Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/33/eaaz8850/DC1

Supplementary Materials for

Altered 3D chromatin structure permits inversional recombination at the *IgH* locus

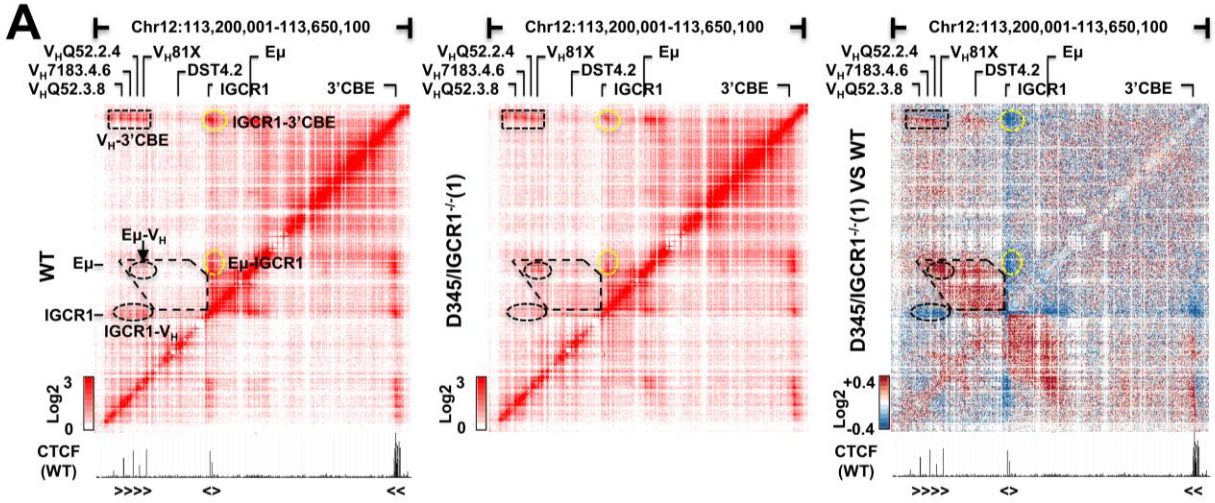
Xiang Qiu, Fei Ma, Mingming Zhao, Yaqiang Cao, Lillian Shipp, Angela Liu, Arun Dutta, Amit Singh, Fatima Zohra Braikia, Supriyo De, William H. Wood III, Kevin G. Becker, Weiqiang Zhou, Hongkai Ji, Keji Zhao, Michael L. Atchison, Ranjan Sen*

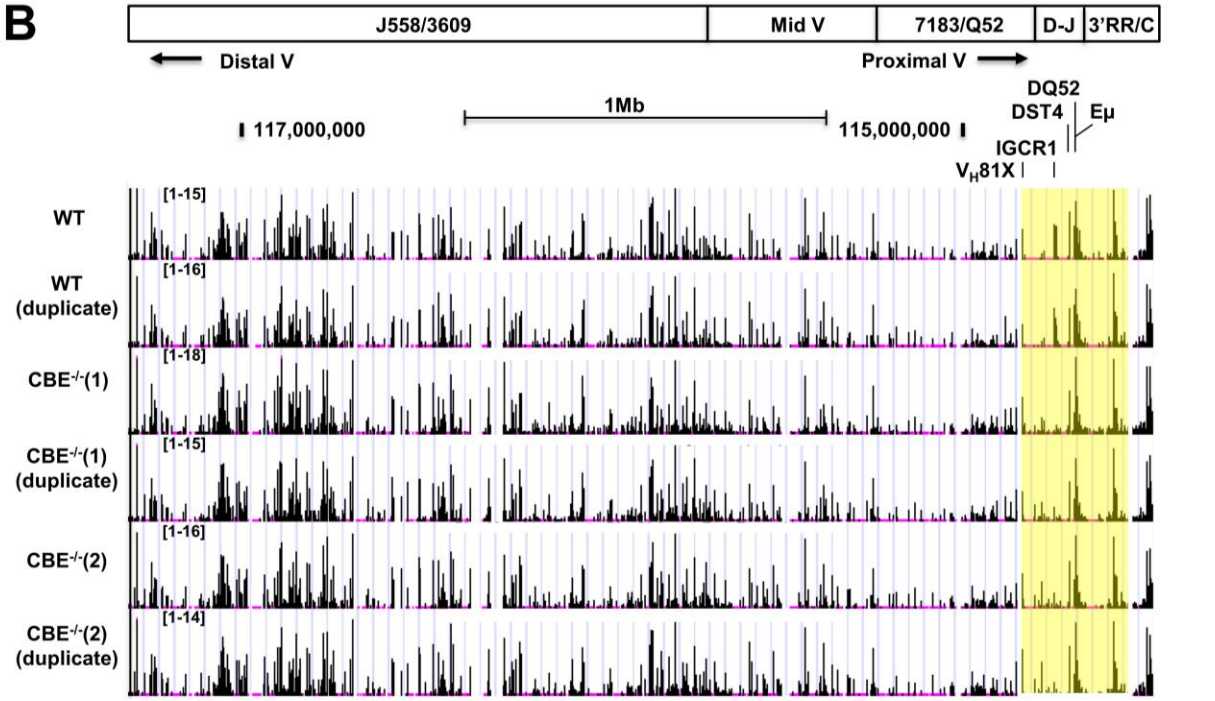
*Corresponding author. Email: senranja@grc.nia.nih.gov

Published 14 August 2020, *Sci. Adv.* **6**, eaaz8850 (2020)
DOI: 10.1126/sciadv.aaz8850

This PDF file includes:

Figs. S1 to S6
Tables S1 to S4


A
Chr12:113,200,001-113,650,100
VHQ52.2.4
VH7183.4.6
VHQ52.3.8
VH81X
DST4.2
IGCR1
Eμ
3'CBE
VH-3'CBE
IGCR1-3'CBE
Eμ-VH
Eμ-IGCR1
IGCR1-VH
WT
Eμ–
IGCR1–
D345/IGCR1-/-(1)
D345/IGCR1-/-(1) VS WT
Log2
CTCF
(WT)


B
J558/3609
Mid V
7183/Q52
D-J
3'RR/C
Distal V
Proximal V
117,000,000
1Mb
115,000,000
DQ52
DST4
Eμ
IGCR1
VH81X
WT
WT (duplicate)
CBE-/-(1)
CBE-/-(1) (duplicate)
CBE-/-(2)
CBE-/-(2) (duplicate)

C

Name	Chromatin 12 localization (mm9)		Log2 transformation						Adjusted p-value		
	Start	End	WT	WT (duplicate)	$CBE^{-/-}$(1)	$CBE^{-/-}$(1) (duplicate)	$CBE^{-/-}$(2)	$CBE^{-/-}$(2) (duplicate)	WT & $CBE^{-/-}$(1)	WT & $CBE^{-/-}$(2)	$CBE^{-/-}$(1) & $CBE^{-/-}$(2)
3'Cγ1	114560190	114560790	6.747	6.418	6.230	6.739	6.503	6.467	0.7527	0.7850	0.9987
Eμ-DQ52	114664638	114670222	10.831	10.843	11.108	10.881	10.963	10.830	0.4498	0.8124	0.6797
DST4	114684249	114684750	7.327	7.243	7.292	7.244	7.495	7.561	0.9337	0.2431	0.2310
5'DFL16.1	114719251	114721083	8.741	8.594	7.327	7.094	7.160	7.088	0.0018	0.0019	0.7246
IGCR1 (CBE2)	114723238	114724090	7.874	7.765	4.882	5.261	5.262	5.106	0.0007	0.0009	0.6900
IGCR1 (CBE1)	114725828	114726376	7.429	7.548	4.786	4.386	4.353	4.873	0.0009	0.0011	0.9481
DST4.2	114763445	114763776	3.509	3.531	4.786	5.060	5.317	5.880	0.0064	0.0023	0.0926
V_H81X	114816190	114816948	7.059	6.839	7.931	7.922	8.140	8.020	0.0057	0.0044	0.5063
V_HQ52.2.4	114825855	114826672	6.348	6.418	7.633	7.689	7.530	7.521	0.0017	0.0030	0.5119
V_H7183.4.6	114835183	114835818	6.880	6.839	7.239	7.171	7.233	7.268	0.0934	0.0815	0.8360

Fig. S1. Chromatin accessibility on WT and IGCR1-mutated *IgH* alleles.

A. Normalized contact frequency for Capture Hi-C of WT (left) and IGCR1-deleted (middle) *IgH* alleles (corresponding to Fig. 1B). Interacting regions discussed in the text are highlighted within dashed lines. Difference interaction map between WT and IGCR1-deleted IgH alleles is shown on the right. Contacts that decrease (blue) or increase (red) on IGCR1-deleted alleles are indicated. Position and orientation of CTCF bound sites are indicated below heat map (*47*).

B. Schematic map of *IgH* locus is shown on the top. ATAC-Seq assays were carried out in duplicate using RAG2-deficient pro-B cell lines with WT or IGCR1-mutated ($CBE^{-/-}$(1), $CBE^{-/-}$(2)) *IgH* alleles. The entire *IgH* locus is shown (chr12:114,452,000-117,310,000, mm9) after normalization and peak calling. The partial *IgH* locus (chr12:114,554,576-114,839,712, mm9) highlighted in yellow is provided in Fig. 1C.

C. Localization, and statistical analysis of ATAC peaks marked within colored rectangles in Fig. 1C. Differential analysis of peak was performed using limma (*54*) based on moderated t-tests. To adjust for multiple testing, *p*-values were adjusted using Benjamini-Hochberg (BH) procedure to obtain FDR. Adjusted *p*-values less than 0.01 were considered to be statistically different for ATAC peaks marked within green or red rectangles in Fig. 1C.

See also Fig. 1.

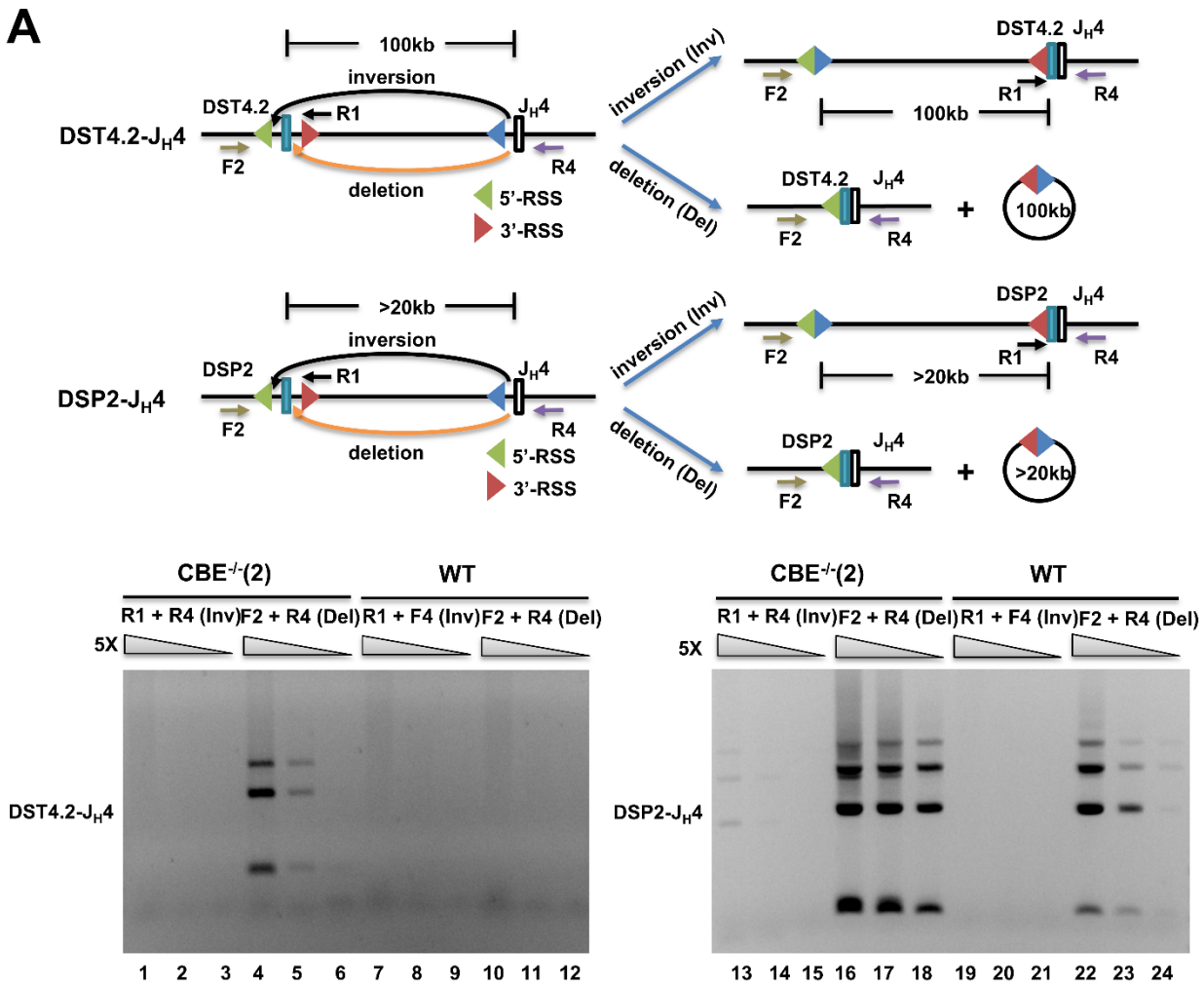

A
100kb
inversion
DST4.2
R1
JH4
DST4.2-JH4
F2
R4
deletion
5'-RSS
3'-RSS
inversion (Inv)
deletion (Del)
DST4.2 JH4
F2
R1
R4
100kb
DST4.2 JH4
F2
R4
100kb
>20kb
inversion
DSP2
R1
JH4
DSP2-JH4
F2
R4
deletion
5'-RSS
3'-RSS
inversion (Inv)
deletion (Del)
DSP2 JH4
F2
R1
R4
>20kb
DSP2 JH4
F2
R4
>20kb
CBE-/-(2)
WT
R1 + R4 (Inv)
F2 + R4 (Del)
R1 + F4 (Inv)
F2 + R4 (Del)
5X
DST4.2-JH4
1 2 3 4 5 6 7 8 9 10 11 12
CBE-/-(2)
WT
R1 + R4 (Inv)
F2 + R4 (Del)
R1 + F4 (Inv)
F2 + R4 (Del)
5X
DSP2-JH4
13 14 15 16 17 18 19 20 21 22 23 24
B
Synthesized oligo
60bp
60bp
R1
inversion
R4
or
60bp
60bp
F2
deletion
R4
DST4.2-JH4
Amplification efficiency
2.0
1.2
0.20
0.12
0.030
0.015
0.000
inversion R1 + R4
deletion F2 + R4
DSP2-JH4
Amplification efficiency
1.2
0.7
0.12
0.07
0.02
0.01
0.00
inversion R1 + R4
deletion F2 + R4
+ 100% oligo
+ 10% oligo
+ 1% oligo
- oligo

Fig. S2. Recombination of DST4.2 and DSP2 on WT and IGCR1-mutated *IgH* alleles.

A. Schematic representation of 5'- and 3'-RSS of DST4.2 or DSP2 recombining to J_H4 gene segment by deletion (orange arrow) or inversion (black arrow), respectively (left). Locations and orientation of primers used to assay recombination are indicated, together with the 12-RSSs of DST4.2 or DSP2 (green and red triangles) and 23-RSS flanking J_H4 gene segment (blue triangle). Products of each form of rearrangement are shown to the right. Signal-end junction by-products consisting of heptamer-heptamer fused RSSs are indicated as back-to-back triangles under inversional products. The corresponding region is deleted as an episome during deletional recombination (circle with fused triangles). Primer combinations specific for each kind of recombination are shown. Genomic DNA purified from IGCR1-mutated ($CBE^{-/-}$(2)) and WT *IgH* allele containing pro-B cell lines expressing RAG2 were amplified using the indicated primer combinations to assay DST4.2-J_H4 or DSP2-J_H4 recombination by deletion or inversion (lanes 1-24). Each set of 3 lanes contain 5-fold increasing amounts of genomic DNA starting at 8 ng (lane 3, 6, 9, 15, 18, 21 and 24). Reaction products were separated by agarose gel electrophoresis and visualized by ethidium bromide staining. Data shown is representative of 2 biological replicate experiments.

B. Amplification efficiency of primers designed to test recombination products of inversional (R1+R4) or deletional (F2+R4) recombination of DST4.2 or DSP2 were tested with 120 nucleotide long synthetic recombination products, with 60 nucleotides around each primer (top). Realtime-PCR was carried out using indicated primers with serially diluted synthesized recombination products (100%, 10% and 1%) mixed with 200ng genomic DNA from a RAG2 deficient pro-B cell line. Reactions containing no synthetic recombination products served as the negative control (– oligo). Amplification efficiency is shown relative to the amount of PCR

product obtained in reactions containing 100% inversional oligo using primers R1 and R4. Data are shown as mean ± SEM of three independent experiments.

See also Fig. 2, Table S4.

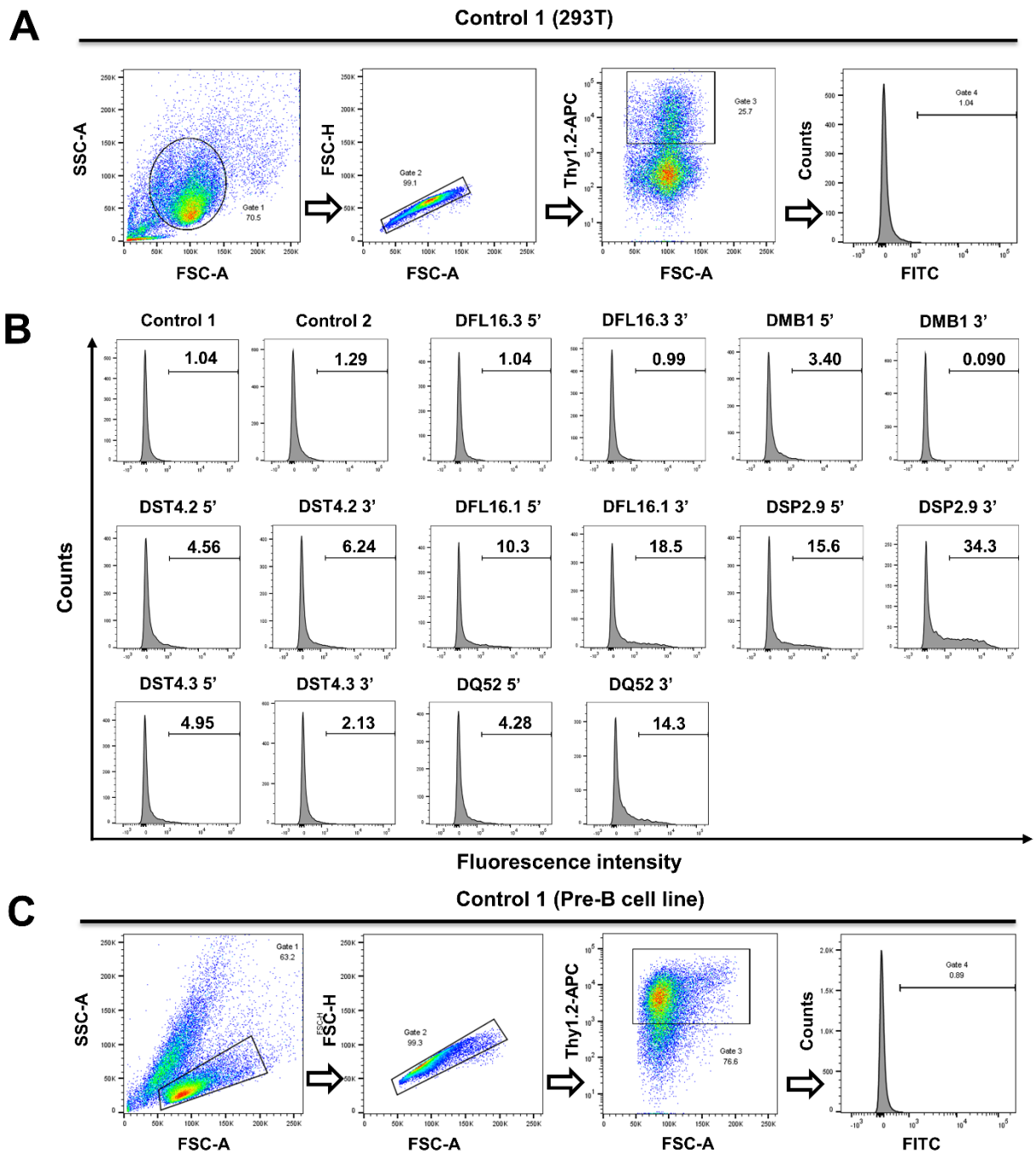

A
Control 1 (293T)
SSC-A
FSC-A
Gate 1
70.5
FSC-H
FSC-A
Gate 2
99.1
Thy1.2-APC
FSC-A
Gate 3
25.7
Counts
FITC
Gate 4
1.04
B
Control 1
1.04
Control 2
1.29
DFL16.3 5'
1.04
DFL16.3 3'
0.99
DMB1 5'
3.40
DMB1 3'
0.090
DST4.2 5'
4.56
DST4.2 3'
6.24
DFL16.1 5'
10.3
DFL16.1 3'
18.5
DSP2.9 5'
15.6
DSP2.9 3'
34.3
DST4.3 5'
4.95
DST4.3 3'
2.13
DQ52 5'
4.28
DQ52 3'
14.3
Counts
Fluorescence intensity
C
Control 1 (Pre-B cell line)
SSC-A
FSC-A
Gate 1
63.2
FSC-H
FSC-A
Gate 2
99.3
Thy1.2-APC
FSC-A
Gate 3
76.6
Counts
FITC
Gate 4
0.89

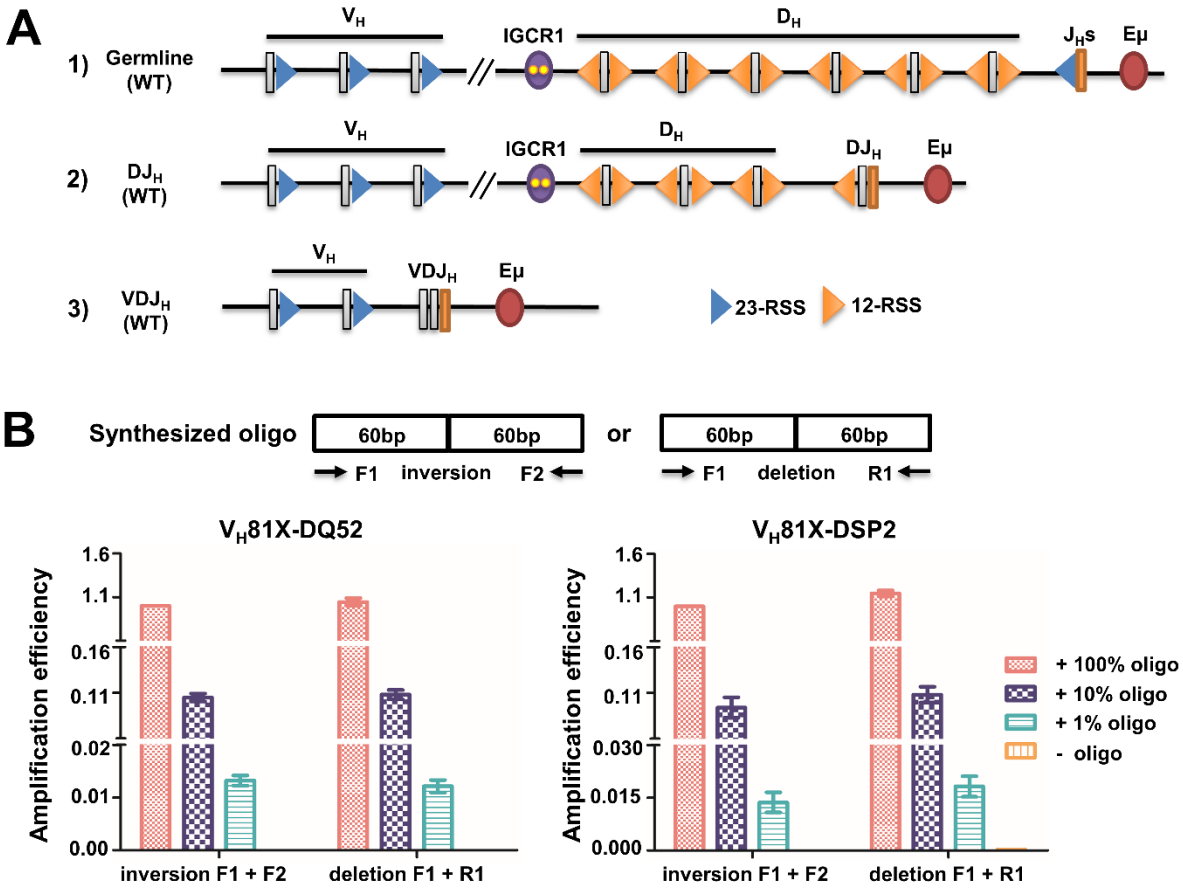
Fig. S3. Fluorescence-activated cell sorting (FACS) analysis of D_H RSS strength.

A. FACS gating strategy and analysis of Control 1 is shown. 293T cells were co-transfected with recombination reporters and expression vectors for RAG1 and RAG2. Non-functional 12-RSS is indicated as Control 1. After two days in culture, Thy1.2 positive cells were gated from single and live cells, and used for analysis of GFP-expressing cells. GFP intensity within $Thy1.2^+$ populations was used as a measure of recombination efficiency.

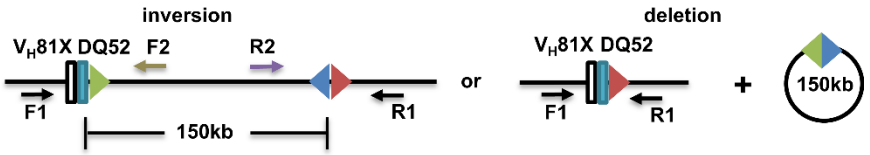
B. 293T cells were co-transfected with recombination reporters and expression vectors for RAG1 and RAG2. Sequences of 12-RSSs are listed in Table S1. Control 2 is indicated as 293T cells transfected with recombination reporter (3'-RSS of DFL16.1) in the absence of RAG1/2. Average results of three independent experiments are shown in Fig. 2C and 2D.

C. A RAG1/2 expressing pre-B cell line was infected with recombination reporter retroviruses as described (*29, 30*). Non-functional 12-RSS is indicated as Control 1. After two days in culture, Thy1.2 positive cells were gated from single and live cells, and used for analysis of GFP intensity. GFP intensity within $Thy1.2^+$ populations was used as a measure of recombination efficiency.

See also Fig. 2, Table S1.


A
1) Germline (WT)
2) DJH (WT)
3) VDJH (WT)
VH
DH
IGCR1
JHs
Eμ
DJH
VDJH
23-RSS
12-RSS
B
Synthesized oligo
60bp
60bp
or
F1
inversion
F2
F1
deletion
R1
VH81X-DQ52
VH81X-DSP2
Amplification efficiency
1.6
1.1
0.16
0.11
0.02
0.01
0.00
0.030
0.015
0.000
inversion F1 + F2
deletion F1 + R1
+ 100% oligo
+ 10% oligo
+ 1% oligo
- oligo

C



D

V_H81X-DQ52 (F1-F2)

3' V_H81X | Reverse DQ52 | Reverse 5' 12-RSS

Heptamer Nonamer

ACAGCCTTGTATTACTGTGCAAGACA GTCCCAGTTAGCACTGTGGTGCTCCGCTTAGTCAAAACC

. . . . . .G. . . . . .

. . . . . .

. . . . . .

. . . . . .C. . . . . .

. . . . . .

E

V_H81X-DQ52 (F1-R1)

3' V_H81X | DQ52 | 3' 12-RSS

Heptamer Nonamer

ACAGCCTTGTATTACTGTGCAAGACACTAACTGGGACCACGGTGACACGTGGCTCAACAAAAACC

. . . . . .

. . . . . .

. . . . . .

. . . . . .

. . . . . .

. . . . . .

F

V_H81X-DQ52 (R2-R1)

23-RSS | 3' 12-RSS

Nonamer Heptamer Heptamer Nonamer

GGTTTTAGTTTGAGCTCACAGTAACTTTTGCTCATTGTGCACGGTGACACGTGGCTCAACAAAAACC

. . . . . .

. . . . . .

. . . . . .

. . . . . .

. . . . . .

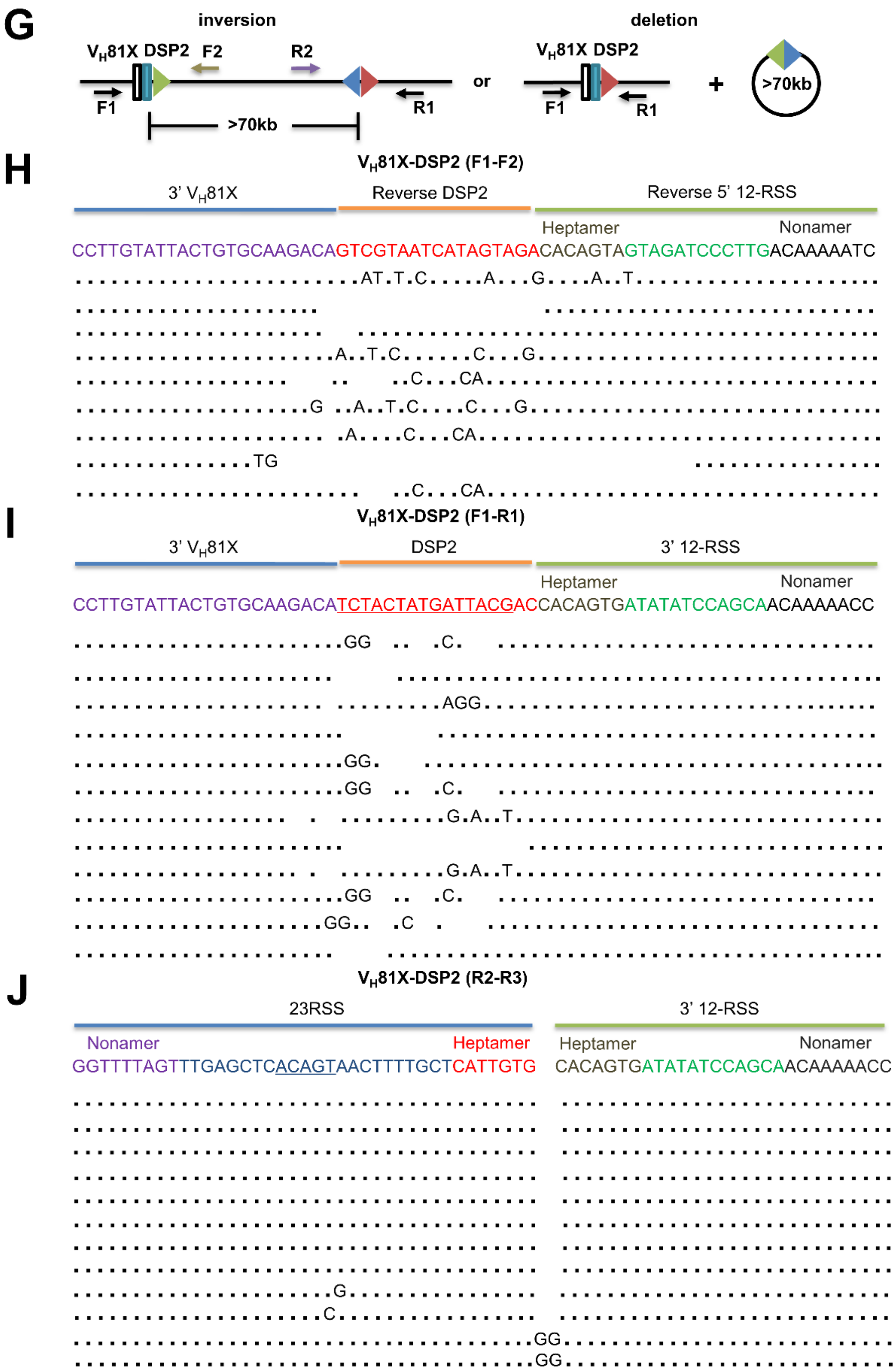

G
inversion
VH81X DSP2 F2 R2
F1 R1
>70kb
or
deletion
VH81X DSP2
F1 R1
+
>70kb
H
VH81X-DSP2 (F1-F2)
3' VH81X
Reverse DSP2
Reverse 5' 12-RSS
Heptamer
Nonamer
CCTTGTATTACTGTGCAAGACAGTCGTAATCATAGTAGACACAGTAGTAGATCCCTTGACAAAAATC
I
VH81X-DSP2 (F1-R1)
3' VH81X
DSP2
3' 12-RSS
Heptamer
Nonamer
CCTTGTATTACTGTGCAAGACATCTACTATGATTACGACCACAGTGATATATCCAGCAACAAAAACC
J
VH81X-DSP2 (R2-R3)
23RSS
3' 12-RSS
Nonamer
Heptamer
Heptamer
Nonamer
GGTTTTAGTTTGAGCTCACAGTAACTTTTGCTCATTGTG
CACAGTGATATATCCAGCAACAAAAACC

Fig. S4. V_H81X-DQ52 and V_H81X-DSP2 recombination junction sequences.

A. Schematic map of germline (lane 1), DJ_H (lane 2) and VDJ_H (lane 3) recombined WT *IgH* alleles. DJ_H alleles can retain several unrearranged D_H gene segments (lane 2). 5' of the DJ_H junction is recombined to V_H to deletes germline D_H gene segments and IGCR1 (lane 3). RSSs are indicated as colored triangles: 23-RSS (blue), 12-RSS (yellow). Recombination regulatory sequences Eμ and IGCR1 are indicated as colored ovals.

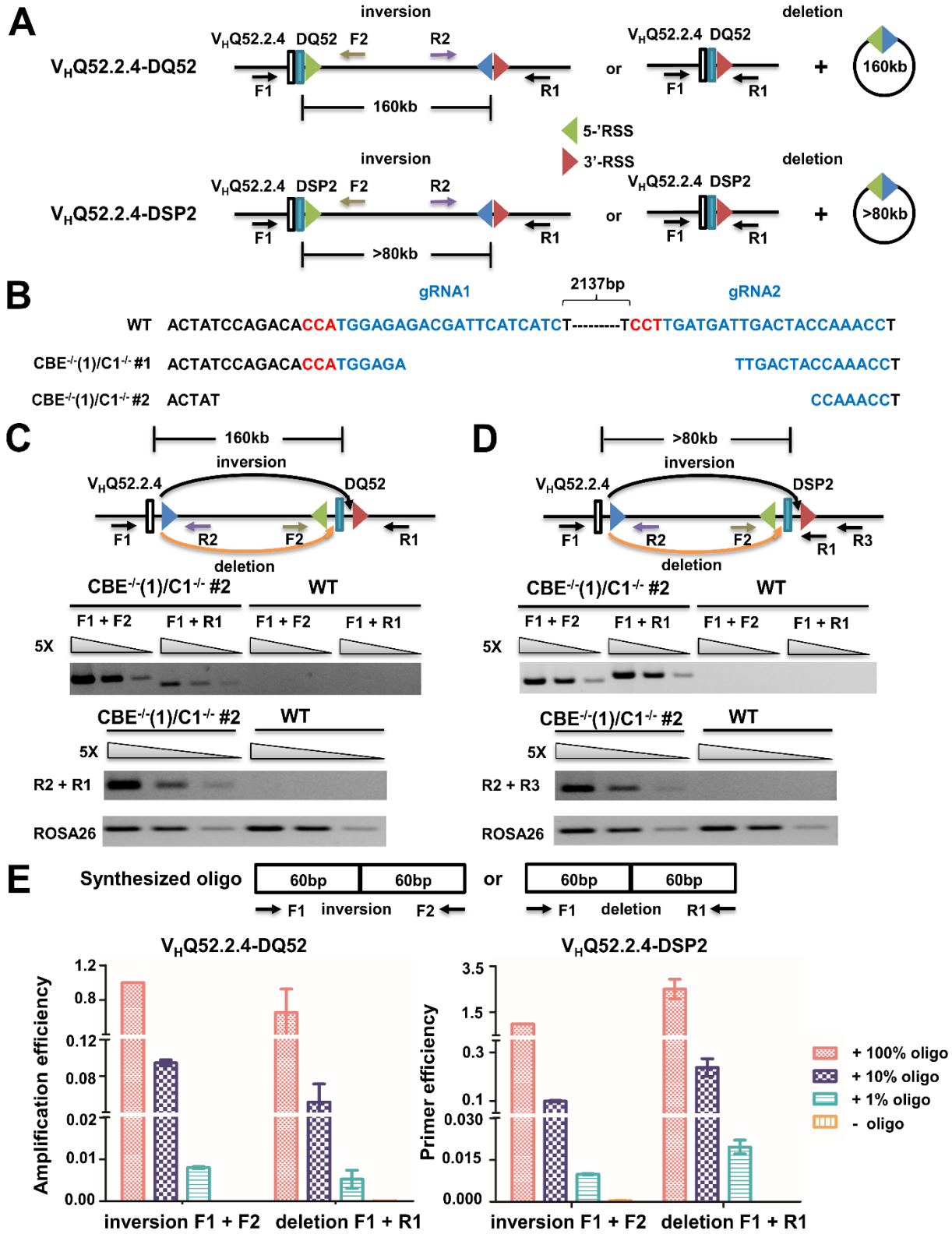
B. Amplification efficiency of primers (V_H81X-DQ52 and V_H81X-DSP2) used to detect inversional or deletional recombination was tested with 120 nucleotide long synthetic recombination products, with 60 nucleotides around each primer (top). Realtime-PCR was carried out using indicated primers with serially diluted synthesized recombination products (100%, 10% and 1%) mixed with 200ng genomic DNA from a RAG2 deficient pro-B cell line and. Reactions containing no synthetic recombination products served as the negative control (–oligo). Amplification efficiency is shown relative to the amount of PCR product obtained in reactions containing 100% inversional oligo. Data are shown as mean ± SEM of three independent experiments.

C and G. Schematic representation of V_H81X rearrangement to 5'- and 3'-RSS of DQ52 (C) or DSP2 (G) gene segment by inversion (left) or deletion (right) was described in Fig. 3A.

D-F and H-J. Genomic DNA purified from IGCR1-mutated pro-B cell line ($CBE^{-/-}$(2)) was amplified with different primer combinations that identified inversional (D and H, using primers F1-F2) or deletional (E and I, using primers F1-R1) recombination products of V_H81X and DQ52 or DSP2. Signal-end junctional sequences (F and J) were obtained after amplification with primers R2 and R1 or R3. Locations and orientation of primers used to assay recombination are indicated above the sequences. Purified amplicons were cloned into pGEM®-

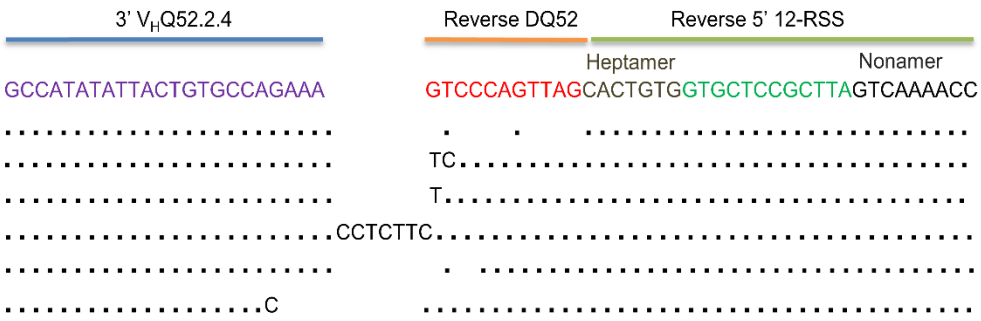
T Vector, followed by sequencing. First line represents predicted junction sequence of V_H81X-DQ52 or DSP2.2a recombination, followed by sequences of independently cloned junctions. DSP2.2a is one of DSP2 family gene segments. Dots identify the same sequence as the top line. Blanks and added nucleotides represent deletion and addition of nucleotides, respectively, during recombination.

See also Fig. 3, Table S4.


A
inversion
deletion
$V_HQ52.2.4$-DQ52
$V_HQ52.2.4$ DQ52 F2 R2
F1 R1
160kb
or
+
160kb
5-'RSS
3'-RSS
$V_HQ52.2.4$-DSP2
$V_HQ52.2.4$ DSP2 F2 R2
>80kb
B
gRNA1
2137bp
gRNA2
WT ACTATCCAGACACCATGGAGAGACGATTCATCATCT---------TCCTTGATGATTGACTACCAAACCT
CBE-/-(1)/C1-/- #1 ACTATCCAGACACCATGGAGA TTGACTACCAAACCT
CBE-/-(1)/C1-/- #2 ACTAT CCAAACCT
C
160kb
inversion
$V_HQ52.2.4$ DQ52
F1 R2 F2 R1
deletion
CBE-/-(1)/C1-/- #2
WT
F1 + F2
F1 + R1
5X
R2 + R1
ROSA26
D
>80kb
$V_HQ52.2.4$ DSP2
F1 R2 F2 R1 R3
R2 + R3
E
Synthesized oligo
60bp
60bp
or
F1 inversion F2
F1 deletion R1
$V_HQ52.2.4$-DQ52
Amplification efficiency
1.2
0.8
0.12
0.08
0.02
0.01
0.00
inversion F1 + F2
deletion F1 + R1
$V_HQ52.2.4$-DSP2
Primer efficiency
3.5
1.5
0.3
0.1
0.030
0.015
0.000
+ 100% oligo
+ 10% oligo
+ 1% oligo
- oligo

F

V_HQ52.2.4-DQ52 (F1-F2)

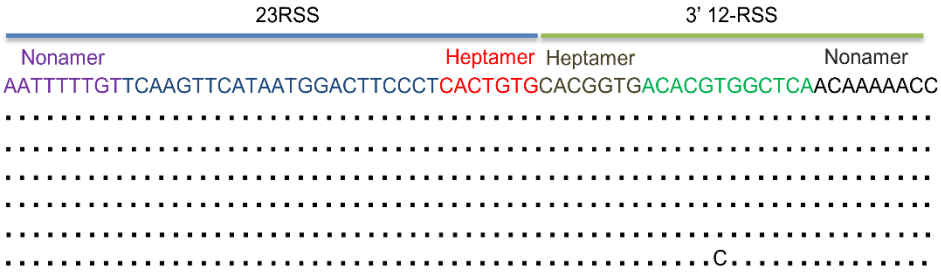


G

V_HQ52.2.4-DQ52 (F1-R1)

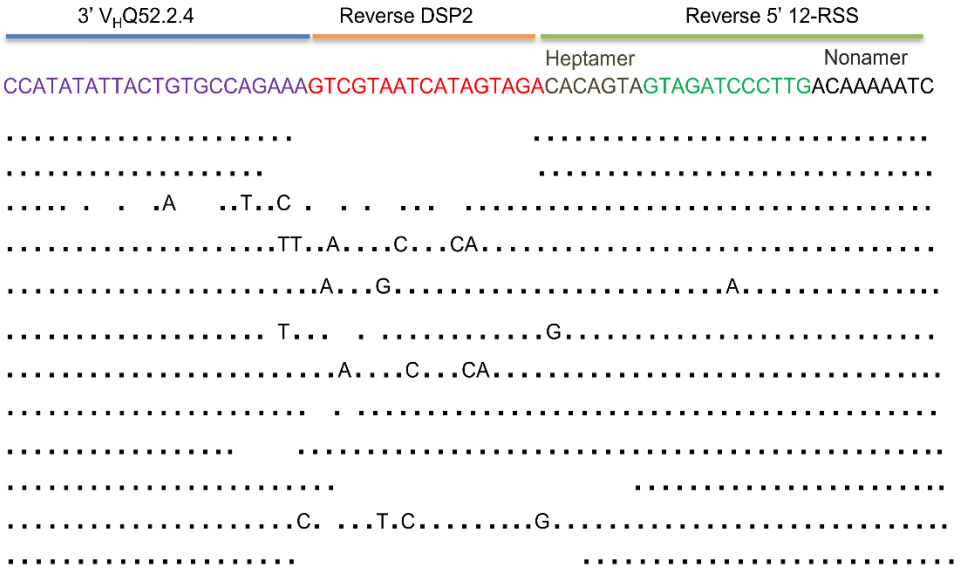
H

V_HQ52.2.4-DQ52 (R2-R1)



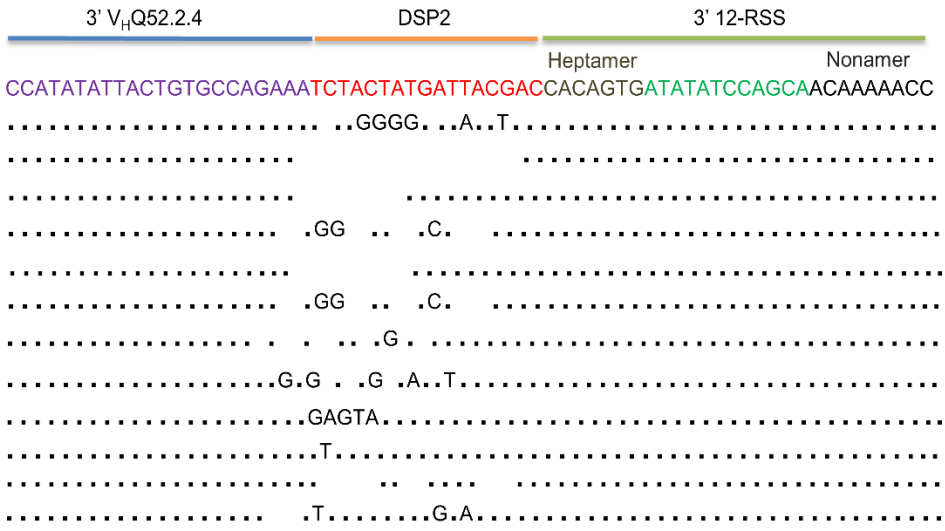
I

V_HQ52.2.4-DSP2 (F1-F2)



J

V_HQ52.2.4-DSP2 (F1-R1)



K

V_HQ52.2.4-DSP2 (R2-R3)

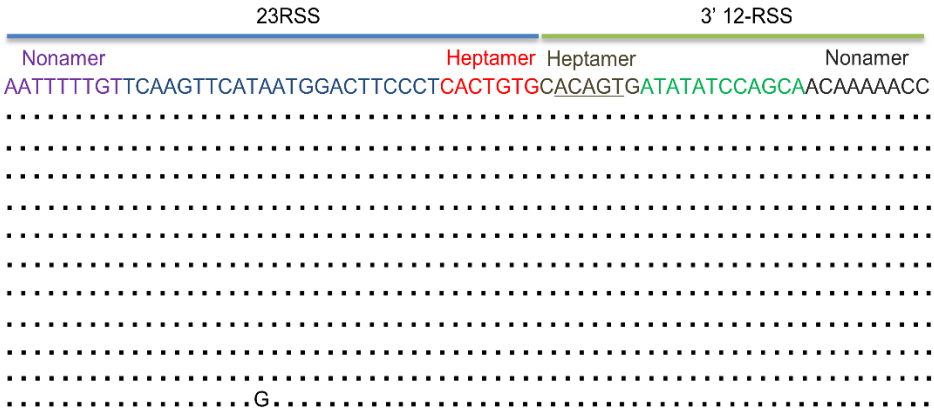


Fig. S5. $V_HQ52.2.4$-DQ52 and $V_HQ52.2.4$-DSP2 recombination junction sequences.

A. Schematic representation of $V_HQ52.2.4$ rearrangement to 5'- and 3'-RSS of DQ52 (top) and DSP2 (bottom) gene segments by inversion or deletion, respectively. Locations and orientation of primers used to assay recombination are indicated. Products of each form of rearrangements are shown. Signal-end junction by-products consisting of heptamer-heptamer fused RSSs, are indicated as back-to-back triangles under inversional products. The corresponding region is deleted as an episome during deletional recombination (circle with fused triangles). Primer combinations specific for each kind of recombination are shown.

B. $CBE^{-/-}(1)/C1^{-/-}$ #1 and $CBE^{-/-}(1)/C1^{-/-}$ #2 are two different subclones derived from IGCR1-mutated pro-B cell line $CBE^{-/-}(1)$ with V_H81X and CBE deletion using CRISPR/Cas9 (C1) (*24*). Top line represents germline sequence around C1, followed by sequence of $CBE^{-/-}(1)/C1^{-/-}$ #1 and $CBE^{-/-}(1)/C1^{-/-}$ #2. PAM, gRNA1 and gRNA2 sequences are indicated in red and blue, respectively.

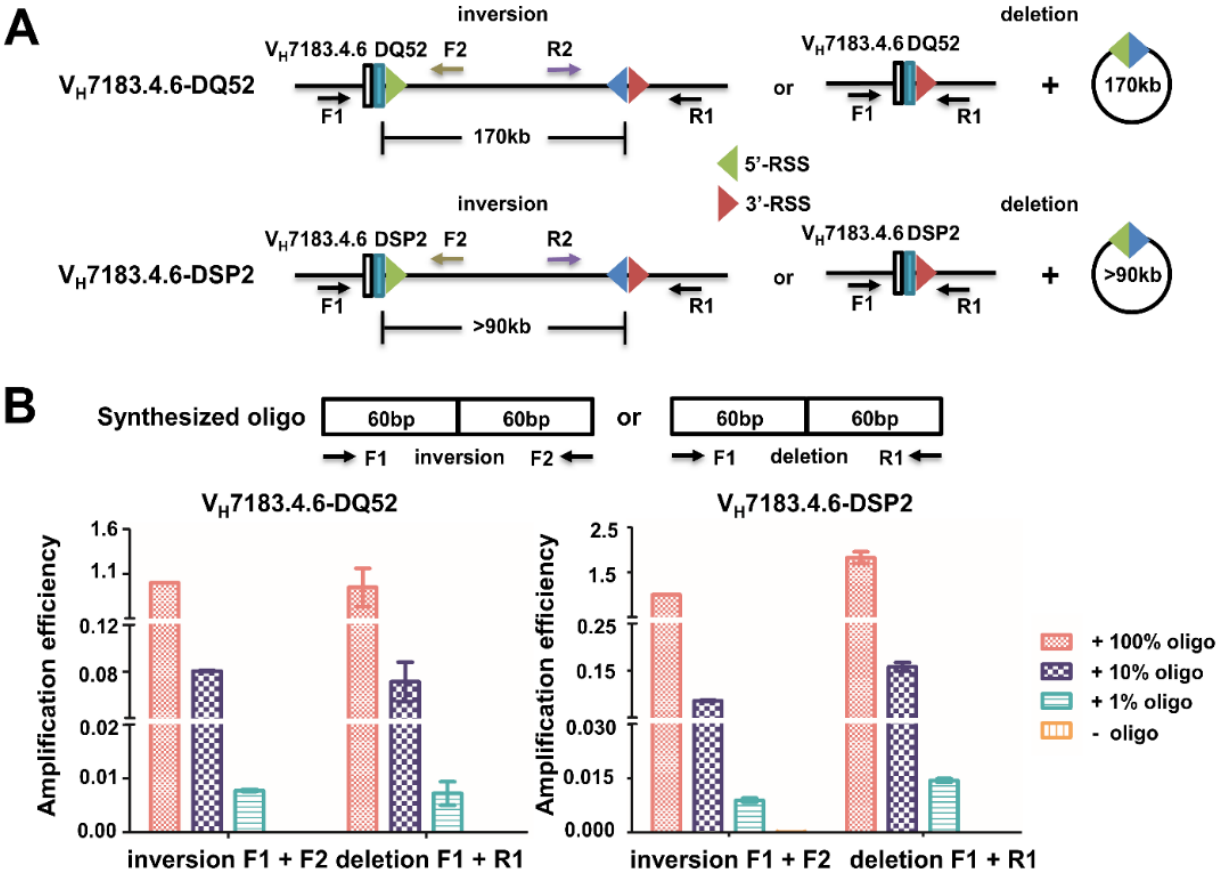
C and D. Recombination of $V_HQ52.2.4$-DQ52 or $V_HQ52.2.4$-DSP2 in $CBE^{-/-}(1)/C1^{-/-}$ #1 are shown in Fig. 5B and 5C. Recombination of $V_HQ52.2.4$-DQ52 or $V_HQ52.2.4$-DSP2 in $CBE^{-/-}(1)/C1^{-/-}$ #2 are shown in part C and D, as described in Fig. 5.

E. Amplification efficiency analyses were carried out as described in Fig. S2B to assay inversion versus deletion for $V_HQ52.2.4$-DQ52 (left) or $V_HQ52.2.4$-DSP2 (right), respectively. Data are shown as mean ± SEM of three independent experiments.

F-K. Genomic DNA purified from IGCR1-mutated pro-B cell line with C1 deletion ($CBE^{-/-}(1)/C1^{-/-}$ #1) was amplified with different primer combinations that identified inversional (F and I, using primers F1-F2) or deletional (G and J, using primers F1-R1) recombination products of $V_HQ52.2.4$ and DQ52 or DSP2. Signal-end junctional sequences (H and K) were obtained after

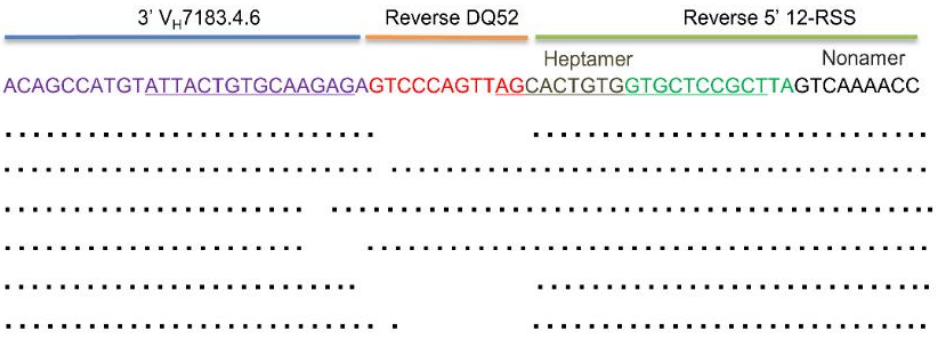
amplification with primers R2 and R1 or R3. Locations and orientation of primers used to assay recombination are indicated above the sequences. Purified amplicons were cloned into pGEM®-T Vector, followed by sequencing. First line represents predicted junction sequence of V_HQ52.2.4-DQ52 or DSP2.2a recombination, followed by sequences of independently cloned junctions. DSP2.2a is one of DSP2 family gene segments. Dots identify the same sequence as the top line. Blanks and added nucleotides represent deletion and addition of nucleotides, respectively, during recombination.

See also Fig. 5 and Table S4.


A
inversion
deletion
VH7183.4.6-DQ52
VH7183.4.6 DQ52
F2
R2
F1
R1
170kb
or
+
5'-RSS
3'-RSS
VH7183.4.6-DSP2
VH7183.4.6 DSP2
>90kb
B
Synthesized oligo
60bp
inversion
deletion
Amplification efficiency
inversion F1 + F2
deletion F1 + R1
+ 100% oligo
+ 10% oligo
+ 1% oligo
- oligo

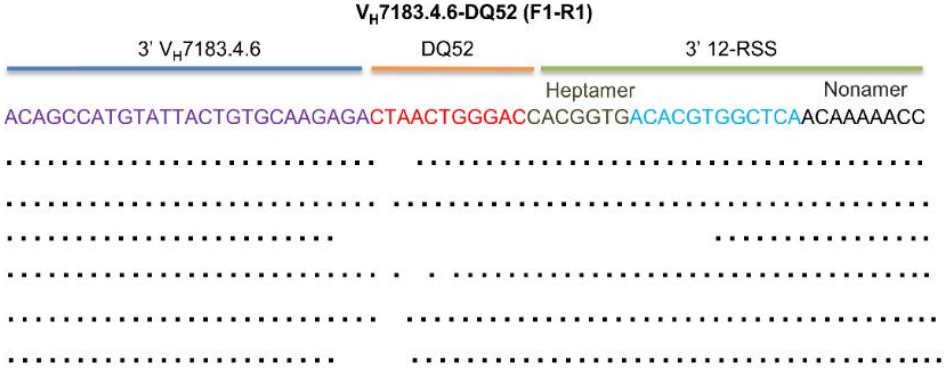
C

V_H7183.4.6-DQ52 (F1-F2)



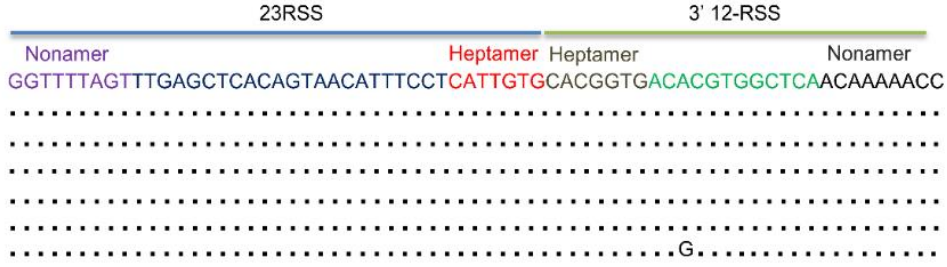
D

V_H7183.4.6-DQ52 (F1-R1)



E

V_H7183.4.6-DQ52 (R2-R1)



F

V_H7183.4.6-DSP2 (F1-F2)

G

V_H7183.4.6-DSP2 (F1-R1)

H

V_H7183.4.6-DSP2 (R2-R3)

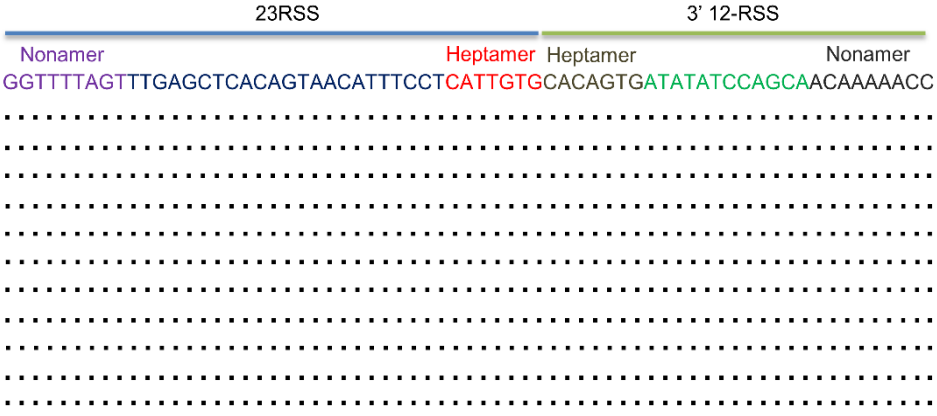


Fig. S6. V_H7183.4.6-DQ52 and V_H7183.4.6-DSP2 recombination junction sequences.

A. Schematic representation of V_H7183.4.6 rearrangements to 5'- and 3'-RSS of DQ52 (top) and DSP2 (bottom) gene segments by inversion or deletion, respectively. Locations and orientation of primers used to assay recombination are indicated. Products of each form of rearrangements are shown. Signal-end junction by-products consisting of heptamer-heptamer fused RSSs, are indicated as back-to-back triangles under inversional products. The corresponding region is deleted as an episome during deletional recombination (circle with fused triangles). Primer combinations specific for each kind of recombination are shown.

B. Amplification efficiency analyses were carried out as described in Fig. S2B to assay inversion versus deletion for V_H7183.4.6-DQ52 (left) or V_H7183.4.6-DSP2 (right), respectively. Data are shown as mean ± SEM of three independent experiments.

C-H. Genomic DNA purified from IGCR1-mutated pro-B cell line with C1 and C2 deletion ($CBE^{-/-}$(1)/$C1^{-/-}$ $C2^{-/-}$) was amplified with different primer combinations that identified inversional (C and F, using primers F1-F2) or deletional (D and G, using primers F1-R1) recombination products of V_H7183.4.6 and DQ52 or DSP2. Signal-end junctional sequences (E and H) were obtained after amplification with primers R2 and R1 or R3. Locations and orientation of primers used to assay recombination are indicated above the sequences. Purified amplicons were cloned into pGEM®-T Vector, followed by sequencing. First line represents predicted junction sequence of V_H7183.4.6-DQ52 or DSP2.2a recombination, followed by sequences of independently cloned junctions. DSP2.2a is one of DSP2 family gene segments. Dots identify the same sequence as the top line. Blanks and added nucleotides represent deletion and addition of nucleotides, respectively, during recombination.

See also Fig. 5 and Table S4.

Table S1. D_H 12-RSS recombination information content (RIC) scores and strength.

D_H genes	12-RSS sequence	RIC score	RIC pass/fail
DFL16.3	cacagtagtagagcctttctccaaaaac	-32.4731235371929	PASS
	cacaatggtgtatctagtagcaaaagtt	-37.456359861778	PASS
DMB1	cacagtgagagaggcagatactgaattc	-33.8741979229379	PASS
	cacagtgatacagagcatccatggaaga	-33.0860539045419	PASS
DST4.2	cactgtgacagtaacttgttcaaaatcc	-27.8193402641352	PASS
	cactgtaagaaaagctcaaaccaaaact	-32.9230344632939	PASS
DFL16.1	cacagtagtagatcccttcacaaaaagc	-15.3467711676282	PASS
	cacagtgctatatccatcagcaaaaacc	-16.9821406634171	PASS
DSP2.9	cacagtagtagatcccttgacaaaaatc	-15.0838006588257	PASS
	cacagtgatatatccagctacaaaaacc	-14.0122887731066	PASS
DST4.3	cactgtgacaataccttgttcaaaatcc	-28.776801859089	PASS
	cagctagggctgtcactgaaagaaaagc	-50.5463936019675	FAIL
DQ52	cactgtggtgctccgcttagtcaaaacc	-25.0935566077474	PASS
	cacggtgacacgtggctcaacaaaaacc	-20.724575878483	PASS
Control 1	ggatccggatccggatccggatccggat		

Above table contains RSS sequences and predicted RIC scores by DNAgrab (http://www.itb.cnr.it/rss/analyze.html) for D_H (129 strain mice) genes. A PASS score for an RSS is RIC >-40.

	293T (% of EGFP cells)			Pre-B cell line (% of EGFP cells)	
12-RSS	Rep1	Rep2	Rep3	Rep1	Rep2
Control 1	0.51	1.04	0.850	0.89	0.83
Control 2	0.96	1.29	0.990		
DFL16.3 5'-RSS	1.610	1.040	1.720		
DFL16.3 3'-RSS	1.75	0.99	1.730		
DMB1 5'-RSS	2.160	3.400	3.000		
DMB1 3'-RSS	0.10	0.09	0.062	0.82	0.71
DST4.2 5'-RSS	3.430	4.560	6.130	4.83	3.03
DST4.2 3'-RSS	7.69	6.24	7.000	6.21	6.57
DFL16.1 5'-RSS	11.800	10.300	10.700	17.60	14.30
DFL16.1 3'-RSS	19.20	18.50	20.300	21.60	20.50
DSP2.9 5'-RSS	17.800	15.600	16.900		
DSP2.9 3'-RSS	36.70	34.30	32.400		
DST4.3 5'-RSS	5.500	4.950	4.980		
DST4.3 3'-RSS	1.01	2.13	1.310		
DQ52 5'-RSS	3.140	4.280	3.020		
DQ52 3'-RSS	15.90	14.30	15.700		

Rep stands for replication.

Table S2. Deep sequencing results with DQ52 as bait.

Sample (BARCODE)	Del (raw reads)	Del (percentage)	Inv (raw reads)	Inv (percentage)
RAG2$^{-/-}$ Rep1 (CGCATTAA)	106420	99.97463526%	27	0.025364736%
RAG2$^{-/-}$ Rep2 (CTTCGCGC)	85852	99.96274044%	32	0.037259559%
RAG2$^{-/-}$ average	96136	99.96868785%	30	0.031312147%
CBE$^{-/-}$ (1) Rep1 (AACACCTA)	754213	67.78096216%	358508	32.21903784%
CBE$^{-/-}$ (1) Rep2 (ACTAACTG)	525475	73.57728643%	188706	26.42271357%
CBE$^{-/-}$ (1) average	639844	70.6791243%	273607	29.3208757%
CBE$^{-/-}$ (2) Rep1 (ATCGCCAG)	485834	64.9904287%	261713	35.0095713%
CBE$^{-/-}$ (2) Rep2 (CATTCCAA)	513956	63.36882209%	297099	36.63117791%
CBE$^{-/-}$ (2) average	499895	64.17962539%	279406	35.82037461%

Del, Inv and Rep stand for deletion, inversion and replication, respectively.

Table S3. Deep sequencing results with DSP2 as bait.

Sample (BARCODE)	Del (raw reads)	Del (percentage)	Inv (raw reads)	Inv (percentage)
RAG2$^{-/-}$ Rep1 (CGAACTGT)	1154	100%	0	0%
RAG2$^{-/-}$ Rep2 (CTCTGTCT)	803	100%	0	0%
RAG2$^{-/-}$ average	979	100%	0	0%
CBE$^{-/-}$ (1) Rep1 (CTTAAGAT)	5276	90.68408388%	542	9.315916122%
CBE$^{-/-}$ (1) Rep2 (GCAAGTAG)	9842	91.29023282%	939	8.709767183%
CBE$^{-/-}$ (1) average	7559	90.98715835%	741	9.012841653%
CBE$^{-/-}$ (2) Rep1 (ATATAGGA)	25446	92.22572578%	2145	7.77427422%
CBE$^{-/-}$ (2) Rep2 (CACGTGTT)	29915	94.01615387%	1904	5.98384613%
CBE$^{-/-}$ (2) average	27681	93.12093983%	2025	6.879060175%

Del, Inv and Rep stand for deletion, inversion and replication, respectively.

Table S4. Primer list.

cDNA real-time PCR		
Name	Sequence (5'-3')	Reference
DST4.2--F	ACAAGTTACTGTCACAGTGGGC	
DST4.2-R	ACACAAAGCCAGAAAGGGAATAG	
3'DST4.2-F	ACAGGCAGATAGGATCAGTCTT	
3'DST4.2-R	AGGGGTAGTATGGTGTCTGTTA	
DQ52-F	TGGTGCAAGGTTTTGACTAAGC	(*5*)
DQ52-R	CCAAACAGAGGGTTTTTGTTGAG	(*5*)
γ-actin-F	GGTGTCCGGAGGCACTCTT	(*5*)
γ-actin-R	TGAAAGTGGTCTCATGGATACCA	(*5*)
Eμ-F	AATACCCGAAGCATTTACAGTGACT	(*24*)
Eμ-R	AAGATTTGTGAAGCCGTTTTGACCA	(*24*)
DJ/VD recombination assay		
DST4.2-F (P2) (Fig.2)	ACTCAGCAGGAAGGCTGTGAAGTC	
DST4.2-R (P3) (Fig. 2)	ATAAAGGACTTGATATTCAGGAGC	
V_H81X-F (P1) (F1)	CTGAAACTCTCCTGTGAATCCAA	(*24*)
DSP2-F (P4)(F2) (Fig, 2-3)	ATGGCCCCTGACACTCTGCACTGCTA	(*40*)
J_H4-R	GGGTCTAGACTCTCAGCCGGCTCCCTCAGGG	(*40*)
ROSA26-F	AAAGTCGCTCTGAGTTGTTAT	
ROSA26-R	GGAGCGGGAGAAATGGATATG	
DQ52-F (F2) (Fig. 3 and 5)	GCGACTGTTTTGAGAGAAATCATTG	(*40*)
DQ52-F (R1) (Fig. 3 and 5)	CATCCACCCTTCTGATGCTTGCATT	(*8*)
V_H81X-R (R2) (Fig. 3)	GGCTGGTGTCTGGTCTACCATT	
DSP2-R1 (Fig. 3 and 5)	TGGGTTTTTGTTGCTGGATATATC	(*40*)
DSP2-R3 (Fig. 3 and 5)	TAGTGTGCTTTCACCTGTCTGTG	
V_HQ52.2.4-F1 (Fig. 5)	AGCTCACACTAAGCTGAGAAGCT	

V_HQ52.2.4-R2 (Fig. 5)	TTGCAGCTCTGTAGGATGAACAA	
V_H7183.4.6-F1 (Fig. 5)	GGAGGCCGGTCCTGGATTCGAGTT	
V_H7183.4.6-R2 (Fig. 5)	AATGCATCAGTAGGACAGAGTAC	
Synthesized oligo		
DST4.2-F2-J_H4	ACTCAGCAGGAAGGCTGTGAAGTCTCTGACAGGATTT TGAACAAGTTACTGTCACAGTGGAGCATTGCAGACTA ATCTTGGATATTTGTCCCTGAGGGAGCCGGCTGAGAG TCTAGACCC	
DST4.2-F2-J_H4 (R)	GGGTCTAGACTCTCAGCCGGCTCCCTCAGGGACAAAT ATCCAAGATTAGTCTGCAATGCTCCACTGTGACAGTA ACTTGTTCAAAATCCTGTCAGAGACTTCACAGCCTTCC TGCTGAGT	
DST4.2-R1-J_H4	ATAAAGGACTTGATATTCAGGAGCACCTTGCTGGTTG GGTATGTATACACATATACATGCAGCATTGCAGACTA ATCTTGGATATTTGTCCCTGAGGGAGCCGGCTGAGAG TCTAGACCC	
DST4.2-R1-J_H4 (R)	GGGTCTAGACTCTCAGCCGGCTCCCTCAGGGACAAAT ATCCAAGATTAGTCTGCAATGCTGCATGTATATGTGTA TACATACCCAACCAGCAAGGTGCTCCTGAATATCAAG TCCTTTAT	
DSP2-F2-J_H4	ATGGCCCCTGACACTCTGCACTGCTACCTCTGGCCCCA CCAGACCATGTTCCTGCATAACAGCATTGCAGACTAA TCTTGGATATTTGTCCCTGAGGGAGCCGGCTGAGAGT CTAGACCC	
DSP2-F2-J_H4 (R)	GGGTCTAGACTCTCAGCCGGCTCCCTCAGGGACAAAT ATCCAAGATTAGTCTGCAATGCTGTTATGCAGGAACA TGGTCTGGTGGGGCCAGAGGTAGCAGTGCAGAGTGTC AGGGGCCAT	
DSP2-R1-J_H4	TGGGTTTTTGTTGCTGGATATATCACTGTGGTCGTAAT CATAGTAGACACAGTAGTAGATAGCATTGCAGACTAA TCTTGGATATTTGTCCCTGAGGGAGCCGGCTGAGAGT CTAGACCC	
DSP2-R1-J_H4 (R):	GGGTCTAGACTCTCAGCCGGCTCCCTCAGGGACAAAT ATCCAAGATTAGTCTGCAATGCTATCTACTACTGTGTC TACTATGATTACGACCACAGTGATATATCCAGCAACA AAAACCCA	
V_H81X-DSP2-F2	CTGAAACTCTCCTGTGAATCCAATGAATACGAATTCCC TTCCCATGACATGTCTTGGGTCGTTATGCAGGAACATG GTCTGGTGGGGCCAGAGGTAGCAGTGCAGAGTGTCAG GGGCCAT	

V_H81X-DSP2-F2 (R):	ATGGCCCCTGACACTCTGCACTGCTACCTCTGGCCCCACCAGACCATGTTCCTGCATAACGACCCAAGACATGTCATGGGAAGGGAATTCGTATTCATTGGATTCACAGGAGAGTTTCAG	
V_H81X-DSP2-R1	CTGAAACTCTCCTGTGAATCCAATGAATACGAATTCCCTTCCCATGACATGTCTTGGGTCATCTACTACTGTGTCTACTATGATTACGACCACAGTGATATATCCAGCAACAAAAACCCA	
V_H81X-DSP2-R1 (R)	TGGGTTTTTGTTGCTGGATATATCACTGTGGTCGTAATCATAGTAGACACAGTAGTAGATGACCCAAGACATGTCATGGGAAGGGAATTCGTATTCATTGGATTCACAGGAGAGTTTCAG	
V_H81X-DQ52-F2:	CTGAAACTCTCCTGTGAATCCAATGAATACGAATTCCCTTCCCATGACATGTCTTGGGTCTCCCCAATCTGCCAGTCATCTCTTGAGTCAGGGACCAATGATTTCTCTCAAAACAGTCGC	
V_H81X-DQ52-F 2 (R)	GCGACTGTTTTGAGAGAAATCATTGGTCCCTGACTCAAGAGATGACTGGCAGATTGGGGAGACCCAAGACATGTCATGGGAAGGGAATTCGTATTCATTGGATTCACAGGAGAGTTTCAG	
V_H81X-DQ52-R1	CTGAAACTCTCCTGTGAATCCAATGAATACGAATTCCCTTCCCATGACATGTCTTGGGTCCTGTGGACAGGTCTTAGATGGGGAAAGAATGAGCAAATGCAAGCATCAGAAGGGTGGATG	
V_H81X-DQ52-R 1 (R)	CATCCACCCTTCTGATGCTTGCATTTGCTCATTCTTTCCCCCATCTAAGACCTGTCCACAGGACCCAAGACATGTCATGGGAAGGGAATTCGTATTCATTGGATTCACAGGAGAGTTTCAG	
V_HQ52.2.4-DSP2-F2	AGCTCACACTAAGCTGAGAAGCTCCATCCTCTTCTCATAGAGCCTCCATCAGAGCATGGCGTTATGCAGGAACATGGTCTGGTGGGGCCAGAGGTAGCAGTGCAGAGTGTCAGGGGCCAT	
V_HQ52.2.4-DSP2-F2 (R)	ATGGCCCCTGACACTCTGCACTGCTACCTCTGGCCCCACCAGACCATGTTCCTGCATAACGCCATGCTCTGATGGAGGCTCTATGAGAAGAGGATGGAGCTTCTCAGCTTAGTGTGAGCT	
V_HQ52.2.4-DSP2-R1	AGCTCACACTAAGCTGAGAAGCTCCATCCTCTTCTCATAGAGCCTCCATCAGAGCATGGCATCTACTACTGTGTCTACTATGATTACGACCACAGTGATATATCCAGCAACAAAAACCCA	
V_HQ52.2.4-DSP2-R1 (R)	TGGGTTTTTGTTGCTGGATATATCACTGTGGTCGTAATCATAGTAGACACAGTAGTAGATGCCATGCTCTGATGGAGGCTCTATGAGAAGAGGATGGAGCTTCTCAGCTTAGTGTGAGCT	
V_HQ52.2.4-DQ52-F2	AGCTCACACTAAGCTGAGAAGCTCCATCCTCTTCTCATAGAGCCTCCATCAGAGCATGGCTCCCCAATCTGCCAG	

	TCATCTCTTGAGTCAGGGACCAATGATTTCTCTCAAAA CAGTCGC	
V_HQ52.2.4-DQ52-F2 (R)	GCGACTGTTTTGAGAGAAATCATTGGTCCCTGACTCA AGAGATGACTGGCAGATTGGGGAGCCATGCTCTGATG GAGGCTCTATGAGAAGAGGATGGAGCTTCTCAGCTTA GTGTGAGCT	
V_HQ52.2.4-DQ52-R1	AGCTCACACTAAGCTGAGAAGCTCCATCCTCTTCTCAT AGAGCCTCCATCAGAGCATGGCCTGTGGACAGGTCTT AGATGGGGAAAGAATGAGCAAATGCAAGCATCAGAA GGGTGGATG	
V_HQ52.2.4-DQ52-R1 (R)	CATCCACCCTTCTGATGCTTGCATTTGCTCATTCTTTCC CCATCTAAGACCTGTCCACAGGCCATGCTCTGATGGA GGCTCTATGAGAAGAGGATGGAGCTTCTCAGCTTAGT GTGAGCT	
V_H7183.4.6-DSP2-F2	GGAGGCCGGTCCTGGATTCGAGTTCCTCACATTCAGT GATGAGCACTGAACACGGACCCCGTTATGCAGGAACA TGGTCTGGTGGGGCCAGAGGTAGCAGTGCAGAGTGTC AGGGGCCAT	
V_H7183.4.6-DSP2-F2 (R)	ATGGCCCCTGACACTCTGCACTGCTACCTCTGGCCCCA CCAGACCATGTTCCTGCATAACGGGGTCCGTGTTCAGT GCTCATCACTGAATGTGAGGAACTCGAATCCAGGACC GGCCTCC	
V_H7183.4.6-DSP2-R1	GGAGGCCGGTCCTGGATTCGAGTTCCTCACATTCAGT GATGAGCACTGAACACGGACCCCATCTACTACTGTGT CTACTATGATTACGACCACAGTGATATATCCAGCAAC AAAAACCCA	
V_H7183.4.6-DSP2-R1 (R)	TGGGTTTTTGTTGCTGGATATATCACTGTGGTCGTAAT CATAGTAGACACAGTAGTAGATGGGGTCCGTGTTCAG TGCTCATCACTGAATGTGAGGAACTCGAATCCAGGAC CGGCCTCC	
V_H7183.4.6-DQ52-F2	GGAGGCCGGTCCTGGATTCGAGTTCCTCACATTCAGT GATGAGCACTGAACACGGACCCCTCCCCAATCTGCCA GTCATCTCTTGAGTCAGGGACCAATGATTTCTCTCAAA ACAGTCGC	
V_H7183.4.6-DQ52-F2 (R)	GCGACTGTTTTGAGAGAAATCATTGGTCCCTGACTCA AGAGATGACTGGCAGATTGGGGAGGGGTCCGTGTTCA GTGCTCATCACTGAATGTGAGGAACTCGAATCCAGGA CCGGCCTCC	
V_H7183.4.6-DQ52-R1	GGAGGCCGGTCCTGGATTCGAGTTCCTCACATTCAGT GATGAGCACTGAACACGGACCCCCTGTGGACAGGTCT TAGATGGGGAAAGAATGAGCAAATGCAAGCATCAGA AGGGTGGATG	
Deep sequencing for recombination		
$RAG2^{-/-}$ rep1 (DQ52 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGCAT TAACTAGTGTGAGGTTTAAGCCT	

RAG2$^{-/-}$ rep2 (DQ52 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTCG CGCCTAGTGTGAGGTTTAAGCCT	
CBE$^{-/-}$ (1) rep1 (DQ52 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACAC CTACTAGTGTGAGGTTTAAGCCT	
CBE$^{-/-}$ (1) rep2 (DQ52 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTAA CTGCTAGTGTGAGGTTTAAGCCT	
CBE$^{-/-}$ (2) rep1 (DQ52 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATCGC CAGCTAGTGTGAGGTTTAAGCCT	
CBE$^{-/-}$ (2) rep2 (DQ52 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATTC CAACTAGTGTGAGGTTTAAGCCT	
RAG2$^{-/-}$ rep1 (DSP2 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGAAC TGTCCCCTGACACTCTGCACTGCTA	
RAG2$^{-/-}$ rep2 (DSP2 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTCTG TCTCCCCTGACACTCTGCACTGCTA	
CBE$^{-/-}$ (1) rep1 (DQ52 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTAA GATCCCCTGACACTCTGCACTGCTA	
CBE$^{-/-}$ (1) rep2 (DSP2 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCAAG TAGCCCCTGACACTCTGCACTGCTA	
CBE$^{-/-}$ (2) rep1 (DSP2 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATATA GGACCCCTGACACTCTGCACTGCTA	
CBE$^{-/-}$ (2) rep2 (DSP2 bait)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCACGT GTTCCCCTGACACTCTGCACTGCTA	
Bio-DQ52	/5BiosG/GATCAGAATACCCATACTCT	
Bio-DSP2	/5BiosG/TGGGGAGATAGAATCCCAGGAG	