

ASHG Communications Workshop

Helga Thorvaldsdóttir, Broad Institute

Jim Robinson, UC San Diego

Integrative Genomics Viewer (IGV)

ASHG 2018, San Diego

ASHG Communications Workshop

Helga Thorvaldsdóttir, Broad Institute

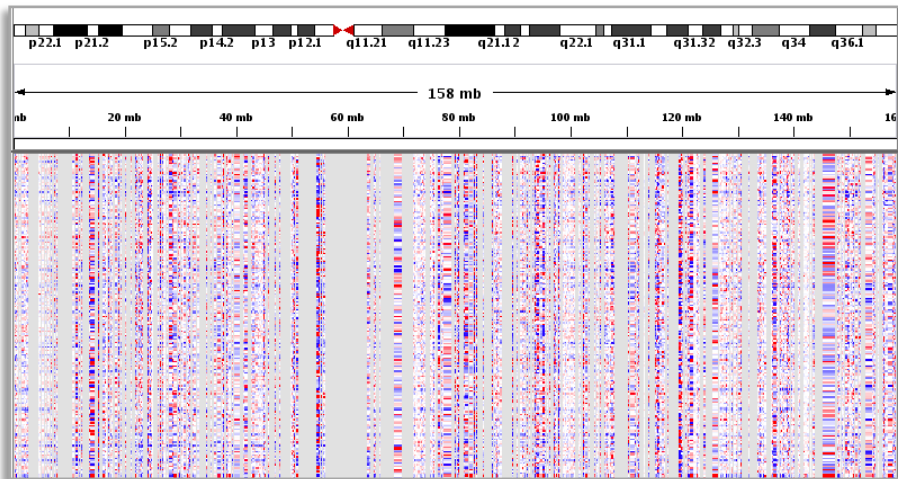
Jim Robinson, UC San Diego

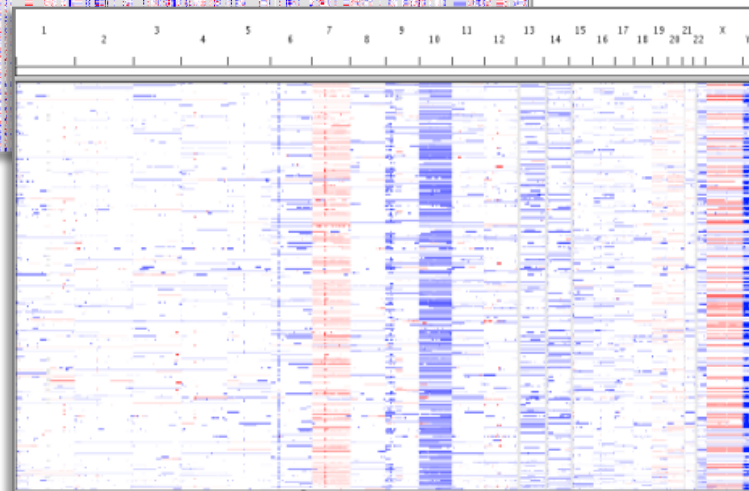
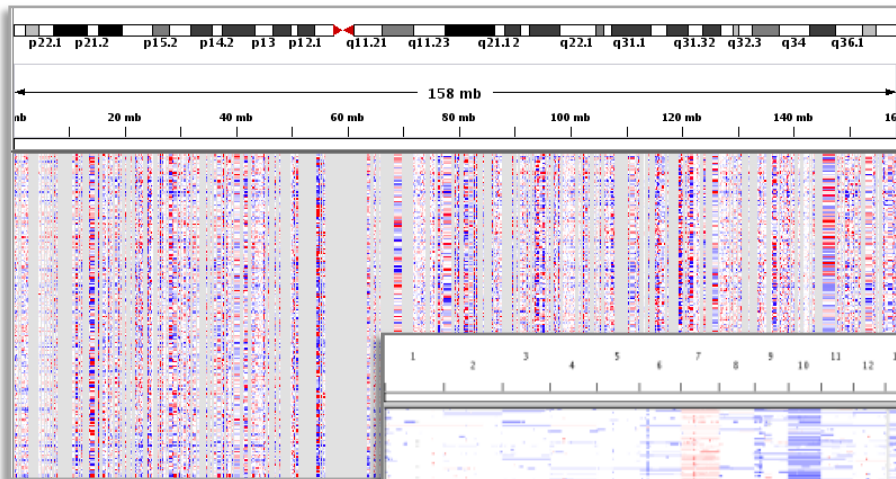
Integrative Genomics Viewer (IGV)

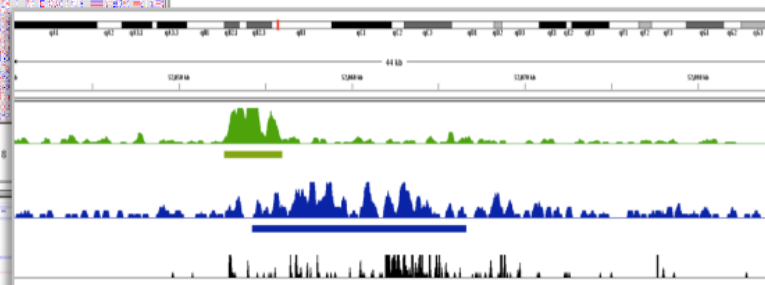
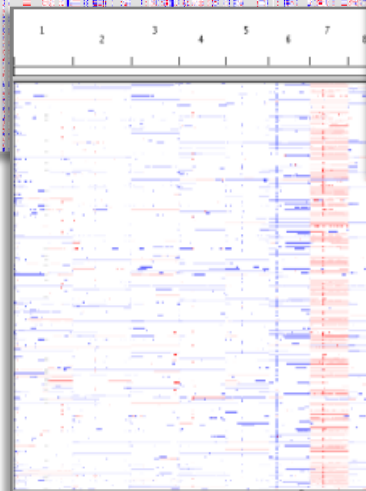
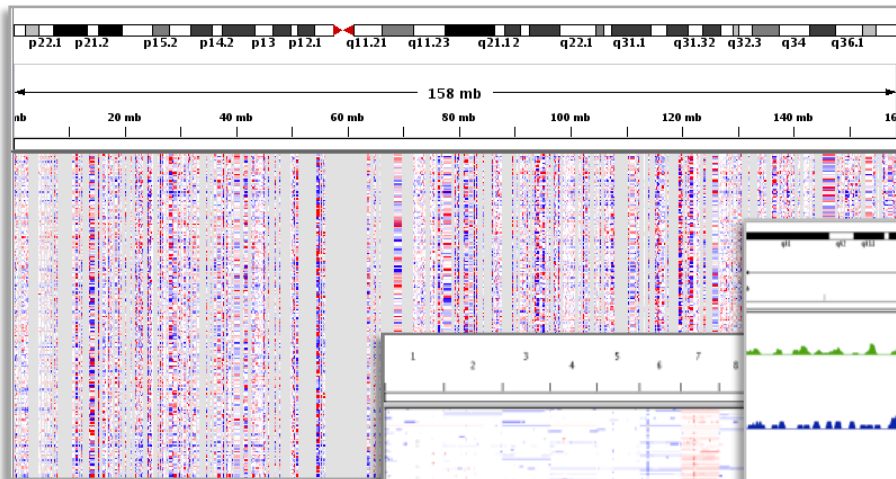
ASHG 2018, San Diego

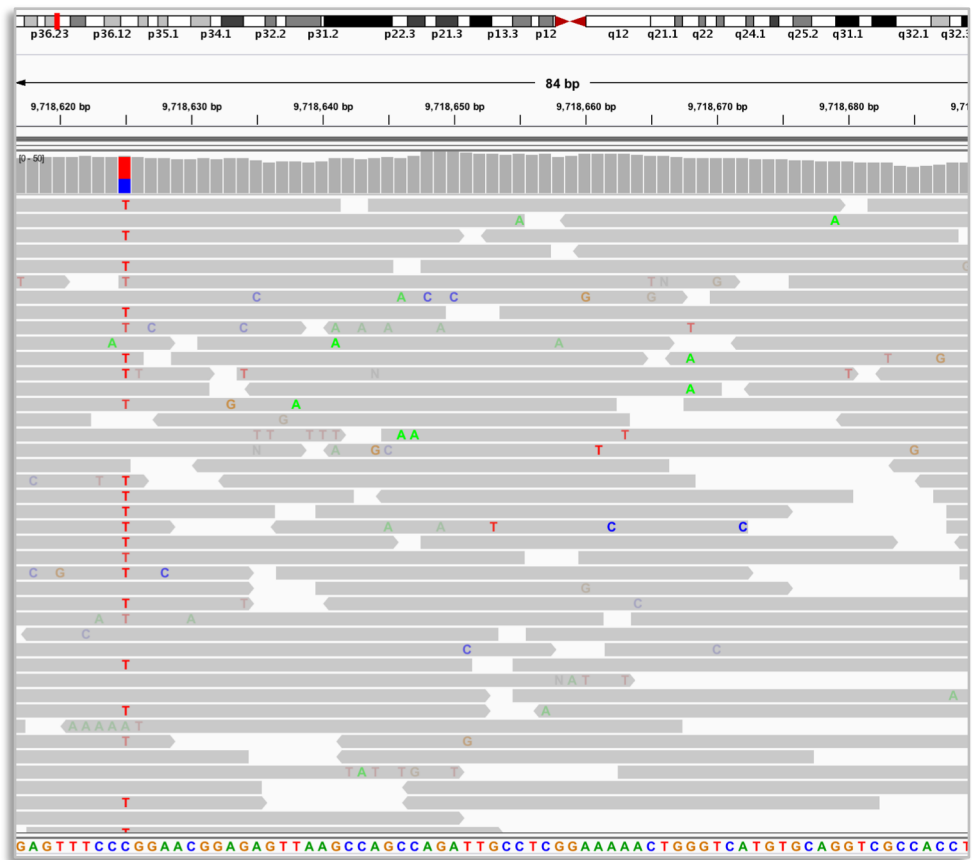
1. Visualizing genome variants in NGS data
2. Interactive figures

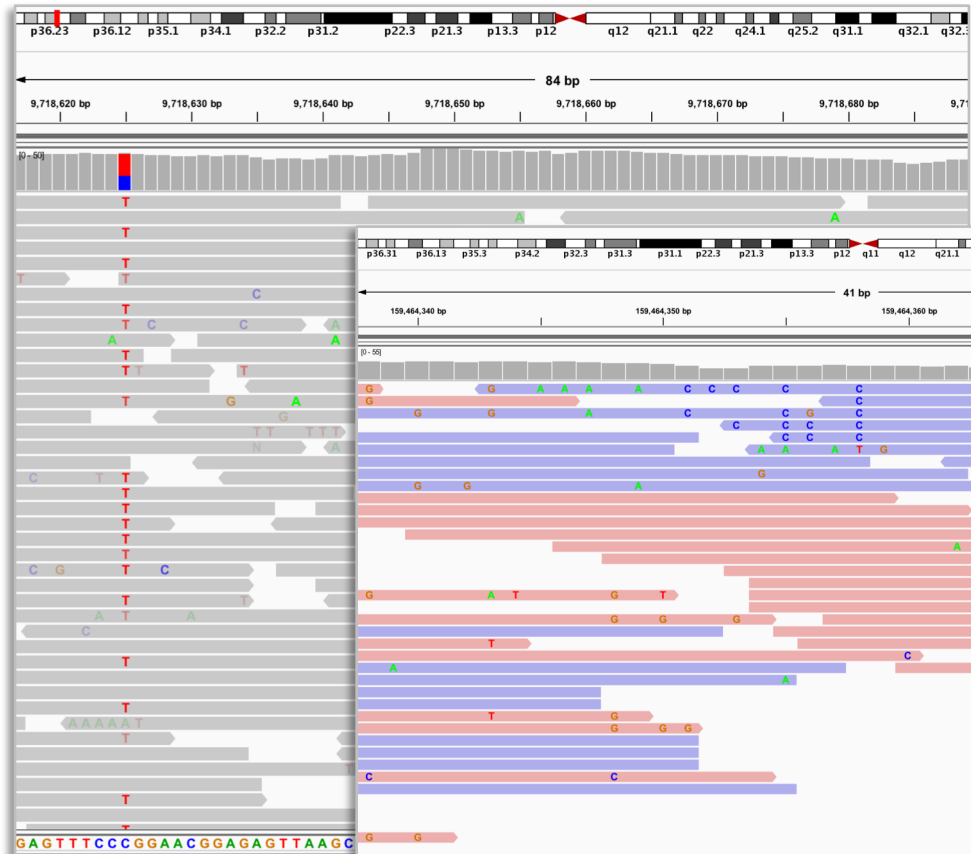
1. Data visualization: aligned sequence data
2. Interactive figures

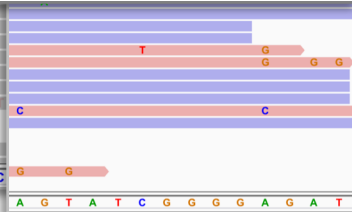
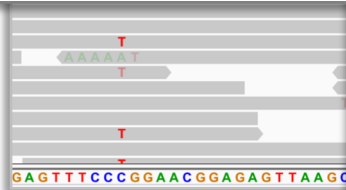
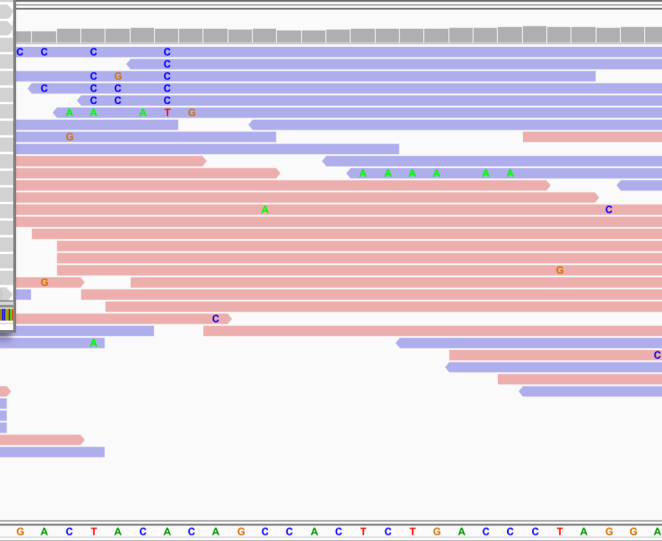
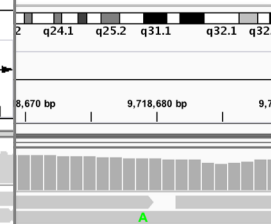
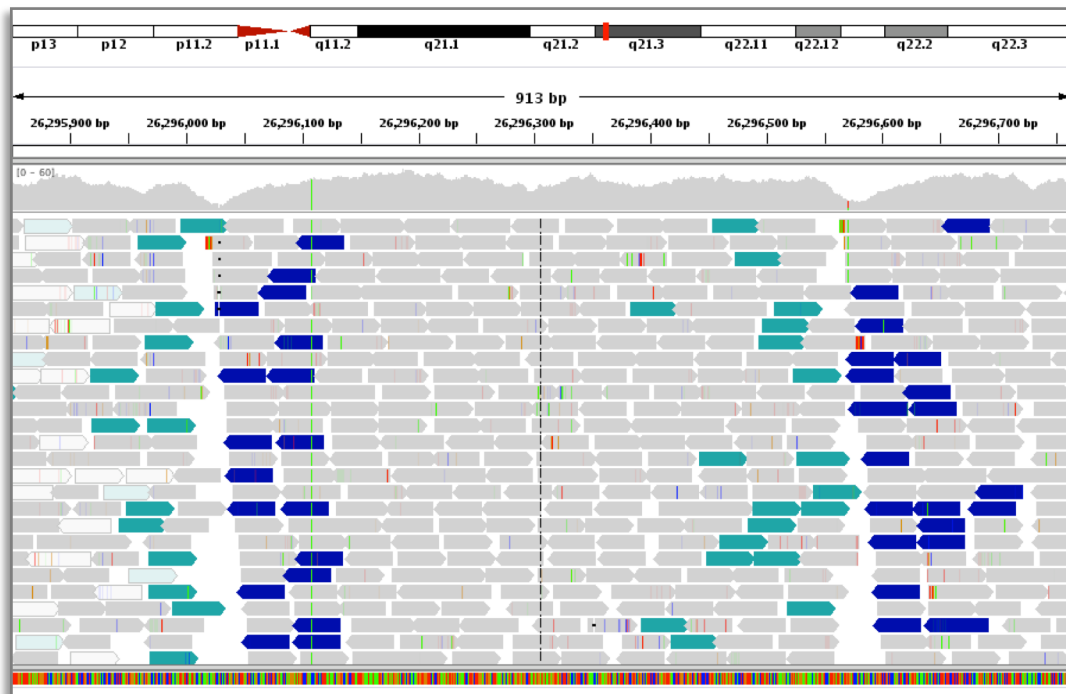


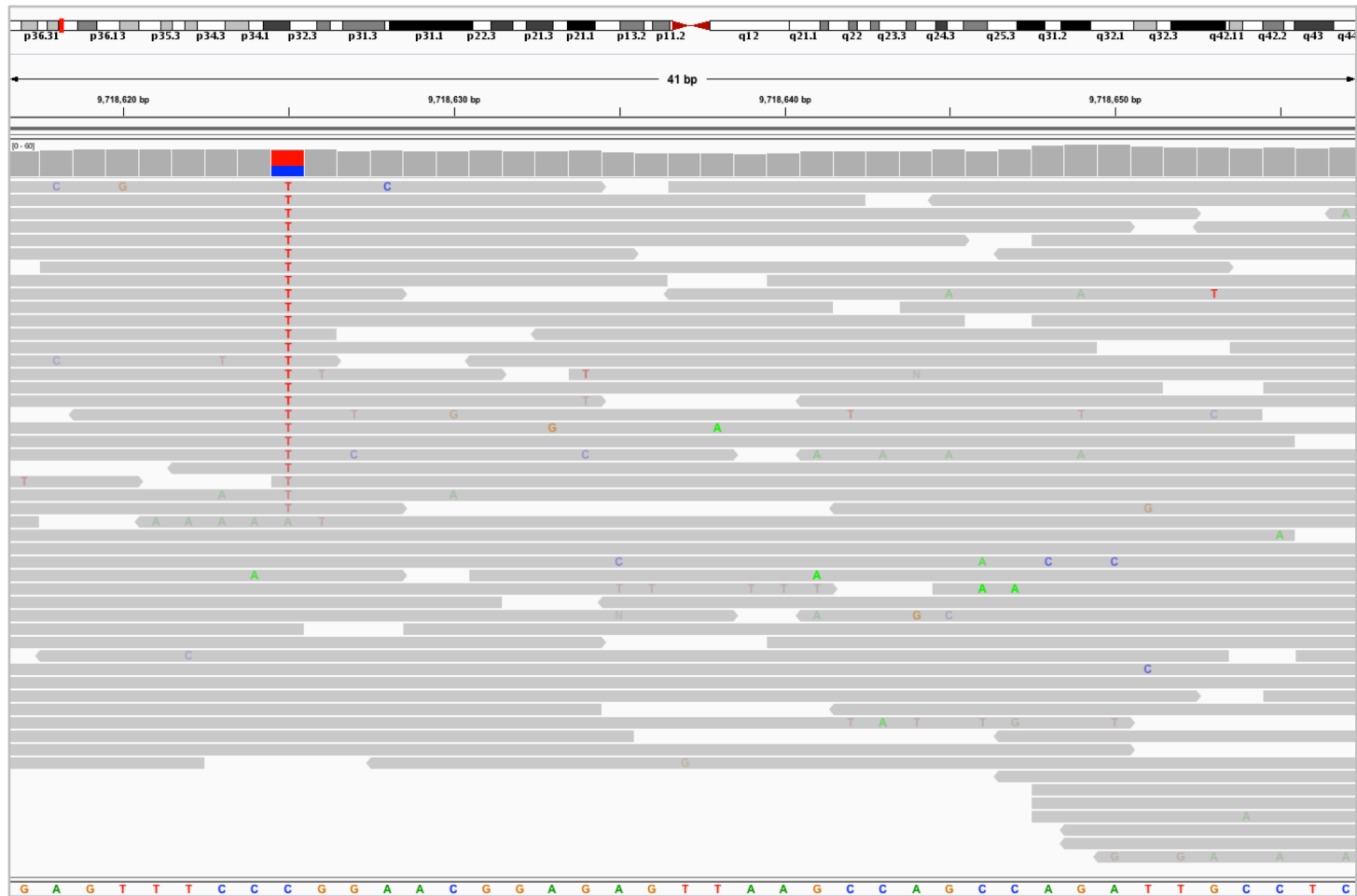


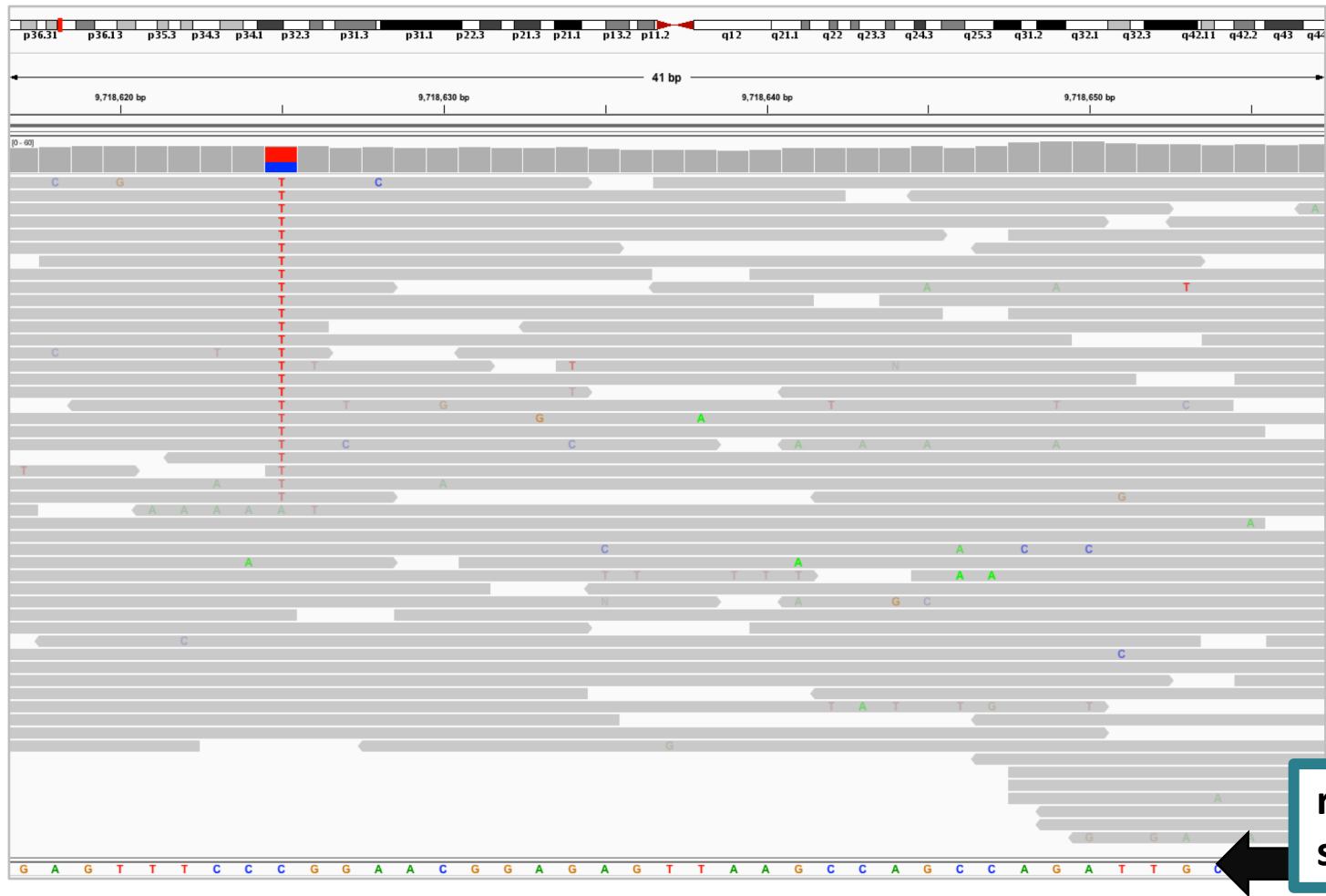




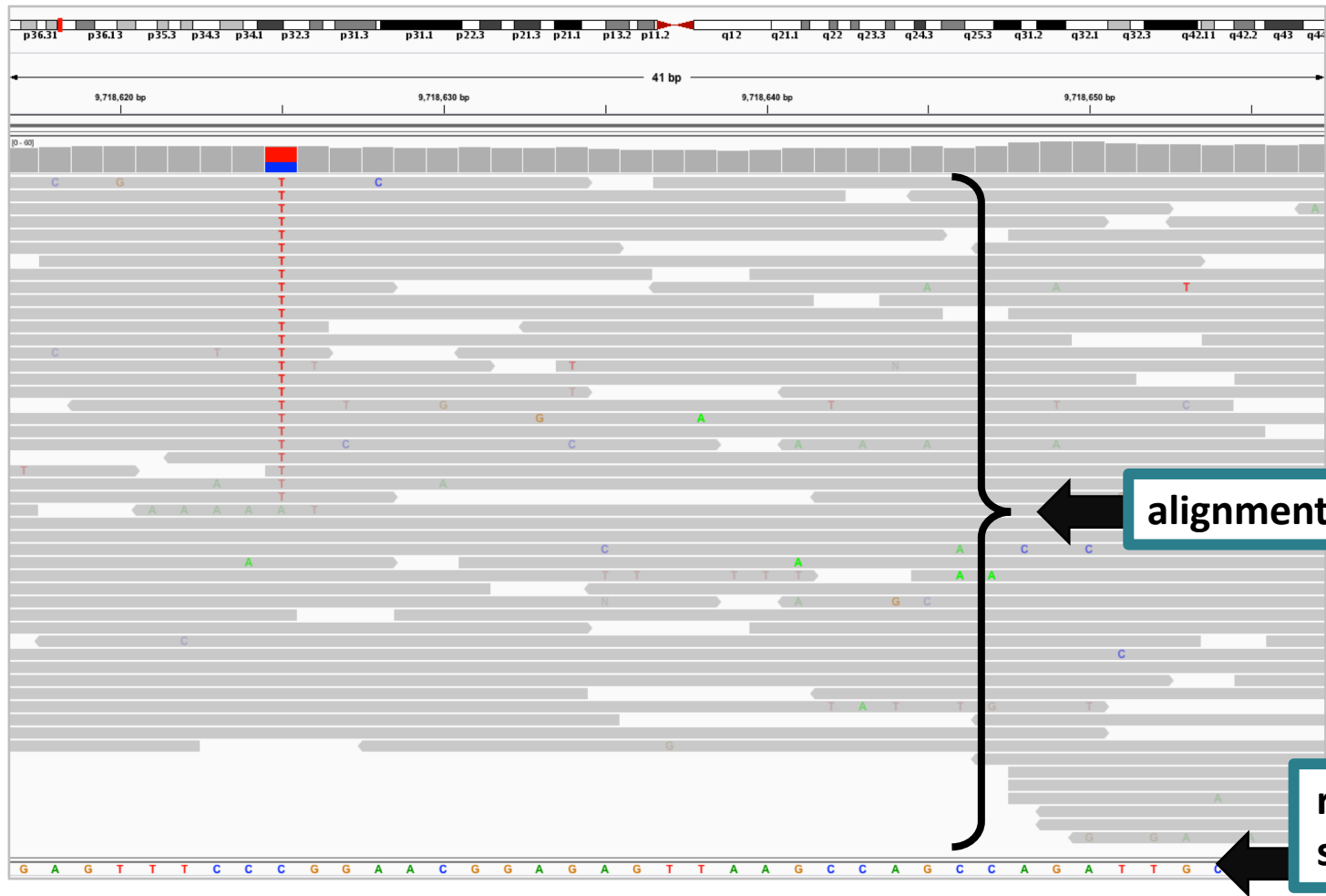






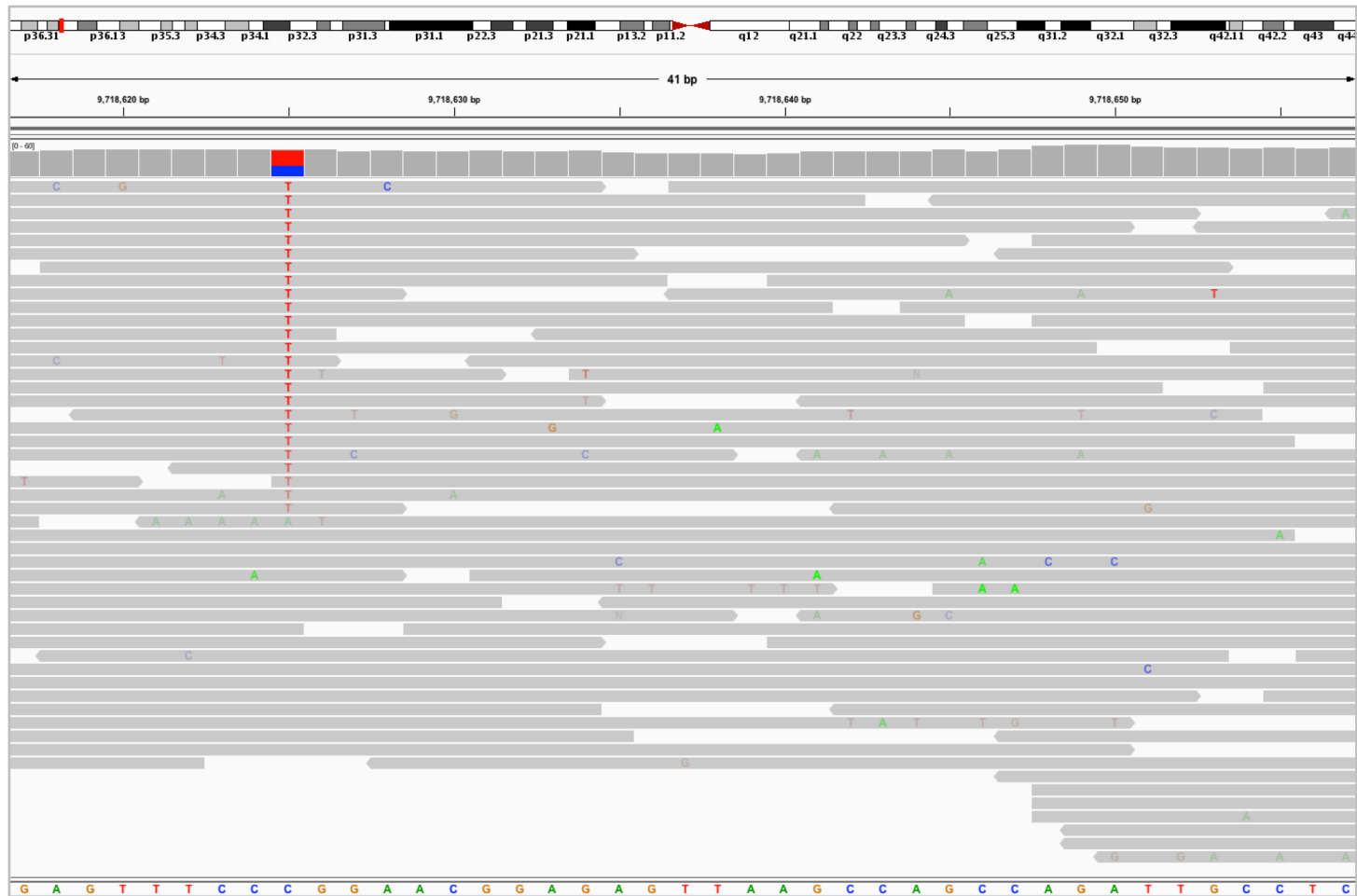


reference
sequence









[illegible]

10	123_33:1:255:636:182	99	1	9718591	20	36M	=	9718757	202	AGACAGGCGATCTCCCTCTATACACGCGTTTCTG	B???:3-2?A@3.4A,5>6-2<<,?%*;	RG:Z:ERR001769	MF:i:130	Aq:i:20	NM:i:5	UQ:i:77	H0:i:0	H1:i:0
----	----------------------	----	---	---------	----	-----	---	---------	-----	------------------------------------	-----------------------------------	----------------	----------	---------	--------	---------	--------	--------

[illegible]

IL23_337:1:255:636:182

99

1

9718591

20

36M

=

9718757

202

AGACAGGCGATCTCCTTCTTATACACGCGTTTTCTG

B??4;;3-2?A@3.4A,5>,6=2<-<,#0'27"%;*

RG:Z:ERR001769

MF:i:130

Aq:i:20

NM:i:5

UQ:i:77

H0:i:0

H1:i:0

```
IL23_337:1:255:636:182
99
1
9718591
20
36M
=
9718757
202
AGACAGGCGATCTCCTTCTTATACACGCGTTTTCTG
B??4;;3-2?A@3.4A,5>,6=2<-<,#0'27"%;*
RG:Z:ERR001769
MF:i:130
Aq:i:20
NM:i:5
UQ:i:77
H0:i:0
H1:i:0
```

NAME

sam – Sequence Alignment/Map file format

DESCRIPTION

Sequence Alignment/Map (SAM) format is TAB-delimited. Apart from the header lines, which are started with t

1	QNAME	Query template/pair NAME
2	FLAG	bitwise FLAG
3	RNAME	Reference sequence NAME
4	POS	1-based leftmost POSition/coordinate of clipped sequence
5	MAPQ	MAPping Quality (Phred-scaled)
6	CIGAR	extended CIGAR string
7	MRNM	Mate Reference sequence NaMe ('=' if same as RNAME)
8	MPOS	1-based Mate POSition
9	TLEN	inferred Template LENgth (insert size)
10	SEQ	query SEquence on the same strand as the reference
11	QUAL	query QUALity (ASCII-33 gives the Phred base quality)
12+	OPT	variable OPTional fields in the format TAG:VTYPE:VALUE

Each bit in the FLAG field is defined as:

0x0001	p	the read is paired in sequencing
0x0002	P	the read is mapped in a proper pair
0x0004	u	the query sequence itself is unmapped
0x0008	U	the mate is unmapped
0x0010	r	strand of the query (1 for reverse)
0x0020	R	strand of the mate
0x0040	1	the read is the first read in a pair
0x0080	2	the read is the second read in a pair
0x0100	s	the alignment is not primary
0x0200	f	the read fails platform/vendor quality checks
0x0400	d	the read is either a PCR or an optical duplicate
0x0800	S	the alignment is supplementary

where the second column gives the string representation of the FLAG field.

IL23_337:1:255:636:182

99

1

9718591

20

36M

=

9718757

202

AGACAGGCGATCTCCTTCTTATACACGCGTTTTCTG

B??4;;3-2?A@3.4A,5>,6=2<-<,#0'27"%;*

RG:Z:ERR001769

MF:i:130

Aq:i:20

NM:i:5

UQ:i:77

H0:i:0

H1:i:0

NAME

sam – Sequence Alignment/Map file format

DESCRIPTION

Sequence Alignment/Map (SAM) format is TAB-delimited. Apart from the header lines, which are started with t

1	QNAME	Query template/pair NAME
2	FLAG	bitwise FLAG
3	RNAME	Reference sequence NAME
4	POS	1-based leftmost POSition/coordinate of clipped sequence
5	MAPQ	MAPping Quality (Phred-scaled)
6	CIGAR	extended CIGAR string
7	MRNM	Mate Reference sequence NaMe ('=' if same as RNAME)
8	MPOS	1-based Mate POSition
9	TLEN	inferred Template LENgth (insert size)
10	SEQ	query SEquence on the same strand as the reference
11	QUAL	query QUALity (ASCII-33 gives the Phred base quality)
12+	OPT	variable OPTional fields in the format TAG:VTYPE:VALUE

Each bit in the FLAG field is defined as:

0x0001	p	the read is paired in sequencing
0x0002	P	the read is mapped in a proper pair
0x0004	u	the query sequence itself is unmapped
0x0008	U	the mate is unmapped
0x0010	r	strand of the query (1 for reverse)
0x0020	R	strand of the mate
0x0040	1	the read is the first read in a pair
0x0080	2	the read is the second read in a pair
0x0100	s	the alignment is not primary
0x0200	f	the read fails platform/vendor quality checks
0x0400	d	the read is either a PCR or an optical duplicate
0x0800	S	the alignment is supplementary

where the second column gives the string representation of the FLAG field.

IL23_337:1:255:636:182

99

1

9718591

20

36M

=

9718757

202

AGACAGGCGATCTCCTTCTTATACACGCGTTTTCTG

B??4;;3-2?A@3.4A,5>,6=2<-<,#0'27"%;*

RG:Z:ERR001769

MF:i:130

Aq:i:20

NM:i:5

UQ:i:77

H0:i:0

H1:i:0

NAME

sam – Sequence Alignment/Map file format

DESCRIPTION

Sequence Alignment/Map (SAM) format is TAB-delimited. Apart from the header lines, which are started with t

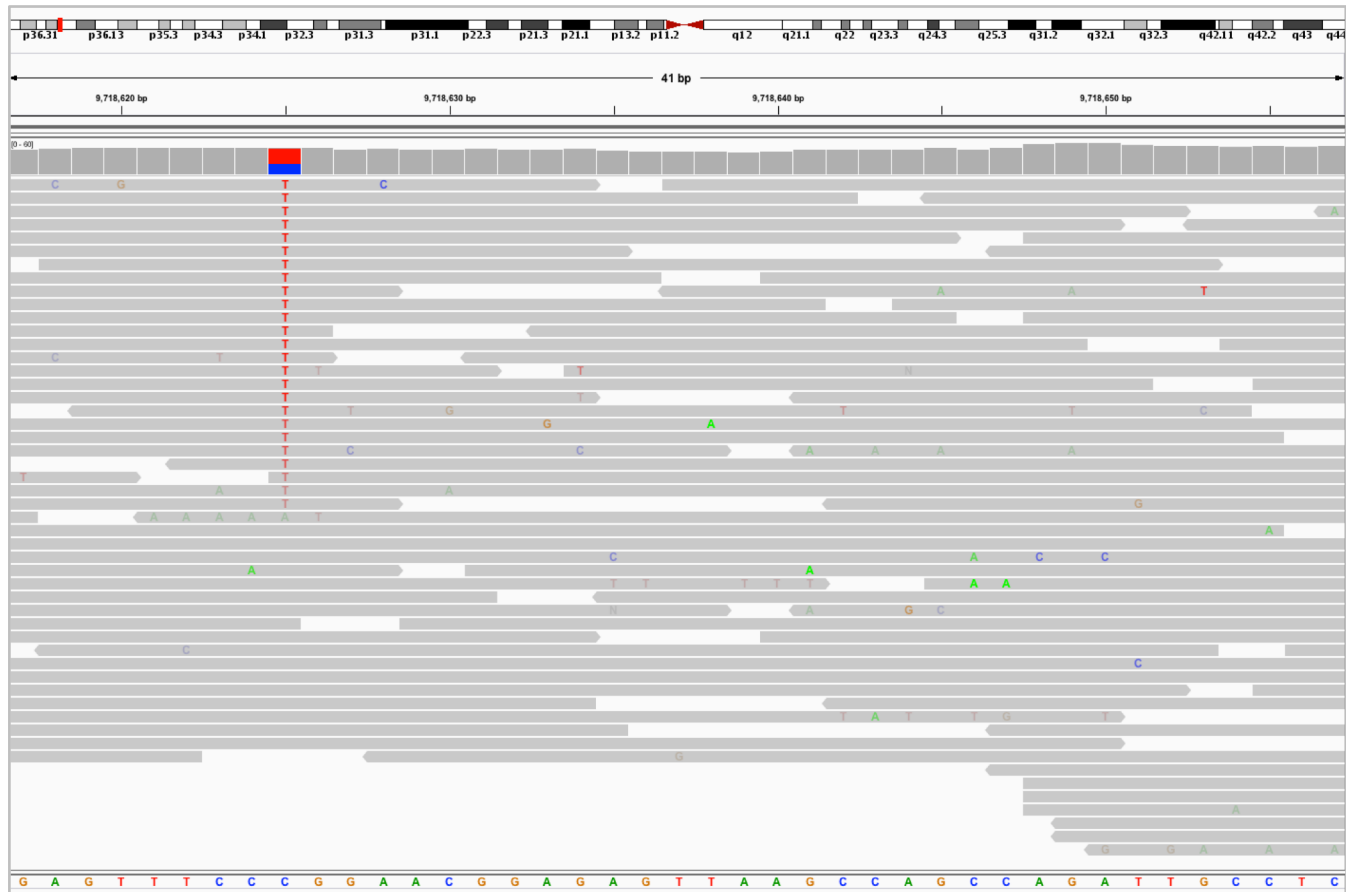
1	QNAME	Query template/pair NAME
2	FLAG	bitwise FLAG
3	RNAME	Reference sequence NAME
4	POS	1-based leftmost POSition/coordinate of clipped sequence
5	MAPQ	MAPping Quality (Phred-scaled)
6	CIGAR	extended CIGAR string
7	MRNM	Mate Reference sequence NaMe ('=' if same as RNAME)
8	MPOS	1-based Mate POSition
9	TLEN	inferred Template LENgth (insert size)
10	SEQ	query SEquence on the same strand as the reference
11	QUAL	query QUALity (ASCII-33 gives the Phred base quality)
12+	OPT	variable OPTional fields in the format TAG:VTYPE:VALUE

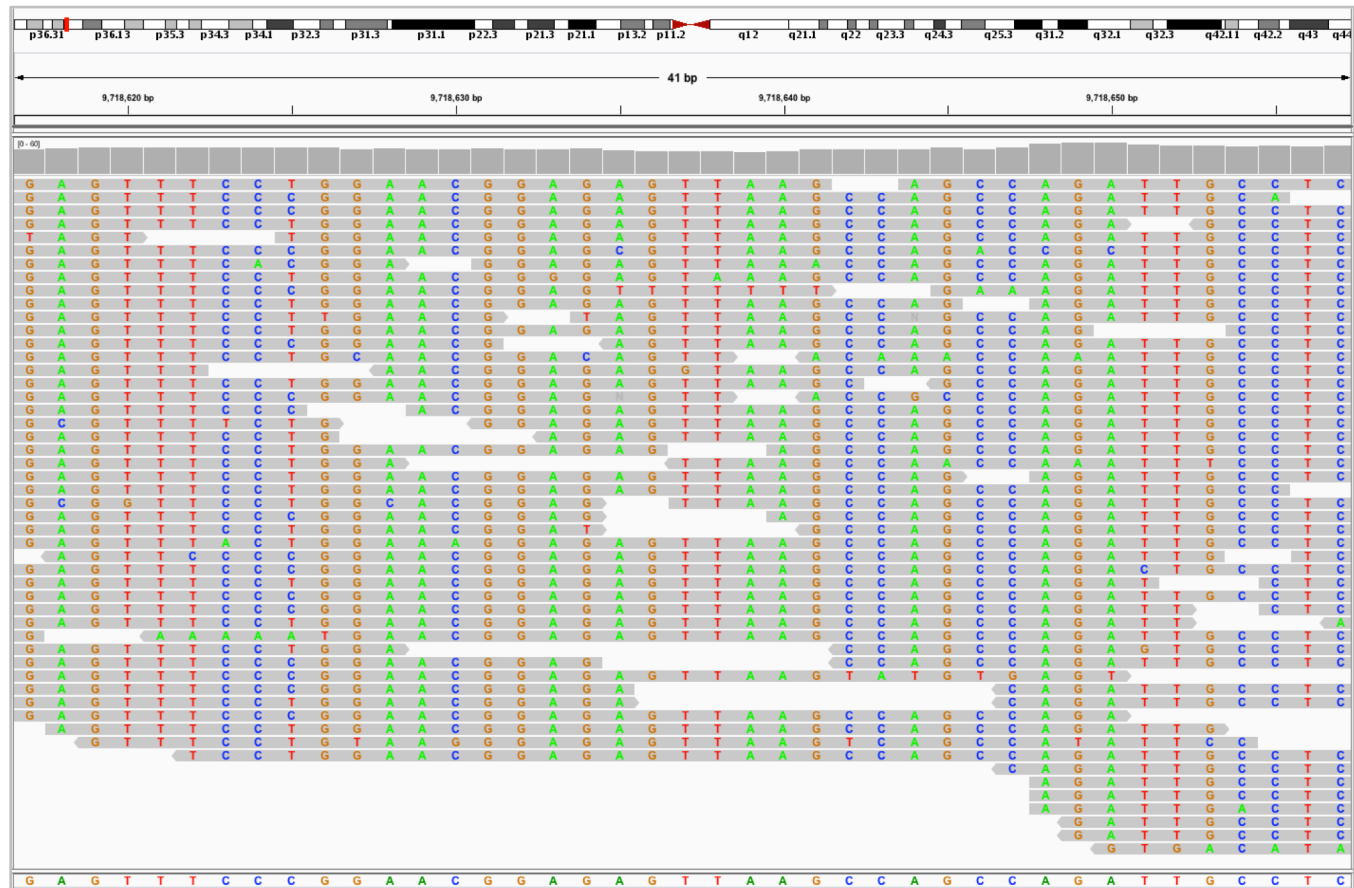
Each bit in the FLAG field is defined as:

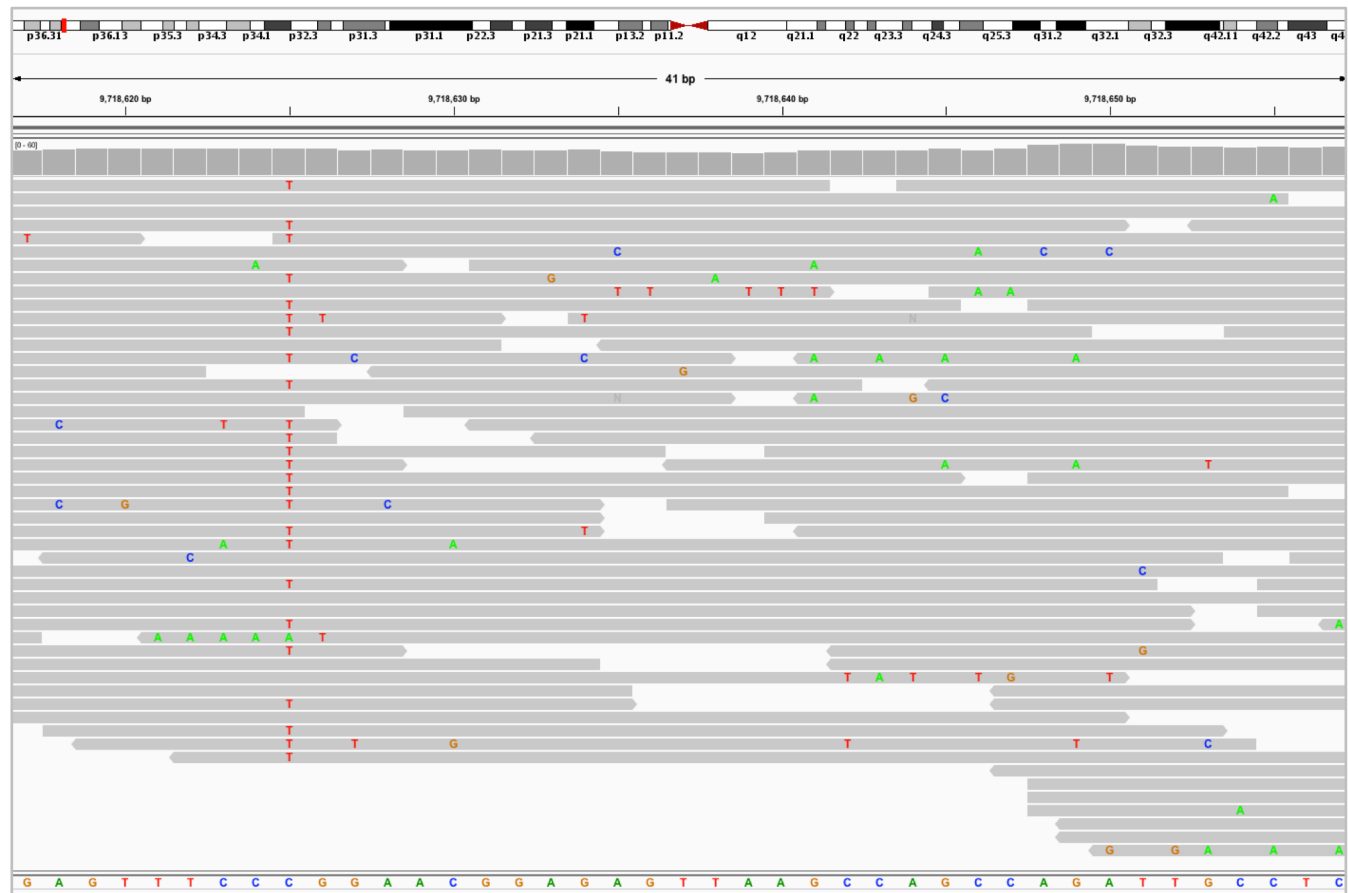
0x0001	p	the read is paired in sequencing
0x0002	P	the read is mapped in a proper pair
0x0004	u	the query sequence itself is unmapped
0x0008	U	the mate is unmapped
0x0010	r	strand of the query (1 for reverse)
0x0020	R	strand of the mate
0x0040	1	the read is the first read in a pair
0x0080	2	the read is the second read in a pair
0x0100	s	the alignment is not primary
0x0200	f	the read fails platform/vendor quality checks
0x0400	d	the read is either a PCR or an optical duplicate
0x0800	S	the alignment is supplementary

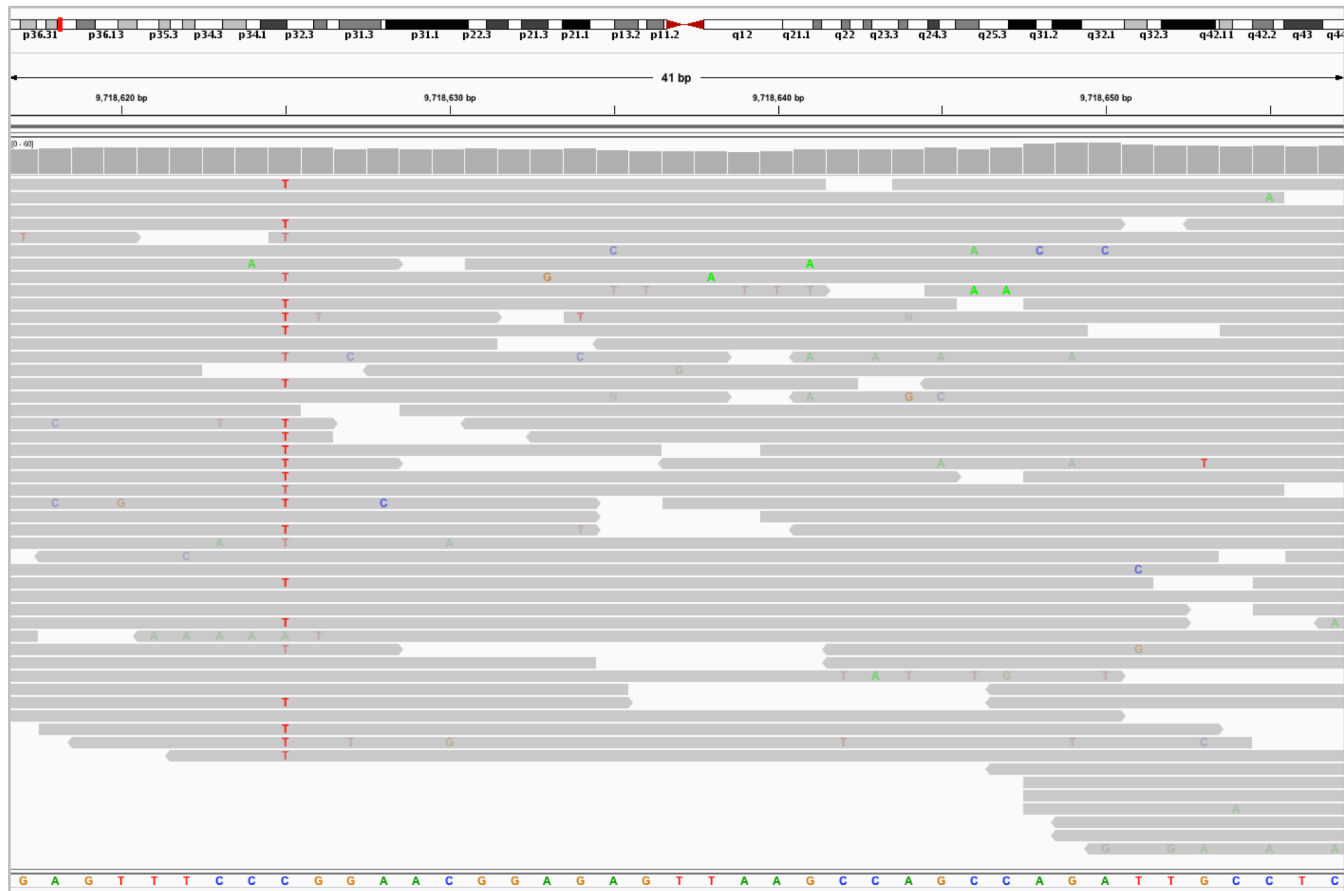
where the second column gives the string representation of the FLAG field.

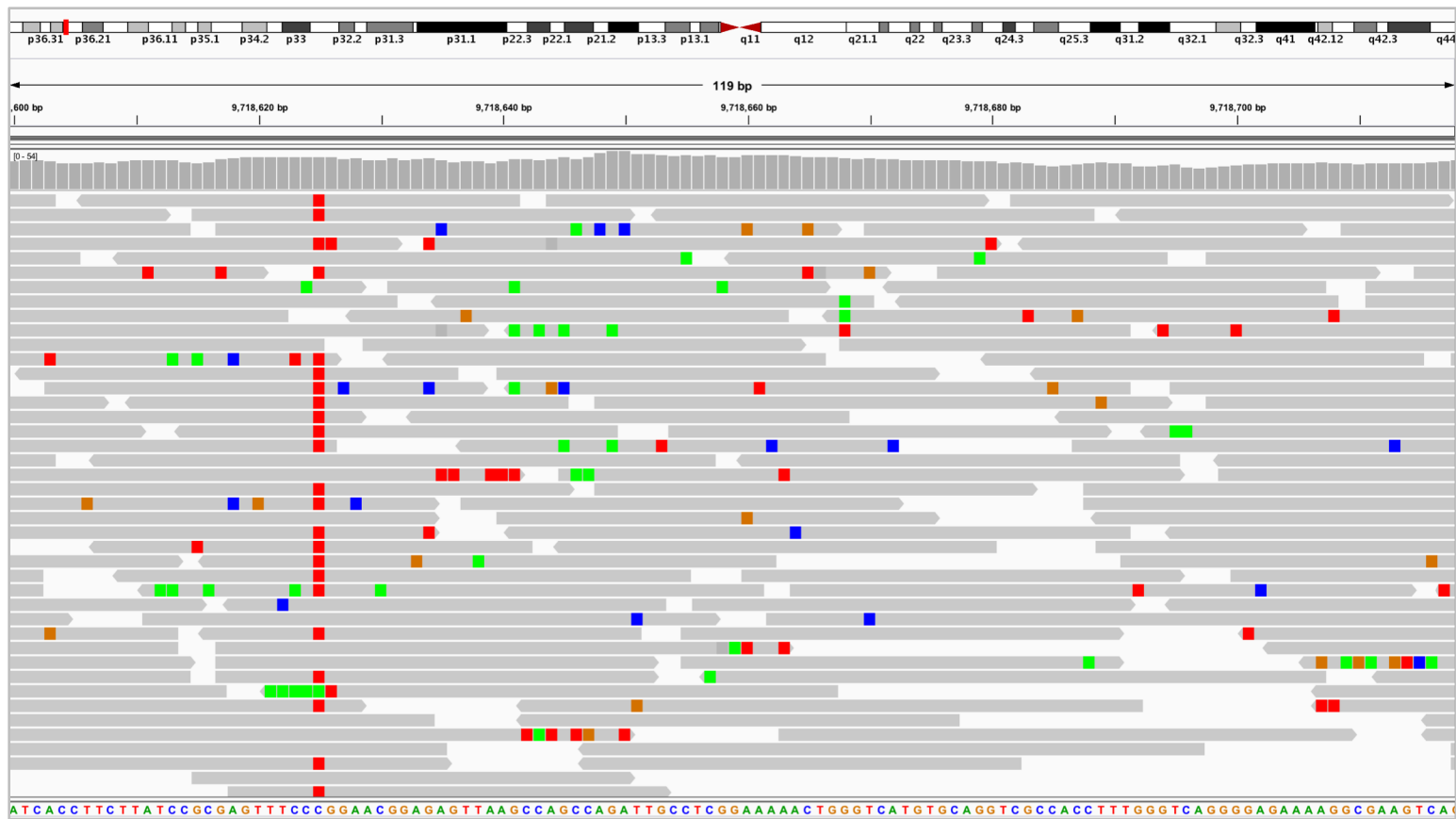
[illegible]

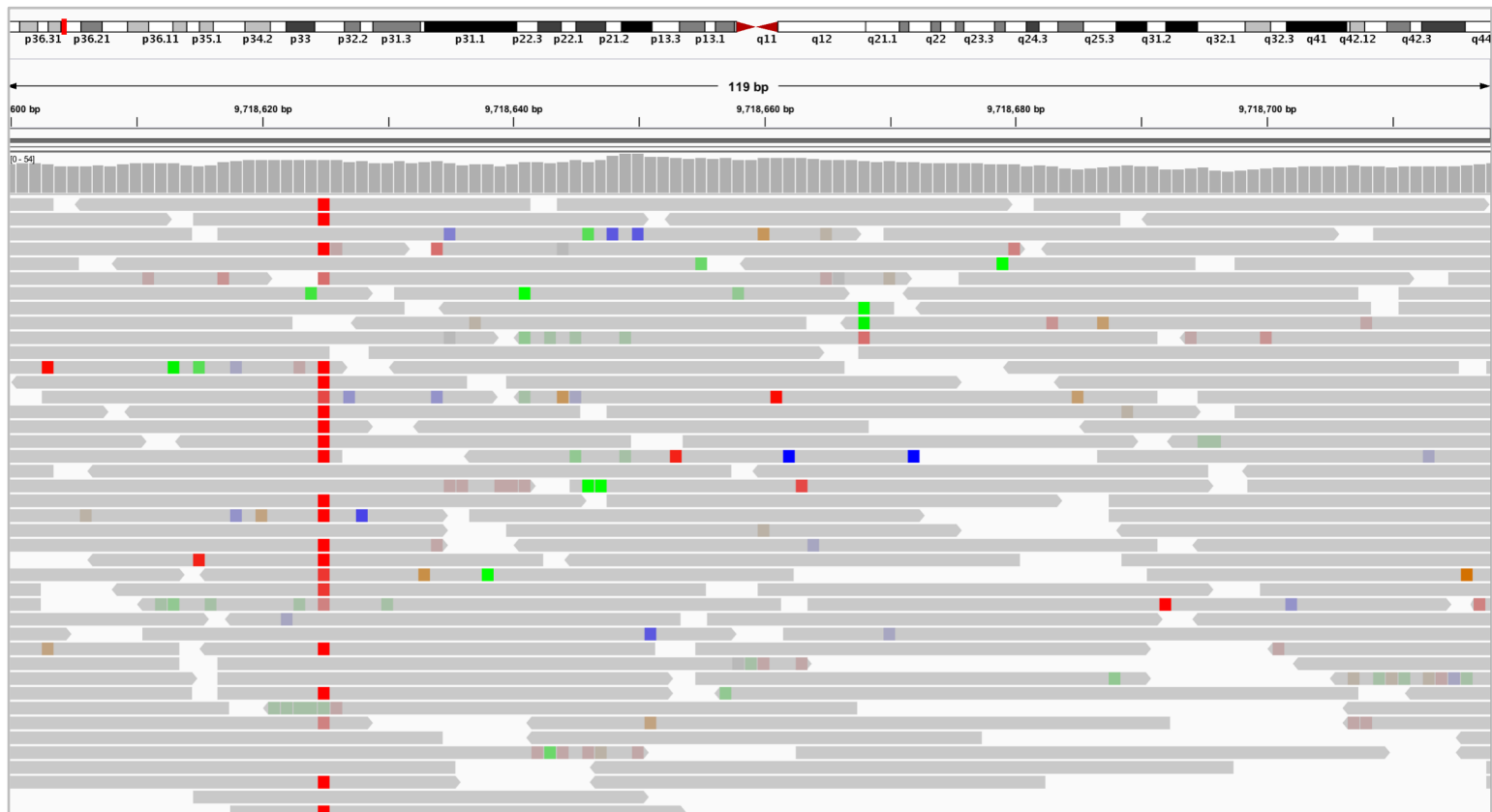


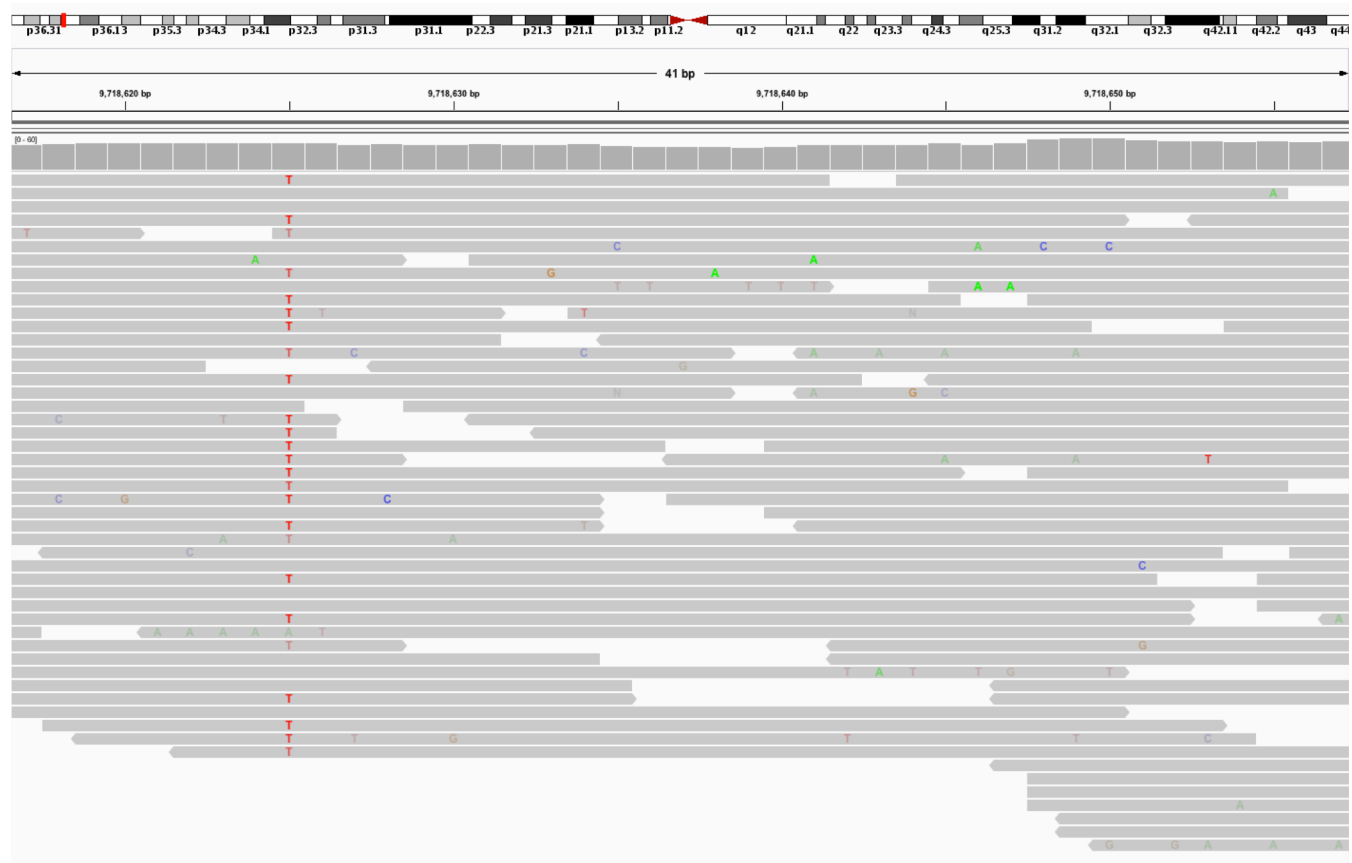


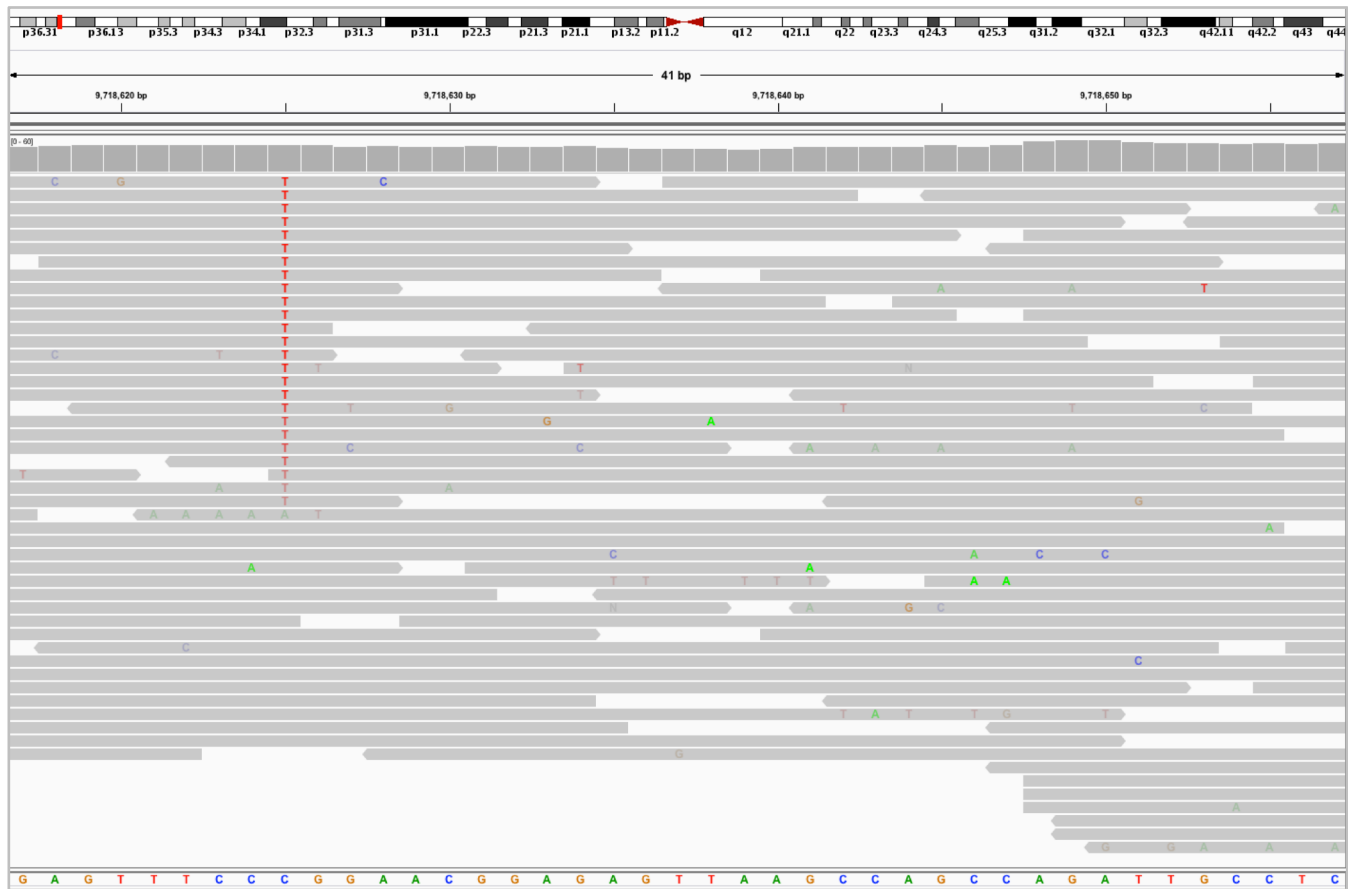


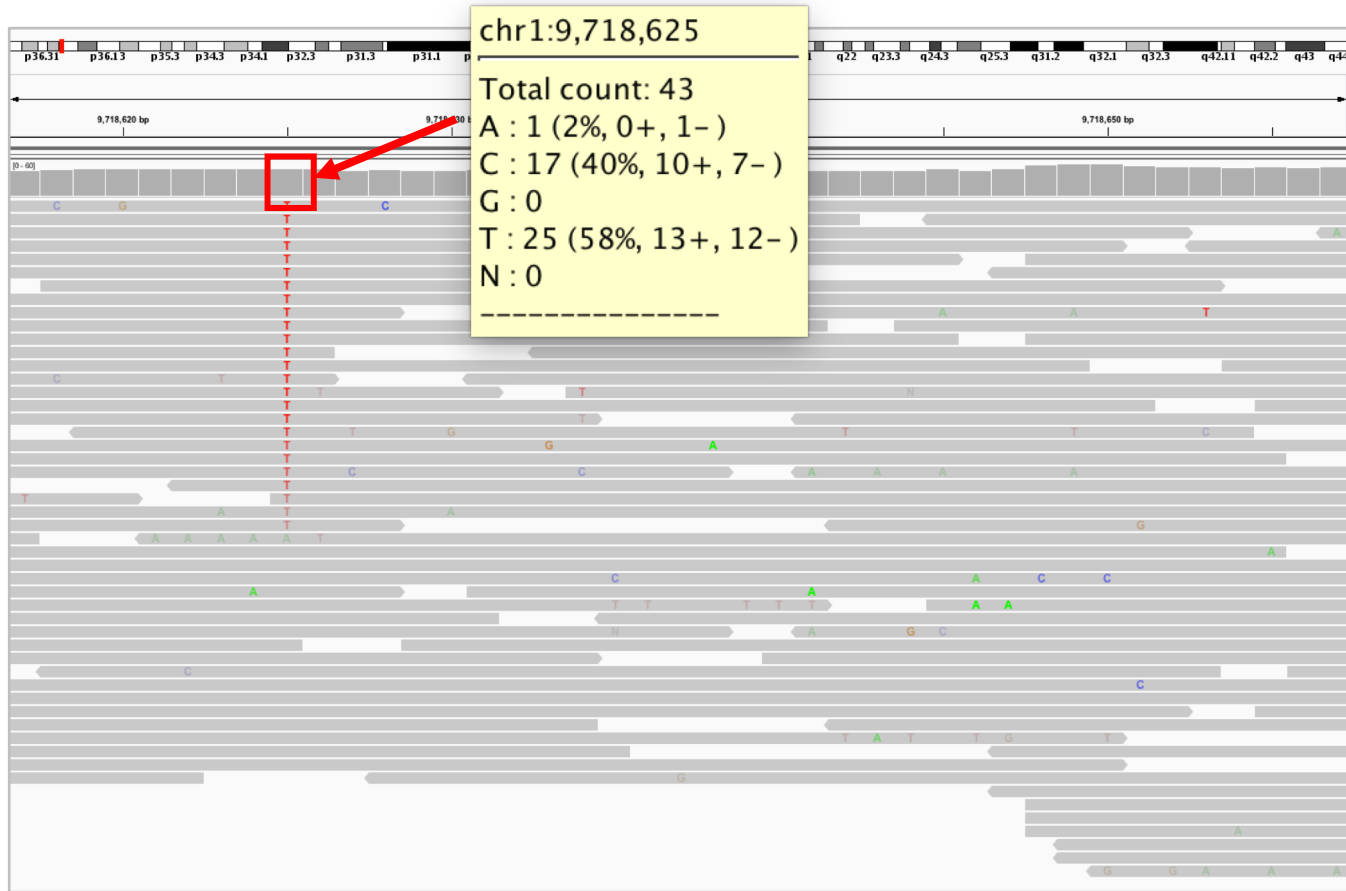


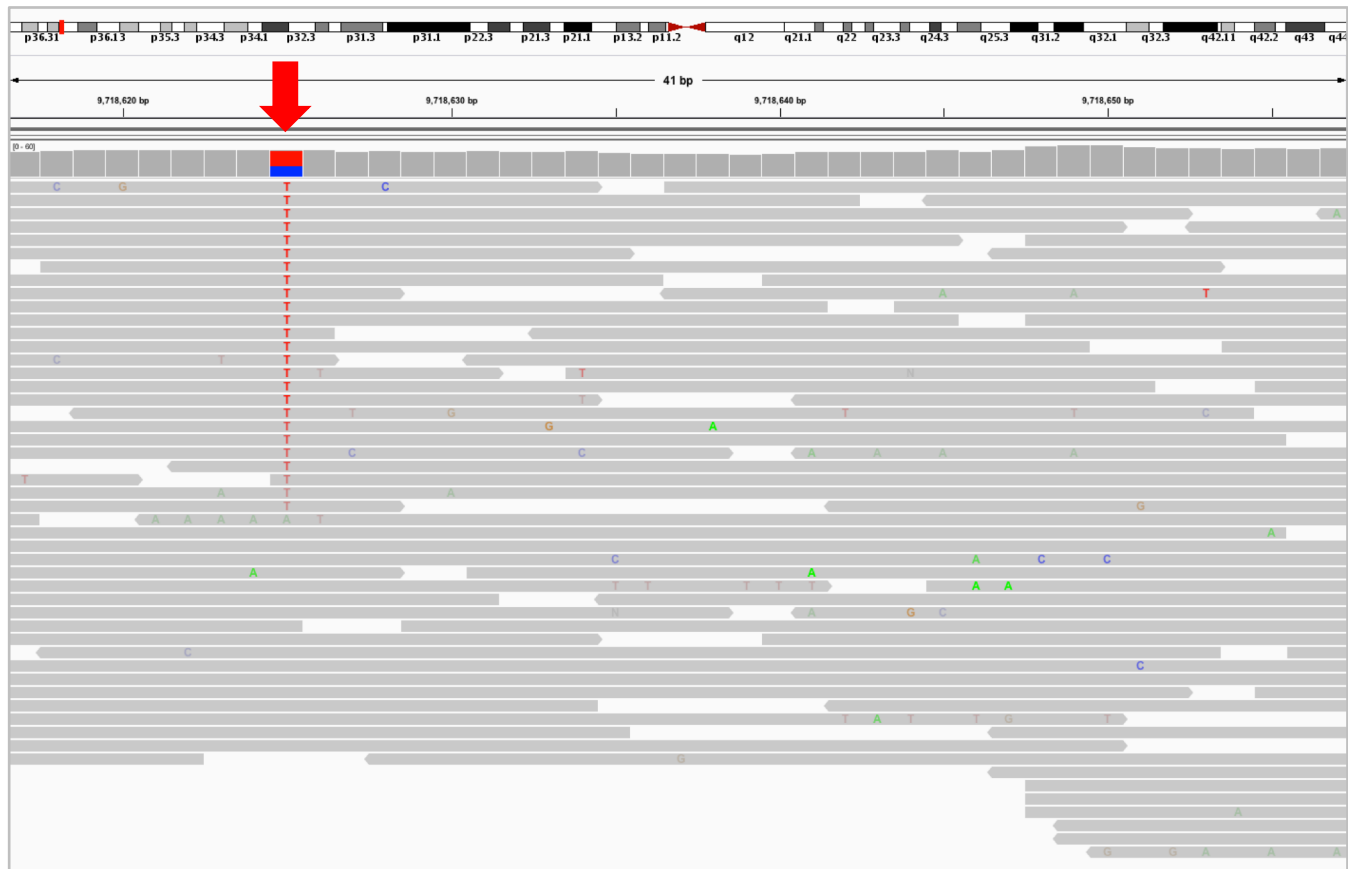




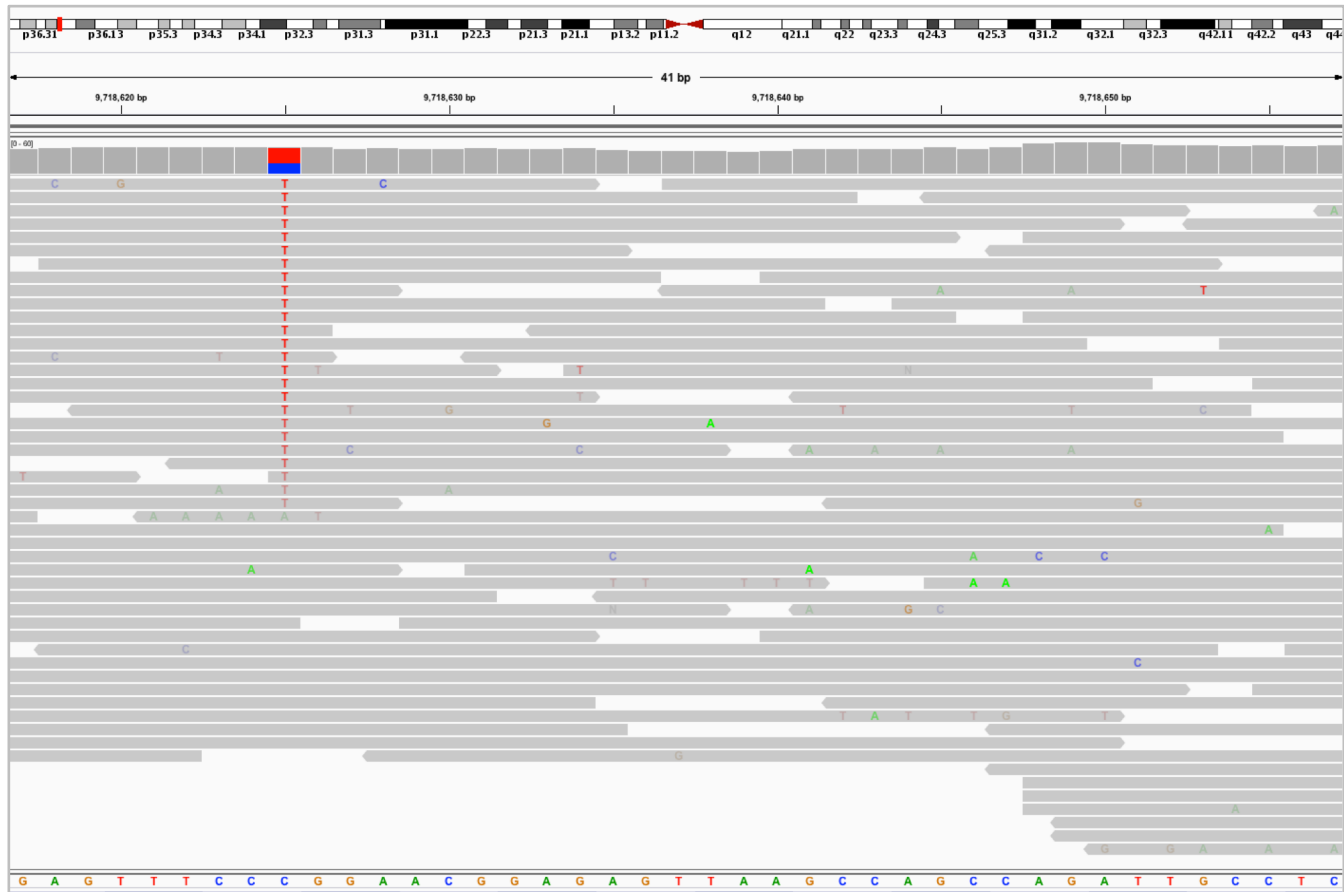


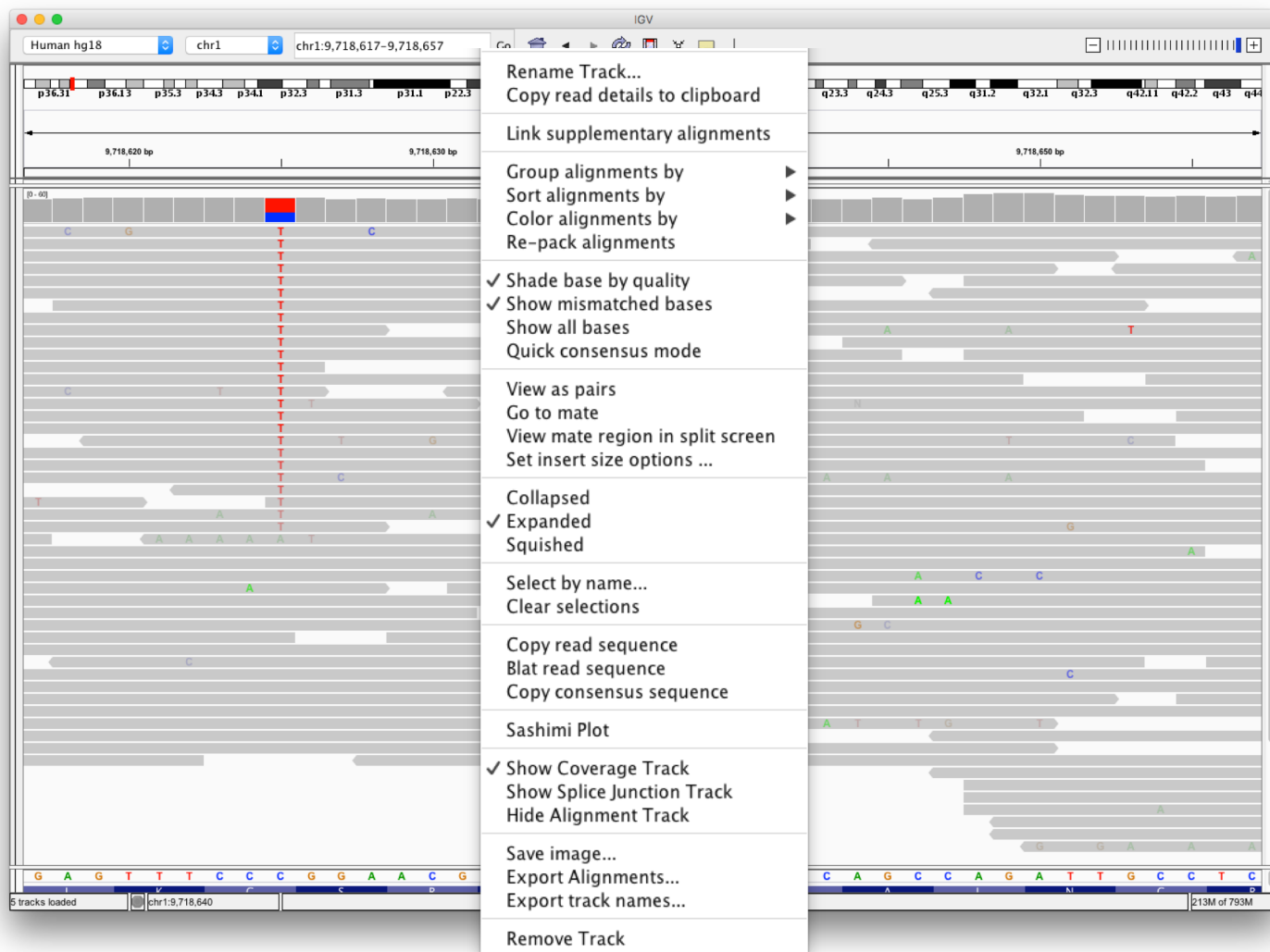






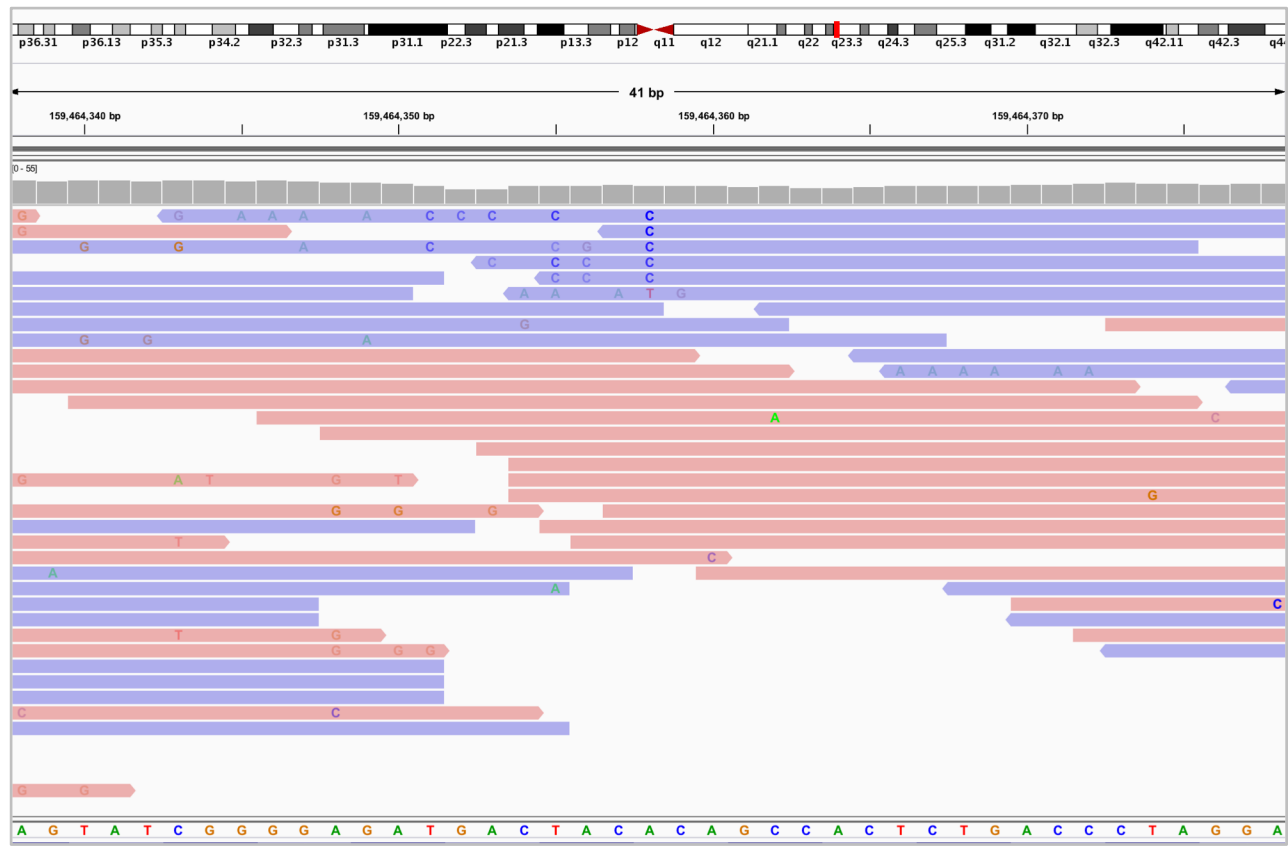


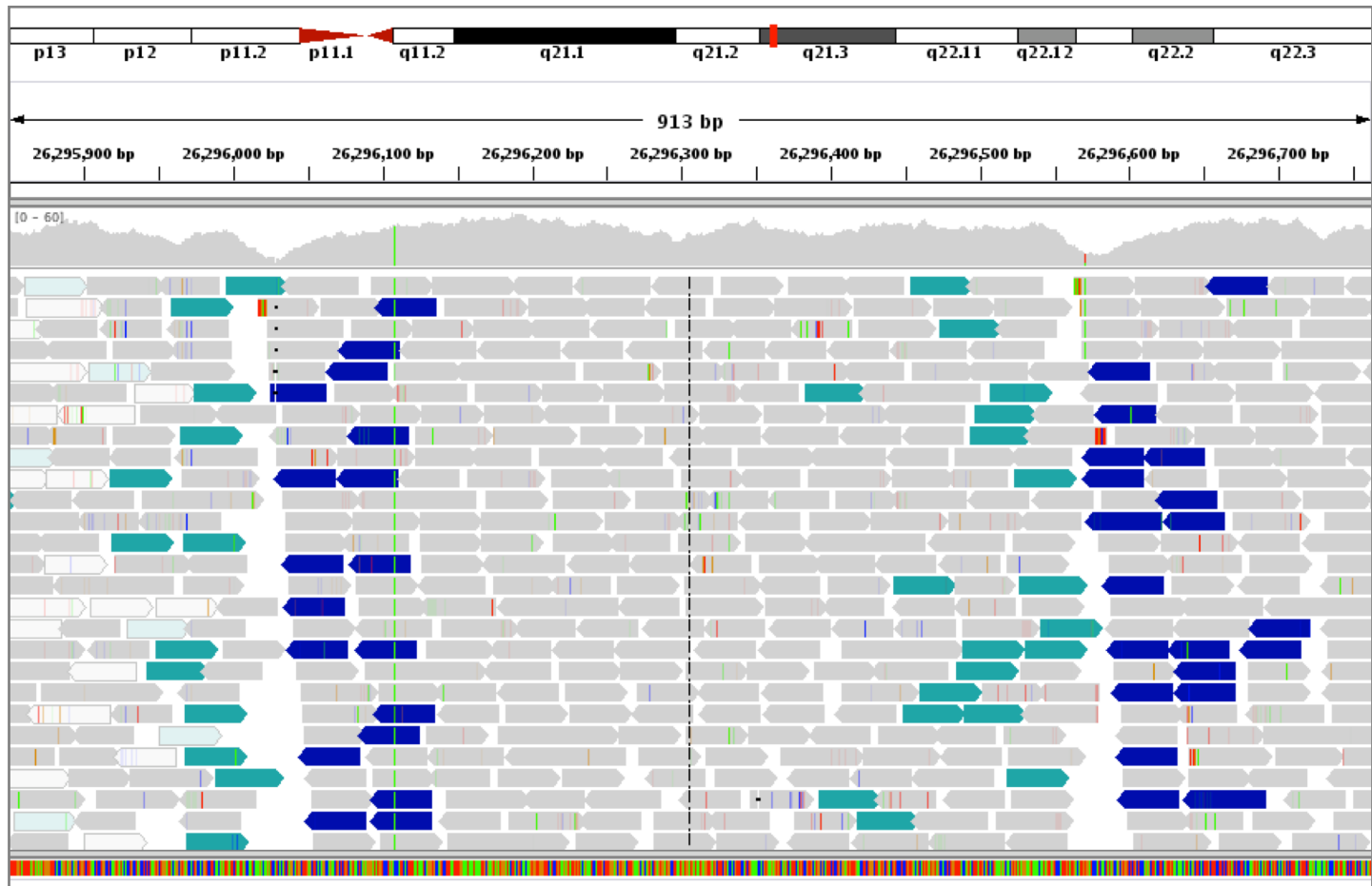












Paired-end sequencing

DNA or
cDNA



fragment



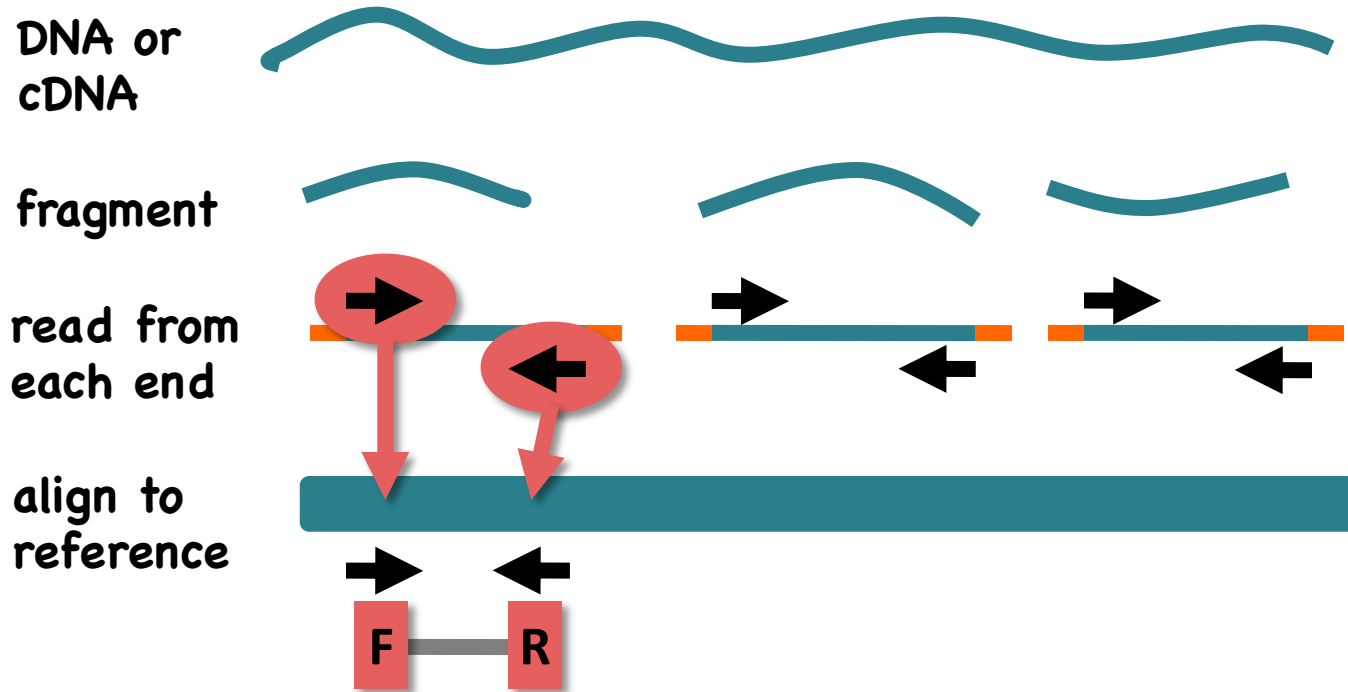
read from
each end



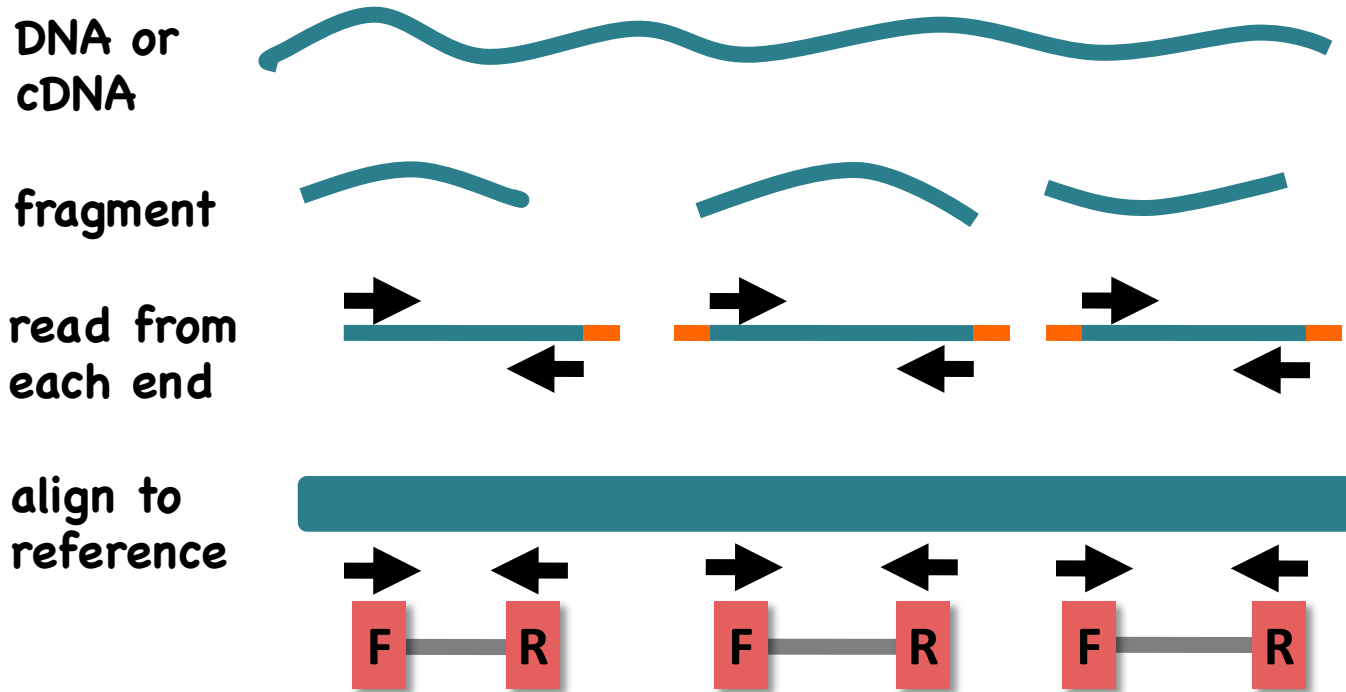
align to
reference



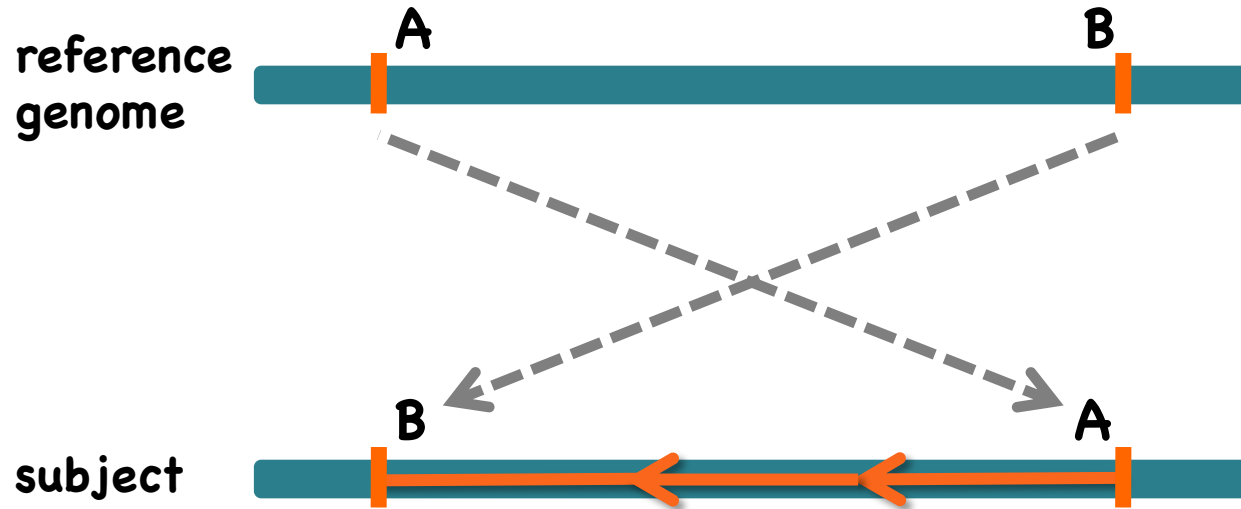
Paired-end sequencing



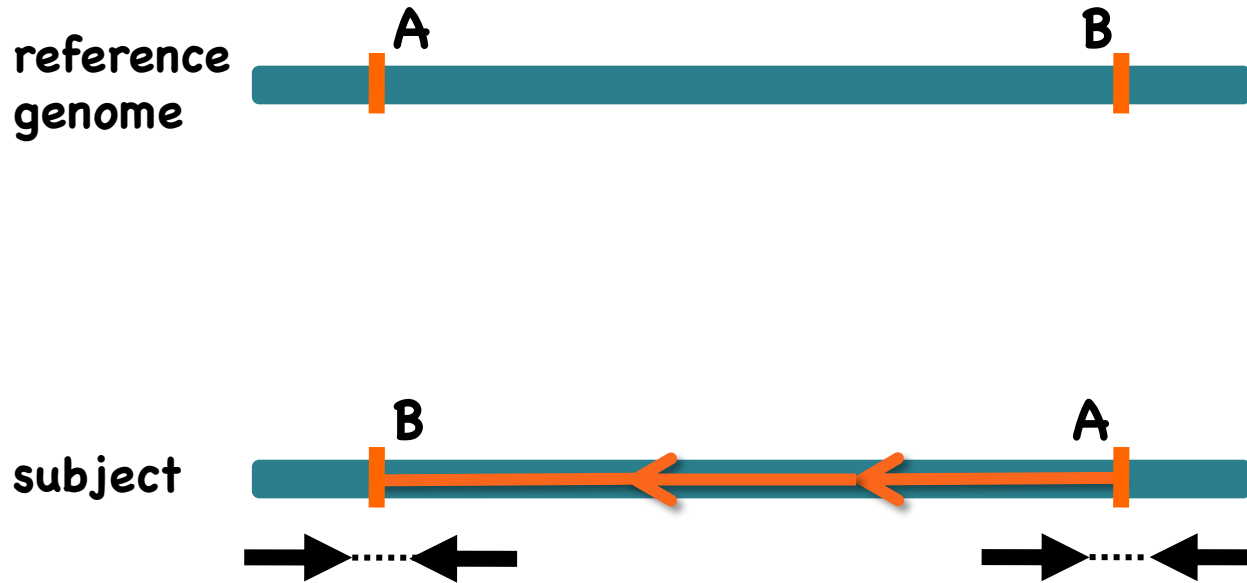
Paired-end sequencing



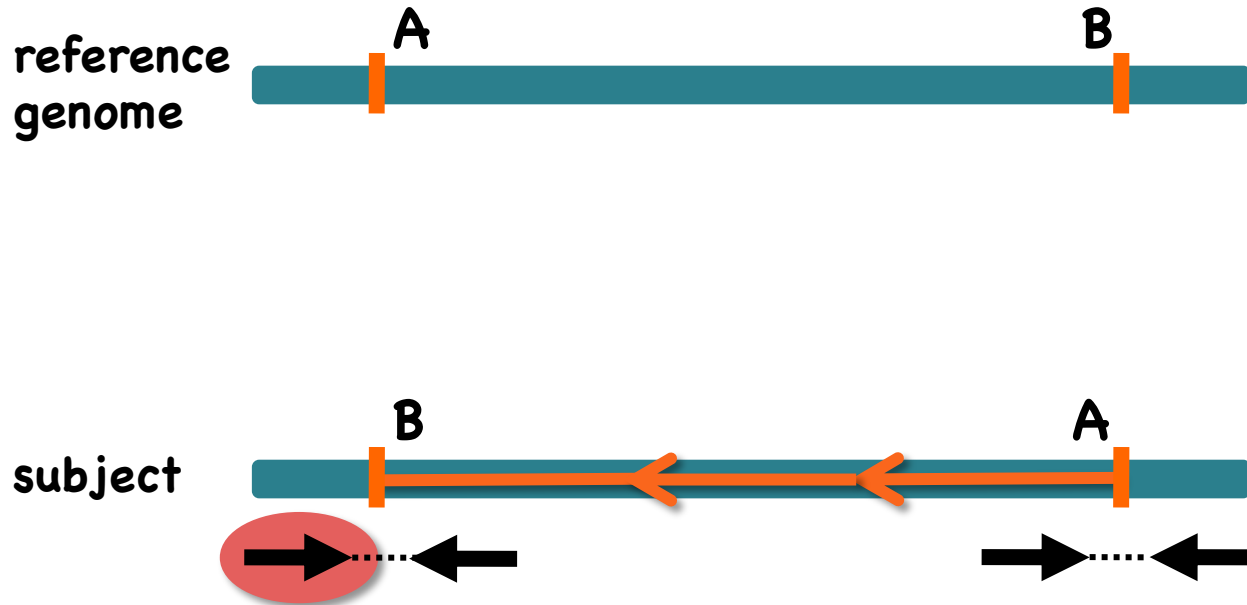
Inversion



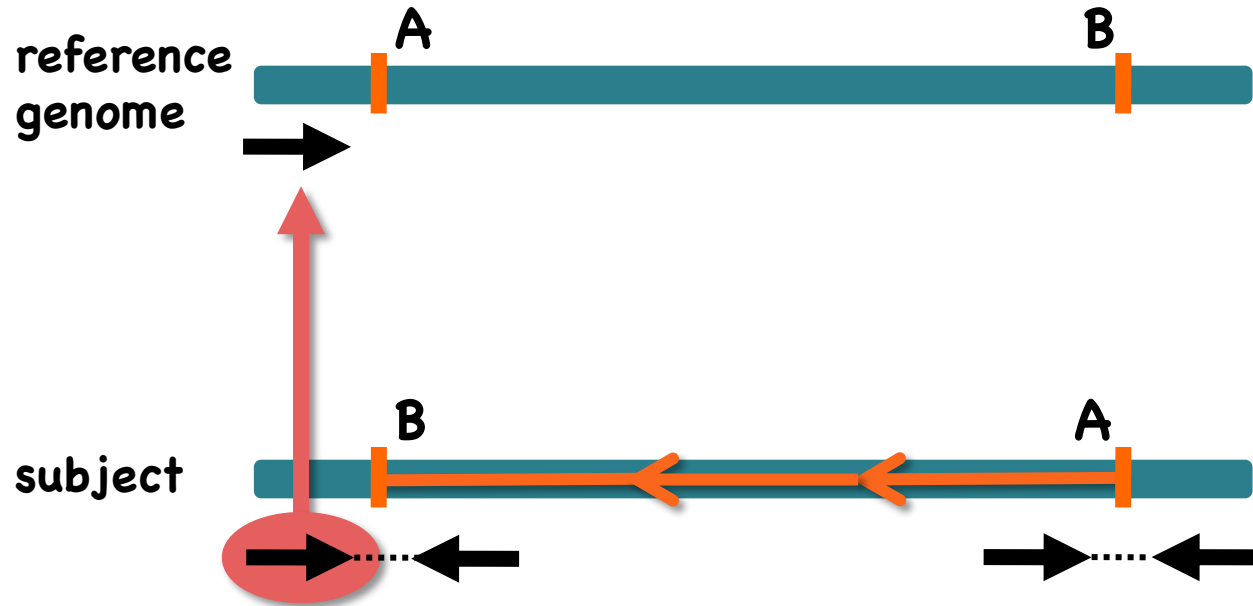
Inversion



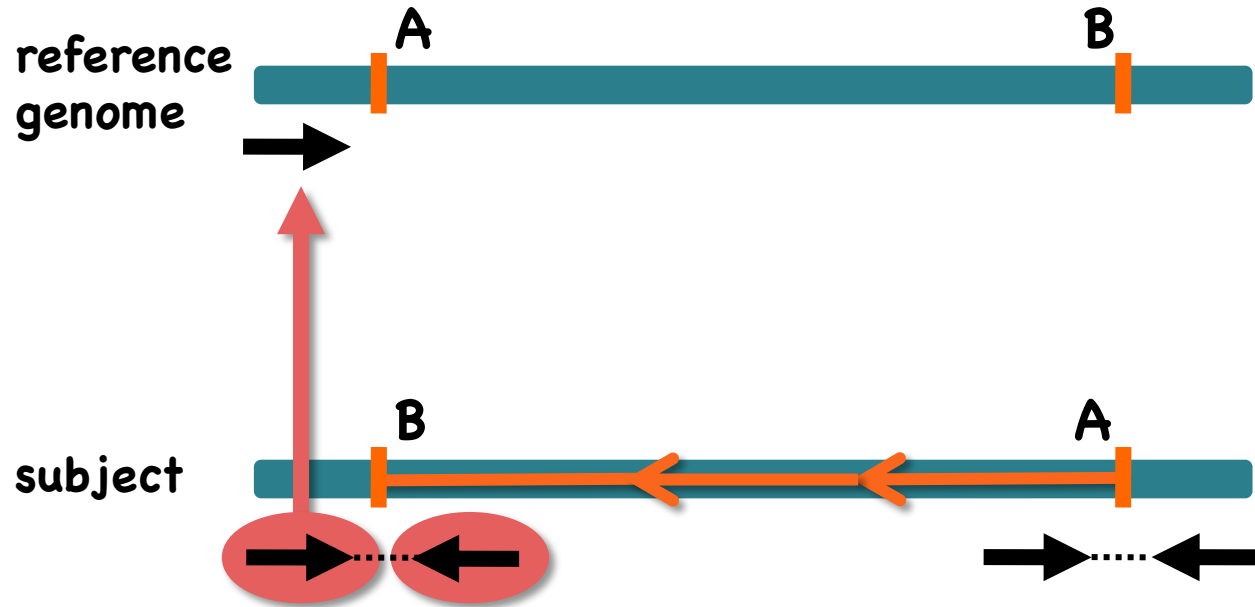
Inversion



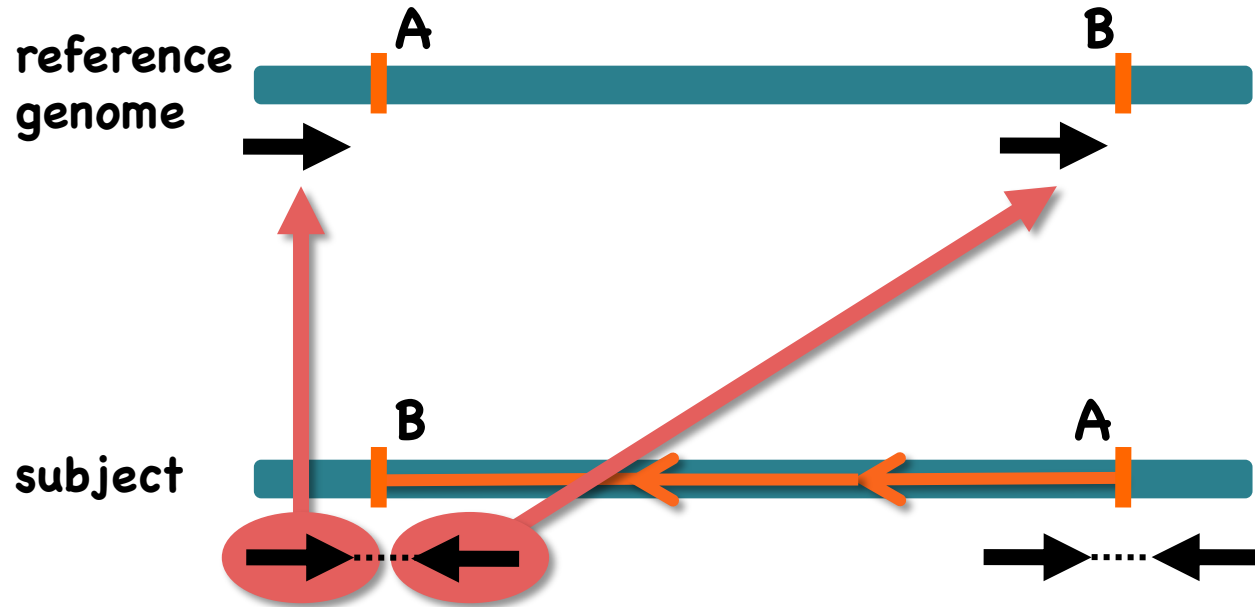
Inversion



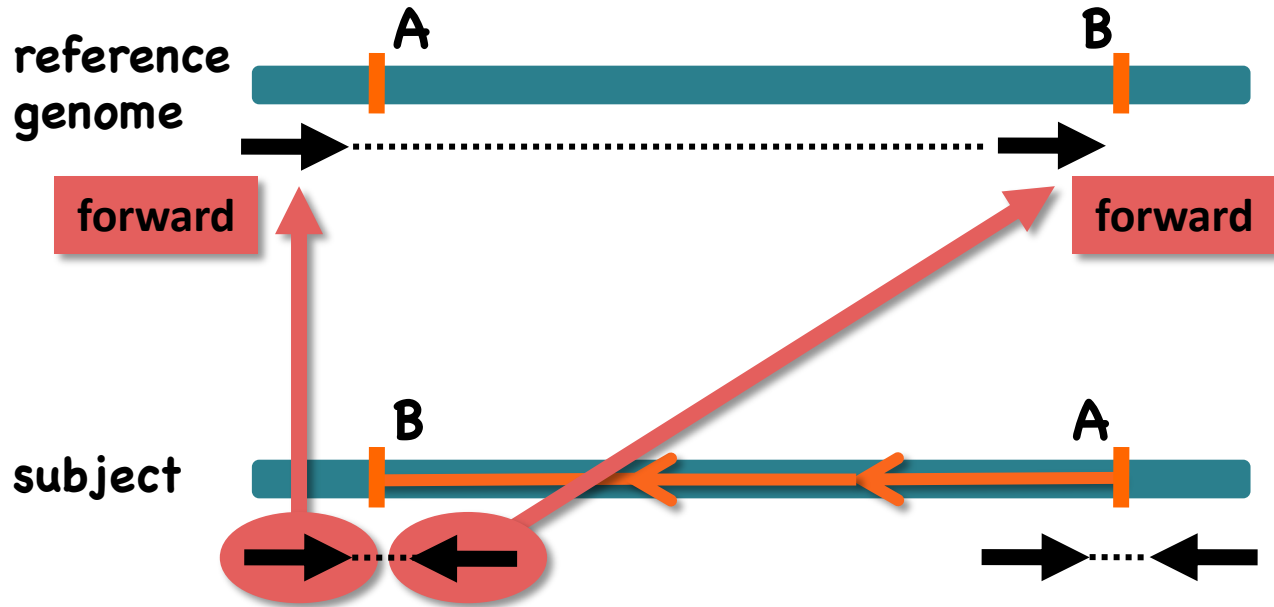
Inversion



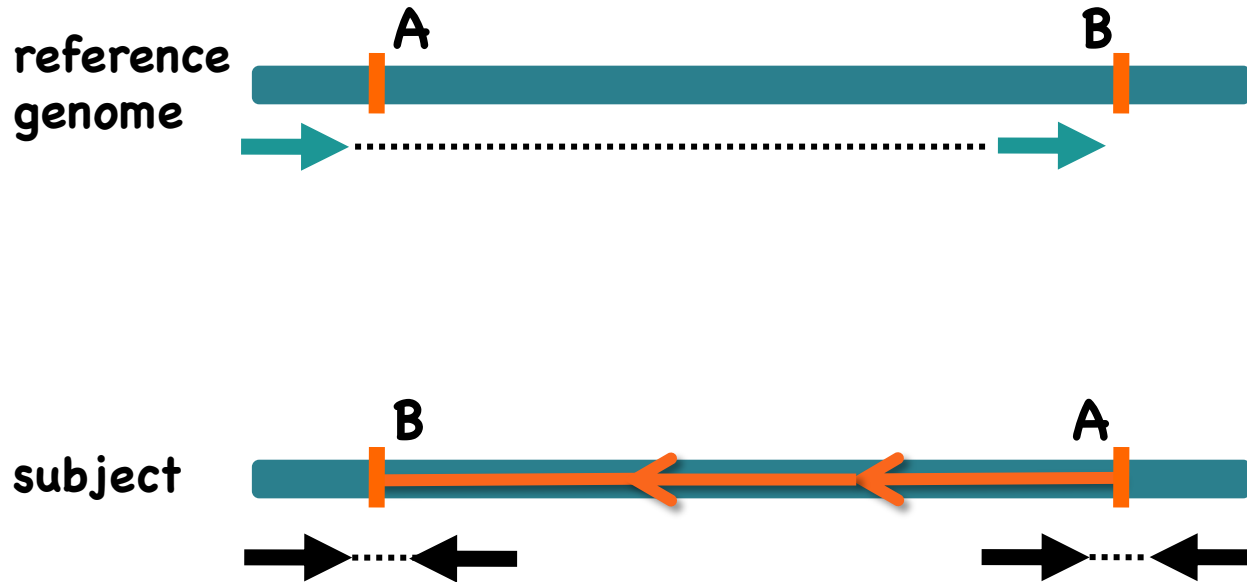
Inversion



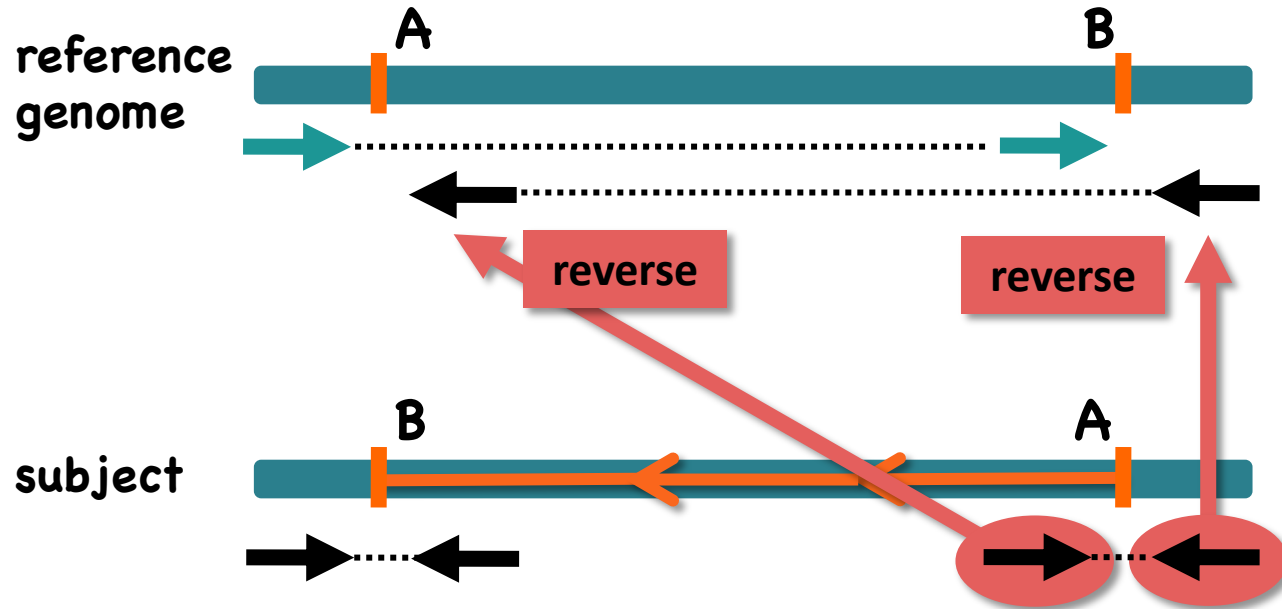
Inversion



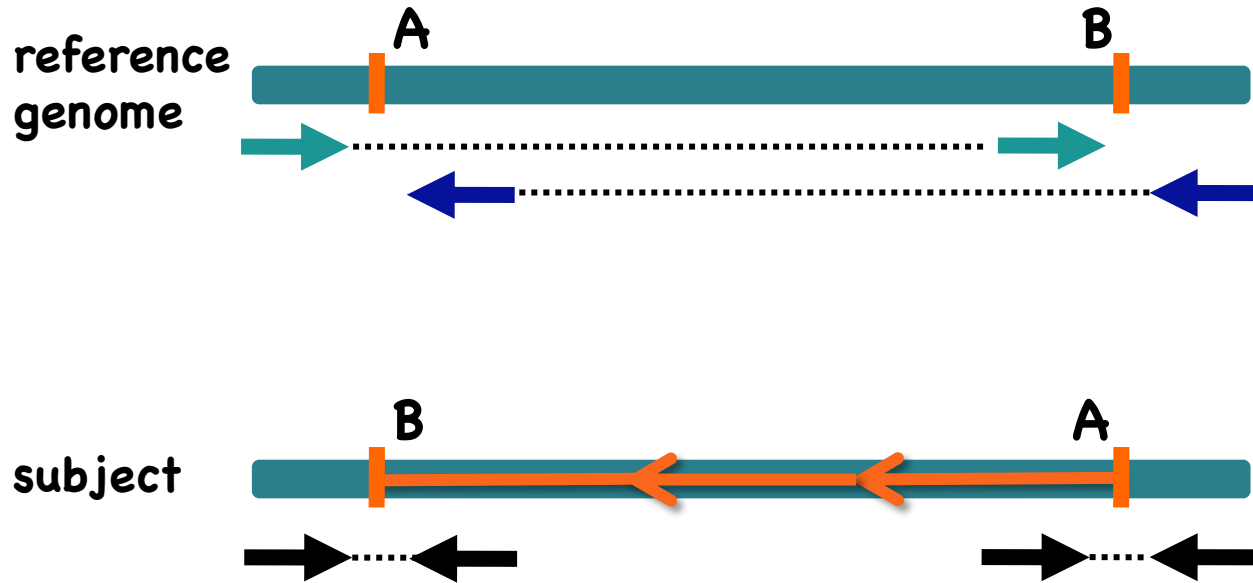
Inversion

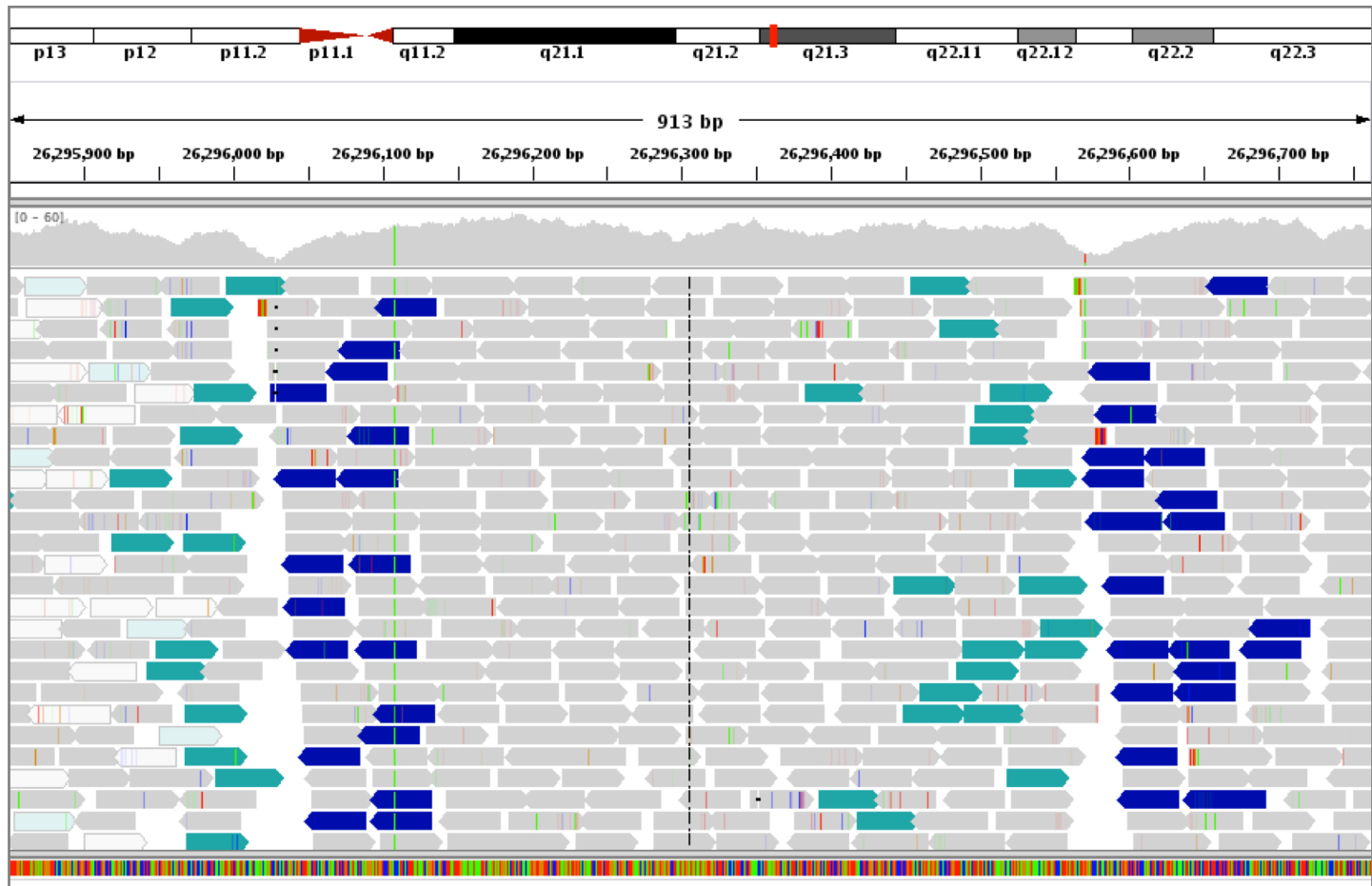


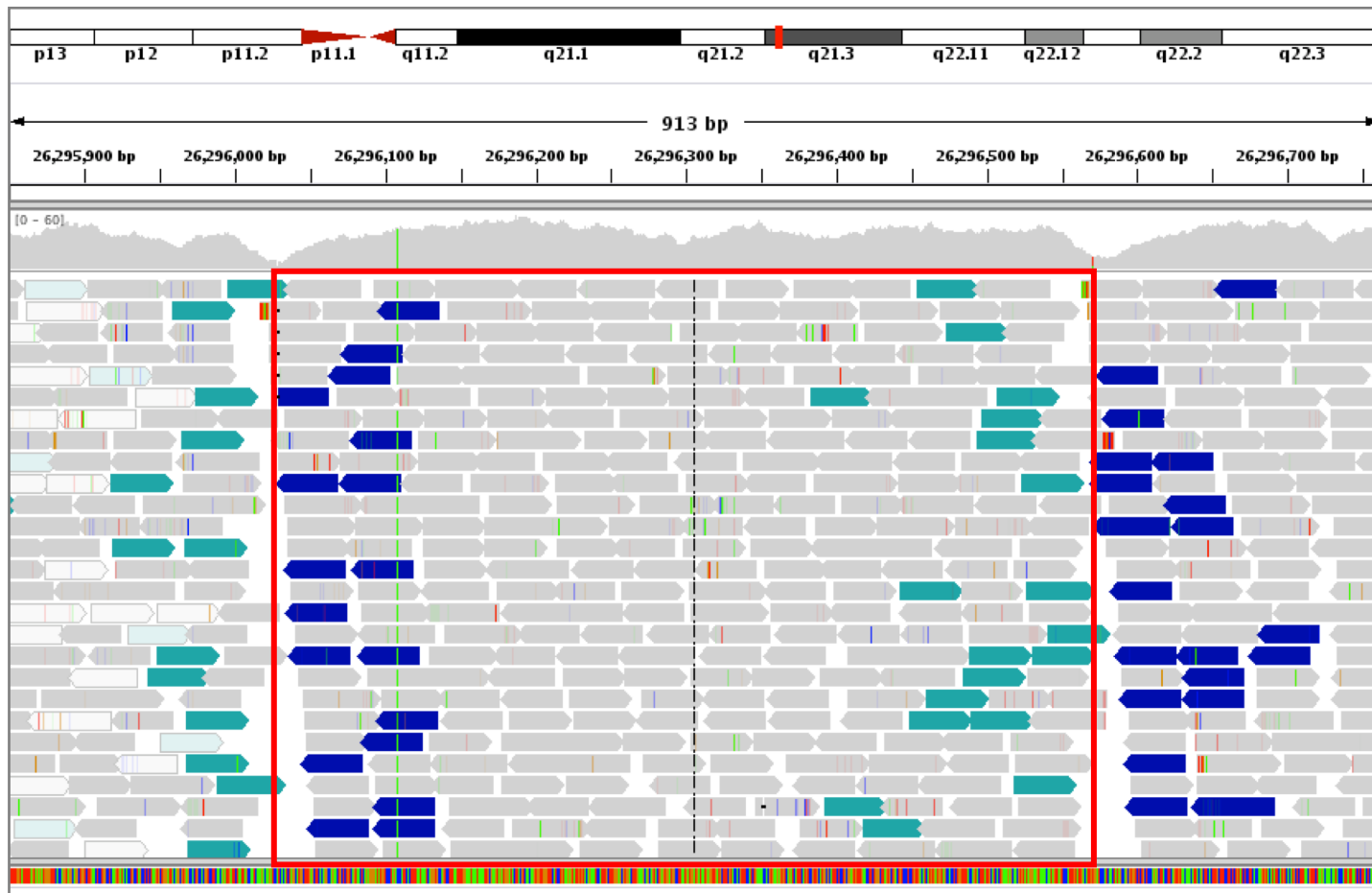
Inversion

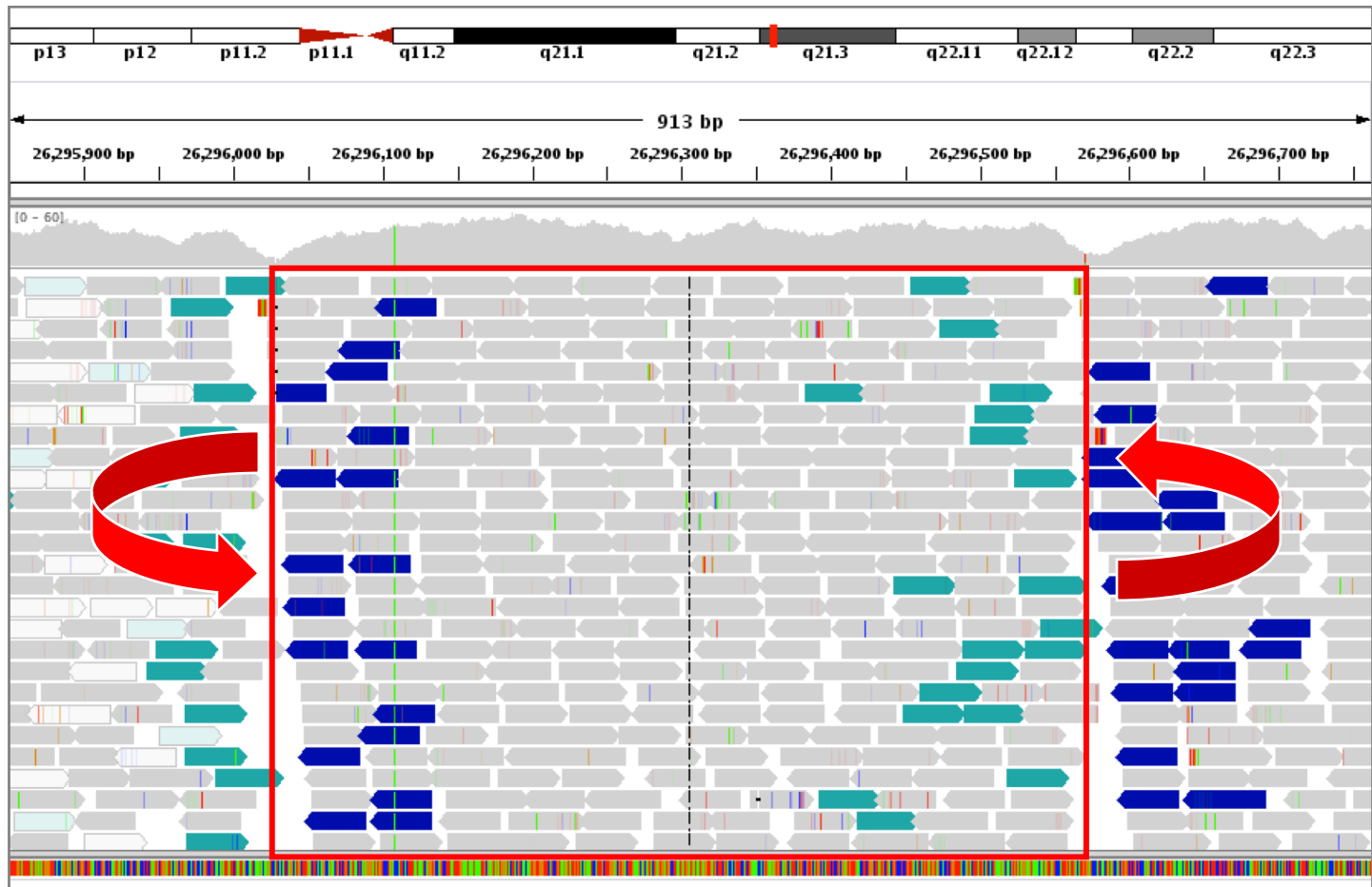


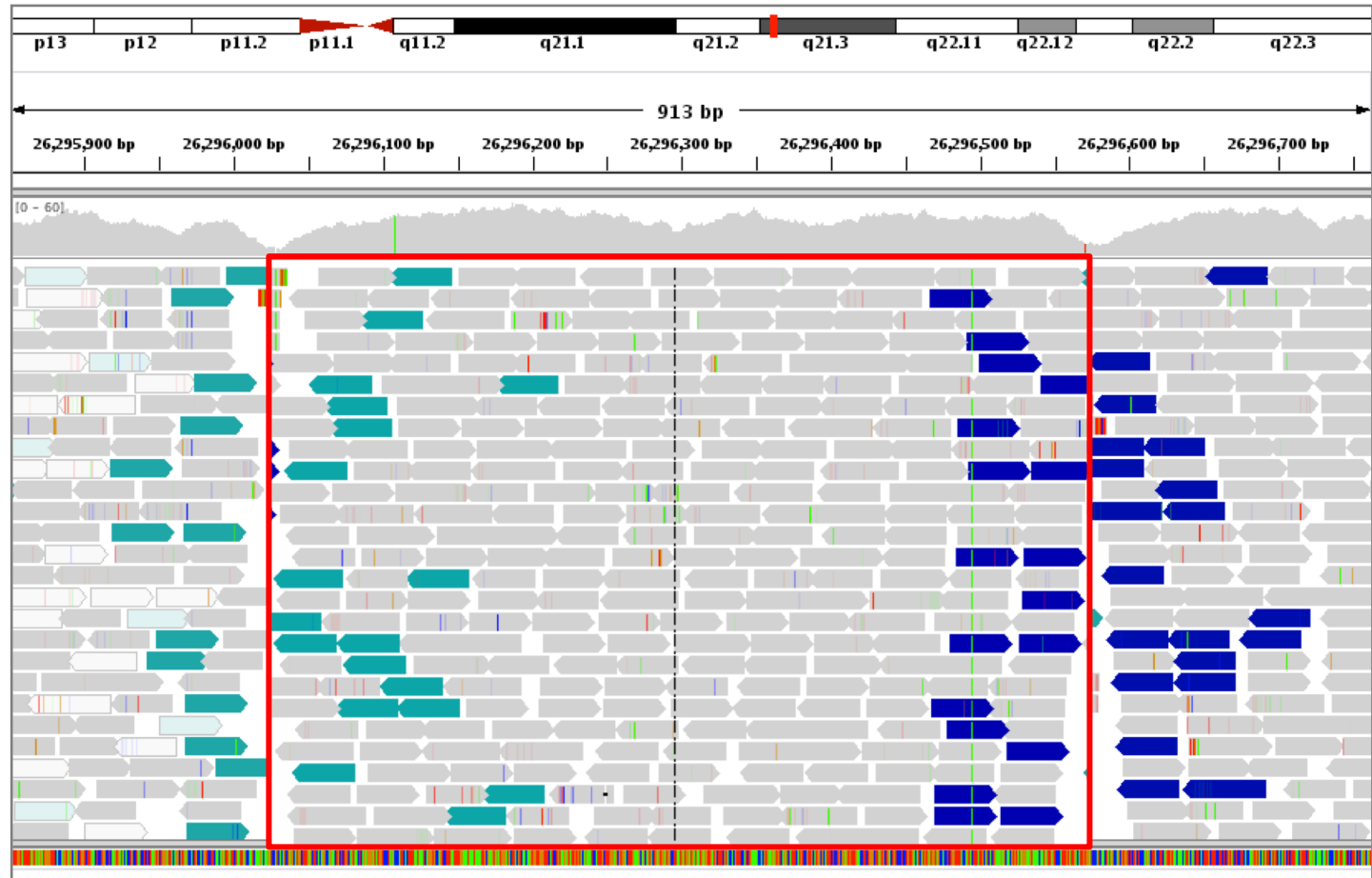
Inversion

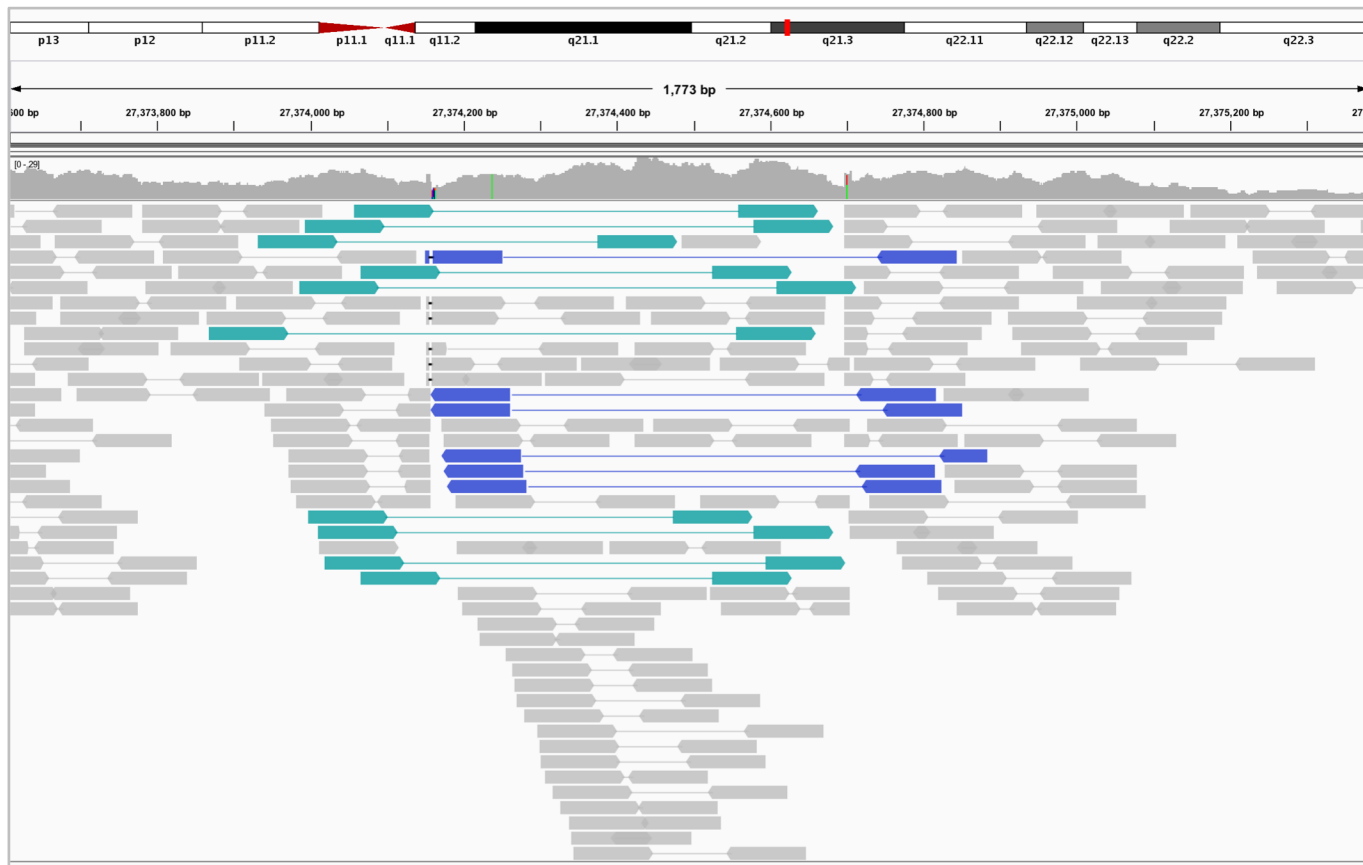


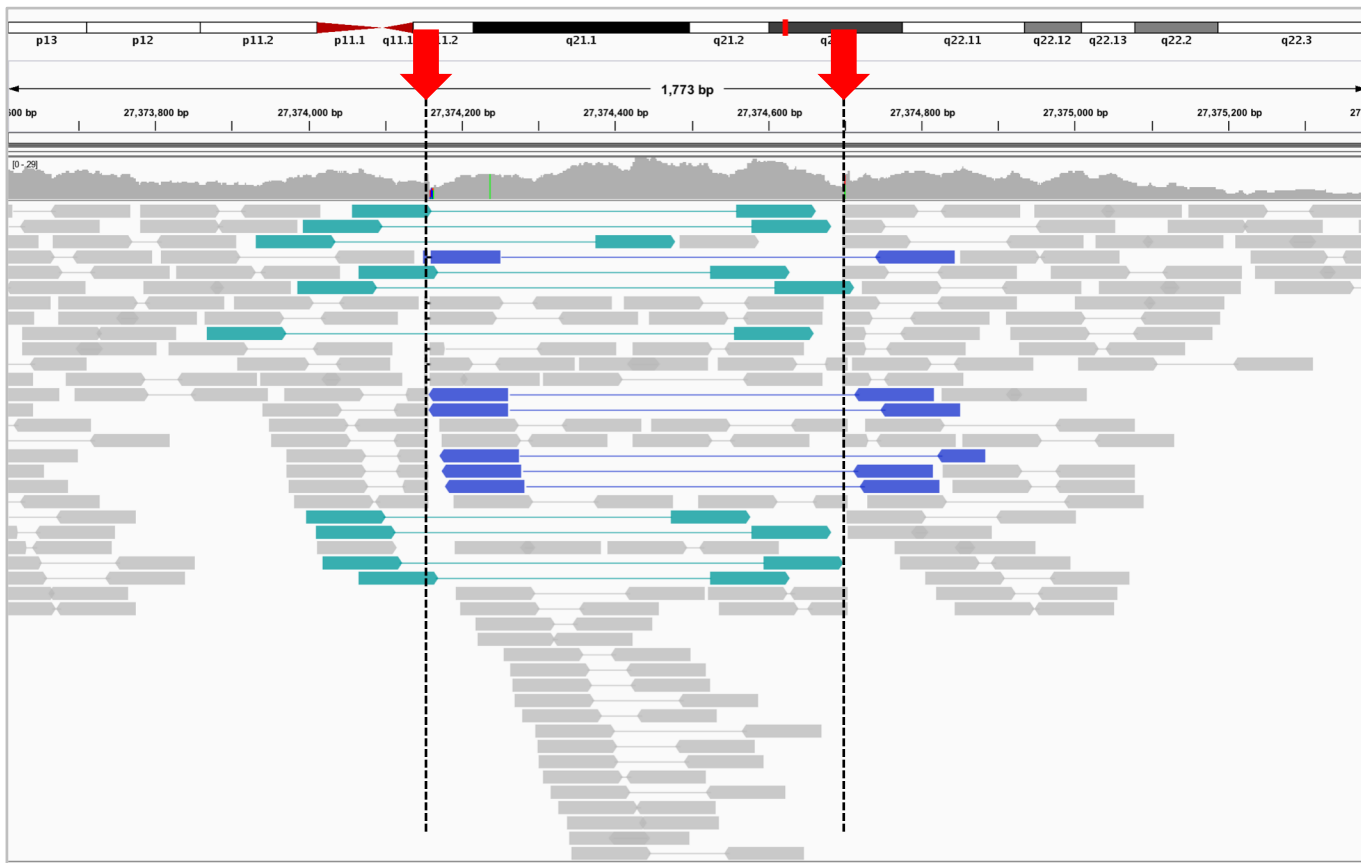


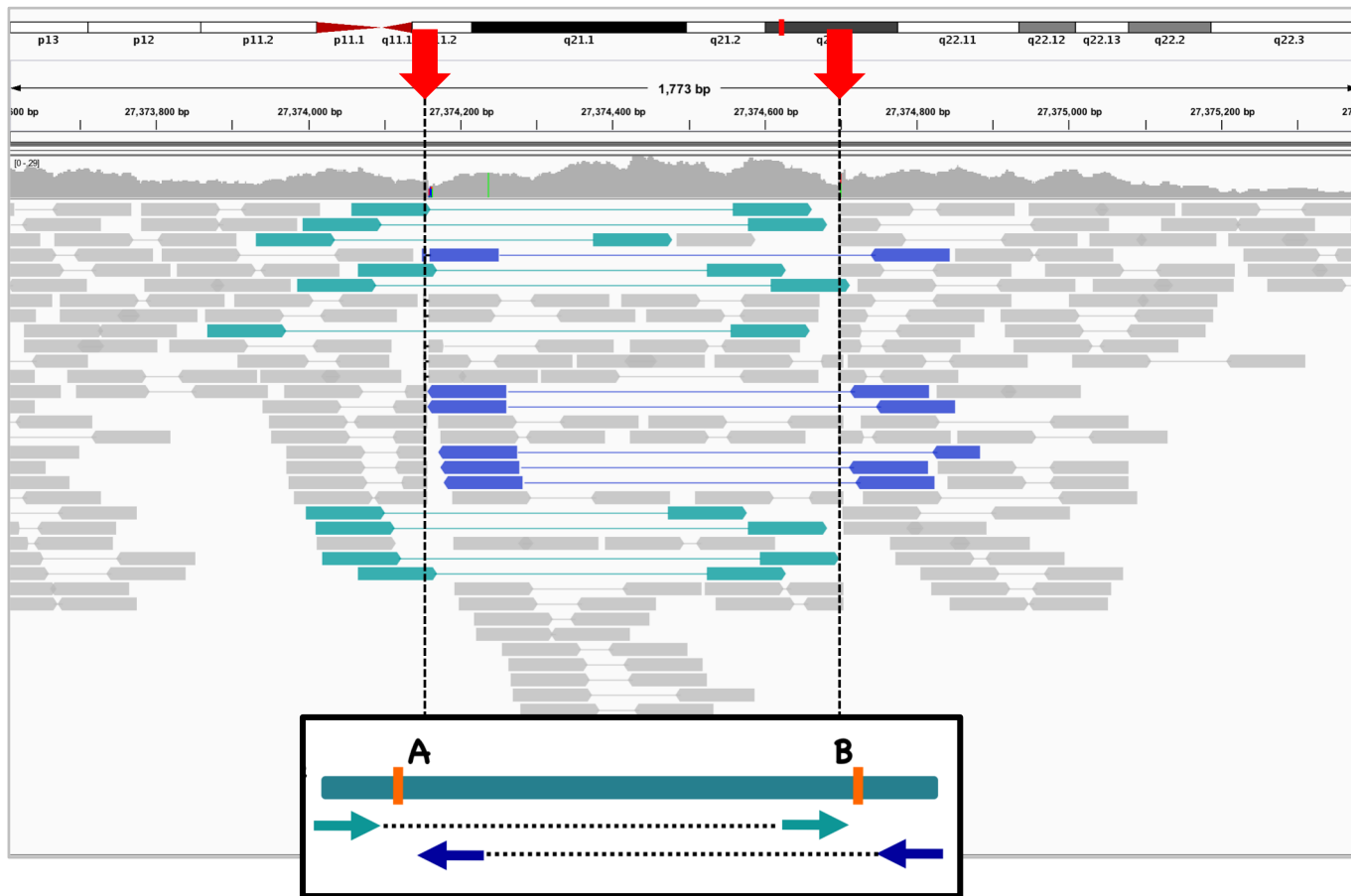




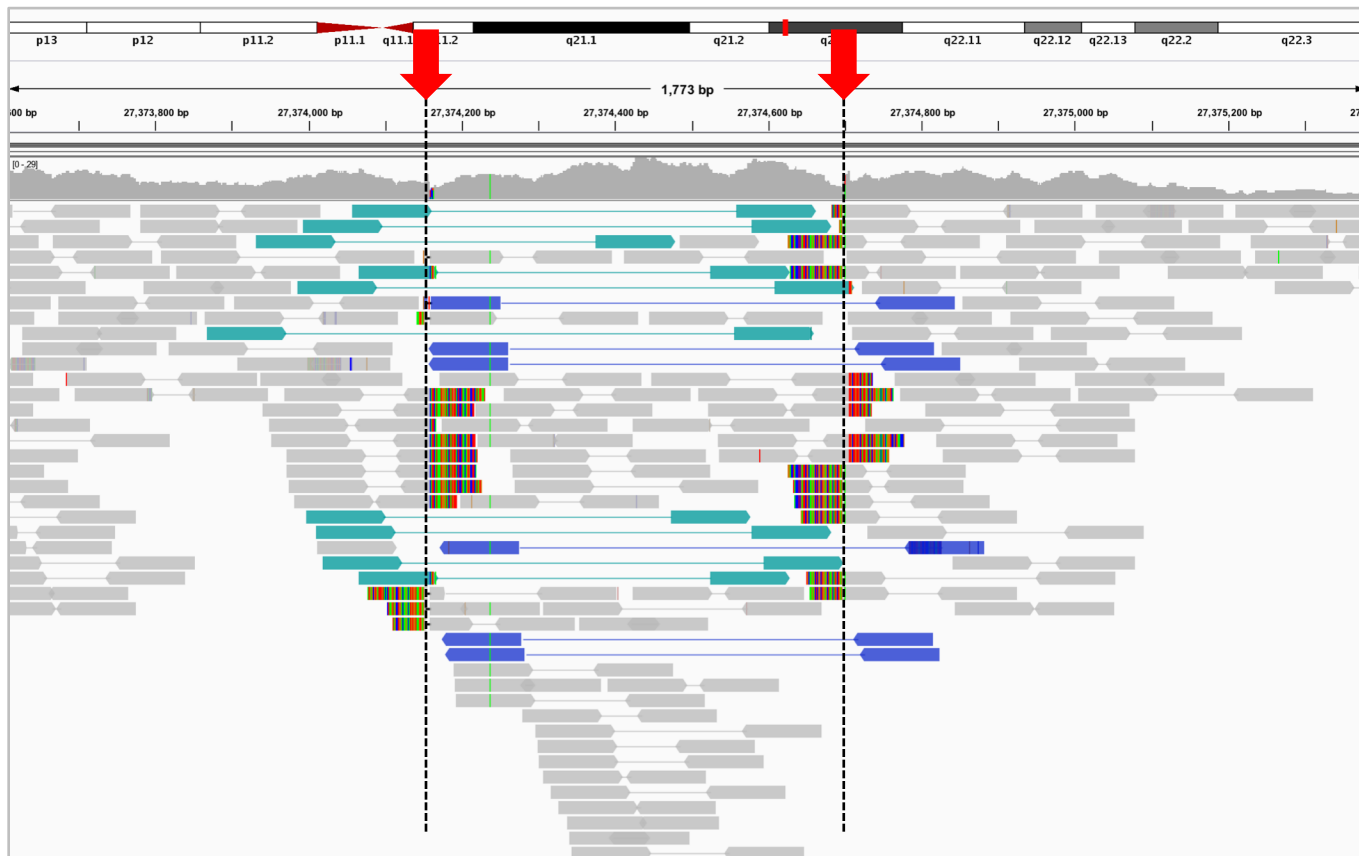














Interactive genome visualizations in your

- Publication
- Web page
- Lab notebook
- Emails, Twitter, ...

Cell

ARTICLE | [VOLUME 171, ISSUE 2, P305-320.E24, OCTOBER 05, 2017](#)

Cohesin Loss Eliminates All Loop Domains

Suhas S.P. Rao • Su-Chen Huang • Brian Glenn St Hilaire • Jesse M. Engreitz • Elizabeth M. Perez •
[Kyong-Rim Kieffer-Kwon](#) • [Adrian L. Sanborn](#) • [Sarah E. Johnstone](#) • [Gavin D. Bascom](#) • [Ivan D. Bochkov](#) •
[Xingfan Huang](#) • [Muhammad S. Shamim](#) • [Jaeweon Shin](#) • [Douglass Turner](#) • [Ziyi Ye](#) • [Arina D. Omer](#) •
[James T. Robinson](#) • [Tamar Schlick](#) • [Bradley E. Bernstein](#) • [Rafael Casellas](#) • [Eric S. Lander](#) •
[Erez Lieberman Aiden](#)  ¹⁸  • [Show less](#) • [Show footnotes](#)

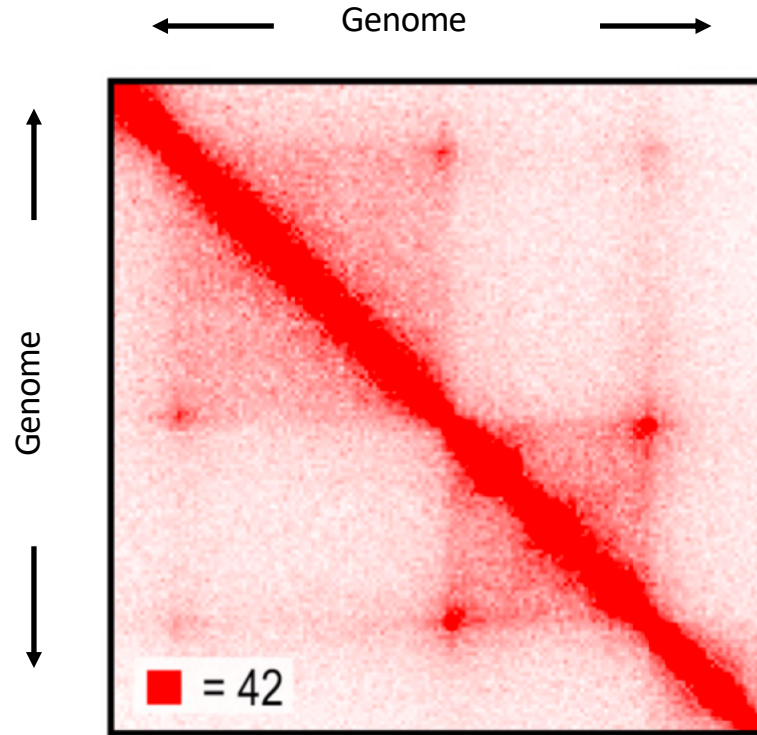
[Open Archive](#) • DOI: <https://doi.org/10.1016/j.cell.2017.09.026> •



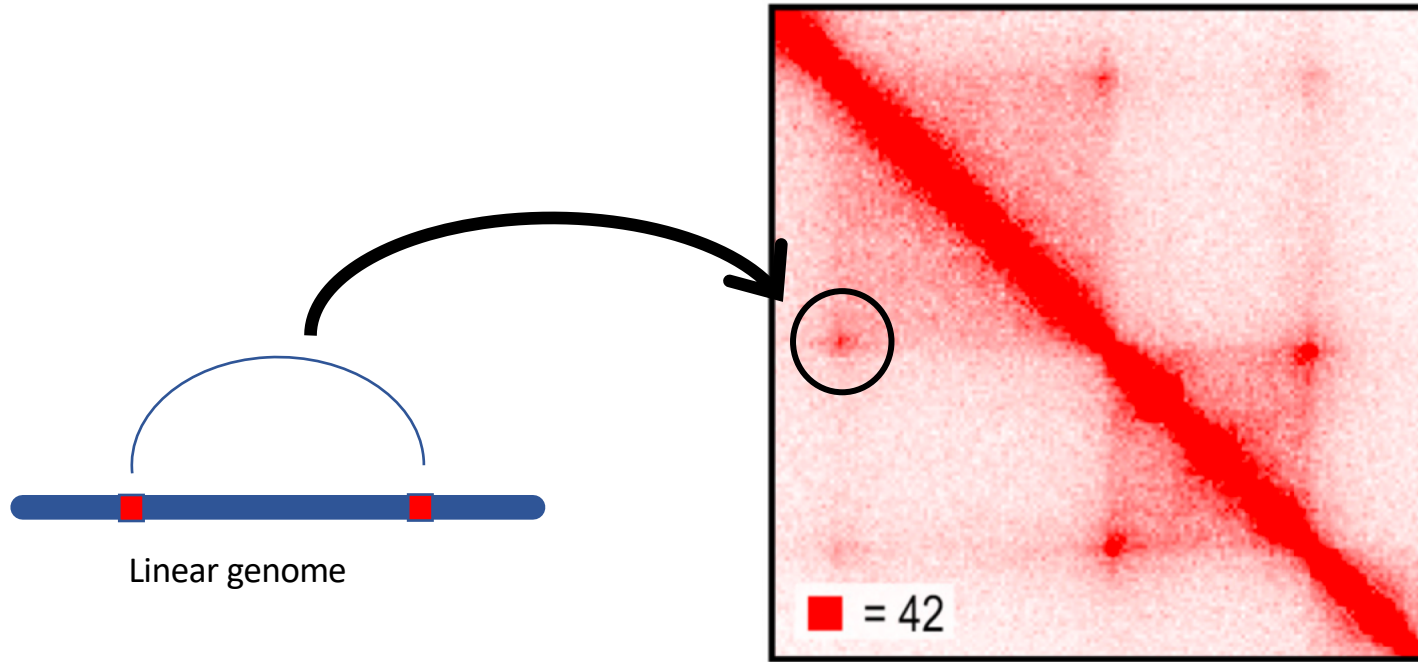
Cohesin-independent
loops and

HIC Contact Map

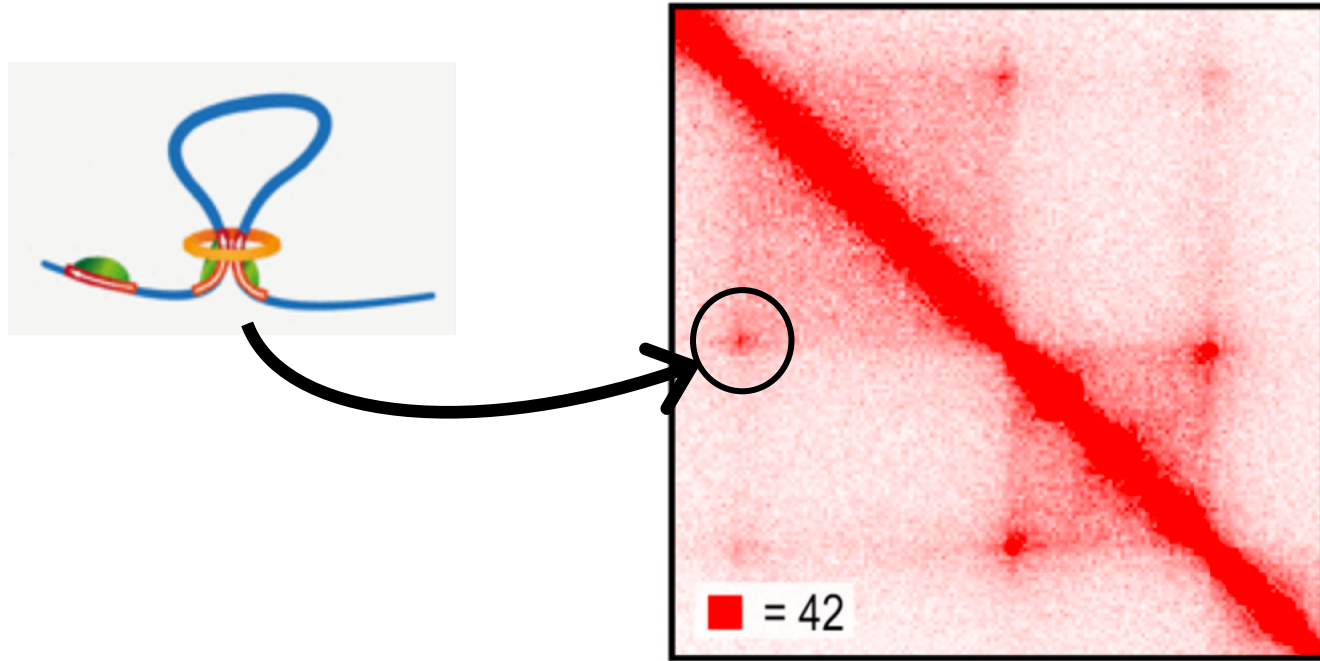
Bright spots highlight
regions of frequent
contact



Interactive figures



Interactive figures



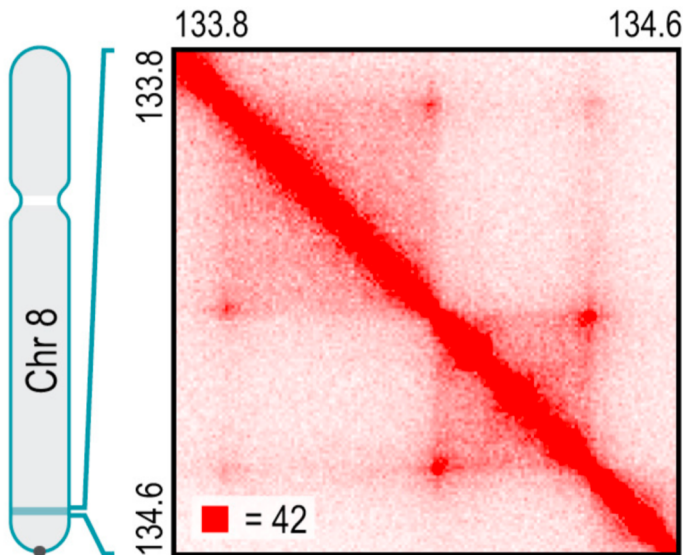
Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins, Wultz et al, The EMBO Journal (2017) e201798004

Interactive figures

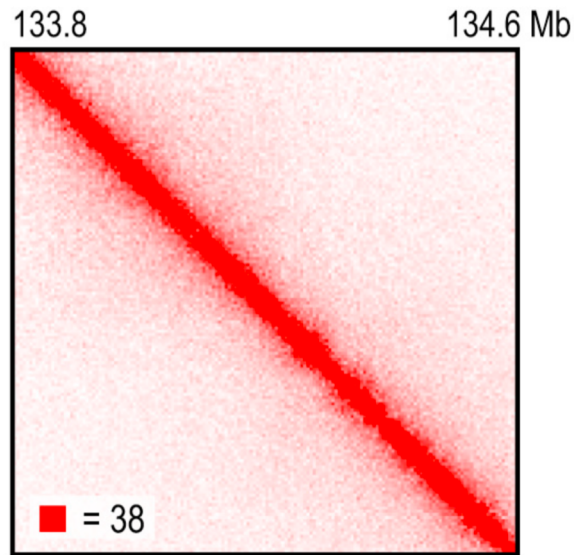
A



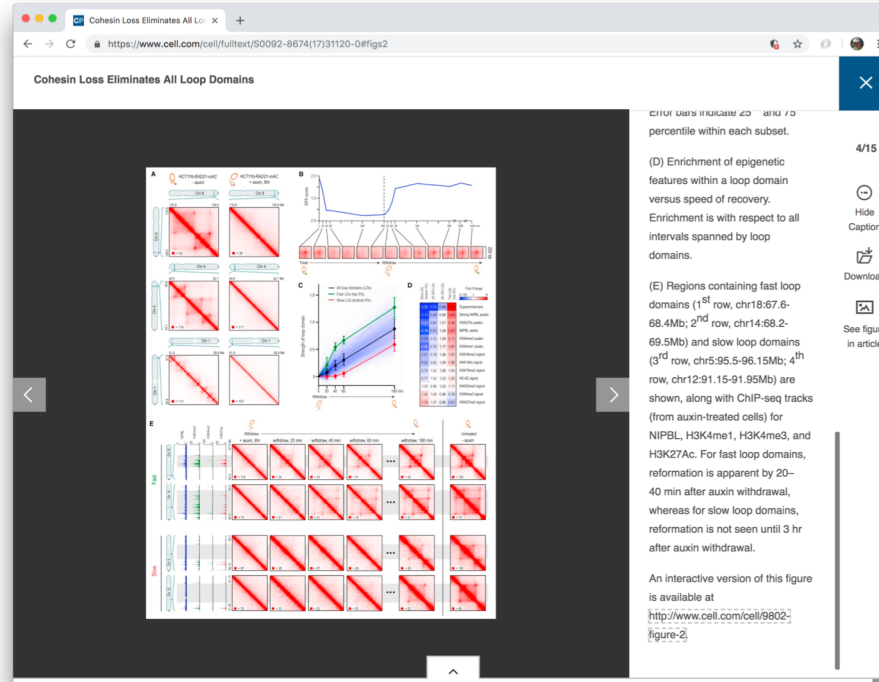
HCT116-RAD21-mAC
- auxin



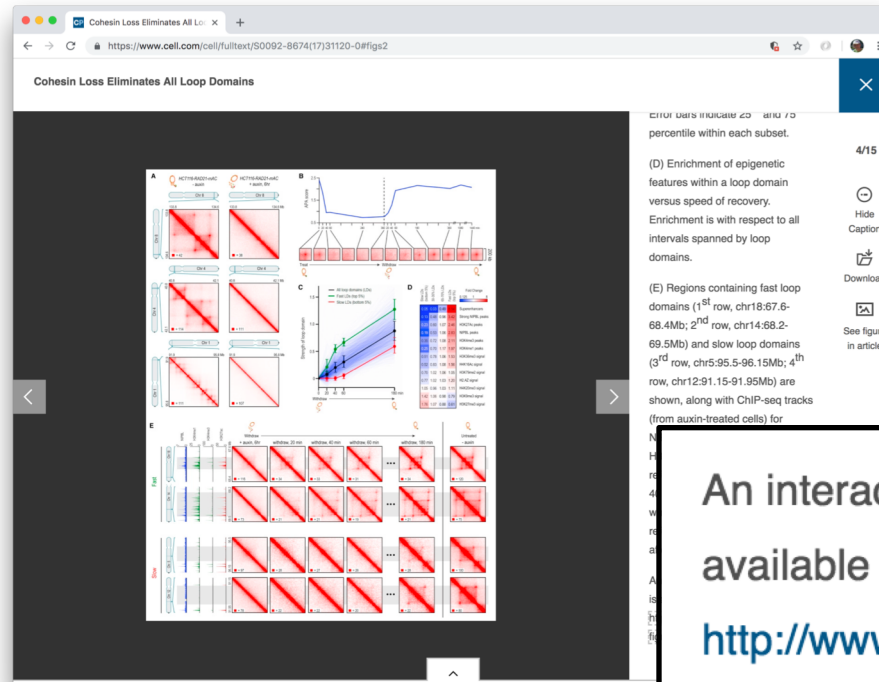
HCT116-RAD21-mAC
+ auxin, 6hr



Interactive figures

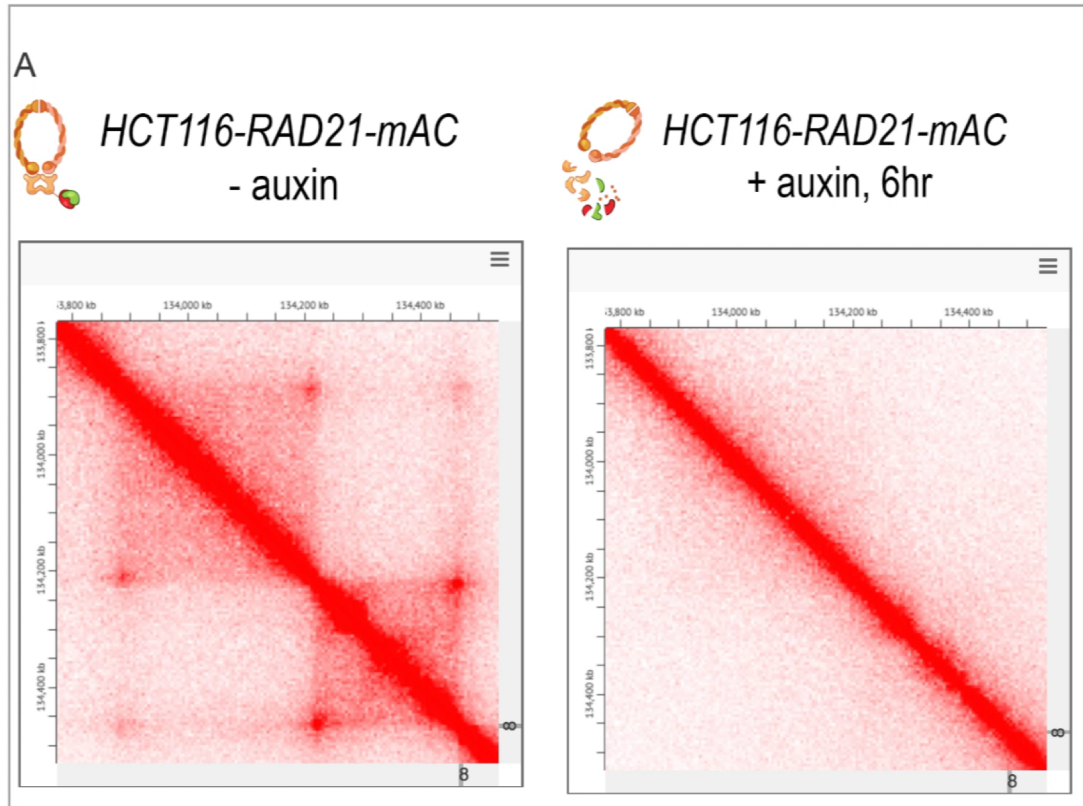


Interactive figures



An interactive version of this figure is available at <http://www.cell.com/cell/9802-figure-2>.

Figure 2. Cohesin Degradation Eliminates Loop Domains



Interactive figures with IGV-Web

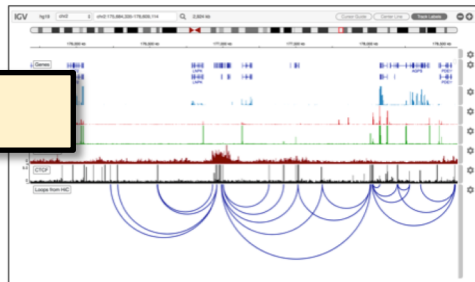
<http://bit.ly/2xDF7Sc>



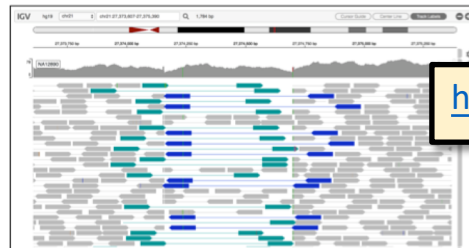
<http://bit.ly/2xBgiq3>



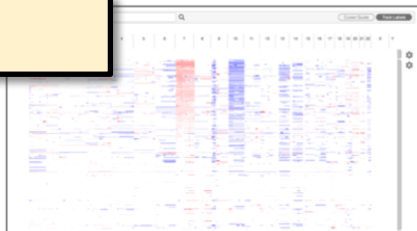
<http://bit.ly/2C8liWN>



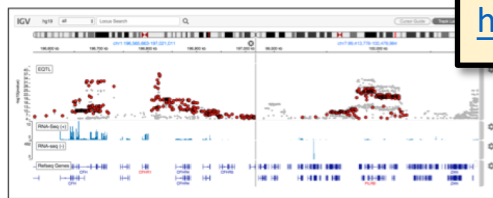
<http://bit.ly/2NFgxdq>



<http://bit.ly/2MmXWhS>



<http://bit.ly/2CCK1DJ>



Demonstration



<http://bit.ly/2PEet3h>

IGV

Not Secure | igv.org/app/

IGV

Genome Load Track Share Bookmark

IGV hg19 all all

Cursor Guide Track Labels

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y

Refseq Genes

Zoom in to see features

Help

Demonstration

1. Load files <http://bit.ly/2J2JTxF>
2. Go to SLC25A3 <http://bit.ly/2CocRa5>
3. Set track color <http://bit.ly/2PMSfMo>
4. Zoom in <http://bit.ly/2PETIKp>
5. Share

IGV currently funded by:

- NCI's Informatics Technology for Cancer Research (ITCR)

Previous funding:

- National Cancer Institute
- Starr Cancer Consortium

Poster #1669

IGV-Web: An application for viewing and sharing genomic datasets

Friday 2pm – 4pm