

# *Visualisation of alignment results*

Irina Pulyakhina

WTCHG, RNA-Seq course, 25-April

# Alignment (1)

...ACTTTTACCGACGCA...

TTTT



# Alignment (2)

```
...ACTTTTACCGAACGTACGTACGTAGCTAGTCGATGCTAGCCGCA...  
      TTTT   CGAA GTAC  ACGT GCTAG   GATG TAGC  
        TTTA           CGTA GTAG       ATGC  GCCG  
      CTTT                TAGC         TGCT  CCGC
```

# Alignment (3)

```

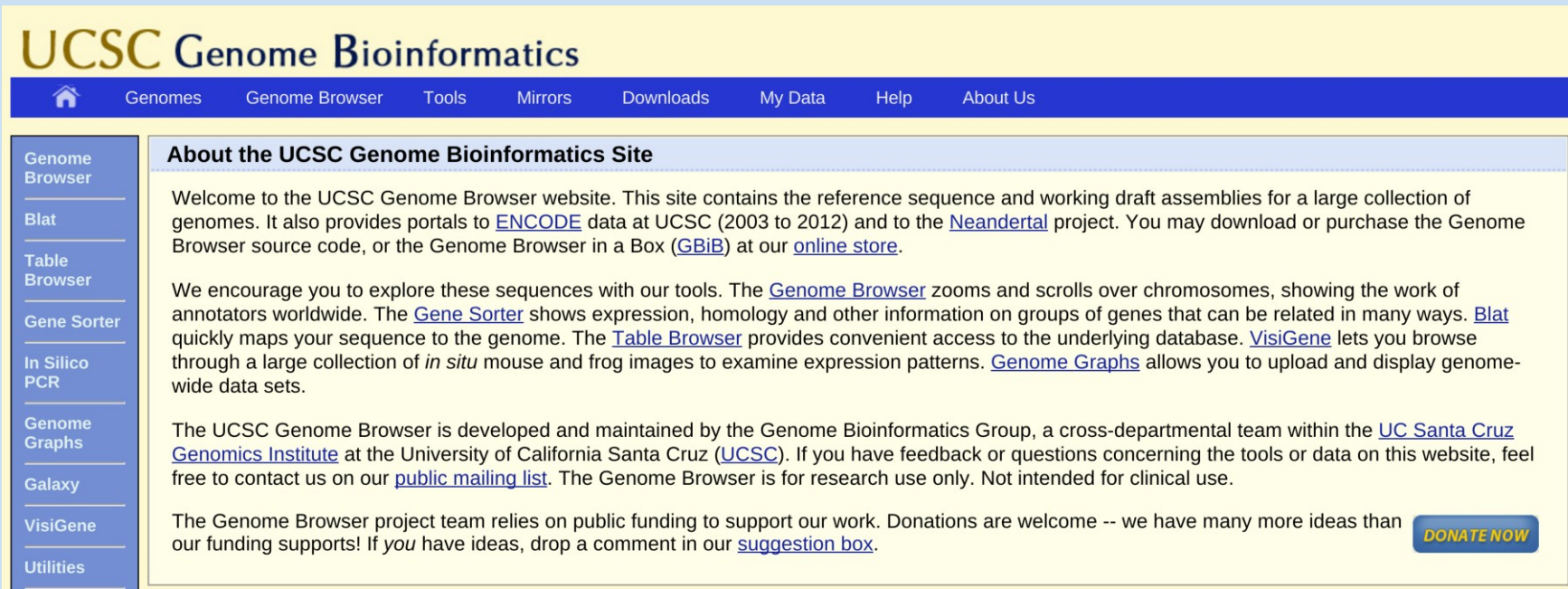
                                     AAAAA
                                     AAAAA
                                     AAAAA
                                     GAAAAA
                                     ATAAATGCAAAAA
                                     AGCATAGATGCAAAAA
                                     AGCATAGATGCAAAAA
                                     AAAAGTACAGGCCAATATACCTAATGAGCATAGATG
                                     AAAAACTACAGGCCAATATACCTAATGAGCATAGAT
                                     CAAAACTACAGGCCAATATACCTAATGAGCATAGAT
                                     AACAAAACCTACAGGCCAATATACCTAATGAGCATAG
                                     AACAAAACCTACAGGCCAATATACCTAATGAGCATAG
                                     AAAAAAGCAAAAGCTACAGGCCAATATACCTAATGAGC
                                     AAAAAAAACAAAACCTGACAGGCCAATATACCTAATGAG
                                     AAAAAAACAAAACCTACAGTCCAATATACCTAATGAG
                                     AAAAAAACAAAACCTACAGGCCAATATACCTAATG
                                     AAAAAAACAAAACCTACAGGCCAATATACCTAATG
                                     AAAAAAACAAAACCTGACAGGCCAATATACCT
                                     AACAAAACAAAACCTACAGGCCAATATACCT
IAAACCAGGAGAGGACATAACAAAAAAACAA
IAAACCAGGAGAGGACATAACAAAAAAACAA
IAAACCAGGAGAGGACAT      AAAAAAACAAAACCTACAGGCCAATATACCTAATGA
IAAACCAGGA      AAAAAAACAAAACCTACAGGCCAATATACCT      GCATAGATGCAAAAA
IAAACCAG      AAAAAAACAAAACCTACAGGCCAATATACCT      GCATAGATGCAAAA
IAAACCAGGAGAGGACATAACAAAAAAACAAAACCTACAGGCCAATA      AATGAGCATAGATGCAAAAA
IAAACCAGGAGAGGACATAACAAAAAAACAAAACCTACAGGCCAATATACCTAATGAGCATAGATGCAAAAA
IAAACCAGGAGAGGAC      AAAAAAACAAAACCTACAGGCCAATATACCTAATG      AGATGCAAAAA
IAAACCAGGAGAGGA      AAAAAAACAAAACCTACAGGCCAATATACCTAAT      TAGATGCAAAAA
IAAACCAGGAG      AAAAAAACAAAACCTACAGGCCAATATAC      GAAACATAGAAAGCAAAAA
IAAACCAGGA      ATAAAAAACAAAACCTACAGGCCAATATAC      GAGCATAGATGCAAAAA
IAAACCAG      GGAATAAAAAACAAAACCTACAGGCCAAT      TAATGAGCATAGATGCAAAAA
IAAACCAGGAGAGGACATAACAAAAAAACAAAACCTACAGGCCAATATACCTAATGAGCATAGATGCAAAAA

```

Need **efficient** ways to visualize alignment data.

# Genome browsers (1)

UCSC genome browser – interactive web service,  
<https://genome-euro.ucsc.edu/index.html>



The screenshot shows the UCSC Genome Bioinformatics website. The header features the site name and a navigation bar with links: Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. A left sidebar lists various tools: Genome Browser, Blat, Table Browser, Gene Sorter, In Silico PCR, Genome Graphs, Galaxy, VisiGene, and Utilities. The main content area, titled 'About the UCSC Genome Bioinformatics Site', contains a welcome message, a description of the site's resources (including ENCODE data and the Neandertal project), a list of tools and their functions, and information about the development team at the UC Santa Cruz Genomics Institute. It also includes a 'DONATE NOW' button.

**UCSC Genome Bioinformatics**

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

**About the UCSC Genome Bioinformatics Site**

Welcome to the UCSC Genome Browser website. This site contains the reference sequence and working draft assemblies for a large collection of genomes. It also provides portals to [ENCODE](#) data at UCSC (2003 to 2012) and to the [Neandertal](#) project. You may download or purchase the Genome Browser source code, or the Genome Browser in a Box ([GBiB](#)) at our [online store](#).

We encourage you to explore these sequences with our tools. The [Genome Browser](#) zooms and scrolls over chromosomes, showing the work of annotators worldwide. The [Gene Sorter](#) shows expression, homology and other information on groups of genes that can be related in many ways. [Blat](#) quickly maps your sequence to the genome. The [Table Browser](#) provides convenient access to the underlying database. [VisiGene](#) lets you browse through a large collection of *in situ* mouse and frog images to examine expression patterns. [Genome Graphs](#) allows you to upload and display genome-wide data sets.

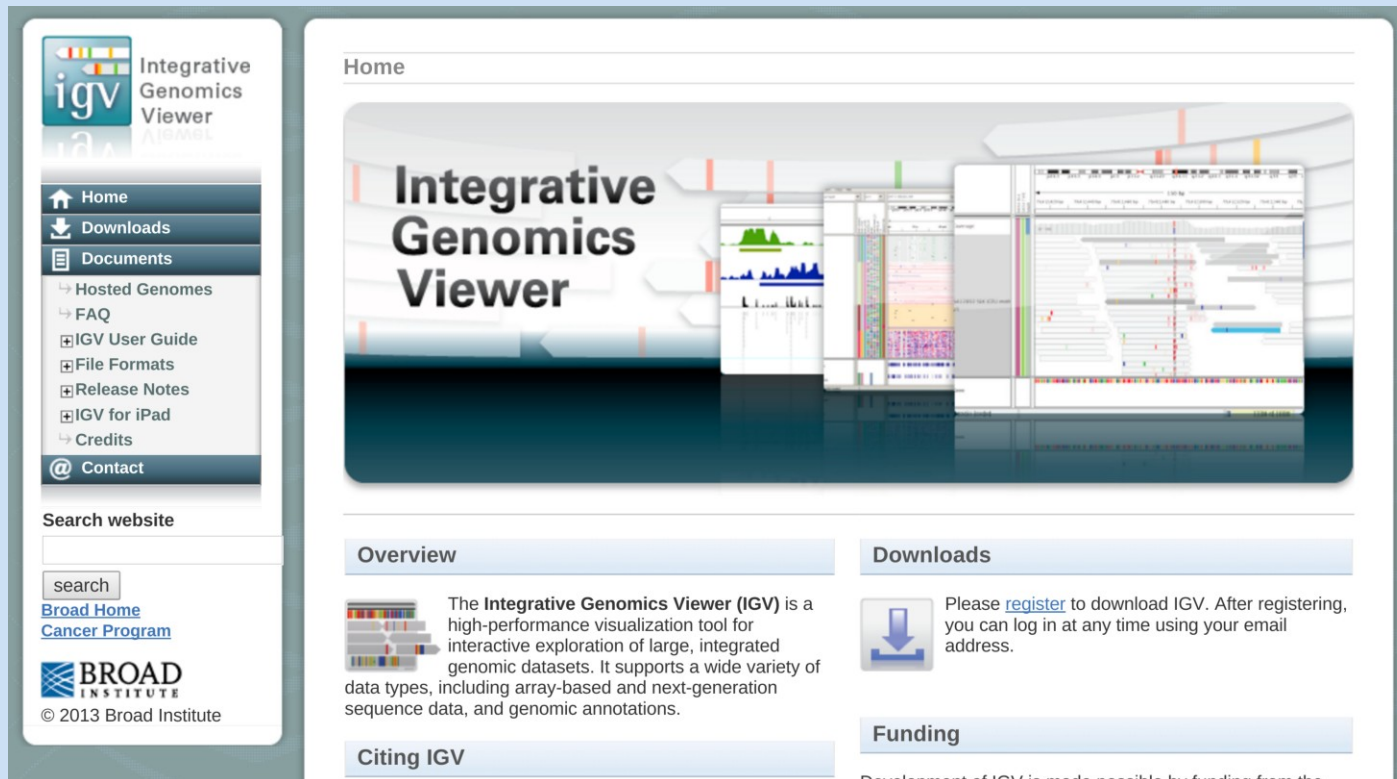
The UCSC Genome Browser is developed and maintained by the Genome Bioinformatics Group, a cross-departmental team within the [UC Santa Cruz Genomics Institute](#) at the University of California Santa Cruz ([UCSC](#)). If you have feedback or questions concerning the tools or data on this website, feel free to contact us on our [public mailing list](#). The Genome Browser is for research use only. Not intended for clinical use.

The Genome Browser project team relies on public funding to support our work. Donations are welcome -- we have many more ideas than our funding supports! If you have ideas, drop a comment in our [suggestion box](#).

[DONATE NOW](#)

# Genome browsers (2)

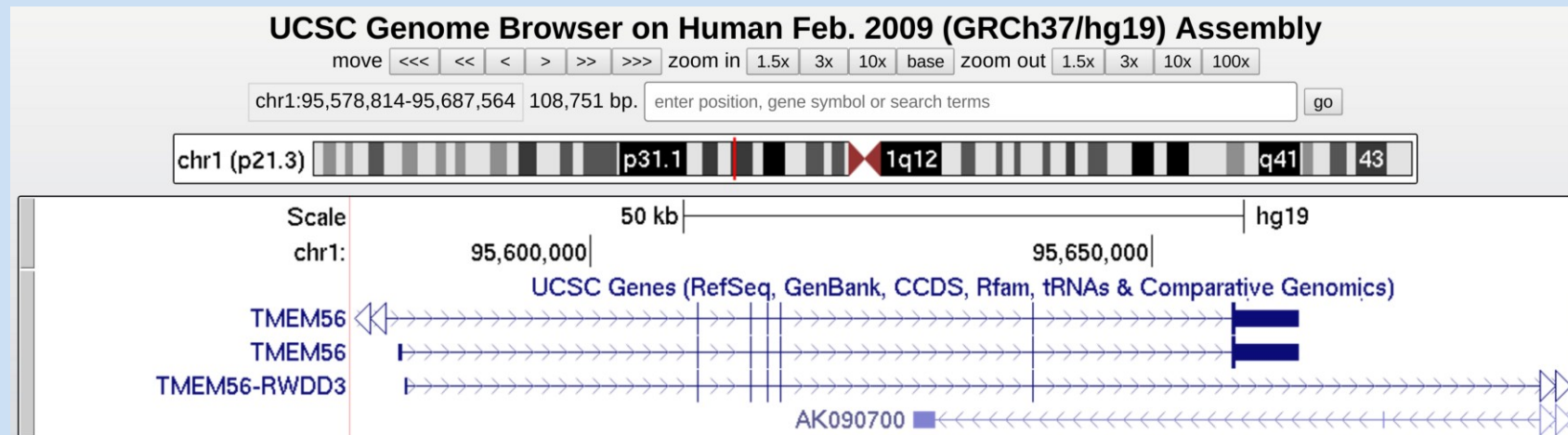
IGV – integrative genome viewer, to install locally  
<https://www.broadinstitute.org/igv/>



The screenshot shows the IGV website homepage. On the left is a navigation sidebar with links: Home, Downloads, Documents, Hosted Genomes, FAQ, IGV User Guide, File Formats, Release Notes, IGV for iPad, Credits, and Contact. Below the sidebar is a search bar and the Broad Institute logo. The main content area features a large banner with the text 'Integrative Genomics Viewer' and a visual representation of the IGV interface showing genomic tracks. Below the banner are sections for 'Overview' (describing IGV as a high-performance visualization tool), 'Downloads' (with a download icon and registration instructions), 'Funding' (mentioning funding from the Broad Institute), and 'Citing IGV'.

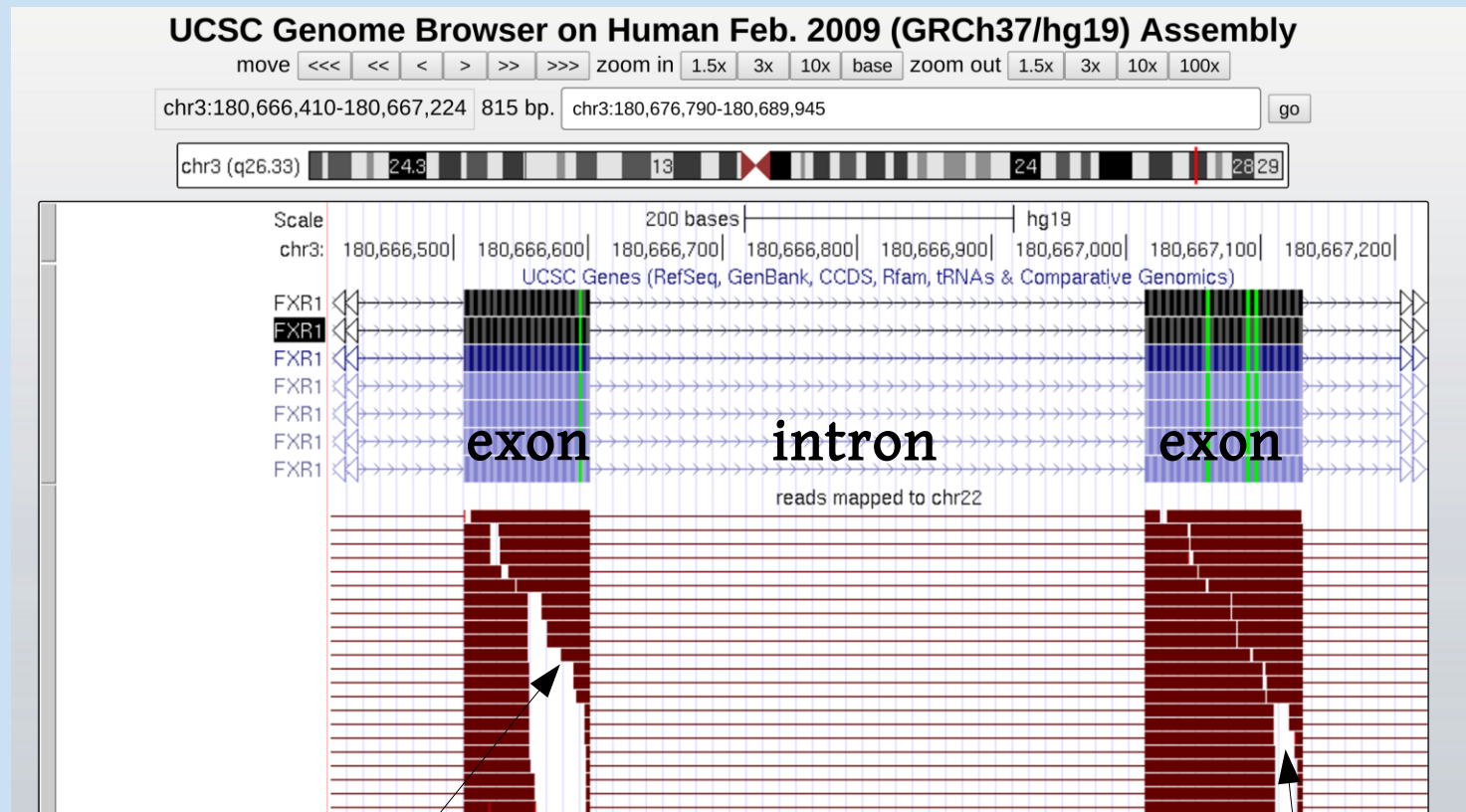
# Genome browsers (3)

Concept – move around the genome to zoom-in on the areas of interest.





# Visualizing BAM files (1)



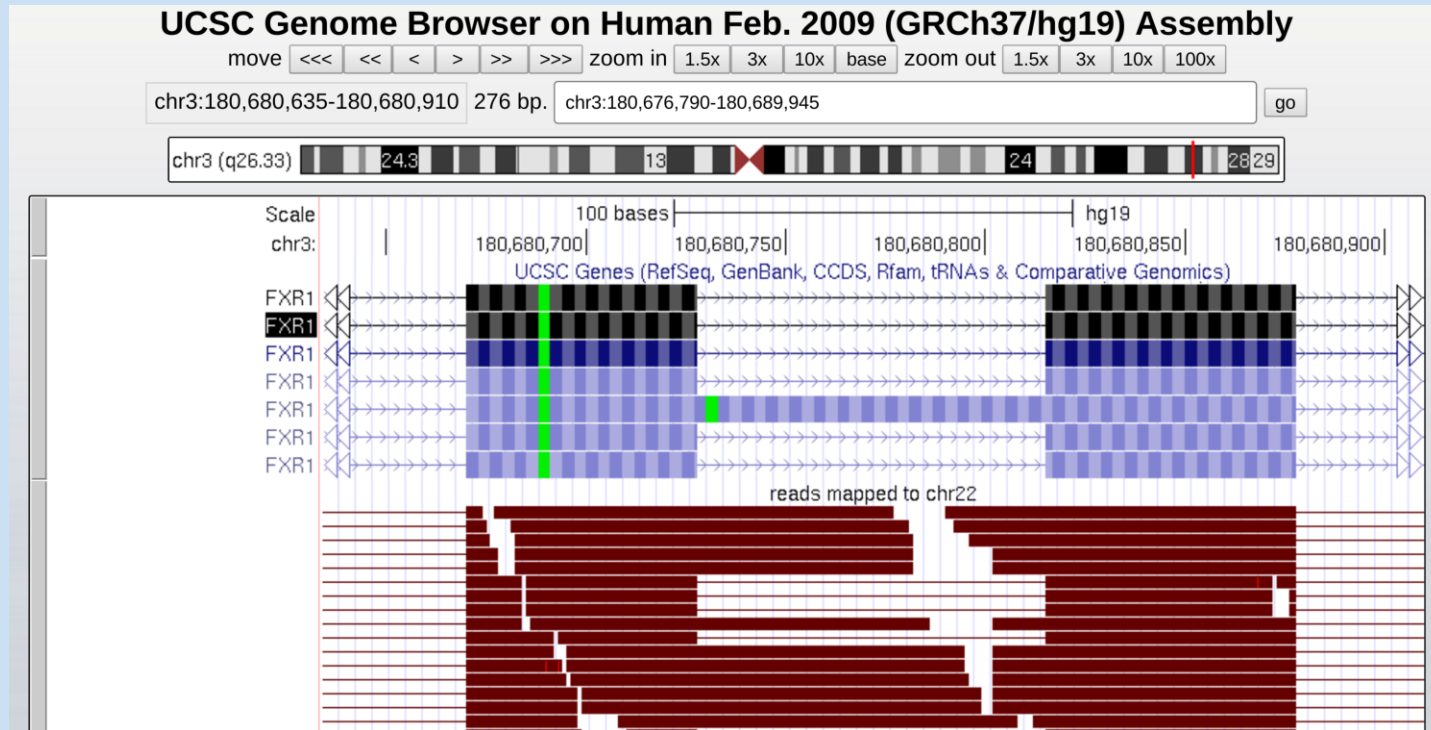
Part of a read  
mapped to an exon

Read was split here

Part of a read  
mapped to an exon



# Visualizing BAM files (2)



Some reads fully map to this intron – indication of a splicing event called intron retention (which can also be seen in the annotation).

# Wiggle and BigWig files

Wiggle file – per position, number of reads mapped to this position a.k.a. **coverage**.

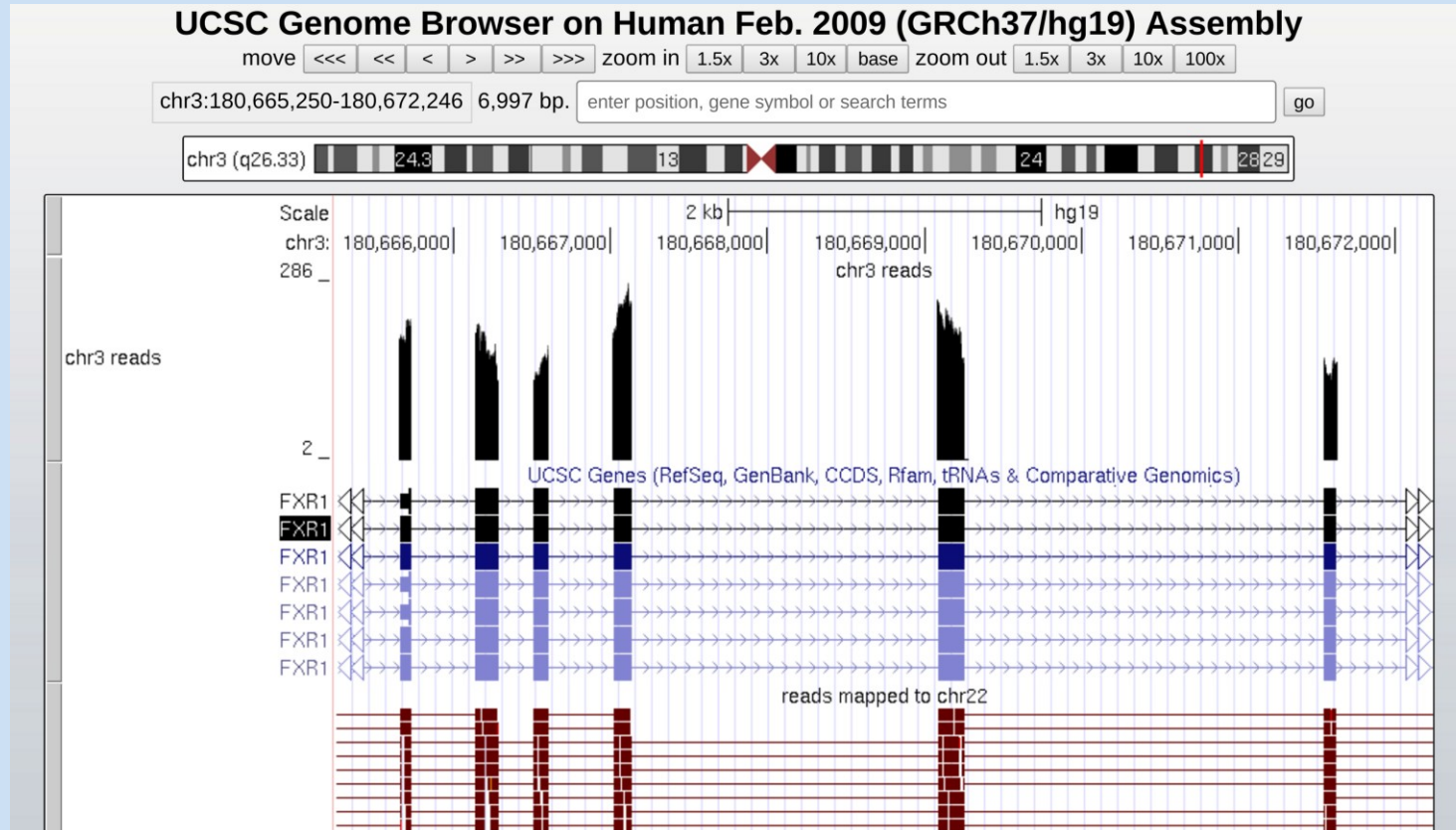
```
      1 2 3 4 5 6 7 8  
...ACTTTTAC...  
      TTTT  
      TTTA  
      CTTT
```

Wiggle file:

```
chromosome=chrN  
1    0  
2    1  
3    2  
4    3  
5    4  
6    2  
7    1  
8    0  
...
```

BigWig – binary Wiggle file.

# Visualising BigWig files



# BED files

```
...ACTTTTACCGAACGTACGTACGTAGCTAGTCGATGCTAGCCGCA...  
TT-CC-----AGCT-----CTAGCC
```

BAM



BED

BED files are rarely used to visualise raw RNA-Seq data.

# BED and BEDPE files

```
      1 2 3 4 5                27                36        41
...ACTTTTACCGAACGTACGTACGTAGCTAGTCGATGCTAGCCGCA...
      TT-CC-----AGCT-----CTAGCC
```

BAM

BED file:

```
chrN    5    27
chrN    36   41
```

BEDPE file:

```
chrN    5    27      chrN    36      41
```

BEDPE file retains the connection between ends, but is not supported by the majority of genome browsers.

# Demo

[https://genome-euro.ucsc.edu/cgi-bin/hgTracks?  
hgS\\_doOtherUser=submit&hgS\\_otherUserName=pulyakhina  
&hgS\\_otherUserSessionName=2016%2DRNAseq\\_course](https://genome-euro.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=pulyakhina&hgS_otherUserSessionName=2016%2DRNAseq_course)