

Figure S1. CTCF binding to the WT *IgH* alleles, related to Figure 1.

Figure S2. Epigenetic features of WT and IGCR1-mutated *IgH* alleles-1, related to Figure 1.

Figure S3. Epigenetic features of WT and IGCR1-mutated *IgH* alleles-2, related to Figure 1.

Figure S4. CTCF- and YY1-dependence of E μ -V_H81X loop, related to Figure 2.

Figure S5. D_H gene segment utilization on WT and IGCR1-deleted or mutated alleles, related to Figure 3.

Figure S6. V_H recombination efficiency of WT and IGCR1-mutated *IgH* alleles, related to Figure 4.

Figure S7. RAG1/2 binding to IGCR1-deleted *IgH* alleles, related to Figure 5.

Figure S8. Long-distance interactions on IGCR1-mutated *IgH* alleles-1, related to Figure 6.

Figure S9. Long-distance interactions on IGCR1-mutated *IgH* alleles-2, related to Figure 6.

Figure S10. Features of pro-B cells with deletion of proximal V_Hs on WT or IGCR1-mutated *IgH* alleles, related to Figure 6.

Table S1. Average spatial distance and chromatin compaction in the *IgH* locus from different cell lines, related to Figure 1, 2 and 6.

Table S2: V_H7183 usage in IGCR1^{-/-} and WT pro-B cells, related to Figure 3.

Table S3. Primer list, related to Figures 1-6, S1-10.

Table S4. Barcode sequence for different cell lines, related to Figure 6.

Table S5. V_H utilization in different cell lines, related to Figure 6.

Figure S1

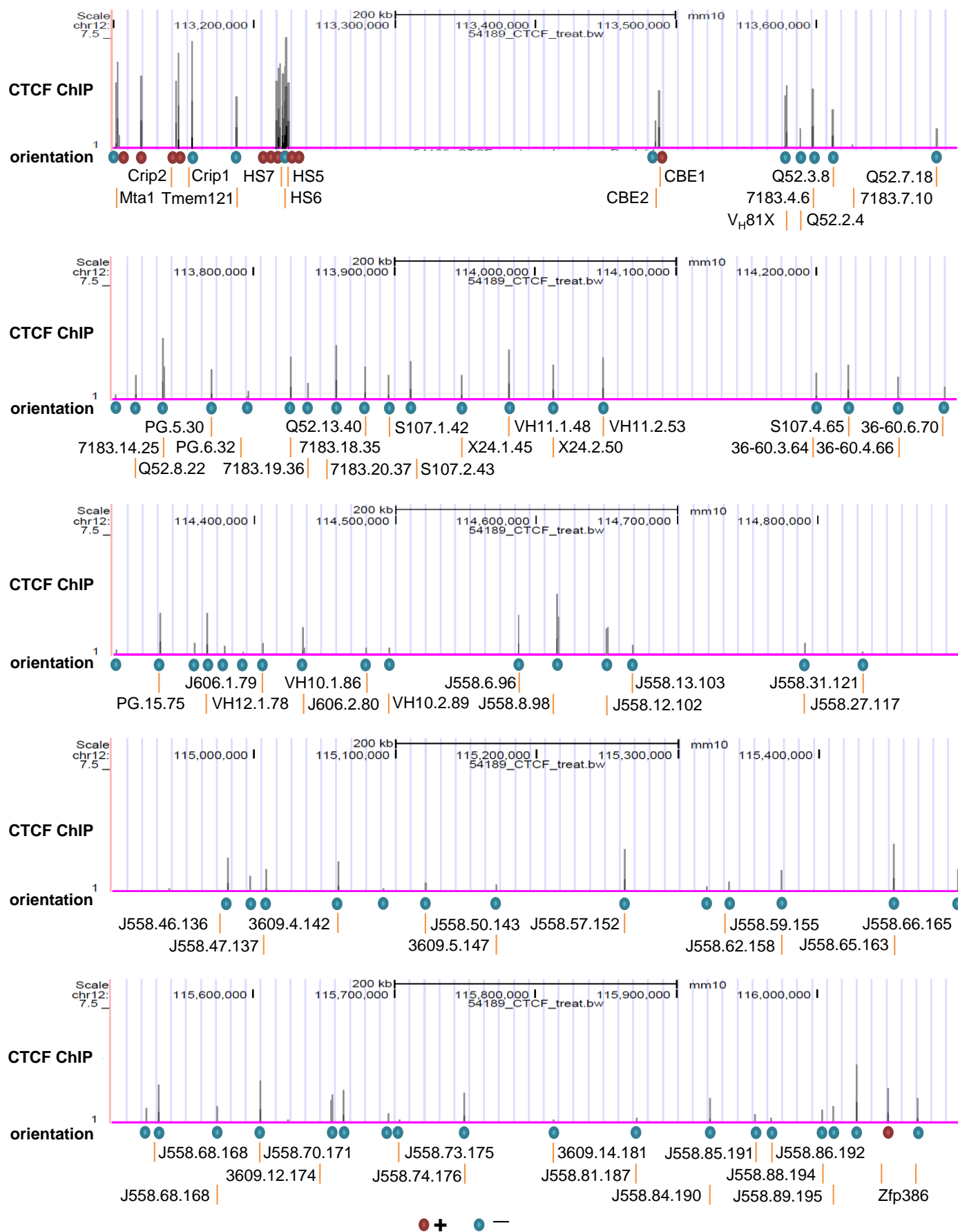
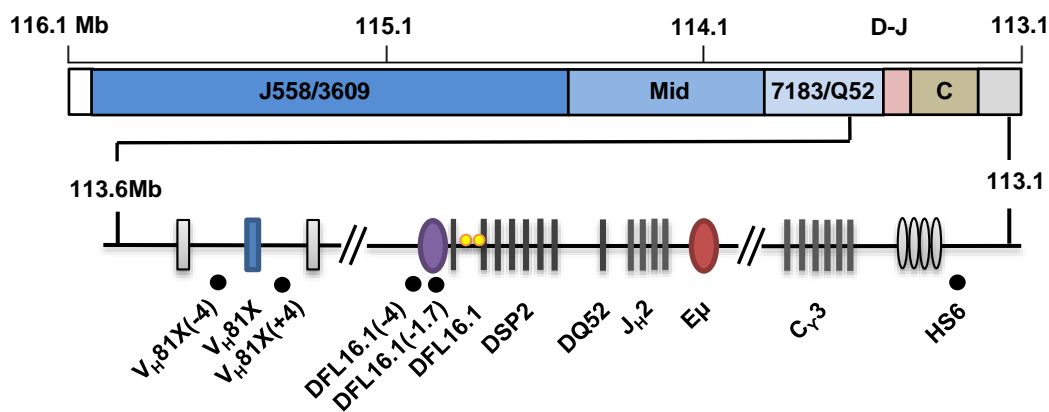


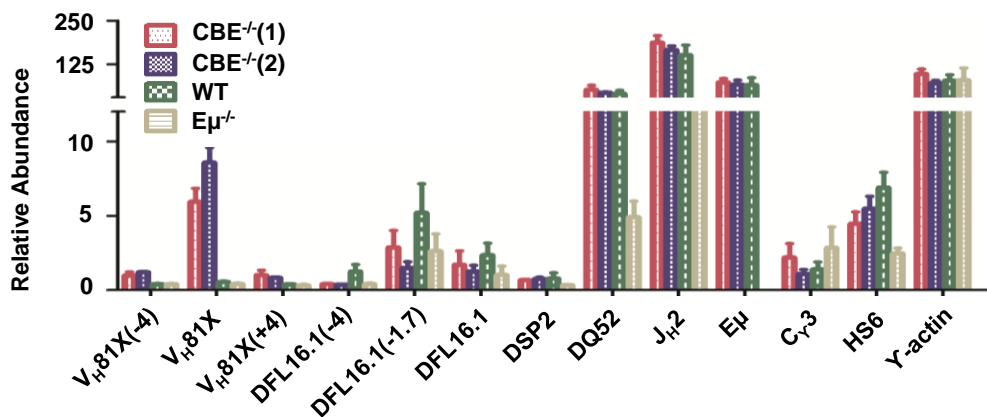
Figure S1. CTCF binding to the WT *IgH* alleles, related to Figure 1.

Orientations of CTCF binding sites in *IgH* locus on WT alleles. CTCF ChIP-Seq was extracted from (Lin et al., 2012) (GEO: GSM987805). CTCF binding to the *IgH* locus in pro-B cells is shown by vertical lines. Orientation of CTCF binding sites is noted by blue and red ovals. 11 sites are oriented with the C-rich core on the top strand (+) (red oval) and 75 are oriented with the G-rich core on the bottom strand (-) (blue oval). V_H gene segments are indicated by vertical pink lines.

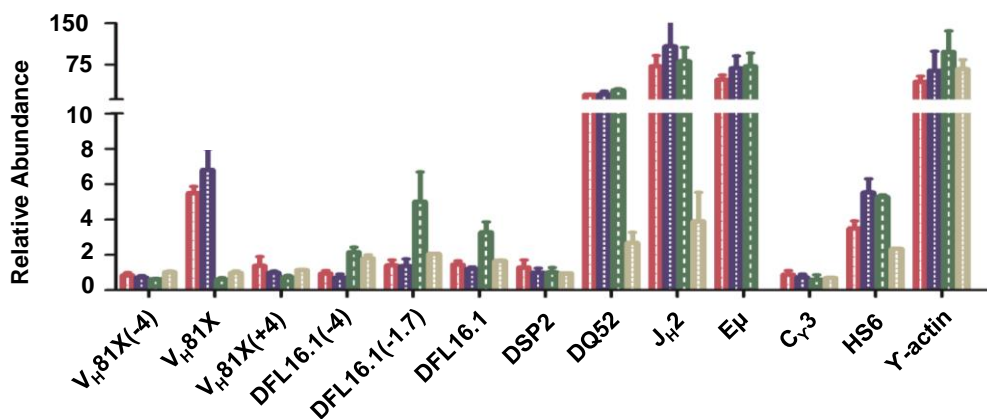
Figure S2



H3K4me3



H3K9ac



H3K9me2

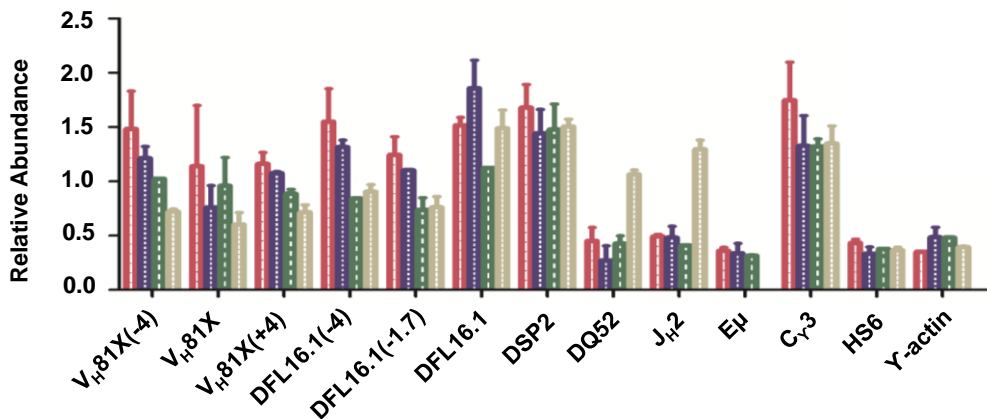


Figure S2. Epigenetic features of WT and IGCR1-mutated *IgH* alleles-1, related to Figure 1.

IGCR1-mutated *IgH* alleles. Four RAG2-deficient pro-B cell lines containing wild-type *IgH* alleles (WT), IGCR1-mutated *IgH* alleles (CBE^{-/-}(1) and (2)), and E μ -deficient *IgH* alleles (E μ ^{-/-}), were used in chromatin immunoprecipitation (ChIP) assays using anti-H3K4me3, anti-H3K9ac and anti-H3K9me2 antibodies as previously described (Chakraborty et al., 2007; Chakraborty et al., 2009; Subrahmanyam et al., 2012). Positions of amplicons are indicated in the schematic on the top line. V_H81X(-4) and V_H81X(+4) are located approximately 4 kb 5' or 3' of V_H81X, respectively. HS6 is located 5 kb from HS4. Relative abundance (Y axis) of each amplicon in the immunoprecipitate was calculated relative to an equal amount (200 pg) of input DNA. Data are shown as mean \pm SEM of two or three independent ChIP experiments in each cell line. ChIP amplicons were assayed in duplicates.

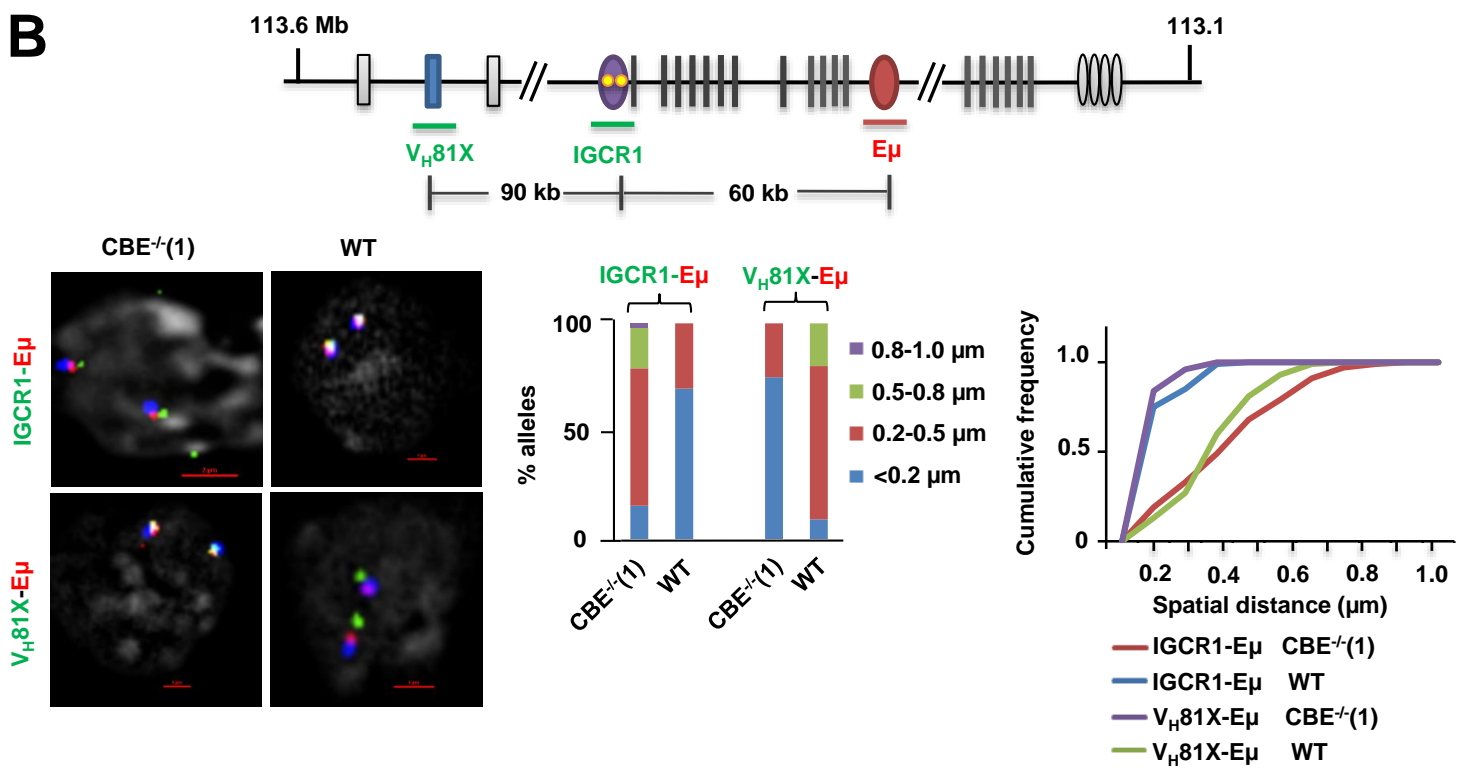
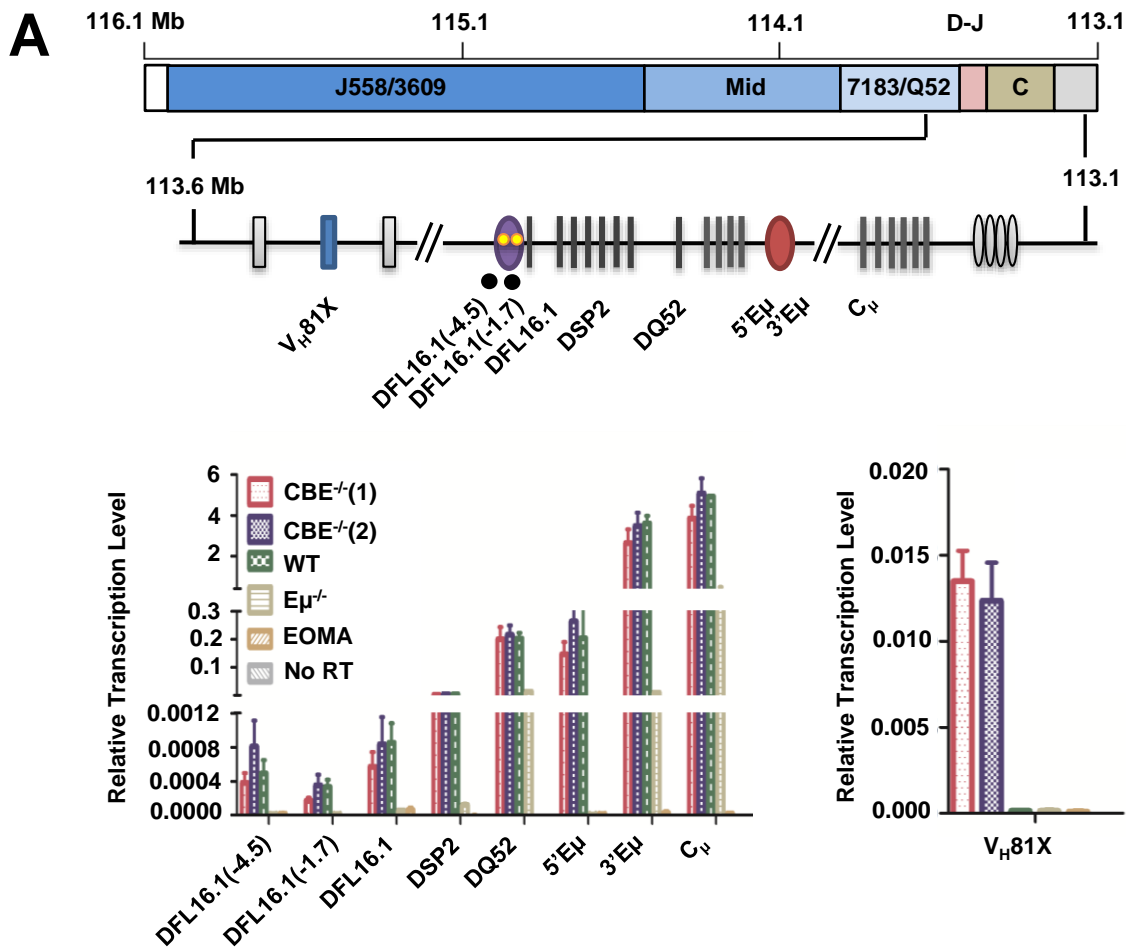


Figure S3. Epigenetic features of WT and IGCR1-mutated *IgH* alleles-2, related to Figure 1.

A. Transcription of WT and IGCR1-mutated *IgH* alleles. Total RNA from 4 cell lines described in Figure S2, together with EOMA (a non-B cell line of endothelial origin from 129 strain), was reverse transcribed with random hexamer primers and assayed by qPCR. No RT served as a control. Positions of amplicons are indicated in the schematic on the top line. DFL16.1 (-4.5) and DFL16.1 (-1.7) are located 4.5 kb and 1.7 kb 5' of DFL16.1, respectively. Relative transcript level (Y axis) was calculated according to the formula $2^{-(Ct_{\text{gene}} - Ct_{\gamma\text{-actin}})}$, with γ -actin mRNA serving to normalize between samples. Data are shown as mean \pm SEM of two independent RNA preparations from each cell line.

B. FISH analyses of WT and IGCR1-mutated alleles. Biological replicate of the experiment described in Figure 1C. Spatial distance measurements and quantitation was done as described in Figure 1C legend.

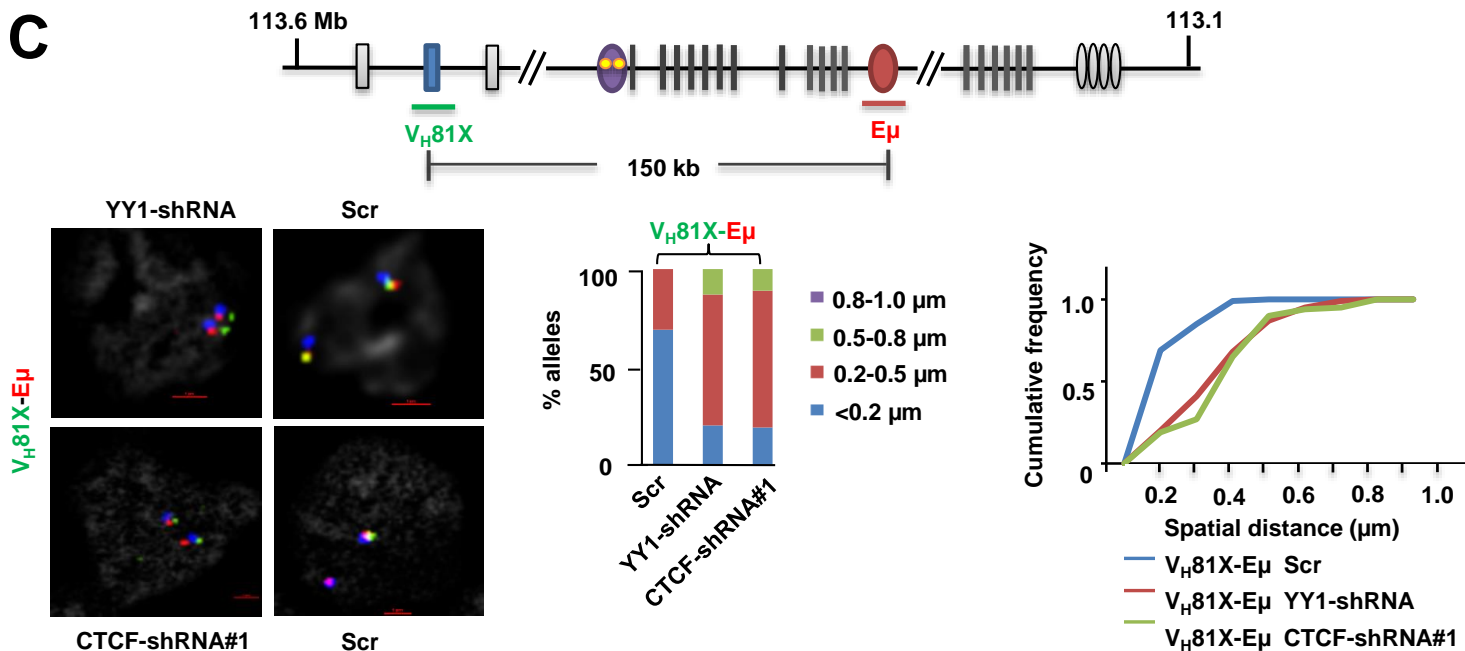
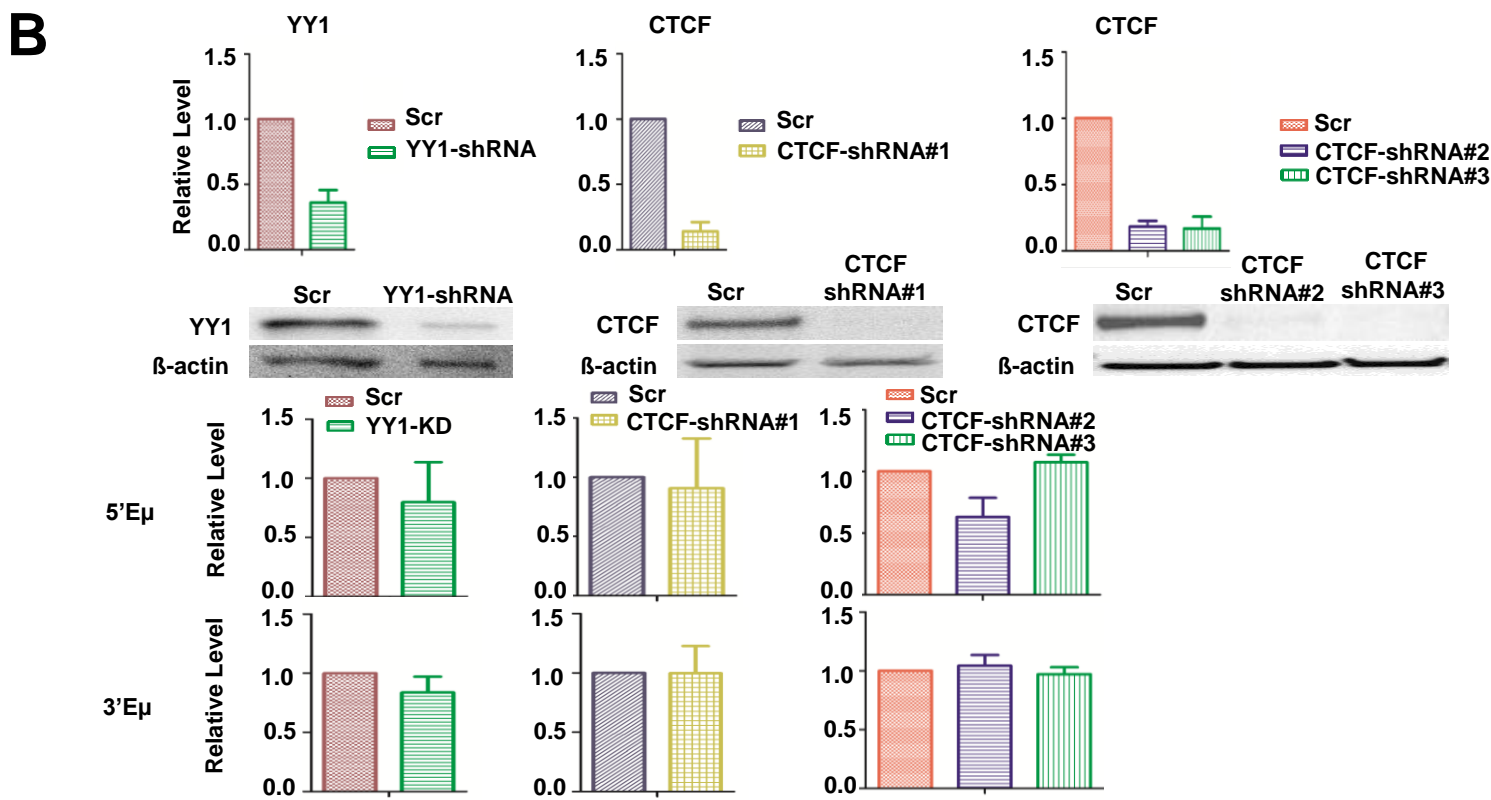
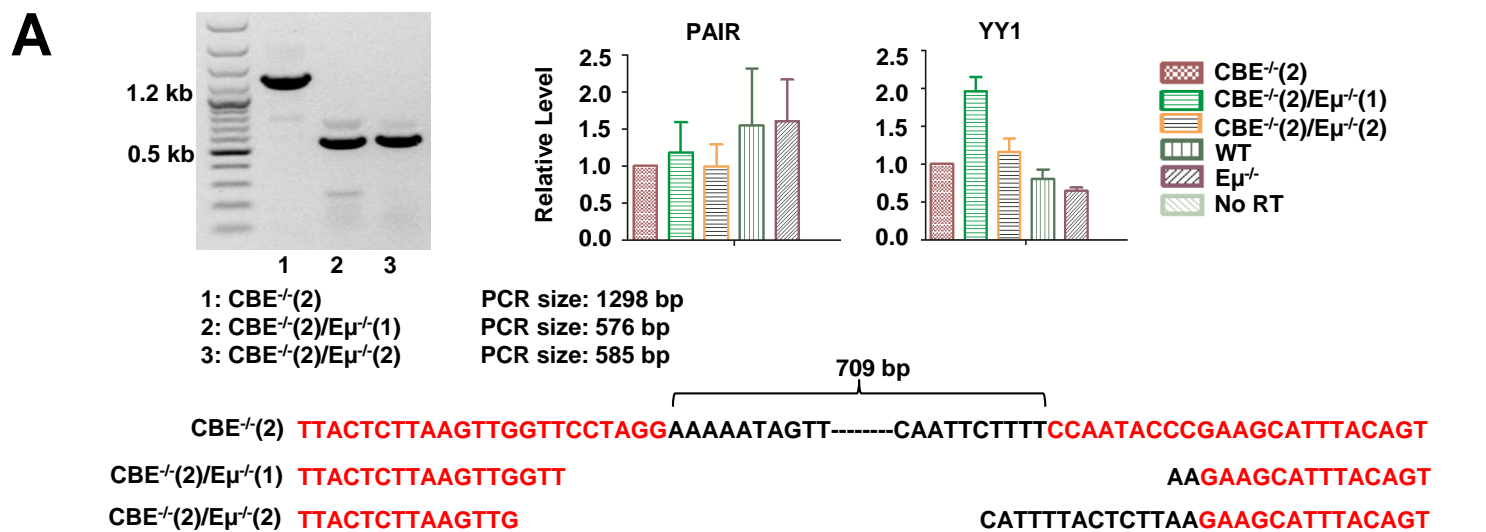


Figure S4. CTCF- and YY1-dependence of E μ -V $_H$ 81X loop, related to Figure 2.

A. Generation and characterization of IGCR1-mutated and E μ -deleted *IgH* alleles. gRNAs (Table S3) were used to delete E μ in the CBE^{-/-}(2) cell line that contains IGCR1-mutated *IgH* alleles. Two subclones were further characterized for functional studies. Homozygous E μ deletion was ascertained by PCR followed by sequencing of the deletion junction (lower panel). Total RNA isolated from CBE^{-/-}(2), CBE^{-/-}(2)/E μ ^{-/-}(1), CBE^{-/-}(2)/E μ ^{-/-}(2), WT or E μ ^{-/-} cells was assayed by quantitative RT-PCR using amplicons from *IgH*-associated PAIR elements (Ebert et al., 2011) and *Yy1* gene. No RT refers to cDNA synthesis without adding reverse transcriptase. Data are shown as mean \pm SEM of two independent RNA preparations from each cell line.

B. Characterization of YY1- and CTCF-knockdown (KD) cells. CBE^{-/-}(2) cells were infected with lentiviruses carrying shRNAs targeted to *Yy1* or *Ctcf* mRNA, or a scrambled (Scr) shRNA to serve as a negative control. After 5 days selection in puromycin, Total RNA isolated from KD or control cells was assayed by qPCR by quantitation of *Yy1*, *Ctcf*, and E μ transcripts (5'E μ and 3'E μ expression). Transcript levels were normalized to γ -actin mRNA and are represented relative to levels in control cells (Y axis). Data are shown as mean \pm SEM of two or three independent knock-down experiments followed by RNA analysis. Whole cell extracts from KD or control cells were assayed by immunoblotting using antibodies against YY1, CTCF or β -actin. β -actin served as a loading control. Data are shown as mean \pm SEM of two independent knockdown experiments.

C. FISH analyses of E μ -V $_H$ 81X loop in KD and control (Scr) cells. Biological replicate experiments were done as described in Figure 2B.

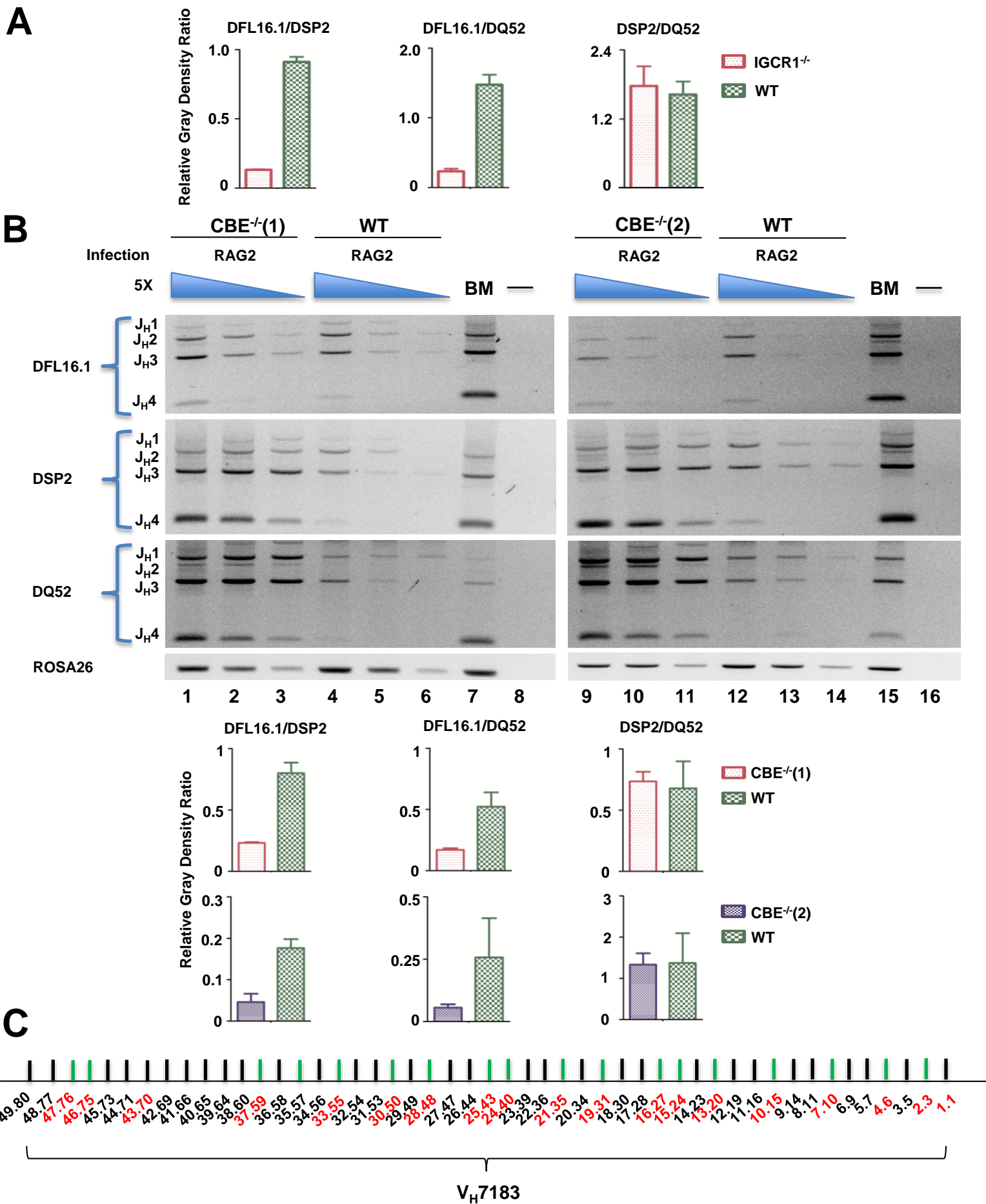


Figure S5. D_H gene segment utilization on WT and IGCR1-deleted or mutated alleles, related to Figure 3.

A. Quantitation of D_H rearrangements in bone marrow pro-B cells carrying WT and IGCR1-deleted *IgH* alleles. D_H-J_H4 rearrangements shown in Figure 3A were scanned and analyzed with gray density analysis software (GeneTools from Syngene). For each D_H gene segment, the total intensity of bands corresponding to rearrangements to J_H1-4 was obtained from lanes 2 and 5 (Figure 3A). Bar graphs show the ratio of DFL16.1 to DSP2, as well as DQ52 rearrangements within each genotype (left and middle); the ratios of DSP2 to DQ52 rearrangements are shown on the right. Data are shown as mean ± SEM of two rearrangement assays from independent pro-B cell sorts.

B. RAG2-deficient pro-B cell lines, CBE^{-/-}(1) and (2), and WT, were infected with *Rag2*-expressing lentivirus or control lentivirus. After 4 or 5 days of selection with puromycin, genomic DNA was prepared to assay rearrangements. D_H-J_H4 rearrangements were amplified and analyzed by agarose gel electrophoresis. 5-fold decreasing amounts of genomic DNA starting at 200ng (lanes 1, 4, 9 and 12) are shown for each genotype. The non-rearranging ROSA26 locus served as a loading control. Upper panel, rearrangements of each D_H gene segment to the four J_H gene segments are shown. Lower panel, gray density (lane 2, 5, 10 and 13) was calculated as described for part A, and ratios of DFL16.1/DSP2, DFL16.1/DQ52 and DSP2/DQ52 are shown. Data are shown as mean ± SEM of two independent infections of each cell line followed by rearrangement assays.

C. Schematic map of V_H7183 gene family. 19 (red) out of total 49 V_H genes are recognized by the V_H7183 primer used in the rearrangement assays. Sequence of V_H7183 primer is listed in Table S3. Green and black lines stand for functional and non-functional genes, respectively (Retter et al., 2007).

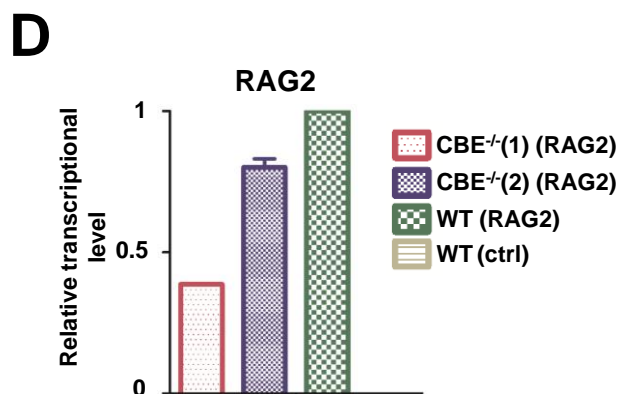
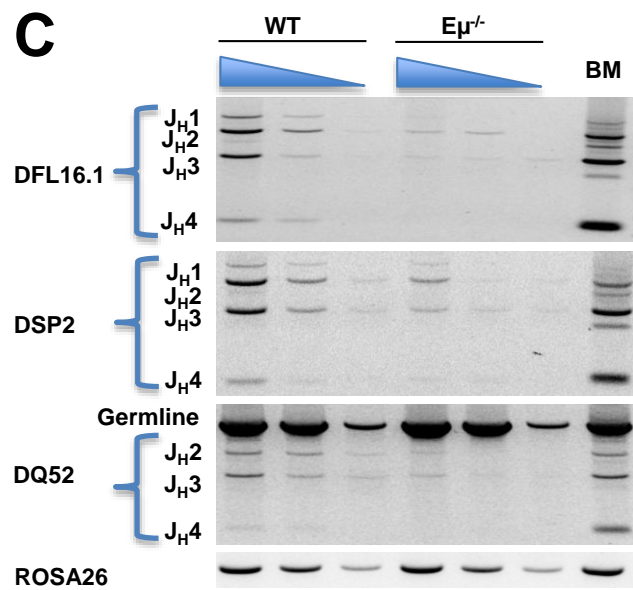
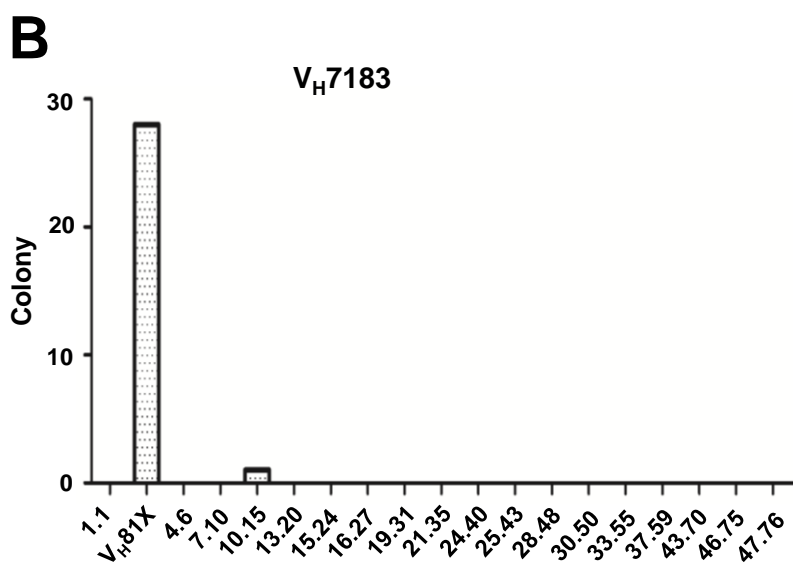
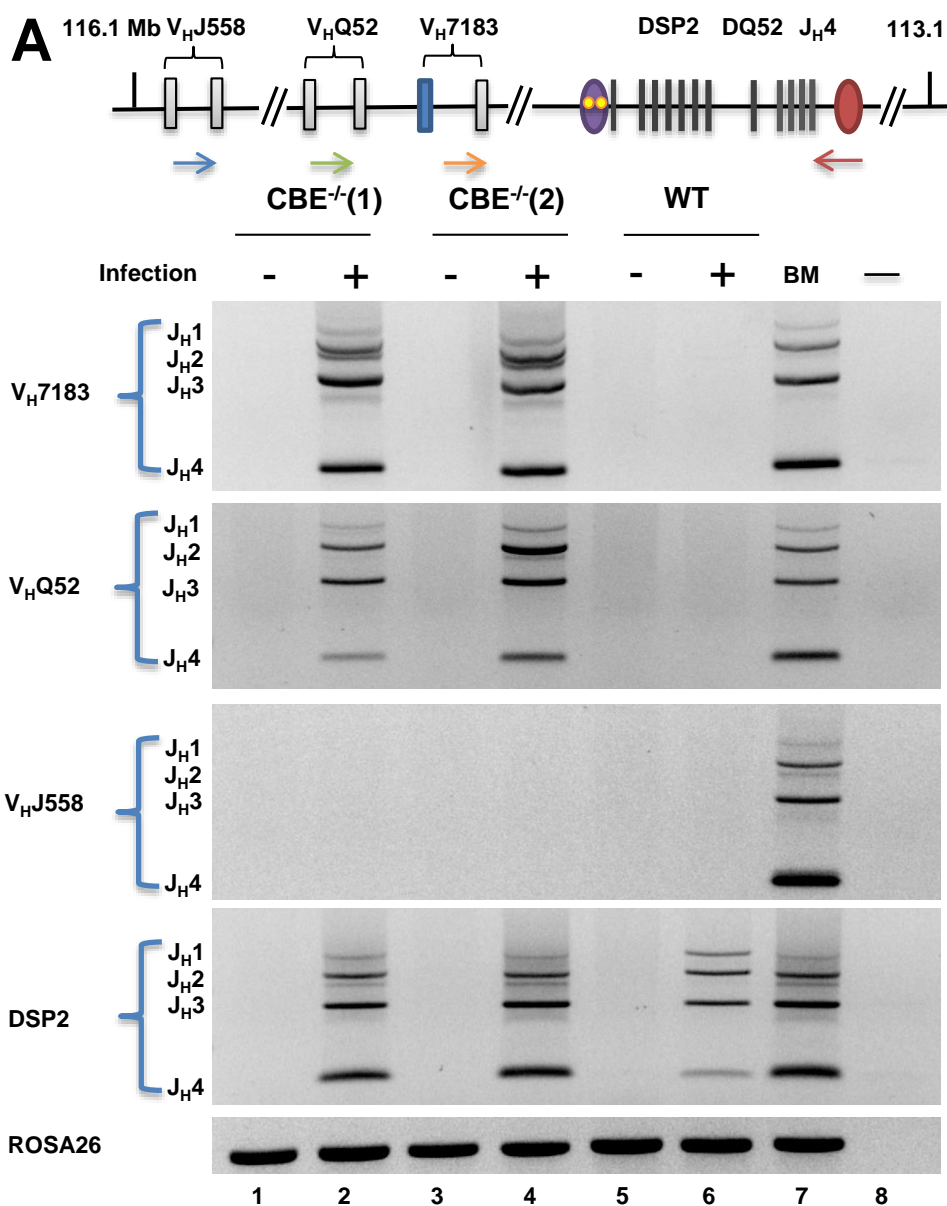


Figure S6. V_H recombination efficiency of WT and IGCR1-mutated *IgH* alleles, related to Figure 4.

A. RAG2-deficient pro-B cell lines carrying IGCR1-mutated alleles (CBE^{-/-}(1) and (2)) or WT *IgH* alleles (WT) were infected with *Rag2*-expressing lentivirus or control lentivirus. After 14 days of selection with puromycin, genomic DNA was prepared for rearrangement. Genomic DNA purified from these cells was used in amplification reactions with a 5' primer that scores for V_HJ558 (blue), V_HQ52 (green), or V_H7183 (orange) together with a 3' primer located beyond J_H4 (red). Amplification products were analyzed by agarose gel electrophoresis. The non-rearranging ROSA26 locus served as a loading control. Lane 7 is positive control using bone marrow genomic DNA; lane 8 contains no genomic DNA. Data shown is representative of two independent infections followed by analyses of recombination.

B. RAG2-deficient pro-B cell line carrying IGCR1-mutated *IgH* alleles (CBE^{-/-}(1)) was infected with *Rag2* expressing lentivirus. After 14 days of selection with puromycin, genomic DNA was prepared for rearrangement. Genomic DNA purified from these cells was used in amplification reactions with a 5' primer that scores for 19 V_H7183 gene segments, or 13 V_HQ52 gene segments together with a 3' primer located beyond J_H1. Amplification products were sub-cloned into pGEM®-T, and 29 (V_H7183) or 30 (V_HQ52) clones were sequenced to estimate V_H and D_H usage. Numbers of clones with V_H7183 (left) and V_HQ52 (right) rearrangements are shown in the graph.

C. RAG2-deficient pro-B cell lines carrying WT *IgH* alleles (WT) or E μ deficient alleles (E μ ^{-/-}) were infected with *Rag2* expressing lentivirus or control lentivirus. After 14 days of selection with puromycin, genomic DNA purified from these cells was used in amplification reactions with a 5' primer that scores for DFL16.1, DSP2 or DQ52 together with a 3' primer located beyond J_H4 (red). Amplification products were analyzed by agarose gel electrophoresis. The non-rearranging ROSA26 locus served as a loading control. Data shown is representative of two independent infections followed by analyses of recombination.

D. Total RNA from cell lines described in part A was reverse transcribed with random hexamer primers, and transcriptional levels of *Rag2* were normalized to γ -actin mRNA and are represented relative to levels in WT cells (Y axis). WT (ctrl) represents the starting RAG2-deficient pro-B cell line with WT *IgH* alleles after transduction with a control virus. Data are shown as mean \pm SEM of two independent experiments.

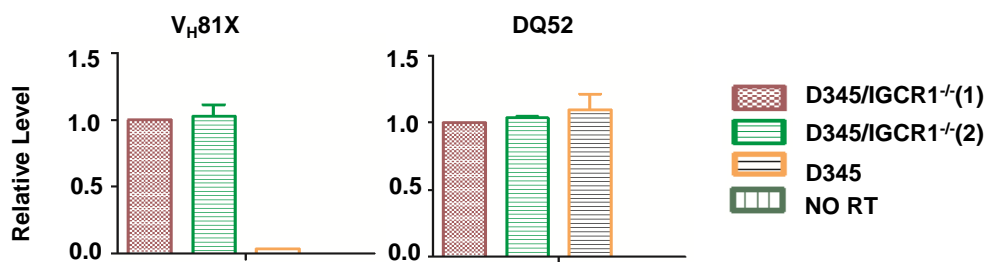
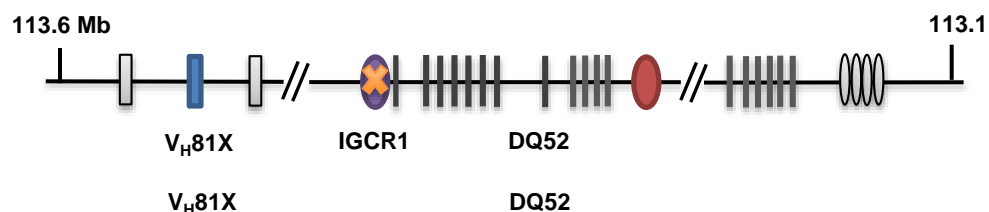
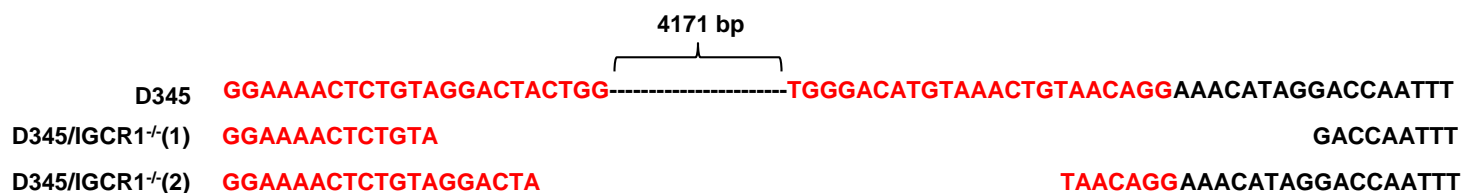
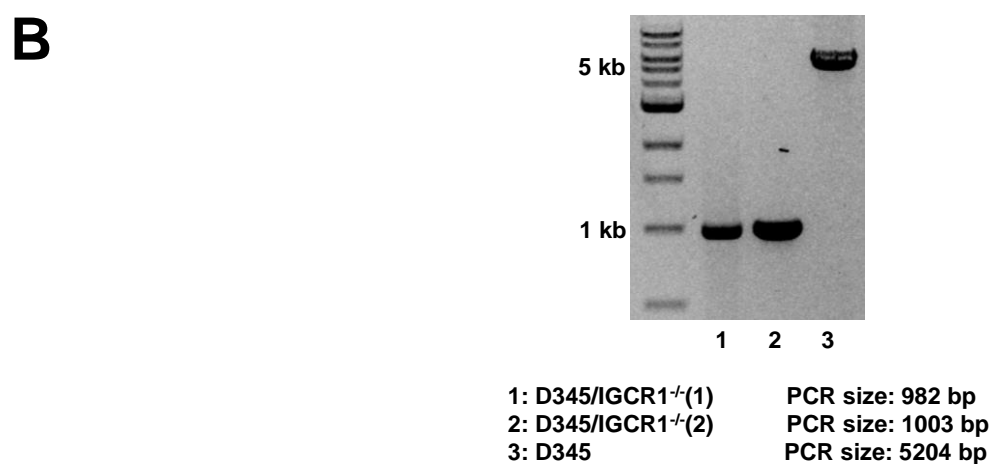
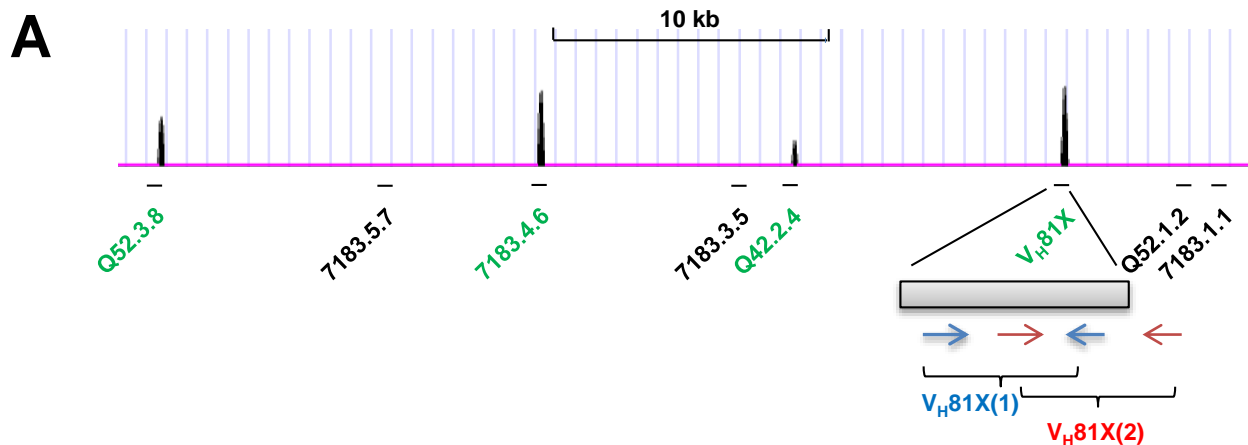


Figure S7. RAG1/2 binding to IGCR1-deleted *IgH* alleles, related to Figure 5.

A. Location of CHIP amplicons and CTCF binding sites within proximal V_H genes. CTCF ChIP-Seq data was derived from previous published data (Lin et al., 2012). Locations of seven of the 3'-most V_H gene segments are shown below the line. V_H Q52.1.2 and V_H 7183.1.1 are pseudogenes; V_H 81X is the 3'-most functional V_H gene segment. Expanded view below V_H 81X shows positioning of primer pairs for V_H 81X(1) and V_H 81X(2) amplicons. The 3' primer of V_H 81X(2) amplicon ensures detection of only unrearranged V_H 81X gene segments.

B. Generation and characterization of IGCR1-deleted derivatives of the pro-B cell line D345 that expresses catalytically inactive RAG1 and endogenous RAG2. gRNAs listed in Table S3 were used to delete IGCR1 using CRISPR/Cas9. PCR was employed to show the deletion of IGCR1. Two subclones, D345/IGCR1^{-/-}(1) and D345/IGCR1^{-/-}(2), were selected for further analyses. PCR assays using primers flanking IGCR1 confirmed homozygous deletion of approximately 4 kb. PCR products were subcloned and sequenced to obtain precise junctional information. Total RNA isolated from D345/IGCR1^{-/-}(1), D345/IGCR1^{-/-}(2) cells was assayed by quantitative RT-PCR to examine V_H 81X and DQ52 transcripts. cDNA synthesis without the addition of reverse transcription polymerase is indicated as No RT. Transcript levels were normalized to γ -actin mRNA and are shown relative to levels in D345/IGCR1^{-/-}(1) cells (Y axis). Data are shown as mean \pm SEM of two independent experiments.

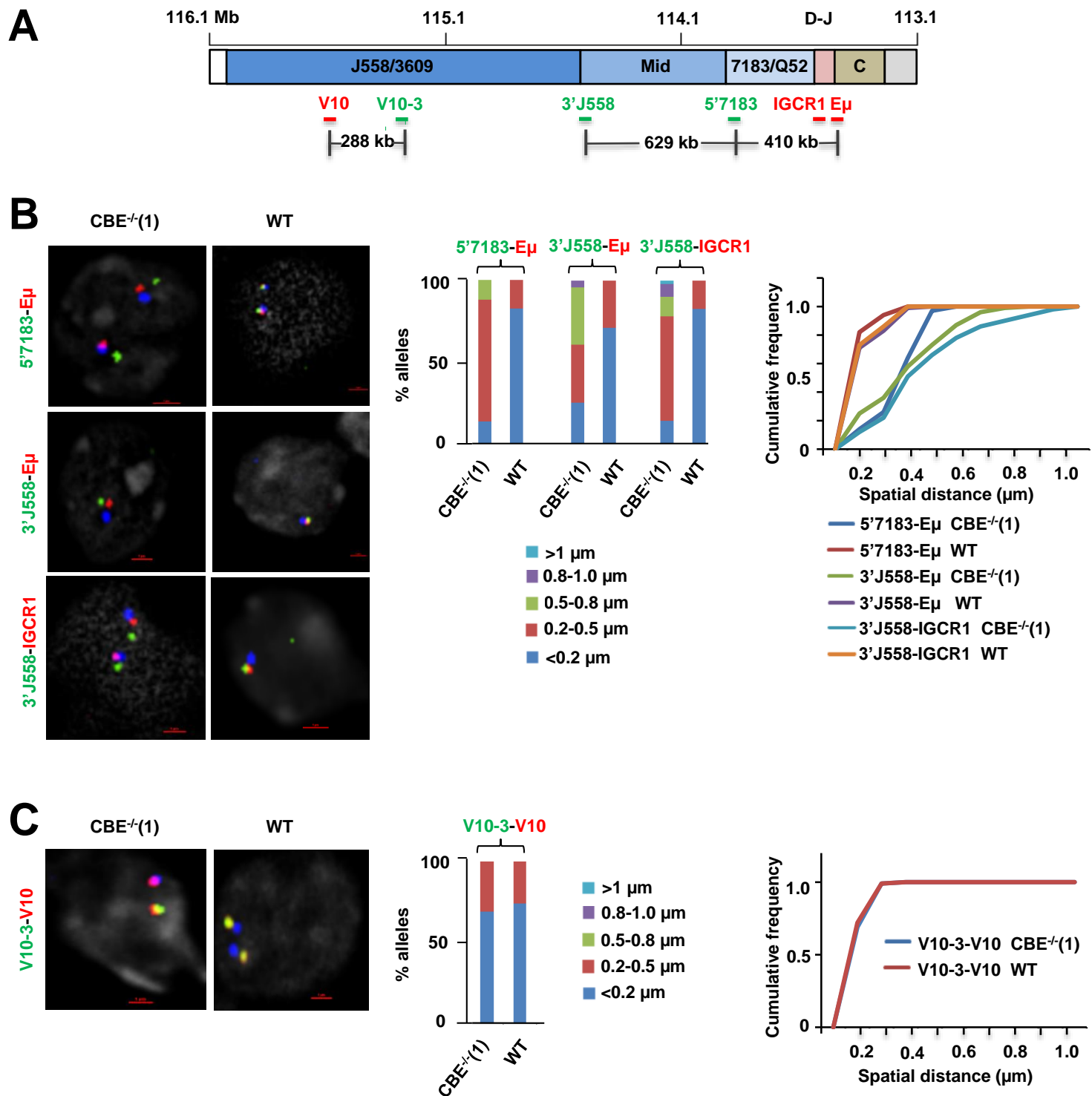


Figure S8. Long-distance interactions on IGCR1-mutated *IgH* alleles-1, related to Figure 6.

A and B. Biological replicates of experiments described in Figure 6.

C. Structure of the V_H locus. A CTCF-dependent interaction in the distal V_H locus was measured using probes V10-3 and V10 in RAG2-deficient pro-B cell lines with WT and IGCR1-mutated ($CBE^{-/-(1)}$) *IgH* alleles (top panel). BAC RP23-201H14 is in blue. Representative nuclei are shown with probe combination indicated on the left. Spatial distance measurements from 100 nuclei of each genotype are summarized in the bar graph and cumulative frequency plots as described in Figure 6B. At least two independent FISH experiments were carried out with each probe combination.

Figure S9

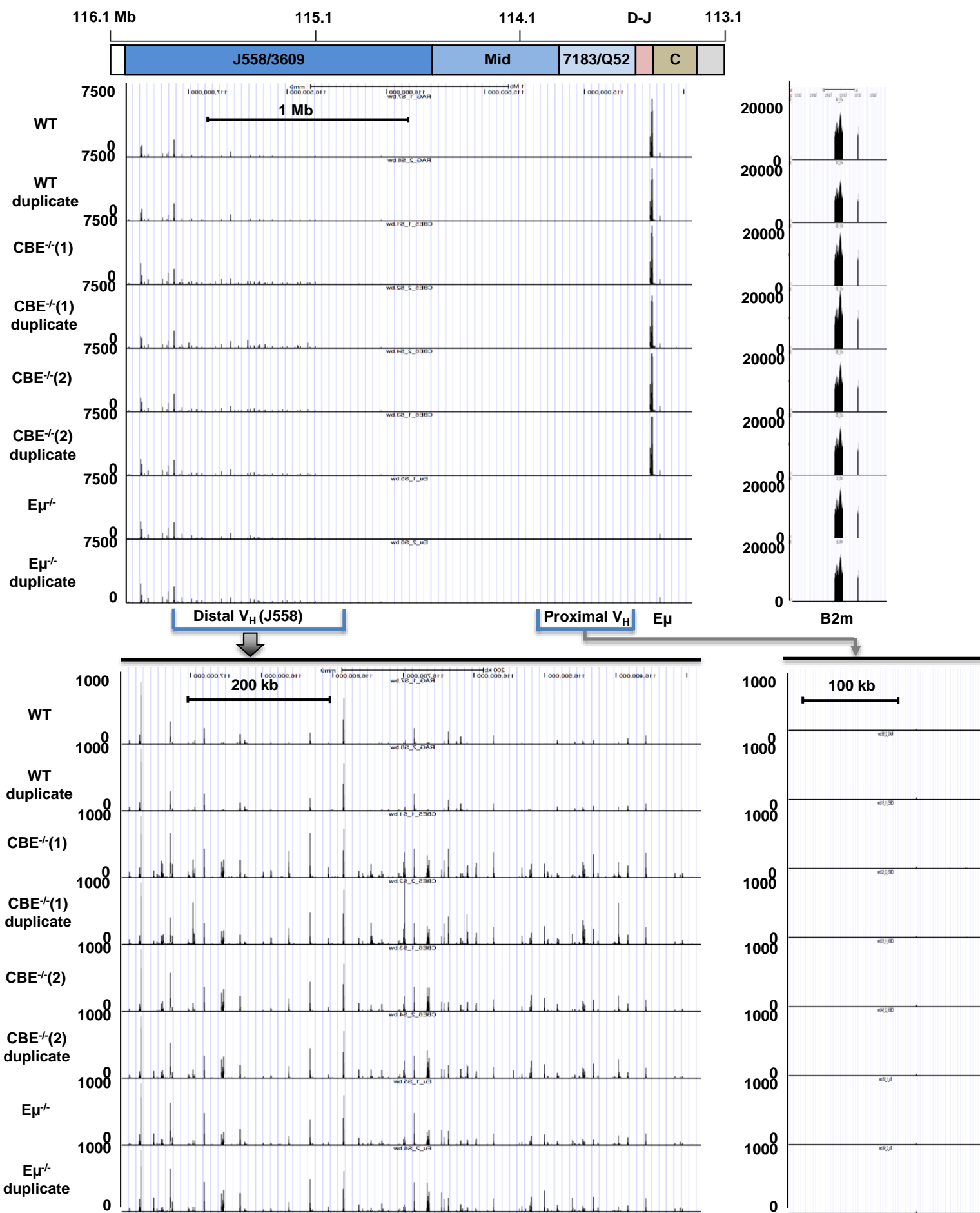


Figure S9. Long-distance interactions on IGCR1-mutated *IgH* alleles-2, related to Figure 6.

RNA-Seq analysis of different cell lines with WT or IGCR1-mutated alleles. CBE^{-/-}(1) and CBE^{-/-}(2) carry mutations in IGCR1 whereas E μ ^{-/-} lacks the intronic enhancer. The entire *IgH* locus (Ch12: 114,452,000 - 117,310,000) in mm9 is shown in the top half. B2m (Chr2: 121,975,206 - 121,978,345) was used as a control. Parts of the locus comprising distal V_H locus (Ch12: 116,276,841-117,096,033) and proximal V_H locus (chr12: 114,800,000 - 115,000,000) are shown in the bottom part of this figure. Duplicate experiments were carried out and results from each replicate are shown.

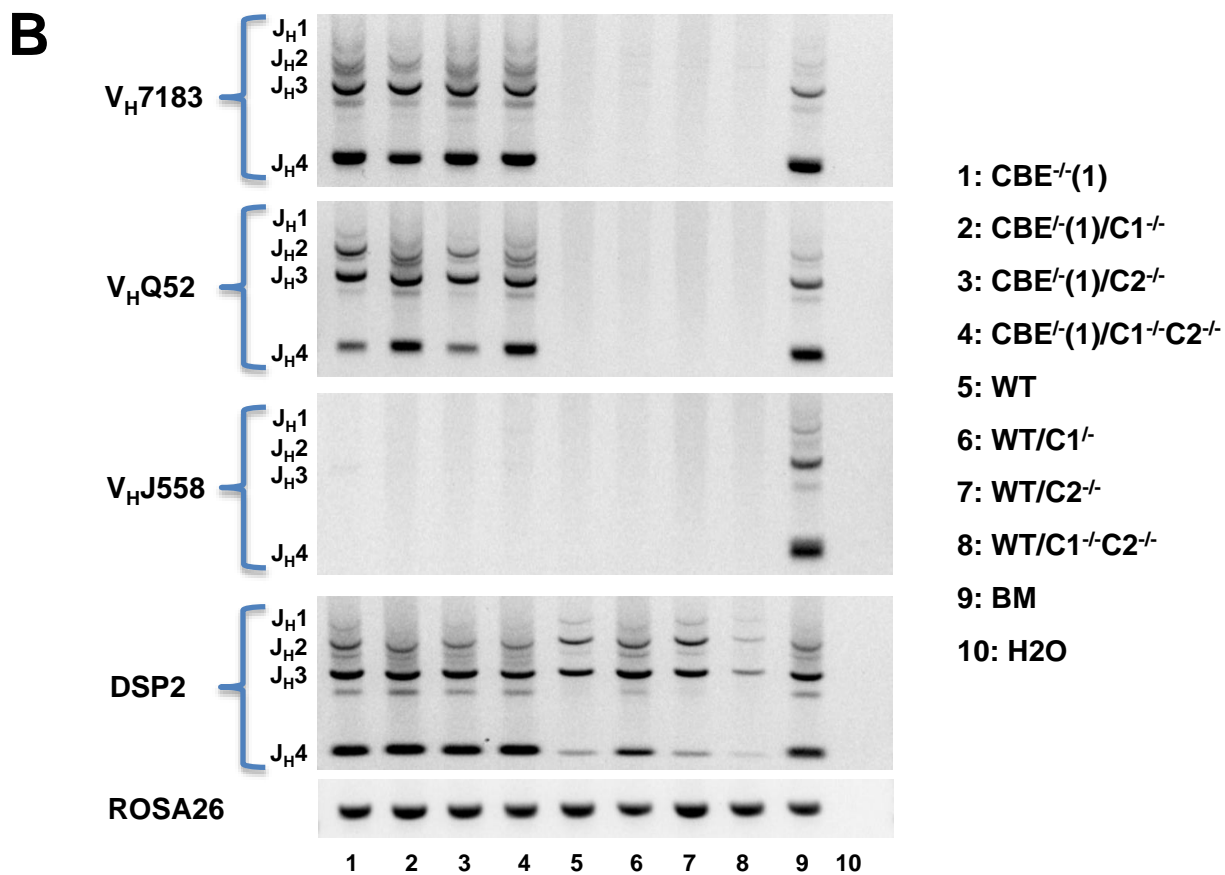
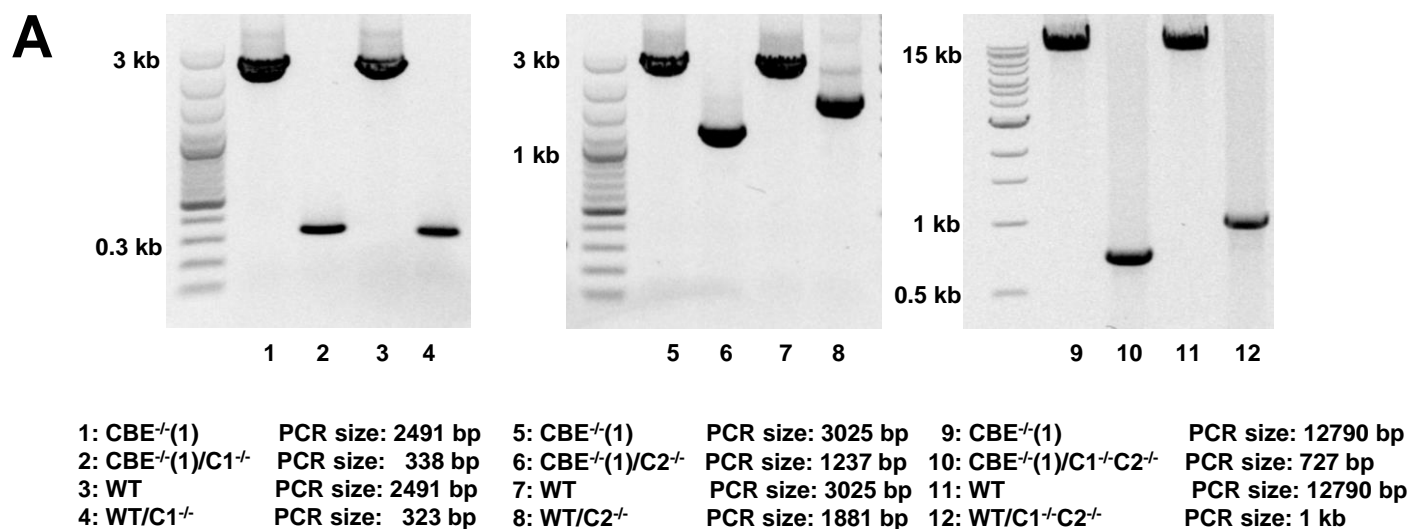


Figure S10. Features of pro-B cells with deletion of proximal V_Hs on WT or IGCR1-mutated *IgH* alleles, related to Figure 6.

A. Generation and characterization of different pro-B cell lines. gRNAs listed in Table S3 were used to delete C1 (left), C2 (middle) or C1andC2 (right) on WT or IGCR1 mutated alleles using CRISPR/Cas9. PCR assays using primers flanking C1, C2 or C1C2 confirmed homozygous deletion. PCR products were subcloned into pGEM®-T and sequenced to obtain precise junctional information. PCR size information of different cell lines are listed below of the gel.

B. Pro-B cell lines with different *IgH* genotypes as labeled were infected with *Rag2* expressing lentivirus. After 10 days of selection with puromycin, genomic DNA purified from these cells was used in amplification reactions with a 5' primer that scores for V_HJ558, V_HQ52, V_H7183 or DSP2 together with a 3' primer located beyond J_H4. Amplification products were analyzed by agarose gel electrophoresis. The non-rearranging ROSA26 locus served as a loading control. Lanes 9 is the positive control using bone marrow genomic DNA; lanes 10 contain no genomic DNA. Data shown is representative of two independent infections followed by analyses of recombination.

Table S1. Average spatial distance and chromatin compaction in the *IgH* locus from different cell lines, related to Figure 1, 2 and 6.

Probes	Genomic Distance (Mb)	Number of Alleles	Spatial Distance± SD in μm	Compaction ^a	Genotype	D-statistic (K-S test)	P-value
IGCR1-E μ	0.06	106	0.319±0.168	64	CBE ^{-/-} (1)		
IGCR1-E μ	0.06	105	0.161±0.067	127	WT	0.49847	8.25E-12
V _H 81X-E μ	0.15	100	0.141±0.042	362	CBE ^{-/-} (1)		
V _H 81X-E μ	0.15	100	0.309±0.206	165	WT	0.53	1.26E-12
V _H 81X-E μ	0.15	100	0.156±0.055	327	Scr		
V _H 81X-E μ	0.15	100	0.350±0.146	146	YY1-shRNA	0.65	< 2.2e-16
V _H 81X-E μ	0.15	100	0.347±0.172	147	CTCF-shRNA#1	0.63	< 2.2e-16
5'7183-E μ	0.41	100	0.467±0.227	299	CBE ^{-/-} (1)		
5'7183-E μ	0.41	100	0.175±0.083	797	WT	0.65	< 2.2e-16
3'J558-E μ	1.04	100	0.390±0.202	907	CBE ^{-/-} (1)		
3'J558-E μ	1.04	100	0.199±0.118	1777	WT	0.51	1.01E-11
3'J558-IGCR1	0.98	100	0.475±0.246	701	CBE ^{-/-} (1)		
3'J558-IGCR1	0.98	100	0.184±0.092	1811	WT	0.66	< 2.2e-16
V10-3-V10	0.50	100	0.140±0.057	1214	CBE ^{-/-} (1)		
V10-3-V10	0.50	100	0.137±0.054	1241	WT	0.09	0.8127

^aCompaction Value: 340 [nm] x genomic length [kb] / spatial distance [nm] x 1 kb

Table S2: V_H7183 usage in IGCR1^{-/-} and WT pro-B cells, related to Figure 3.

V _H 7183 family	IGCR1 ^{-/-} (1)		IGCR1 ^{-/-} (2)		WT(1)		WT(2)	
	Colony	Percentage	Colony	Percentage	Colony	Percentage	Colony	Percentage
V _H 7183.1.1	2	3.6%	4	7.1%				
V _H 7183.2.3	52	94.6%	50	89.3%	19	33.9%	6	10.9%
V _H 7183.4.6			1	1.8%	9	16.1%	7	12.7%
V _H 7183.7.10					4	7.1%	2	3.6%
V _H 7183.10.15							1	1.8%
V _H 7183.13.20					4	7.1%	8	14.5%
V _H 7183.15.24					3	5.4%	6	10.9%
V _H 7183.19.31					3	5.4%	2	3.6%
V _H 7183.21.35			1	1.8%			6	10.9%
V _H 7183.24.40					3	5.4%	1	1.8%
V _H 7183.25.43							2	3.6%
V _H 7183.28.48					2	3.6%		
V _H 7183.30.50					1	1.8%		
V _H 7183.33.55							4	7.3%
V _H 7183.35.57	1	1.8%						
V _H 7183.37.59					1	1.8%	1	1.8%
V _H 7183.46.75					2	3.6%	2	3.6%
V _H 7183.47.76					5	8.9%	7	12.7%
Total	55	100%	56	100%	56	100%	55	100%

Table S3. Primer list, related to Figures 1-6, S1-10.

3C		
Name	Sequence (5'-3')	Reference
7183.4.6	ATTGCACTGAATGTGTAGATTG	
V _H 81X	CTGGCTACTGCTTTTGATTAA	
7183.1.1	GATGTACTTTTGCTAAGTCTGC	
IGCR1	GGATGTGAGTAGCTAGAGGATA	
E μ	GGAACAATTCCACACAAAGACTC	(Guo et al., 2011a)
HS4	GTAGGAAGAGTCATTCTTGACC	
E μ (+20)	ATACATGAACACCTCTCTCTCT	
HS4 (-21)	TCTAAACTCTCAACCCAGAGTA	
HS4 (+19)	CCAAAATTCTCAACCAGGAAAC	
α -amylase-F	GCTTCCATGATACTCTATGTTCTTCCT	(Guo et al., 2011a)
α -amylase-R	GAGATCTTACGTAGGCACTTAGTGGTATAA	(Guo et al., 2011a)
E μ probe	AGCTTTAAGAGCAGCAGCCACAGCT	(Guo et al., 2011a)
V _H 81X probe	TCAGAAGTAAAGCCAAGAGCTTACATC	
α -amylase probe	TTGAATATGTACCGAGTACACATGGATGGTGCAT	(Guo et al., 2011a)
FISH		
V _H 81X-F	GTCTCTGTGTTGTTCTATGTCT	
V _H 81X-R	TAGTAGATGAACTACCATCAG	
IGCR1-F	CATCAGAAACAACAGCATCTCC	
IGCR1-R	GGAACATCGTGTCTGTGTTTGT	
5'7183-F	TTGGCTCACTCTGAGTTGGGATTCCCTC	(Guo et al., 2011a)
5'7183-R	TAAAGCTGAACAAGGACCACAAGACGA	(Guo et al., 2011a)
3'558-F	AAGTCCCTGGGAGCTCTGGGGCAGTCA	(Guo et al., 2011a)
3'558-R	TTGTTTCTAGGAAAAGATAGGCACACAGAT	(Guo et al., 2011a)
V10-3-F	GCACATCTTCATTGTTCTTCTGAAATC	(Guo et al., 2011a)
V10-3-R	CTGACCCAGCCTACTGAAGA GTCAAAC	(Guo et al., 2011a)
V10-F	CACCTCCAATAGCACTCACAGGTTGGC	(Guo et al., 2011a)
V10-R	CAAAGGCTGCCTGCACTAAGACTGGTG	(Guo et al., 2011a)
E μ -F	AGCTCATGGTACTTTGAGGAAATC	
E μ -R	TTGTAGGAGGACTTCCCTAATCTG	
ChIP/RNA/real-time PCR		
YY1-F	AAGATATTGACCATGAAACA	
YY1-R	CATGAGGGCAAGCTATTGTTC	
CTCF-F	GCGCCCGTTCCAGTGCAGTTTGTG	(Verma-Gaur et al., 2012)
CTCF-R	GCATAAGGCTTCTCCCCCGTGTGA	(Verma-Gaur et al., 2012)
V _H 81X-F	ATTAATGGCTGCGACCAACT	(Guo et al., 2011b)
V _H 81X-R	GCTTAGTGCAGCCTGGAGAG	(Guo et al., 2011b)
γ -actin-F (Figure 2)	GGTGTCCGGAGGCACTCTT	(Chakraborty et al., 2007)
γ -actin-R (Figure 2)	TGAAAGTGGTCTCATGGATACCA	(Chakraborty et al., 2007)
DQ52-F (Figure 2)	TGGTGCAAGGTTTTGACTAAGC	(Chakraborty et al., 2007)
DQ52-R (Figure 2)	CCAAACAGAGGGTTTTTGTGAG	(Chakraborty et al., 2007)
5'E μ -F	TTAACCGAGGAATGGGAGTG	(Degner et al., 2011)
5'E μ -R	GGTGGGGCTGGACAGAGTGT	
3'E μ -F	AATACCCGAAGCATTTACAGTGACT	

3'εμ-R	AAGATTTGTGAAGCCGTTTTGACCA	
3'J558-F	GCCAGGCTTTCTACACCTTTTCC	(Guo et al., 2011a)
3'J558-R	CCTTGCCCTTGAACCTTCTGATTG	(Guo et al., 2011a)
Q52.3.8-F	GTCTCAGGGTTCTCATTAACCAG	
Q52.3.8-R	GCTGAATGATAATTTGTGCTCCC	
7183.5.7-F	TAGTGACAGTGTCTGACTCAGG	
7183.5.7-R	GAGAGTATCACGGACCTTTCAG	
7183.4.6-F	AGGCCGGTCCTGGATTGAGTT	
7183.4.6-R	TCACTAACAAAACAACAACAA	
7183.3.5-F	GGCCTGACTGTGCAGGTAGTCAC	
7183.3.5-R	GAAGGATCTCATGTCTATGAATT	
V _H 81X(2)-F	TCCAGAGACAATACCAAGAAG	
V _H 81X(2)-R	TTGAGCTCACAGTAACTTTTG	
V _H 81X(1)-F	ATTAATGGCTGCGACCAACT	(Guo et al., 2011b)
V _H 81X(1)-R	GCTTAGTGCAGCCTGGAGAG	(Guo et al., 2011b)
Q52.1.2-F	CAGAGCATGACTGTGCTAGAG	
Q52.1.2-R	GACGTGTCTTAGTCCCTCTACT	
71823.1.1-F	TGCCCAAGTCTAACTCCAAAG	
71823.1.1-R	GGGCTCCTTTGTCACAACT	
IG1-F	ATCTTGTGAAAATGGATCAGCC	
IG1-R	ACAAACTTTCATCTACTGTTGA	
IG2-F	GCGAAGTAGCATGGTCTTAGTT	
IG2-R	TGCTCAGTCAACAGCCATCAG	
DFL16.1-F (Figure 5)	ACACCTGCAAAACCAGAGACCATA	(Subrahmanyam et al., 2012)
DFL16.1-R (Figure 5)	GATCTGACTTTAGGGTGGTTGGA	(Subrahmanyam et al., 2012)
DSP2-F (Figure 5)	TGTTACCTTACTTGGCAGGGATT	(Subrahmanyam et al., 2012)
DSP2-R (Figure 5)	TGGGTTTTTGTGCTGGATATATC	(Subrahmanyam et al., 2012)
J _H 2-F	TACTTTGACTACTGGGGC	(Chakraborty et al., 2009)
J _H 2-R	CCCTAGTCCTTCATGACC	(Chakraborty et al., 2009)
V _H 81X(-4)-F	GGAACGAGAACAGAAACCCTAA	
V _H 81X(-4)-R	CTTCTTCTCCTGCCCTTACTC	
V _H 81X(+4)-F	GTCCTCCATTTCTCTCAGTGTG	
V _H 81X(+4)-R	GAATAGTTGCAAGTGGGCTTTG	
DFL16.1(-4)-F	CCTCCTTGACAGCCTTGGGG	
DFL16.1(-4)-R	GTCACTCCCTCCTCCCTGTTG	
DFL16.1(-1.7)-F	GGCGAAGAAGCTTGTATCACAT	(Chakraborty et al., 2007)
DFL16.1(-1.7)-R	GTGGGTGGGATCAGTGCCCATCTTA	(Chakraborty et al., 2007)
DFL16.1-F (Figure S1D)	CAAAGCAGCCACCATCCAG	(Chakraborty et al., 2007)
DFL16.1-R (Figure S1D)	GCAGCACGGTTGAGTTTCAG	(Chakraborty et al., 2007)
DSP2-F (Figure S1D)	CAACAAAACCCAGTATGCCCAG	(Chakraborty et al., 2007)
DSP2-R (Figures 5, S1D)	GTGCTTTCACCTGT TGTGGG	(Chakraborty et al., 2007)
DQ52-F (Figures 5, S1D)	CCCTGTGGTCTCTGACTGGTG	(Chakraborty et al., 2007)
DQ52-R (Figure S1D)	GATTTCTCAAGCCTCTCTACTTCCTC	(Chakraborty et al., 2007)
εμ-F	GGAATGGGAGTGAGGCTCTCTC	(Chakraborty et al., 2007)
εμ-R	CTGCAGGTGTTCTGGTTCTGATCGG	(Chakraborty et al., 2007)
Cγ3-F	TGGACAAACAGAAGTAGACATGGGTC	(Subrahmanyam et al., 2012)

C γ 3-R	GGGGTTTAGAGGAGAGAAGGCAC	(Subrahmanyam et al., 2012)
HS6-F	CCGCCCTTCACACCCTGACAAAC	(Subrahmanyam et al., 2012)
HS6-R	CTGGCACTGAGCAAGCAAACCTCT	(Subrahmanyam et al., 2012)
γ -actin-F (Figures 5, S1D)	GACACCCAACCCCGTGACG	(Guo et al., 2011a)
γ -actin-R (Figures 5, S1D)	GCGGCCATCACATCCCAG	(Guo et al., 2011a)
DFL16.1(-4.5)-F	AGGCATCTCATCTCACTCTAAGC	(Subrahmanyam et al., 2012)
DFL16.1(-4.5)-R	TGTGTCCCTCTAAGACGAGTGAAT	(Subrahmanyam et al., 2012)
DFL16.1-F (Figure S1E)	AACTCAACCGTGCTGCCTG	(Chakraborty et al., 2007)
DFL16.1-R (Figure S1E)	CAATGGGTTTTTGCTGATGG	(Chakraborty et al., 2007)
C μ -F	AGAGATCTGCATGTGCCATT	(Bolland et al., 2007)
C μ -R	TGGTGGGACGAACACATTTACA	(Bolland et al., 2007)
RAG2-F	TGTGCTACATCATATATTCTCC	
RAG2-R	TACTGGCAAGTGAGTGTCCCTCC	
PAIR-F	GAGCTCTCCTTCAGTTCTTGAAT	(Ebert et al., 2011)
PAIR-R	TCTGCACTGTGTGACAATGG	(Ebert et al., 2011)
DJ/VDJ recombination assay		
DFL16.1-F	ACA CCT GCA AAA CCA GAG ACC ATA	
DSP-F	ATG GCC CCT GAC ACT CTG CAC TGC TA	
DQ52-F	GCGACTGTTTTGAGAGAAATCATTGG	(Subrahmanyam et al., 2012)
J _H 4-R	GGGTCTAGACTCTCAGCCGGCTCCCTCAGGG	(Subrahmanyam et al., 2012)
ROSA26-F	AAAGTCGCTCTGAGTTGTTAT	
ROSA26-R	GGAGCGGGAGAAATGGATATG	
V _H 7183	GTGGAGTCTGGGGGAGGCTTA	(Guo et al., 2011b)
V _H Q52	CTCACAGAGCCTGTCCATCAC	(Guo et al., 2011b)
V _H J558	ARGCCTGGGRCTTCAGTGAAG	(Guo et al., 2011b)
J _H 1-R	TGAGGAGACGGTGACCGTGGTCCC	
V _H 81X	GGAGTTGGTTCGAGCCATTAATAGT	(Guo et al., 2011b)
DQ52-R	CATCCACCCTTCTGATGCTTGCATT	(Guo et al., 2011b)
PCR		
E μ -F (Figure S2C)	GGACCACCTCTGTGACAGCATT	
E μ -R (Figure S2C)	AGAGGGAGTTCACACAGAGCAT	
IGCR-F (Figure S5B)	CTATCTGGGTACCTGTGAACTCT	
IGCR-F (Figure S5B)	CGTATGAGCCATCTCAGCTCAG	
C1-F (Figure S6E)	CTGAAACTCTCCTGTGAATCCAA	
C1-F (Figure S6E)	CTGGTGCTCAAGCAGATTCA	
C2-F (Figure S6E)	CACTATCTACAGAGATGATGCC	
C2-R (Figure S6E)	ATTCTGAGCTGCCTTTTCTGAA	
gRNA target sequence		
E μ -1	TTACTCTTAAGTTGGTTCCT	
E μ -2	ATACCCGAAGCATTACAGT	
IGCR1-1	GGAAAACCTCTGTAGGACTAC	(Hu et al., 2015)
IGCR1-2	TGGGACATGTAAACTGTAAC	(Hu et al., 2015)
C1-1	TGGAGAGACGATTCATCATC	
C1-2	TGATGATTGACTACCAAACC	
C2-1	GAGCTGGGCTTCCTTGCAAT	
C2-2	CGTTTACCTTACAAAACCC	

Table S4. Barcode sequence for different cell lines, related to Figure 6.

CBE ^{-/-} (1)	CGAACTGT
	CTCTGTCT
	GATGGAAT
CBE ^{-/-} (2)	ATATAGGA
	CACGTGTT
CBE ^{-/-} (1)/C1 ^{-/-}	ATCGCCAG
	CATTCCAA
CBE ^{-/-} (1)/C2 ^{-/-}	CAGGAGGC
	CGACTGGG
CBE ^{-/-} (1)/C1 ^{-/-} C2 ^{-/-}	GGCAGACG
	CACAGTTG

V _H S	CBE ^{-/-} (1)			CBE ^{-/-} (2)		CBE ^{-/-} (1)/C1 ^{-/-}		CBE ^{-/-} (1)/C2 ^{-/-}		CBE ^{-/-} (1)/C1 ^{-/-} C2 ^{-/-}	
V _H 36-60.a1.85	0	0	0	0	0	0	0	0	0	0	0
V _H relic.a5psi.86	0	0	0	0	0	0	0	0	0	0	0
V _H 11.a1.87	0	0	0	1	0	0	0	0	0	0	0
V _H SM7.a2psi.88	0	0	0	0	0	0	0	0	0	0	0
V _H X24.a2.89	0	0	0	0	0	0	0	0	5	0	0
V _H 36-60.a2.90	1	2	0	0	0	0	0	0	0	3	1
V _H relic.a6psi.91	0	0	0	0	0	0	0	0	0	0	0
V _H 11.a2.92	0	0	0	0	0	0	0	1	1	0	0
V _H SM7.a3.93	0	0	0	0	1	0	0	0	0	0	0
V _H relic.a7psi.94	0	0	0	0	0	0	0	0	0	0	0
V _H 3609N.a1psi.95	0	0	0	0	0	0	0	0	0	0	0
V _H GAM3-8.a1.96	0	0	0	0	0	0	0	0	0	0	0
V _H 12.a1psi.97	0	0	0	0	0	0	0	0	0	0	0
V _H GAM3-8.a2.98	0	0	0	0	0	0	0	0	0	0	0
V _H 12.a2psi.99	0	0	0	0	0	0	0	0	0	0	0
V _H GAM3-8.a3.100	0	0	0	0	0	0	0	0	0	0	0
V _H 12.a3.101	0	0	0	0	0	0	0	0	9	0	0
V _H GAM3-8.a4.102	0	0	0	0	0	0	0	0	0	0	0
V _H 12.a4psi.103	0	0	0	0	0	0	0	0	0	0	0
V _H 12.a5psi.104	0	0	0	0	0	0	0	0	0	0	0
V _H GAM3-8.a5.105	0	0	0	3	0	0	0	0	1	0	0
V _H S107.a3.106	0	0	0	14	11	0	0	0	0	2	0
V _H 15.a1psi.107	0	0	0	0	0	0	0	0	0	0	0
V _H SM7.a4.108	0	0	0	1	1	0	0	0	1	0	0
V _H 36-60.a3psi.109	0	1	0	0	0	0	0	0	0	0	0
V _H S107.a4.110	0	0	0	0	0	0	0	0	0	0	0
V _H 36-60.a4.111	0	0	0	0	0	0	0	1	0	0	0
V _H 36-60.a5.112	0	0	0	0	0	0	0	0	0	0	0
V _H 3609N.a2psi.113	0	0	0	0	0	0	0	0	0	0	0
V _H 36-60.a6.114	0	0	0	0	0	1	0	0	0	6	0
V _H GAM3-8.a6.115	0	1	0	0	0	0	0	0	0	0	3
V _H 36-60.a7psi.116	0	0	0	0	0	0	0	0	0	0	0
V _H 36-60.a8.117	0	0	0	0	0	0	0	0	0	0	0
V _H J606.a1psi.118	0	0	0	0	0	0	0	0	0	0	0
V _H J558.a1psi.119	0	0	0	0	0	0	0	0	0	0	0
V _H 3609N.a3.120#	0	0	0	3	10	0	0	0	1	0	0
V _H 36-60.a9.121#	0	0	0	0	0	0	0	0	0	0	0
V _H J606.a2psi.122#	0	0	0	0	0	0	0	0	0	0	0
V _H 36-60.a10.123#	0	0	0	0	0	0	0	0	1	0	0
V _H J606.a3.124#	0	0	0	0	0	0	0	0	0	0	0
V _H J606.a4.125#	0	0	0	0	0	0	0	0	0	0	0
V _H J606.a5psi.126#	0	0	0	0	0	0	0	0	0	0	0
V _H J606.a6.127#	0	0	0	0	0	0	0	0	0	0	0
V _H J606.a7.128#	0	0	0	1	0	0	0	0	0	1	0