Science Immunology
AAAS

Supplementary Materials for

RNA exosome drives early B cell development via noncoding RNA processing mechanisms

Brice Laffleur *et al.*

Corresponding author: Brice Laffleur, laffleur.brice@gmail.com; Uttiya Basu, ub2121@cumc.columbia.edu

Sci. Immunol. **7**, eabn2738 (2022)
DOI: 10.1126/sciimmunol.abn2738

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

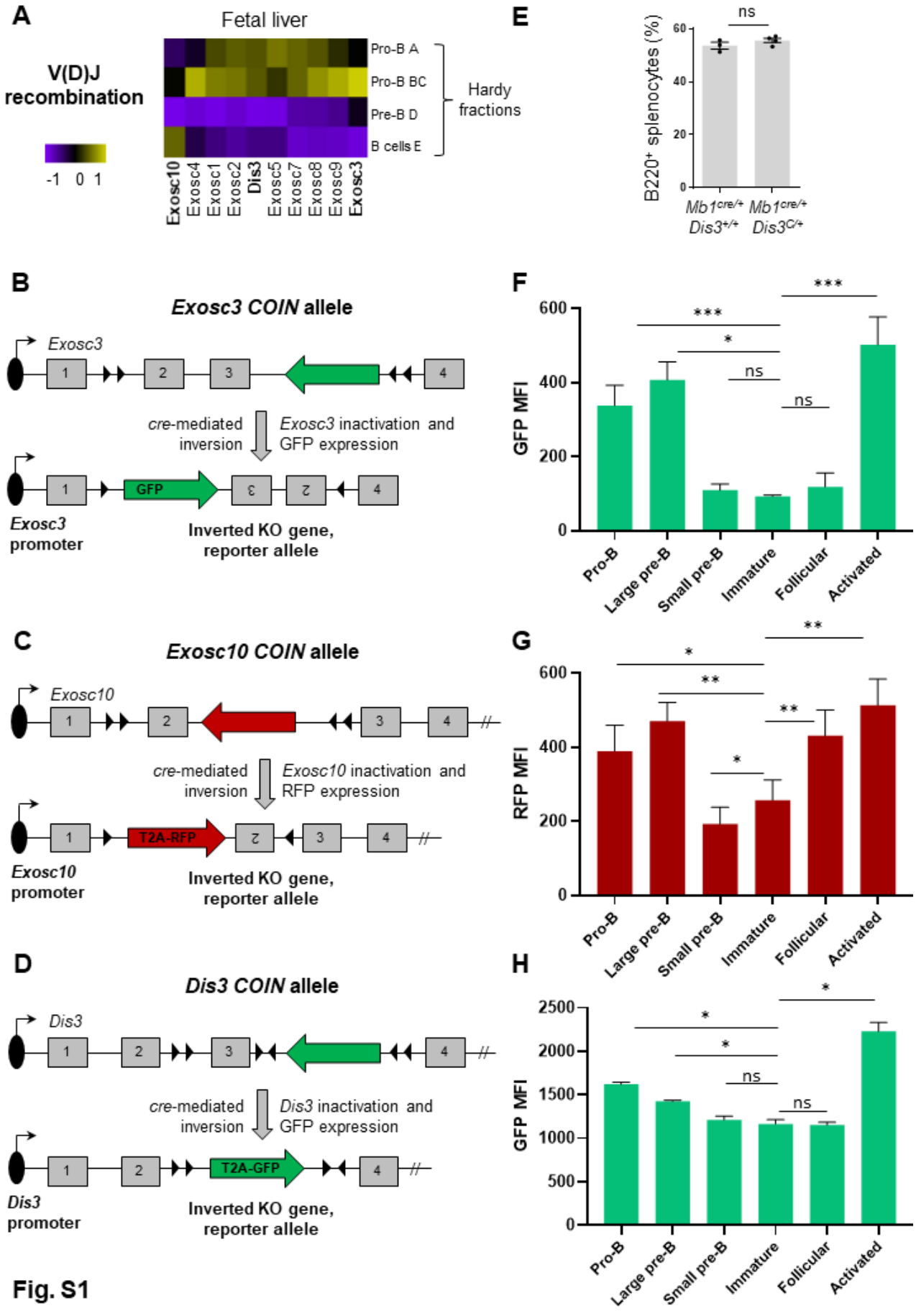
A
Fetal liver
V(D)J recombination
-1 0 1
Pro-B A
Pro-B BC
Pre-B D
B cells E
Hardy fractions
Exosc10 Exosc4 Exosc1 Exosc2 Dis3 Exosc5 Exosc7 Exosc8 Exosc9 Exosc3
E
ns
B220⁺ splenocytes (%)
60
40
20
0
Mb1cre/+ Dis3+/+
Mb1cre/+ Dis3C/+
B
Exosc3 COIN allele
Exosc3
1 2 3 4
cre-mediated inversion
Exosc3 inactivation and GFP expression
1 GFP 3 2 4
Exosc3 promoter
Inverted KO gene, reporter allele
F
GFP MFI
600
400
200
0
***
***
*
ns
ns
Pro-B
Large pre-B
Small pre-B
Immature
Follicular
Activated
C
Exosc10 COIN allele
Exosc10
1 2 3 4 //
cre-mediated inversion
Exosc10 inactivation and RFP expression
1 T2A-RFP 2 3 4 //
Exosc10 promoter
Inverted KO gene, reporter allele
G
RFP MFI
600
400
200
0
**
*
**
*
**
Pro-B
Large pre-B
Small pre-B
Immature
Follicular
Activated
D
Dis3 COIN allele
Dis3
1 2 3 4 //
cre-mediated inversion
Dis3 inactivation and GFP expression
1 2 T2A-GFP 4 //
Dis3 promoter
Inverted KO gene, reporter allele
H
GFP MFI
2500
2000
1500
1000
500
0
*
*
*
ns
ns
Pro-B
Large pre-B
Small pre-B
Immature
Follicular
Activated

Fig. S1

Fig. S1. Dynamic RNA exosome subunits expression during B cell development. (**A**) Heat map showing RNA exosome subunits expression in mouse fetal liver B cells, from “ImmGen” databank. Relative transcriptomic expression of the RNA exosome subunits is shown in the prenatal B cell populations. (**B**) Schematic of the *Exosc3 COIN* allele. The endogenous *Exosc3* exons are shown (grey boxes) and the inverted GFP cassette is represented by the green arrow. After *cre*-mediated allelic inversion using the lox sites (black triangles) the *Exosc3* exons 2 and 3 are inverted and the GFP cassette is able to be expressed. (**C**) Schematic of the *Exosc10 COIN* allele. The endogenous *Exosc10* exons are shown (grey boxes) and the inverted RFP cassette is represented by the red arrow. After *cre*-mediated allelic inversion using the lox sites (black triangles) the *Exosc10* exon 2 is inverted and the GFP cassette is able to be expressed. (**D**) Schematic of the *Dis3 COIN* allele. The endogenous *Dis3* exons are shown (grey boxes) and the inverted GFP cassette is represented by the green arrow. After *cre*-mediated allelic inversion using the lox sites (black triangles) the *Dis3* exon 3 is deleted and the GFP cassette is able to be expressed. Note: the *Exosc3*, *Exosc10* and *Dis3* promoters drive the transcription of the endogenous corresponding alleles or of GFP/RFP after inversion, considering this allele as a reporter system. (**E**) Flow cytometry analyses of splenocytes from $Mb1^{cre/+}$ $Dis3^{+/+}$and $Mb1^{cre/+}$ $Dis3^{C/+}$ mice. B splenocytes ($B220^{+}$) were quantified (bar graphs show the mean of 3 independent experiments +/- SEM, two-tailed unpaired *t*-test). GFP or RFP mean fluorescent intensity (MFI) of GFP^{+} or RFP^{+} cells were determined from pro-B cells, large pre-B cells, small pre-B cells, immature B cells, follicular B cells and activated B cells from $Mb1^{cre/+}$ $Dis3^{C/+}$ mice (**F**), $Mb1^{cre/+}$ $Exosc10^{C/+}$ mice (**G**), and $Mb1^{cre/+}$ $Exosc3^{C/+}$ mice (**H**), using at least 3 mice from 3 independent experiments (two-tailed paired *t*-tests).

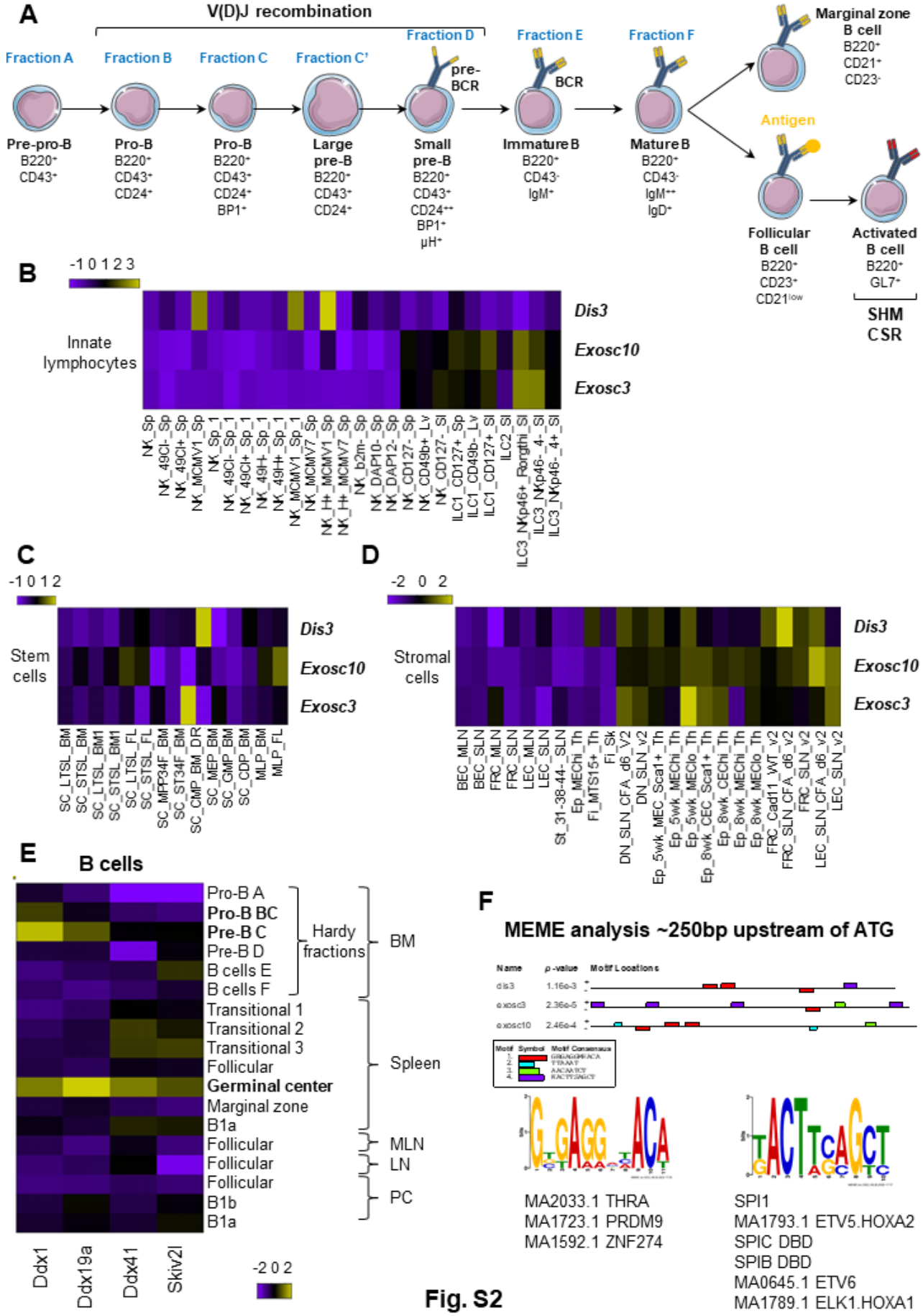
A
V(D)J recombination
Fraction A
Fraction B
Fraction C
Fraction C'
Fraction D
Fraction E
Fraction F
pre-BCR
BCR
Pre-pro-B B220+ CD43+
Pro-B B220+ CD43+ CD24+
Pro-B B220+ CD43+ CD24+ BP1+
Large pre-B B220+ CD43+ CD24+
Small pre-B B220+ CD43+ CD24++ BP1+ μH+
Immature B B220+ CD43- IgM+
Mature B B220+ CD43- IgM++ IgD+
Marginal zone B cell B220+ CD21+ CD23-
Antigen
Follicular B cell B220+ CD23+ CD21low
Activated B cell B220+ GL7+
SHM
CSR
B
-1 0 1 2 3
Innate lymphocytes
Dis3
Exosc10
Exosc3
NK_Sp
NK_49Cl-_Sp
NK_49Cl+_Sp
NK_MCMV1_Sp
NK_Sp_1
NK_49Cl-_Sp_1
NK_49Cl+_Sp_1
NK_49I+_Sp_1
NK_49I+_Sp_1
NK_MCMV1_Sp_1
NK_MCMV7_Sp
NK_I+_MCMV1_Sp
NK_I+_MCMV7_Sp
NK_b2m-_Sp
NK_DAP10-_Sp
NK_DAP12-_Sp
NK_CD127-_Sp
NK_CD49b+_Lv
NK_CD127-_Sl
ILC1_CD127+_Sp
ILC1_CD49b-_Lv
ILC1_CD127+_Sl
ILC2_Sl
ILC3_NKp46+_Rorgfhi_Sl
ILC3_NKp46-_4-_Sl
ILC3_NKp46-_4+_Sl
C
-1 0 1 2
Stem cells
Dis3
Exosc10
Exosc3
SC_LTSL_BM
SC_STSL_BM
SC_LTSL_BM1
SC_STSL_BM1
SC_LTSL_FL
SC_STSL_FL
SC_MFP34F_BM
SC_ST34F_BM
SC_CMP_BM_DR
SC_MEP_BM
SC_GMP_BM
SC_CDP_BM
MLP_BM
MLP_FL
D
-2 0 2
Stromal cells
Dis3
Exosc10
Exosc3
BEC_MLN
BEC_SLN
FRC_MLN
FRC_SLN
LEC_MLN
LEC_SLN
St_31-38-44-_SLN
Ep_MEChi_Th
Fi_MTS15+_Th
Fi_Sk
DN_SLN_CFA_d6_V2
DN_SLN_v2
Ep_5wk_MEC_Sca1+_Th
Ep_5wk_MEChi_Th
Ep_5wk_MEClo_Th
Ep_8wk_CEC_Sca1+_Th
Ep_8wk_CEChi_Th
Ep_8wk_MEChi_Th
Ep_8wk_MEClo_Th
FRC_Cad11_WT_v2
FRC_SLN_CFA_d6_v2
FRC_SLN_v2
LEC_SLN_CFA_d6_v2
LEC_SLN_v2
E
B cells
Pro-B A
Pro-B BC
Pre-B C
Pre-B D
B cells E
B cells F
Hardy fractions
BM
Transitional 1
Transitional 2
Transitional 3
Follicular
Germinal center
Marginal zone
B1a
Spleen
Follicular
MLN
Follicular
LN
Follicular
PC
B1b
B1a
Ddx1
Ddx19a
Ddx41
Skiv2l
-2 0 2
F
MEME analysis ~250bp upstream of ATG
Name
p-value
Motif Locations
dis3
1.16e-3
exosc3
2.36e-5
exosc10
2.46e-4
Motif
Symbol
Motif Consensus
MA2033.1 THRA
MA1723.1 PRDM9
MA1592.1 ZNF274
SPI1
MA1793.1 ETV5.HOXA2
SPIC DBD
SPIB DBD
MA0645.1 ETV6
MA1789.1 ELK1.HOXA1

Fig. S2

Fig. S2. RNA exosome subunits expression and regulation. (**A**) Schematic of B cell development. The Hardy's fractions corresponding to developing B cells are indicated in blue (fractions A to F), the main markers to identify these cells are indicated in black. Marginal zone B cells, follicular B cells and activated B cells are also shown. (**B-D**) Relative transcriptomic expressions of RNA exosome subunits *Exosc3*, *Exosc10* and *Dis3* from ImmGen databank in innate lymphocytes (**B**), stem cells (**C**), and stromal cells (**D**) are shown. (**E**) RNA helicases (*Ddx1*, *Ddx19a*, *Ddx41*, *Skiv2l*) expression during B cell development. (**F**) MEME analyses of *Exosc3*, *Exosc10* and *Dis3* promoters.

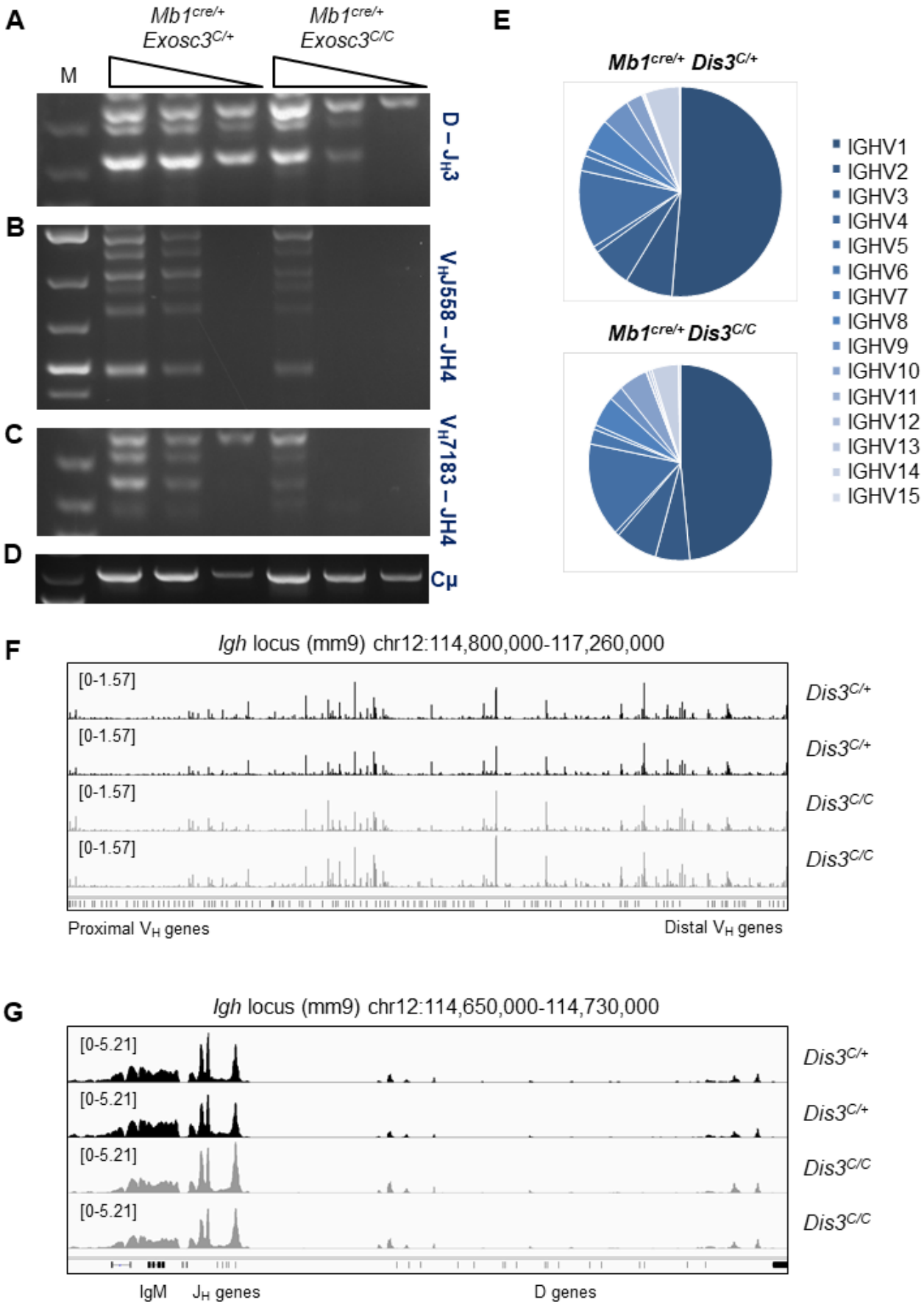
A
Mb1cre/+ Exosc3C/+
Mb1cre/+ Exosc3C/C
M
D – JH3
B
VHJ558 – JH4
C
VH7183 – JH4
D
Cμ
E
Mb1cre/+ Dis3C/+
Mb1cre/+ Dis3C/C
IGHV1
IGHV2
IGHV3
IGHV4
IGHV5
IGHV6
IGHV7
IGHV8
IGHV9
IGHV10
IGHV11
IGHV12
IGHV13
IGHV14
IGHV15
F
Igh locus (mm9) chr12:114,800,000-117,260,000
[0-1.57]
Dis3C/+
Dis3C/C
Proximal VH genes
Distal VH genes
G
Igh locus (mm9) chr12:114,650,000-114,730,000
[0-5.21]
IgM
JH genes
D genes

Fig. S3

Fig. S3. RNA exosome deletion impairs V(D)J recombination. (**A**) D-J_H3 PCR analysis. Genomic DNA was amplified using forward D gene oligonucleotides and J_H3 reverse. PCR were performed on 100 ng, 10 ng and 1 ng of the same DNA samples. (**B**) V_HJ558-J_H4 PCR analysis. Genomic DNA was amplified using forward V_HJ558 family gene oligonucleotides and J_H4 reverse. PCR were performed on 100 ng, 10 ng and 1 ng of DNA. (**C**) V_H7183-J_H4 PCR analysis. Genomic DNA was amplified using forward V_H7183 family gene oligonucleotides and J_H4 reverse. PCR were performed on 100 ng, 10 ng and 1 ng of the same DNA samples. (**D**) A control PCR using constant IgM gene (Cμ) was used as loading control. PCR were performed on 100 ng, 10 ng and 1 ng of the same DNA samples. These PCR were performed on *Mb1*$^{cre/+}$ *Exosc3*$^{C/+}$ and *Mb1*$^{cre/+}$ *Exosc3*$^{C/C}$ genomic DNA. The gel electrophoresis pictures are representative from at least 3 independent experiments, using a total of at least 3 mice per group. (**E**) V_H family usage in *Mb1*$^{cre/+}$ *Dis3*$^{C/+}$ and *Mb1*$^{cre/+}$ *Dis3*$^{C/C}$ mice. Combined data from 3 independent experiments are shown. (**F**) ATAC-seq experiments were performed on sorted pro-B cells from *Mb1*$^{cre/+}$ *Dis3*$^{C/+}$ and *Mb1*$^{cre/+}$ *Dis3*$^{C/C}$ mice, accessibility of the V_H genes is shown (2 independent experiments). (**G**) Accessibility of the J_H and D genes is shown (2 independent experiments).

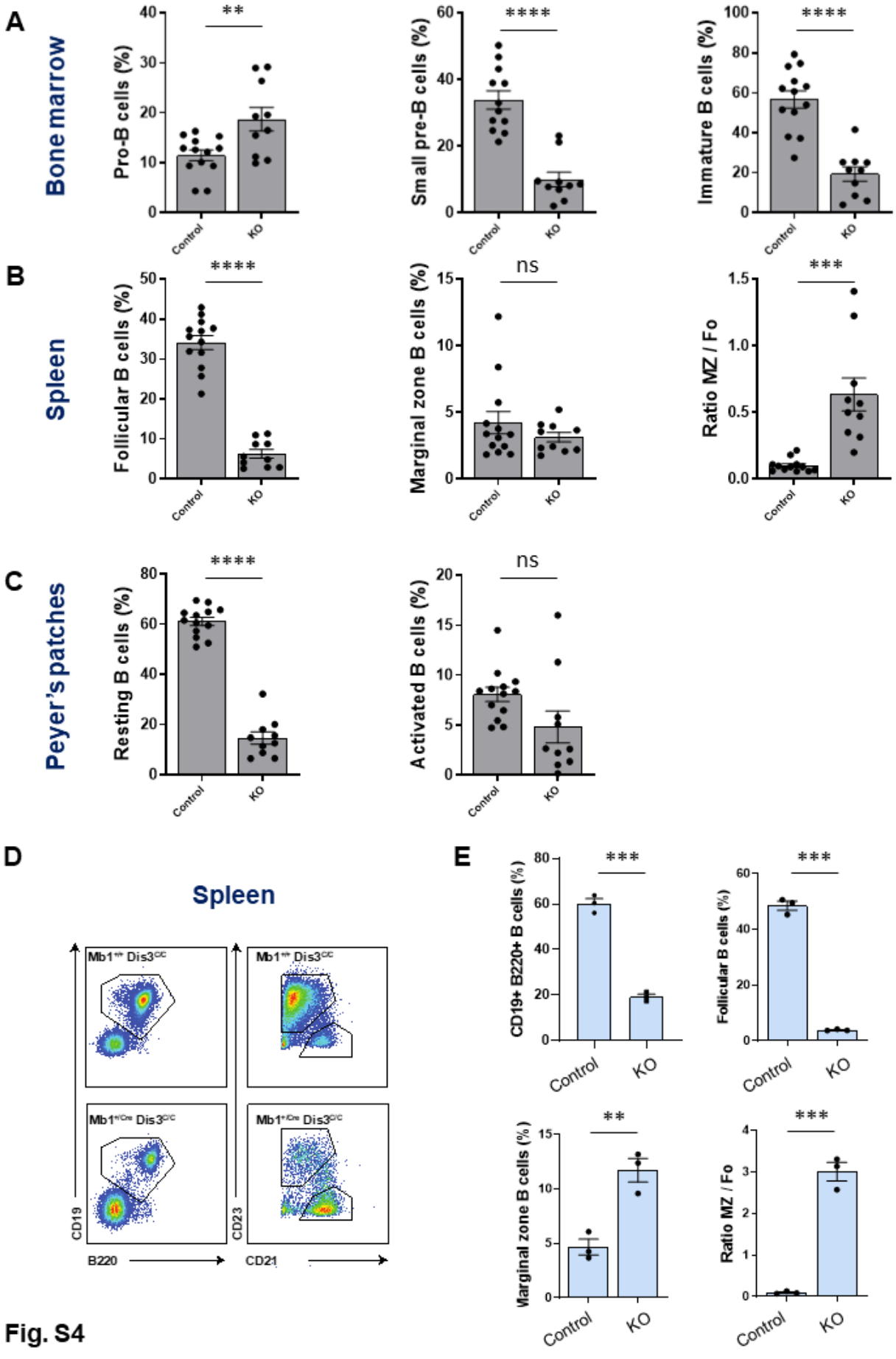
A
Bone marrow
Pro-B cells (%)
**
Small pre-B cells (%)
****
Immature B cells (%)
****
B
Spleen
Follicular B cells (%)
****
Marginal zone B cells (%)
ns
Ratio MZ / Fo
***
C
Peyer's patches
Resting B cells (%)
****
Activated B cells (%)
ns
Control
KO
D
Spleen
Mb1+/+ Dis3C/C
Mb1+/Cre Dis3C/C
CD19
B220
CD23
CD21
E
CD19+ B220+ B cells (%)
***
Follicular B cells (%)
***
Marginal zone B cells (%)
**
Ratio MZ / Fo
***

Fig. S4

Fig. S4. B cell development in RNA exosome-deficient mice. We combined the data from *Mb1*$^{cre/+}$ RNA exosome-proficient mice *(C/+,* referred as control) and *Mb1*$^{cre/+}$ RNA exosome-deficient mice (*C/C,* referred as KO) in (**A**) bone marrow pro-B cells, small pre-B cells and immature B cells. (**B**) Spleen follicular (Fo) B cells, marginal zone (MZ) B cells and MZ/Fo ratio. (**C**) Peyer's patches resting and activated B cells. (**D**) Flow cytometry analyses of B splenocytes in from *Mb1*$^{+/+}$ *Dis3*$^{C/C}$ and *Mb1*$^{cre/+}$ *Dis3*$^{C/C}$ mice. Cells were stained with B cell specific markers CD19 and B220, and double positive cells were stained with CD21 and CD23 antibodies to estimate the proportion of follicular and marginal zone B cells (3 independent experiments). (**E**) Quantification of follicular and marginal zone B cells from total splenocytes in *Mb1*$^{+/+}$ *Dis3*$^{C/C}$ (control) and *Mb1*$^{cre/+}$ *Dis3*$^{C/C}$ (KO) mice (3 independent experiments, two-tailed unpaired *t*-tests).

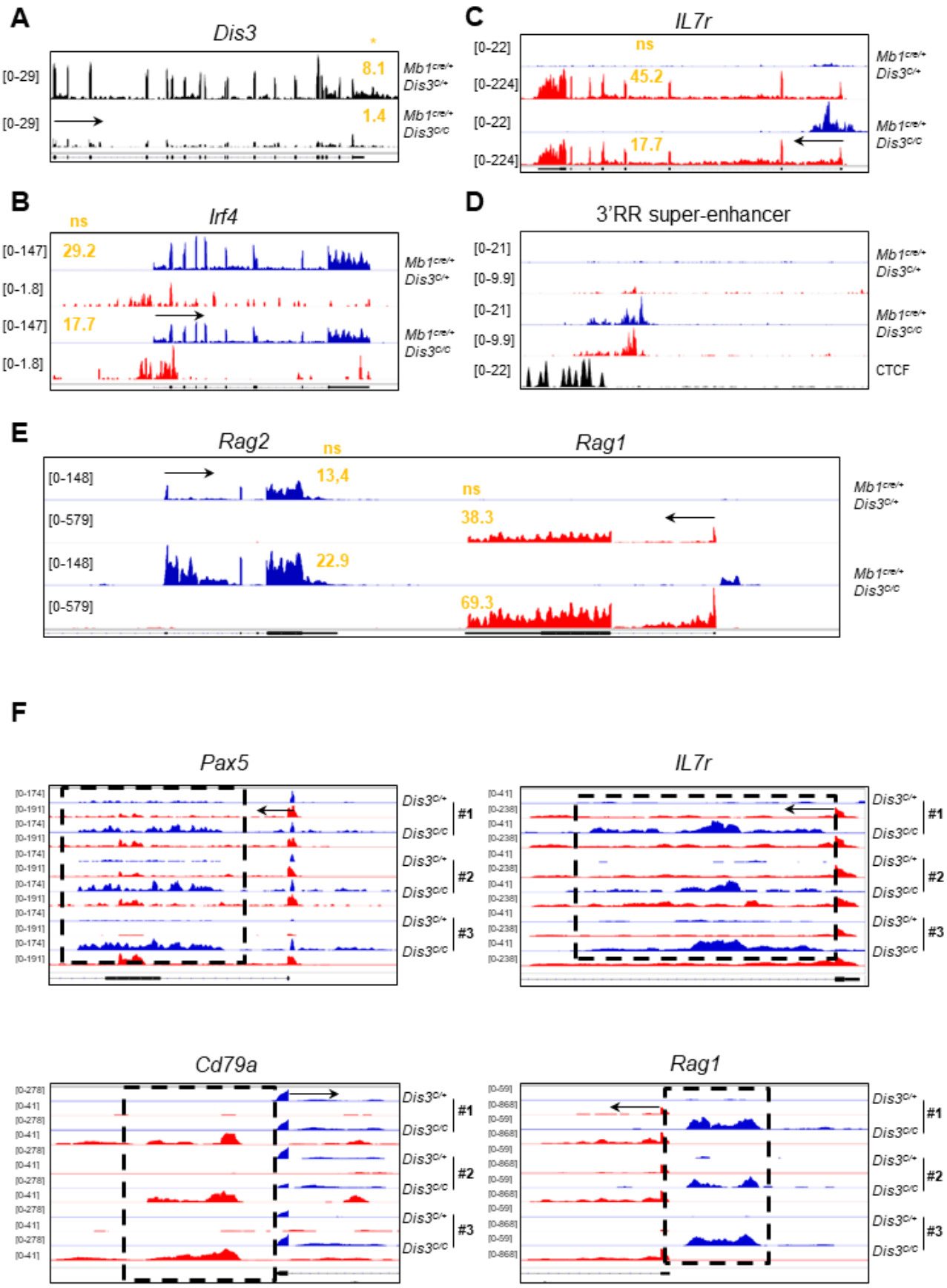
A
Dis3
[0-29]
[0-29]
8.1
1.4
Mb1cre/+ Dis3O/+
Mb1cre/+ Dis3O/O
B
Irf4
ns
[0-147]
[0-1.8]
[0-147]
[0-1.8]
29.2
17.7
C
IL7r
[0-22]
[0-224]
[0-22]
[0-224]
ns
45.2
17.7
D
3'RR super-enhancer
[0-21]
[0-9.9]
[0-21]
[0-9.9]
[0-22]
CTCF
E
Rag2
Rag1
[0-148]
[0-579]
[0-148]
[0-579]
13,4
ns
38.3
22.9
69.3
F
Pax5
IL7r
Cd79a
Rag1
Dis3O/+
Dis3O/O
#1
#2
#3

Fig. S5

Fig. S5. RNA-sequencing reveals RNA exosome substrates from pro-B cells *in vivo*. (**A**) RNA-sequencing IGV tracks showing *Dis3* deletion in $Mb1^{cre/+}$ $Dis3^{C/C}$ pro-B cells ($B220^{+}$, $CD43^{+}$, $CD25^{-}$, IgM^{-}), compare to the control $Mb1^{cre/+}$ $Dis3^{C/+}$ pro-B cells. Note the *Dis3* gene is shown from exon 3 to the last exon, downstream of the inversion sites (see **fig. S1D** for details). Mean RPKM of *Dis3* (excluding exons 1 and 2) is indicated in orange, black arrows indicate the direction of coding transcription, significance was determined from **Table S1**. (**B**) Example of aTSS RNAs and antisense RNAs accumulation at the *Irf4* gene, while mRNA expression is not affected in the absence of DIS3. Mean RPKM of *Irf4* is indicated in orange. (**C**) Example of aTSS RNAs and antisense RNAs accumulation at the *IL7r* gene, while mRNA expression is not affected in the absence of DIS3. Mean RPKM of *IL7r* is indicated in orange. (**D**) Example of eRNAs accumulation in the absence of DIS3, at the 3'RR super-enhancer from the *Igh* locus. (**E**) *Rag1* and *Rag2* gene expression is not affected by the absence of DIS3. Means RPKM of *Rag1* and *Rag2* are indicated in orange. (**F**) Example of ncRNAs reproducibly accumulating at different loci in the absence of *Dis3*, from 3 independent RNA-sequencing experiments.

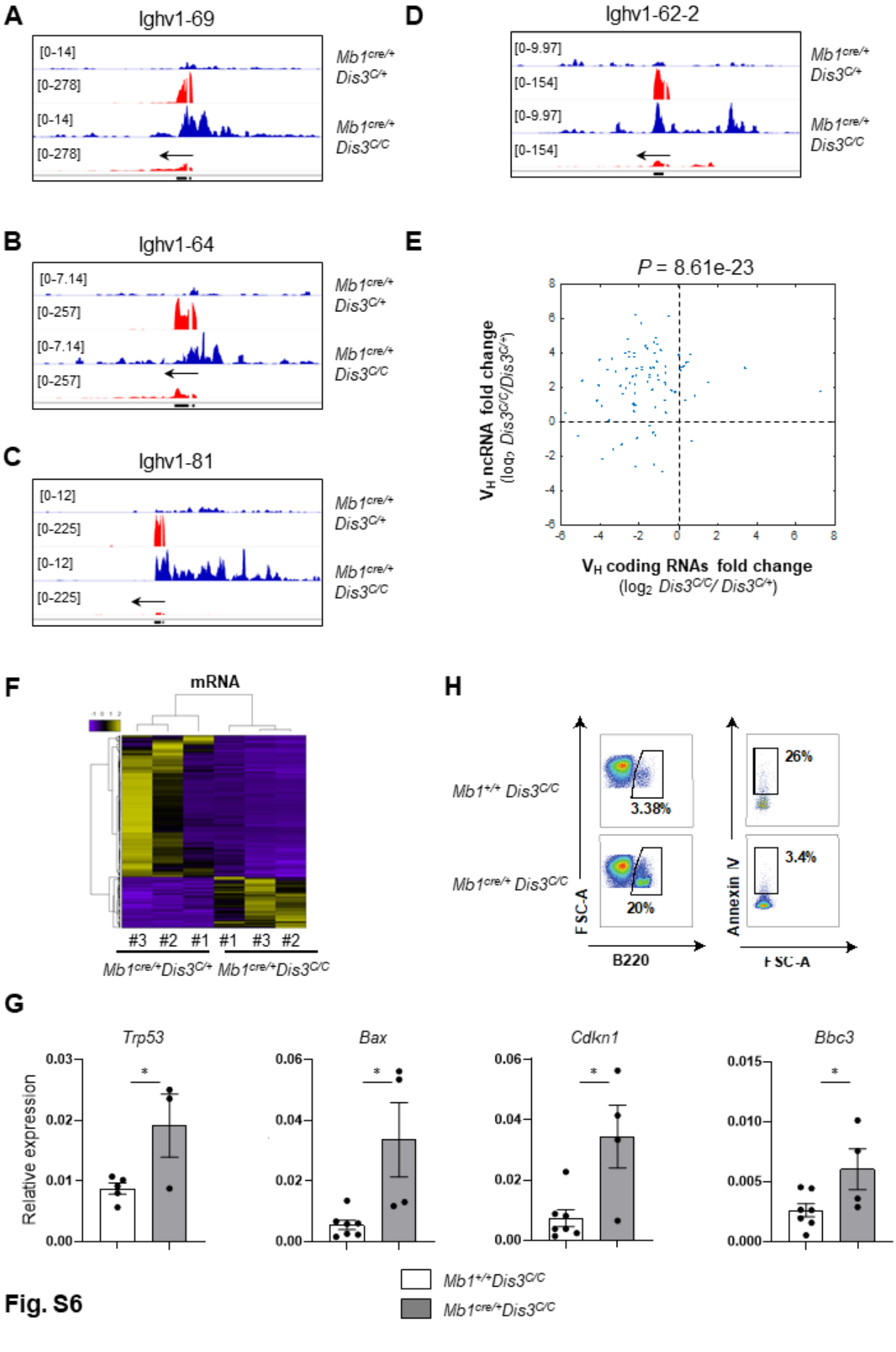
A
Ighv1-69
[0-14]
[0-278]
[0-14]
[0-278]
Mb1cre/+ Dis3C/+
Mb1cre/+ Dis3C/C
B
Ighv1-64
[0-7.14]
[0-257]
[0-7.14]
[0-257]
Mb1cre/+ Dis3C/+
Mb1cre/+ Dis3C/C
C
Ighv1-81
[0-12]
[0-225]
[0-12]
[0-225]
Mb1cre/+ Dis3C/+
Mb1cre/+ Dis3C/C
D
Ighv1-62-2
[0-9.97]
[0-154]
[0-9.97]
[0-154]
Mb1cre/+ Dis3C/+
Mb1cre/+ Dis3C/C
E
P = 8.61e-23
VH ncRNA fold change
(log2 Dis3C/C/Dis3C/+)
VH coding RNAs fold change
(log2 Dis3C/C/ Dis3C/+)
F
mRNA
#3 #2 #1 #1 #3 #2
Mb1cre/+Dis3C/+ Mb1cre/+Dis3C/C
H
Mb1+/+ Dis3C/C
Mb1cre/+ Dis3C/C
3.38%
20%
26%
3.4%
FSC-A
B220
Annexin V
FSC-A
G
Trp53
Bax
Cdkn1
Bbc3
Relative expression
Mb1+/+Dis3C/C
Mb1cre/+Dis3C/C

Fig. S6

Fig. S6. RNA exosome-deficient pro-B cells accumulate ncRNAs overlapping V_H genes. (**A**) Example of non-coding RNA accumulation at V_H (Ighv1-69) gene in the absence of RNA exosome activity, while coding gene expression is decreased. (**B**) Example of non-coding RNA accumulation at V_H (Ighv1-64) gene in the absence of RNA exosome activity, while coding gene expression is decreased. (**C**) Example of non-coding RNA accumulation at V_H (Ighv1-81) gene in the absence of RNA exosome activity, while coding gene expression is decreased. (**D**) Example of non-coding RNA accumulation at V_H (Ighv1-62-2) gene in the absence of RNA exosome activity, while coding gene expression is decreased. (**E**) Inverse correlation between non-coding RNA accumulation and V_H coding expression in the absence of DIS3. We calculated coding and non-coding RNA level from *Mb1*$^{cre/+}$ *Dis3*$^{C/+}$ and *Mb1*$^{cre/+}$ *Dis3*$^{C/C}$ pro-B cells RNA-sequencing data and show the result as $\log_2$ fold change (one-sample proportion test). (**F**) Heat map showing differentially expressed genes (see also **Table S1**). (**G**) RT-qPCR evaluation of p53 pathway-related genes in pro-B cells, using at least 3 mice from 3 independent experiments, mean is shown +/- SEM, two-sided unpaired *t*-tests. (**H**) Flow cytometry analysis of apoptosis evaluated by Annexin V, one experiment.

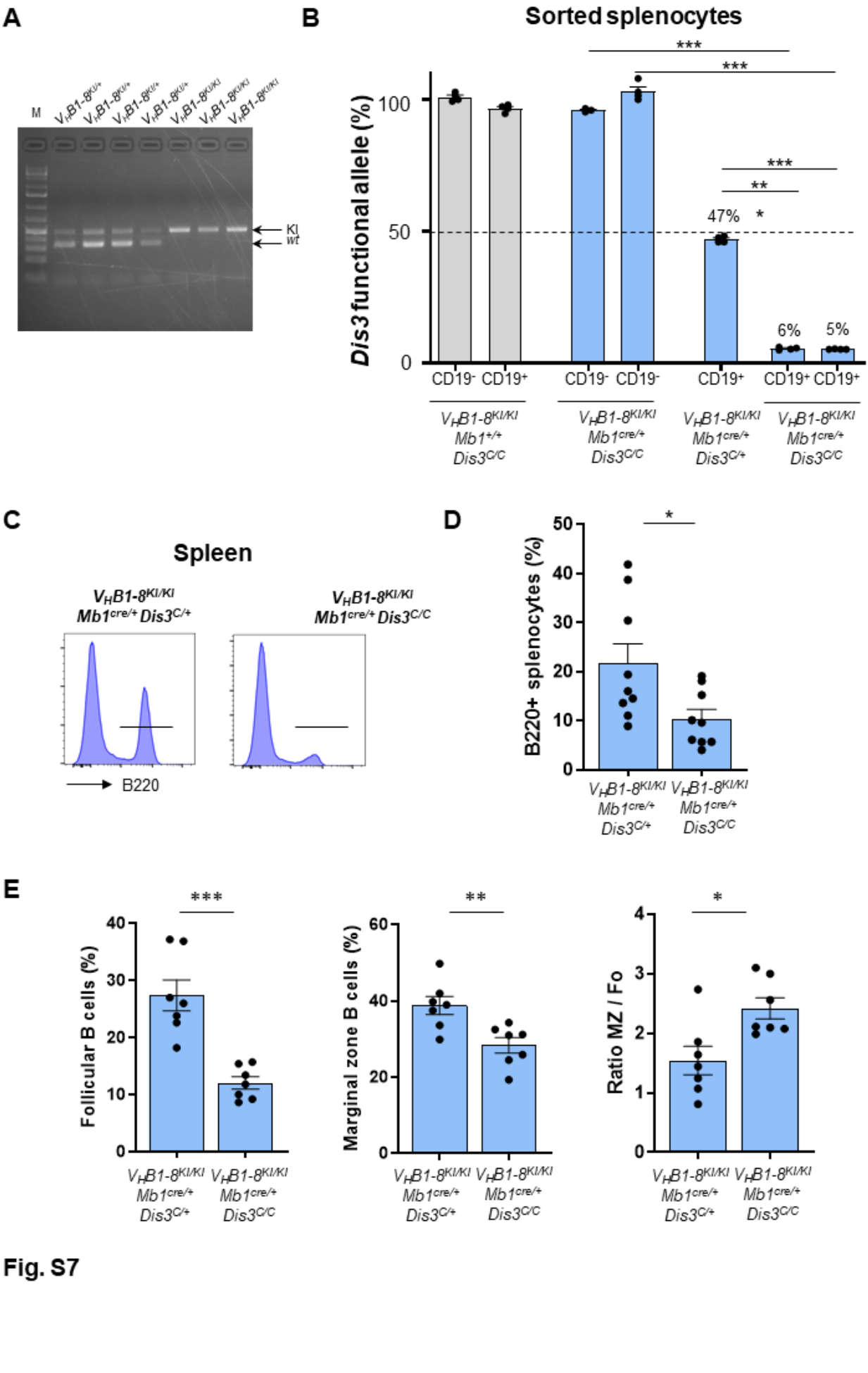
A
M
KI
wt
B
Sorted splenocytes
Dis3 functional allele (%)
100
50
0
***
***
***
**
*
47%
6%
5%
CD19- CD19+
CD19- CD19-
CD19+
CD19+ CD19+
VHB1-8KI/KI Mb1+/+ Dis3C/C
VHB1-8KI/KI Mb1cre/+ Dis3C/C
VHB1-8KI/KI Mb1cre/+ Dis3C/+
VHB1-8KI/KI Mb1cre/+ Dis3C/C
C
Spleen
VHB1-8KI/KI Mb1cre/+ Dis3C/+
VHB1-8KI/KI Mb1cre/+ Dis3C/C
B220
D
B220+ splenocytes (%)
50
40
30
20
10
0
*
VHB1-8KI/KI Mb1cre/+ Dis3C/+
VHB1-8KI/KI Mb1cre/+ Dis3C/C
E
Follicular B cells (%)
40
30
20
10
0
***
Marginal zone B cells (%)
60
40
20
0
**
Ratio MZ / Fo
4
3
2
1
0
*
VHB1-8KI/KI Mb1cre/+ Dis3C/+
VHB1-8KI/KI Mb1cre/+ Dis3C/C

Fig. S7

Fig. S7. Alteration of peripheral B cells distribution in the absence of DIS3. (**A**) PCR for genotyping *V_HB1-8* mice. The KI allele (*V_HB1-8^{KI}*) band size is 500 base pairs (bp) and the *wt* allele (*V_HB1-8^{+}*) band size is 337 bp. M = DNA marker (1kb plus). (**B**) DNAs were collected from the different mice as indicated, including various controls. Cells were sorted as $CD19^+$ and $CD19^-$, and *Dis3 COIN* inversion was evaluated by qPCR. Each bar graph shows the mean +/- SEM from one experiment done in triplicate. Two independent experiments were performed with similar results for the *Dis3 COIN* deletion in *$Mb1^{cre/+}$ $Dis3^{C/C}$* $CD19^+$ B cells. (**C**) Flow cytometry analysis of $B220^+$ splenocytes from *V_HB1-$8^{KI/KI}$ $Mb1^{cre/+}$ $Dis3^{C/+}$* and *V_HB1-$8^{KI/KI}$ $Mb1^{cre/+}$ $Dis3^{C/C}$* mice. (**D**) Quantification of $B220^+$ splenocytes from *V_HB1-$8^{KI/KI}$ $Mb1^{cre/+}$ $Dis3^{C/+}$* (n=9), and *V_HB1-$8^{KI/KI}$ $Mb1^{cre/+}$ $Dis3^{C/C}$* (n=9) mice. 6 independent experiments, mean is shown +/- SEM, two-tailed unpaired *t*-test. (**E**) Follicular and marginal zone (MZ) B cells frequencies were determined from these mice, along with the subsequent ratio MZ/Fo (n=7). 6 independent experiments, mean is shown +/- SEM, two-tailed unpaired *t*-test.

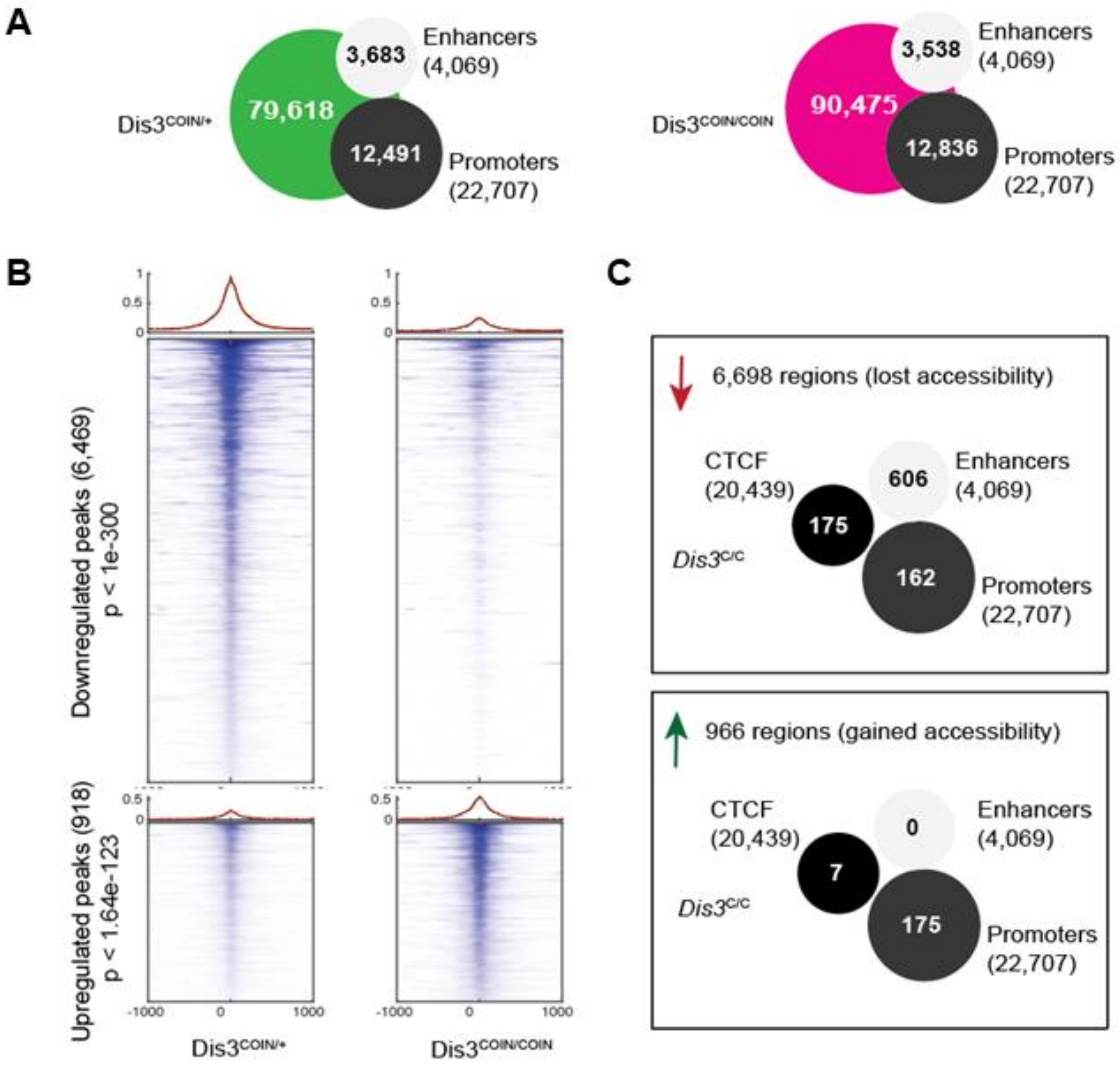
A
Dis3COIN/+
79,618
3,683
Enhancers
(4,069)
12,491
Promoters
(22,707)
Dis3COIN/COIN
90,475
3,538
Enhancers
(4,069)
12,836
Promoters
(22,707)
B
Downregulated peaks (6,469)
p < 1e-300
Upregulated peaks (918)
p < 1.64e-123
-1000
0
1000
Dis3COIN/+
Dis3COIN/COIN
C
6,698 regions (lost accessibility)
CTCF
(20,439)
Dis3C/C
175
606
Enhancers
(4,069)
162
Promoters
(22,707)
966 regions (gained accessibility)
CTCF
(20,439)
Dis3C/C
7
0
Enhancers
(4,069)
175
Promoters
(22,707)
D
Mb1cre/+ Dis3C/+
Mb1cre/+ Dis3C/C
IGKV1
IGKV2
IGKV3
IGKV4
IGKV5
IGKV6
IGKV8
IGKV9
IGKV10
IGKV11
IGKV12
IGKV13
IGKV14
IGKV15
IGKV16
IGKV17
IGKV18
IGKV19
E
Mb1cre/+ Dis3C/+
Mb1cre/+ Dis3C/C
IGLV1
IGLV2
IGLV3

Fig. S8

Fig. S8. Genome-wide changes in chromatin accessibility in the absence of RNA processing. (**A**) ATAC-seq showing the total number of peaks in control cells (79,618) and DIS3-deficient pro-B cells (90,475). The numbers of peaks at enhancers and promoters are shown. (**B**) ATAC-seq peaks at sites were accessibility was decreased (top) or increased (bottom). (**C**) ATAC-seq showing the distribution of peaks with lost accessibility (top) or gained accessibility (bottom). (**D**) Vκ family usage in *Mb1*$^{cre/+}$ *Dis3*$^{C/+}$ and *Mb1*$^{cre/+}$ *Dis3*$^{C/C}$ mice. Combined data from 3 independent experiments are shown. (**E**) Vλ family usage in *Mb1*$^{cre/+}$ *Dis3*$^{C/+}$ and *Mb1*$^{cre/+}$ *Dis3*$^{C/C}$ mice. Combined data from 3 independent experiments are shown.

Table S1. Differentially expressed coding and non-coding RNAs in DIS3-proficient and DIS3-deficient pro-B cells. RNA-sequencing data were analyzed from *Mb1*$^{cre/+}$ *Dis3*$^{C/C}$ pro-B cells (n=3), compare to the control *Mb1*$^{cre/+}$ *Dis3*$^{C/+}$ pro-B cells (n=3) for coding (mRNAs) and non-coding RNAs (lncRNAs, eRNAs, and aTSS-RNAs).

Table S2. Differential ATAC-seq peaks in DIS3-proficient and DIS3-deficient pro-B cells. ATAC-sequencing data were analyzed from *Mb1*$^{cre/+}$ *Dis3*$^{C/C}$ pro-B cells (n=2), compare to the control *Mb1*$^{cre/+}$ *Dis3*$^{C/+}$ pro-B cells (n=2) for up-regulated and down-regulated peaks.

Table S3. Raw data and statistical analyses.