

Supplemental Information of

cLoops2: a full-stack comprehensive analytical tool for chromatin interactions

Yaqiang Cao ^{1*}, Shuai Liu ^{1*}, Gang Ren ^{1*}, Qingsong Tang¹, and Keji Zhao ^{1#}

¹ Laboratory of Epigenome Biology, Systems Biology Center, National Heart, Lung and Blood Institute, NIH, Bethesda, MD, 20892, USA.

*: These authors contributed equally to this work.

#: Corresponding Author, Dr. Keji Zhao, National Institutes of Health, Bethesda, MD, USA, E-Mail: zhaok@nhlbi.nih.gov

Public data re-analyzed

Used datasets were summarized as follows. GM12878 and K562 Hi-TrAC data were obtained from *Liu. et al.* NHGRI-EBI GWAS Catalog data (1) were downloaded from <https://www.ebi.ac.uk/gwas/api/search/downloads/full> on November 11, 2020.

Accession	Data Type	Cell Type	Factor	Reference
ENCLB620YVM, ENCLB779ZWL, (ENCODE)	ChIA-PET (FASTQ)	K562	RAD21	(2)
GSM2705045, GSM2705044, GSM2705043, (GEO)	HiChIP (FASTQ)	K562	H3K27ac	(3)
GSM765405 (GEO)	RNA-seq (FASTQ)	K562	None	(4)
ENCLB784HEF,	ChIA-PET	GM12878	RAD21	(2)

ENCLB535GER, (ENCODE)	(FASTQ)			
GSM2705042, GSM2705041, (GEO)	HiChIP (FASTQ)	GM12878	H3K27ac	(3)
GSM758559 (GEO)	RNA-seq (FASTQ)	GM12878	None	(4)
GSM1551550, GSM1551551, GSM1551552, GSM1551553, GSM1551554, GSM1551555, GSM1551556, GSM1551557, GSM1551558, GSM1551559, GSM1551560, GSM1551561, GSM1551562, GSM1551563, GSM1551564, GSM1551565, GSM1551566, GSM1551567, GSM1551568,	Hi-C (FASTQ)	GM12878	None	(5)

GSM1551569, GSM1551570, GSM1551571, GSM1551572, GSM1551573, GSM1551574, GSM1551575, GSM1551576, GSM1551577, GSM1551578, GSM1551579, GSM1551580, GSM1551581, GSM1551582, GSM1551592, GSM1551592, GSM1551593, GSM1551594, GSM1551595, GSM1551596, GSM1551597, GSM1551598, (GEO)				
GSM1872886 (GEO)	ChIA-PET (FASTQ)	GM12878	CTCF	(6)
GSM2326178	Trac-looping	Resting CD4+	None	(7)

GSM2326179 GSM2326180 GSM2782295 GSM2782296 (GEO)	(FASTQ)			
GSM2326181 GSM2326182 GSM2326183 (GEO)	Trac-looping (FASTQ)	Activated CD4+	None	(7)
GSM2326184 GSM2326185 (GEO)	RNA-seq	Resting CD4+	None	(7)
GSM2326186 GSM2326187 (GEO)	RNA-seq	Activated CD4+	None	(7)
ENCFF356LIU (ENCODE)	ChIP-seq (peaks)	GM12878	CTCF	(8)
ENCFF367KIF (ENCODE)	ChIP-seq (peaks)	GM12878	H3K27ac	(8)
ENCFF228GWY (ENCODE)	ChIP-seq (peaks)	GM12878	H3K4me3	(8)
ENCFF002DDJ (ENCODE)	ChIP-seq (peaks)	K562	CTCF	(8)
ENCFF044JNJ (ENCODE)	ChIP-seq (peaks)	K562	H3K27ac	(8)
ENCFF322IFF	ChIP-seq	K562	H3K27me3	(8)

(ENCODE)	(peaks)			
ENCFF624XRN (ENCODE)	ChIP-seq (peaks)	K562	H2A.Z	(8)
ENCFF183UQD (ENCODE)	ChIP-seq (peaks)	K562	H3K4me1	(8)
GSM3609747 (GEO)	CUT&RUN (aligned reads BED format)	K562	CTCF	(9)
GSM3609748 (GEO)	CUT&RUN (aligned reads BED format)	K562	CTCF	(9)
GSM3609749 (GEO)	CUT&RUN (aligned reads BED format)	K562	CTCF	(9)
GSM3609750 (GEO)	CUT&RUN (aligned reads BED format)	K562	CTCF	(9)
GSM3609755 (GEO)	CUT&RUN (aligned reads BED format)	K562	H3K27ac	(9)
GSM3609756 (GEO)	CUT&RUN (aligned reads BED format)	K562	H3K27ac	(9)
GSM3609757 (GEO)	CUT&RUN (aligned reads BED format)	K562	H3K27ac	(9)
GSM3609758 (GEO)	CUT&RUN (aligned reads BED format)	K562	H3K27ac	(9)
GSM3609759	CUT&RUN	K562	H3K27ac	(9)

(GEO)	(aligned reads BED format)			
GSM3770938 (GEO)	CUT&RUN (aligned reads BED format)	K562	H3K27me3	(9)
GSM3770939 (GEO)	CUT&RUN (aligned reads BED format)	K562	H3K27me3	(9)
GSM3770940 (GEO)	CUT&RUN (aligned reads BED format)	K562	H3K27me3	(9)
GSM3609773 (GEO)	CUT&RUN (aligned reads BED format)	K562	IgG	(9)
GSM3560260 (GEO)	CUT&Tag (aligned reads BED format)	K562	H2A.Z	(10)
GSM3536516 (GEO)	CUT&Tag (aligned reads BED format)	K562	H3K4me1	(10)
GSM3536514 (GEO)	CUT&Tag (aligned reads BED format)	K562	H3K27ac	(10)
GSM3560264 (GEO)	CUT&Tag (aligned reads BED format)	K562	IgG	(10)

Supplemental Methods

Pre-processing of CUT&RUN and CUT&TAG data

Pre-processed BED files mapped to hg19 were downloaded from GEO and converted to hg38 by liftOver for all following analyses. The same parameters were set for peak

calling by other tools as described in part on the comparison of ChIC-seq peak-callings. cLoops2 callPeaks with settings of -eps 100,200 -minPts 5 -sen for CTCF and H3K27ac, -eps 1000, 2000 -minPts 5 -sen for H3K27me3 were used to call peaks from CUT&RUN data. cLoops2 callPeaks with settings of -eps 100,200 -minPts 5 -sen for H2A.Z and H3K27ac, -eps 1000,2000 -minPts 5 -sen for H3K4me1, were used to call peaks from CUT&TAG data.

Processing of Hi-C, HiChIP, and ChIA-PET data

Pre-processing of Hi-C, H3K27ac HiChIP data, and RAD21 ChIA-PET data from raw FASTQ files to mapped BEDPE files was described as *Liu et al.*

Sub-samplings of 150 million PETs from H3K27ac HiChIP data were used for all downstream analyses. H3K27ac HiChIP loops were called by cLoops2 callLoops module with parameters settings of -eps 1000,2000,3000,4000,5000 -minPts 10,20,30,40,50 -w -j -cut 10000. The cLoops2 callDiffLoops module called differentially enriched loops with default parameters.

Sub-samplings of 17 million PETs from RAD21 ChIA-PET data were used for all downstream analyses. RAD21 ChIA-PET loops were called by cLoops2 callLoops module with parameters settings of -eps 500,1000,2000 -minPts 3,5. The cLoops2 callDiffLoops module called differentially enriched loops with default parameters.

Comparison of loops for GM12878 H3K27ac HiChIP data

GM12878 H3K27ac HiChIP data were used to compare the loops called by cLoops2, FitHiChIP (11), HiCCUPS (12), hichipper (13), mango (14), and MAPS (15). HiCCUPS called loops were obtained from reference (3). FitHiChIP called loops were obtained from reference (11). Loops called by hichipper, mango, and MAPS were obtained from the supplemental data reported by the MAPS paper at <https://doi.org/10.1371/journal.pcbi.1006982.s027>. Loops in hg19 coordinates were converted to hg38 for comparison by UCSC liftOver. All PETs were used for cLoops2 to call loops with the key parameters of -eps 1000,2000,3000,4000,5000 -minPts 10,20,30,40,50 -cut 10000.

HiCCUPS called the fewest loops (6,385) while achieving the highest enrichment score comparing loop nearby regions (**Supplemental Figure 12A**). Other tools called the

variable number of loops. From the aggregation analysis of H3K27ac HiChIP data (**Supplemental Figure 12A**) and in situ Hi-C data (**Supplemental Figure 12B**), no tool outperforms both than others. Loop distance distribution was similar with a median distance around 100kb (**Supplemental Figure 12C**), and loop anchor sizes of hichipper, FitHiChIP, and MAPS were fixed as 5kb, meanwhile results from cLoops2, mango, and HiCCUPS were dependent on selected parameters (**Supplemental Figure 12D**).

As global indicators did not provide a straightforward performance conclusion, we further checked some examples of the loops and anchors (merged) by randomly browsing chromosome 21 and selected two examples that may represent typical loops called from each tool (**Supplemental Figure 13**). In the first example (**Supplemental Figure 13A**), we marked the visible peaks as “a” to “l” for the H3K27ac HiChIP 1D data. Most of the loops detected by cLoops2 were supported by strong interaction signals in the region. Meanwhile, as specifically designed for Hi-C data, HiCCUPS miss all the loops. As peaks d and e were not seen in ENCODE H3K27ac ChIP-seq data, therefore methods depend on ChIP-seq first to get the peaks as anchors, such as hichipper and mango, may miss the loops formed by c and e, and d and e. Due to fixed anchor size, or inaccurate calling of peaks from the HiChIP data, FitHiChIP, hichipper, and MAPS, called multiple loops between peak c and peak f, and led to an extensive number of total loops as shown in **Supplemental Figure 12A**. Mango seems to call loops for every two peaks within the region, and therefore the highest number of loops were detected by mango (493,675). Also, due to the specific design of the peak-calling process of FitHiChIP and MAPS, they identified a loop with one anchor located between peak e and peak f, which is not visible either from the HiChIP 1D signal or ChIP-seq data and thus may not be a reliable loop.

The second example (**Supplemental Figure 13B**) should be an ideal genomic region to show enhancer-promoter and promoter-promoter loops detected from H3K27ac ChIP-seq data. However, comparison of the results from loop calling by various methods indicates that the loop prediction is much more complex than expected. We marked the visible peaks from the ChIP-seq as “a” to “g”, by which peak a was not visible, and f and g are weak peaks in H3K27ac HiChIP 1D. We expected loops with strong interactions of b-c, c-d, c-e, c-g, d-e, d-g from the interaction heatmap. cLoops2 detected the expected loops, also called the loop from the peak a to peak b and one

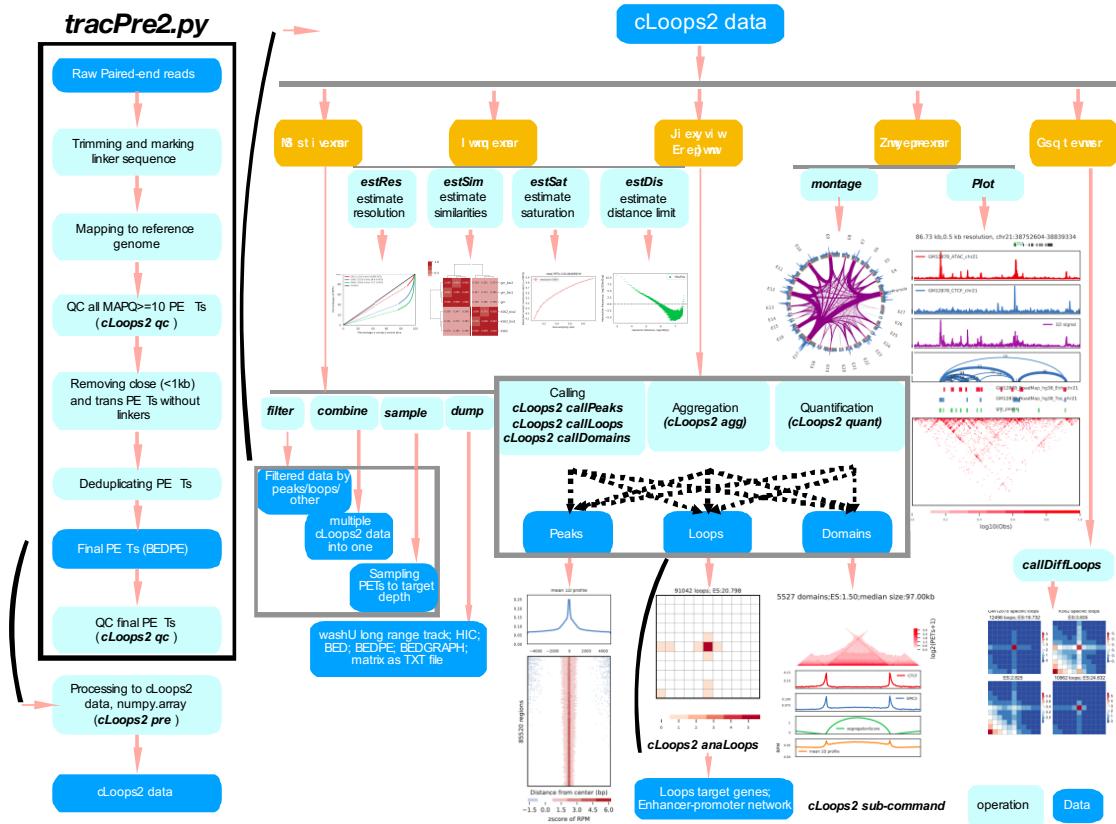
between b and c. FitHiChIP, however, seems to have a bias for anchors such as peak b and peak c, called many loops from these two strong anchors but missed the one between peak b and peak c. Again, as HiCCUPS was not designed for HiChIP data, it missed all loops in this region. Hichipper missed long-distance loops in this region and only called multiple loops between peak d and peak e as splitting them into multiple peaks. Mango, called loops for every two peaks as it called, same to its performance in the first example. MAPS' performance is similar to FitHiChIP, called multiple loops from strong anchors of peak b and peak c, but missed the visible loops between peak b and peak c.

Based on these comparisons, we think no tool can be thought as the best for calling loops from H3K27ac HiChIP data. It is still a challenging computational task to model the HiChIP data and call accurate loops.

Calling domains from Hi-TrAC data

Domains were called from Hi-TrAC data based on a variant of insulation score (16), termed segregation score. A sliding window with a fixed size (as a parameter) around each bin was used to obtain the submatrix for calculating the segregation score for each bin. The contact matrix is further \log_2 transformed and calculated as a correlation matrix. For the up-right corner sub-matrix of the correlation matrix, all < 0 values are assigned as 0, and the mean value of the matrix is assigned as a segregation score for the bin. After calculating segregation scores for all bins in one chromosome, z-score normalization was performed to the segregation scores. Finally, bins with > 0 segregation scores were stitched together as candidate domains. This algorithm is implemented in cLoops2 callDomains module.

Supplemental Figure 1

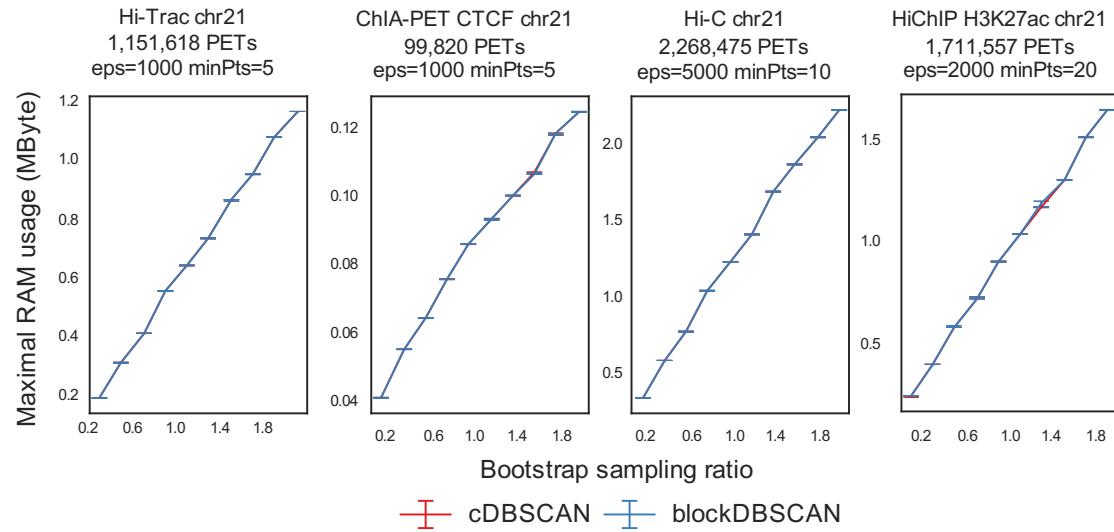


Supplemental Figure 1. Overview of cLoops2 main modules architecture

The main functions of cLoops2 were integrated into the main command `cLoops2`; meanwhile, extra analysis scripts were located in the script directory, such as Hi-TrAC data pre-processing analysis script `tracPre2.py`. In `tracPre2.py`, FASTQ files were first trimmed of linker sequence CTGTCTCTTATACACATCT for both ends. Only PETs with both ends length $\geq 10\text{bp}$ were kept. Trimmed PETs were mapped to the target reference genome with Bowtie2 (17). Mapped PETs with MAPQ ≥ 10 were converted to BEDPE files. PETs $< 1\text{kb}$ without linker sequence in any end were further filtered. PCR replicates of PETs were filtered if the locations of both ends were identical. In `cLoops2` main functions, modules were classified into five categories: IO operation, estimation, features analysis, visualization, and comparison. In IO operation category, `cLoops2 pre` module processes input BEDPE files into `cLoops2` data; `cLoops2 filterPETs` module filters PETs by overlapping with peaks/loops or applies blockDBSCAN noise remove process; `cLoop2 combine` module combines multiple

cLoops2 data directories; cLoops2 sample module does sampling PETs to target sequencing depth and cLoops2 dump module converts cLoops2 data into other popular data formats such as BED, BEDPE, and HIC. In the estimation category, cLoops2 estRes module estimates interaction resolutions; cLoops2 estSim module estimate different datasets interaction similarities; cLoops2 estSat module estimates sequencing saturation for sequencing depth and cLoops2 estDis module estimates significant interaction distance limitations. The feature analysis category, features calling results from cLoops2 callPeaks module, callLoops module, and callDomains module can be further processed with cLoops2 agg module for aggregation analysis to draw global conclusions and cLoops2 quant module to obtain digital quantification for each feature to perform detailed analysis. An extra module called anaLoops was used to annotate loops target genes and obtain the potential enhancer-promoter interaction networks. In the visualization category, the cLoops2 montage module generates a Rehoboam plot for showing interactions for selected interest regions, and the cLoops2 plot module generates interactions heatmaps/arches/scatter plot together with genes genomic regions, loops, and other 1D profile data. In the comparison category, cLoops2 callDiffLoops performs differentially enriched loops analysis to find sample-specific loops, and the loops can be annotated using cLoops2 anaLoops to find potential specific target genes.

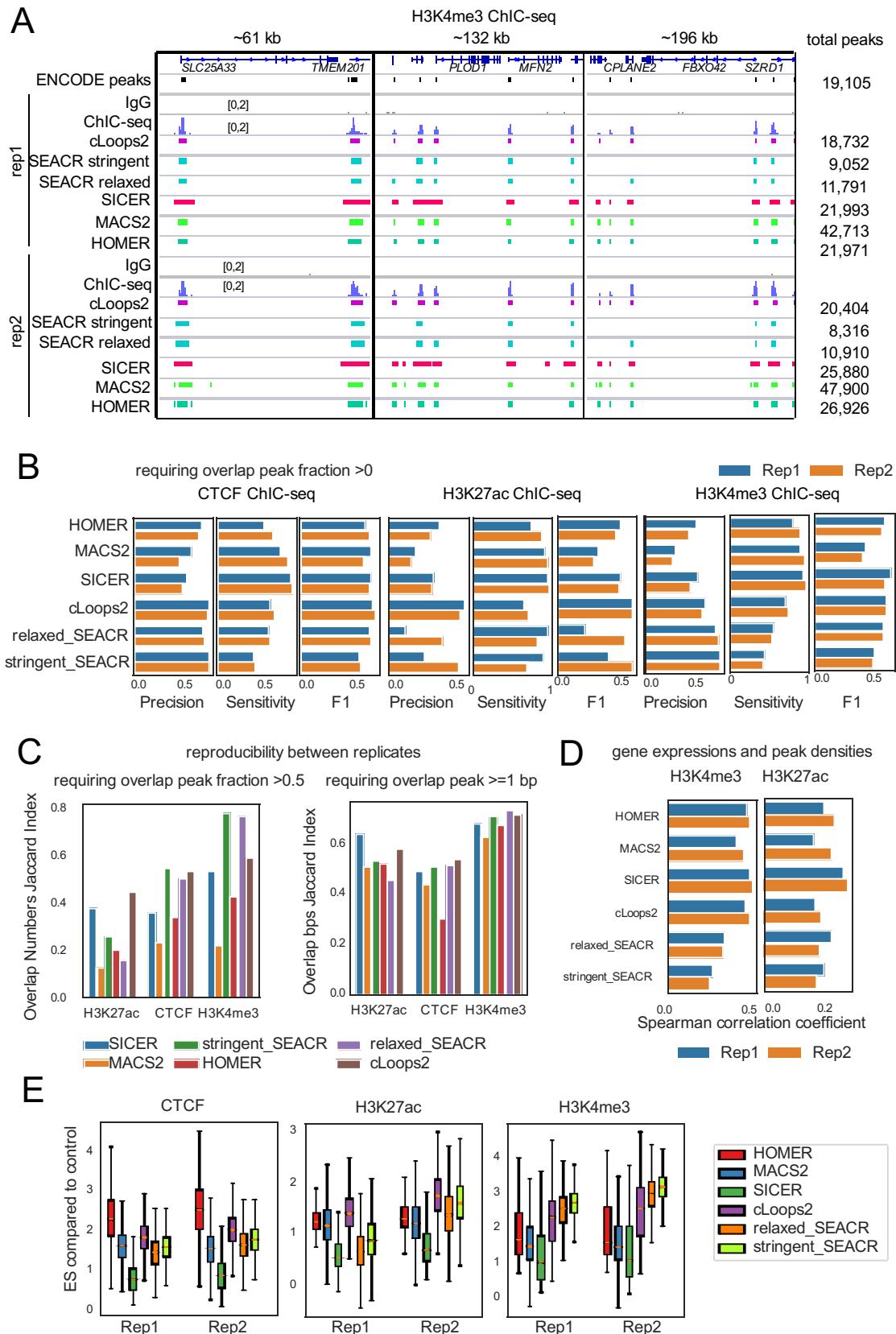
Supplemental Figure 2



Supplemental Figure 2. Comparison of memory usage between cDBSCAN and blockDBSCAN

Comparison of memory usage using actual interaction data of Hi-TrAC, CTCF ChIA-PET, Hi-C, and H3K27ac HiChIP data from GM12878 cells, based on five repeats sampling. The results from cDBSCAN and blockDBSCAN were highly similar; therefore, most parts of the curves were overlapped.

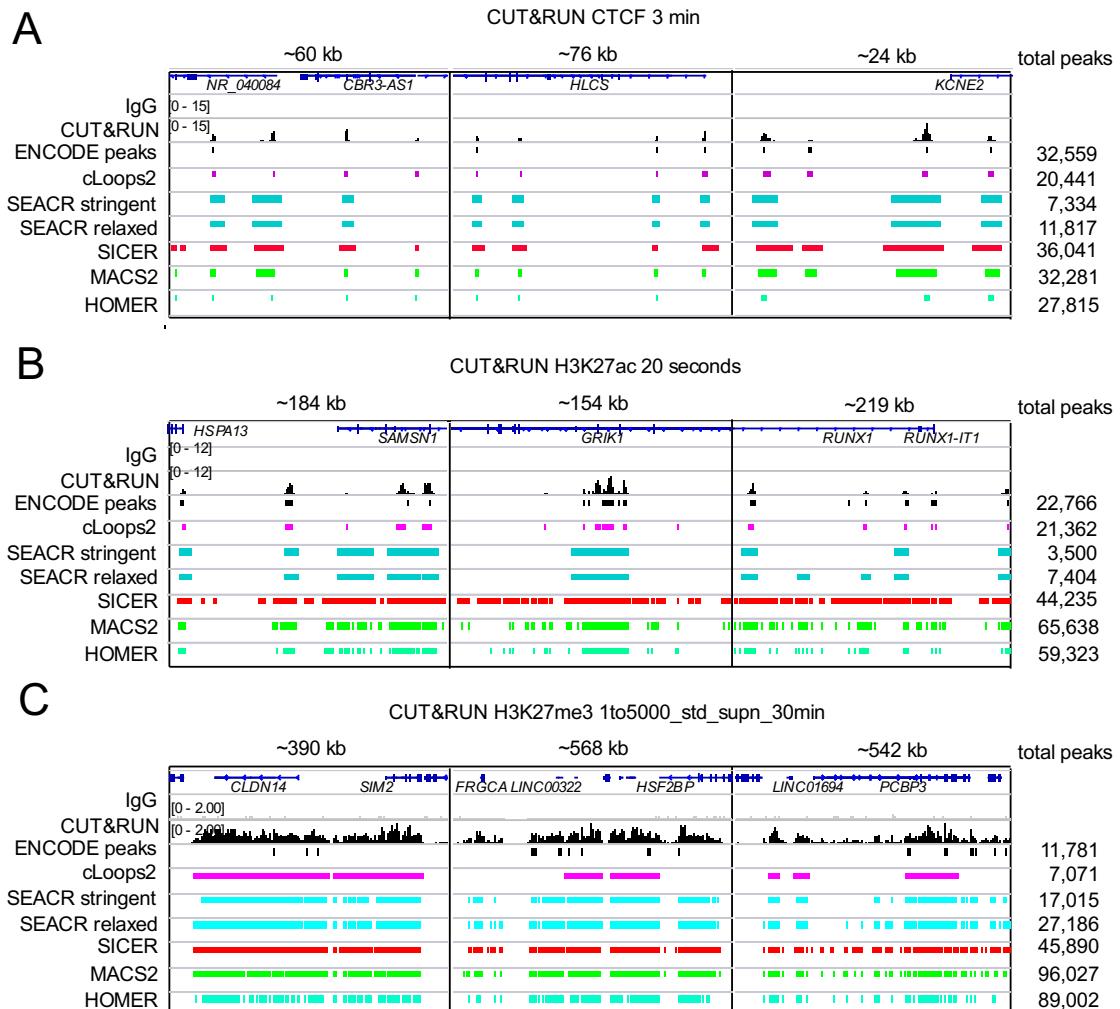
Supplemental Figure 3



Supplemental Figure 3. Comparison of ChIC-seq peak calling by various tools

- (A) Examples of H3K4me3 ChIC-seq profiles and peaks called by different tools.
- (B) Precision, sensitivity, and F1 scores for peaks called by different tools comparing with ENCODE peaks, requiring peaks overlap ≥ 1 bp.
- (C) Reproducibility between ChIC-seq replicates. The left panel requiring peaks overlap fraction ≥ 0.5 , and the right panel requiring peaks overlap ≥ 1 bp.
- (D) Correlation between gene expression and ChIC-seq peak densities.
- (E) ChIC-seq signal enrichment score (ES) of called peak regions comparing to IgG control.

Supplemental Figure 4



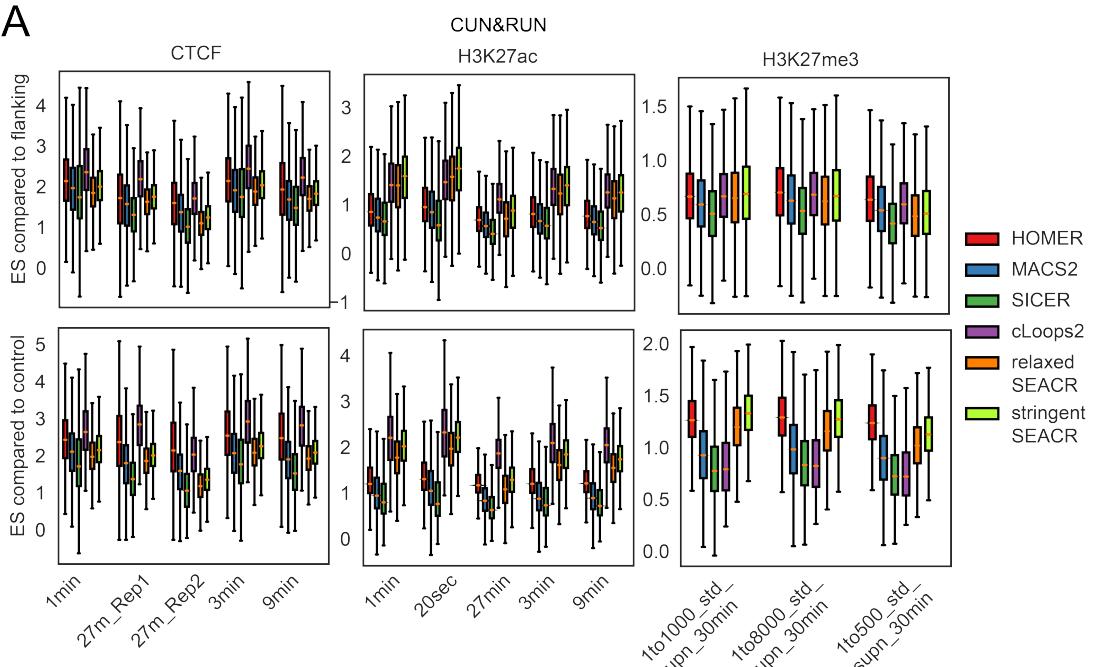
Supplemental Figure 4. Examples of peaks called by cLoops2 and other tools from CUT&RUN data in K562 cells

- (A) Examples of CTCF CUT&RUN binding profiles and peaks called by different tools.

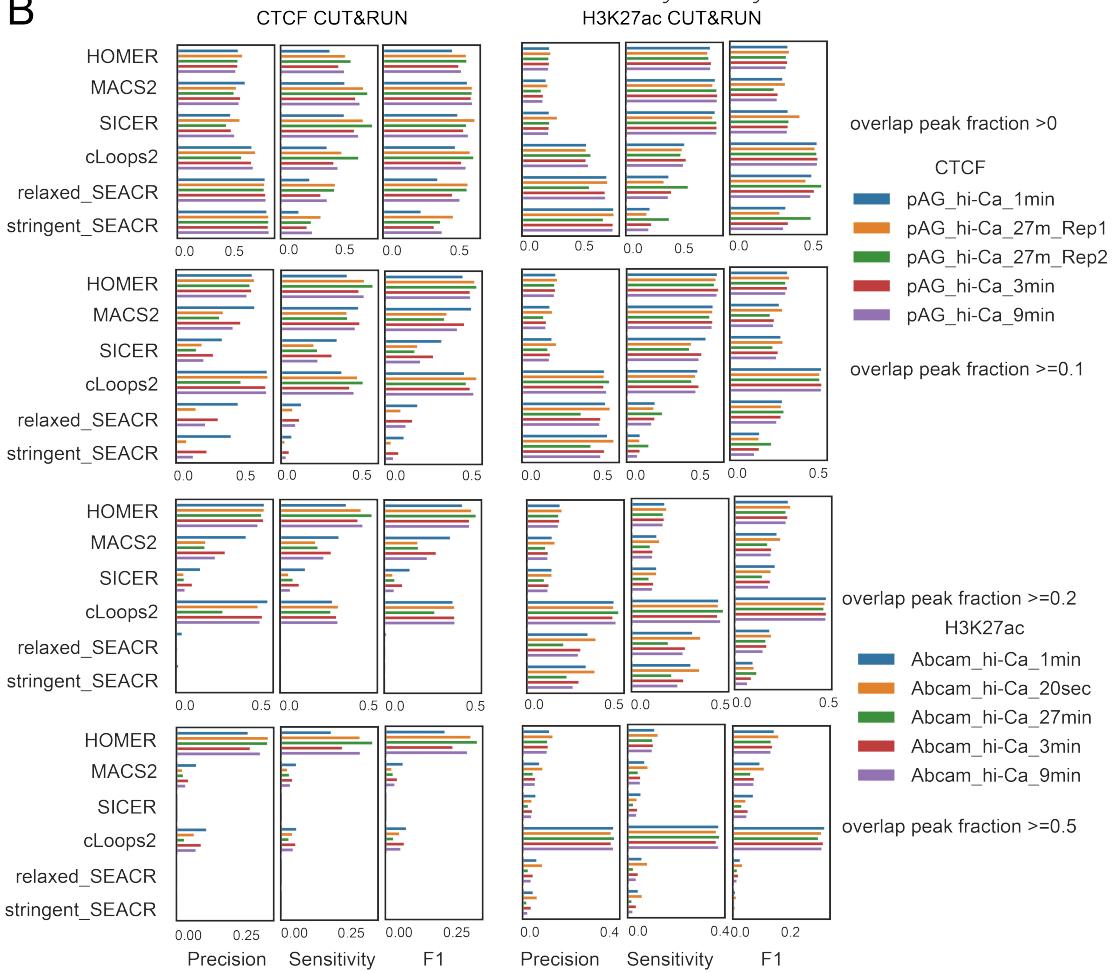
- (B) Examples of H3K27ac CUT&RUN profiles and peaks called by different tools.
- (C) Examples of H3K27me3 CUT&RUN profiles and peaks called by different tools.

Supplemental Figure 5

A



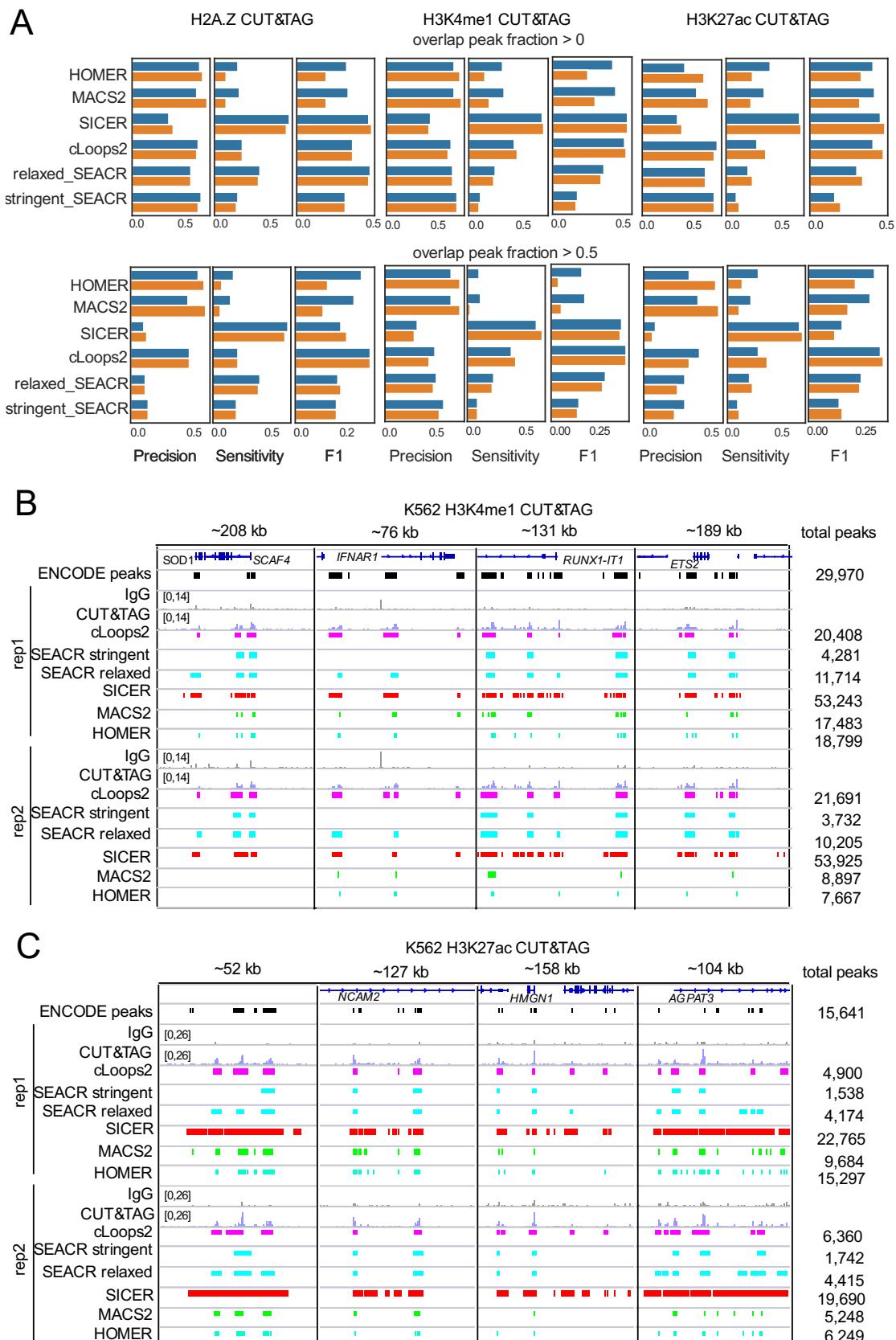
B



Supplemental Figure 5. Global comparison of peaks called by cLoops2 and other tools from CUT&RUN data

- (A) The signal enrichment scores (ES) of peak regions compared to peak-flanking regions and IgG control. The peaks were called by cLoops2 and other tools from CTCF, H3K27ac, and H3K27me3 CUT&RUN data.
- (B) Precision, sensitivity, and F1 scores for peaks called by different tools compared to the ENCODE peaks.

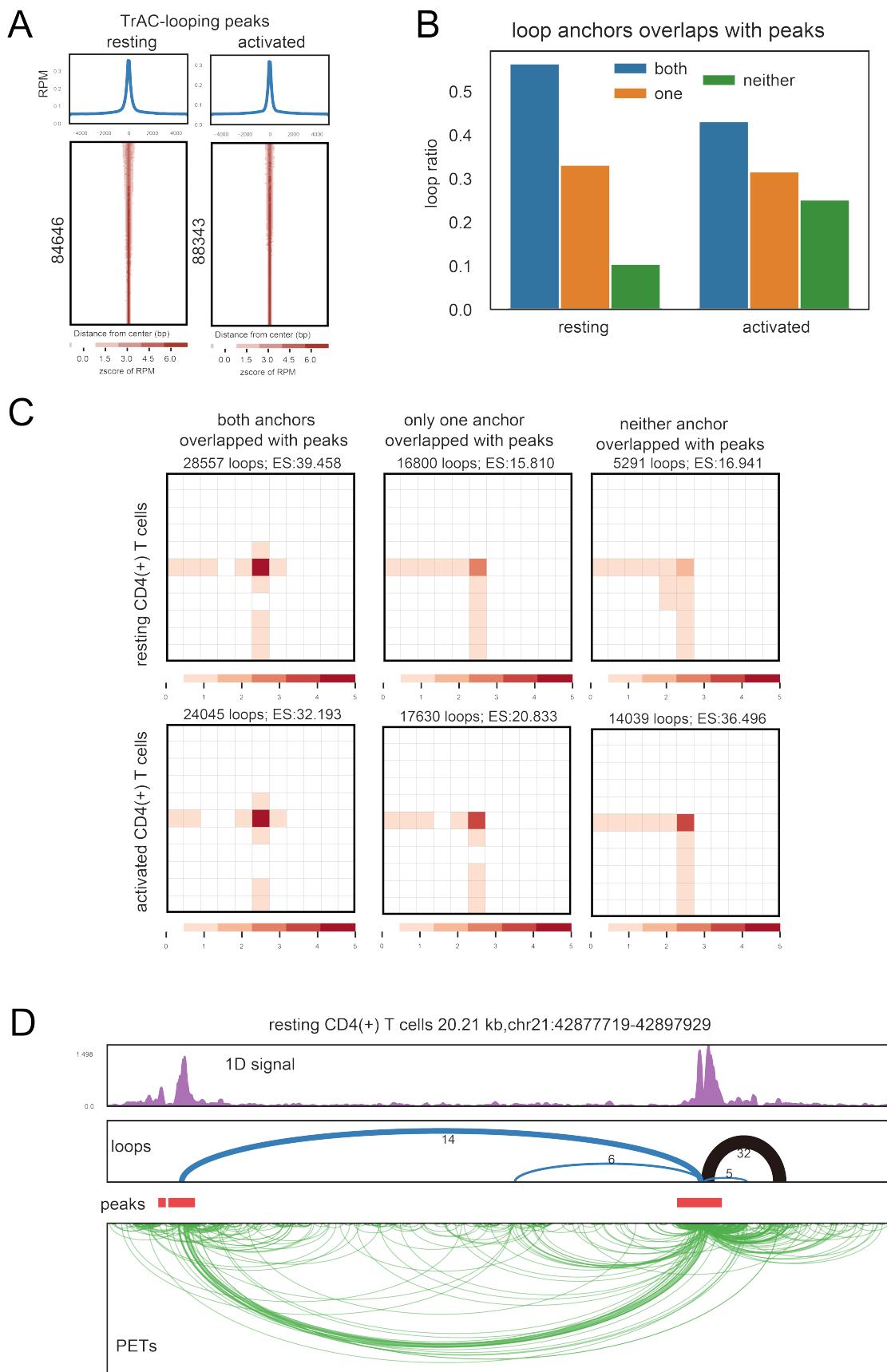
Supplemental Figure 6



Supplemental Figure 6. Global comparison of peaks called by cLoops2 and other tools from CUT&TAG data

- (A) Precision, sensitivity, and F1 scores for peaks called by different tools in comparison with the ENCODE peaks.
- (B) Examples of H3K4me1 CUT&TAG profiles and peaks called by different tools.
- (C) Examples of H3K27ac CUT&TAG profiles and peaks called by different tools.

Supplemental Figure 7

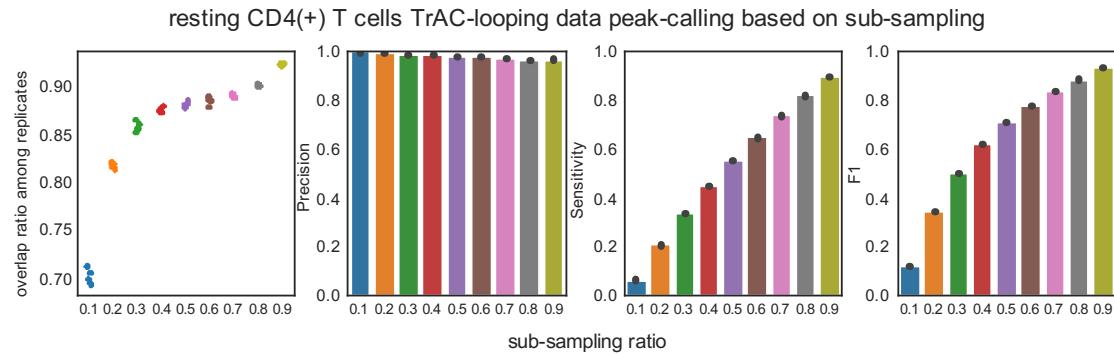


Supplemental Figure 7. cLoops2 is sensitive to detect non-peak loop anchors for interaction data

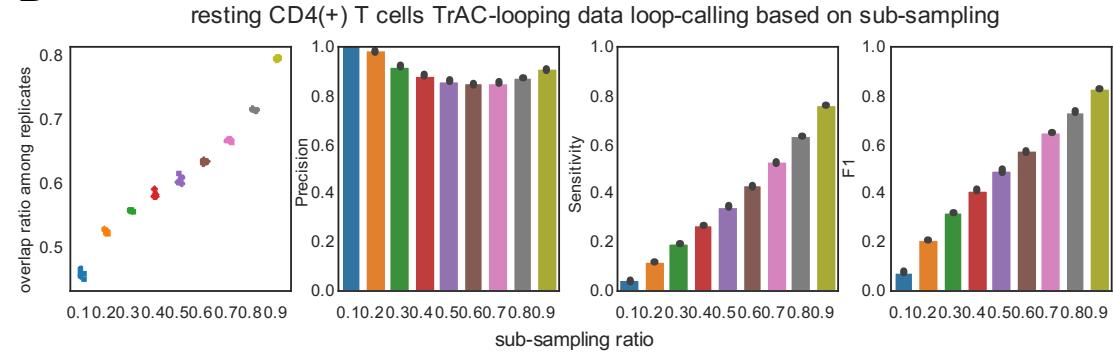
- (A) 1D signal aggregation analysis for peaks called by cLoops2 from resting and activated CD4+ cells TrAC-looping data.
- (B) Overlaps of peaks and loops anchors called by cLoops2 from TrAC-looping data. In figure legends, “both” marks loops with both anchors overlapped with called peaks, “one” marks loops with only one anchor overlapped with peaks, and “neither” marks loops without any overlaps with peaks.
- (C) Interaction signal aggregation analysis for loops classified by anchors overlapped with peaks.
- (D) Randomly selected example of loops not both anchors are peaks. A strong loop with 32 PETs (marked by black) only has one anchor overlapping with peaks, and from the 1D signal, the non-peak-overlapping anchor is not a peak.

Supplemental Figure 8

A



B

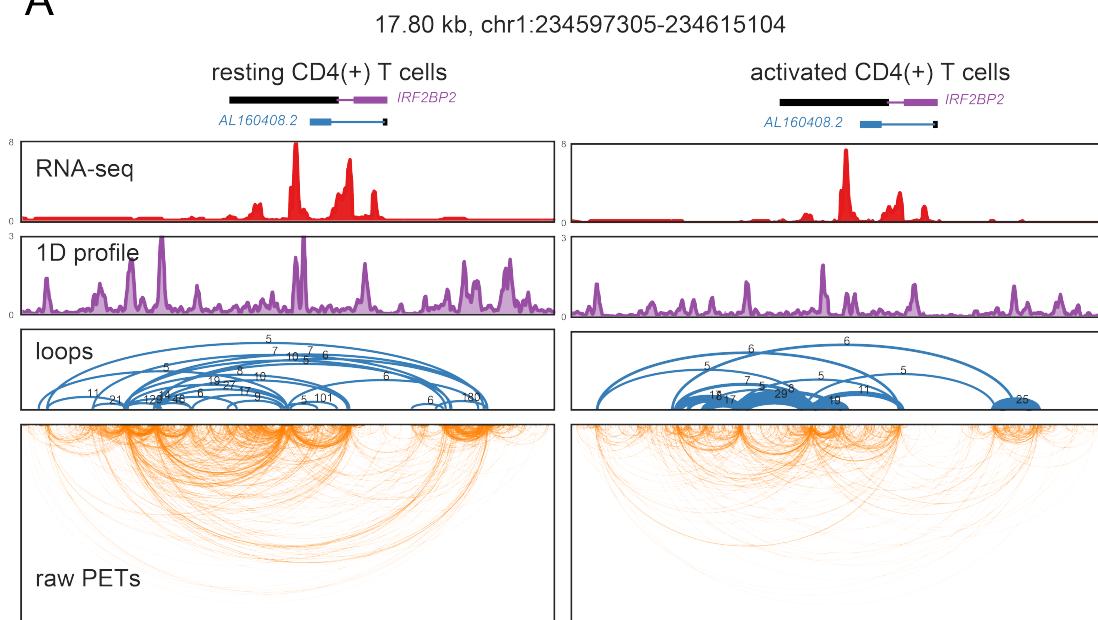


Supplemental Figure 8. cLoops2 is sensitive to library sequencing depth for peak-calling and loop-calling performance of interaction data

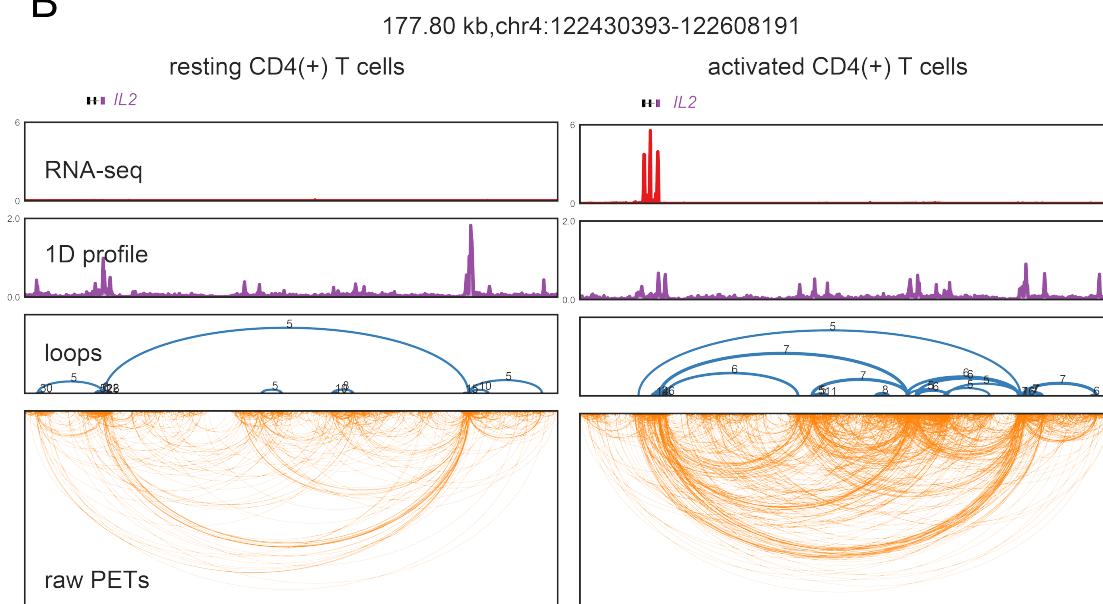
- (A) Sub-sampling with 5 replicates at a ratio from 0.1 to 0.9 was performed for the resting CD4+ T cells TrAC-looping data to benchmark the performance of cLoops2 callPeaks. Reproducibility among sub-sampling replicates was measured by the mean overlapping ratio for one replicate to all others. Precision was calculated as overlapped peaks ratio comparing to peaks called from all reads. Sensitivity was obtained as detected peaks ratio from peaks called from all reads. F1 score was calculated from the precision and sensitivity.
- (B) Sub-sampling was performed the same way as panel (A) to benchmark the performance of the cLoops2 callLoops module.

Supplemental Figure 9

A



B



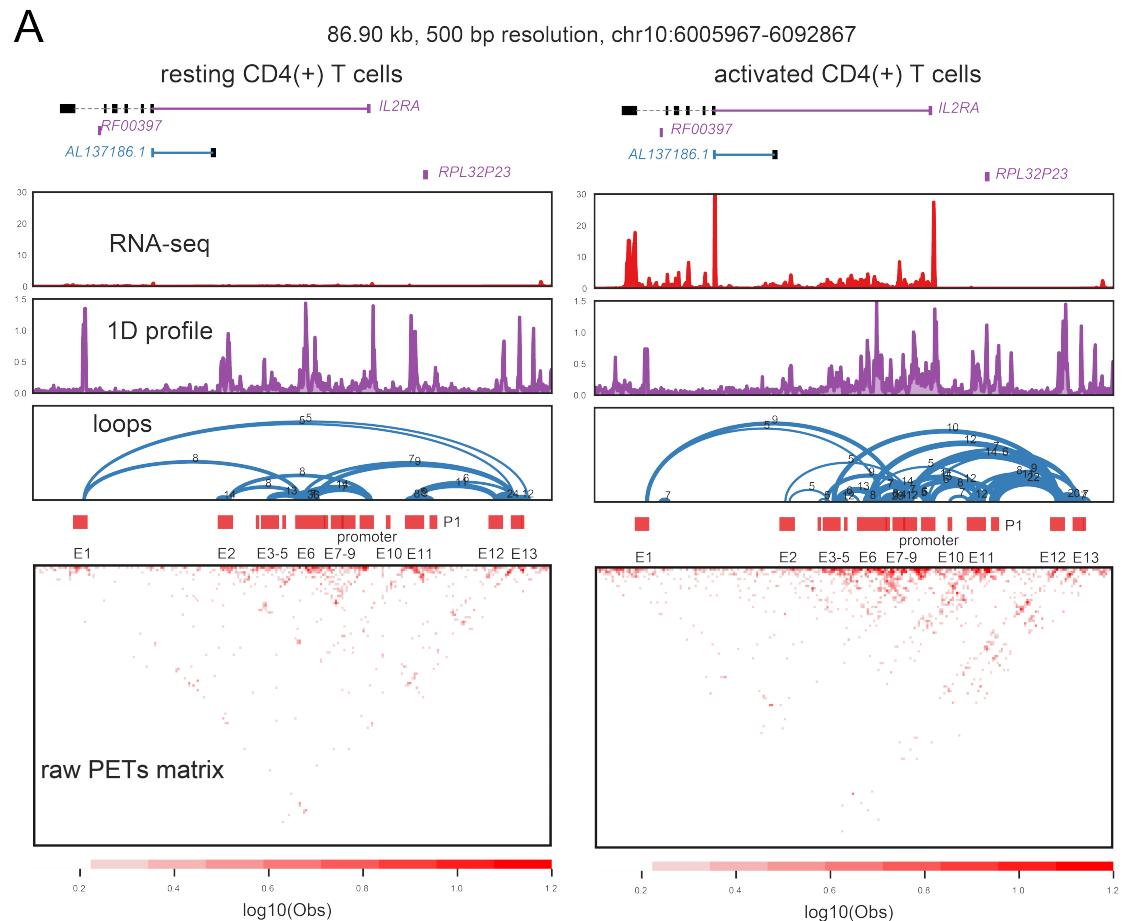
Supplemental Figure 9. Examples of loops called by cLoops2 from CD4 T cell TrAC-looping data

(A) Loops called by cLoops2 from TrAC-looping data around the *IRF2BP2* gene.

The plots were generated by the cLoops2 plot module with -arch parameter. The blue color marked the first exon of the gene in the positive strand, and the purple color marked the first exon of the gene in the negative strand.

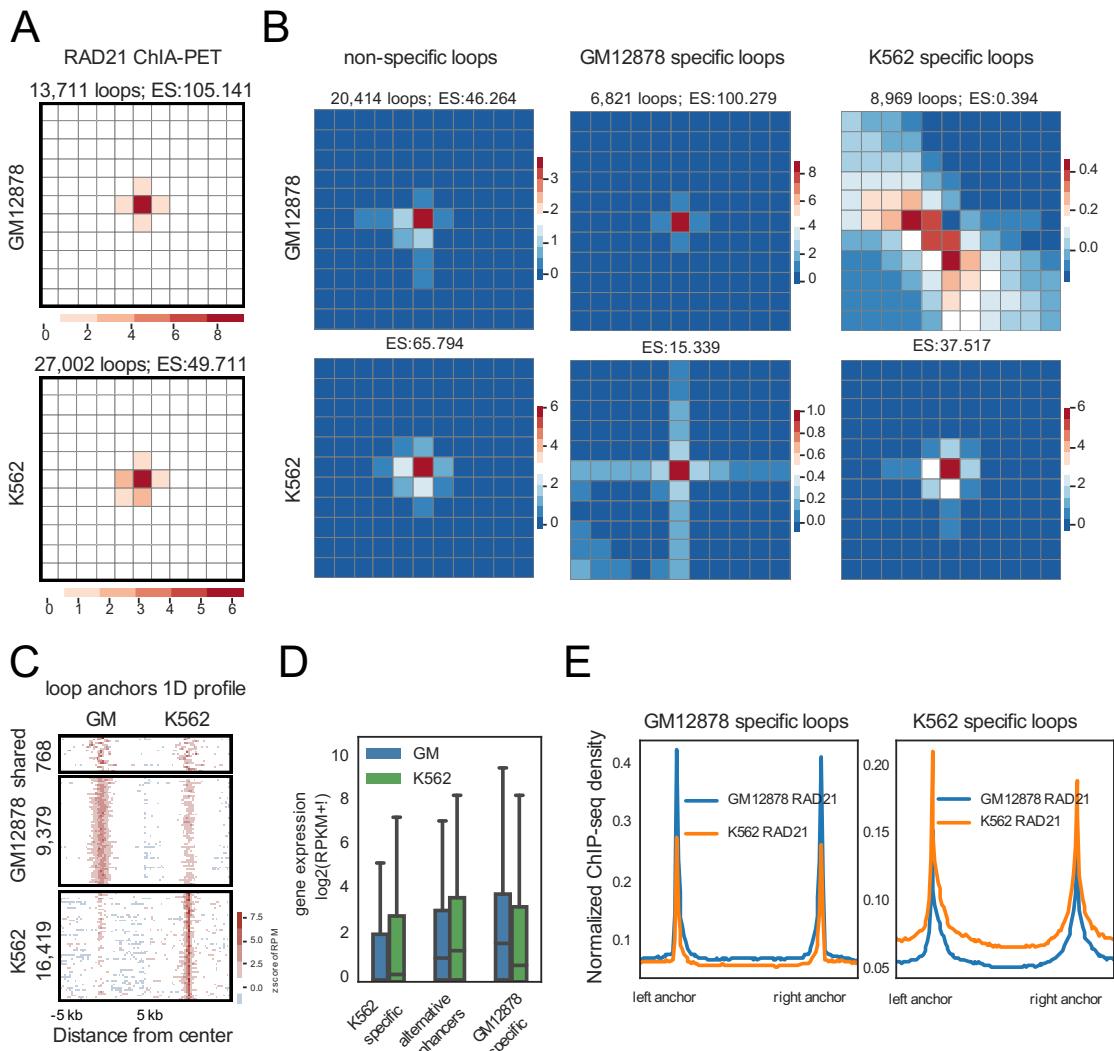
(B) Loops called by cLoops2 from TrAC-looping data around the *IL2* gene. The plots were generated by the cLoops2 plot module with -arch parameter.

Supplemental Figure 10



Supplemental Figure 10. The loops called by cLoops2 around the *IL2RA* gene promoter in resting and activated CD4 T cells

Supplemental Figure 11

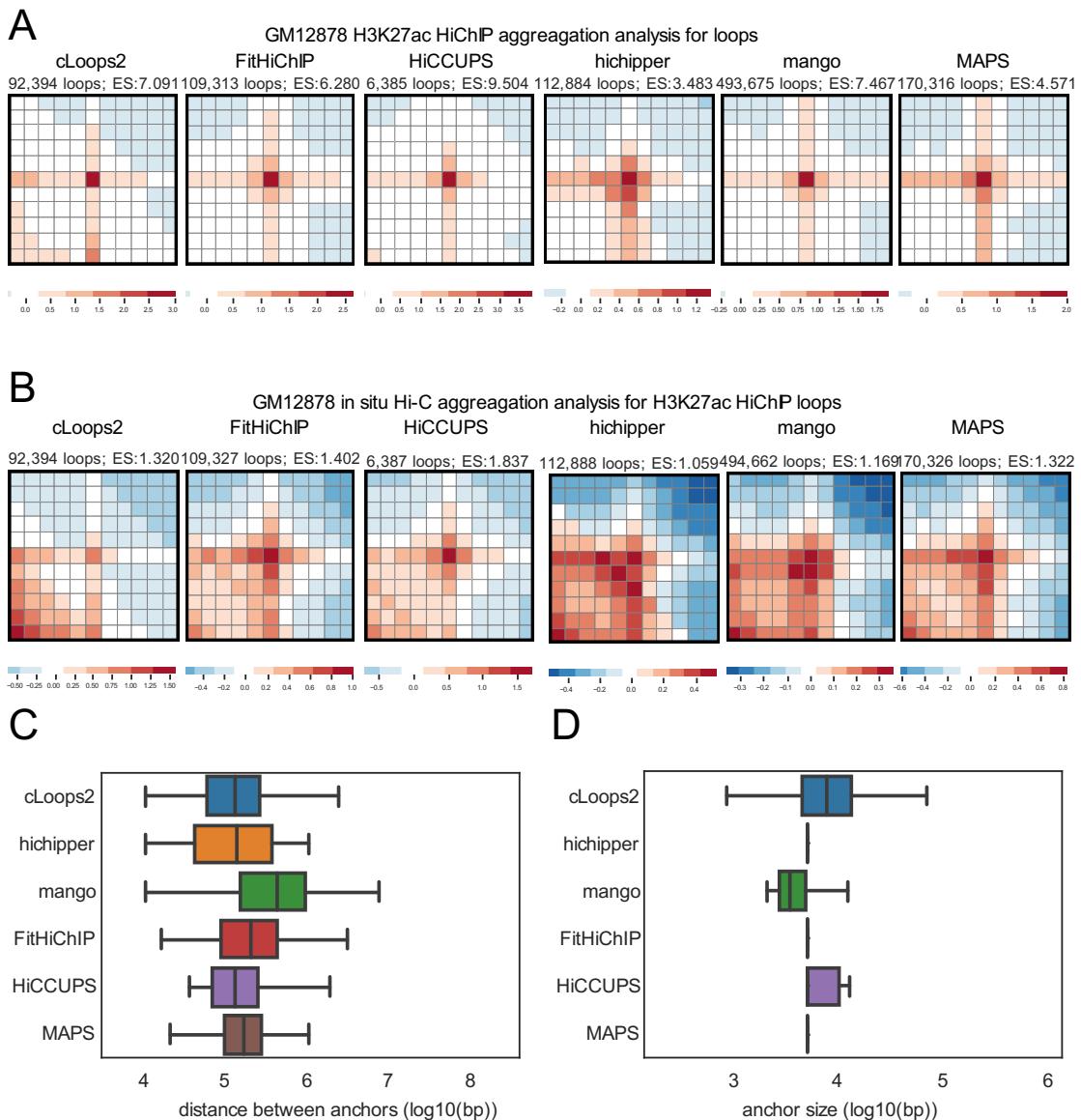


Supplemental Figure 11. Differentially enriched loops from RAD21 ChIA-PET data called by cLoops2

- (A) Aggregation analysis of loops called by cLoops2 from the RAD21 ChIA-PET data from GM12878 and K562 cells.
- (B) Aggregation analysis of the non-specific and cell-specific RAD21 guided loops between GM12878 and K562.
- (C) Aggregation analysis of ChIA-PET 1D signals of the combined anchors from the non-specific and cell-specific loops.
- (D) Gene expression association with cell-specific RAD21 guided loops. Alternative enhancers referred to the genes whose promoters were looped to different enhancers in the two different cell types.

(E) Aggregated RAD21 ChIP-seq signals on cell-specific RAD21-guided loops. ChIP-seq data were obtained from ENCODE.

Supplemental Figure 12



Supplemental Figure 12. Comparison of loops called for GM12878 H3K27ac HiChIP data

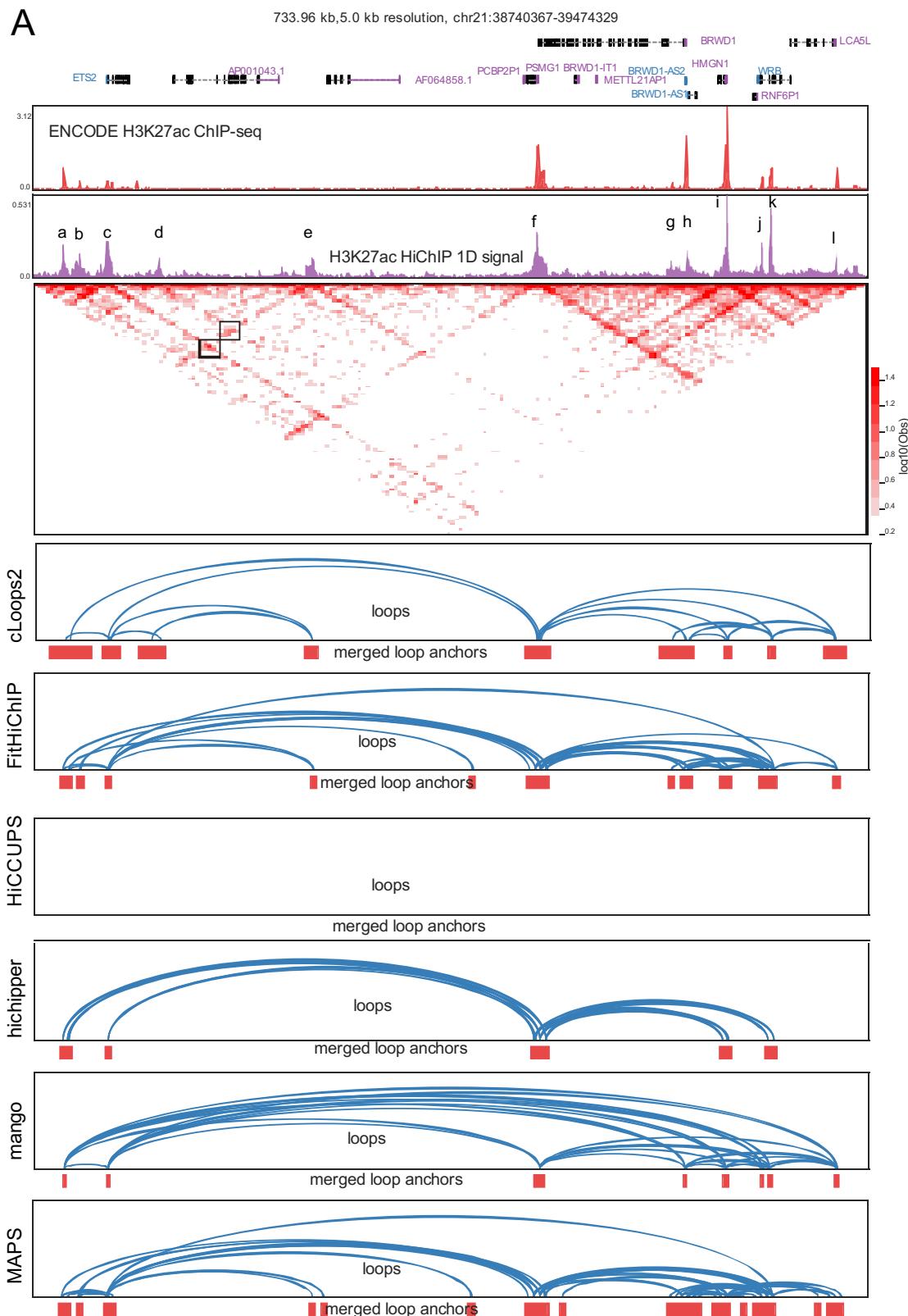
(A) Interaction aggregation analysis of GM12878 H3K27ac HiChIP data for loops called by different tools. The numbers of called loops and the mean enrichment scores of the center to rest of the matrix were annotated as the title for each heatmap.

(B) Interaction aggregation analysis of GM12878 Hi-C data for loops called from H3K27ac HiChIP data.

(C) Distribution of loop anchor distances.

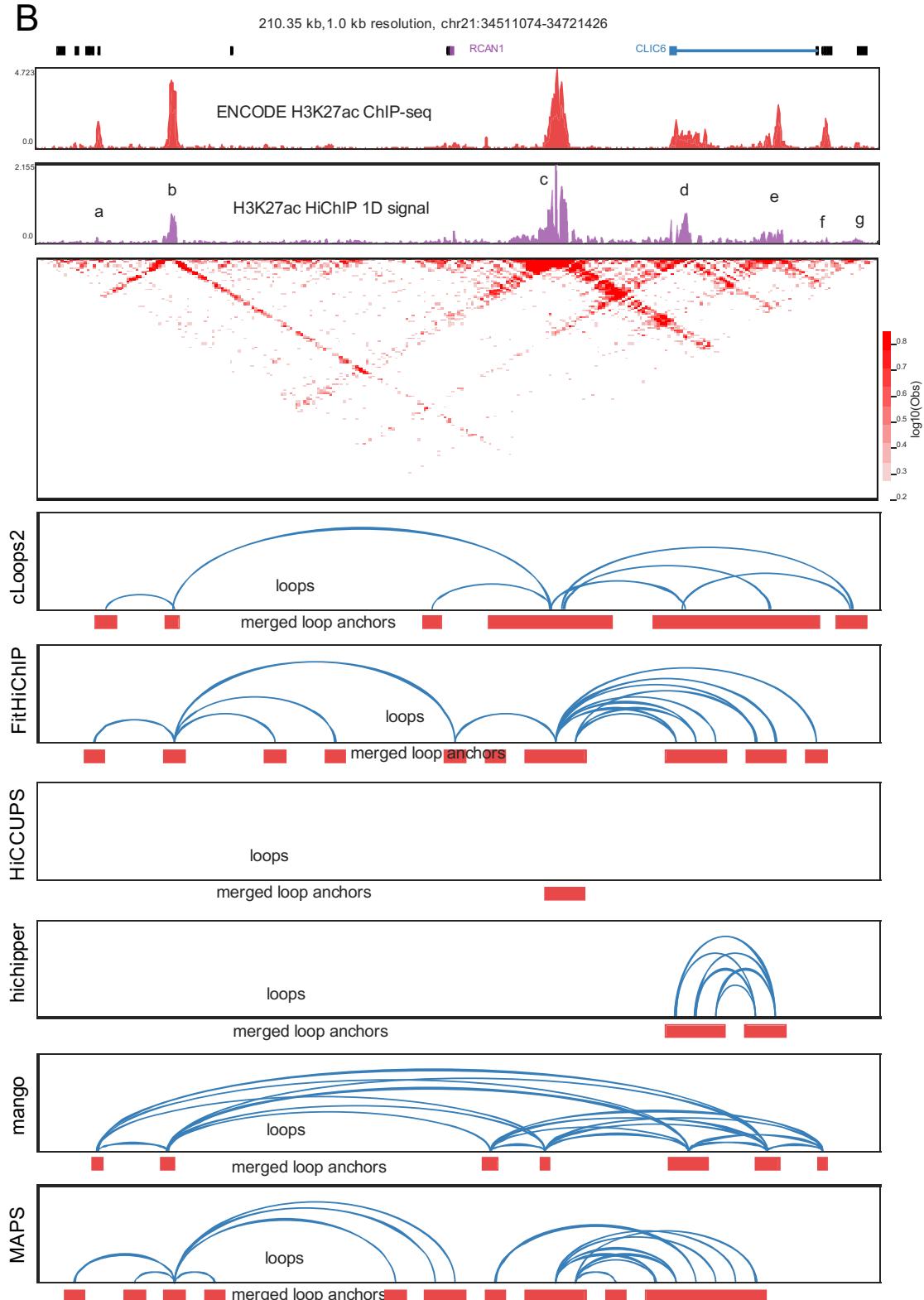
(D) Distribution of loop anchor sizes.

Supplemental Figure 13



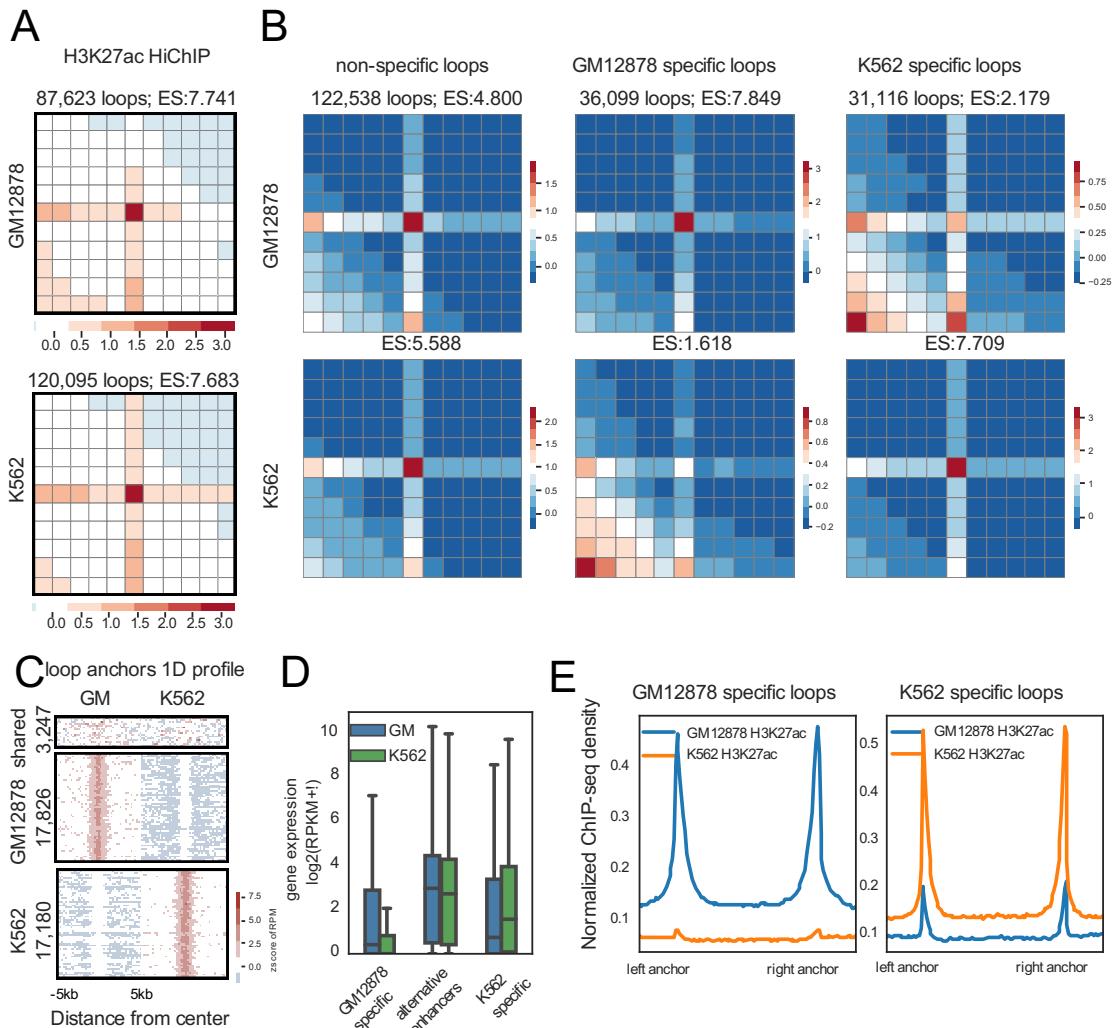
Supplemental Figure 13

B



Supplemental Figure 13. Examples of called loops for GM12878 H3K27ac HiChIP data

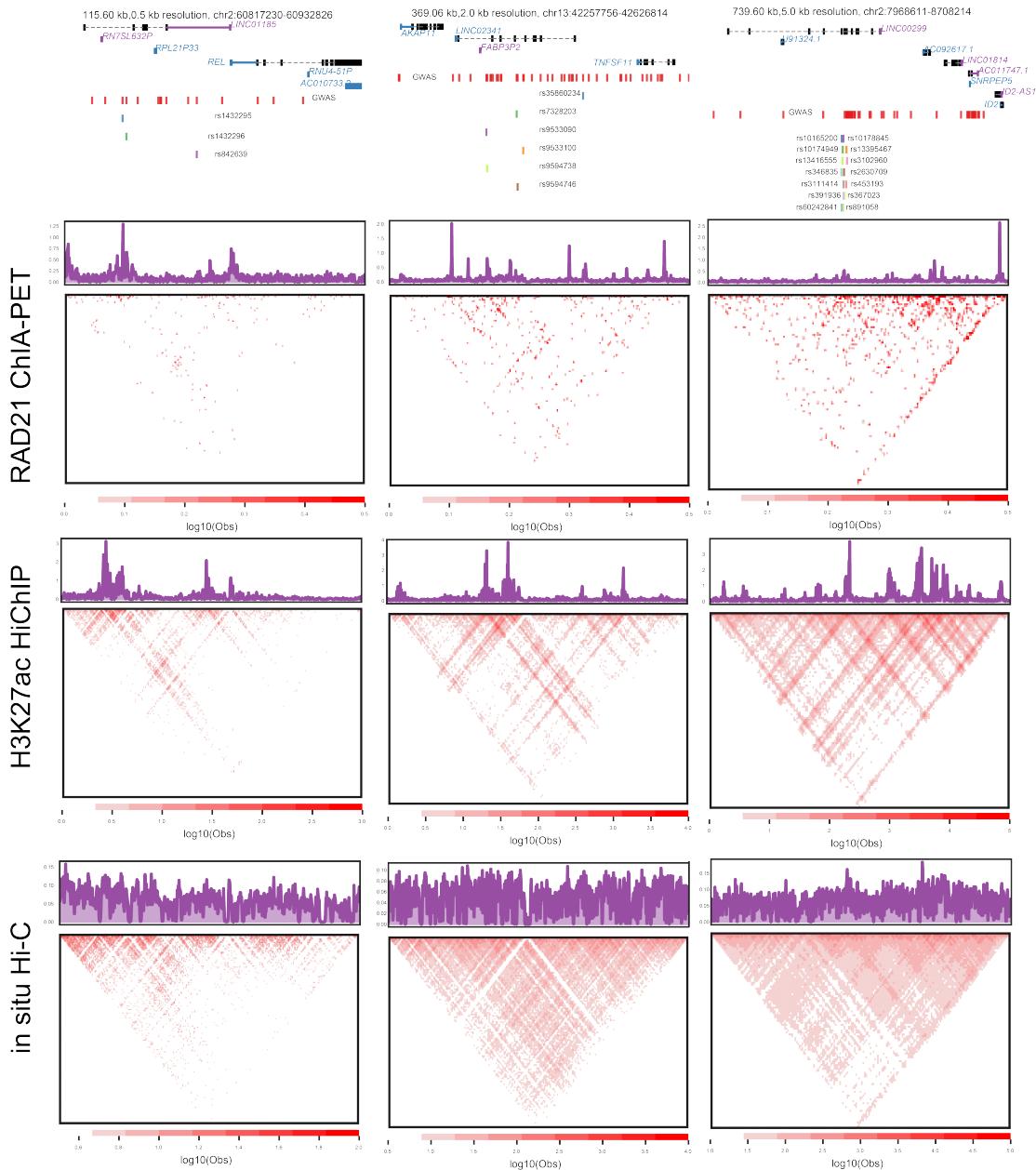
Supplemental Figure 14



Supplemental Figure 14. Differentially enriched loops from H3K27ac HiChIP data called by cLoops2

- (A) Aggregation analysis of loops called by cLoops2 from the H3K27ac HiChIP data from GM12878 and K562 cells.
- (B) Aggregation analysis of the non-specific and cell-specific H3K27ac associated loops between GM12878 and K562.
- (C) Aggregation analysis of HiChIP 1D signals of the combined anchors from the non-specific and cell-specific loops.
- (D) Gene expression is associated with cell-specific H3K27ac associated loops.
- (E) Aggregated H3K27ac ChIP-seq signals on the cell-specific H3K27ac-associated loops. ChIP-seq data were obtained from ENCODE.

Supplemental Figure 15



Supplemental Figure 15. GM12878 RAD21 ChIA-PET, H3K27ac HiChIP, and Hi-C data for the three lncRNAs shown in Figure 5C-E

Reference

1. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E. *et al.* (2019) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*, **47**, D1005-D1012.
2. Grubert, F., Srivas, R., Spacek, D.V., Kasowski, M., Ruiz-Velasco, M., Sinnott-Armstrong, N., Greenside, P., Narasimha, A., Liu, Q., Geller, B. *et al.* (2020)

- Landscape of cohesin-mediated chromatin loops in the human genome. *Nature*, **583**, 737-743.
3. Mumbach, M.R., Satpathy, A.T., Boyle, E.A., Dai, C., Gowen, B.G., Cho, S.W., Nguyen, M.L., Rubin, A.J., Granja, J.M., Kazane, K.R. *et al.* (2017) Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat Genet*, **49**, 1602-1612.
 4. Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F. *et al.* (2012) Landscape of transcription in human cells. *Nature*, **489**, 101-108.
 5. Rao, Suhas S.P., Huntley, Miriam H., Durand, Neva C., Stamenova, Elena K., Bochkov, Ivan D., Robinson, James T., Sanborn, Adrian L., Machol, I., Omer, Arina D., Lander, Eric S. *et al.* (2014) A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*, **159**, 1665-1680.
 6. Tang, Z., Luo, Oscar J., Li, X., Zheng, M., Zhu, Jacqueline J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B. *et al.* (2015) CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell*, **163**, 1611-1627.
 7. Lai, B., Tang, Q., Jin, W., Hu, G., Wangsa, D., Cui, K., Stanton, B.Z., Ren, G., Ding, Y., Zhao, M. *et al.* (2018) Trac-looping measures genome structure and chromatin accessibility. *Nat Methods*, **15**, 741-747.
 8. Consortium, E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57-74.
 9. Meers, M.P., Bryson, T.D., Henikoff, J.G. and Henikoff, S. (2019) Improved CUT&RUN chromatin profiling tools. *Elife*, **8**.
 10. Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G., Ahmad, K. and Henikoff, S. (2019) CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun*, **10**, 1930.
 11. Bhattacharyya, S., Chandra, V., Vijayanand, P. and Ay, F. (2019) Identification of significant chromatin contacts from HiChIP data by FitHiChIP. *Nat Commun*, **10**, 4221.
 12. Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M.H., Lander, E.S. and Aiden, E.L. (2016) Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst*, **3**, 95-98.
 13. Lareau, C.A. and Aryee, M.J. (2018) hichipper: a preprocessing pipeline for calling DNA loops from HiChIP data. *Nat Methods*, **15**, 155-156.
 14. Phanstiel, D.H., Boyle, A.P., Heidari, N. and Snyder, M.P. (2015) Mango: a bias-correcting ChIA-PET analysis pipeline. *Bioinformatics*, **31**, 3092-3098.
 15. Juric, I., Yu, M., Abnousi, A., Raviram, R., Fang, R., Zhao, Y., Zhang, Y., Qiu, Y., Yang, Y., Li, Y. *et al.* (2019) MAPS: Model-based analysis of long-range chromatin interactions from PLAC-seq and HiChIP experiments. *PLoS Comput Biol*, **15**, e1006982.
 16. Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J. and Meyer, B.J. (2015) Condensin-driven remodelling of

- X chromosome topology during dosage compensation. *Nature*, **523**, 240-244.
17. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357-U354.