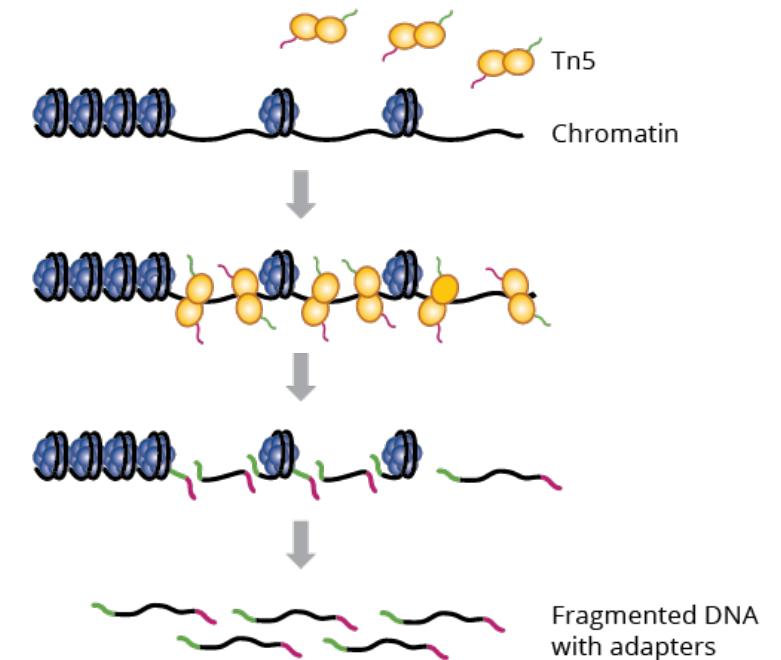
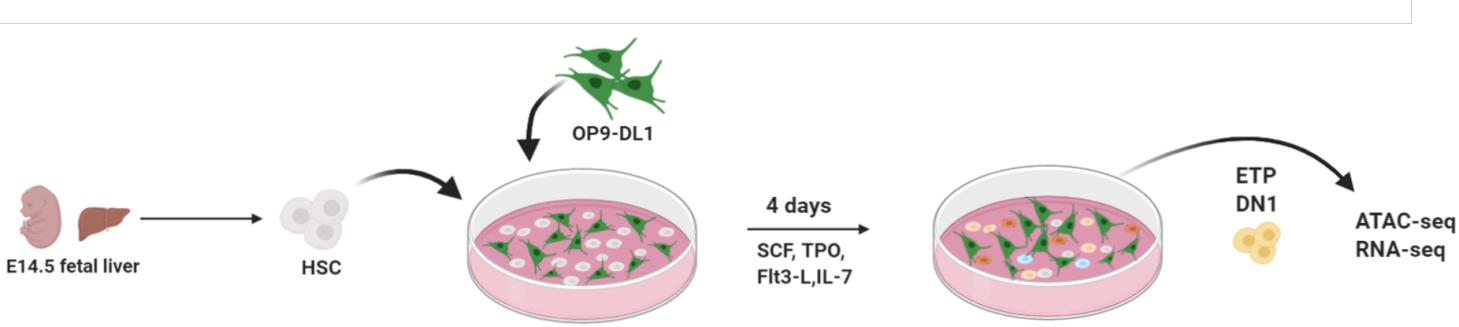
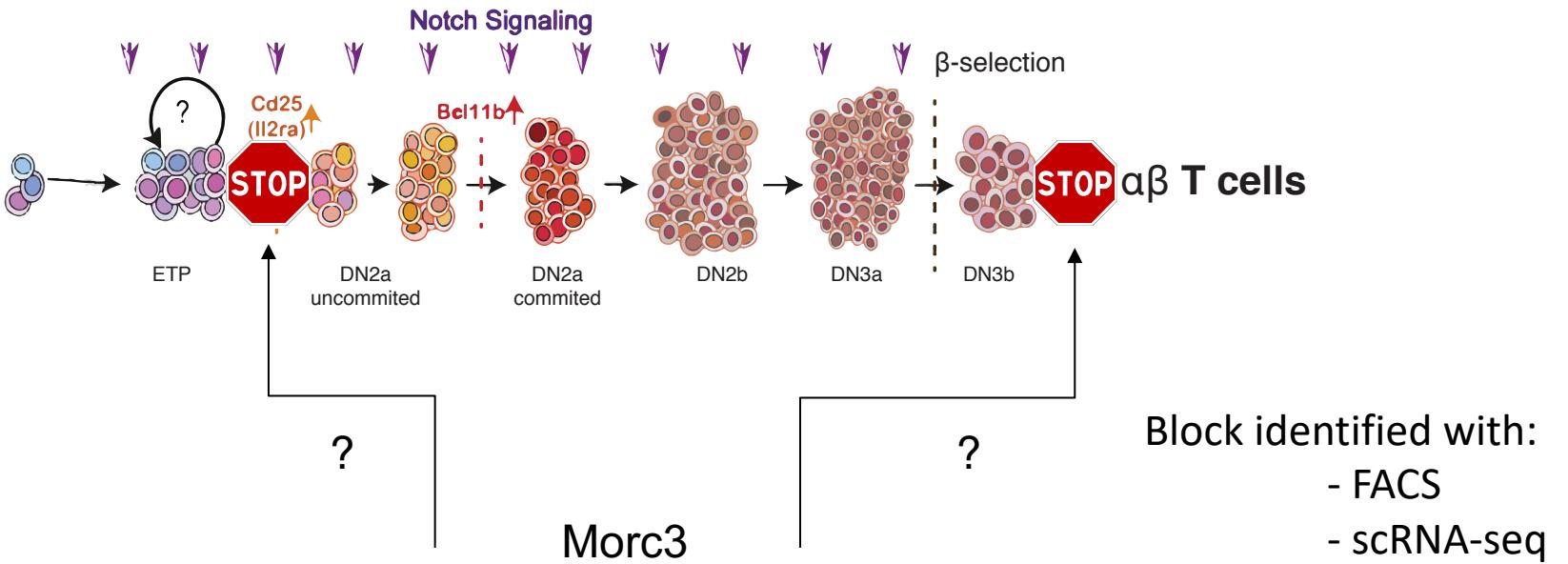


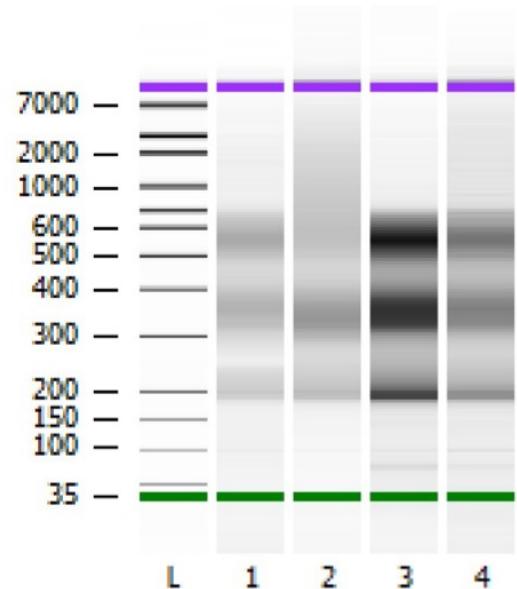
DN1 ATAC-seq

Veronica

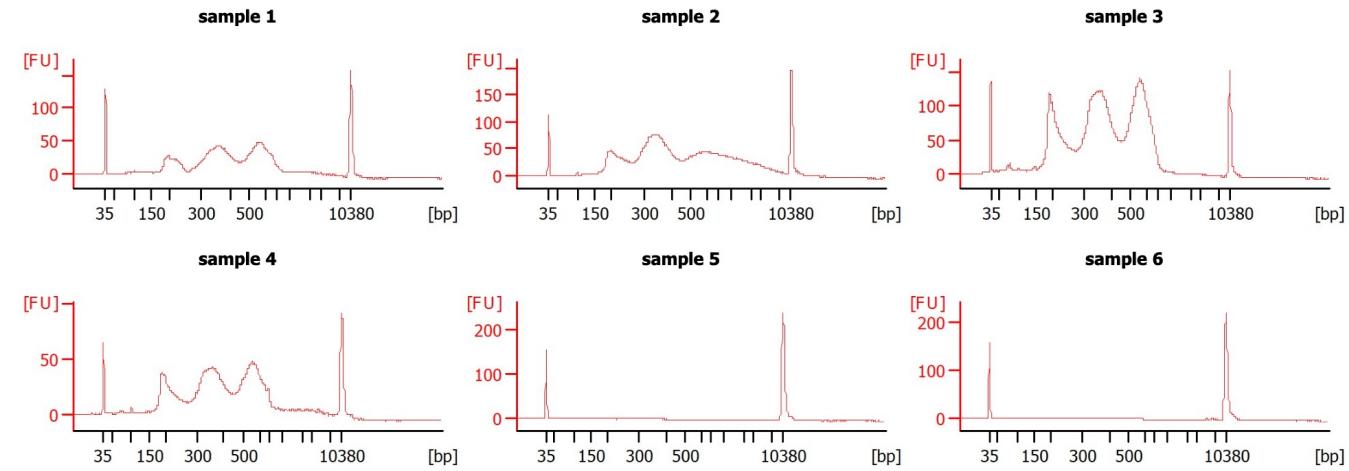
Background



- 50000 DN1 cells from OP9:
 - 2 x WT
 - 2 x KO

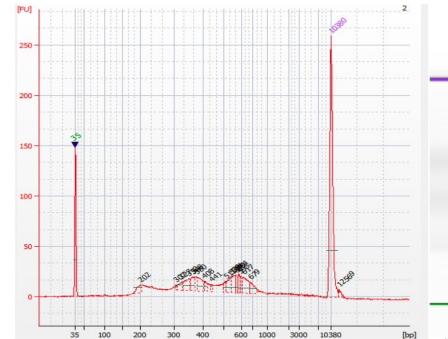


sample nb	name	primer 1	primer 2	Concentration Qubit	Cycles added to second PCR
1	WT1	S507	N702	1,28	12
2	WT2	S507	N704	1,98	13
3	KO1	S507	N705	3,15	10
4	KO2	S507	N707	1,97	8

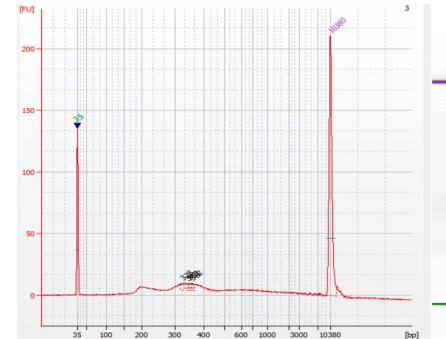


#	Library Name	Library Type	Conc. (ng/uL)	Conc. (nM)	Size (bp)	Result*	
1	DN1_WT_rep1	ETC	0.2	0.72	434	Fail	Low Quantity to (Run or Capture)
2	DN1_WT_rep2	ETC	0.26	0.97	409	Fail	Low Quantity to (Run or Capture)
3	DN1_KO_rep1	ETC	0.85	3.26	403	Fail	Low Quantity to (Run or Capture)
4	DN1_KO_rep2	ETC	0.24	0.88	417	Fail	Low Quantity to (Run or Capture)

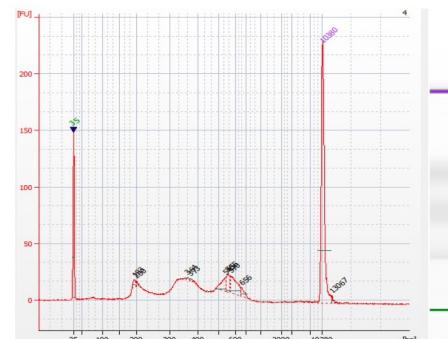
Click to Enlarge =>1:Library : DN1_WT_rep1



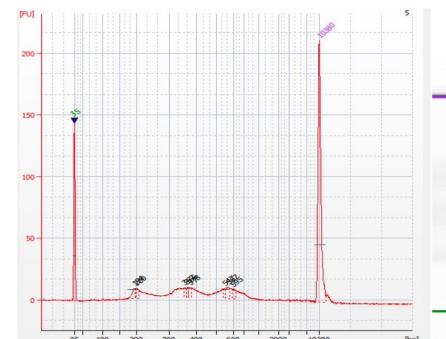
Click to Enlarge =>2:Library : DN1_WT_rep2



Click to Enlarge =>3:Library : DN1_KO_rep1



Click to Enlarge =>4:Library : DN1_KO_rep2



QC reports provided by Macrogen

Client Name	Macrogen Europe
Company / Institution	Macrogen Europe
Order Number	HN00155645
Type of Read	Paired-end
Read Length	151
Number of Samples	3
Type of Sequencer	Illumina platform

Download link	File size	md5sum
DN1_KO_rep1_1.fastq.gz	1.0G	c2095e9bc993755383be9247d830e2e3
DN1_KO_rep1_2.fastq.gz	991.1M	8d654303f1407119a14ccf671cd4d375
DN1_WT_rep1_1.fastq.gz	347.0M	31260d4e330390984afb1aabf1f39e60
DN1_WT_rep1_2.fastq.gz	322.8M	ff40f5e78e477fc6632206f1cc6ece8f
DN1_WT_rep2_1.fastq.gz	295.6M	339faea0788651328629e0dd3722b12f
DN1_WT_rep2_2.fastq.gz	283.9M	ee417d9b45ea135496b10b6edebf49ce

Sample ID	Total read bases (bp)	Total reads	GC(%)	AT(%)	Q20(%)	Q30(%)
DN1_KO_rep1	5,031,883,230	33,323,730	51.04	48.96	94.66	89.32
DN1_WT_rep1	1,474,346,182	9,763,882	51.46	48.54	94.92	89.57
DN1_WT_rep2	1,013,035,444	6,708,844	51.28	48.72	89.67	77.98

Table 1. Raw data Stats (maximum 20 samples)

Sample ID	Total read bases (bp)	Total reads	GC(%)	AT(%)	Q20(%)	Q30(%)
DN1_KO_rep1	5,031,883,230	33,323,730	51.04	48.96	94.66	89.32
DN1_WT_rep1	1,474,346,182	9,763,882	51.46	48.54	94.92	89.57
DN1_WT_rep2	1,013,035,444	6,708,844	51.28	48.72	89.67	77.98

Analysis

- Done with our [Snakemake-ATAC-seq pipeline](#)

Quality controls

General Statistics

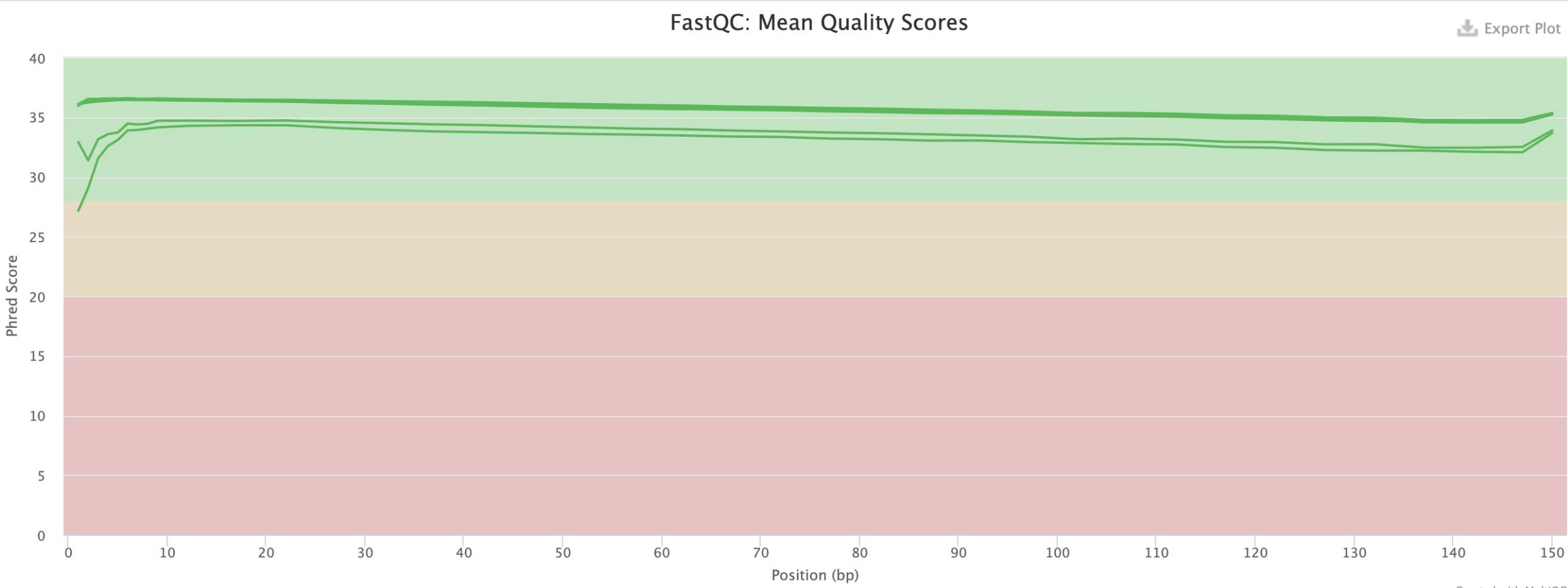
Copy table Configure Columns Plot Showing 9/9 rows and 7/9 columns.

Sample Name	% Aligned	Insert Size	% Dups	% Dups	% GC	Length	M Seqs
DN1KO1	92%	168 bp	23.4%				
DN1KO1_R1				25.3%	47%	111 bp	16.6
DN1KO1_R2				25.7%	47%	111 bp	16.6
DN1WT1	93%	148 bp	19.8%				
DN1WT1_R1				18.2%	48%	110 bp	4.9
DN1WT1_R2				18.6%	48%	110 bp	4.9
DN1WT2	95%	178 bp	11.2%				
DN1WT2_R1				6.9%	49%	123 bp	3.4
DN1WT2_R2				6.9%	49%	123 bp	3.4

Quality controls

FastQC: Mean Quality Scores

 Export Plot

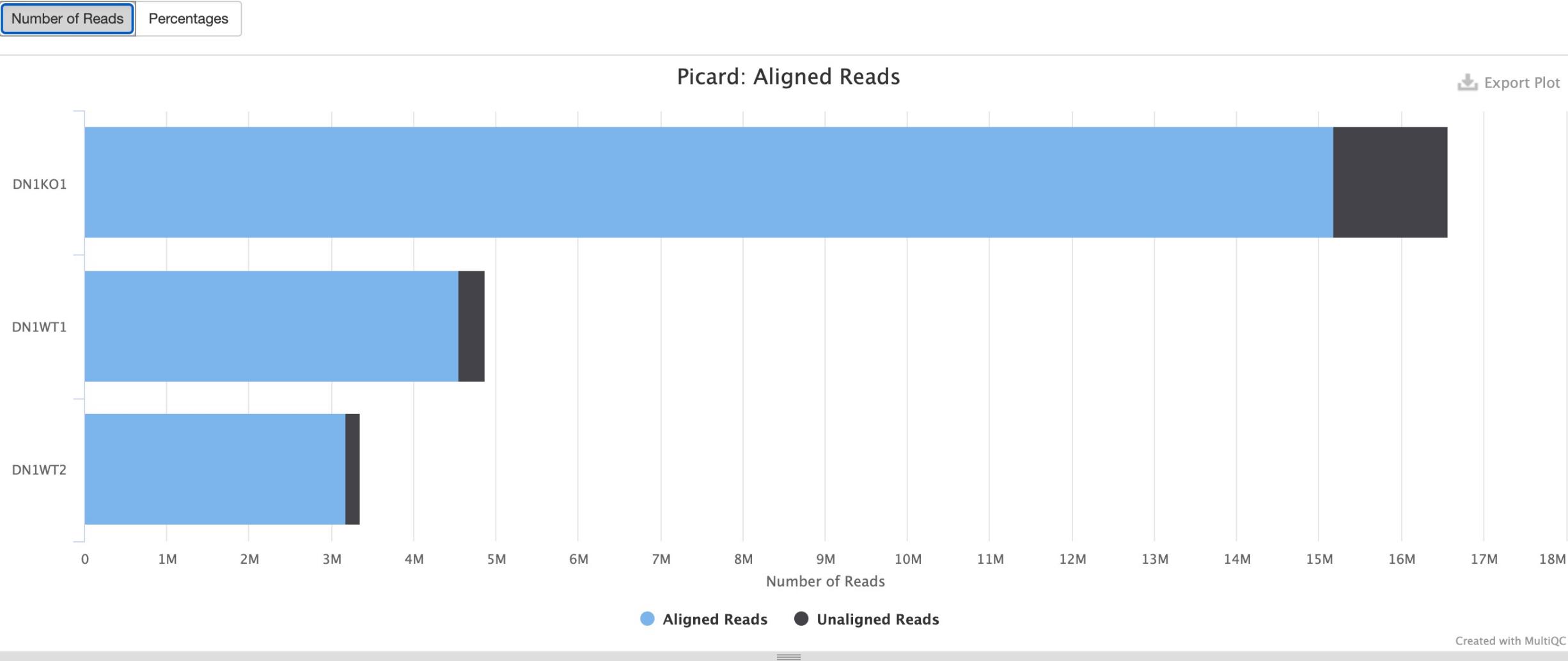


© 2024 QIAGEN BioIT

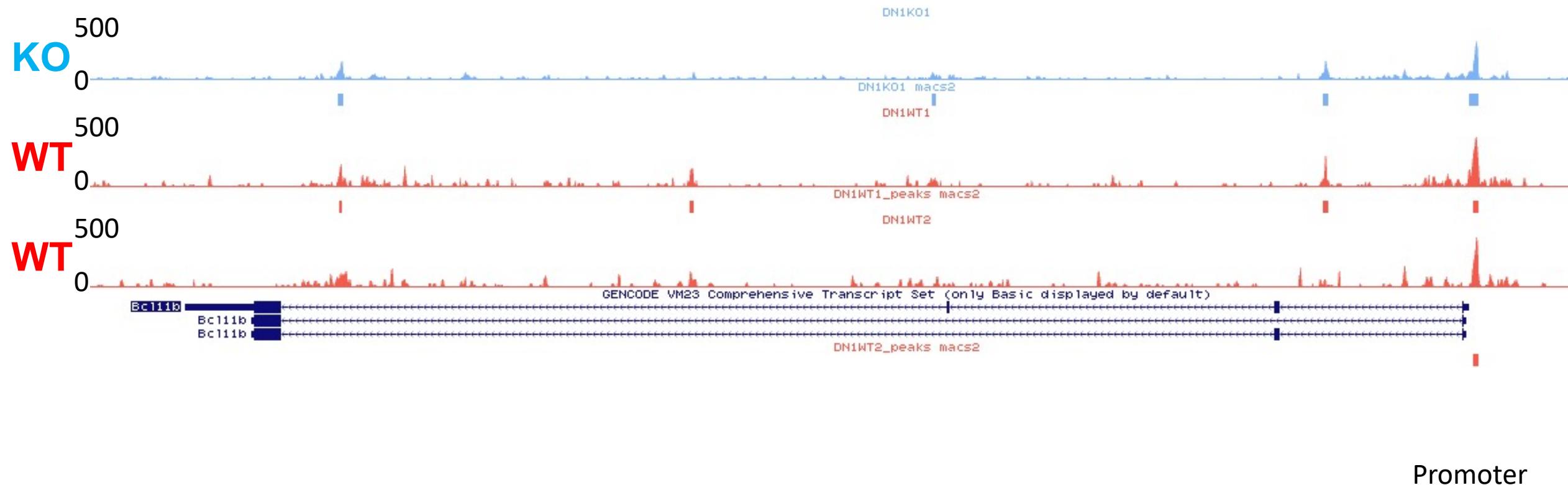
Quality controls

Alignment Summary

Please note that Picard's read counts are divided by two for paired-end data.

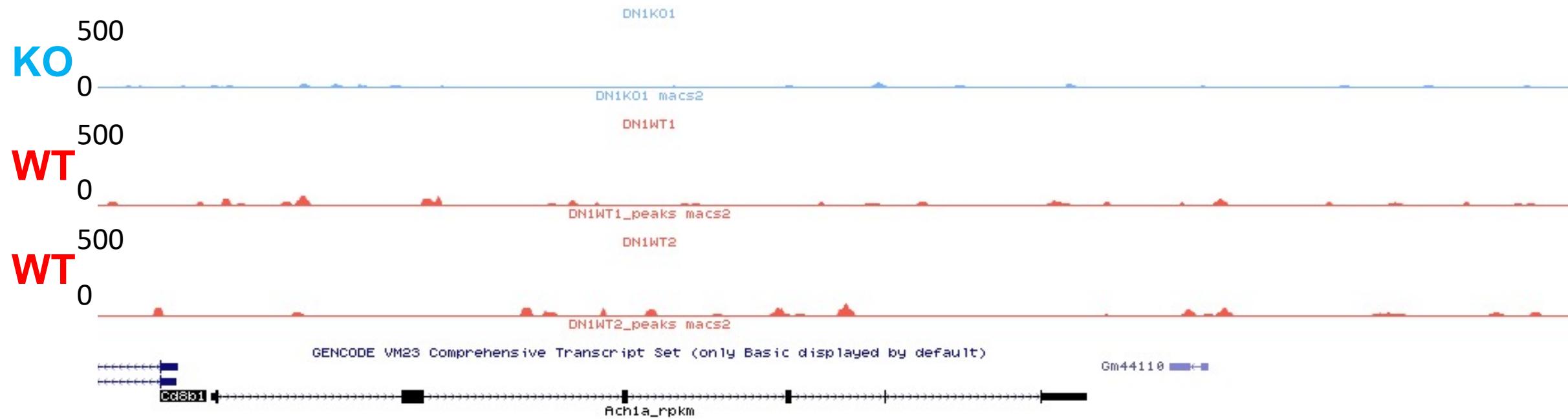


Bcl11b



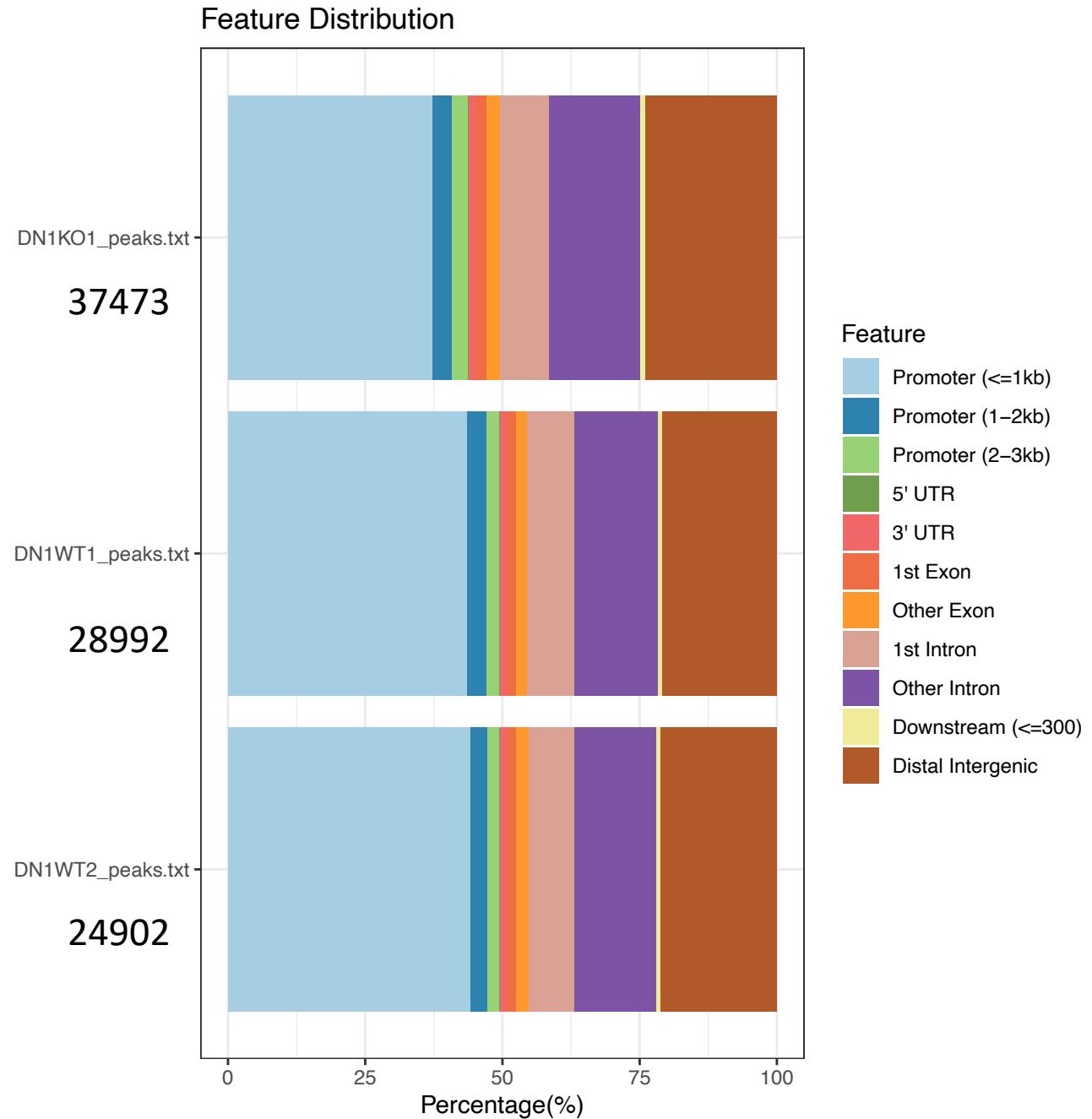
Promoter

Cd8b1



Peak calling: MACS2

name	peaks
WT1	28992
WT2	24902
KO1	37473



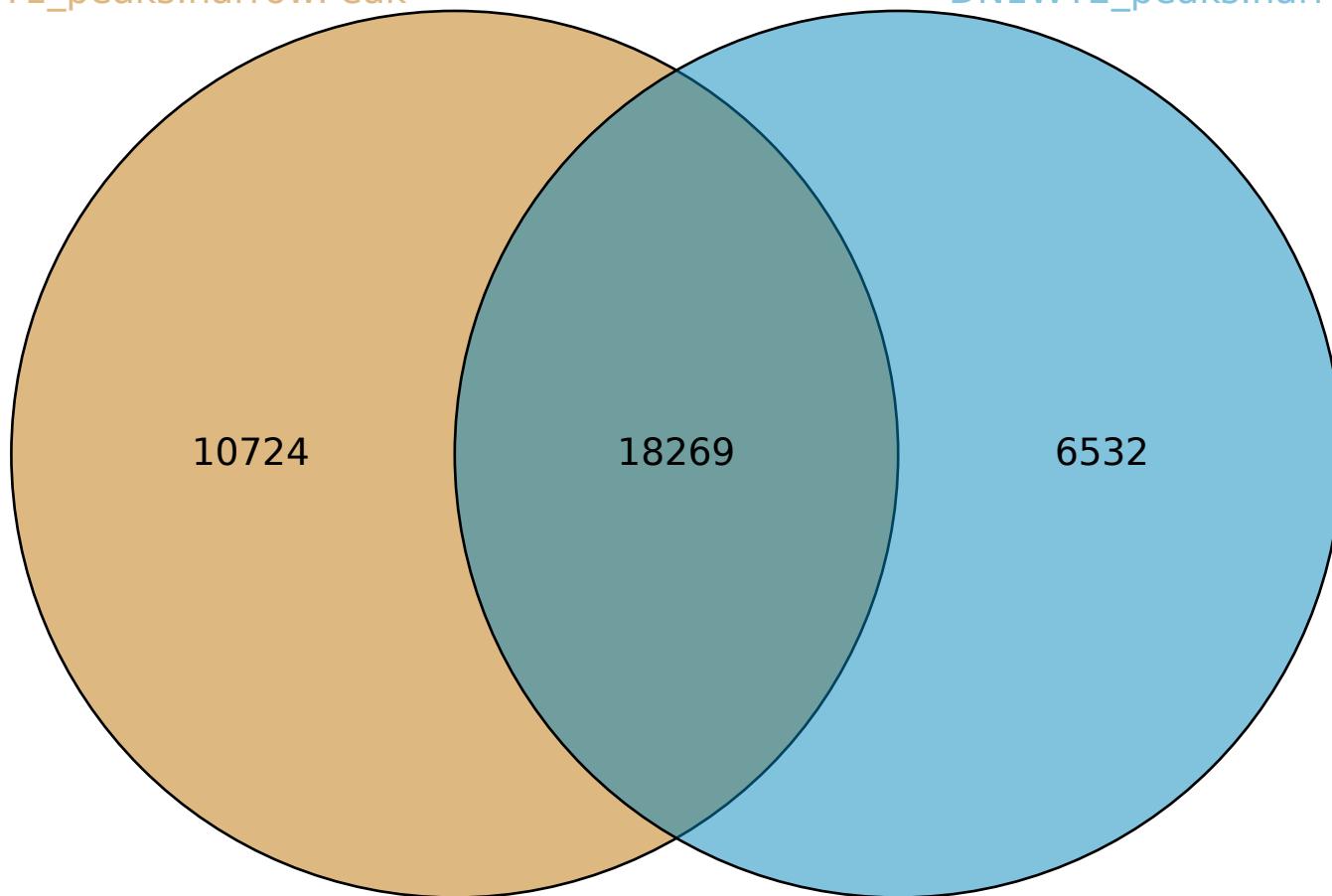
Overlap WT1 and WT2

DN1WT1_peaks.narrowPeak
DN1WT2_peaks.narrowPeak

DN1WT1_peaks.narrowPeak DN1WT2_peaks.narrowPeak

63% of WT1 peaks overlap
with WT2

73% of WT2 peaks overlap
with WT1



Differential peak analysis

2xWT versus 1xKO

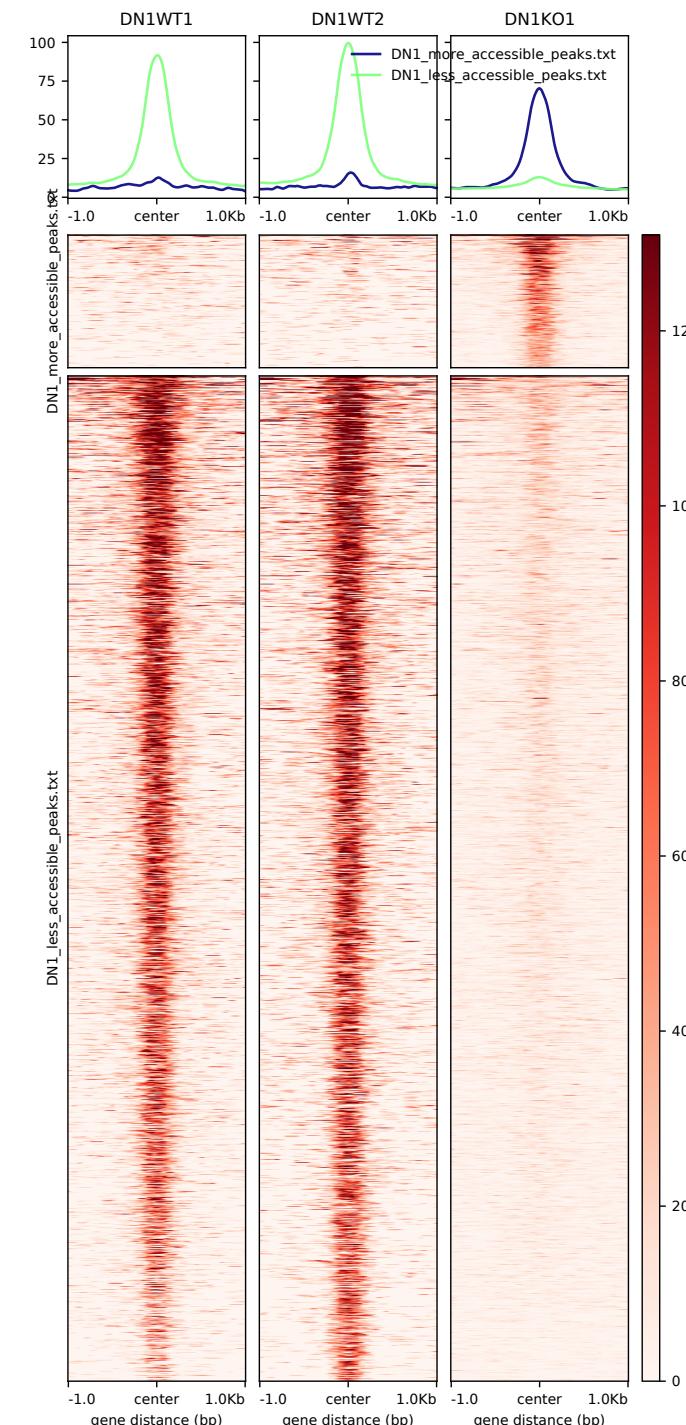
Homer **getDifferentialPeaks FC > 4 / p-value <0.05**

Relative to WT:

Less accessible in KO	2046
More accessible in KO	269

Average size less accessible peaks: 320bp
(SD=113bp)

Average size more accessible peaks: 284bp
(SD=145bp)



The *cis*-Regulatory Atlas of the Mouse Immune System

2x DN1 analyzed

99% Aligned reads (10M reads)

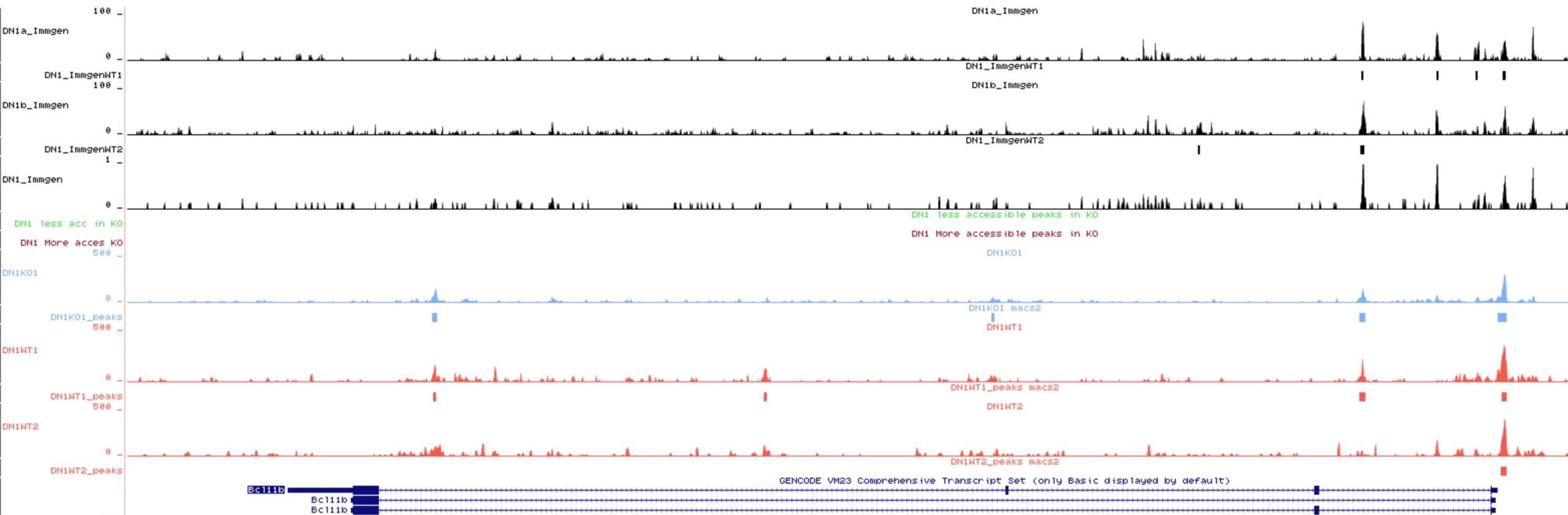
Phred > 35

Samples	Peaks
ImmgenWT1	58066
ImmgenWT2	42862

Authors

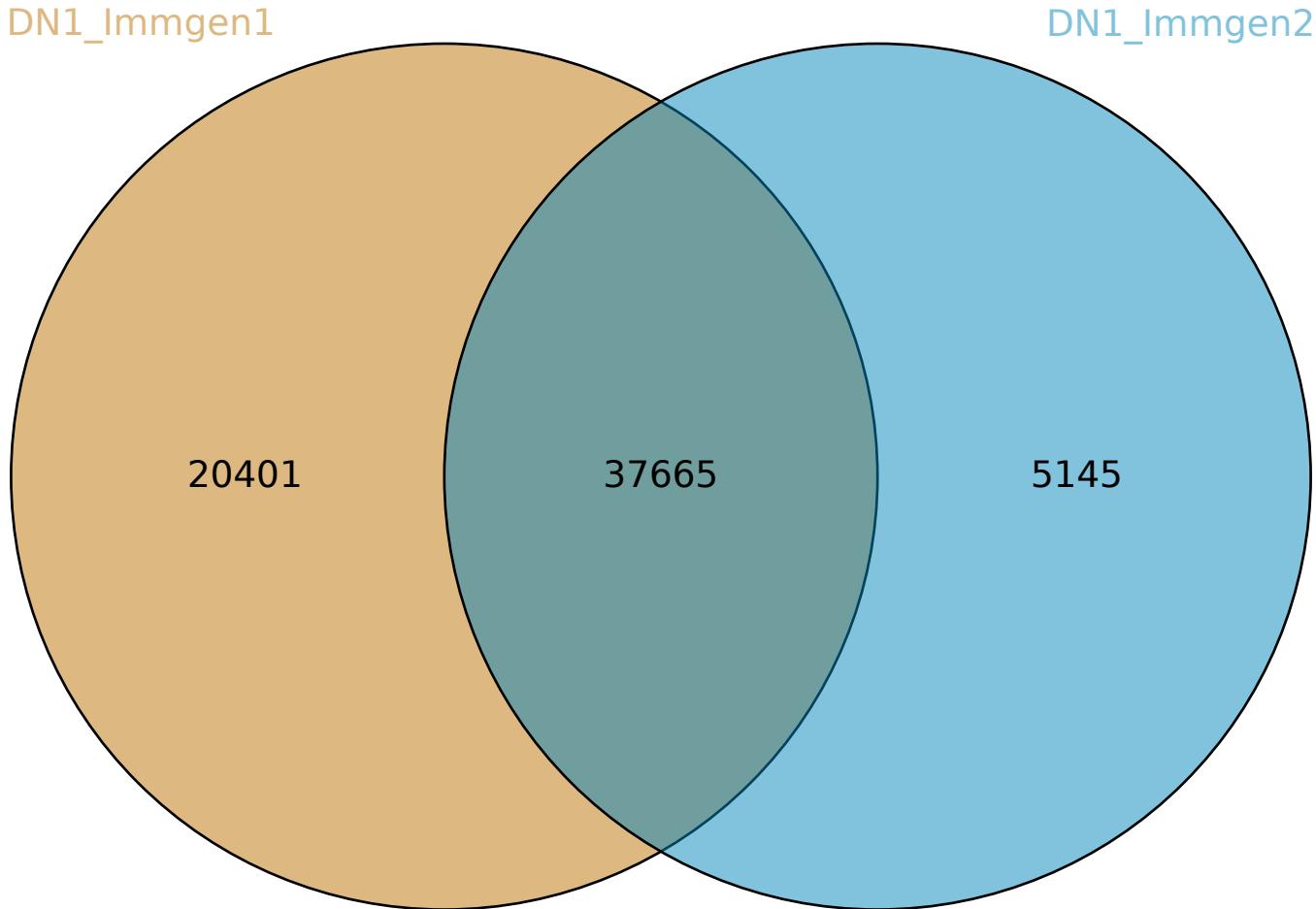
Hideyuki Yoshida, Caleb A. Lareau,
Ricardo N. Ramirez, ...,
Jason D. Buenrostro, Christophe Benoist,
the Immunological Genome Project

Bcl11b



Overlap Immgen Replicates

DN1_Immgen1
DN1_Immgen2

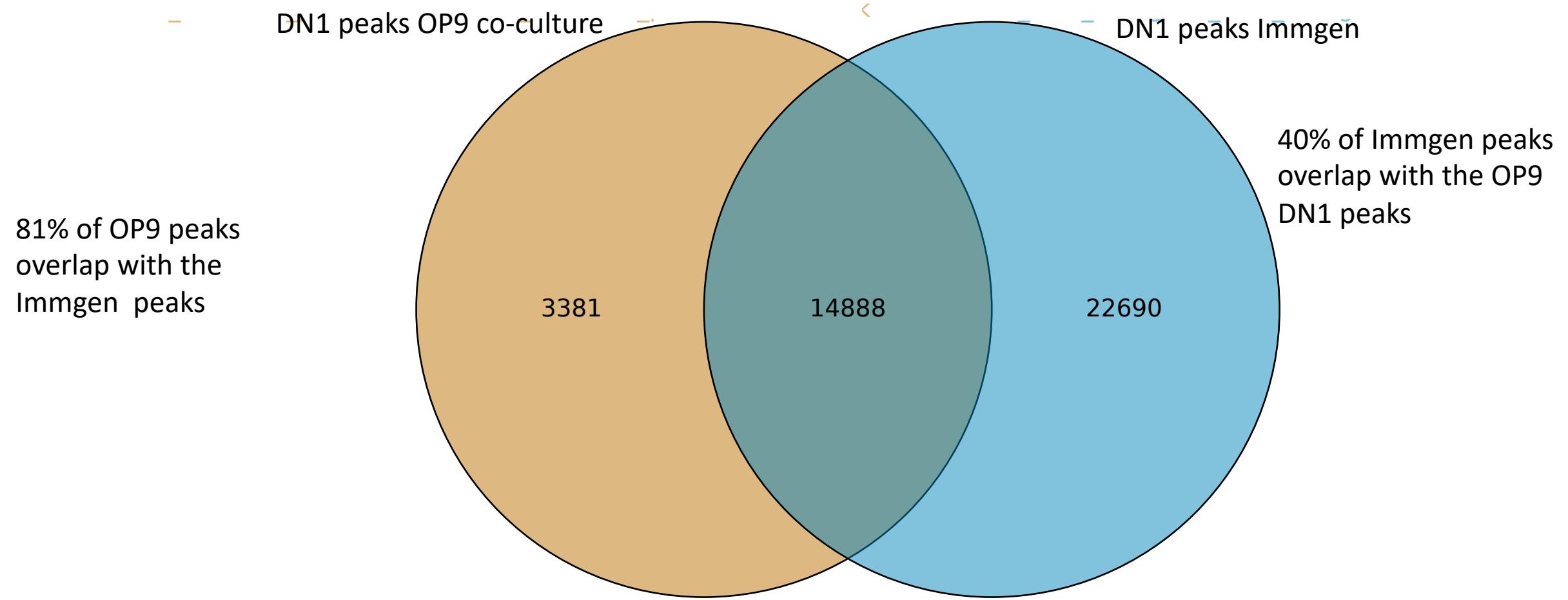


64% of WT1 overlaps with WT2

88% of WT2 overlaps with WT1

Overlap Immgen and OP9 peaks

11_DN1WT1_peaks.narrowPeak_DN1WT2_peaks.narrowPeak
11_DN1_Immgen1_DN1_Immgen2



The OP9 ATAC-seq detects less peaks than the Immgen DN1 but most of them are overlap. With our dataset we might be missing some change in accessibility at certain loci.

Motif analysis

Differential peaks are mostly found away from the gene promoters

Number of Peaks

Feature Distribution

2046

DN1_less_accessible_peaks.txt

269

DN1_more_accessible_peaks.txt

37473

DN1KO1_peaks.txt

28992

DN1WT1_peaks.txt

24902

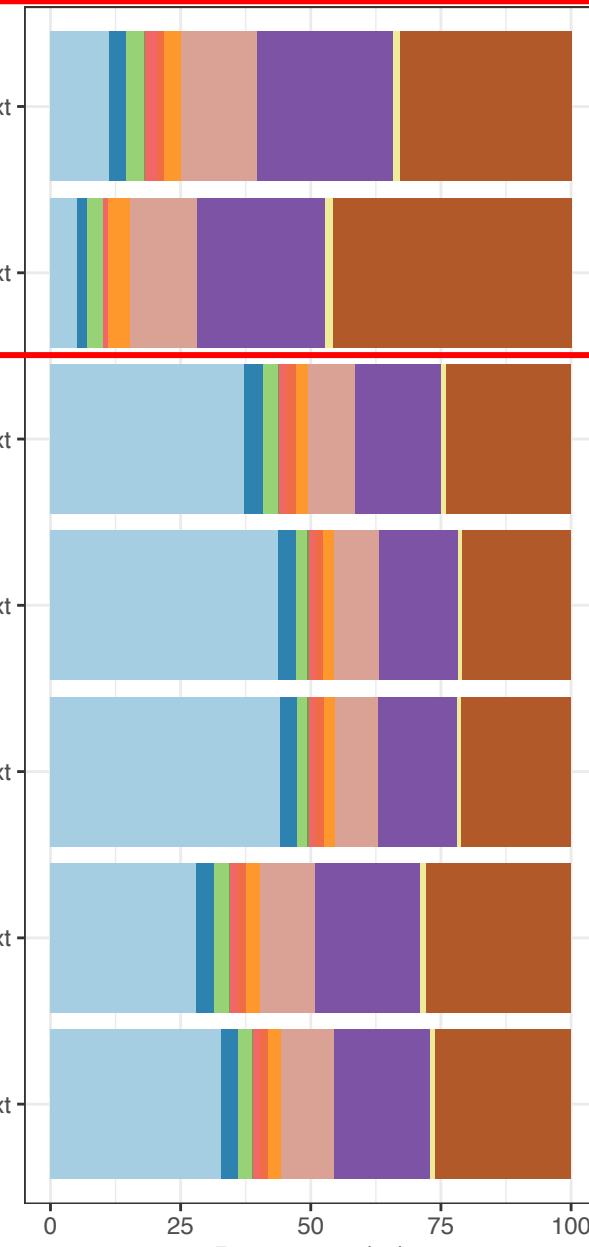
DN1WT2_peaks.txt

58065

ImmegenDN1A_peaks.txt

42861

ImmegenDN1B_peaks.txt

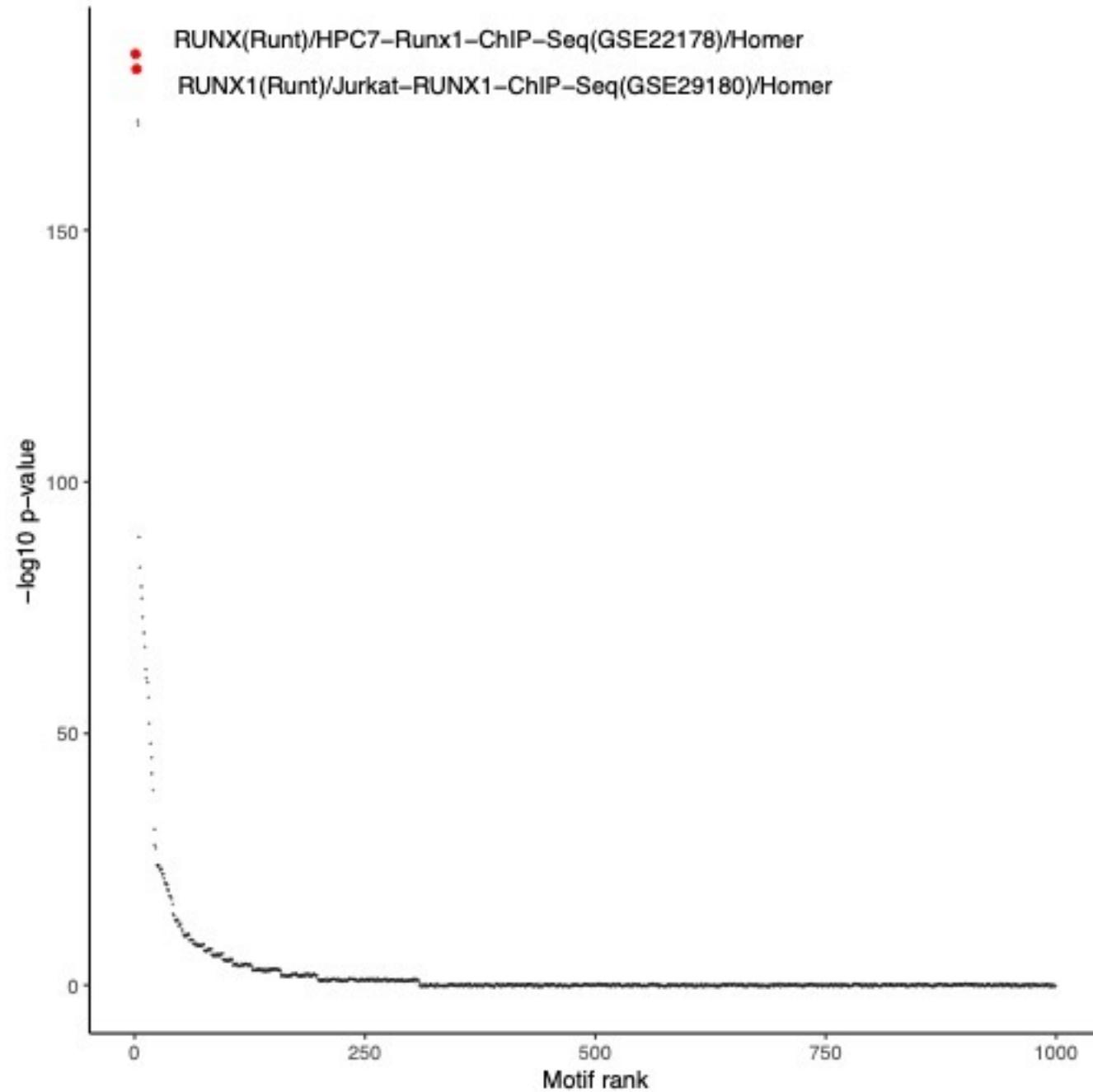


Feature

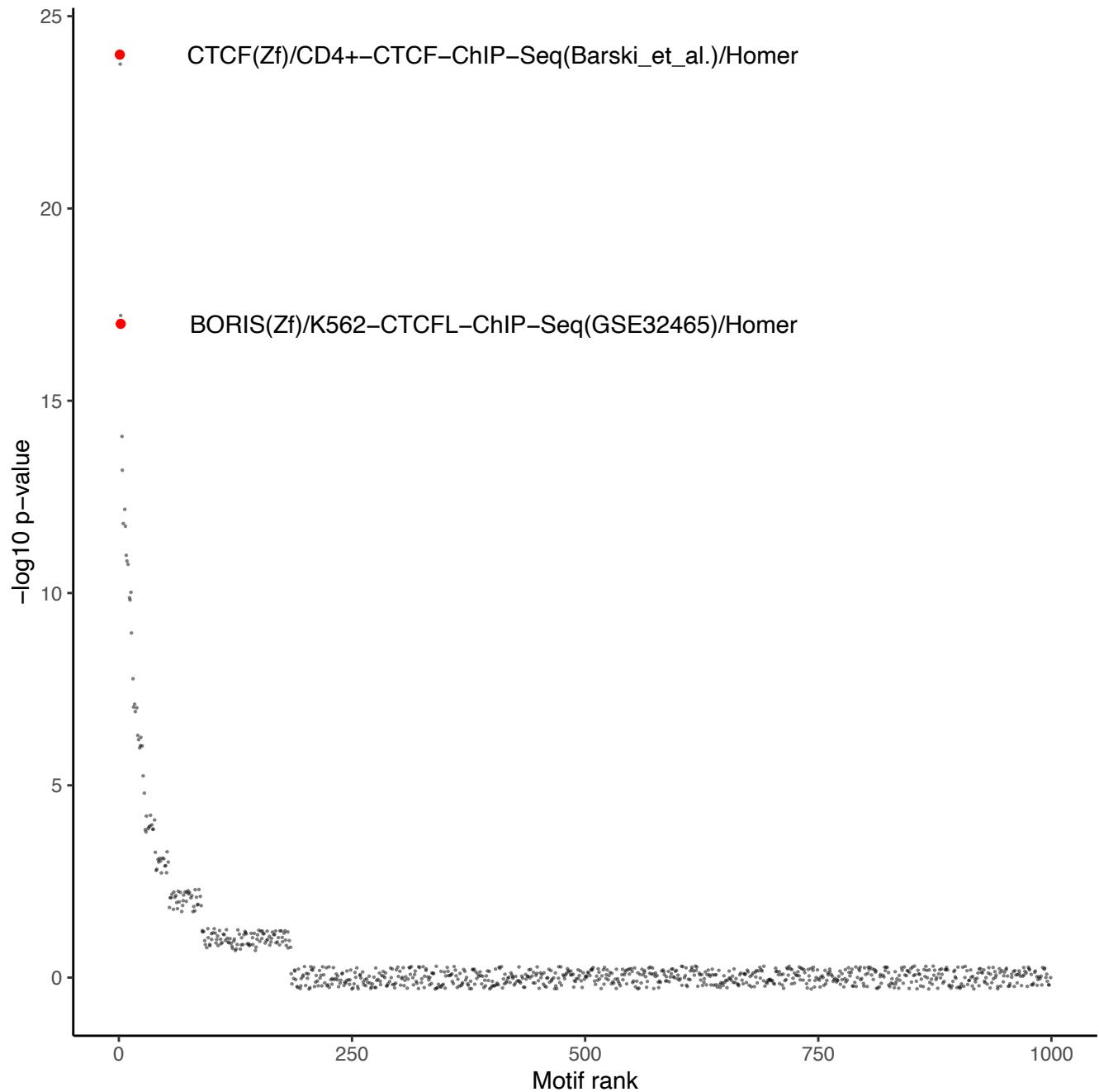
- Promoter (<=1kb)
- Promoter (1-2kb)
- Promoter (2-3kb)
- 5' UTR
- 3' UTR
- 1st Exon
- Other Exon
- 1st Intron
- Other Intron
- Downstream (<=300)
- Distal Intergenic

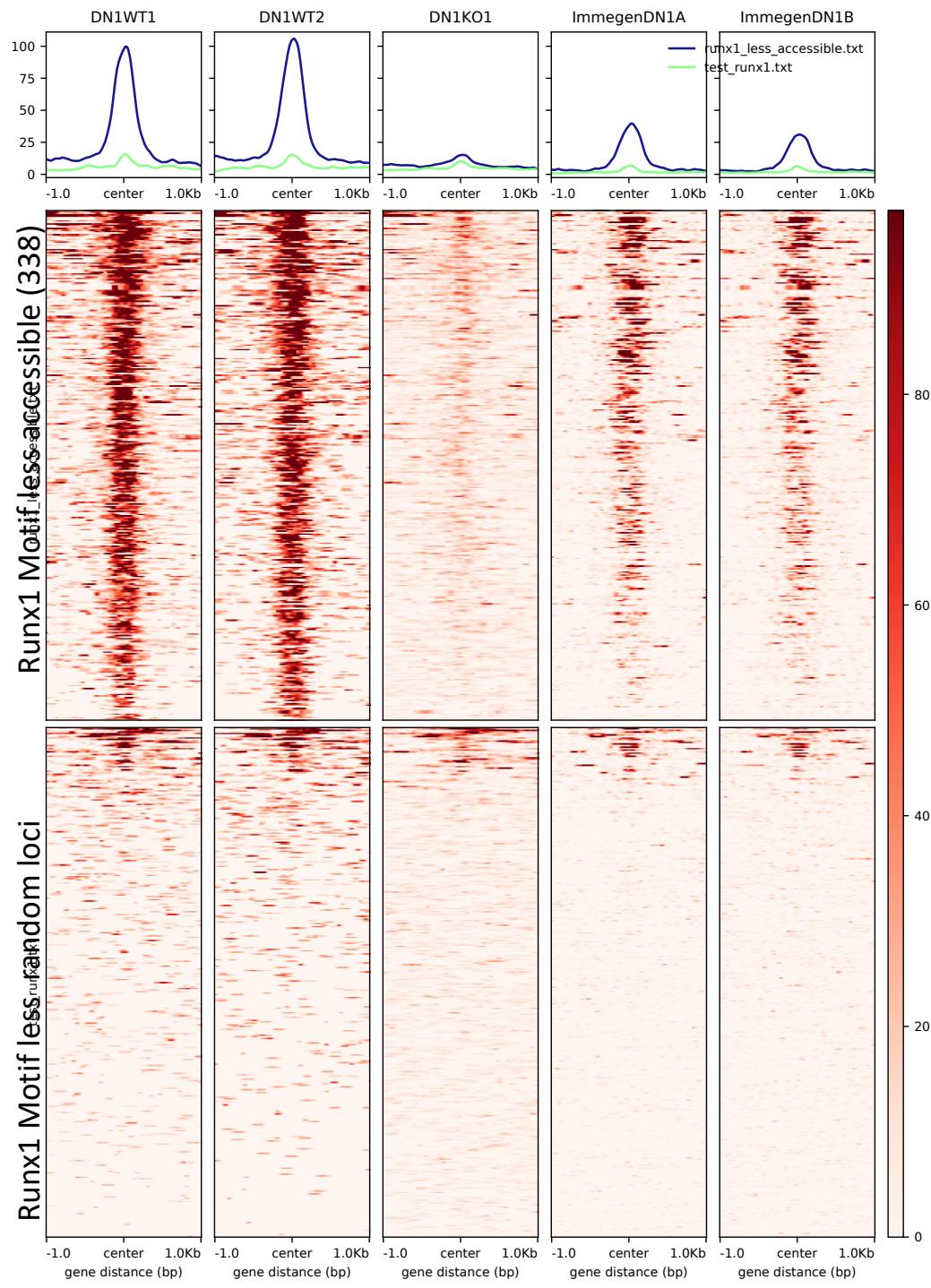
Percentage(%)

2046 LESS accessible loci in the mutant compared to the WT



269 MORE accessible loci in the mutant compared to the WT

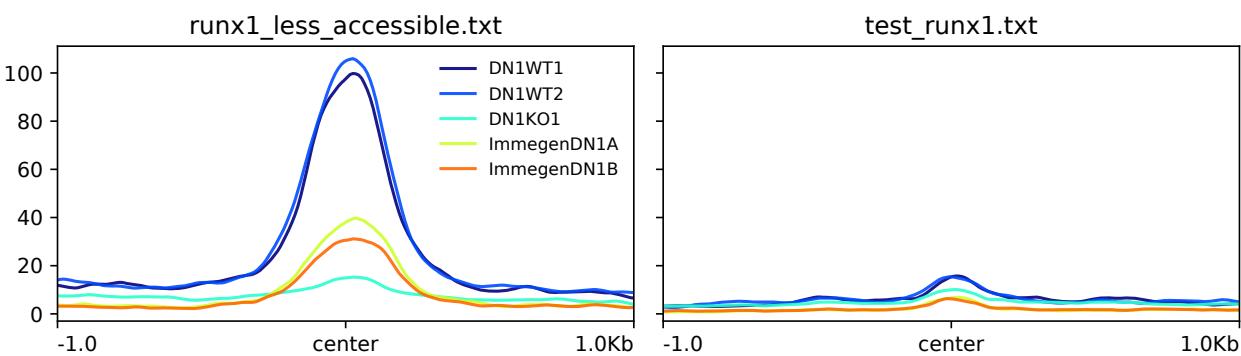




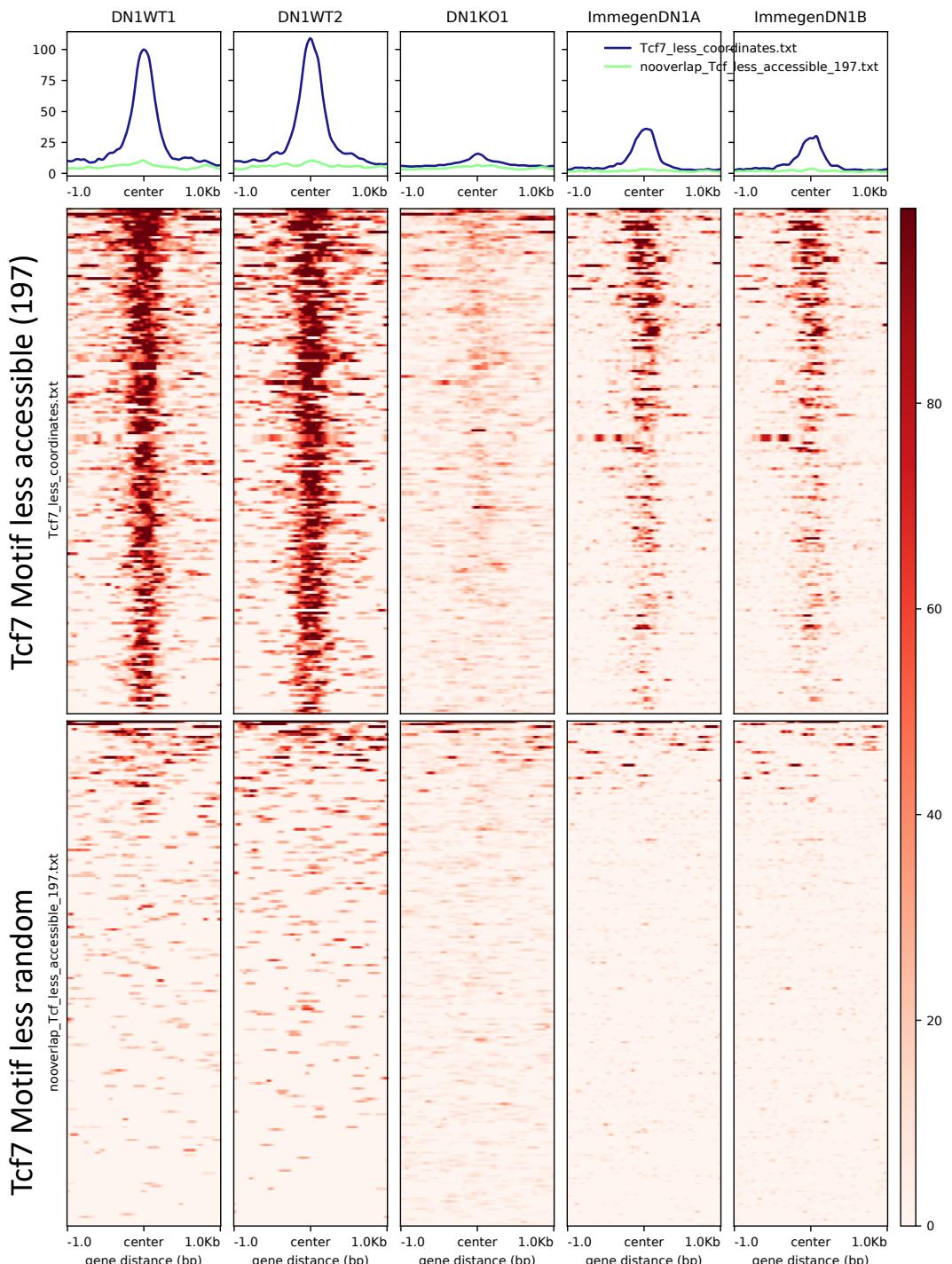
Runx1 motif: CWAACCACAR

Occurrence genome-wide :85689

Overlap with Less accessible peaks: 338 (17%)



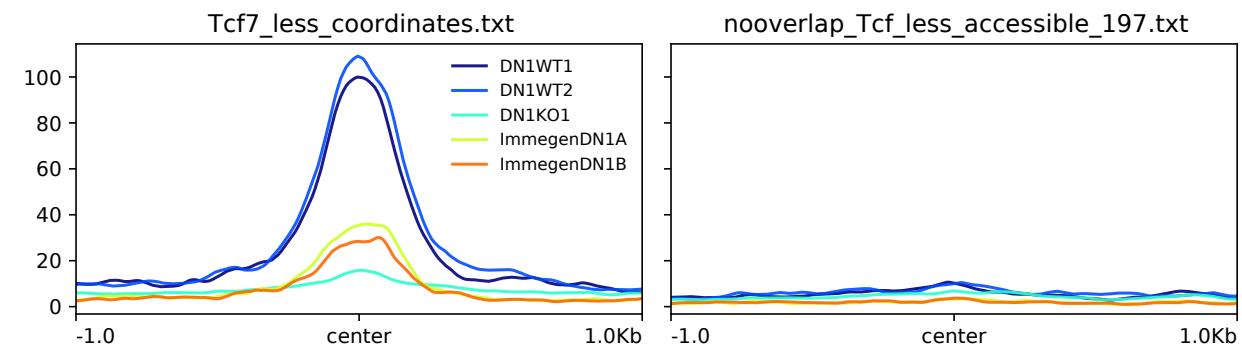
Runx1 motif are more accessible in WT than in mutant.
Also true from Immgen



Tcf7 motif: CCACATCAAAGG

Occurrence genome-wide : 70110

Overlap with Less accessible peaks: 197 (10%)



Summary

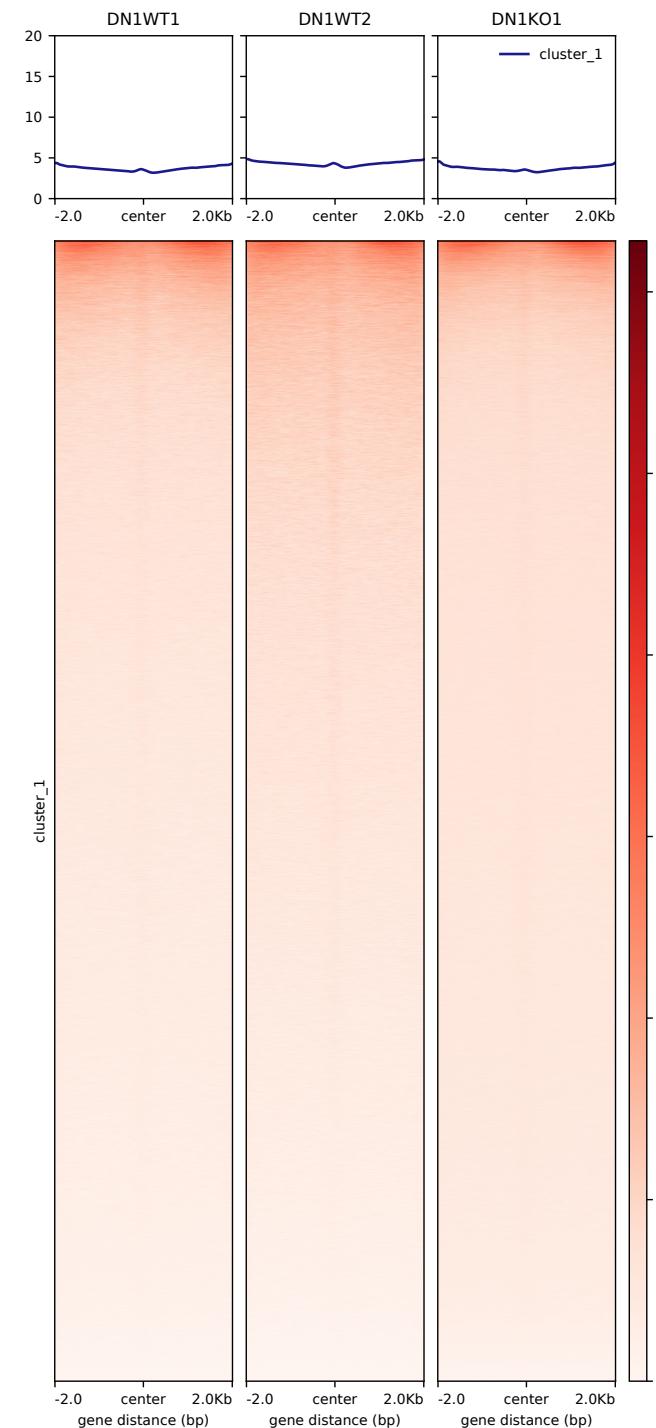
- Despite low coverage, we can identify accessible chromatin in OP-9 ATAC-seq
- Most peaks identified are located in promoter region
- Identification of Differentially accessible loci (DALs):
 - 2046 are less accessible in the mutant than in WT
 - 269 are more accessible in the mutant than in WT
- The motif analysis shows that potential binding sites for important TF regulators of T-cell commitment are impacted
- Are those DALs enhancers?
- To do: Differential peak analysis between WT OP9 and WT Immgen.

Analysis of coverage at TEs

Coverage at **LTR**

Coverage at LTR does not change between WT and KO

Both are lowly covered with ATAC-seq reads.

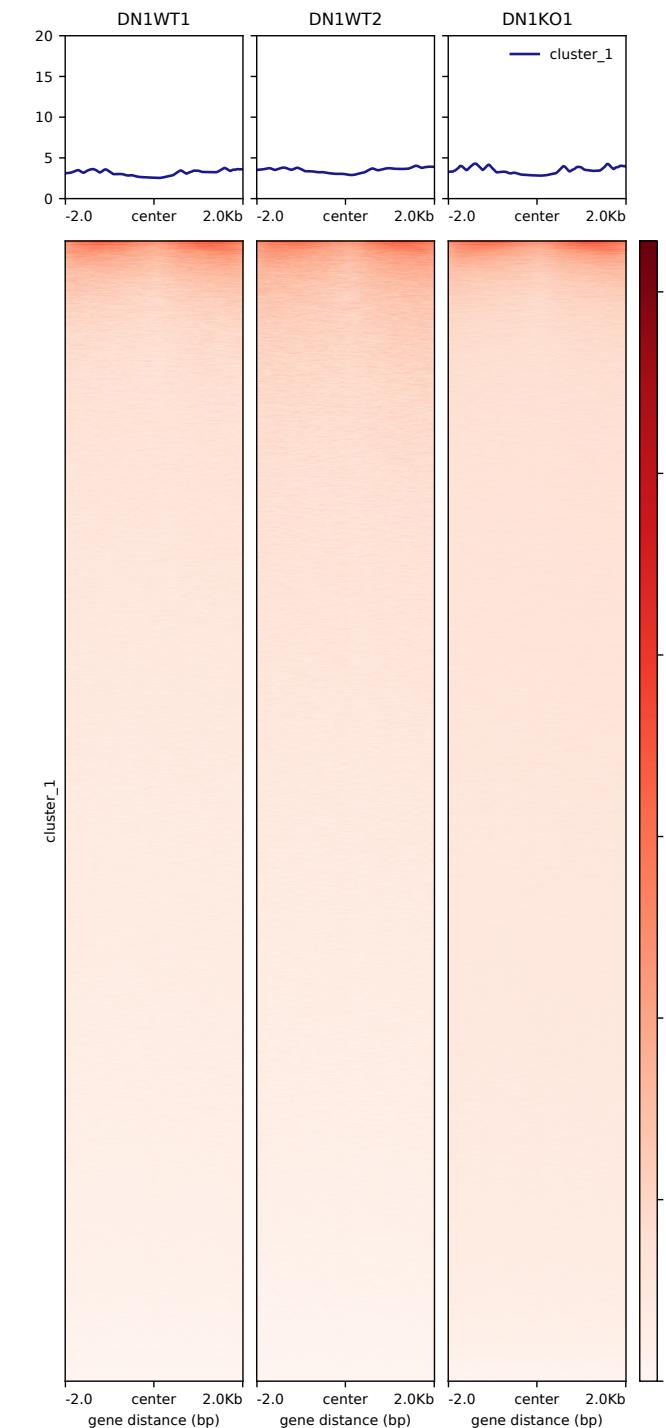


Analysis of coverage at TEs

Coverage at **LINE**

Coverage does not change between WT and KO

Both are lowly covered with ATAC-seq reads.

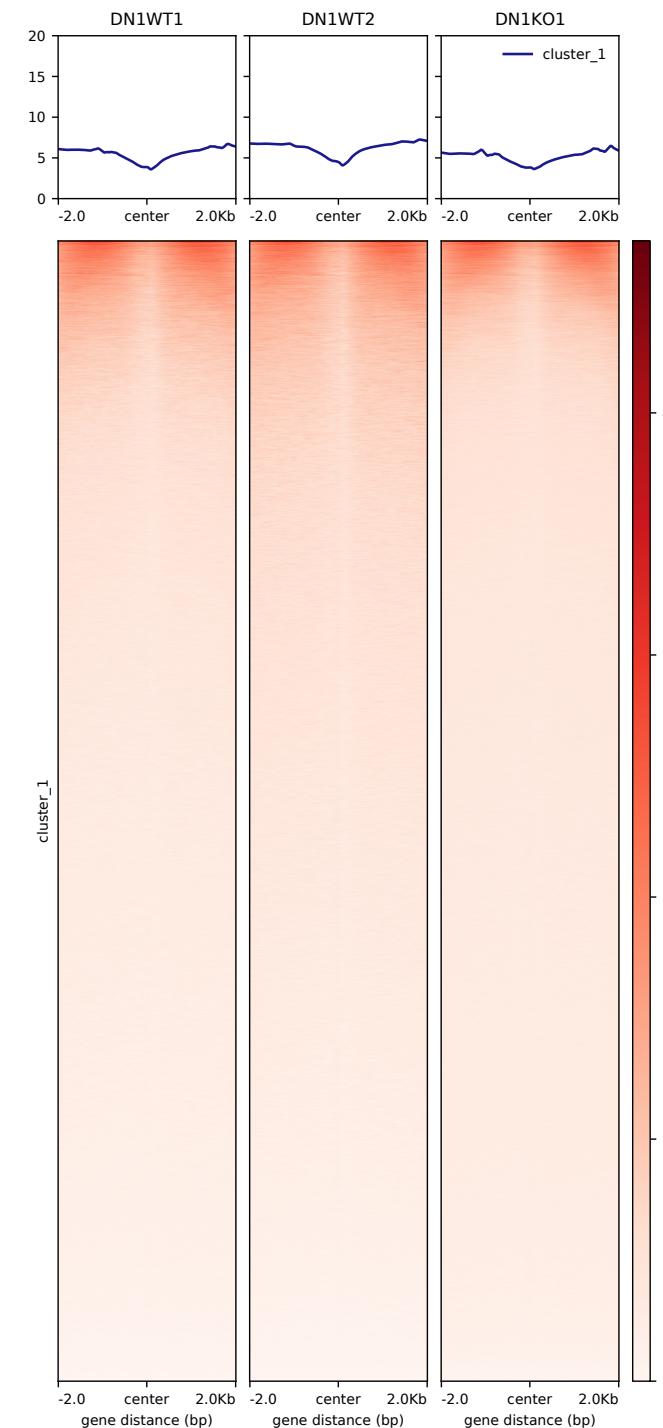


Analysis of coverage at TEs

Coverage at **SINE**

Coverage does not change between WT and KO

Both are lowly covered with ATAC-seq reads.

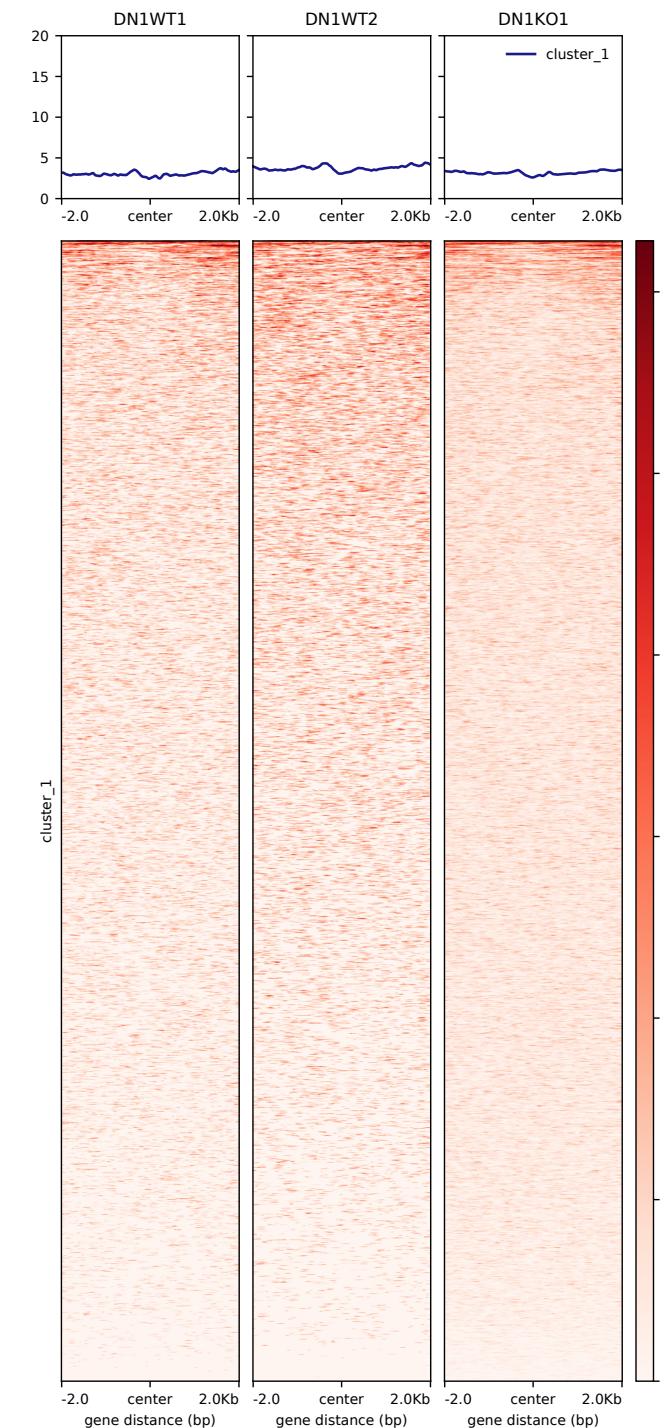


Analysis of coverage at TEs

Coverage at **MERVL**

Coverage does not change between WT and KO

Both are lowly covered with ATAC-seq reads.

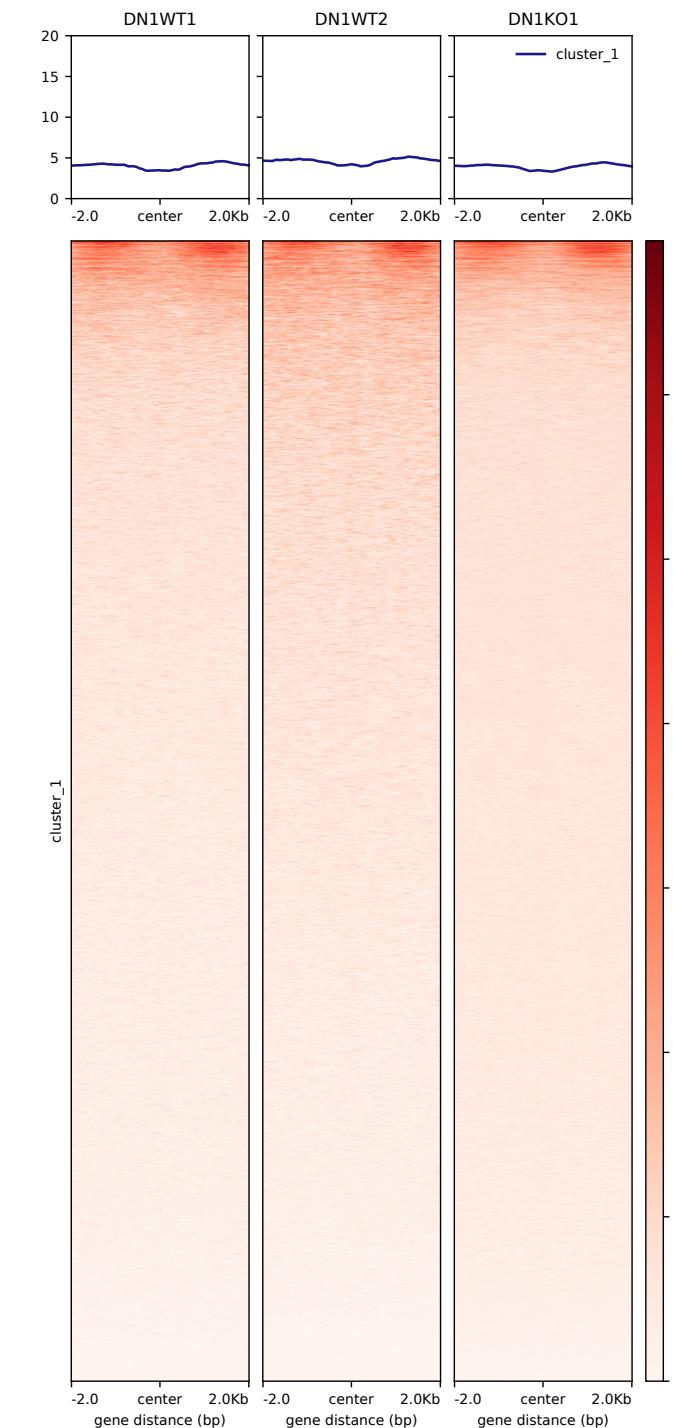


Analysis of coverage at TEs

Coverage at **ERV1**

Coverage does not change between WT and KO

Both are lowly covered with ATAC-seq reads.

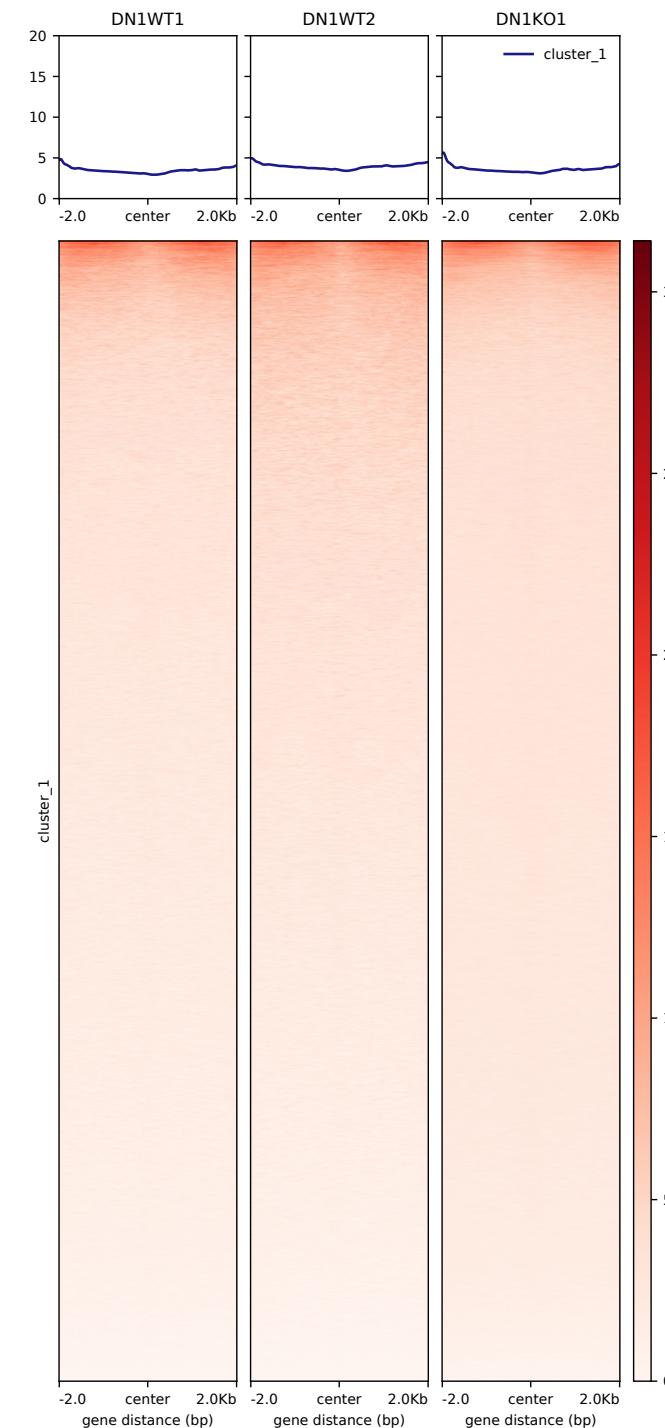


Analysis of coverage at TEs

Coverage at **ERVK**

Coverage does not change between WT and KO

Both are lowly covered with ATAC-seq reads.

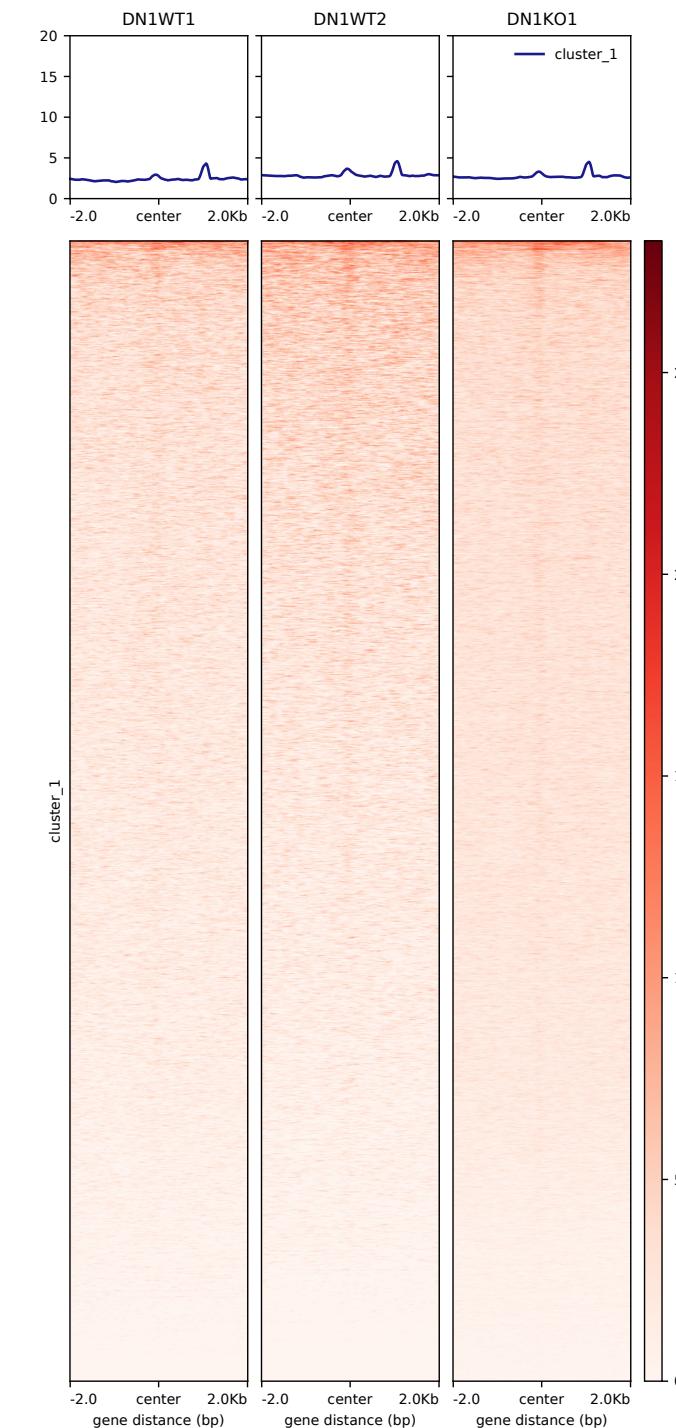


Analysis of coverage at TEs

Coverage at **IAP**

Coverage does not change between WT and KO

Both are lowly covered with ATAC-seq reads.



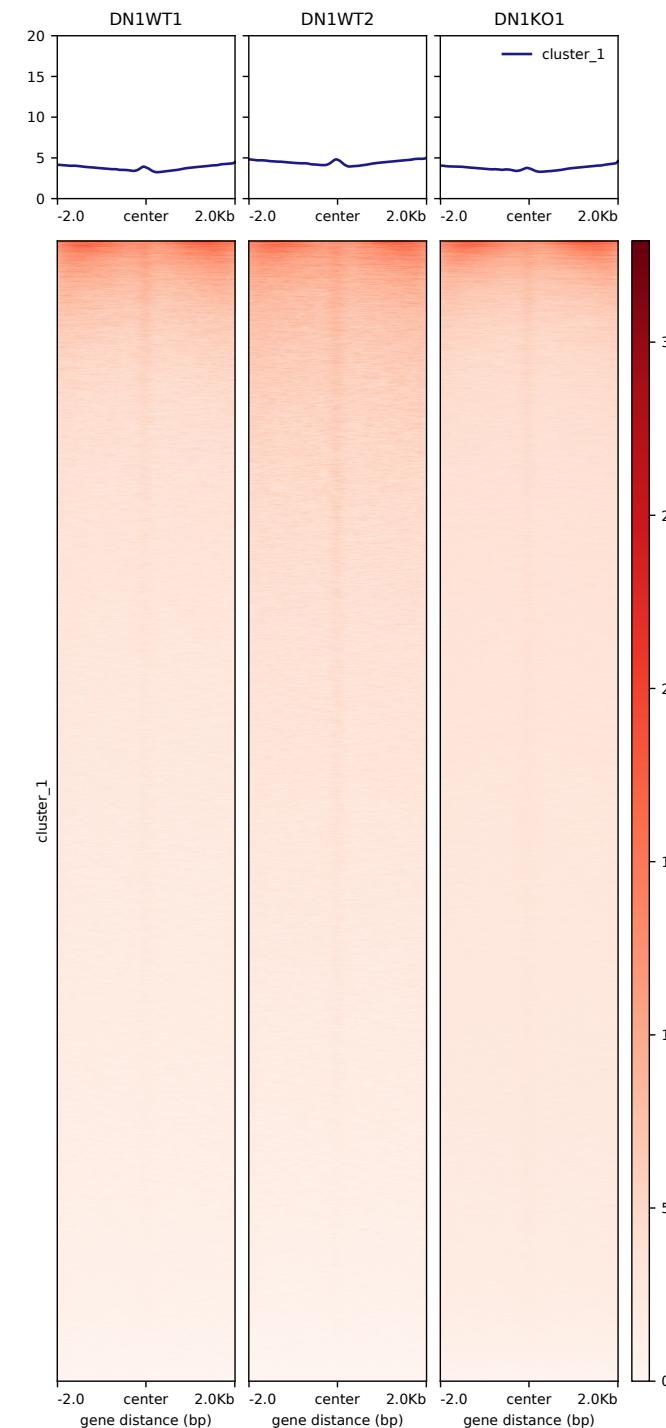
Analysis of coverage at TEs

Coverage at **ERVL**

Coverage does not change between WT and KO

Both are lowly covered with ATAC-seq reads.

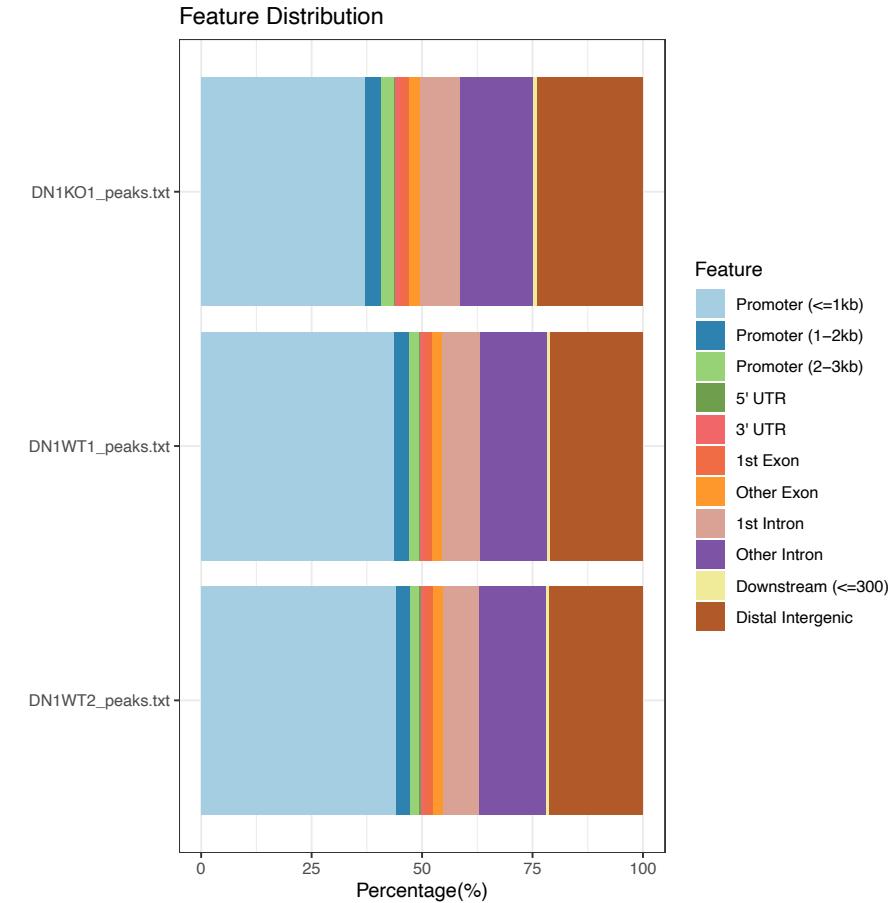
ERVL show a slight coverage of accessibility in the center of the element



Results DN1 ATAC-seq July 2021

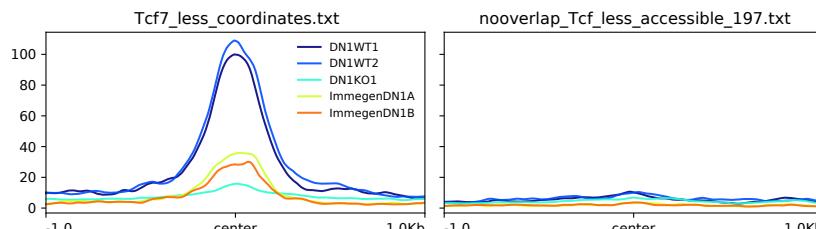
- Sequenced 2xWT and 1x mutant
- Despite low coverage, we can identify accessible chromatin in OP-9 ATAC-seq
- Most peaks identified are in promoter region

name	peaks
WT1	28992
WT2	24902
KO1	37473



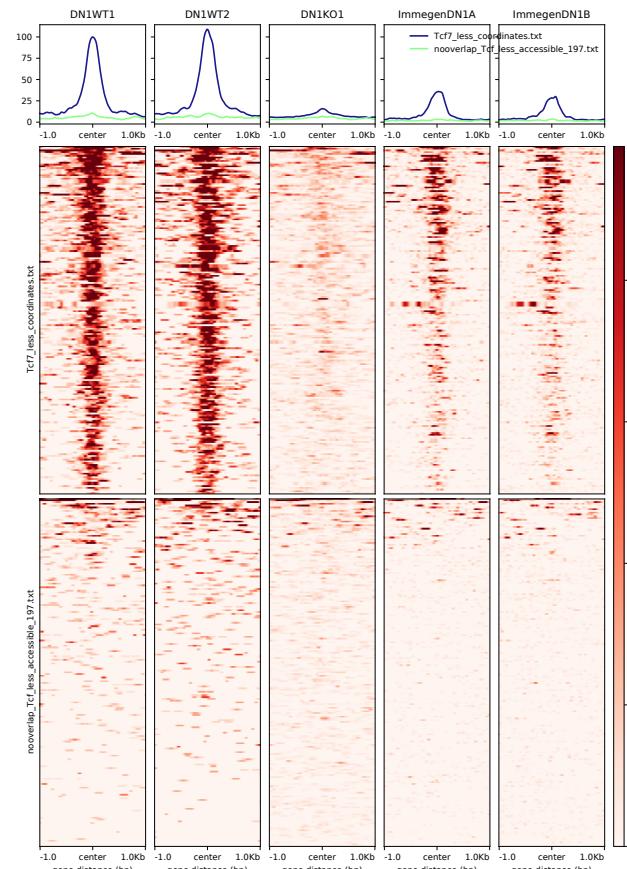
Results DN1 ATAC-seq July 2021

- Identification of Differentially accessible loci (DALs):
 - 2046 are less accessible in the mutant than in WT
 - 269 are more accessible in the mutant than in WT
- The motif analysis shows that potential binding sites for important TF regulators of T-cell commitment are impacted
- E.g Specific Tcf7 potential binding sites are less accessible in Morc3 mutant than in WT at DN1 (Similar for Runx1)



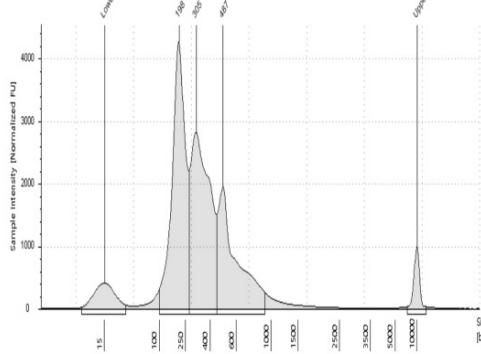
Occurrence genome-wide : 70110

Overlap with Less accessible peaks: 197 (10%)

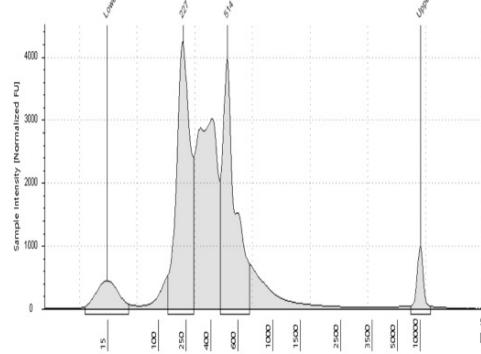


Overview DN1 ATAC 2nd experiment

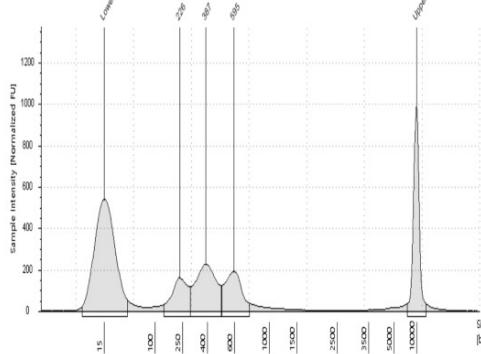
[Click to Enlarge =>1:Library : DN1_WT_Dec1](#)



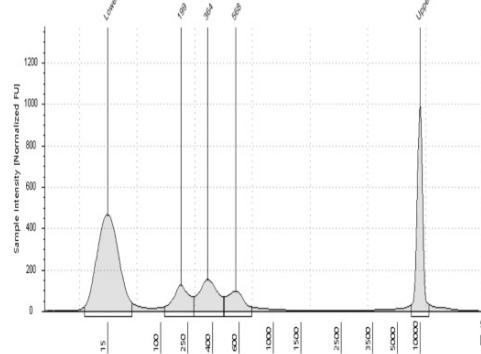
[Click to Enlarge =>2:Library : DN1_KO_Dec1](#)



[Click to Enlarge =>3:Library : DN1_KO_Jul1](#)



[Click to Enlarge =>4:Library : DN1_KO_Jul2](#)



SampleID	p5 Adapter	i5 Index	p7 Adapter	i7 Index	Sample Conc. (ng/ul)	Volume (uL)
DN1_WT_Dec1	S522	TTATGCGA	N714	TCATGAGC	5,168	12
DN1_KO_Dec1	S522	TTATGCGA	N715	CCTGAGAT	6,144	12
DN1_KO_Jul1	S505	GTAAGGAG	N705	AGGAGTCC	1,12	10
DN1_KO_Jul2	S505	GTAAGGAG	N707	GTAGAGAG	1,52	10

150bp
PE sequencing
Macrogen

Fastp: Filtered Reads

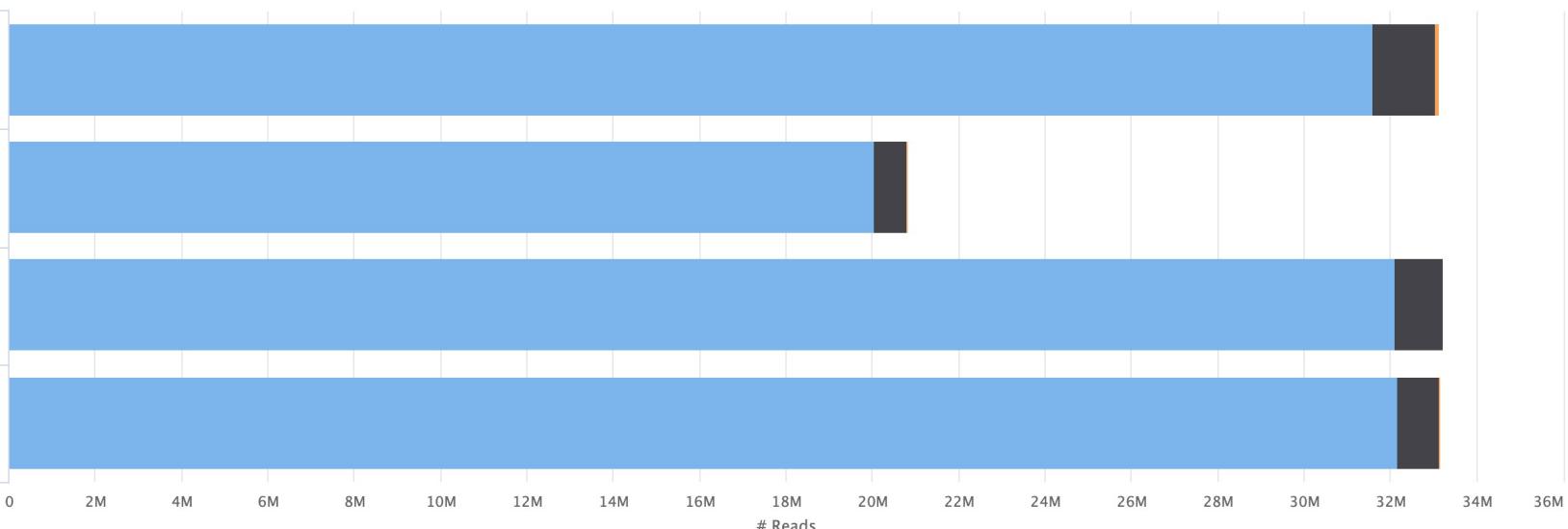
[Export Plot](#)

DN1KO1Jul_fastp

DN1KO2Jul_fastp

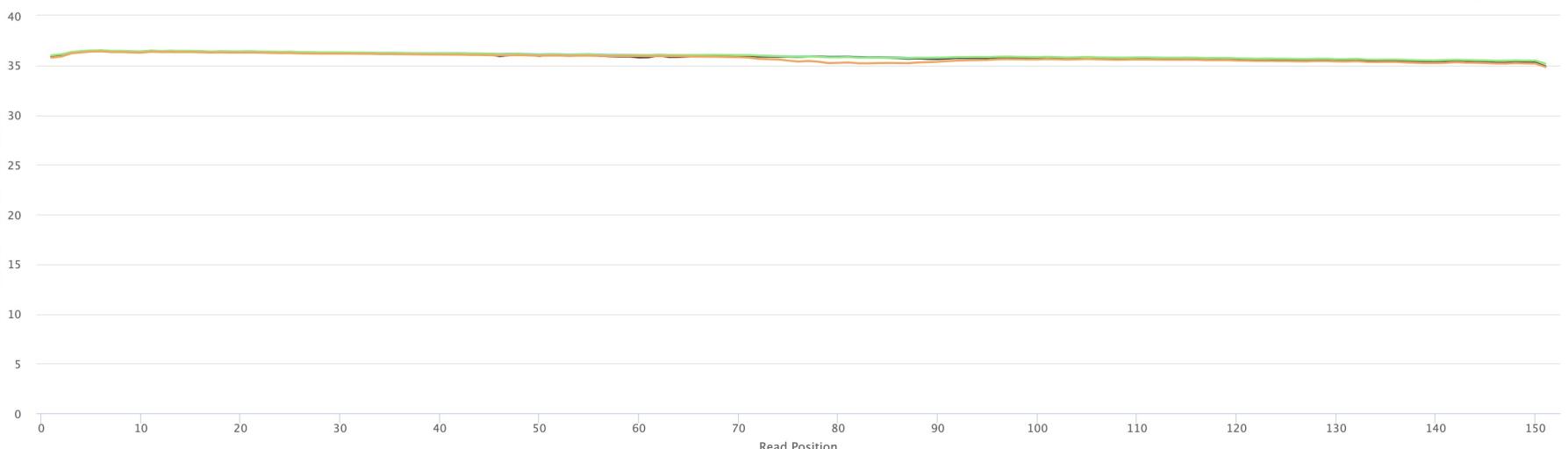
DN1KODec1_fastp

DN1WTDec1_fastp

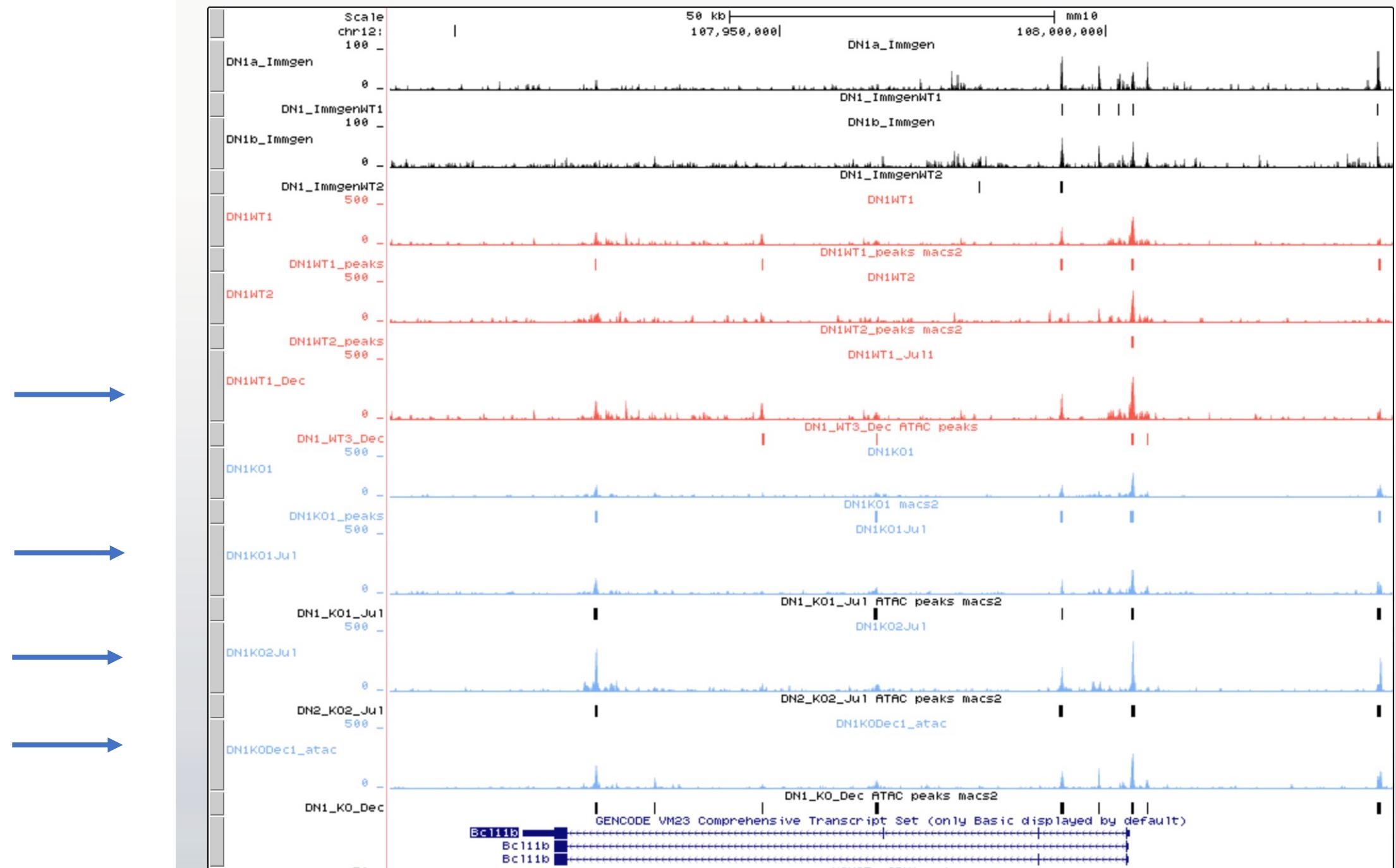


Created with MultiQC

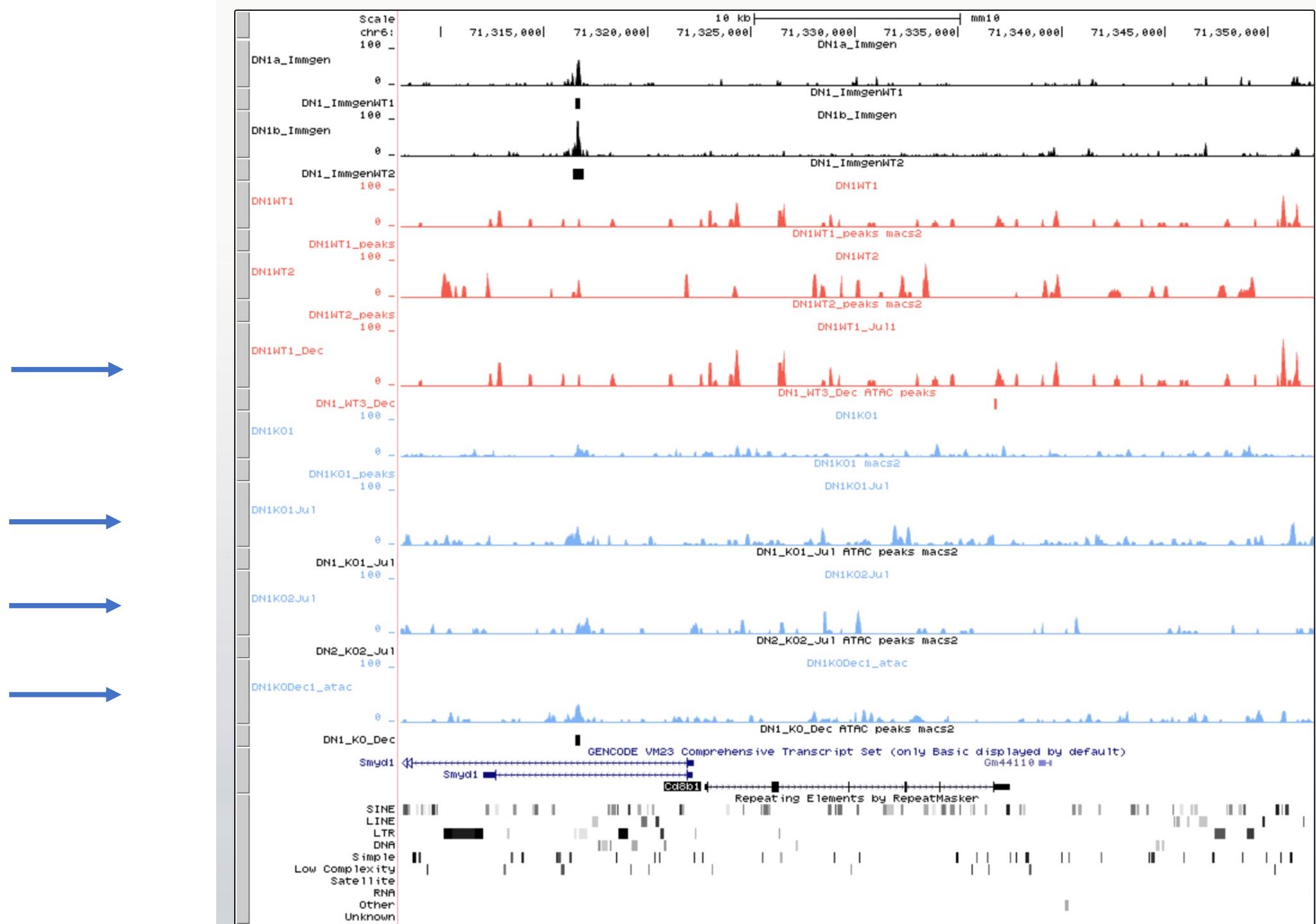
Fastp: Sequence Quality

[Export Plot](#)

Bcl11b



Cd8b1

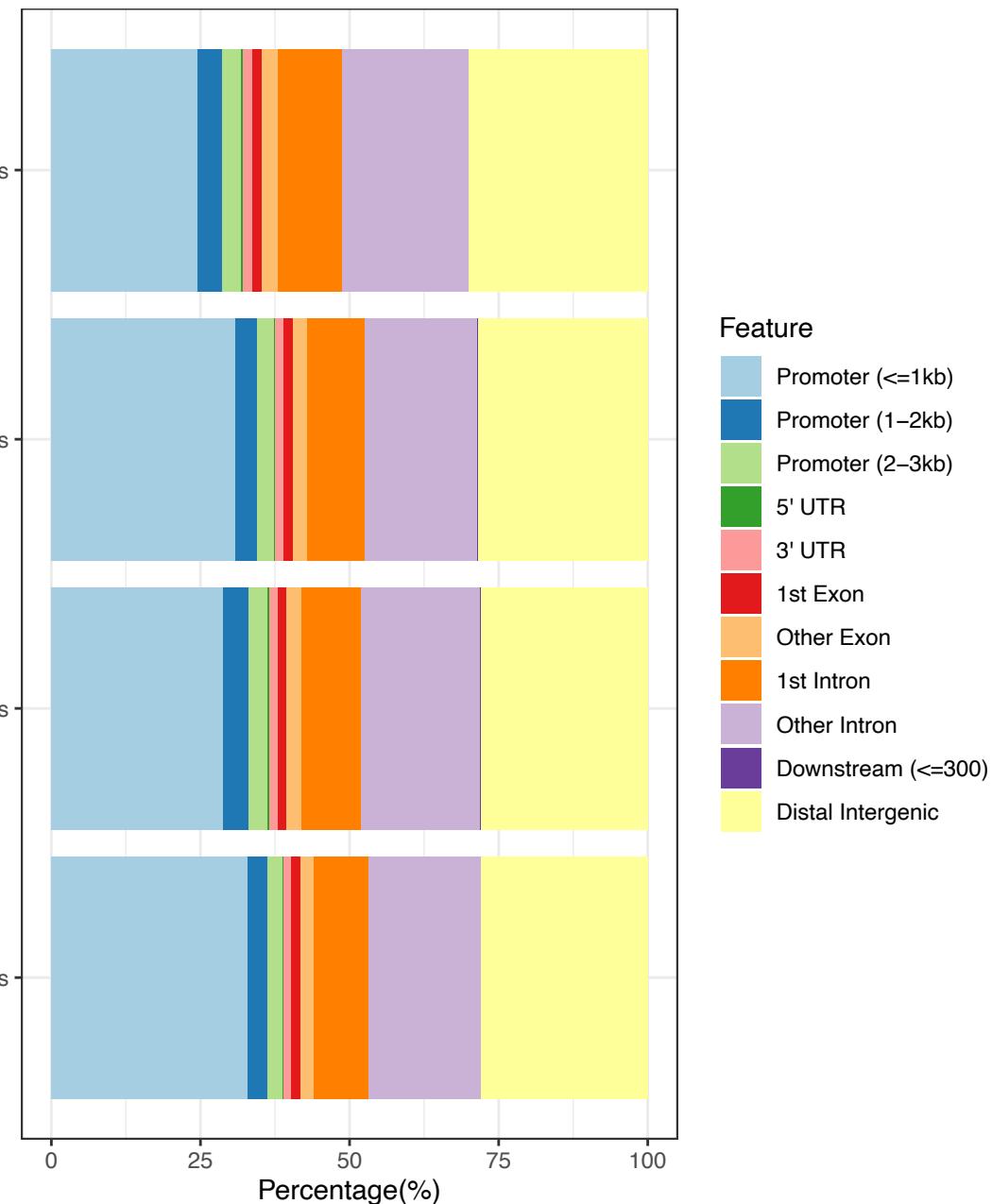


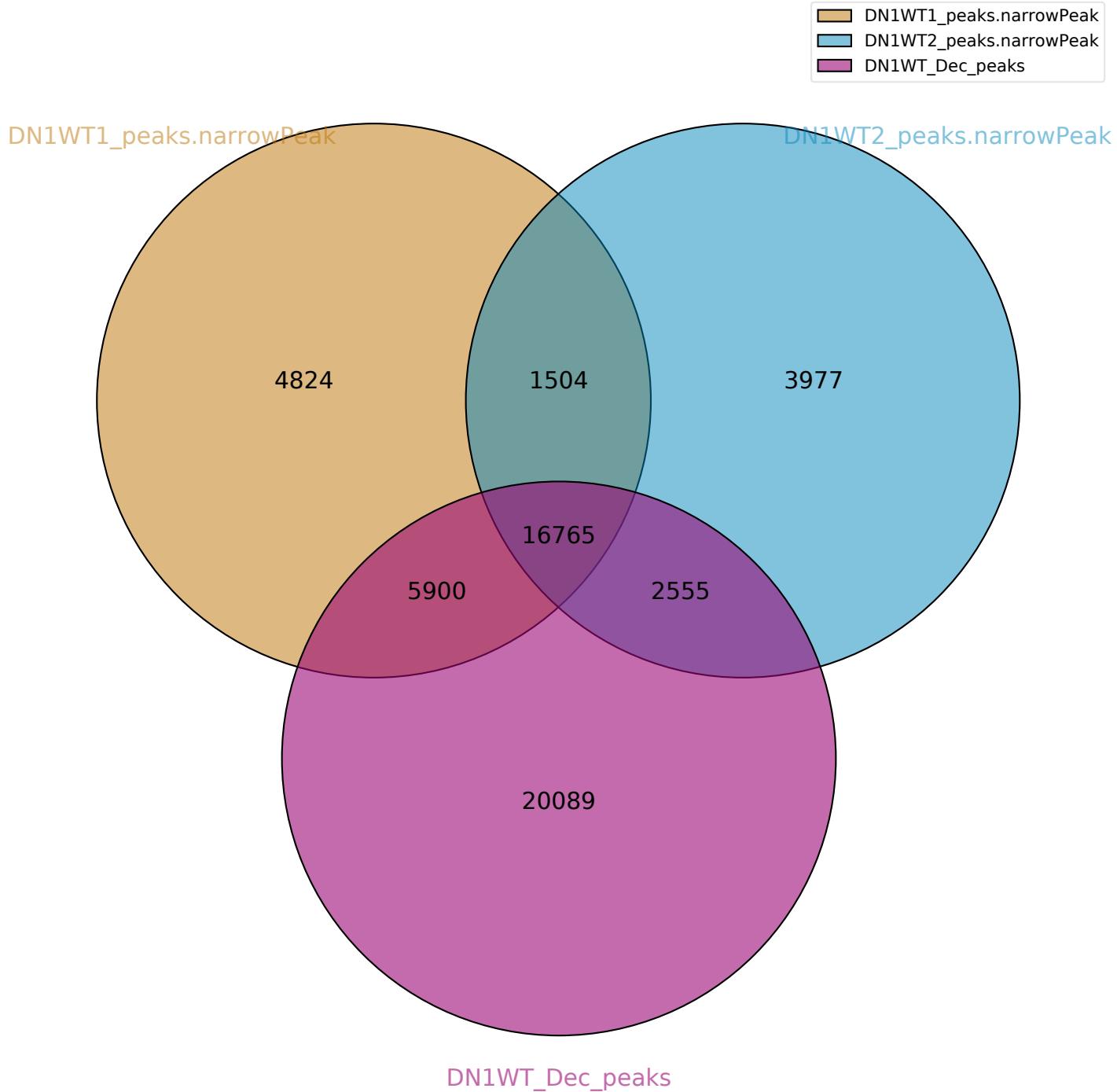
Annotation peaks DN1 ATAC Dec 2021

name	peaks	Experiment
WT1	28992	
WT2	24902	
KO1	37473	July 2021
WTDec	46140	
KODec	76930	
KOJul1	45832	
KOJul2	49211	Dec 2021

With a higher coverage we detect more peaks.
 Peaks are mostly at promoters or distal intergenic regions

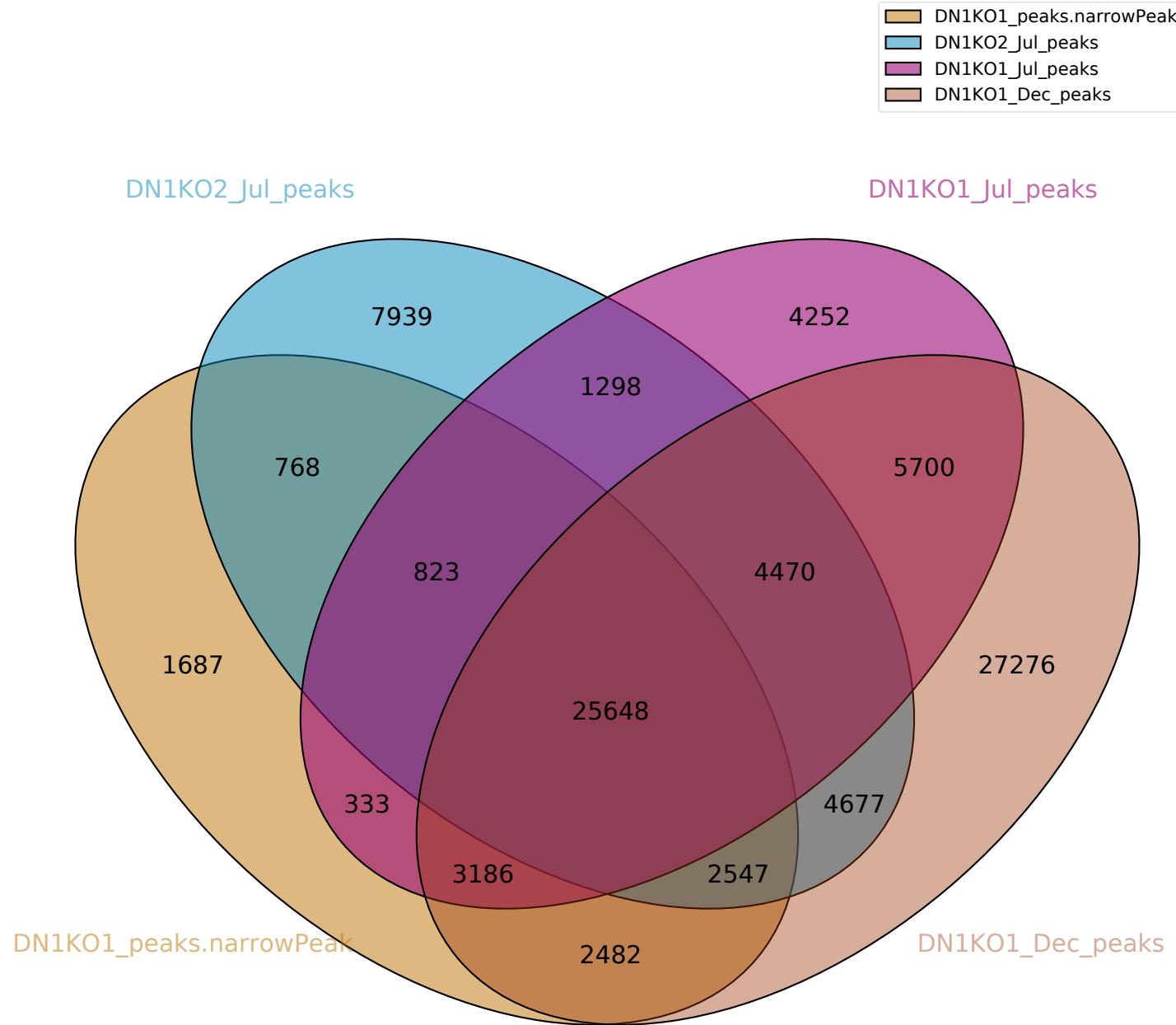
Feature Distribution





The ATAC-seq of december get many peaks that were not detected previously

I have the coordinates of those peaks and can annotate them.



The ATAC-seq of december get many peaks that were not detected previously

I have the coordinates of those peaks and can annotate them.

name	peaks	Experiment
WT1	28992	
WT2	24902	
KO1	37473	July 2021
WTDec	46140	
KODec	76930	
KOJul1	45832	
KOJul2	49211	Dec 2021