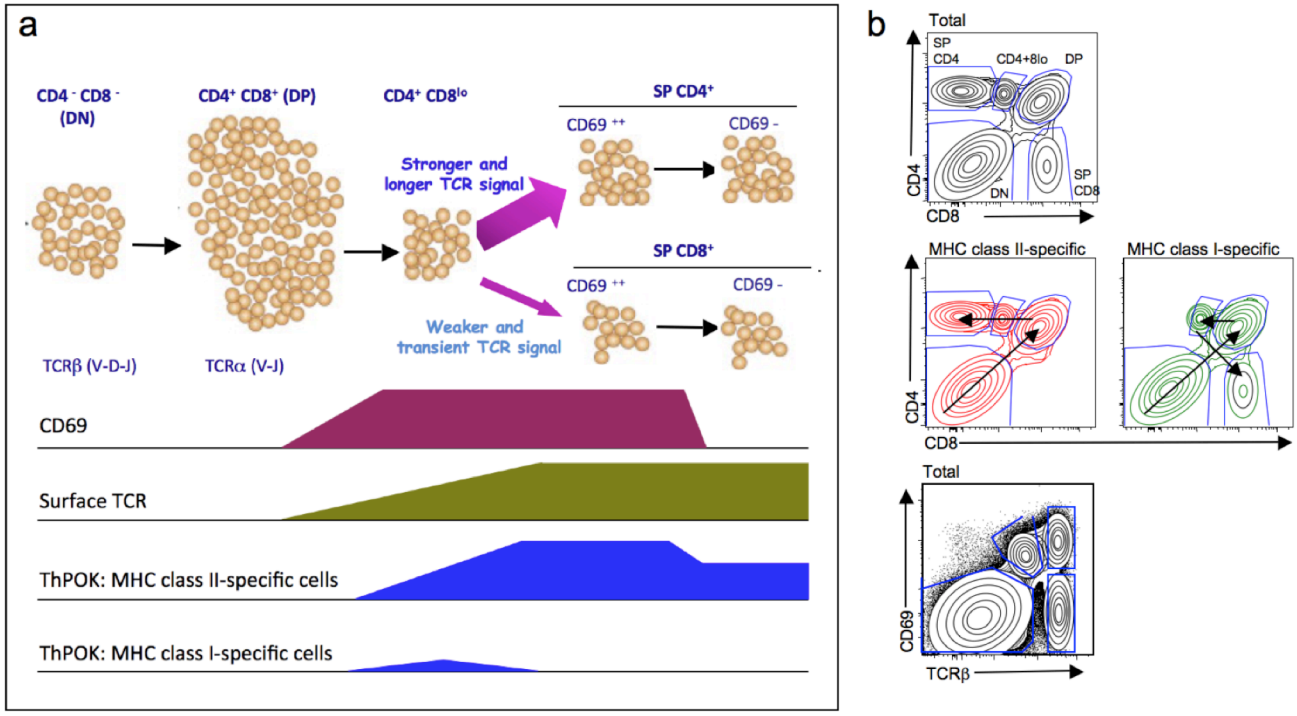
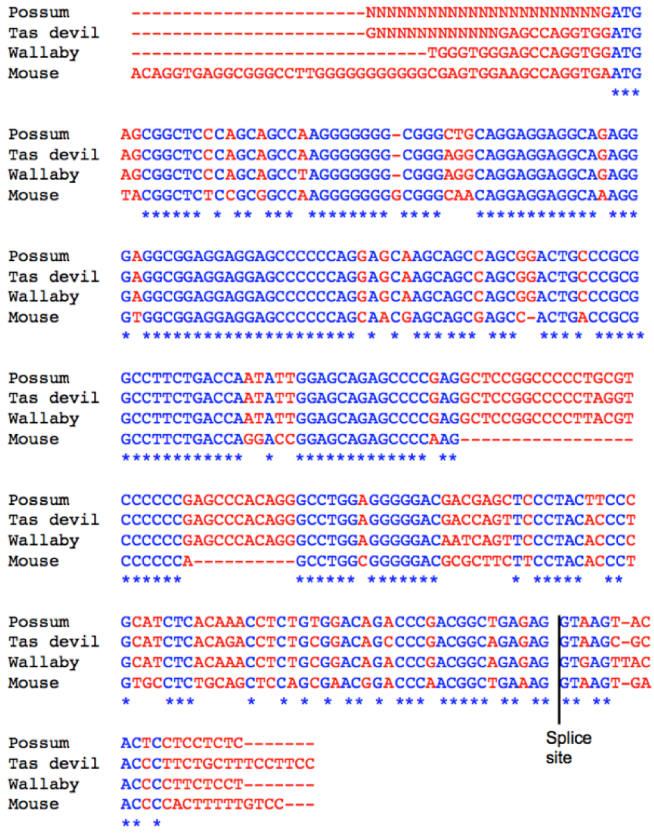
Suppl. Fig. 1

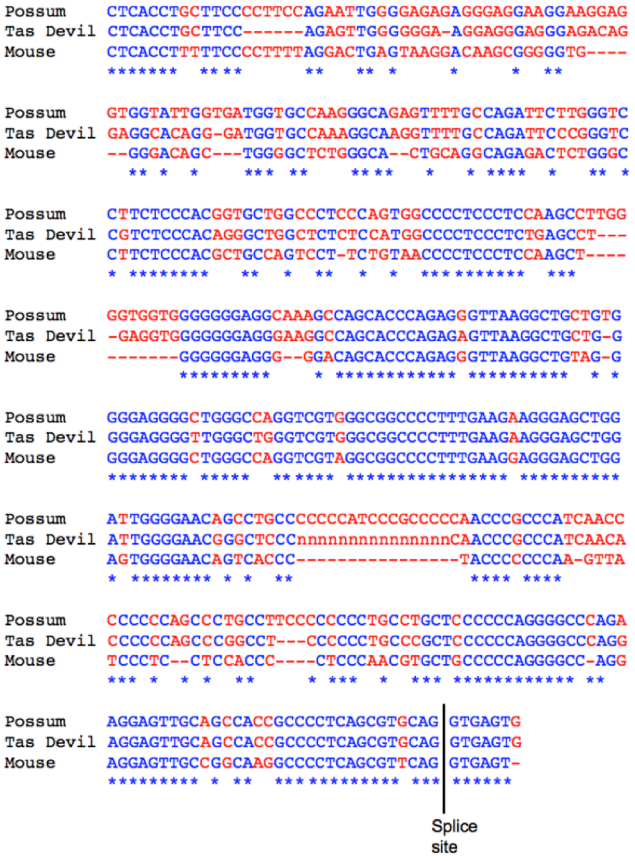

a
CD4⁻ CD8⁻ (DN)
CD4⁺ CD8⁺ (DP)
CD4⁺ CD8ˡᵒ
SP CD4⁺
CD69 ⁺⁺
CD69 -
Stronger and longer TCR signal
SP CD8⁺
CD69 ⁺⁺
CD69 -
Weaker and transient TCR signal
TCRβ (V-D-J)
TCRα (V-J)
CD69
Surface TCR
ThPOK: MHC class II-specific cells
ThPOK: MHC class I-specific cells
b
Total
SP CD4
CD4+8lo
DP
DN
SP CD8
CD4
CD8
MHC class II-specific
MHC class I-specific
CD4
CD8
Total
CD69
TCRβ

Suppl. Fig. 2

a.



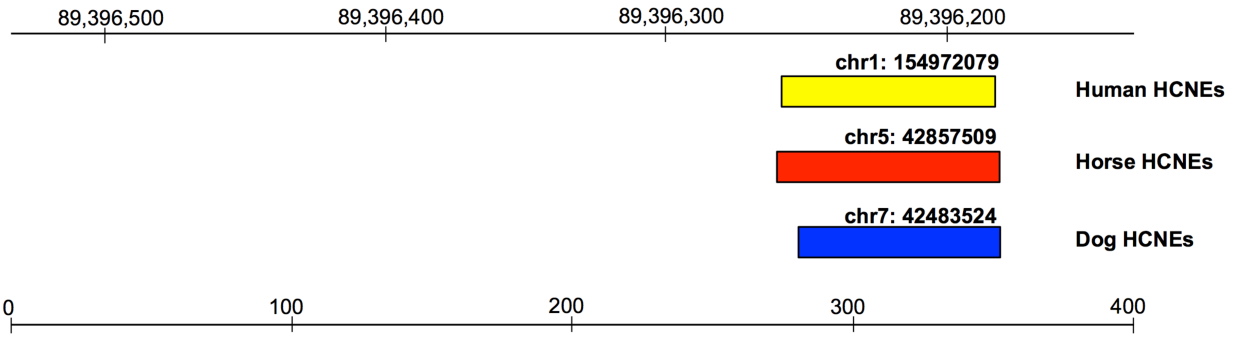
b.



Suppl. Fig. 3 a

Suppl. Fig. 3b

Suppl. Fig. 3c



Organism	Genome assembly	Location	Size
Mouse	NCBI Build 38 (mm10)	chr3:89,396,182..89,396,257	76 bp
Human	NCBI Build 37 (hg19)	chr1:154,972,079..154,972,154	76 bp

Alignment

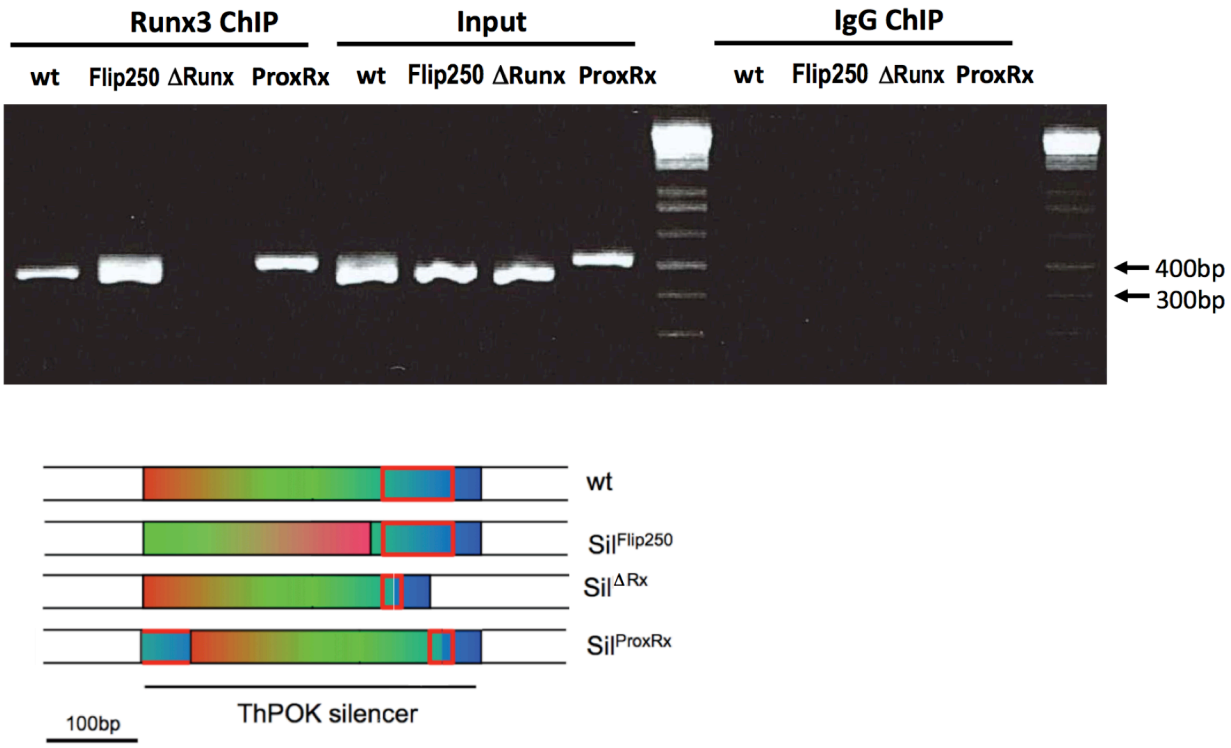
```
Mouse   CCGCGGCTGCCTCCGCTTCCCTCGAAAAACCCGCCGCCCAGGAAACCGCTGGGGGCGCTC
Human   CCTCCGCTGCCTCCGCTTCCCTCGAAAAACCCGCCGCCCAGGAAACCGCTGGGGGCGCTC
        ** * ************************************************************

Mouse   AGACCGCAGCCCCCTC
Human   AGACCGCAGACCCCCC
        ********* **** *
```

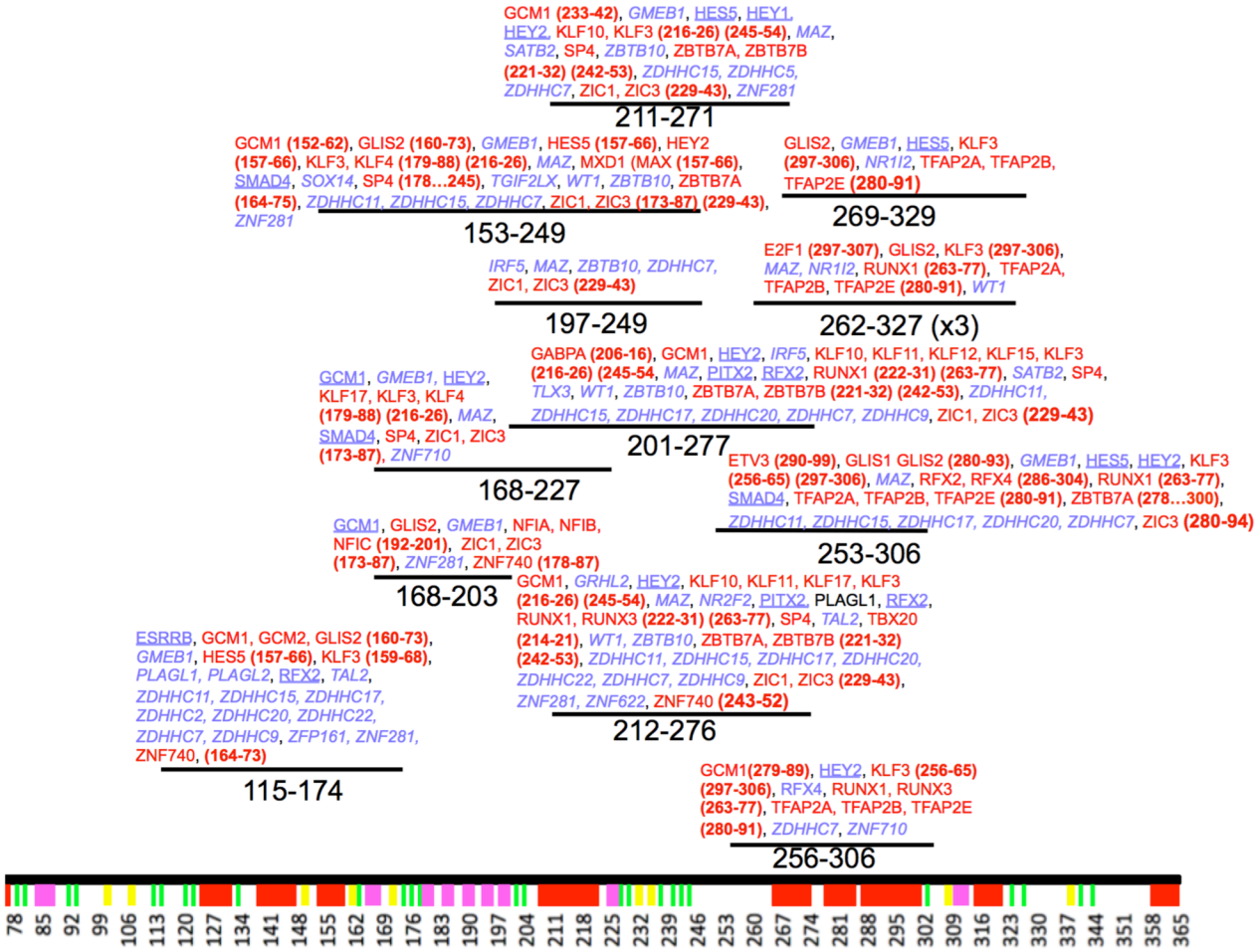
Note: the human sequence is reverse complemented in the alignment.

Similarity: 72 / 76 identical columns = 94.74%

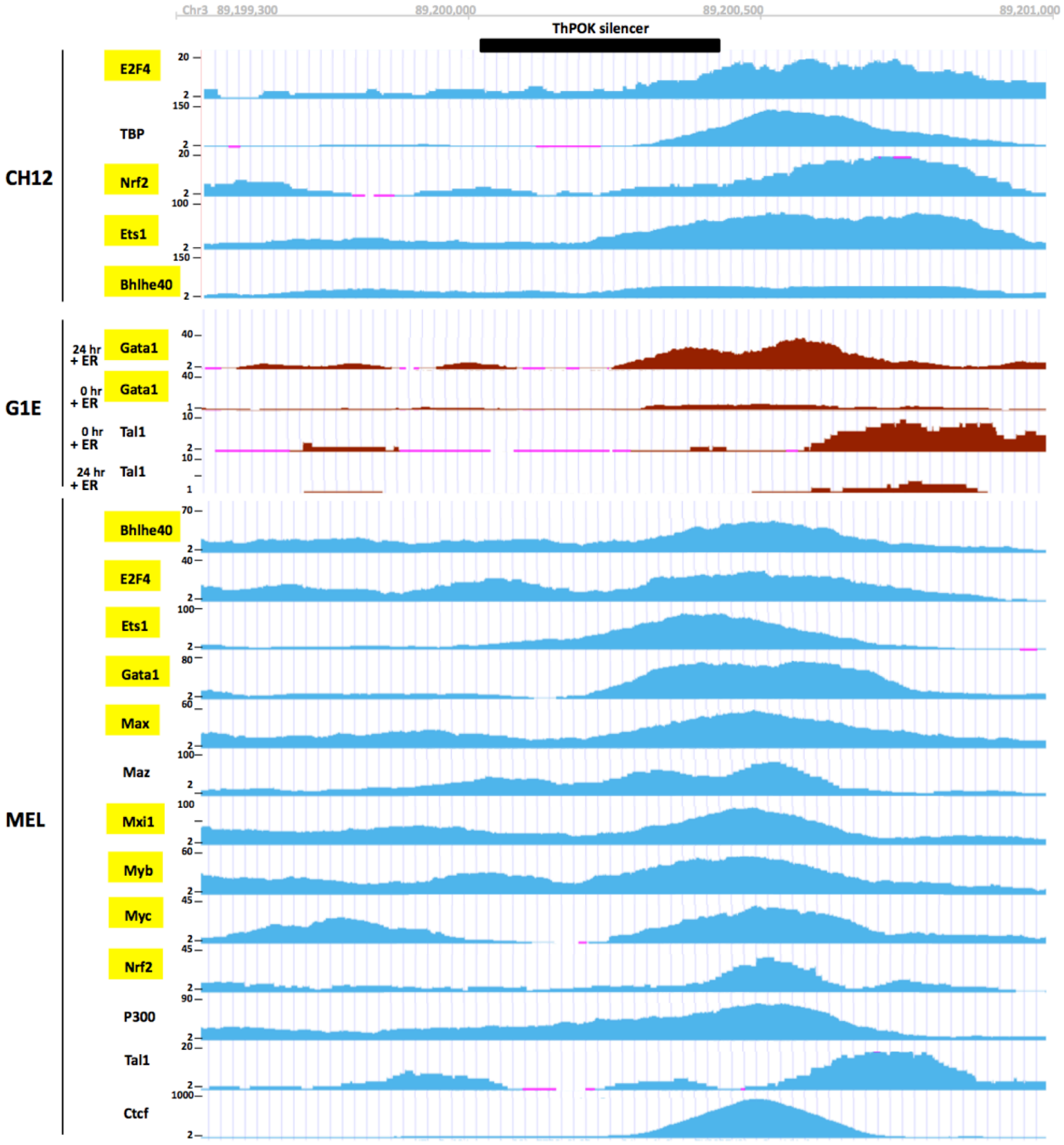
Suppl. Fig. 3d


Runx3 ChIP
Input
IgG ChIP
wt
Flip250
ΔRunx
ProxRx
wt
Flip250
ΔRunx
ProxRx
wt
Flip250
ΔRunx
ProxRx
400bp
300bp
wt
Sil^Flip250
Sil^ΔRx
Sil^ProxRx
100bp
ThPOK silencer

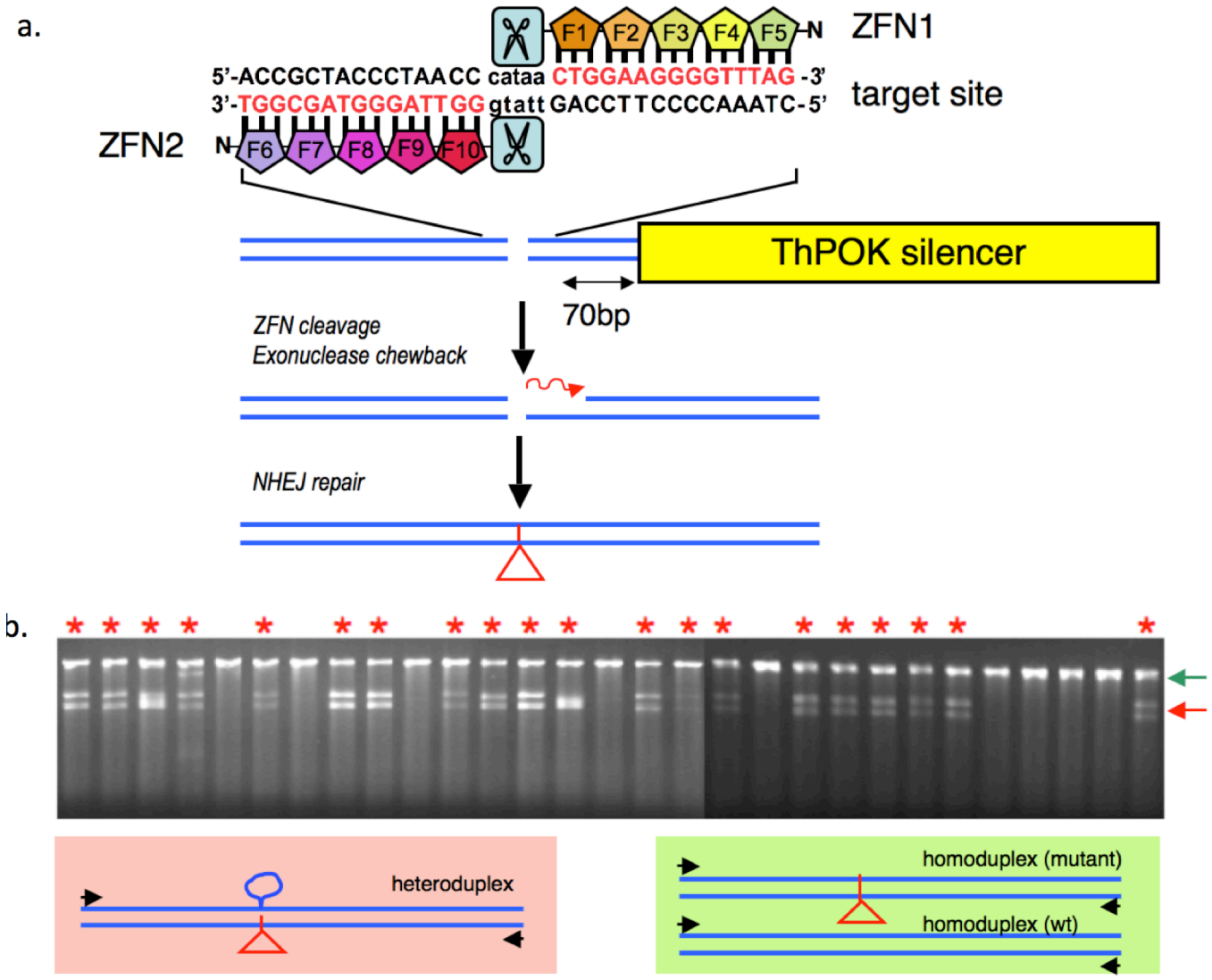
Suppl. Fig. 4a


GCM1 (233-42), GMEB1, HES5, HEY1, HEY2, KLF10, KLF3 (216-26) (245-54), MAZ, SATB2, SP4, ZBTB10, ZBTB7A, ZBTB7B (221-32) (242-53), ZDHHC15, ZDHHC5, ZDHHC7, ZIC1, ZIC3 (229-43), ZNF281
211-271
GCM1 (152-62), GLIS2 (160-73), GMEB1, HES5 (157-66), HEY2 (157-66), KLF3, KLF4 (179-88) (216-26), MAZ, MXD1 (MAX (157-66), SMAD4, SOX14, SP4 (178...245), TGIF2LX, WT1, ZBTB10, ZBTB7A (164-75), ZDHHC11, ZDHHC15, ZDHHC7, ZIC1, ZIC3 (173-87) (229-43), ZNF281
153-249
GLIS2, GMEB1, HES5, KLF3 (297-306), NR1I2, TFAP2A, TFAP2B, TFAP2E (280-91)
269-329
IRF5, MAZ, ZBTB10, ZDHHC7, ZIC1, ZIC3 (229-43)
197-249
E2F1 (297-307), GLIS2, KLF3 (297-306), MAZ, NR1I2, RUNX1 (263-77), TFAP2A, TFAP2B, TFAP2E (280-91), WT1
262-327 (x3)
GCM1, GMEB1, HEY2, KLF17, KLF3, KLF4 (179-88) (216-26), MAZ, SMAD4, SP4, ZIC1, ZIC3 (173-87), ZNF710
168-227
GABPA (206-16), GCM1, HEY2, IRF5, KLF10, KLF11, KLF12, KLF15, KLF3 (216-26) (245-54, MAZ, PITX2, RFX2, RUNX1 (222-31) (263-77), SATB2, SP4, TLX3, WT1, ZBTB10, ZBTB7A, ZBTB7B (221-32) (242-53), ZDHHC11, ZDHHC15, ZDHHC17, ZDHHC20, ZDHHC7, ZDHHC9, ZIC1, ZIC3 (229-43)
201-277
ETV3 (290-99), GLIS1 GLIS2 (280-93), GMEB1, HES5, HEY2, KLF3 (256-65) (297-306), MAZ, RFX2, RFX4 (286-304), RUNX1 (263-77), SMAD4, TFAP2A, TFAP2B, TFAP2E (280-91), ZBTB7A (278...300), ZDHHC11, ZDHHC15, ZDHHC17, ZDHHC20, ZDHHC7, ZIC3 (280-94)
253-306
GCM1, GLIS2, GMEB1, NFIA, NFIB, NFIC (192-201), ZIC1, ZIC3 (173-87), ZNF281, ZNF740 (178-87)
168-203
GCM1, GRHL2, HEY2, KLF10, KLF11, KLF17, KLF3 (216-26) (245-54), MAZ, NR2F2, PITX2, PLAGL1, RFX2, RUNX1, RUNX3 (222-31) (263-77), SP4, TAL2, TBX20 (214-21), WT1, ZBTB10, ZBTB7A, ZBTB7B (221-32) (242-53), ZDHHC11, ZDHHC15, ZDHHC17, ZDHHC20, ZDHHC22, ZDHHC7, ZDHHC9, ZIC1, ZIC3 (229-43), ZNF281, ZNF622, ZNF740 (243-52)
212-276
ESRRB, GCM1, GCM2, GLIS2 (160-73), GMEB1, HES5 (157-66), KLF3 (159-68), PLAGL1, PLAGL2, RFX2, TAL2, ZDHHC11, ZDHHC15, ZDHHC17, ZDHHC2, ZDHHC20, ZDHHC22, ZDHHC7, ZDHHC9, ZFP161, ZNF281, ZNF740, (164-73)
115-174
GCM1(279-89), HEY2, KLF3 (256-65) (297-306), RFX4, RUNX1, RUNX3 (263-77), TFAP2A, TFAP2B, TFAP2E (280-91), ZDHHC7, ZNF710
256-306
78 85 92 99 106 113 120 127 134 141 148 155 162 169 176 183 190 197 204 211 218 225 232 239 246 253 260 267 274 281 288 295 302 309 316 323 330 337 344 351 358 365

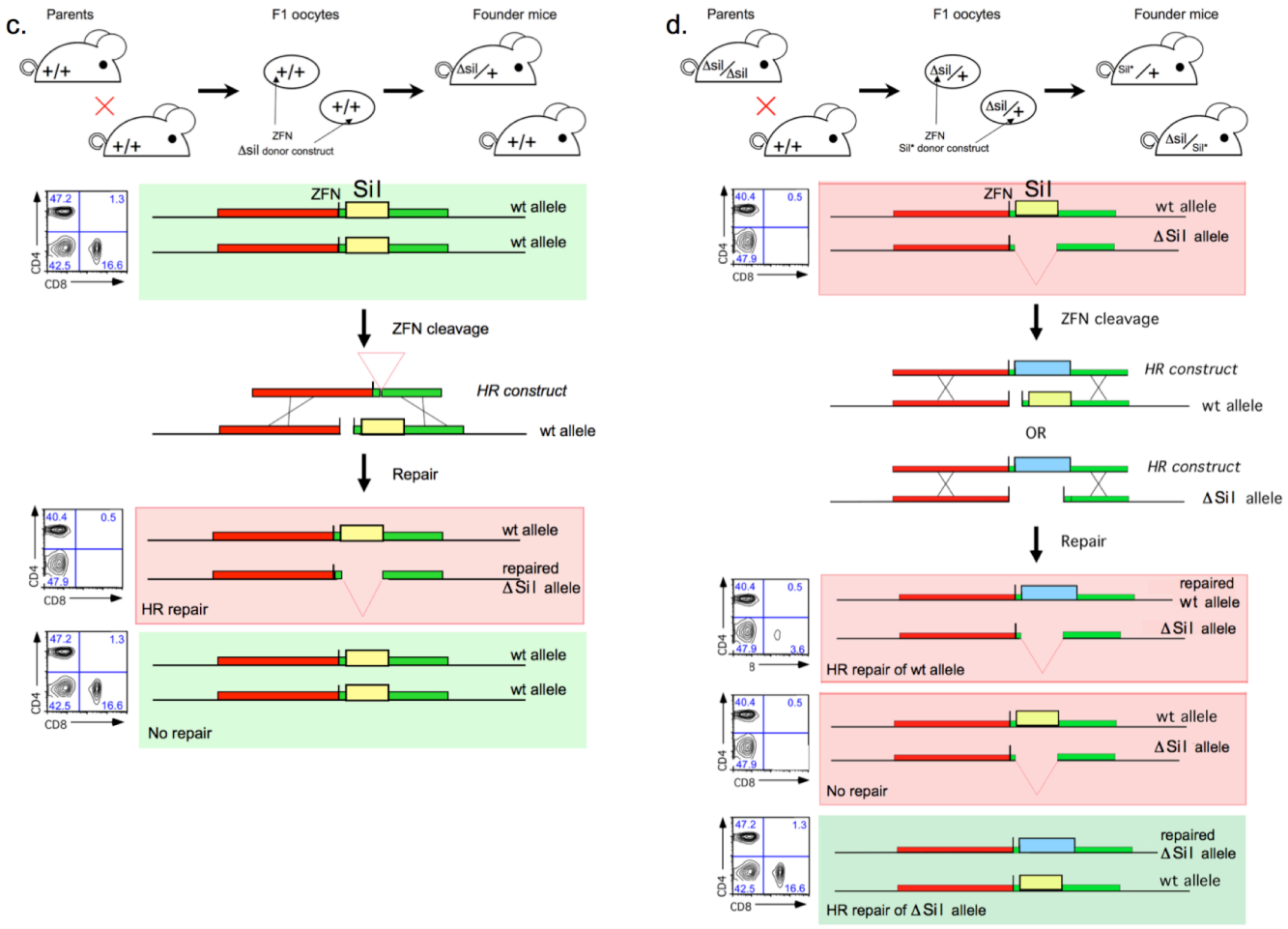
Suppl. Fig. 4b


Chr3 89,199,300
89,200,000
89,200,500
89,201,000
ThPOK silencer
CH12
E2F4
TBP
Nrf2
Ets1
Bhlhe40
G1E
24 hr + ER
Gata1
0 hr + ER
Gata1
0 hr + ER
Tal1
24 hr + ER
Tal1
MEL
Bhlhe40
E2F4
Ets1
Gata1
Max
Maz
Mxi1
Myb
Myc
Nrf2
P300
Tal1
Ctcf

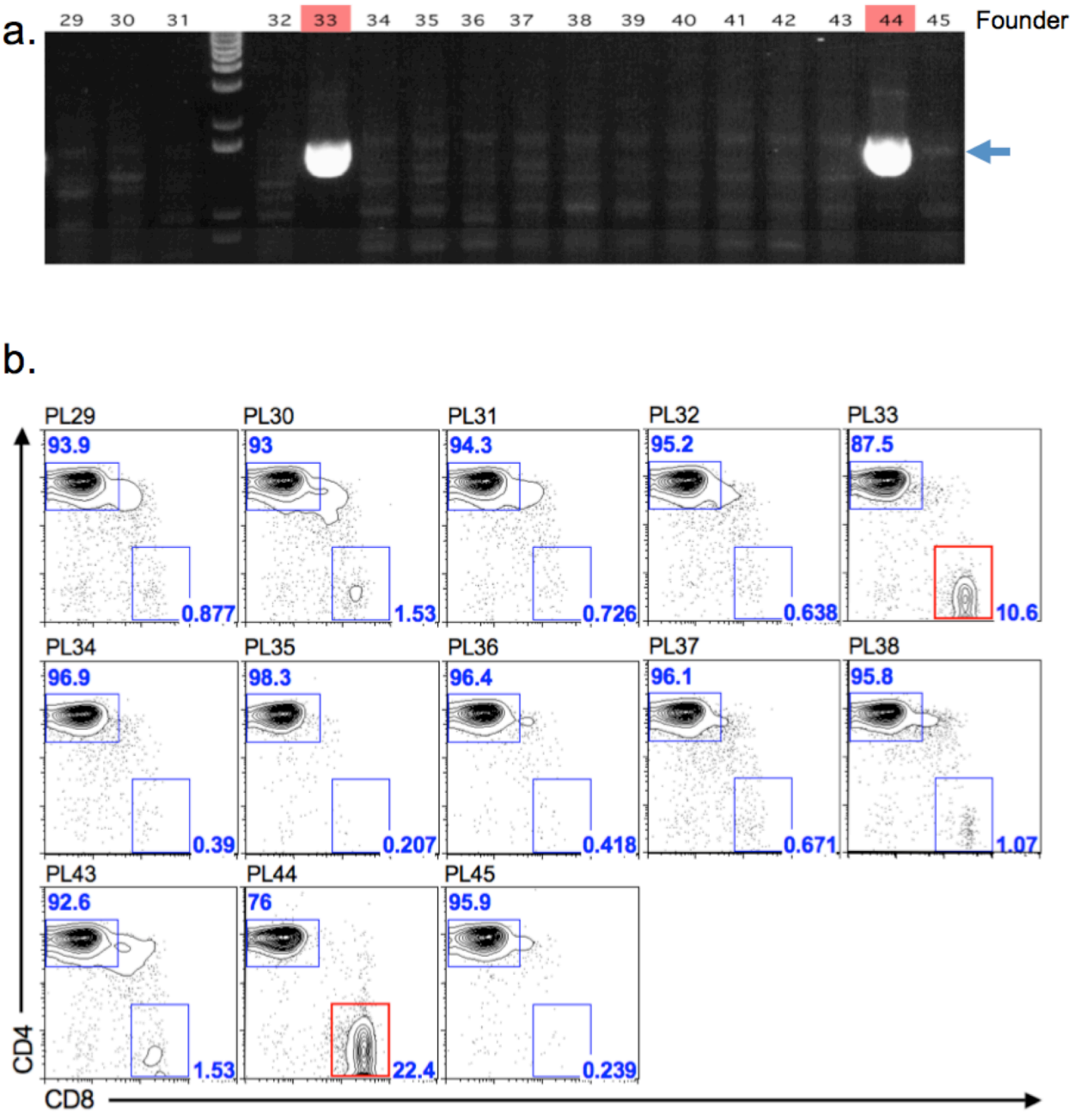
Suppl. Fig. 5


a.
ZFN1
F1 F2 F3 F4 F5 N
5'-ACCGCTACCCTAACC cataa CTGGAAGGGGTTTAG -3'
3'-TGGCGATGGGATTGG gtatt GACCTTCCCCAAATC-5'
target site
ZFN2
N F6 F7 F8 F9 F10
ThPOK silencer
70bp
ZFN cleavage
Exonuclease chewback
NHEJ repair
b.
heteroduplex
homoduplex (mutant)
homoduplex (wt)

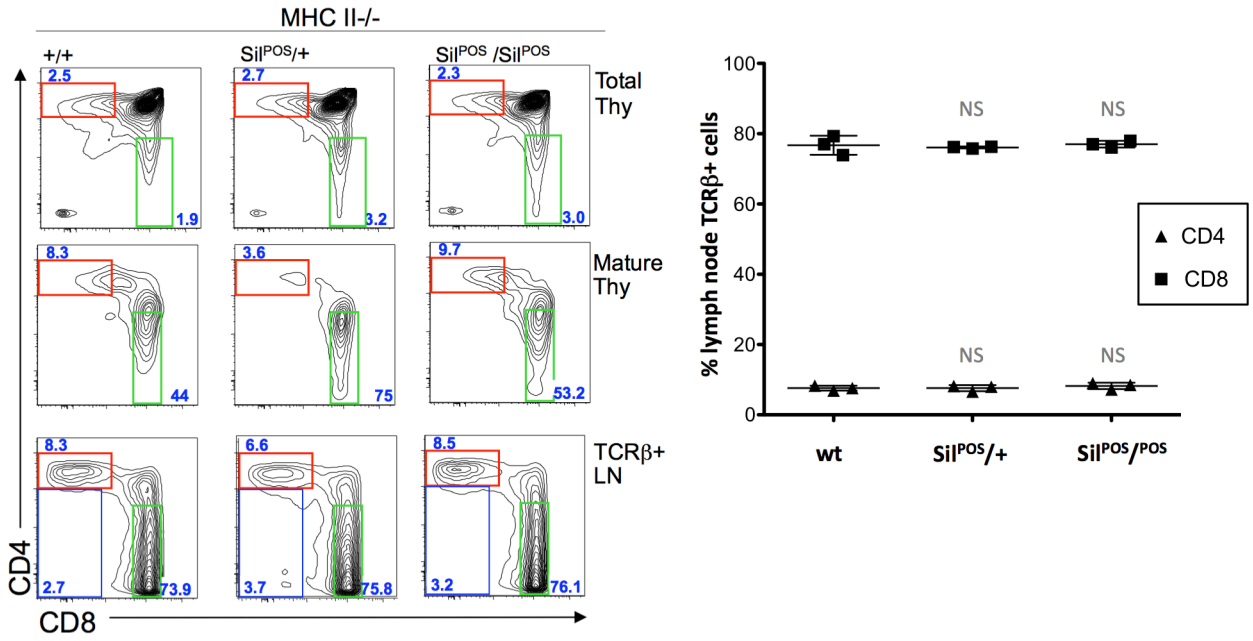
Suppl. Fig. 5


c.
Parents
F1 oocytes
Founder mice
+/+
+/+
+/+
+/+
ΔSil donor construct
ΔSil/+
+/+
ZFN
Sil
wt allele
wt allele
ZFN cleavage
HR construct
wt allele
Repair
wt allele
repaired ΔSil allele
HR repair
wt allele
wt allele
No repair
CD4
CD8
47.2
1.3
42.5
16.6
40.4
0.5
47.9
d.
Parents
F1 oocytes
Founder mice
ΔSil/ΔSil
+/+
ΔSil/+
Sil*/+
Sil* donor construct
ΔSil/Sil*
wt allele
ΔSil allele
ZFN cleavage
HR construct
wt allele
OR
HR construct
ΔSil allele
Repair
repaired wt allele
ΔSil allele
HR repair of wt allele
wt allele
ΔSil allele
No repair
repaired ΔSil allele
wt allele
HR repair of ΔSil allele
3.6
8

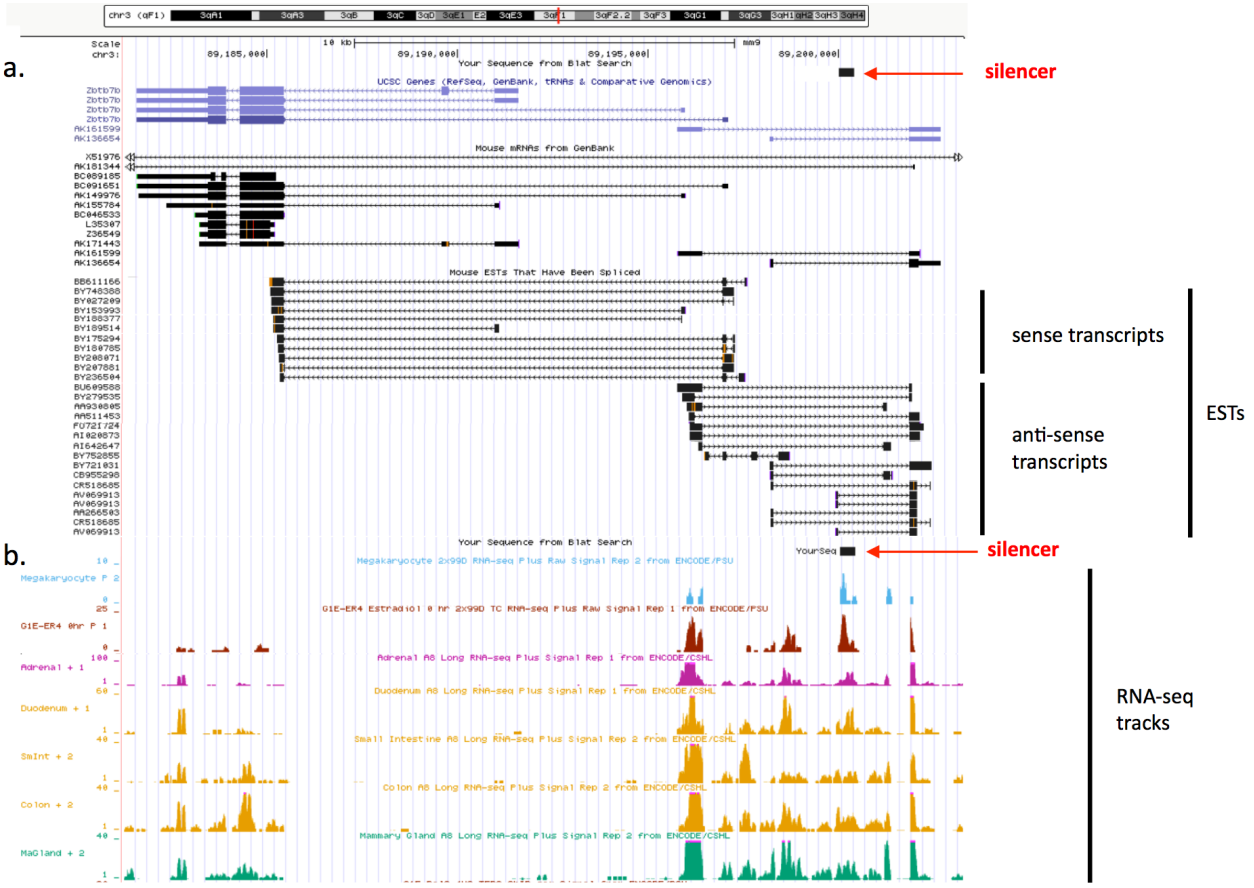
Suppl. Fig. 6


a.
29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
Founder
b.
PL29
93.9
0.877
PL30
93
1.53
PL31
94.3
0.726
PL32
95.2
0.638
PL33
87.5
10.6
PL34
96.9
0.39
PL35
98.3
0.207
PL36
96.4
0.418
PL37
96.1
0.671
PL38
95.8
1.07
PL43
92.6
1.53
PL44
76
22.4
PL45
95.9
0.239
CD4
CD8

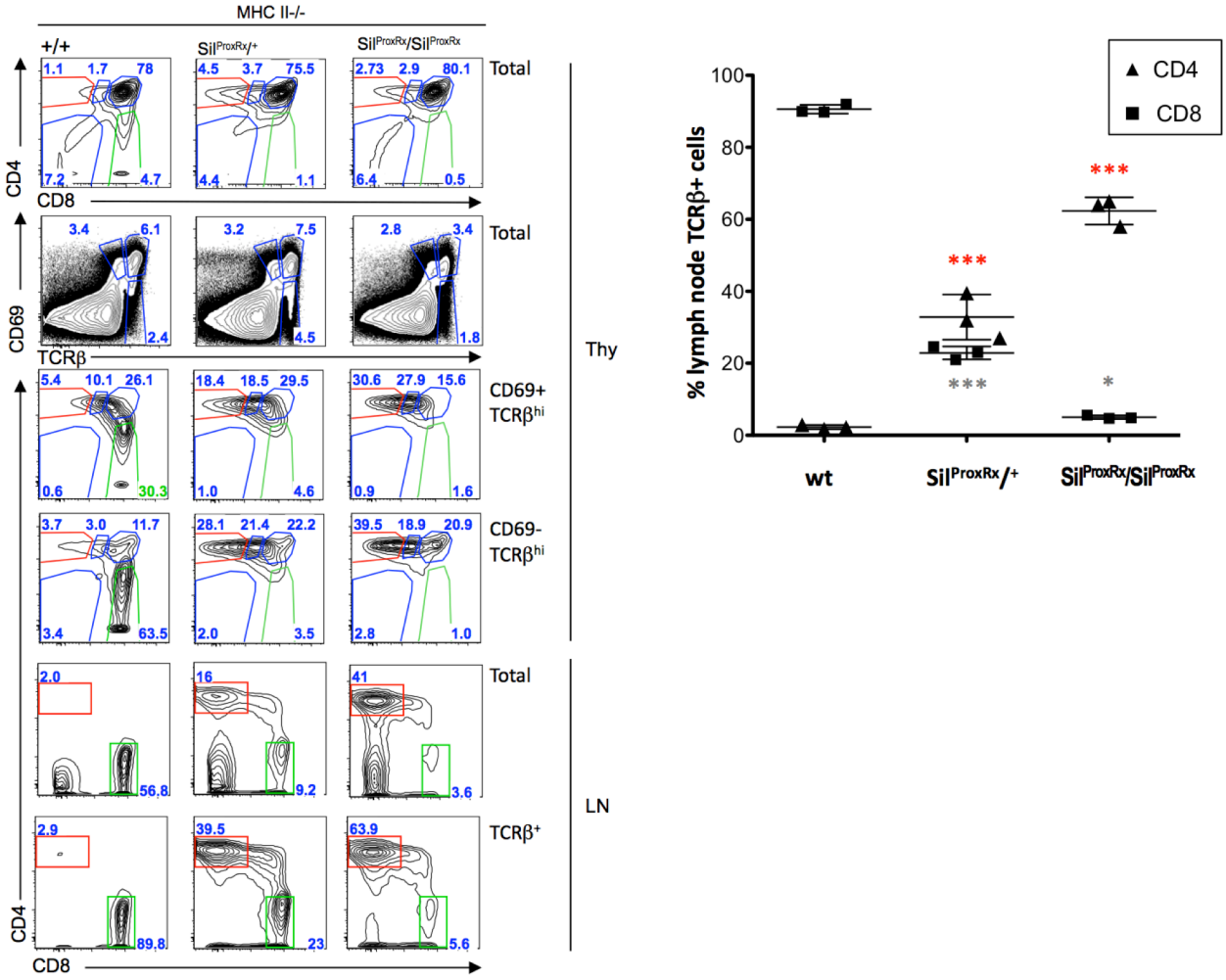
Suppl. Fig. 7


MHC II-/-
+/+
SilPOS/+
SilPOS /SilPOS
Total Thy
Mature Thy
TCRβ+ LN
CD4
CD8
2.5
1.9
2.7
3.2
2.3
3.0
8.3
44
3.6
75
9.7
53.2
8.3
2.7
73.9
6.6
3.7
75.8
8.5
3.2
76.1
% lymph node TCRβ+ cells
100
80
60
40
20
0
NS
NS
NS
NS
CD4
CD8
wt
SilPOS/+
SilPOS/POS

Suppl. Fig. 8


a.
silencer
sense transcripts
anti-sense transcripts
ESTs
b.
silencer
RNA-seq tracks

Suppl. Fig. 9


MHC II-/-
+/+
SilProxRx/+
SilProxRx/SilProxRx
Total
CD4
CD8
CD69
TCRβ
CD69+ TCRβhi
CD69- TCRβhi
Thy
LN
TCRβ+
% lymph node TCRβ+ cells
CD4
CD8
wt
SilProxRx/+
SilProxRx/SilProxRx

SUPPLEMENTARY FIGURE LEGENDS

Suppl. Fig. 1. **T lymphocyte development.** a) Diagram of stages in thymic development, including, in sequential order, double negative (DN), double positive (DP), intermediate CD4+8lo, and finally single-positive (SP) CD4 and CD8. Thymic development is also marked by changes in expression of TCR and CD69 surface markers. TCR and CD69 surface markers are both upmodulated as thymocytes mature from the DP to the CD4+8lo stage and even further during subsequent transition to the SP CD4 and CD8 stages. Finally, just before exiting the thymus, SP thymocytes lose CD69 expression, but still maintain high TCR expression. Most peripheral T lymphocytes, from the blood, lymph nodes and spleen, exhibit the same TCR+ CD69- surface phenotype. At the CD4+8lo stage, MHC class II- and class I-specific cells express high or low levels of ThPOK, respectively, and diverge into alternate CD4 and CD8 lineages. b) Surface expression pattern of CD4, CD8, CD69 and TCR by total thymocytes, or by MHC class I- or II-specific thymocytes, as indicated. Diagrams are based on flow cytometric analysis of wt mouse thymocytes.

Suppl. Fig. 2. Interspecies comparison of ThPOK non-coding exons. a) Distal non-coding exons. b) Proximal non-coding exons. Dashes denote sequences that are genuinely lacking in the indicated species, whereas n's indicate artificial sequencing gaps due to incompleteness of genomic data. In the wallaby sequence, the entire proximal exon falls into such a sequencing gap. Blue and red type indicate identity between mouse and all marsupials or difference between mouse and at least one marsupial, respectively. Start site of each mouse exon and 3' splice sites are based on publicly available EST sequences.

Suppl. Fig. 3. High conservation of ThPOK silencers between marsupials and between placental mammal species. CLUSTAL alignments of ThPOK silencers from **a)** 2 placental mammals (mouse vs human), and **b)** 2 marsupials (opossum vs. Tasmanian devil). ThPOK silencers of human and mouse show 87% identity, and opossum (S. American) and Tasmanian devil (Australia) marsupials show 85% identity. Highly conserved syntenic block region is indicated by red box. **c)** Top panel shows alignment of highly conserved non-coding element (HCNE) for indicated placental mammal species with respect to mouse ThPOK silencer, according to ANCORA database (http://ancora.genereg.net/cgi-bin/gbrowse/mm10/?name=chr3:89377644..89394776). Genomic coordinates at top of panel refer to Mouse (mm10) genome assembly, coordinates at bottom to mouse ThPOK silencer numbering given in Fig. 2, above. Bottom panel shows alignment of indicated mouse and human HCNE sequences (note ANCORA analysis is for reverse complement compared to Figure 2). d) Anti-Runx3 chromatin immunoprecipitation (ChIP) of sorted peripheral CD4 T cells from wt or indicated homozygous mutant mice. The PCR primers used span the ThPOK silencer. Results are typical of 3 independent experiments.

Suppl. Fig. 4. Experimental validation of predicted TF binding to mouse ThPOK silencer. **a)** 12 different mouse ThPOK silencer fragments (indicated by thin black bars) were tested as bait in Y1H analyses against an array of 1,086 different mammalian TFs and chromatin binding factors. Factors that bind to each fragment are listed above the corresponding line. Factors that are also predicted to bind to a particular fragment by JASPAR algorithm are marked in red (along with predicted coordinates in brackets). **b)** ChIP analysis with antibodies against indicated transcription factors for 3 mouse cell lines (CH12, B-cell lymphoma; G1E, erythroid line; MEL,

CML-like leukemia cell line). Yellow indicates factors whose binding is predicted by JASPAR analysis. ChIP data is derived from UCSC Genome Browser (NCBI37/mm9 mouse genome assembly).

Suppl. Fig. 5. Strategy for in vivo mutagenesis of ThPOK silencer. a), b), c) Generation of Sil^{Δ} allele. Site-specific ZFN and homologous repair (HR) construct that lacks the ThPOK silencer are coinjected into single-cell +/+ oocytes (a), leading to generation of a double-stranded break at the ZFN site located 50bp upstream of the ThPOK silencer, which undergoes HR-mediated repair to generate mice carrying the Sil^{Δ} allele (b). The Sil^{Δ} allele causes dominant constitutive expression of ThPOK, leading to exclusive generation of CD4 T cells in Sil^{Δ}/+ mice, as revealed by FACS analysis of peripheral blood (c). d) Generation and phenotyping of mice containing mutant Silencer alleles. Site-specific ZFN and homologous repair (HR) construct that contains variant ThPOK silencer are coinjected into single-cell Sil^{Δ} /+ oocytes, leading to generation of a double-stranded break at the ThPOK silencer, which undergoes HR-mediated repair to generate mice carrying the mutant Silencer allele. If the mutant Silencer construct is inserted into the Sil^{Δ} allele, it will potentially correct dominant constitutive expression of ThPOK, leading to restoration of CD8 T cells in Sil^{Δ} /+ mice (bottom right panel in d). In panels c and d, yellow boxes indicate the mouse ThPOK silencer, while red and green boxes indicate adjacent arms of homology included in targeting constructs. The blue box indicates the opossum ThPOK silencer.

Suppl. Fig. 6. Generation of ThPOK-Sil^{POS} mice. a) PCR typing of panel of founder mice derived from oocytes injected with the opossum-silencer knockin construct. One of PCR primers used is specific to results with primers specific to the opossum silencer. Note that 2 of 17 mice show a specific band (whole panel of 30 founders yielded 3 positives). b) FACS analysis of peripheral blood samples from partial set of 13 founders mice derived from oocytes injected with the opossum-silencer knockin construct. Note that two PCR positive mice from panel a (PL33 and PL44) show restoration of CD8 T cells, indicating that ΔSil allele was corrected by homologous recombination.

Suppl. Fig. 7. Marsupial ThPOK silencer supports normal development of MHC class I-restricted thymocytes. FACS analysis of CD4, and CD8a expression in indicated thymocyte and lymph node subsets from mice lacking expression of MHC class II. Plot at right shows % of SP CD4 and CD8 cells within gated TCRβ+ lymph node fraction (n = 3, for each strain). Error bars represent standard deviations. Significant differences between mutant and wt mice were determined by paired T test, and indicated by asterisks (* $p>0.01$; ** $p> 0.005$; *** $p> 0.001$).

Suppl. Fig. 8. ThPOK silencer is transcribed in opposite orientation to ThPOK gene in multiple different cell types. a) Organization of ThPOK (Zbtb7b) gene and reverse transcripts (including AK161599, AK136654) that originate near 5' end of ThPOK gene. b) Transcripts defined by RNA-seq analysis for indicated cell types, revealing reverse transcripts overlapping with the ThPOK silencer. MaGland = mammary gland; SmInt = small intestine. Modified from USCS genome browser (http://genome.ucsc.edu/).

Suppl. Fig. 9. Altering motif grammar of ThPOK silencer disrupts silencing function. FACS analysis of CD4, and CD8a expression in indicated thymocyte and lymph node subsets from wt (+/+), heterozygous and homozygous Sil^{ProxRx} mice lacking expression of MHC class II. Plot at right shows % of SP CD4 and CD8 cells within gated TCRβ+ lymph node fraction (n = 3, for each strain). Error bars represent standard deviations. Significant differences between mutant and wt mice were determined by paired T test, and indicated by asterisks (* p>0.01; ** p> 0.005; *** p> 0.001).