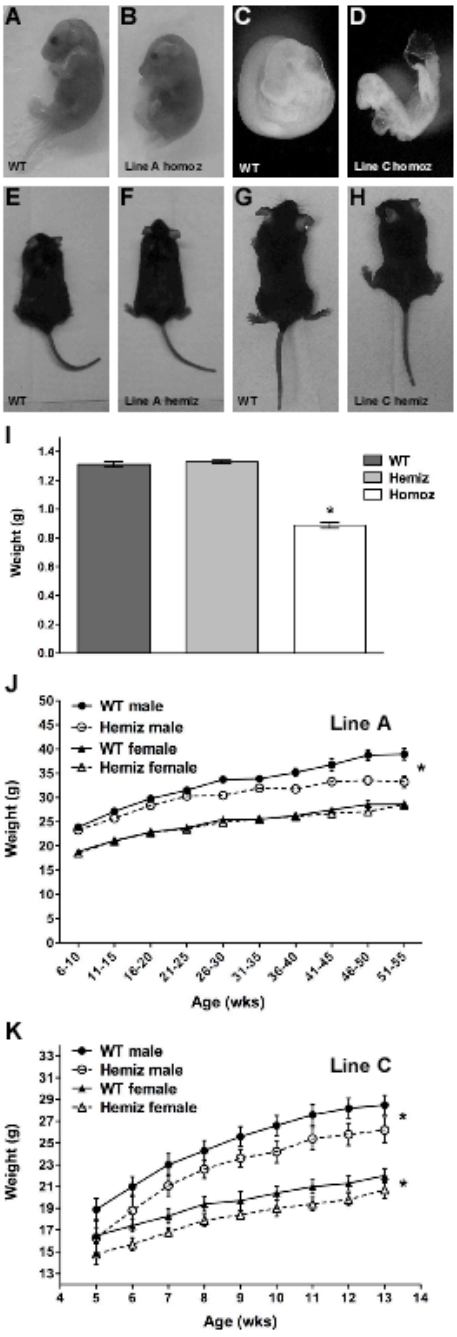
Supplementary Information

DNMT3B7, a truncated DNMT3B isoform expressed in human tumors, disrupts embryonic development and accelerates lymphomagenesis

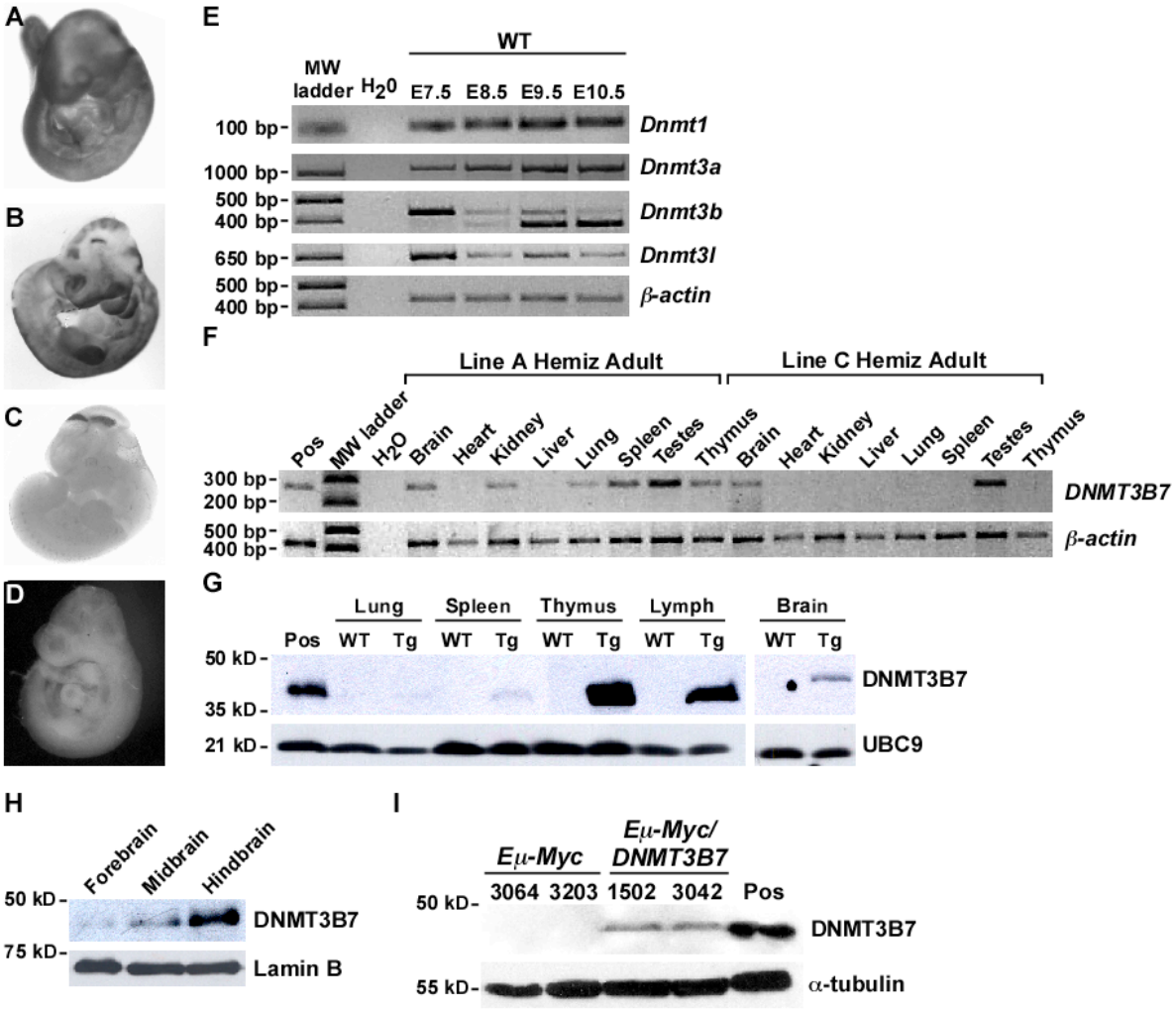
Mrinal Y. Shah, Aparna Vasanthakumar, Natalie Y. Barnes, Maria E. Figueroa, Anna Kamp, Christopher Hendrick, Kelly R. Ostler, Elizabeth M. Davis, Shang Lin, John Anastasi, Michelle M. Le Beau, Ivan Moskowitz, Ari Melnick, Peter Pytel, and Lucy A. Godley

Supplementary Figure S1



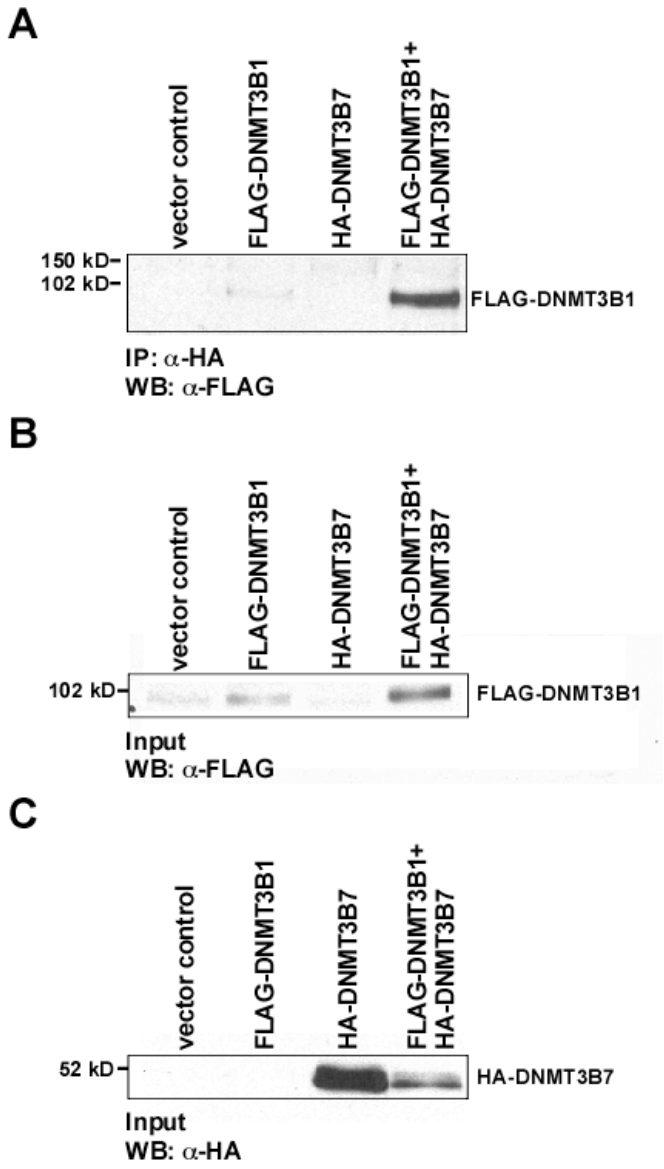
Supplementary Figure S1. Gross embryonic and adult phenotypes of *DNMT3B7* transgenic mice. (A-D) Gross photographs of wild-type versus homozygous transgenic embryos. WT, wild-type; homoz, homozygous. Line A (A) WT and (B) homozygous embryos, E17.5. Line C (C) WT and (D) homozygous embryos, E10.5. (E-H) Gross photographs of wild-type versus hemizygous transgenic animals. Hemiz, hemizygous. Line A (E) WT and (F) hemizygous males, 25 weeks of age. Line C (G) WT and (H) hemizygous females, 6 weeks of age. (I) Weights of Line A wild-type (n = 27), hemizygous (n = 62), and homozygous (n = 22) mice at P0. * denotes P<0.02 between wild-type and homozygous animals using the two-tailed Student's t-test. (J) Weights of Line A wild-type and hemizygous mice over time (n≥10 animals at each timepoint). * denotes P<0.02 after 25 weeks for wild-type and hemizygous male mice using the two-tailed Student's *t*-test. Line A hemizygous animals were the same size as their non-transgenic littermates initially, but stopped gaining weight at 25 weeks of age. (K) Weights of Line C wild-type and hemizygous mice over time (n≥7 animals at each timepoint). * denotes P<0.02 for all timepoints between wild-type and hemizygous males and females using the two-tailed Student's *t*-test. Values represent mean ± s.e.m.

Supplementary Figure S2

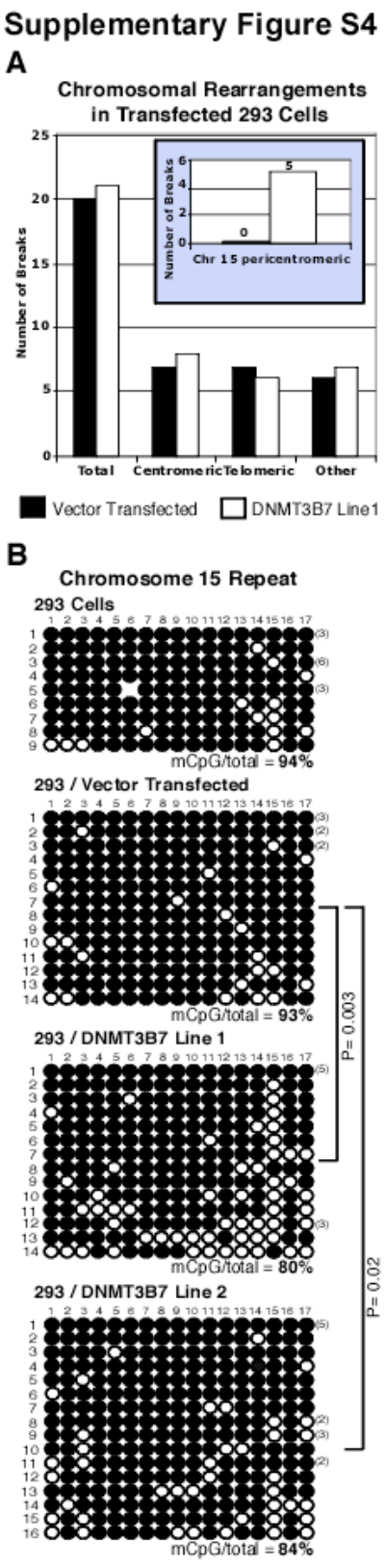


Supplementary Figure S2. Expression of endogenous *Dnmts* and *DNMT3B7* in mouse embryos, adult organs, and mediastinal tumors. (A-D) *DNMT3B7* expression by RNA *in situ* hybridization. Whole mount *in situ* hybridizations demonstrate the location of *DNMT3B7* expression in Line A hemizygous embryos at (A) E9.5, (B) E10.5, and (C) E11.5, versus a wild-type embryo at (D) E10.5. (E) RT-PCR for expression of *Dnmt1*, *Dnmt3a*, *Dnmt3b*, and *Dnmt3l* in embryos from E7.5-E10.5. (F) RT-PCR for *DNMT3B7* expression in organs from Line A and Line C hemizygous male mice at 32 weeks. (G) Western blot of nuclear protein extracts derived from Line A tissues isolated from a female mouse at 16 weeks. The migration of the molecular weight markers is given to the left. WT, wild-type; Tg, transgenic. The source of each protein extract is given along the top. Pos, positive control; Lymph, lymph nodes. In each lane, 60 μg of protein was loaded, and a Western blot for DNMT3B was performed, with UBC9 as the loading control. (H) A Western blot for DNMT3B of nuclear protein extracts derived from the brain of a Line A homozygous embryo at E14.5. A Western blot for lamin B served the loading control. Sixty μg of protein was loaded in each lane. (I) Western blot for DNMT3B in cell lines derived from *Eμ-Myc* or *Eμ-Myc/DNMT3B7* mediastinal lymphomas. The animal from which cells were derived is given at the top. A Western blot for α-tubulin served as the loading control. Fifteen μg of protein was loaded in each lane.

Supplementary Figure S3

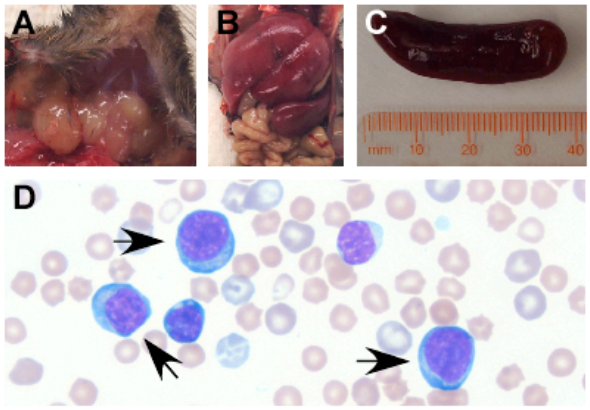


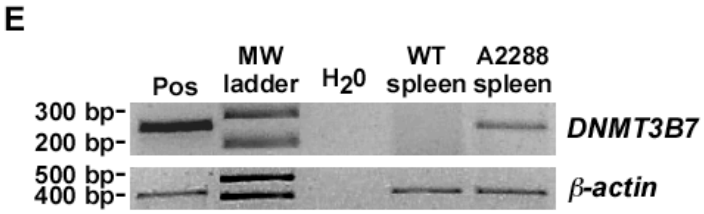
Supplementary Figure S3. Interaction of DNMT3B and DNMT3B7 in cells. (A) Co-immunoprecipitation using anti-HA antibody from the lysates of 293 cells transfected with expression vectors for FLAG-DNMT3B1 and HA–DNMT3B7 used alone or in combination. The precipitates were analysed by immunoblotting with an anti-FLAG antibody. IP, immunoprecipitation; WB, Western blot. Western blots for (B) FLAG-DNMT3B1 and (C) HA-DNMT3B7 of input lysates.

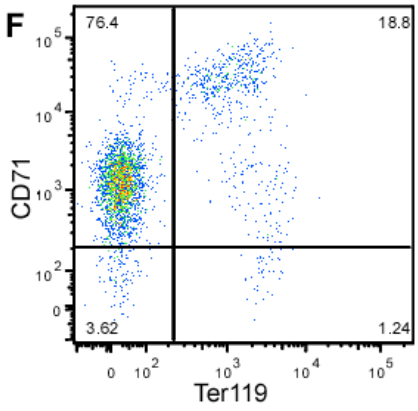


Supplementary Figure S4. Expression of DNMT3B7 causes chromosomal rearrangements and hypomethylation of repetitive DNA sequences. (A) Expression of DNMT3B7 in 293 cells results in an increased number of structural rearrangements involving the pericentromeric region of chromosome 15. Twenty metaphase cells were examined for each cell line (293; vector-transfected; DNMT3B7-expressing Lines 1 and 2) at various passages and were completely karyotyped to identify chromosomal aberrations and structural rearrangements. The graph illustrates the frequency of the total number of rearrangements, the centromeric rearrangements, the telomeric rearrangements, and other chromosomal rearrangements for vector-transfected (black bars) vs. DNMT3B7-expressing cells (white bars) at passage 10. The inset shows the frequency of rearrangements of the pericentromeric region of chromosome 15. Each rearrangement was an independent new abnormality. (B) Methylation state of 17 individual CpG dinucleotides from a chromosome 15-specific repetitive DNA element in 293 cells, as determined by sodium bisulfite analysis. Methylated CpG dinucleotides are represented by filled-in black circles, and unmethylated CpG dinucleotides are represented by open circles. Each numbered row represents an individual clone, and the CpG dinucleotide number is given across the top of each section. The P values are given for comparisons between the vector-transfected cells and each transfected cell line. Depictions of the methylation state of; vector-transfected cells; and DNMT3B7-expressing Lines 1 and 2. DNMT3B7-expressing cells were hypomethylated relative to both parental and vector-transfected cells. The P values for the comparisons to the parental cells were: parental 293 cells vs. Line 1, P=0.002; parental 293 cells vs. Line 2, P=0.01. The P values for the comparisons to the vector-transfected cells are given in the figure.

Supplementary Figure S5







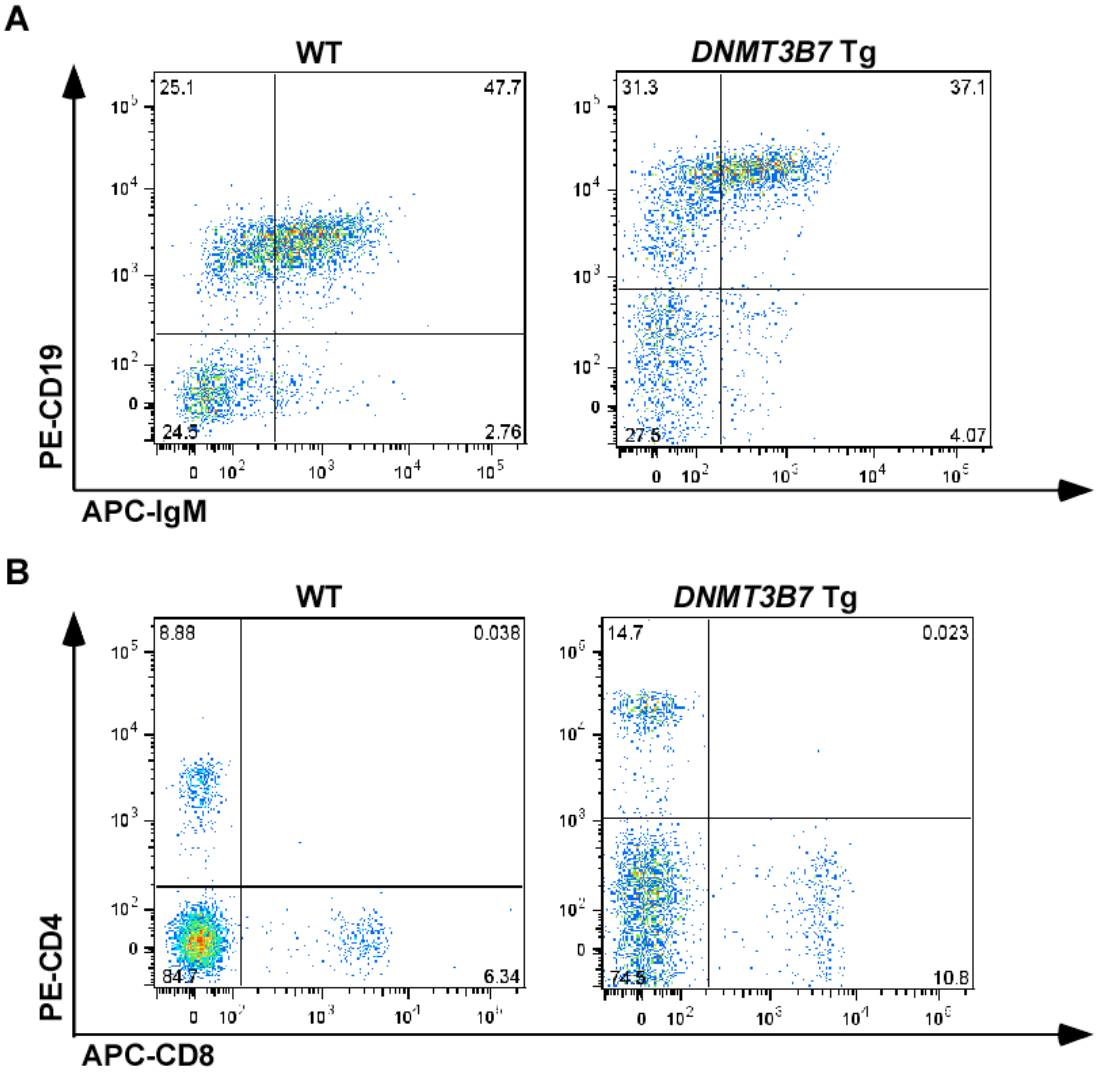
G

Karyotype:

Normal: 40,XX [2/10]
Clone 1: 41,X,-X,+9,t(14;19)(D;D1),+15,dup(17)(CE2) [6/10]
Non-clonal abnormality 1: 39,X,-X,t(14;19)(D;D1),dup(17)(CE2) [1/10]
Non-clonal abnormality 2: 40,X,-X,+9,t(14;19)(D;D1),dup(17)(CE2) [1/10]

Supplementary Figure S5. Erythroleukemia in a *DNMT3B7* hemizygous mouse at 43 weeks. (A-C) Gross photographs of involved organs in a *DNMT3B7* hemizygous mouse with erythroleukemia: (A) lymph nodes; (B) liver; and (C) spleen. (D) Hemotoxylin and eosin-stained peripheral blood cells. Black arrows indicate erythroblasts. (E) RT-PCR expression of DNMT3B7 in the spleen. Pos, positive control; WT, wild-type mouse; A2288, DNMT3B7 hemizygous mouse with erythroleukemia. (F) Flow cytometric analysis of peripheral blood cells for Ter119 and CD71 markers. (G) Karyotype of the erythroleukemic cells.

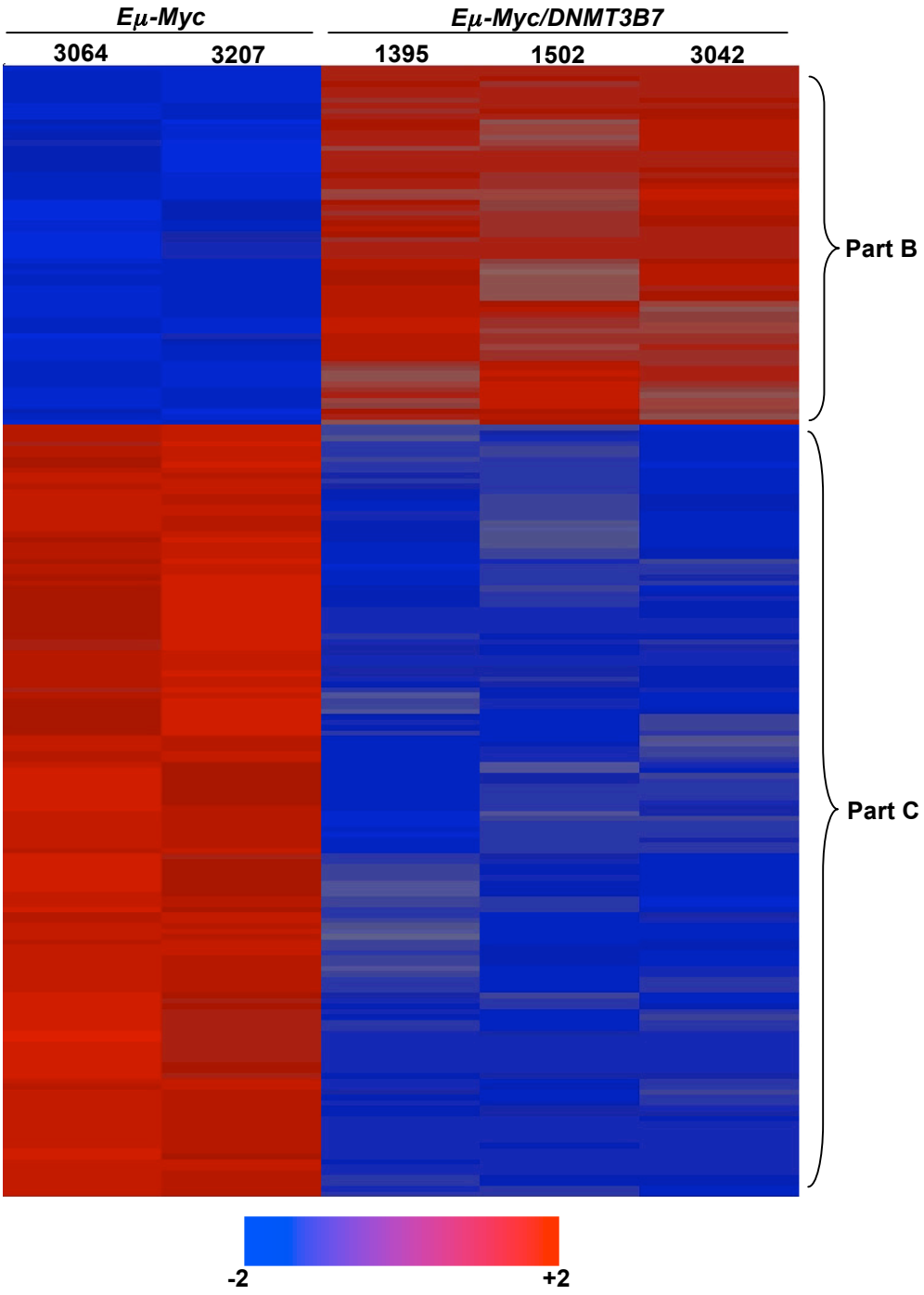
Supplementary Figure S6



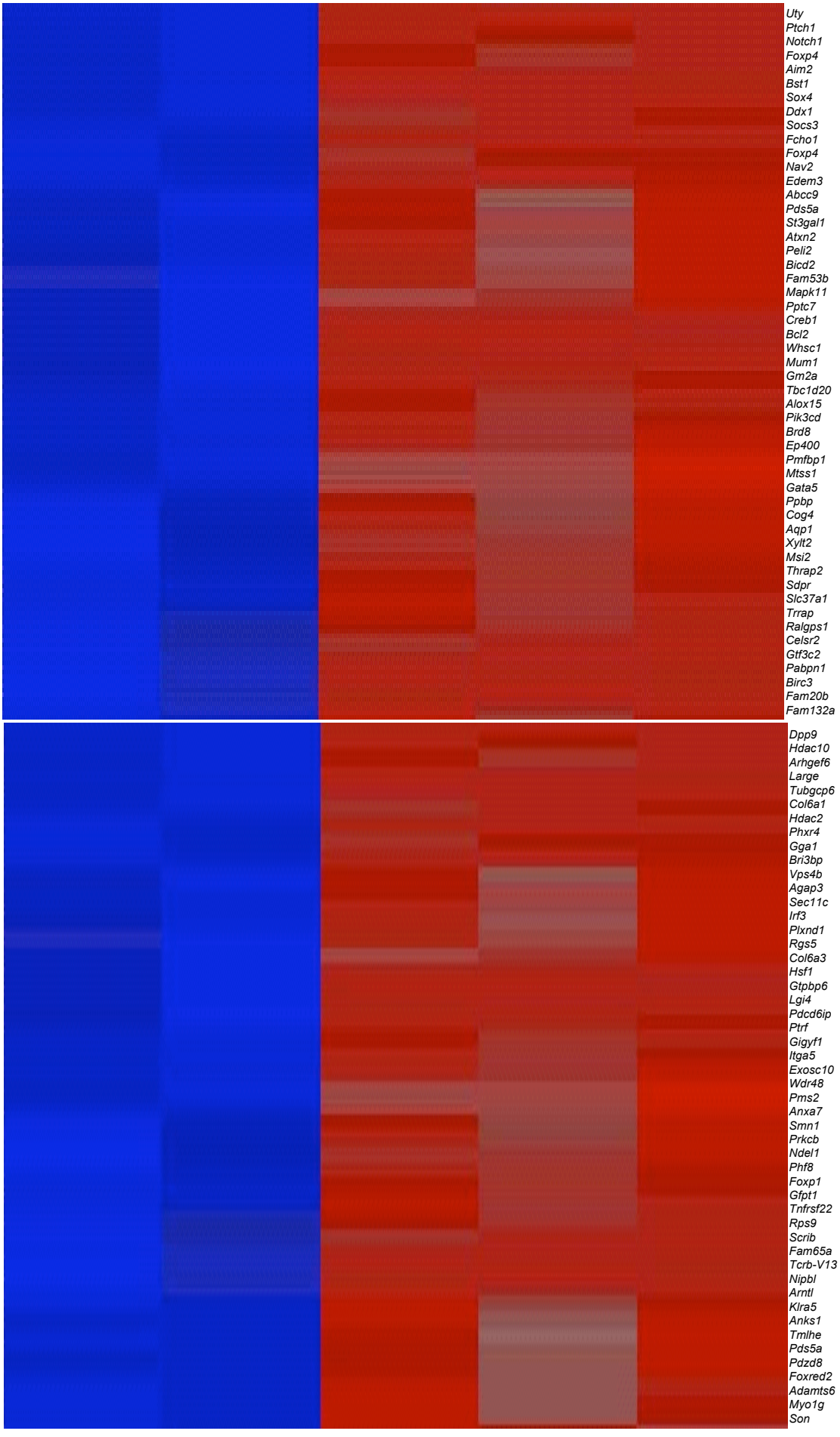
Supplementary Figure S6. Characterization of peripheral blood from 50-week old wild-type (WT) and *DNMT3B7* transgenic (Tg) mice. FACS analysis on peripheral blood for expression of the B lineage markers (A) CD19 and IgM and the T lineage markers (B) CD4 and CD8. Shown are representative data from a single WT and a single *DNMT3B7* transgenic animal. There is no statistically significant difference in peripheral blood between WT and *DNMT3B7* mice.

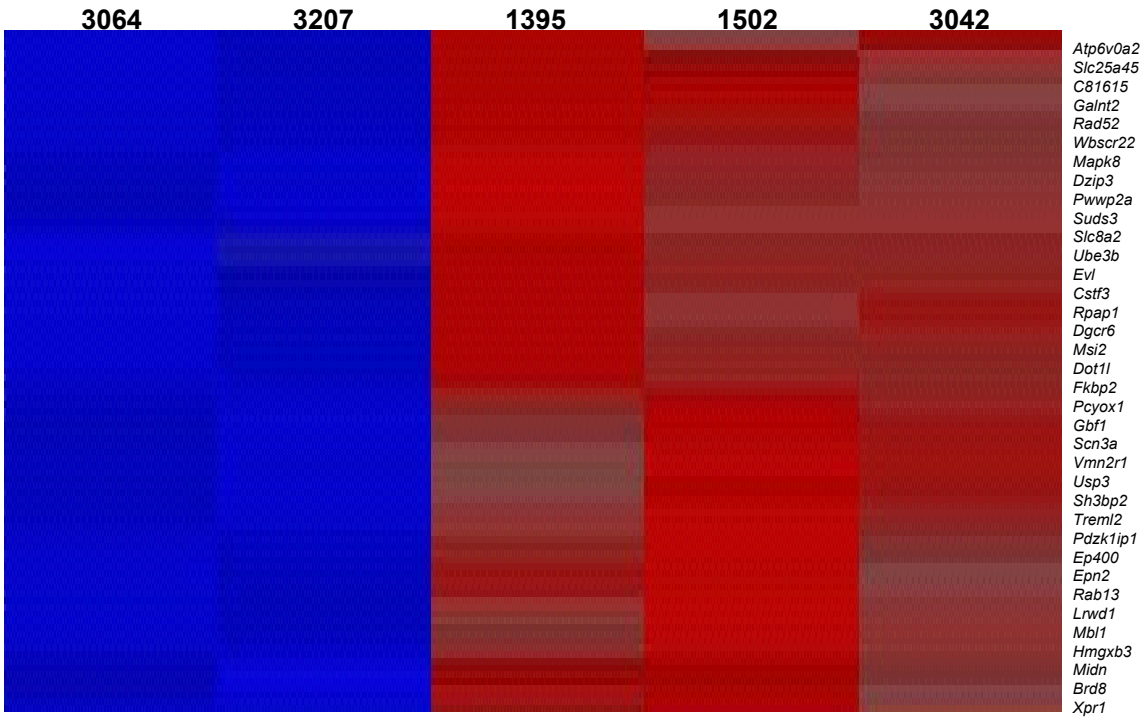
Supplementary Figure S7

A

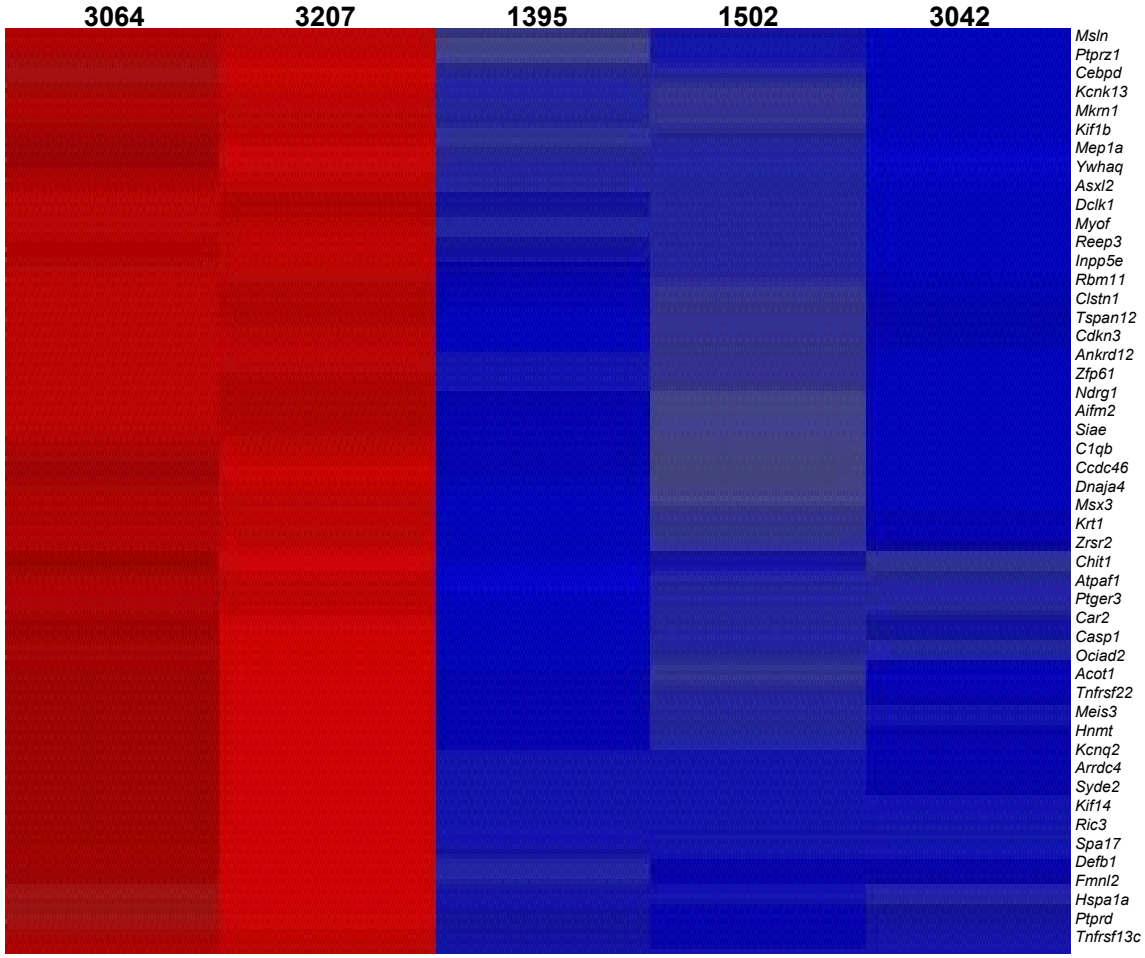


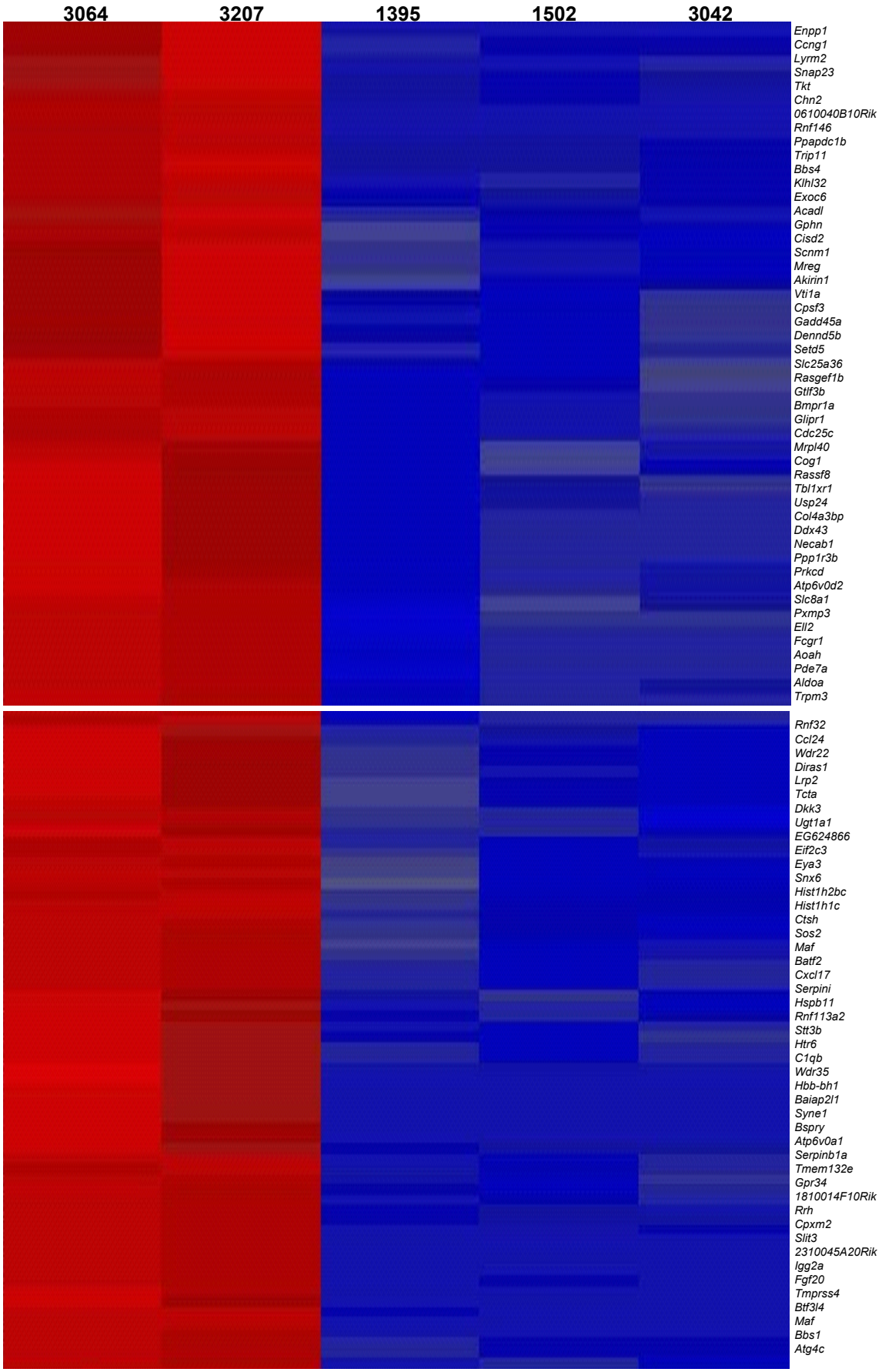
B

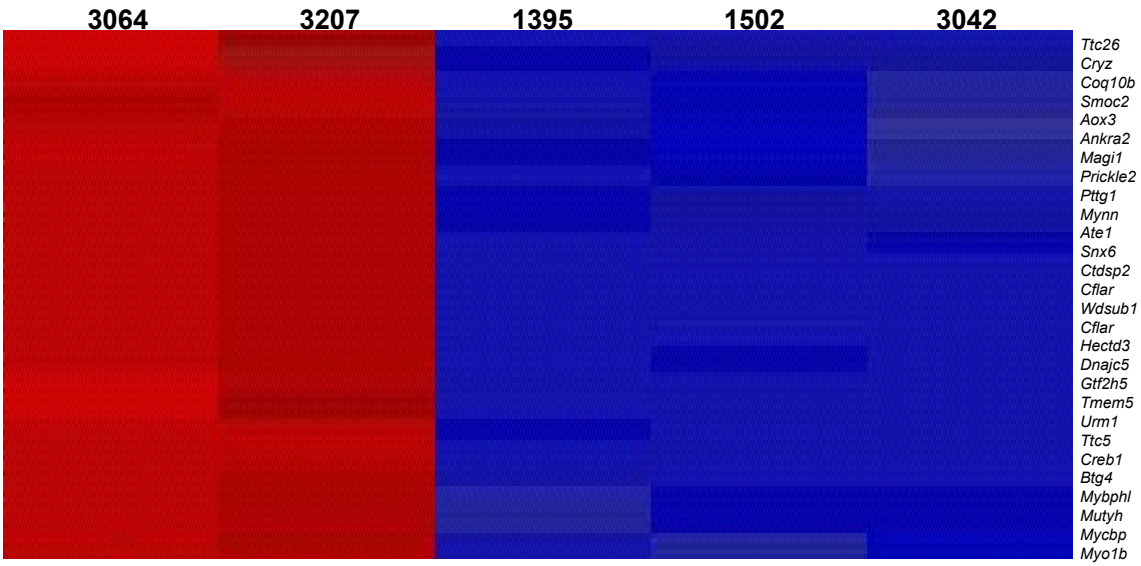

3064
3207
1395
1502
3042
Uty
Ptch1
Notch1
Foxp4
Aim2
Bst1
Sox4
Ddx1
Socs3
Fcho1
Foxp4
Nav2
Edem3
Abcc9
Pds5a
St3gal1
Atxn2
Peli2
Bicd2
Fam53b
Mapk11
Pptc7
Creb1
Bcl2
Whsc1
Mum1
Gm2a
Tbc1d20
Alox15
Pik3cd
Brd8
Ep400
Pmfbp1
Mtss1
Gata5
Ppbp
Cog4
Aqp1
Xylt2
Msi2
Thrap2
Sdpr
Slc37a1
Trrap
Ralgps1
Celsr2
Gtf3c2
Pabpn1
Birc3
Fam20b
Fam132a
Dpp9
Hdac10
Arhgef6
Large
Tubgcp6
Col6a1
Hdac2
Phxr4
Gga1
Bri3bp
Vps4b
Agap3
Sec11c
Irf3
Plxnd1
Rgs5
Col6a3
Hsf1
Gtpbp6
Lgi4
Pdcd6ip
Ptrf
Gigyf1
Itga5
Exosc10
Wdr48
Pms2
Anxa7
Smn1
Prkcb
Ndel1
Phf8
Foxp1
Gfpt1
Tnfrsf22
Rps9
Scrib
Fam65a
Tcrb-V13
Nipbl
Arntl
Klra5
Anks1
Tmlhe
Pds5a
Pdzd8
Foxred2
Adamts6
Myo1g
Son


3064
3207
1395
1502
3042
Atp6v0a2
Slc25a45
C81615
Galnt2
Rad52
Wbscr22
Mapk8
Dzip3
Pwwp2a
Suds3
Slc8a2
Ube3b
Evl
Cstf3
Rpap1
Dgcr6
Msi2
Dot1l
Fkbp2
Pcyox1
Gbf1
Scn3a
Vmn2r1
Usp3
Sh3bp2
Treml2
Pdzk1ip1
Ep400
Epn2
Rab13
Lrwd1
Mbl1
Hmgxb3
Midn
Brd8
Xpr1

C

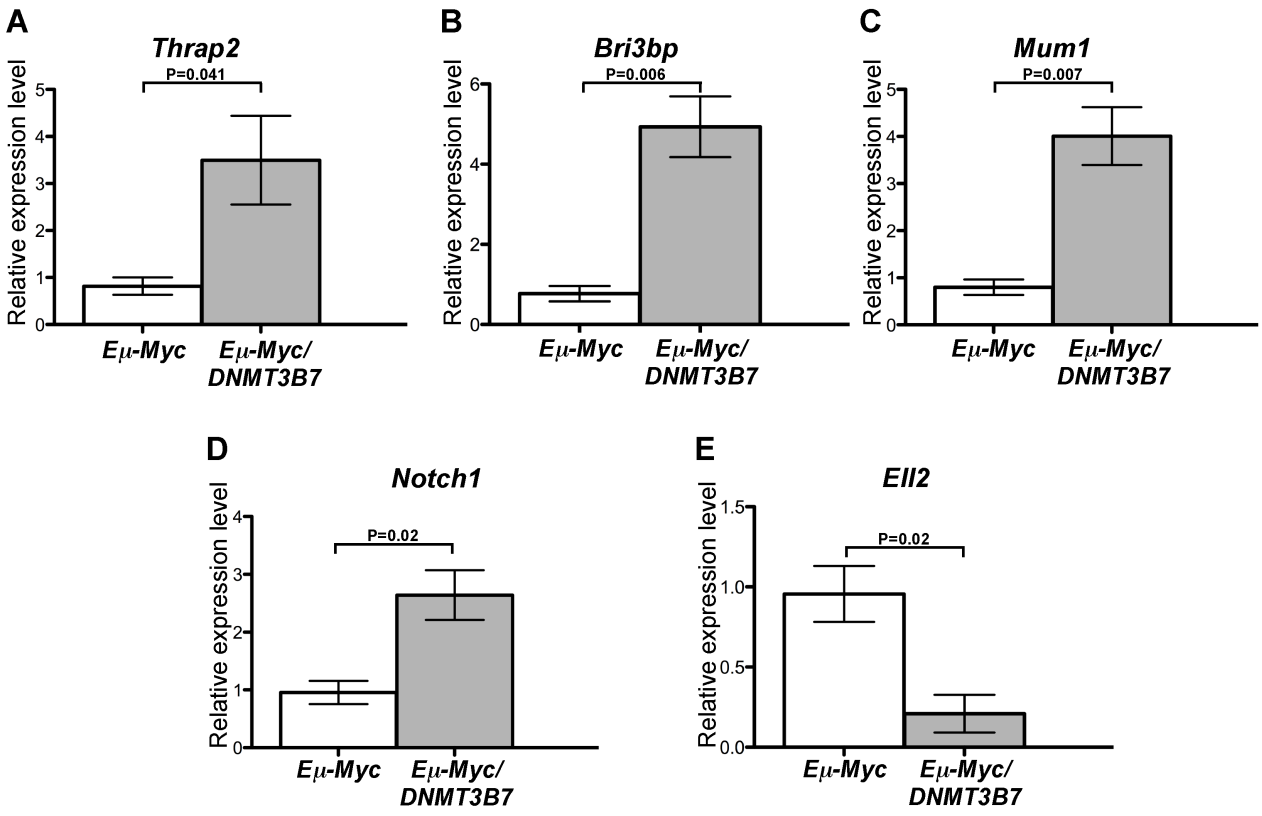

3064
3207
1395
1502
3042
Msln
Ptprz1
Cebpd
Kcnk13
Mkrn1
Kif1b
Mep1a
Ywhaq
Asxl2
Dclk1
Myof
Reep3
Inpp5e
Rbm11
Clstn1
Tspan12
Cdkn3
Ankrd12
Zfp61
Ndrg1
Aifm2
Siae
C1qb
Ccdc46
Dnaja4
Msx3
Krt1
Zrsr2
Chit1
Atpaf1
Ptger3
Car2
Casp1
Ociad2
Acot1
Tnfrsf22
Meis3
Hnmt
Kcnq2
Arrdc4
Syde2
Kif14
Ric3
Spa17
Defb1
Fmnl2
Hspa1a
Ptprd
Tnfrsf13c


3064
3207
1395
1502
3042
Enpp1
Ccng1
Lyrm2
Snap23
Tkt
Chn2
0610040B10Rik
Rnf146
Ppapdc1b
Trip11
Bbs4
Klhl32
Exoc6
Acadl
Gphn
Cisd2
Scnm1
Mreg
Akirin1
Vti1a
Cpsf3
Gadd45a
Dennd5b
Setd5
Slc25a36
Rasgef1b
Gtlf3b
Bmpr1a
Glipr1
Cdc25c
Mrpl40
Cog1
Rassf8
Tbl1xr1
Usp24
Col4a3bp
Ddx43
Necab1
Ppp1r3b
Prkcd
Atp6v0d2
Slc8a1
Pxmp3
Ell2
Fcgr1
Aoah
Pde7a
Aldoa
Trpm3
Rnf32
Ccl24
Wdr22
Diras1
Lrp2
Tcta
Dkk3
Ugt1a1
EG624866
Eif2c3
Eya3
Snx6
Hist1h2bc
Hist1h1c
Ctsh
Sos2
Maf
Batf2
Cxcl17
Serpini
Hspb11
Rnf113a2
Stt3b
Htr6
C1qb
Wdr35
Hbb-bh1
Baiap2l1
Syne1
Bspry
Atp6v0a1
Serpinb1a
Tmem132e
Gpr34
1810014F10Rik
Rrh
Cpxm2
Slit3
2310045A20Rik
Igg2a
Fgf20
Tmprss4
Btf3l4
Maf
Bbs1
Atg4c

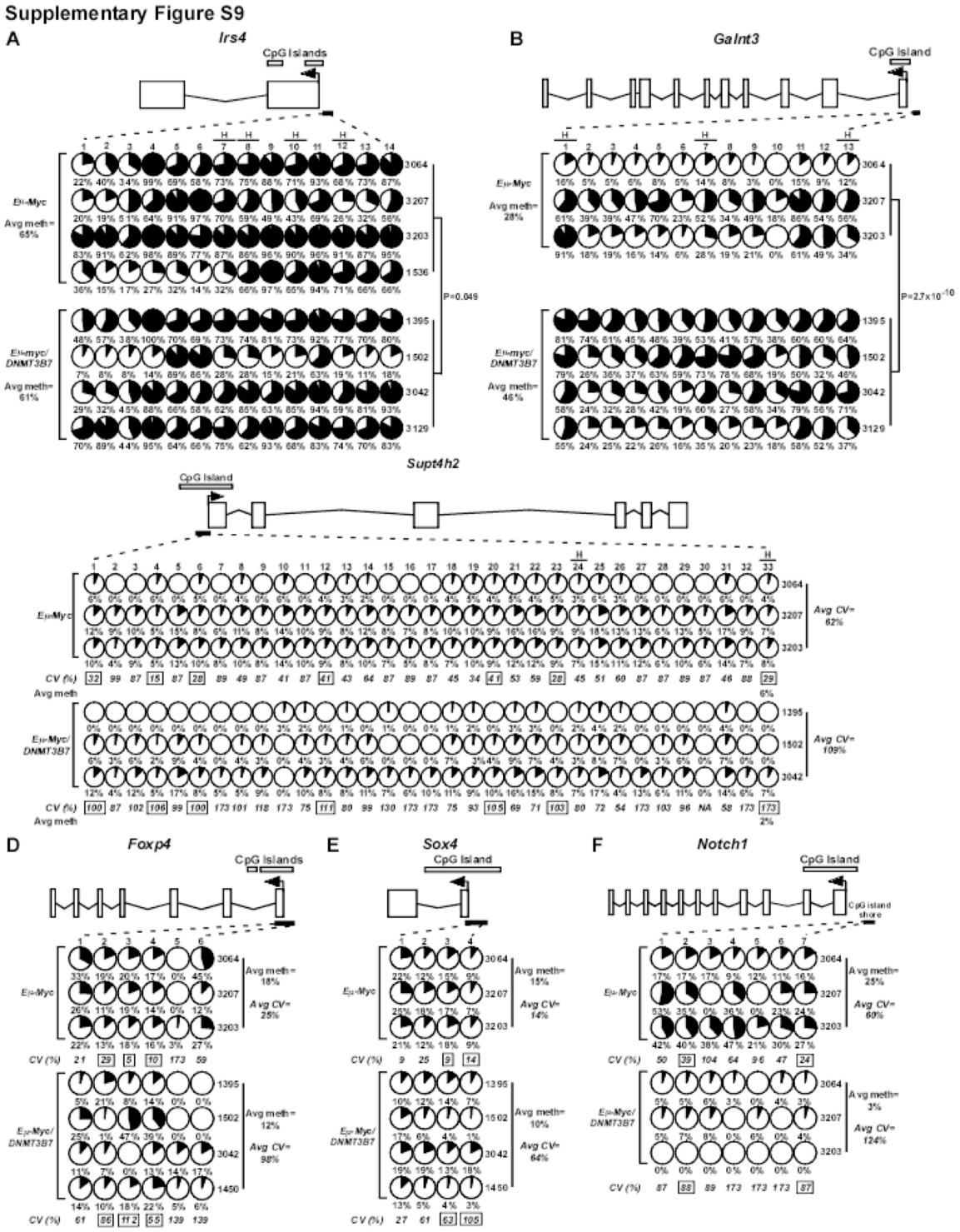


Supplementary Figure S7. Gene expression profiles derived from mediastinal lymphomas of *Eμ-Myc* and *Eμ-Myc/DNMT3B7* transgenic mice. (A) Supervised hierarchical clustering of mediastinal lymphomas from *Eμ-Myc* and *Eμ-Myc/DNMT3B7* transgenic mice. The heat map of the supervised hierarchical clustering was derived using Partek. Each row represents the relative level of expression for a single gene, and each column shows the gene expression levels for a single sample. Mouse identification numbers and genotypes are given at the top. The expression level is indicated by color, with red showing fold overexpression and blue showing fold underexpression. The color scale is indicates fold change. Brackets on the right indicate which sections are depicted in Parts B and C of the figure. (B-C) Enlarged sections of the heat map. Gene names are shown on the right.

Supplementary Figure S8



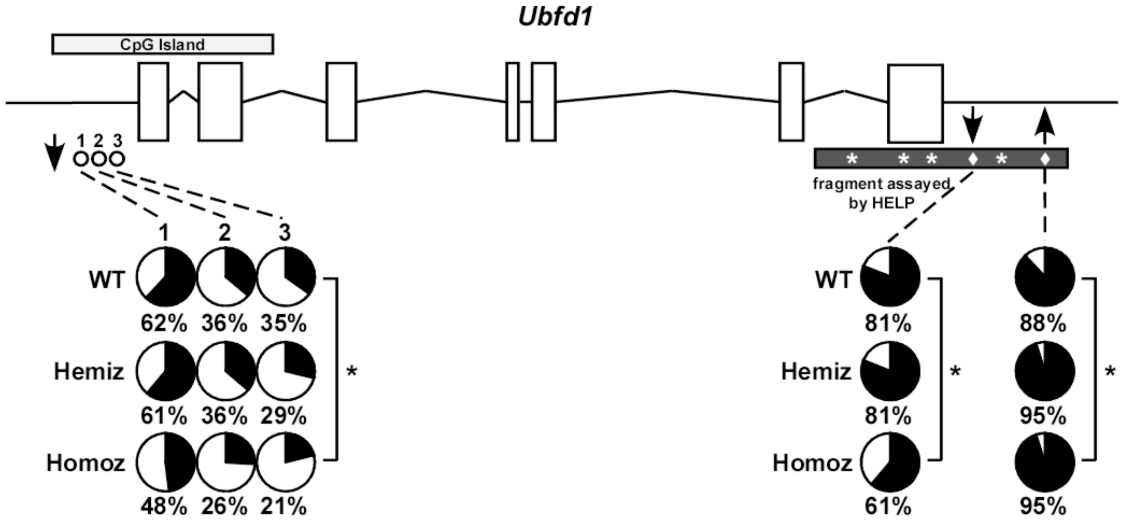
Supplementary Figure S8. Validation of Affymetrix gene expression data by quantitative real-time PCR analyses. The expression of (A) *Thrap2*, (B) *Bri3bp*, (C) *Mum1*, (D) *Notch1* and (E) *Ell2* in *Eμ-Myc/DNMT3B7* tumor cell lines was determined by real-time PCR using Power SYBR Green PCR Master Mix (Applied Biosystems), and is shown in proportion to the expression of *β-actin*. Expression was normalized to the *Eμ-Myc* tumor cell line 3064. Values are expressed as mean ± standard deviation (n = 3). P values were determined using the two-tailed Student's *t*-test.



Supplementary Figure S9. Variability of DNA methylation changes in mediastinal lymphomas of *Eμ-Myc/DNMT3B7* mice. Methylation changes were assessed by bisulfite sequencing in the CpG island (A) *Irs4*, (B) *Galnt3*, (C) *Supt4h2*, (D) *Foxp4*, (E) *Sox4*, and (F)

the CpG island shore of *Notch1* in lymphomas of *Eμ-Myc* and *Eμ-Myc/DNMT3B7* mice. Schematic diagrams for each gene are shown (not to scale). Exons are represented by vertical rectangles, and the location of the CpG island is shown with a horizontal, shaded rectangle. The black arrow indicates the location of the transcriptional start site (TSS). The smaller, black horizontal rectangle indicates the location of the CpGs analyzed for changes in DNA methylation. The numbers across the top indicate specific CpG dinucleotides in a region of the CpG island. Changes in methylation of specific CpG dinucleotides are indicated by the small, shaded circles. Shading indicates the amount of DNA methylation at each specific CpG, and the number below the circles represents the percent methylated cytosine. The average percent methylation for each genotype is indicated to the right, except in panel B, where the average is indicated underneath only for the significantly hypomethylated *Hpa*II site at position 33. In panels A-E, each row represents methylation data obtained from a single lymphoma, indicated to the right. In panels A and B, CpG positions with an "H" over the number indicate the location of a HpaII site. In panels C-E, the average coefficient of variance (CV) is given as a percentage underneath each CpG position. CV values with a black box around the number denote $P \leq 0.05$ for *Eμ-Myc* versus *Eμ-Myc/DNMT3B7* lymphomas at specific CpGs. The average CV across all positions for each genotype is indicated to the right.

Supplementary Figure S10



Supplementary Figure S10. Measurement of locus-specific DNA methylation changes in genomic DNA from brains of E15.5-P0 Line A versus wild-type animals. Schematic of the gene *Ubfd1* and validation of the HELP assay results by bisulfite sequencing. Exons are represented by vertical rectangles, with the location of the CpG island, and the fragment interrogated by the HELP assay indicated by horizontal rectangles. Asterisks indicate the location of *Hpa*II sites in the HELP fragment. Diamonds indicate the location of *Hpa*II sites that had significant changes in methylation. Hypermethylation is indicated by an upward arrow, whereas hypomethylation is indicated by a downward arrow. The open circles refer to a region of the CpG island that is significantly hypomethylated. Each shaded circle represents an individual CpG dinucleotide. Shading indicates the average amount of DNA methylation at each specific CpG, and the number below the circle represents percent methylated cytosine. The DNA methylation state of the *Hpa*II fragments within the *Ubfd1* gene and found to be significant by the HELP assay was tested using PCR amplification and sequencing of bisulfite-treated brain DNA. The percentage of DNA methylation is given for the *Hpa*II site (chromosome 7, position 109,377,238 in the left panel and position 109,380,237 in the right panel on the Mouse May 2005 assembly, UCSC Genome Project). Results shown are averages of n = 4 for each genotype. * denotes $P \leq 0.02$ for comparisons to wild-type animals using the two-tailed Student's *t*-test.

Supplementary Tables

Supplementary Table S1: Primer sequences used for RT-PCR amplification.

Primer name	5'-3' sequence	Annealing temperature
(m) *ApoB100* real time F	GCACGTGGGCTCCAG	60°C
(m) *ApoB100* real time R	GTCATTTCTGCCTTTGCG	
(h) *DNMT3B7* real time F	AAACCCAACAACACGCAAC	60°C
(h) *DNMT3B7* real time R	GGTAGGAGGGTCCAGAGAG	
(h) *DNMT3B7* 2260-2278	CCCGTATTCTCTCTGGACC	62°C
(h) *DNMT3B7* 3371-3350	CAGTTTAGTAGTTGGACTTAGG	
(m) *Dnmt1* 1135-1154	ATCAACTCACCAAAGTGCCC	58°C
(m) *Dnmt1* 1236-1218	CAACATCTGGGGTTCATCC	
(m) *Dnmt3a* 1893-1912	AGACCCCTGGAACTGCTACA	57°C
(m) *Dnmt3a* 2924-2905	ACTACTTCAGTTTGCCCCCA	
(m) *Dnmt3b* ex9F' 1144-1166	GCACTTTAATCTGGCTACCTTCA	59°C
(m) *DNMT3b* ex13R' 1520-1499	CTTCCAGATTGCCCTTGTTGTT	
(m) *Dnmt3L* 956-975	CGGGTACTGAGCCTTTTTAG	63°C
(m) *Dnmt3L* 1602-1583	GAAGAAATCTGCCCTGGTTC	
(m) β-*actin* F'	GGCATTGTTACCAACTGGGACG	54°C
(m) β-*actin* R'	TTTGATGTCACGCACGATTTCC	
(m) *Thrap2* real time F'	CAT TGT TGG TGG CGT CCG GA	56°C
(m) *Thrap2* real time R'	TCT GGC TCC TCT TAG ACT GGG	
(m) *Bri3bp* real time F'	GGC TCA GAA GTT ACT GTC CAG G	56°C
(m) *Bri3bp* real time R'	GAC AGG TTG GAC ACG TCA AGC	
(m) *Mum1* real time F'	AGC ACG CAC TGC TGG ACA GA	47°C
(m) *Mum1* real time R'	TCA GGC AGC ACC TTC CTG CAT G	
(m) *Notch1* real time F'	GTG CTC TGG GTG CCA ACC CTT	60°C
(m) *Notch1* real time R'	ACA GGT GCC CGT TGA AGC CTT T	
(m) *Ell2* real time F'	TCT CTA GCT CTG GAG CCT CCC A	60°C
(m) *Ell2* real time R'	GGA TTC AGA TTG GCC ACC TGT TG	

(m) mouse, (h), human

Supplementary Table S2: Primer sequences used for bisulfite PCR amplification.

Primer name	5'-3' sequence	Annealing temperature
(m) *Ubfd1* F2771-2789	AAGGTTTTTTGTTTTGAGG	52°C
(m) *Ubfd1* R2852-2833	AAAATTCTAAAAAACTTCCC	
(m) *Ubfd1* F5728-5748	ATGTTGTATTTTTGTTAGTTA	52°C
(m) *Ubfd1* R5835-5815	AAAAATTCTAATCATATCCAA	
(m) *Thrap2* F'	GGTTTGGAGATAAATAAGG	49°C
(m) *Thrap2* R'	AATAACAATCCTCCAAACTC	
(m) *Bri3bp* F'	GTTTTTAGGAGGGGTTTGAGAA	57°C
(m) *Bri3bp* R'	CCTCCCCTATACCCTTAAAAC	
(m) *Mum1* F'	GAAGGTATTAGAAGTTTGTGAGGAGTT	60°C
(m) *Mum1* R'	ACATCATAACCCCTCCCACAAA	
(m) *Foxp4* F'	GGGGGGGGTTTTGATTATTTTT	54°C
(m) *Foxp4* R'	CACAATAAACTAAAAACCCATCC	
(m) *Sox4* F'	AGTTTTTTTTTTTGTAGGAGGGAGGG	61°C
(m) *Sox4* R'	TTTCCCCCACCCCTTTCCCTA	
(m) *Notch1* F'	TAGTTGTTAGAGTTGAGAGTTAGAG	58°C
(m) *Notch1* R'	CCCAAAAAACAACAAAAAACCAAAACC	
(m) *Ell2* Bis2 F	GAGATTTATTAGAGTTATAAGGTGAG	56°C
(m) *Ell2* Bis2 R	AACCAAACAAACCAAAAACCCC	
(m) *Mnt* Bis2 F	GTTTGGTGATGTAGTTTAATAGGATG	58°C
(m) *Mnt* Bis2 R	CCTATAATCTAACCCCTAAATCTCTT	
(m) *Supt4h2* Bis2 F	GAAGAGTATTTAGGTTAGAGTATTTTG	55°C
(m) *Supt4h2* Bis2 R	TAAACAAACCATTACCCAAAATACC	
(m) *Irs4* Bis2 F	TGTTTATGTTTTTGTTTTTTTAGGTAGT	54°C
(m) *Irs4* Bis2 R	TCAACAACAACAACAACCTCC	
(m) *Galnt3* Bis1 F	TAGGGGTATTAGGGGGATTTT	51°C
(m) *Galnt3* Bis1 R	TACCCTACCCACCTACCAA	
(m) *Notch1* isl-1	GTATTAGTTTTTGGGGAGTT	52°C
(m) *Notch1* isl-2	TAAAAAAAATCCTAAAACAAAAA	

(m) mouse

Supplementary Table S3: Survival statistics in Lines A and C identifies early lethality.

Age	Genotype	Animals alive at indicated time	
		Line A x Line A	Line C x WT
P0	WT	36/148 (24%)	24/47 (51%)
	Hemiz	88/148 (59%)	23/47 (49%)
	Homoz	24/148 (16%) *	N/A
3 wks	WT	26/92 (28%)	322/474 (68%)
	Hemiz	66/92 (72%)	152/474 (32%) **
	Homoz	0/92 (0%) **	N/A

* $P \leq 0.03$
** $P \leq 0.0001$

Supplementary Table S4: Craniofacial abnormalities in *DNMT3B7* transgenic mice.

Age	Line	Cleft palate	Eye abnormalities*	
			Gross examination	Histologic examination
E14.5-E17.5	Line C hemizygous	ND	0/39 (0%)	ND
	Line A homozygous	2/5 (40%)	19/65 (29%)	
E18.5-P0	Line C hemizygous	3/8 (38%)	2/35 (6%)	1/8 (13%)
	Line A homozygous	6/10 (60%)	14/60 (23%)	6/10 (60%)

* Eye abnormalities include: no lens; abnormal attachment; abnormal fibrovascular tissue; deep vestigial structure

ND=not done

Supplementary Table S5: Cardiac abnormalities in *DNMT3B7* transgenic mice.

Embryonic Age	Line	Side-by-side great vessels	VSD	Thin myocardium
E14.5	WT	N/A	3/11 (27%)	0/11 (0%)
	Line C hemizygous	N/A	7/13 (54%)	1/13 (8%)
	Line A homozygous	N/A	7/8 (88%)	2/8 (25%)
P0	WT	0/2 (0%)	0/2 (0%)	N/A
	Line C hemizygous	0/4 (0%)	2/9 (22%)	2/11 (18%)
	Line A homozygous	4/12 (33%) *	6/11 (55%) †	0/12 (0%)

* Of 4 animals, 1 had right-sided aorta.
† Of 6 animals, 2 had side-by-side great vessels; 1 had side-by-side great vessels and right-sided aorta.

Supplementary Table S6: Comparison of craniofacial and cardiac abnormalities in Line A and Line C.

Genotype	Mouse ID#	Age	Craniofacial Abnormalities			Cardiac Abnormalities		
			Cleft Palate	R eye defect	L eye defect	VSD	Side-by-side GV	Thin myocardium
Line A homoz	111-5	E14.5	No	✓	No	✓	No	No
	113-4	E14.5	No	✓	No	✓	No	✓
	113-5	E14.5	No	No	No	✓	No	✓
	85-1	P0	✓	✓	✓	✓	No	No
	95-3	P0	No	No	No	No	No	No
	150-2	P0	No	No	No	No	No	No
	150-3	P0	✓	✓	✓	No	✓	No
	27-2	E15.5	✓	✓	No	ND		
	27-6	E15.5	✓	✓	✓			
	24-3	P0	✓	No	No			
	75-6	P0	No	✓	No			
	75-7	P0	No	No	No			
	95-3	P0	No	No	No			
	150-3	P0	✓	✓	✓			
	239-1	P0	✓	✓	✓			
	239-2	P0	✓	✓	✓			
	239-3	P0	✓	✓	✓			
Line C hemiz	138-1	P0	✓	No	No	✓	No	✓
	145-1	P0	✓	No	No	No	No	No
	159-1	P0	No	No	No	No	No	No
	159-2	P0	No	No	No	✓	No	No
	172-1	P0	No	No	No	No	No	No
	172-3	P0	No	✓	No	No	No	No
	16-1	P0	✓	✓	No	ND		
	16-2	P0	No	No	No			
	138-1	P0	✓	No	No			
	145-1	P0	✓	No	No			
	159-1	P0	No	No	No			
	159-2	P0	No	No	No			
	172-1	P0	No	No	No			
	172-3	P0	No	No	No			

ND=not done

Supplementary Table S7: Number of circulating lymphocytes in *DNMT3B7* transgenic mice.

Age (weeks)	Females				Males			
	WT		Transgenic		WT		Transgenic	
	N	Lymphocytes (x 1000/mL)	N	Lymphocytes (x 1000/mL)	N	Lymphocytes (x 1000/mL)	N	Lymphocytes (x 1000/mL)
8-16	35	11.9 ±2.8	37	11.6±2.4	31	10.1±2.8	48	11.4 ±2.4
17-24	27	10.6±3.2	24	10.7±2.5	20	9.1±2.8	26	9.5±2.2
25-32	30	9.6±2.8	32	9.6±2.3	29	8.4±2.1	39	9.3±2.4
33-40	36	9.2±2.1	31	9.2±2.3	25	7.7±2.0	31	8.4±2.5
41-48	39	8.4±2.3	31	8.4±2.6	26	8.2±2.3	35	10.1±3.4
>48	26	8.7±2.3	16	8.7±2.4	27	9.7±4.4	30	11.6 ±2.7

Results are presented as average lymphocyte number ± S.D.

Supplementary Table S8: Genes consistently altered in *Eμ-Myc/DNMT3B7* mediastinal lymphomas relative to *Eμ-Myc* mediastinal lymphomas.

Gene ID	P-value	Fold change
Uty	0.0381	13.08
Ddx1	0.0387	11.10
Socs3	0.0378	7.86
Aim2	0.0206	6.43
Bst1	0.0228	7.42
Sox4	0.0442	6.79
Fcho1	0.0298	6.76
Foxp4	0.0346	6.01
Nav2	0.0187	5.98
Edem3	0.0066	4.79
Abcc9	0.0364	4.60
Pds5a	0.0366	4.53
St3gal1	0.0097	4.43
Atxn2	0.0250	4.35
Peli2	0.0468	3.96
Bicd2	0.0491	3.95
Fam53b	0.0283	3.85
Mapk11	0.0485	3.77
Alox15	0.0064	3.74
Pik3cd	0.0178	3.66
Brd8	0.0475	3.64
Ep400	0.0136	3.63
Pmfbp1	0.0458	3.62
Mtss1	0.0190	3.60
Gata5	0.0428	3.56
Ppbp	0.0267	3.54
Cog4	0.0469	3.54
Aqp1	0.0492	3.51
Foxp4	0.0192	3.37
Xylt2	0.0427	3.32
Msi2	0.0369	3.26
Thrap2	0.0130	3.25
Sdpr	0.0482	3.24
Notch1	0.0350	3.19
Slc37a1	0.0365	3.19

Trrap	0.0430	3.18
Ralgps1	0.0032	3.14
Celsr2	0.0063	3.13
Dot1l	0.0424	3.11
Xpr1	0.0476	3.06
Fam132a	0.0052	3.03
Dpp9	0.0480	3.03
Hdac10	0.0253	3.01
Pabpc1l	0.0375	2.99
Arhgef6	0.0240	2.94
Large	0.0002	2.87
Tubgcp6	0.0292	2.86
Col6a1	0.0151	2.84
Hdac2	0.0412	2.83
Phxr4	0.0269	2.83
Gga1	0.0460	2.81
Vps4b	0.0493	2.79
Agap3	0.0454	2.78
Sec11c	0.0382	2.77
Irf3	0.0466	2.76
Plxnd1	0.0212	2.74
Rgs5	0.0378	2.74
Col6a3	0.0434	2.72
Hsf1	0.0392	2.71
Gtpbp6	0.0099	2.71
Lgi4	0.0202	2.70
Pdcd6ip	0.0383	2.70
Gtf3c2	0.0116	2.69
Pabpn1	0.0470	2.67
Birc3	0.0472	2.66
Fam20b	0.0444	2.65
Brd8	0.0401	2.57
Ptrf	0.0285	2.57
Ptch1	0.0096	2.54
Gigyf1	0.0272	2.53
Itga5	0.0336	2.51
Exosc10	0.0457	2.48
Wdr48	0.0206	2.45
Pms2	0.0303	2.44
Anxa7	0.0454	2.43

Smn1	0.0209	2.43
Prkcb	0.0391	2.39
Ndel1	0.0482	2.33
Phf8	0.0491	2.33
Foxp1	0.0385	2.32
Gfpt1	0.0286	2.31
Tnfrsf22	0.0267	2.30
Rps9	0.0391	2.28
Mett11d1	0.0224	2.26
Scrib	0.0229	2.26
Fam65a	0.0341	2.24
Tcrb-V13	0.0115	2.23
Nipbl	0.0499	2.17
Arntl	0.0426	2.08
Klra5	0.0019	2.06
Anks1	0.0302	2.04
Tmlhe	0.0098	2.04
Pds5a	0.0121	2.04
Pdzd8	0.0096	2.03
Foxred2	0.0109	2.02
Adamts6	0.0236	2.02
Myo1g	0.0433	2.01
Son	0.0225	2.01
Atp6v0a2	0.0281	2.00
Slc25a45	0.0025	1.99
C81615	0.0125	1.98
Galnt2	0.0380	1.97
Rad52	0.0455	1.96
Whsc1	0.0448	1.96
Mum1	0.0438	1.91
Wbscr22	0.0409	1.91
Mapk8	0.0388	1.91
Dzip3	0.0221	1.91
Pwwp2a	0.0189	1.91
Suds3	0.0301	1.87
Slc8a2	0.0050	1.87
Ube3b	0.0272	1.86
Evl	0.0266	1.86
Cstf3	0.0269	1.85
Rpap1	0.0393	1.84

Dgcr6	0.0262	1.84
Msi2	0.0196	1.83
Bcl2	0.0342	1.83
Bri3bp	0.0269	1.75
Gm2a	0.0486	1.62
Tbc1d20	0.0027	1.62
Pptc7	0.0075	1.62
Fkbp2	0.0380	1.61
Pcyox1	0.0074	1.60
Gbf1	0.0098	1.60
Scn3a	0.0041	1.56
Vmn2r1	0.0117	1.55
Usp3	0.0092	1.41
Sh3bp2	0.0147	1.38
Creb1	0.0232	1.38
Treml2	0.0280	1.37
Pdzk1ip1	0.0253	1.35
Ep400	0.0316	1.32
Epn2	0.0343	1.24
Rab13	0.0447	1.24
Lrwd1	0.0486	1.22
Mbl1	0.0398	1.21
Hmgxb3	0.0271	1.20
Midn	0.0416	1.20
Mybphl	0.0457	-1.20
Mutyh	0.0344	-1.20
Mycbp	0.0121	-1.20
Myo1b	0.0442	-1.21
Gtlf3b	0.0422	-1.21
Bmpr1a	0.0082	-1.21
Glipr1	0.0316	-1.21
Cdc25c	0.0481	-1.22
Mrpl40	0.0261	-1.22
Cog1	0.0283	-1.22
Rassf8	0.0330	-1.22
Tbl1xr1	0.0262	-1.23
Usp24	0.0487	-1.23
Col4a3bp	0.0030	-1.23
Rnf32	0.0386	-1.24
Ccl24	0.0172	-1.24

Wdr22	0.0261	-1.24
Diras1	0.0166	-1.25
Lrp2	0.0273	-1.25
Tcta	0.0447	-1.25
Dkk3	0.0174	-1.25
Ugt1a1	0.0219	-1.25
EG624866	0.0137	-1.26
Eif2c3	0.0330	-1.26
Eya3	0.0366	-1.26
Rbm11	0.0278	-1.26
Rbm11	0.0278	-1.26
Clstn1	0.0348	-1.26
Tspan12	0.0051	-1.26
Cdkn3	0.0134	-1.27
Ankrd12	0.0482	-1.27
Zfp61	0.0467	-1.27
Snx6	0.0418	-1.27
Ctdsp2	0.0140	-1.28
Cflar	0.0190	-1.29
Wdsub1	0.0390	-1.28
Cflar	0.0190	-1.29
Hectd3	0.0154	-1.29
Dnajc5	0.0055	-1.29
Gtf2h5	0.0032	-1.29
Tmem5	0.0288	-1.31
Urm1	0.0086	-1.31
Ttc5	0.0235	-1.31
Creb1	0.0278	-1.34
Btg4	0.0164	-1.34
Msln	0.0204	-1.34
Ptprz1	0.0273	-1.34
Cebpd	0.0485	-1.34
Kcnk13	0.0308	-1.34
Mkrn1	0.0176	-1.35
Kif1b	0.0056	-1.35
Mep1a	0.0259	-1.35
Ywhaq	0.0160	-1.36
Asxl2	0.0307	-1.36
Dclk1	0.0178	-1.37
Myof	0.0100	-1.37

Reep3	0.0017	-1.39
Inpp5e	0.0401	-1.39
Ttc26	0.0227	-1.40
Cryz	0.0462	-1.40
Coq10b	0.0296	-1.40
Smoc2	0.0231	-1.40
Aox3	0.0370	-1.40
Ankra2	0.0109	-1.44
Magi1	0.0325	-1.44
Prickle2	0.0158	-1.44
Pttg1	0.0095	-1.45
Mynn	0.0282	-1.45
Ate1	0.0357	-1.45
Lyrm2	0.0381	-1.45
Snap23	0.0347	-1.45
Tkt	0.0270	-1.46
Chn2	0.0410	-1.48
0610040B10Rik	0.0417	-1.48
Rnf146	0.0154	-1.48
Ppapdc1b	0.0451	-1.49
Trip11	0.0329	-1.49
Bbs4	0.0165	-1.50
Klhl32	0.0390	-1.50
Exoc6	0.0475	-1.50
Acadl	0.0342	-1.50
Gphn	0.0071	-1.50
Cisd2	0.0359	-1.51
Scnm1	0.0057	-1.51
Mreg	0.0347	-1.51
Akirin1	0.0454	-1.54
Vti1a	0.0270	-1.60
Cpsf3	0.0470	-1.64
Gadd45a	0.0026	-1.65
Dennd5b	0.0002	-1.65
Setd5	0.0496	-1.74
Slc25a36	0.0243	-1.75
Rasgef1b	0.0295	-1.75
Snx6	0.0280	-1.75
Hist1h2bc	0.0392	-1.76
Serpini1	0.0464	-1.84

Rnf113a2	0.0222	-1.84
Stt3b	0.0123	-1.84
Htr6	0.0305	-1.85
C1qb	0.0338	-1.86
Wdr35	0.0494	-1.86
Hbb-bh1	0.0017	-1.88
Baiap2l1	0.0132	-1.88
Hspa1a	0.0176	-1.88
Ptprd	0.0174	-1.89
Tnfrsf13c	0.0259	-1.90
Enpp1	0.0365	-1.91
Ccng1	0.0471	-1.93
Hist1h1c	0.0003	-1.93
Ctsh	0.0402	-1.94
Sos2	0.0179	-1.95
Maf	0.0292	-1.99
Batf2	0.0126	-2.11
Cxcl17	0.0016	-2.13
Ndrg1	0.0479	-2.16
Aifm2	0.0390	-2.16
Siae	0.0025	-2.16
C1qb	0.0345	-2.17
Ccdc46	0.0417	-2.17
Dnaja4	0.0009	-2.18
Msx3	0.0463	-2.18
Krt1	0.0156	-2.22
Zrsr2	0.0474	-2.25
Chit1	0.0298	-2.26
Atpaf1	0.0356	-2.26
Ptger3	0.0339	-2.26
Ddx43	0.0004	-2.30
Necab1	0.0326	-2.34
Ppp1r3b	0.0219	-2.37
Prkcd	0.0300	-2.38
Atp6v0d2	0.0493	-2.38
Slc8a1	0.0426	-2.39
Pxmp3	0.0218	-2.39
Ell2	0.0298	-2.40
Fcgr1	0.0283	-2.41
Aoah	0.0220	-2.44

Car2	0.0429	-2.51
Casp1	0.0352	-2.52
Pde7a	0.0010	-2.53
Aldoa	0.0060	-2.53
Trpm3	0.0428	-2.54
Tmprss4	0.0408	-2.55
Btf3l4	0.0416	-2.55
Maf	0.0466	-2.56
Bbs1	0.0171	-2.69
Atg4c	0.0430	-2.65
Meis3	0.0426	-2.67
Hnmt	0.0354	-2.67
Kcnq2	0.0334	-2.65
Arrdc4	0.0347	-2.66
Syde2	0.0373	-2.60
Kif14	0.0384	-2.74
Ric3	0.0218	-2.74
Spa17	0.0453	-2.76
Fmnl2	0.0348	-2.76
Syne1	0.0291	-2.77
Bspry	0.0185	-2.81
Atp6v0a1	0.0236	-2.84
Serpinb1a	0.0096	-2.87
Tmem132e	0.0155	-3.02
Gpr34	0.0280	-3.03
1810014F10Rik	0.0100	-3.07
Rrh	0.0225	-3.13
Cpxm2	0.0149	-3.20
Slit3	0.0174	-3.40
2310045A20Rik	0.0173	-3.43
Igg2a	0.0175	-3.57
Ociad2	0.0238	-3.77
Acot1	0.0098	-3.81
Tnfrsf22	0.0097	-3.87
Ccl12	0.0157	-3.94
Hbb-bh1	0.0440	-4.08
Fgf20	0.0015	-4.10
Hspb1	0.0036	-6.56
Defb1	0.0368	-19.60

Supplementary Materials and Methods

Generation and monitoring of transgenic mice

The *DNMT3B7* cDNA (1) was amplified using a high-fidelity *Taq* polymerase (Roche) and ligated to the pEμ/pBSVE6BK vector and sequenced completely. The *DNMT3B7* transgenic construct was purified twice by CsCl density ultracentrifugation and injected into one-cell C57Bl/6 embryos using standard techniques. Founder animals were identified by Southern blotting, and each founder was bred into a line of mice.

Mice carrying the *Eμ-Myc* transgene were monitored by physical examination three times a week for the development of peripheral and mediastinal lymphomas. Mediastinal lymphomas were identified by respiratory compromise introduced when the neck was constrained. All mice were housed in a pathogen-free barrier facility maintained under IACUC guidelines and an approved protocol.

Genotyping of the animals was performed by isolating genomic DNA from yolk sac or tail and testing for the presence of the transgene using either Southern blot analysis or quantitative real-time PCR relative to the single-copy *ApoB* gene. The *Eμ-Myc* cassette was detected by PCR as previously described (2).

RNA in situ hybridizations

An RNA probe specific for *DNMT3B7* was generated using a 610 base-pair fragment inserted into pBluescriptII SK+. Linearization with *Asp*718 and transcription with the T3 polymerase (Promega) produced the antisense probe, whereas linearization

with *Sac*I and transcription with the T7 polymerase (Promega) produced the sense probe. RNA *in situ* hybridization*s* were performed using standard techniques (3).

Gene expression studies

RNA quality was assessed using an Agilent 2100 Bioanalyzer and RNAs with an RNA integrity number ≥7.0 were hybridized to Affymetrix Mouse Genome 430 2.0 gene expression microarrays. The raw data were subjected to quality control assessment by using GC-RMA (Gene Chip-Robust Multiarray Averaging). Partek Genomics Suite 6.3 (Partek Inc.) was used to perform supervised hierarchical clustering and ANOVA analysis to identify the genes that were highly represented.

Real time PCR analyses

The expression of genes identified by the Affymetrix gene expression array in *Eμ-Myc/DNMT3B7* tumor cell lines was analyzed by real time PCR on the Applied Biosystems 7500 Fast PCR system using the Power SYBR Green PCR Master Mix (Applied Biosystems). Primer sequences are listed in Supplementary Table S1. Analysis of each gene was performed by the ΔΔCt method using the endogenous control *β-actin* and shown as a fold change relative to expression of the gene in the *Eμ-Myc* tumor cell lines. Fold change was normalized to the obtained value in the *Eμ-Myc* tumor cell line 3064. The two-tailed Student's *t*-test was used to assess significance.

DNA methylation analysis

The HELP assay was carried out as previously published with slight modifications (4). One microgram of genomic DNA was digested overnight with either *Hpa*II or *Msp*I (NEB). On the following day the reactions were extracted once with phenol-chloroform and resuspended in 11 μL of 10 mM Tris-HCl, pH 8.0, and the digested DNA was used to set up an overnight ligation of the JHpaII adapter using T4 DNA ligase. The adapter-ligated DNA was used to carry out the PCR amplification of the HpaII and MspI-digested DNA as previously described (4). Samples were labeled and hybridized onto a custom mouse oligonucleotide array at the Roche-NimbleGen Service Laboratory (Roche NimbleGen, Design name: 2006-10-26_MM5_HELP_Promoter, Design ID: 4803). The HELP assay array design can be found at http://www.ncbi.nlm.nih.gov/geo/ under accession number GPL10283.

The array was custom designed to cover 25,626 **H**paII **a**mplifiable **f**ragments (HAF) located at gene promoters and imprinted regions. HAF are defined as genomic sequences between two flanking HpaII sites found within 200-2,000 bp from each other. HAF were aligned to the MM5 build of the mouse genome and then annotated to the nearest transcription start site (TSS), allowing for a maximum distance of 5 kb from the TSS. Scanning was performed using a GenePix 4000B scanner (Molecular Devices) (5). Quality control and preprocessing of HELP microarrays was performed as previously described (6). After intra-array normalization, background noise was determined for each channel as the 2.5 median absolute deviations from the median of the random probes' $\log_2$ signal for that channel. Each channel was centered by

subtracting this noise threshold from its $\log_2$-transformed signal intensities. The *Hpa*II/*Msp*I (unmethylated/reference) ratio was then determined for each probe set on the array.

PCR amplification from bisulfite-treated DNA was performed using genomic DNA treated with sodium bisulfite as described (7), and Zymo*Taq* Premix (Zymo Research) or AmpliTaq Gold polymerase (Applied Biosystems), using the primers in Supplementary Table S2. PCR products were gel purified and sequenced at The University of Chicago Cancer Research Center DNA Sequencing Facility. The percent methylation of each CpG was determined by submitting the entire PCR product for "bulk" DNA sequencing analysis. The sequencing results were analyzed with the program 4Peaks, and the percent DNA methylation determined for a single CpG was calculated by measuring the C vs. T peak heights at that particular position on the chromatogram. The percent methylation was determined by the equation:
(C peak height) / (C peak height + T peak height) x 100

Variability in methylation was quantified by calculating the coefficient of variance (CV) for percent methylation at each CpG dinucleotide, as well as the average CV for each set of tumors. Statistical significance was assessed by performing the F test.

Supplementary References

1. Ostler, K. R., Davis, E. M., Payne, S. L., Gosalia, B. B., Exposito-Cespedes, J., Beau, M. M., and Godley, L. A. Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins. Oncogene, *26:* 5553-5563, 2007.
2. Schmitt, C. A., Fridman, J. S., Yang, M., Baranov, E., Hoffman, R. M., and Lowe, S. W. Dissecting p53 tumor suppressor functions in vivo. Cancer Cell, *1:* 289-298, 2002.
3. Correia, K. M. and Conlon, R. A. Whole-mount in situ hybridization to mouse embryos. Methods, *23:* 335-338, 2001.
4. Khulan, B., Thompson, R. F., Ye, K., Fazzari, M. J., Suzuki, M., Stasiek, E., Figueroa, M. E., Glass, J. L., Chen, Q., Montagna, C., Hatchwell, E., Selzer, R. R., Richmond, T. A., Green, R. D., Melnick, A., and Greally, J. M. Comparative isoschizomer profiling of cytosine methylation: the HELP assay. Genome Res, *16:* 1046-1055, 2006.
5. Selzer, R. R., Richmond, T. A., Pofahl, N. J., Green, R. D., Eis, P. S., Nair, P., Brothman, A. R., and Stallings, R. L. Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. Genes Chromosomes Cancer, *44:* 305-319, 2005.
6. Thompson, R. F., Reimers, M., Khulan, B., Gissot, M., Richmond, T. A., Chen, Q., Zheng, X., Kim, K., and Greally, J. M. An analytical pipeline for genomic representations used for cytosine methylation studies. Bioinformatics, *24:* 1161-1167, 2008.

7. Clark, S. J., Harrison, J., Paul, C. L., and Frommer, M. High sensitivity mapping of methylated cytosines. Nucleic Acids Res, *22:* 2990-2997, 1994.