Figure S1, related to Figure 1

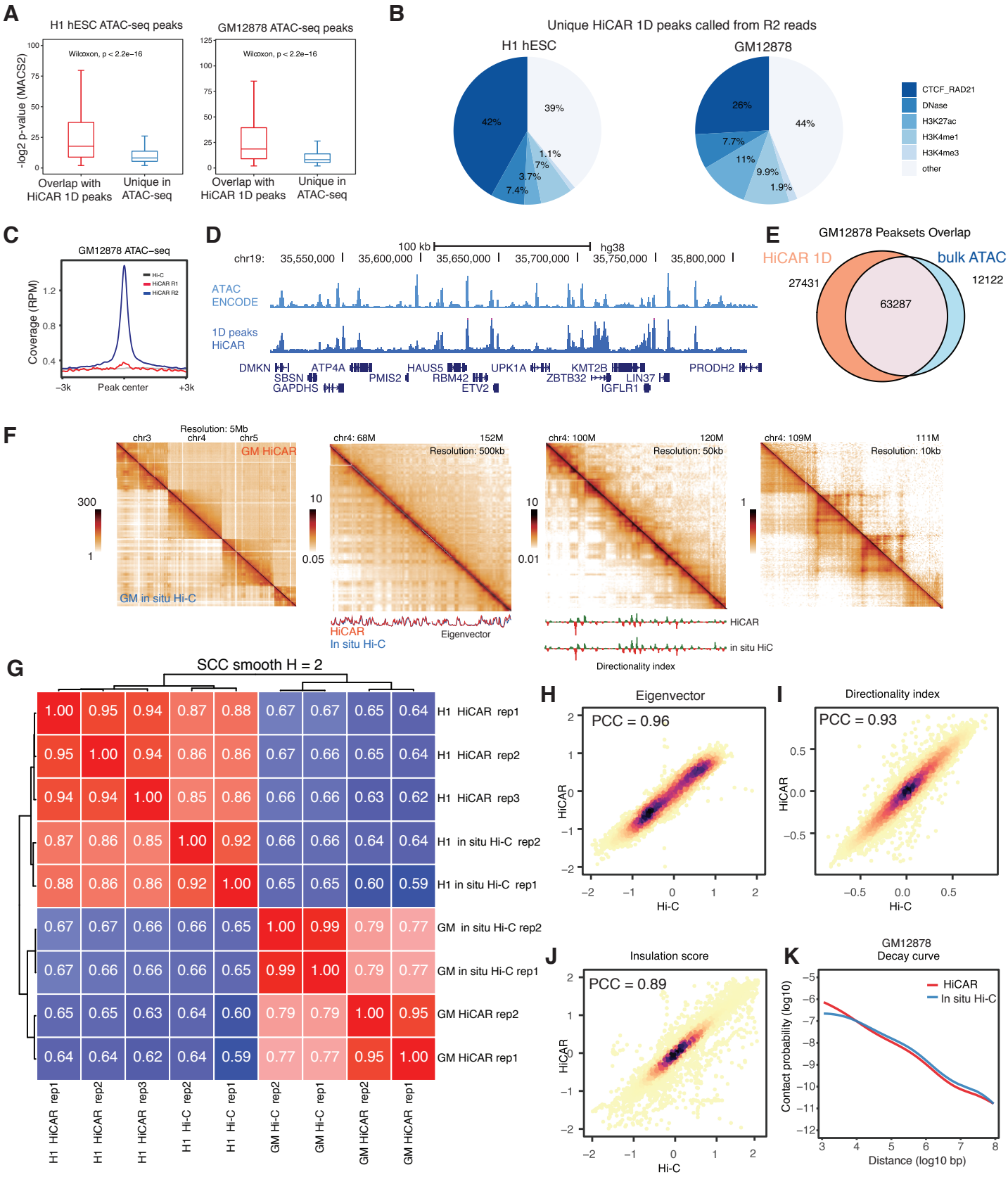


Figure S1. HiCAR identifies high-confident chromatin accessibility and genome architecture in both H1 hESC and GM12878 cells. Related to Figure 1.

(A) We compared the open chromatin peaks called by MAPS2 using HiCAR R2 reads and regular ATAC-seq data in H1 hESC (left) and GM12878 (right) cells. Boxplot showing the distribution of the MACS2 P-value of the overlapping peaks shared by HiCAR and ATAC-seq (red box), and the peaks unique to ATAC-seq (blue box). Wilcoxon rank-sum test was used for statistical analysis to compute P value. **(B)** The HiCAR unique 1D peaks overlap with DHS sites (DNase) and ChIP-seq peaks of CTCF, RAD21, H3K27ac, H3K4me3, H3K4me1 in H1 hESC (left) and GM12878 cells (right). **(C)** We count the number of HiCAR R2 reads (blue), R1 reads (red), and *in situ* Hi-C (black) reads within +/- 3kb window centered at GM12878 ATAC-seq peaks. The HiCAR R1, R2 and Hi-C reads are normalized against sequence depth (reads per million, RPM). The average reads count per ATAC-seq peak was plotted as signal coverage surrounding the +/- 3kb window of ATAC-seq peaks in y-axis. (**D)** A representative genome browser view showing the reads signals of the public GM12878 ATAC-seq data (top, light blue) and GM12878 HiCAR R2 reads (bottom, dark blue). **(E)** Venn diagram showing open chromatin peaks called by MAPS2 using GM12878 HiCAR R2 reads (HiCAR 1D peaks, pink) and GM12878 ATAC-seq (blue). **(F)** The sequence depth normalized contact matrices of HiCAR (top right, above the diagonal) and in situ Hi-C (bottom left, below the diagonal) data from GM12878 at successive zoom-in views. The GM12878 *in situ* Hi-C data was obtained from the 4DN data portal. The color represents sequence depth normalized reads signal (counts per million mapped reads). The Compartment Score, Directionality index, computed from in situ Hi-C and HiCAR data were plotted underneath the contact matrices as indicated. **(G)** HiCRep was employed to compute the similarity of chromatin contact matrices generated by HiCAR and in situ Hi-C from H1 hESC and GM12878. SCC values are computed using HiCrep. **(H-J)** Scatter plots and Pearson correlation coefficient (PCC) comparing the compartment scores **(H)**, directionality index **(I)**, and Insulation score **(J)** calculated from GM12878 HiCAR versus those from GM12878 *in situ* Hi-C. PCC: **(K)** Plot showing chromatin

contact frequency (y-axis) as a function of linear genomic distance (x-axis) measured by HiCAR (red curve) and *in situ* Hi-C (blue curve) in GM12878 cells.

Figure S2, related Figure 1

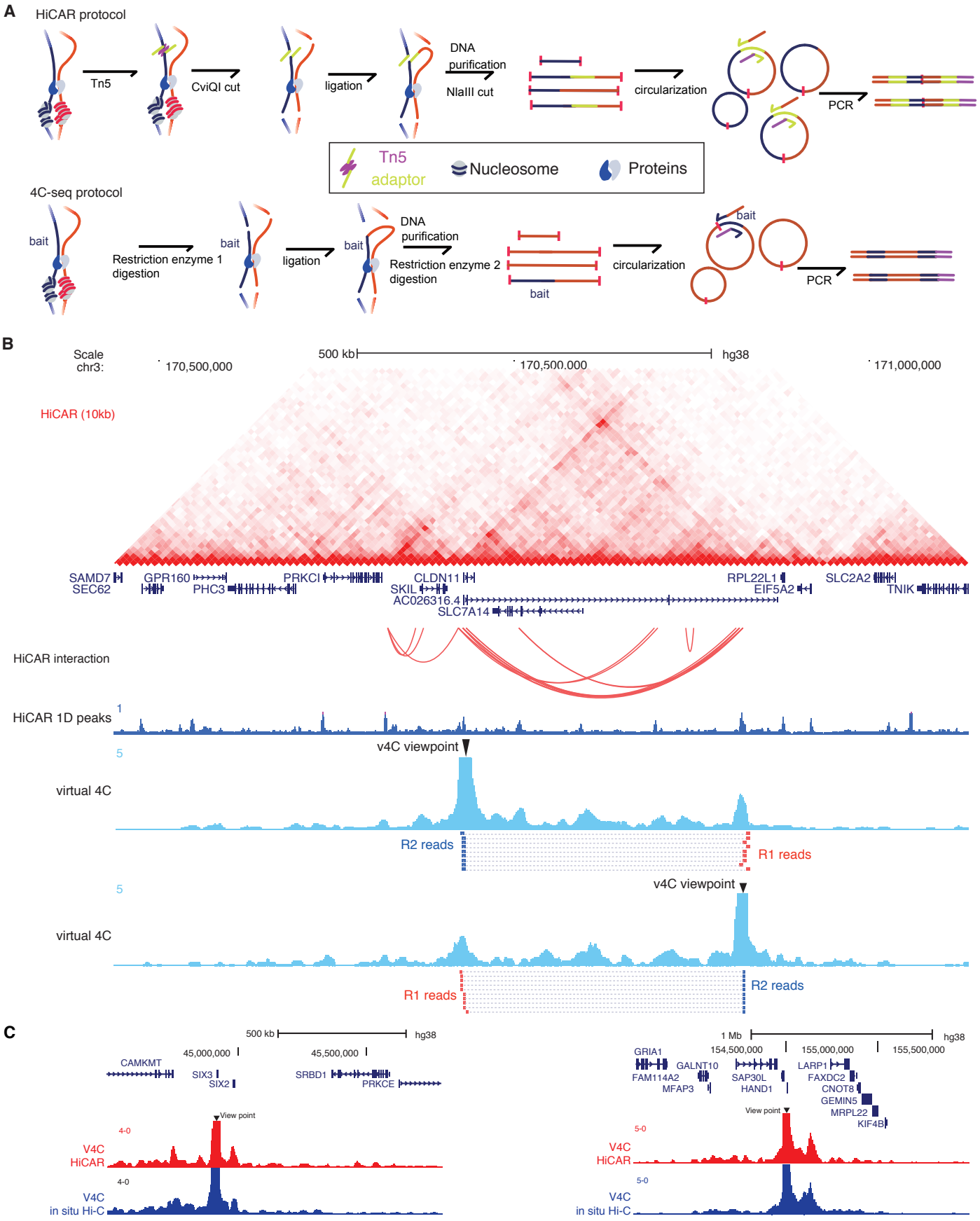
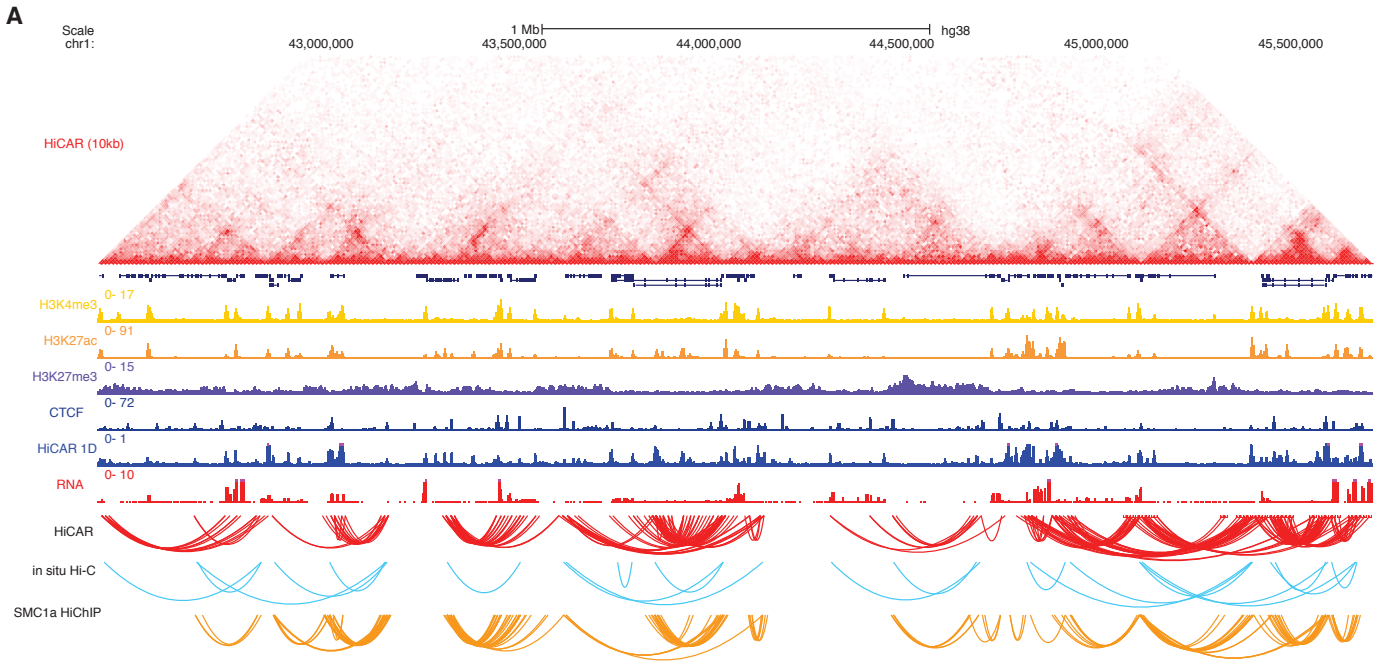
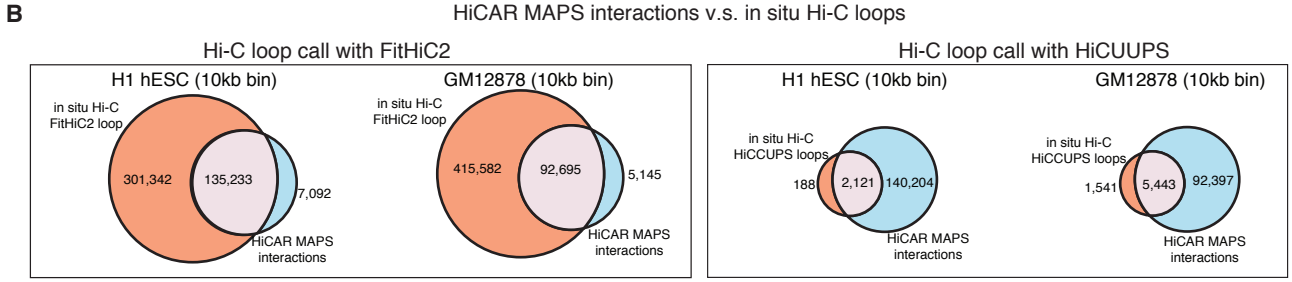


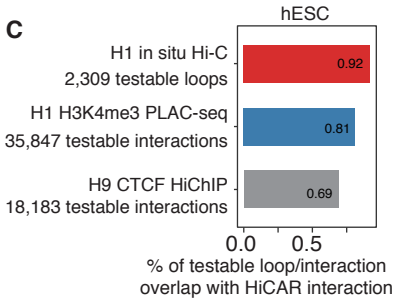
Figure S2. The virtual 4C analysis of HiCAR data. Related to Figure 1.

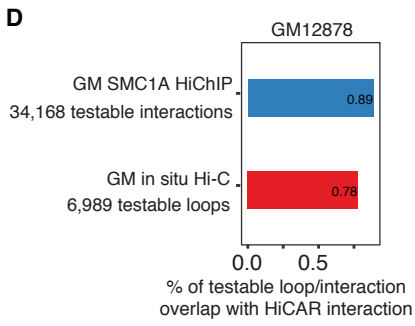
(A) Schematic illustration comparing HiCAR (top) and 4C-seq (bottom) experimental protocols. After the Tn5 tagmentation in HiCAR, both HiCAR and 4C-seq require two rounds of restriction enzyme digestion followed by ligation and circularization steps, to produce the circularized DNA template for PCR library amplification. 4C-seq requires a pair of primers recognizing the specific "bait" sequence, while HiCAR utilizes a pair of primers annealing to the Tn5 adaptor sequence. **(B)** The genome browser screenshot showing the virtual 4C (V4C) contact profile extracted from HiCAR data centered on two open chromatin peaks in H1 hESC. The H1 hESC HiCAR interaction matrices are shown as heatmaps along with HiCAR interaction arch tracks. The dashed lines represent the randomly sampled PET with R2 reads (blue) overlapping with the "viewpoint" open chromatin peaks (2kb). The red dots denote the R1 reads from the sampled PETs. **(C)** The v4C contact profile derived from HiCAR (red tracks) and in situ Hi-C (blue tracks) data of H1 hESC centered on SIX3 (left) and HAND1 (right) locus, respectively.

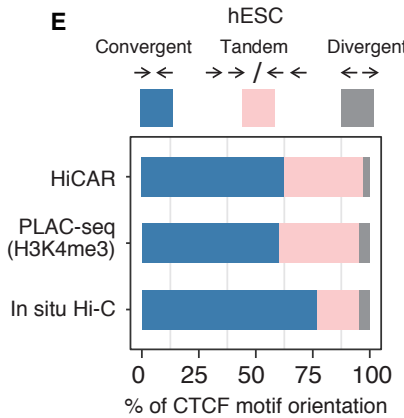
Figure S3, related to Figure 3











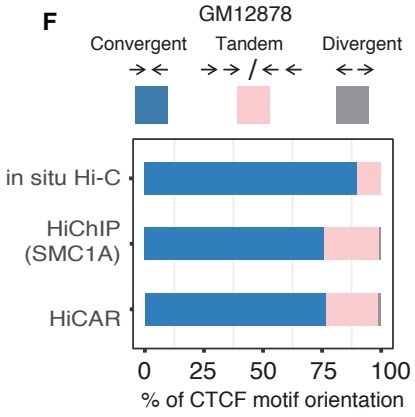
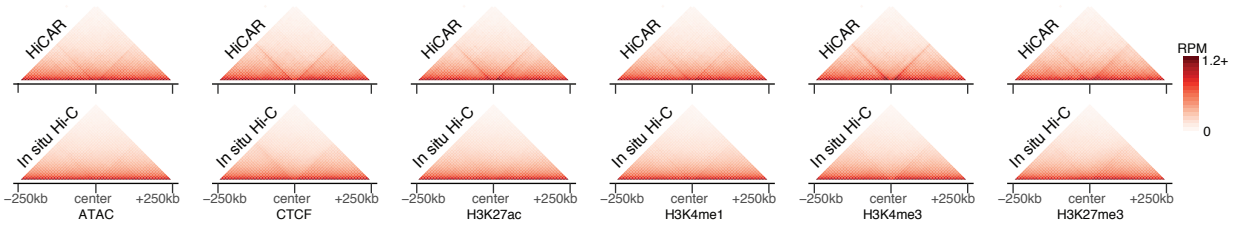


Figure S3. HiCAR is a robust and sensitive method to identify open chromatin anchored cRE interactions in H1 hESC and GM12878 cells. Related to Figure 3.

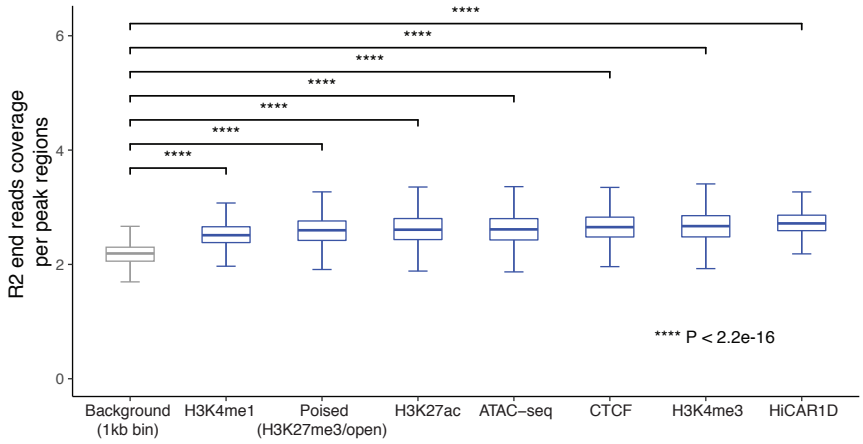
(A) GM12878 HiCAR contact matrices are shown on the top as heatmap along with genome browser tracks of H3K4me3, H3K27ac, H3K27me3, CTCF ChIP-Seq, RNA-seq and HiCAR R2 reads (1D open chromatin profile) from GM12878. The arch tracks represent the chromatin loops and interactions identified from HiCAR, in situ Hi-C and SMC1a HiChIP data from GM12878. **(B)** Venn diagram of HiCAR MAPS interactions compared to in situ Hi-C loops called by FitHiC2 (left box) and HiCCUPS (right box) in both H1 hESC and GM12878 cells. **(C, D)** We examined the orientation of the CTCF motifs located on the pairwise anchors of each chromatin loop and interactions. The length of the color bar indicates the percentage of convergent (blue), tandem (pink) and divergent (grey) CTCF motif pairs among the tested HiCCUPS loops and MAPS interactions in **(C)** hESC and **(D)** GM12878 cells. **(E, F)** The Hi-C, HiChIP, and PLAC-seq loops and interactions with at least one anchor overlapping with HiCAR 1D open chromatin peaks are defined as the “testable” loops/interactions. The percentage of the “testable” loops/interactions that overlap with HiCAR interaction were calculated to estimate the sensitivity of HiCAR interaction calling in **(E)** hESC and **(F)** GM12878 cells.

Figure S4, related to Figure 3

A



B



C

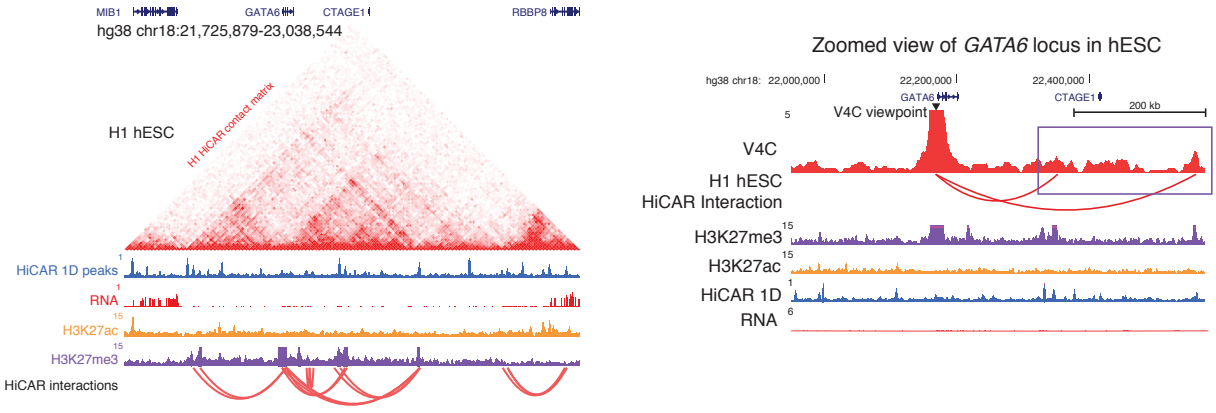


Figure S4: HiCAR can effectively enrich open chromatin sequences associated with distinct epigenetic marks. Related to Figure 3.

(A) Using sequence depth normalized H1 hESC HiCAR (top) and in situ Hi-C (bottom) contact matrix, the paired-end reads counts from the indicated sub chromatin contact matrix (10kb bin) were aggregated. The sub chromatin contact matrices are entered on the indicated ATAC-seq and ChIP-seq peaks of H1 hESC, and extend +/- 250kb window. Color key: normalized reads counts (RPM). **(B)** Boxplot showing that the coverage of H1 hESC HiCAR library R2 reads across the human genome at indicated marks at 1kb fixed bins. The number of HiCAR R2 reads overlapping with 1kb ChIP-seq peaks of CTCF, H3K4me1, H3K4me3, H3K27ac, ATAC-seq and HiCAR 1D peaks were calculated and shown as boxplot. The “poised” peaks were defined as H3K27me3 ChIP-seq peaks overlapping with HiCAR 1D peaks. All the 1kb bins not overlapping with any of the above mentioned peaks are selected as the background. The Wilcox test were performed between background bins and bins overlapped with ChIP-seq, ATAC-seq or HiCAR 1D peaks. **(C)** Genome browser screenshot showing HiCAR contact matric, 1D peaks, RNA signal, and ChIP-seq signals of H3K27ac and H3K27me3 on GATA6 locus in H1 hESC. Right: The zoomed view including virtual 4C (V4C) tracks centered on 2kb GATA6 TSS.

Figure S5, related to Figure 4

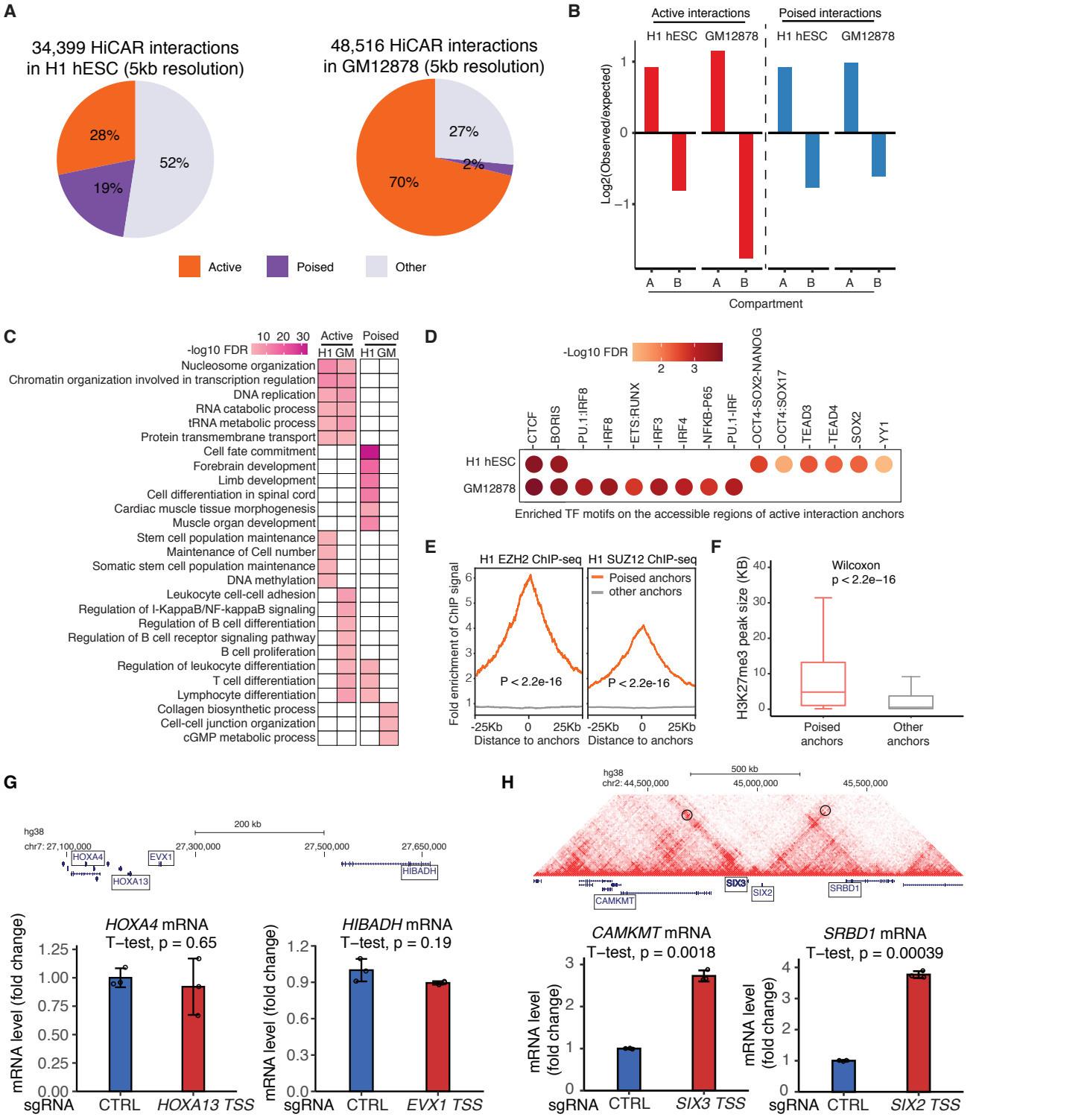


Figure S5. HiCAR captures open chromatin anchored "active" versus "poised" interactions in H1 hESC and GM12878 cells. Related to Figure 4.

(A) Pie chart showing the percentage of active (orange) and "poised" interactions (purple) called in H1 hESC (left) and GM12878 (right) HiCAR data at 5kb resolution. **(B)** Based on the genome-wide distribution of A/B compartment in H1 hESC and GM12878 cells, we calculated the "expected" number of active or "poised" interactions that are located within A or B compartment in H1 hESC and GM12878. We also counted the actual number of active (left) and "poised" (right) interactions detected in compartment A or B, and compared to that of "expected" numbers in H1 hESC and GM12878. Y-axis: log2 transformed fold change of observed versus expected active (red) and "poised" (blue) interactions in H1 hESC and GM12878 cells. **(C)** We selected the genes with promoters overlapping with the anchor sequences of active or "poised" interactions detected in H1 hESC and GM12878 cells. The resulting gene lists were subjected to Gene Ontology analysis. Color key: P value from clusterProfile hypergeometric test. **(D)** We selected the HiCAR 1D peaks overlapping with anchor sequences of active or "poised" interactions detected in H1 hESC and GM12878 cells. Transcription-factor (TF) motif enrichment analysis was performed using HOMER on the open chromatin sequences (HiCAR 1D peaks) overlapping with interaction anchors. Color key: P value output from HOMER. **(E, F)** We selected the "poised" interactions from H1 hESC, and defined the rest of interactions as "other" interactions. **(E)** We expanded +/- 25kb from the center of the "poised" and "other" anchors, and calculated the ChIP-seq signal enrichment of EZH2 (left) and SUZ12 (right) surrounding these anchor regions. The ChIP-seq reads within +/- 25kb of anchors (100bp bin) was calculated, and compared to the average ChIP-seq signal outside 25kb away from anchors. The orange and grey curves represent ChIP-seq signal enrichment on the "poised" and "other" anchor, respectively. Wilcoxon test P value < 2.2e-16. **(F)** The size (kilobase, kb) of H3K27me peaks overlapping with "poised" versus "other" anchors of HiCAR interactions in H1 hESC were shown as box plot. Wilcoxon test P value < 2.2e-16. **(G, H)** CRISPRa experiments were carried to induce activation of promoter sequences of

SIX2, SIX3, EVX1, and *HOXA13*. The non-targeting sgRNA was used as negative control. H1 hESC were infected by lentivirus co-expressing VP64-dCas9-VP64 and sgRNA, selected by Puromycin for 3 days, and collected for total RNA extraction 10-days post infection. The mRNA level of the indicated genes were quantified by RT-qPCR analysis. The data was collected from three biological replicates. P-values: two-tailed Student's t-test. The H1 hESC HiCAR contact matrix is shown (**H**, top panel).

Methods S1. Detailed HiCAR protocol (for 30,000 to 100,000 input cells), related to STAR Methods.

- **Cell crosslinking (0.5h)**
  1. Prepare single cell suspension (by Accutase or Trypsin digestion, or FACS sorting).
  2. Wash the cells once by PBS. Spin down the cells by centrifugation at 850g for 5min.
  3. Resuspend the cells in 1ml PBS.
  4. Crosslink the cells by adding formaldehyde to the final concentration of 1% formaldehyde. Incubate the cells at room temperature for 10mins.
  5. Add 2.5M Glycine to the final concentration of 0.2M. Mix and incubate at room temperature for 10mins to quench formaldehyde.
  6. Spin down the cells (850g, 5min). Wash the cells once by PBS.
- **Nuclei preparation and Tn5 tagmentation (3.5h)**
  7. Add 400μl NPB buffer (PBS containing 5% BSA, 1mM DTT, 0.2% IGEPAL, and Roche Complete Proteinase Inhibitor, RNase Inhibitor), incubate at 4°C for 15min.
  8. Centrifuge at 600g for 5min. The pellet contains nuclei while the supernatant contains cytoplasm RNA. Save the supernatant for the optional RNA-seq library construction using the same input cells.
  9. Add 350μl 2x TB buffer (66mM Tris-AC pH 7.8, 132mM K-AC, 20mM Mg-AC, 32% DMF), 335μl molecular biology water, and 15μl assembled Tn5 (BfaI-truseqR1-pmeI-nextera7, table). Rotate at 37°C for 1.5h for tagmentation.
  10. Add 350μl EDTA (40mM) to stop the reaction
  11. Centrifuge at 850g for 5min, discard the supernatant.
  12. Wash nuclei once with 600μl PBS containing 0.075% BSA.
  13. Resuspend the nuclei in 32.5μl biology water, 5μl 10x NEBuffer 3.1(NEB, Cat# B7203S), 12.5μl 2% SDS, and incubated at 62ºC for 10 minutes
  14. Centrifuge at 850g for 5min. Remove the supernatant. The supernatant contains nuclei RNA and can be combined with the supernatant collected in step 8 for the optional RNA-seq library construction. The pellet will be used in the following step for HiCAR library preparation.
  15. Resuspend the nuclei pellet by adding 100μl H_2O, 14μl 10x NEBuffer3.1, 25ul 10% Triton X-100, mix and rotate at 37°C for 15min.
- **CviQI digestion and in situ ligation (6.5h)**

16. Spin down the nuclei at 4°C (850g, 5min). Wash the nuclei pellet once by 300μl 1.1x NEBuffer 3.1.
17. Resuspend the pellet in 90μl 1.1x NEBuffer 3.1.
18. Add 100U CviQI (NEB, #R0639L), 1μl splint oligo (20μM, oligo TruseqR1, Table S1) into the nuclei in 90μl 1.1x NEBuffer 3.1. Adjust the volume to 100μl. Mix and incubate at room temperature for 2h.
19. Perform in situ ligation by adding the following reagents.

Stock	Volume
10x T4 Ligation Buffer	48 μl
T4 DNA ligase (NEB, #M0202S)	6 μl
20mg/ml BSA (NEB, #B9000S)	2.4 μl
10% Triton X-100	40 μl
ddH_2O	283.6 μl

20. In situ ligation for 4h at room temperature, or overnight at 16°C.

- **Reverse crosslink and genomic DNA purification (3h).**

21. Centrifuge at 2000g for 5min, discard the supernatant and keep the pellet.
22. Resuspend the pellet in 200μl of 2x RCB buffer (100 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.4% SDS). Mix and incubate at 68°C for at least 1.5h.
23. Add 2μl glycogen (Invitrogen, #10814010), 20μl NaAC (3M, pH 5.4), 500μl Ethanol. Mix and incubate at -80°C for at least 30min.
24. Centrifuge at 14000g for 20min at 4°C to precipitate DNA. Discard the supernatant
25. Wash the DNA pellet twice by 75% cold Ethanol.
26. Dissolve the dried DNA in 21μl 10mM Tris-HCl (pH8.0).
27. Measure DNA concentration using Qubit.

- **Fill-in and second digestion using NlaIII (2.5h)**

28. Prepare fill-in mixture

Stock	Volume (μl)
DNA	19
10mM dNTP	4
10X Cutsmart buffer	5
T4 DNA polymerase (NEB, # M0203L)	1.5
ddH2O	20.5

29. Mix and incubate at room temperature for 30min.
30. Incubate at 75°C for 20min to inactive T4 DNA polymerase
31. Add 1μl NlaIII (NEB, # R0125L).
32. Mix and incubate at 37°C for 1h.
33. Add 0.9x (45μl) volume SPRI beads (BECKMAN, # B23319) to the sample.

34. Mix and incubate at room temperature for 5min.
35. Put the tubes on the magnet rack separator. Discard supernatant after the solution is clear.
36. Wash pellet twice by 80% Ethanol.
37. Elute DNA by 80μl 10mM Tris-HCl (pH8.0).
38. Measure DNA concentration by Qubit.

- **Circularization (2.5h)**

39. Prepare circularization mixture, adjust the DNA concentration to 1ng/μl in molecular biology water.

Stock	Volume (μl)
DNA in water (1ng/μl)	469
10x T4 Ligation Buffer	25
T4 DNA ligase 400U/μl	6

40. Mix and incubate at room temperature for 2h or overnight at 16°C.
41. Purify DNA by DNA clean & concentrator kit (Zymo, #D4013), elute DNA in 16μl H_2O.

- **PmeI digestion and library amplification PCR (3h)**

42. Prepare PmeI digestion mixture. PmeI will digest the Tn5 adaptor and linearize the circularized DNA.

Stock	Volume (μl)
DNA	14
10x CutSmart buffer	1.7
PmeI (NEB, R0560L)	1.3

43. Mix by vortex and incubate at 37°C for 1h.
44. Prepare PCR mixture

Stock	Volume (μl)
DNA	16
5x Q5 buffer	16
dNTP (10mM)	1.6
Nextera-pcr-i7-10-L (Table S1)	1.6
NEB primer i501 (Table S1)	1.6
Q5 Polymerase (NEB, # m0491L)	0.8
H_2O	58.4

45. Mix by vortex and perform PCR program as following

Temperature (°C)	Time (sec)	
72	240	
98	30	
98	10	12 cycles
59	30	
72	45	
72	300	
4	forever	

46. Purify library DNA (300-750bp) using DNA recovery kit (Zymo, #D4002) or SPRI beads, following manufacture's instruction. This HiCAR library is ready for next generation sequencing. We recommend to sequencing each HiCAR library to 200-500 million reads, paired-end, 100 cycles. If the sequencing lane only contains HiCAR libraries, need to do 3 chemistry-only cycles (no imaging occurs) in the first 3bp of the R1 sequencing.