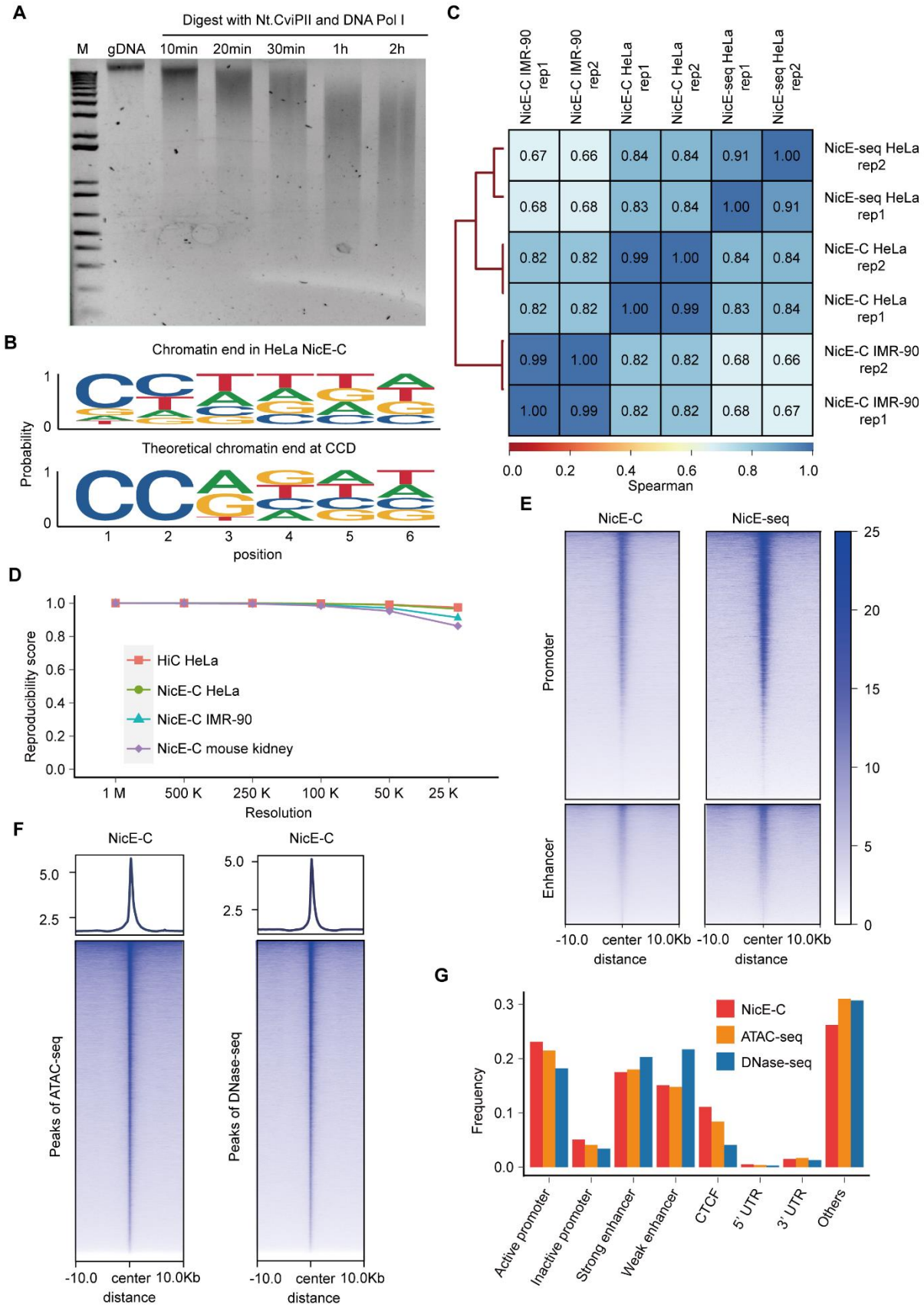
Supplementary Material

NicE-C efficiently reveals open chromatin-associated chromosome interactions at high resolution

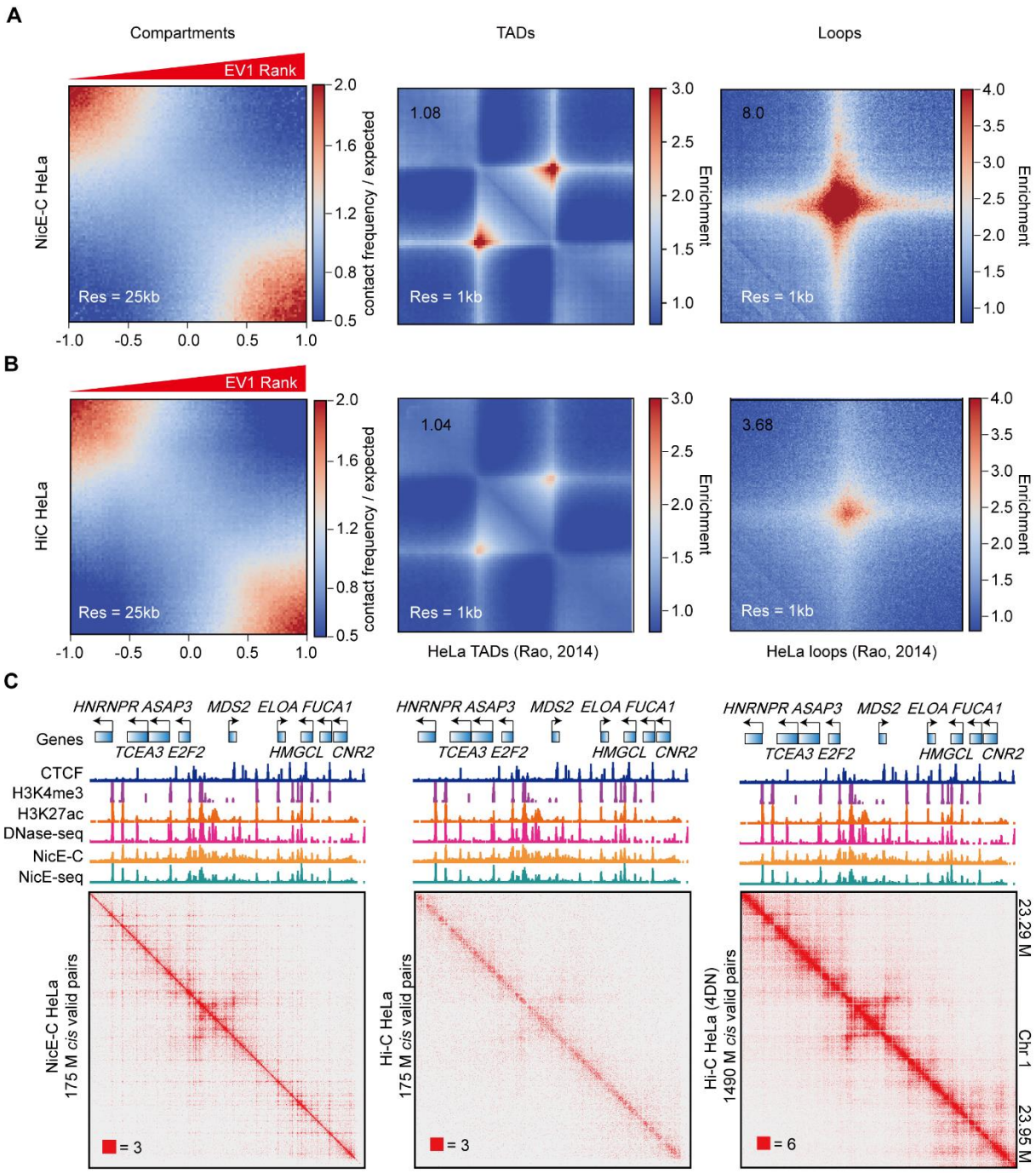
Zhengyu Luo[1, 6], Ran Zhang[2, 6], Tengfei Hu[1, 6], Yuting Zhu[1], Yueming Wu[1], Wenfei Li[3], Zhi Zhang[4], Xuebiao Yao[5], Haiyi Liang[2], Xiaoyuan Song[1]

[1]MOE Key Laboratory for Cellular Dynamics, CAS Key Laboratory of Brain Function and Disease, School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China
[2]CAS Key Laboratory of Mechanical Behavior and Design of Materials, Department of Modern Mechanics, University of Science and Technology of China, Hefei, China
[3]Affiliated Psychological Hospital of Anhui Medical University, Hefei Fourth People's Hospital, Anhui Mental Health Center
[4]CAS Key Laboratory of Brain Function and Disease, School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China
[5]MOE Key Laboratory for Cellular Dynamics, Hefei, China, Anhui Key Laboratory for Cellular Dynamics and Chemical Biology, Hefei, China 230027
[6]These authors contributed equally
Corresponding authors: songxy5@ustc.edu.cn

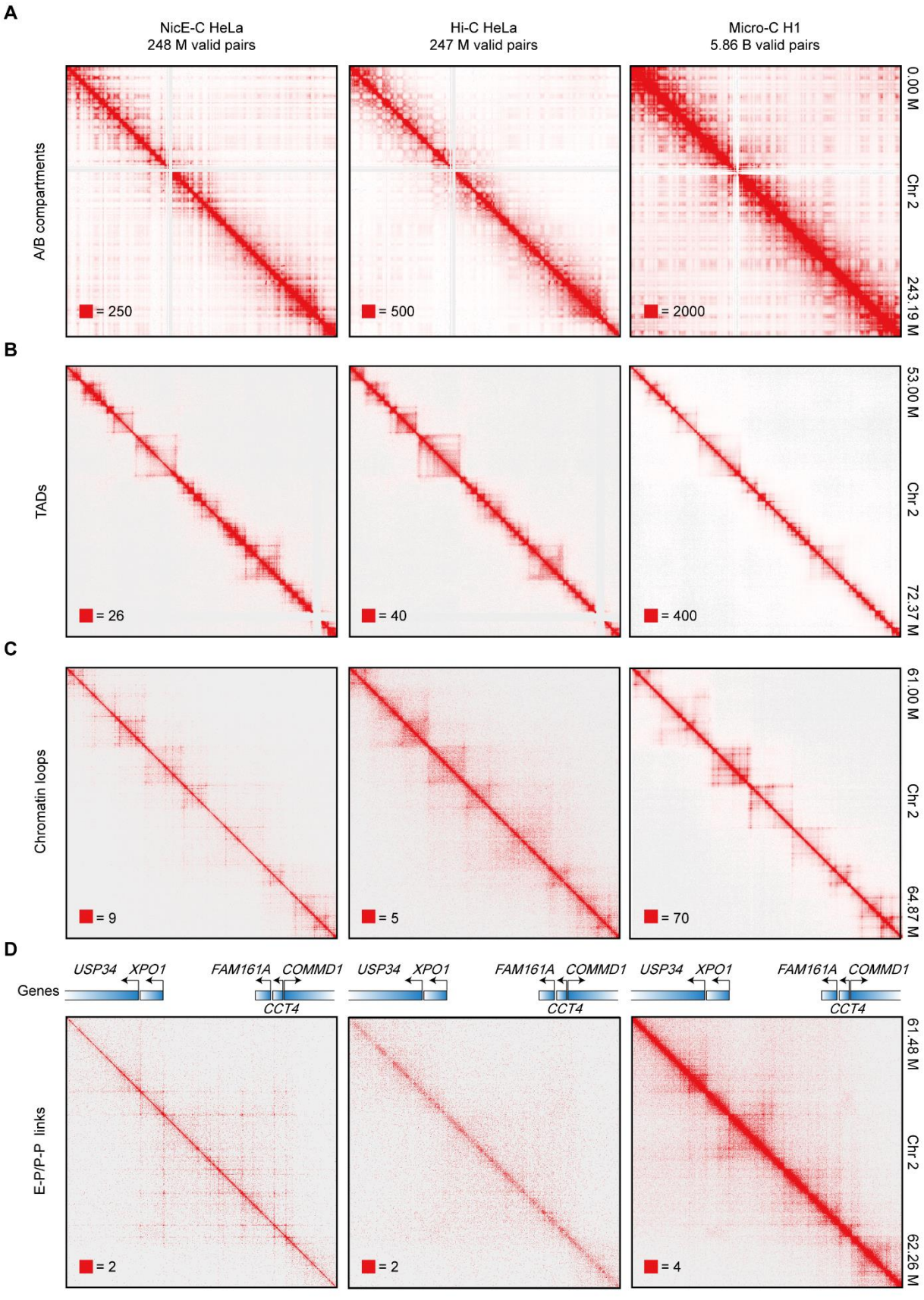
TABLE OF CONTENTS


A
Digest with Nt.CviPII and DNA Pol I
M
gDNA
10min
20min
30min
1h
2h
B
Chromatin end in HeLa NicE-C
Probability
Theoretical chromatin end at CCD
position
C
NicE-C IMR-90 rep1
NicE-C IMR-90 rep2
NicE-C HeLa rep1
NicE-C HeLa rep2
NicE-seq HeLa rep1
NicE-seq HeLa rep2
0.67 0.66 0.84 0.84 0.91 1.00
0.68 0.68 0.83 0.84 1.00 0.91
0.82 0.82 0.99 1.00 0.84 0.84
0.82 0.82 1.00 0.99 0.83 0.84
0.99 1.00 0.82 0.82 0.68 0.66
1.00 0.99 0.82 0.82 0.68 0.67
Spearman
D
Reproducibility score
HiC HeLa
NicE-C HeLa
NicE-C IMR-90
NicE-C mouse kidney
Resolution
E
NicE-C
NicE-seq
Promoter
Enhancer
distance
F
NicE-C
Peaks of ATAC-seq
Peaks of DNase-seq
distance
G
Frequency
NicE-C
ATAC-seq
DNase-seq
Active promoter
Inactive promoter
Strong enhancer
Weak enhancer
CTCF
5'UTR
3'UTR
Others

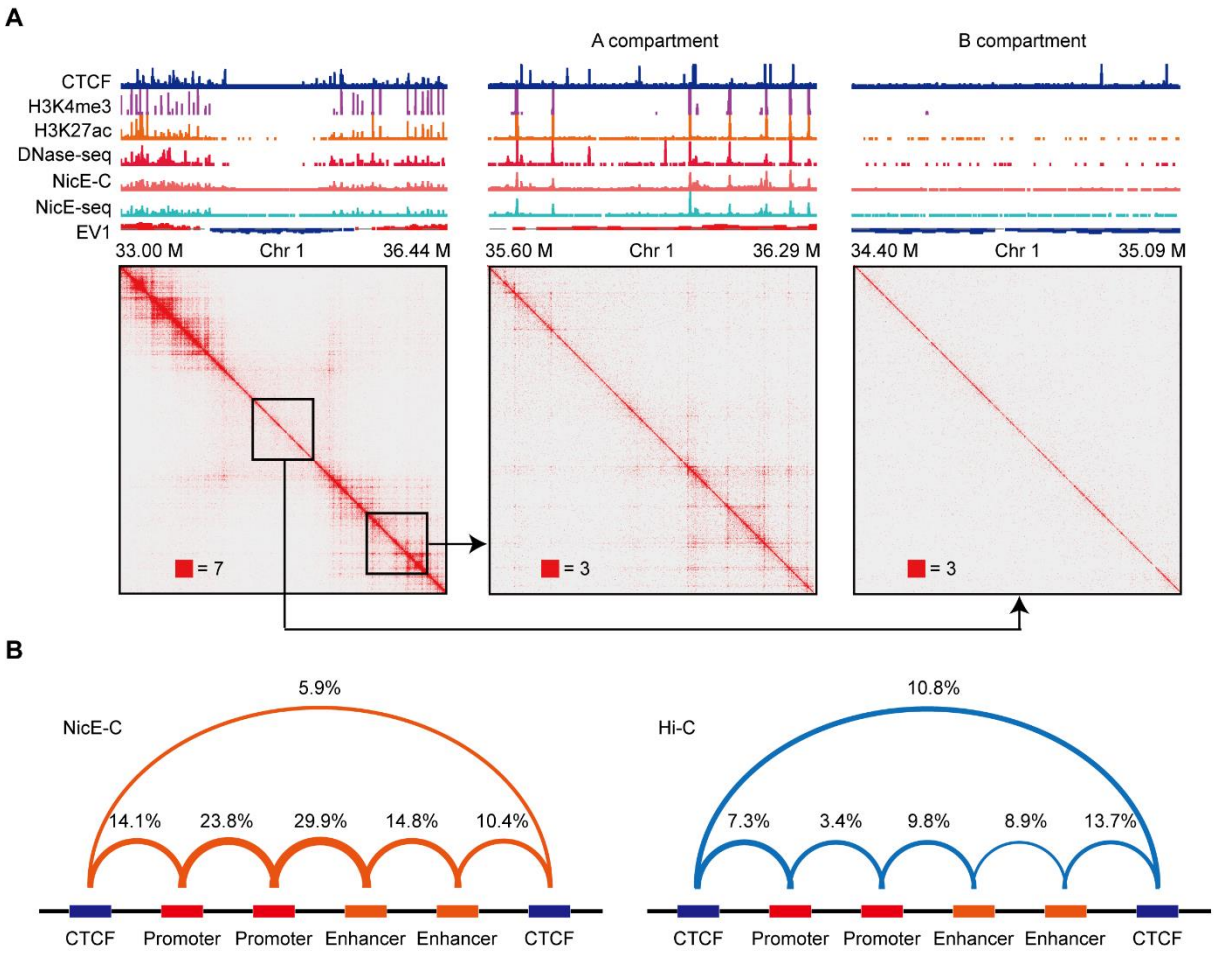
Supplemental Figure S1. Chromatin fragmentation, reproducibility analysis, and open chromatin enrichment of the NicE-C method. (A) An agarose gel showing chromatin fragmentation of crosslinked HeLa cells in NicE-C digestion reaction with the indicated reaction times. (B) The sequences of chromatin ends identified from HeLa NicE-C sequencing data (top), and the theoretical sequences of chromatin ends of hg19 genome when the chromatin ends generated by NicE-C was CCD (D=A/G/T) (bottom). (C) Spearman's correlation between open chromatin part of NicE-C and NicE-seq data. (D) Reproducibility analysis of NicE-C replicates (Hi-C part), reproducibility scores were calculated by HiC-Rep at different resolutions. (E) Heatmaps of HeLa NicE-C and NicE-seq signal enrichment at promoter and enhancer regions. (F) A heatmap showing the enrichment of HeLa NicE-C data around HeLa ATAC-seq and DNase-seq open chromatin regions (data from PMID 29731168 and PMID 21854988). (G) The distribution of HeLa NicE-C, ATAC-seq, and DNase-seq open chromatin peaks in the indicated types of genomic regions.


A
Compartments
TADs
Loops
EV1 Rank
NicE-C HeLa
Res = 25kb
contact frequency / expected
1.08
Res = 1kb
Enrichment
8.0
Res = 1kb
Enrichment
B
EV1 Rank
HiC HeLa
Res = 25kb
contact frequency / expected
1.04
Res = 1kb
Enrichment
3.68
Res = 1kb
Enrichment
HeLa TADs (Rao, 2014)
HeLa loops (Rao, 2014)
C
HNRNPR ASAP3 MDS2 ELOA FUCA1
Genes
TCEA3 E2F2 HMGCL CNR2
CTCF
H3K4me3
H3K27ac
DNase-seq
NicE-C
NicE-seq
NicE-C HeLa
175 M cis valid pairs
= 3
Hi-C HeLa
175 M cis valid pairs
= 3
Hi-C HeLa (4DN)
1490 M cis valid pairs
= 6
23.29 M
Chr 1
23.95 M

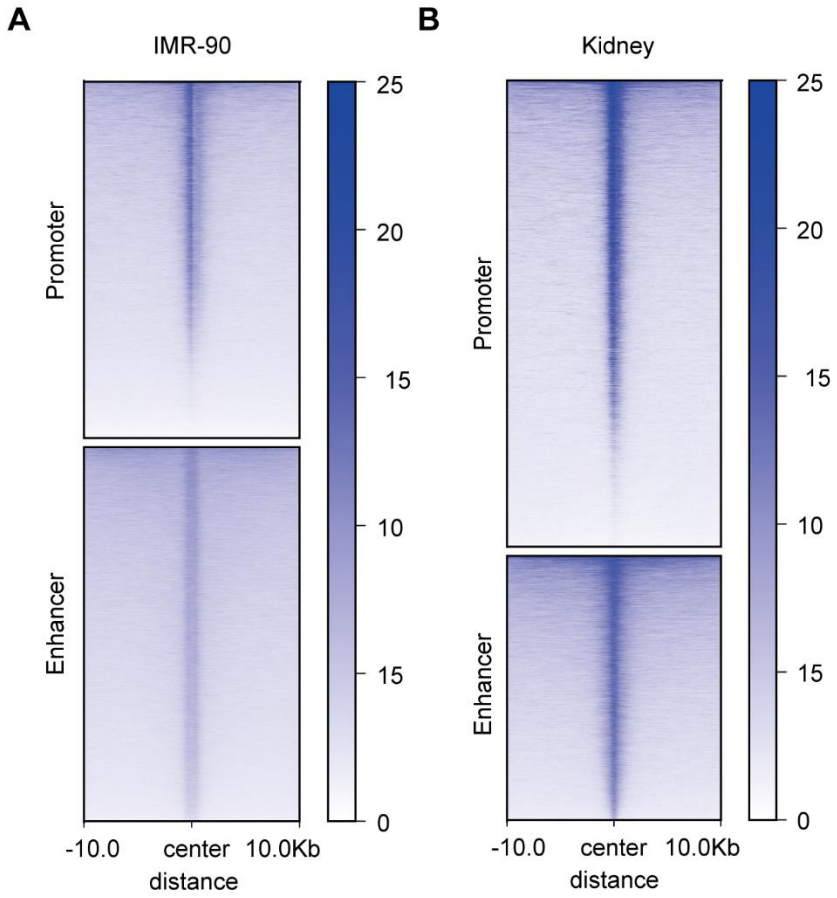
Supplemental Figure S2. Overview of chromatin organization in NicE-C and *in situ* Hi-C data. (A, B) Overview of chromatin organization in NicE-C and *in situ* Hi-C data. A/B compartments, TADs, and loops revealed by HeLa NicE-C data (A, 175M *cis* pairs) and *in situ* Hi-C data (B, 175M *cis* pairs). Left: Saddle plot of compartmentalization strength. The upper-left and bottom-right represent the contact frequency between B-B and A-A compartments, and the upper-right and bottom-left show the contact frequency between A and B compartments. Middle: Pile-up analysis of TADs (published HeLa TADs from Rao et al, 2014 were used) with NicE-C data. Right: Pile-up analysis of loop enrichment (published HeLa chromatin loops from Rao et al, 2014 were used) with NicE-C data. (C) HeLa NicE-C and Hi-C chromatin contact maps of an example region on chromosome 1 at 1-kb resolution. Left: NicE-C data with 175 M *cis* pairs. Middle: down-sampled Hi-C data with 175 M *cis* pairs. Right: Hi-C data with about 1, 490 M *cis* pairs (High-resolution Hi-C datasets from 4DN: 4DNESCMX7L58). Snapshots of 1D chromatin tracks (ChIP-seq of CTCF, H3K4me3 and H3K27ac, DNase-seq (see Methods for references), open chromatin signal of NicE-C and NicE-seq) in this region are also shown. Numbers below the interaction maps correspond to the maximum signal in the matrix.


A
NicE-C HeLa
248 M valid pairs
Hi-C HeLa
247 M valid pairs
Micro-C H1
5.86 B valid pairs
A/B compartments
= 250
= 500
= 2000
0.00 M
Chr 2
243.19 M
B
TADs
= 26
= 40
= 400
53.00 M
Chr 2
72.37 M
C
Chromatin loops
= 9
= 5
= 70
61.00 M
Chr 2
64.87 M
D
Genes
USP34
XPO1
FAM161A
COMMD1
CCT4
E-P/P-P links
= 2
= 2
= 4
61.48 M
Chr 2
62.26 M

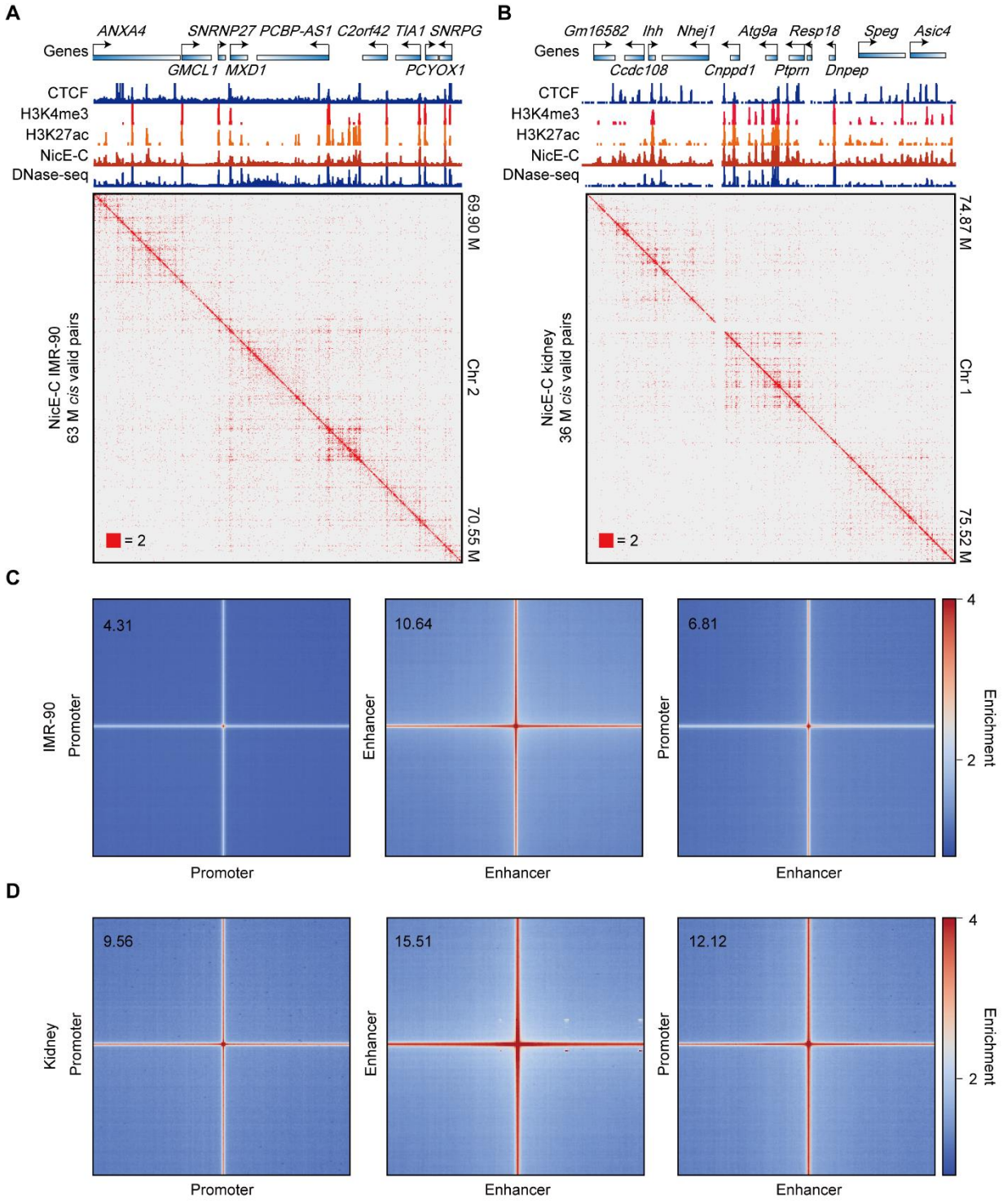
Supplemental Figure S3. Comparison of NicE-C, Hi-C, and Micro-C data. (A-C) Comparison of chromatin contact heatmaps of NicE-C (HeLa cells), Hi-C (HeLa cells), and Micro-C data (H1 cells, Nils Krietenstein *et al*, 2020.) at different resolutions. NicE-C recapitulated the chromatin structures including A/B compartments (A, whole Chromosome 2), TADs (B, 25-kb resolution), and chromatin loops (C, 5-kb resolution). (D) When zooming into a region on Chromosome 2 at 1-kb resolution, NicE-C could detect enhancer-promoter (E-P) or promoter-promoter (P-P) loops and stripes similar to Micro-C data, while Hi-C data could not provide detailed information at this scale.



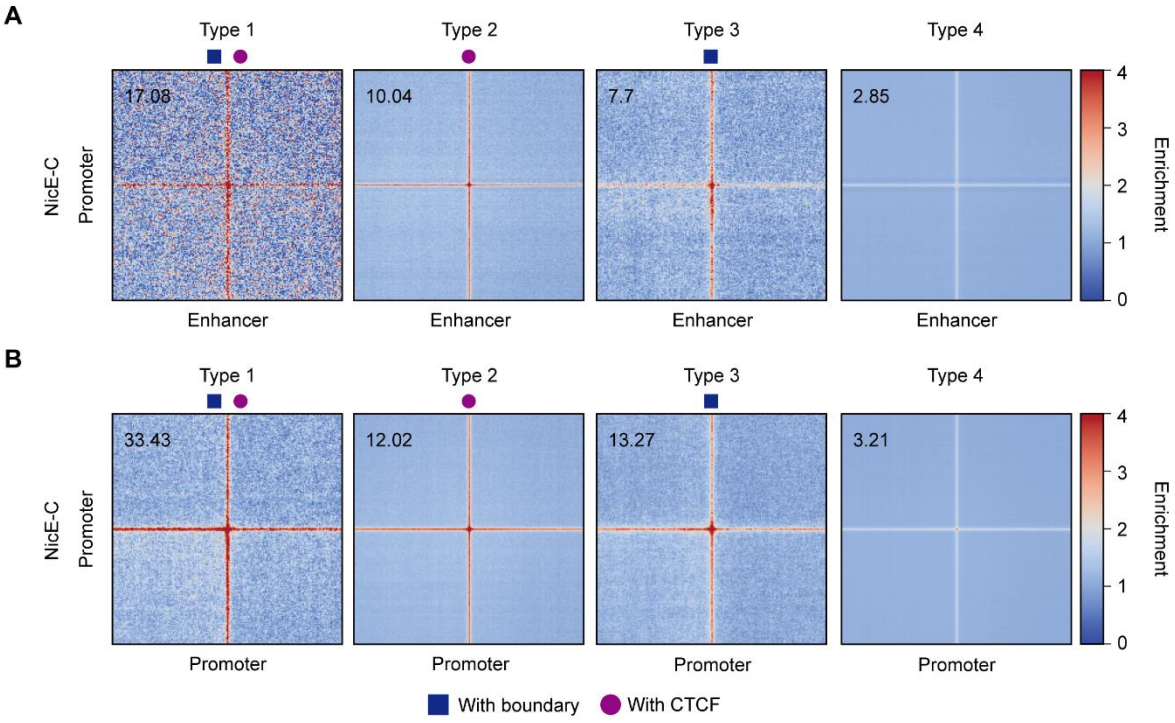
Supplemental Figure S4. NicE-C generally occupies open chromatin associated interactions. (A) Left: Example regions of HeLa NicE-C chromatin contact maps on Chromosome 1 at 5-kb resolution. Active chromatin markers and NicE-C chromatin interactions are enriched in compartment A, but are depleted in compartment B. Snapshots of one dimensional (1D) chromatin tracks (ChIP-seq of CTCF, H3K4me3 and H3K27ac, open chromatin signals of NicE-C and DNase-seq, the first eigenvector (EV1) of compartments) in this region are also shown. (B) Summaries of the major classes of high confidence interactions identified with NicE-C and *in situ* Hi-C (Rao et al, 2014) in HeLa cells.



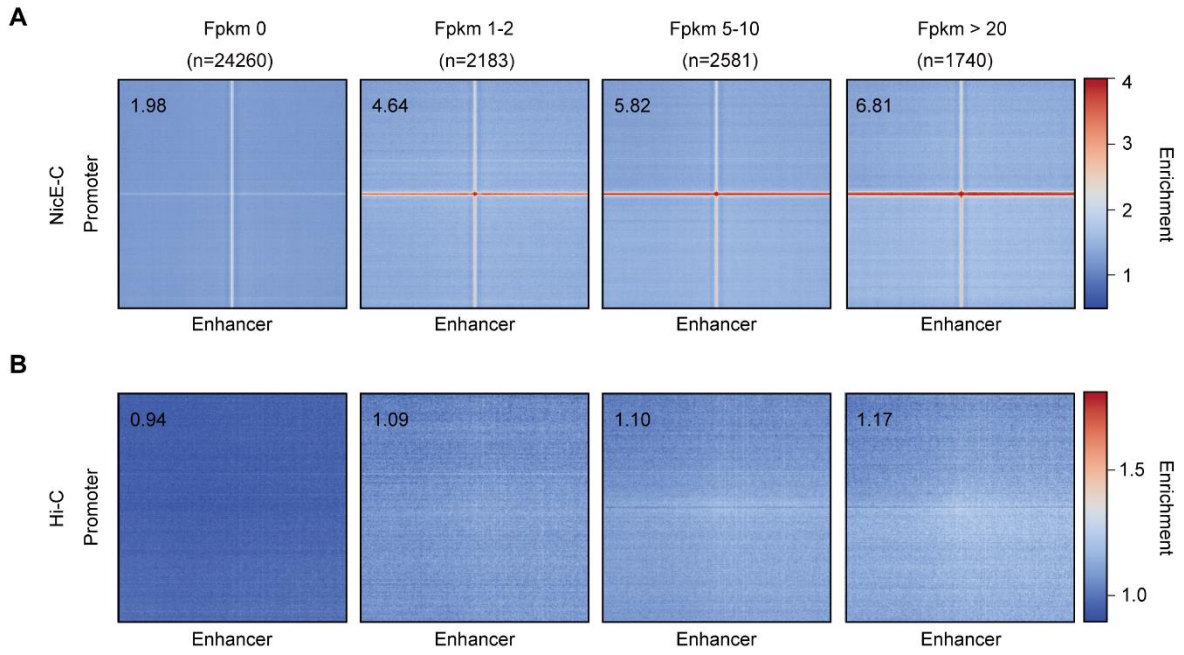
Supplemental Figure S5. NicE-C peaks enriched at promoter and enhancer regions. (A) Heatmaps show signal enrichment at promoter and enhancer regions of IMR-90 cells NicE-C data. (B) Heatmaps show signal enrichment at promoter and enhancer regions of adult (6 months) female mouse kidney cells NicE-C data.


A
ANXA4
SNRNP27
PCBP-AS1
C2orf42
TIA1
SNRPG
Genes
GMCL1
MXD1
PCYOX1
CTCF
H3K4me3
H3K27ac
NicE-C
DNase-seq
NicE-C IMR-90
63 M cis valid pairs
69.90 M
Chr 2
70.55 M
= 2
B
Gm16582
Ihh
Nhej1
Atg9a
Resp18
Speg
Asic4
Genes
Ccdc108
Cnppd1
Ptprn
Dnpep
CTCF
H3K4me3
H3K27ac
NicE-C
DNase-seq
NicE-C kidney
36 M cis valid pairs
74.87 M
Chr 1
75.52 M
= 2
C
IMR-90
Promoter
4.31
10.64
6.81
Promoter
Enhancer
Promoter
Enhancer
Enhancer
4
2
Enrichment
D
Kidney
Promoter
9.56
15.51
12.12
Promoter
Enhancer
Promoter
Enhancer
Enhancer
4
2
Enrichment

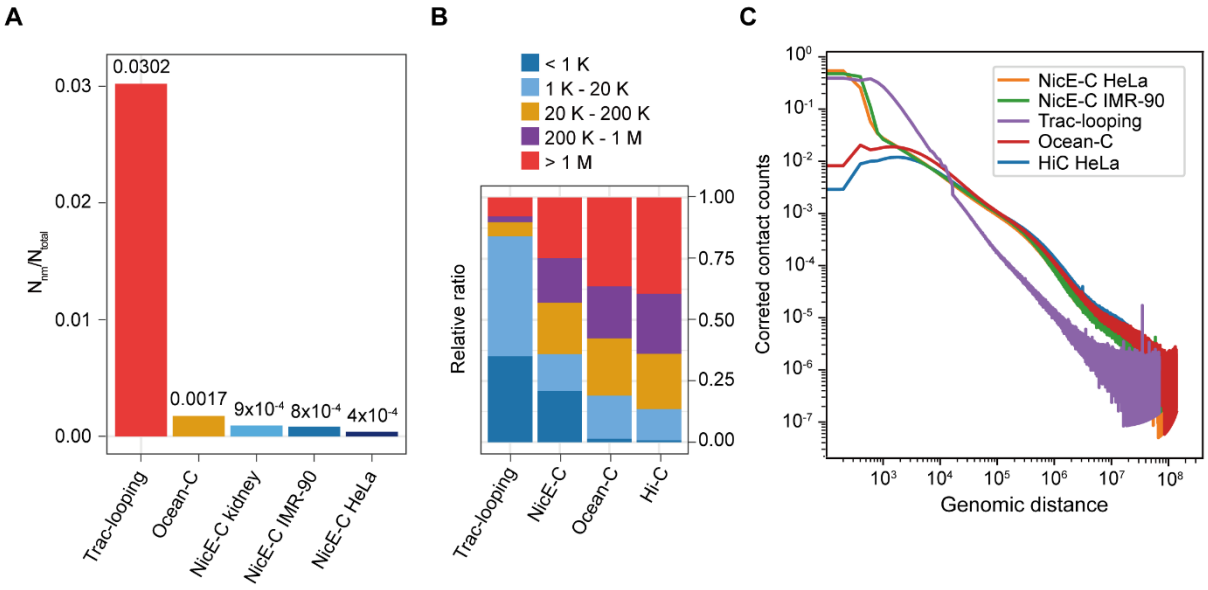
Supplemental Figure S6. NicE-C profiles open chromatin and chromatin interactions in IMR-90 cells and mouse kidney cells. (A) IMR-90 NicE-C chromatin contact maps of an example region on Chromosome 2 at 1-kb resolution. Snapshots of one dimension (1D) chromatin tracks (ChIP-seq of CTCF, H3K4me3 and H3K27ac, open chromatin signals of NicE-C and DNase-seq) in this region are also shown. (B) Mouse kidney cells NicE-C chromatin contact maps of an example region on Chromosome 2 at 1-kb resolution. Snapshots of 1D chromatin tracks (ChIP-seq of CTCF, H3K4me3 and H3K27ac, open chromatin signals of NicE-C and DNase-seq) in this region are also shown. Numbers below the interaction maps in A and B correspond to maximum signal in the matrix. (C) Genome-wide averaged pile-up matrices of NicE-C (IMR-90 cells) P-P, E-E (enhancer-enhancer) and E-P loops and stripes were plotted at 1-kb resolution (windows=200 kb). (D) Genome-wide averaged pile-up matrices of NicE-C (mouse kidney cells) P-P, E-P, and E-E loops and stripes were plotted at 1-kb resolution (windows=200 kb).



Supplemental Figure S7. Promoter-enhancer interactions are enhanced by CTCF and TAD boundary. (A) Genome-wide averaged pile-up matrices of loops and stripes between different types of promoters and enhancers (E-P) in HeLa cells NicE-C data (175 M *cis* valid pairs) were plotted at 1-kb resolution (windows=200 kb). Type 1 promoters and enhancers were overlapped with CTCF bound TAD boundaries, type 2 promoters and enhancers were overlapped with CTCF but not TAD boundaries, type 3 promoters and enhancers were overlapped with TAD boundaries without CTCF binding, type 4 promoters and enhancers were overlapped with neither CTCF nor TAD boundaries. (B) Genome-wide averaged pile-up matrices of loops and stripes between different types of promoters (P-P) in HeLa cells Hi-C data (175 M *cis* valid pairs) were plotted at 1-kb resolution (windows=200 kb). Different types of promoters were same as described in A.



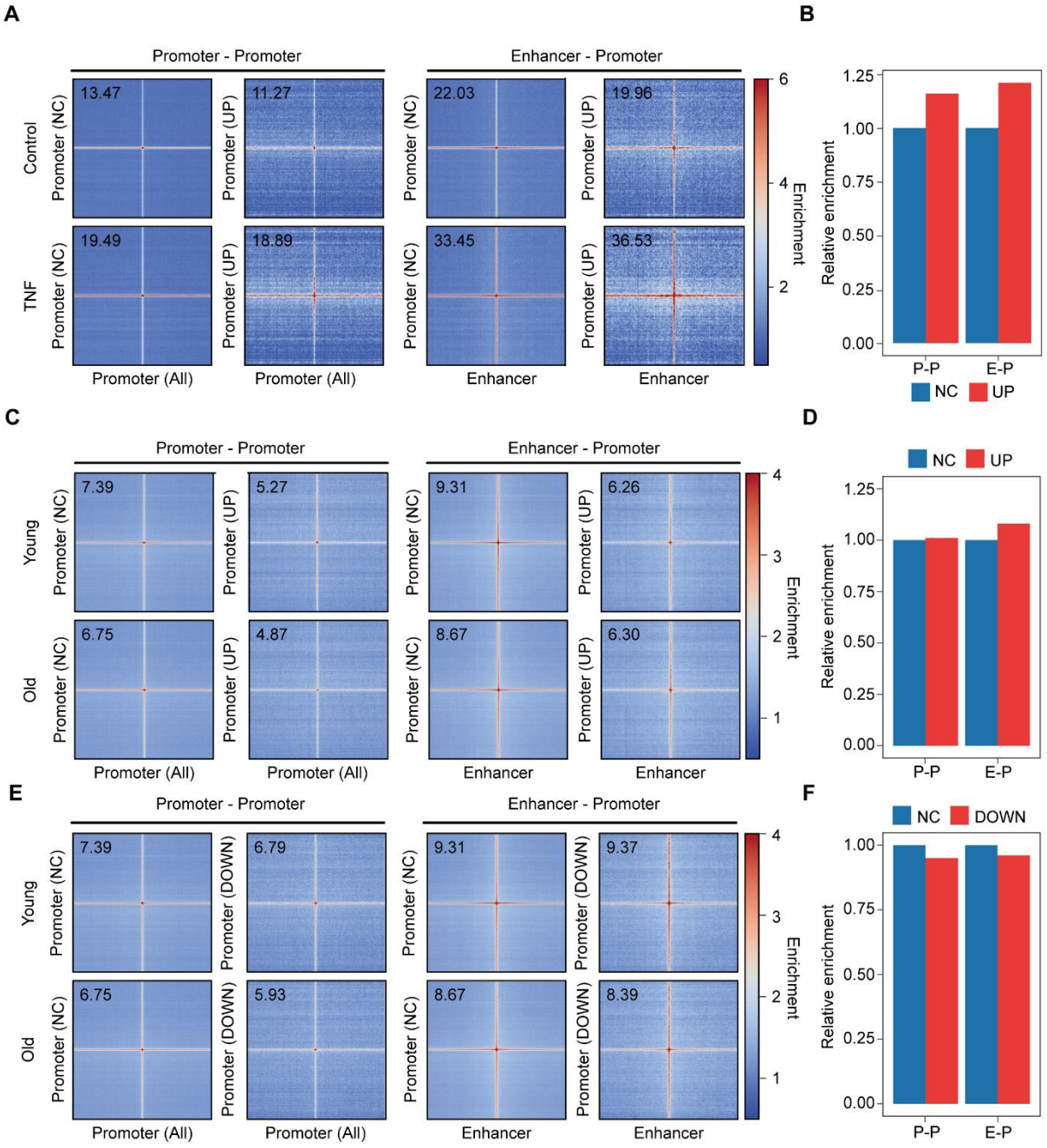
Supplemental Figure S8. Transcription-associated promoter-enhancer interactions identified by NicE-C. (A) Genome-wide averaged pile-up matrices of loops and stripes between promoters (of genes with different expression levels) and enhancers in HeLa cells NicE-C data (175 M *cis* valid pairs) were plotted at 1-kb resolution (windows=200 kb). (B) Genome-wide averaged pile-up matrices of loops and stripes between promoters (of genes with different expression levels) and enhancers in HeLa cells Hi-C data (175 M *cis* valid pairs) were plotted at 1-kb resolution (windows=200 kb).



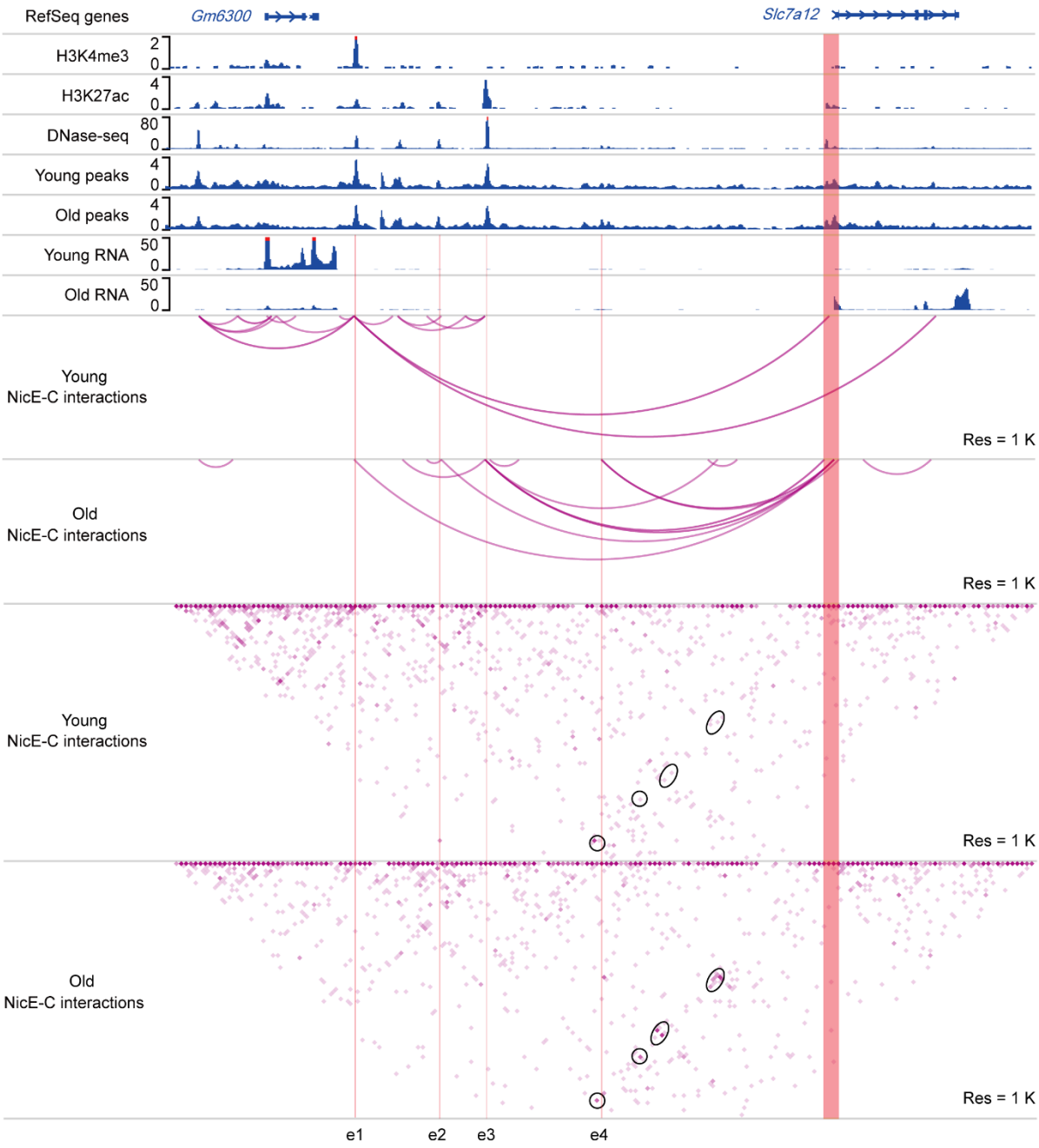
D

Statistic	HeLa NicE-C This study	IMR-90 NicE-C This study	HeLa Hi-C This study	GM12878 Ocean-C Li et al. 2018	CD4+ Trac-looping Lai et al. 2018
Cell number	1M	1M	1M	2-5M	100M
Valid Pairs	248M	85M	247M	274M	86M
Trans Pairs(%)	28%	21%	23%	21%	26%
Cis Pairs(%)	72%	79%	77%	79%	74%
Cis long-range pairs	63.9% (>20kb)	64.6% (>20kb)	86.4% (>20kb)	80.9% (>20kb)	15.7% (>20kb)
Cis short-range pairs	36.1%	35.4%	13.6%	19.1%	84.3%
Processing time	2 Days	2 Days	2 Days	3-4 Days	4-5 Days

Supplemental Figure S9. Comparison of NicE-C, OCEAN-C, and Trac-looping data. (A) The N_{nm} (reads consisting of genomic DNA and mitochondrial DNA)/N_{total} (total reads) of Trac-looping, OCEAN-C, and NicE-C data. (B) Distribution of *cis* valid pairs of Trac-looping, NicE-C, OCEAN-C, and Hi-C data. (C) The chromatin contact probability relative to genomic distance measured for NicE-C (HeLa cells), NicE-C (IMR-90 cells), Trac-looping ($CD4^+$ T cells), OCEAN-C (GM12878 cells), and Hi-C (HeLa cells) are shown. (D) Comparison of the data output of NicE-C, Hi-C, Ocean-C, and Trac-looping.


A
Promoter - Promoter
Enhancer - Promoter
Control
TNF
Promoter (NC)
Promoter (UP)
13.47
11.27
22.03
19.96
19.49
18.89
33.45
36.53
Promoter (All)
Enhancer
Enrichment
6
4
2
B
Relative enrichment
1.25
1.00
0.75
0.50
0.25
0.00
P-P
E-P
NC
UP
C
Young
Old
7.39
5.27
9.31
6.26
6.75
4.87
8.67
6.30
D
E
Promoter (DOWN)
6.79
9.37
5.93
8.39
F
DOWN

Supplemental Figure 10. Dynamic enhancer-promoter interactions detected by NicE-C. (A) Genome-wide averaged pile-up matrices (plotted at 1-kb resolution, windows = 200 kb) of HeLa-S3 enhancer-promoter (E-P) loops and stripes for no changed genes (NC) and up-regulated genes (UP) after TNF stimulation. (B) The plot showed the relative enrichment of genome-wide chromatin interactions formed between the promoters of up-regulated genes and other promoters or enhancers, enrichment data were from the pile-up plot in A. (C) Genome-wide averaged pile-up matrices (plotted at 1-kb resolution, windows=200 kb) of young (2 months) and old (18 months) mouse kidney cells NicE-C data. Loops and stripes between promoters of 310 up-regulated genes (log2 fold change > 2, $p < 0.05$) in old kidney and other promoters or enhancers were plotted. We also selected 1,838 genes with almost no expression changes as a control (log2 fold change between -0.1 to 0.1). (D) The plot showed the relative enrichment of genome-wide chromatin interactions formed between the promoters of up-regulated genes and other promoters or enhancers, enrichment data were from the pile-up plot in C. (E) Genome-wide averaged pile-up matrices (plotted at 1-kb resolution, windows=200 kb) of young (2 months) and old (18 months) mouse kidney cells NicE-C data. Loops and stripes between promoters of 257 down-regulated genes (log2 fold change < -2, $p < 0.05$) in old kidney and other promoters or enhancers were plotted. 1,838 genes with almost no expression changes were used as a control (log2 fold change between -0.1 to 0.1). (F) The plot showed the relative enrichment of genome-wide chromatin interactions formed between the promoters of down-regulated genes and other promoters or enhancers, enrichment data were from the pile-up plot in E.



Supplemental Figure 11. An example of increased E-P loops in old mouse kidney cells compared to those in young mouse kidney cells around the *Slc7a12* locus (up-regulated gene in old mouse kidney cells). Red arrows pointing to e1 to e4 show the putative enhancers based on ChIP-seq, DNase-seq, and our NicE-C peaks. Region in orange represents the gene promoter. Ovals indicate examples of increased interactions associated with indicated gene promoter in old kidney compared to young kidney.