

Introduction to the Integrative Genomics Viewer (IGV)

Helga Thorvaldsdóttir
Broad Institute of MIT and Harvard
March 2017



Acknowledgments

IGV is developed in the Mesirov Lab

- Initially at the Broad Institute, Cambridge, MA
- Now at UC San Diego and Broad Institute



IGV creator and lead developer – Jim Robinson

Development of IGV has been made possible by funding from:

- National Cancer Institute (NCI) (www.cancer.gov)
- NCI's Information Technology for Cancer Research (ITCR) (itcr.nci.nih.gov)
- Starr Cancer Consortium (www.starrcancer.org)
- And
 - Stand Up 2 Cancer
 - National Institute of General Medical Sciences (NIGMS) (www.nigms.nih.gov)
 - IGV participates in GenomeSpace (www.genomespace.org), which is funded by the National Human Genome Research Institute (NHGRI) (www.genome.gov).

The logo for the Integrative Genomics Viewer (IGV) consists of the letters "IGV" in a bold, white, sans-serif font, centered within a dark gray square.

IGV

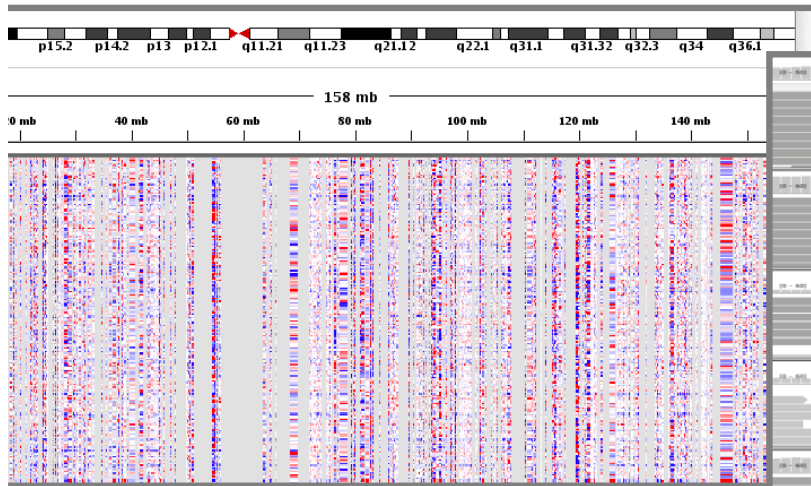
**Interactive
visual exploration of
integrated
genomic data**

The logo for the Integrative Genomics Viewer (IGV) consists of the letters "IGV" in a bold, white, sans-serif font, centered within a dark gray square.

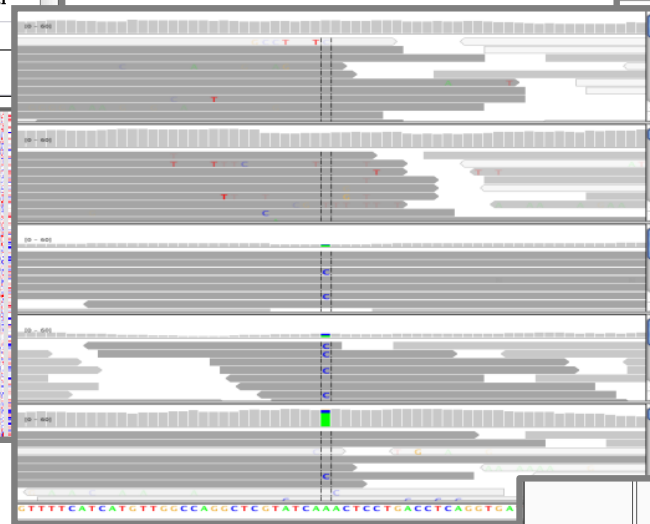
IGV

**Interactive
visual exploration of
integrated
genomic data**

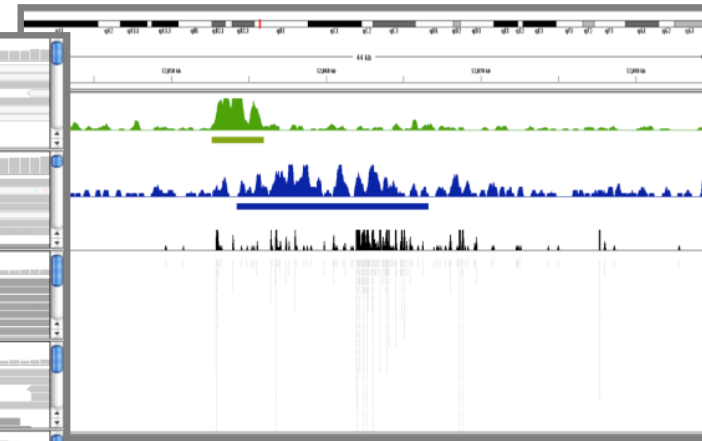
Microarray gene expression



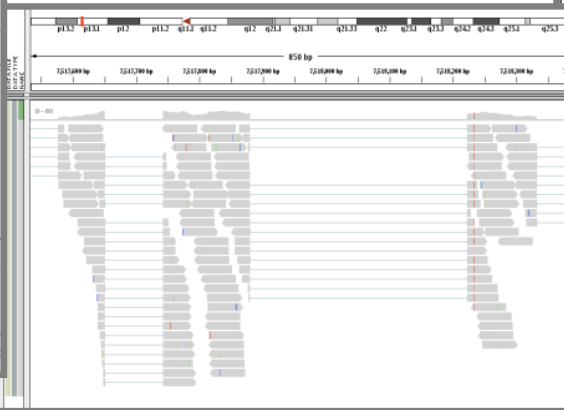
NGS alignments



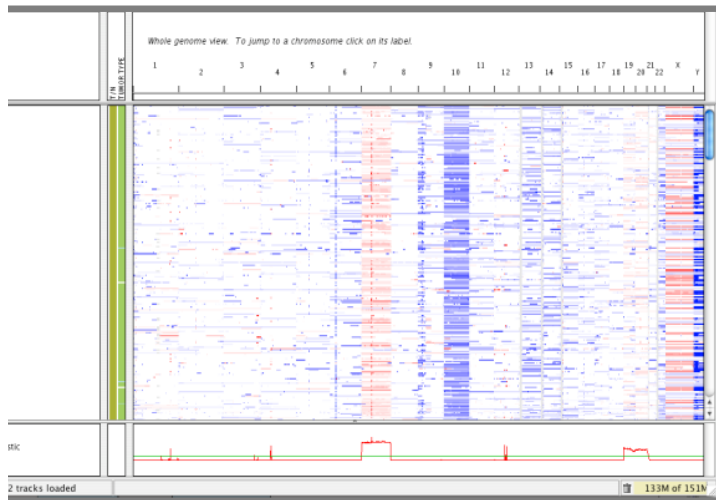
Epigenomics



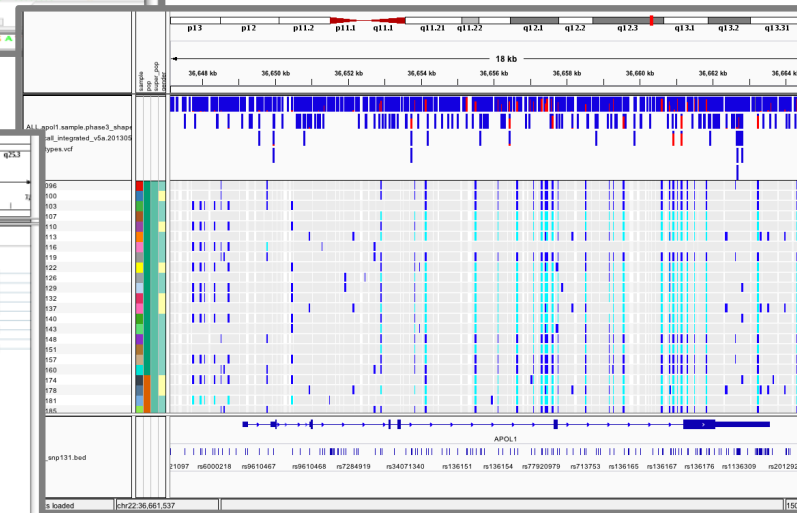
RNA-seq



Copy number



Variants and genotypes



Gene annotations



DCMWSGFSAAKLVSSEKLASYQAARKDSGSPNPARGHSVCSTSSSLYLQDLSAAAASECIDPSVVVF

IGV supports many different file formats

- [BAM](#)
- [BED](#)
- [BedGraph](#)
- [bigBed](#)
- [bigWig](#)
- [Birdsuite Files](#)
- [broadPeak](#)
- [CBS](#)
- [CN](#)
- [Custom File Formats](#)
- [Cytoband](#)
- [FASTA](#)
- [GCT](#)
- [genePred](#)
- [GFF/GTF](#)
- [GISTIC](#)
- [Goby](#)
- [GWAS](#)
- [IGV](#)
- [LOH](#)
- [MAF \(Multiple Alignment Format\)](#)
- [MAF \(Mutation Annotation Format\)](#)
- [Merged BAM File](#)
- [MUT](#)
- [narrowPeak](#)
- [PSL](#)
- [SAM](#)
- [Sample Info \(Attributes\) file](#)
- [SEG](#)
- [SNP](#)
- [TAB](#)
- [TDF](#)
- [Track Line](#)
- [Type Line](#)
- [VCF](#)
- [WIG](#)
- [chrom.sizes](#)

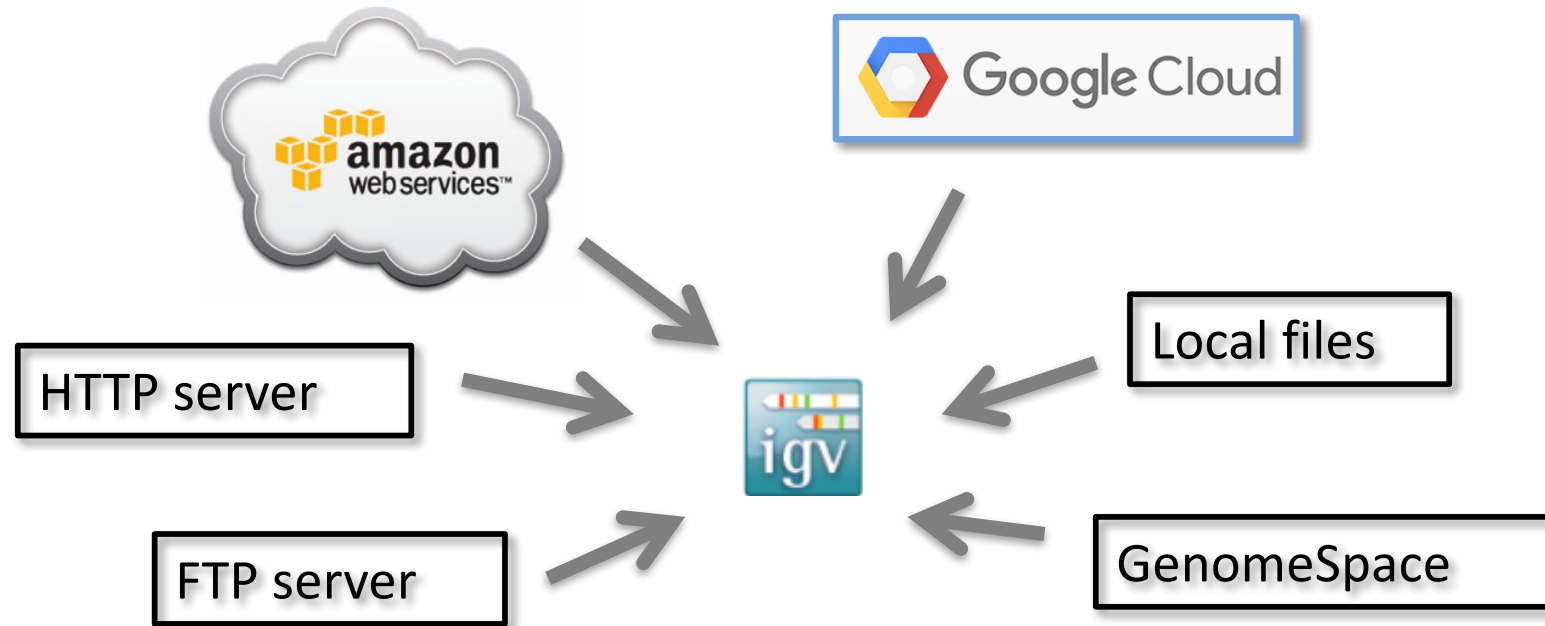
For more info see: [*www.broadinstitute.org/igv/FileFormats*](http://www.broadinstitute.org/igv/FileFormats)

The IGV logo consists of the letters 'IGV' in a bold, white, sans-serif font, centered within a dark gray square.

IGV

**Interactive
visual exploration of
integrated
genomic data**

IGV data sources



View **local** files without uploading

View **remote** files without downloading whole dataset

The IGV logo consists of the letters "IGV" in a bold, white, sans-serif font, centered within a dark gray square.

IGV

Interactive

visual exploration of

integrated

genomic data

The logo for the Integrative Genomics Viewer (IGV) consists of the letters "IGV" in a bold, white, sans-serif font, centered within a dark gray square.

IGV

Interactive

visual exploration of

integrated

genomic data

The logo for the Integrative Genomics Viewer (IGV) consists of the letters "IGV" in a bold, white, sans-serif font, centered within a dark gray square.

IGV

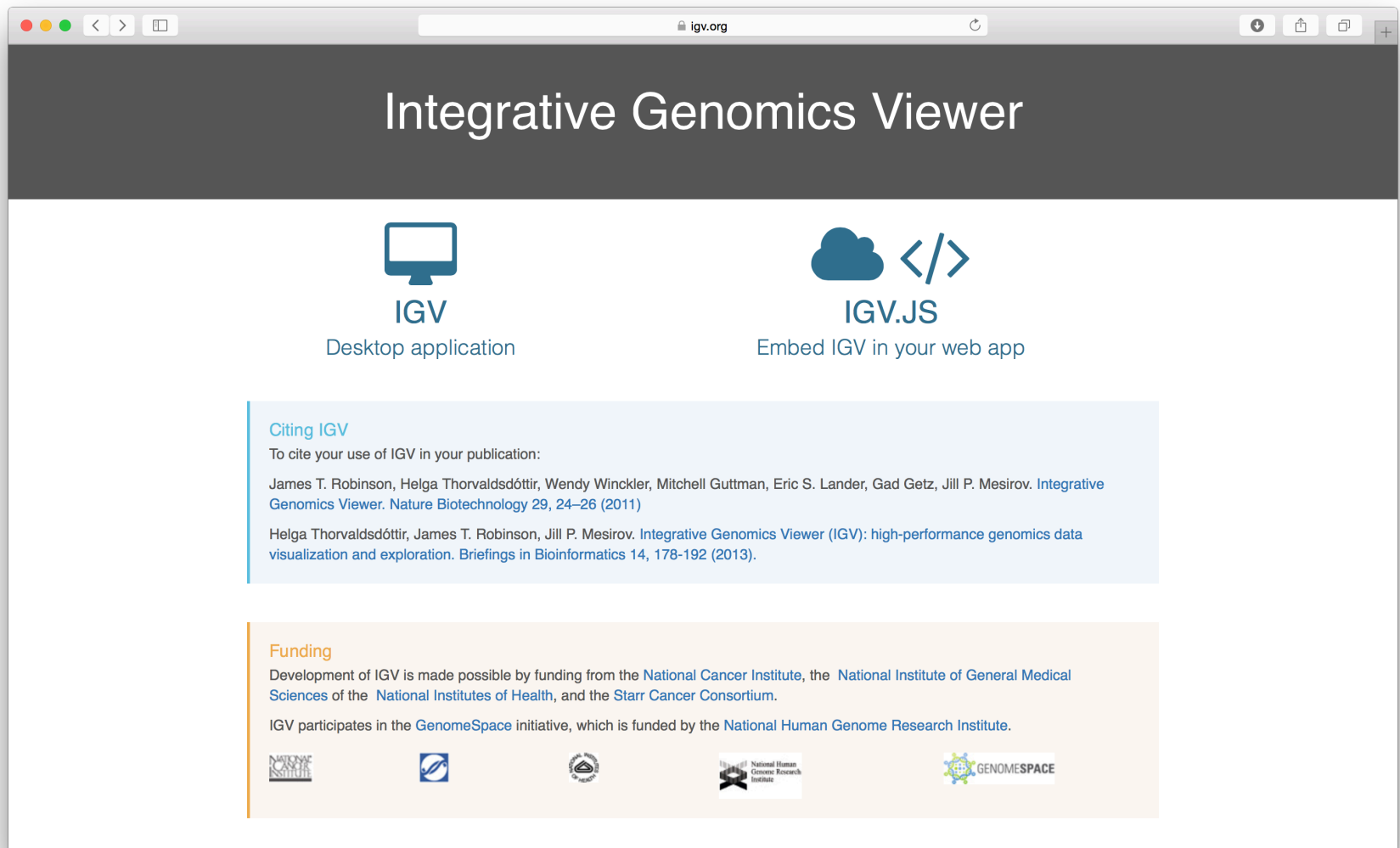
Interactive

visual exploration of

integrated

genomic data

www.igv.org





www.igv.org

The screenshot shows the IGV website homepage in a browser window. The browser's address bar displays 'igv.org'. The page has a dark grey header with the title 'Integrative Genomics Viewer' in white. Below the header, the main content area is white. On the left, the text 'Desktop application' is written in red. In the center, there is a blue icon of a computer monitor with the text 'IGV' and 'Desktop application' below it; this entire section is circled in red. To the right of the center, there is a blue cloud icon with a code symbol '</>' and the text 'IGV.JS' and 'Embed IGV in your web app' below it. Below these options, there are two light blue boxes. The first box is titled 'Citing IGV' and contains text about citing the software in publications, with two references listed. The second box is titled 'Funding' and contains text about the funding sources for the development of IGV, with logos for the National Cancer Institute, National Institute of General Medical Sciences, National Institutes of Health, Starr Cancer Consortium, and the National Human Genome Research Institute (GenomeSpace) at the bottom.

Integrative Genomics Viewer

Desktop application


IGV
Desktop application


IGV.JS
Embed IGV in your web app

Citing IGV

To cite your use of IGV in your publication:


James T. Robinson, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, Jill P. Mesirov. [Integrative Genomics Viewer](#). *Nature Biotechnology* 29, 24–26 (2011)

Helga Thorvaldsdóttir, James T. Robinson, Jill P. Mesirov. [Integrative Genomics Viewer \(IGV\): high-performance genomics data visualization and exploration](#). *Briefings in Bioinformatics* 14, 178-192 (2013).


Funding

Development of IGV is made possible by funding from the [National Cancer Institute](#), the [National Institute of General Medical Sciences](#) of the [National Institutes of Health](#), and the [Starr Cancer Consortium](#).

IGV participates in the [GenomeSpace](#) initiative, which is funded by the [National Human Genome Research Institute](#).



www.broadinstitute.org/igv




Integrative
Genomics
Viewer

- Home
- Downloads
- Documents
 - Hosted Genomes
 - FAQ
 - IGV User Guide
 - File Formats
 - Release Notes
 - IGV for iPad
 - Credits
- Contact

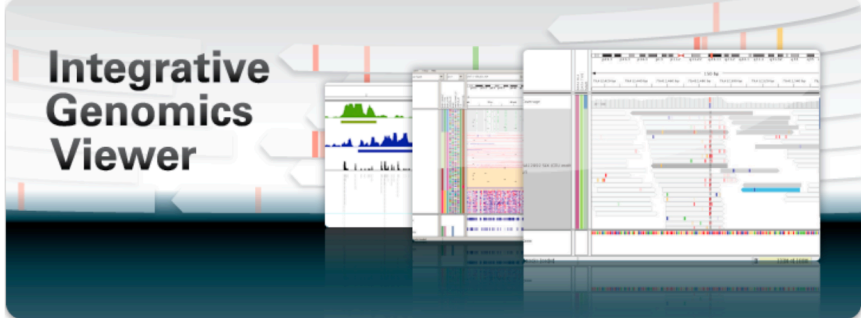
Search website

search



© 2013-2016 Broad Institute

Home



Integrative Genomics Viewer

Overview

The **Integrative Genomics Viewer (IGV)** is a high-performance visualization tool for interactive exploration of large, integrated genomic datasets. It supports a wide variety of data types, including array-based and next-generation sequence data, and genomic annotations.





Downloads

Download the IGV desktop application and igvtools.

Funding

Development of IGV is made possible by funding from the [National Cancer Institute](#), the [National Institute of General Medical Sciences of the National Institutes of Health](#), and the [Starr Cancer Consortium](#).

IGV participates in the [GenomeSpace](#) initiative, which is funded by the [National Human Genome Research Institute](#).



Citing IGV

To cite your use of IGV in your publication:


James T. Robinson, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, Jill P. Mesirov. [Integrative Genomics Viewer](#). *Nature Biotechnology* 29, 24–26 (2011)


Helga Thorvaldsdóttir, James T. Robinson, Jill P. Mesirov. [Integrative Genomics Viewer \(IGV\): high-performance genomics data visualization and exploration](#). *Briefings in Bioinformatics* 14, 178-192 (2013).

www.igv.org

The screenshot shows the IGV website homepage. At the top, the title "Integrative Genomics Viewer" is displayed in a large, white font on a dark grey background. Below this, there are two main options: "IGV Desktop application" with a monitor icon, and "IGV.JS" with a cloud and code icon. The "IGV.JS" option is circled in red, and the text "JavaScript version" is written in red to its right. Below these options, there is a light blue box containing the "Citing IGV" section, which provides information on how to cite the software in publications. At the bottom, there is a light orange box containing the "Funding" section, which lists the organizations that fund the development of IGV. The footer of the page features logos for the National Cancer Institute, National Institute of General Medical Sciences, National Institutes of Health, Starr Cancer Consortium, National Human Genome Research Institute, and GENOMESPACE.

Integrative Genomics Viewer


IGV
Desktop application


IGV.JS
Embed IGV in your web app

JavaScript version

Citing IGV

To cite your use of IGV in your publication:






James T. Robinson, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, Jill P. Mesirov. [Integrative Genomics Viewer](#). *Nature Biotechnology* 29, 24–26 (2011)

Helga Thorvaldsdóttir, James T. Robinson, Jill P. Mesirov. [Integrative Genomics Viewer \(IGV\): high-performance genomics data visualization and exploration](#). *Briefings in Bioinformatics* 14, 178-192 (2013).

Funding

Development of IGV is made possible by funding from the [National Cancer Institute](#), the [National Institute of General Medical Sciences](#) of the [National Institutes of Health](#), and the [Starr Cancer Consortium](#).

IGV participates in the [GenomeSpace](#) initiative, which is funded by the [National Human Genome Research Institute](#).


    


www.igv.org

The screenshot shows the IGV website homepage in a browser window. The browser's address bar displays 'igv.org'. The page has a dark grey header with the title 'Integrative Genomics Viewer' in white. Below the header, the main content area is white. On the left, the text 'Desktop application' is written in red. In the center, there is a blue icon of a computer monitor with the text 'IGV' and 'Desktop application' below it; this entire section is circled in red. To the right of the center, there is a blue cloud icon with a code symbol '</>' and the text 'IGV.JS' and 'Embed IGV in your web app' below it. Below these options, there are two light blue boxes. The first box is titled 'Citing IGV' and contains text about citing the software in publications, with two references listed. The second box is titled 'Funding' and contains text about the funding sources for the development of IGV, followed by logos for the National Cancer Institute, National Institute of General Medical Sciences, National Institutes of Health, Starr Cancer Consortium, and the GenomeSpace initiative.

Integrative Genomics Viewer

Desktop application


IGV
Desktop application


IGV.JS
Embed IGV in your web app

Citing IGV

To cite your use of IGV in your publication:


James T. Robinson, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, Jill P. Mesirov. [Integrative Genomics Viewer](#). *Nature Biotechnology* 29, 24–26 (2011)

Helga Thorvaldsdóttir, James T. Robinson, Jill P. Mesirov. [Integrative Genomics Viewer \(IGV\): high-performance genomics data visualization and exploration](#). *Briefings in Bioinformatics* 14, 178-192 (2013).

Funding

Development of IGV is made possible by funding from the [National Cancer Institute](#), the [National Institute of General Medical Sciences](#) of the [National Institutes of Health](#), and the [Starr Cancer Consortium](#).

IGV participates in the [GenomeSpace](#) initiative, which is funded by the [National Human Genome Research Institute](#).



www.broadinstitute.org/igv

Integrative Genomics Viewer

Home

Downloads

Documents

Hosted Genomes

FAQ

IGV User Guide

File Formats

Release Notes

IGV for iPad

Credits

Contact

Search website

BROAD INSTITUTE

© 2013-2016 Broad Institute

Integrative Genomics Viewer

Overview

The Integrative Genomics Viewer (IGV) is a high-performance visualization tool for interactive exploration of large, integrated genomic datasets. It supports a wide variety of data types, including array-based and next-generation sequence data, and genomic annotations.

Downloads

Download the IGV desktop application and igvtools.

Funding

Development of IGV is made possible by funding from the [National Cancer Institute](#), the [National Institute of General Medical Sciences of the National Institutes of Health](#), and the [Starr Cancer Consortium](#).

IGV participates in the [GenomeSpace](#) initiative, which is funded by the [National Human Genome Research Institute](#).

Citing IGV

To cite your use of IGV in your publication:

James T. Robinson, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, Jill P. Mesirov. [Integrative Genomics Viewer](#). *Nature Biotechnology* 29, 24–26 (2011)

Helga Thorvaldsdóttir, James T. Robinson, Jill P. Mesirov. [Integrative Genomics Viewer \(IGV\): high-performance genomics data visualization and exploration](#). *Briefings in Bioinformatics* 14, 178-192 (2013).

NATIONAL CANCER INSTITUTE
NATIONAL INSTITUTE OF GENERAL MEDICAL SCIENCES
NATIONAL HUMAN GENOME RESEARCH INSTITUTE
GENOMESPACE

IGV is freely available under an MIT open-source license.

Install IGV

Options for installing and running IGV:

1. (Mac only) Download and run the Mac application; or
2. (Windows) Download and run the self-extracting archive; or
3. (All systems) Use the Java Web Start buttons (Mac users: see below for limitations); or
4. (All systems) Download the binary distribution and run IGV from the command line.

Note: IGV 2.3.x requires Java 7. Users with Java 6 (JRE 1.6) should first try to upgrade Java to the latest version. If this is not possible you will need to run a 2.2.x version available in the [archive](#).

Mac

Download and unzip the Mac App archive, then double-click the IGV application to run it. The application can be moved to the "Applications" folder, or anywhere else.

Download
Mac App

Download
Mac App

or

Download
Windows Package

Windows

NOTE: for 64-bit windows only. For 32-bit windows use the 750MB launch button below.

1. Download the Windows package and execute the self-extracting archive.
2. It will prompt you for a location to extract the folder, choose anywhere you like (e.g. your home folder).
3. On completion, open the new folder.
4. Double-click the file "igv.bat", it might appear as just "igv" depending on your settings.

Download
Windows Package

Java Web Start (All Platforms)

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 does not work for some users due to security settings. The recommended solution is to use the bundled Mac App from the link above. Alternatively you can try to work around this by right-clicking on the buttons and saving the "jnlp" file, then right-clicking on the saved "jnlp" file and select "Open With > Java Web Start".

Chrome: Chrome does not automatically launch the Java Web Start 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, or if on a Mac right-click and select "Open With > Java Web Start"

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.



Launch

Java Web Start

or

Download
Binary Distribution

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

Install IGV

Options for installing and running IGV:

1. (Mac only) Download and run the Mac application; or
2. (Windows) Download and run the self-extracting archive; or
3. (All systems) Use the Java Web Start buttons (Mac users: see below for limitations); or
4. (All systems) Download the binary distribution and run IGV from the command line.

Note: IGV 2.3.x requires Java 7. Users with Java 6 (JRE 1.6) should first try to upgrade Java to the latest version. If this is not possible you will need to run a 2.2.x version available in the [archive](#).

Mac

Download and unzip the Mac App archive, then double-click the IGV application to run it. The application can be moved to the "Applications" folder, or anywhere else.

Download
Mac App

Download
Mac App

or

Download
Windows Package

Windows

NOTE: for 64-bit windows only. For 32-bit windows use the 750MB launch button below.

1. Download the Windows package and execute the self-extracting archive.
2. It will prompt you for a location to extract the folder, choose anywhere you like (e.g. your home folder).
3. On completion, open the new folder.
4. Double-click the file "igv.bat", it might appear as just "igv" depending on your settings.

Download
Windows Package

Java Web Start (All Platforms)

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 does not work for some users due to security settings. The recommended solution is to use the bundled Mac App from the link above. Alternatively you can try to work around this by right-clicking on the buttons and saving the "jnlp" file, then right-clicking on the saved "jnlp" file and select "Open With > Java Web Start".

Chrome: Chrome does not automatically launch the Java Web Start 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, or if on a Mac right-click and select "Open With > Java Web Start"

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.



Launch

Java Web Start

or

Download
Binary Distribution

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".

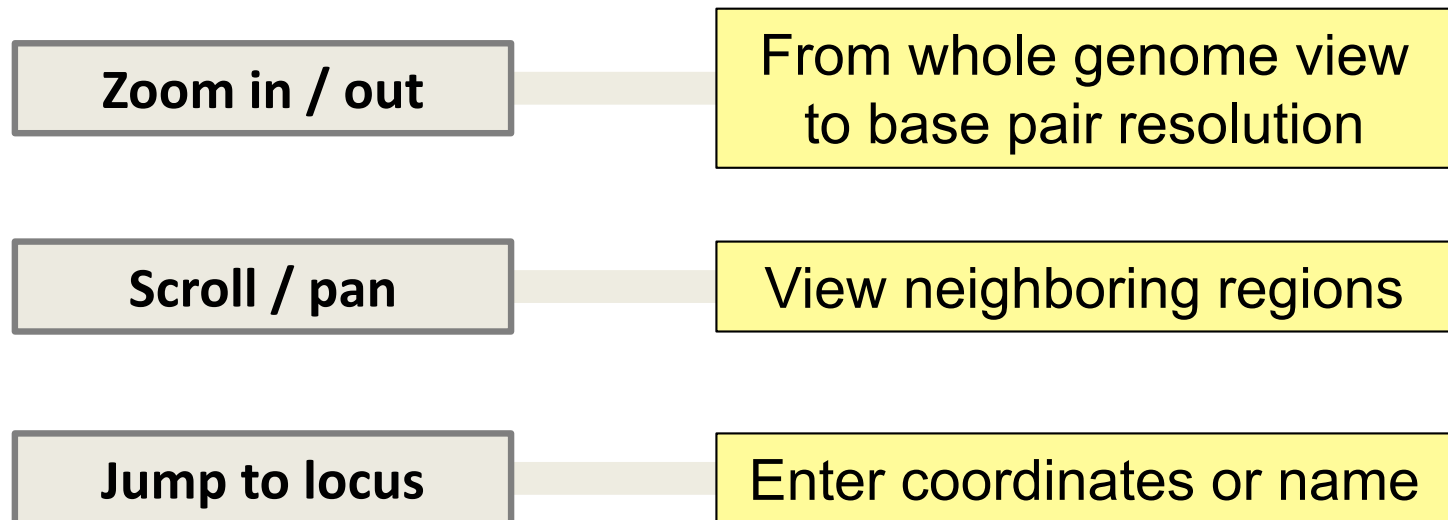
Download
Binary Distribution

Using IGV – The Basics

IGV user interface basics

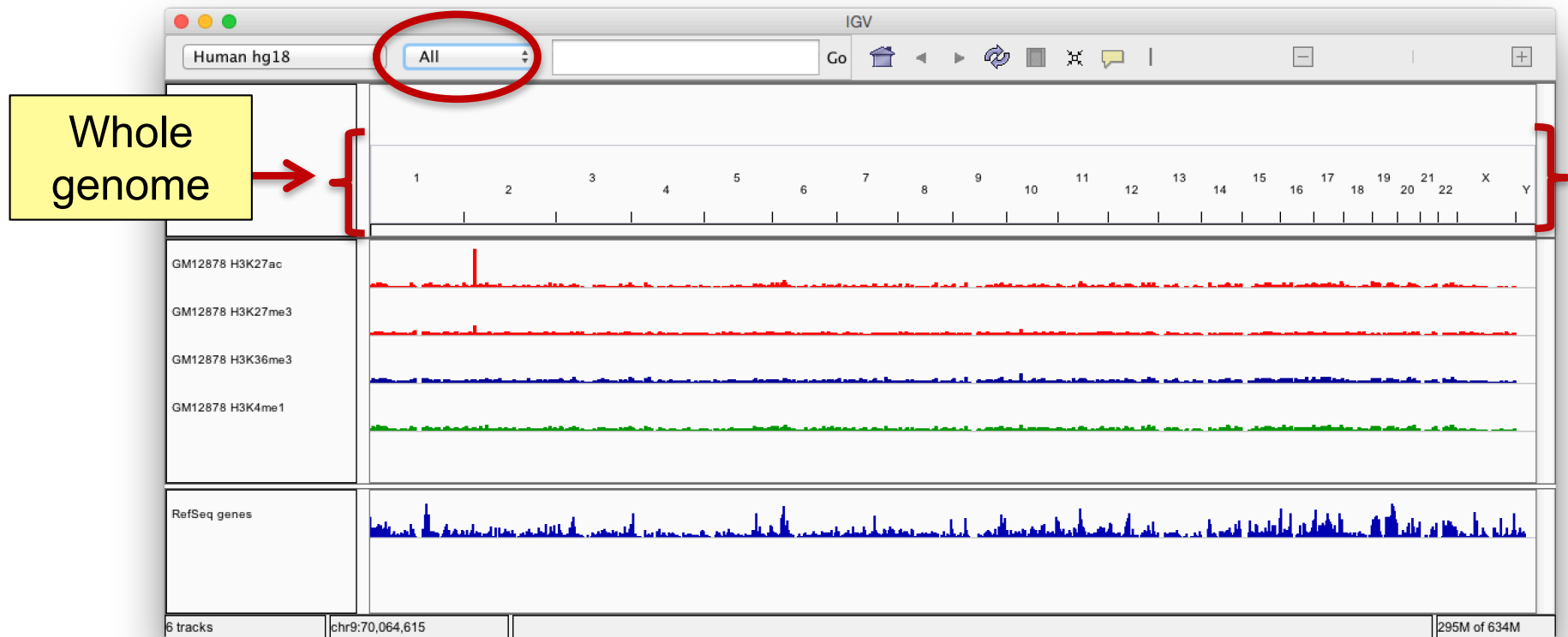
Hands-on exercise

- Select reference genome
- Load data from IGV hosted data server
- Navigate



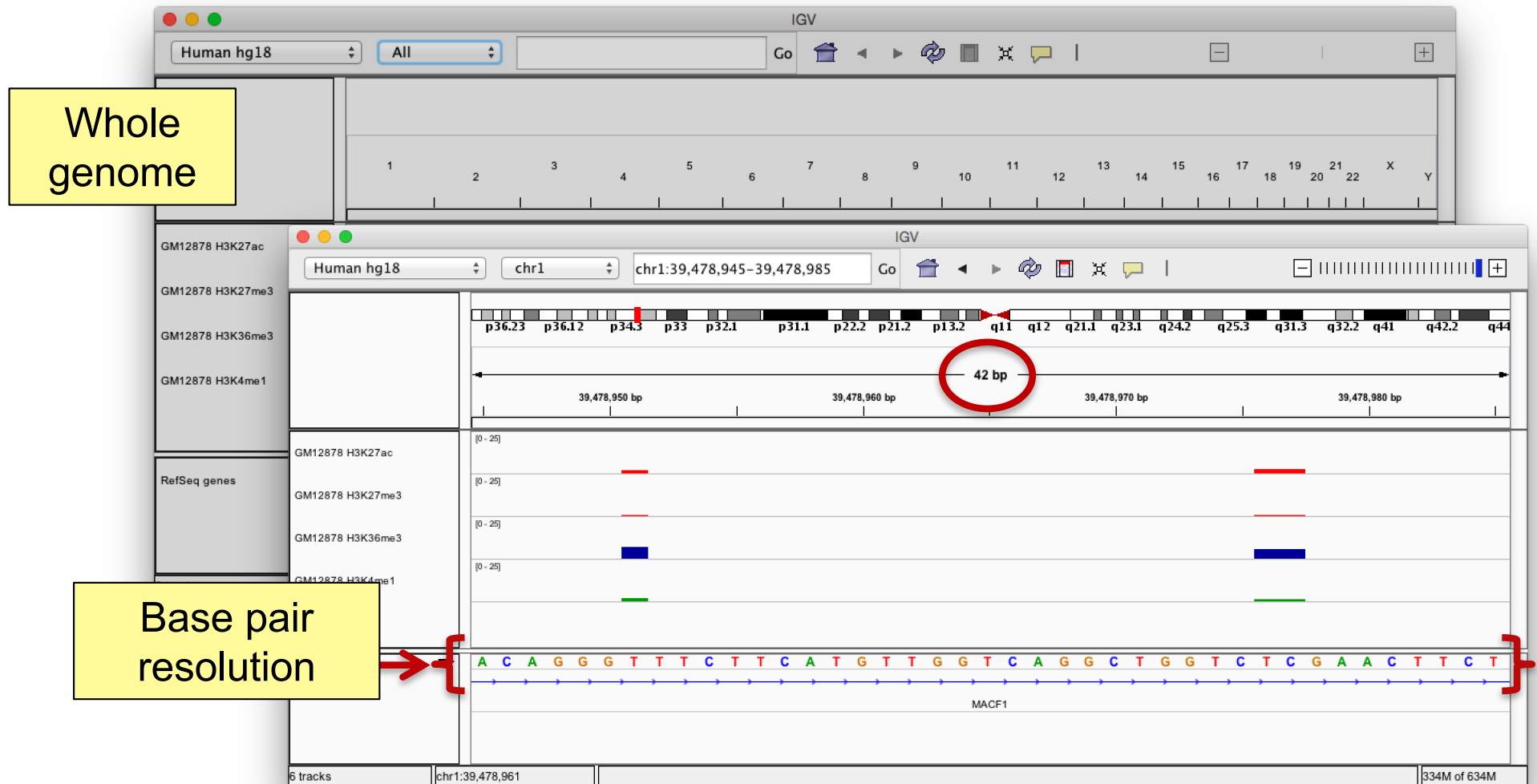
IGV user interface basics

Hands-on exercise



IGV user interface basics

Hands-on exercise



IGV user interface basics

Hands-on exercise

See handout

More basics – loading data

Other common ways of loading data

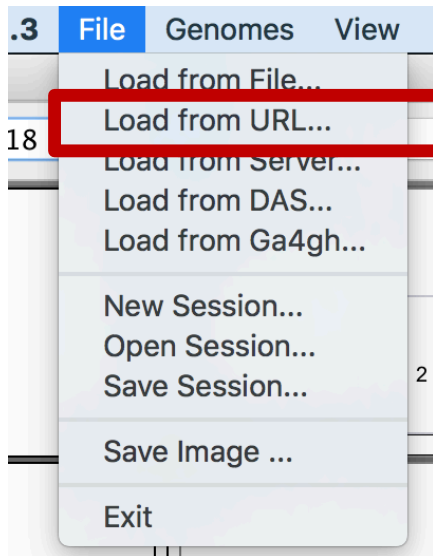
- Files > Load from File
- Files > Load from URL

More basics – loading data

Other common ways of loading data

- Files > Load from File
- Files > Load from URL

Try loading from URL:



1) Select *Files > Load from URL*

2) Enter

<http://www.igv.org/workshops/March2017/signal.tdf>

3) Click “OK”

OR

1) In a web browser, navigate to

<http://www.igv.org/workshops/March2017>

2) Drag the file ***signal.tdf*** and drop it onto the IGV window

More basics – track options

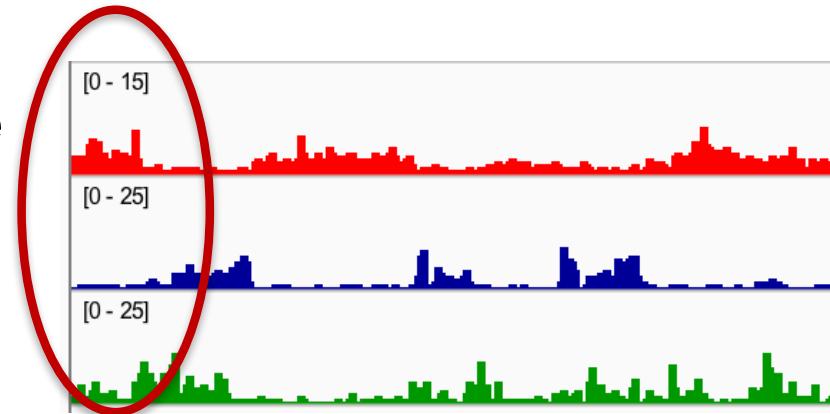
Right-click popup menu = track specific options menu

Common options

- Change track name, color, height
- Remove track

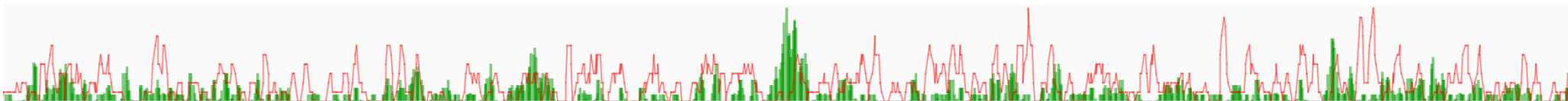
For quantitative data tracks: **Data Range**

- Set to an explicit range
- Enable log scale
- Enable autoscale
- Select multiple tracks and enable “group” autoscale



For quantitative data tracks: **Data Display:**

- Change between bar chart, line plot, point plot, heatmap
- Select multiple tracks and enable “overlay tracks”



More basics – track options

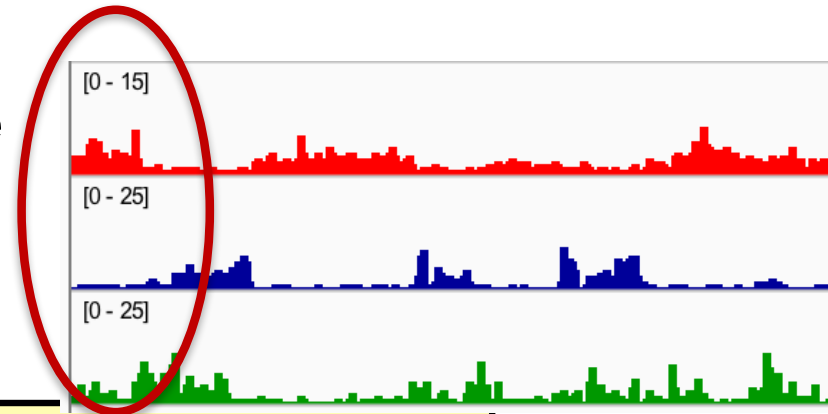
Right-click popup menu = track specific options menu

Common options

- Change track name, color, height.
- Remove track

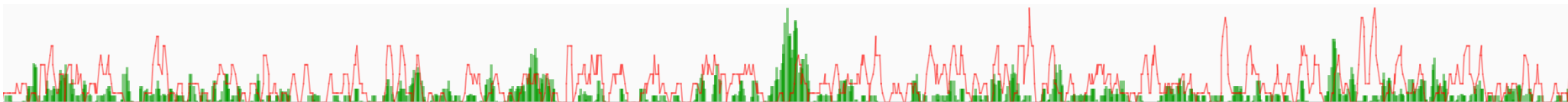
For quantitative data tracks: **Data Range**

- Set to an explicit range
- Enable log scale
- Enable autoscale
- Select multiple tracks and enable “group” autoscale

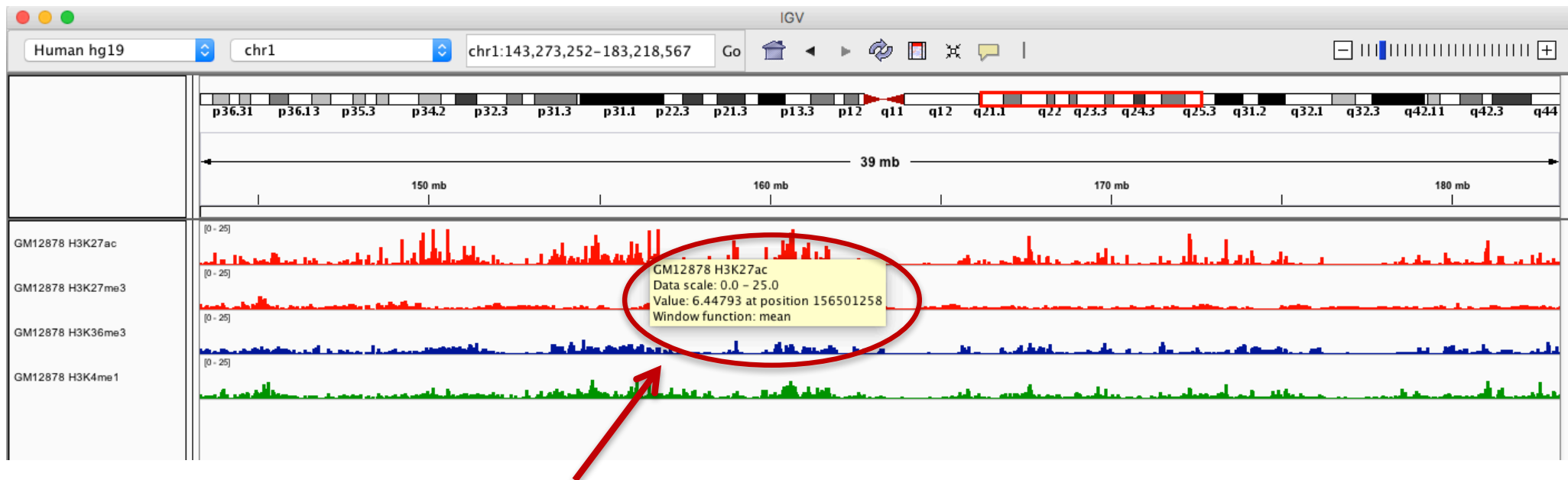


For quantitative data tracks: **Data Display:**

- Change between bar chart, line plot, point plot, heatmap
- Select multiple tracks and enable “overlay tracks”



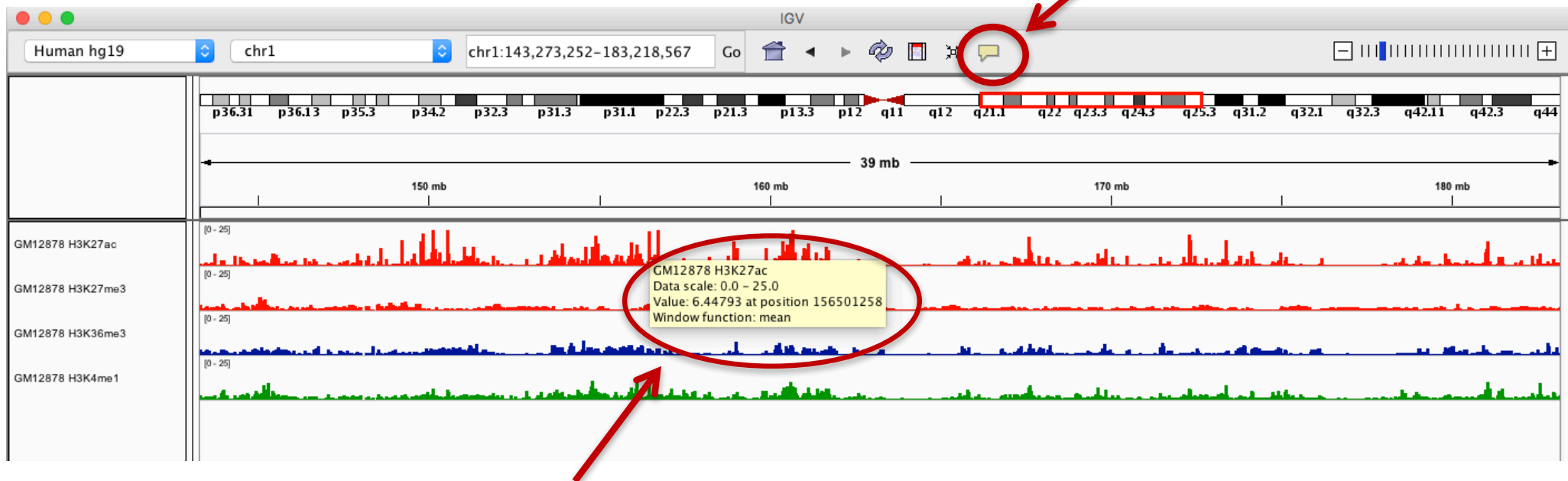
More basics – data details popup



By default, a “data detail” window pops up as the the cursor hovers over a data element in a track

More basics – data details popup

Change the popup behavior by clicking on the yellow balloon icon.
Options are (1) ***On hover*** (2) ***On click*** (3) ***Never***



By default, a “data detail” window pops up as the the cursor hovers over a data element in a track

More basics – the *igv* folder

In your home folder, IGV creates a folder named **igv** to store:

igv.log file

- Logs IGV activity and error messages
- If you send us a bug report or help question, we may ask for a copy of this file
- To start a fresh log, you can delete or rename this file (but not while IGV is running)

prefs.properties file

- Stores all the preference you set in *View > Preferences* so they are remembered across IGV sessions
- You can reset all preferences by deleting or renaming this file (but not while IGV is running)

genomes folder

- Stores info about the genomes you've loaded – for the *genomes dropdown menu*.

Viewing Multiple Regions

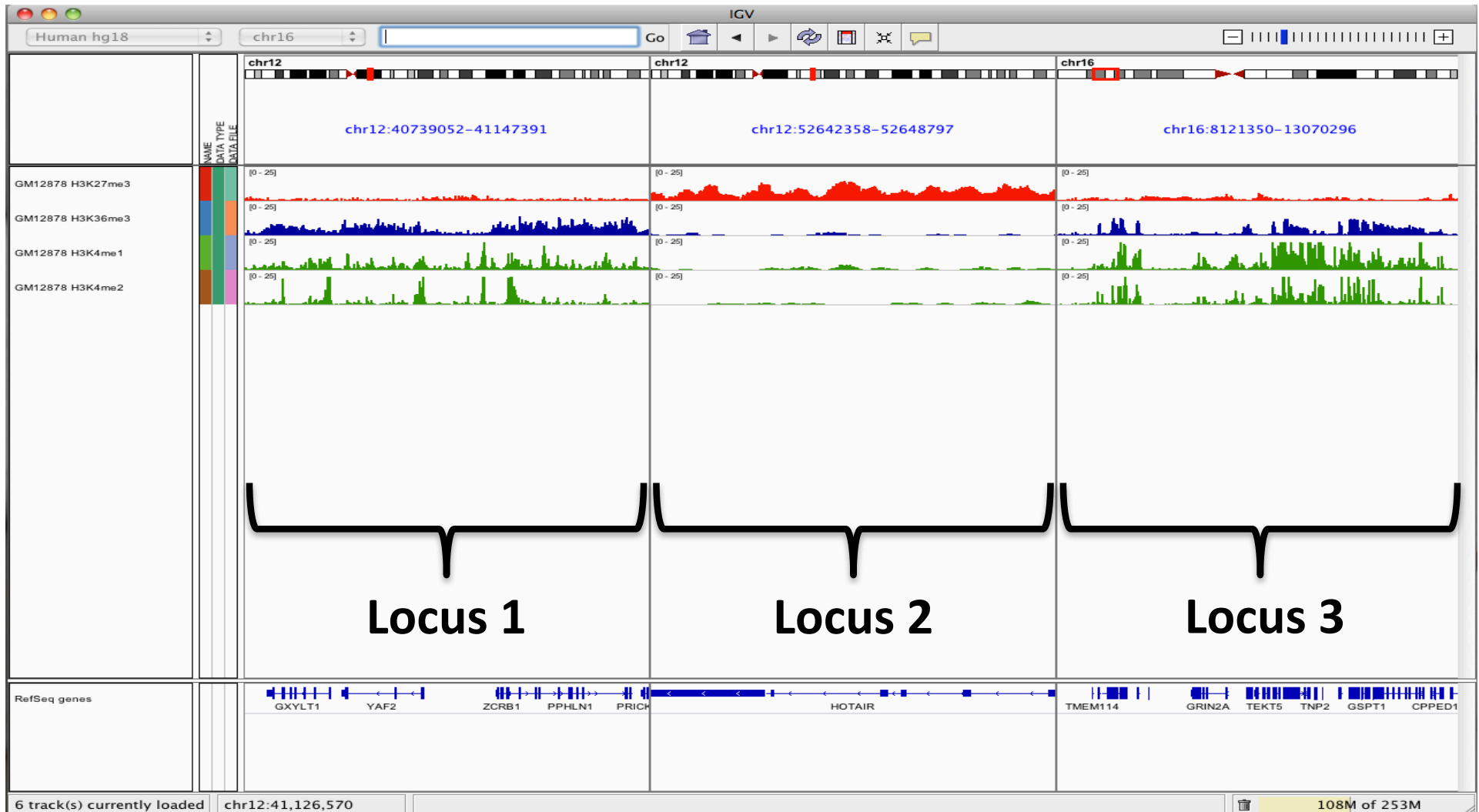
Viewing multiple regions

multi-locus or split-screen view



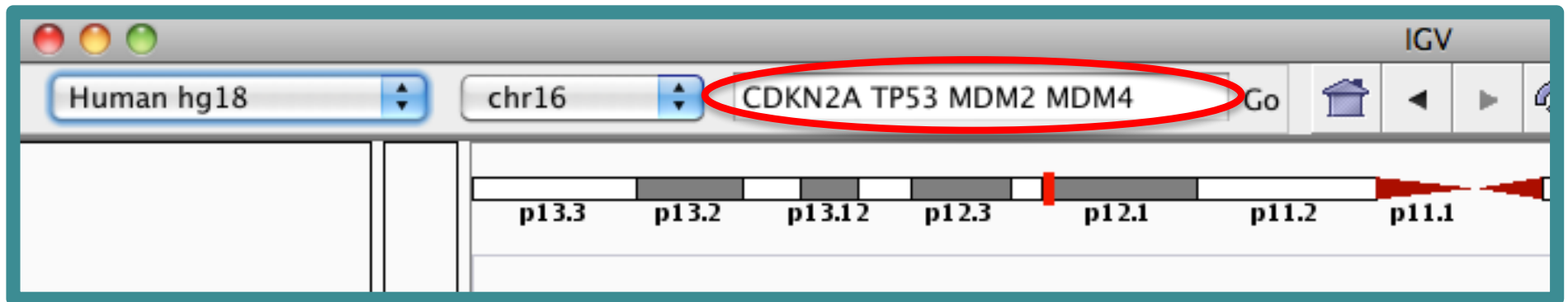
Viewing multiple regions

multi-locus or split-screen view



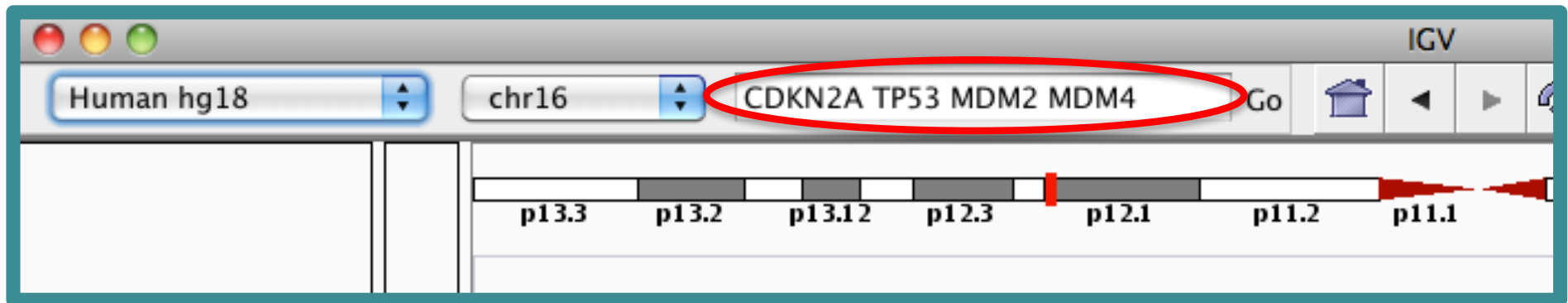
Multiple regions – search box

Enter multiple loci or features in the search box



Multiple regions – search box

Enter multiple loci or features in the search box



For a more persistent list of regions, use ***Regions of Interest***

Multiple regions – “regions of interest” tools

For a more persistent list of regions, use *Regions of Interest*

To **define** regions:

(1) Use the Region of Interest tool

- Click on the tool icon in the tool bar
- Click twice in the data panel to specify the region boundaries



Multiple regions – “regions of interest” tools

For a more persistent list of regions, use *Regions of Interest*

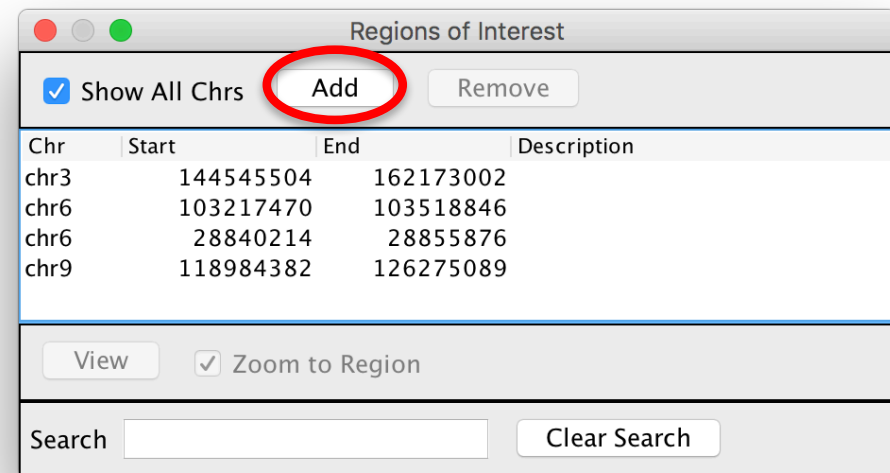
To **define** regions:

(1) Use the Region of Interest tool

- Click on the tool icon in the tool bar
- Click twice in the data panel to specify the boundaries

(2) Use the Region Navigator

- Navigate to the new region
- Select *Regions > Region Navigator*
- Click **Add** to create a region from the current view



Multiple regions – “regions of interest” tools

For a more persistent list of regions, use *Regions of Interest*

To **define** regions:

(1) Use the Region of Interest tool

- Click on the tool icon in the tool bar
- Click twice in the data panel to specify the boundaries

(2) Use the Region Navigator

- Navigate to the new region
- Select *Regions > Region Navigator*
- Click **Add** to create a region from the current view

(3) **Import a file** of pre-defined regions

- Prepare a .BED file containing the region definitions
- Select *Regions > Import Regions*

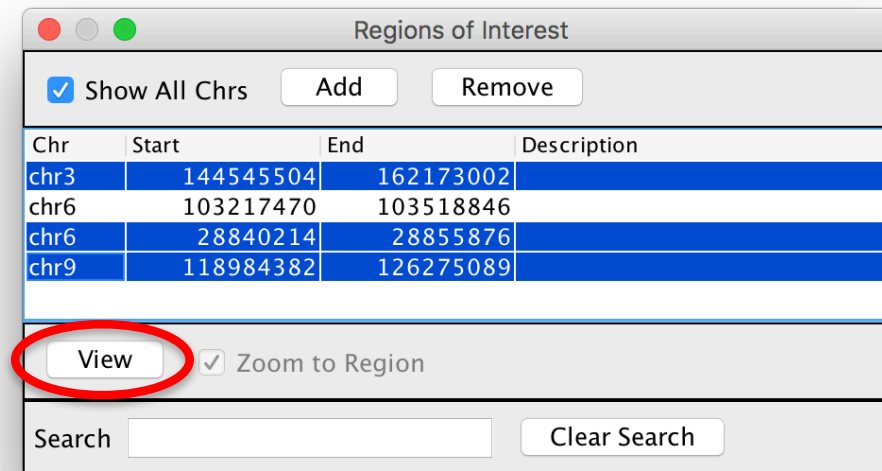
Multiple regions – “regions of interest” tools

For a more persistent list of regions, use *Regions of Interest*

To **view** multiple regions of interest you defined:

Use the Region Navigator

- Select *Regions* > *Region Navigator*
- Select 2 or more regions
- Click **View**



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**

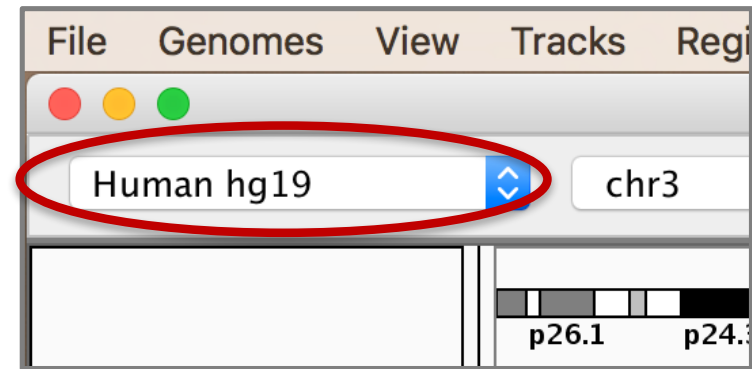


more about loading

Reference Genomes

Genome menu

The genome drop-down menu is in the upper left corner of the IGV window.



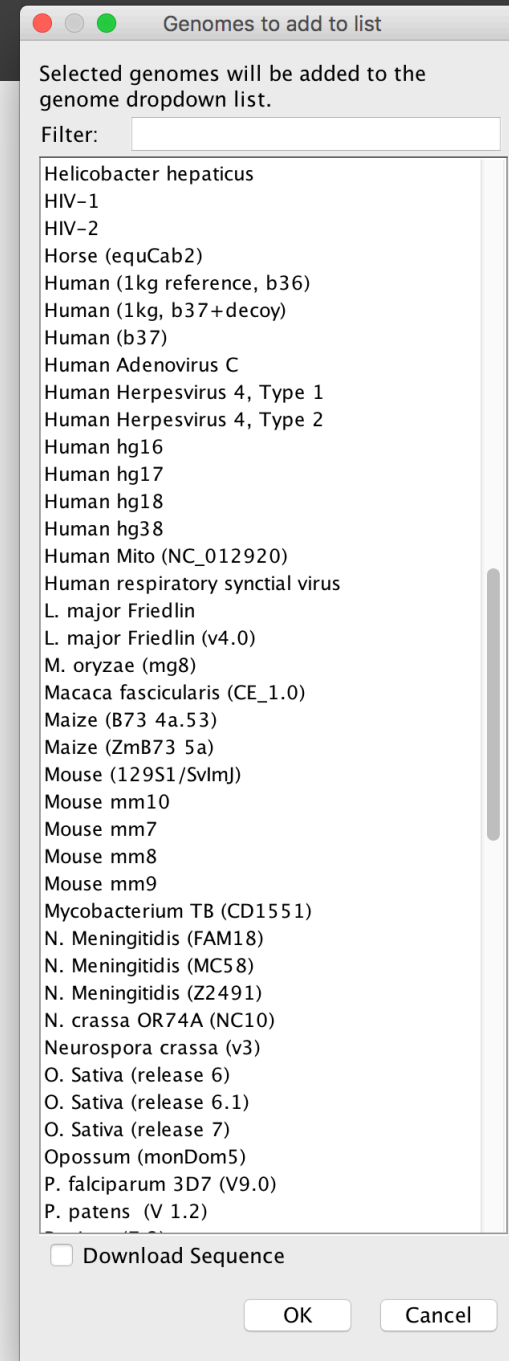
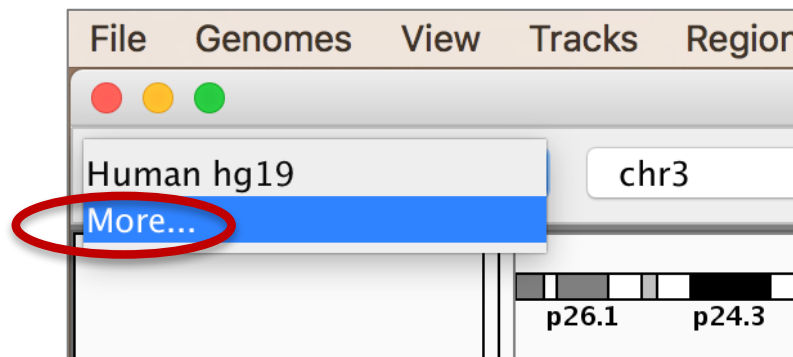
The genome menu is short when you first install IGV.

But IGV **hosts dozens of genomes** and you can **load other genomes** too.

More hosted genomes

To get more genomes from the IGV server:

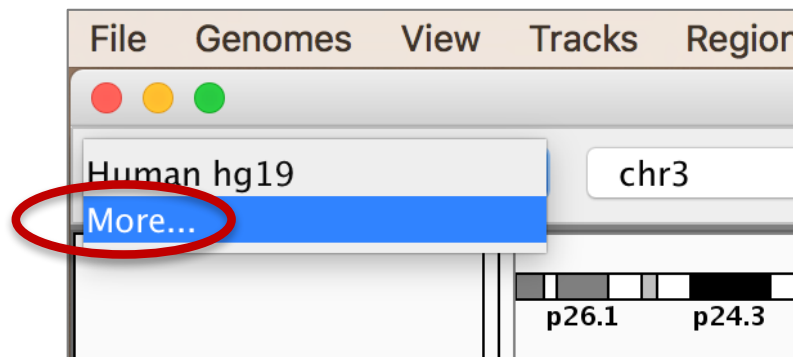
- (1) Select **More** from the genome menu to see the complete list



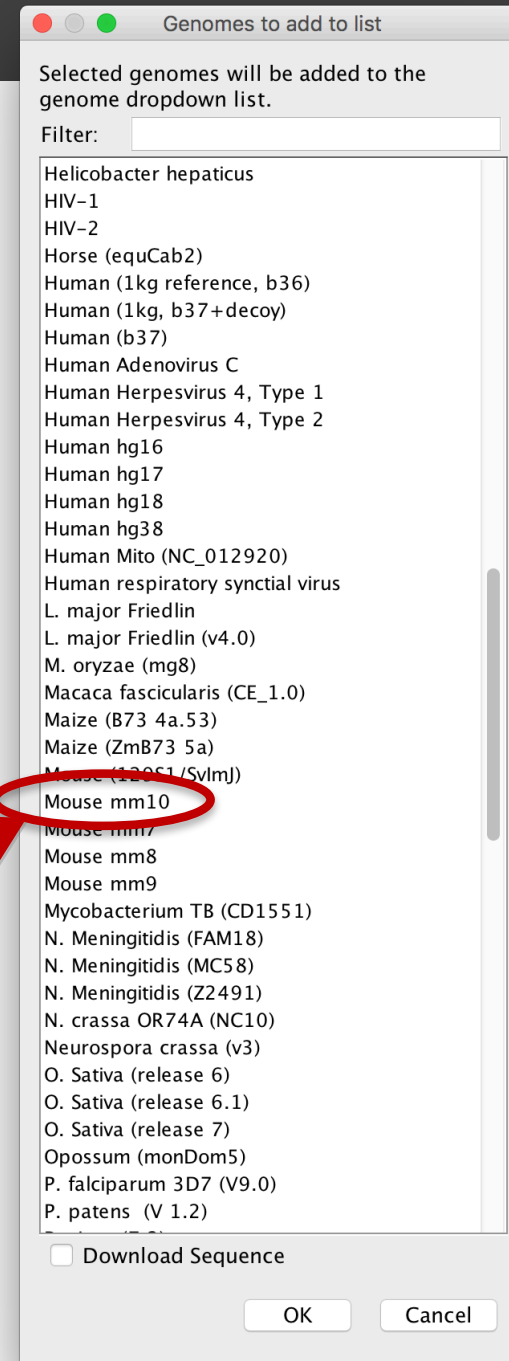
More hosted genomes

To get more genomes from the IGV server:

- (1) Select **More** from the genome menu to see the complete list



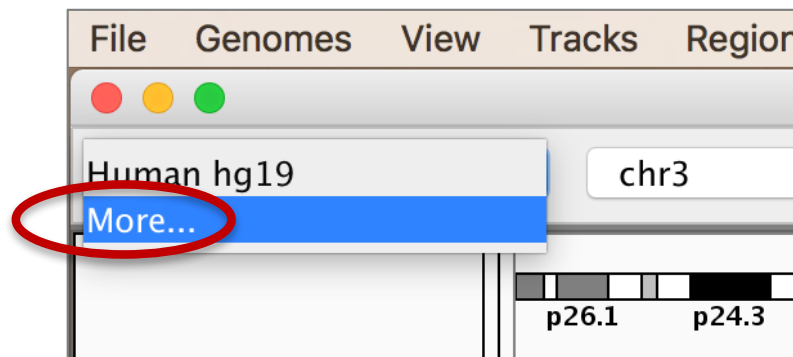
- (2) Click on entry in the list, e.g. **Mouse mm10**



More hosted genomes

To get more genomes from the IGV server:

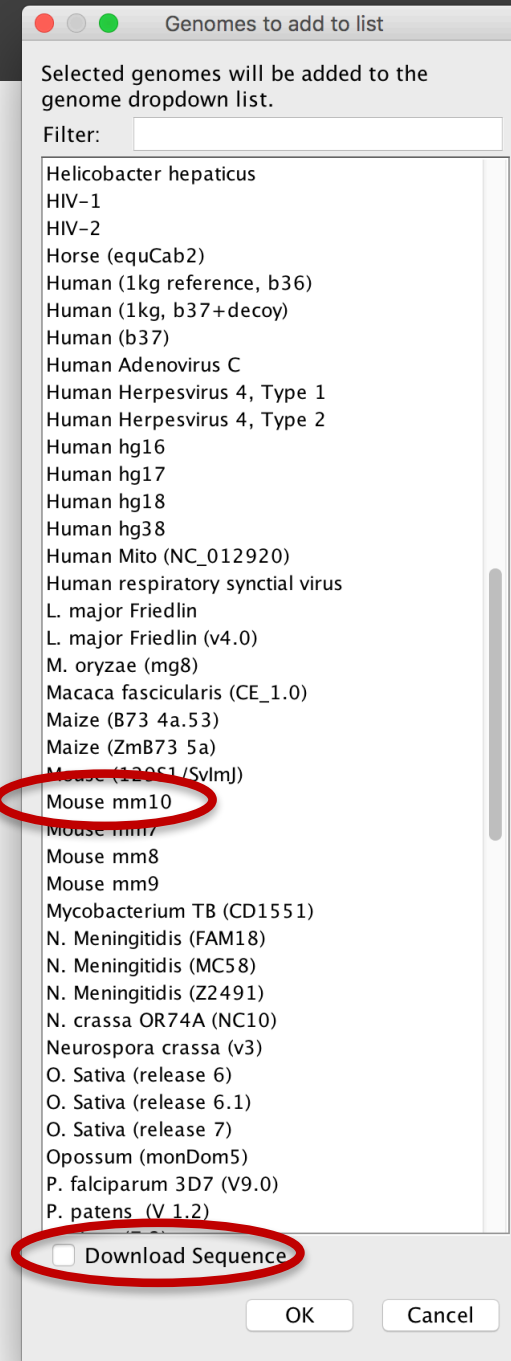
- (1) Select **More** from the genome menu to see the complete list



- (2) Click on entry in the list, e.g. **Mouse mm10**

- (3) Optionally select **Download Sequence**

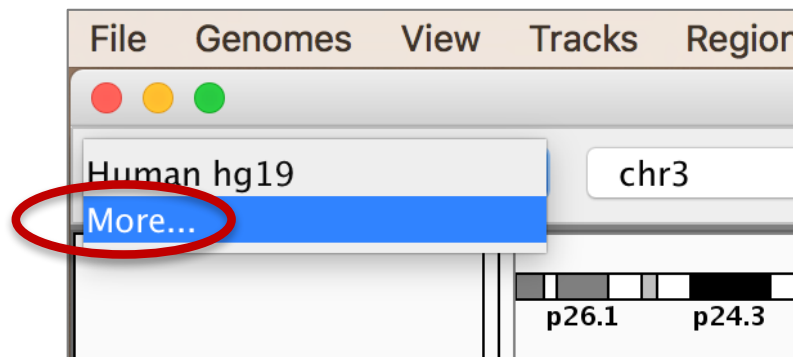
Do **not** download the sequence today!



More hosted genomes

To get more genomes from the IGV server:

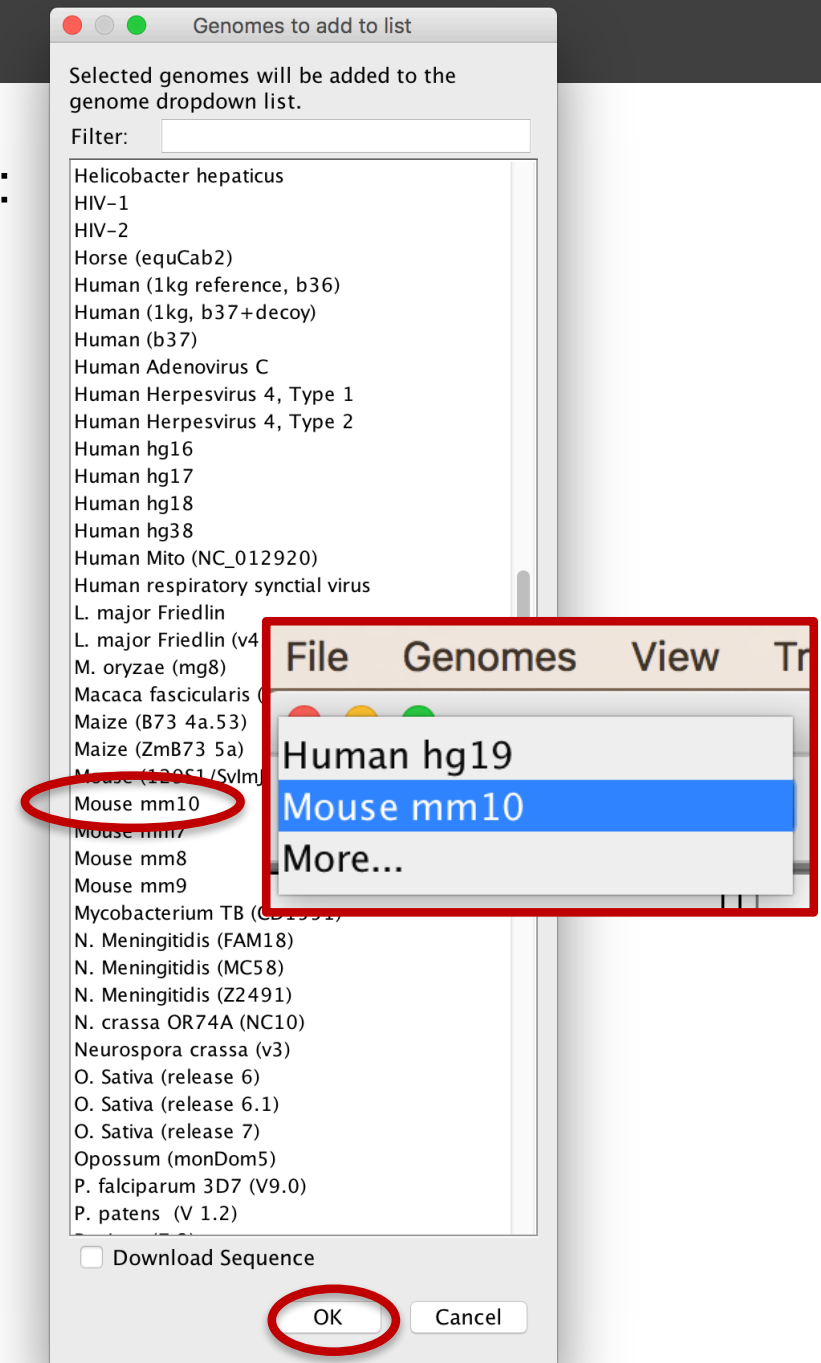
(1) Select **More** from the genome menu to see the complete list



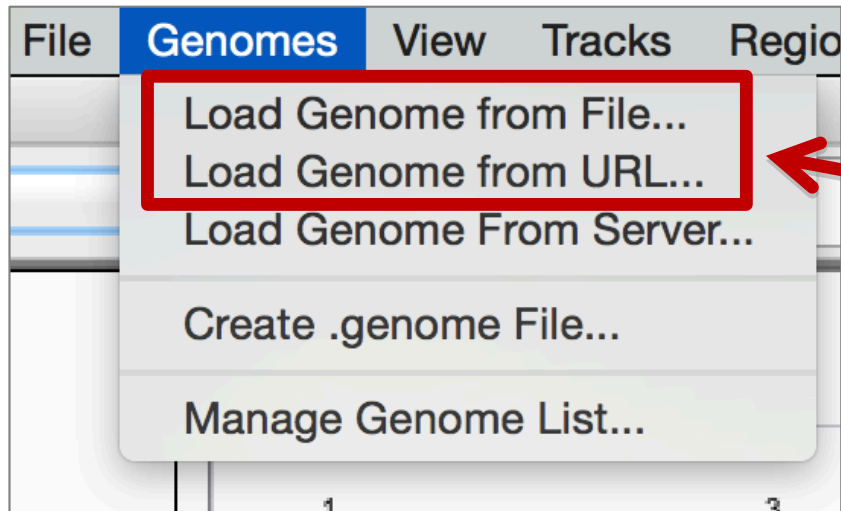
(2) Click on entry in the list, e.g. **Mouse mm10**

(3) Optionally select **Download Sequence**

(4) Click **OK**

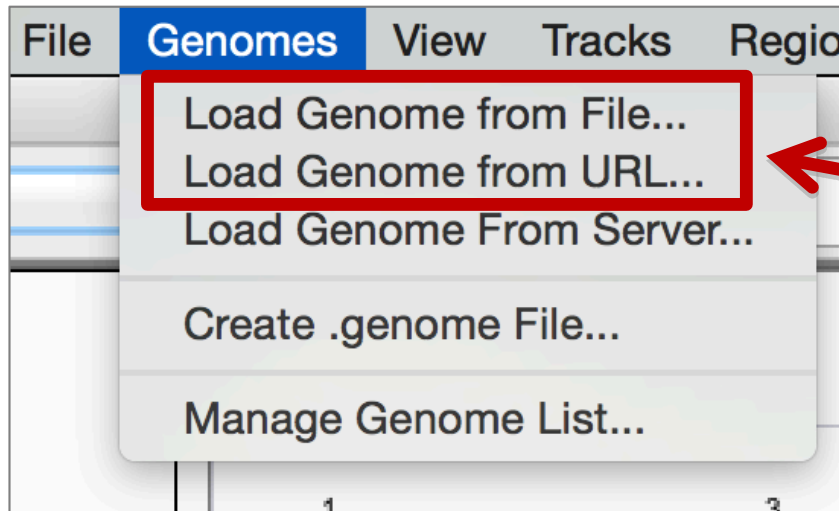


Loading other genomes



Use **Genomes** menu to load genome from an indexed FASTA file

Loading other genomes



Use **Genomes** menu to load genome from an indexed FASTA file

- (1) Use *samtools faidx* command to create .fai index file
See github.com/samtools/samtools for info about samtools
- (2) If no index file found, IGV will try to create the index upon load.

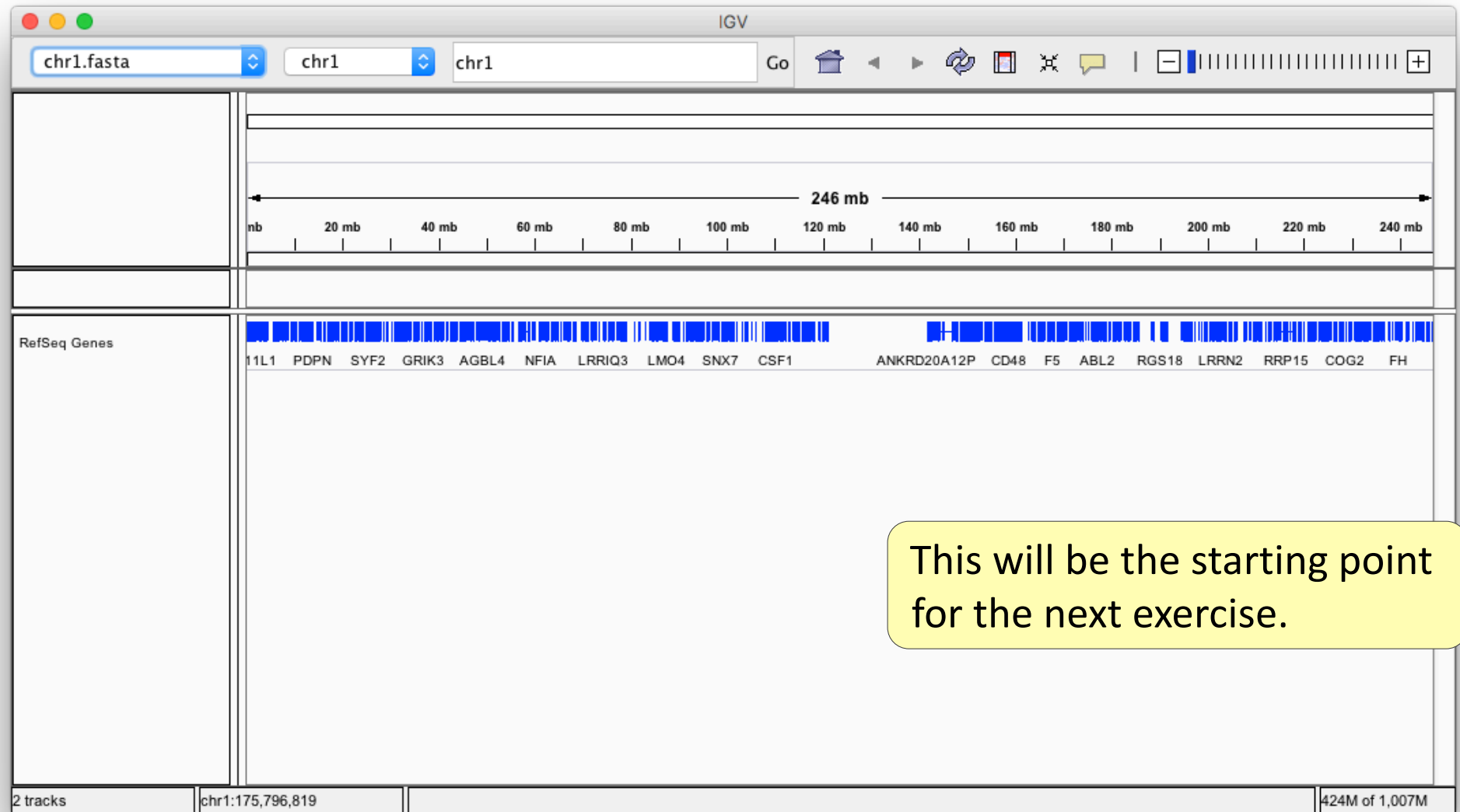
Load genome from file

Hands-on exercise

See handout

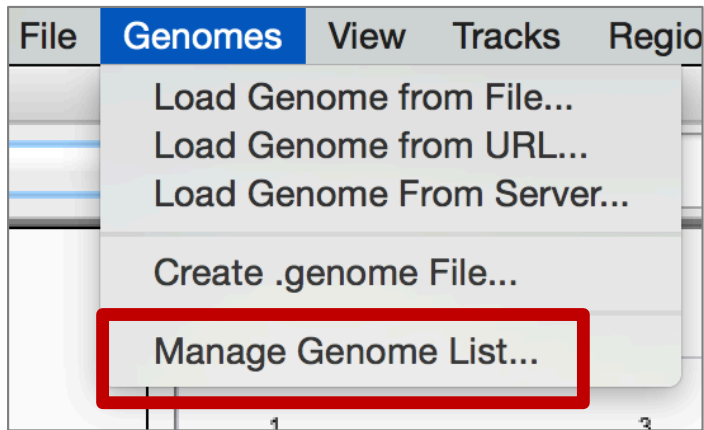
Load genome from file

Hands-on exercise



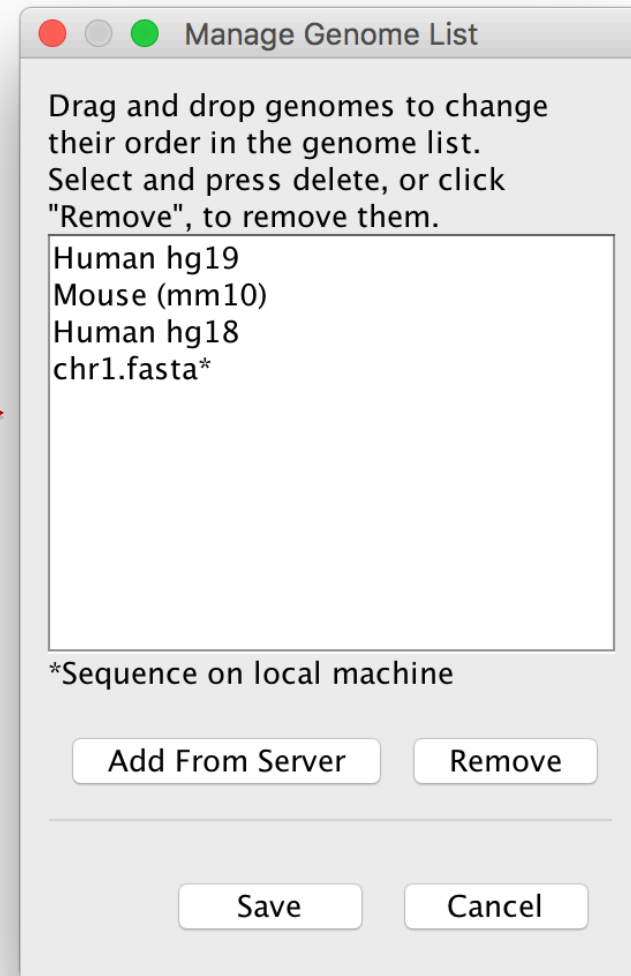
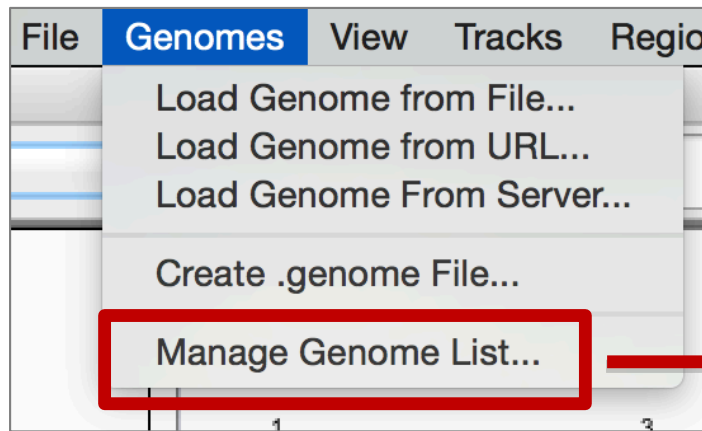
Removing genome from list

Genomes > Manage Genome List



Removing genome from list

Genomes > Manage Genome List

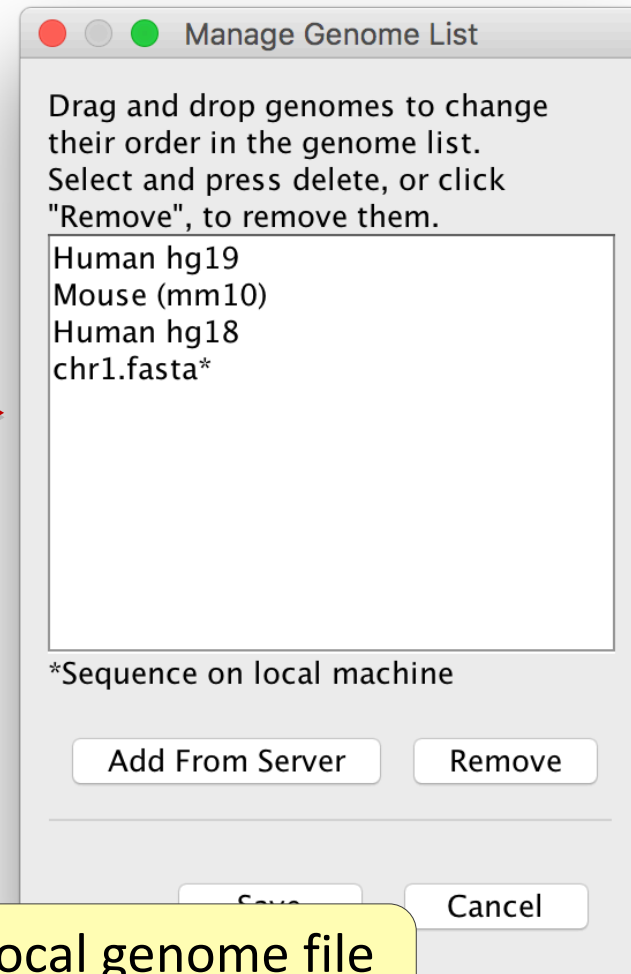
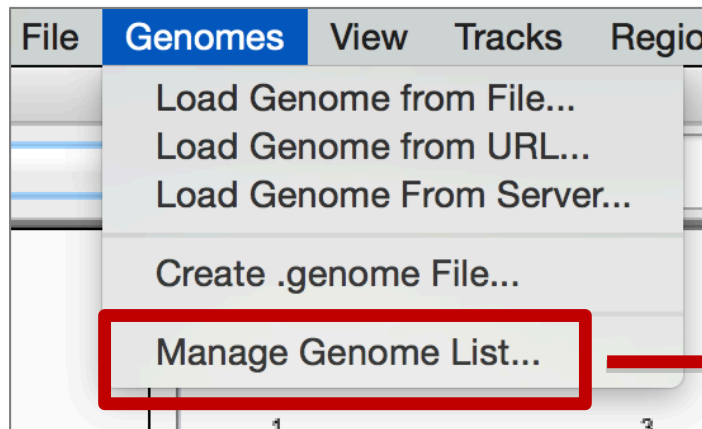


- Remove genomes from list
- Rearrange list
- Add to list

Do **not** remove chr1.fasta now.
We're using it for the next exercise.

Removing genome from list

Genomes > Manage Genome List



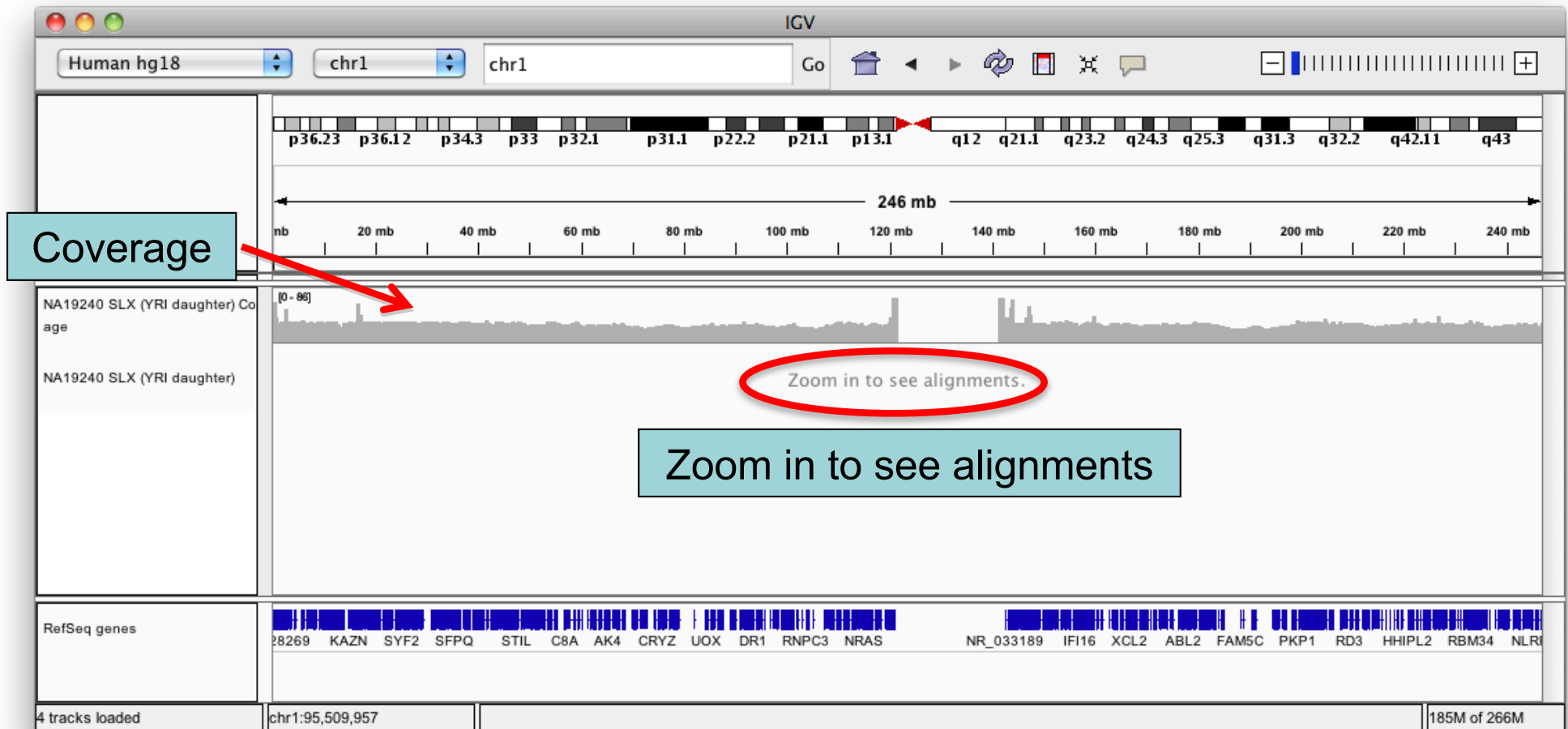
- Remove genomes from list
- Rearrange list
- Add to list

Note: IGV can get confused if you delete a local genome file from the file system but not from the IGV menu

Viewing Next Generation Sequencing (NGS) Data

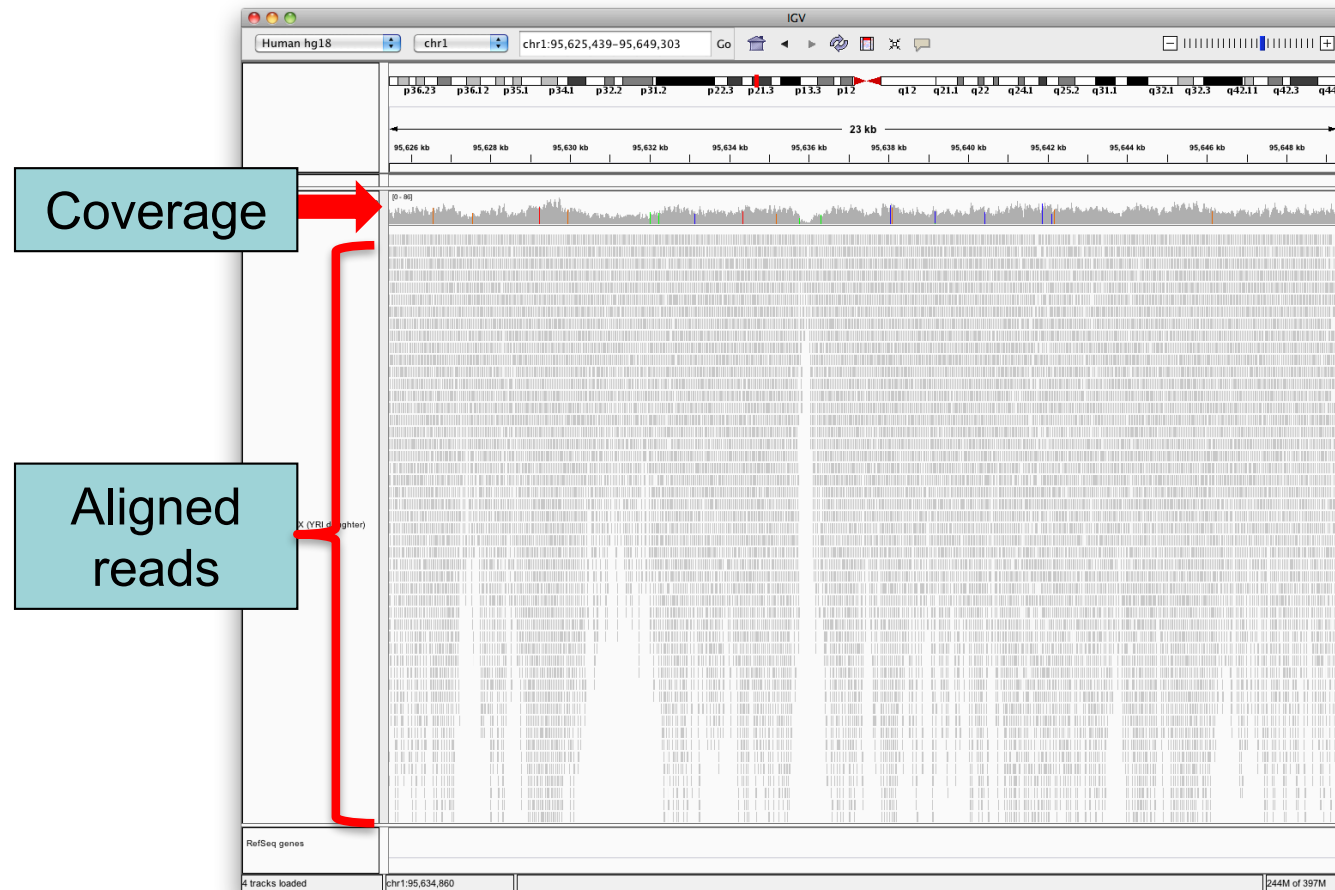
Viewing alignments

Whole chromosome view



Viewing alignments

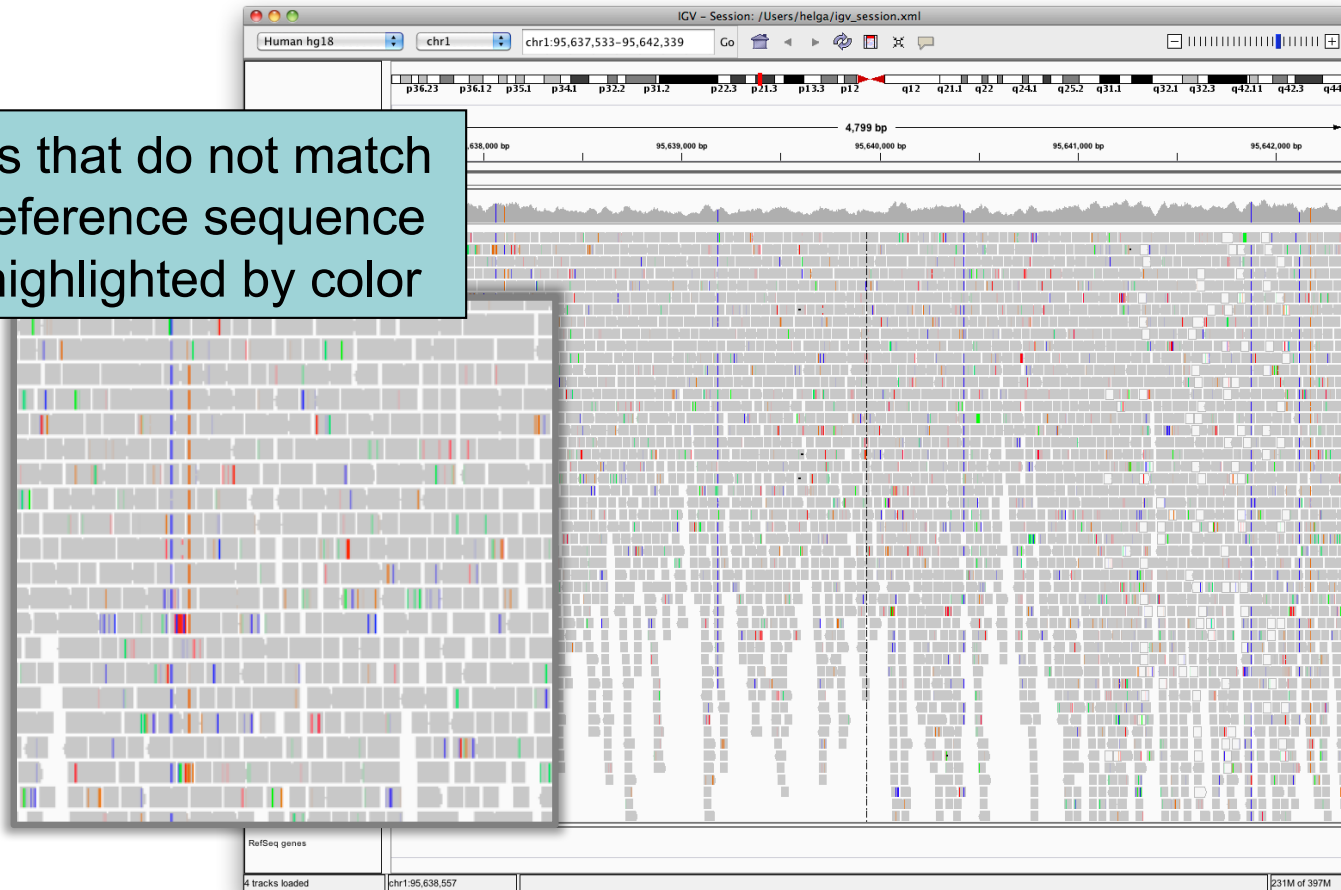
Zoom in to view alignments



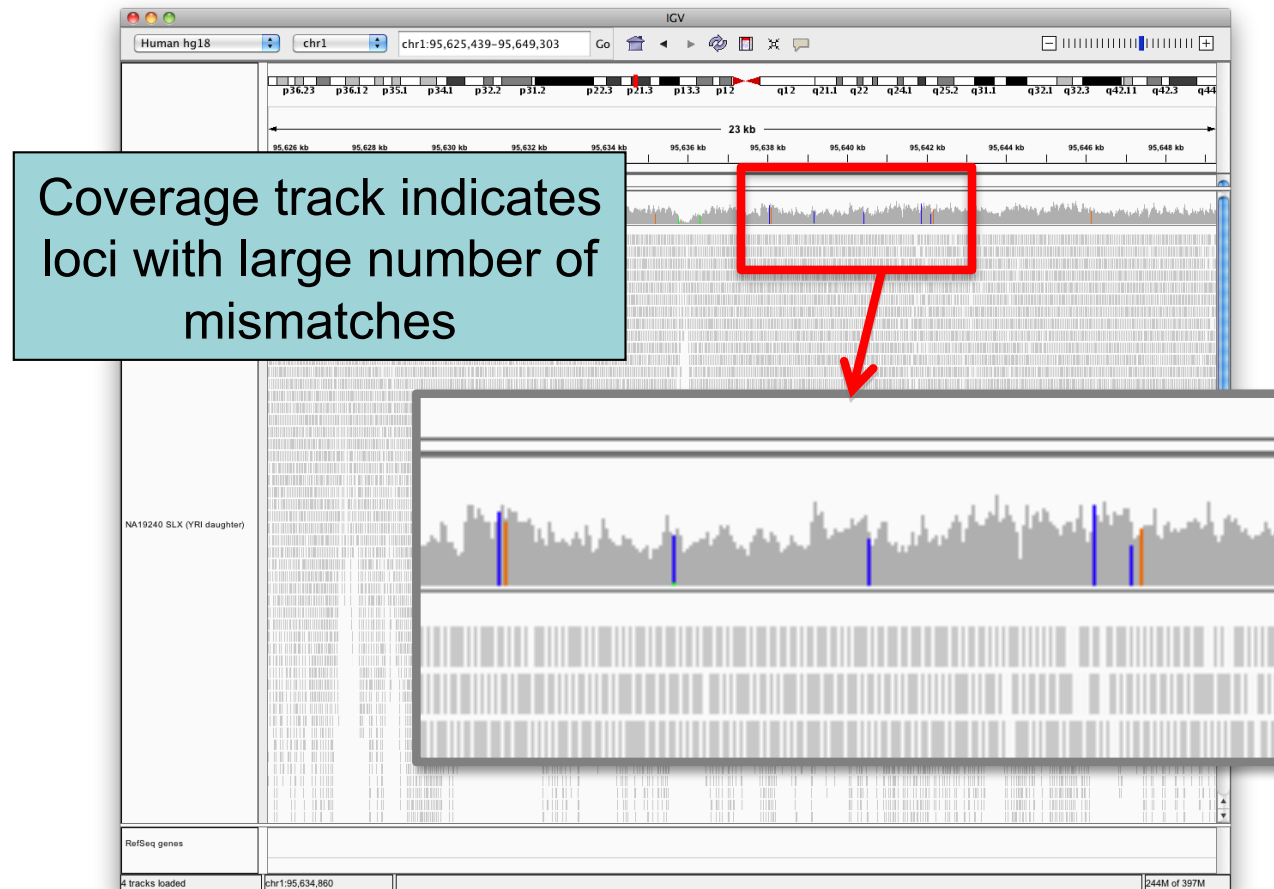
Viewing alignments

Zoom in to see more detail

Bases that do not match the reference sequence are highlighted by color



Viewing alignments

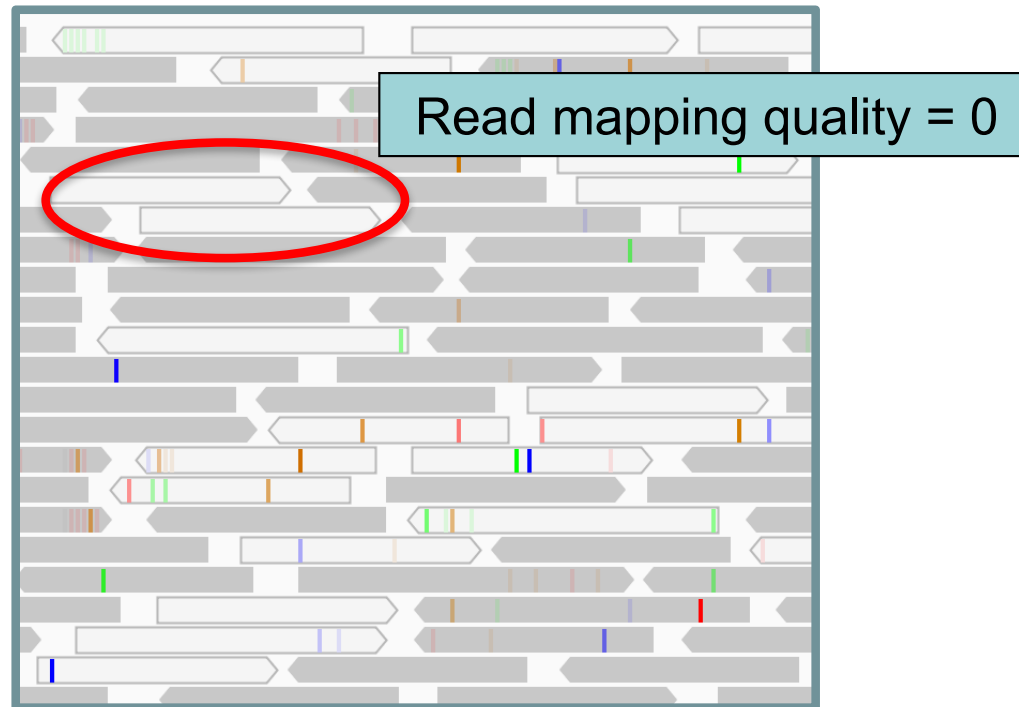


Viewing alignments

Zoom in to see more detail



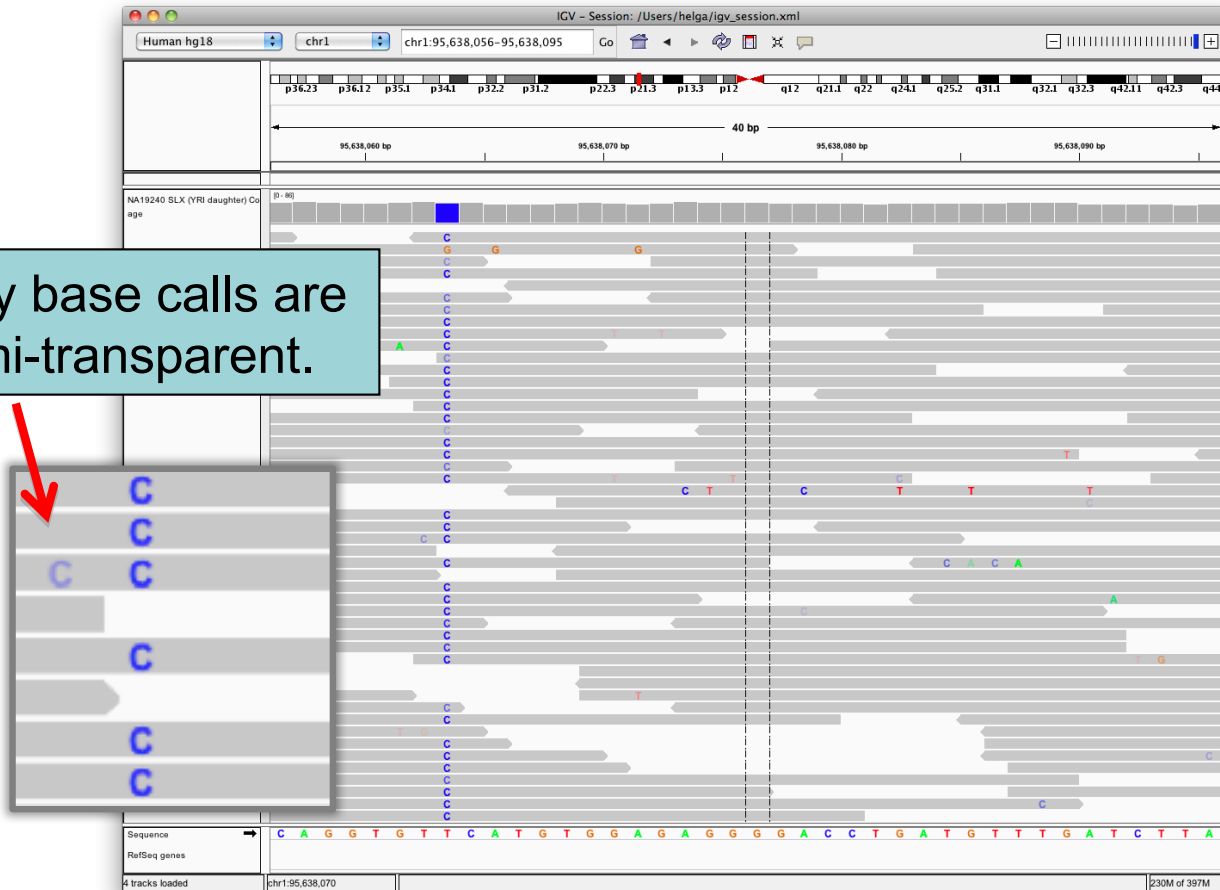
Viewing alignments



Viewing alignments

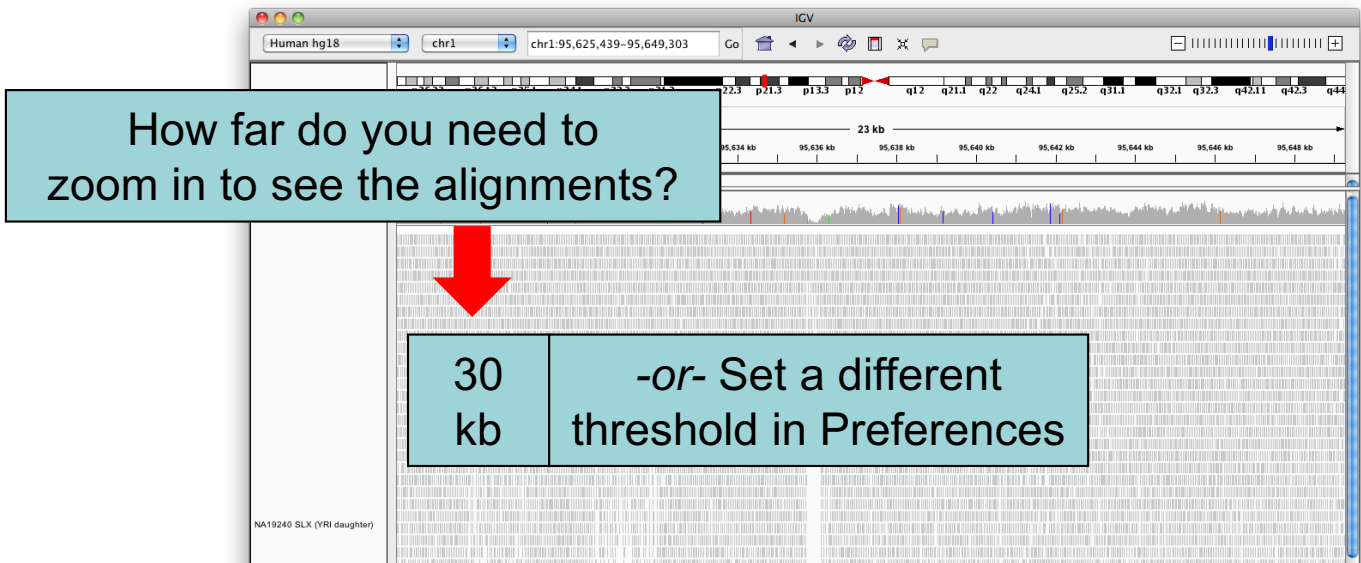
Zoom in to see more detail

Low-quality base calls are faint, semi-transparent.



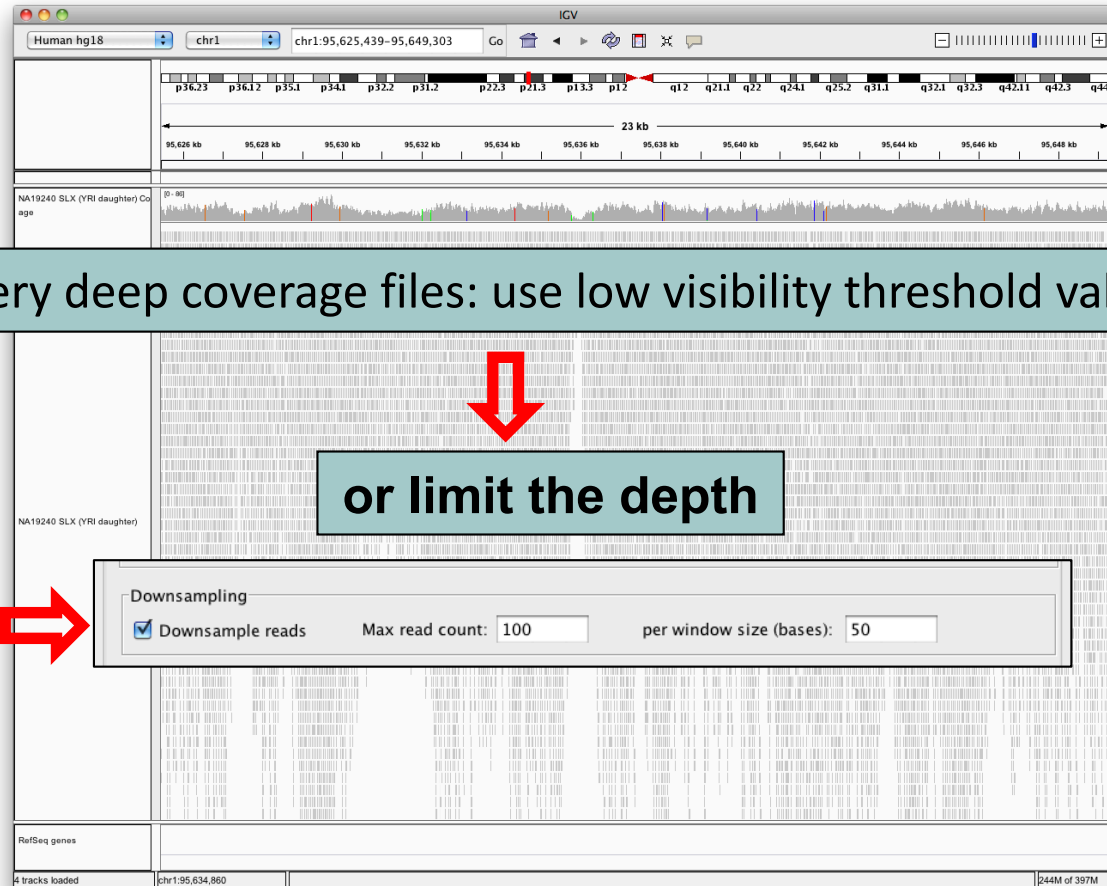
Viewing alignments

Zoom in to view alignments



- Higher value → requires more memory
- Low coverage files → ok to use higher value
- Very deep coverage files → use lower value

Viewing alignments



Viewing alignments



Viewing SNPs

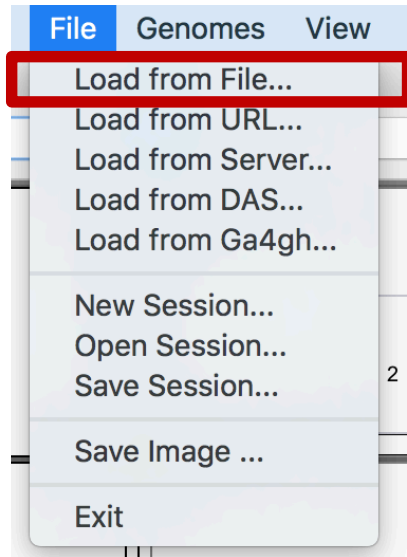
Hands-on exercise

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

Follow along

Viewing SNPs

Hands-on exercise



Click *File > Load from File*

Navigate to IGV folder that was pre-installed on the computer, and then to **Data / snps**

Open

NA12878.SLX.sample.bam

and

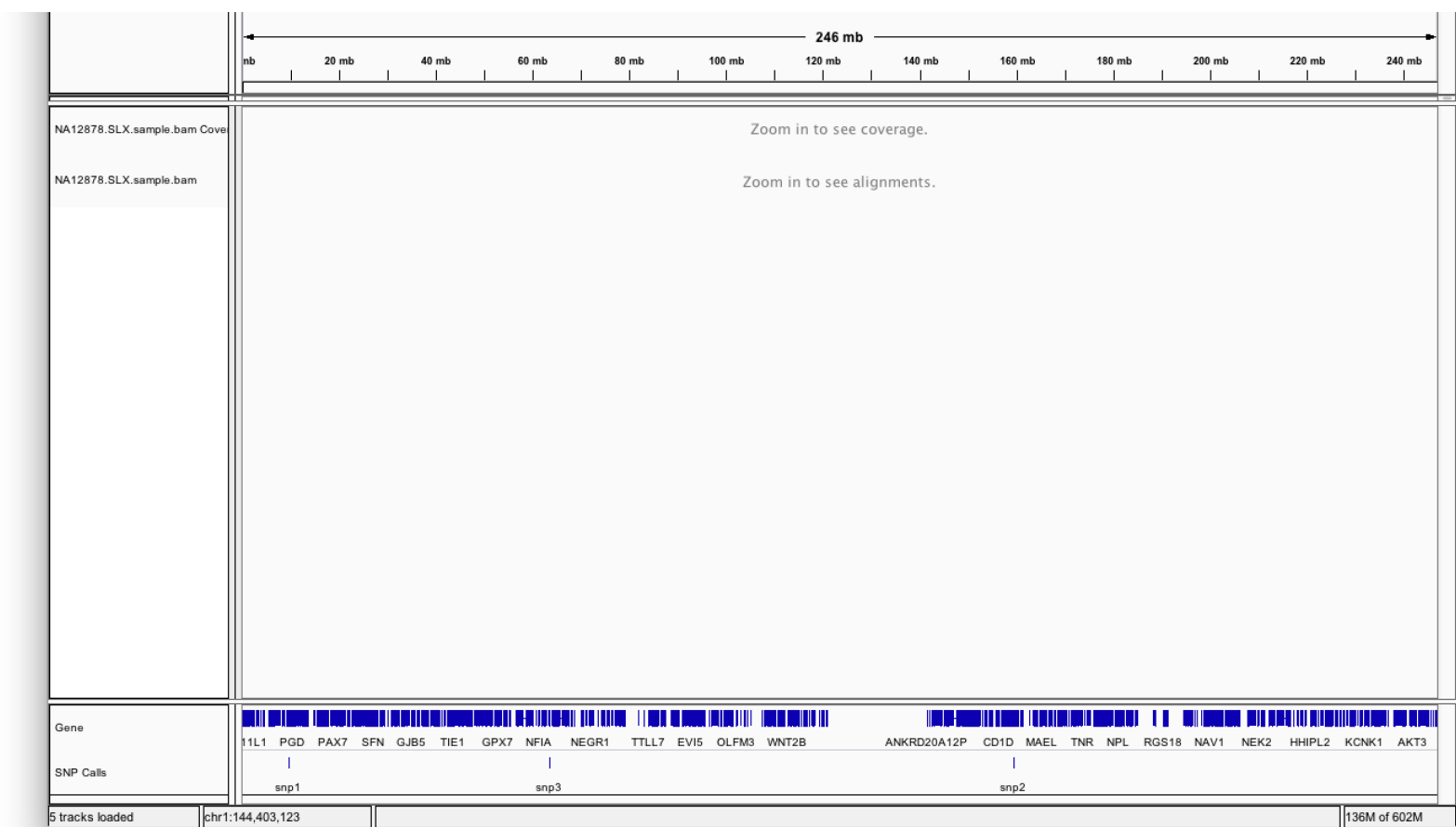
snp_calls.bed

not the .bai file

Note: This assumes you successfully loaded the genome in the previous exercise.

Viewing SNPs

Hands-on exercise

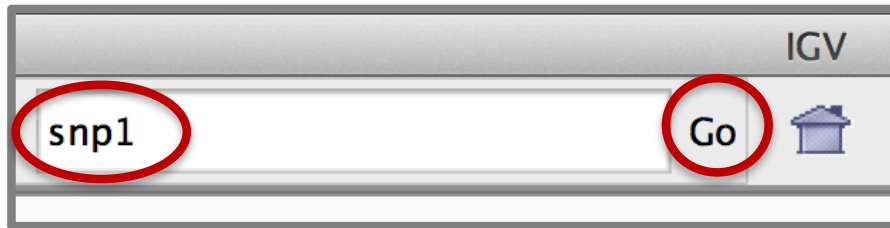


Viewing SNPs

Hands-on exercise

Navigate to first putative SNP locus

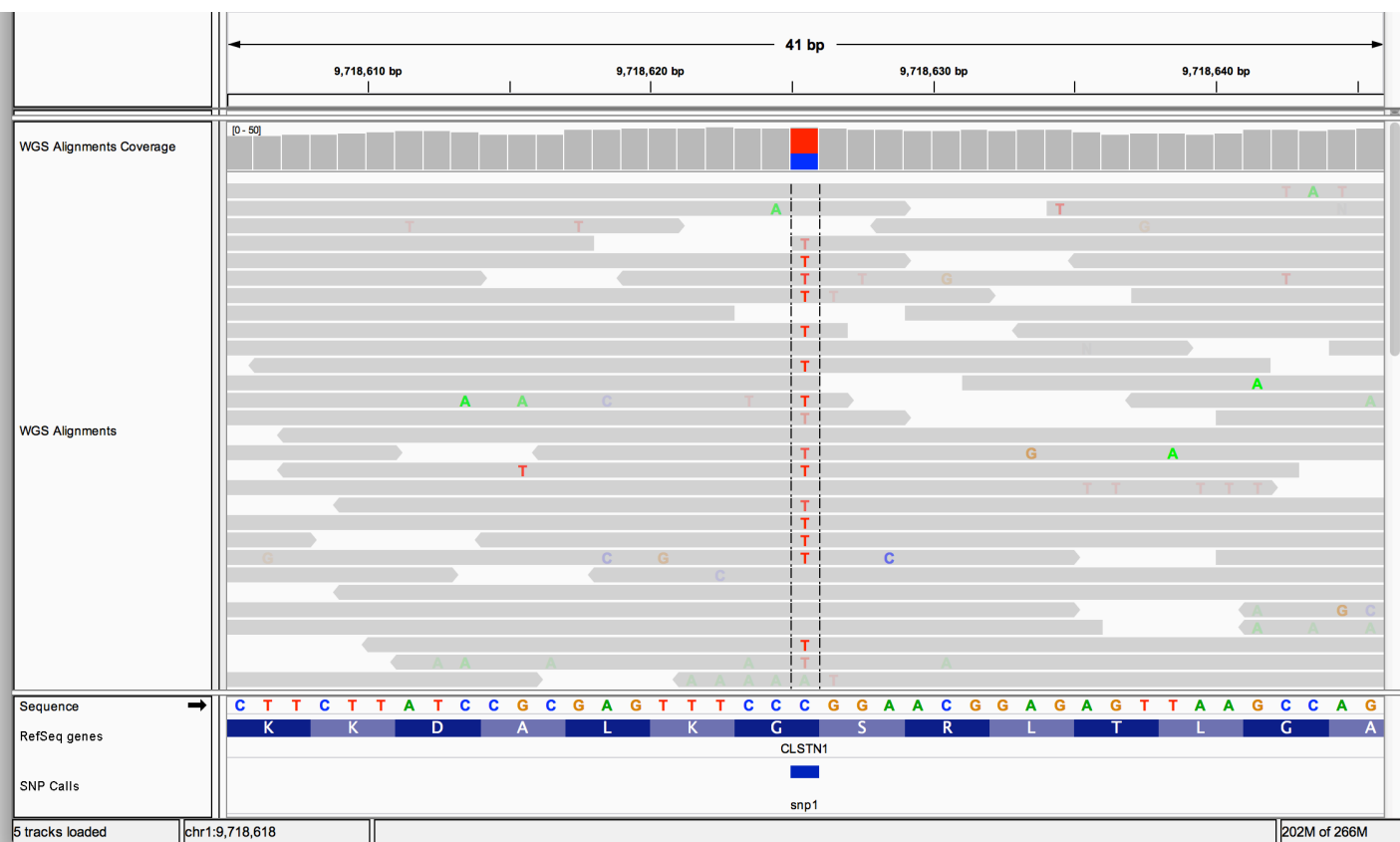
Type “snp1” in the **Search Box** and click **Go**



The image shows a screenshot of the IGV (Integrative Genomics Viewer) search interface. The title bar at the top right says "IGV". Below it is a search bar containing the text "snp1". To the right of the search bar is a button labeled "Go" and a home icon. Both the search bar and the "Go" button are circled in red to indicate they should be used for the exercise.

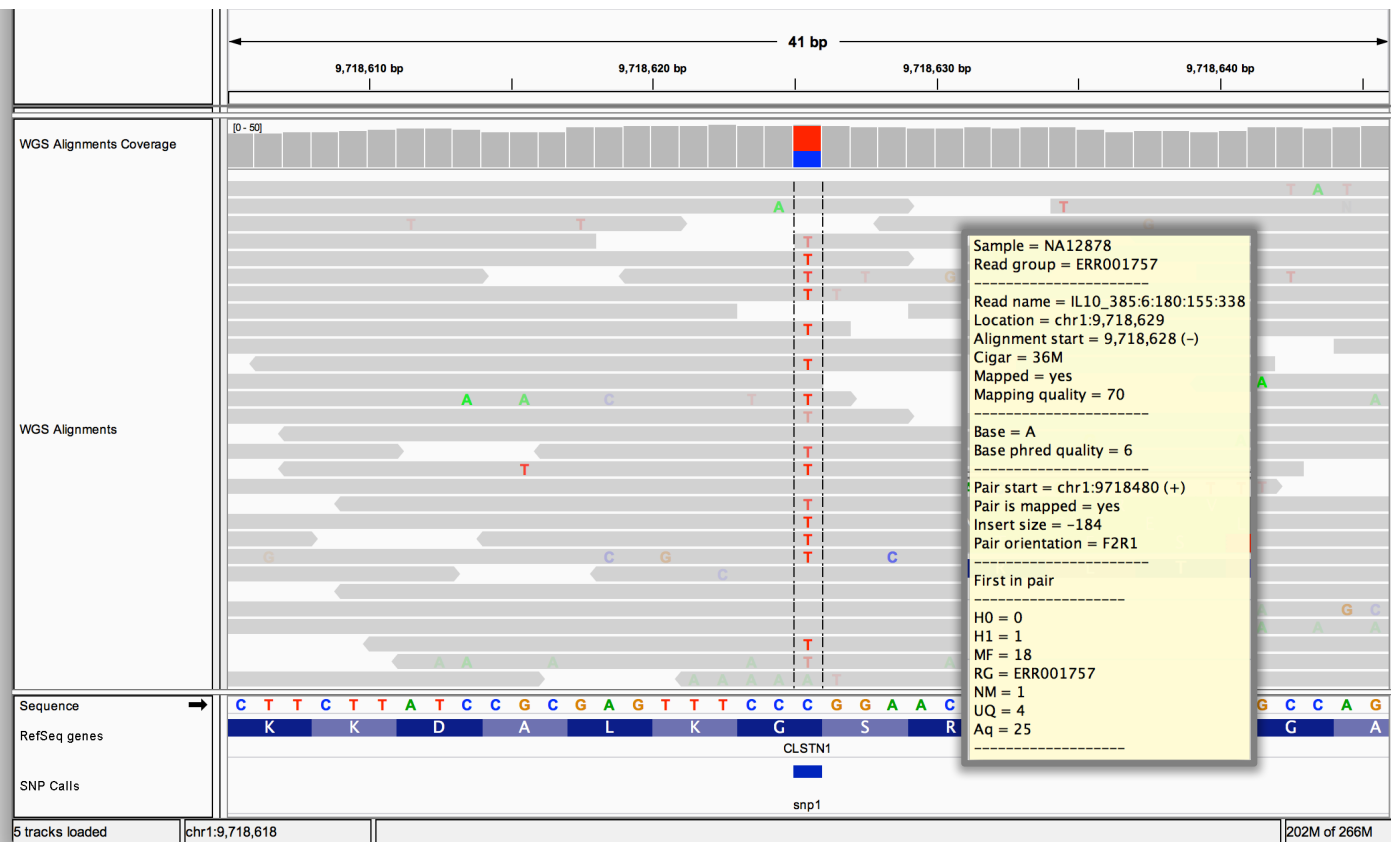
Viewing SNPs

Hands-on exercise



Viewing SNPs

Hands-on exercise

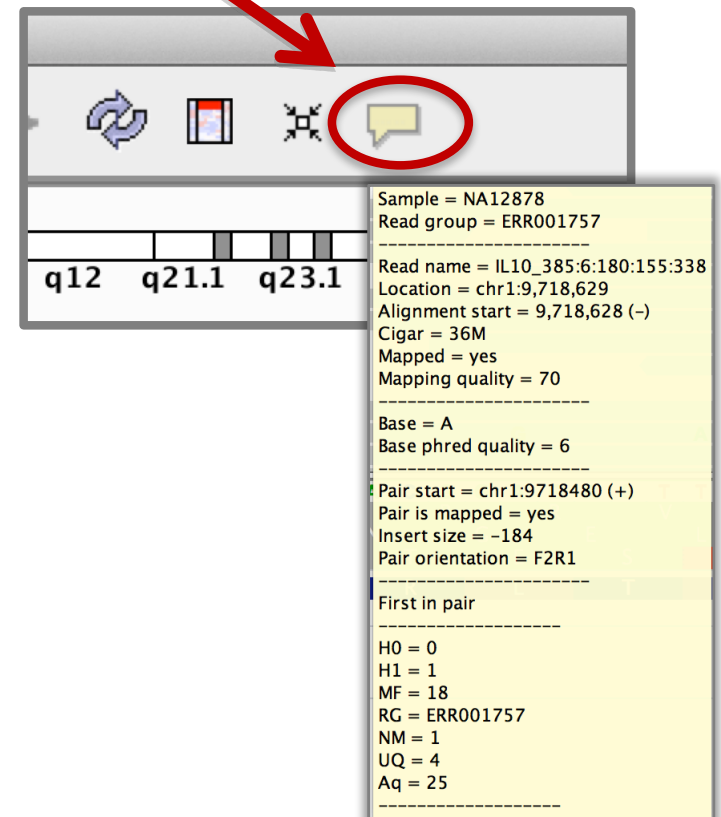
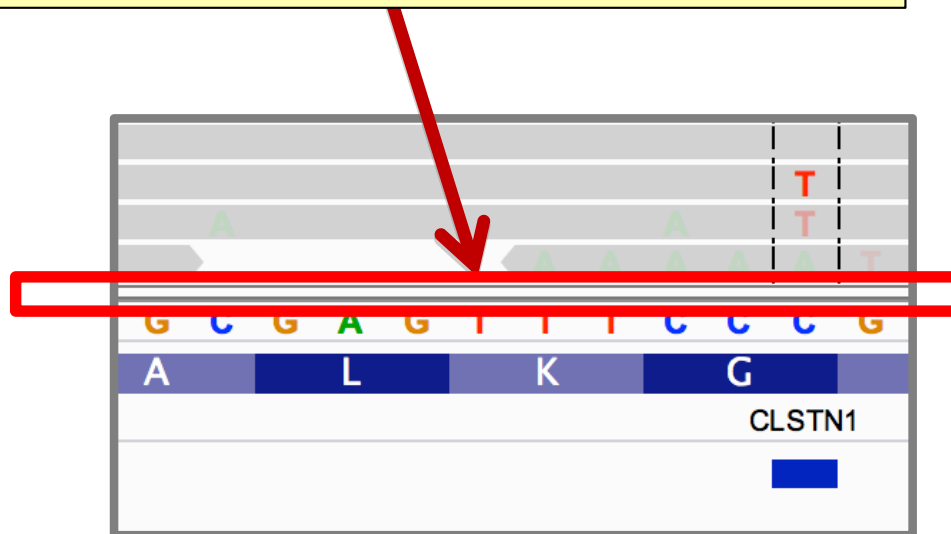


Viewing SNPs

Hands-on exercise

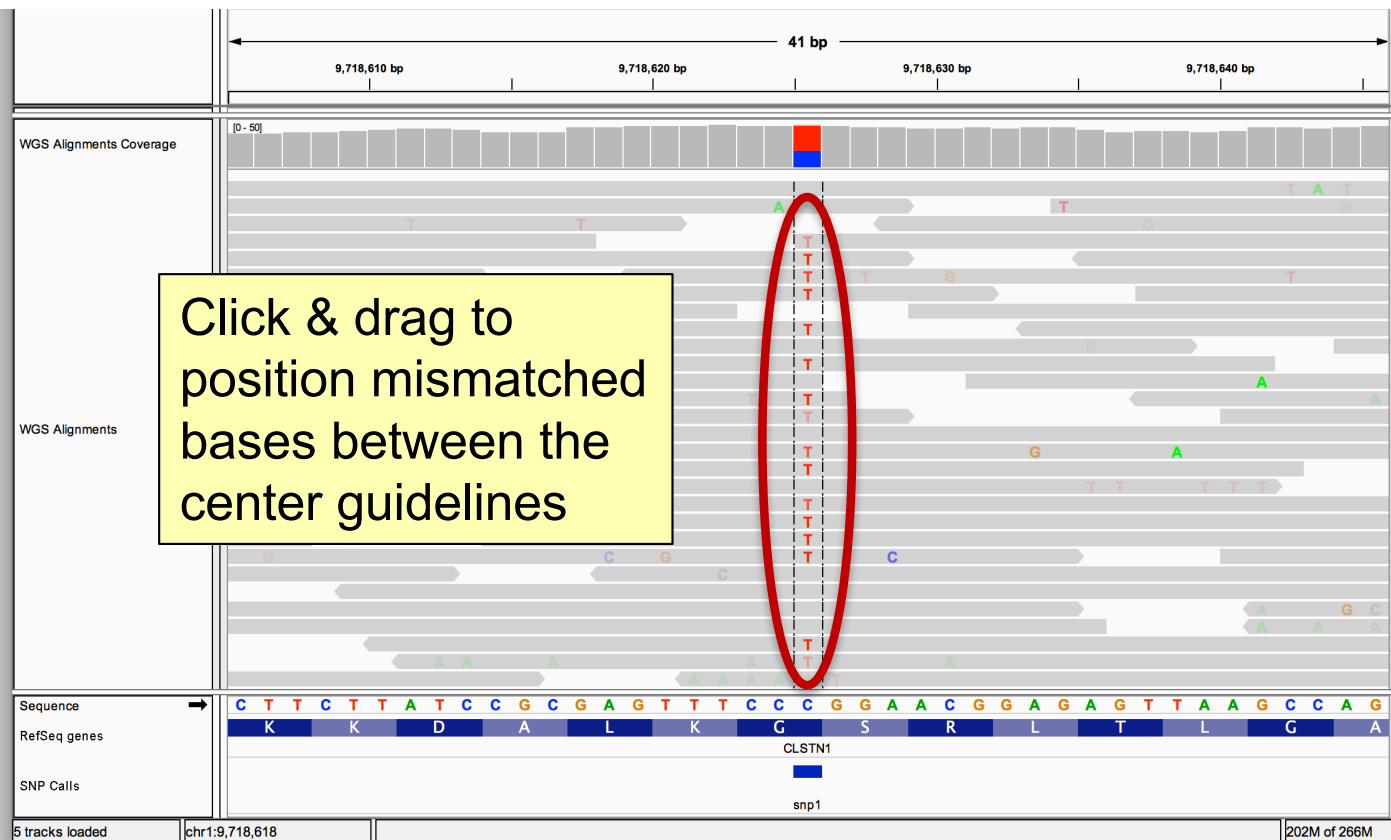
To modify the info popup behavior:
Click yellow balloon icon in toolbar

For a larger data panel:
Click & drag the window divider



Viewing SNPs

Hands-on exercise



Viewing SNPs

Hands-on exercise

The screenshot displays a genomic browser interface with several tracks. The 'WGS Alignments' track is highlighted, and a context menu is open over it. The menu options are:

- Rename Track...
- Copy read details to clipboard
- Group alignments by
- Sort alignments by** (highlighted with a red arrow)
- Color alignments by
- ✓ Shade base by quality
- ✓ Show mismatched bases
- Show all bases

The 'Sort alignments by' sub-menu is open, showing the following options:

- start location
- read strand
- first-of-pair strand
- base** (highlighted with a red arrow)
- mapping quality
- sample
- read group
- insert size
- chromosome of mate
- tag

Below the menu, a yellow box contains the text: **Right-click on alignments and select Sort alignments by > base**.

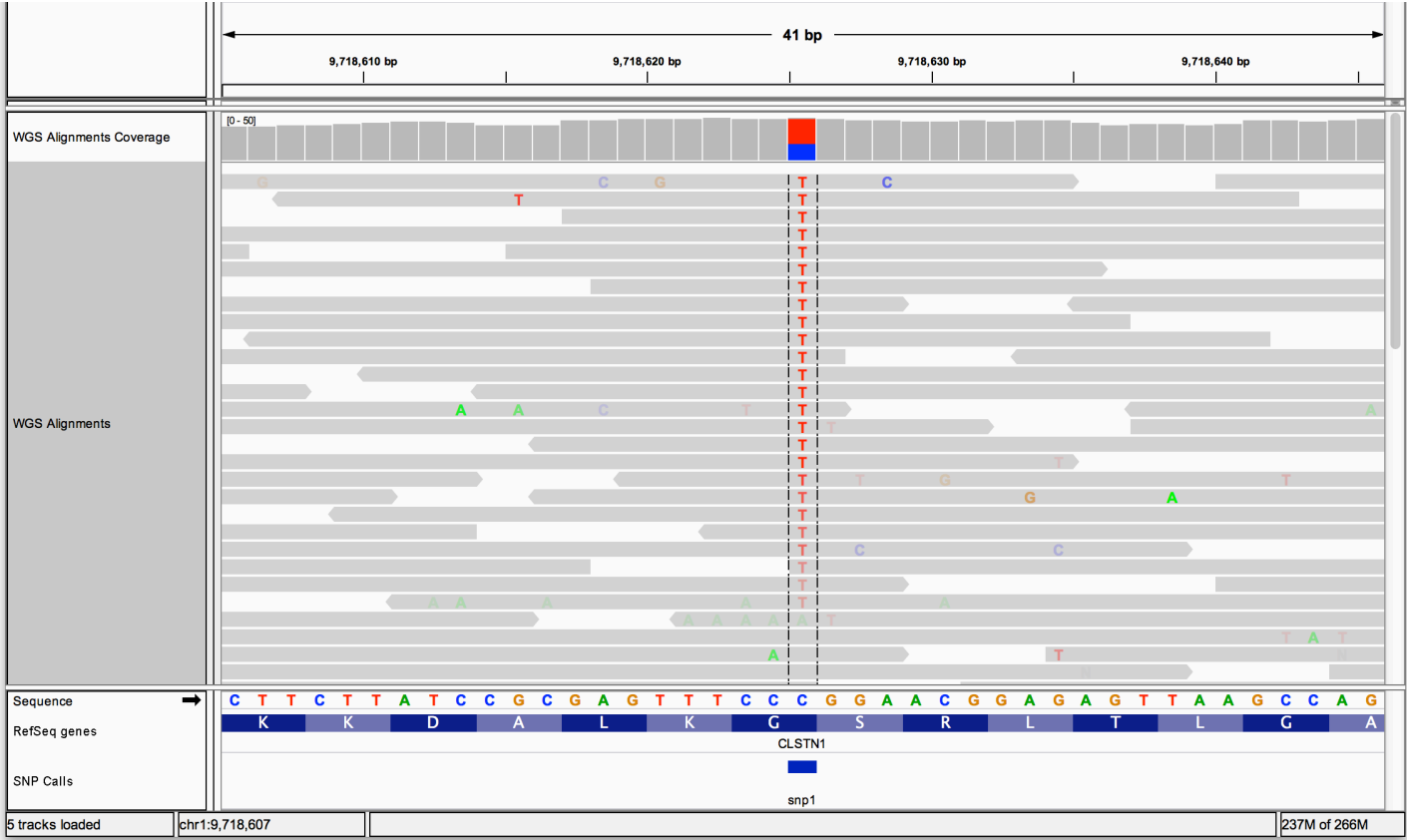
Another yellow box contains the text: **On Mac: Right-click = Control-click**.

The background shows the 'WGS Alignments Coverage' track, the 'Sequence' track with the reference sequence (C T T C T T A T C C G G A G T T T C C C G G A A C G G A G A G T T A A G C C A G), and the 'RefSeq genes' track showing the gene 'CLSTN1'. The 'SNP Calls' track shows a single SNP call 'snp1'.

At the bottom, the status bar indicates '5 tracks loaded', 'chr1:9,718,618', and '202M of 266M'.

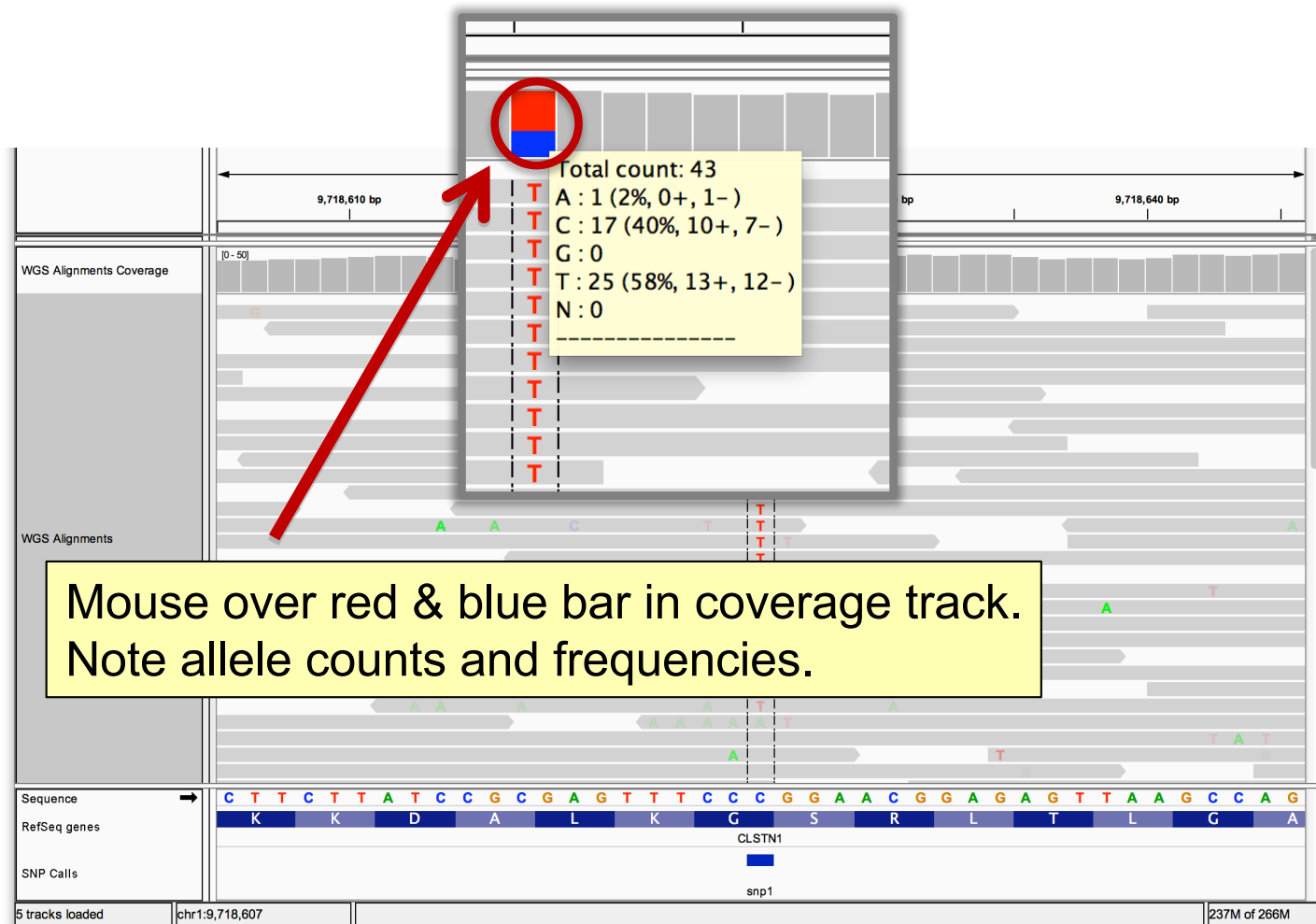
Viewing SNPs

Hands-on exercise



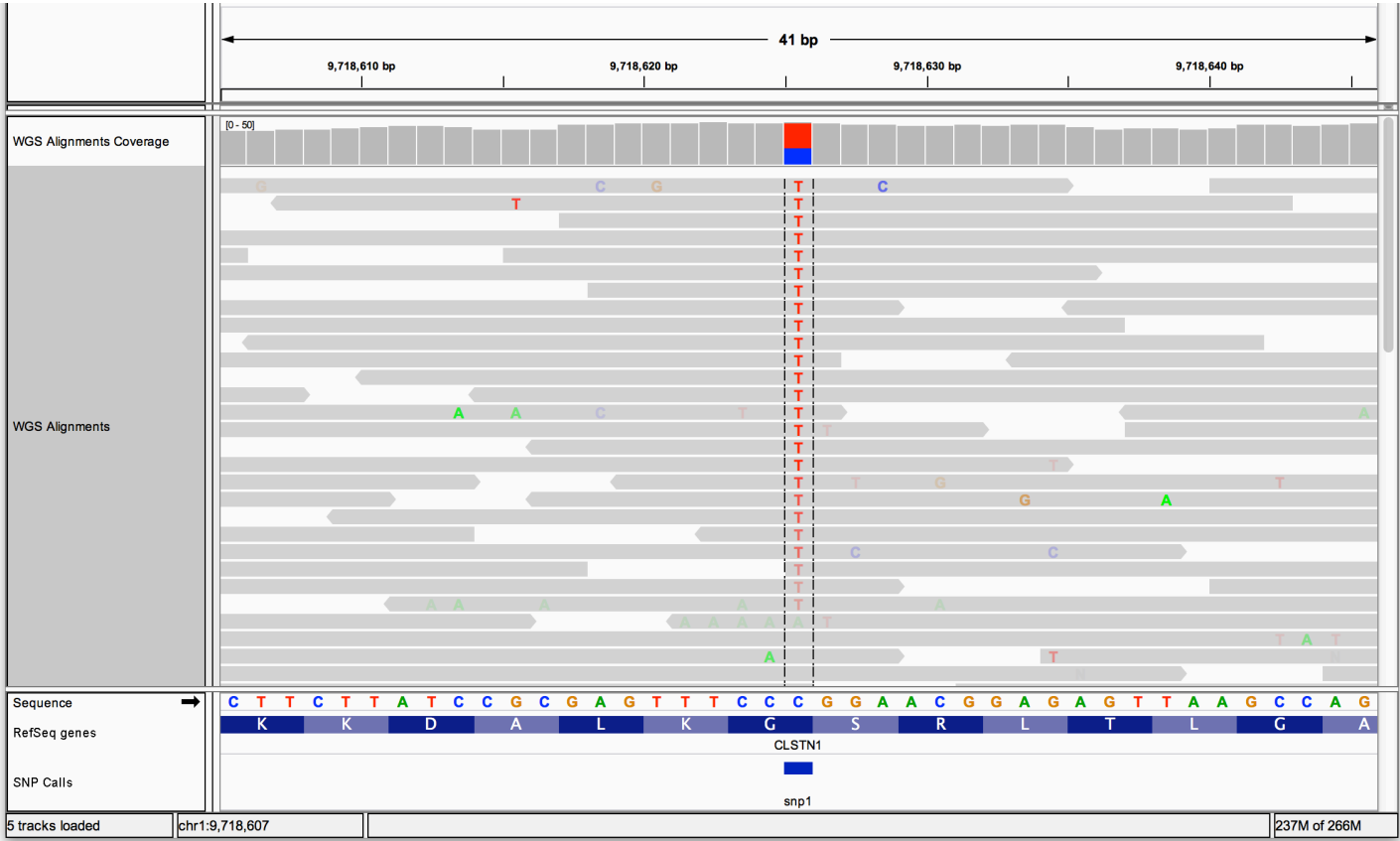
Viewing SNPs

Hands-on exercise



Viewing SNPs

Hands-on exercise

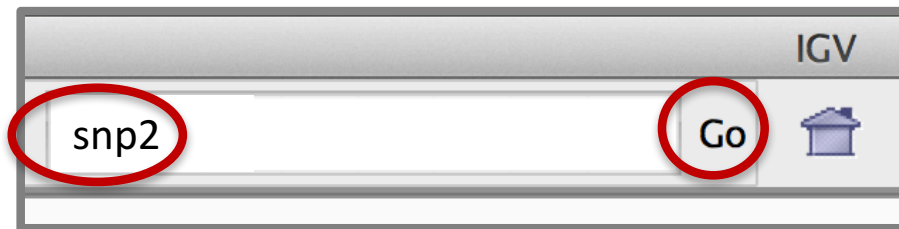


Viewing SNPs

Hands-on exercise

Navigate to next putative SNP locus

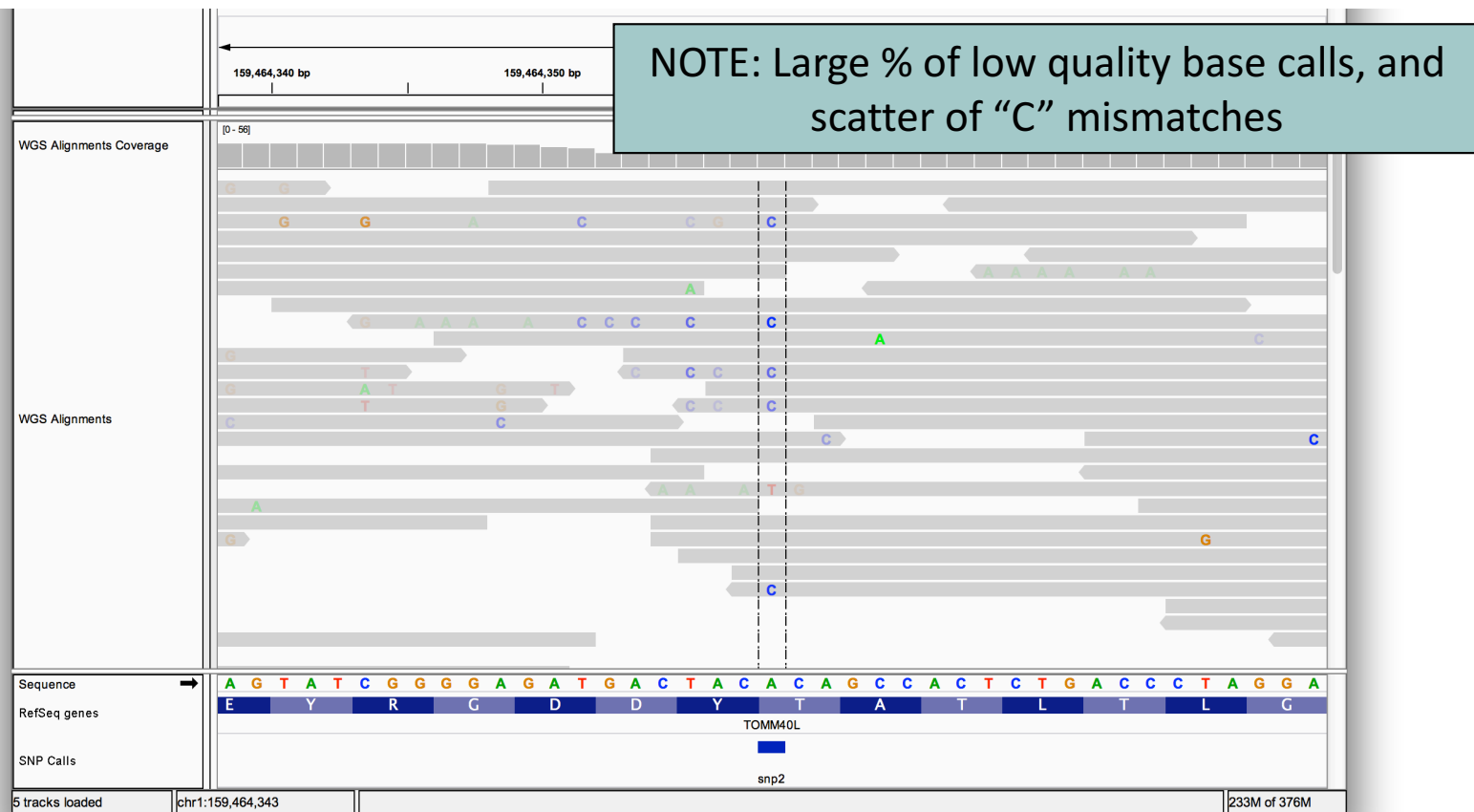
Type “snp2” in the **Search Box** and click **Go**



A screenshot of the IGV (Integrated Genomics Viewer) search interface. The interface is a horizontal bar with a light gray background. On the right side, the text "IGV" is displayed. Below this, there is a search bar containing the text "snp2". To the right of the search bar is a button labeled "Go". Both the search bar and the "Go" button are circled in red. To the right of the "Go" button is a small blue house icon.

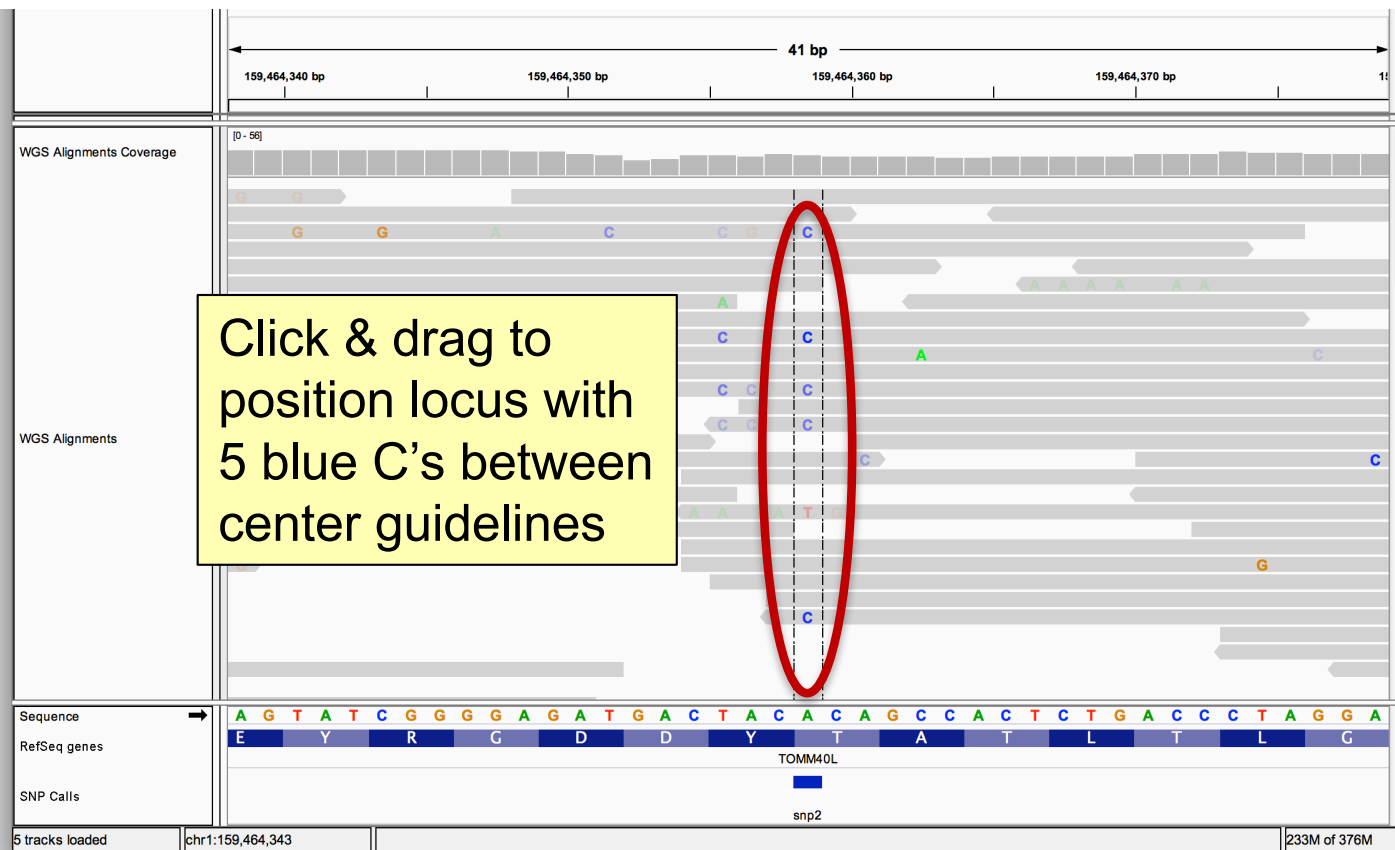
Viewing SNPs

Hands-on exercise



Viewing SNPs

Hands-on exercise



Viewing SNPs

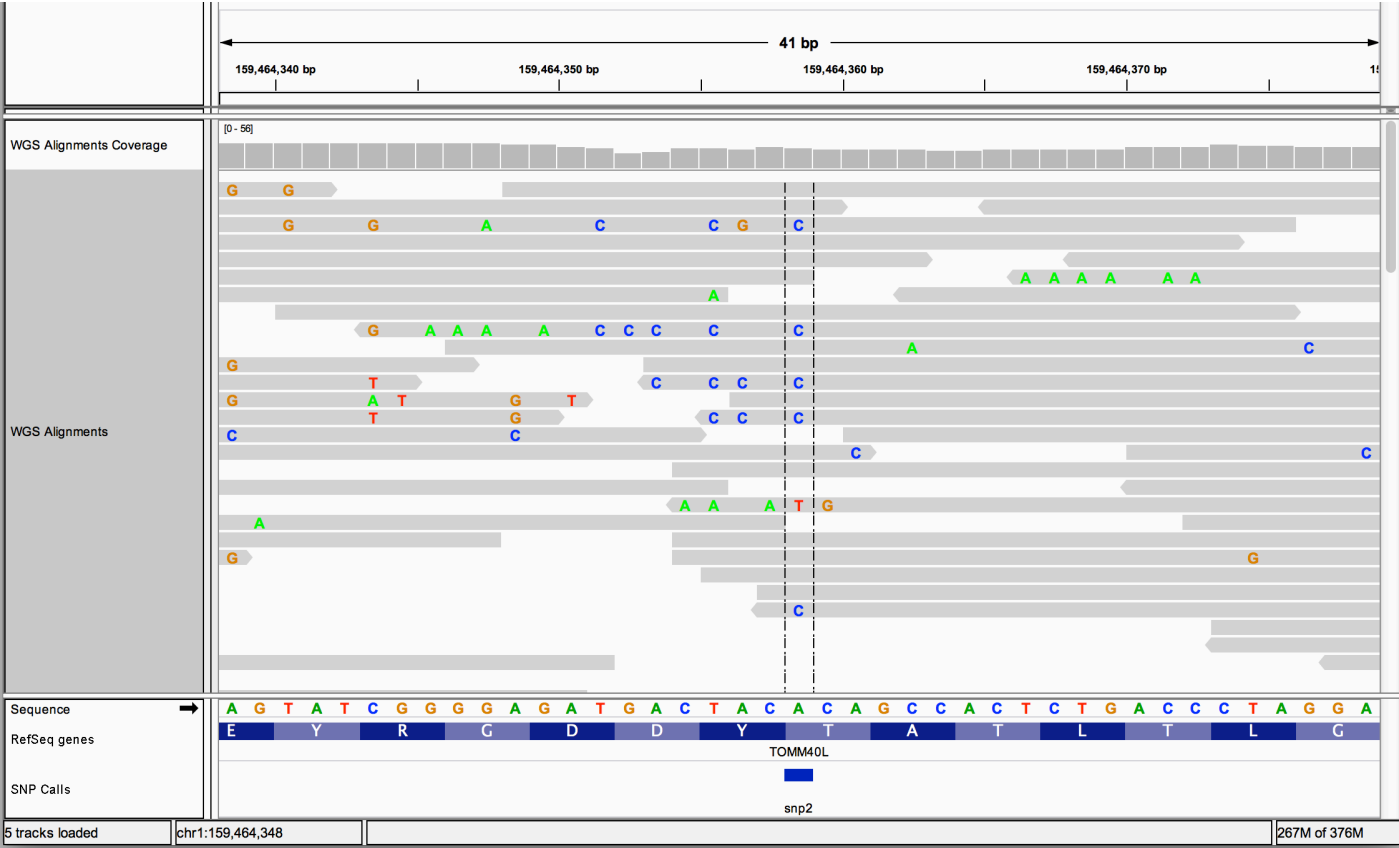
Hands-on exercise

The screenshot displays a genomic browser interface. At the top, a scale bar indicates a 41 bp region from 159,464,340 bp to 159,464,360 bp. Below this, the 'WGS Alignments Coverage' track shows a histogram of read coverage. The 'WGS Alignments' track displays individual sequencing reads with colored bases (A, G, C, T) and arrows indicating their orientation. A red arrow points from a read in this track to a context menu. The menu, titled 'WGS Alignments', contains the following options: 'Rename Track...', 'Copy read details to clipboard', 'Group alignments by', 'Sort alignments by', 'Color alignments by', '✓ Shade base by quality' (highlighted in blue), '✓ Show mismatched bases', 'Show all bases', 'View as pairs', 'Go to mate', 'View mate region in split screen', 'Set insert size options ...', and 'Re-pack alignments'. At the bottom of the browser, the 'Sequence' track shows the reference sequence: A G T A T C G G G G A G A T G A C T A C A C A G C C A C T C T G A C C C T A G G A. Below this, the 'RefSeq genes' track shows the gene 'TOMM40L' and the 'SNP Calls' track shows a variant 'snp2'. The status bar at the bottom indicates '5 tracks loaded', 'chr1:159,464,343', and '233M of 376M'.

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

Viewing SNPs

Hands-on exercise



Viewing SNPs

Hands-on exercise

The screenshot displays a genomic browser interface with several tracks. The top track is 'WGS Alignments Coverage', showing a bar chart of coverage. Below it is the 'WGS Alignments' track, which displays individual sequencing reads as horizontal bars. A red arrow points from a read in this track to a context menu. The context menu has a title 'WGS Alignments' and contains the following options: 'Rename Track...', 'Copy read details to clipboard', 'Group alignments by' (with a submenu arrow), 'Sort alignments by' (highlighted in blue), 'Color alignments by' (with a submenu arrow), 'Shade base by quality' (checked), 'Show mismatched bases' (checked), 'Show all bases', 'View as pairs', 'Go to mate', and 'View mate region in split screen'. The 'Sort alignments by' submenu is open, showing options: 'start location', 'read strand' (highlighted in blue), 'first-of-pair strand', 'base', 'mapping quality', 'sample', 'read group', 'insert size', 'chromosome of mate', and 'tag'. A yellow text box with a black border is overlaid on the image, containing the text: 'Right-click on alignments and select **Sort alignments by > read strand**'. At the bottom of the browser, there are tracks for 'Sequence', 'RefSeq genes', and 'SNP Calls'. The 'Sequence' track shows the reference sequence: A G T A T C G G G G A G A T G A C T A C A C A G C C A C T C T G A C C C T A G G A. The 'RefSeq genes' track shows the gene 'TOMM40L' with exons represented by blue boxes and introns by lines. The 'SNP Calls' track shows a single SNP call 'snp2' at position chr1:159,464,343. The bottom status bar indicates '5 tracks loaded', 'chr1:159,464,343', and '233M of 376M'.

WGS Alignments

- Rename Track...
- Copy read details to clipboard
- Group alignments by
- Sort alignments by**
- Color alignments by
- ✓ Shade base by quality
- ✓ Show mismatched bases
- Show all bases
- View as pairs
- Go to mate
- View mate region in split screen

start location
read strand
first-of-pair strand
base
mapping quality
sample
read group
insert size
chromosome of mate
tag

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

Sequence
RefSeq genes
SNP Calls

5 tracks loaded | chr1:159,464,343 | 233M of 376M

Viewing SNPs

Hands-on exercise

The screenshot displays a genomic browser interface with several tracks. The 'WGS Alignments Coverage' track shows a bar chart of coverage. The 'WGS Alignments' track shows individual sequencing reads. A context menu is open over the alignments, with the 'Color alignments by' option selected. A secondary menu is open for 'Color alignments by', with 'read strand' selected. A yellow callout box contains the text: 'Right-click on alignments and select Color alignments by > read strand'. The bottom tracks show the reference sequence (A G T A T C G G G A G A T G A C T A C A C A G C C A C T C T G A C C C T A G G A), RefSeq genes (E Y R G D D Y T A T L T L G), and SNP Calls (snp2). The bottom status bar indicates '5 tracks loaded', 'chr1:159,464,343', and '233M of 376M'.

WGS Alignments

- Rename Track...
- Copy read details to clipboard
- Group alignments by
- Sort alignments by
- Color alignments by
 - no color
 - insert size
 - pair orientation
 - ✓ insert size and pair orientation
 - read strand
 - first-of-pair strand
 - read group
 - sample
 - tag
 - bisulfite mode
- ✓ Shade base by quality
- ✓ Show mismatched bases
- Show all bases
- View as pairs
- Go to mate
- View mate region in split screen
- Set insert size options...

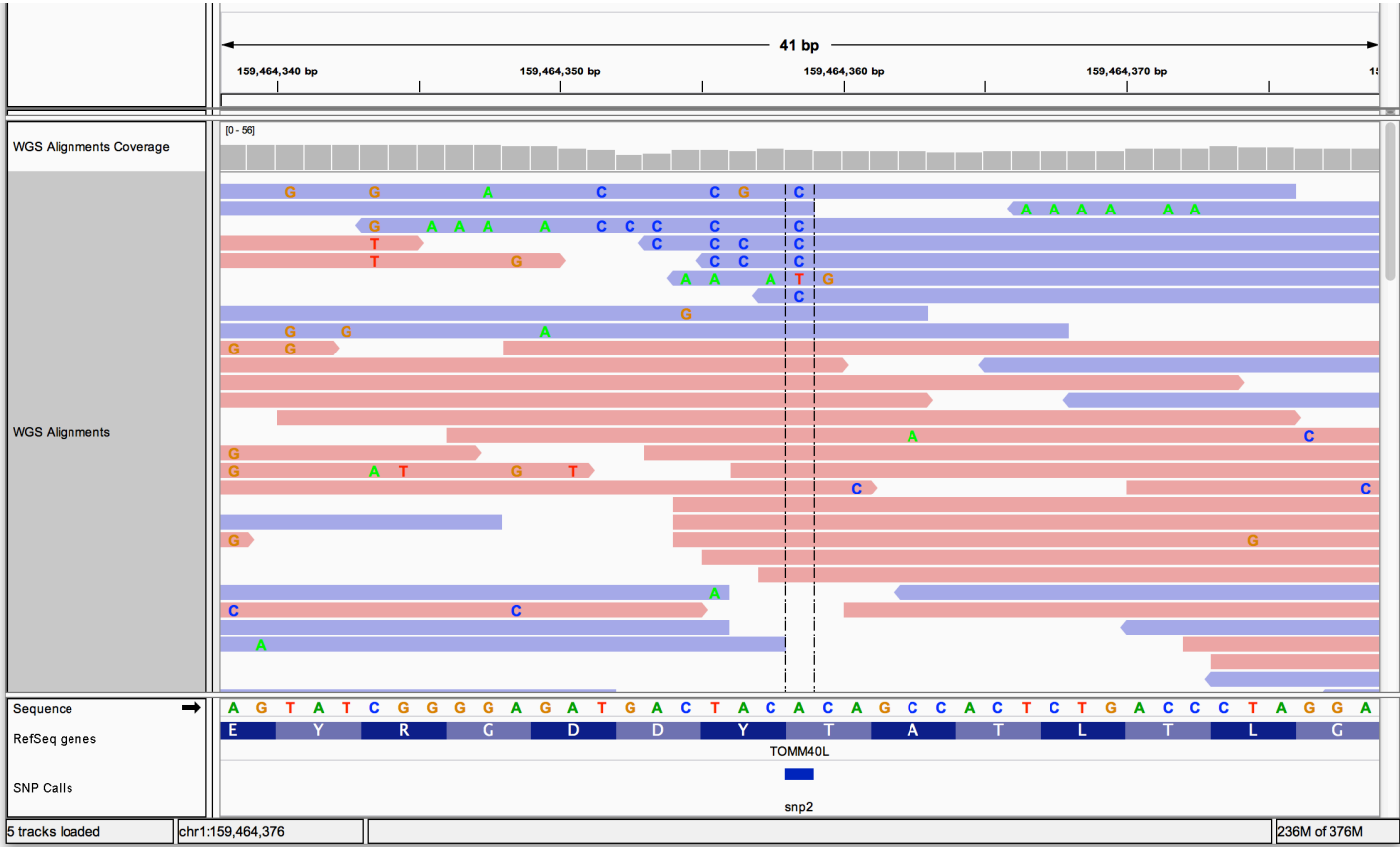
Right-click on alignments and select **Color alignments by > read strand**

Sequence
RefSeq genes
SNP Calls

5 tracks loaded | chr1:159,464,343 | 233M of 376M

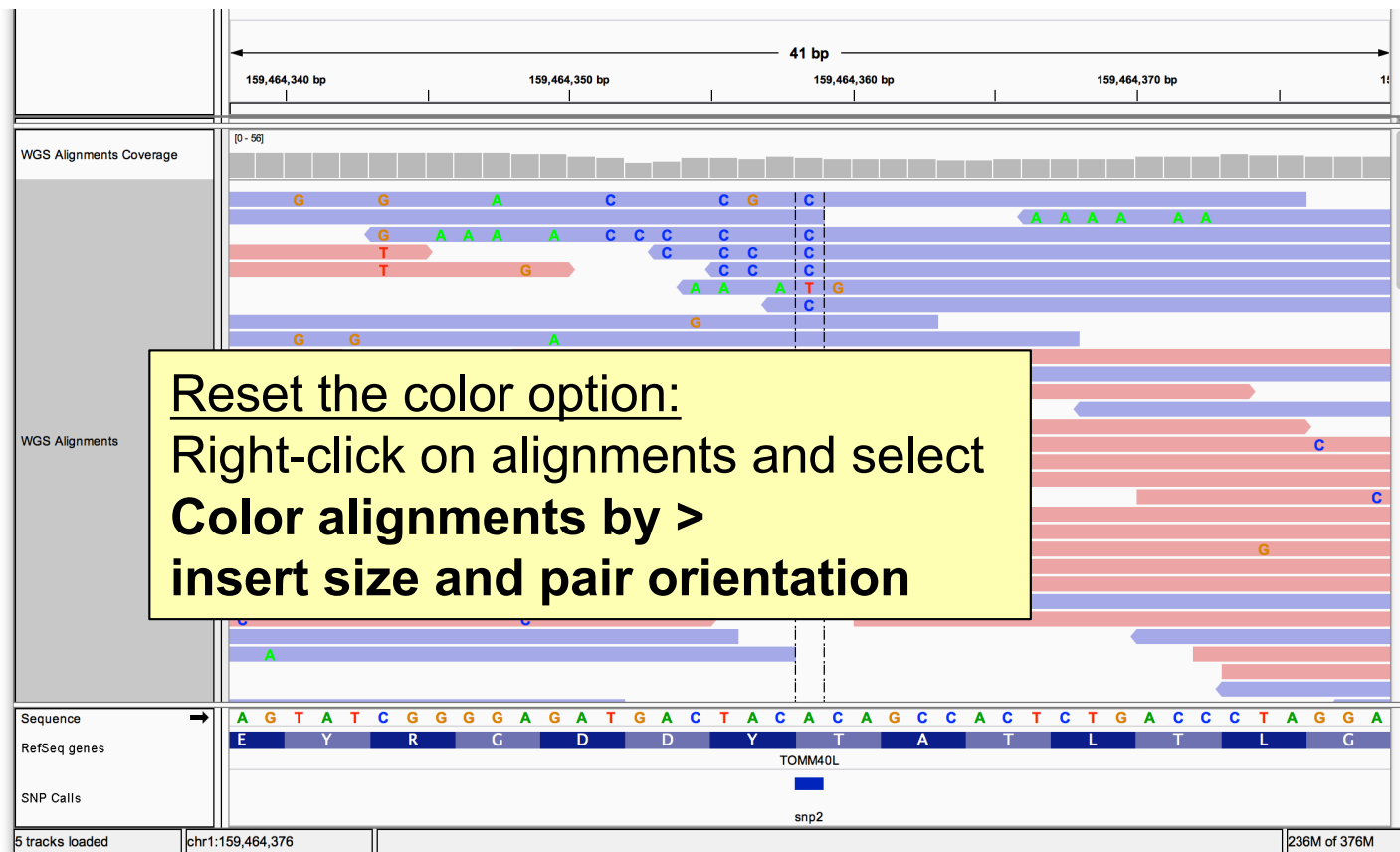
Viewing SNPs

Hands-on exercise



Viewing SNPs

Hands-on exercise



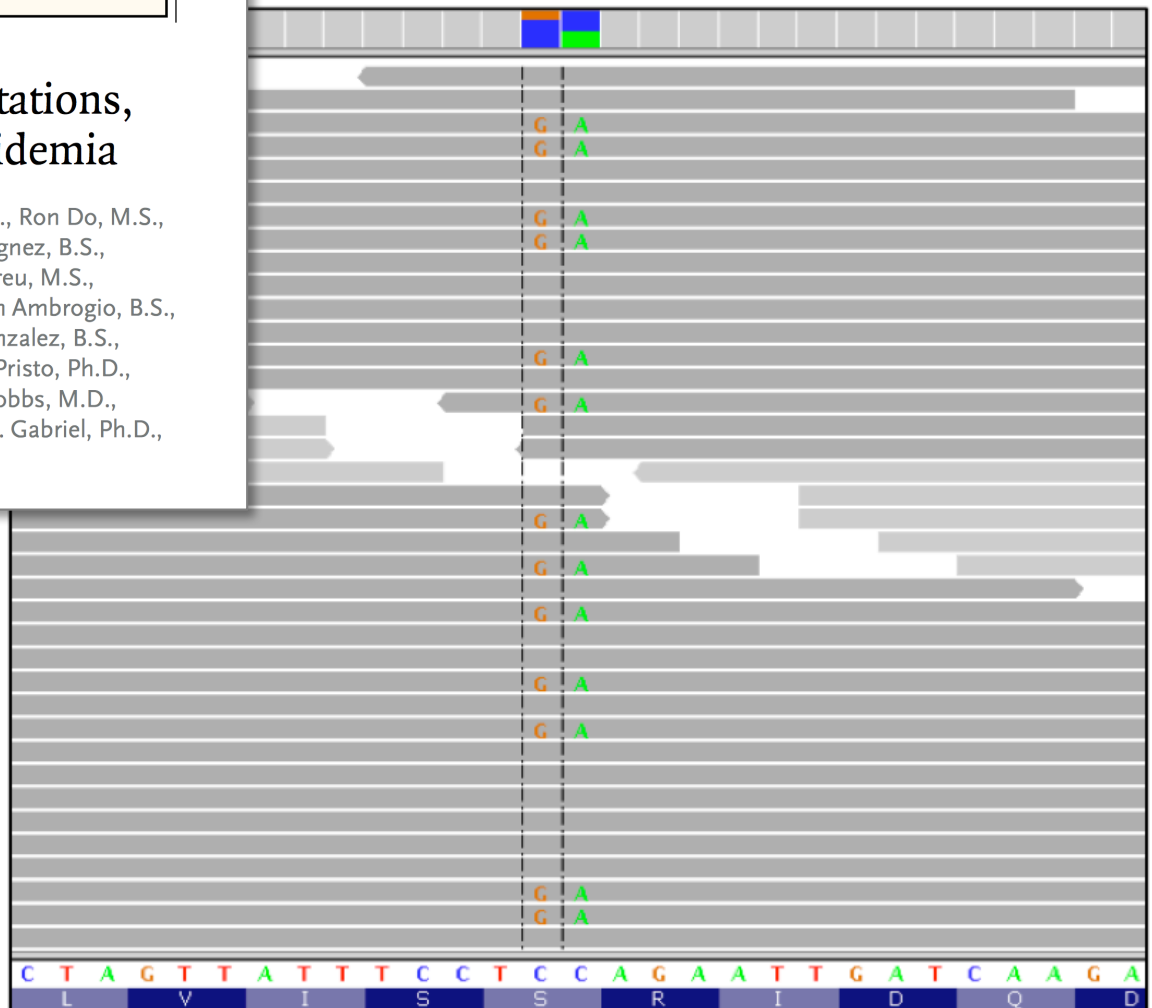
Viewing SNPs

The NEW ENGLAND JOURNAL of MEDICINE

BRIEF REPORT

Exome Sequencing, *ANGPTL3* Mutations, and Familial Combined Hypolipidemia

Kiran Musunuru, M.D., Ph.D., M.P.H., James P. Pirruccello, B.S., Ron Do, M.S.,
Gina M. Peloso, M.S., Candace Guiducci, B.S., Carrie Sougnez, B.S.,
Kiran V. Garimella, M.S., Sheila Fisher, M.L.A., Justin Abreu, M.S.,
Andrew J. Barry, B.S., Tim Fennell, B.S., Eric Banks, Ph.D., Lauren Ambrogio, B.S.,
Kristian Cibulskis, B.S., Andrew Kernysky, Ph.D., Elena Gonzalez, B.S.,
Nicholas Rudzicz, M.S., James C. Engert, Ph.D., Mark A. DePristo, Ph.D.,
Mark J. Daly, Ph.D., Jonathan C. Cohen, Ph.D., Helen H. Hobbs, M.D.,
David Altshuler, M.D., Ph.D., Gustav Schonfeld, M.D., Stacey B. Gabriel, Ph.D.,
Pin Yue, Ph.D., and Sekar Kathiresan, M.D.



Viewing SNPs



BED file used to define “bookmarks” in SNP exercise

snp_calls.bed

```
track name="SNP Calls"
```

1	9718624	9718625	snp1
---	---------	---------	------

1	159464357	159464358	snp2
---	-----------	-----------	------

1	63650225	63650226	snp3
---	----------	----------	------

igvtools

igvtools

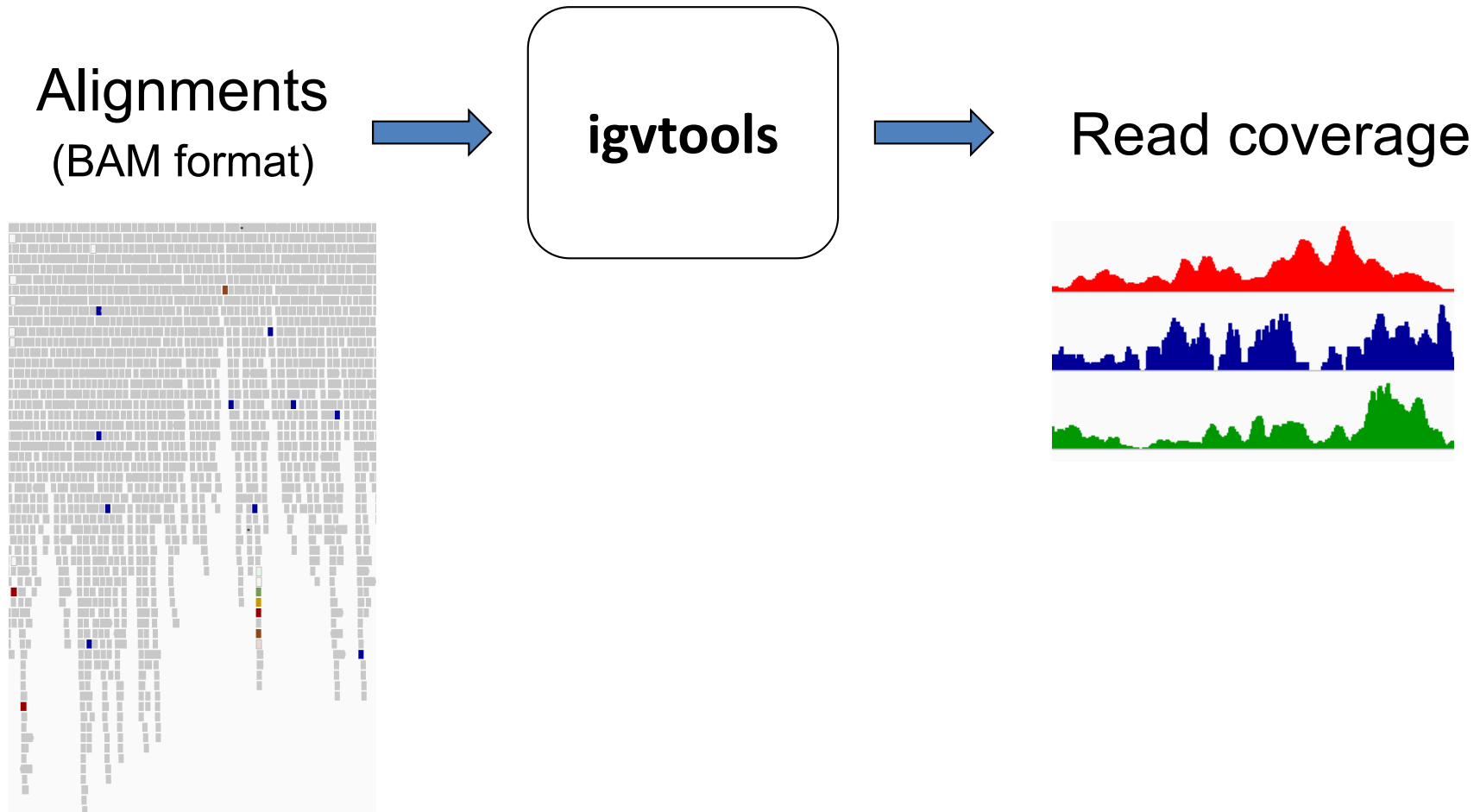
A set of utilities for preparing large files for efficient display.

count	<ul style="list-style-type: none">– Computes alignment coverage from BAM files– Produces TDF or WIG files
toTDF	<ul style="list-style-type: none">– Converts sorted data file to binary tiled data (TDF)– Supported file formats: WIG, bedGraph
sort	<ul style="list-style-type: none">– Sorts file by genomic start position.– Supported file formats: BED, GFF, GTF, PSL, SAM, BAM, VCF
index	<ul style="list-style-type: none">– Creates index for large genomic annotation files and alignments.– Supported file formats: BED, GFF, GTF, PSL, SAM, BAM, VCF

igvtools: count

count

transforms alignment files to
read coverage files in TDF or WIG format.



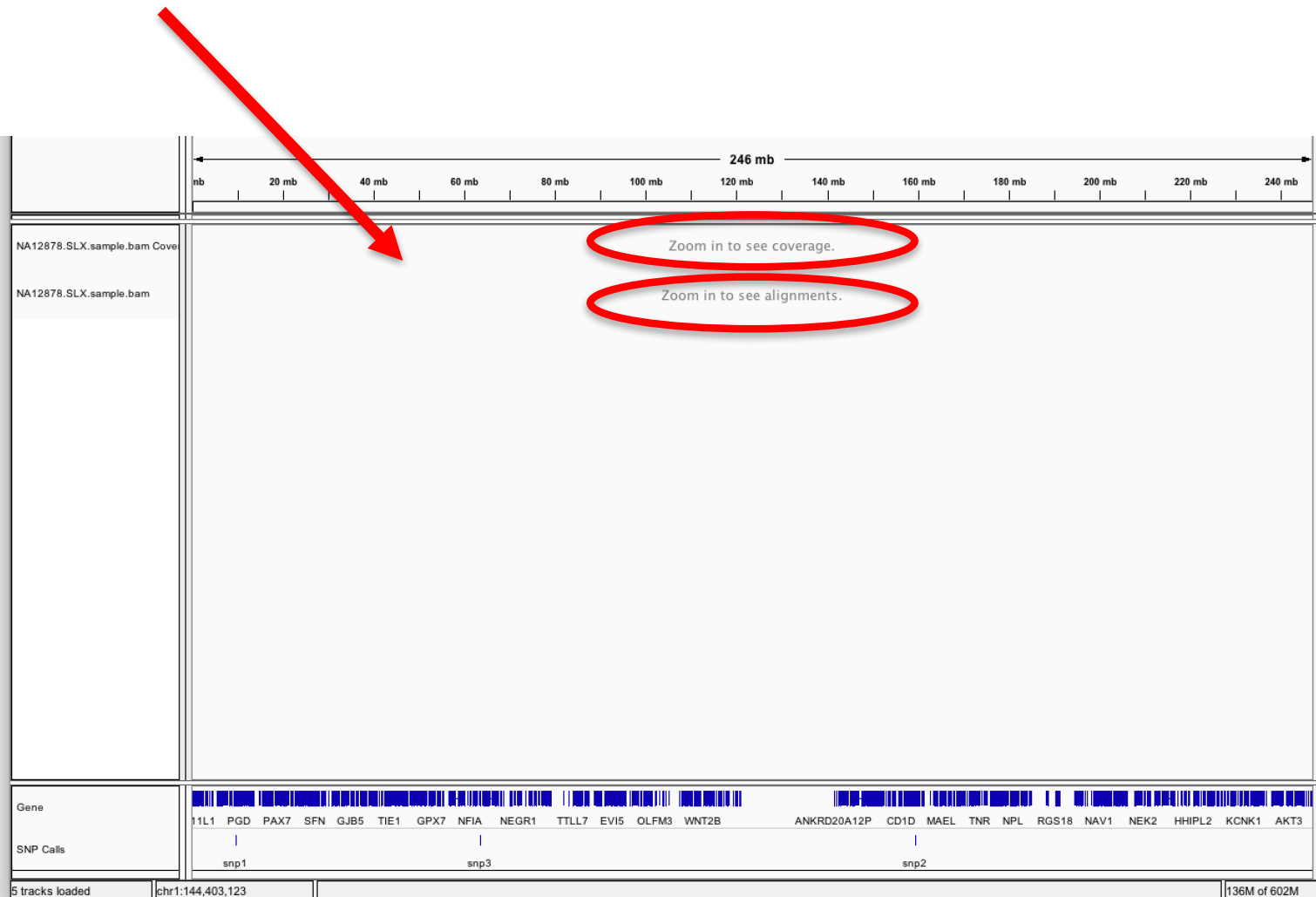
Coverage track

Coverage track visible before alignments



Coverage track

In our hands-on exercise:
Coverage track was **not** visible before alignments



igvtools

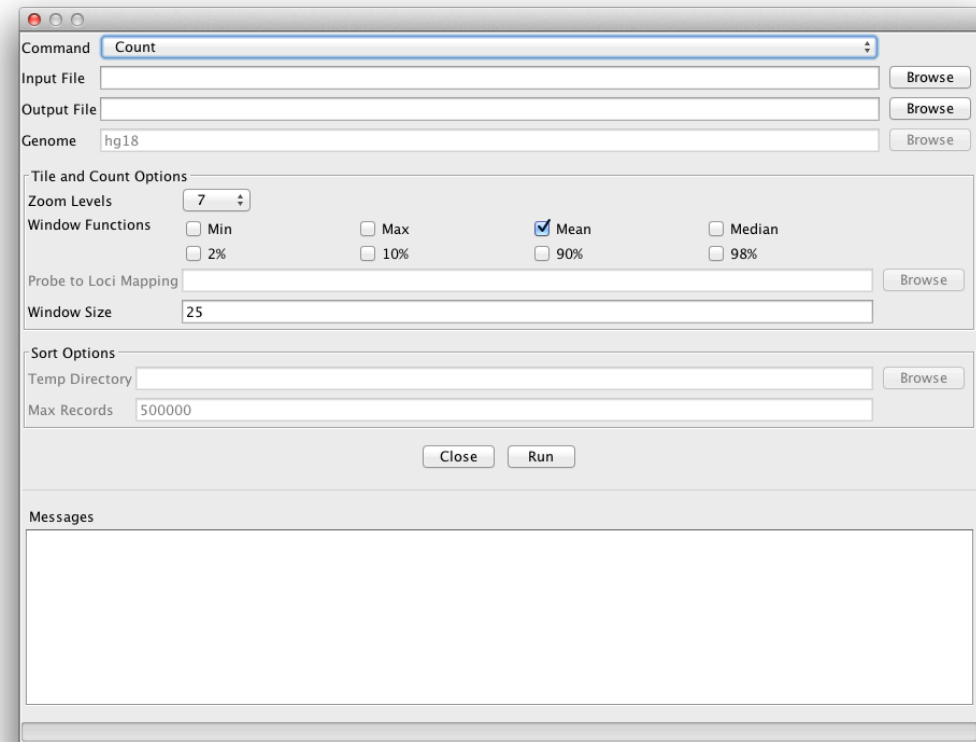
(A) Launch from the IGV menu

Tools > Run igvtools...

or

(B) Run from the command line

```
[GM31D-693:~] jrobinso%  
[GM31D-693:~] jrobinso% igvtools count NA12878.bam NA12878.bam.tdf hg18
```



Pre-compute coverage track

Hands-on exercise

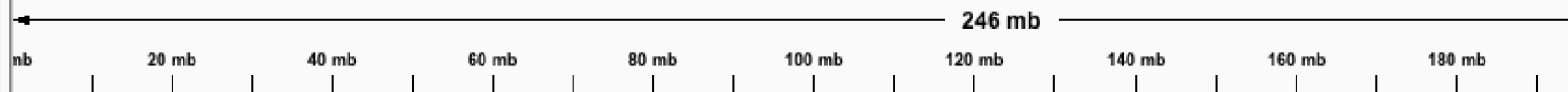
- Use ***igvtools*** to compute the coverage for a .bam file
- Associate the new file with the coverage track

Follow along

Coverage track

Hands-on exercise

Click on the left-most tick on the zoom ruler to zoom all the way out



NA12878.SLX.sample.bam Cover

Zoom in to see coverage.

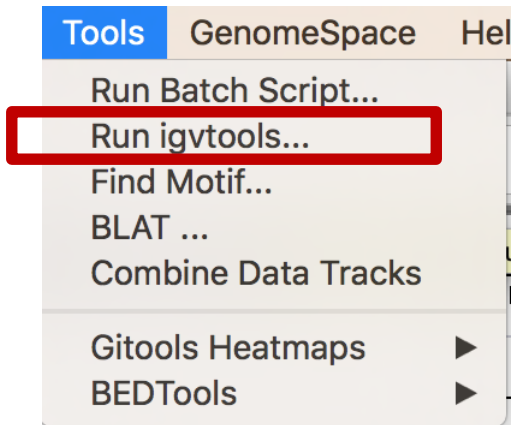
NA12878.SLX.sample.bam

Zoom in to see alignments.

Pre-compute coverage track

Hands-on exercise

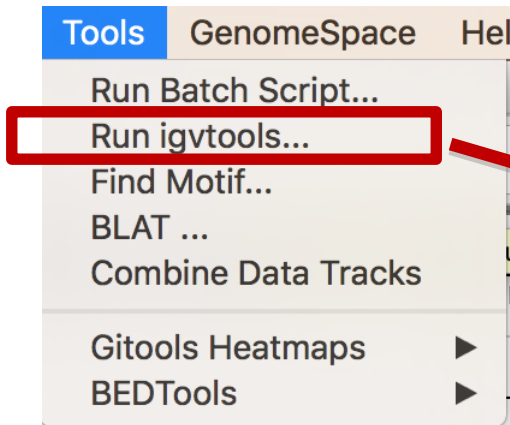
Click ***Tools > Run igvtools***



Pre-compute coverage track

Hands-on exercise

Click **Tools > Run igvtools**



Select the **Count** command from the dropdown menu

A screenshot of the 'igvtools' dialog box. The 'Command' dropdown menu is open, and 'Count' is selected and circled with a red oval. Below the dropdown, there are fields for 'Input File', 'Output File', and 'Genome'. The 'Genome' field contains the path '/Users/helga/Documents/IGVMarch2017/Exercises/genome/chr1.fasta'. There are 'Browse' buttons next to each of these fields. Below these fields, there are sections for 'TDF and Count options', 'Sort Options', and 'Messages'. The 'TDF and Count options' section includes 'Zoom Levels' (set to 7), 'Window Functions' (with checkboxes for Min, Max, Mean, Median, 2%, 10%, 90%, 98%), 'Probe to Loci Mapping' (with a 'Browse' button), 'Window Size' (set to 25), 'Extension Factor', and a checkbox for 'Count as Pairs'. The 'Sort Options' section includes 'Temp Directory' (with a 'Browse' button) and 'Max Records' (set to 500000). At the bottom of the dialog are 'Close' and 'Run' buttons. The 'Messages' section at the very bottom is empty.

Pre-compute coverage track

Hands-on exercise

1. Input file: Browse to the *Data / snps_morning* folder and select the **.bam** file

Command

2. Output file: Defaults to same folder, and same name + .tdf suffix

Zoom Levels

Window Functions ☐ Min ☐ Max ☒ Mean ☐ Median
☐ 2% ☐ 10% ☐ 90% ☐ 98%

Probe to Loci Mapping

Window Size

Extension Factor

☐ Count as Pairs

Sort Options

Temp Directory

Max Records

3. Other values: Keep defaults

4. Click *Run*

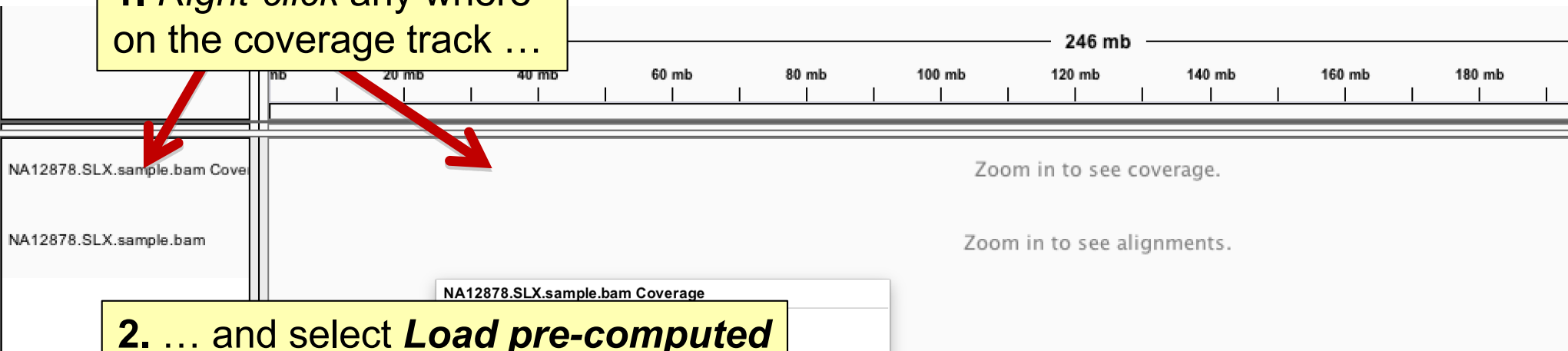
Messages
100.0%
Done

5. Close window after you see “Done” in the **Messages area**

Coverage track

Hands-on exercise

1. Right-click anywhere on the coverage track ...



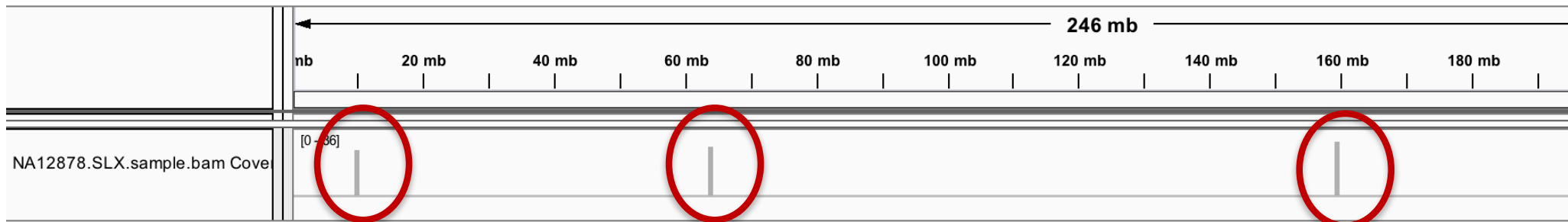
2. ... and select **Load pre-computed coverage data** from the popup menu

- Set Data Range...
- Log scale
- ☒ Autoscale
- ☒ Show Data Range
- Set allele frequency threshold...
- Load pre-computed coverage data...**
- Copy Details to Clipboard
- ☒ Show Alignment Track
- Show Splice Junction Track
- Hide Track
- Save image...
- Export track names...

3. Browse to the **Data / snps_morning** folder and select the file **NA12878.SLX.sample.bam.tdf**

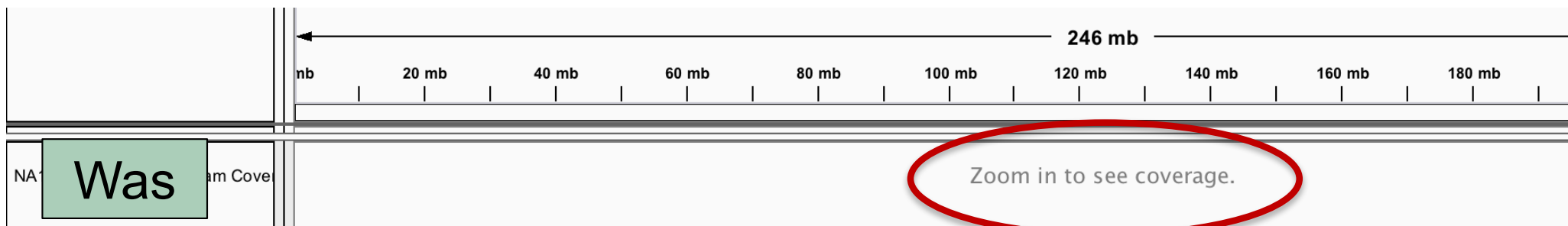
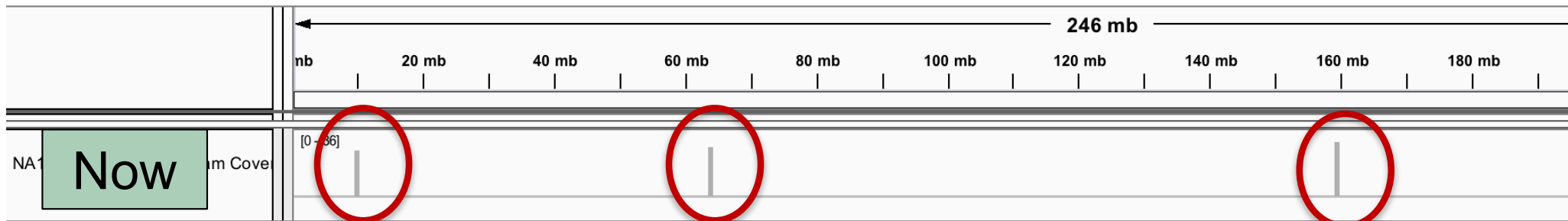
Coverage track

Hands-on exercise



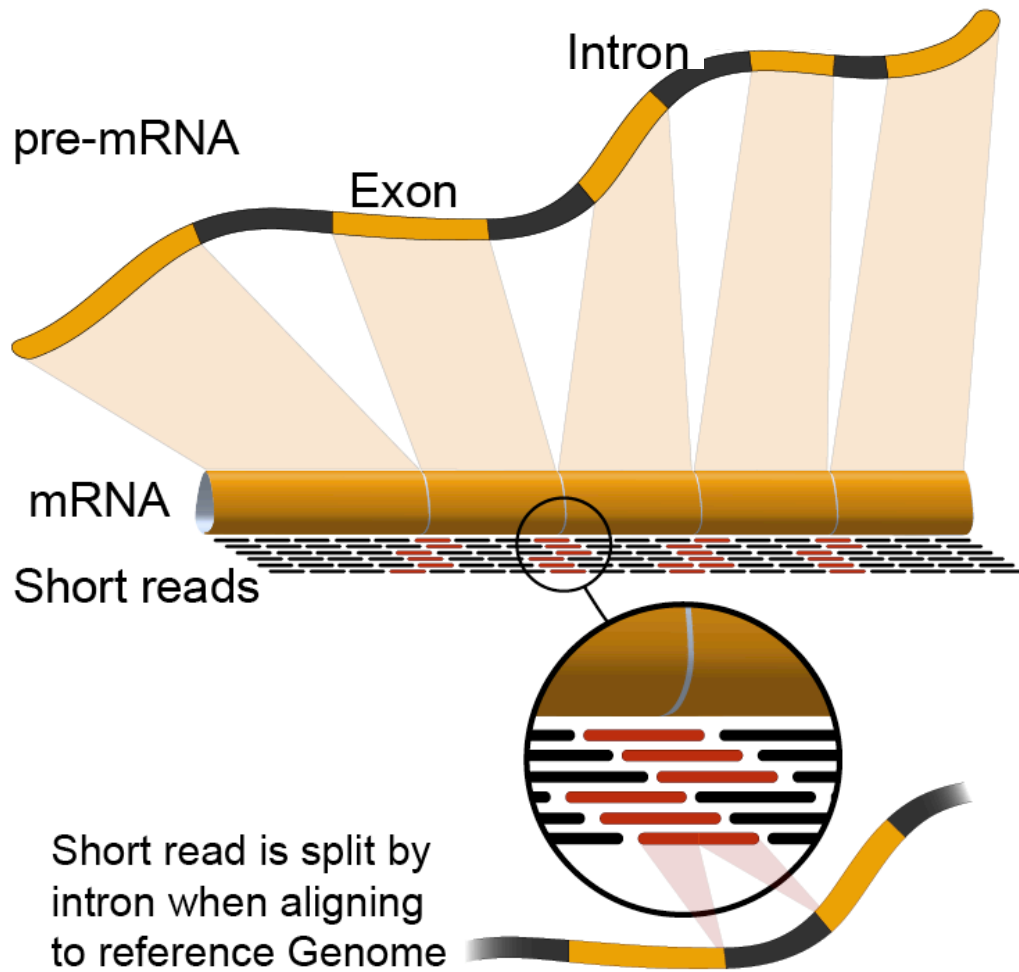
Coverage track

Hands-on exercise

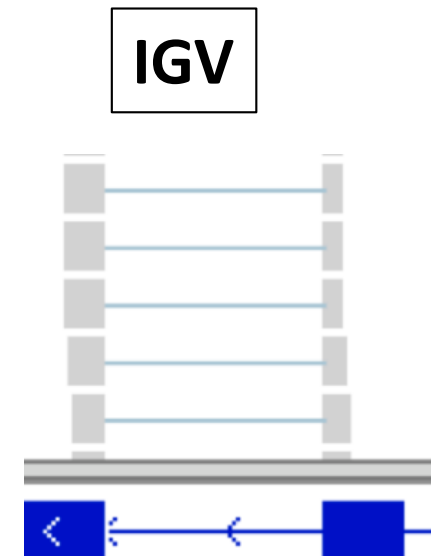
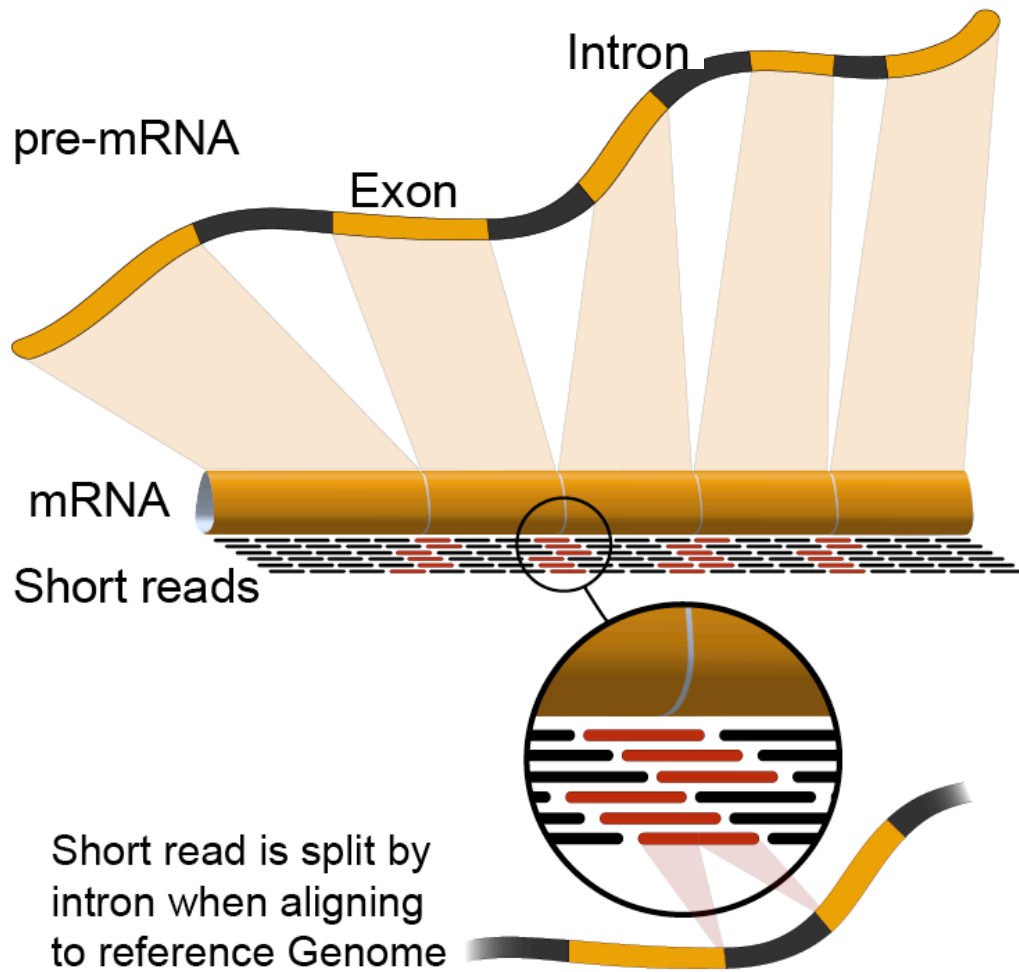


Viewing RNA-seq Data

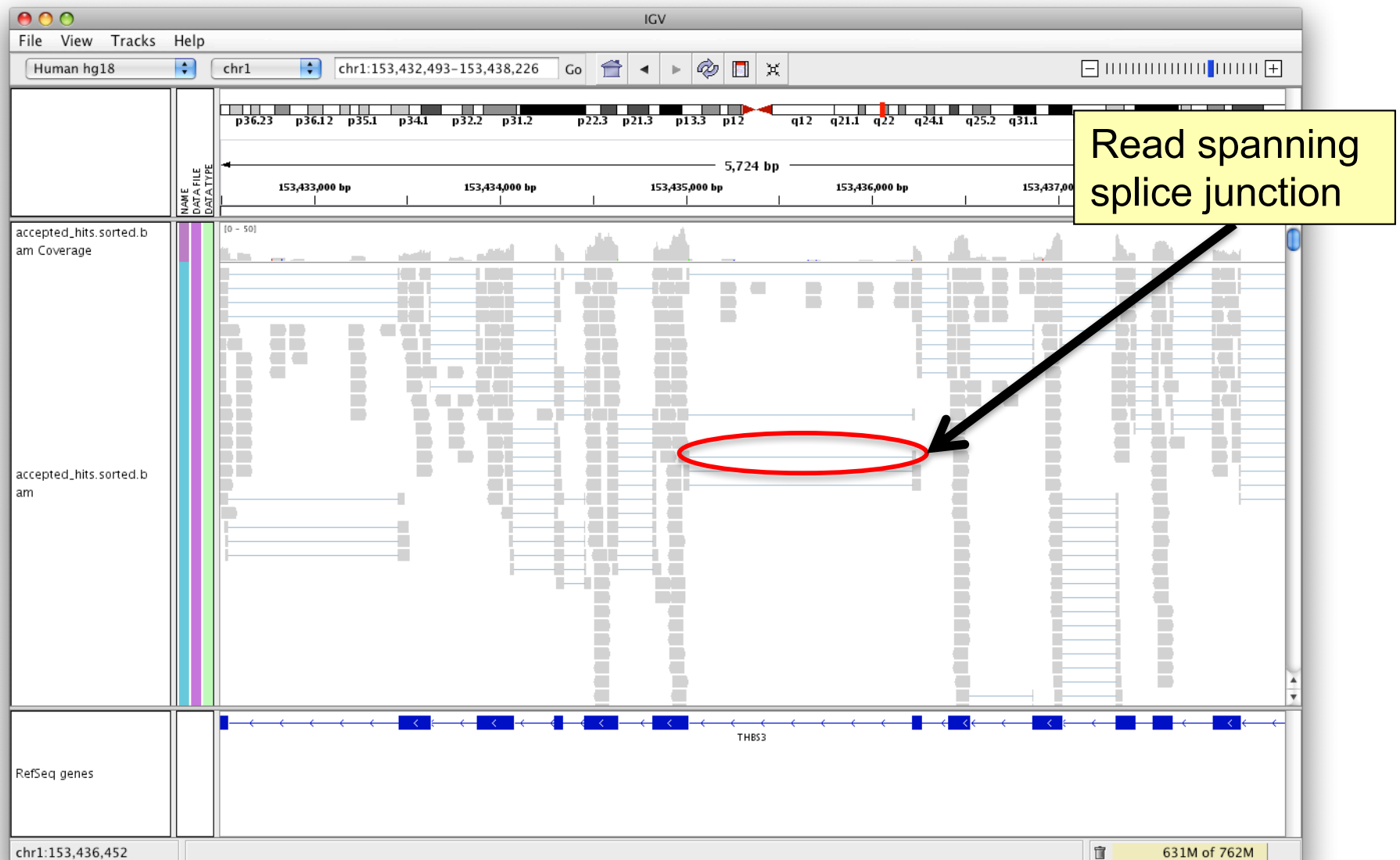
RNA-seq



RNA-seq

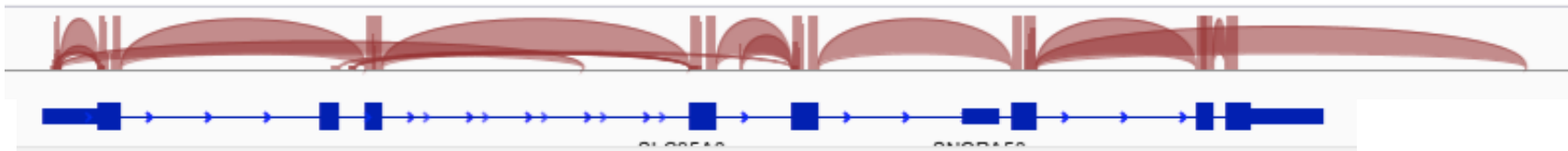


RNA-seq alignments



RNA-seq alignments

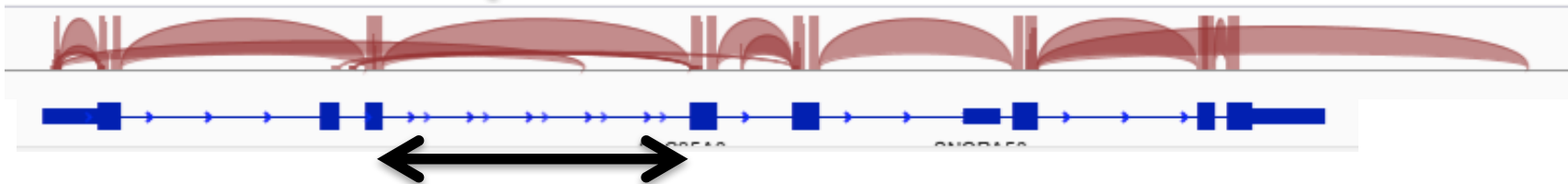
Splice junction track



RNA-seq alignments

Splice junction track

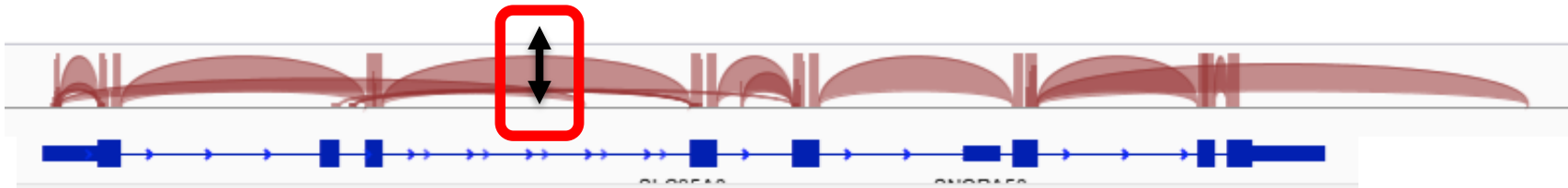
Arcs represent reads that span exon junctions



RNA-seq alignments

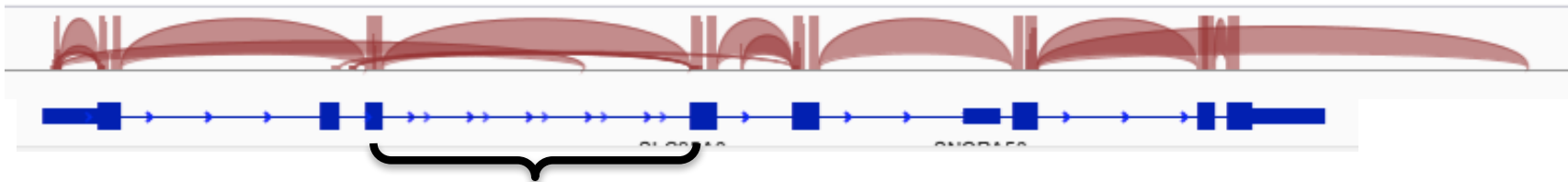
Splice junction track

Height is proportional
to number of reads



RNA-seq alignments

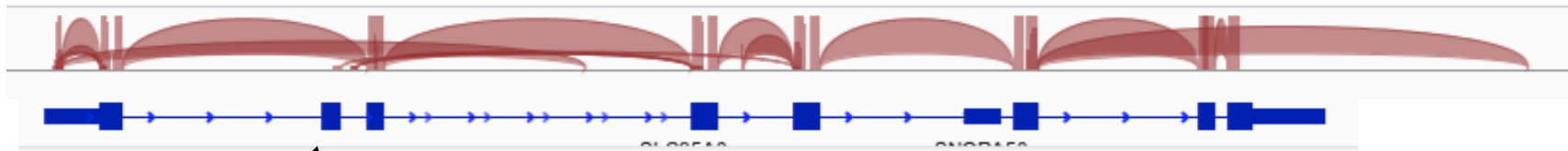
Splice junction track



Many reads connecting
these exons

RNA-seq alignments

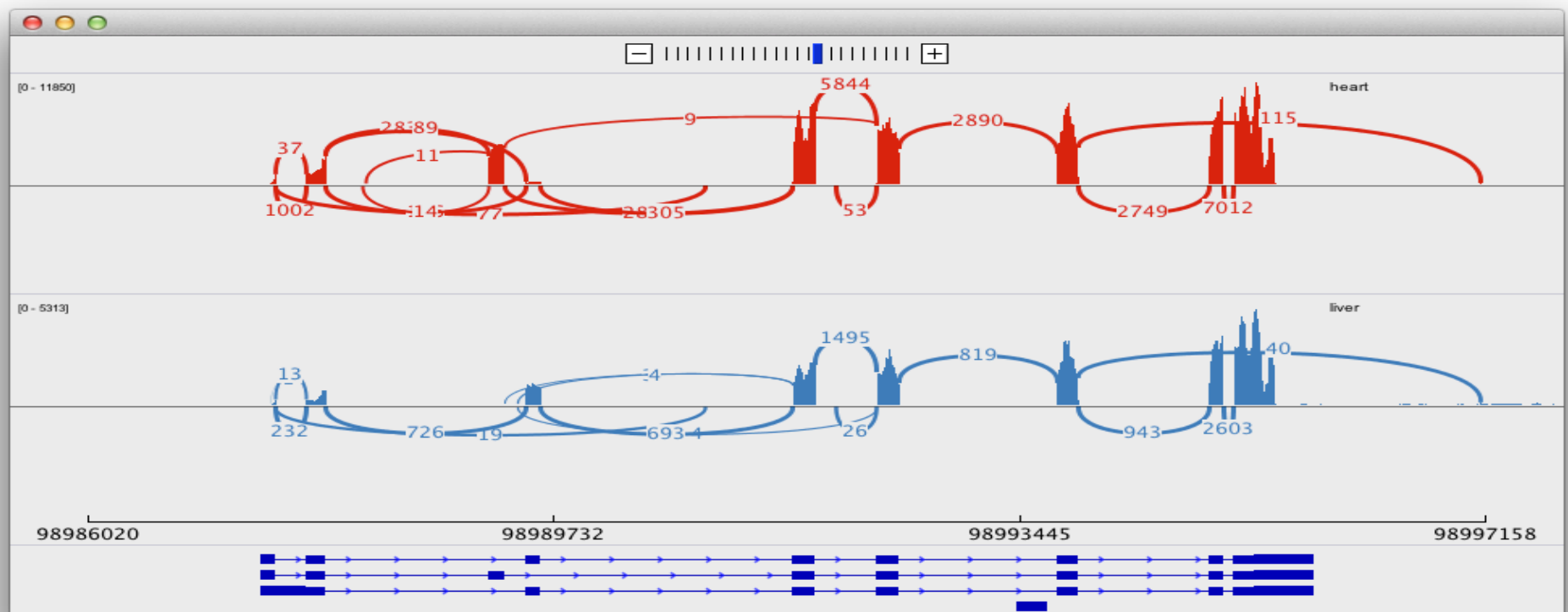
Splice junction track



Relatively few
spanning to this exon

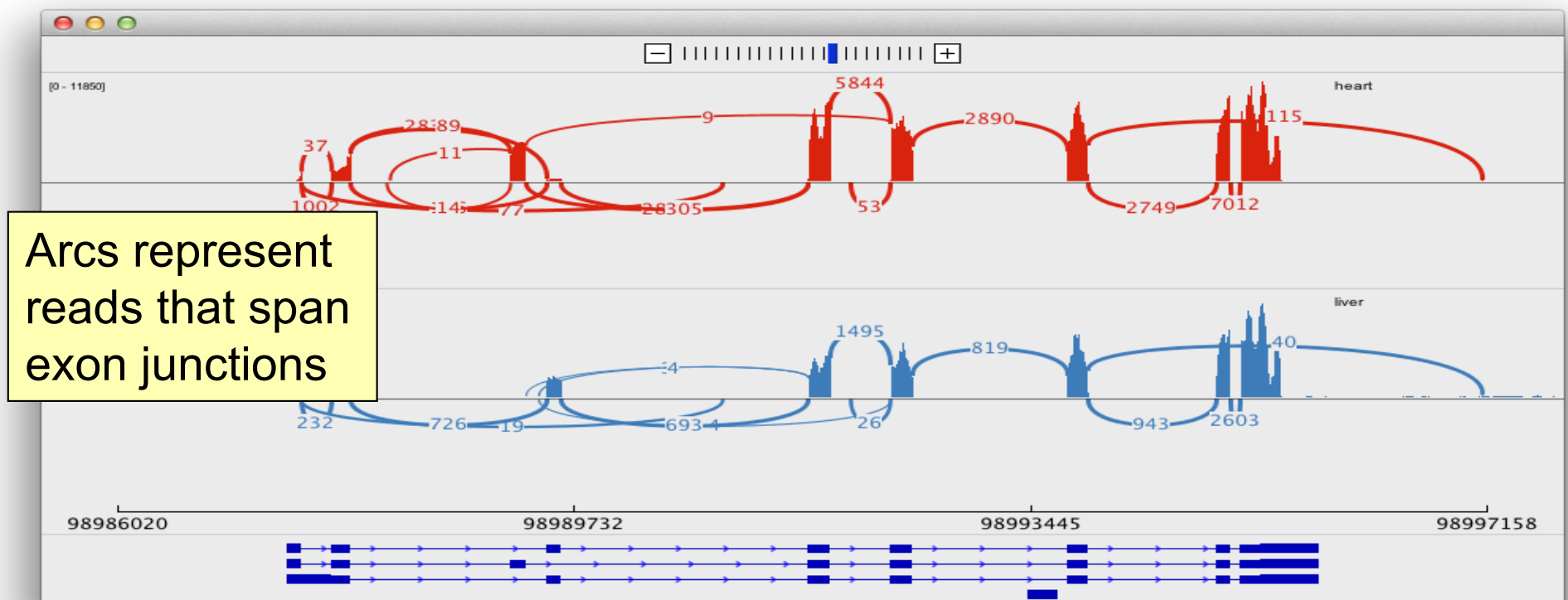
RNA-seq alignments

Sashimi Plot



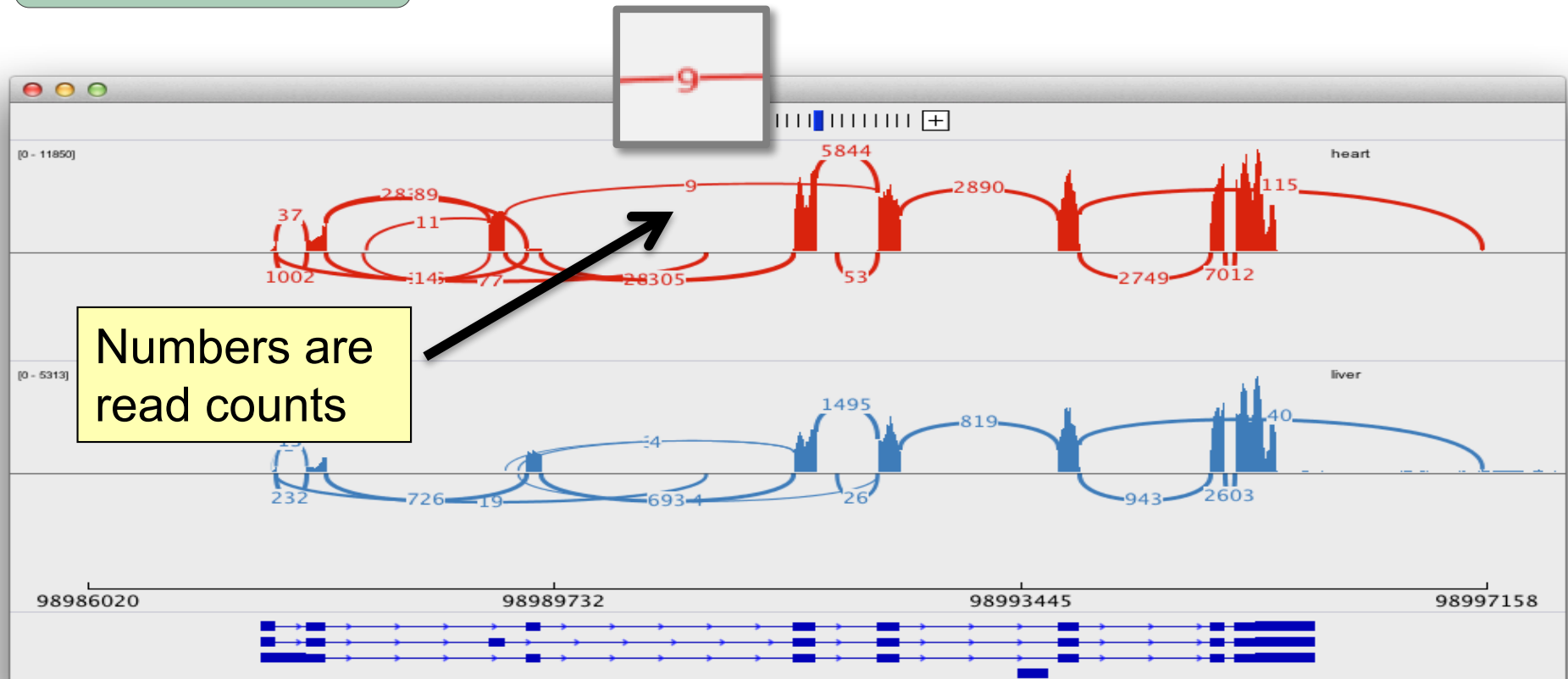
RNA-seq alignments

Sashimi Plot



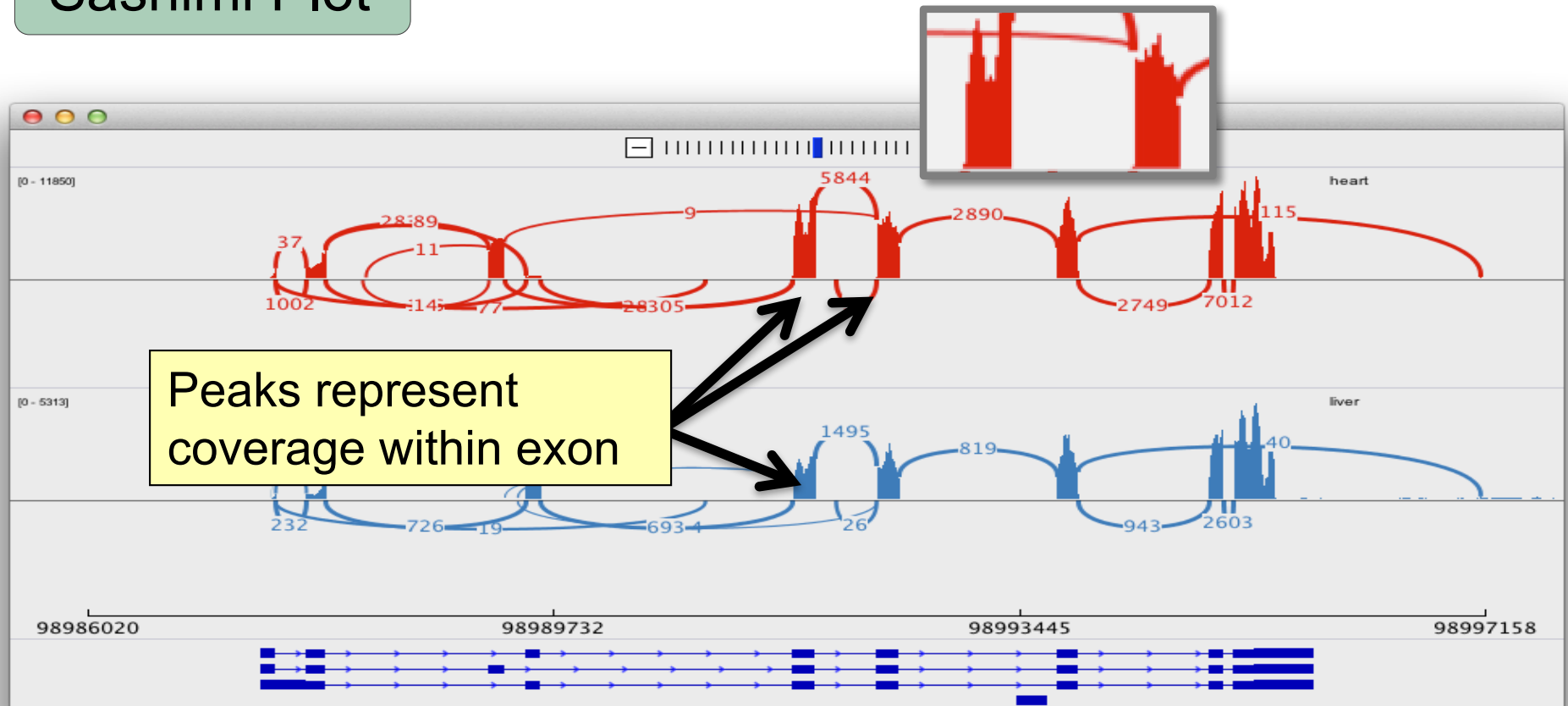
RNA-seq alignments

Sashimi Plot



RNA-seq alignments

Sashimi Plot



Sashimi plot

Viewing RNA splicing with Sashimi Plots

Reference:

Katz et al.

***Quantitative visualization of alternative exon expression
from RNA-seq data***

Bioinformatics (2015)

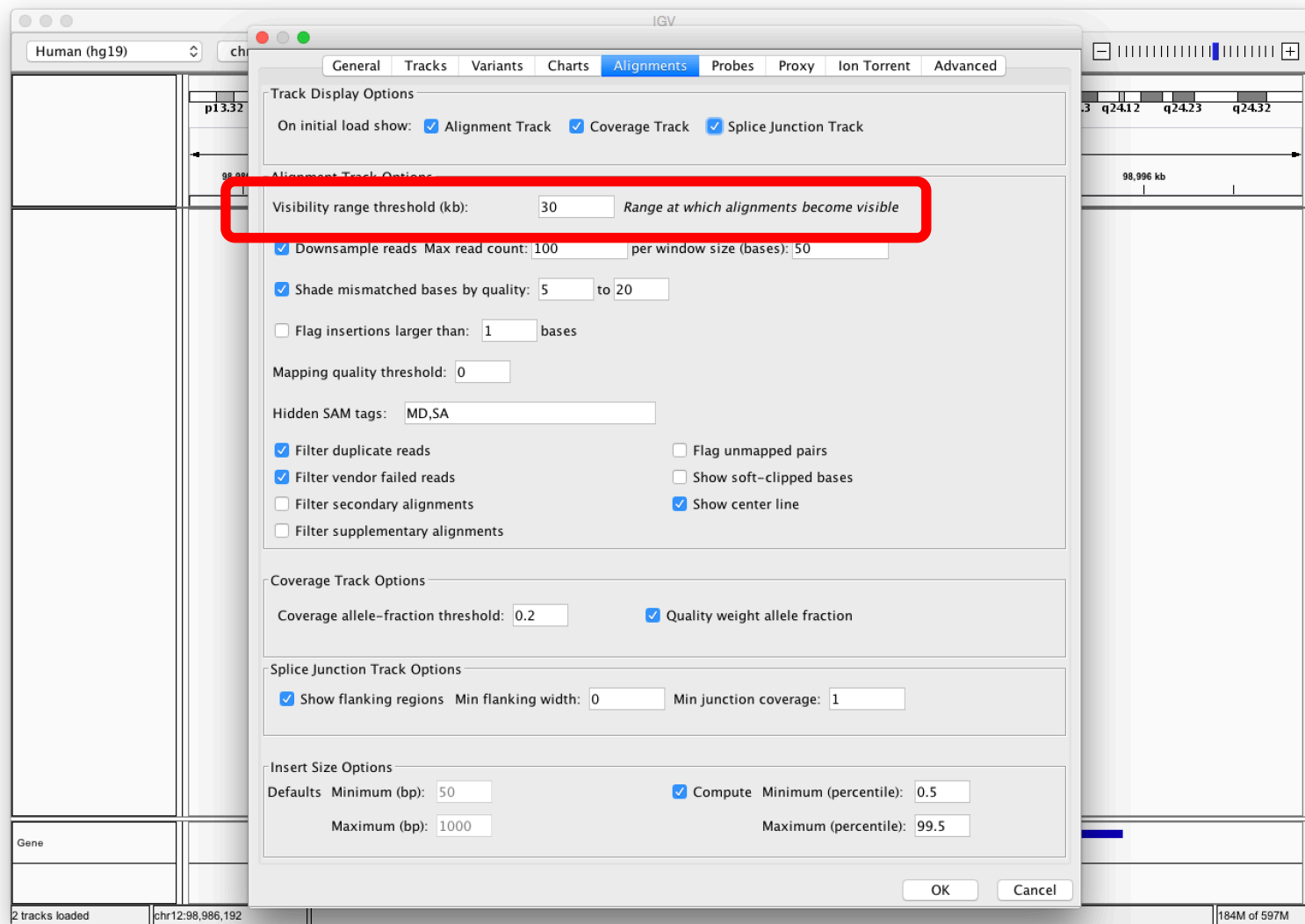
Viewing RNA-seq data

Hands-on exercise

Examine tissue-specific alternative splicing.

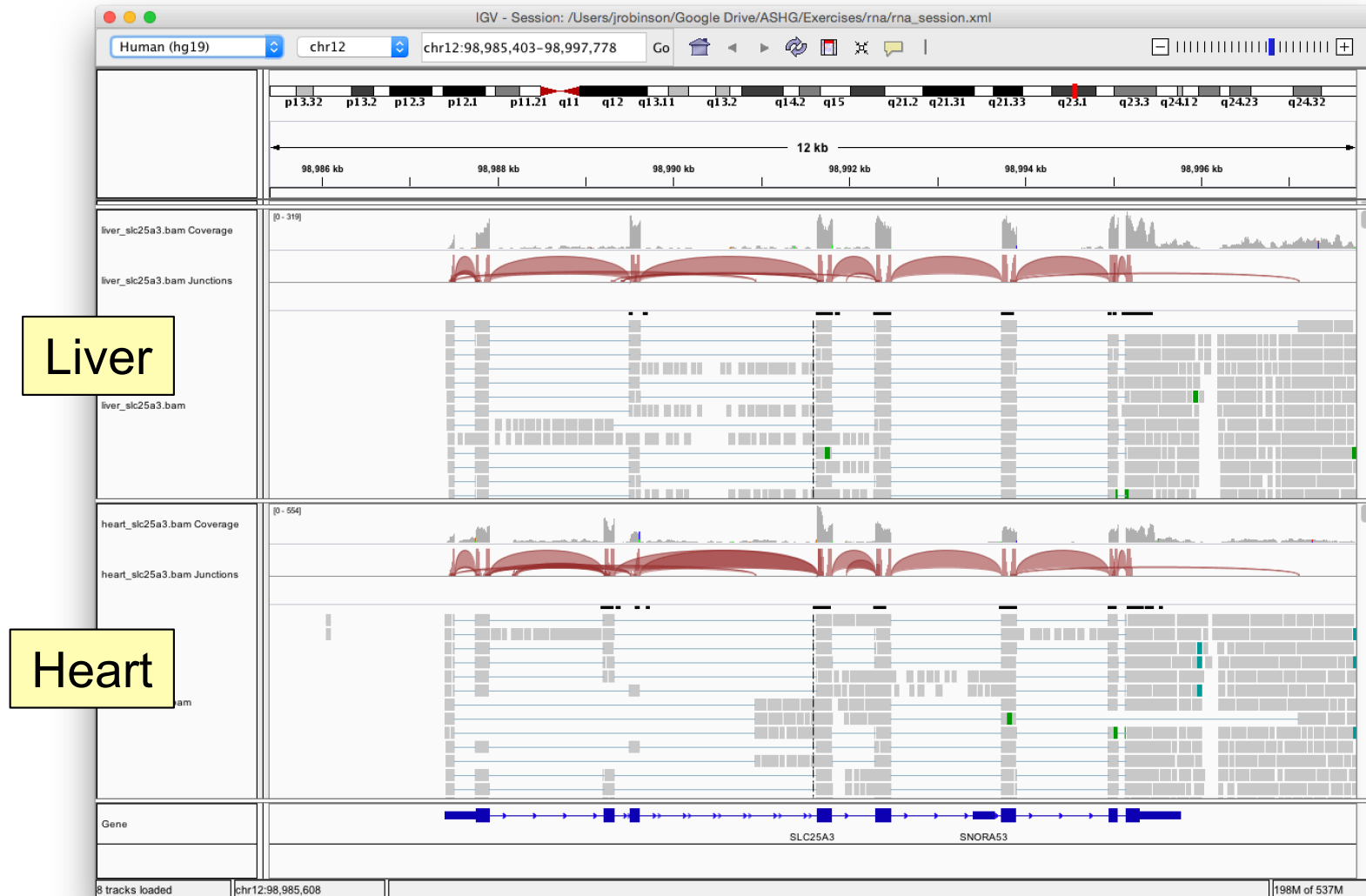
Data: Illumina BodyMap 2.0

See handout



Viewing RNA-seq data

Hands-on exercise



Viewing RNA-seq data

Hands-on exercise

Coverage



Viewing RNA-seq data

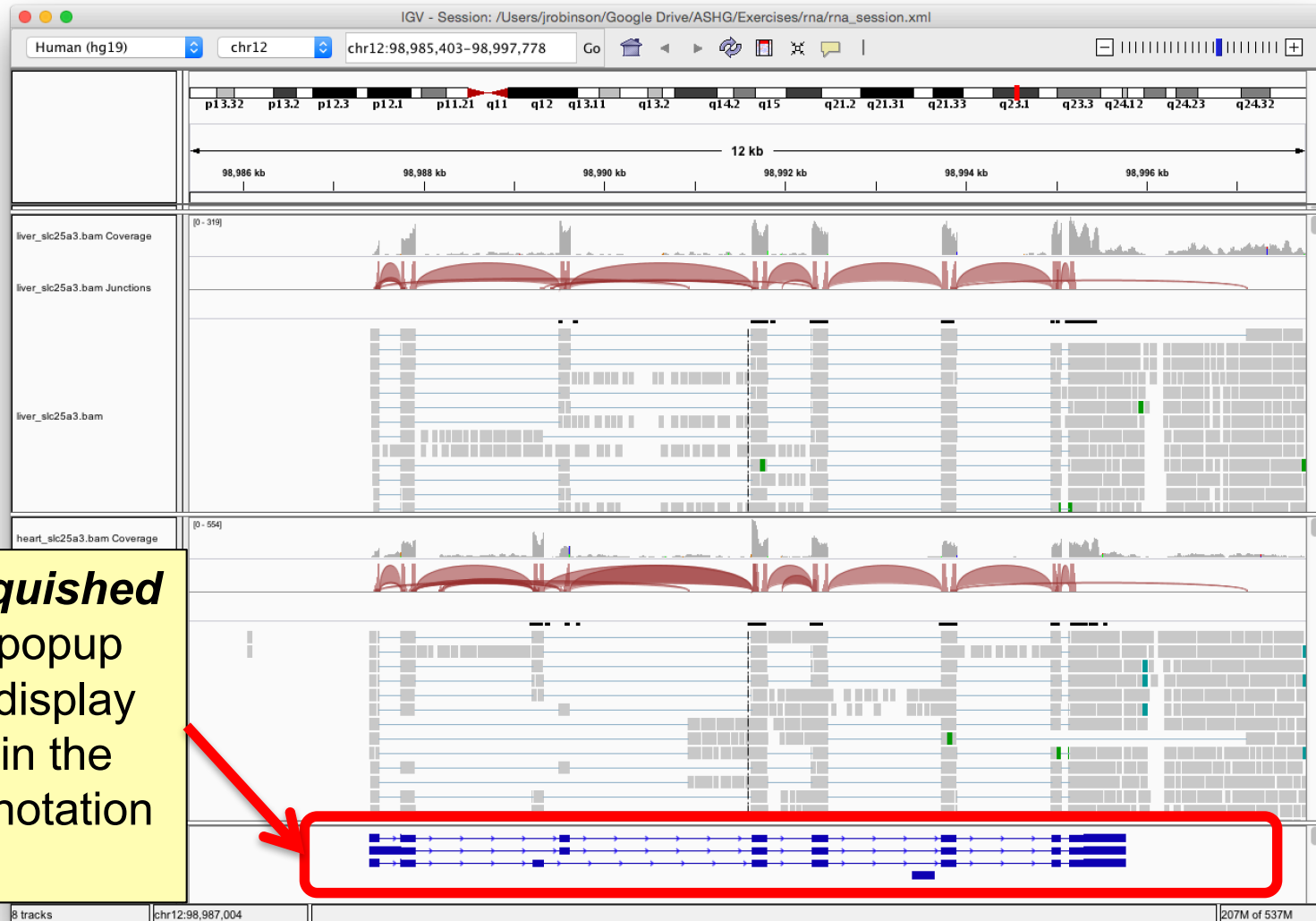
Hands-on exercise

Junction Coverage



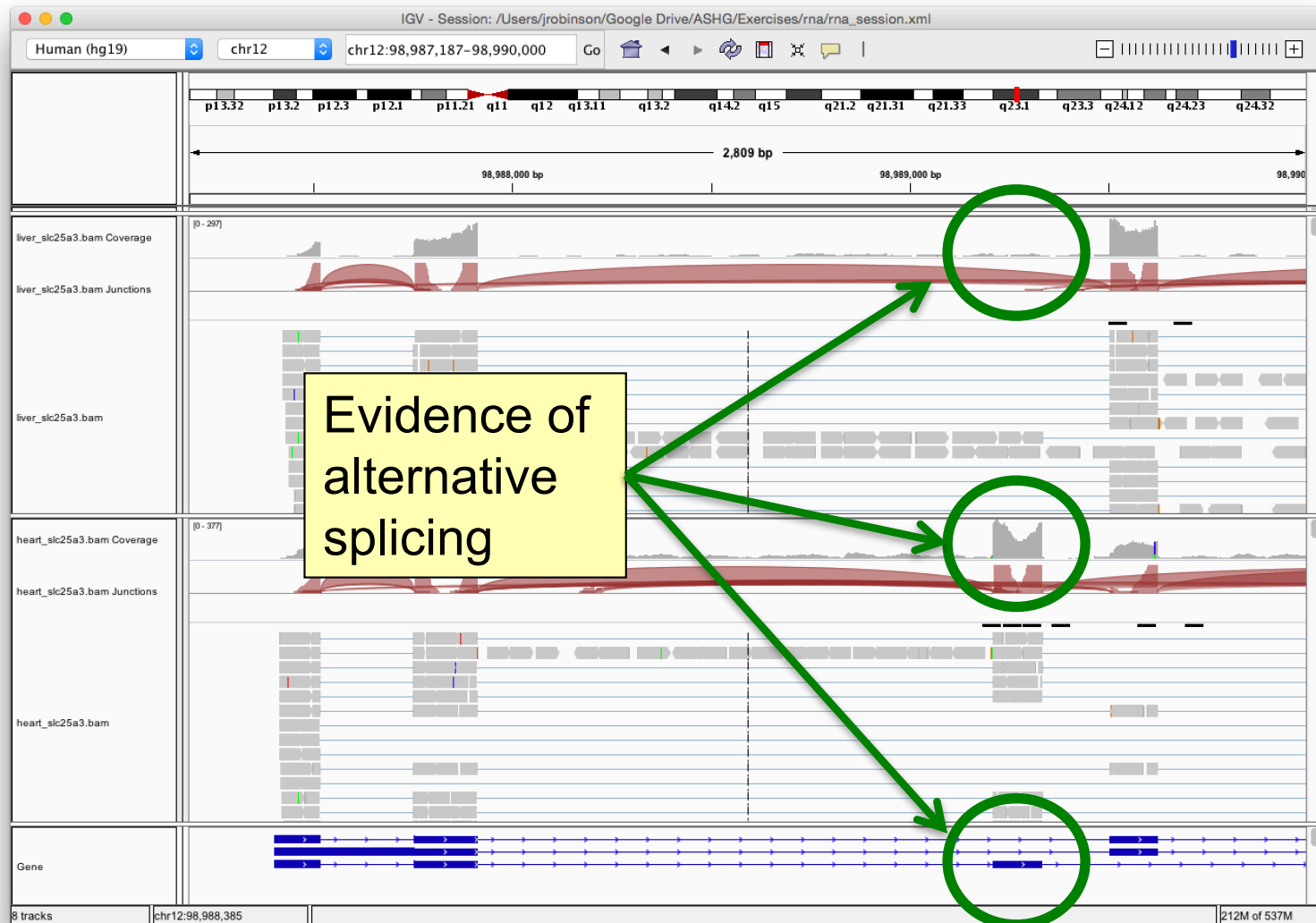
Viewing RNA-seq data

Hands-on exercise



Viewing RNA-seq data

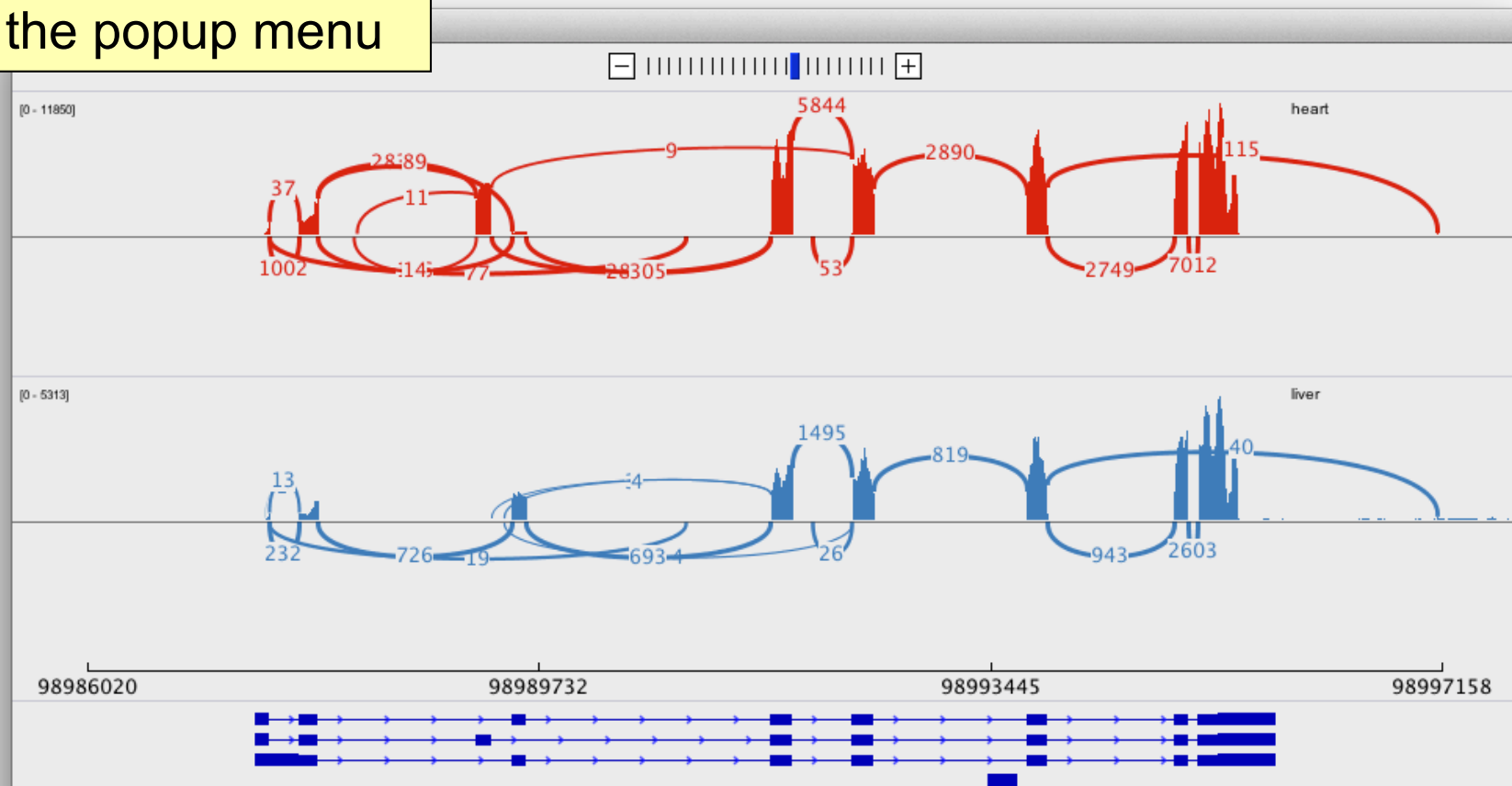
Hands-on exercise



Viewing RNA-seq data

Hands-on exercise

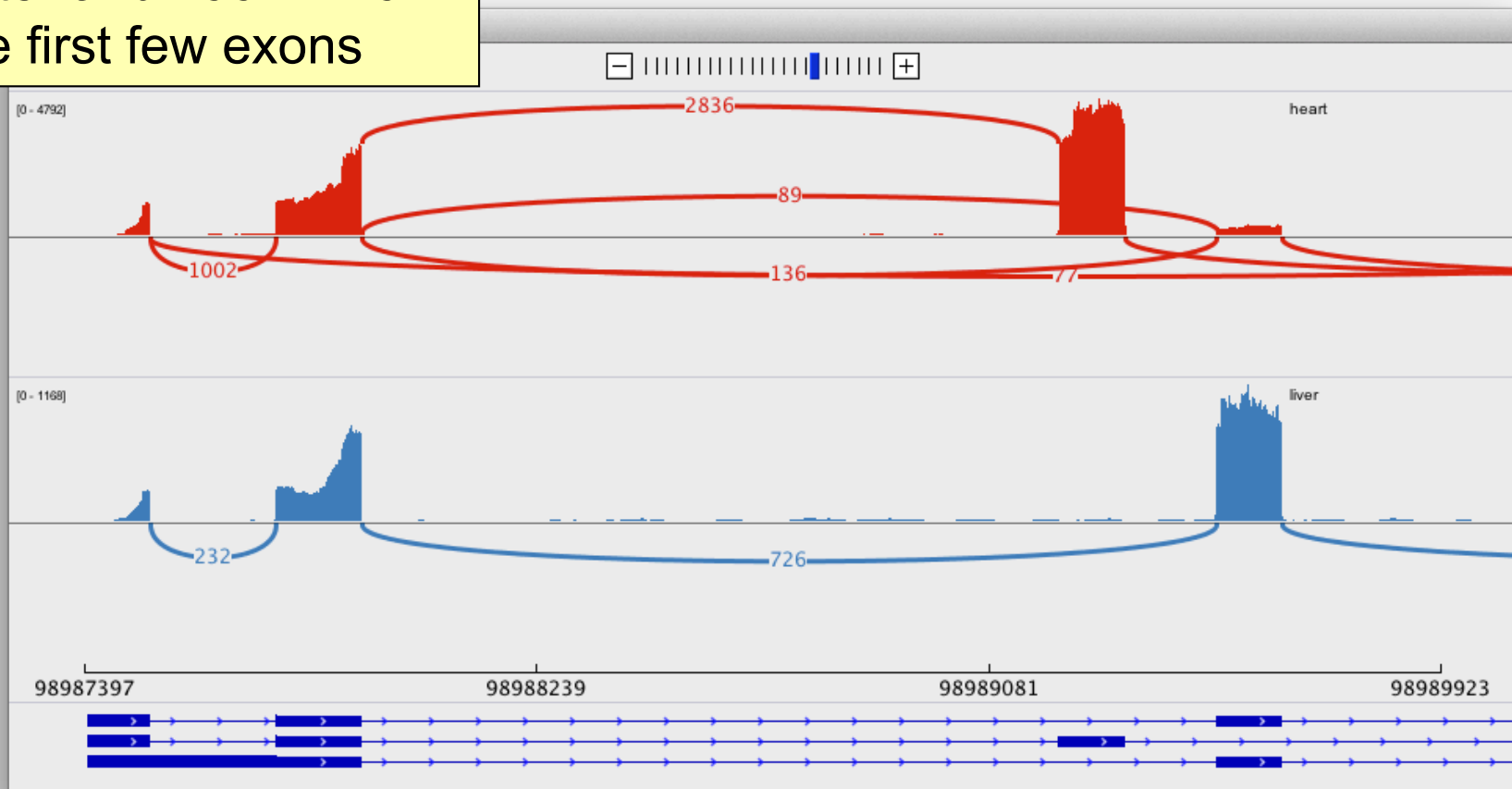
Launch **Sashimi Plot**
from the popup menu



Viewing RNA-seq data

Hands-on exercise

Filter and zoom in on
the first few exons



Viewing Structural Events

Structural events

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

Paired-end sequencing

DNA or
cDNA



fragment



read from
each end



align to
reference



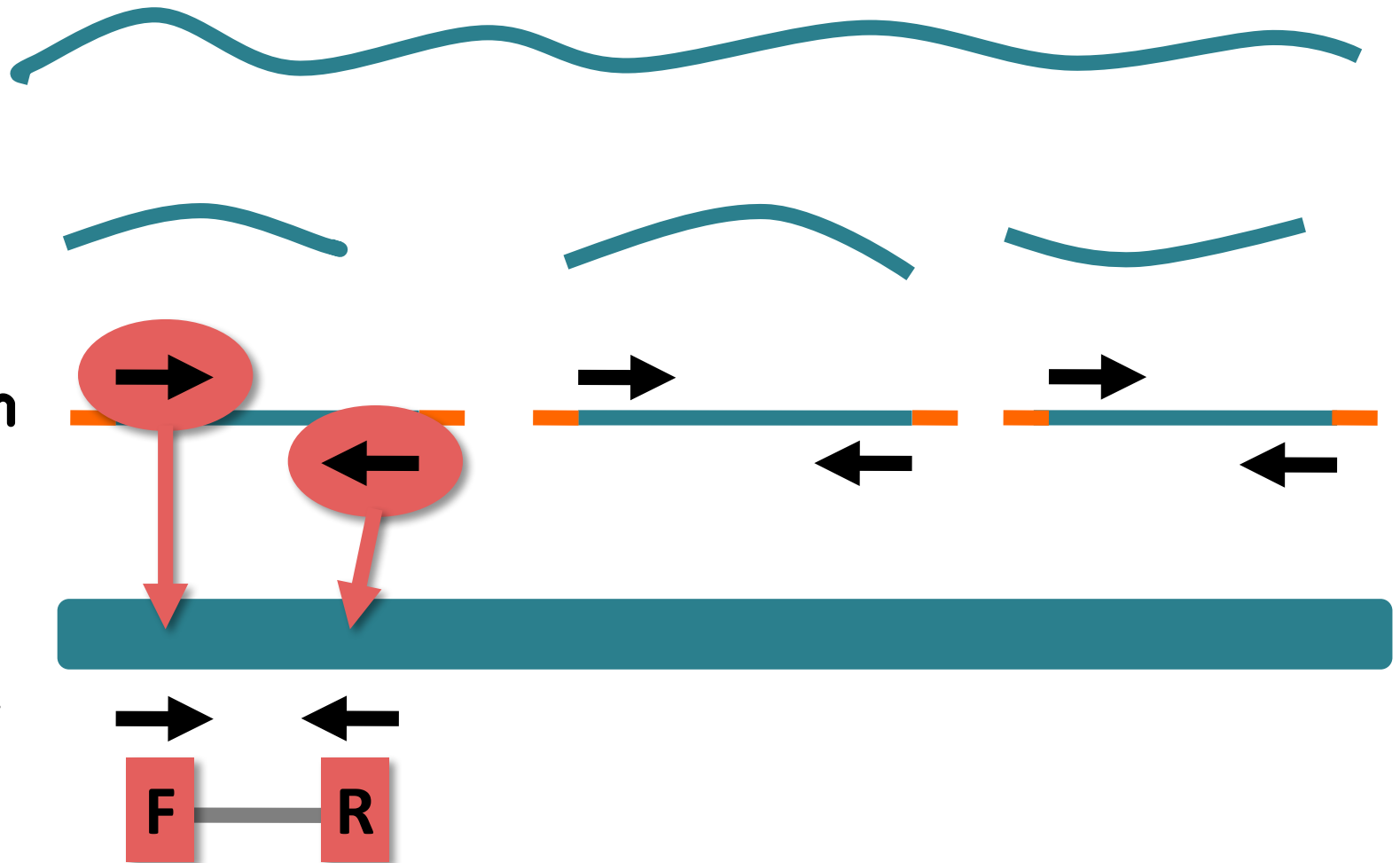
Paired-end sequencing

DNA or
cDNA

fragment

read from
each end

align to
reference



Paired-end sequencing

DNA or
cDNA



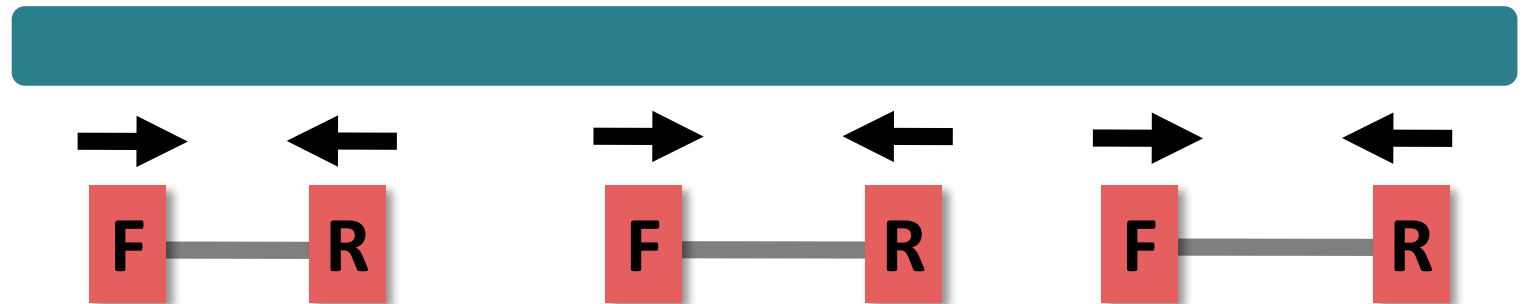
fragment



read from
each end



align to
reference



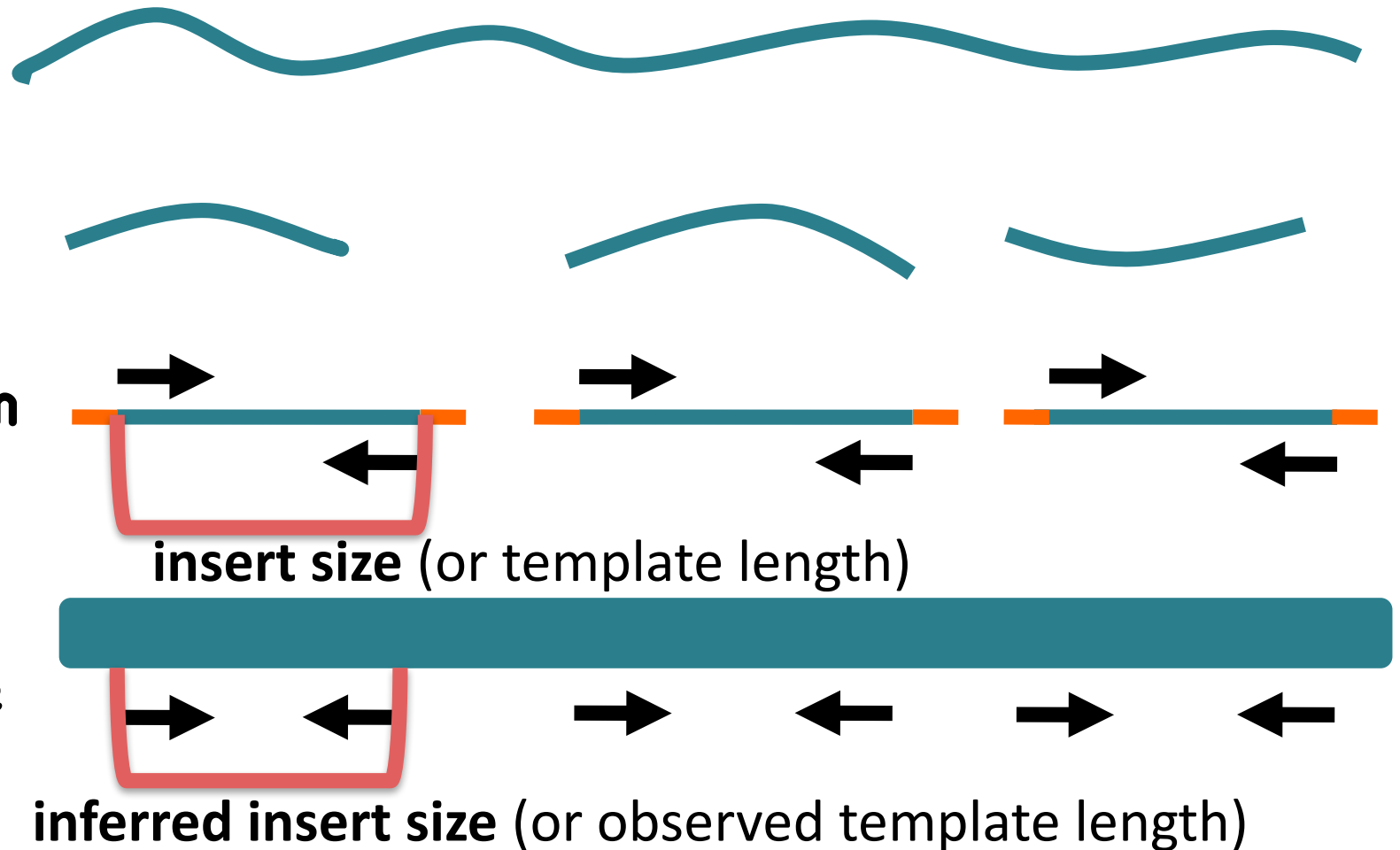
Paired-end sequencing

DNA or
cDNA

fragment

read from
each end

align to
reference



Interpreting inferred insert size

Differences between *insert size* and *inferred insert size** can be used to detect evidence of structural variants, including

- Deletions
- Insertions
- Inter-chromosomal rearrangements

* or *template length* and *observed template length*

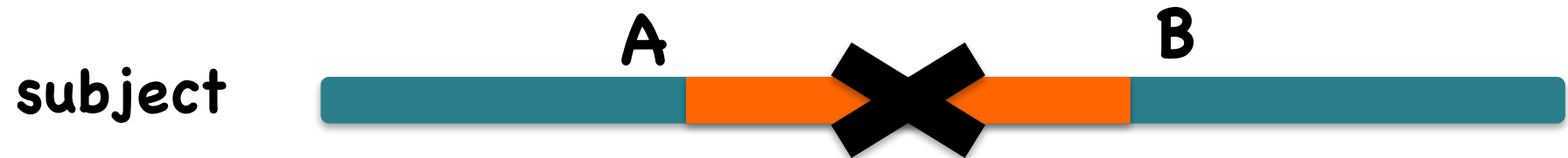
Interpreting inferred insert size

Differences between *insert size* and *inferred insert size** can be used to detect evidence of structural variants, including

- Deletions
- Insertions
- Inter-chromosomal rearrangements

* or *template length* and *observed template length*

Deletion



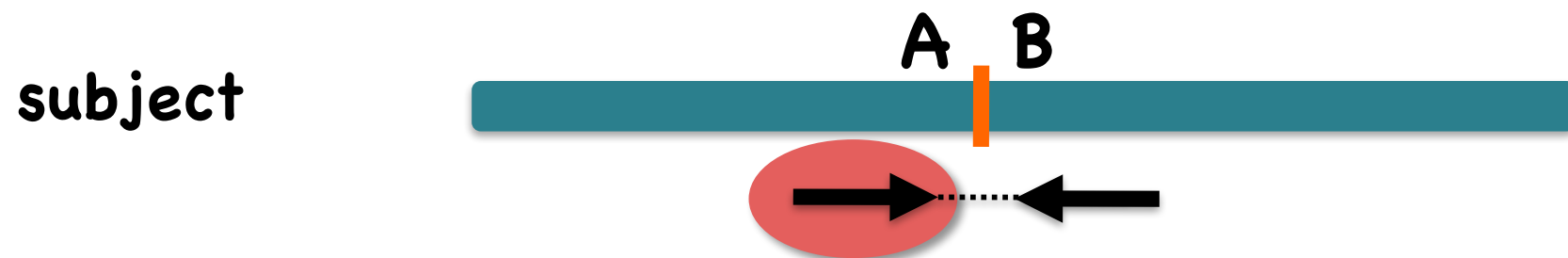
Deletion



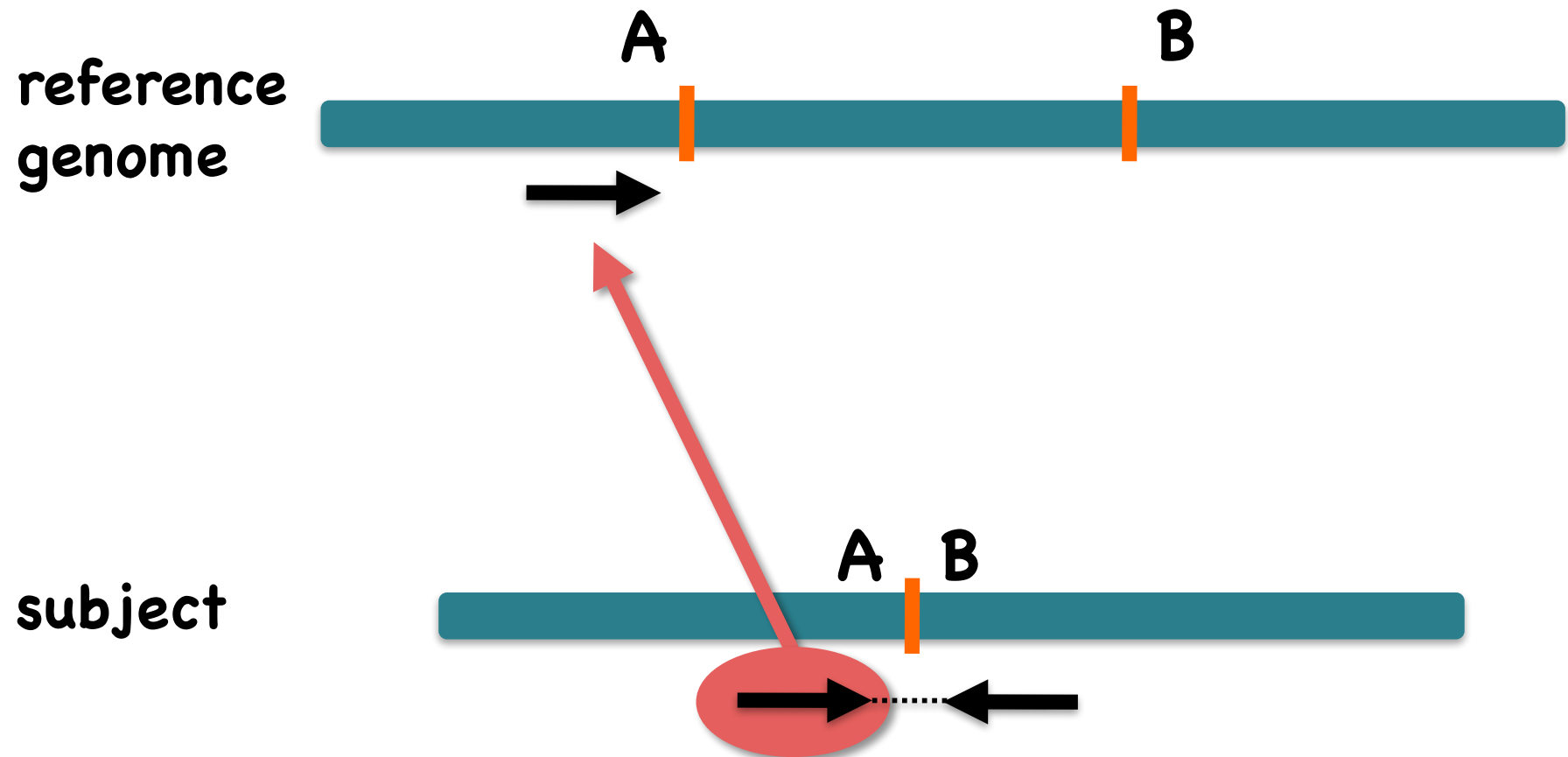
Deletion



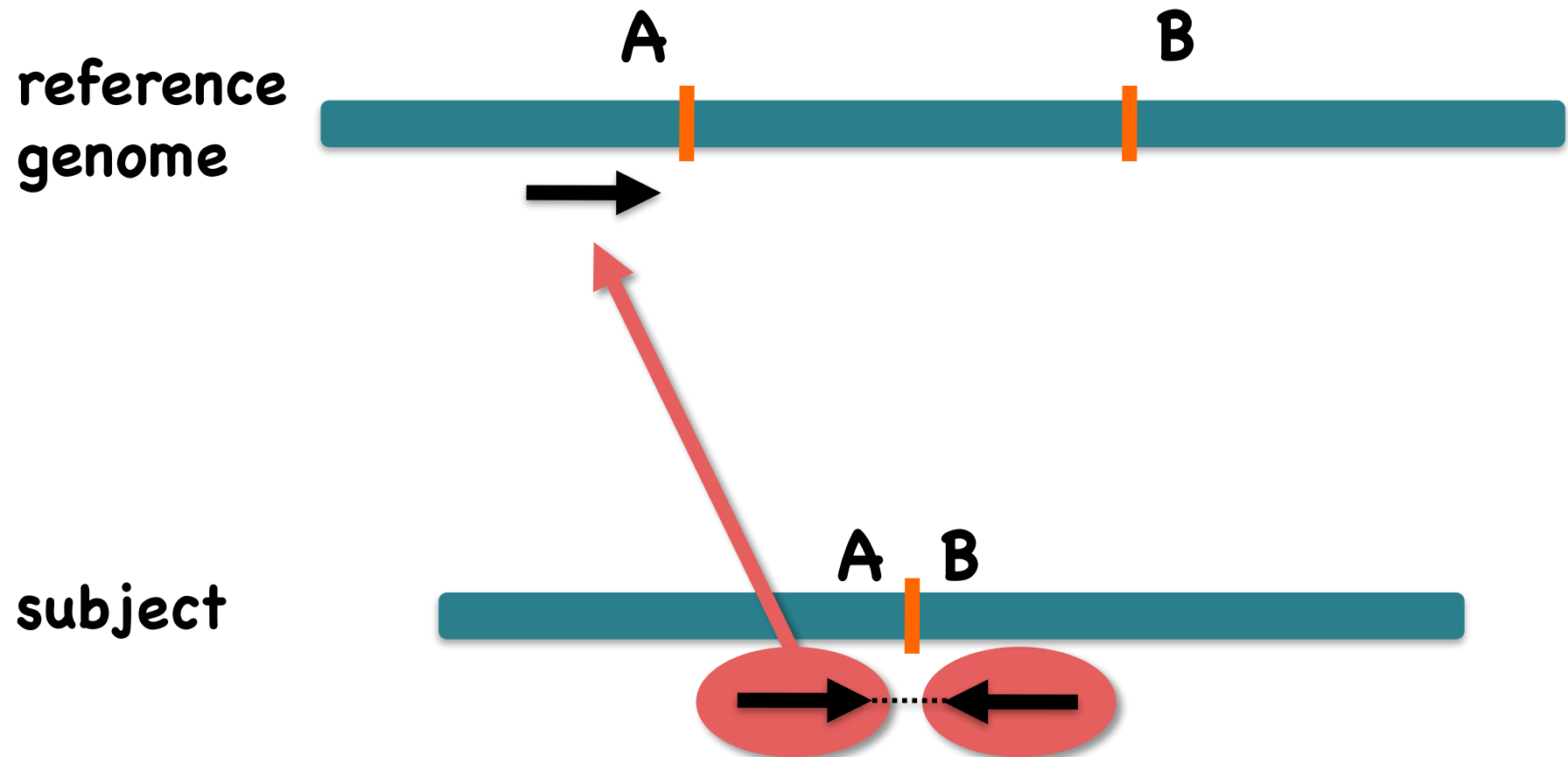
Deletion



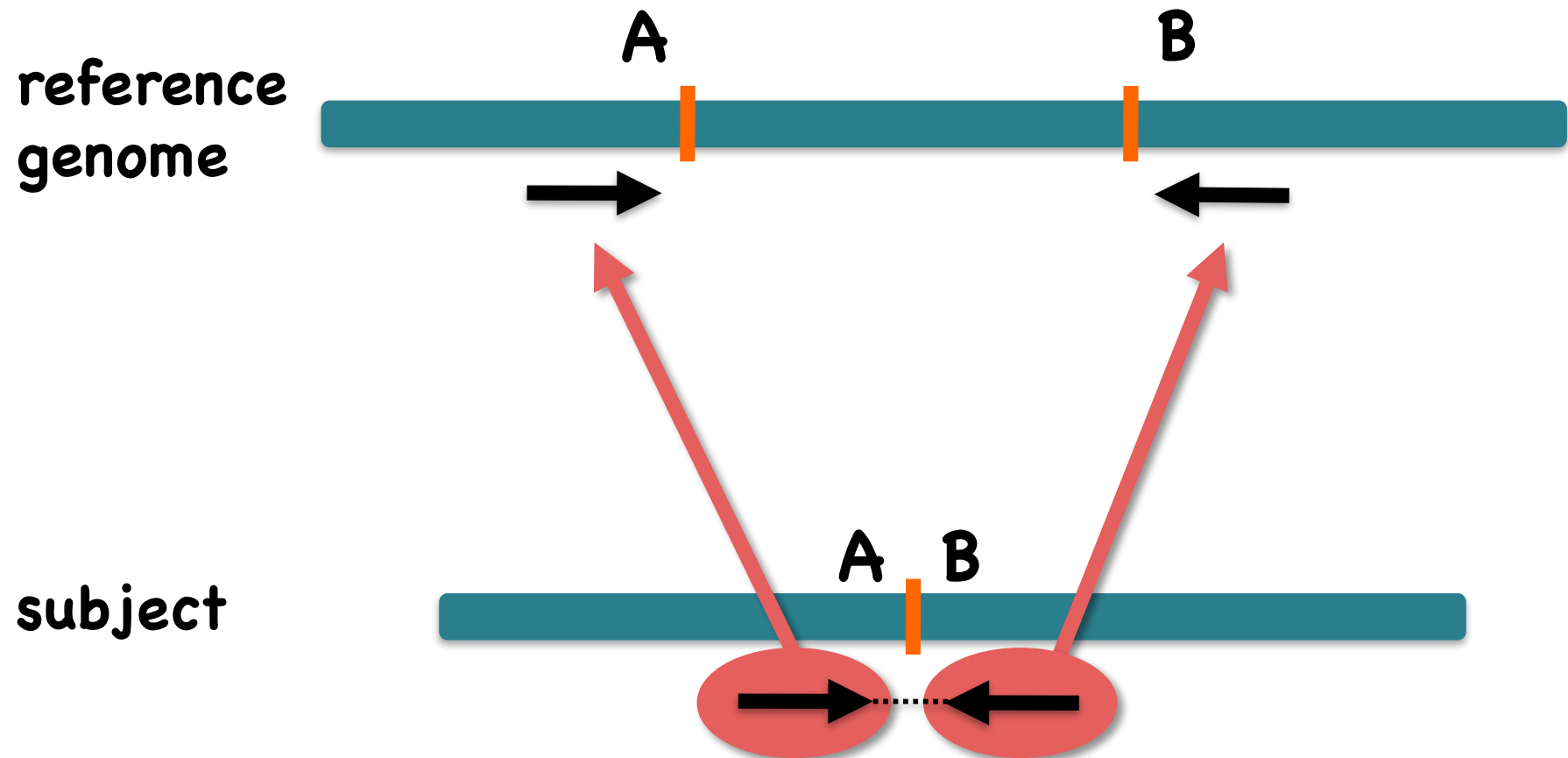
Deletion



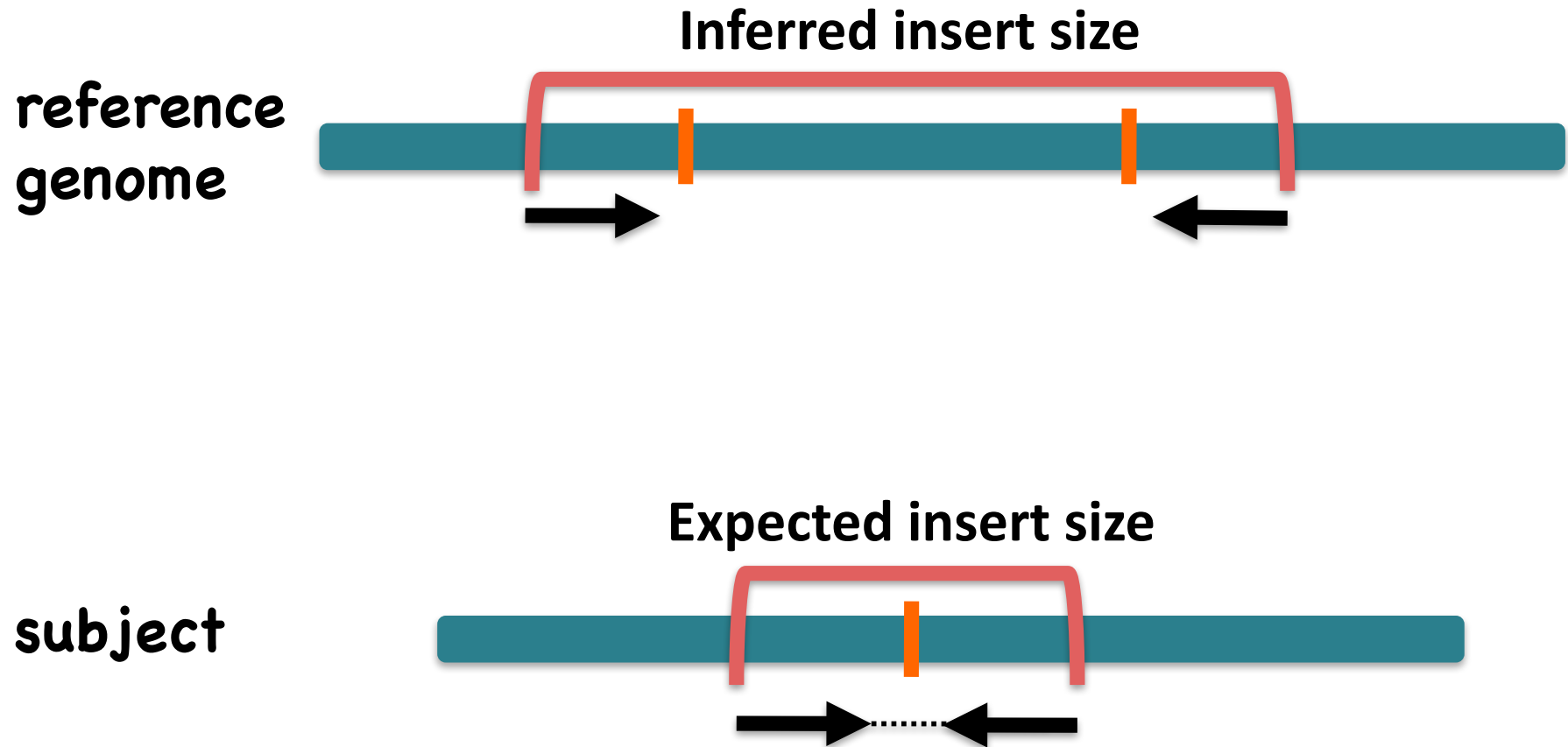
Deletion



Deletion



Deletion



Inferred insert size is greater than expected value

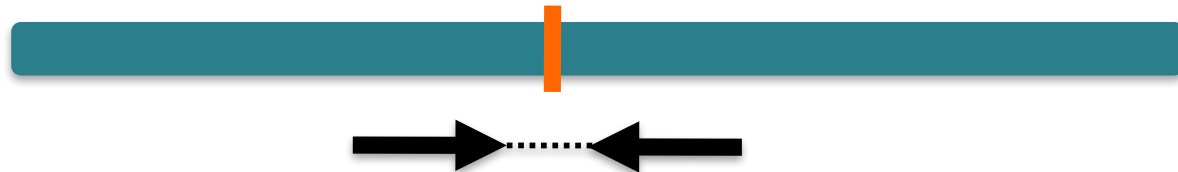
Deletion

Pairs with larger than expected insert size are colored red

**reference
genome**

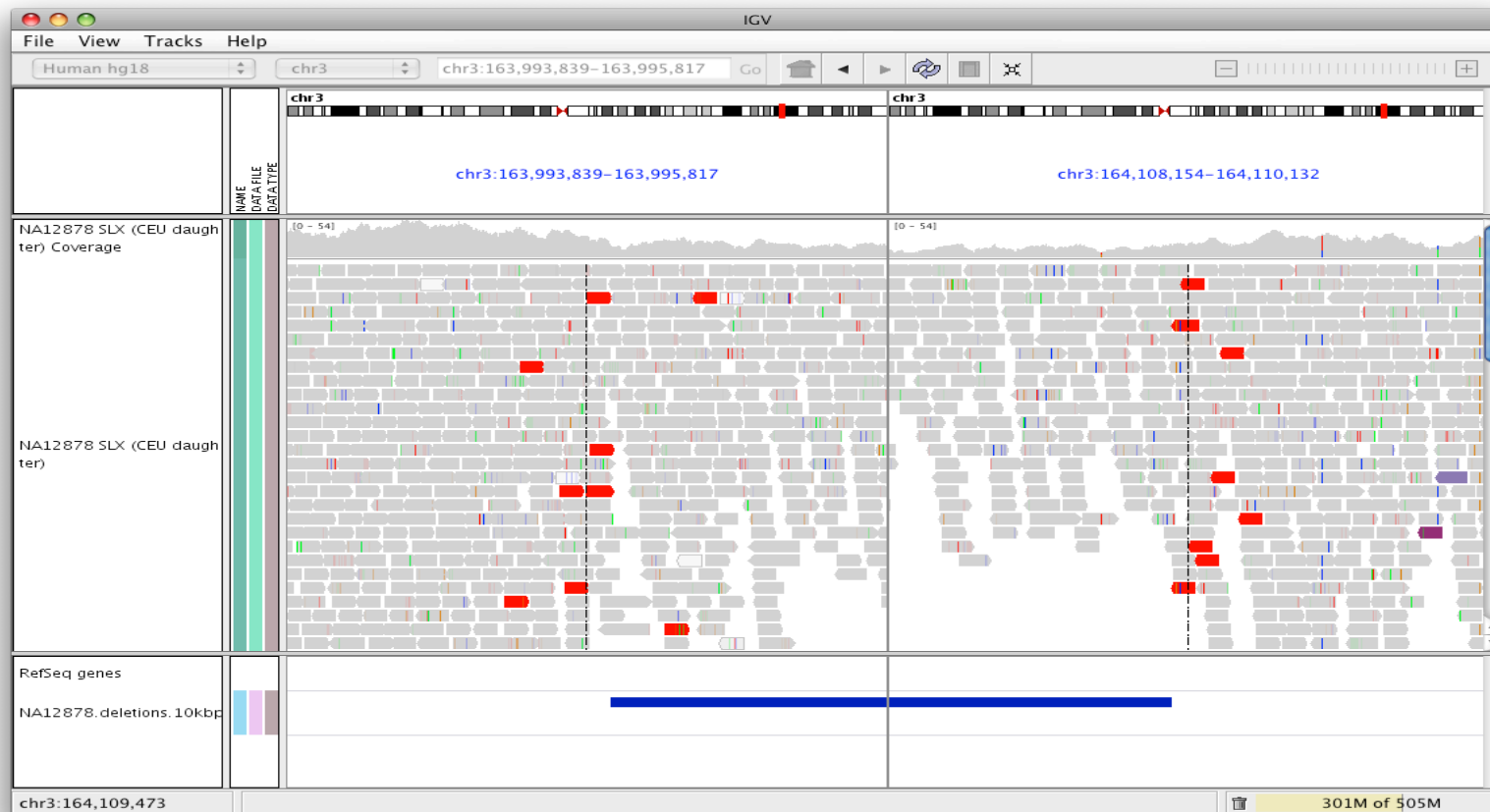


subject

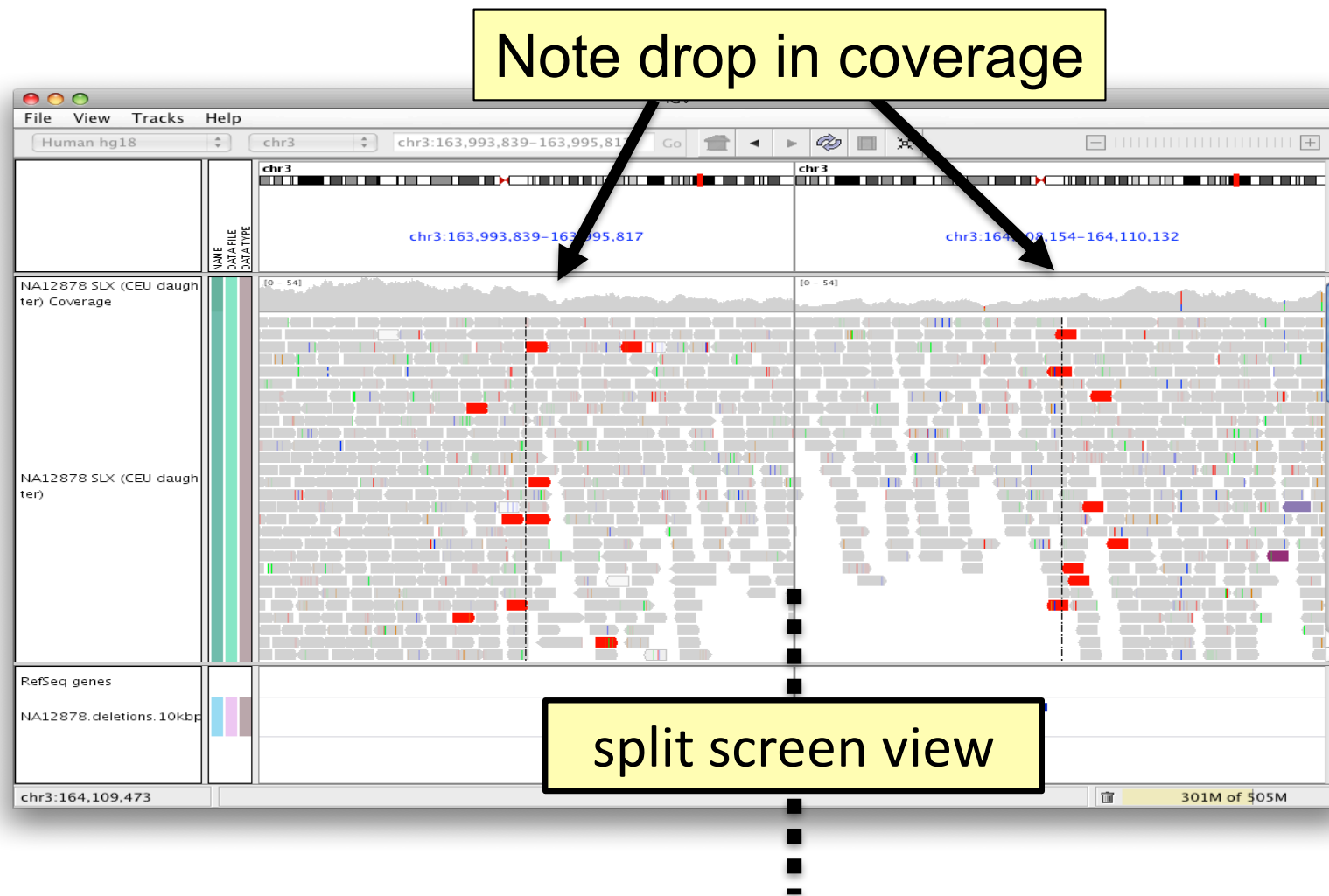


Deletion

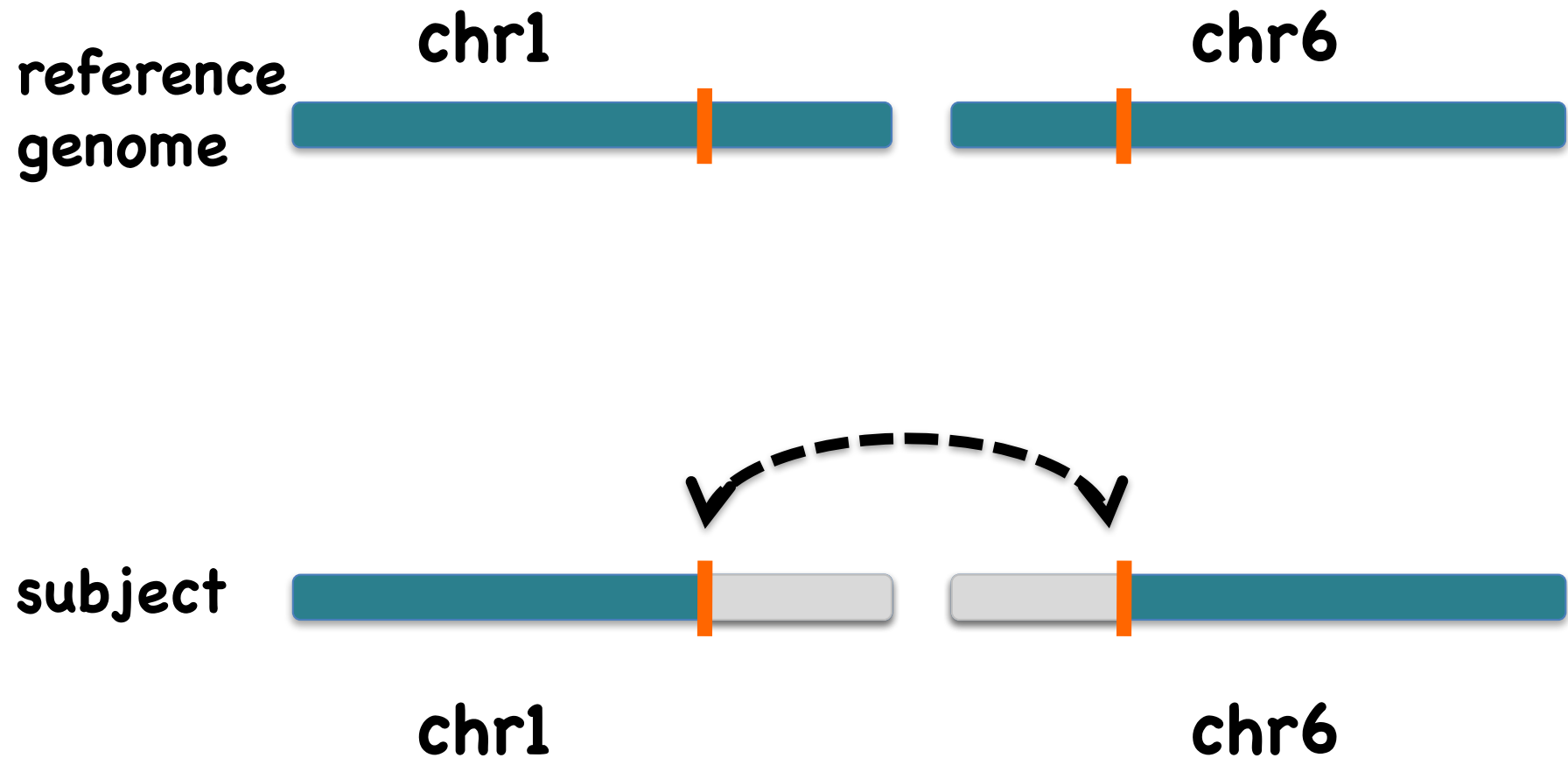
Pairs with larger than expected insert size are colored red



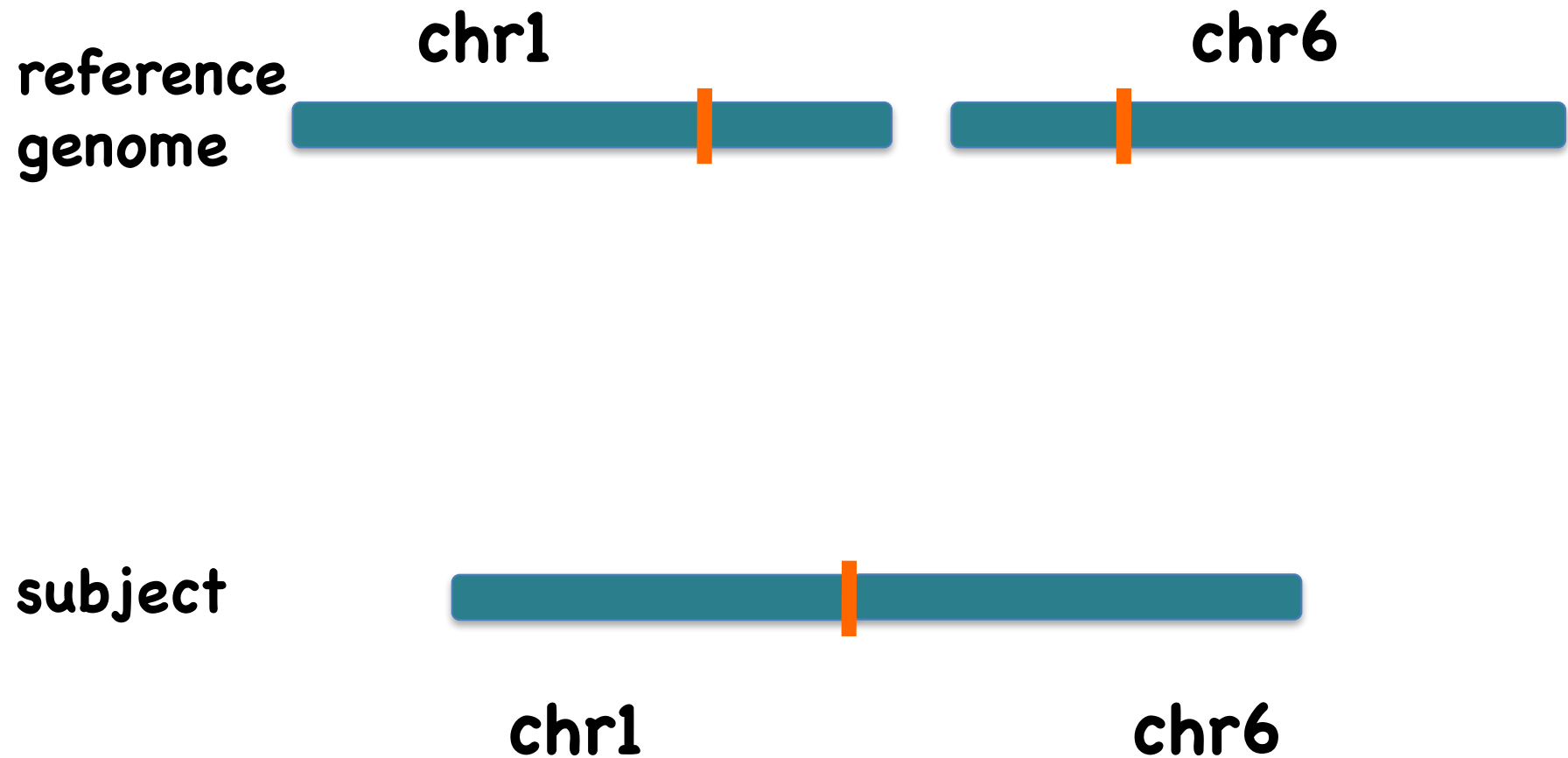
Deletion



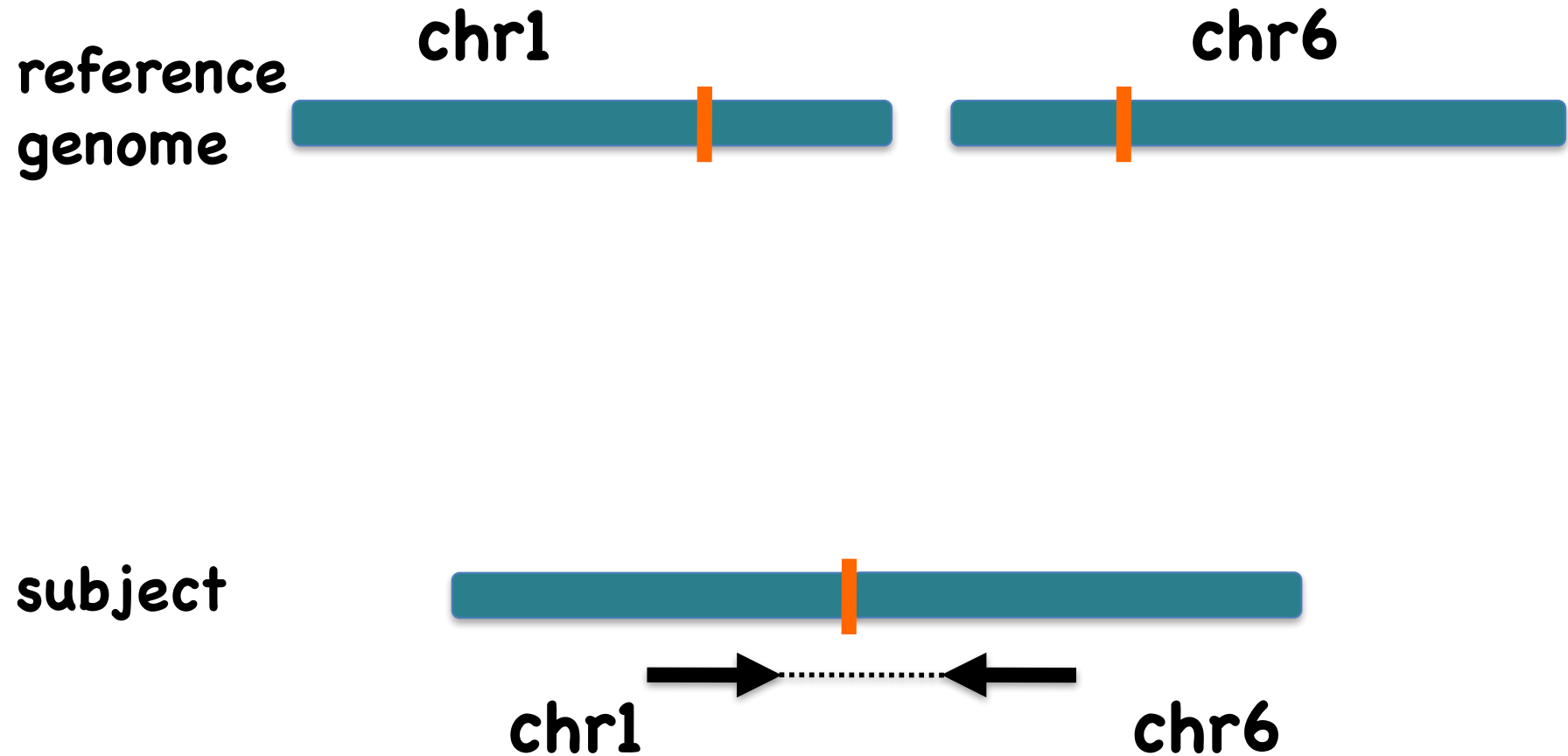
Inter-chromosomal fusion



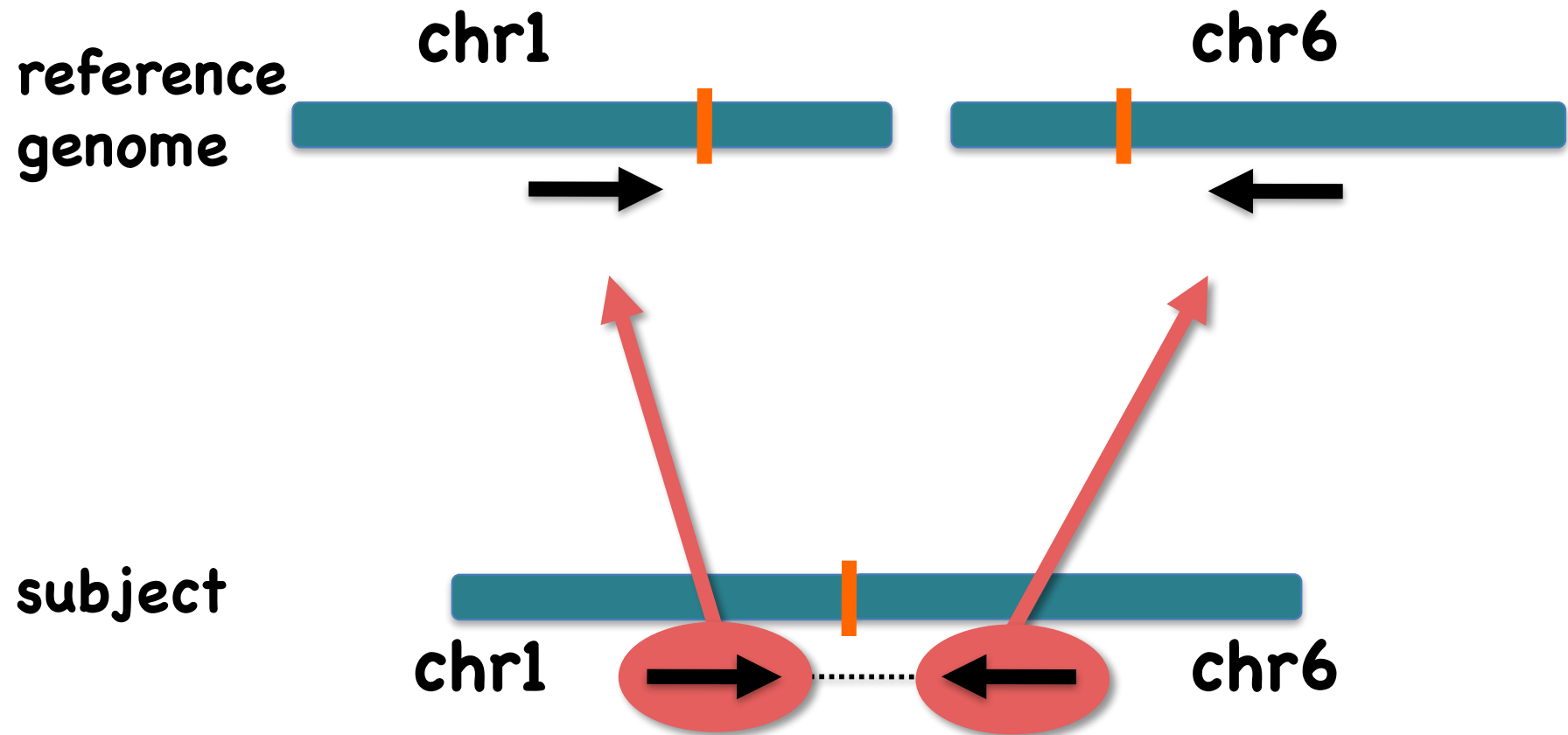
Inter-chromosomal fusion



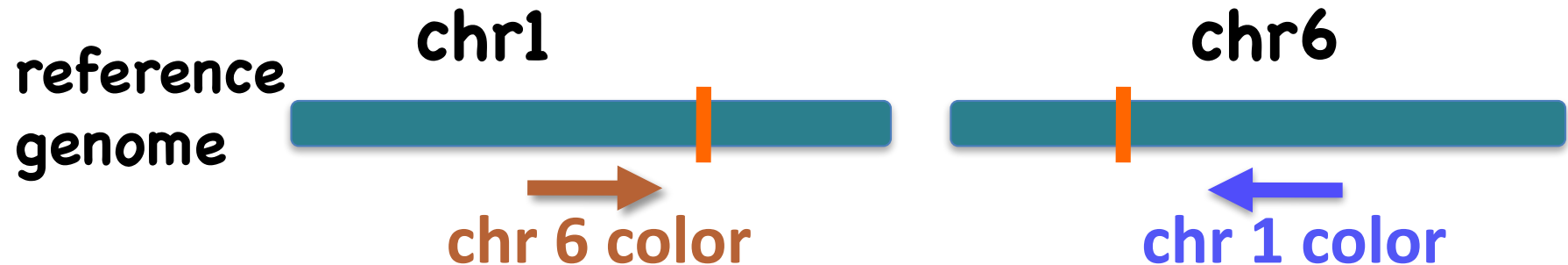
Inter-chromosomal fusion



Inter-chromosomal fusion

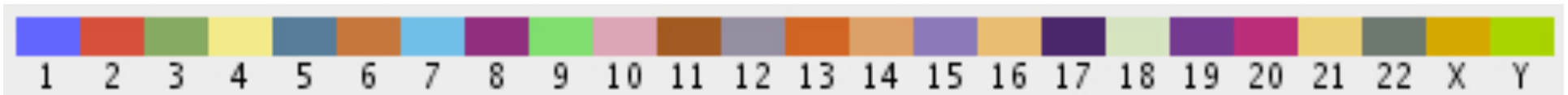


Inter-chromosomal fusion

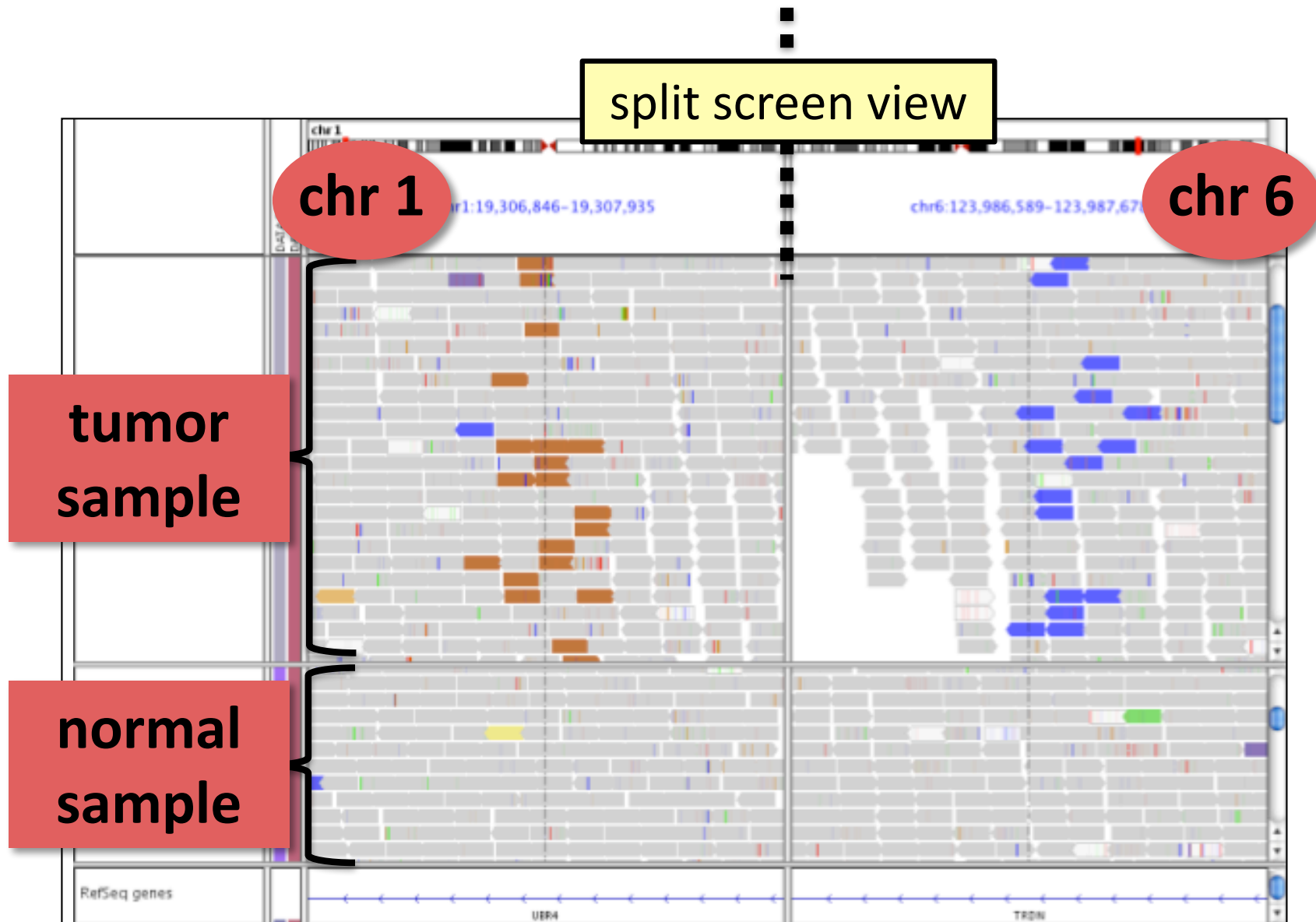


Each read is colored by the chromosome of its mate

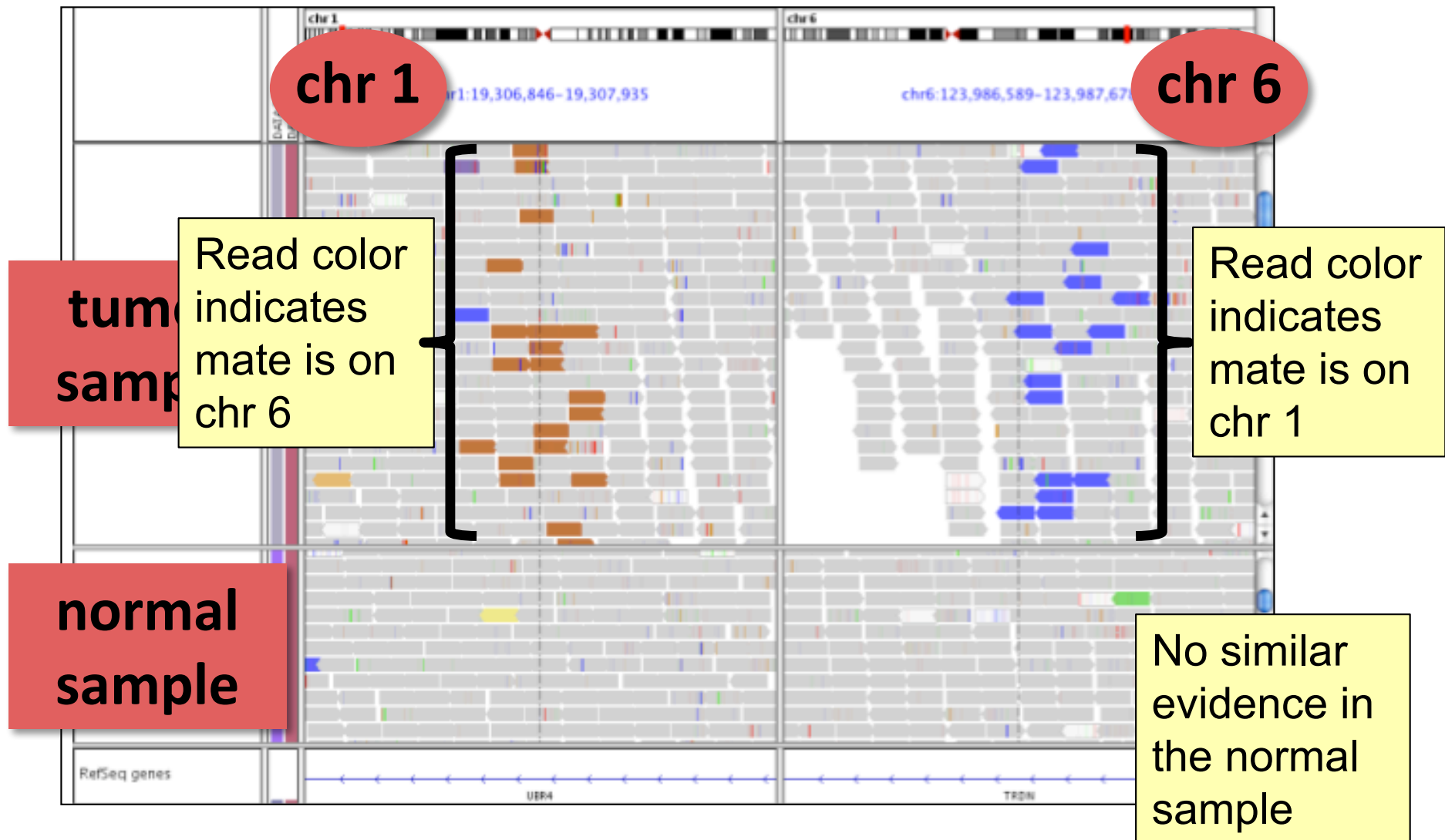
Chromosome colors:



Inter-chromosomal fusion



Inter-chromosomal fusion



Interpreting pair orientations

Orientation of paired reads can reveal evidence of structural events, including:

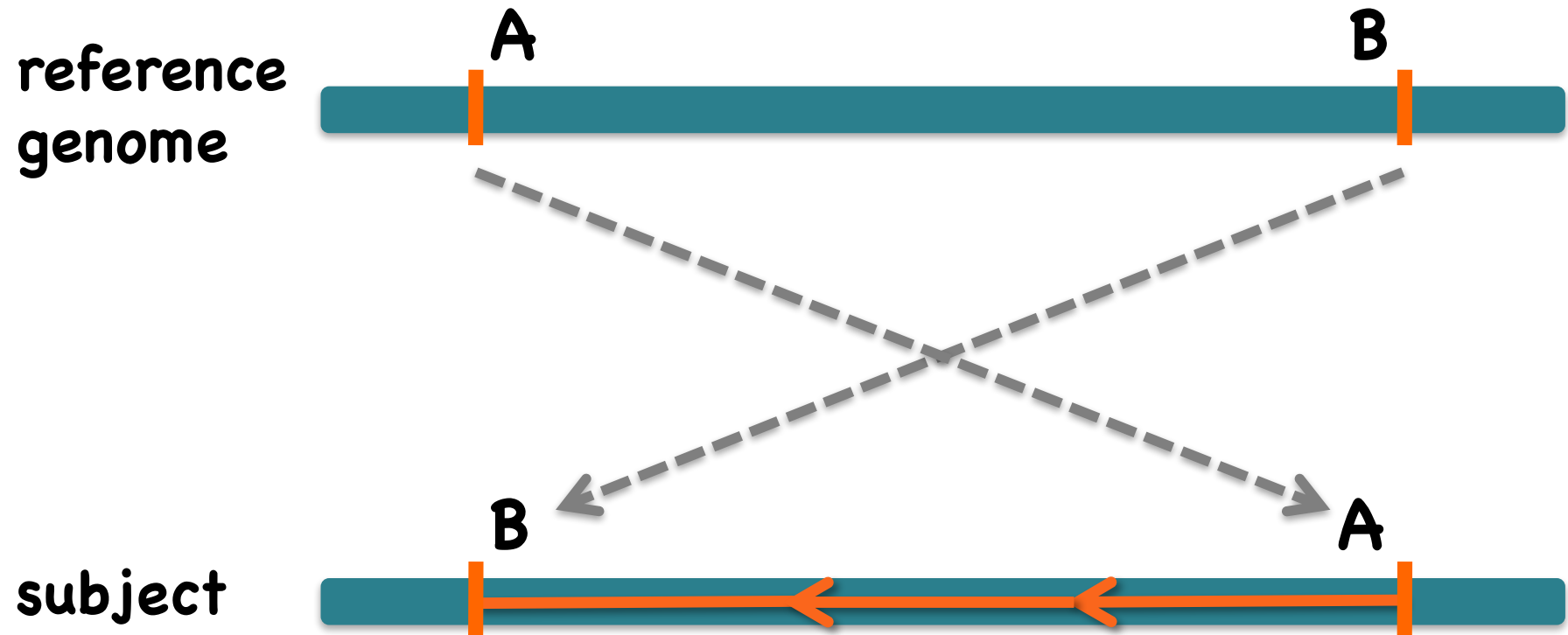
- Inversions
- Duplications
- Translocations

Interpreting pair orientations

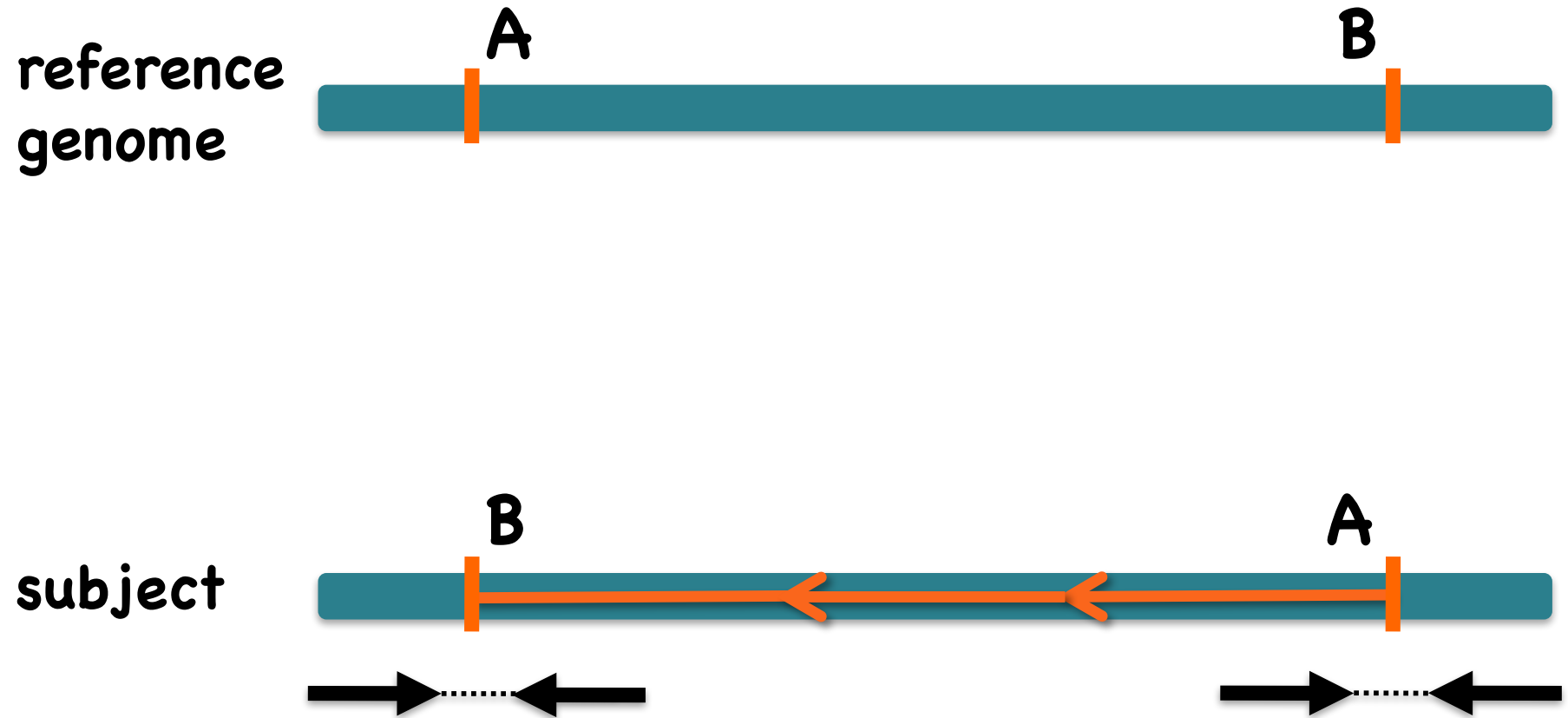
Orientation of paired reads can reveal evidence of structural events, including:

- Inversions
- Duplications
- Translocations

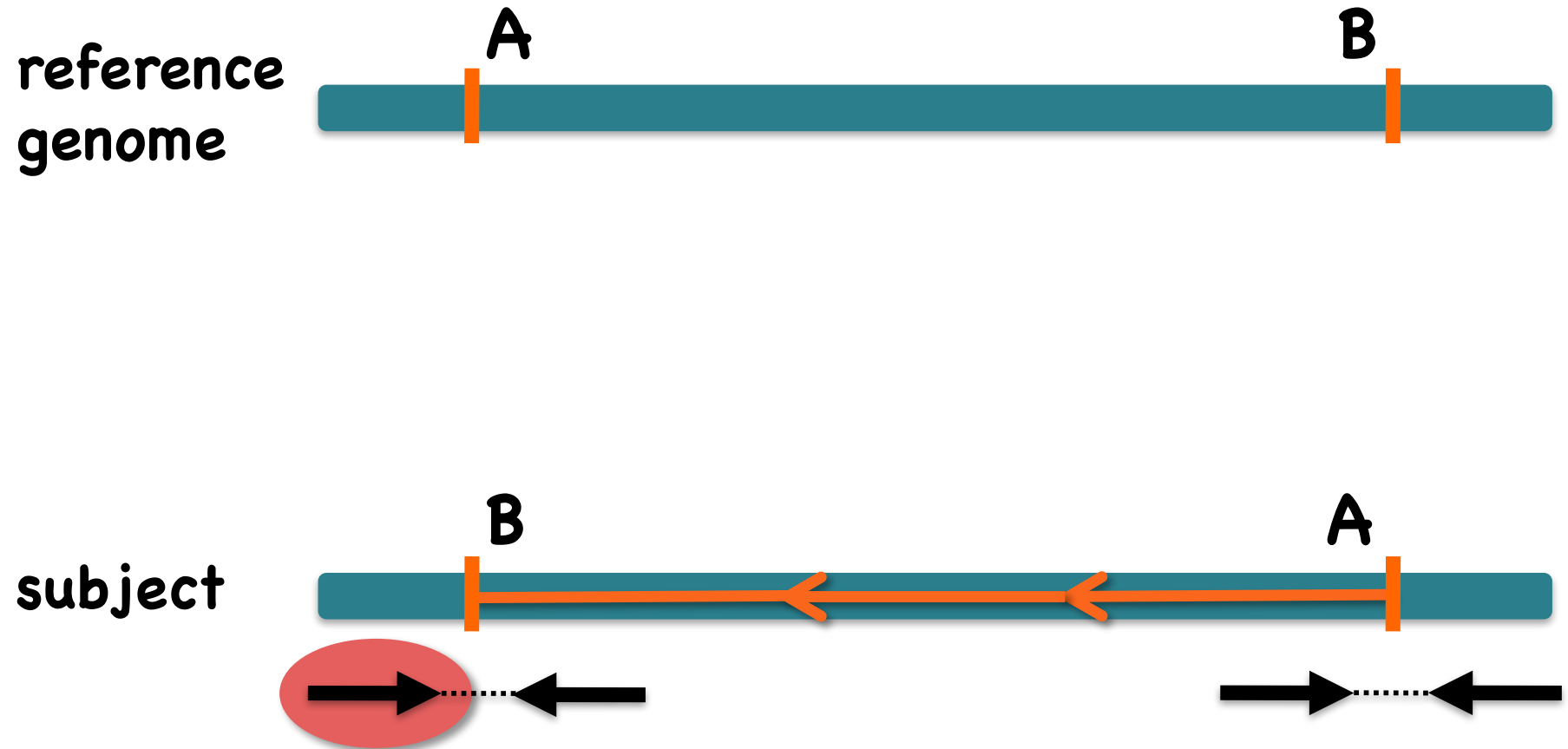
Inversion



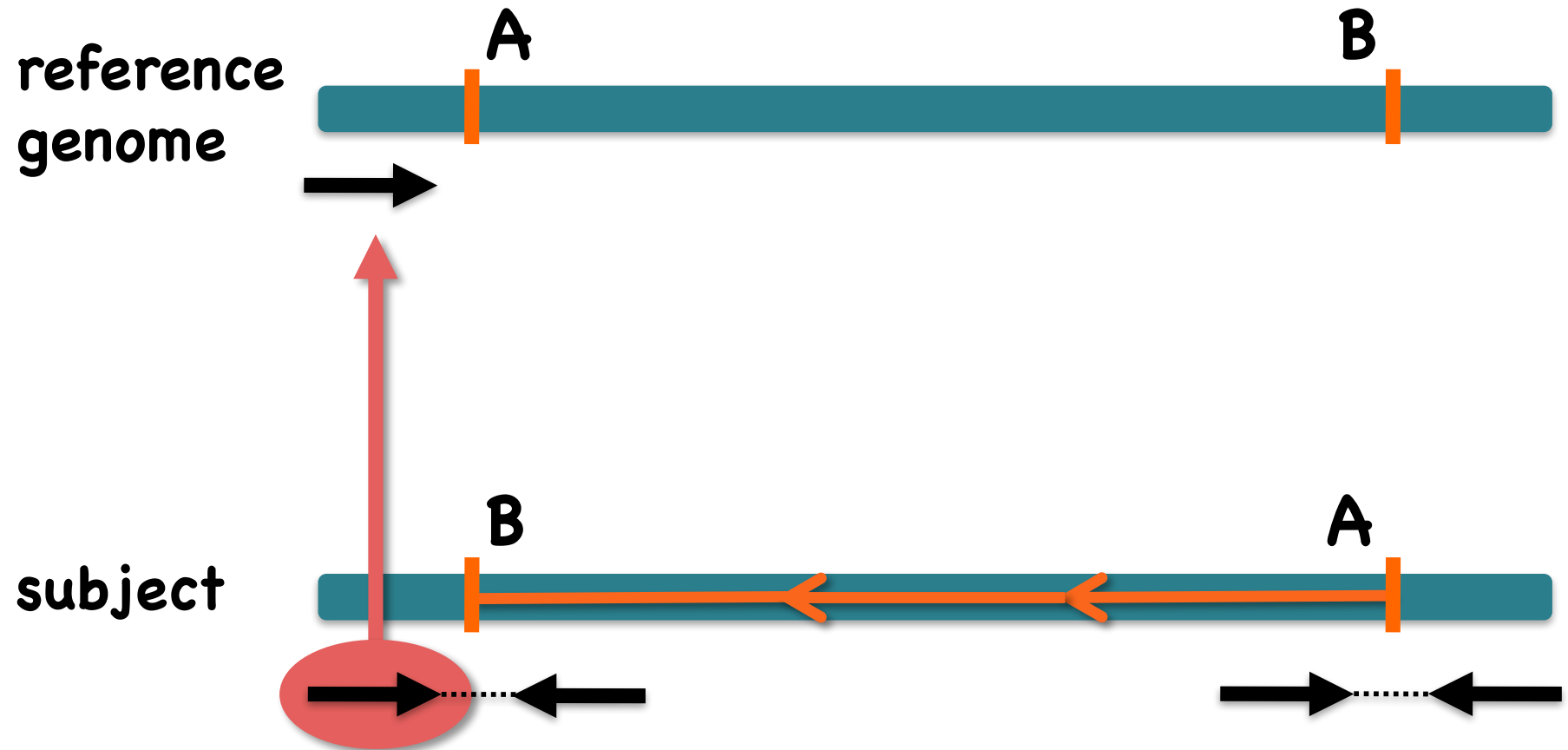
Inversion



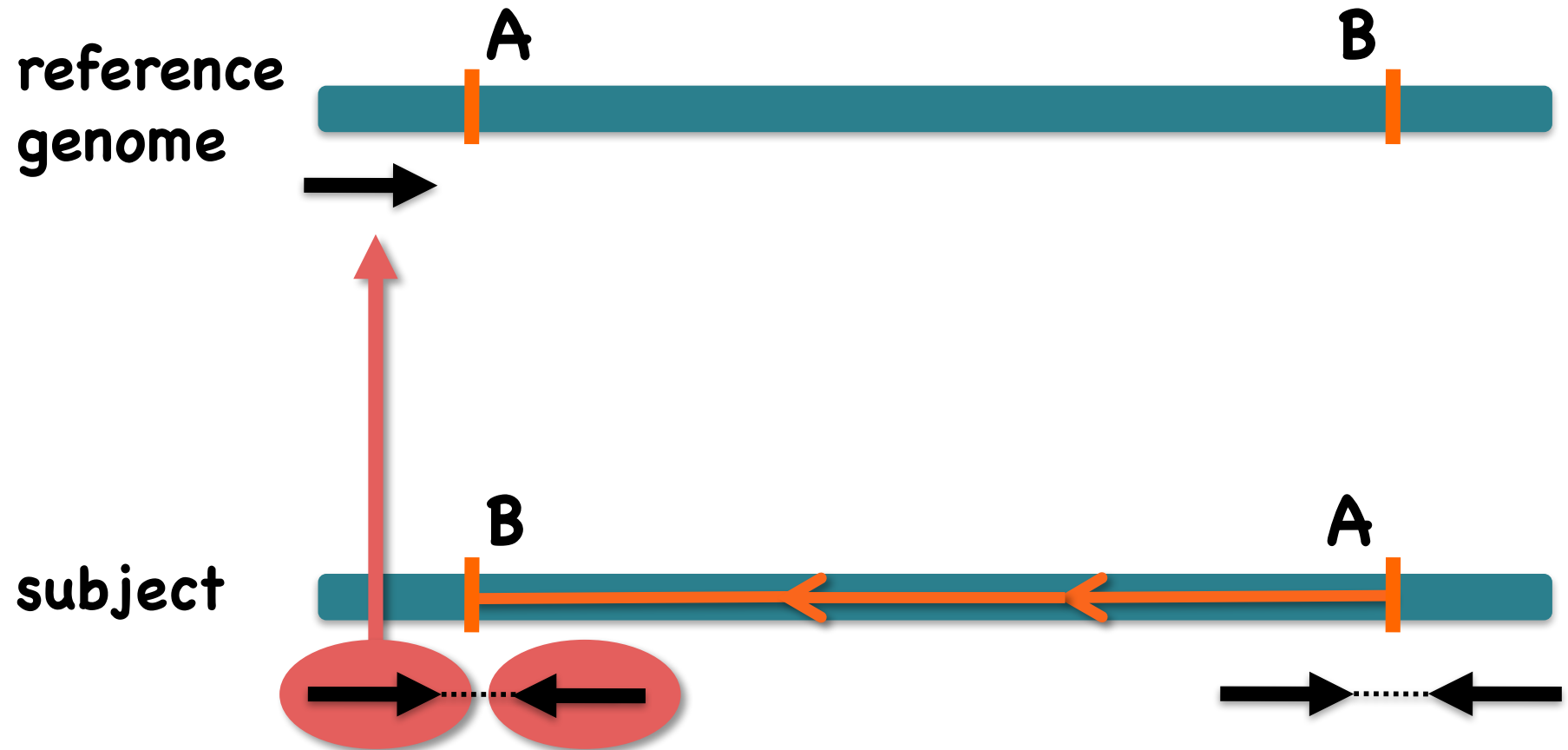
Inversion



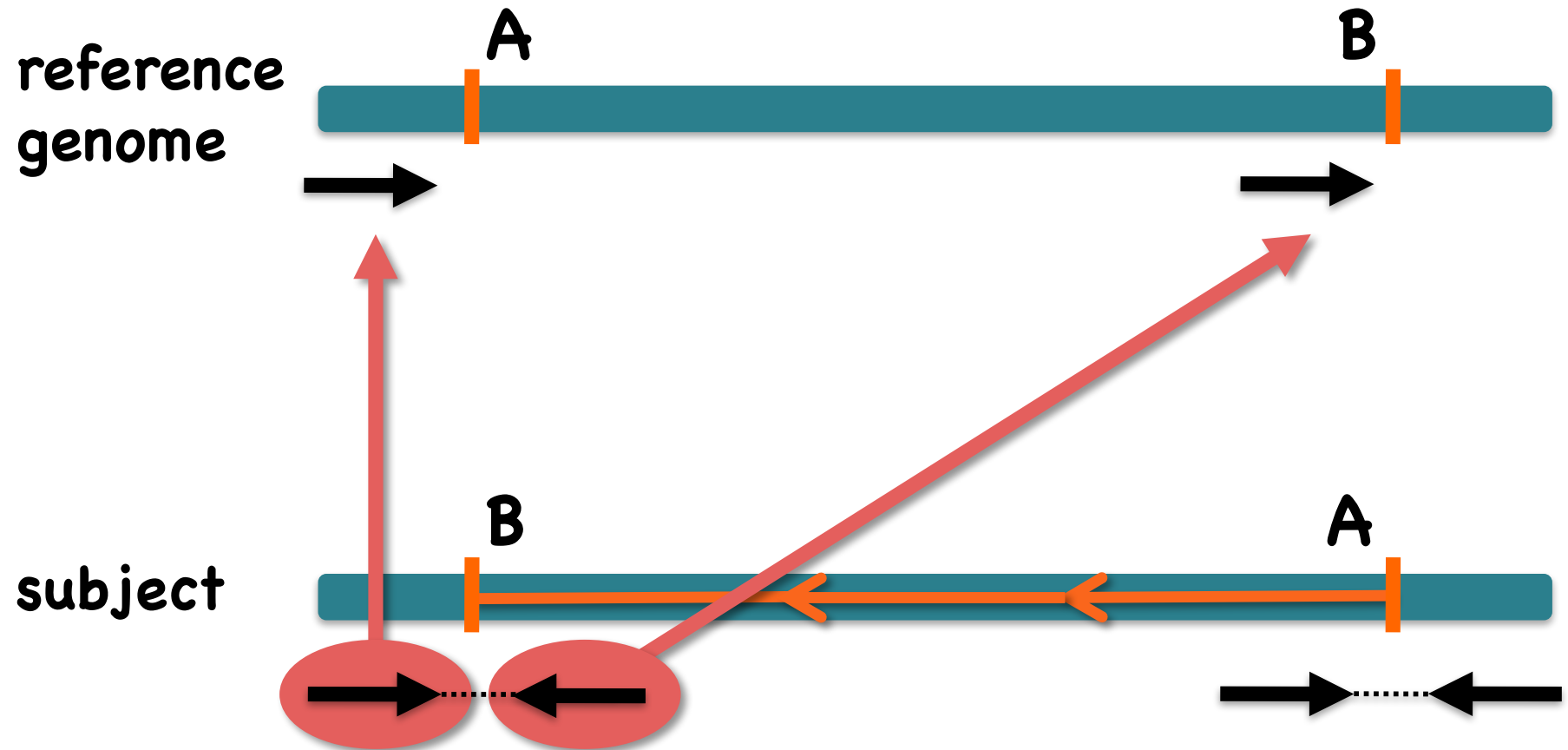
Inversion



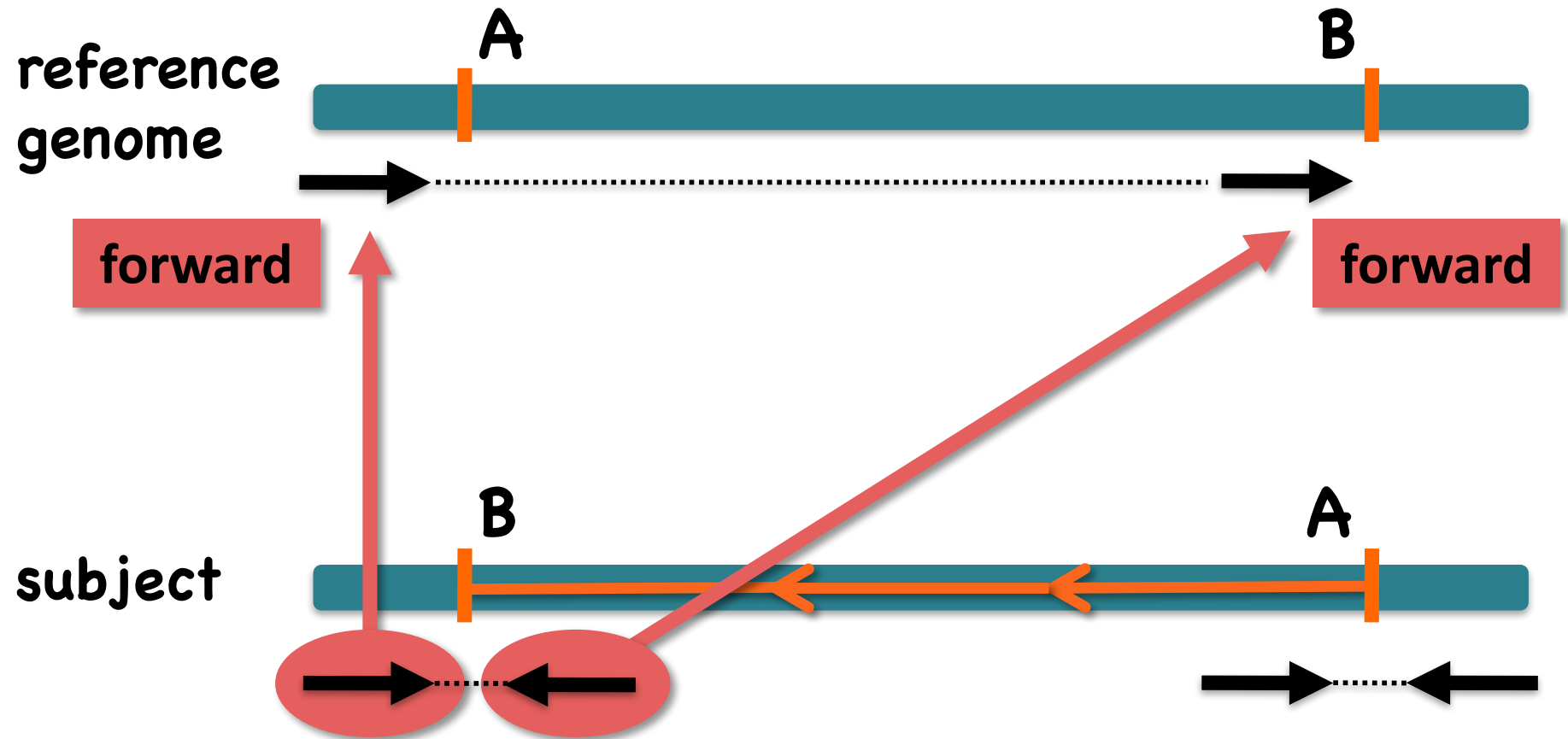
Inversion



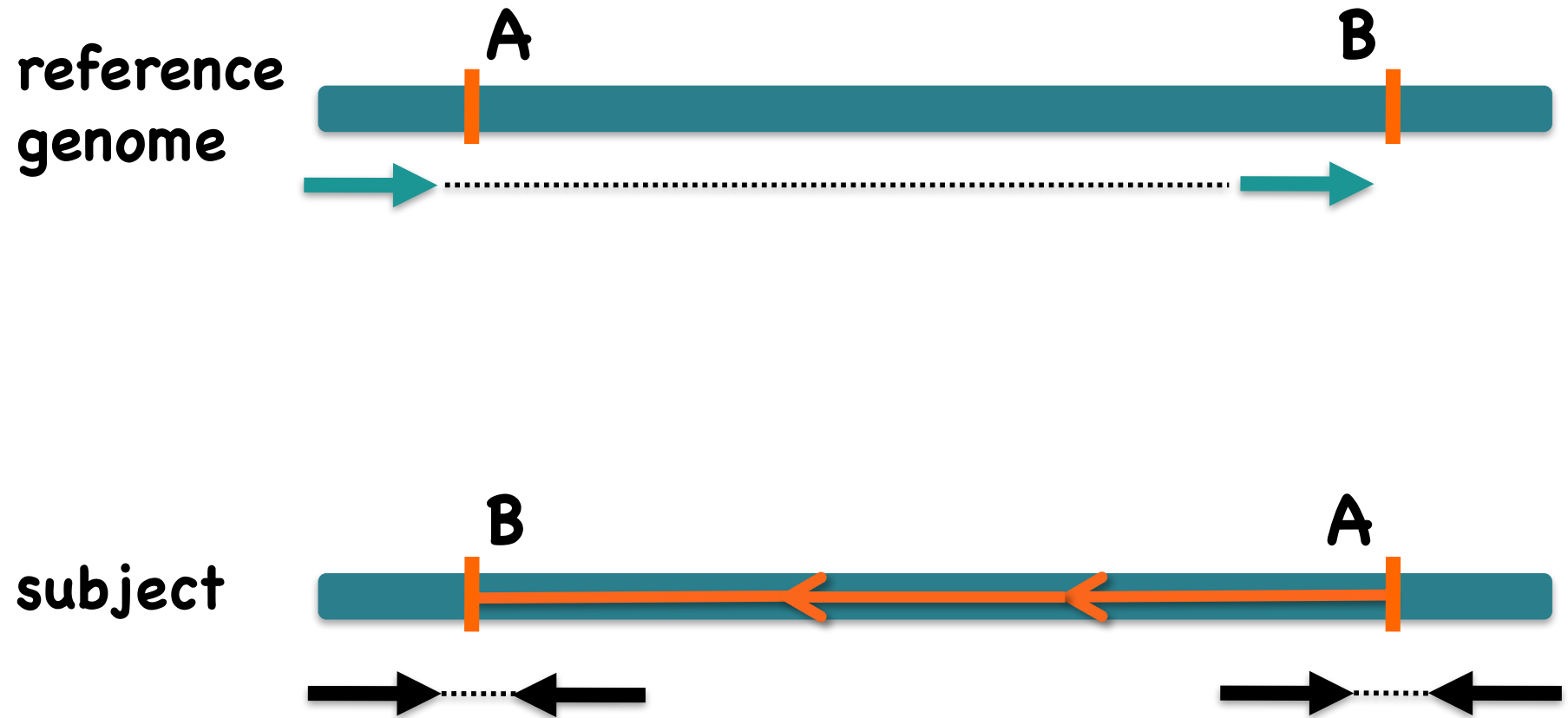
Inversion



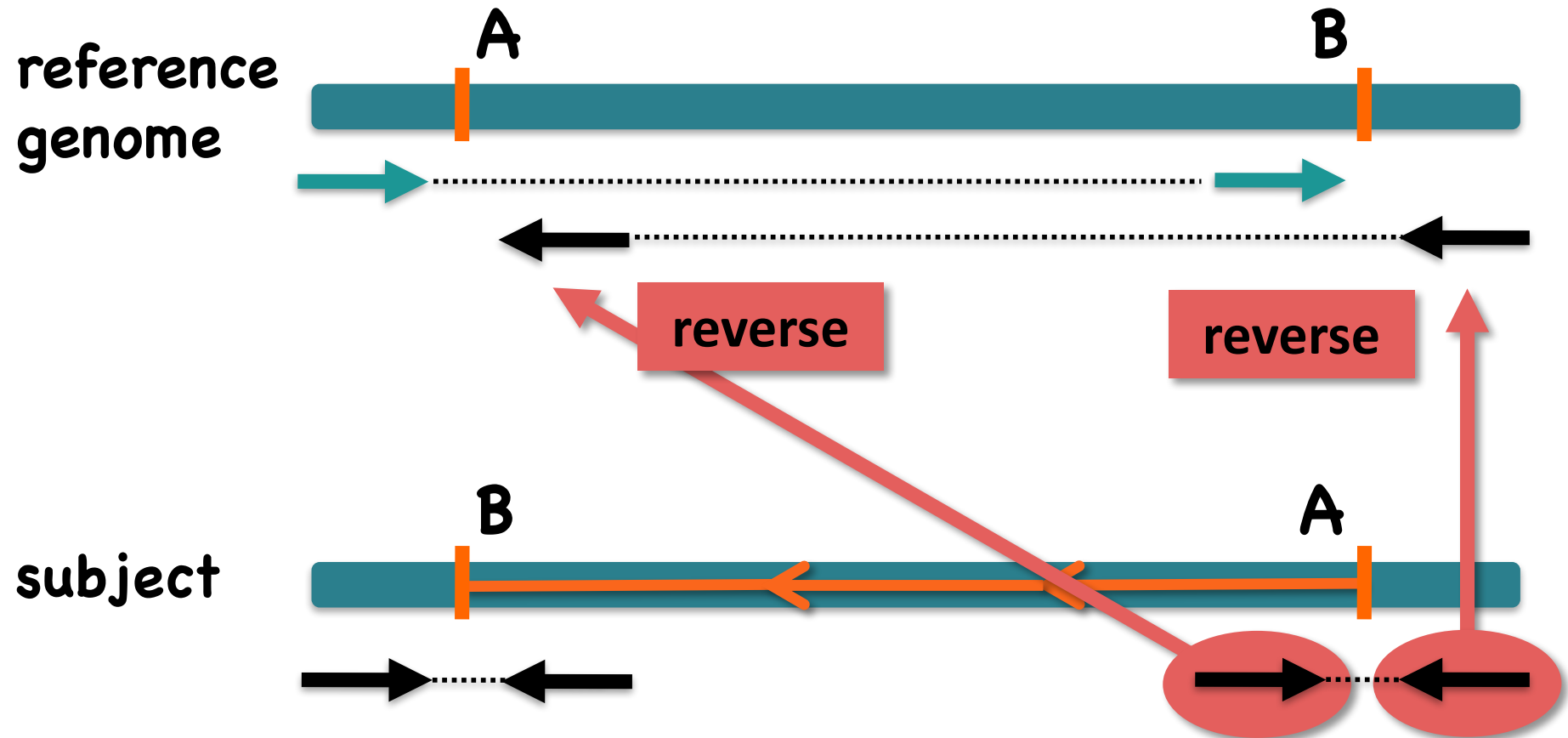
Inversion



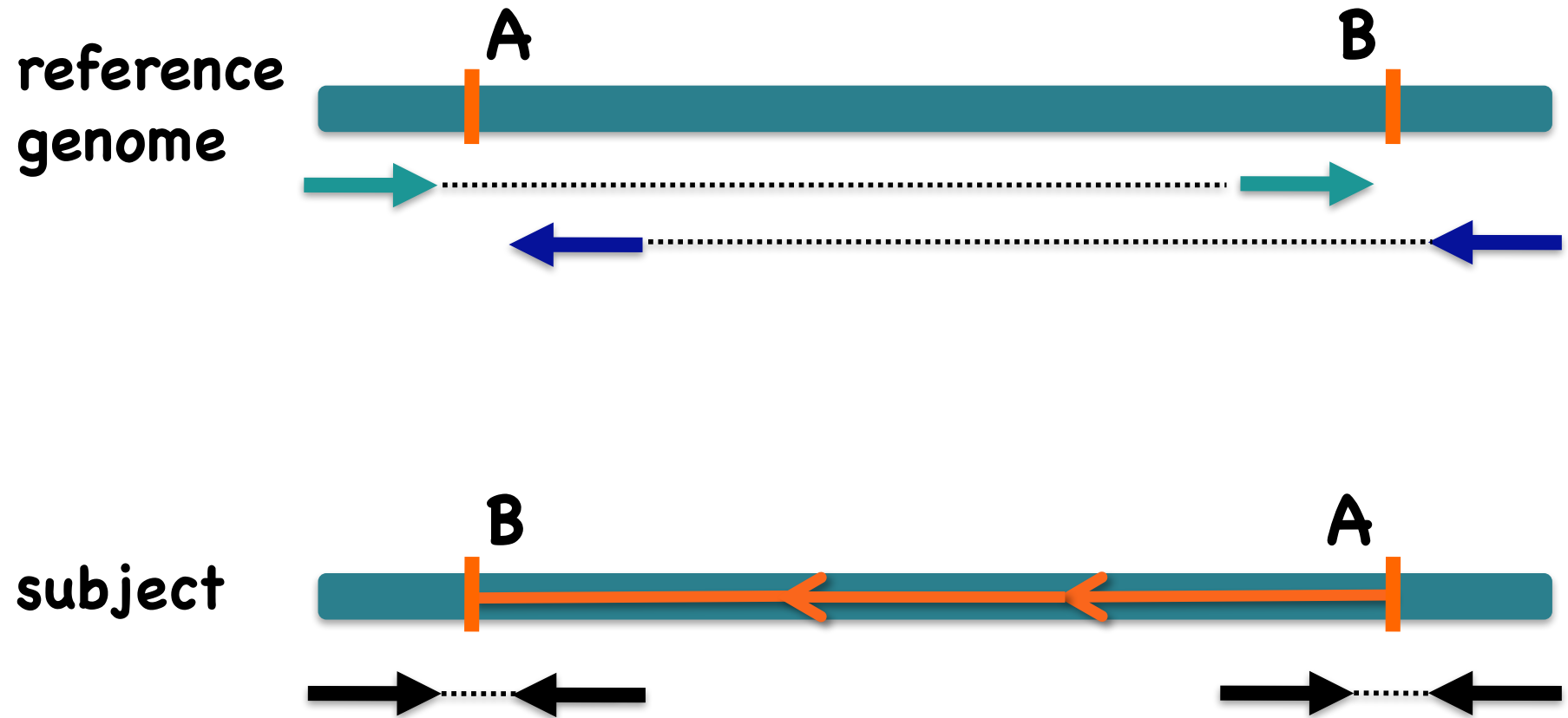
Inversion



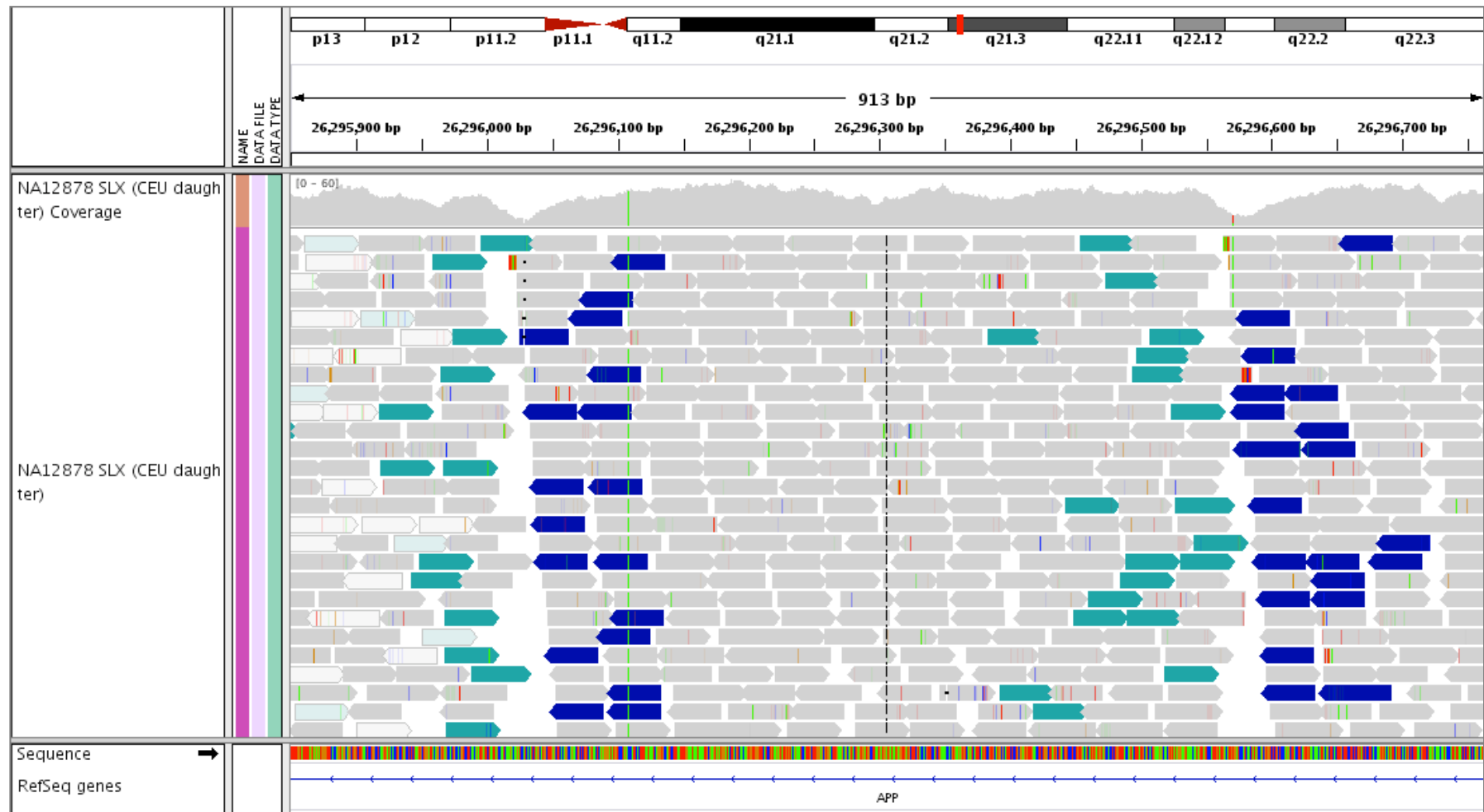
Inversion



Inversion

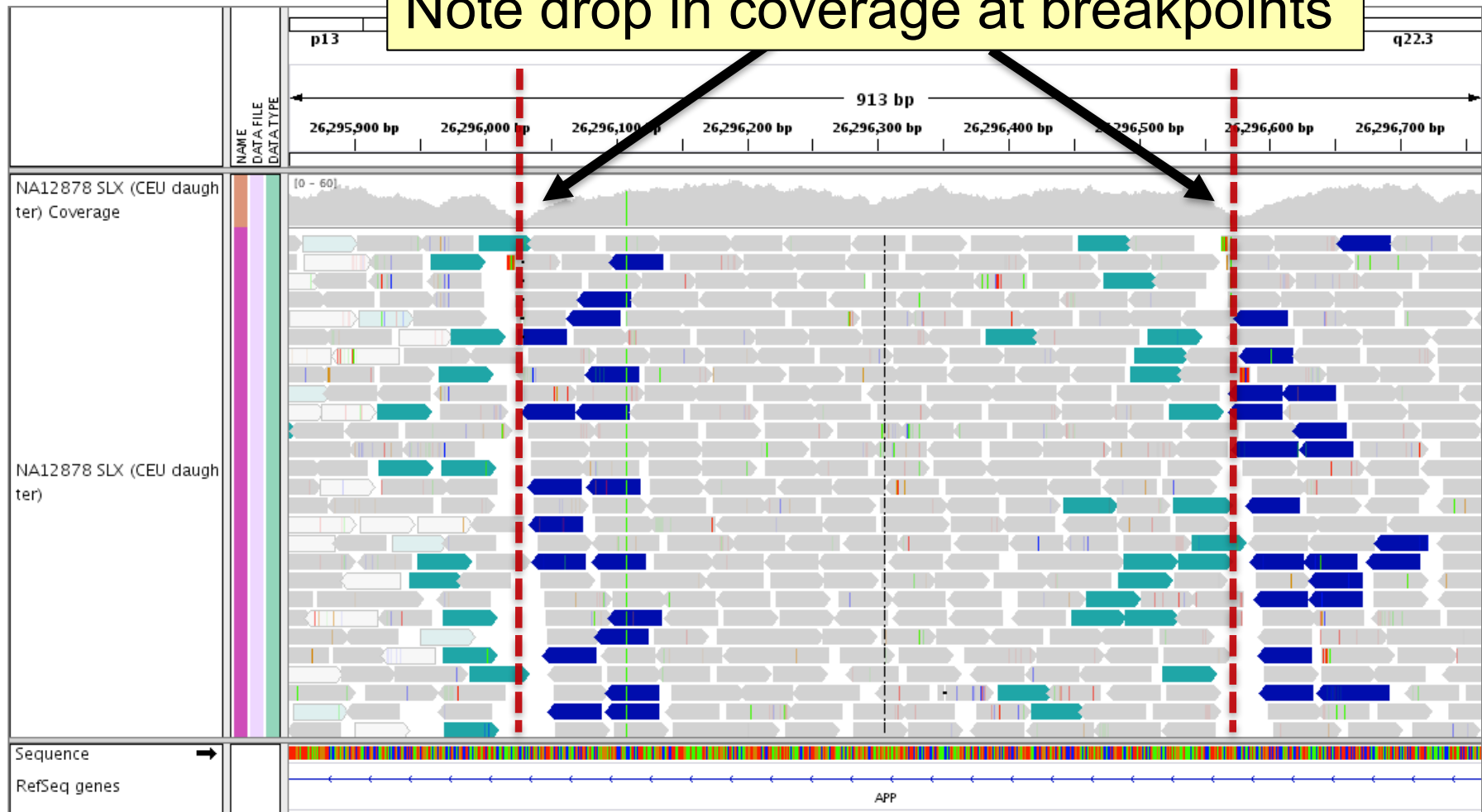


Inversion



Inversion

Note drop in coverage at breakpoints



Viewing structural events

Hands-on exercise

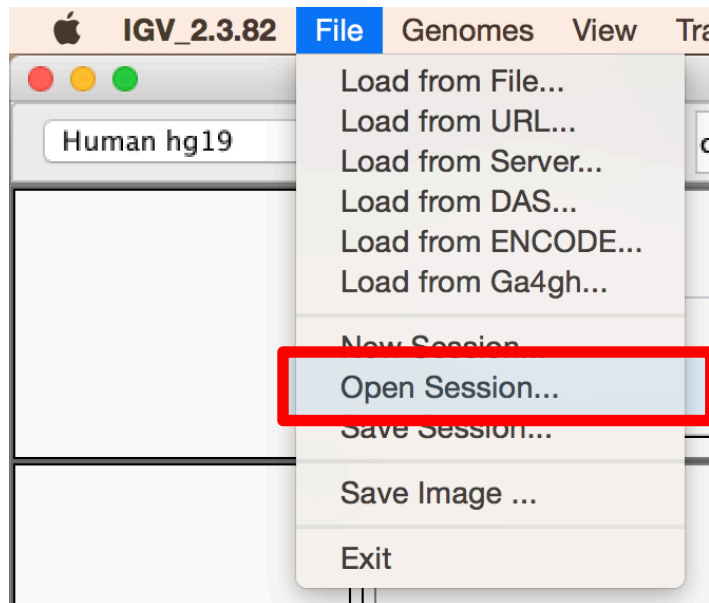
- Use color and viewing options to examine an inversion
- Data from 1000 Genomes

Follow along

Viewing structural events

Hands-on exercise

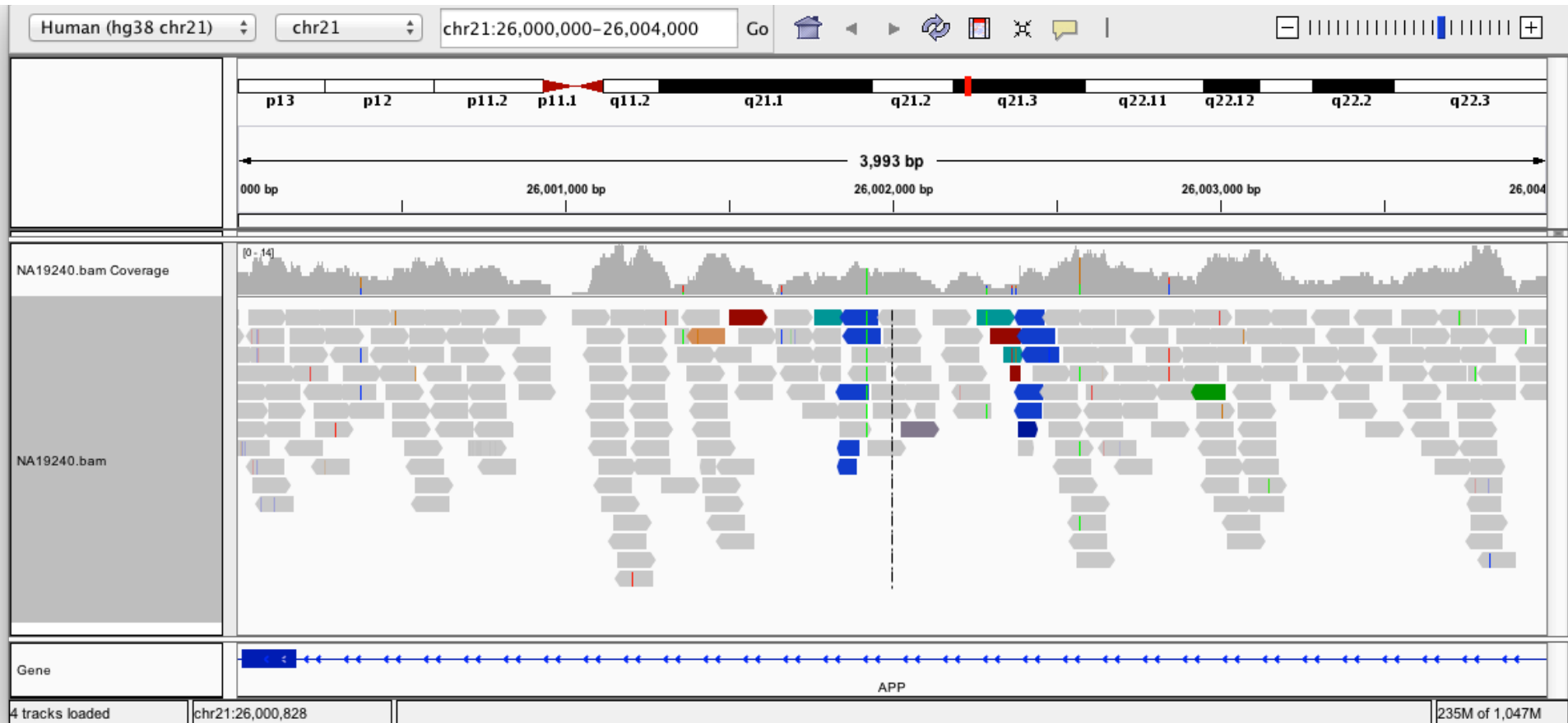
Click **File > Open Session**



Navigate to the folder IGV workshop folder / **Data / sv**s and select file **svs_session.xml**

Viewing structural events

Hands-on exercise



Viewing structural events

Hands-on exercise

Right-click on the alignments and select
Color alignments by > insert size

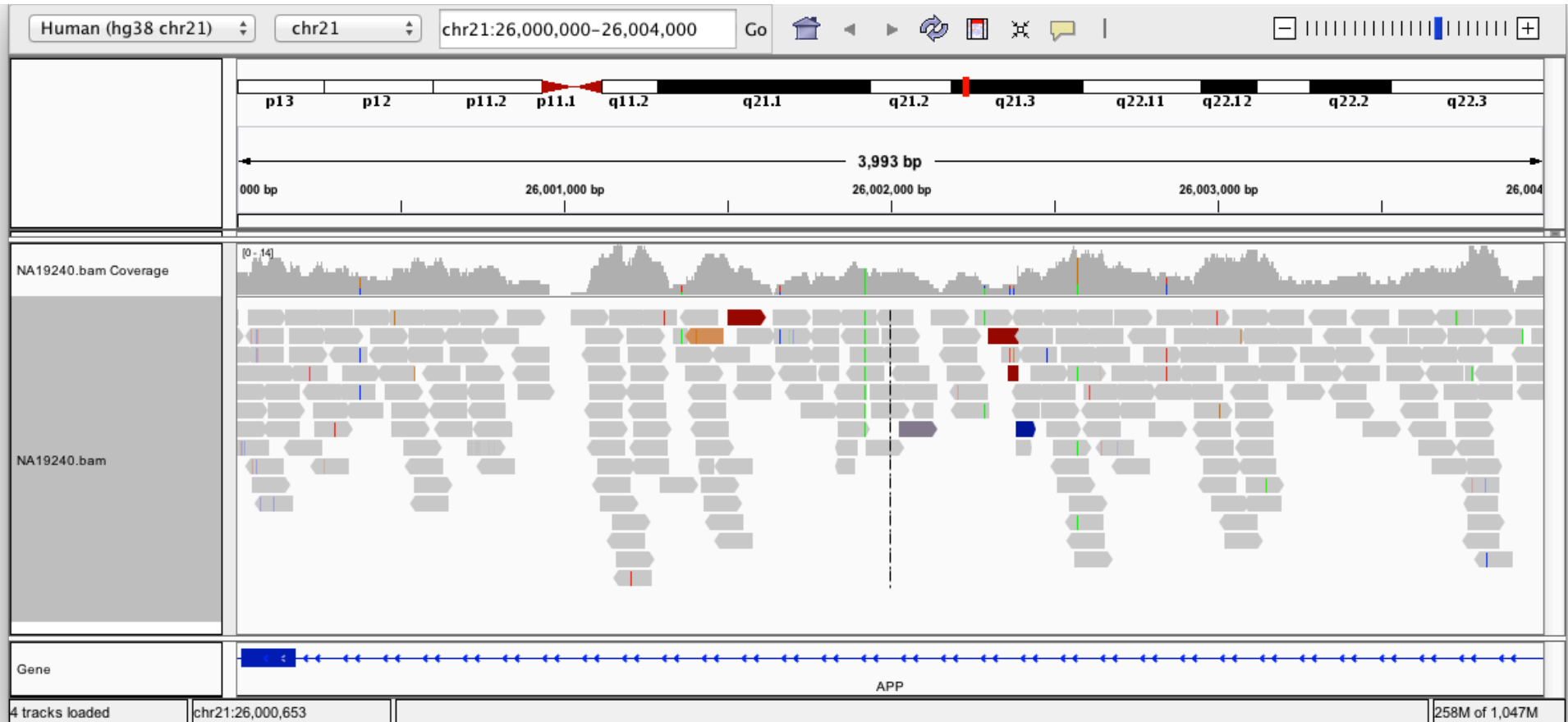
The screenshot displays a genomic browser interface with a right-click context menu open over the alignment tracks. The menu options are as follows:

- Rename Track...
- Copy read details to clipboard
- Group alignments by
- Sort alignments by
- Color alignments by** (selected, leading to a sub-menu)
 - no color
 - insert size** (highlighted)
 - pair orientation
 - ✓ insert size and pair orientation
 - read strand
 - first-of-pair strand
 - read group
 - sample
 - library
 - tag
 - bisulfite mode
- Re-pack alignments
- ✓ Shade base by quality
- ✓ Show mismatched bases
- Show all bases
- View as pairs
- Go to mate
- View mate region in split screen
- Set insert size options ...
- Collapsed
- ✓ Expanded
- Squished
- Select by name...
- Clear selections
- Copy read sequence
- Blat read sequence
- Copy consensus sequence
- Sashimi Plot
- ✓ Show Coverage Track
- Show Splice Junction Track
- Hide Track
- Save image...
- Export Alignments...
- Export track names...
- Remove Track

The background interface shows a genomic track with coverage and alignments. The left sidebar lists tracks: NA19240.bam Coverage and NA19240.bam. The bottom status bar indicates 4 tracks loaded and the current position is chr21:26,000,828.

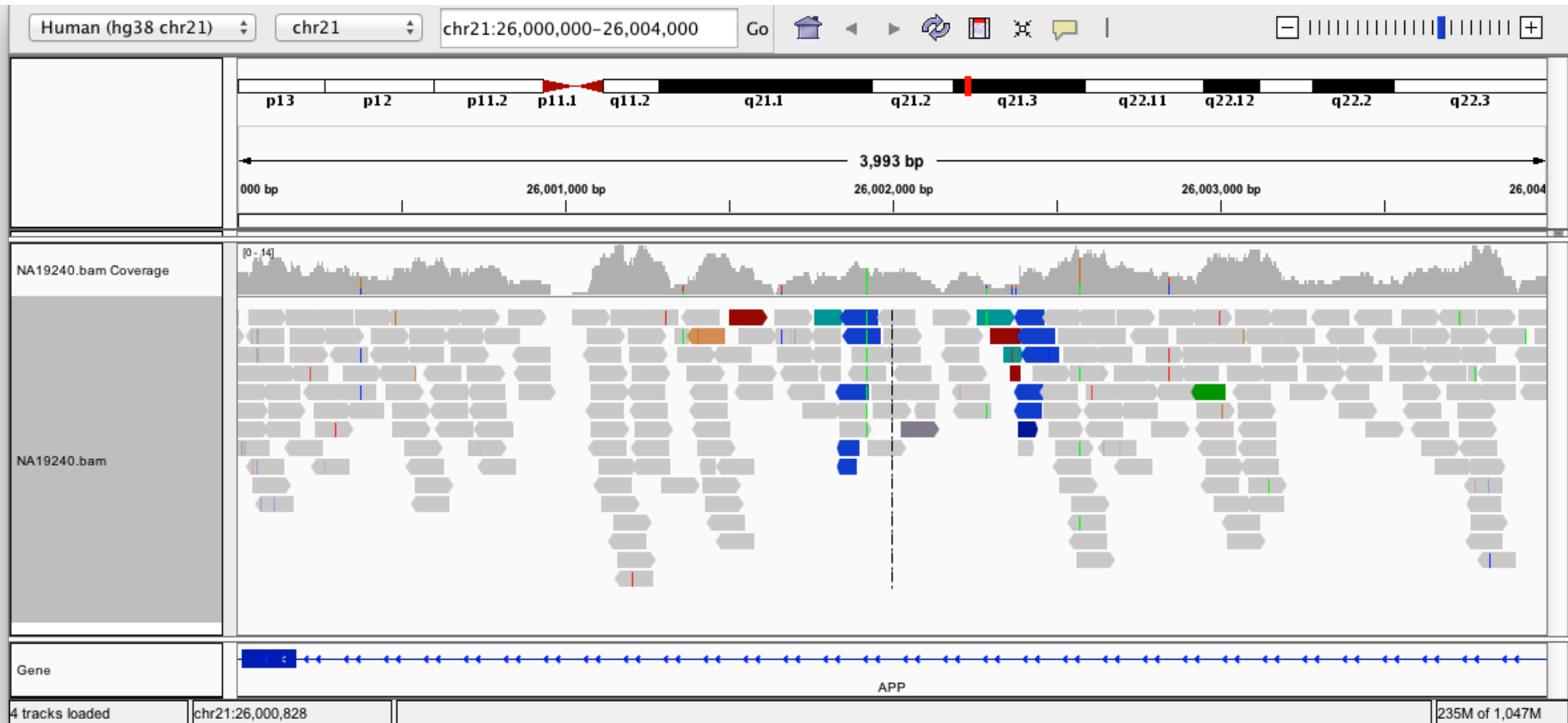
Viewing structural events

Hands-on exercise



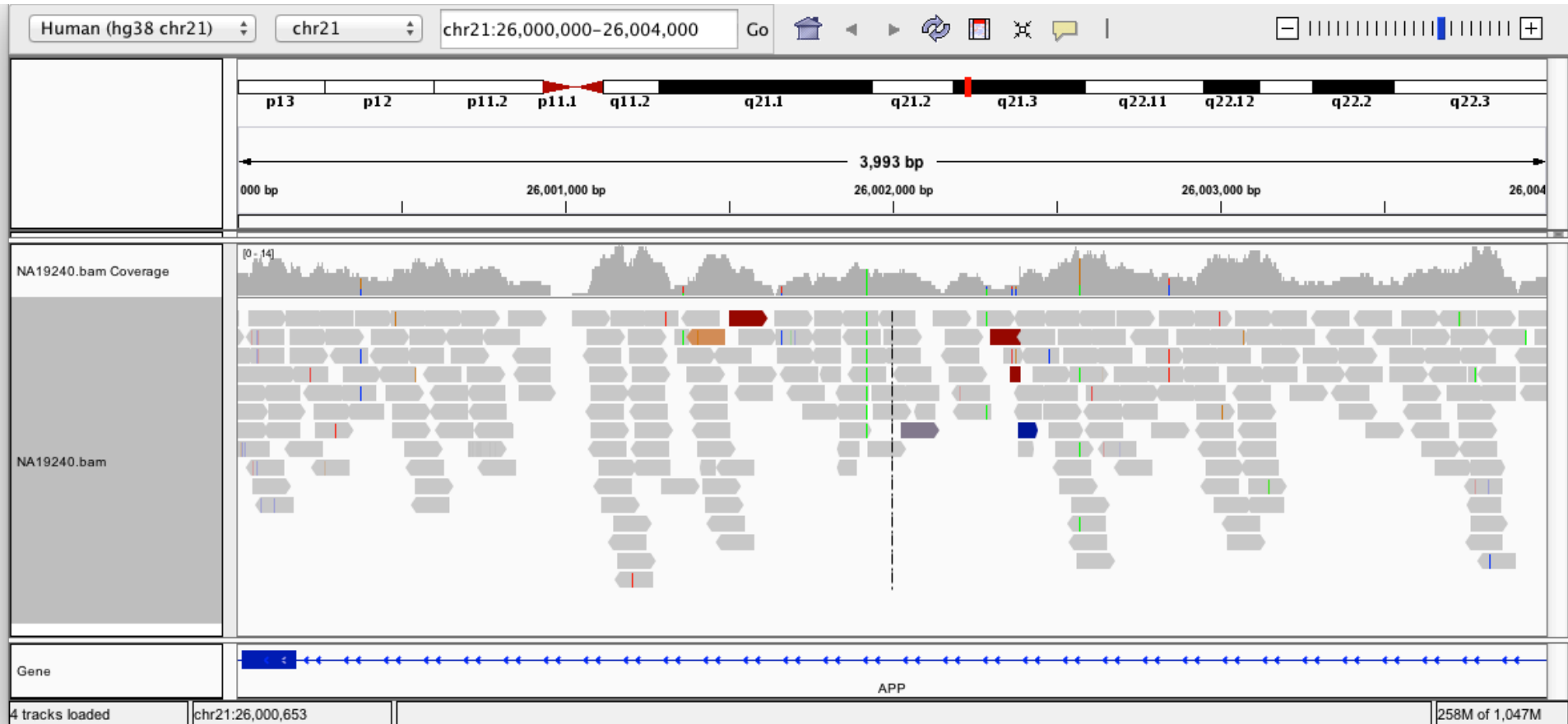
Viewing structural events

Hands-on exercise



Viewing structural events

Hands-on exercise



Viewing structural events

Hands-on exercise

Right-click on the alignments and select
Color alignments by > pair orientation

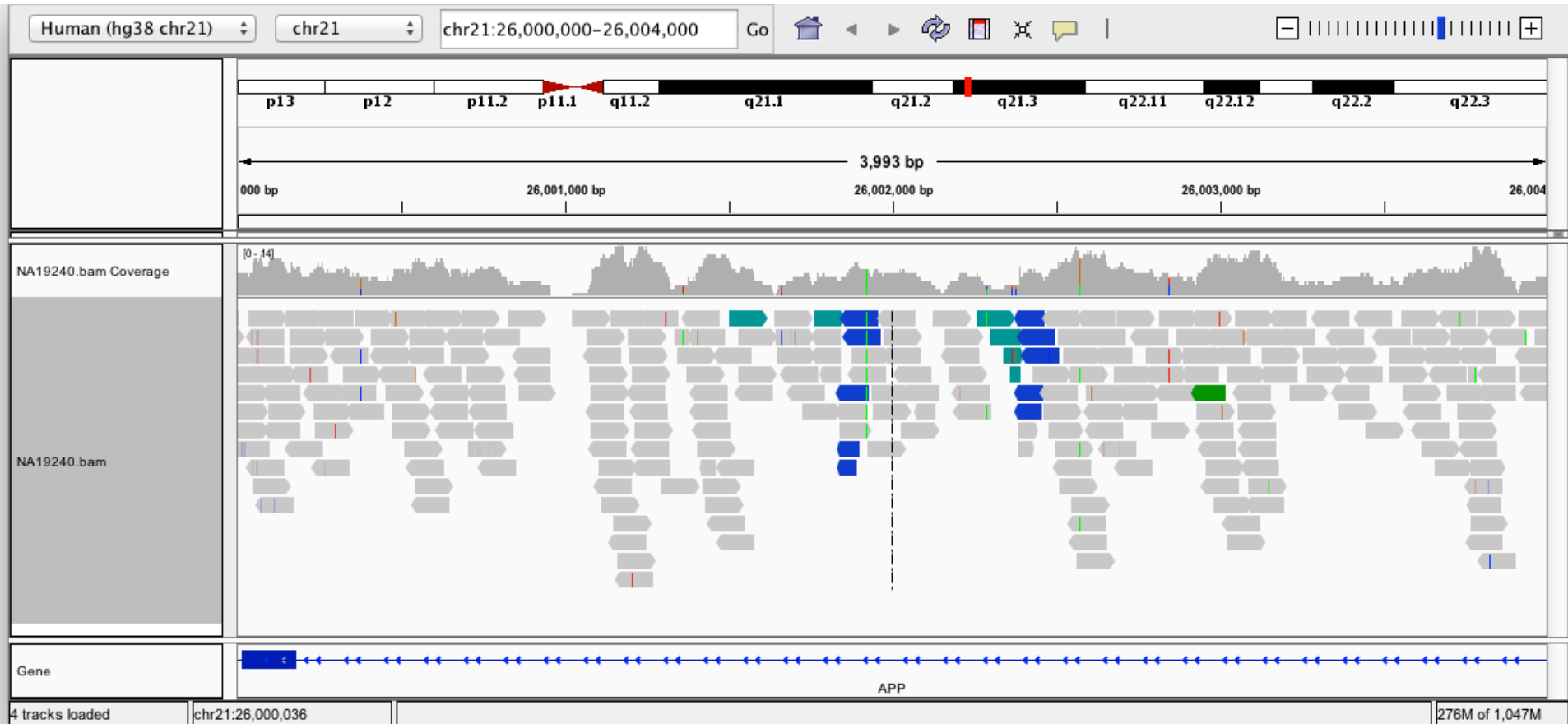
The screenshot displays a genomic browser interface with a right-click context menu open over a track of alignments. The menu options are as follows:

- Rename Track...
- Copy read details to clipboard
- Group alignments by ▶
- Sort alignments by ▶
- Color alignments by ▶**
 - no color
 - ✓ insert size
 - pair orientation**
 - insert size and pair orientation
 - read strand
 - first-of-pair strand
 - read group
 - sample
 - library
 - tag
 - bisulfite mode
- Re-pack alignments
- ✓ Shade base by quality
- ✓ Show mismatched bases
- Show all bases
- View as pairs
- Go to mate
- View mate region in split screen
- Set insert size options ...
- Collapsed
- ✓ Expanded
- Squished
- Select by name...
- Clear selections
- Copy read sequence
- Blat read sequence
- Copy consensus sequence
- Sashimi Plot
- ✓ Show Coverage Track
- Show Splice Junction Track
- Hide Track
- Save image...
- Export Alignments...
- Export track names...
- Remove Track

The background interface includes tracks for 'NA19240.bam Coverage', 'NA19240.bam', and 'Gene'. The 'Gene' track shows a blue box representing a gene model. The bottom status bar indicates '4 tracks loaded' and 'chr21:26,000,653'.

Viewing structural events

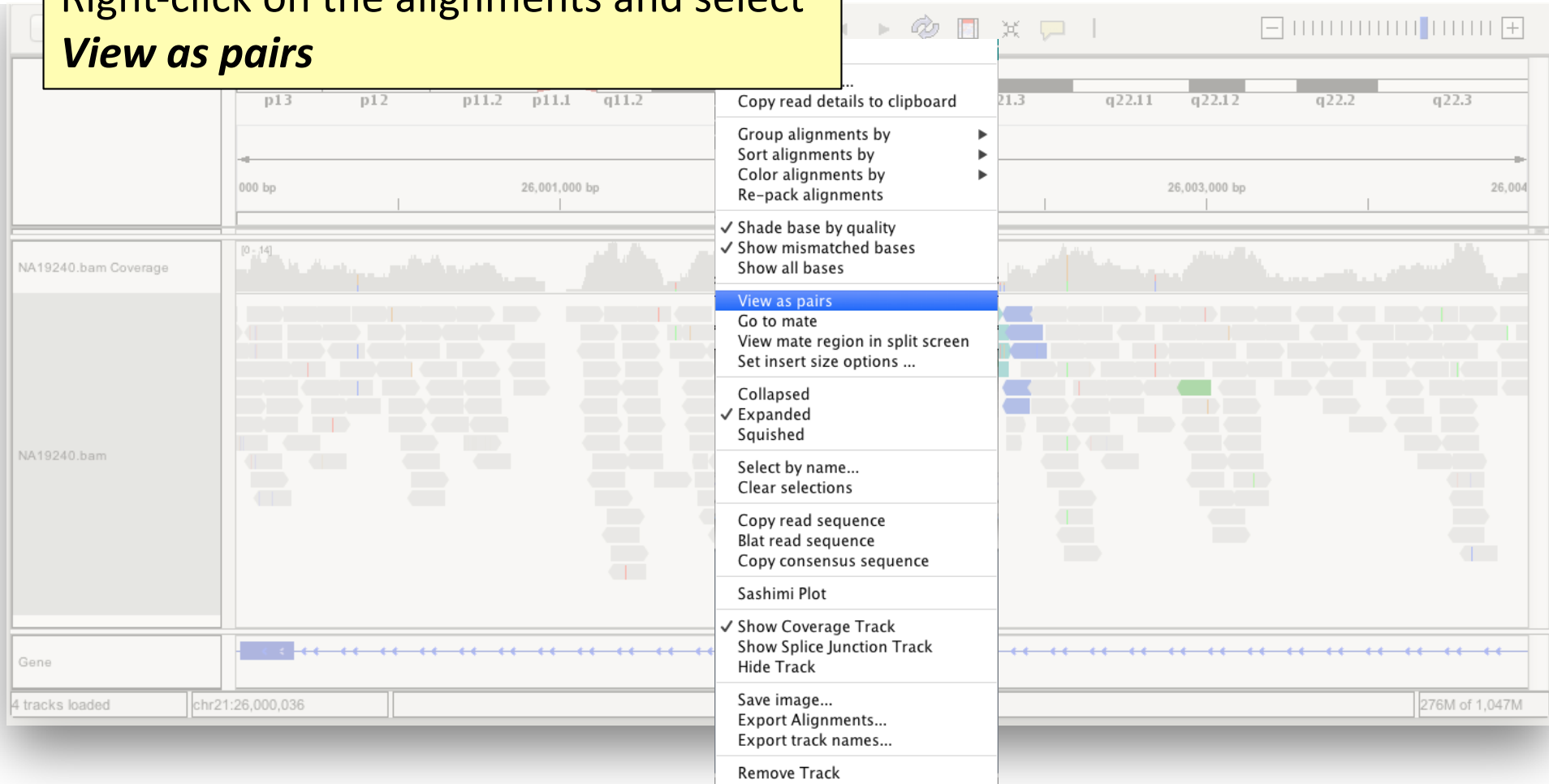
Hands-on exercise



Viewing structural events

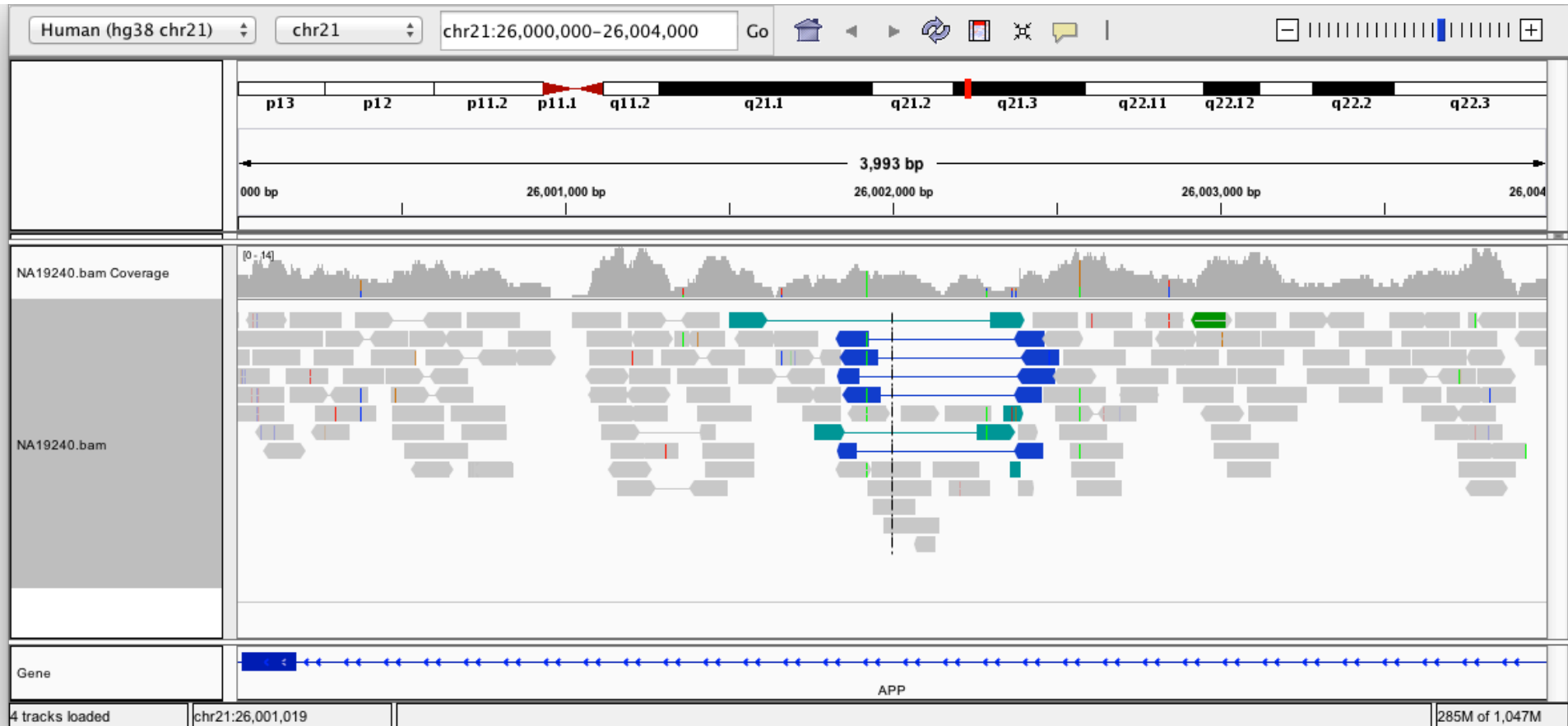
Hands-on exercise

Right-click on the alignments and select
View as pairs



Viewing structural events

Hands-on exercise



Viewing Structural Events



Availability

Features available in **IGV 3.0 Beta**

See **Other Versions** on the IGV Downloads page

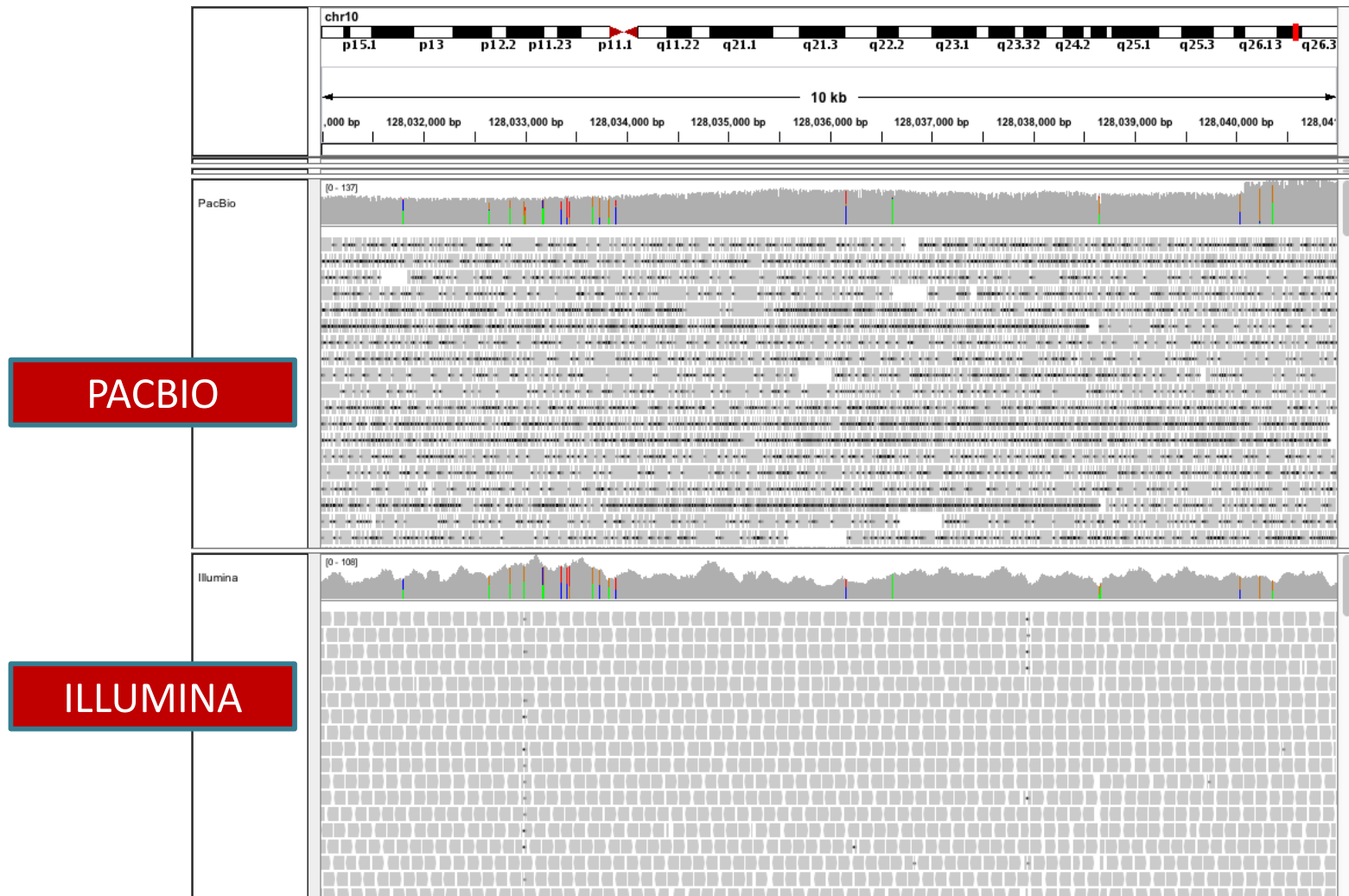
Third Gen Sequencing

- Long reads:
~10,000 bp up to ~100,00 bp
- High raw error rate in individual reads
(high consensus accuracy with sufficient coverage)
- Platforms:
 - PacBio, Oxford Nanopore

IGV Enhancements

- Filtering raw errors to reveal consensus variants
- Extended support for split alignments
 - Linking split reads
 - Support for “SA” tag

Filtering raw errors – small indels

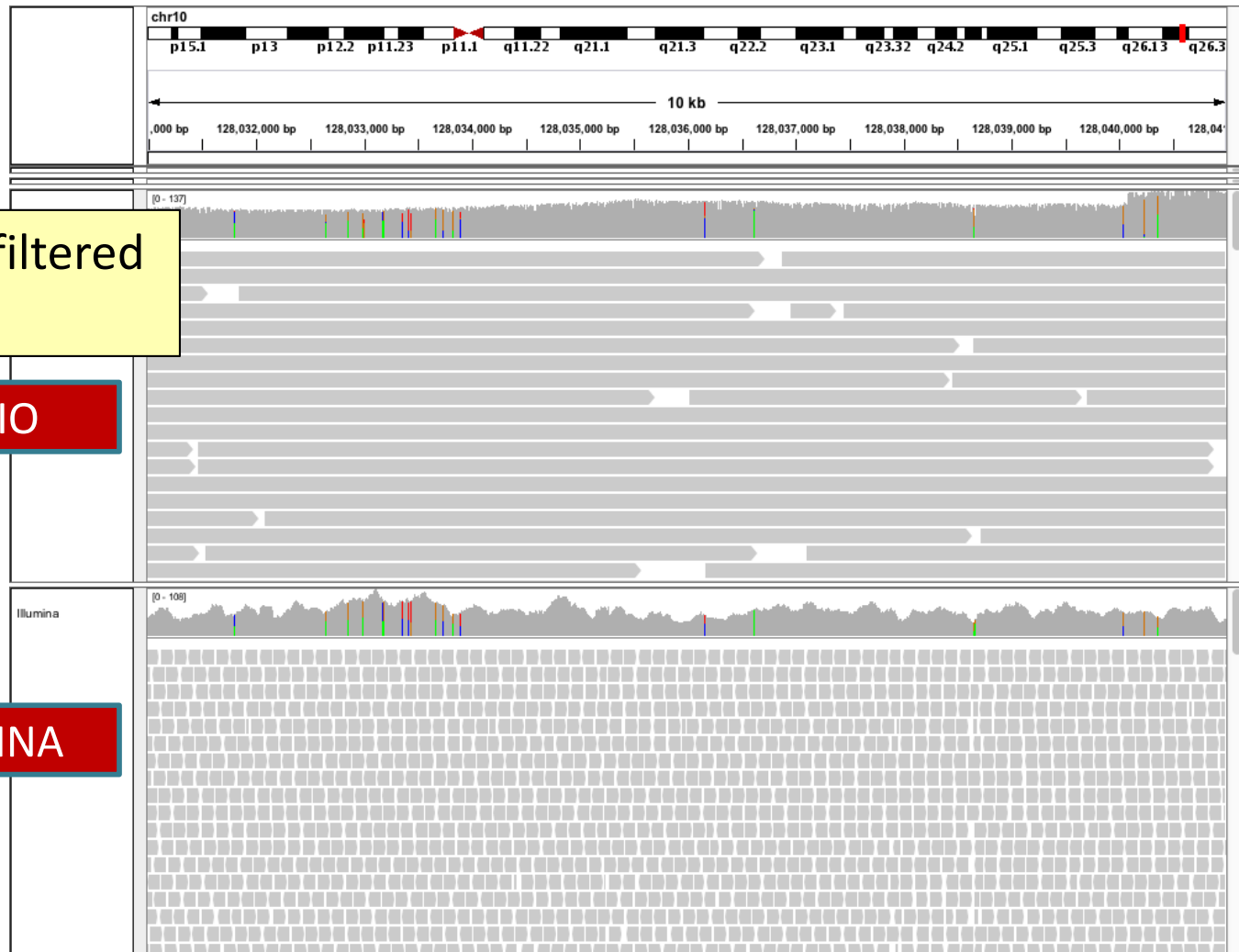


Filtering raw errors - indels

Small indels filtered
from view

PACBIO

ILLUMINA

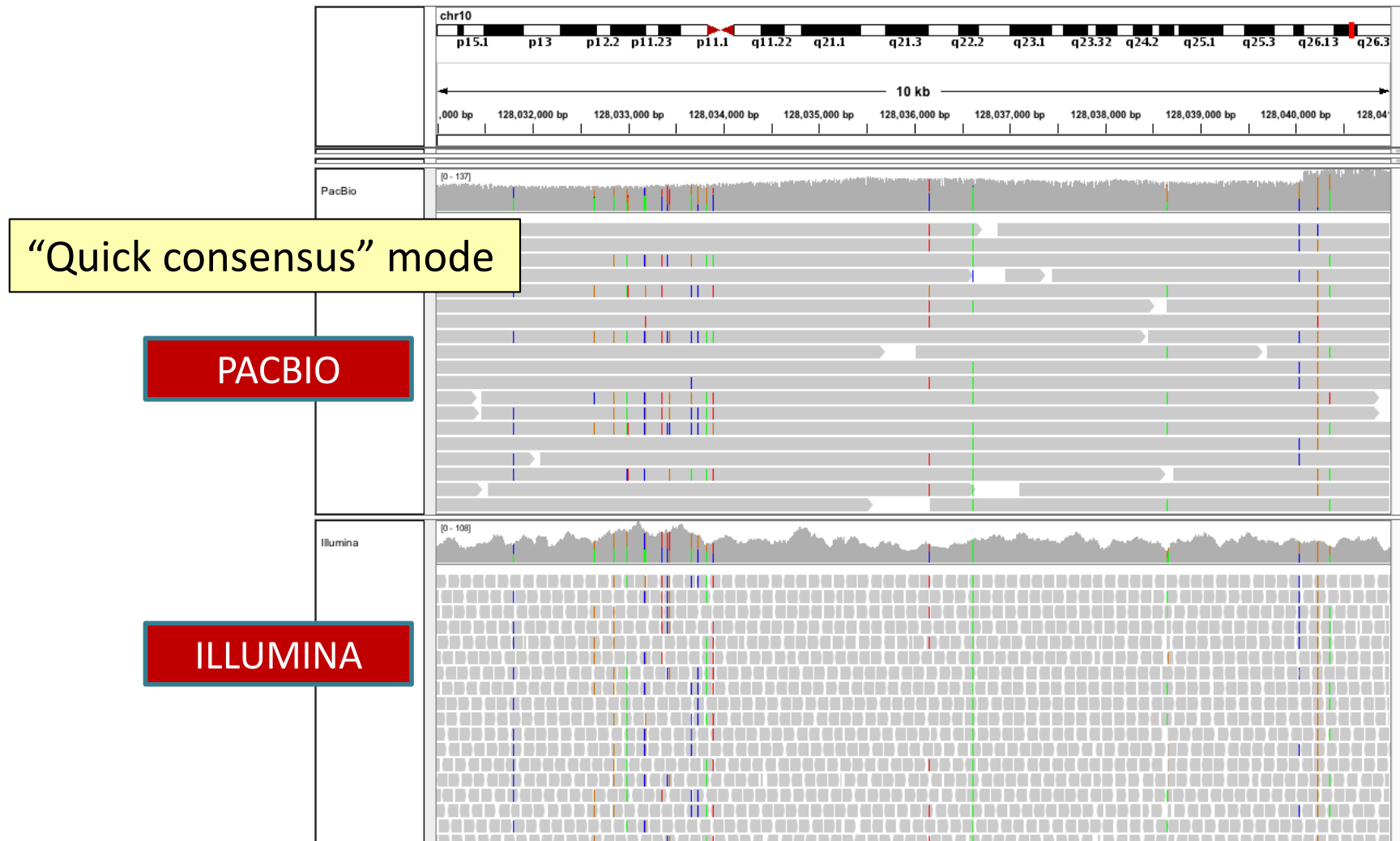


Aaron Wenger, PacBio

Filtering raw errors - nucleotides



Filtering raw errors - nucleotides

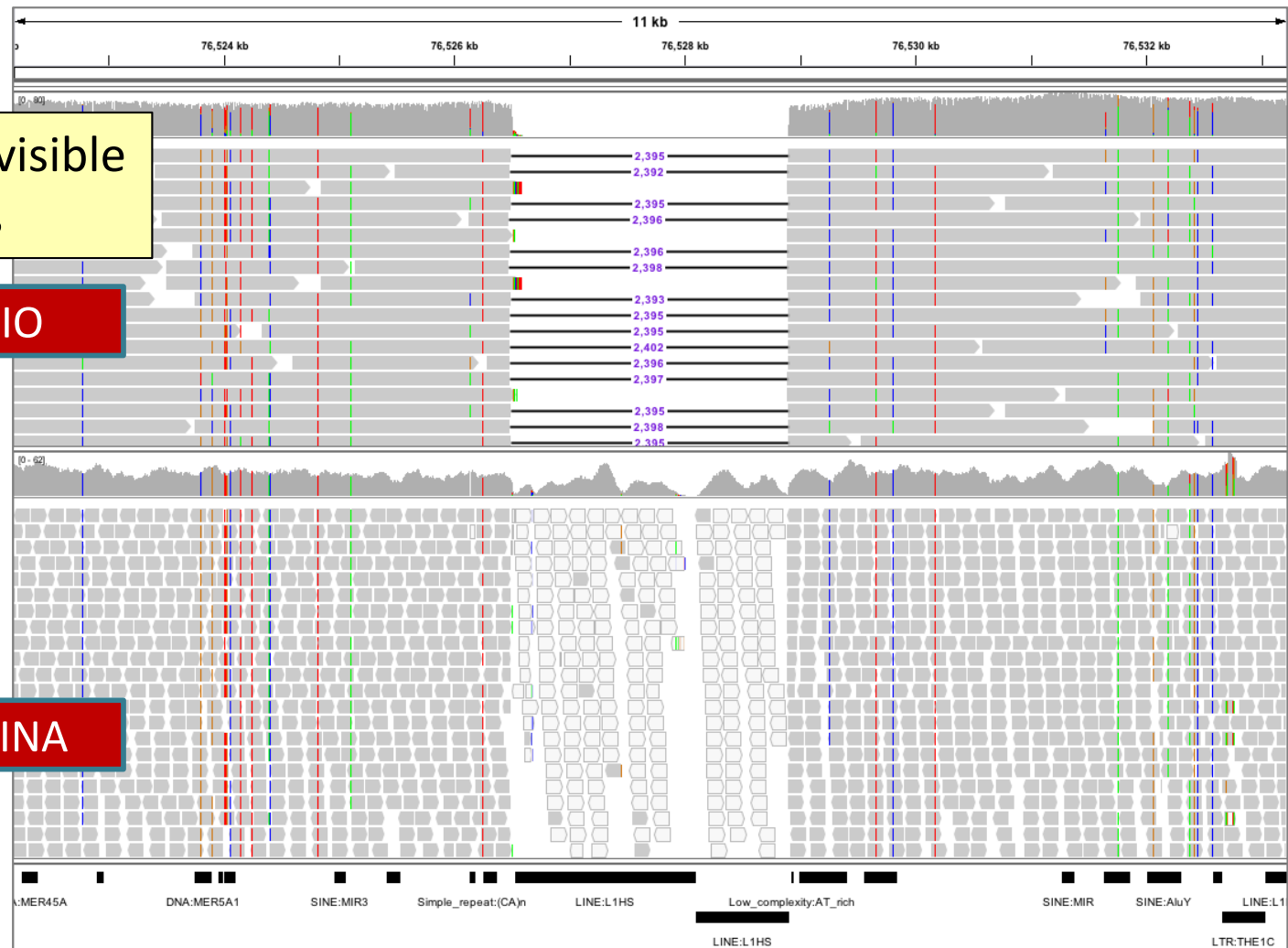


Structural variant example - deletion

Deletion clearly visible
in the long reads

PACBIO

ILLUMINA



Aaron Wenger, PacBio

View > Preferences

Alignments tab

Purple insertion labels

Hide small indels

Only show consistent mismatches

General Tracks Variants Charts **Alignments** Probes Proxy IonTorrent Advanced

Track Display Options

On initial load show: ☒ Alignment Track ☒ Coverage Track ☐ Splice Junction Track

Alignment Track Options

Visibility range threshold (kb): 500 *Range at which alignments become visible*

☒ Downsample reads Max read count: 2000 per window size (bases): 50

☒ Shade mismatched bases by quality: 5 to 20

Mapping quality threshold: 20

☒ Label indels > 5 bases

☒ Hide indels < 50 bases

☐ Hide indels < 1 pixels

☐ Filter duplicate reads

☒ Filter vendor failed reads

☒ Filter secondary alignments

☐ Show center line

Hidden SAM tags: MD,SA,XA

☐ Flag unmapped pairs

☐ Show soft-clipped bases

☒ Quick consensus mode

☒ Filter supplementary alignments

Coverage Track Options

Coverage allele-fraction threshold: 0.25 ☐ Quality weight allele fraction

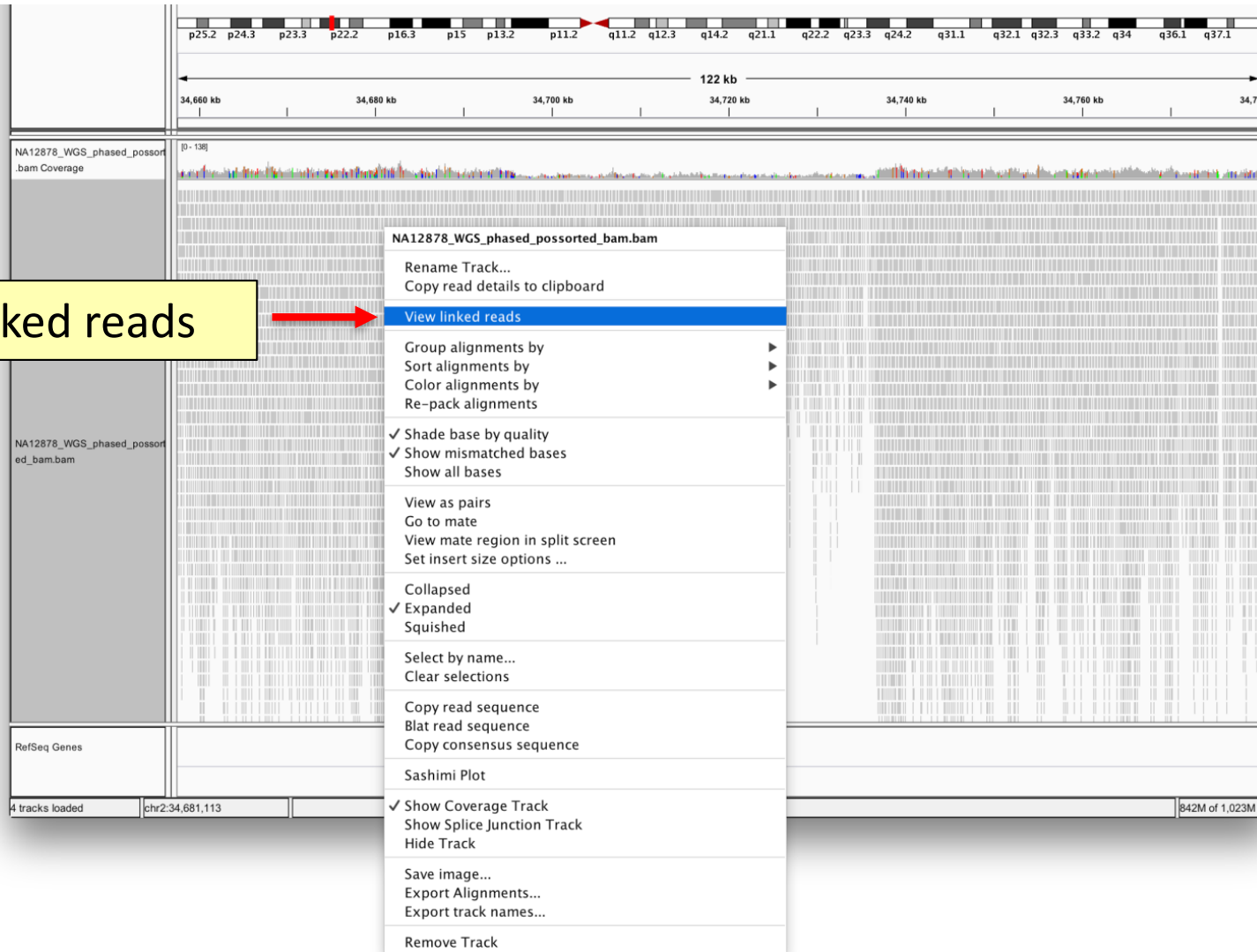
Splice Junction Track Options

10X

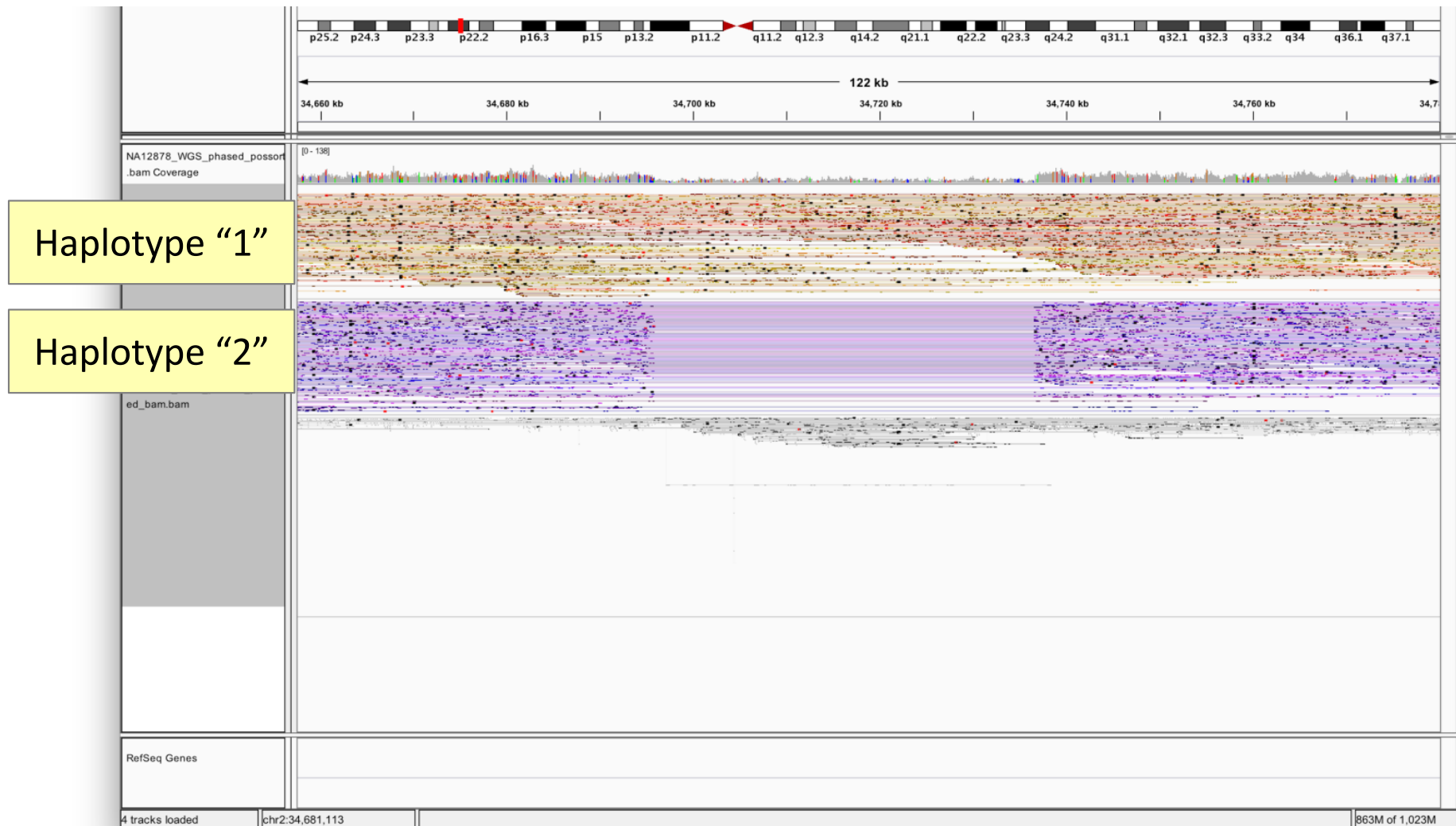
- Long molecules are separated into compartments.
All molecules in same compartment share a barcode.
- Barcodes can be used to reconstruct molecules from short read alignments (Illumina)

IGV enhancements for 10X

View linked reads



IGV enhancements for 10X



Viewing Variants

Variant Call Format (VCF):

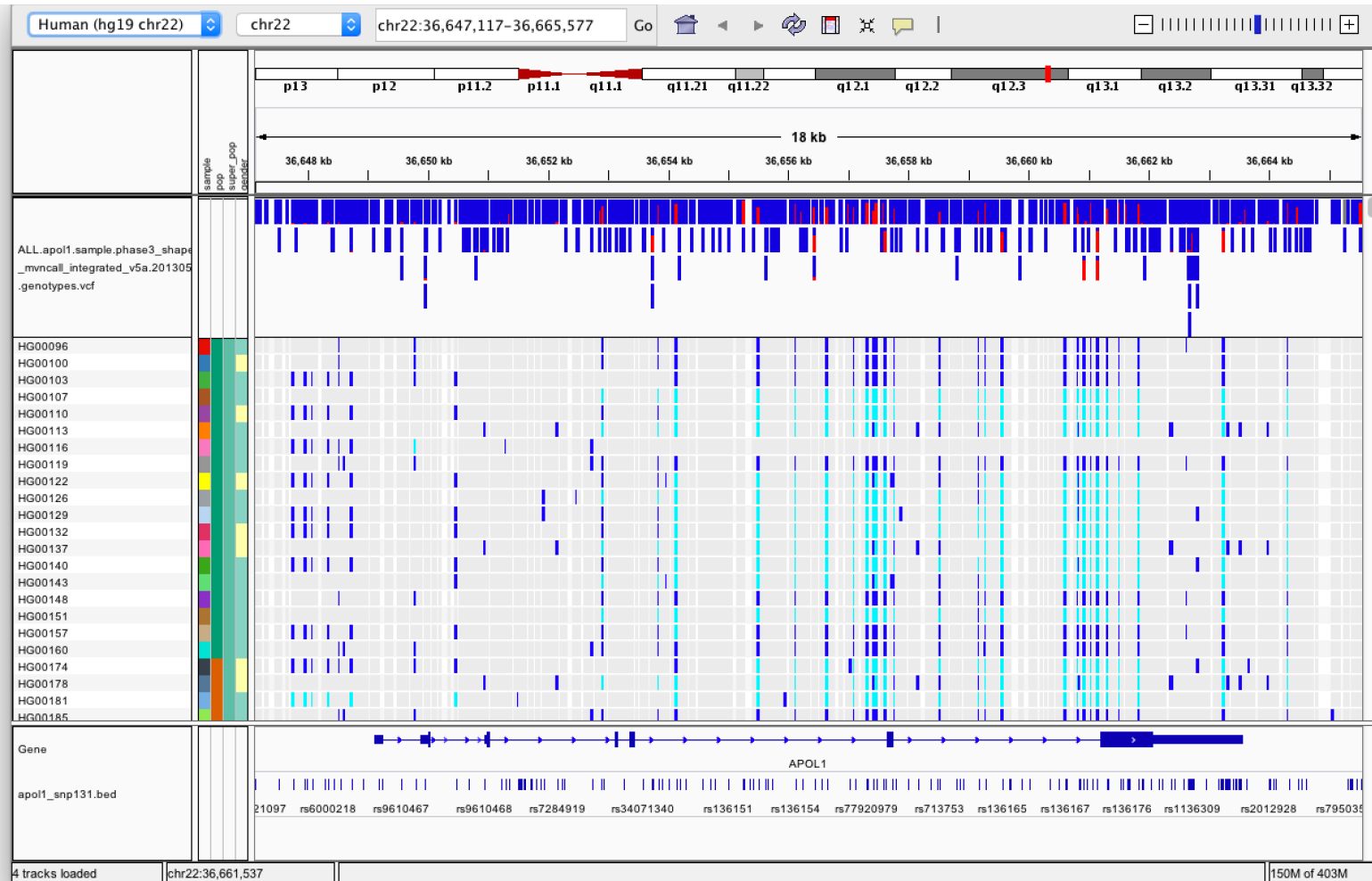
File format for encoding variant sites and genotypes

The VCF specification is now maintained by the
Global Alliance for Genomics and Health (GA4GH)

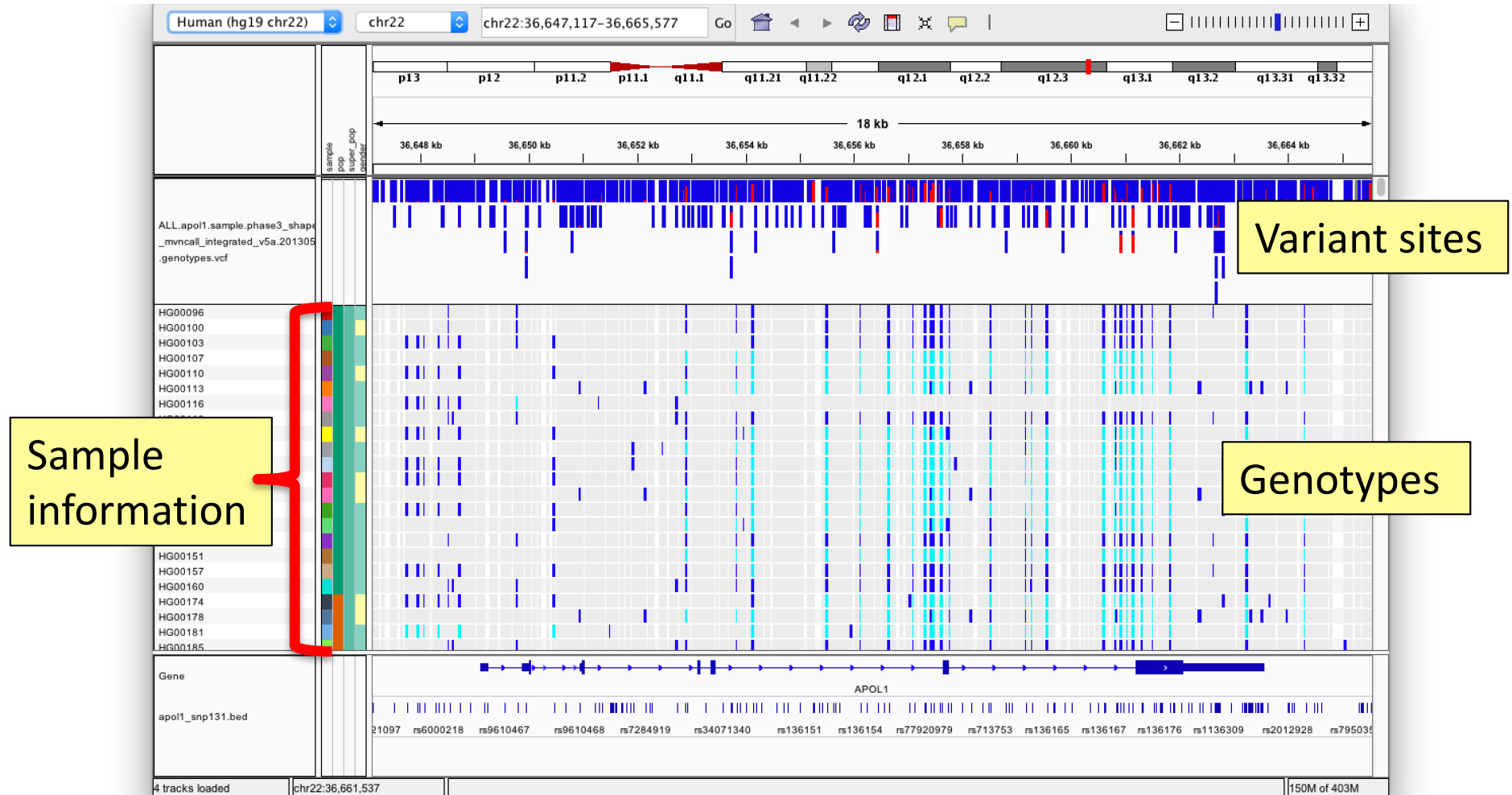
→ *Data Working Group* → *File Formats Task Team*

(<http://ga4gh.org/#/fileformats-team>)

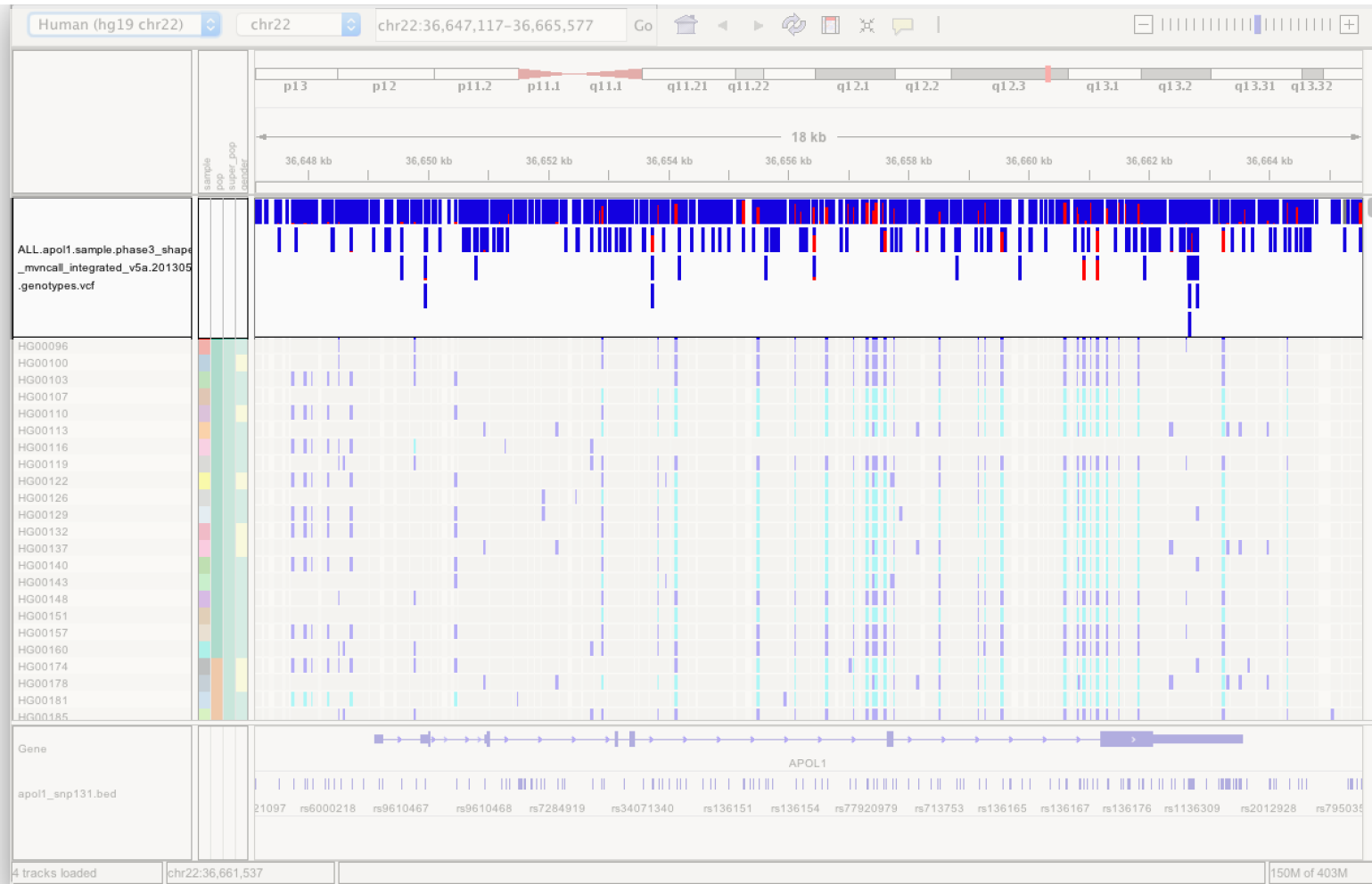
Viewing variants (VCF file)



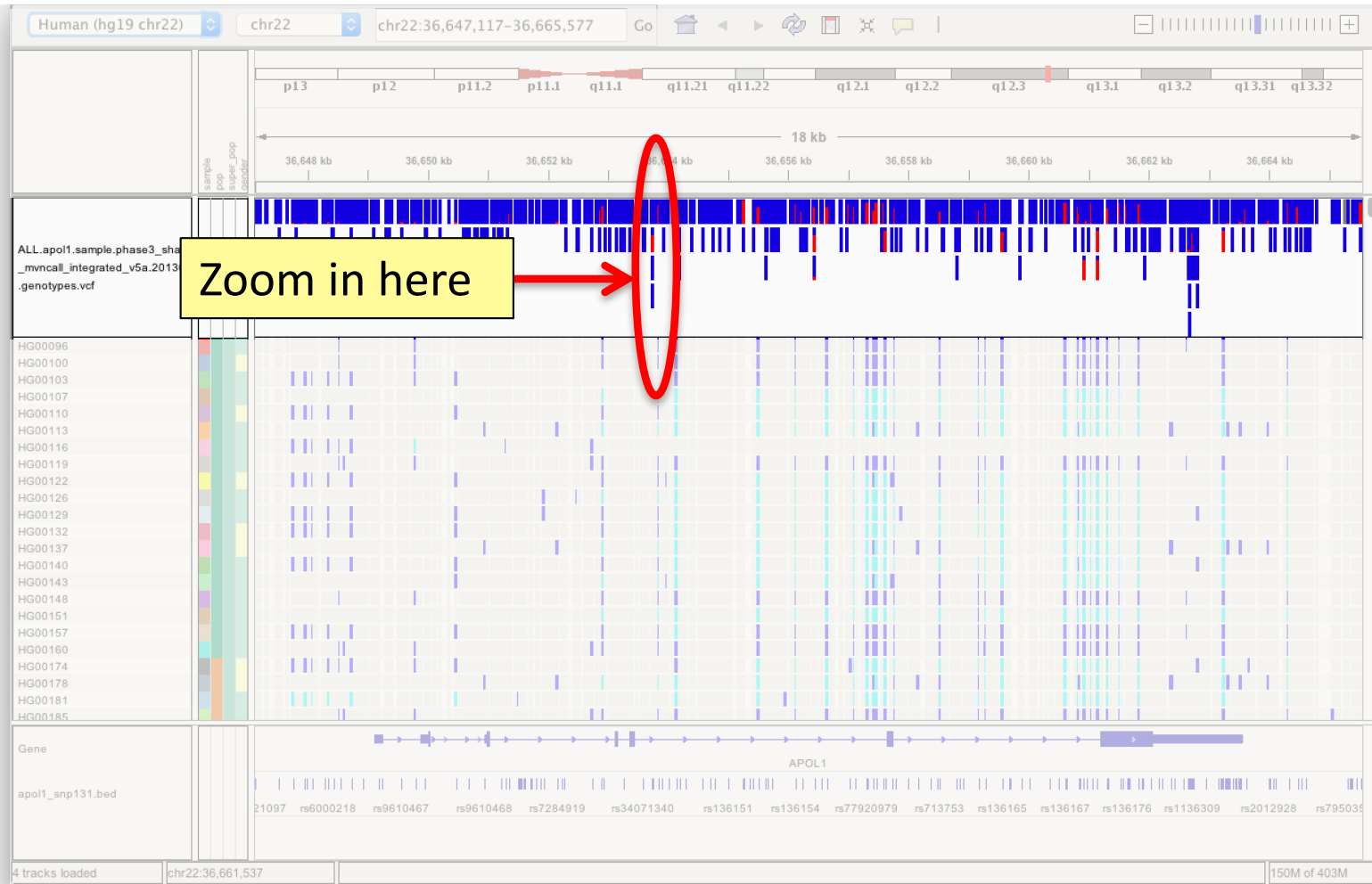
Viewing variants (VCF file)



Viewing variants (VCF file)

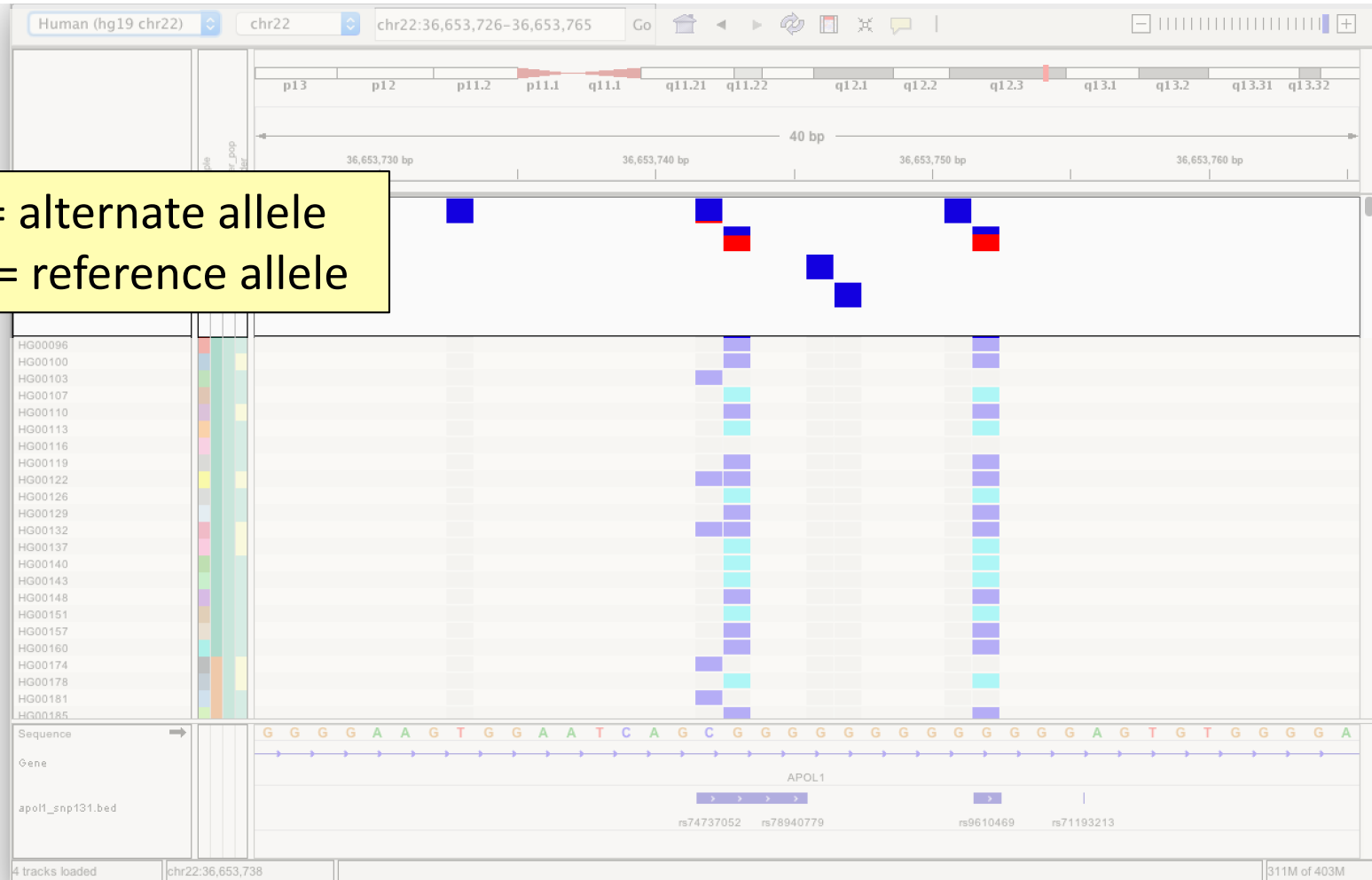


Viewing variants (VCF file)

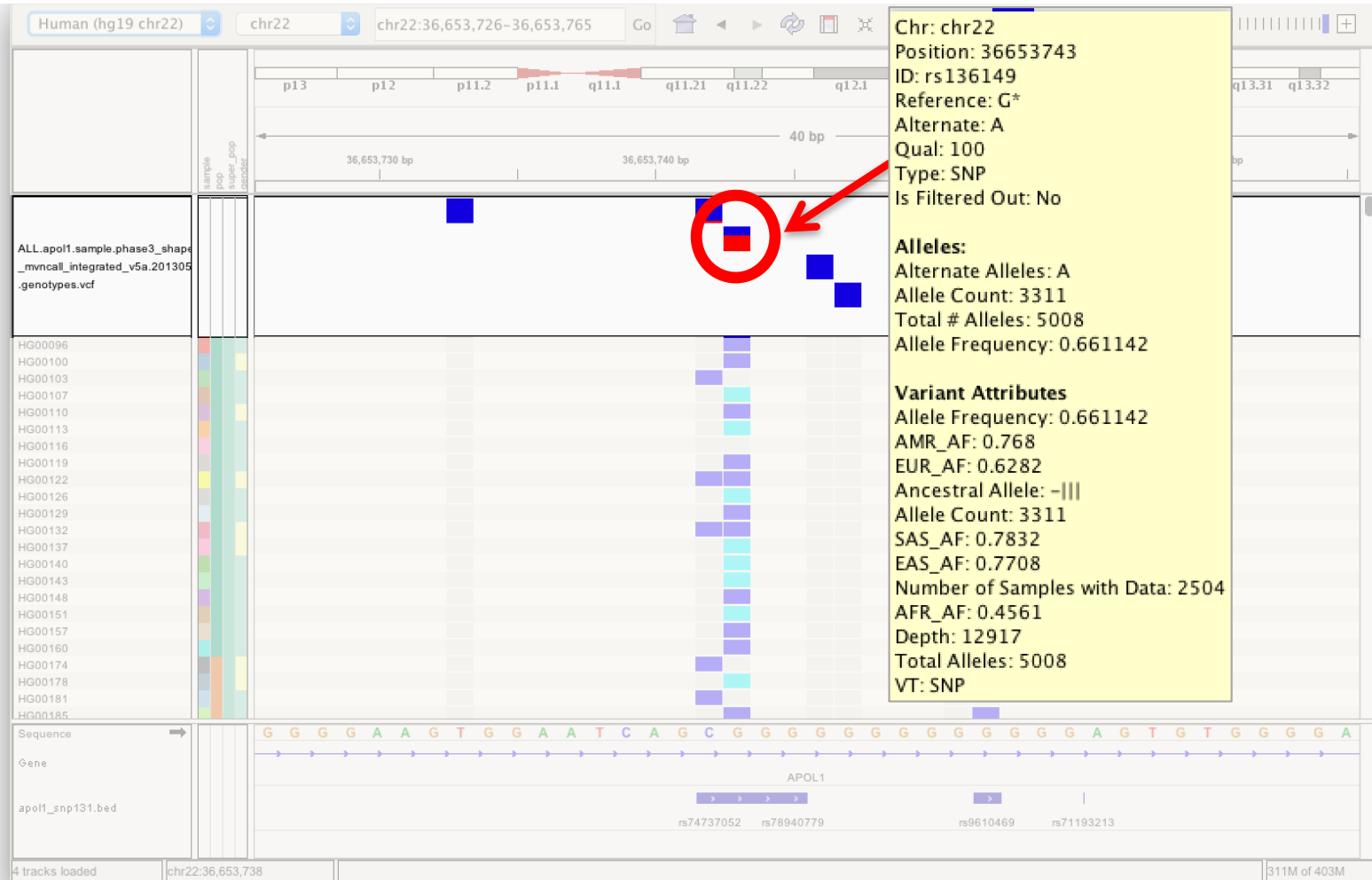


Viewing variants (VCF file)

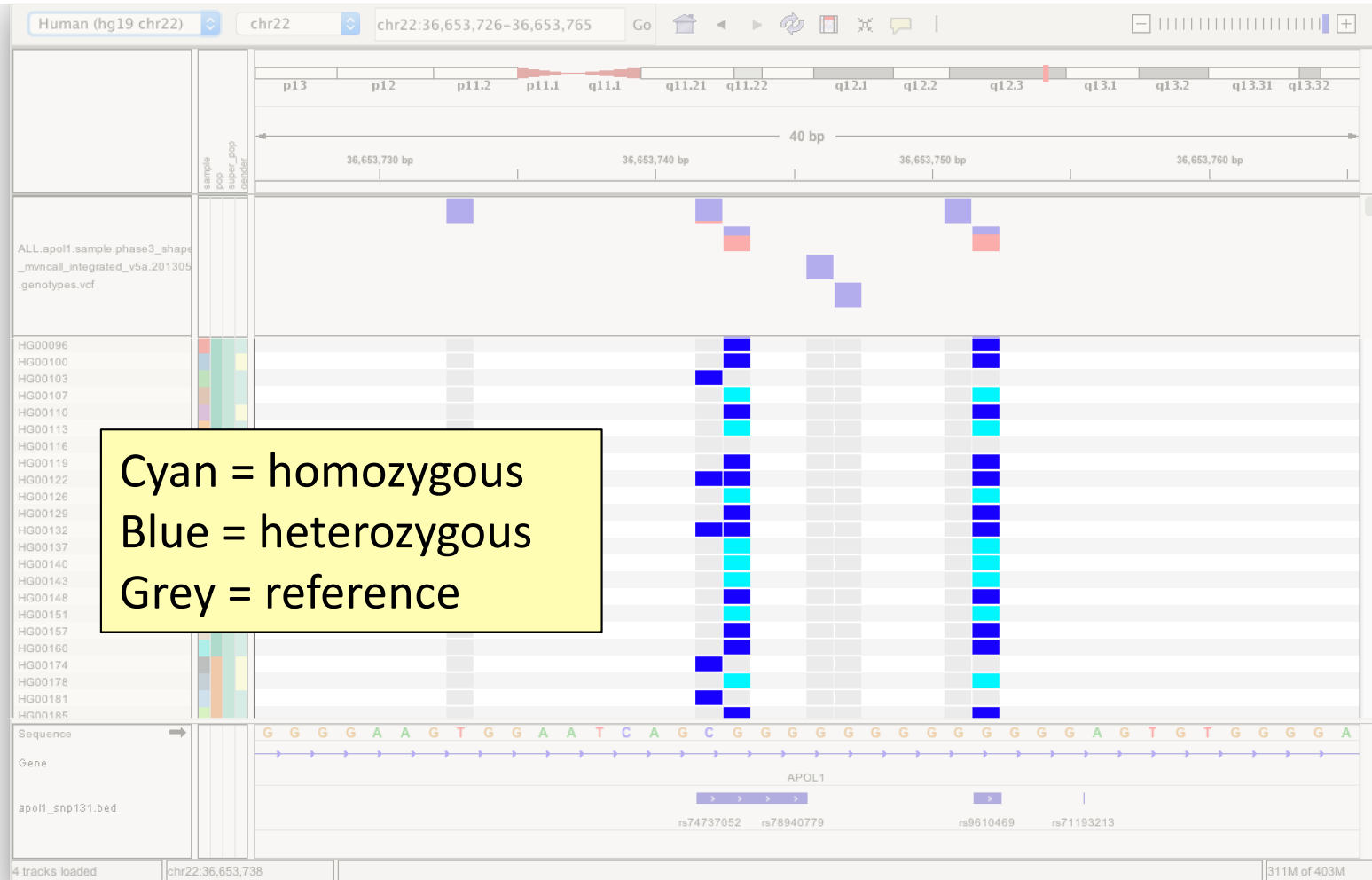
Red = alternate allele
Blue = reference allele



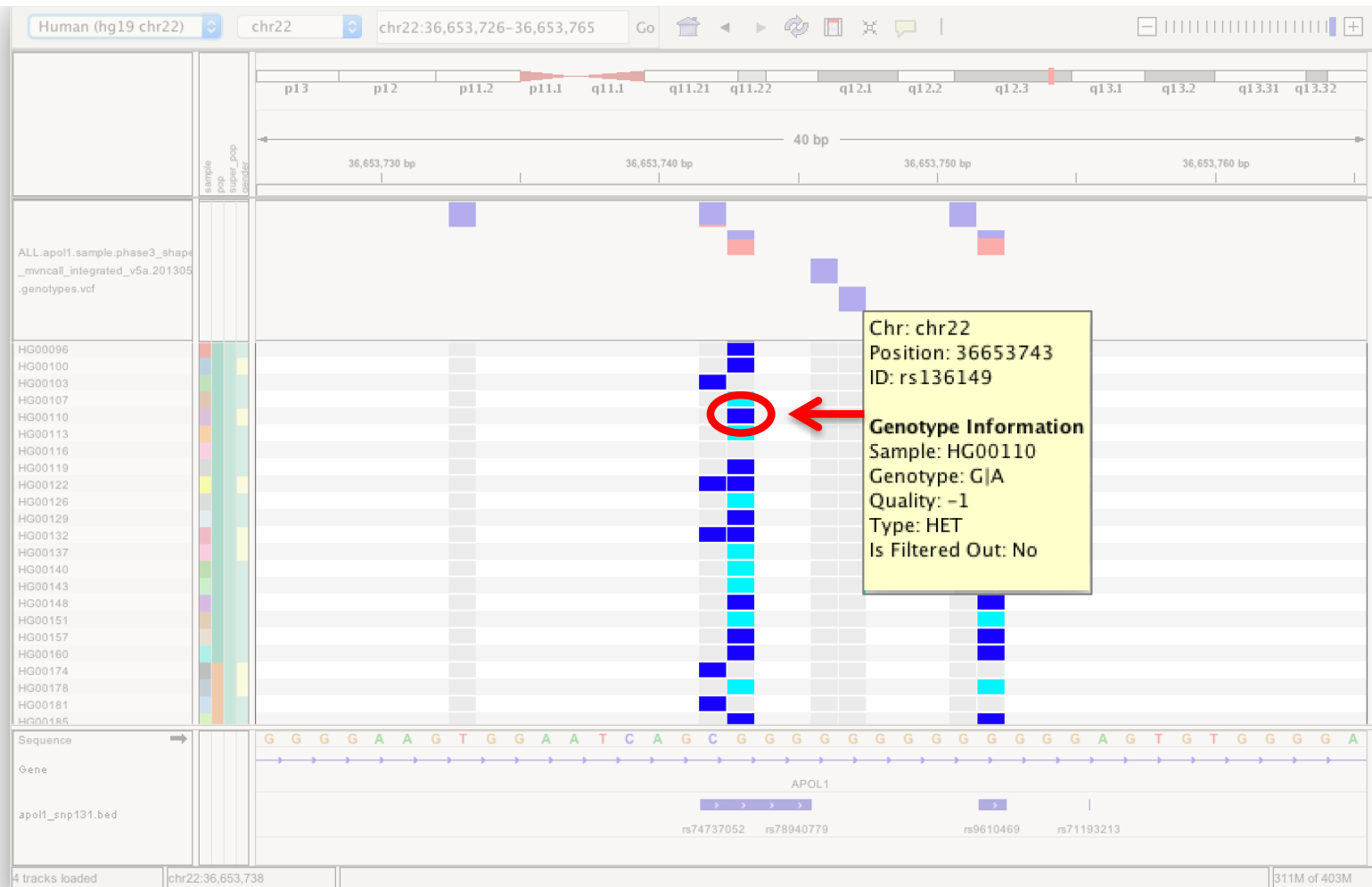
Viewing variants (VCF file)



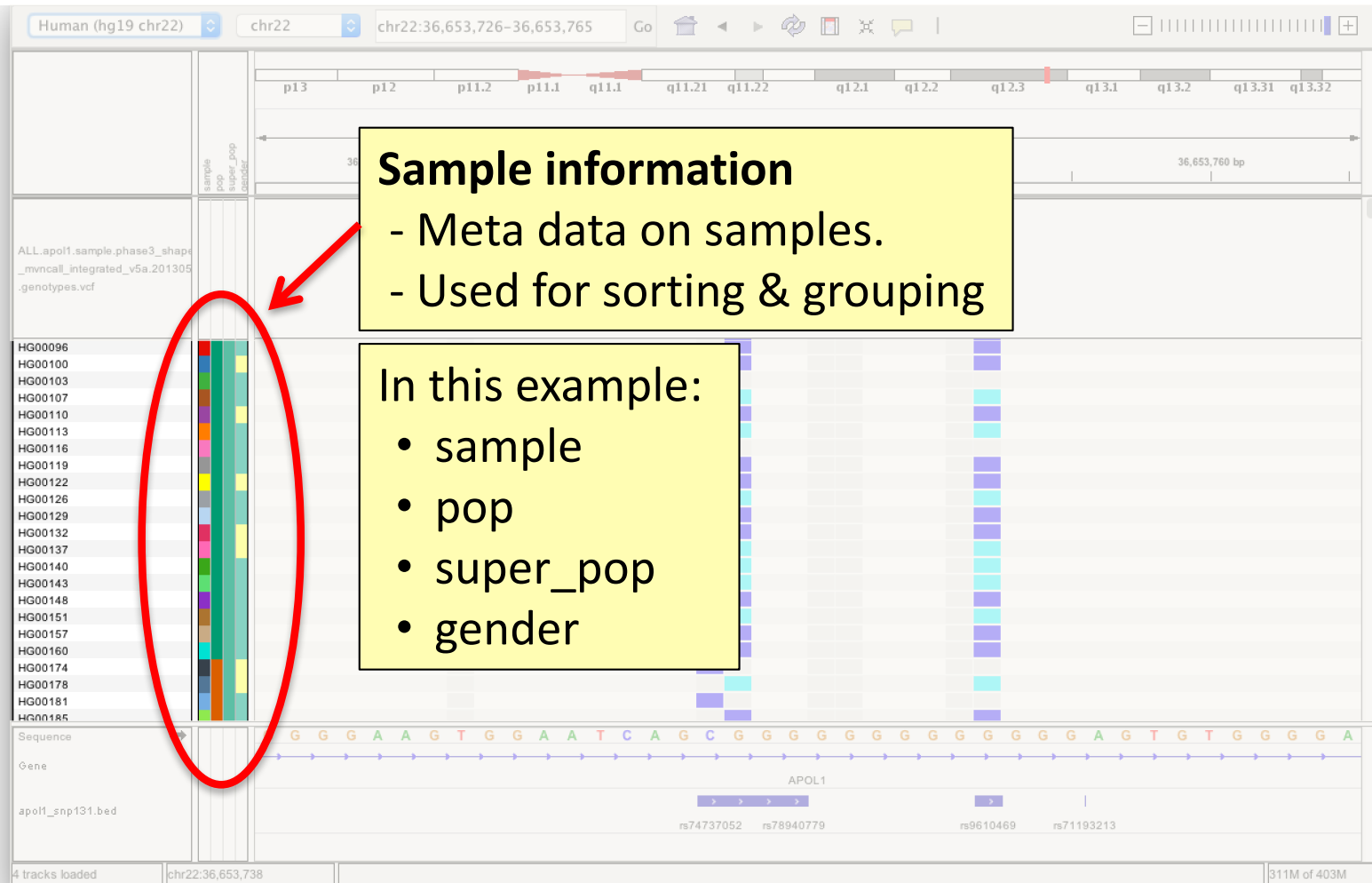
Viewing variants (VCF file)



Viewing variants (VCF file)



Viewing variants – sample information



Viewing variants – sample information

Sample information is defined in an auxiliary tab-delimited file.

- First **row** defines the meta data column names
- First **column** contains sample names
(must match VCF sample names)

Example:

sample	pop	super_pop	gender
HG00096	GBR	EUR	male
HG00097	GBR	EUR	female
HG00099	GBR	EUR	female
HG00100	GBR	EUR	female
HG00101	GBR	EUR	male
HG00102	GBR	EUR	female
HG00103	GBR	EUR	male
HG00105	GBR	EUR	male

VCF sample names

Meta data column names

Viewing variants (VCF file)

Hands-on exercise

Use sample information to examine a variant that segregates by population.

Reference:

Giulio Genovese et al.

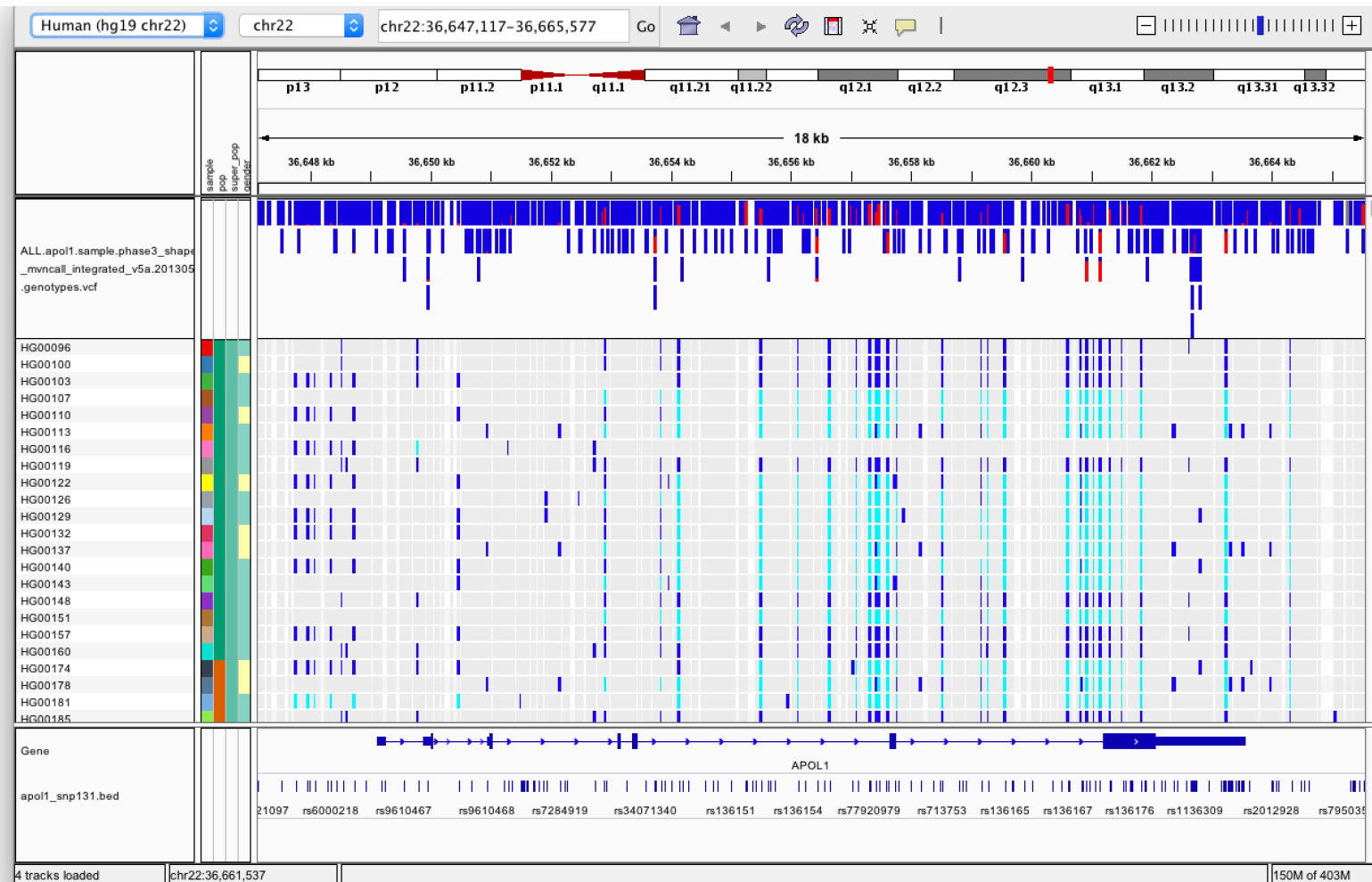
Association of Trypanolytic ApoL1 Variants with Kidney Disease in African Americans,

Science 329, 841 (2010); DOI: 10.1126/science.1193032

See handout

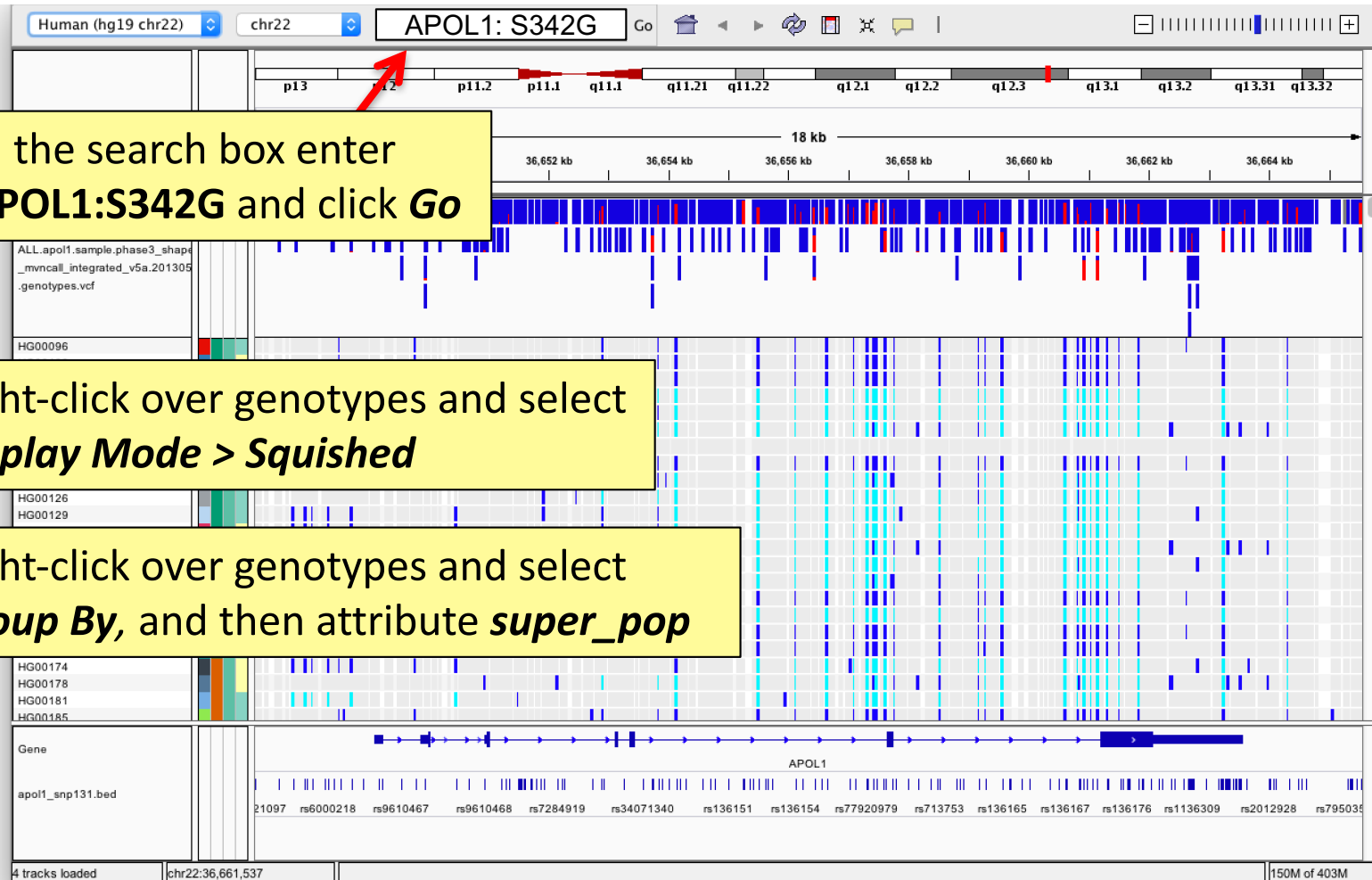
Viewing variants (VCF file)

Hands-on exercise



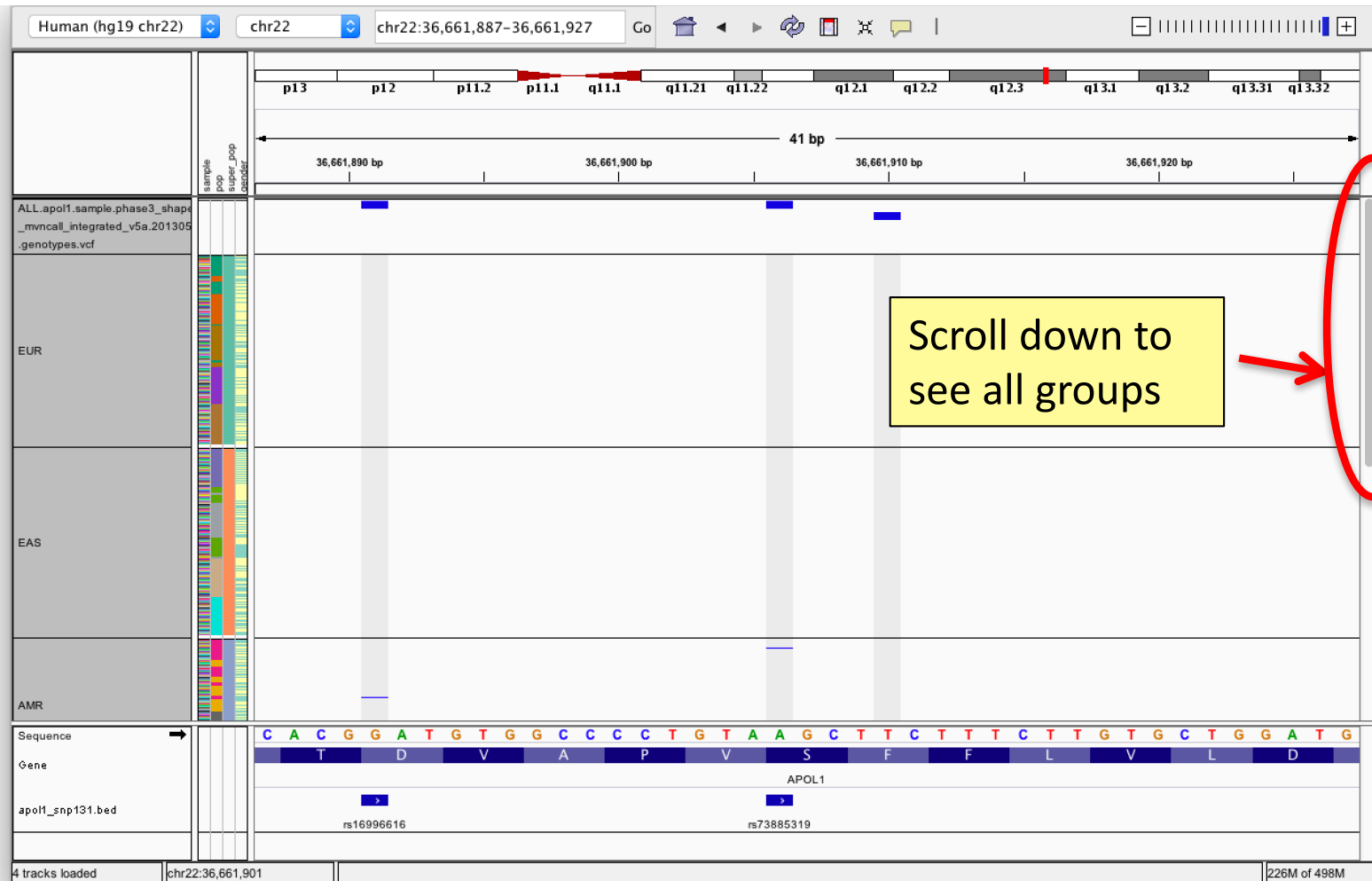
Viewing variants (VCF file)

Hands-on exercise



Viewing variants (VCF file)

Hands-on exercise



Viewing variants (VCF file)

Hands-on exercise

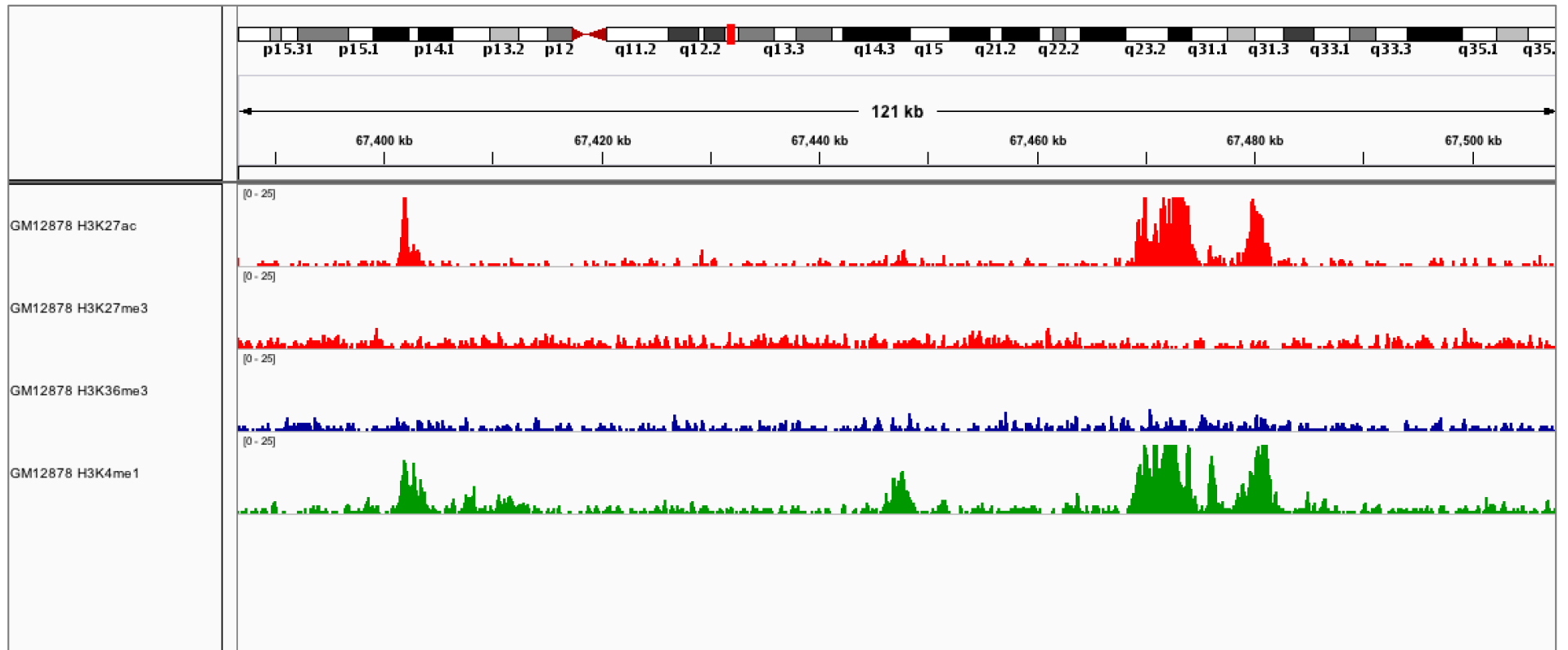


Other Tools

Tools

- Combining numeric data tracks
- BLAT

Combining numeric data tracks



Combining numeric data tracks

Select Tools > Combine Data Tracks

Tools

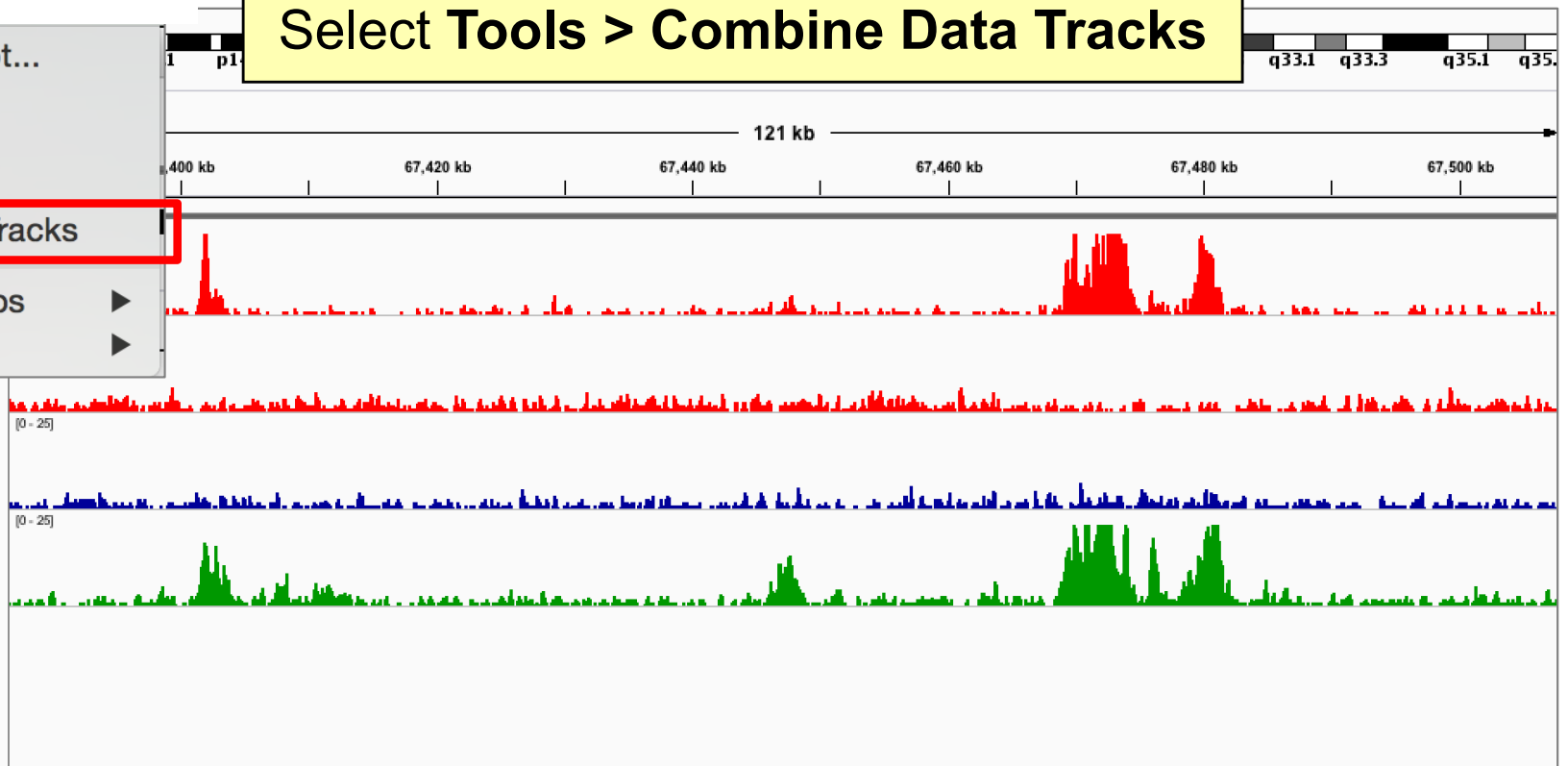
Run Batch Script...
Run igvtools...
Find Motif...
BLAT ...

Combine Data Tracks

Gitools Heatmaps ▶
BEDTools ▶

GM12878 H3K36me3

GM12878 H3K4me1



Combining numeric data tracks

1. Select two tracks from the dropdown menus

2. Select the operation:

- Add
- Subtract
- Multiply
- Divide

3. Name the new combined track

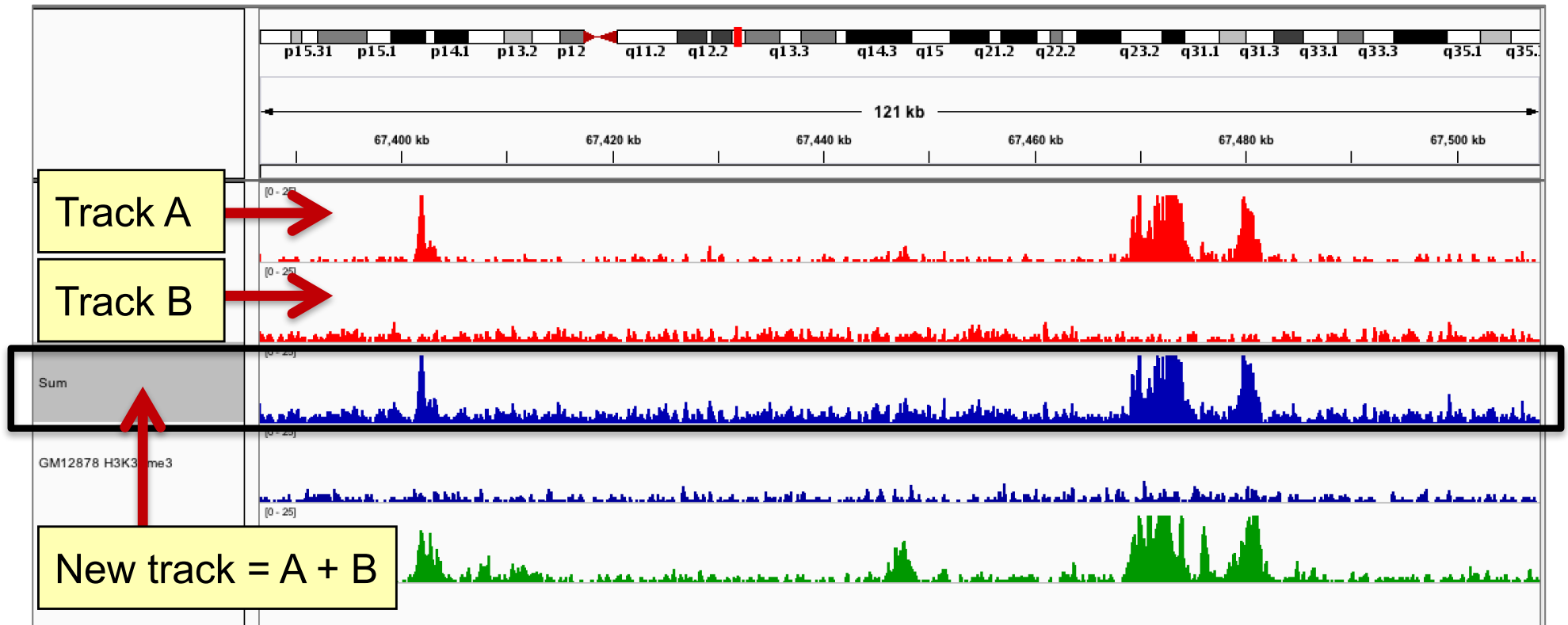
The image shows a 'Combine tracks' dialog box with the following fields and controls:

- Track A:** A dropdown menu currently showing 'GM12878 H3K27ac'.
- Operation:** A dropdown menu currently showing 'Add'.
- Track B:** A dropdown menu currently showing 'GM12878 H3K27me3'.
- Result Track Name:** A text input field containing the text 'Sum'.
- Buttons:** 'OK' and 'Cancel' buttons at the bottom.

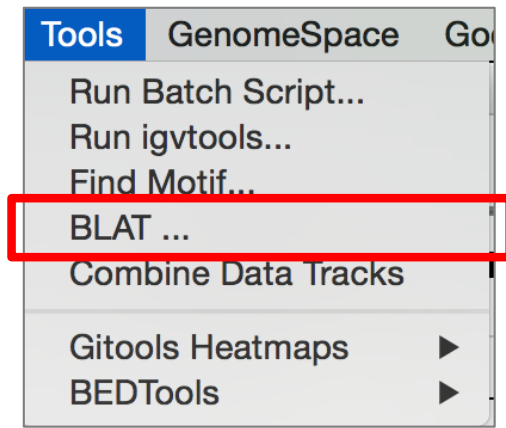
Three red arrows point from yellow instructional boxes to the dialog:

- One arrow points from the 'Track A' dropdown to the first instruction box.
- Another arrow points from the 'Operation' dropdown to the second instruction box.
- A third arrow points from the 'Result Track Name' text field to the third instruction box.

Combining numeric data tracks

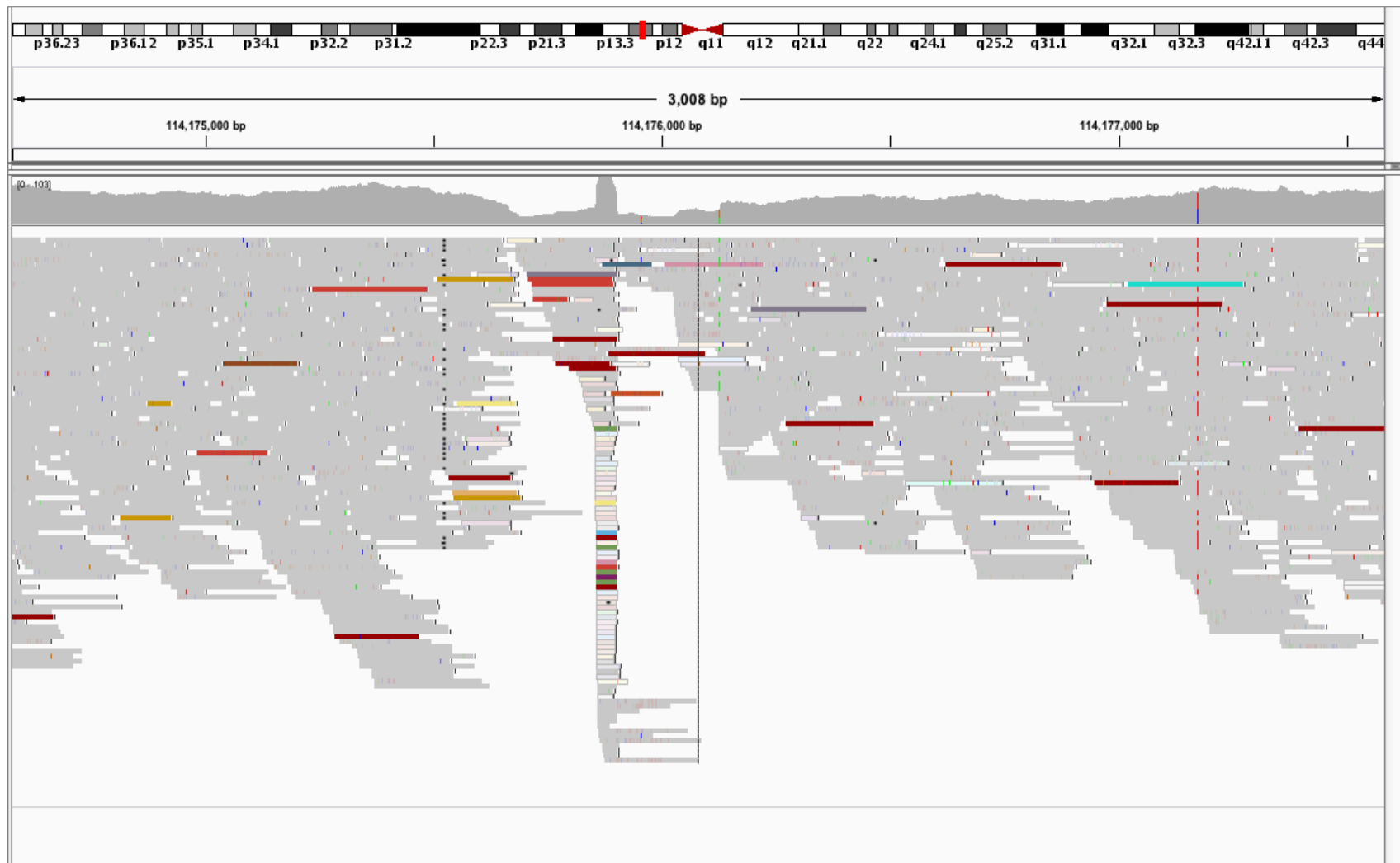


BLAT

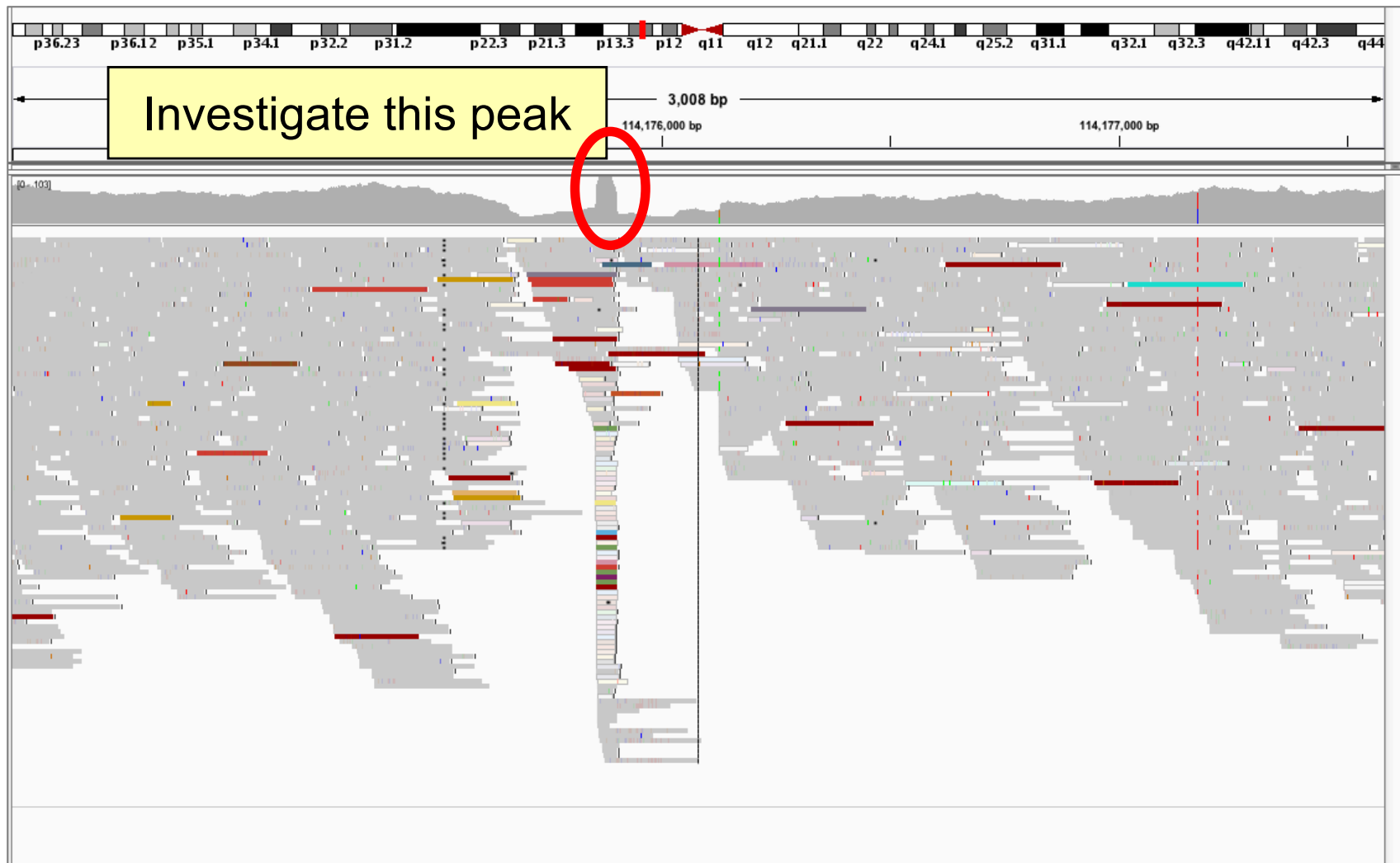


- Use a BLAT server to do search for alternative alignments for
 - any feature or sequence read, or
 - user-entered sequence.
- User-specified server or UCSC (volume limits apply).

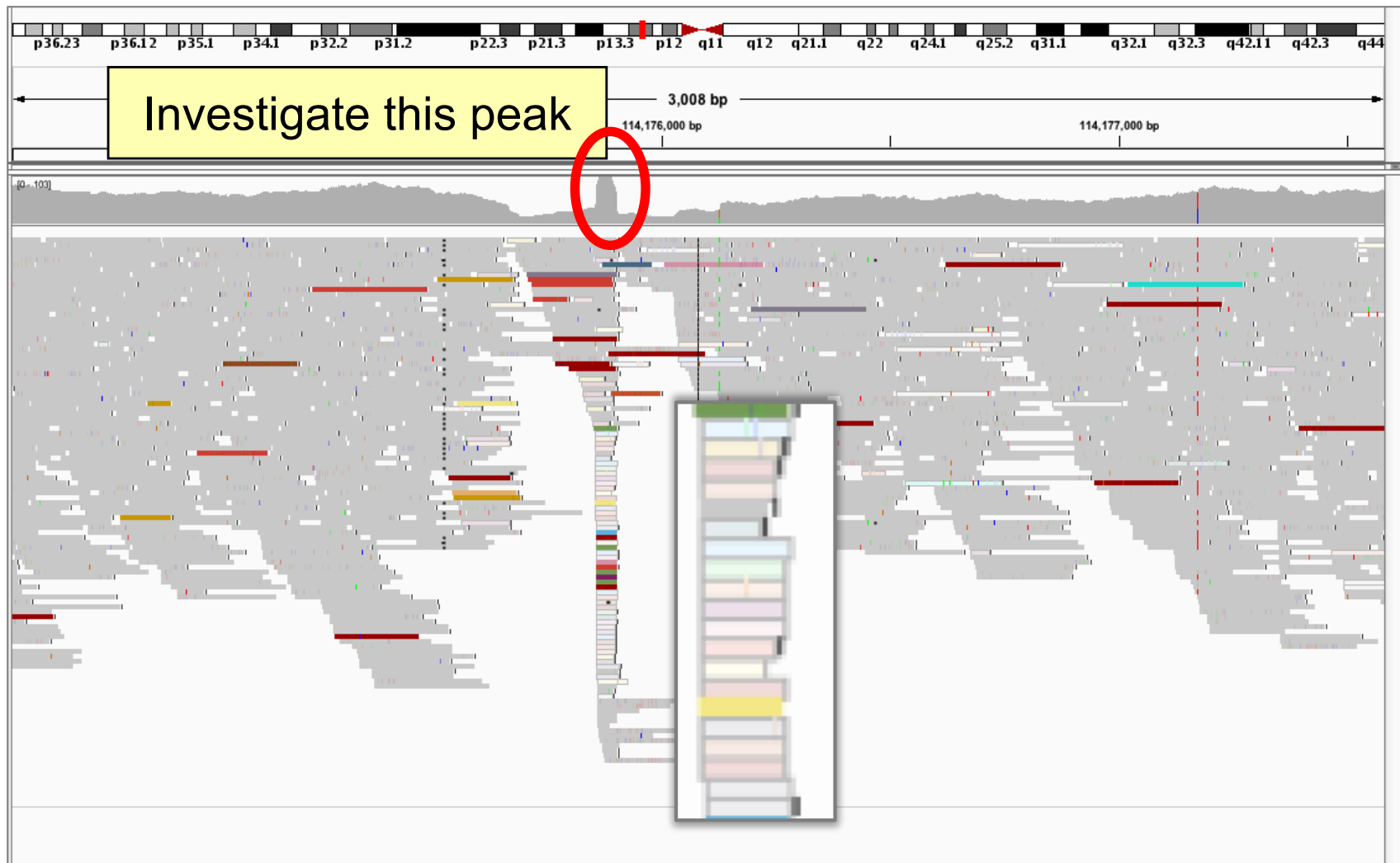
BLAT example



BLAT example

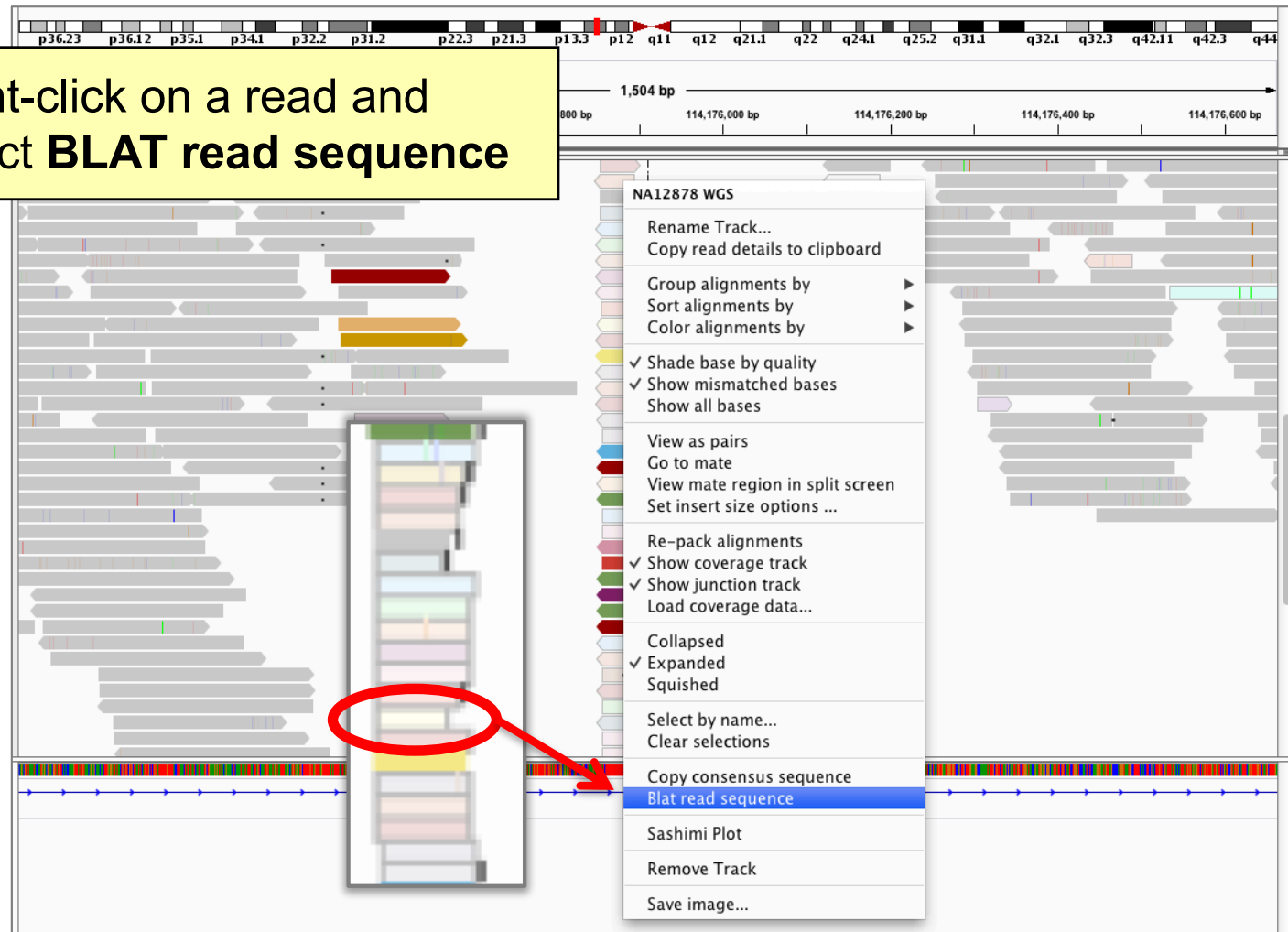


BLAT example



BLAT example

Right-click on a read and
select **BLAT read sequence**



BLAT example

File												
BLAT result for query sequence: TAATGCACACGAACTCCATTAGGTATTCATATGGGTTGAAAAGGGGGGTAACTTGGCCTAAATAGAAAATGTCTTCTCATCATTTTTTT												
<i>Click on a row to go to alignment</i>												
chr	start	end	strand	score	match	mis-match	rep. match	N's	Q gap co...	Q gap ba...	T gap co...	T gap ba...
chr6	1552157	1552307	-	536	142	8	0	0	0	0	0	0
chr1	401076...	401078...	-	492	147	19	0	0	2	9	3	16
chr2	715213...	715215...	+	456	137	22	0	0	0	0	1	1
chr13	525933...	525935...	+	444	135	23	0	0	0	0	1	1
chr17	5468054	5468211	+	440	128	16	0	0	1	12	1	13
chr10	933987...	933989...	+	436	132	22	0	0	0	0	1	1
chr15	956086...	956088...	-	428	129	20	0	0	0	0	2	4
chr5	166236...	166238...	+	424	126	18	0	0	0	0	2	4
chr3	102995...	102996...	+	424	124	15	0	0	1	5	2	9
chr17	708930...	708932...	-	420	130	24	0	0	0	0	1	1
chr12	571299...	571301...	-	420	135	27	0	0	1	1	2	4
chr5	665088...	665089...	+	420	136	31	0	0	0	0	0	0
chr9	915725	915886	+	416	127	21	0	0	0	0	2	13
chr2	223622...	223622...	+	400	120	17	0	0	1	11	2	13
chr19	583386...	583388...	-	396	124	24	0	0	0	0	1	1
chr16	574045...	574047...	+	392	118	16	0	0	2	29	2	8
chr1	171497...	171497...	+	388	114	16	0	0	0	0	1	1
chr10	101042...	101042...	-	384	111	13	0	0	0	0	2	7
chr2	228547...	228547...	+	384	116	19	0	0	0	0	1	1
chr1	202874...	202875...	+	384	120	23	0	0	0	0	1	1
chr1	549438...	549439...	+	384	119	22	0	0	0	0	1	1
chr12	570020...	570021...	-	380	113	17	0	0	0	0	1	1
chr4	878470...	878471...	+	376	119	22	0	0	0	0	3	6
chr2	145618...	145619...	+	376	110	11	0	0	3	38	2	40
chr1	7404808	7404937	+	372	109	15	0	0	0	0	1	5
chr1	247194...	247194...	-	368	106	12	0	0	0	0	2	6
chrX	147458...	147458...	-	356	108	17	0	0	1	35	1	36

BLAT example

Many alternative alignments
with similar scores

GTTGAAAAGGGGGGTAACTTGGCCTAAATAGAAAATGTCTTCTCATCTTTTTT

chr	start	end	strand	score	match	mis-match	rep. match	N's	Q gap co...	Q gap ba...	T gap co...	T gap ba...
chr6	1552157	1552307	-	536	142	8	0	0	0	0	0	0
chr1	401076...	401078...	-	492	147	19	0	0	2	9	3	16
chr2	715213...	715215...	+	456	137	22	0	0	0	0	1	1
chr13	525933...	525935...	+	444	135	23	0	0	0	0	1	1
chr17	5468054	5468211	+	440	128	16	0	0	1	12	1	13
chr10	933987...	933989...	+	436	132	22	0	0	0	0	1	1
chr15	956086...	956088...	-	428	129	20	0	0	0	0	2	4
chr5	166236...	166238...	+	424	126	18	0	0	0	0	2	4
chr3	102995...	102996...	+	424	124	15	0	0	1	5	2	9
chr17	708930...	708932...	-	420	130	24	0	0	0	0	1	1
chr12	571299...	571301...	-	420	135	27	0	0	1	1	2	4
chr5	665088...	665089...	+	420	136	31	0	0	0	0	0	0
chr9	915725	915886	+	416	127	21	0	0	0	0	2	13
chr2	223622...	223622...	+	400	120	17	0	0	1	11	2	13
chr19	583386...	583388...	-	396	124	24	0	0	0	0	1	1
chr16	574045...	574047...	+	392	118	16	0	0	2	29	2	8
chr1	171497...	171497...	+	388	114	16	0	0	0	0	1	1
chr10	101042...	101042...	-	384	111	13	0	0	0	0	2	7
chr2	228547...	228547...	+	384	116	19	0	0	0	0	1	1
chr1	202874...	202875...	+	384	120	23	0	0	0	0	1	1
chr1	549438...	549439...	+	384	119	22	0	0	0	0	1	1
chr12	570020...	570021...	-	380	113	17	0	0	0	0	1	1
chr4	878470...	878471...	+	376	119	22	0	0	0	0	3	6
chr2	145618...	145619...	+	376	110	11	0	0	3	38	2	40
chr1	7404808	7404937	+	372	109	15	0	0	0	0	1	5
chr1	247194...	247194...	-	368	106	12	0	0	0	0	2	6
chrX	147458...	147458...	-	356	108	17	0	0	1	35	1	36

Closing

For further information and help

IGV web site: www.igv.org

User forum: groups.google.com/group/igv-help

Source code: github.com/igvteam

Please cite your use of IGV

Robinson et al.

Integrative Genomics Viewer. Nature Biotechnology (2011).

Thorvaldsdóttir, Robinson, and Mesirov.

Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.

Briefings in Bioinformatics (2012).

