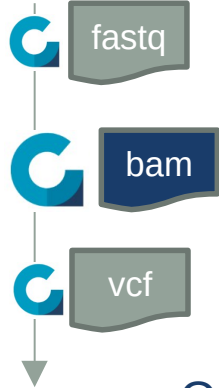


Viewing a BAM file

- With user graphical interface
IGV – Integrative Genome Viewer
- Software <http://software.broadinstitute.org/software/igv/>
- Data <https://software.broadinstitute.org/gatk/blog?id=9044>
2017 Feb workshop presentation slides and tutorial materials
Germline Data Bundle (Day 2)



Viewing a BAM file

IGV

- Open IGV
- Genomes
 - Load Genome from File
- File
 - Load from file
 - Load from file
 - Load from file
- Go
- Zoom in

`data/ref/ref.fasta`

`data/bams/son.bam`

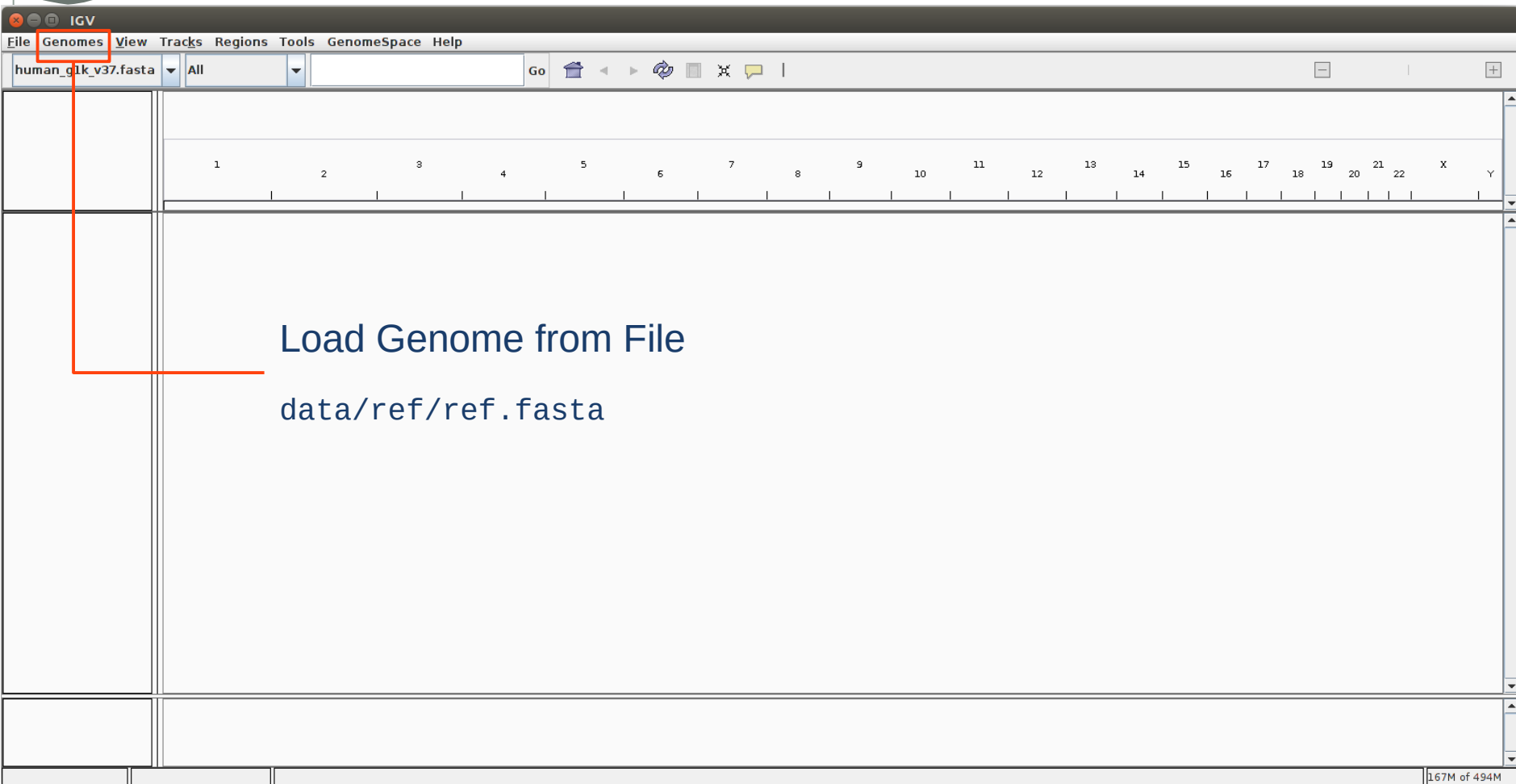
`data/bams/mother.bam`

`data/bams/father.bam`

`20:10,020,000-10,050,000`



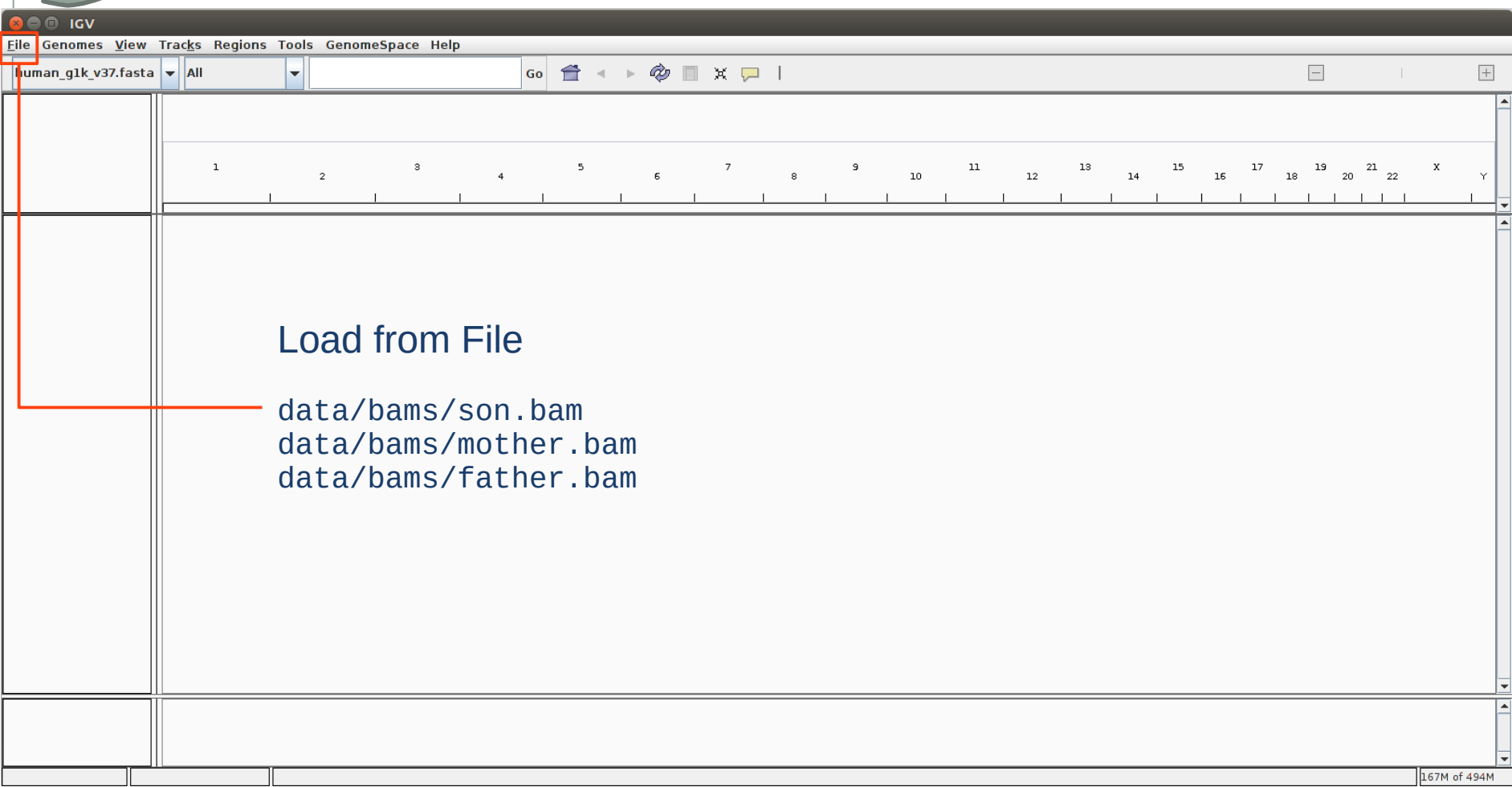
Viewing a BAM file IGV



Load Genome from File
data/ref/ref.fasta

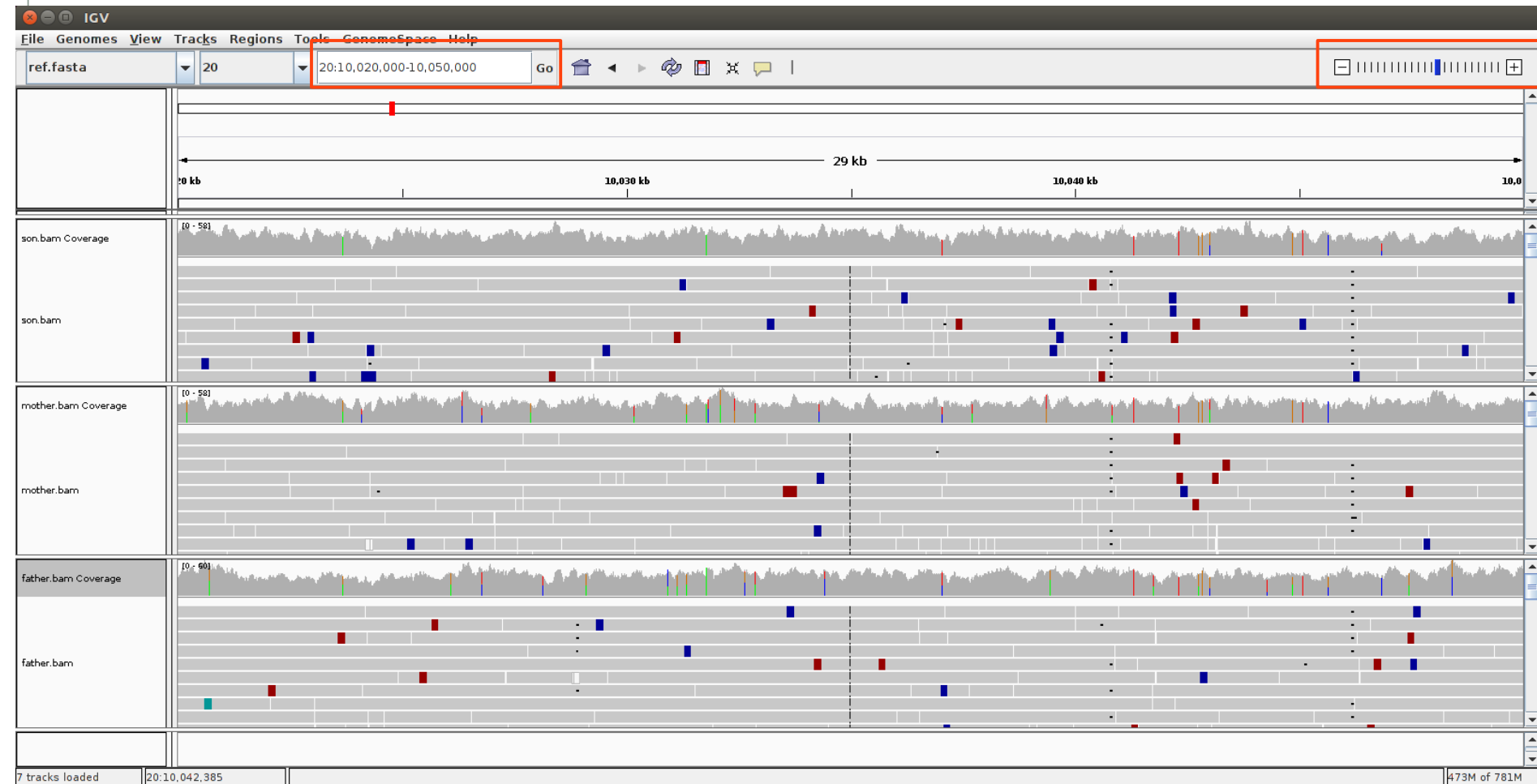


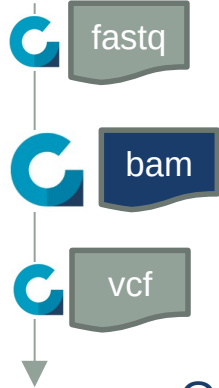
Viewing a BAM file IGV





Viewing a BAM file IGV





Viewing a BAM file

IGV

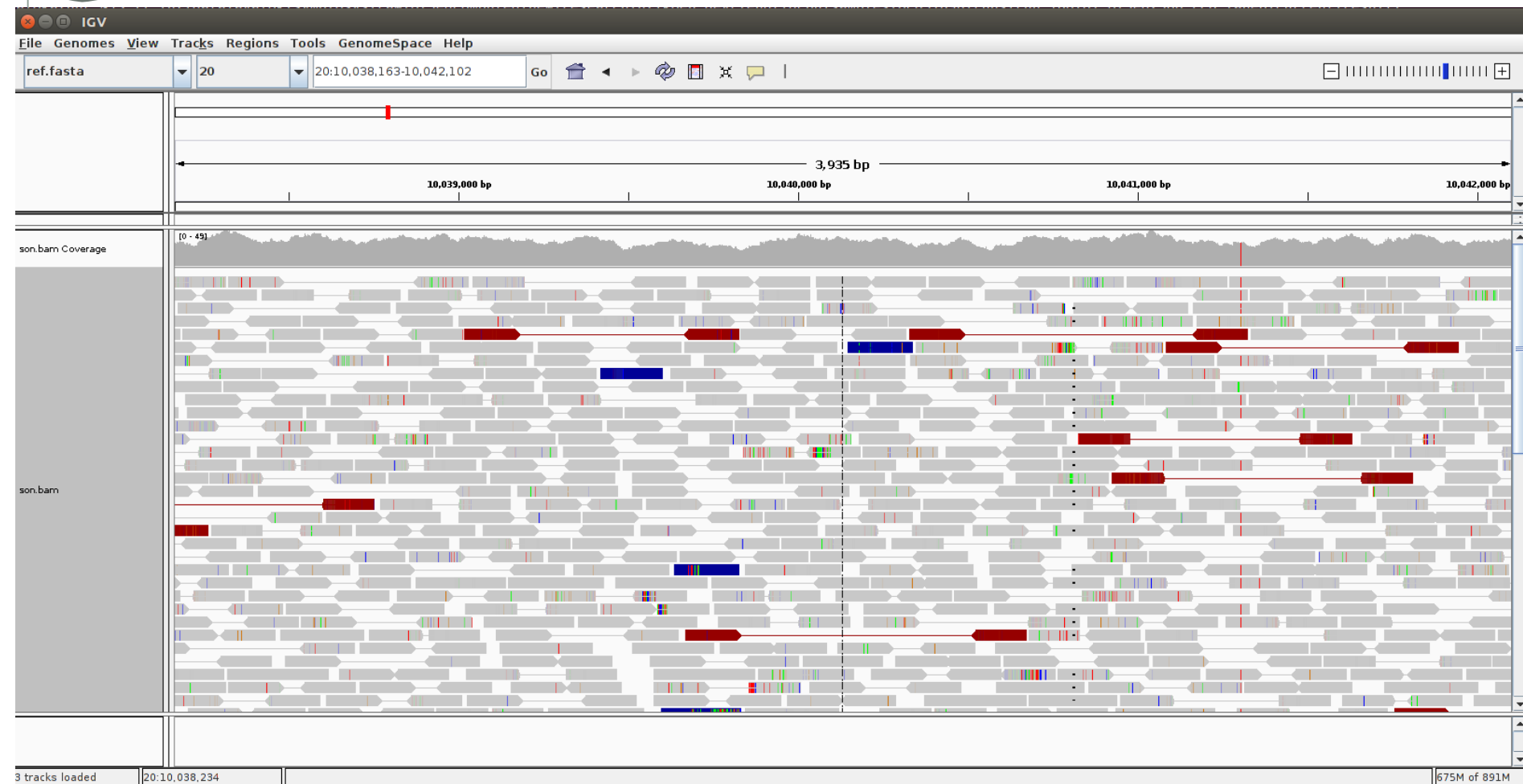
- Questions

- Do we look at Paired End (PE) or Single End (SE) data?
Tip: Put the mouse pointer on the coverage track
- What is the coverage?
Tip: Put the mouse pointer on the colored reads
- Does it look like Whole Genome Sequencing (WGS) or targeted sequencing?
- What's about other BAM files available from the mother?
Tip: Go to 20:15,870,000-15,876,500



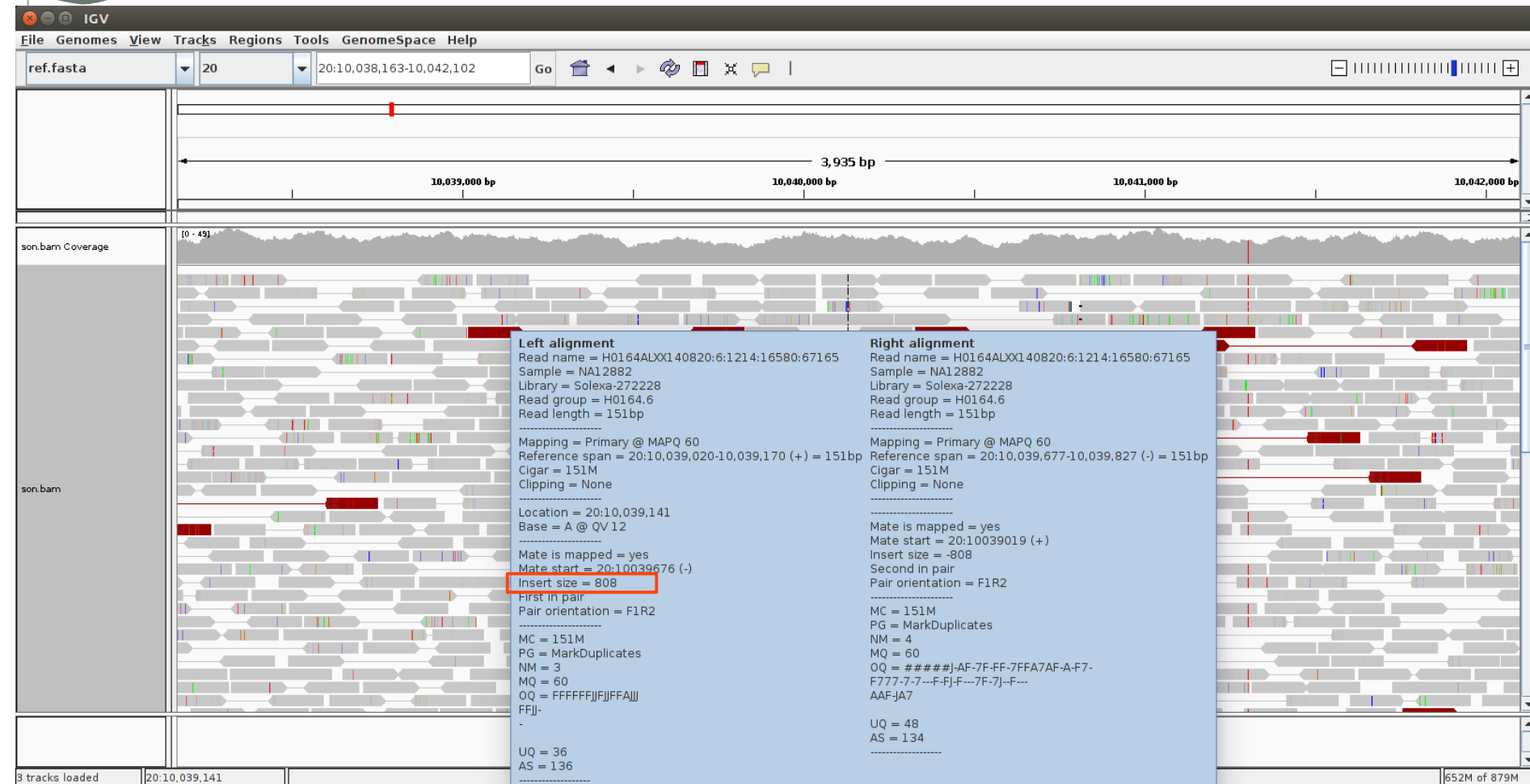
Viewing a BAM file

Read in pairs



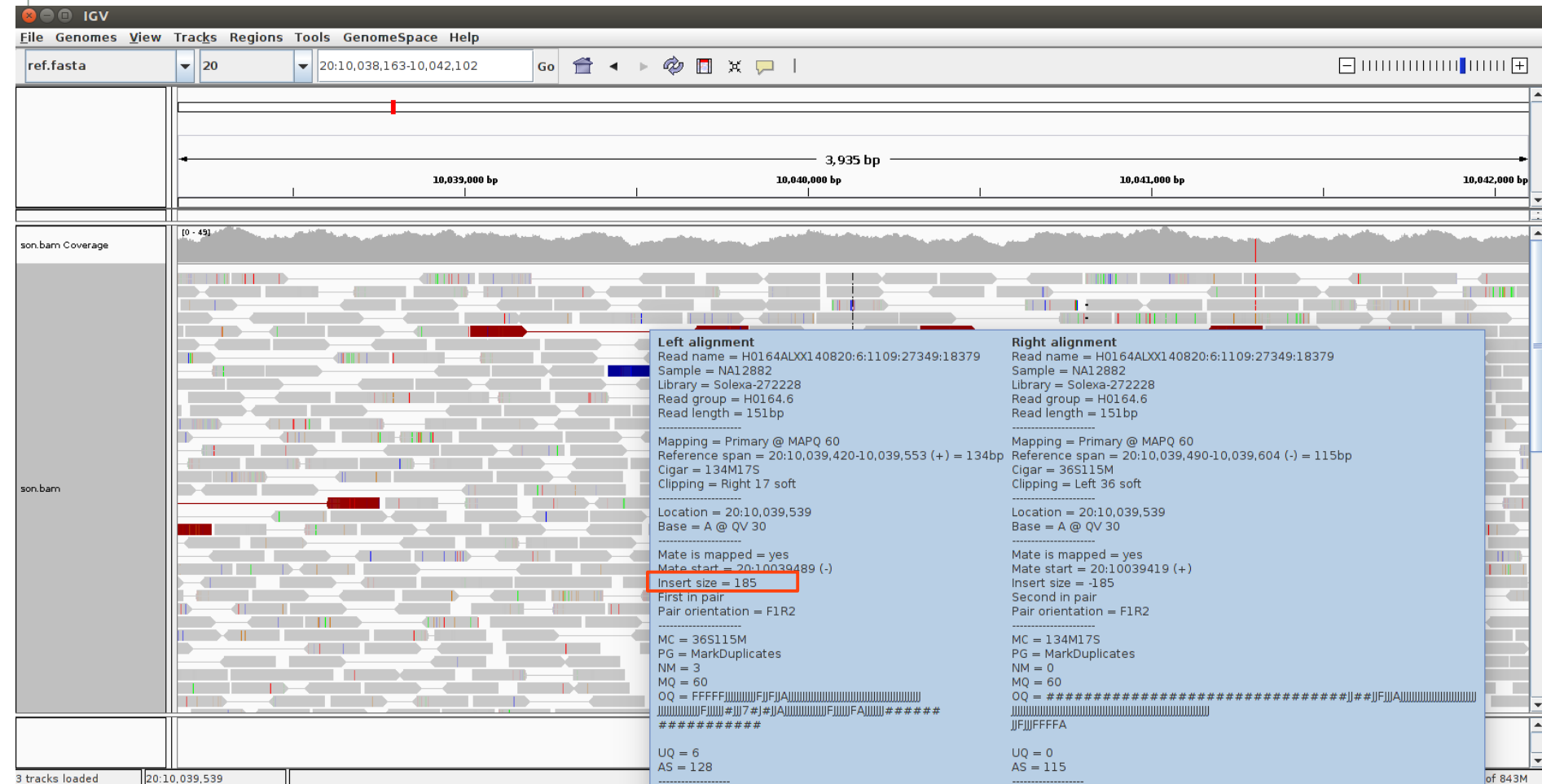
Viewing a BAM file

Large insert size



Viewing a BAM file

Small insert size





Viewing a BAM file

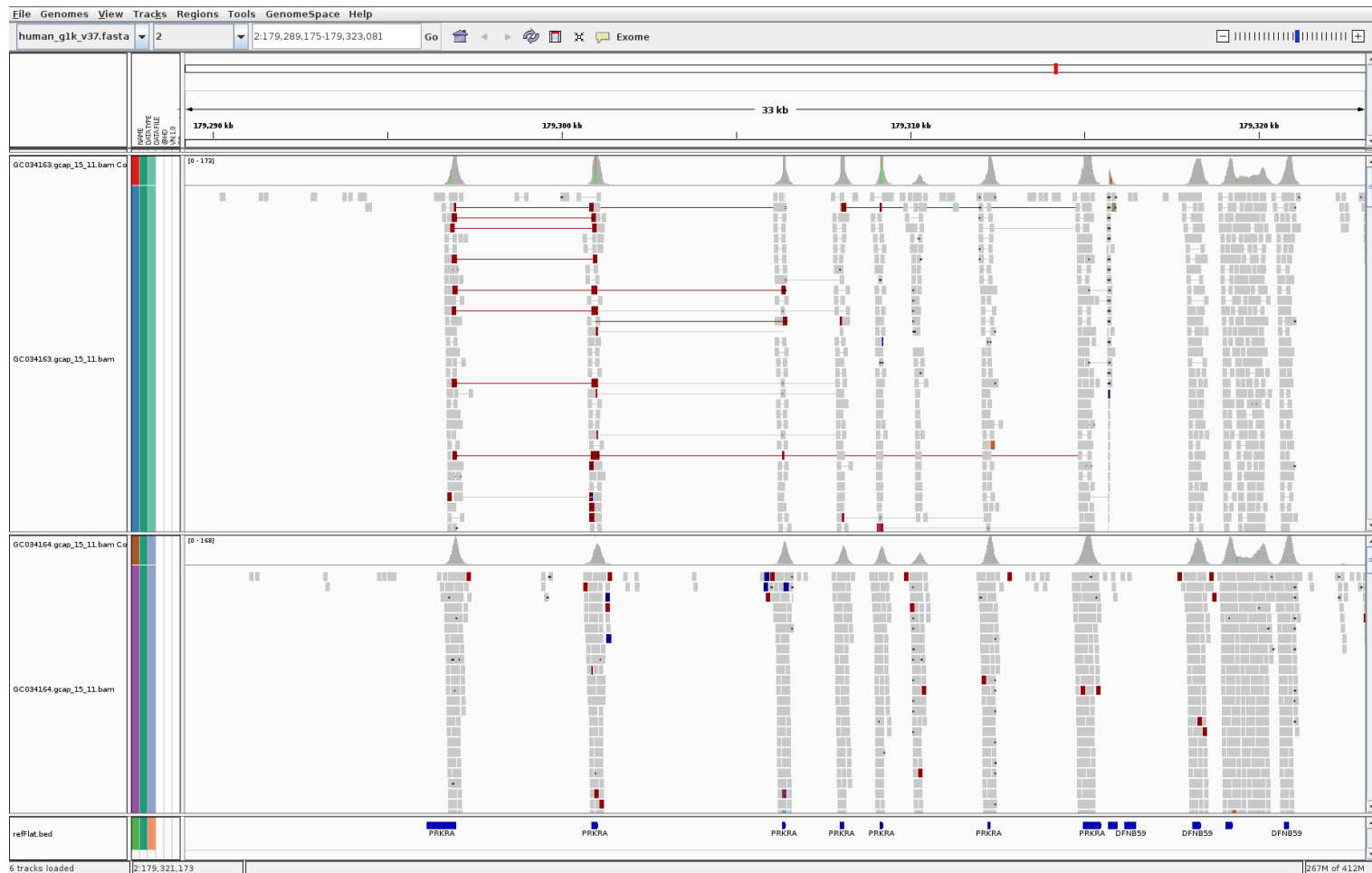
Soft-clipped bases

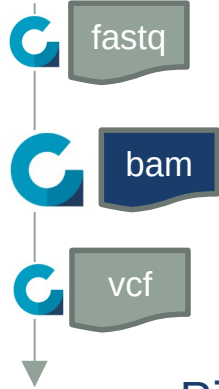




Viewing a BAM file

Special cases





BAM file quality control

- Picard

Set of (command line) tools to

- Manipulate NGS data
 - SAM/BAM/CRAM
 - VCF/BCF
- Compute metrics

<http://broadinstitute.github.io/picard/>

Two examples

- Example 1

Finding the genetic cause of a disease

- Targeted assay (Nimblegen capture)
- ~ 5,800 genes linked to Mendelian disorders
- Illumina PE 126bp

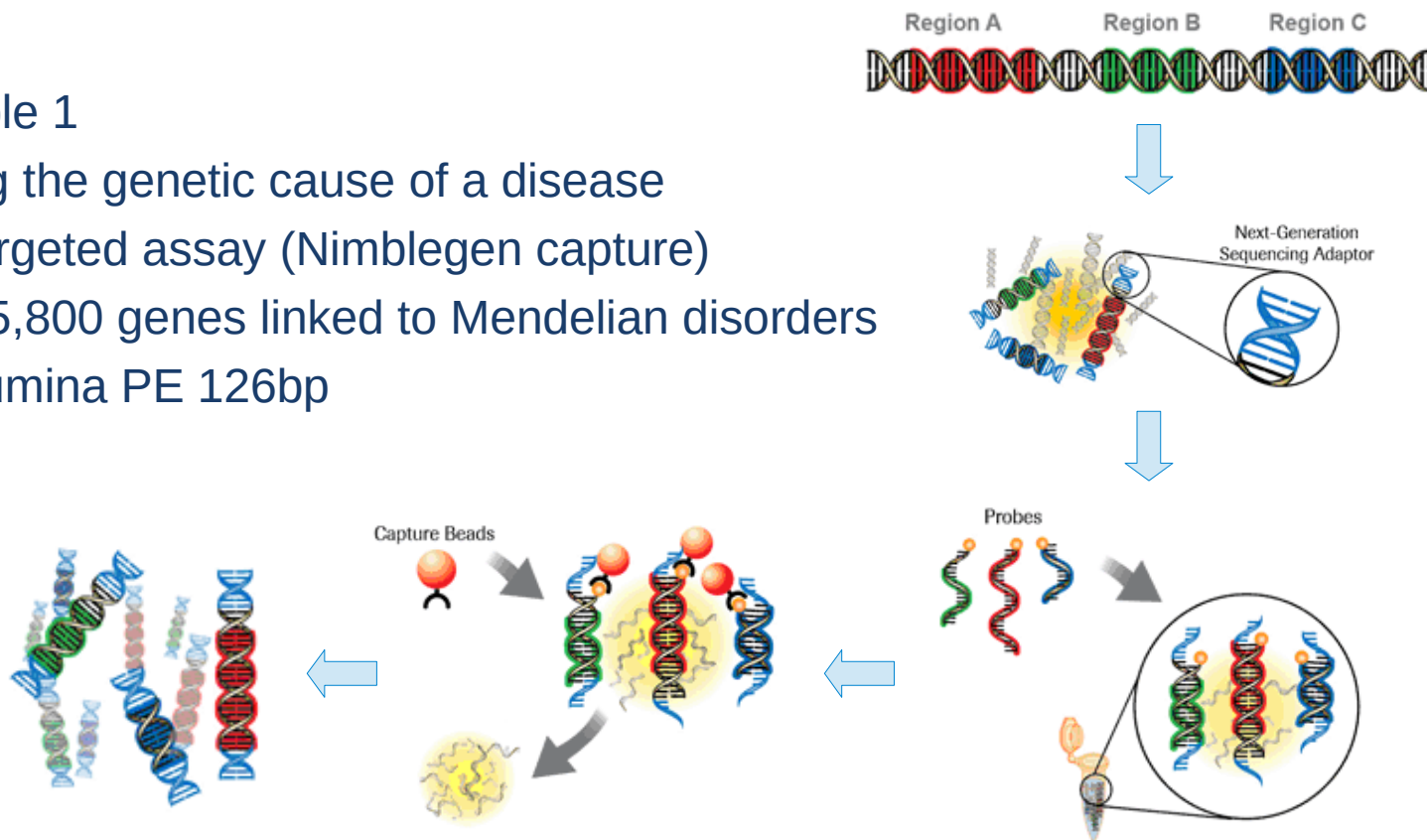
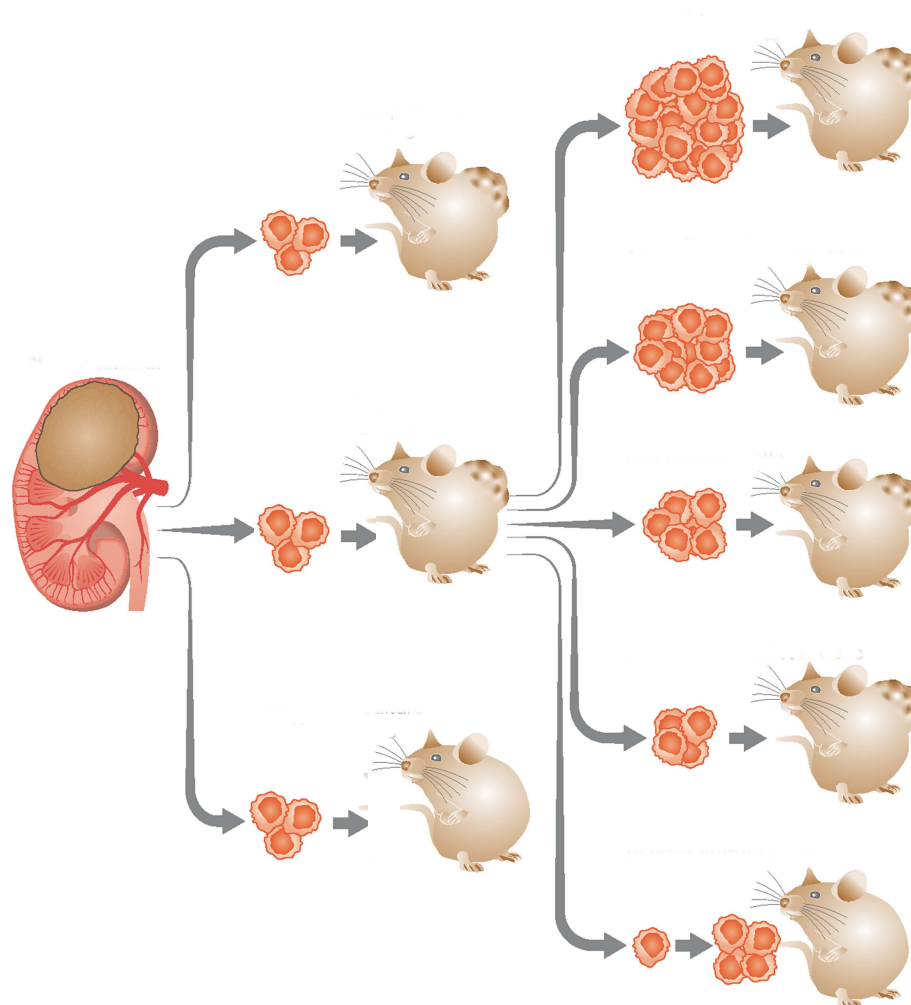


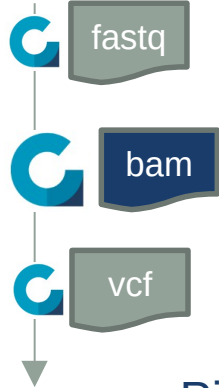
Figure adapted from Nimblegen

Two examples

- Example 2
- Fingerprinting of xenocrafts
 - Amplicon sequencing
 - 31 SNPs
 - Illumina PE 151bp



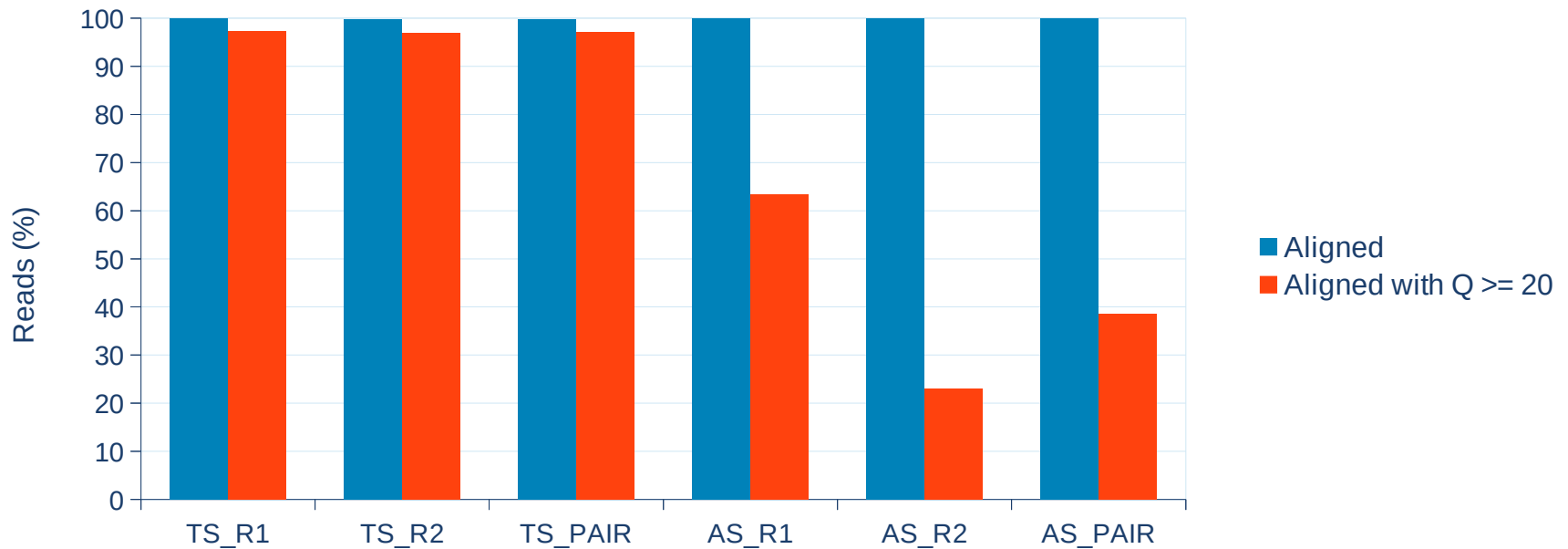
Trace platform, <http://www.uzleuven-kuleuven.be/lki/trace/>
Figure adapted from Peter Hohenstein, EMBO Molecular Medicine: 5 (1), 2013

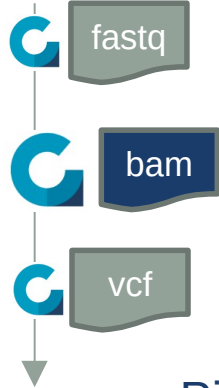


BAM file quality control

- Picard CollectAlignmentSummaryMetrics

```
java -jar ~/bin/picard.jar CollectInsertSizeMetrics  
I=sample.bam  
O=sample_sum_metrics.txt  
R=ref.fasta
```





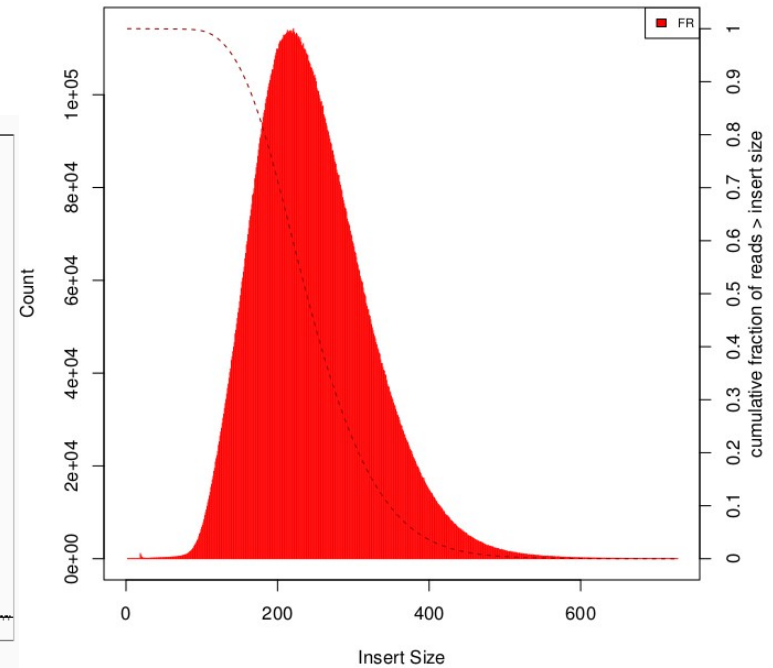
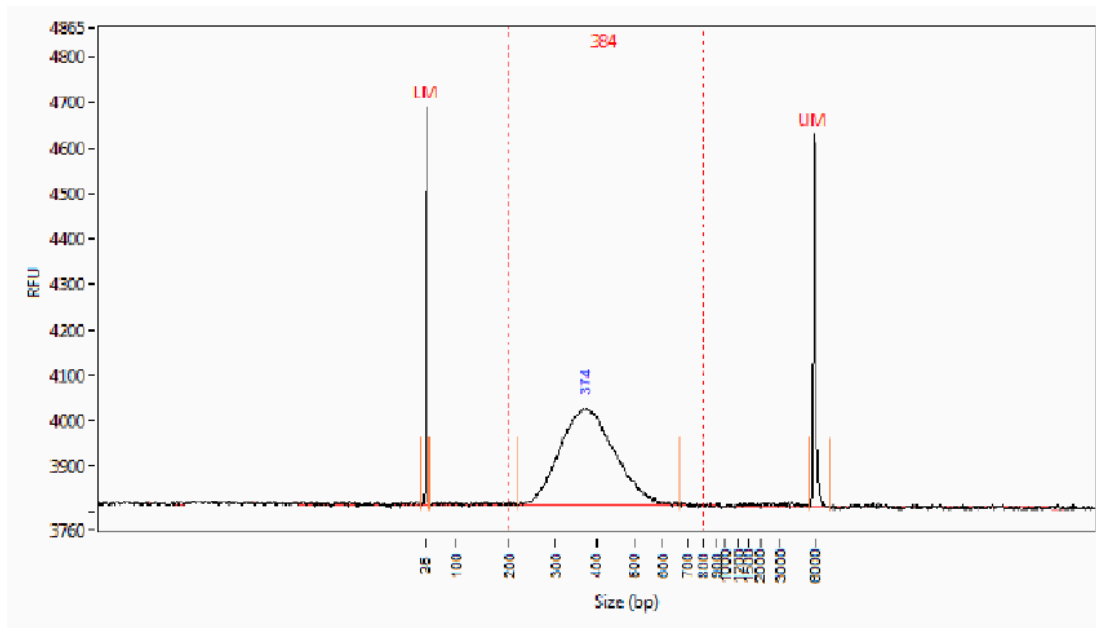
Insert size

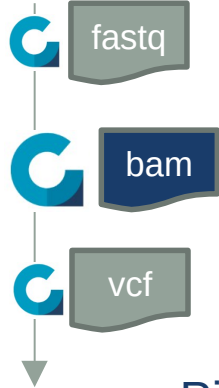
- Picard CollectInsertSizeMetrics

```

java -jar ~/bin/picard.jar CollectInsertSizeMetrics
I=sample.bam
O=sample_insert_size_metrics.txt
H=sample_insert_size_histogram.pdf
  
```

- Example 1





Metrics

Targeted sequencing

- Picard HS metrics

```
java -jar ~/bin/picard.jar CollectHsMetrics  
I=sample.bam  
O=sample_hs_metrics.txt  
R=reference.fasta  
BAIT_INTERVALS=bait.intervals_list  
TARGET_INTERVALS=target.intervals_list
```

- Example 1

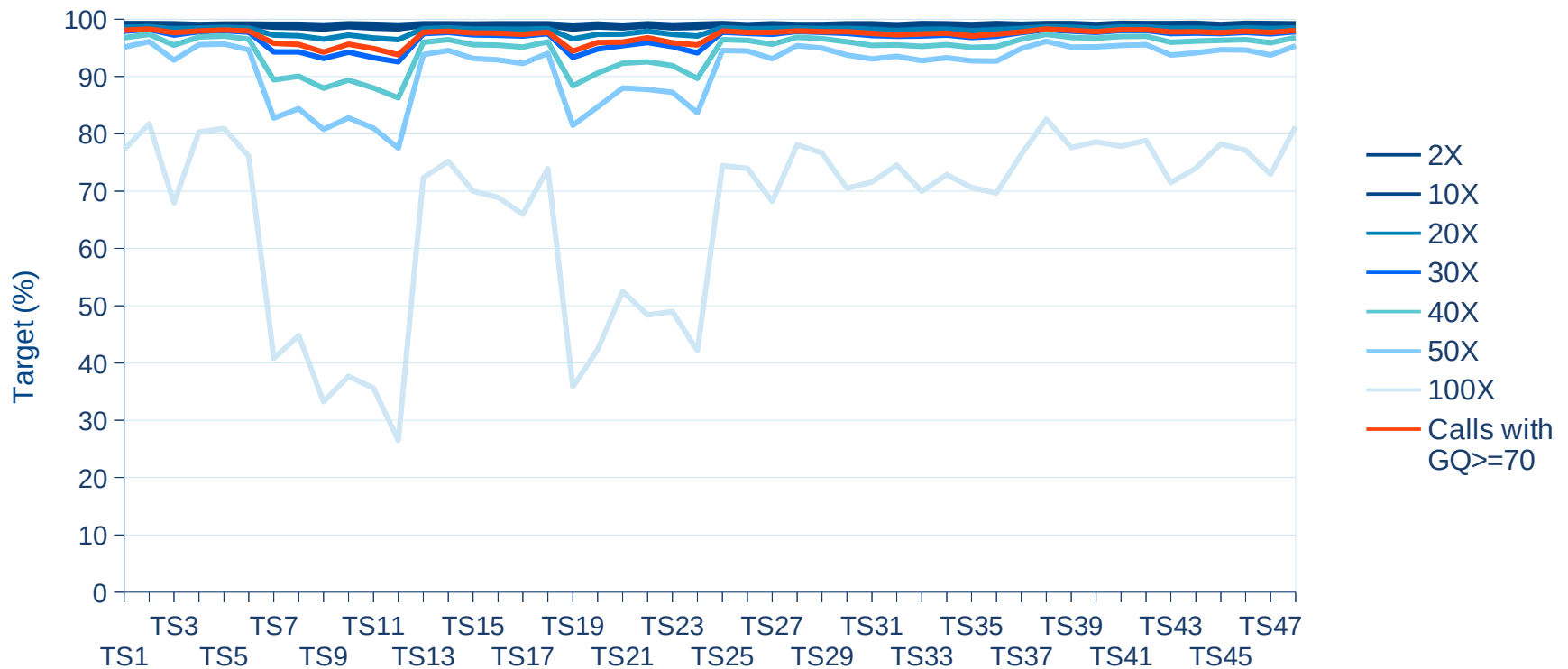
	Mean target coverage	Median target coverage	Selected bases	Fold enrichment
TS_1	120.70	114	0.798	86.93

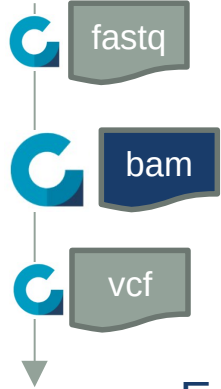


Metrics

Targeted sequencing

- Example 1

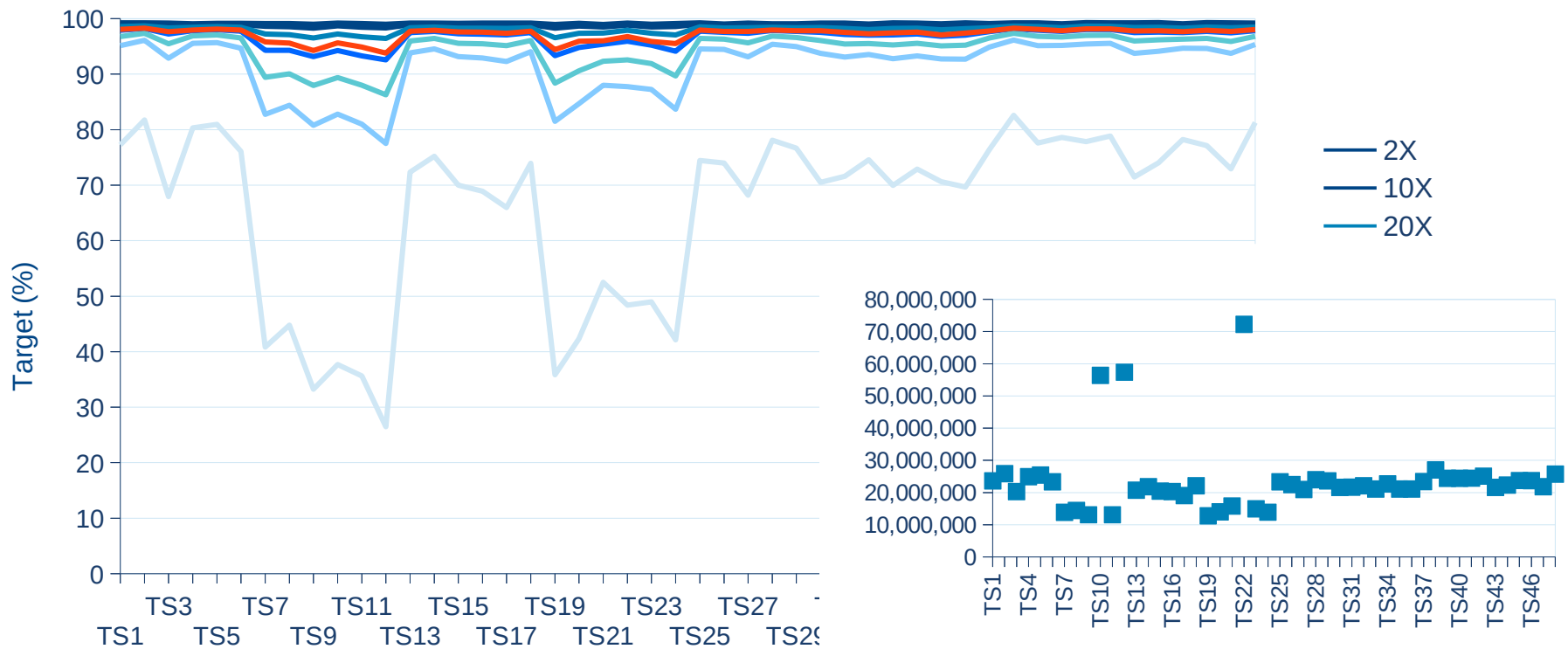


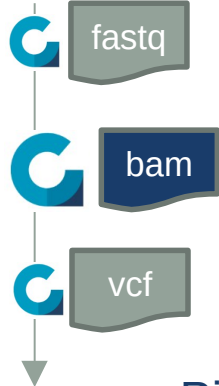


Metrics

Targeted sequencing

- Example 1





Metrics

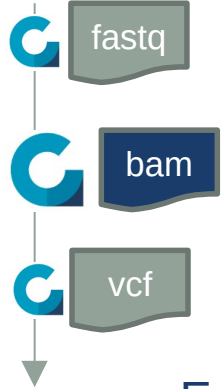
Amplicon sequencing

- Picard PCR metrics

```
java -jar ~/bin/picard.jar CollectInsertSizeMetrics  
I=sample.bam  
O=sample_ipcr_metrics.txt  
R=ref.fasta  
AMPLICON_INTERVALS=amplicon.intervals_list  
TARGET_INTERVALS=target.intervals_list
```

- Example 2

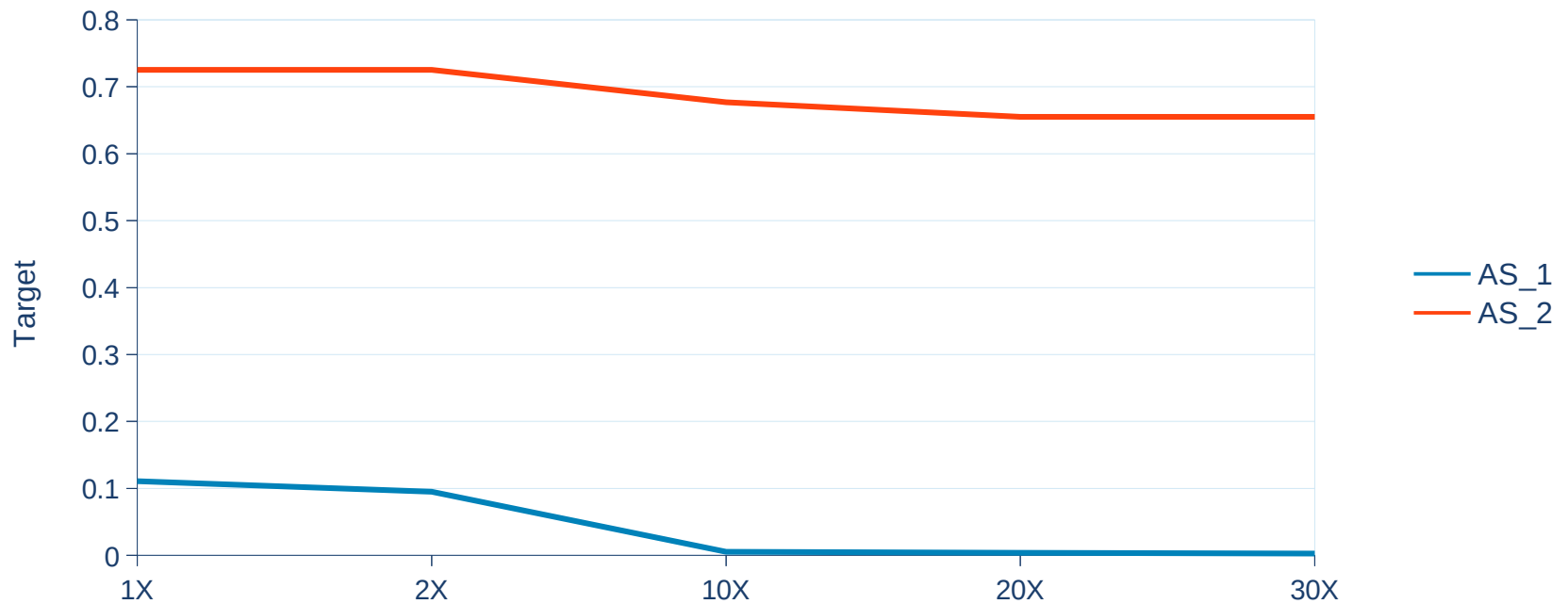
	Mean target coverage	Median target coverage	Fold enrichment
AS_1	0.78	0	5.43
AS_2	359.9	97	44.25

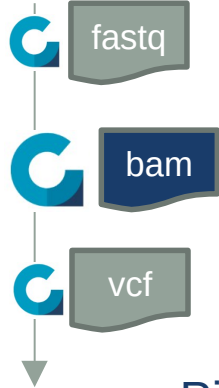


Metrics

Amplicon sequencing

- Example 2





Metrics

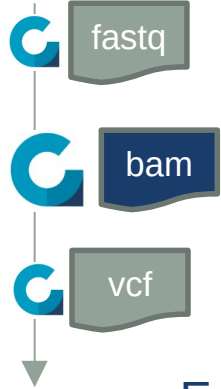
Whole genome sequencing

- Picard WGS metrics

```
java -jar ~/bin/picard.jar CollectWgsMetrics.jar  
I=sample.bam  
O=sample_wgs_metrics.txt  
R:ref.fasta
```

- Example 3

	Mean coverage	Median coverage
WGS_1	30.32	31
WGS_2	30.69	31
WGS_3	30.76	31



Metrics

Whole genome sequencing

- Example 3

