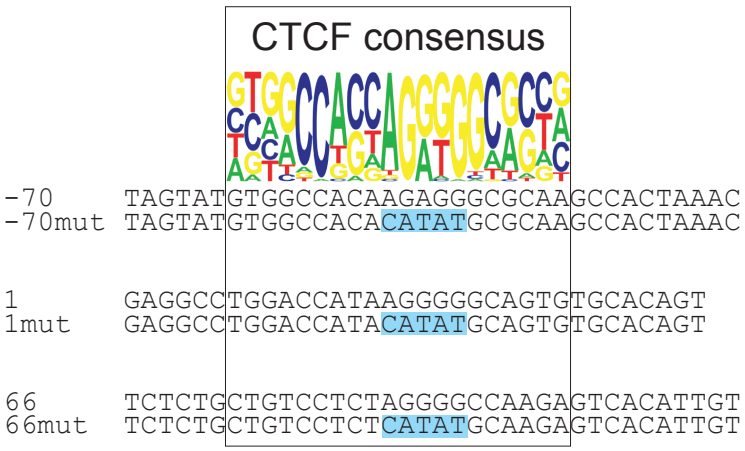
Supplementary Information

Supplementary Figures

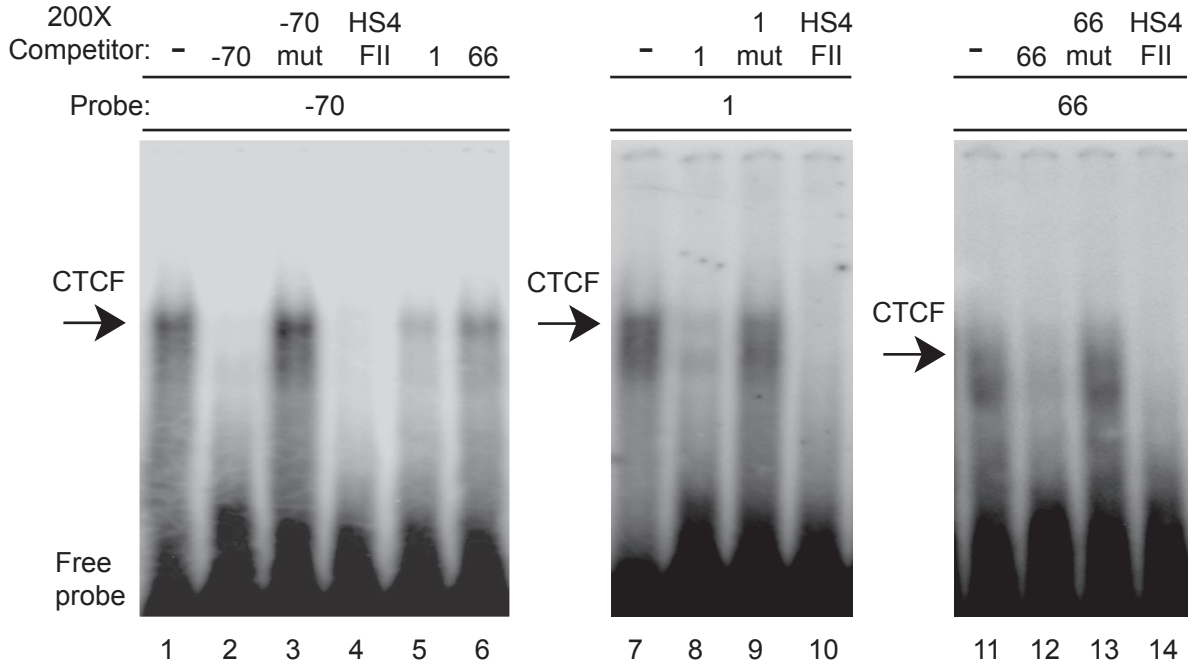
A

```
-70 kb site
Mouse    CCACAAGAGGGCGCAA
Human    CCACAGGAGGGAGCCA
         || | :| |||:  :|
Chicken CCGCTAGGGGGCAGCA
(5'HS4 FII)

1 kb site
Mouse    CCATAAGGGGGCAGTG
Human    CCATTGGGGGGCAGTA
         ||  ::|||||||| :
Chicken CCGCTAGGGGGCAGCA
(5'HS4 FII)

66 kb site
Mouse    CCTCTAGGGGCCAAGA
Human    CCTCTAGGGGCCATTA
         || ||||||| ||  |
Chicken CCGCTAGGGGGCAGCA
(5'HS4 FII)
```

B



C



Supplementary Fig. 1. Electromobility shift assays of CTCF binding to sequences in the *Ifng* locus.
(**A**) Aligments of conserved CTCF sites in the *Ifng* locus. Sixteenbases of the Chicken 5'HS4 FII CTCF site are aligned with conserved mouse *Ifng* and human *IFNG* sequences. Vertical lines indicate matches with both the human and mouse sequences, while dots indicate a match with only one of these species. (**B**) Sequences of the oligonucleotides used for EMSA probes are shown below the 20 bp human CTCF consensus determined by Kim *et al*. (Kim *et al*., 2007). The CTCF mutants used for mutant EMSA probes are highlighted. (**C**) EMSA using mouse recombinant CTCF protein and the -70, +1 or +66 kb CTCF-binding sites as probes. In competition assays, a 200-fold molar excess of the indicated wildtype competitior oligonucleotides or mutant oliginucleotides was added to the binding mixture.

CTCF-70

>mm8_dna range=chr10:117773383-117774057
GAGAAGGAAG AATTTAACAT GGAGGGGAAG AGGTGAGCCC TGTGGAGGAC
CCTTGGAGAG TTGCCAGAAA GATAGCTGCT CTGGGTCGTT TAAAATGGGA
ATTTGGAACA TGTCCTCAAT ATTTTCTTCT CATGCAGGAA TCTTCAGTTT
CTGTTCCGAT GCATCTCCTC TGTCAGTTTT GGAGCTTTTG TTGGTCTAGC
TCCACCCACA AGGTGGGGCA GTGGTATGCT GAATAGTGTT GCAGTTTAGT
GGCTTGCGCC CTCTTGTGGC CACATACTAG CATAGCACCG ACCAATCTCC
TTCCACCTCA AGCCTGCTTC GGTAAATTTG TTCCACAAAC CATGGCCAGC
GCTCTGGTAT AAAGCAATTT CCCCGGGAAG TAGAACGTCA CAGCATCGGA
CTGATATTTG TATTGCCTCC TCAGGCATTG CGGTGGAAAA TTAGACTCTG
GCTACCTGTT TAATTTCCCT CTGATCATAA GTTTATATTA TTTGACAGAA
AACTTGGTAG CAGTTAAACA TAAATTTTAT TGAGCATTTT CTGTACAAGA
AATTCTTATT TTTATTCGGG TCAGAGAGTA TCAGAGAGAT TTTATAAGAC
ATAGAATGTG GCATTAATTT AATGTTTCCT TGATAAAAAC AGCTACATAG
GGGTAAAACA AACAGTATAT GCTGT

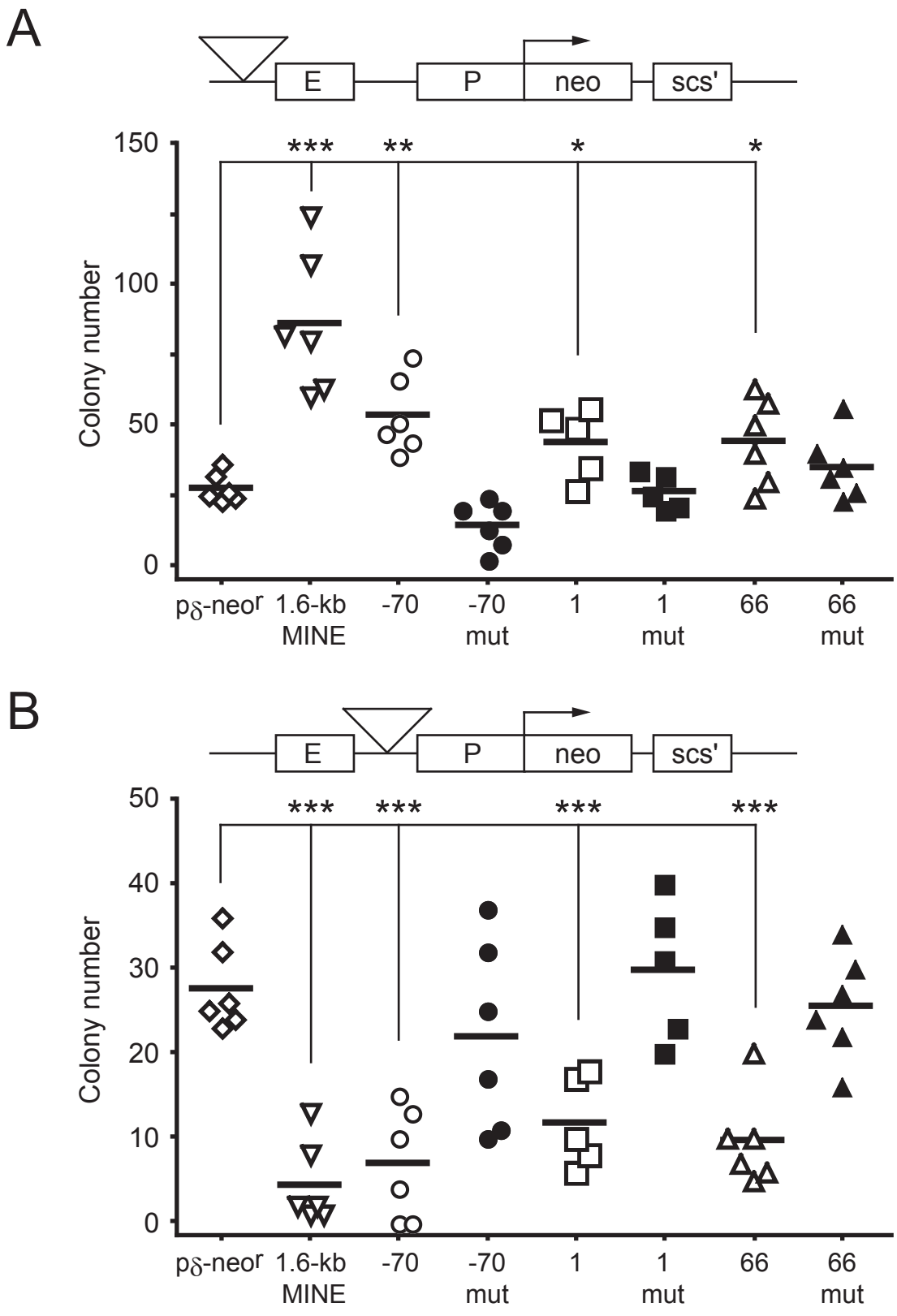
CTCF+1

>mm8_dna range=chr10:117844645-117845247
GGGAAACAGC TTATCACTTA AAGTAATGGG AGGCAAAGTT CTTAATGTTT
CTGTTTAAGA AATGCCCCCC TCCCCCGAAG TTGTTTTACC CTGAGGAGTA
ATTTTGAAAT TCTACTACAA TCGAATCAAA GCCAGCCGCT TGGTGAGGTT
TAATGAGACT CTATATTTCT CAGAGTTTTC AGGATTTTCT CTTTTGTGAT
TTTTTTTTTA AATTAATGGT TAACAGAAGC AAATTTTATA TCCTTAGAAT
TTTATAATTT GTCCCAAATG TATATACTAG TATACATATA GCTATTAGCA
TAATGTTCAA GCATCTACGG TCAATCCTCT CCTCACAGTG TAAATCAAGC
TGCCTCCCGT ATGTGTTTGG AGCTATTTTT TAAAGTAGCA ATGAAGCCCT
ATTACAGCAC AGACTGATGT TTCAGAGGCC TGGACCATAA GGGGGCAGTG
TGCACAGTGG GATGATTCTA ATGATTTTCT CTCTCCCTTT CTTAATTTGA
GAGTAAGTTT CTTTTTAGCT CACTATGGTT ACTTGAGGGT ACTTGAAACC
TCGTCAGCTT TGTTGAGTTT ATTTGTGGCC TTGCAGTTTC AAGACTGGGC
CAG

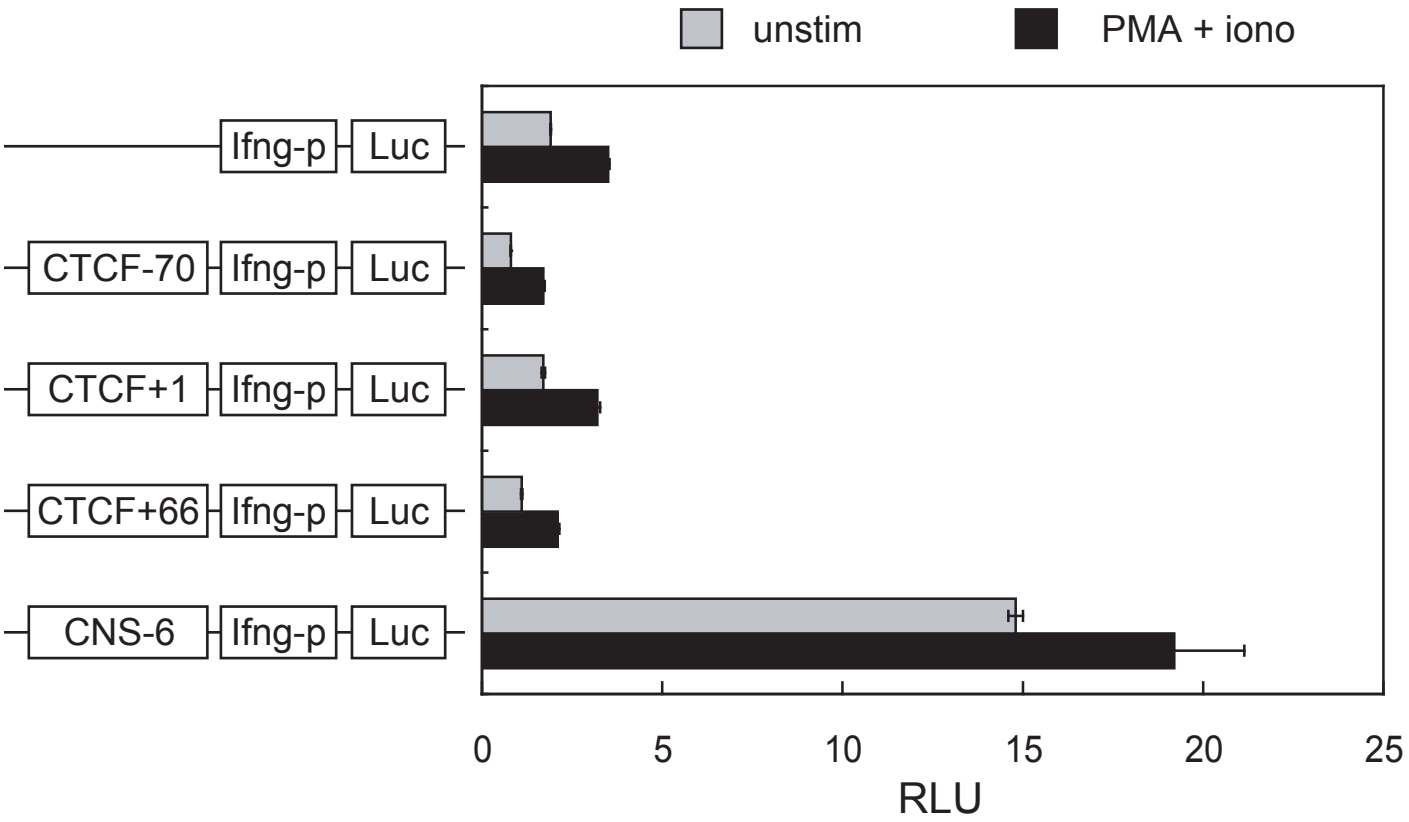
CTCF+66

>mm8_dna range=chr10:117910510-117911268
GGTCACAGGG AGACAGGGCA GCACAGTCCT TAGCCAAGGT CAACTGGACA
GACCTCCGGT TCCGCCTTCT CAGAAGTGTG TCTCACAGCC ATGCCTGGAT
CCTTCCTATC AGGGCGCTTG CCCATGTCCT AACTTTCCTG ACACTGATTA
TTGGGGCTAT TATAATGACC TCCCTAACTA CAGTGTGTAA TGTTCCTGGT
TCAGCTTCCA TGTTGCTTGG CAAGTAGCAG CTTTCTCCTC ACCTGCTTTC
TCCTTGATAT CTGATACGAA TCTGTCACTG TGCATGCCTG TAATCATCTC
TCTGCTGTCC TCTAGGGGCC AAGAGTCACA TTGTTCCACT TGTATCCCTG
AGGTGGCAGA CAAGAAGTGG GAGAGGCTCA CTTCCAGGAA AGAACTATAC
TGAATGCTGC ACAGCATCCA TGTCCTCTCT TGCCCAGGAC AGTCCAGTGA
GGCACTTGTT AGTATTTGCT CCACCTTAAA GATCAGACAG CCAAGTCTGA
AAACTGAGTA CCTTACCCAA GCAATGCCTC ACCCCGTGGA AGGAAGGCAA
GGTAGCCTTG TTTAGAGTTT GTGCTGAGTG ATTTCTCTTT GGTGGAGTGG
GGAAGCGAGG TAGGGTTGGC ATCAGGGCAA AATATGTTTA TTTGCATAAT
TTATTTCTCA AGATGGGTTT TACCTTGTAT TTTGCTTACA CACCAGCTGC
CTAACTGGTC TCCGTTGATA GTCTTTTTTC TTCCCAAGCT GTTATTCAGG
AGCTAGAGG

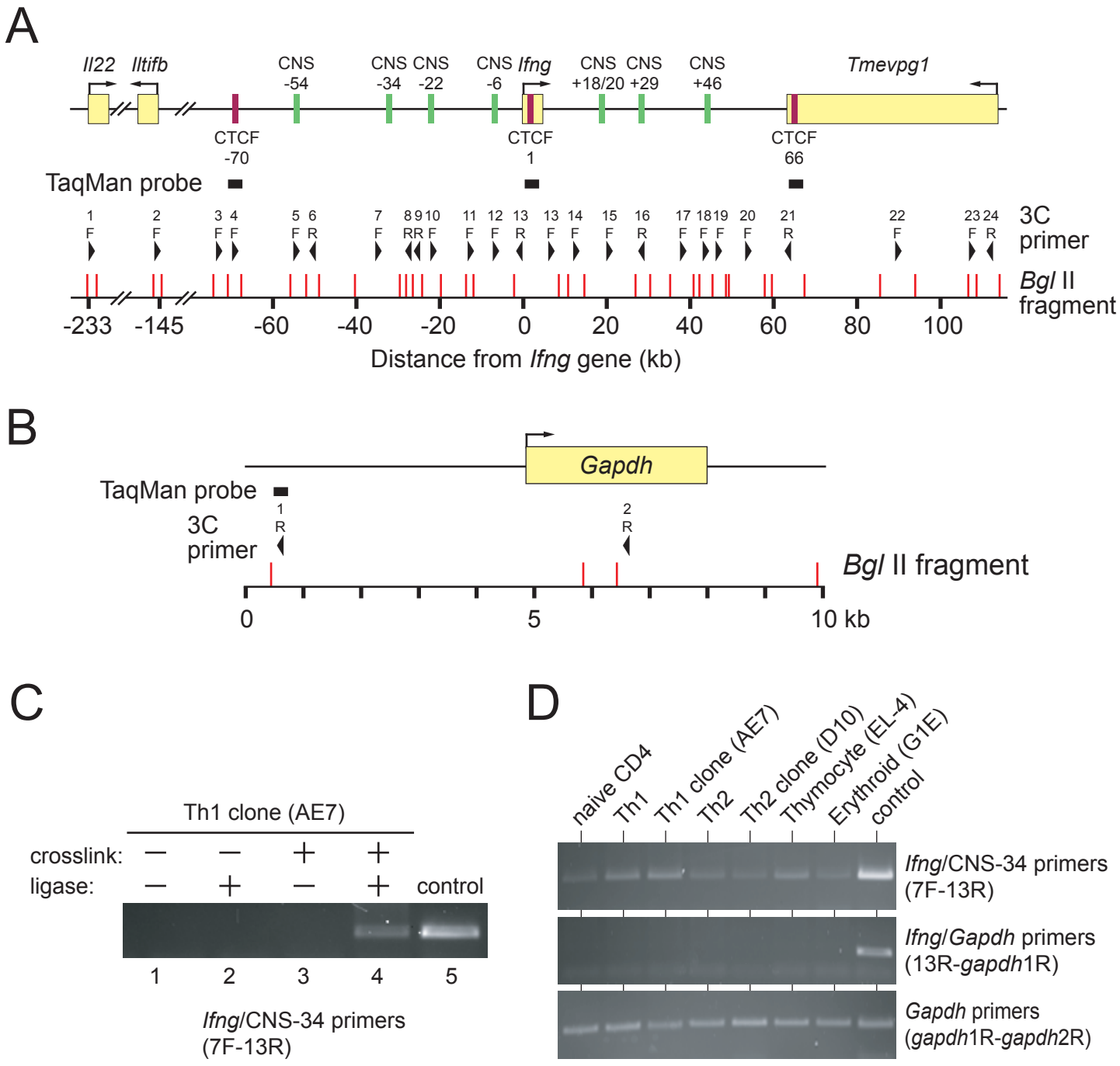
Supplementary Fig. 2. Sequence of 3 CTCF binding elements in the mouse *Ifng* locus.
The CTCF binding sites (yellow) and the CpG dinucleotides (blue) are highlighted, and the annealing sites of the bisulfite sequencing primers (Supplementary Table 8) are underlined.



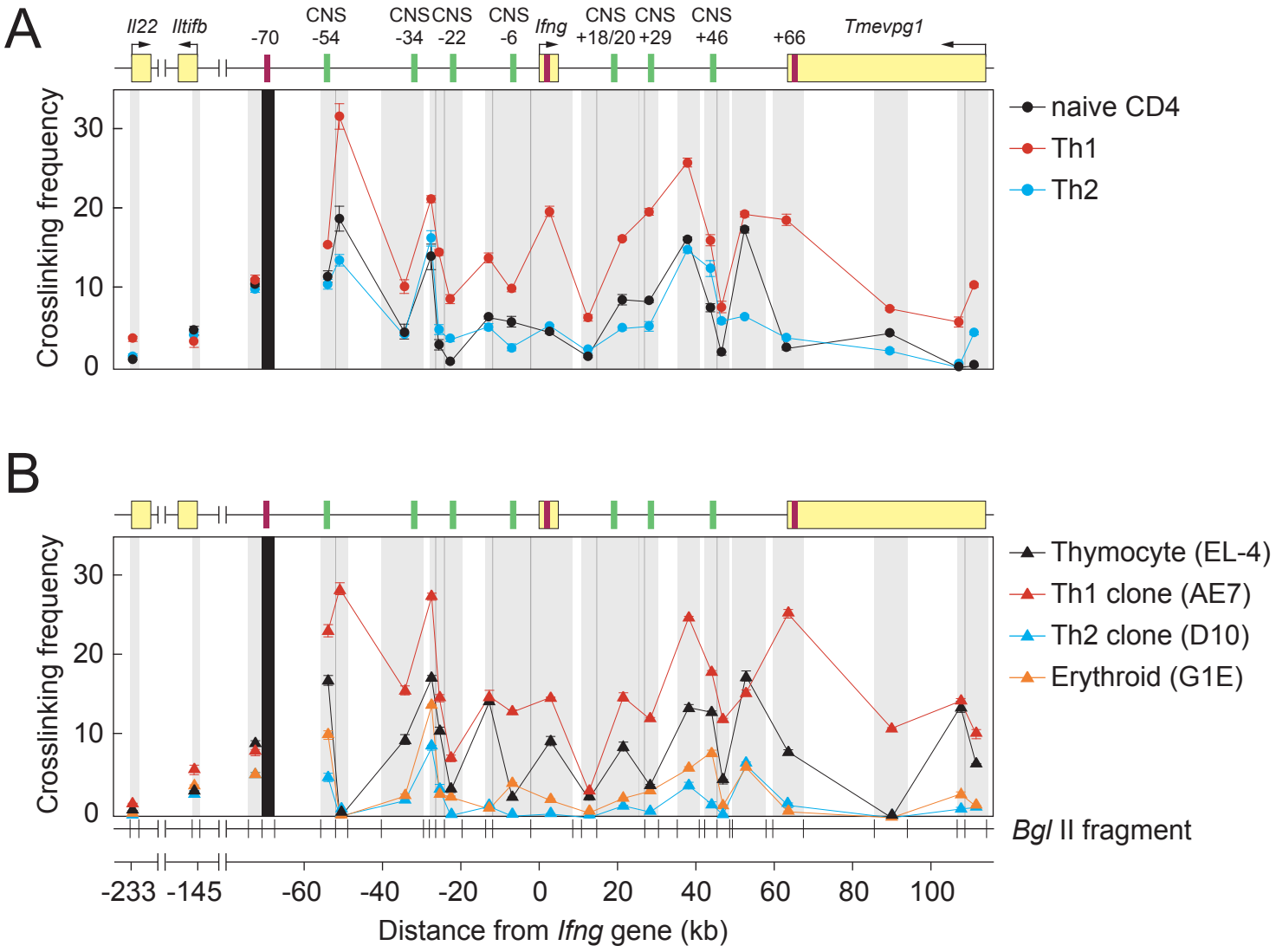
Supplementary Fig. 3. Insulator function of the CTCF-binding sites. EL-4 cells were transfected with reporter consrtucts (Pδ-neo^r) containing the bacterial neomycin resisitant gene (neo) under control of the promoter (P) and enhancer (E) of the human T cell receptor δ gene, flanked 3' by the drosophila insulator elememt (scs'). The number of G418-resistant colonies was determined by soft agar assay as described (Schoenborn *et al*., 2007). Positions where test fragments are introduced are marked by triangles. The human *c-MYC* insulator element (1.6-kb MINE) was used as a postive control. The fragments containing mutated CTCF sites (shown in Supplementary Fig. 1B) are shown as filled symbols. (**A**) Boundary element assay. (**B**) Enhancer-blocking assay. Small horizontal bars indicate the mean. Data represent a compilation ofthree independent experiments. ***, $p<0.0008$, **, $p<0.002$, and *, $p<0.03$, versus thePδ-neo^r control.



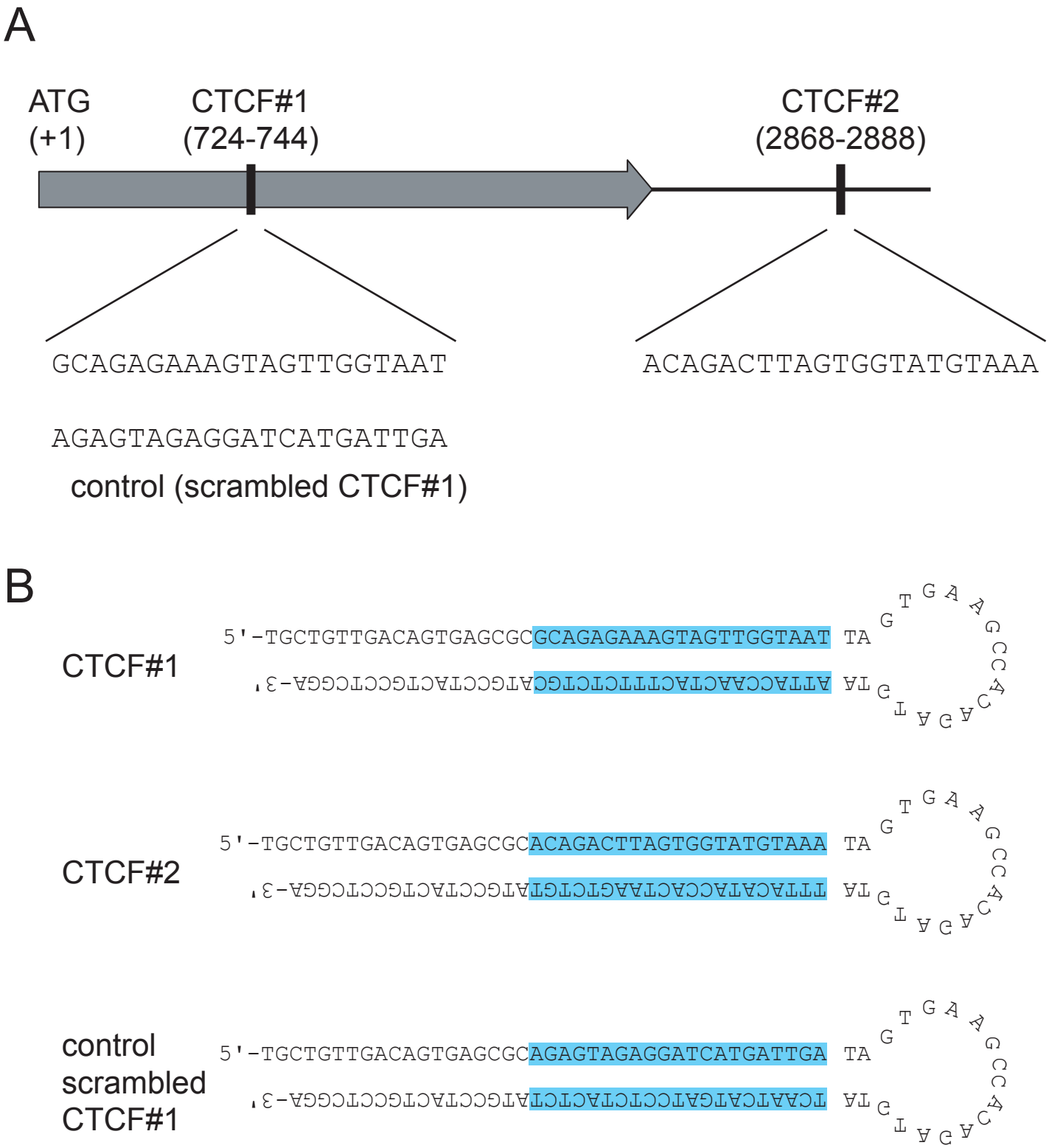
Supplementary Fig. 4. The CTCF binding elements exhibit no enhancer activity for the *Ifng* promoter. Luciferase assay of EL-4 cells transfected with reporter constructs containing the *Ifng* promoter and the indicated CTCF binding elements or *Ifng*CNS-6, assessed by dual-luciferase assay in cells unstimulated (gray bars) or stimulated with PMA plus ionomycin (filled bars), presented as firefly luciferase units normalized to renilla luciferase units. Data (mean ± s.d. of triplicate samples) are from one representative experiment of three independent experiments.



Supplementary Fig. 5. Strategy used to assess the 3-dimensional conformation of the mouse *Ifng* locus by 3C and examples of control experiments used to validate the specificity of 3C products. (**A**) *Bgl*II restriction map of the *Ifng* locus spanning a region of 350 kb and a scheme showing positions of regulatory elements. TaqMan probes used for the 3C assay are shown by rectangles. Numbered arrowheads indicate the site and direction of the primers used for the 3C assay. The red vertical lines above the scale show positions of *Bgl*II restriction sites. (**B**) The *Gapdh* locus. Schematic diagram shows the primers used in the assay and the enzyme sites (red vertical lines). TaqMan probes used for the 3C assay are shown by rectangles. Numbered arrowheads indicate the site and direction of the primers. (**C**) Example of electrophoretic analyses of 3C PCR products generated using primer pairs for fragments CNS-34 and *Ifng* on non-crosslinked genomic DNA and crosslinked genomic DNA with or without ligation. A 3C product was observed in AE7 Th1 cells when formaldehyde crosslinking and T4 DNA ligase were included during the assay (lane 4) and in the positive control consisting of ligated BAC DNA (lane 5). There are no random intramolecular ligation products without formaldehyde crosslinking and T4 DNA ligase (lanes 1-3). (**D**) Example of electrophoretic analysis of 3C PCR products using primer pairs for fragments containing CNS-34 and *Ifng* using as templates DNA from the cell types shown or control BAC DNA (top). Primer pairs for the *Ifng* CNS-34 fragment and one of the *Bgl*II *Gapdh* fragments resulted in no product (middle). *Gapdh* products obtained with primer pairs for the two *Bgl*II fragments on the *Gapdh* locus were used for normalizing signals (bottom).



Supplementary Fig. 6. 3C assay results using the -70kb CTCF element as anchor fragment. Relative cross-linking frequencies were measured between a fixed anchor fragment bearing the -70kb CTCF binding site and other *Bgl*II fragments on primary CD4+ T cells (**A**) and mouse cell lines (**B**). The format of the figure panels and methods of assessment are as in Fig. 3.



Supplementary Fig. 7. miR-30 based shRNA design for CTCF knockdown. (**A**) Schematic representation of CTCF hairpin targeting sequences (CTCF#1 and CTCF#2) and the control, which consisted of scrambled CTCF#1 sequence. Numbers represent the nucleotide positions of cDNA relative to the *Ctcf* translation start site. (**B**) Hairpin representations of the targeting sequences. The sense and antisense strands of the CTCF targeting sequences are highlighted on the hairpins.

SUPPLEMENTARY TABLES

Supplementary Table 1

Primers used for 3C assay

Location*	Name	Primer sequence (5' to 3')	mm8 coordinates
-235	1F	GGATGAGAAAGATGTTTGGAAC	chr10: 117609108-117609129
-145	2F	GACACTGGAGAAAGGGCTCA	chr10: 117698571-117698590
-71	3F	AGGACCTATACGCTGGCAGT	chr10: 117772724-117772743
-69	4F (anchor)	GCCCATACAGACTTTGAAGC	chr10: 117775291-117775310
-53	5F	AATGGTTCAGGTCTGGTTGA	chr10: 117791193-117791212
-51	6R	TCATAGCCCCATATTTCCAG	chr10: 117792885-117792904
-30	7F	GAGAGTGTGGGCTCCCTTCT	chr10: 117813682-117813701
-28	8R	GAGACACCTCCACTCTGCAT	chr10: 117816479-117816498
-27	9R	GCTTCTCTAATGAACTGTAATGTGC	chr10: 117817149-117817173
-24	10R	GCTTTGATCATAGCTTCATTATC	chr10: 117819704-117819726
-12	11F	TACACTCATTCCTGCCTGGT	chr10: 117832272-117832291
-3	12F	AGAATGTGCCATTAGAAGTC	chr10: 117841486-117841505
2	13R (anchor)	CAGATGTAAGATGGGATCTC	chr10: 117841650-117841669
7	13F (anchor)	GTGTATGCTCCGTGGCTAGT	chr10: 117850802-117850821
15	14F	GAAGCCTGTAGGGGAGAGAC	chr10: 117859041-117859060
27.8	15F	AGAATTCTGGCCTGTGATTG	chr10: 117871836-117871855
27.9	16R	TGAGCTCCAAACCTCTGTTC	chr10: 117871988-117872007
41	17F	GGTTTAGGAAAGCATTTGTCTG	chr10: 117884890-117884911
48	18F	CCTGGGTCTTTGTATGGTCTT	chr10: 117891637-117891657
51	19F	GACCCATAGCTCTTGCCTCT	chr10: 117895490-117895509
63	20F	GACATGTCCTGTGTTGTGTTC	chr10: 117907377-117907397
66	21R (anchor)	CCTCTGACAAATCTGCCTCT	chr10: 117910402-117910421
98	22F	CACTGAGGAGCCACCAGTAA	chr10: 117941599-117941618
111.2	23F	GTCCATCCTCTGCCTCAGTA	chr10: 117955258-117955277
111.3	24R	CTGGTGAATCTTCCCACCTA	chr10: 117955400-117955419
	Gapdh1R (anchor)	ACACAGGCAAAATACCAATG	chr6: 125135230-125135249
	Gapdh2R	GAATGCTTGGATGTACAACC	chr6: 125129297-125129316

*Location corresponds to the distance in kilobases from *Ifng* transcription start site (chr 10: 177844040 UCSC mm8, UCSC Genome Browser: http://www.genome.ucsc.edu/). F or R designate forward or reverse primers as compared to direction of *Ifng* gene or *Gapdh* gene transcription.

Supplementary Table 2

TaqMan probes used for 3C assay

Name	Probe sequence (5' to 3')
IFNG13R	6FAM-CCATAGTGAAAAGTCACATGGCTGAGAAACACTTC-TAMRA
IFNG13F	6FAM-AGATCTCTCAGACACTGACTGAGCC-TAMRA
GAPDH1R	6FAM-AGATCTTAATTCCTGGTCCCTT-TAMRA
-70CTCF4F	6FAM-AGTCACGTGGCCTGCCGTCCCTAC-TAMRA
66CTCF21R	6FAM-AACATAGTAAGTGGGGCCTATACCCTGTGCC-TAMRA

Probe names are assigned as in Supplementary Table 1.

Supplementary Table 3

Oligonucleotides used for shRNA retroviral construction

Name	Sequence (5’ to 3’)
CTCF#1	TGCTGTTGACAGTGAGCGCGCAGAGAAAGTAGTTGGTAATTAGTGAAGCCACAGATGTAATTA CCAACTACTTTCTCTGCATGCCTACTGCCTCGGA
CTCF#2	TGCTGTTGACAGTGAGCGAACAGACTTAGTGGTATGTAAATAGTGAAGCCACAGATGTATTTAC ATACCACTAAGTCTGTGTGCCTACTGCCTCGGA
scrambled CTCF#1	TGCTGTTGACAGTGAGCGCAGAGTAGAGGATCATGATTGATAGTGAAGCCACAGATGTATCAA TCATGATCCTCTACTCTATGCCTACTGCCTCGGA
mir30-f	CAGAAGGCTCGAGAAGGTATATTGCTGTTGACAGTGAGCG
mir30-r	CTAAAGTAGCCCCTTGAATTCCGAGGCAGTAGGCA

Supplementary Table 4

Primer sequences used for gene expression analysis

Name	Primer sequence (5’ to 3’)
CTCF-f	GACCACAAATCTAGAACCAAAGAAC
CTCF-r	GTTGGCTTCGGAGGCTTCATATTACC
actin-f	TCCTTCGTTGCCGGTCCAC
actin-r	ACCAGCGCAGCGATATCGTC

Supplementary Table 5

Probes used for EMSA

Name	Probe sequence (5’ to 3’)
HS4FII	CCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGGCAGCA
CTCF-70	TAGTATGTGGCCACAAGAGGGCGCAAGCCACTAAAC
CTCF-70mut*	TAGTATGTGGCCACACATATGCGCAAGCCACTAAAC
CTCF1	GAGGCCTGGACCATAAGGGGGCAGTGTGCACAGT
CTCF1mut*	GAGGCCTGGACCATACATATGCAGTGTGCACAGT
CTCF+66	ACAATGTGACTCTTGGCCCCTAGAGGACAGCAGAGA
CTCF+66mut*	ACAATGTGACTCTTGCATATGAGAGGACAGCAGAGA

* Oligonucleotides are also used to change the CTCF-binding sites of boundary element constructs.

Supplementary Table 6

Primers used to generate luciferase-reporter constructs

CTCF element	Primer sequence (5’ to 3’)
CTCF-70	F: TGGGTACCTTTGTTGGTCTAGCTCCACCC
	R: TTCTCGAGTTAACTGCTACCAAGTTTTCTG
CTCF+1	F: TGGGTACCTAGCATAATGTTCAAGCATCTACGG
	R: TTCTCGAGAGTCTTGAAACTGCAAGGCCACA
CTCF+66	F: TGGGTACCATTGGGGCTATTATAATGACC
	R: TTCTCGAGCAGACTTGGCTGTCTGATCTTT

Supplementary Table 7

Primers used to generate boundary element constructs

Region	Element	Primer sequence (5’ to 3’)
CTCF-70	Insulator	F: TTGCGGCCGCTTTGTTGGTCTAGCTCCACCC
		R: TTTCTAGATTAACTGCTACCAAGTTTTCTG
CTCF-70	Enhancer Blocker	F: CAATCGATTTGTTGGTCTAGCTCCACCC
		R: GGGTCGACTTAACTGCTACCAAGTTTTCTG
CTCF+1	Insulator	F: CGCGGCCGCTAGCATAATGTTCAAGCATCTACGG
		R: TTTCTAGAAGTCTTGAAACTGCAAGGCCACA
CTCF+1	Enhancer Blocker	F: CCATCGATAGCATAATGTTCAAGCATCTACGG
		R: TTGTCGACAGTCTTGAAACTGCAAGGCCACA
CTCF+66	Insulator	F: CGCGGCCGCATTGGGGCTATTATAATGACC
		R: TTCTAGACAGACTTGGCTGTCTGATCTTT
CTCF+66	Enhancer Blocker	F: CCATCGATTGGGGCTATTATAATGACC
		R: TTGTCGACAGACTTGGCTGTCTGATCTTT

Supplementary Table 8

BIS-PCR primers used for CpG methylation analysis

Region	mm8 Begin	mm8 End	Forward	Reverse
CTCF-70	117773383	117774057	BISCTCF-70F	BISCTCF-70R
CTCF+1	117844645	117845247	BISCTCF+1F	BISCTCF+1R
CTCF+66	117910510	117911268	BISCTCF+66F	BISCTCF+66R

Sequence of BIS-PCR primers

Name	Primer sequence (5’ to 3’)
BISCTCF-70F	GAGAAGGAAGAATTTAATATGGAGGGGAAGAGG
BISCTCF-70R	ACAACATATACTATTTATTTTACCCCTATA
BISCTCF+1F	GGGAAATAGTTTATTATTTAAAGTAATGGGAGGT
BISCTCF+1R	CTAACCCAATCTTAAAACTACAAAACCAC
BISCTCF+66F	GGTTATAGGGAGATAGGGTAGTATAGTTTTTAGT
BISCTCF+66R	CCTCTAACTCCTAAATAACAACTTA

Supplementary Table 9

Primers used for Chromatin Immunoprecipitation (ChIP) assay

Mouse

Location*	Forward Primer (5’ to 3’)	Reverse Primer (5’ to 3’)	UCSC mm8 coordinates
-70488	CTGTCAGTTTTGGAGCTTTTGTTG	GCAATACAAATATCAGTCCGATGC	117773552-117773798
-21836	GGTGATCCACAGGAAGGAGA	GAGCAGAAATTTGGCCTCTT	117822204-117822316
-49	GAGGAGCCTTCGATCAGGTA	CAAGTCAGAGGGTCCAAAGG	117843991-117844101
993	CAATGAAGCCCTATTACAGCACAG	TCTTTTAGCTGCAGGATGTACTGG	117845033-117845267
18436	TGGTCCCATACCTGTGAGTATTTG	TAAGACCACACCTCCACATAGTGC	117862476-117862727
29855	CGCTCAGTATACAGCCAGTCACTT	AAGCTCTAGCTGCCCTGATTAAAA	117873895-117874147
66684	GCTTGGCAAGTAGCAGCTTT	GGAAGTGAGCCTCTCCCACT	117910724-117910894

* Relative to the mouse *Ifng* transcription start site (chr10: 177844040 UCSC mm8)

Human

Location*	Forward Primer (5’ to 3’)	Reverse Primer (5’ to 3’)	UCSC hg17 coordinates
-63648	TCCTAAAGTGGGTGGGACAG	TTACCCCAGGGCATAAACTG	66903218-66903436
-31	AAAATGGAAGTGCCAACGAT	TGGTTTTGCATTCAGCACTT	66869952-66870153
-22	AACCATGGCTGAACTTCCTG	GGTGTAGACGGCCTAGAGCA	66860708-66860857
-16	GGGAGAAGAGCGATTTCCTT	CAAGGTGCCTTCCATCTCTC	66855735-66855900
-4	TCTTCCGTAGGTTTGGCTTG	CCTTGGTTTCCCTTTTCCTC	66843940-66844119
-16112	GGGAGAAGAGCGATTTCCTT	CAAGGTGCCTTCCATCTCTC	66855735-66855900
1069	ACCCTACAGGAAGGCACAGT	TGATACTCCAAAGGTCCCAAA	66838559-66838719
22	ATTTGCACACTGCTCACCAA	TCTACCGCTTTCCCACTGAC	66817457-66817635
39728	ACCCGAGAAGATGCCTACCT	CCACTTCCATCCAGTGCTTT	66799811-66800060
80	AAACATGGGAAGGGGAAACT	ACCAGCCCACCACTACAGAG	66760806-66761006
119009	GATTCACAGTGCCCTTCTCC	CCCCCAAGCATAAAGGATCT	66720638-66720779

* Relative to the human *IFNG* transcription start site (chr12: 66839788 UCSC hg17)

Supplementary Table 10

Primers used to generate a bacterial expression vector for the recombinant CTCF protein

Name	Primer sequence (5’ to 3’)
CTCFzf11-f	AGGGATCCAAAAAGGTGTAAAGAAAACATTC
CTCFzf11-r	GGGAGCTCTTTTCCCCCTCTACGCCATCTG