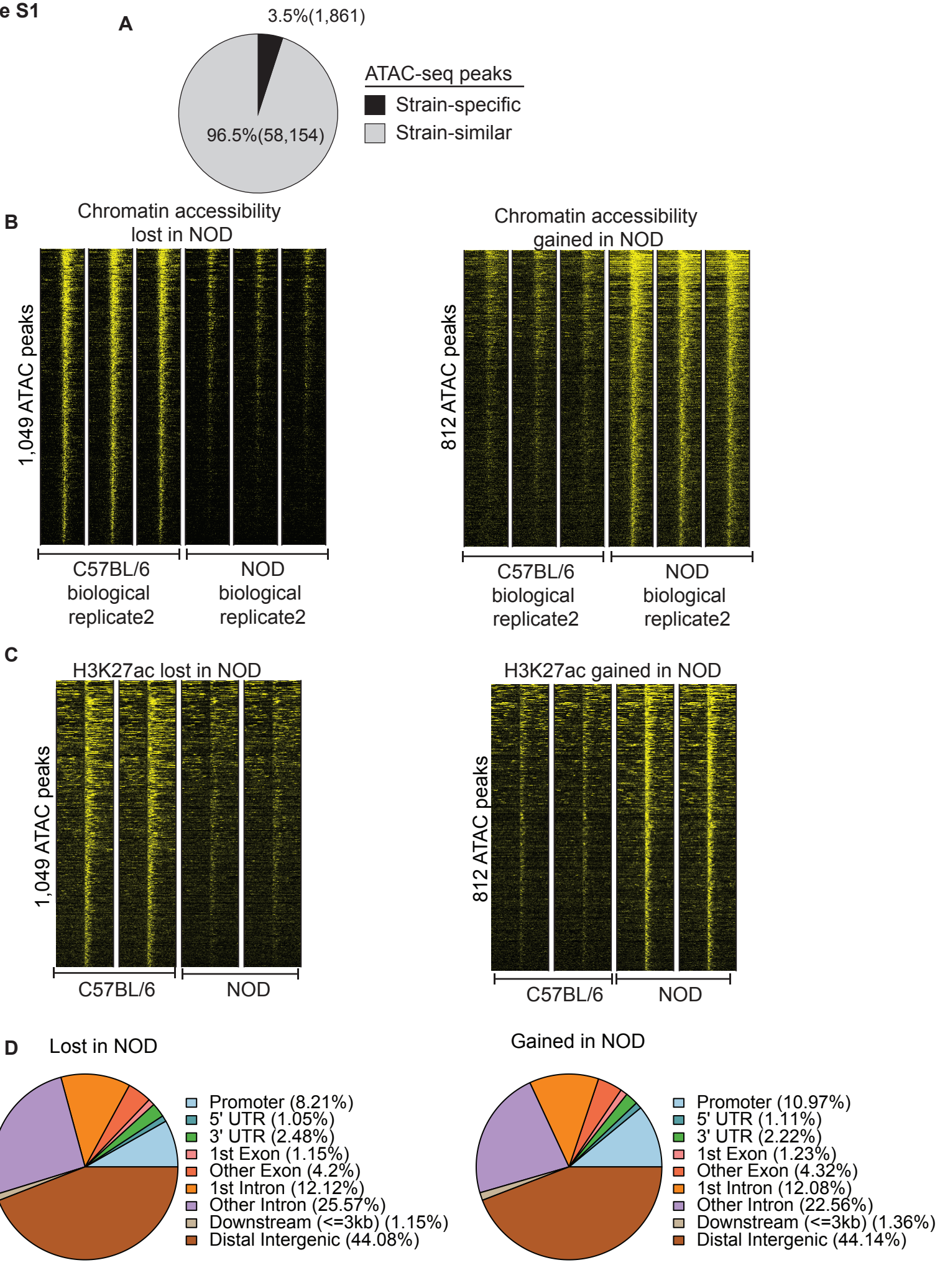
Figure S1


A
3.5%(1,861)
96.5%(58,154)
ATAC-seq peaks
Strain-specific
Strain-similar
B
Chromatin accessibility lost in NOD
1,049 ATAC peaks
C57BL/6 biological replicate2
NOD biological replicate2
Chromatin accessibility gained in NOD
812 ATAC peaks
C57BL/6 biological replicate2
NOD biological replicate2
C
H3K27ac lost in NOD
1,049 ATAC peaks
C57BL/6
NOD
H3K27ac gained in NOD
812 ATAC peaks
C57BL/6
NOD
D
Lost in NOD
Promoter (8.21%)
5' UTR (1.05%)
3' UTR (2.48%)
1st Exon (1.15%)
Other Exon (4.2%)
1st Intron (12.12%)
Other Intron (25.57%)
Downstream (<=3kb) (1.15%)
Distal Intergenic (44.08%)
Gained in NOD
Promoter (10.97%)
5' UTR (1.11%)
3' UTR (2.22%)
1st Exon (1.23%)
Other Exon (4.32%)
1st Intron (12.08%)
Other Intron (22.56%)
Downstream (<=3kb) (1.36%)
Distal Intergenic (44.14%)

Supplementary Figure 1. Genomic features of strain-specific open chromatin regions (related to Figure 1).

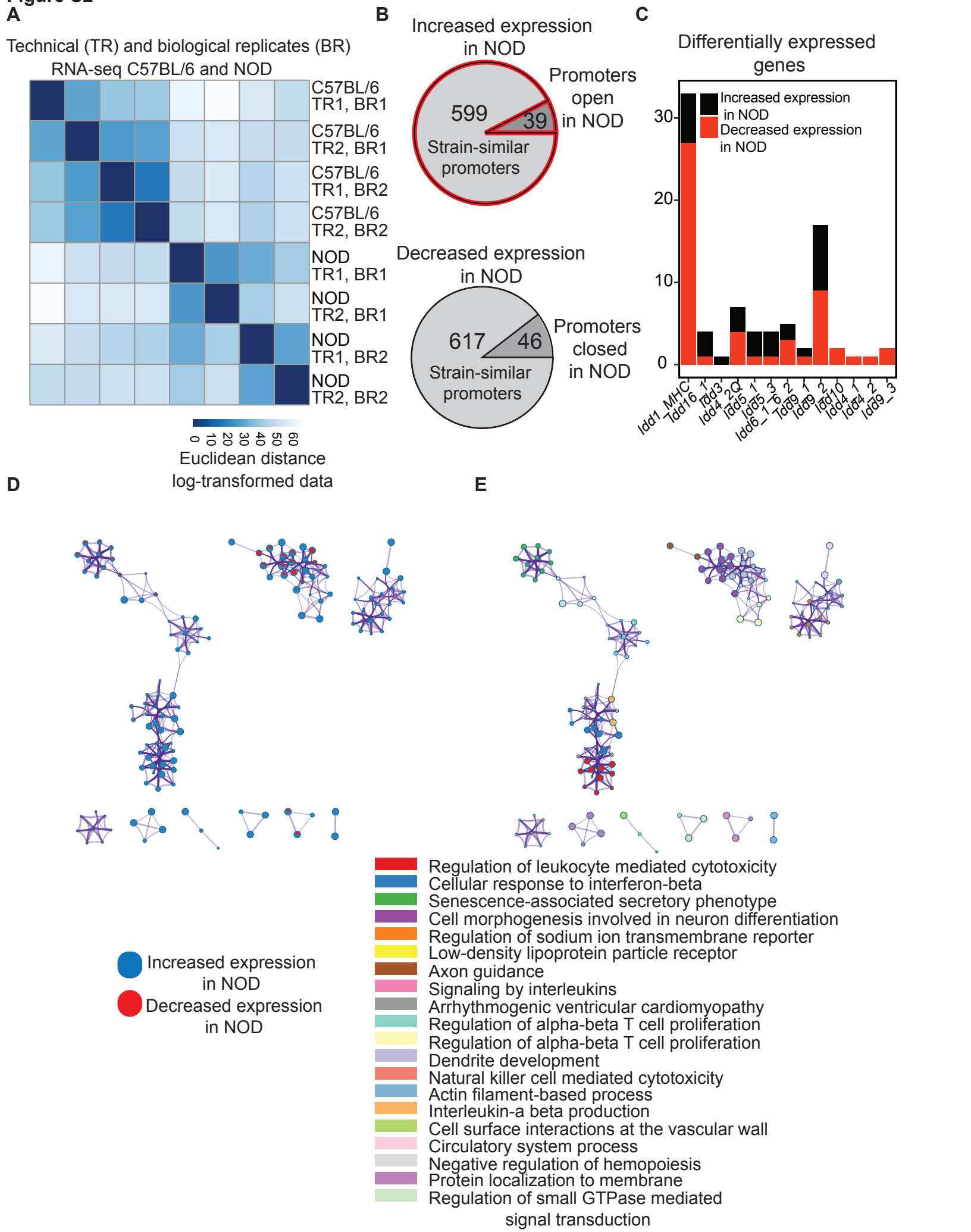
(A) Pie chart depicts the percentage of strain-specific regulatory elements measured by ATAC-seq. Three technical and two biological replicates were generated for each strain. Together, we identified 60,015 open chromatin regions in T cells, with 96% of them (58,154) demonstrating similar levels of accessibility between the two genetically distinct animals (Figure S1A). Using $\log_2$FoldChange>1 and *padj*<1e-4 threshold, we found 1,049 accessible chromatin regions that were unique to T cells of C57BL/6 mice (referred to as 'lost in NOD') while 812 regulatory regions were unique to T cells of NOD mice (referred to as 'gained in NOD')

(B) Strain-specific ATAC-seq peaks are reproduced in biological replicates.

(C) Strain-specific ATAC-seq peaks overlap with strain-specific H3K27ac.

(D) The strain-specific regulatory elements were enriched at noncoding genomic regions. Promoters were defined as +/-1000bp of transcriptional start sites (TSS).

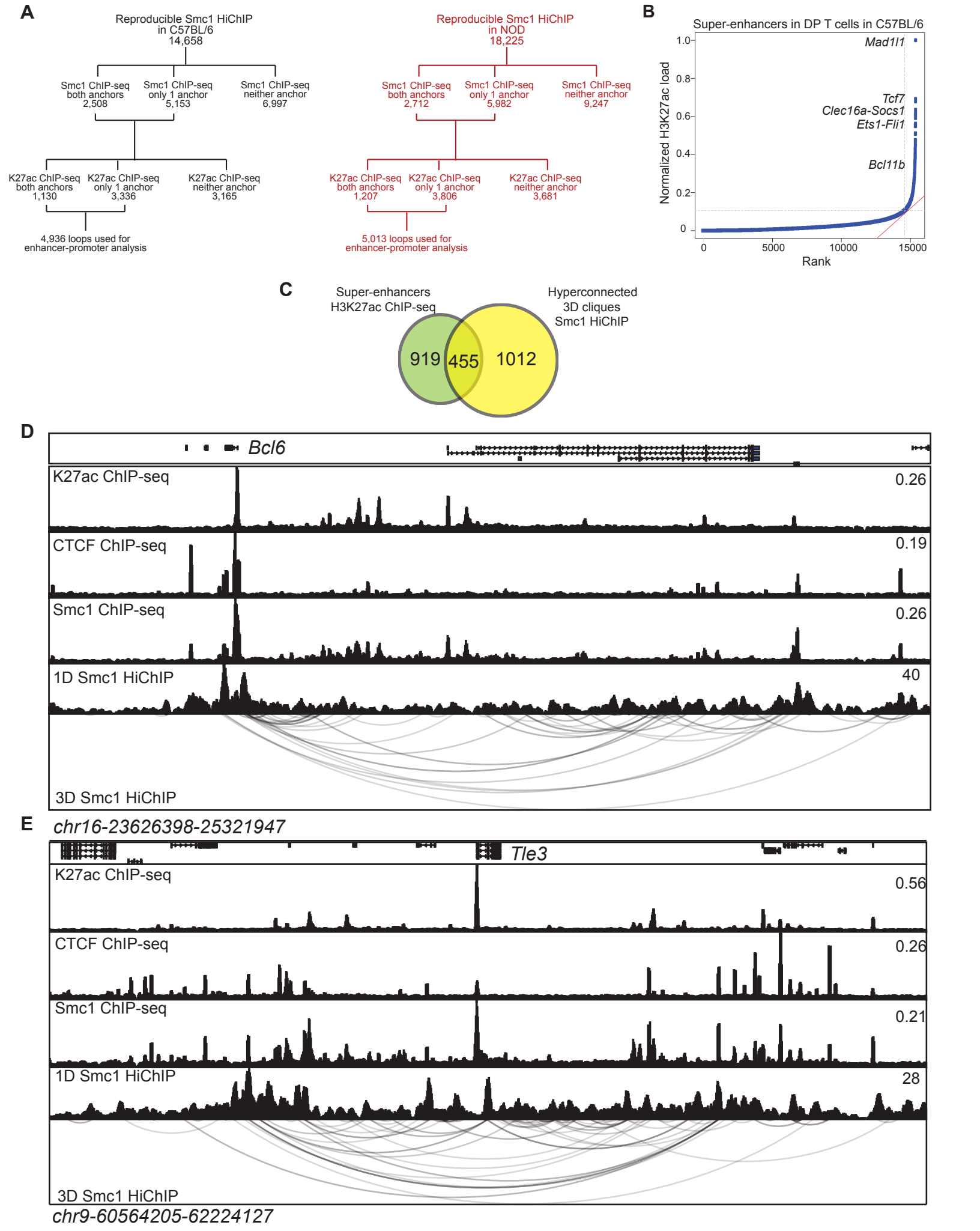
Figure S2


A
Technical (TR) and biological replicates (BR)
RNA-seq C57BL/6 and NOD
C57BL/6 TR1, BR1
C57BL/6 TR2, BR1
C57BL/6 TR1, BR2
C57BL/6 TR2, BR2
NOD TR1, BR1
NOD TR2, BR1
NOD TR1, BR2
NOD TR2, BR2
0 10 20 30 40 50 60
Euclidean distance
log-transformed data
B
Increased expression in NOD
599
39
Strain-similar promoters
Promoters open in NOD
Decreased expression in NOD
617
46
Strain-similar promoters
Promoters closed in NOD
C
Differentially expressed genes
Increased expression in NOD
Decreased expression in NOD
30
20
10
0
Idd1_MHC
Idd16_1
Idd3
Idd4_2Q
Idd5_1
Idd5_3
Idd5_1_6
Idd9_2
Idd9_1
Idd9_2
Idd10
Idd4_1
Idd4_2
Idd9_3
D
Increased expression in NOD
Decreased expression in NOD
E
Regulation of leukocyte mediated cytotoxicity
Cellular response to interferon-beta
Senescence-associated secretory phenotype
Cell morphogenesis involved in neuron differentiation
Regulation of sodium ion transmembrane reporter
Low-density lipoprotein particle receptor
Axon guidance
Signaling by interleukins
Arrhythmogenic ventricular cardiomyopathy
Regulation of alpha-beta T cell proliferation
Regulation of alpha-beta T cell proliferation
Dendrite development
Natural killer cell mediated cytotoxicity
Actin filament-based process
Interleukin-a beta production
Cell surface interactions at the vascular wall
Circulatory system process
Negative regulation of hemopoiesis
Protein localization to membrane
Regulation of small GTPase mediated signal transduction

Supplementary Figure 2. Increased expression of genes in T lymphocytes of NOD mice are associated with type 1 diabetes (related to Figure 2).

(A) Heatmap depicts euclidean distance for log transformed count data between technical and biological replicates for RNA-seq experiments in DP T cells of C57BL/6 and NOD mice.
(B) Pie charts depict differentially expressed genes with differentially accessible promoters between strains. Promoters are defined as +/-1000bp TSS.
(C) Barplot demonstrates the number of strain-specific genes within each *Idd* interval.
(D-E) Gene ontology analysis for differentially expressed genes using *metascape* suggests the 'regulation of leukocyte mediated cytotoxicity' associated genes to be enriched at genes with increased expression in NOD. While genes within each ontology is depicted as connected nodes in **(E)**, the same genes increased or decreased expression in NOD are shown in **(D)**.

Figure S3


A
Reproducible Smc1 HiChIP in C57BL/6 14,658
Smc1 ChIP-seq both anchors 2,508
Smc1 ChIP-seq only 1 anchor 5,153
Smc1 ChIP-seq neither anchor 6,997
K27ac ChIP-seq both anchors 1,130
K27ac ChIP-seq only 1 anchor 3,336
K27ac ChIP-seq neither anchor 3,165
4,936 loops used for enhancer-promoter analysis
Reproducible Smc1 HiChIP in NOD 18,225
Smc1 ChIP-seq both anchors 2,712
Smc1 ChIP-seq only 1 anchor 5,982
Smc1 ChIP-seq neither anchor 9,247
K27ac ChIP-seq both anchors 1,207
K27ac ChIP-seq only 1 anchor 3,806
K27ac ChIP-seq neither anchor 3,681
5,013 loops used for enhancer-promoter analysis
B
Super-enhancers in DP T cells in C57BL/6
Mad1l1
Tcf7
Clec16a-Socs1
Ets1-Fli1
Bcl11b
Normalized H3K27ac load
Rank
C
Super-enhancers H3K27ac ChIP-seq
Hyperconnected 3D cliques Smc1 HiChIP
919
455
1012
D
Bcl6
K27ac ChIP-seq
CTCF ChIP-seq
Smc1 ChIP-seq
1D Smc1 HiChIP
3D Smc1 HiChIP
E
chr16-23626398-25321947
Tle3
K27ac ChIP-seq
CTCF ChIP-seq
Smc1 ChIP-seq
1D Smc1 HiChIP
3D Smc1 HiChIP
chr9-60564205-62224127

Supplementary Figure 3. Genes with prominent roles in T cell development are enriched at hyperconnected 3D cliques of C57BL/6 (related to Figures 3 and 4).

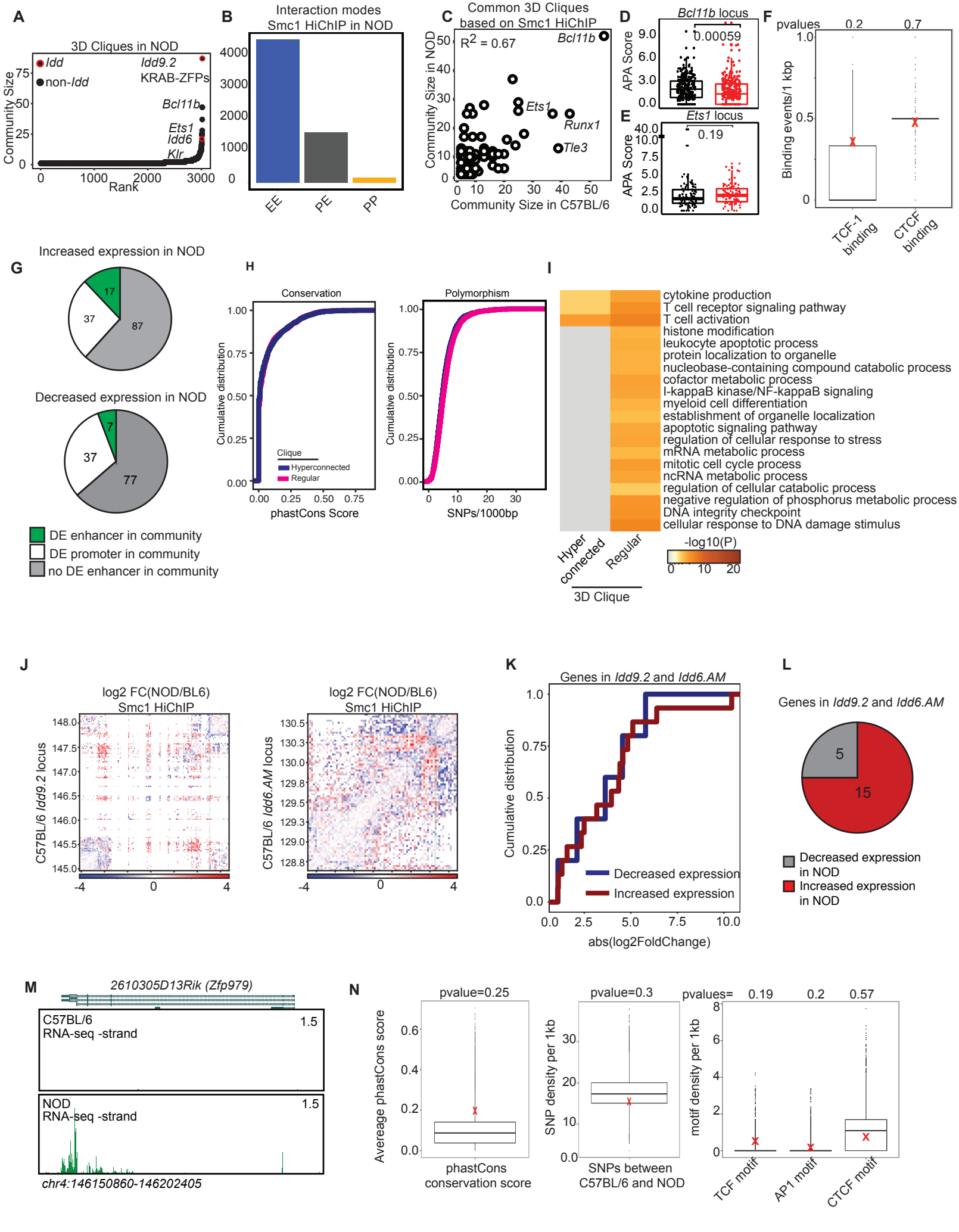
(A) We performed Smc1 HiChIP in replicates with ~500 million sequencing reads for each sample. Only significant loops that were detected in both replicates of a sample with an FDR ≤ 0.05 were retained for further analysis (14,658 reproducible loops in C57BL/6 and 18,225 reproducible loops in NOD). Furthermore, only reproducible Smc1-mediated loops with Smc1 ChIP-seq peaks deposited at least at one anchor were considered for further analysis providing 7,661 interactions (52%) in C57BL/6 and 8,694 significant interactions (47%) in NOD. Significant loops were further filtered with signals from H3K27ac ChIP-seq (4,936 in C57BL/6 and 5,013 in NOD mice).

(B) Super-enhancers defined by H3K27ac load in DP T cells of C57BL/6 mice.

(C) The Venn diagram demonstrates overlapping enhancers marked as super-enhancers with nodes of hyperconnected 3D cliques in C57BL/6.

(D-E) The genome-browser views of hyperconnected 3D cliques overlapping *Bcl6* and *Tle3* genes**.** HiChIP interactions are reproducible in two biological replicates.

Figure S4


A
3D Cliques in NOD
Idd
non-Idd
Idd9.2
KRAB-ZFPs
Bcl11b
Ets1
Idd6
Klr
Community Size
Rank
B
Interaction modes
Smc1 HiChIP in NOD
EE
PE
PP
C
Common 3D Cliques
based on Smc1 HiChIP
R2 = 0.67
Bcl11b
Ets1
Runx1
Tle3
Community Size in NOD
Community Size in C57BL/6
D
Bcl11b locus
APA Score
0.00059
E
Ets1 locus
0.19
F
pvalues
0.2
0.7
Binding events/1 kbp
TCF-1 binding
CTCF binding
G
Increased expression in NOD
17
37
87
Decreased expression in NOD
7
37
77
DE enhancer in community
DE promoter in community
no DE enhancer in community
H
Conservation
Polymorphism
Cumulative distribution
Clique
Hyperconnected
Regular
phastCons Score
SNPs/1000bp
I
cytokine production
T cell receptor signaling pathway
T cell activation
histone modification
leukocyte apoptotic process
protein localization to organelle
nucleobase-containing compound catabolic process
cofactor metabolic process
I-kappaB kinase/NF-kappaB signaling
myeloid cell differentiation
establishment of organelle localization
apoptotic signaling pathway
regulation of cellular response to stress
mRNA metabolic process
mitotic cell cycle process
ncRNA metabolic process
regulation of cellular catabolic process
negative regulation of phosphorus metabolic process
DNA integrity checkpoint
cellular response to DNA damage stimulus
Hyper connected
Regular
-log10(P)
3D Clique
J
log2 FC(NOD/BL6)
Smc1 HiChIP
C57BL/6 Idd9.2 locus
C57BL/6 Idd6.AM locus
K
Genes in Idd9.2 and Idd6.AM
Cumulative distribution
Decreased expression
Increased expression
abs(log2FoldChange)
L
Genes in Idd9.2 and Idd6.AM
5
15
Decreased expression in NOD
Increased expression in NOD
M
2610305D13Rik (Zfp979)
C57BL/6
RNA-seq -strand
NOD
RNA-seq -strand
chr4:146150860-146202405
N
pvalue=0.25
Avereage phastCons score
phastCons conservation score
pvalue=0.3
SNP density per 1kb
SNPs between C57BL/6 and NOD
pvalues=
0.19
0.2
0.57
motif density per 1kb
TCF motif
AP1 motif
CTCF motif

Supplementary Figure 4. Shared and distinct hyperconnected 3D cliques (related to Figures 3-5).

(A) Hyperconnected 3D cliques in DP T cells of NOD mice.

(B) The number of enhancer-enhancer (EE), enhancer-promoter (EP), and promoter-promoter (PP) interactions in double-positive thymocytes measured by H3K27ac HiChIP interactions reproduced in two biological replicates in NOD mice.

(C) Pearson correlation for the number of interactions in hyperconnected 3D cliques shared in the two strains. Genomic regions scored as hyperconnected in two strains were selected and the number of interactions for overlapping 3D cliques are shown in the dot plot.

(**D-E**) APA scores of significant loops in *Bcl11b* and *Ets1* suggest that while interactions at Ets1 locus (**E**) are similar in both strains, T cells of C57BL/6 demonstrates slightly higher *Smc1*-mediated interactions at the *Bcl11b* locus (**D**).

(F) Boxplot represents the distribution of TCF-1 and CTCF binding events in DP T cells measured by ChIP-seq at nodes of 1,000,000 permutated resilient 3D cliques. Red X represents the average binding events at nodes of 17 resilient hyperconnected 3D cliques.

(G) Pie chart depicts number of genes with increased expression in NOD overlapping HiChIP loop anchors with gain in enhancer or promoter elements in NOD compared with C57BL/6. “DE enhancer in community” demonstrates the presence of a differentially accessible enhancer in a 3D community containing the differentially expressed gene. Pie chart depicts number of genes with decreased expression in NOD overlapping HiChIP loop anchors with loss in enhancer or promoter elements in NOD compared with C57BL/6. “DE enhancer in community” demonstrates the presence of a differentially accessible enhancer in a 3D community containing the differentially expressed gene.

(H) Cumulative distribution function for phastCons score and SNPs at hyperconnected (blue) and regular (pink) 3D cliques in NOD.

(I) Gene ontology analysis using *metascape* suggests ‘cytokine production’, ‘TCR signaling pathway’, and ‘T cell activation’ to be associated with hyperconnected 3D cliques in NOD mice.

(J) LogFC in contact frequency matrix for Smc1 HiChIP between two strains at *Idd* loci.

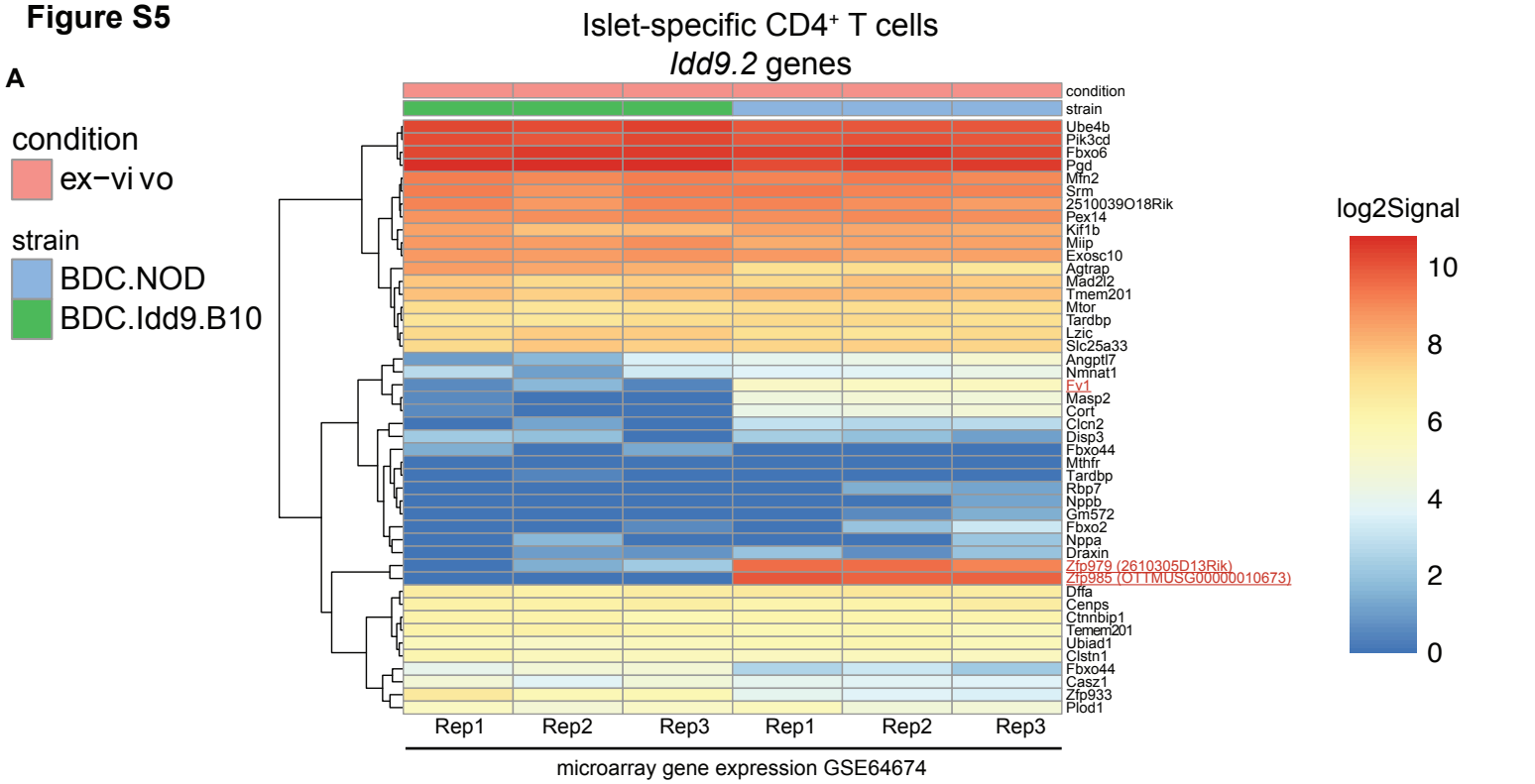
(K) Cumulative distribution depicts absolute log fold-change between two strains for genes in 3D cliques spanning *Idd9.2* and *Idd6.AM.*

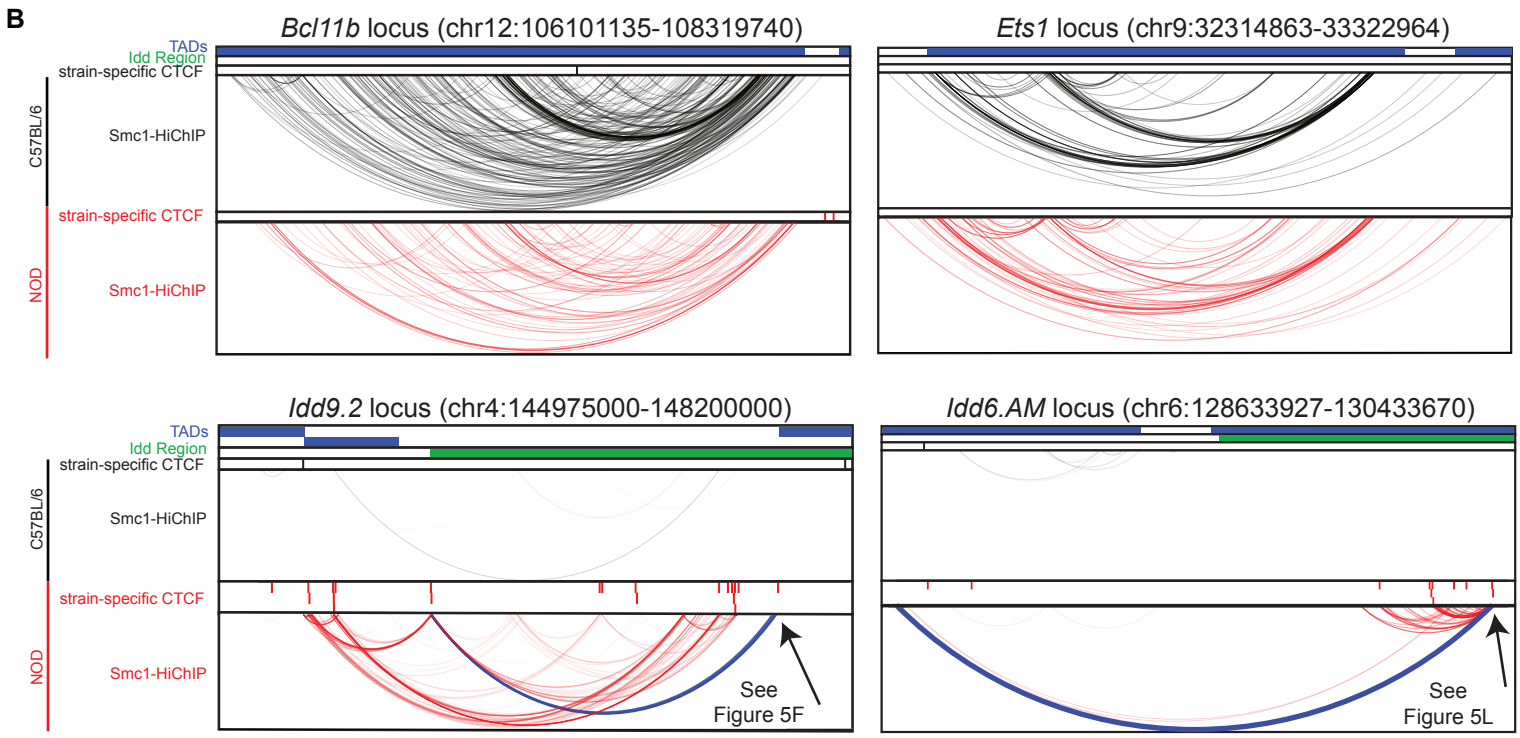
(L) Pie chart demonstrates the number of differentially expressed genes in 3D cliques spanning *Idd9.2* and *Idd6.AM* domains.

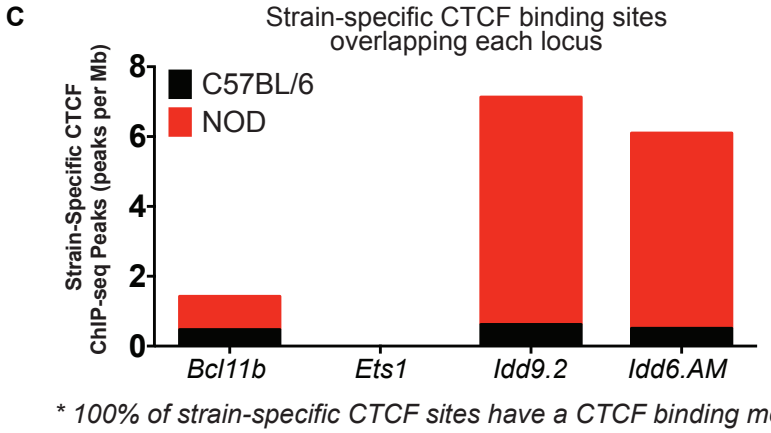
(M) Genome-browser demonstrates negative strand coverage profile of RNA-seq data in C57BL/6 and NOD mice at *2610305D13Rik* (*Zfp979*) gene. All bigwig files are tag-per-million normalized. A representative of 4 RNA-seq experiments is shown.

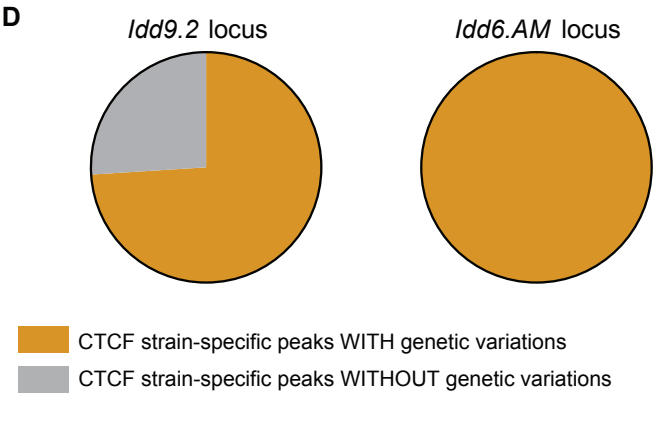
(N) Boxplot represents the distribution of TCF, AP-1 and CTCF motif density per 1kbp at nodes of 100,000 permuted hyperconnected 3D cliques in NOD strain. Red X represents the average density of TCF, AP-1 and CTCF motif per 1kbp at nodes of resilient hyperconnected 3D cliques.

Figure S5


A
Islet-specific CD4+ T cells
Idd9.2 genes
condition
ex−vi vo
strain
BDC.NOD
BDC.Idd9.B10
condition
strain
Ube4b
Pik3cd
Fbxo6
Pgd
Mfn2
Srm
2510039O18Rik
Pex14
Kif1b
Miip
Exosc10
Aqtrap
Mad2l2
Tmem201
Mtor
Tardbp
Lzic
Slc25a33
Angptl7
Nmnat1
Fv1
Masp2
Cort
Clcn2
Disp3
Fbxo44
Mthfr
Tardbp
Rbp7
Nppb
Gm572
Fbxo2
Nppa
Draxin
Zfp979 (2610305D13Rik)
Zfp985 (OTTMUSG00000010673)
Dffa
Cenps
Ctnnbip1
Tmem201
Ubiad1
Clstn1
Fbxo44
Casz1
Zfp933
Plod1
Rep1
Rep2
Rep3
Rep1
Rep2
Rep3
microarray gene expression GSE64674
log2Signal
10
8
6
4
2
0


B
Bcl11b locus (chr12:106101135-108319740)
Ets1 locus (chr9:32314863-33322964)
TADs
Idd Region
strain-specific CTCF
C57BL/6
Smc1-HiChIP
strain-specific CTCF
NOD
Smc1-HiChIP
Idd9.2 locus (chr4:144975000-148200000)
Idd6.AM locus (chr6:128633927-130433670)
See Figure 5F
See Figure 5L

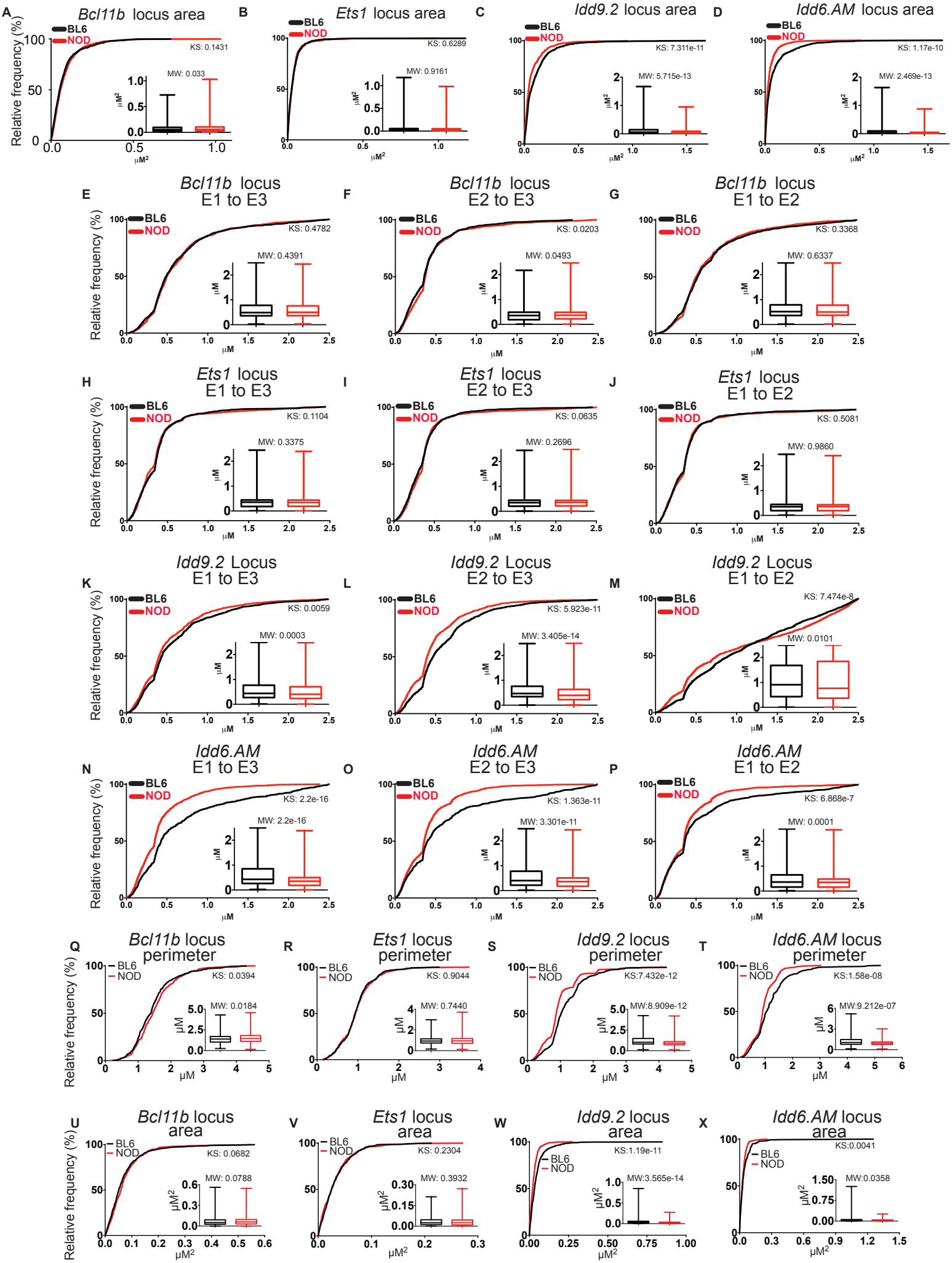

C
Strain-specific CTCF binding sites overlapping each locus
Strain-Specific CTCF ChIP-seq Peaks (peaks per Mb)
C57BL/6
NOD
8
6
4
2
0
Bcl11b
Ets1
Idd9.2
Idd6.AM
* 100% of strain-specific CTCF sites have a CTCF binding motif


D
Idd9.2 locus
Idd6.AM locus
CTCF strain-specific peaks WITH genetic variations
CTCF strain-specific peaks WITHOUT genetic variations

Supplementary Figure 5. Investigation of gene expression and CTCF binding at the two *Idd* regions with hyperconnected 3D cliques (related to Figure 5).

(A) Heatmap demonstrates expression levels of genes in the *Idd9.2* locus in $CD4^+$ T cells of NOD mice (noted as BDC.NOD) in comparison with those in diabetes-resistent congenic *Idd9.2* mice (noted as BDC.Idd9.B10) where both strains carry a transgenic TCR derived from a diabetogenic NOD T cell clone (BDC). As described in (Berry et al., 2015), ex vivo cells from the spleen were stained with anti-CD4 and anti-TCRVβ4, and subsequently sorted for CD4+TCRVβ4+ (transgenic) T cells. Analyzing the publically available microarray data (Berry et al., 2015), we found the genes with increased expression in NOD (in red) within the *Idd* domain harboring the hyperconnected 3D clique, *Fv1*, *Zfp979* and *Zfp985,* in $CD4^+$ T cells of BDC.NOD mice compared to BDC.*Idd9*.B10 that are protected from diabetes. All genes in *Idd9.2* with comparable probe ids in the microarray are shown.
(B) Genome-browser views of the *Bcl11b*, *Ets1*, *Idd9.2*, and *Idd6.AM* loci with the demarcation of TADs (blue) (Dixon et al., 2012), *Idd* regions (green), C57BL/6 CTCF ChIP-seq peaks (black lines), C57BL/6 H3K27ac-HiChIP (black loops), NOD CTCF ChIP-seq peaks (red lines), and NOD H3K27ac-HiChIP (red loops). Yellow boxes highlight regions with increased CTCF binding events in the *Idd* regions.
(C) Bar graph of strain-specific CTCF binding sites that overlap each locus (peaks per 1Mb).
(D) Pie charts indicating the proportion of CTCF strain-specific peaks that overlap with genetic variations within the two *Idd* regions.

Figure S6


A
Bcl11b locus area
KS: 0.1431
MW: 0.033
B
Ets1 locus area
KS: 0.6289
MW: 0.9161
C
Idd9.2 locus area
KS: 7.311e-11
MW: 5.715e-13
D
Idd6.AM locus area
KS: 1.17e-10
MW: 2.469e-13
E
Bcl11b locus E1 to E3
KS: 0.4782
MW: 0.4391
F
Bcl11b locus E2 to E3
KS: 0.0203
MW: 0.0493
G
Bcl11b locus E1 to E2
KS: 0.3368
MW: 0.6337
H
Ets1 locus E1 to E3
KS: 0.1104
MW: 0.3375
I
Ets1 locus E2 to E3
KS: 0.0635
MW: 0.2696
J
Ets1 locus E1 to E2
KS: 0.5081
MW: 0.9860
K
Idd9.2 Locus E1 to E3
KS: 0.0059
MW: 0.0003
L
Idd9.2 Locus E2 to E3
KS: 5.923e-11
MW: 3.405e-14
M
Idd9.2 Locus E1 to E2
KS: 7.474e-8
MW: 0.0101
N
Idd6.AM E1 to E3
KS: 2.2e-16
MW: 2.2e-16
O
Idd6.AM E2 to E3
KS: 1.363e-11
MW: 3.301e-11
P
Idd6.AM E1 to E2
KS: 6.868e-7
MW: 0.0001
Q
Bcl11b locus perimeter
KS: 0.0394
MW: 0.0184
R
Ets1 locus perimeter
KS: 0.9044
MW: 0.7440
S
Idd9.2 locus perimeter
KS:7.432e-12
MW:8.909e-12
T
Idd6.AM locus perimeter
KS:1.58e-08
MW:9.212e-07
U
Bcl11b locus area
KS: 0.0682
MW: 0.0788
V
Ets1 locus area
KS: 0.2304
MW: 0.3932
W
Idd9.2 locus area
KS:1.19e-11
MW:3.565e-14
X
Idd6.AM locus area
KS:0.0041
MW:0.0358
BL6
NOD
Relative frequency (%)
μM
μM²

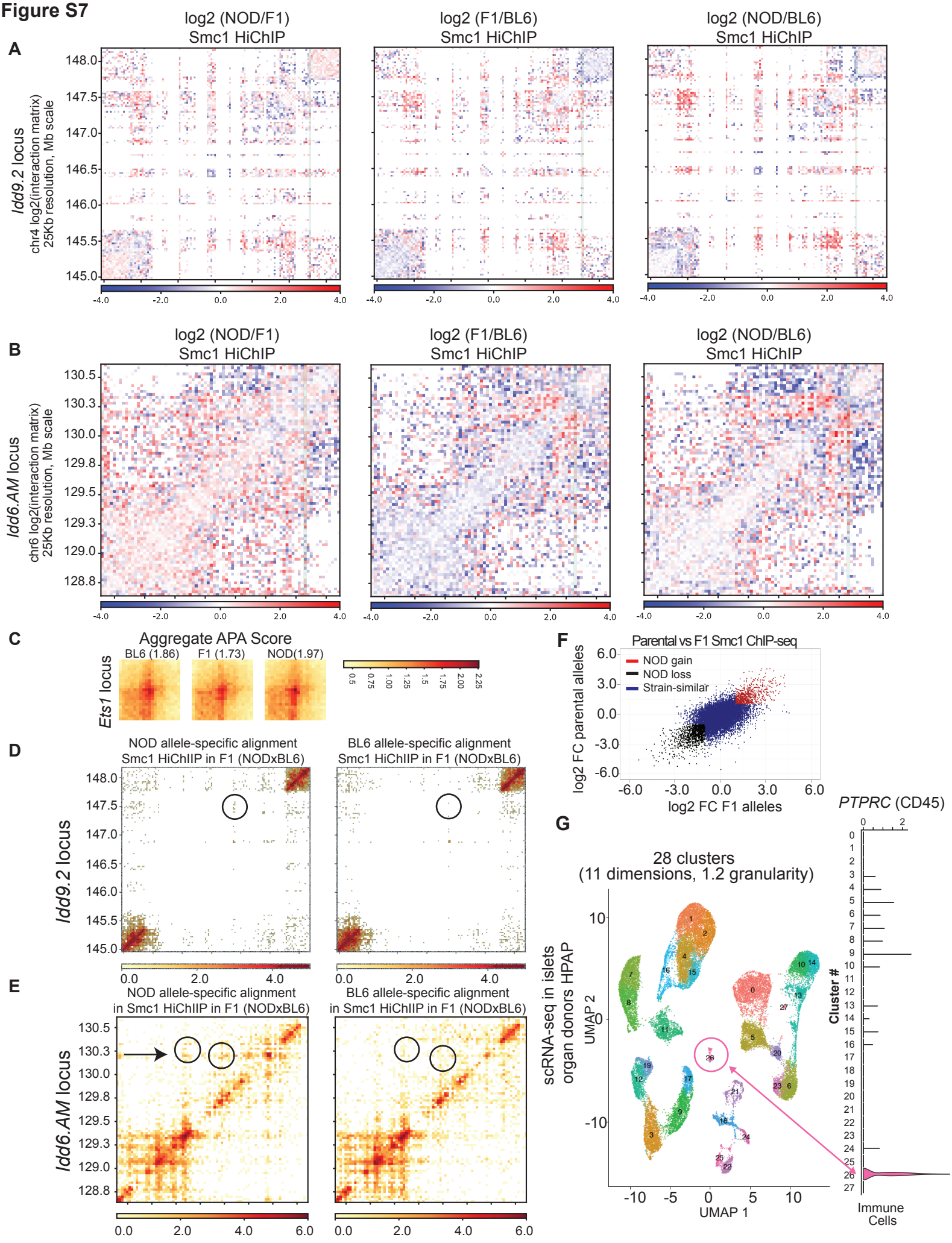
Supplementary Figure 6. Oligopaint 3D FISH corroborates the formation of strain-similar and diabetes-specific hyperconnected 3D cliques (related to Figure 6).

(**A-D**) Cumulative distribution plots and box plots of the spatial area formed by the three probes in ~500 cells (see Materials & Methods for exact numbers) per strain per locus (KS: Kolmogorov-Smirnov test & MW: Mann-Whitney test) at the Bcl11b locus, Ets1 locus, Idd9.2 locus, and Idd6.AM locus.

(**E-P**) Pairwise distances between probes in ~500 cells (see Materials & Methods for exact numbers) per strain per locus (KS: Kolmogorov-Smirnov test & MW: Mann-Whitney test) at the *Bcl11b* locus, *Ets1* locus, *Idd9.2* locus, and *Idd6.AM* locus.

(Q-X) Oligopaint 3D FISH in second biological replicate affirms the findings from the first Oligopaint replicate. (**Q-T**) Cumulative distribution plots and box plots of the spatial perimeter formed by the three probes in ~400 cells (see Materials & Methods for exact numbers) per strain per locus (KS: Kolmogorov-Smirnov test & MW: Mann-Whitney test) at the *Bcl11b* locus, *Ets1* locus, *Idd9.2* locus, and *Idd6.AM* locus. (**U-X**) Cumulative distribution plots and box plots of the spatial area formed by the three probes in ~400 cells (see Materials & Methods for exact numbers) per strain per locus (KS: Kolmogorov-Smirnov test & MW: Mann-Whitney test) at the Bcl11b locus, Ets1 locus, Idd9.2 locus, and Idd6.AM locus.

Figure S7


A
log2 (NOD/F1) Smc1 HiChIP
log2 (F1/BL6) Smc1 HiChIP
log2 (NOD/BL6) Smc1 HiChIP
Idd9.2 locus
chr4 log2(interaction matrix) 25Kb resolution, Mb scale
148.0
147.5
147.0
146.5
146.0
145.5
145.0
-4.0 -2.0 0.0 2.0 4.0
B
log2 (NOD/F1) Smc1 HiChIP
log2 (F1/BL6) Smc1 HiChIP
log2 (NOD/BL6) Smc1 HiChIP
Idd6.AM locus
chr6 log2(interaction matrix) 25Kb resolution, Mb scale
130.5
130.3
130.0
129.8
129.5
129.3
129.0
128.8
-4.0 -2.0 0.0 2.0 4.0
C
Aggregate APA Score
Ets1 locus
BL6 (1.86)
F1 (1.73)
NOD(1.97)
0.5 0.75 1.00 1.25 1.5 1.75 2.00 2.25
D
NOD allele-specific alignment Smc1 HiChIIP in F1 (NODxBL6)
BL6 allele-specific alignment Smc1 HiChIIP in F1 (NODxBL6)
Idd9.2 locus
148.0
147.5
147.0
146.5
146.0
145.5
145.0
0.0 2.0 4.0
E
NOD allele-specific alignment in Smc1 HiChIIP in F1 (NODxBL6)
BL6 allele-specific alignment in Smc1 HiChIIP in F1 (NODxBL6)
Idd6.AM locus
130.5
130.3
130.0
129.8
129.5
129.3
129.0
128.8
0.0 2.0 4.0 6.0
F
Parental vs F1 Smc1 ChIP-seq
NOD gain
NOD loss
Strain-similar
log2 FC parental alleles
log2 FC F1 alleles
6.0
3.0
0.0
-3.0
-6.0
G
28 clusters
(11 dimensions, 1.2 granularity)
scRNA-seq in islets organ donors HPAP
UMAP 2
UMAP 1
10
0
-10
-10 -5 0 5 10
PTPRC (CD45)
Cluster #
0 2
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Immune Cells

Supplementary Figure 7. 3D chromatin misfolding at diabetes-associated loci in NOD mice is mediated in *cis* and is linked to human T1D (related to Figure 7).

(**A, B**) Log2 fold-change of contact frequencies in Smc1 HiChIP for the region harboring *Idd9.2* (A) and *Idd6.AM* (B) between each pair of strains NOD/F1 (left), F1/BL6 (middle) and NOD/BL6 offspring (middle).

(**C**) APA score of Smc1 HiChIP loop was used to quantify the strength in the strain-similar *Ets1* region between parental and F1 strains.

(**D-E**) Contact matrix for Smc1 HiChIP in F1 for allele-specific alignment at Idd9.2 (**D**) and Idd6.AM loci (**E**). We used the allele-specific alignment of HiCPro ("ALLELE_SPECIFIC_SNP" providing NOD genome vcf file) and detected NOD-specific interactions at *Idd* regions.

(F) Ratio-ratio plot of Smc1 ChIP-seq parental log2 fold change in double-positive thymocytes derived from C57BL/6 and NOD mice versus allele-specific log2 fold change in double-positive thymocytes derived from (NOD x C57BL/6) F1 mice.

(**G**) Pancreatic islets were procured from the HPAP consortium under Human Islet Research Network (https://hirnetwork.org/) and single-cell RNA-seq was generated using 10xGenomics. The dimensionality reduction UMAP demonstrates the immune cell population in the islets of these donors that were used for the GSEA analysis.