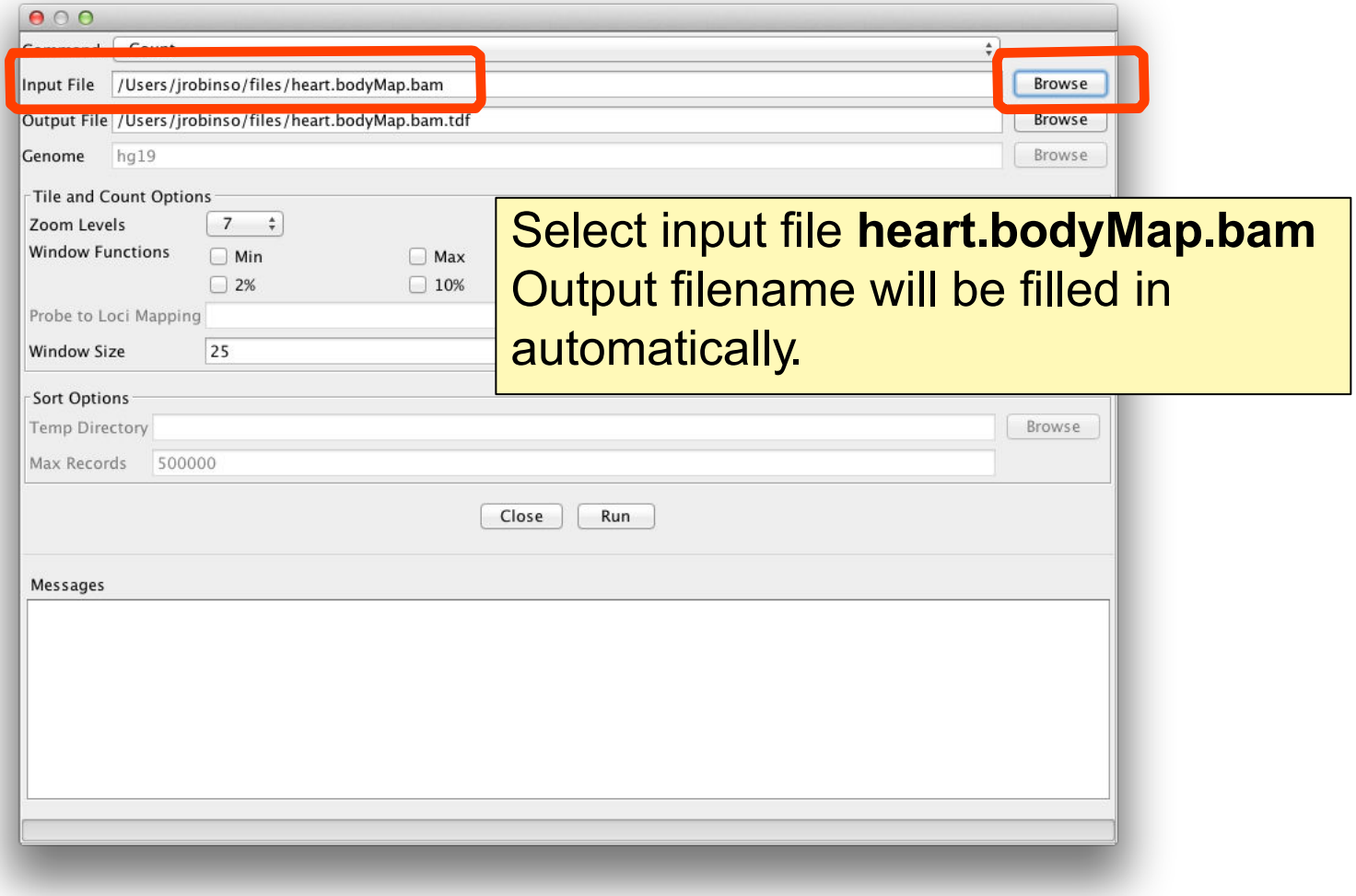
# Computing coverage: igvtools



# Computing coverage: igvtools

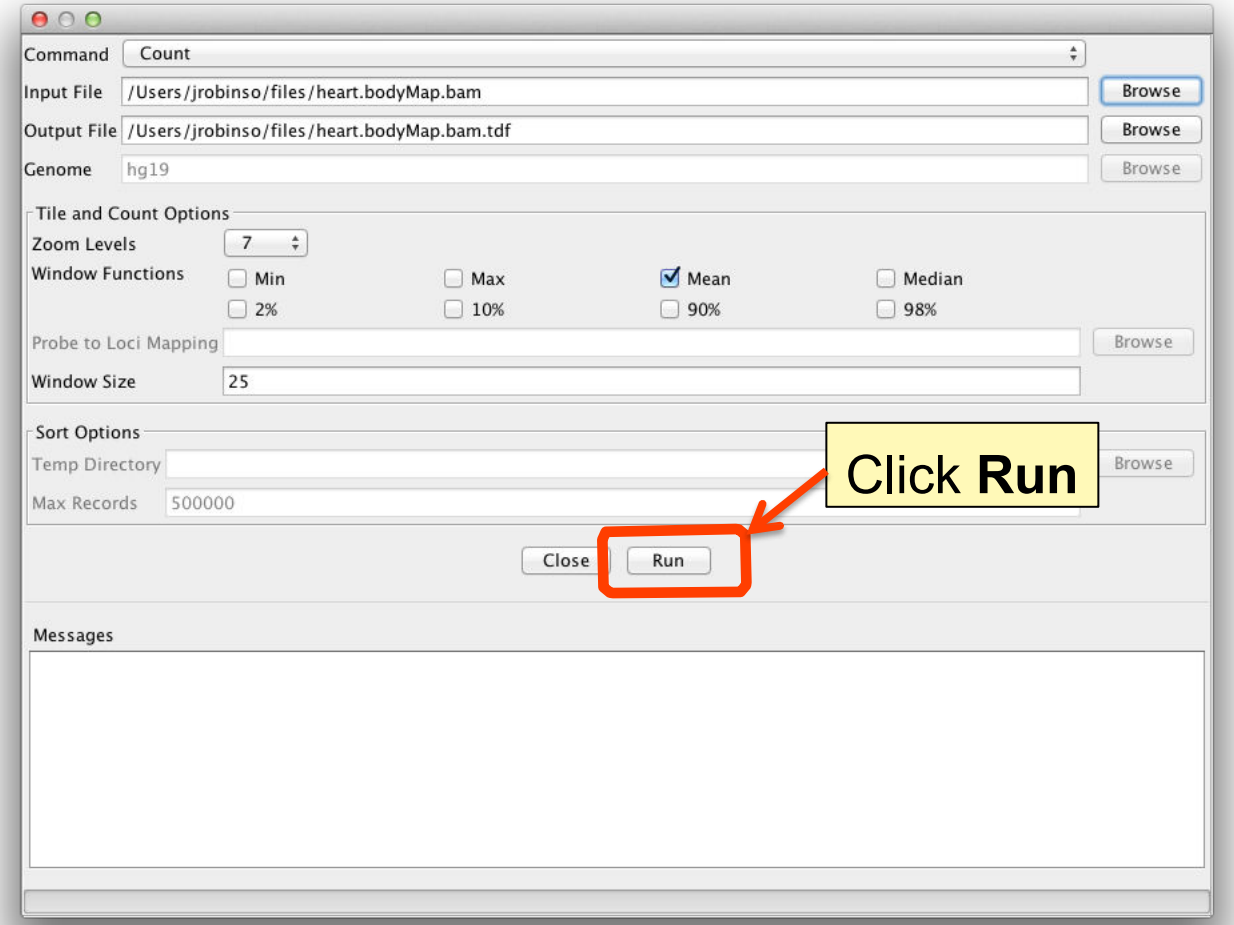


# Computing coverage: igvtools

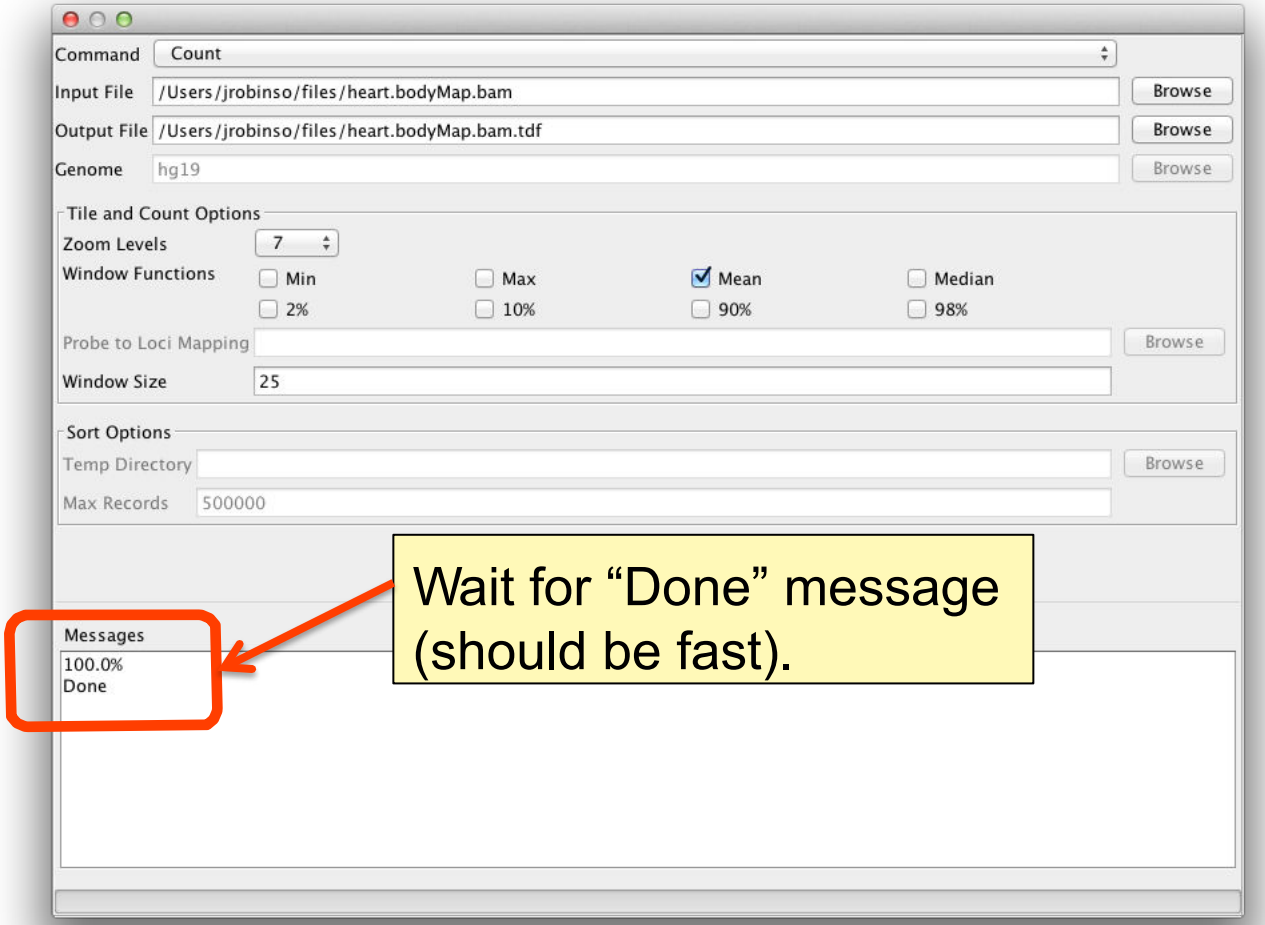



Select input file **heart.bodyMap.bam**
Output filename will be filled in automatically.

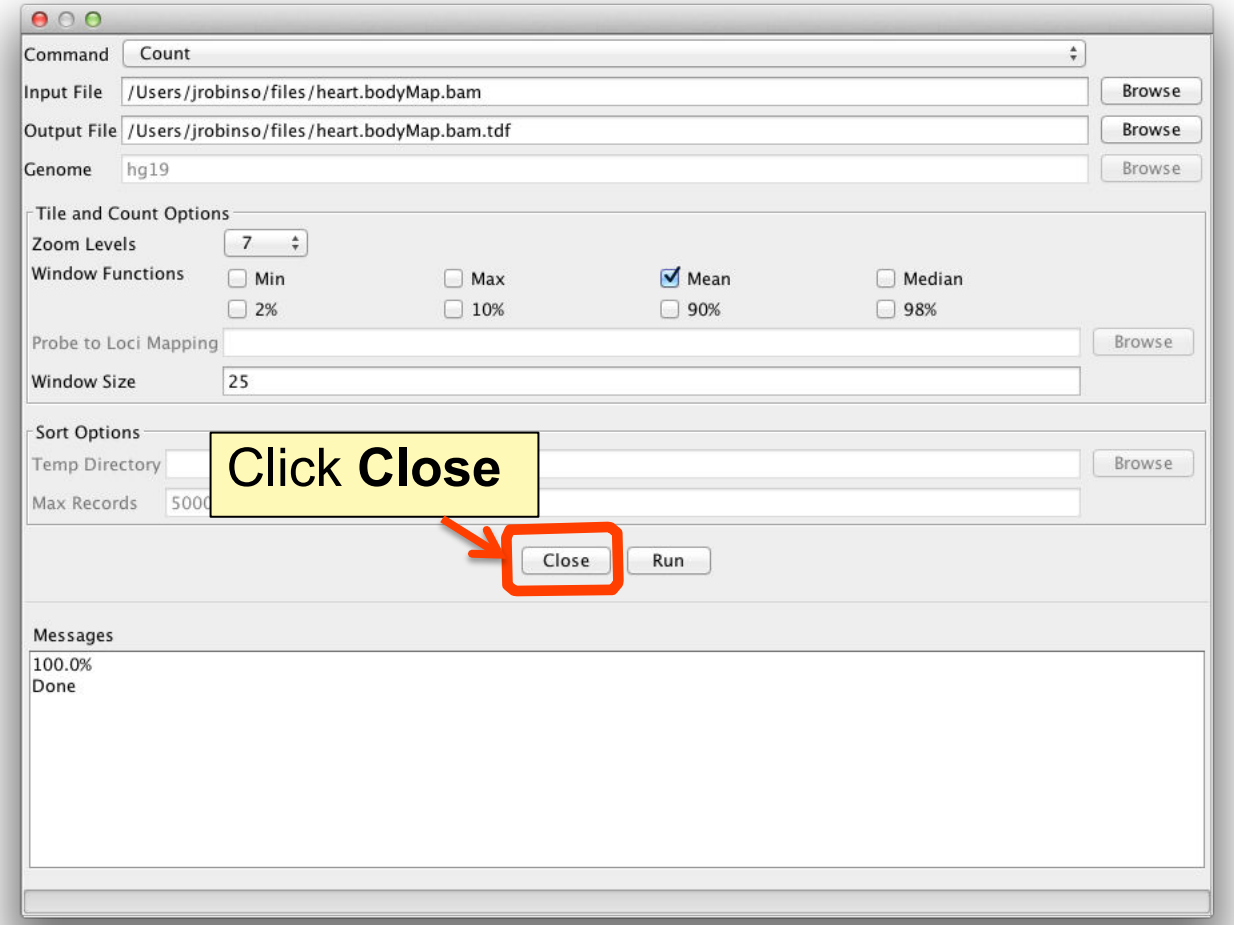
# Computing coverage: igvtools



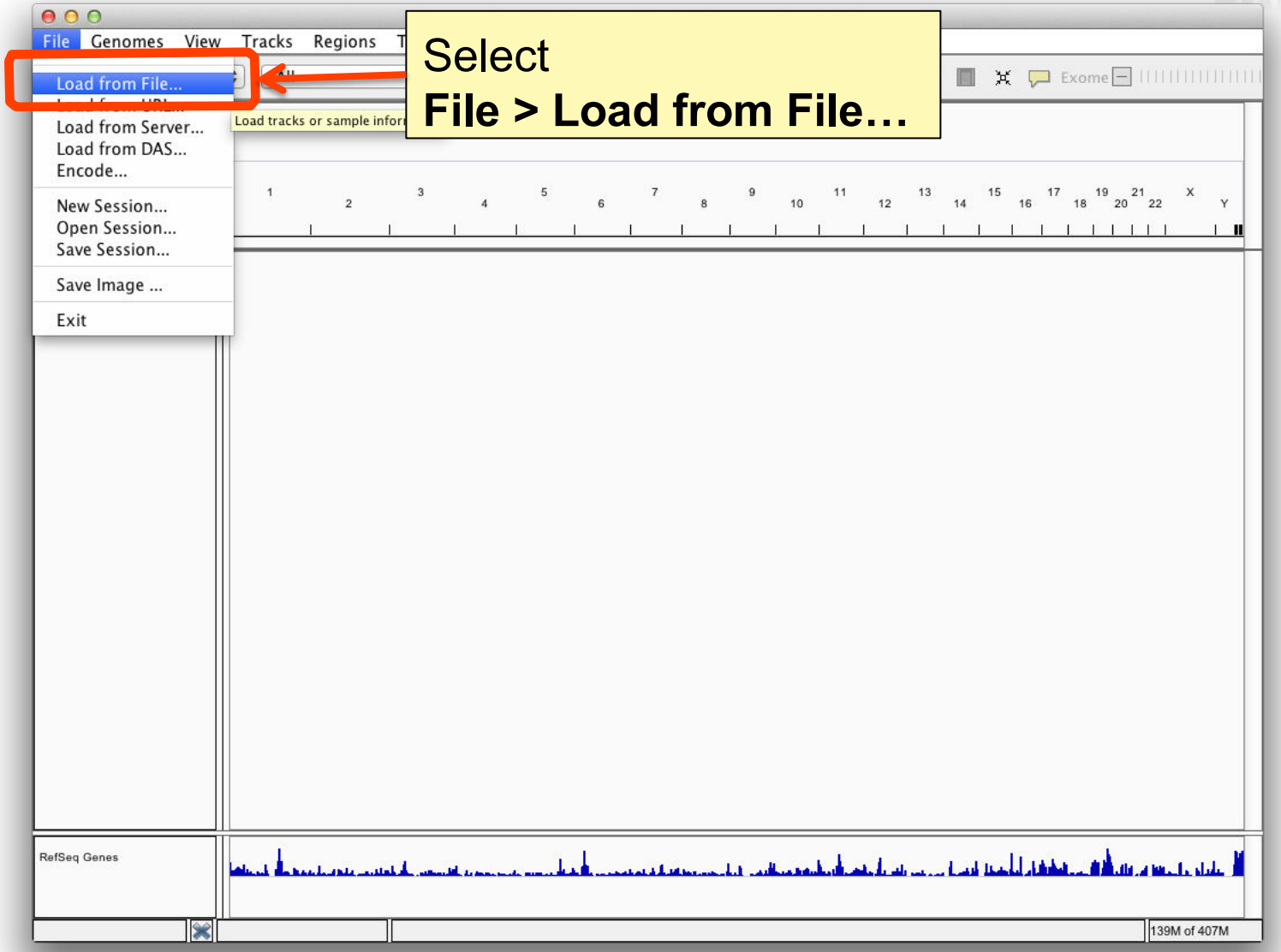
# Computing coverage: igvtools



# Computing coverage: igvtools

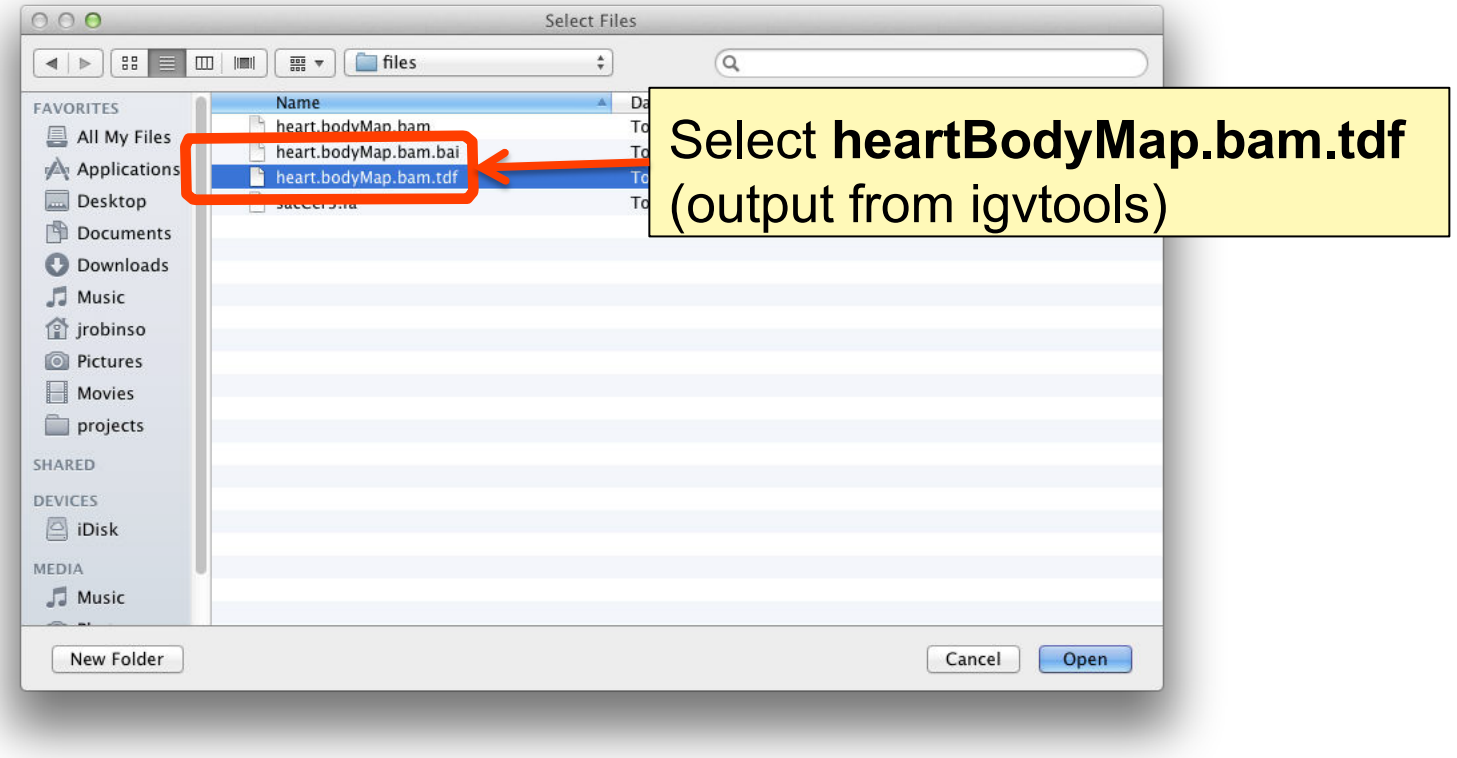


# Computing coverage: igvtools

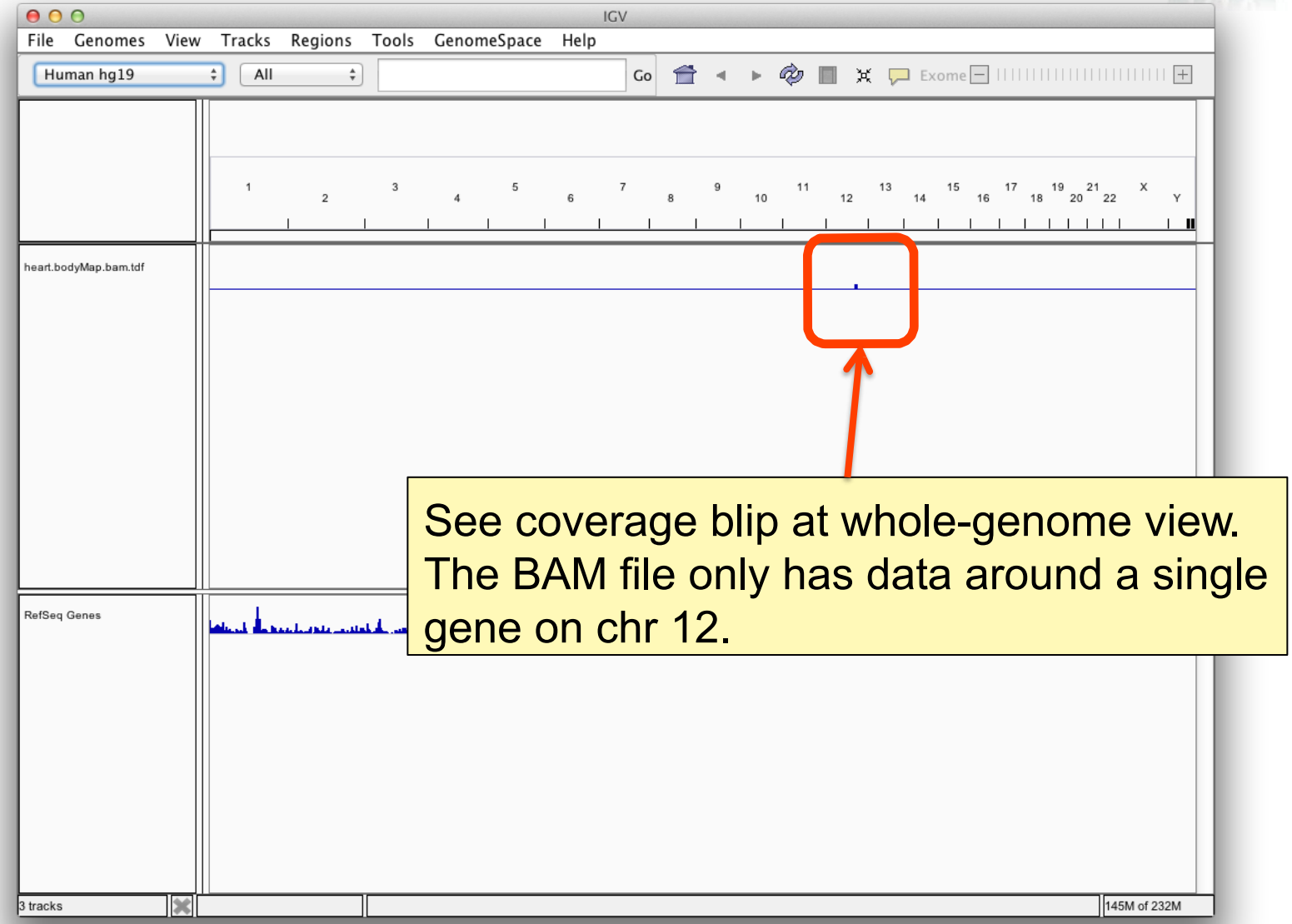



Select
**File > Load from File…**

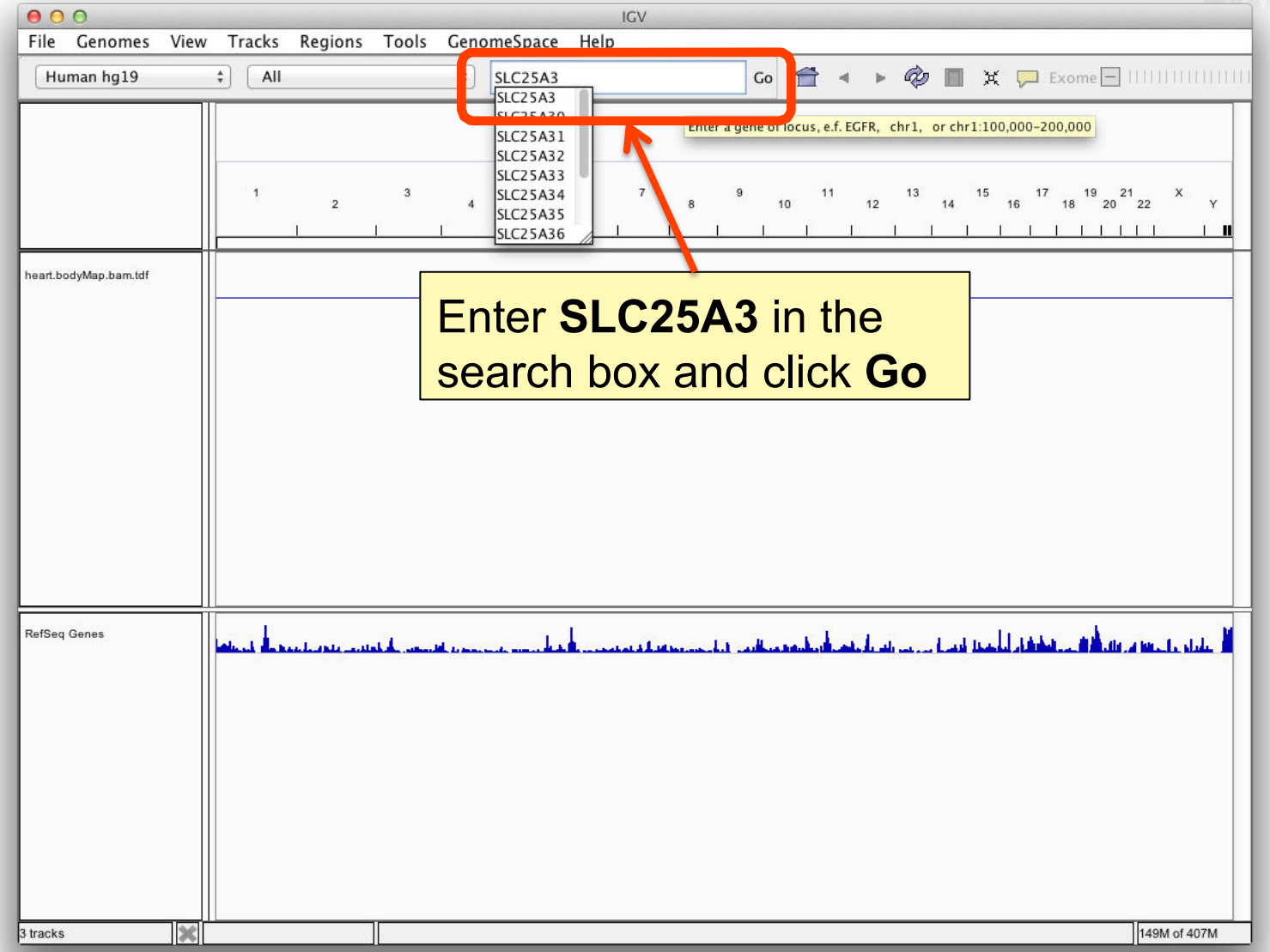
# Computing coverage: igvtools



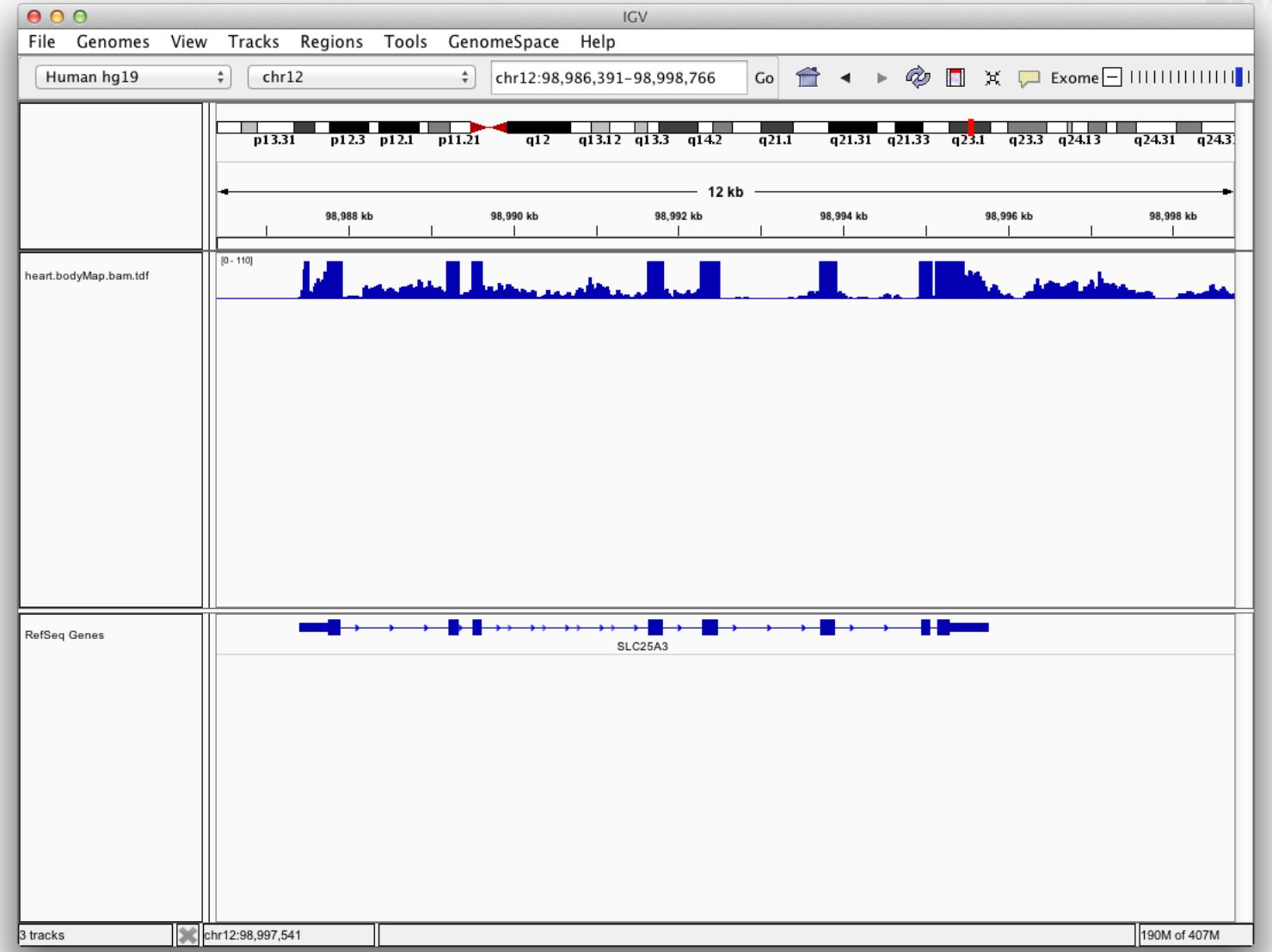
# Computing coverage: igvtools



# Computing coverage: igvtools



# Computing coverage: igvtools



# More about reference genomes


IGV doesn't host the genome you need?

Use any genome you want, if you have the sequence in FASTA format.

Optionally, package genome annotations with the sequence.

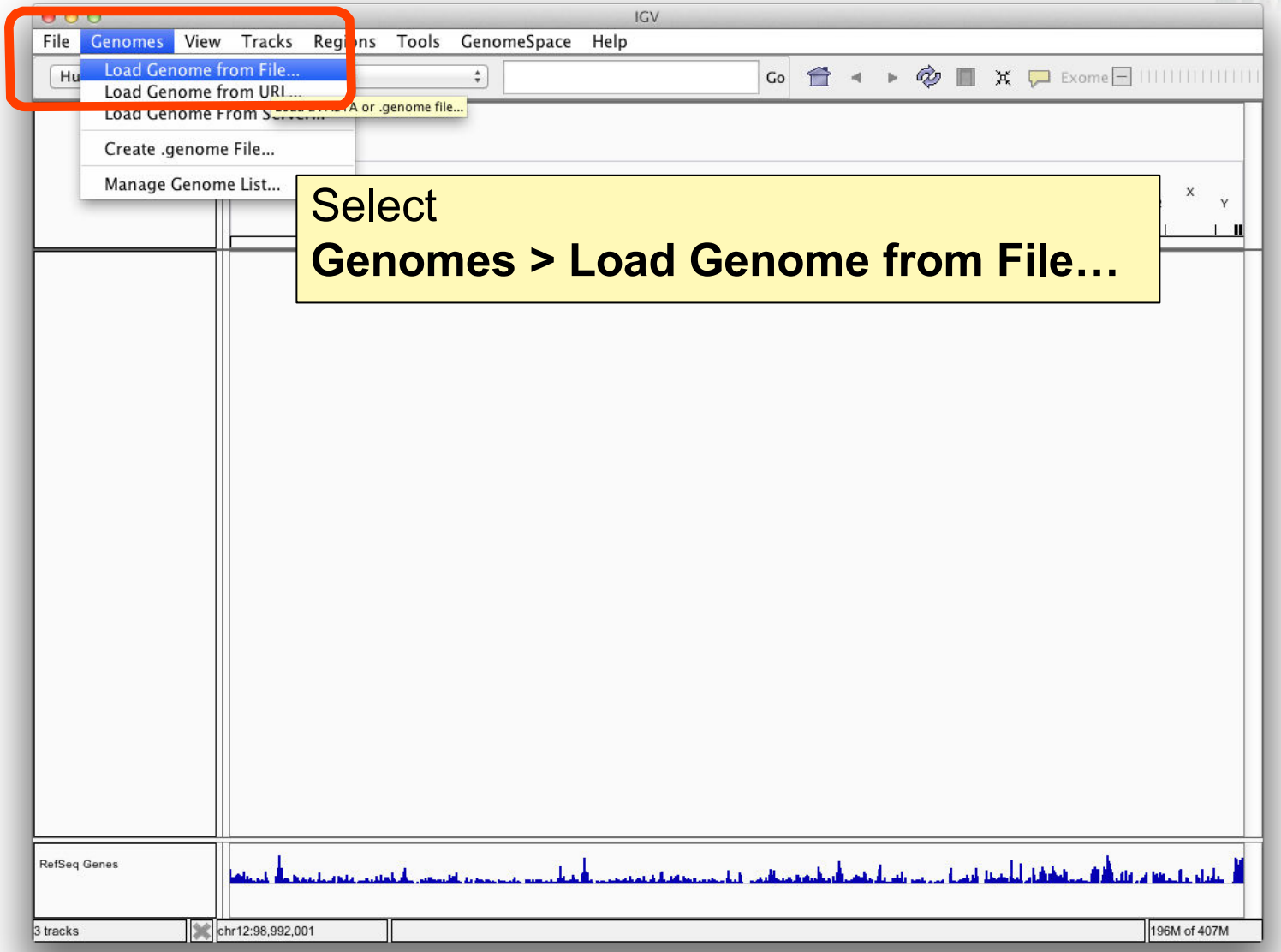
BROAD INSTITUTE

# Loading a genome

Hands-on exercise


BROAD INSTITUTE

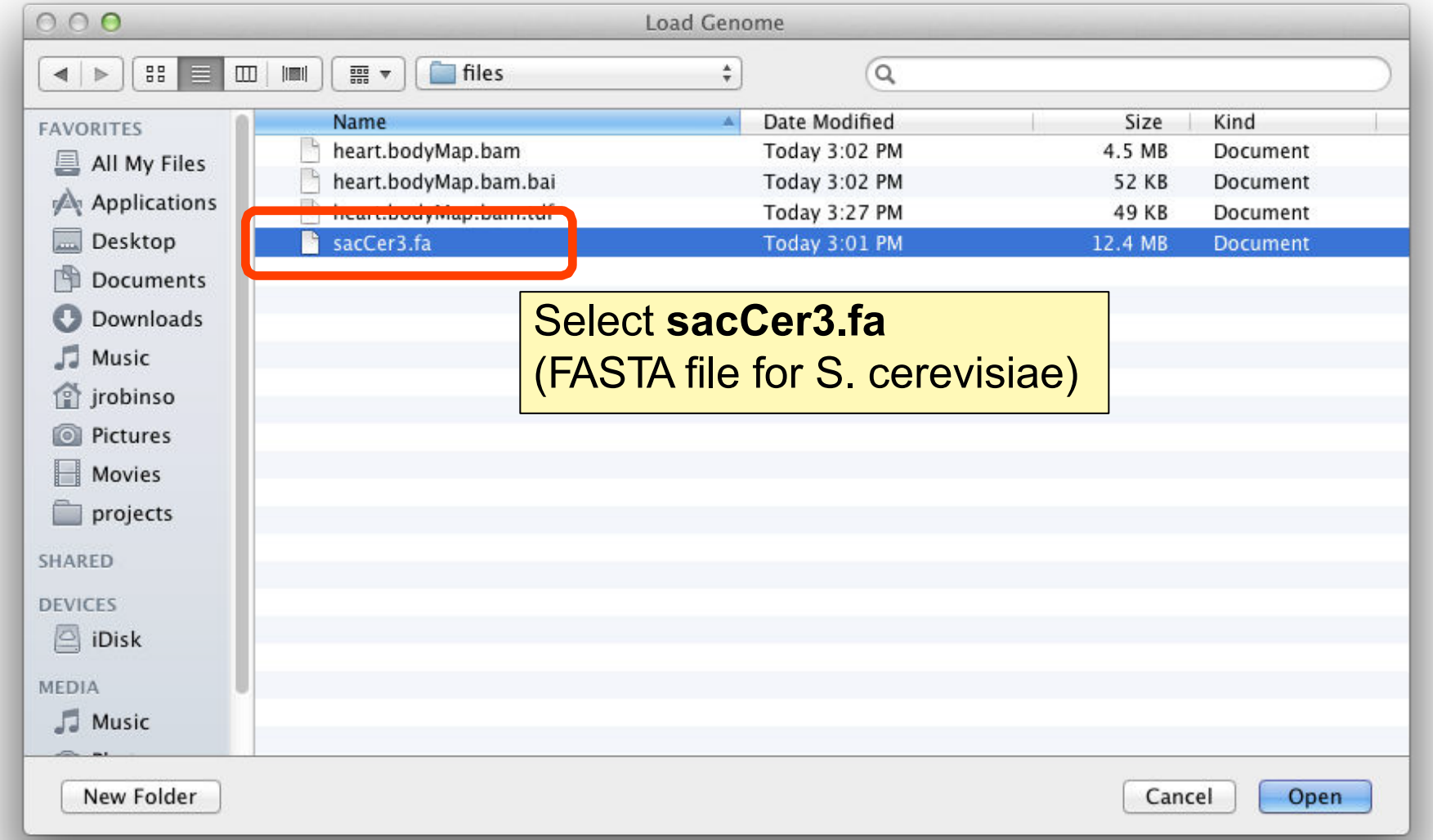
# Loading a genome




Select
**Genomes > Load Genome from File…**

# Loading a genome

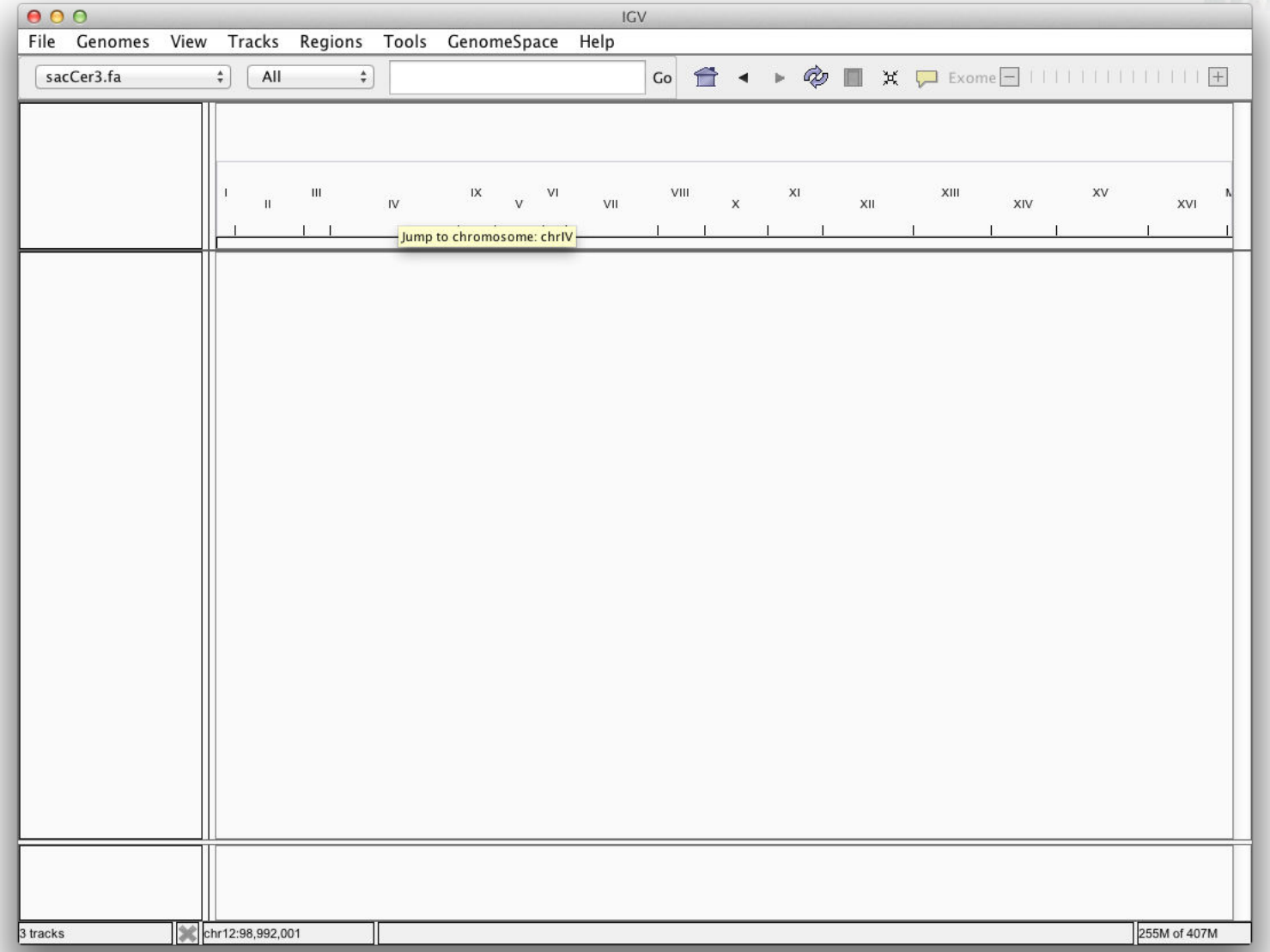
Integrative Genomics Viewer




BROAD INSTITUTE

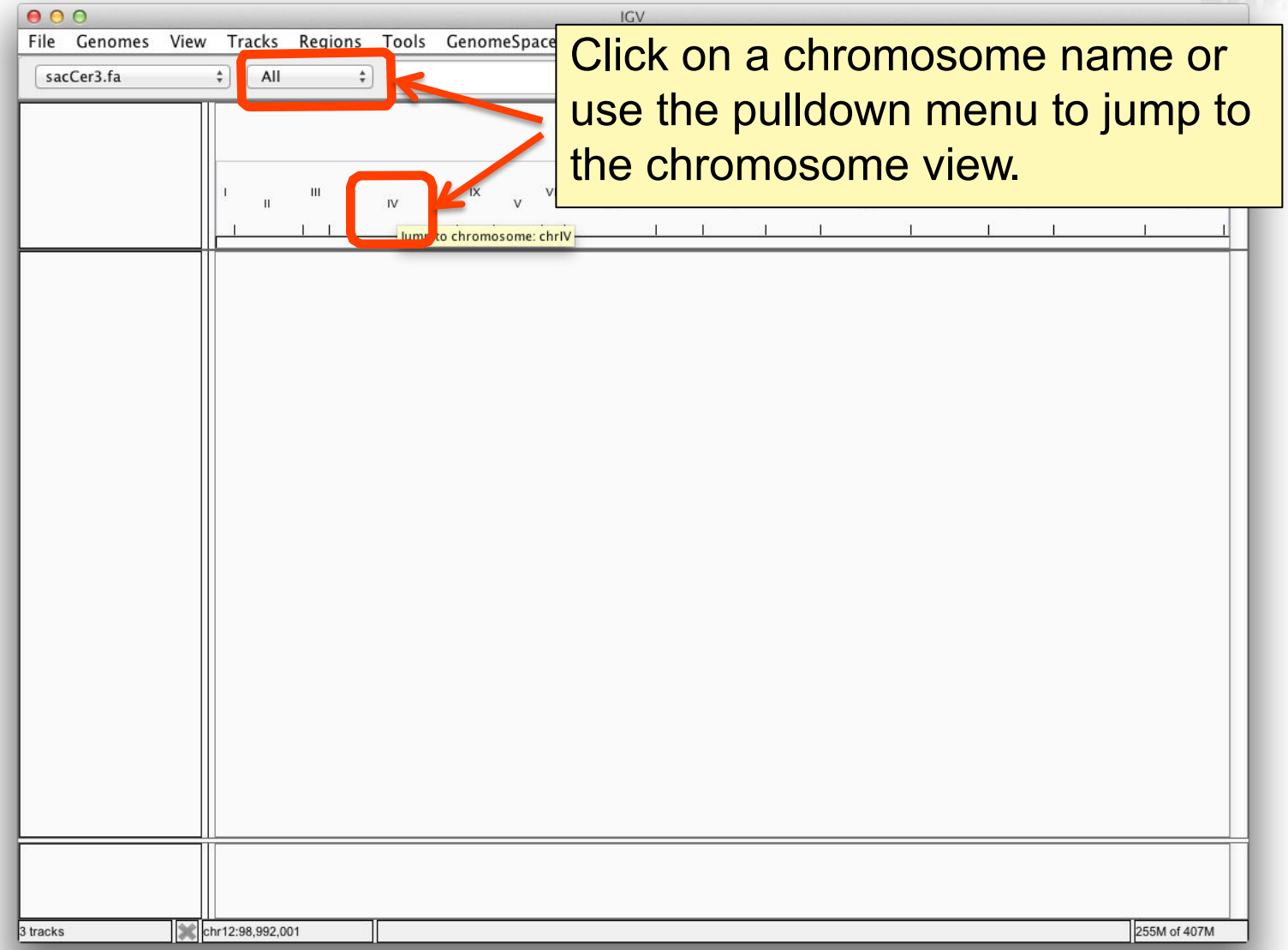
# Loading a genome

Integrative Genomics Viewer




BROAD INSTITUTE

# Loading a genome

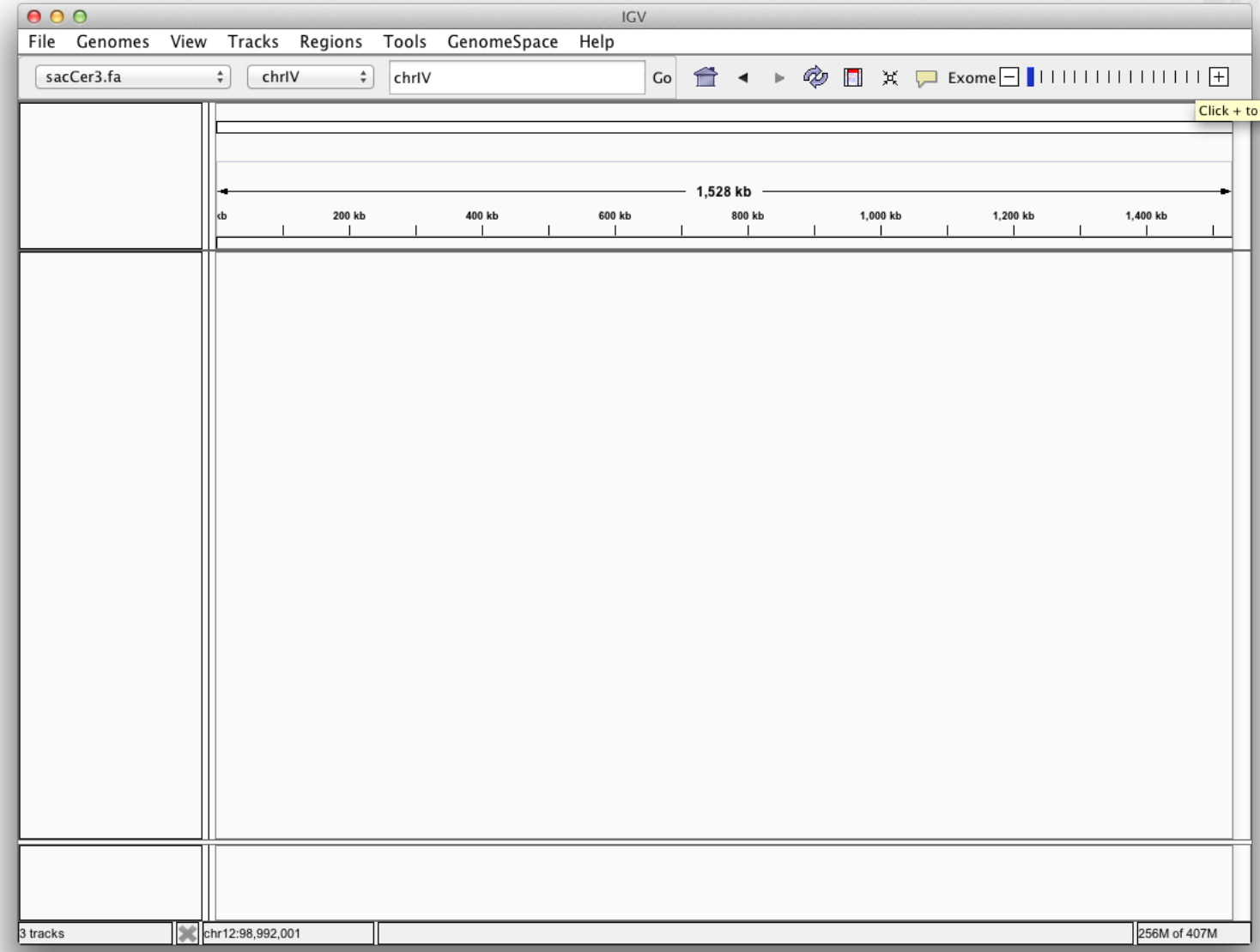



Click on a chromosome name or use the pulldown menu to jump to the chromosome view.

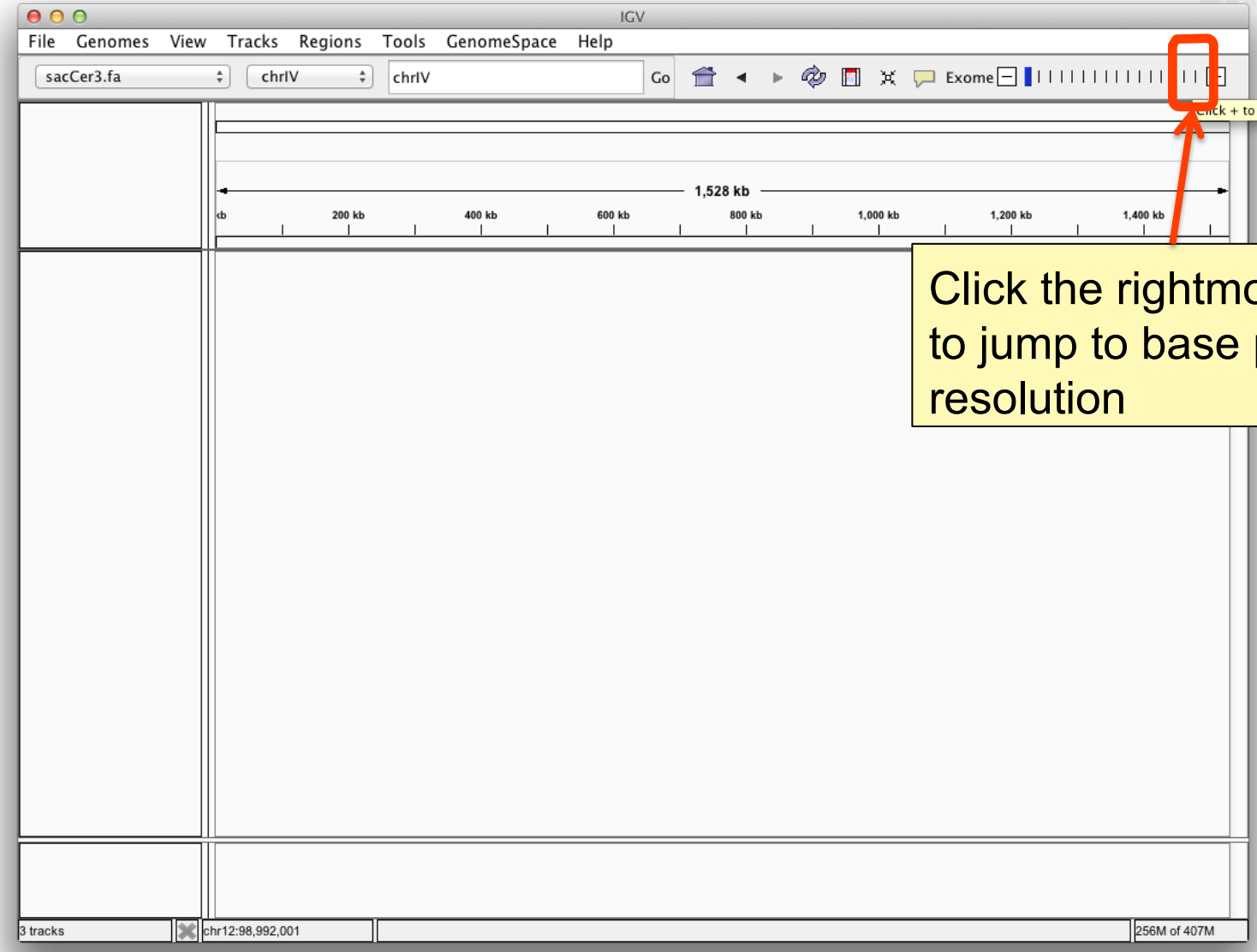
Integrative Genomics Viewer

BROAD INSTITUTE

# Loading a genome



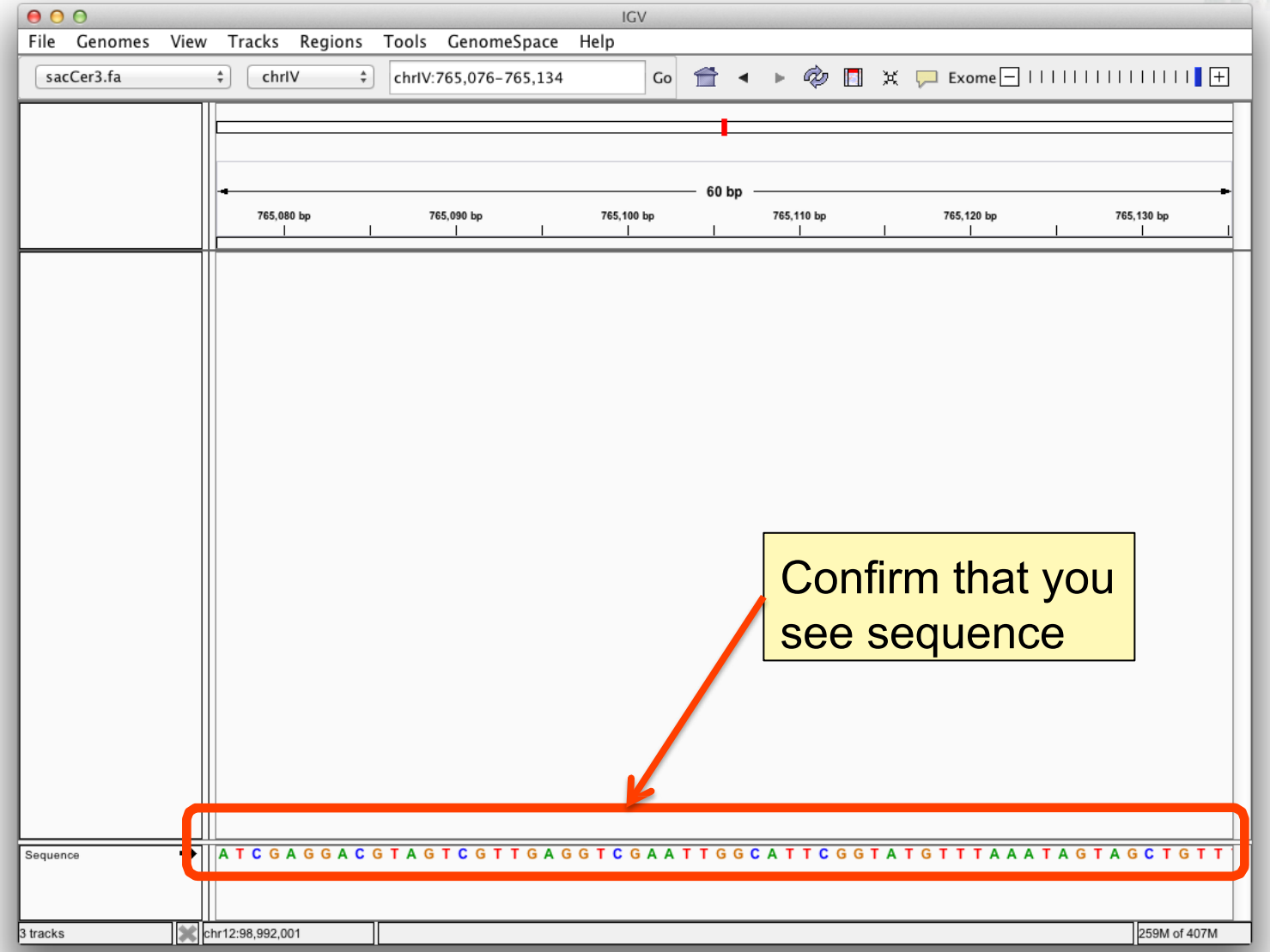
# Loading a genome




Click the rightmost "tick" to jump to base pair resolution

# Loading a genome

Integrative Genomics Viewer




BROAD INSTITUTE

# Acknowledgments



## IGV Team

Jim Robinson, Jacob Silterra, Helga Thorvaldsdóttir, Jill Mesirov (PI)

## Funding

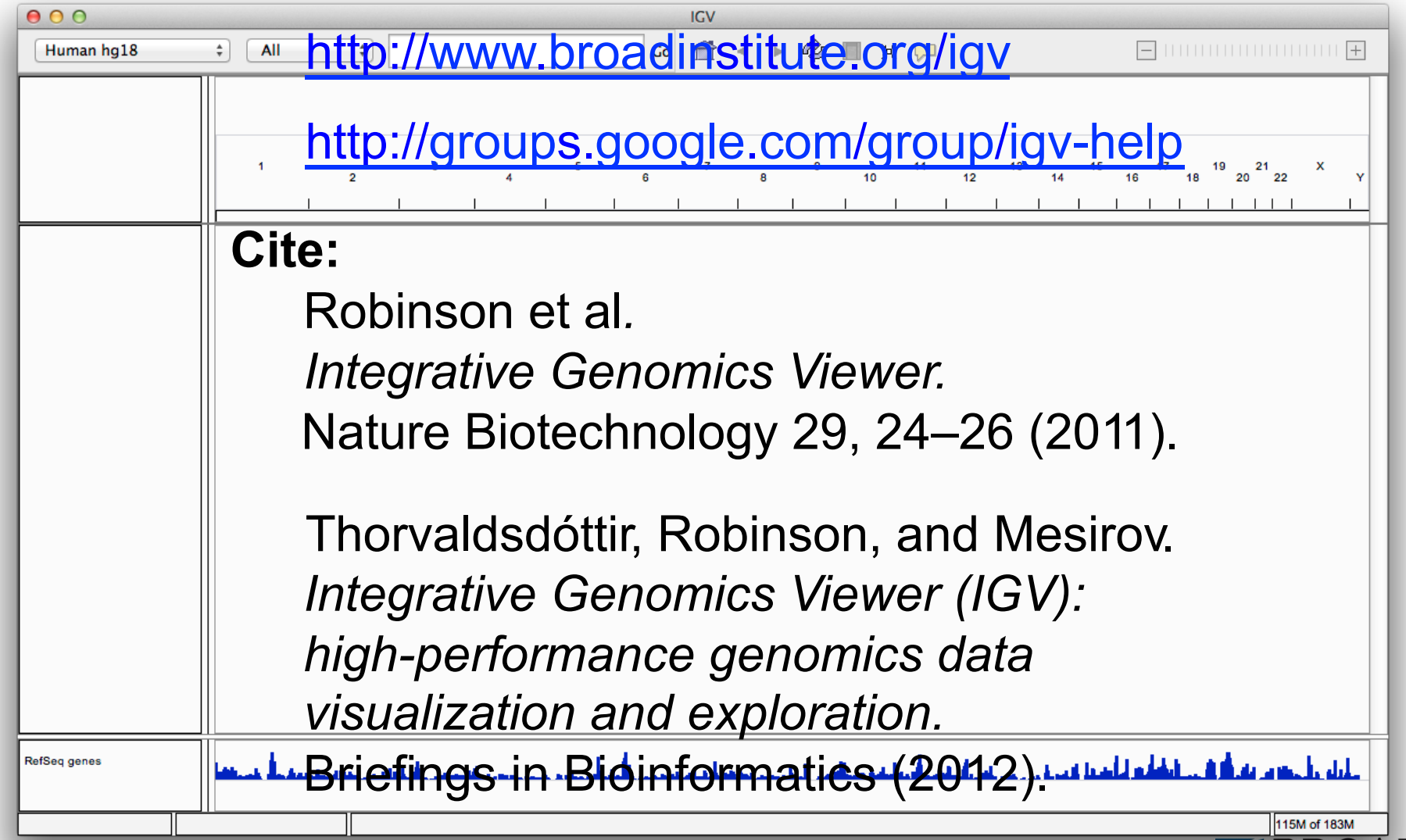
IGV development has been made possible with funding from:

- National Cancer Institute (NCI) http://cancer.gov/
- Starr Cancer Consortium http://www.starrcancer.org/
- National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health http://www.nigms.nih.gov/
- IGV participates in GenomeSpace http://genomespace.org/, which is funded by the the National Human Genome Research Institute (NHGRI) http://www.genome.gov/


BROAD INSTITUTE

Integrative Genomics Viewer

## For further information and help:




http://www.broadinstitute.org/igv

http://groups.google.com/group/igv-help

**Cite:**

Robinson et al.
*Integrative Genomics Viewer.*
Nature Biotechnology 29, 24–26 (2011).

Thorvaldsdóttir, Robinson, and Mesirov.
*Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.*
Briefings in Bioinformatics (2012).

BROAD INSTITUTE