Visualization of mapped reads

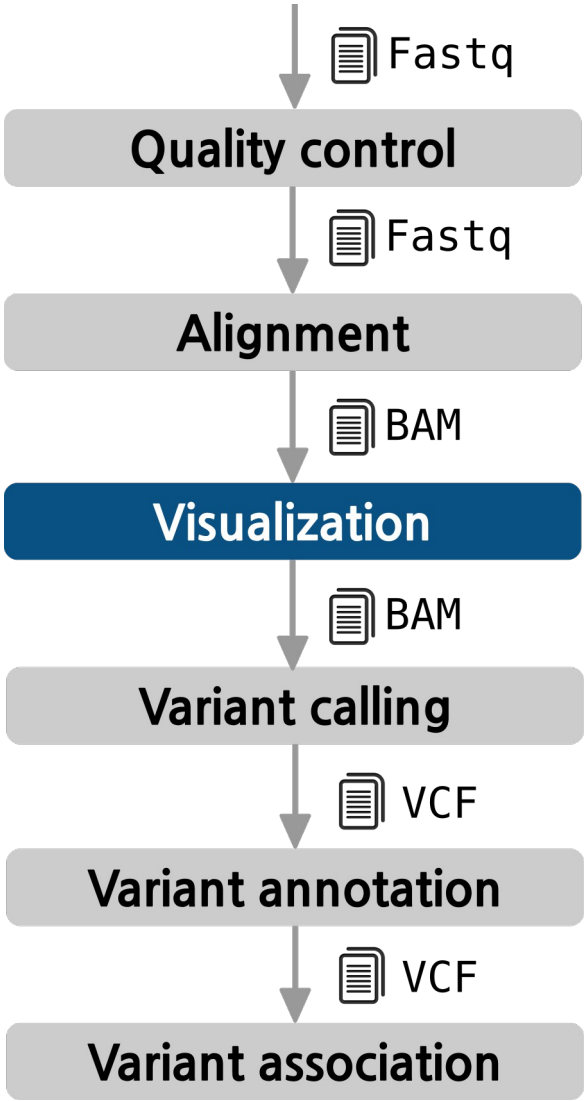
Integrative Genomics Viewer (IGV)

University of Edinburgh
Edinburgh, UK
22nd October 2015

Marta Bleda Latorre
mb2033@cam.ac.uk
Research Assistant at the Department of Medicine
University of Cambridge
Cambridge, UK

UNIVERSITY OF CAMBRIDGE

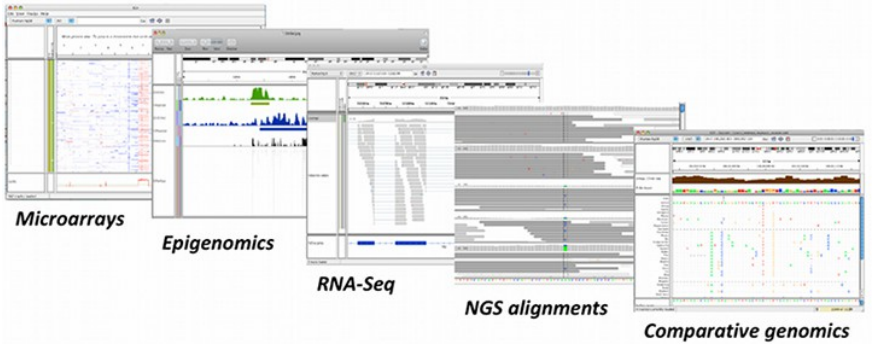
The pipeline



Introduction

- **Large** quantities of genomic **data** (NGS, array based methods, etc)
- **Human interpretation** and judgment using visualization can help complex biological relationships

- Integrative Genomics Viewer (**IGV**)
  - **Integrate** different data types simultaneously
  - View **large datasets** easily
  - Faster navigation or browsing
  - Runs **locally** on your desktop
  - Used by large-scale projects
  - Open source and **freely available**



Helga Thorvaldsdóttir, James T. Robinson, and Jill P. Mesirov
Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration
Brief Bioinform (2013) 14 (2): 178-192

Data types

- Any data related to **genome coordinates**
- **Sample** annotations or attributes
- **Genome** annotations

Recommended file formats

Source data	Recommended File Formats
Sequence alignment data	SAM (must be sorted/indexed) BAM (must be indexed)
Genome annotations	GFF or GFF3 format BED format
Variant data	VCF
Any numeric data	IGV format, TAB format WIG format
Gene expression data	GCT format RES format

Indexing a BAM file

- BAM format: Binary **SAM** file → Reduces disk space and time
- BAM/SAM files need to be **indexed** (using **samtools**) → SAM files will be sorted by start position and indexed
- Index files must reside in the **same directory** as the BAM or SAM file

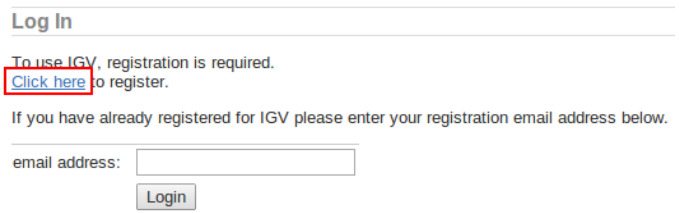
Indexing BAM file

```
samtools index example.bam
```

Registration and download

1. Be sure that **Java 6 or later** is installed on your machine
2. Go to the IGV website:

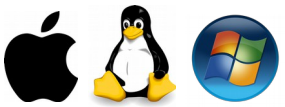
   http://www.broadinstitute.org/igv/home
3. Click **Downloads** at the left panel
4. Click to register and fill the form



5. Download the most suitable file for your system

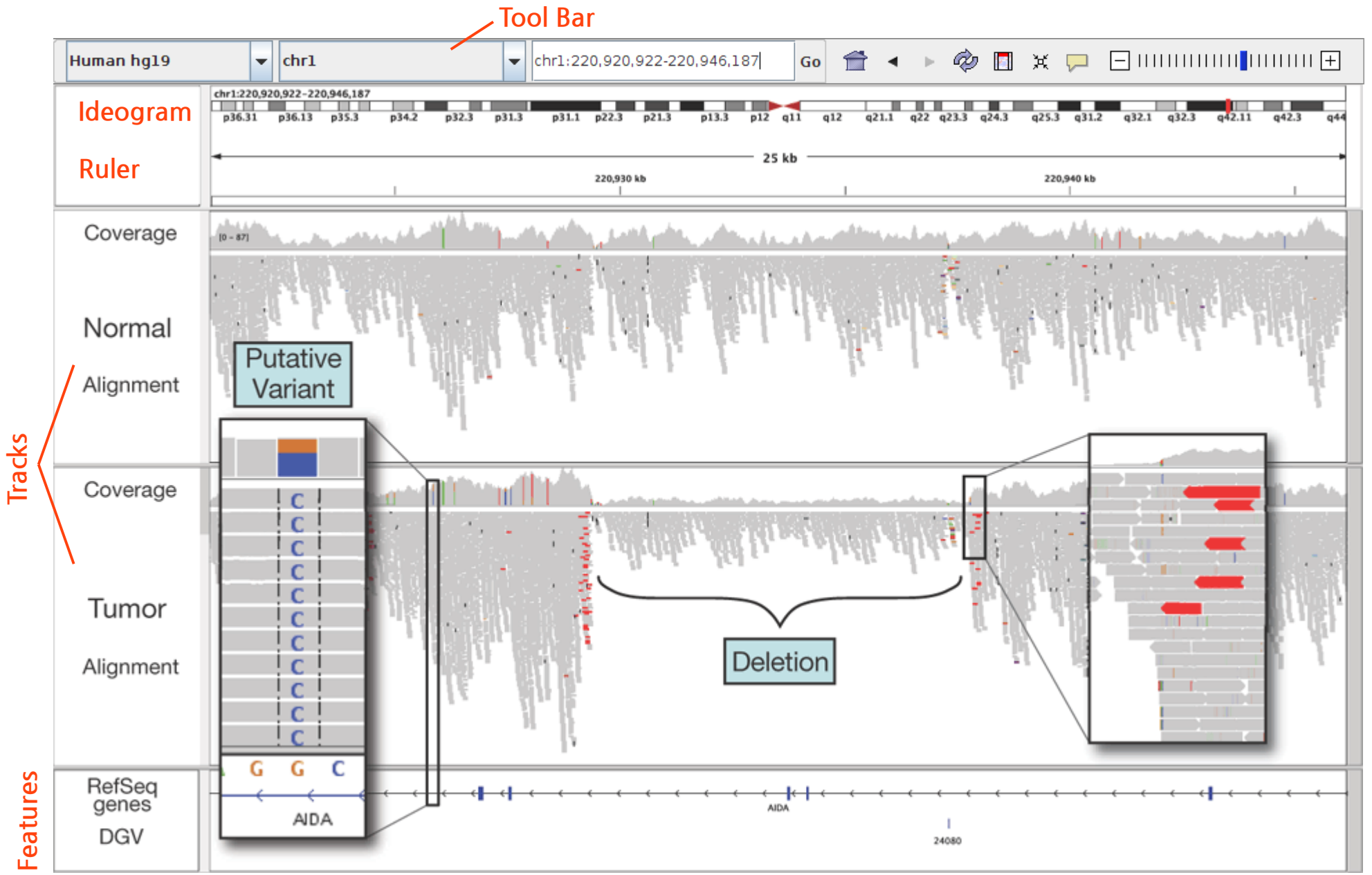
Download Mac App

* Requires Java 7

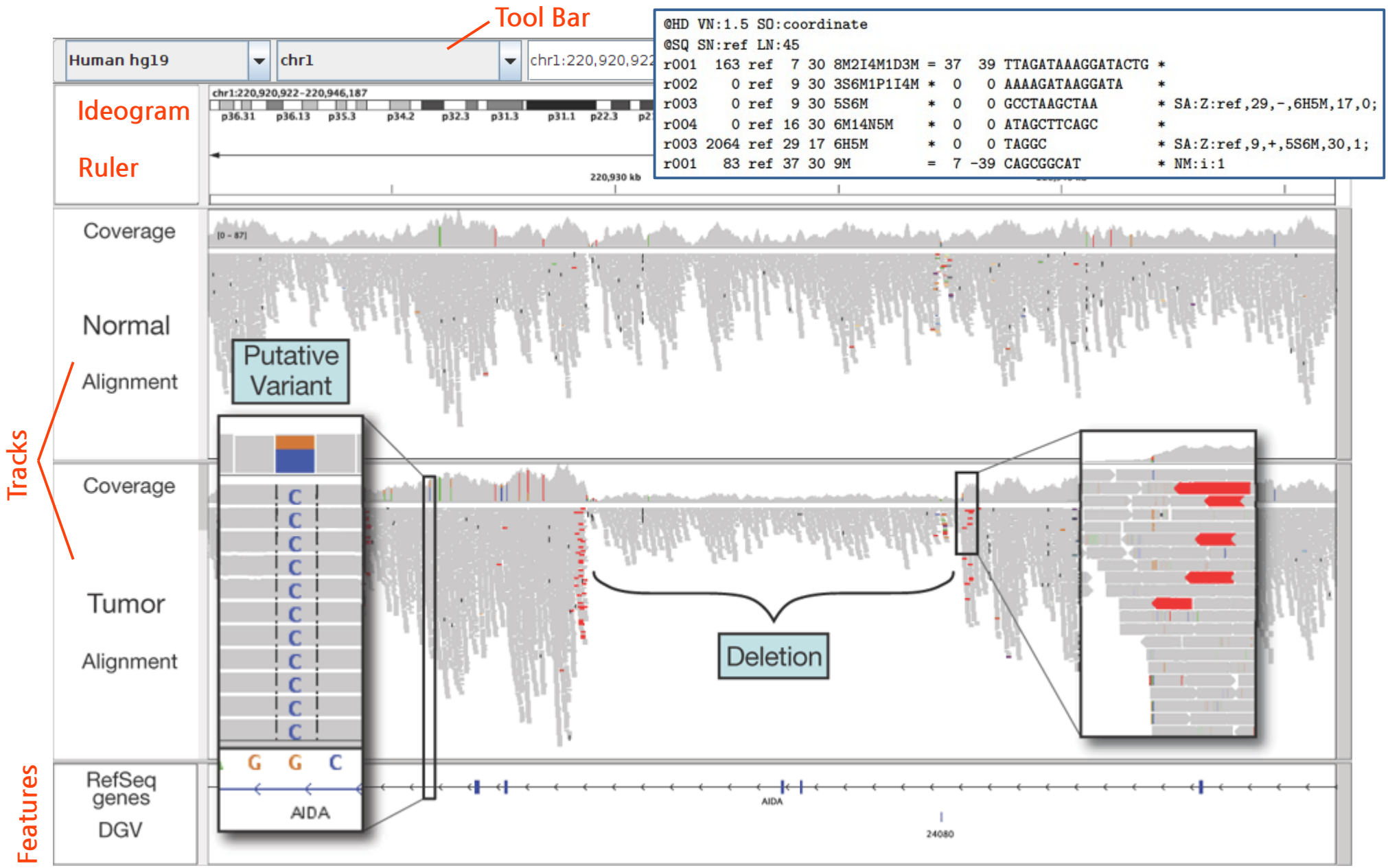


Download Binary Distribution

IGV interface



IGV interface



Hands on!

Download genome

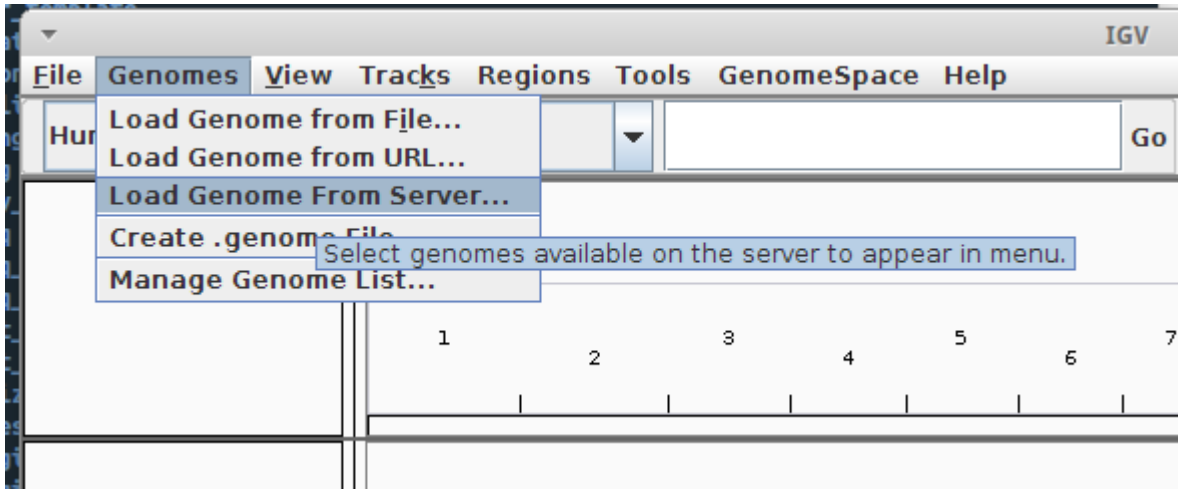
Current version of IGV has "**Human hg18**" and "**Human hg19**" already loaded

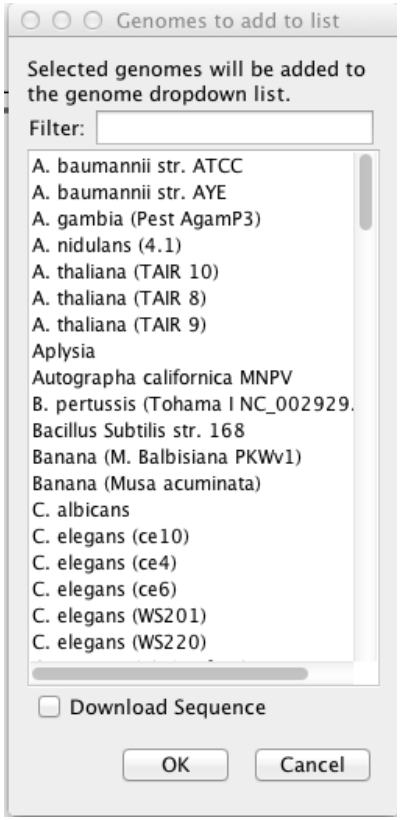
IGV provides **a number of genomes that are hosted** on a server at the Broad Institute

List of genomes hosted: http://www.broadinstitute.org/software/igv/Genomes

- Genomes → Load genome from server…

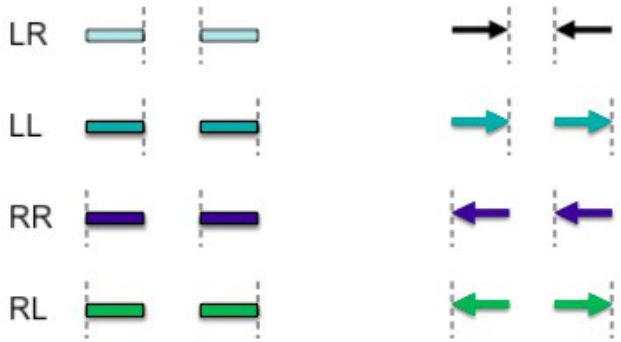
  Select **Human hg19**





IGV: interpreting colored reads

Read pair orientation



LR	Normal reads. 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.

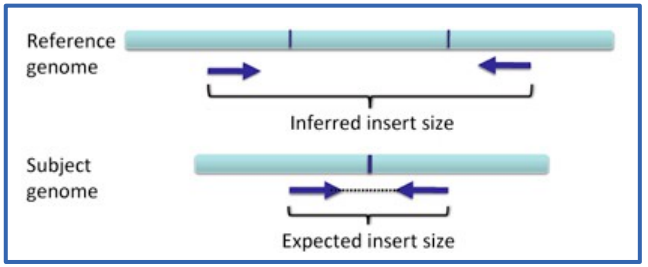
Insert size

Larger than expected (Deletion)

Smaller than expected (Insertion)

Mate of paired end reads that map to other chromosomes

Deletion



Insertion