Supplementary Data

Nuclear positioning rather than contraction controls ordered rearrangements of immunoglobulin loci

Magdalena B. Rother, Robert-Jan Palstra, Suchit Jhunjhunwala, Kevin A.M. van Kester, Wilfred F.J. van IJcken, Rudi W. Hendriks, Jacques J.M. van Dongen, Cornelis Murre and Menno C. van Zelm

Table S1. Primer sequences used for 3C-Seq

Locus	Viewpoint name	Closest gene or region	Forward primer sequence (5'- 3')	Reverse primer sequence (5'- 3')
Igh	CT7-526A21 distal V_H	IghV1-43	TGGATACCCATTAGTTAAGTGCAG	TCTGGACCTGAGCTGGTGAA
Igh	RP23-24I12 proximal V_H	IghV11-2	AGGAAAAAGAGCTGAGAAGA	GAGGTGTGTTGGTAGTAAGG
Igh	iEμ	iEμ	AGTGTTGTTCTGGTAGTTCC	TACGGAGGCAAAAACAAAGA
Igh	3'RR	3'RR	ACTATTTCCACCAGTGTCTACATTT	TGCCCCCAGGAGCAGAT
Igκ	RP23-234A12 distal Vκ	IgκV1-132	TGTTATCATTACCAGAAGATCT	TGTGTACAGTTTATTGGGAGTGT
Igκ	RP24-475M8 proximal Vκ	IgκV12-47	TACCCTGATGGTCTTTTCACTTC	TGGCCATTAAGAAACCTGCTG
Igκ	iEκ *	iEκ	GAAAGTATGACTGCTTGCCATG	CTGCAGTCAGACCCAGATCT
Igκ	3'Eκ *	3'Eκ	CCATACCAGACTGGTTATTGA	GCCTCTTAGTGACCAGATCT

These primers were ordered as oligonucleotides containing 5' adapters necessary for sequencing on the Illumina platform: A1- AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCT (BglII) and A2- CAAGCAGAAGACGGCATACGA (NlaIII). *, described previously: (1)

Table S2. Expression levels of *Igh* and *Igκ* germline transcripts in Rag1$^{-/-}$ pro-B and Rag1$^{-/-}$ VH81X pre-B

Transcript Cluster ID	Transcript Cluster Location	Gene	Expression level	
			pro-B	pre-B
Igh				
10402991	chr12:114891504-114891935	IghV2-3 - IghV5-6	169.47	131.26
10403015	chr12:115920840-115921276	IghV1-14 - IghV1-15	111.08	67.52
10403018	chr12:116026580-116027016	IghV1-22 - IghV1-24	191.53	78.98
10403034	chr12:116532274-116532573	IghV1-56	180.37	109.98
10403043	chr12:116699210-116699642	IghV1-62-2	348.06	111.20
10403048	chr12:116745673-116996669	IghV1-64 - IghV1-71	381.69	150.94
10403063	chr12:116886163-116886598	IghV1-67 - IghV1-69	135.92	72.08
Igκ				
10538871	chr6:67560201-67560536	IgkV1-135	55.31	541.74
10538880	chr6:67966766-68071824	IgkV9-120 - IgkV1-122	8.94	70.78
10538903	chr6:68206398-70676960	Igkk14-111; IgkV3-1 - IgkV6-17	78.39	156.43
10545173	chr6:68581954-68582421	IgkV10-96	19.31	48.00
10545175	chr6:68654498-68654788	IgkV10-95	76.75	191.45
10545198	chr6:69388210-69388497	IgkV4-59	88.24	67.72
10545215	chr6:69714513-69714734	IgkV5-48	18.93	64.32
10545247	chr6:70356459-70356749	IgkV8-16	41.89	97.88
10545252	chr6:70648805-70672948	IgkJ1 - IgkV3-5	147.38	150.72

Table S3. Expression levels of B-cell transcription factors and architectural proteins in Rag1$^{-/-}$ pro-B and Rag1$^{-/-}$ VH81X pre-B

	Transcript Cluster ID	Gene	Expression level	
			pro-B	pre-B
Transcription factors / B-cell commitment	10370837	E2A (TCF3)	1951.47	3255.22
	10375358	EBF1	1257.92	1426.53
	10512669	Pax5	2162.76	4154.53
	10374333	Ikaros (IKZF1)	978.41	2387.14
	10562812	PU.1 (SPI1)	1110.90	3526.14
Pre-BCR + signaling	10567863	CD19	2180.29	3441.24
	10438064	Vpre-B	3401.96	594.74
	10438060	λ5 (IGLL1)	7083.47	1967.32
	10427628	IL-7R	1946.21	3724.03
B-cell specific phenotypic markers	10358224	B220 (PTPRC)	1242.62	1603.59
	10568174	CD43 (SPN)	543.40	458.12
	10500677	CD2	85.73	2129.42
Genomic architectural proteins	10574812	CTCF	1626.17	2136.29
	10398288	YY1	989.08	1123.23
DNA remodeling	10544501	Ezh2	3489.61	3195.72

Table S4. Primer sequences used for RQ-PCR

Target	Forward primer	Primer sequence 5'-3'	Reverse primer	Primer sequence 5'-3'	Probe no
GAPDH *	GAPDH_F	AGCTTGTCATCAACGGGAAG	GAPDH_R	TTTGATGTTAGTGGGGTCTCG	9
iEμ	iEμ_F	AAGACATTTGGGAAGGACTGAC	iEμ_R	CTTCTGCCACCCATCCAC	20
iEκ 1.1	iEκ1.1_F	CTTCAAGGCAGCTCTCAGACAA	iEκ1.1_R	GACTGAGGCACCTCCAGATGTT	-
iEκ 0.8	iEκ0.8_F	CACTGTGGTGGACGTTCGG	iEκ0.8_R	GACTGAGGCACCTCCAGATGTT	-
Probe common iEκ1.1/ iEκ0.8	iEκ_probe	CTGATGCTGCACCAACTGTATCC ATCTTCC	-	-	-

*, described previously: (2)

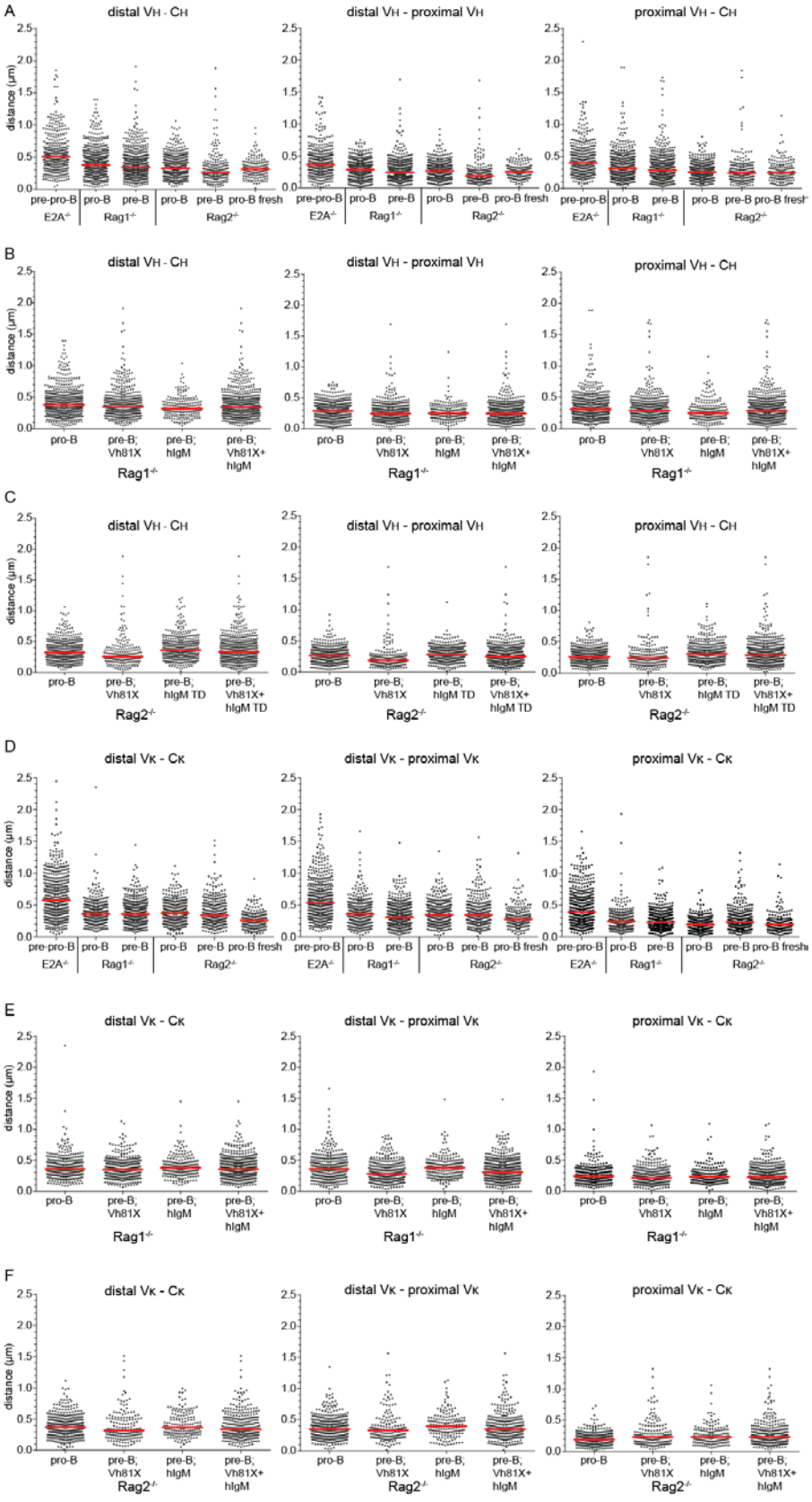
A
distal VH - CH
distal VH - proximal VH
proximal VH - CH
distance (µm)
pre-pro-B
pro-B
pre-B
pro-B fresh
E2A-/-
Rag1-/-
Rag2-/-
B
pro-B
pre-B; Vh81X
pre-B; hIgM
pre-B; Vh81X+ hIgM
Rag1-/-
C
pre-B; hIgM TD
pre-B; Vh81X+ hIgM TD
Rag2-/-
D
distal Vκ - Cκ
distal Vκ - proximal Vκ
proximal Vκ - Cκ
E
F

Figure S1. *Igh* and *Igκ* loci are contracted in Rag1$^{-/-}$ and in Rag2$^{-/-}$ committed B-cell progenitors. Scatter plots show the distances in micrometers (y-axis) separating distal VH, proximal VH, and CH (**A-C**) as well as distal Vκ, proximal Vκ, and Cκ regions (**D-F**) in cultured E2A$^{-/-}$ pre-pro-B, Rag1$^{-/-}$ and Rag2$^{-/-}$ pro-B, Rag1$^{-/-}$ VH81X and Rag1$^{-/-}$ hIgM pre-B as well as Rag2$^{-/-}$ VH81X and Rag2$^{-/-}$ hIgM TD (transduced) pre-B, and in freshly isolated Rag2$^{-/-}$ pro-B cells. For each condition 2-3 mice were used and >100 alleles were analyzed per population. Red horizontal lines represent median distances.

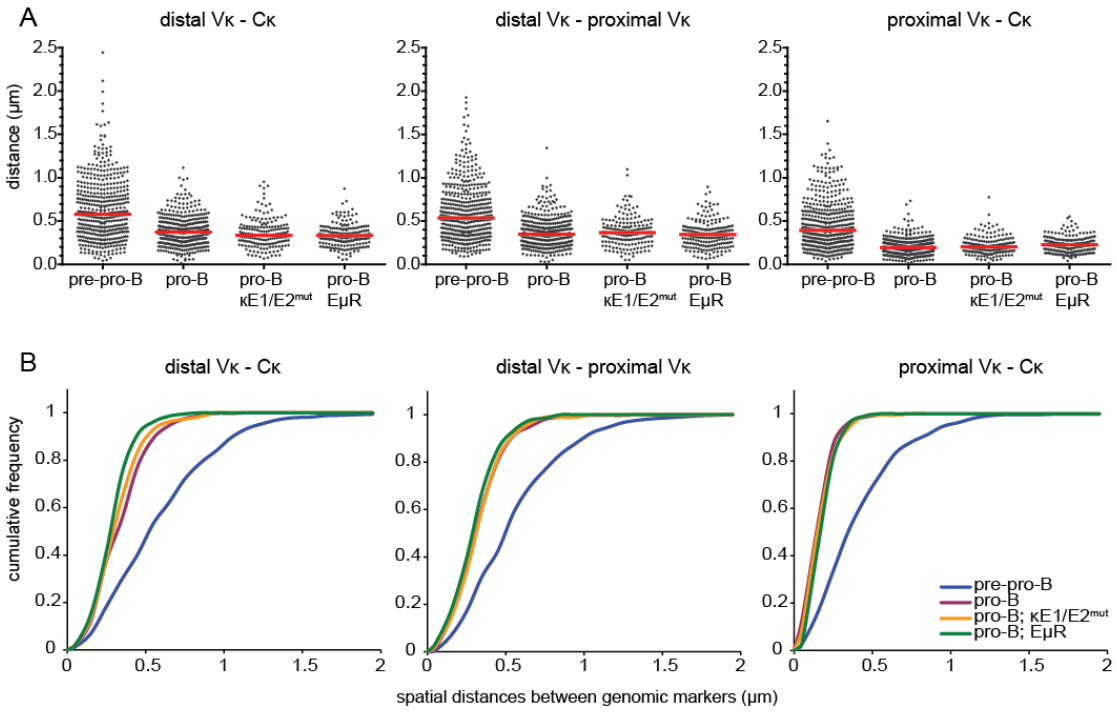

Figure S2. iEκ does not mediate the *Igκ* locus contraction in pro-B cells. (**A**) Scatter plots show the distances in micrometers (y-axis) separating distal Vκ, proximal Vκ, and Cκ regions in cultured E2A$^{-/-}$ pre-pro-B, Rag2$^{-/-}$ pro-B, Rag2$^{-/-}$ κE1\E2^{mut} pro-B and Rag2$^{-/-}$ EμR pro-B cells. For each condition 2-3 mice were used and >100 alleles were analyzed per population. Red horizontal lines represent median distances. (**B**) Cumulative frequency plots show the distribution of spatial distances between the distal Vκ, proximal Vκ, and Cκ region probes in 4 precursor-B cell subsets: E2A$^{-/-}$ pre-pro-B (blue line), Rag2$^{-/-}$ pro-B (purple line), Rag2$^{-/-}$ κE1\E2^{mut} pro-B (orange line) and Rag2$^{-/-}$ EμR pro-B (green line).

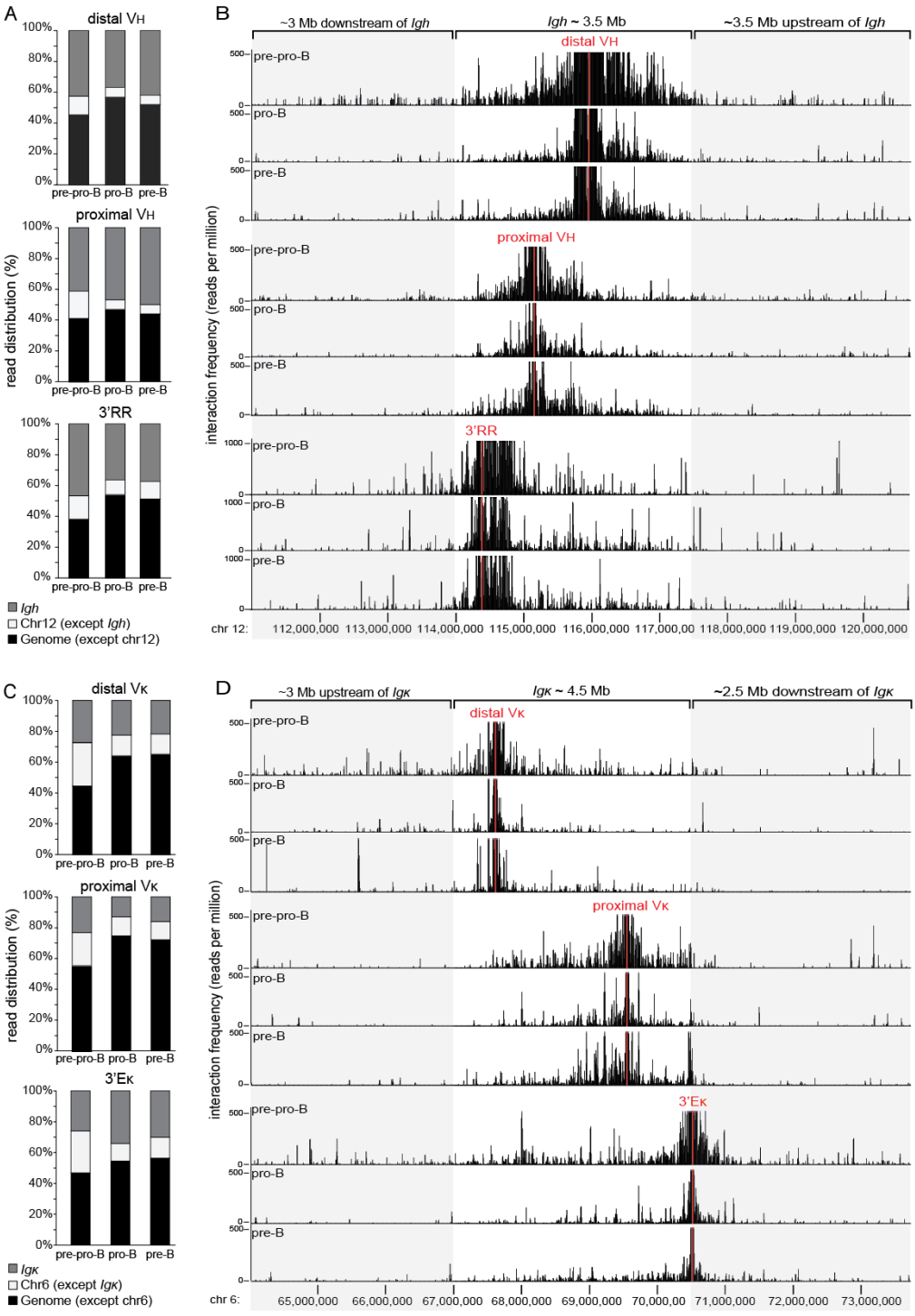

Figure S3.Long-range interactions within the Ig loci. (**A)** Bar graphs showing the relative distribution of 3C-Seq reads in the *Igh* locus, the rest of chromosome 12 (chr12 except *Igh*) and rest of the genome (genome except chr12). Read distributions across the genome were analyzed

for 3 *Igh* viewpoints located in the distal VH region, the proximal VH and 3'RR enhancer viewpoint in cultured E2A$^{-/-}$ pre-pro-B, Rag1$^{-/-}$ pro-B and Rag1$^{-/-}$ VH81X pre-B cells. Indicated cell fractions were obtained from 2 mice. Data represent 1 biological replicate (**B)** 3C-Seq long-range interactions of the 3 *Igh* viewpoints along ~10 Mb range of chromosome 12 were plotted as reads per million for E2A$^{-/-}$ pre-pro-B, Rag1$^{-/-}$ pro-B and Rag1$^{-/-}$ VH81X pre-B cells (**C)** Bar graphs showing the relative distribution of 3C-Seq reads in the *Igκ* locus, the rest of chromosome 6 (chr6 except *Igκ*) and rest of the genome (genome except chr6). Read distributions across the genome were analyzed for 3 *Igκ* viewpoints located in the distal Vκ region, the proximal Vκ and 3'Eκ enhancer viewpoint in cultured E2A$^{-/-}$ pre-pro-B, Rag1$^{-/-}$ pro-B and Rag1$^{-/-}$ VH81X pre-B cells. Indicated cell fractions were obtained from 2 mice. Data represent 1 biological replicate (**D)** 3C-Seq long-range interactions of the 3 *Igκ* viewpoints along ~10 Mb range of chromosome 6 were plotted as reads per million for E2A$^{-/-}$ pre-pro-B, Rag1$^{-/-}$ pro-B and Rag1$^{-/-}$ VH81X pre-B cells.

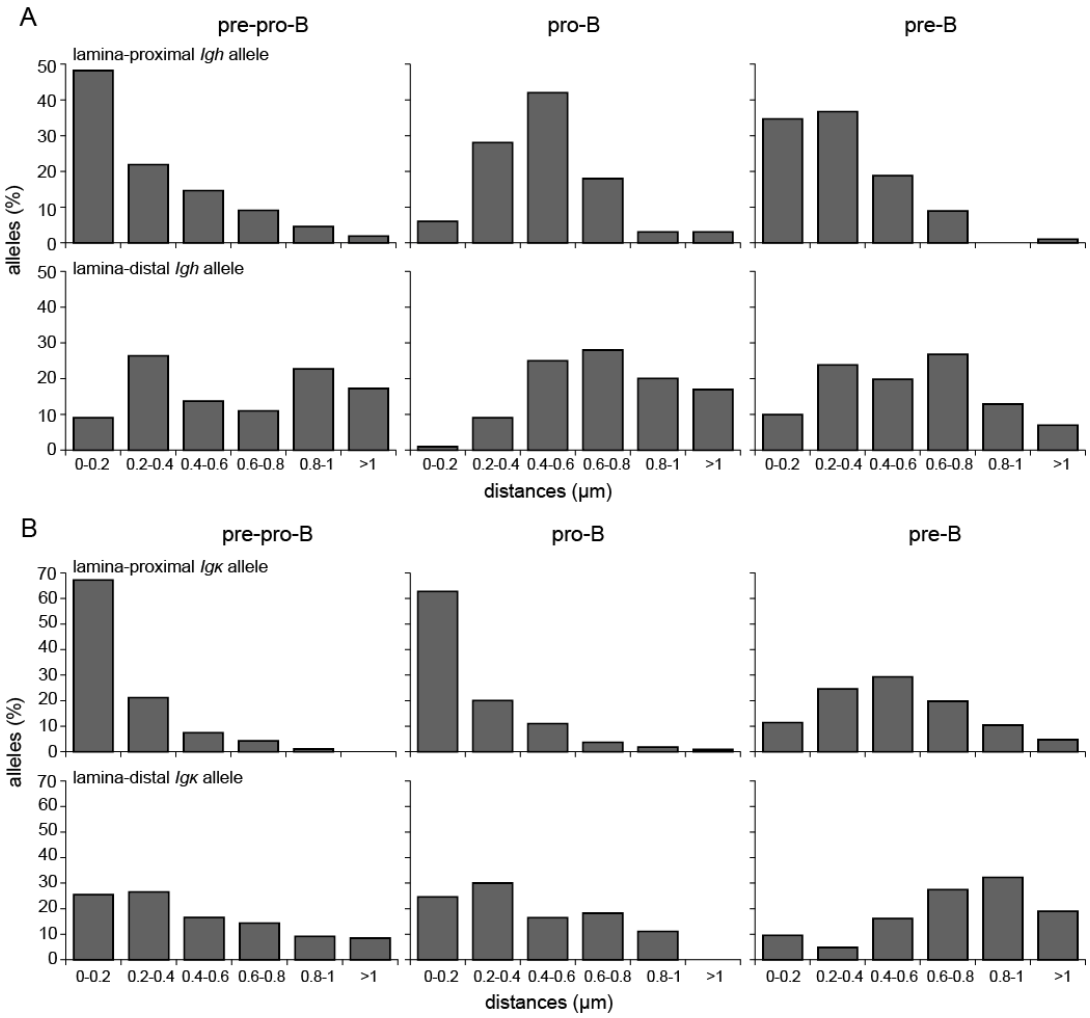

Figure S4. Both *Igh* and *Igκ* alleles re-localize in the nucleus during B-cell development. Bar graphs showing the frequencies of lamina-proximal and lamina-distal *Igh* (**A**) and *Igκ* (**B**) alleles positioned within a distance range from the nuclear lamina in cultured E2A$^{-/-}$ pre-pro-B, Rag1$^{-/-}$ pro-B and Rag1$^{-/-}$ hIgM pre-B cells. The lamina-proximal and lamina-distal *Igh* and *Igκ* alleles were determined based on the average distance between lamina and 3D FISH probes, i.e. distal VH, proximal VH, CH for *Igh*, and distal Vκ, proximal Vκ, Cκ for *Igκ*.

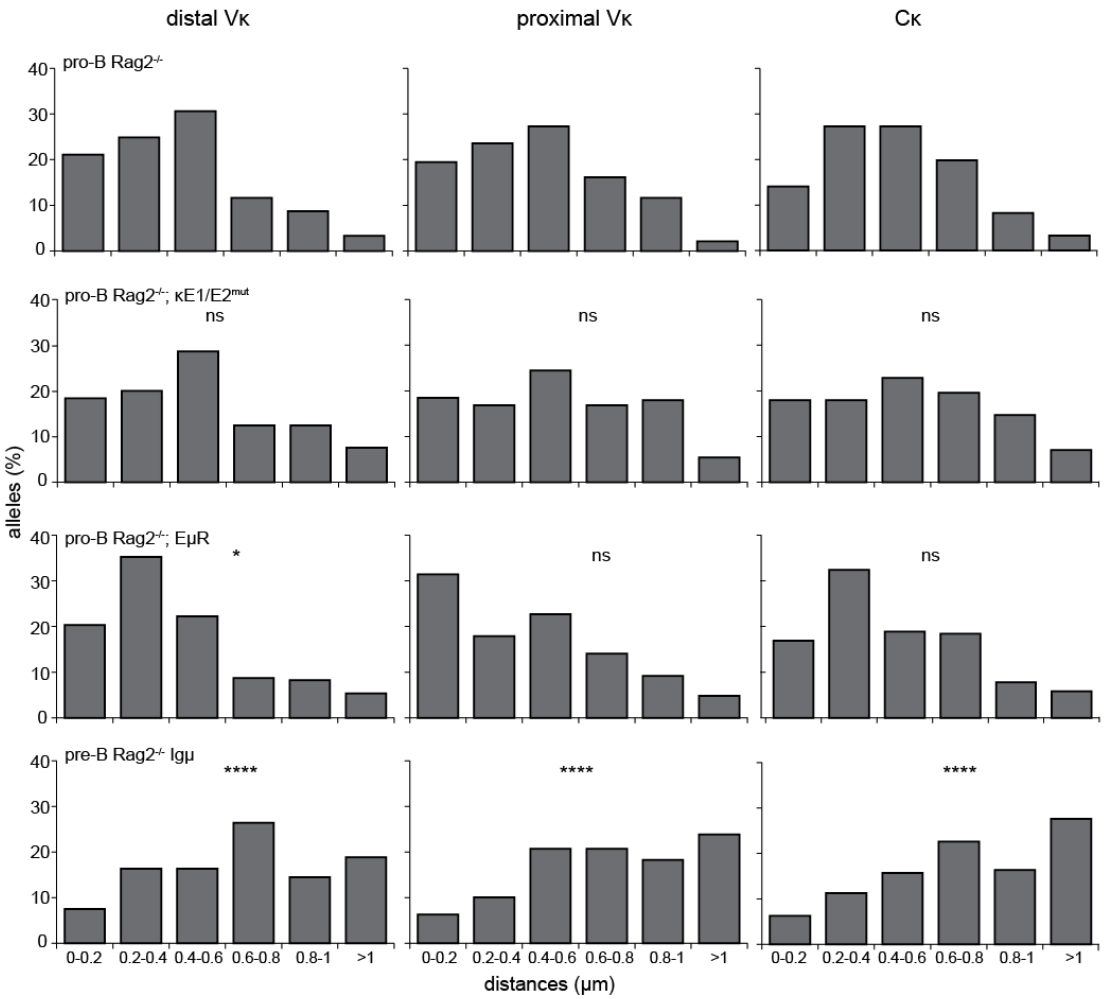

Figure S5. iEκ does not mediate the nuclear positioning of *Igκ* in pro-B cells. Bar graphs show percentages (y-axis) of *Igκ* alleles positioned within certain ranges of distances (x-axis) determined between *Igκ* 3D FISH BAC probes and lamina depicted for cultured $Rag2^{-/-}$ pro-B, $Rag2^{-/-}$ κE1\E2^{mut} pro-B, $Rag2^{-/-}$ EμR pro-B and $Rag2^{-/-}$ Igμ pre-B cells. Statistical significance was calculated with the χ^2 test between $Rag2^{-/-}$ pro-B and each of the remaining subset based on the sum of alleles located <0.4 μm from the lamia and the sum of alleles located >0.4 μm of the lamina. ns, not significant; *, P<.05; ****, P< .0001.

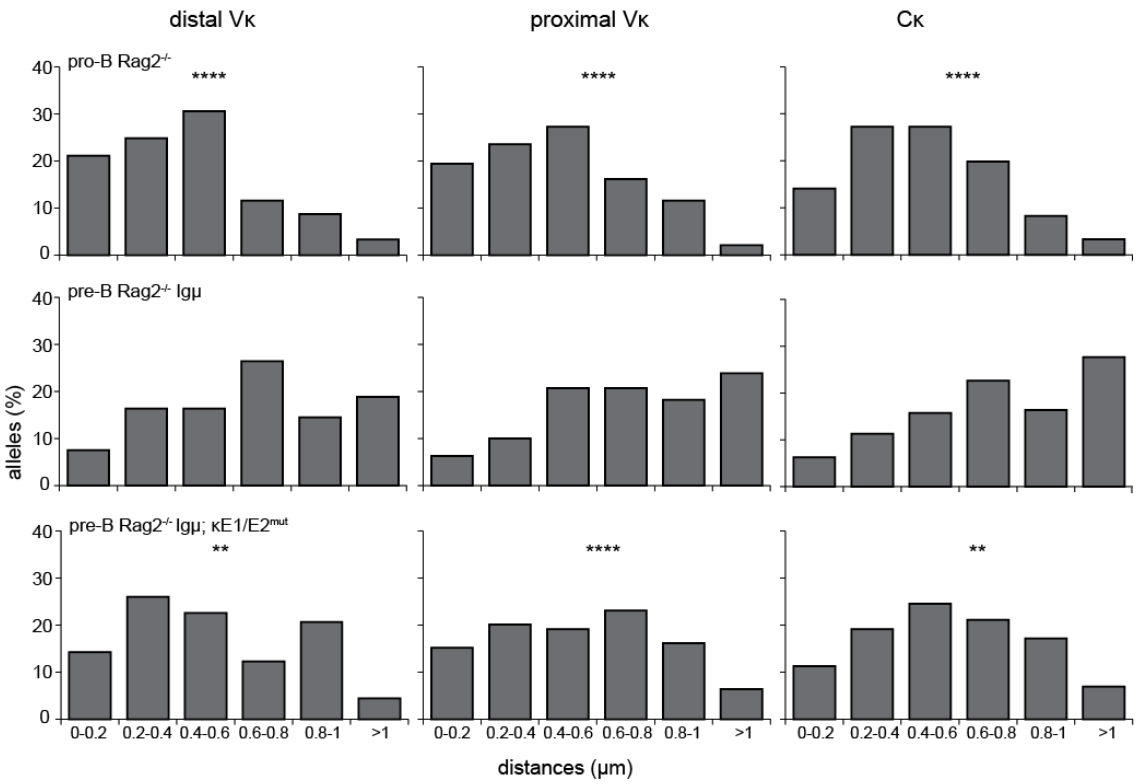

Figure S6. iEκ is likely not involved in the nuclear positioning of *Igκ* in pre-B cells. Bar graphs show percentages (y-axis) of *Igκ* alleles positioned within certain ranges of distances (x-axis) determined between *Igκ* 3D FISH BAC probes and lamina depicted for cultured Rag2$^{-/-}$ pro-B, Rag2$^{-/-}$ Igμ pre-B and Rag2$^{-/-}$ Igμ κE1\E2^{mut} pre-B cells. Statistical significance was calculated with the χ^2 test between Rag2$^{-/-}$ Igμ pre-B and the remaining subsets based on the sum of alleles located <0.4 μm from the lamina and the sum of alleles located >0.4 μm of the lamina. **, P<.01; ****, P< .0001.

Supplementary Materials and Methods

Probe preparation

BAC clones CT7-526A21, RP23-24I12, CT7-34H6 for detecting regions in the murine *Igh* locus (3), and RP23-234A12, RP24-475M8, RP23-435I4 recognizing regions within murine *Igκ* locus (all from BACPAC Resources) were grown in E. coli. Plasmid DNA was purified with the NucleoBond BAC 100 Kit (Macherey-Nagel GmbH&Co.KG). Probes were either directly labeled with Chromatide Alexa Fluor 488-5 dUTP or Chromatide Alexa Fluor 568-5 dUTP (Invitrogen) using Nick Translation Mix (Roche Diagnostics GmbH); or they were indirectly labeled using DIG-Nick Translation Mix (Roche Diagnostics GmbH). Just prior to use, the probes were precipitated and a hybridization cocktail was prepared containing 600 ng of each labeled probe, 4 μg of mouse Cot-1 DNA (Invitrogen), 5 μg of salmon sperm DNA dissolved in 2×SSC + 50% formamide + 10% dextran sulfate. The probes were denatured at 75°C for 5 min prior to hybridization.

3D DNA fluorescence in situ hybridization (FISH)

3D DNA FISH was performed as described previously (3-5). In short, 100 μl of a 1×10^6 cells/ml suspension of cultured cells was directly attached to poly-L-lysine-coated coverslips. The cells were fixed in 4% paraformaldehyde, and permeabilized in a PBS, 0.1% Triton X-100, 0.1% saponin solution, incubated in PBS with 20% glycerol, and subjected to liquid nitrogen immersion. The nuclear membrane was stained using goat anti-Lamin A and B specific antibodies (Santa Cruz) diluted to 2 μg/ml in 1×PBS + 0.1% Triton X-100 + 5% bovine serum albumin for 30 min at 37°C. The nuclear membranes were permeabilized in PBS, 0.5% Triton X-

100, 0.5% saponin prior to hybridization with the DNA probe cocktail for 5 min in a HYBrite™ machine at 75ºC. The coverslips were sealed with nail polish and incubated for 48hr at 37°C. Subsequently, the coverslips were washed and incubated with Cy5-conjugated mouse anti-DIG antibodies (Jackson Immunoresearch) to detect dig-labeled probes. Finally, the coverslips were washed and mounted on slides with 10 µl of Prolong gold anti-fade reagent (Invitrogen) and sealed with nail polish.

Image acquisition and distance calculations

Pictures were captured with a Leica SP5 confocal microscope (Leica Microsystems). Using a 63× lens (NA 1.4), we acquired images of ~70 serial optical sections spaced by 0.15 µm. The data sets were deconvolved and analyzed with Huygens Professional software (Scientific Volume Imaging). The 3D coordinates of the center of mass of each probe were input into Microsoft Excel, and the distances separating each probe were calculated using the equation: $\sqrt{(Xa-Xb)^2+(Ya-Yb)^2+(Za-Zb)^2}$, where X, Y, Z are the coordinates of object a or b. Similarly, the distances between probes and lamins were calculated using 3D coordinates of the center of mass of each probe and 3D coordinates of the point in the lamins determined as the closest to the probe. At least 100 alleles were analyzed per subset.

Circular chromosome conformation capture with high-throughput sequencing (3C-Seq)

The 3C library was prepared as described previously (1,6,7). Briefly, a single cell suspension containing between 1-10 million cells was cross-linked for 10 minutes using 2% formaldehyde. Nuclei were lysed and permeabilized with SDS, which was subsequently neutralized with Triton-X100. Chromatin was digested overnight using 400U BglII restriction enzyme (New England

Biolabs). Digested chromatin was diluted 14-fold in ligation buffer followed by ligation using 100U T4 ligase (Roche Diagnostics). After ligation, cross-links were reversed and the DNA was phenol-chloroform extracted. The DNA was subsequently digested with a second restriction enzyme, NlaIII (New England Biolabs). After phenol-chloroform extraction, the DNA was diluted and re-ligated. Finally, the DNA was extracted with phenol-chloroform and purified on columns using Gel extraction kit (Qiagen). DNA fragments interacting with four *Igh* and four *Igκ* viewpoints were PCR amplified using viewpoint-specific reverse primers (**Table S1**). Primers for the iEκ and 3'Eκ viewpoint-specific PCR were described previously (1) and the rest of the primers were newly designed. The obtained 4C library was single-read sequenced on an Illumina Hi-Seq 2000 platform and 100 bp reads were generated.

3C-Seq data analysis

The downstream processing of generated sequences was performed as described previously (6,7). The obtained reads were mapped against the mm9 (NCBI build 37.1) assembly of the mouse genome. The interaction frequencies between the viewpoints and genomic regions were measured as sequence read counts within individual BglII restriction fragment. To normalize the differences in sequencing depth between samples, the read counts were calculated per million total mapped reads and referred as reads per million (RPM). Data were then visualized using a local UCSC genome browser. For quantitative analysis, V regions of the *Igh* and *Igκ* loci were described as regions of chr12: 114,810,329- 117,244,460 (869 fragments) and chr6: 67,493,424-70,659,328 (1,221 fragments), respectively. BglII restriction fragments <100 bp, as well as the viewpoint and both adjacent fragments were excluded from the quantitative analysis. The latter fragments are very abundant, which is characteristic for 3C-based experiments, and can

introduce large variation in downstream analyses. VH and Vκ gene coordinates were extracted from the IMGT and Ensembl mouse genome databases, and assigned to the BglII restriction fragments.

Quantitative RT-PCR Analysis

Total RNA was extracted from E2A$^{-/-}$ pre-pro-B using the miRNeasy Mini Kit (Qiagen) and reverse-transcribed with SuperScript II reverse transcriptase (Invitrogen). For cDNA amplification, 15 µl reaction mix was used with Taq-Man Universal MasterMix (Applied Biosystems), 670 nM of each primer and 100 nM of FAM-TAMRA-labeled probe (8). Primers were designed with the ProbeFinder software (Roche Applied Science) and probes were from the Universal ProbeLibrary (Roche Applied Science) or designed manually (iEκ) and purchased from Eurogentec (**Table S4**). Duplicate reactions were performed for each cDNA sample. Gene expression was analyzed with a StepOnePlus Real-Time PCR system and StepOne Software version 2.3 (Applied Biosystems). Cycle-threshold levels were calculated and normalized to values obtained for the endogenous reference gene *GAPDH*.

Supplementary References

1. Ribeiro de Almeida, C., Stadhouders, R., de Bruijn, M.J., Bergen, I.M., Thongjuea, S., Lenhard, B., van Ijcken, W., Grosveld, F., Galjart, N., Soler, E. *et al.* (2011) The DNA-binding protein CTCF limits proximal Vkappa recombination and restricts kappa enhancer interactions to the immunoglobulin kappa light chain locus. *Immunity*, **35**, 501-513.
2. Cornelissen, F., Asmawidjaja, P.S., Mus, A.M.C., Corneth, O., Kikly, K. and Lubberts, E. (2013) IL-23 Dependent and Independent Stages of Experimental Arthritis: No Clinical Effect of Therapeutic IL-23p19 Inhibition in Collagen-induced Arthritis. *PLoS One*, **8**.
3. Sayegh, C.E., Jhunjhunwala, S., Riblet, R. and Murre, C. (2005) Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. *Genes Dev*, **19**, 322-327.
4. Nodland, S.E., Berkowska, M.A., Bajer, A.A., Shah, N., de Ridder, D., van Dongen, J.J., LeBien, T.W. and van Zelm, M.C. (2011) IL-7R expression and IL-7 signaling confer a distinct phenotype on developing human B-lineage cells. *Blood*, **118**, 2116-2127.
5. Jensen, K., Rother, M.B., Brusletto, B.S., Olstad, O.K., Aass, H.C.D., van Zelm, M.C., Kierulf, P. and Gautvik, K.M. (2013) Increased ID2 Levels in Adult Precursor B Cells as Compared with Children Is Associated with Impaired Ig Locus Contraction and Decreased Bone Marrow Output. *J Immunol*, **191**, 1210-1219.
6. Stadhouders, R., Kolovos, P., Brouwer, R., Zuin, J., van den Heuvel, A., Kockx, C., Palstra, R.J., Wendt, K.S., Grosveld, F., van Ijcken, W. *et al.* (2013) Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions. *Nat Protoc*, **8**, 509-524.
7. Stadhouders, R., de Bruijn, M.J., Rother, M.B., Yuvaraj, S., Ribeiro de Almeida, C., Kolovos, P., Van Zelm, M.C., van Ijcken, W., Grosveld, F., Soler, E. *et al.* (2014) Pre-B Cell Receptor Signaling Induces Immunoglobulin kappa Locus Accessibility by Functional Redistribution of Enhancer-Mediated Chromatin Interactions. *PLoS Biol*, **12**, e1001791.
8. van Zelm, M.C., Bartol, S.J., Driessen, G.J., Mascart, F., Reisli, I., Franco, J.L., Wolska-Kusnierz, B., Kanegane, H., Boon, L., van Dongen, J.J. *et al.* (2014) Human CD19 and CD40L deficiencies impair antibody selection and differentially affect somatic hypermutation. *J Allergy Clin Immunol*, **134**, 135-144.