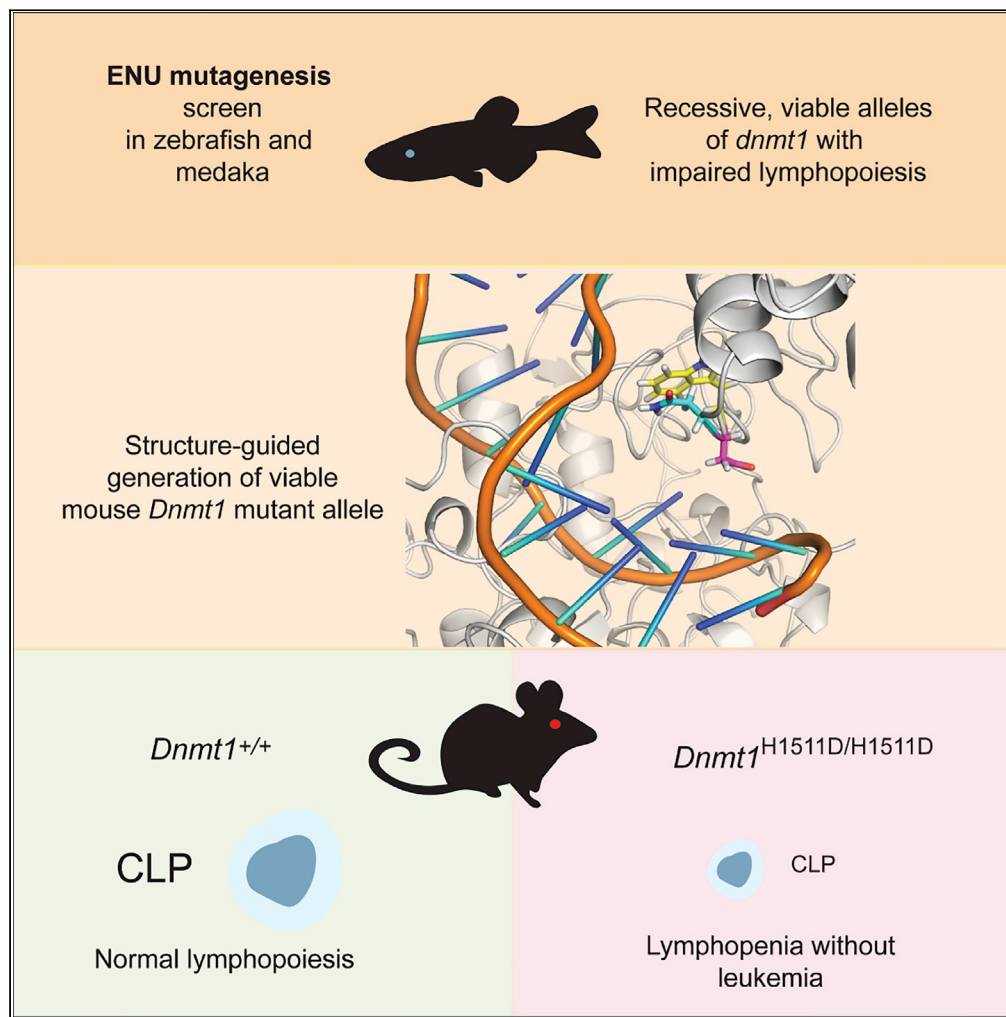


Article

Epigenetic Protection of Vertebrate Lymphoid Progenitor Cells by *Dnmt1*



Norimasa
Iwanami, Kohei
Takeshita, Divine-
Fondzenyuy
Lawir, ...,
Yasushige
Yonezawa,
Michael Schorpp,
Thomas Boehm

iwanami@cc.utsunomiya-u.ac.
jp (N.I.)
boehm@ie-freiburg.mpg.de
(T.B.)

HIGHLIGHTS

Genetic screens identified
recessive viable missense
alleles of *dnmt1* in teleosts

A viable mouse *Dnmt1*
mutant generated by
structure-guided
precision mutagenesis

Missense mutations
distort the catalytic pocket
and reduce enzymatic
activity

DNA hypomethylation
consistently affects
development of the
lymphoid lineage

Iwanami et al., iScience 23,
101260
July 24, 2020 © 2020 The
Author(s).
[https://doi.org/10.1016/
j.isci.2020.101260](https://doi.org/10.1016/j.isci.2020.101260)

Article

Epigenetic Protection of Vertebrate Lymphoid Progenitor Cells by *Dnmt1*

Norimasa Iwanami,^{1,9,*} Kohei Takeshita,² Divine-Fondzenyuy Lawir,^{1,10} Isao Suetake,³ Shoji Tajima,³ Katarzyna Sikora,¹ Inês Trancoso,¹ Connor ÓMeara,¹ Iliana Siamishi,¹ Yousuke Takahama,^{4,5} Makoto Furutani-Seiki,⁶ Hisato Kondoh,⁷ Yasushige Yonezawa,⁸ Michael Schorpp,¹ and Thomas Boehm^{1,11,*}

SUMMARY

DNA methylation is a universal epigenetic mechanism involved in regulation of gene expression and genome stability. The DNA maintenance methylase DNMT1 ensures that DNA methylation patterns are faithfully transmitted to daughter cells during cell division. Because loss of DNMT1 is lethal, a pan-organismic analysis of DNMT1 function is lacking. We identified new recessive *dnmt1* alleles in medaka and zebrafish and, guided by the structures of mutant proteins, generated a recessive variant of mouse *Dnmt1*. Each of the three missense mutations studied here distorts the catalytic pocket and reduces enzymatic activity. Because all three DNMT1 mutant animals are viable, it was possible to examine their phenotypes throughout life. The consequences of genome-wide hypomethylation of DNA of somatic tissues in the *Dnmt1* mutants are surprisingly mild but consistently affect the development of the lymphoid lineage. Our findings indicate that developing lymphocytes in vertebrates are sensitive to perturbations of DNA maintenance methylation.

INTRODUCTION

DNA methylation is a key mechanism of epigenetic control that is required for development and survival by regulating gene expression and genome stability (Bergman and Cedar, 2013; Smith and Meissner, 2013). Methylation of cytosines in DNA is established by *de novo* methylases Dnmt3a and Dnmt3b, whereas its propagation after DNA replication and repair depends on the maintenance of methylase Dnmt1. In the latter process, Dnmt1 recognizes the hemimethylated DNA duplex and copies the methylation pattern of the parental strand to the newly synthesized DNA strand. Considerable information is available concerning the structure and function of the various domains of the Dnmt1 protein, providing the basis to interpret the functional consequences of structural variants (Cheng et al., 2015; Song et al., 2012; Takeshita et al., 2011; Zhang et al., 2015). In accordance with its central cellular function (Smith and Meissner, 2013), mice lacking *Dnmt1* die at around day 9.5 of embryonic development (Li et al., 1992). Similarly, mutations of zebrafish *dnmt1* predicted to impair the function of the catalytic domain die at 8 days postfertilization (dpf) (Anderson et al., 2009). This early embryonic lethality conceals a possible tissue-specific function of this protein in the adult organism.

Lymphocytes represent the cellular underpinning of the adaptive immune system and are, as all blood lineages, descendants of hematopoietic stem cells. The molecular basis of their differentiation, proliferation, selection, and maturation has been extensively studied (Rothenberg et al., 2016; Cumano et al., 2019; Clark et al., 2014; Medvedovic et al., 2011), particularly with respect to the functional roles of certain transcription factors. By contrast, the intricacies of the epigenetic control of lymphopoiesis are only beginning to be explored. Nonetheless, previous work has indicated that DNA methylation is important for normal hematopoietic development (Jeong and Goodell, 2014; Guillamot et al., 2016; Cedar and Bergman, 2011); for instance, low levels of Dnmt1 protein are associated with failure of normal lymphocyte development (Broske et al., 2009) and leukemogenesis (Gaudet et al., 2003). However, so far, hypomorphic germline mutations compatible with a normal life span and selective failure of lymphoid differentiation have not been described in vertebrates.

Here, using a forward genetics approach with medaka and zebrafish, we have identified missense mutations in functionally important regions of *dnmt1* that result in reduced enzymatic activities. The

¹Department of Developmental Immunology, Max Planck Institute of Immunobiology and Epigenetics, Stuebeweg 51, 79108 Freiburg, Germany

²RIKEN SPring-8 Center, Sayo, Hyogo 679-5148, Japan

³Laboratory of Epigenetics, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita-shi, Osaka 565-0871, Japan

⁴Thymus Biology Section, Experimental Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

⁵Division of Experimental Immunology, Institute of Advanced Medical Sciences, University of Tokushima, 3-18-15 Kuramoto, Tokushima 770-8503, Japan

⁶Systems Biochemistry in Pathology and Regeneration, Yamaguchi University Graduate School of Medicine, Ube, Yamaguchi 755-8505, Japan

⁷Faculty of Life Sciences, Kyoto Sangyo University, Motoyama, Kamigamo, Kita-ku, Kyoto 603-8555, Japan

⁸High Pressure Protein Research Center, Institute of Advanced Technology, Kindai University, 930 Nishimitani, Kinokawa, Wakayama 649-6493, Japan

⁹Present addresses: Center for Bioscience Research and Education, Utsunomiya University, Utsunomiya, Tochigi 321-8505, Japan

¹⁰Present addresses: Institute of Zoology, Developmental Biology Unit, University of Cologne, Zulpicher STrasse 47B, 50674 Cologne, Germany

¹¹Lead Contact

Continued



identification of the structural consequences of these missense mutations has allowed us to generate a mouse model of a recessive and viable *Dnmt1* missense mutation that also gives rise to an enzyme of reduced activity. For the first time, these viable mutations provide a pan-organismic view of DNMT1 function in vertebrates that are separated by several million years of independent evolution. Remarkably, in all three species, DNMT1 mutations are associated with impaired lymphoid development. This demonstrates that hematopoietic precursors poised to feed the lymphoid lineages are uniquely sensitive to perturbations of the DNA methylation process. We thus conclude that an ancient epigenetic control mechanism was deployed to enable the development of lymphocytes, which represents a key innovation in the vertebrate immune system (Boehm and Swann, 2014).

RESULTS

Previously, we have conducted large-scale forward genetic screens in zebrafish (Boehm et al., 2003; Iwanami et al., 2016; Schorpp et al., 2006), and medaka (Furutani-Seiki et al., 2004; Iwanami et al., 2004, 2008, 2009) aimed at identifying mutations affecting vertebrate T cell development. As a result, several dozens of mutants exhibiting a noticeable reduction of developing T cells in the thymus in early larval development were identified. Here, we describe the first two viable recessive alleles of teleost *dnmt1*, overcoming the previous limitations on phenotypic analyses imposed by lethal *dnmt1* alleles.

A Zebrafish *dnmt1* Mutation Specifically Affecting Lymphopoiesis

In line IY071, the extent of T cell development in the thymus at 5 days after fertilization (dpf) is severely reduced, as determined by whole mount RNA *in situ* hybridization (Iwanami et al., 2016). The mutation in the IY071 line was identified as a missense mutation (T to G at nucleotide position 3:54,337,492 [GRCz10]) in *dnmt1* causing the substitution of the asparagine residue 1391 to lysine (N1391K) (Iwanami et al., 2016) (Figure 1A). The amino acid replacement occurred in a region exhibiting an exceptional degree of evolutionary conservation (Figure 1A). In the mouse protein, this residue is equivalent to N1510 and is situated in the target recognition domain (TRD) of Dnmt1 (Song et al., 2012). The mutants of both sexes reached adulthood, and apart from a slight size difference (Figure 2A), fish appeared grossly normal. Adult fish exhibited impaired development of T cells (Figure 2B) and B cells (Figure 2C), as determined by RNA *in situ* hybridization and subsequent qPCR-based gene expression analyses in whole kidney marrow cells; other hematopoietic cell lineages appear to be less affected (Figure 2D). Next, we examined whether the hematopoietic environment in *dnmt1* mutants was capable of supporting normal blood cell development. To do this, we transplanted whole kidney marrow cells of wild-type *ikzf1:egfp* transgenic zebrafish into non-irradiated adult *dnmt1* mutants as recipients; the reporter transgene marks the lymphocyte lineage in zebrafish (Bajoghli et al., 2009; Hess and Boehm, 2012). When assayed at 8 days after transplantation, donor cell engraftment was detectable in the thymus (data not shown) and kidney (Figure S1A) of 50% (7/14) of *dnmt1* mutant recipients; by contrast, donor cells were not found in wild-type and heterozygous recipients (0/16; difference significant at $p = 0.002$, X^2 -test), presumably as a consequence of MHC mismatches between donor and recipient. Collectively, these results suggest that the non-hematopoietic microenvironment of *dnmt1* mutants is capable of supporting the differentiation of wild-type lymphocytes. We attribute the failure of engraftment in some *dnmt1* mutants to the residual presence of lymphocytes, which is evident from lymphocyte-specific transcripts in kidney marrow cells (Figure S1B); this indicates that the block of lymphocyte differentiation in the *dnmt1* zebrafish mutant is incomplete. However, it is also possible that the variable outcome of the transplantation experiments is due to different fitness levels of mutant recipient and wild-type donor progenitor cells. When kidney marrow cells of successfully reconstituted *dnmt1* mutants were secondarily transplanted into *c-myb* mutants (which lack both hematopoietic progenitor and mature blood cells and thus serve as universal recipients (Hess et al., 2013)), EGFP⁺ cells were found to be present in the kidney (Figure S1C) and thymus (Figure S1D) of secondary recipients. This finding suggests that long-term reconstituting hematopoietic precursor cells survive in the *dnmt1* mutant recipients and supports the notion of the hematopoietic origin of failing lymphopoiesis in *dnmt1* mutants.

A Medaka *dnmt1* Mutation Specifically Affecting Embryonic Lymphopoiesis

The *gyokuro* (*gkr*, strain *j48-12B*) mutant line exhibited a reduction of *rag1*-expressing immature thymocytes at stage 35 (equivalent to 5 days after fertilization [dpf]) (Iwamatsu, 2004) in whole mount RNA *in situ* hybridization assays (Iwanami et al., 2004) (Figure 3A); this contrasts with normal expression of *foxn1*, a marker of

*Correspondence:
iwanami@cc.utsunomiya-u.ac.jp (N.I.),
boehm@ie-freiburg.mpg.de (T.B.)
<https://doi.org/10.1016/j.isci.2020.101260>

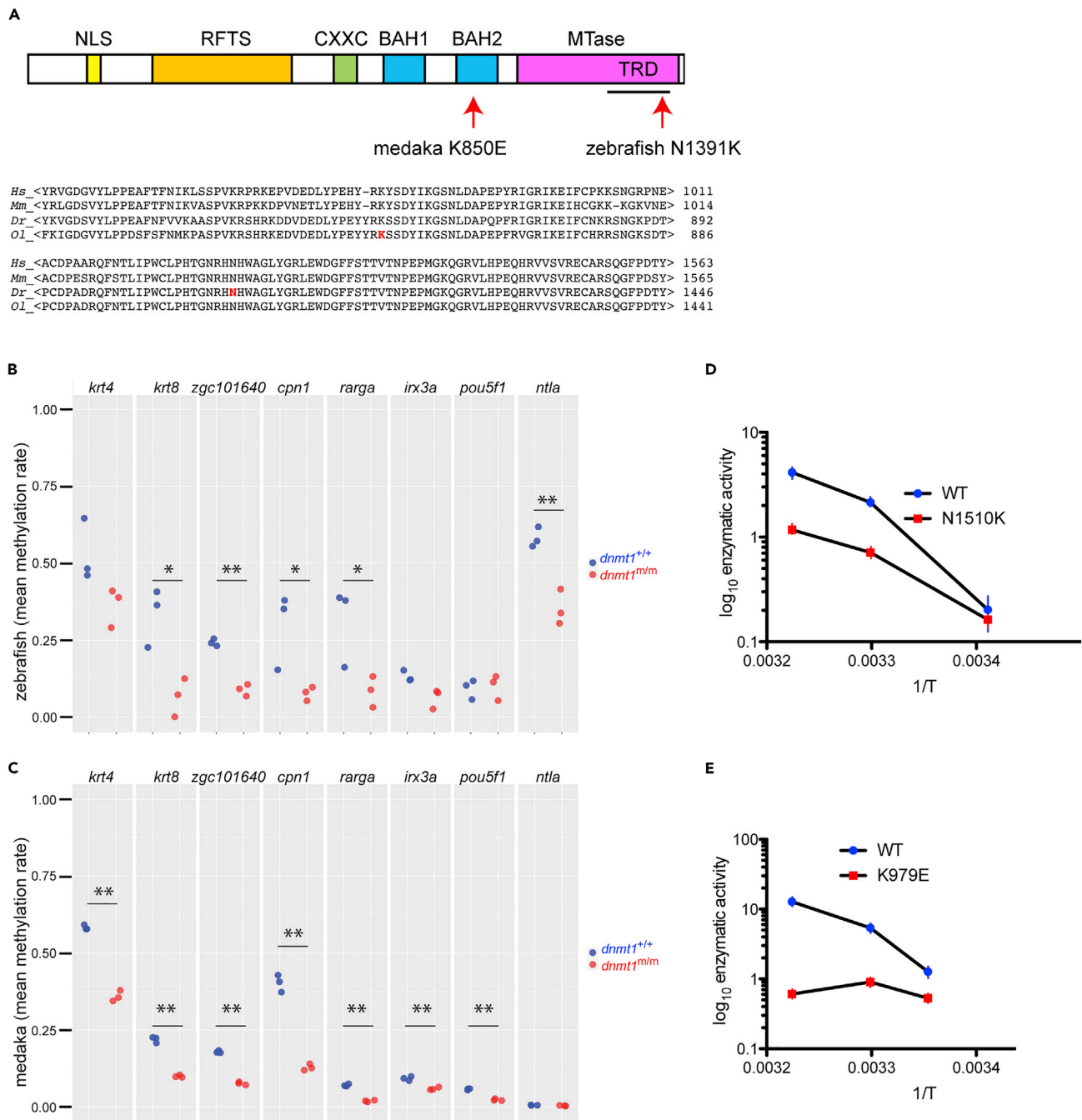


Figure 1. Identification of Missense Mutations in Medaka and Zebrafish *dnmt1* Genes

(A) Schematic structure of the DNMT1 protein with relevant protein domains indicated; NLS, nuclear localization signal; RFTS, replication foci targeting site; CXXC, cysteine-rich domain; BAH, bromo-adjacent homology domains 1 and 2; MTase, catalytic domain. (Arrow) Approximate position of the amino acid replacement in the target recognition domain, TRD (top panel). Alignments of amino acid sequences of human (*Hs*), mouse (*Mm*), zebrafish (*Dr*), and medaka (*Ol*) *dnmt1* proteins at the region close to the medaka (red, middle panel) and zebrafish (red, bottom panels) missense mutations.

(B) Average methylation levels of CpGs in indicated CpG islands of total body DNA in zebrafish *dnmt1* mutants and siblings at 18 dpf. **p* < 0.05, ***p* < 0.01; *t* test, two-tailed.

(C) Average methylation levels of CpGs in indicated CpG islands of total body DNA in medaka *dnmt1* mutants and siblings at 18 dpf. **p* < 0.05, ***p* < 0.01; *t* test, two-tailed.

(D) Enzymatic activities (mol CH₃ incorporated/h/mol of Dnmt1) of mouse wild-type and N1510K variant Dnmt1 proteins as a function of temperature *T* measured in Kelvin displayed as Arrhenius plots; mean ± SEM (*n* = 3).

Figure 1. Continued

(E) Enzymatic activities (mol CH₃ incorporated/h/mol of Dnmt1) of mouse wild-type and K979E variant Dnmt1 proteins as a function of temperature T measured in Kelvin displayed as Arrhenius plots; mean ± SEM (n = 3). See also Figures S1, 2, and 3.

thymic epithelium (Schorpp et al., 2002) (Figure 3B). A combinatorial approach consisting of linkage analysis and genome sequencing (Table S1) identified a missense mutation (Figure 3C) in the second BAH domain of *dnmt1*, K850E (Figure 1A). This region is well conserved in vertebrate DNMT1 enzymes; in the mouse protein, this residue is equivalent to K979E (Figure 1A). Detailed characterization of the *gkr* mutant suggested that the *dnmt1* missense mutation predominantly affects the lymphoid lineage of mutant embryos, whereas myeloid and erythroid differentiation pathways—as determined by RNA *in situ* hybridization using *mpx*- and *gata1*-specific probes—appear to be unimpaired (Figures 3D and 3E). By contrast, adult hematopoiesis was found to be essentially normal in the *gkr* mutants. This is evident from the normal cellularity of the thymus as revealed by analysis of *gkr* mutants transgenic for a *rag1:egfp* reporter (Li et al., 2007) (Figure 3F) and also in histological studies and RNA *in situ* hybridizations of thymus (Figure 3G) and kidney (Figure 3H) with a *rag1*-specific probe; moreover, analysis of kidney marrow cells by flow cytometry (Figure 3I) showed no abnormalities. Collectively, these results suggest that aberrations in the larval stages of T cell development are followed by an appreciable recovery of lymphopoiesis in adult medaka *dnmt1* mutants, which exhibit a normal life-span.

DNA Hypomethylation in *dnmt1* Mutants

The non-conservative amino acid replacements in zebrafish and medaka *dnmt1* mutants (Figure 1A) predict reduced enzymatic activities for the altered proteins. This was examined in two ways. First, we used bisulfite sequencing using DNA extracted from the bodies of 18 dpf wild-type and mutant zebrafish and medaka fish. We focused our analysis on eight CpG islands, which were previously shown to be differentially methylated between sperm and oocyte DNA of zebrafish (five CpG islands were hypermethylated and three were hypomethylated in sperm) (Potok et al., 2013). In zebrafish, the average methylation levels of six CpG islands were substantially lower in mutants, whereas little change was observed in two CpG islands already exhibiting low methylation levels (Figure 1B). When the eight orthologous CpG islands in the medaka genome were examined, we found that the average methylation levels of seven CpG islands were also significantly lower in medaka mutants (Table S2; Figure 1C). These results suggest that the *dnmt1* missense mutations in zebrafish and medaka identified here affect the extent of maintenance of DNA methylation in similar ways and eventually cause hypomethylation in somatic tissues. Second, we studied the consequences of replacing the two evolutionarily conserved residues on the *in vitro* activities of the mouse Dnmt1 enzyme. To this end, we mutated the mouse *Dnmt1* gene to generate the orthologous mutants (K979E for the medaka mutation and N1510K for the zebrafish mutation); we then expressed the wild-type and the two mutant forms in a baculovirus system and used the purified proteins for *in vitro* methylation assays using a hemimethylated DNA substrate. Both mutants exhibited drastically reduced maintenance methylation activities; in addition, the K979E mutant exhibited a notable temperature sensitivity at 37°C (Figures 1D and 1E).

Structural Basis of Reduced Enzymatic Activities of Dnmt1 Mutants

Since the recessive missense mutations in zebrafish and medaka *dnmt1* genes identified here represent the only viable vertebrate mutant *dnmt1* alleles so far described, we set out to determine the structural basis of impaired maintenance methylation. We first examined the available information on the structure of the mouse Dnmt1 protein (Cheng et al., 2015; Song et al., 2012; Takeshita et al., 2011; Zhang et al., 2015) to locate the corresponding residues that are mutated in zebrafish and medaka *dnmt1* proteins. This analysis was facilitated by the high degree of evolutionary conservation of the regions containing the two mutations (Figure 1A). Remarkably, we found that the medaka mutation in the BAH2 domain is positioned in close proximity to the zebrafish mutation in the TRD of the enzyme. In the mouse Dnmt1 protein, residues Y980 and H1511 (both situated very close to the altered residues, *i.e.*, equivalent to mouse residues K979 and N1510, respectively) directly interact with each other (Figure 4A) (Song et al., 2012; Takeshita et al., 2011). These structural features suggested a common mechanism leading to the impairment of DNMT1s enzymatic activity and prompted us to explore the structural consequences of the two mutations by molecular dynamics simulation. When the zebrafish mutation is modeled by replacing asparagine 1510 with a lysine residue, new hydrogen bonds are formed with bases of the substrate DNA near the hemimethylated site (Figures 4B and 4C); we propose that these changes interfere with the processivity of the

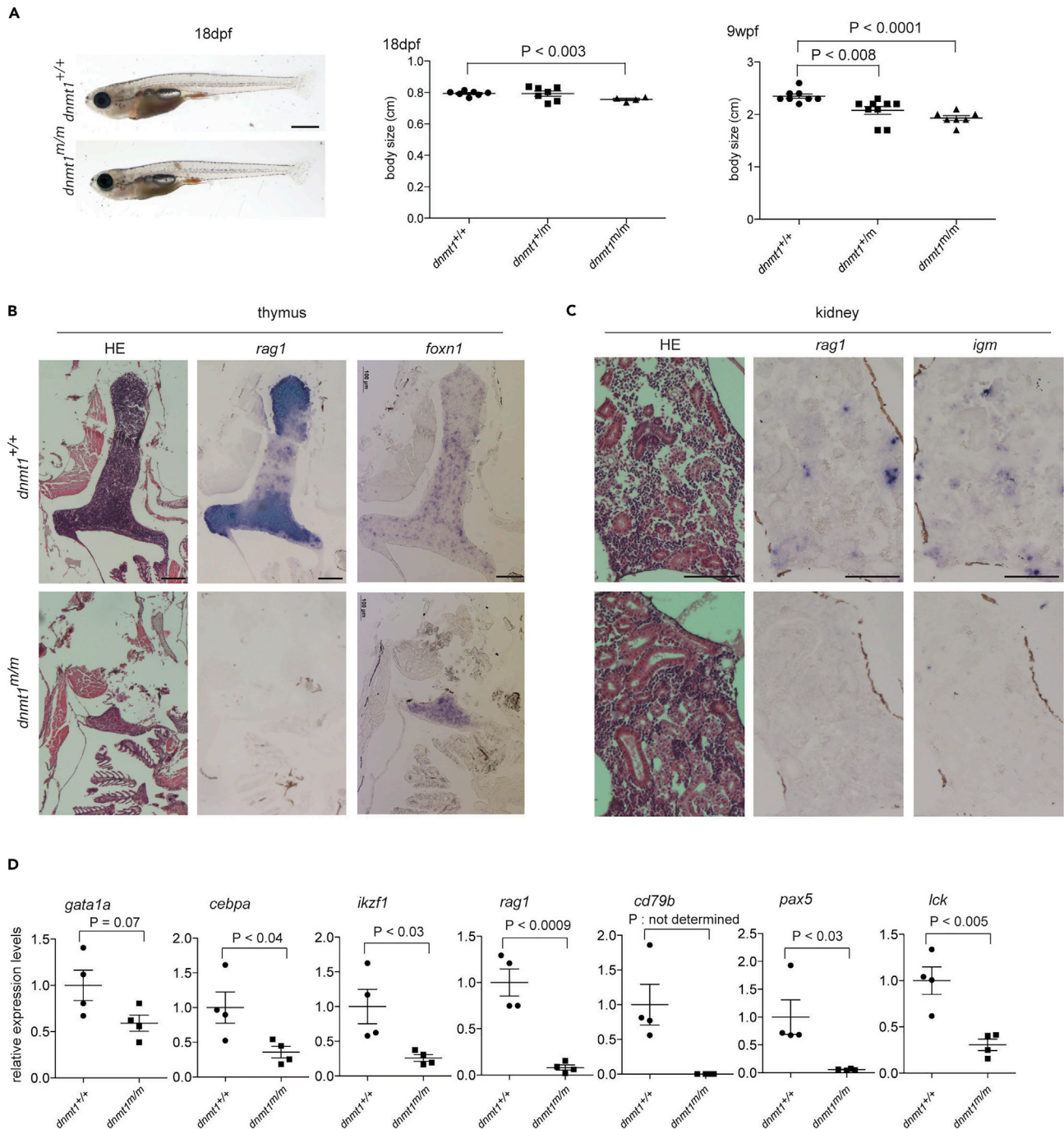


Figure 2. Both T Cell and B Cell Development Are Predominantly Affected in Zebrafish Adult *dnmt1* Mutant

(A) Gross appearances of *dnmt1* mutants at 18 dpf. Scale bar, 1mm (left panel). Standard lengths of *dnmt1* mutant fish at 18 dpf. *dnmt1*^{+/+} (n = 7), *dnmt1*^{+/m} (n = 7), *dnmt1*^{m/m} (n = 4) (middle panel). Standard lengths of *dnmt1* mutant fish at 9 wpf. *dnmt1*^{+/+} (n = 8), *dnmt1*^{+/m} (n = 9), *dnmt1*^{m/m} (n = 7) (right panel). (B) Hematoxylin/eosin (HE) staining and RNA *in situ* hybridization using *rag1* and *foxn1* probes of paraffin sections of thymic region of *dnmt1* mutant at 9 wpf. Scale bars, 100 μ m.

(C) H&E staining and RNA *in situ* hybridization using *rag1* and *igm* probes of paraffin sections of kidney region of *dnmt1* mutant at 9 wpf. Scale bars, 100 μ m. (D) Gene expression levels for indicated genes in whole kidney marrow samples of *dnmt1*^{+/+} and *dnmt1*^{m/m} at 18 wpf as assessed by qPCR; mean \pm SEM. The values for wild-type controls are set to 1.0.

See also [Figures S1, 2, and 3](#).

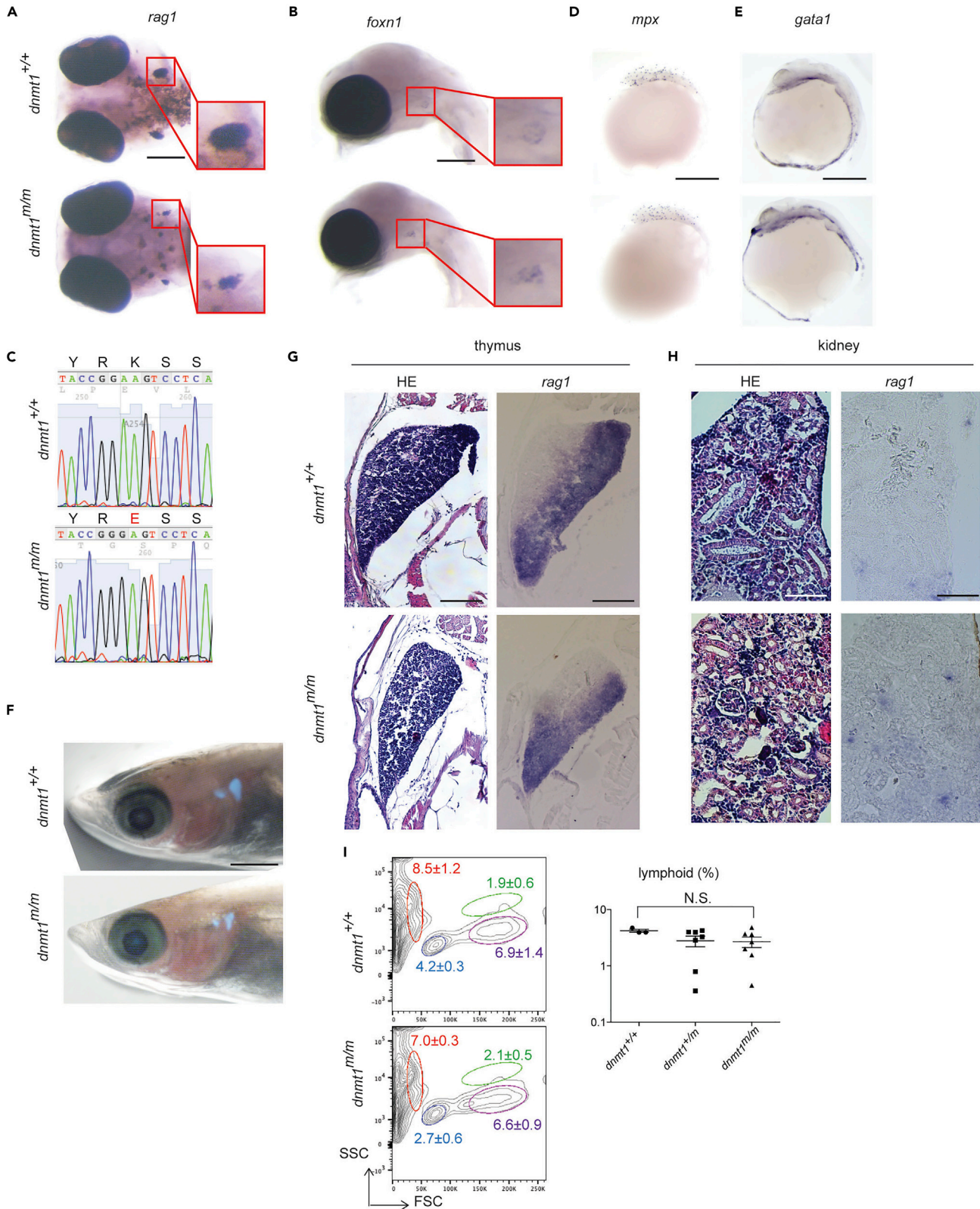


Figure 3. Impaired Lymphocyte Development in Medaka *dnmt1* Mutant

- (A) Whole mount RNA *in situ* hybridization pattern in *gkr* mutant using a *rag1* probe at stage 35 (5 dpf). Ventral view. Right panels display enlarged thymic areas marked in left panels. Scale bar, 200 μ m.
- (B) Expression of the thymic epithelia cell marker *foxn1* in *dnmt1* mutant at stage 35. Right panels display enlarged thymic areas of left panels. Scale bar, 200 μ m.
- (C) Representative sequence traces indicating the A > G transversion at nucleotide position 8,624,154 (MEDAKA1 in Ensembl release 72) on chromosome 8. The conceptual translation in three-letter code is also shown.
- (D) Expression of macrophage marker *mpx* in *dnmt1* mutants at stage 22 (1.5 dpf). Scale bar, 500 μ m.
- (E) Expression of erythroid marker *gata1* in *dnmt1* mutant at stage 25 (2 dpf). Scale bar, 500 μ m.
- (F) Expression of *rag1:egfp* transgene in *dnmt1* mutant at 7 wpf. Scale bar, 2 mm.
- (G) H&E staining and RNA *in situ* hybridization using a *rag1* probe on paraffin sections of the thymic region of *dnmt1* mutant at 7 wpf. Scale bars, 100 μ m.
- (H) H&E staining and RNA *in situ* hybridization using a *rag1* probe on paraffin sections of the kidney region of *dnmt1* mutant at 7 wpf. Scale bars, 50 μ m. For (F)-(H), figures are representative of 1 fish per genotype.
- (I) Flow cytometric profiles of whole kidney marrow cells of *dnmt1* mutant at 7 wpf. Light scatter profiles with percentages of erythroid (red), lymphoid (light blue), myeloid (green), and precursor (purple) indicated; mean \pm SEM is shown; *dnmt1*^{+/+} (n = 3), *dnmt1*^{m/m} (n = 7) (left panels). Percentages of lymphoid population, mean \pm SEM is shown; *dnmt1*^{+/+} (n = 3), *dnmt1*^{+/m} (n = 7), *dnmt1*^{m/m} (n = 7) (right panel). N.S., not significant. Panels A, B, D, and E representative of at least 8 fish per genotype. See also [Figure S3](#).

Dnmt1 enzyme. When the structure of zebrafish *dnmt1* protein is modeled on the basis of the mouse Dnmt1 crystal structure ([Figure S2A](#)), the emergence of new hydrogen bonds between K1391 and the substrate DNA is again observed ([Figure S2B](#)), suggesting that the model of the zebrafish protein calculated from molecular dynamics simulation is an accurate representation of its three-dimensional structure. Further simulations aimed at examining the structural consequences of the medaka mutation indicated that replacing lysine 979 with glutamic acid in mouse Dnmt1 causes a structural rearrangement that repositions Y980 next to D1438; as a result, the substrate DNA is dislodged from the catalytic cleft ([Figures 4D and 4E](#)). These structural changes are mirrored in the model of the medaka K850E mutant protein ([Figures S3A and S3B](#)). Collectively, the structural studies explain the hypomorphic nature of the two *dnmt1* missense mutations through specific structural rearrangements of the catalytic site of the enzyme.

Embryonic Lethality of N1510K and K979E Mouse Mutants

The functional and structural studies reported above implied that the introduction of the N1510K and K979E mutations into the mouse germline should cause an impaired activity of the mutant Dnmt1 proteins *in vivo*. To confirm this, we generated the appropriate nucleotide changes in the mouse *Dnmt1* gene by CRISPR/Cas9-mediated nucleotide replacement and examined the phenotype of the resulting mutant lines. Mice homozygous for either mutation exhibited embryonic lethality, similar to *Dnmt1*-deficient mice ([Li et al., 1992](#)). The phenotype of heterozygous mutants was found to be indistinguishable from that of wild-type siblings, indicating the recessive nature of the two mutant alleles. Based on the non-linear temperature-dependent loss of enzymatic activities of mutant proteins ([Figures 1D and 1E](#)), we propose that owing to the higher body temperature of mice ($\sim 37^\circ$) compared with fish ($\sim 28^\circ\text{C}$), the DNA methylase activities of both mutants fall below the threshold of activities required to sustain normal mouse development.

Structure-Guided Generation of a Viable *Dnmt1* Mutant

Although the orthologous zebrafish and medaka mutations in *Dnmt1* were not compatible with extended survival of homozygous mutant mice, our structural studies nevertheless provided a useful starting point to explore other mutations in Dnmt1. To this end, we focused on the H1511 residue, which was shown to be important in configuring the catalytic site of Dnmt1 ([Song et al., 2012](#); [Takeshita et al., 2011](#)) ([Figure 4A](#)). Here, we report on a mutation that converts histidine 1511 to aspartic acid, H1511D. The outcome of molecular dynamics simulation indicated that the H1511D mutation causes the side chain of N1510 to reposition near to the indol group of tryptophane 1512 ([Figure 5A](#)). As expected, this structural rearrangement in the catalytic center is accompanied by reduced enzymatic activity *in vitro*; the H1511D mutation reduces maintenance activity to about 50% of the wild-type version ([Figure 5B](#)). The residual activity of the H1511D mutant is higher than those observed for the N1510K and K979E mutations (about 30% and 5%, respectively) ([Figures 1D and 1E](#)); this result indicates that we achieved our goal of generating a hypomorphic version of the enzyme with intermediate level of activity. Next, we used CRISPR/Cas9 mutagenesis to introduce the H1511D mutation into the germline of mice. Mice homozygous for the H1511D mutation are viable, providing *in vivo* evidence that the phenotypic consequences of the H1511D mutation are less severe than those of the N1510K and K979E mutations. Several phenotypic features of H1511D homozygous

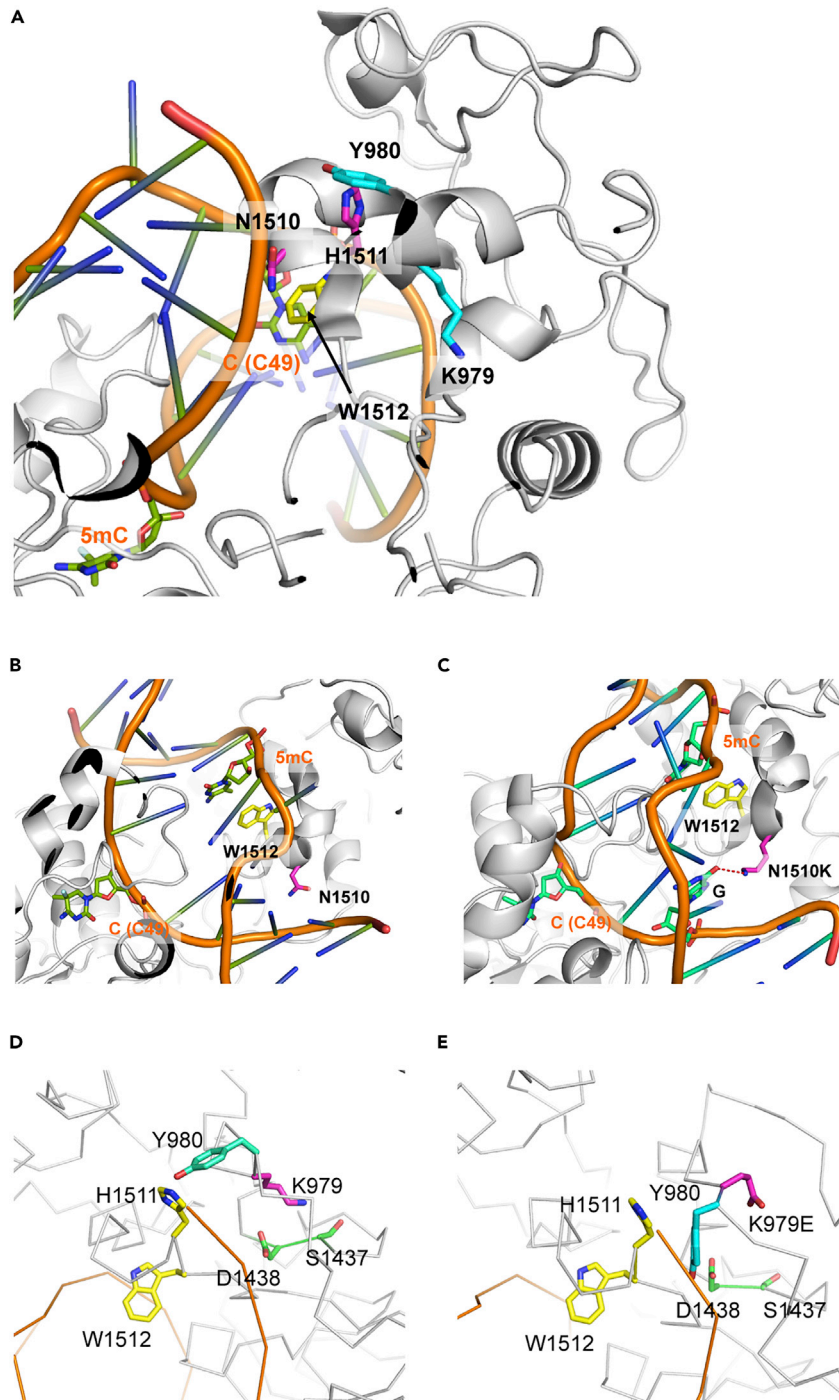


Figure 4. Structure of Wild-type and Variant Dnmt1 Proteins

(A) Structure of the catalytic pocket of the mouse *Dnmt1* protein (Song et al., 2012), highlighting the spatial proximity of the two mutated residues; zebrafish missense mutation corresponding to mouse residue N1510 and medaka missense mutation corresponding to mouse residue K979. The hemimethylated DNA substrate is indicated in orange, with the position of the 5-methylcytosine (5mC) indicated.

(B and C) Comparison of the catalytic pockets of mouse wild-type (B) and N1510K mutant (C) *Dnmt1* proteins as determined by molecular dynamics simulation; note the formation of hydrogen bonds between the lysine residue and the guanine (G) base (red dotted line).

Figure 4. Continued

(D and E) Comparison of the catalytic pockets of mouse wild-type (D) and K979E mutant (E) Dnmt1 proteins as determined by molecular dynamics simulation; note the distortion of the catalytic pocket in the mutant variant. See also [Figure S3](#).

mutant mice are notable. The animals are smaller and have a reduced body mass (2/3 of normal weight at 7 weeks of age); these findings hint at pleiotropic effects of the mutation that do, however, not compromise viability. Considering the observations made in the two teleost models, we focused on the hematopoietic system of H1511D/H1511D mice. With respect to the cellularity in lymphoid organs, we found that the number of bone marrow cells is reduced to half, whereas the numbers of thymocytes and splenocytes are reduced even further to about 40% of wild-type levels ([Figure 5C](#)). In stark contrast to the previously described model of low Dnmt1 activity ([Broske et al., 2009](#)), the Dnmt1^{H1511D/H1511D} mice did not develop hematological malignancies. Both male and female homozygous mutants proved to be fertile, indicating that differentiation of gametes in the mutants is unimpaired. We conclude that the H1511D mutation represents a unique mouse model with which enables us to examine the physiological role of Dnmt1 in development and in adult organ function.

Hematopoietic Abnormalities in Dnmt1^{H1511D} Mutant Mice

Among the individual subsets of hematopoietic precursors in the bone marrow ([Figures S4A, S5A, and S5B](#)), we noted several-fold reductions in the numbers of HSC, MPP1, and MPP4 subsets but little change of MPP2 and MPP3 fractions ([Figure 5D](#)); interestingly, whereas near-normal numbers of CMP, GMP, and MEP precursors were found, CLPs were reduced about 100-fold ([Figures 5E and S5A](#)). Collectively, these findings indicate that the H1511D mutation predominantly affects the differentiation of the lymphoid lineage, reminiscent of the observations in the medaka and zebrafish mutants.

In order to examine the cell-intrinsic nature of the hematopoietic abnormalities in Dnmt1^{H1511D} mutant mice, we carried out competitive bone marrow transplantation experiments. Bone marrow cells of wild-type and Dnmt1-mutant cells calibrated to contain 100 HSCs each were mixed and transplanted into irradiated wild-type recipients and their hematopoietic compartment analyzed after 12 weeks. The results indicated a poor contribution of mutant cells to the reconstituted hematopoietic compartment of the host animals. In the bone marrow of reconstituted mice, we found that wild-type cells outnumbered mutant cells by three orders of magnitude; prepro-B cells, pro-B cells, and pre-B cells were particularly affected ([Figure 6A](#)), in line with the differential reductions seen in the marrow of the mutant donors ([Figures S4A and S5B](#)). In the thymus of reconstituted mice, CD4/CD8 double-positive (DP) cells were almost entirely of wild-type origin, whereas DN1 thymocytes ([Figure 6D](#)) and mature CD4⁺ and CD8⁺ single-positive (SP) cells ([Figure 6B](#)) exhibited greater competitive potential; again, this outcome is reminiscent of the situation of the thymus in mutant donors ([Figure S5C](#)). In the spleen of reconstituted mice, the particular sensitivity of the B cell lineage to the H1511D mutation was readily detectable; of note, mutant CD4⁺ cells were less competitive than CD8⁺ cells ([Figure 6C](#)). In splenocytes of mutant donors, lymphocyte subsets are reduced to equal extent ([Figure S5D](#)). Interestingly, the contribution of mutant granulocytes and neutrophils to the reconstituted peripheral compartment was also small ([Figure 6C](#)); this suggests that although the lymphoid lineage is particularly sensitive to the effects of the H1511D mutation, the hematopoietic aberrations are not entirely lymphoid specific. Indeed, in secondary transplantation experiments no mutant cell engraftment was detectable (data not shown), suggesting that the H1511D mutation has detrimental effects also on the maintenance and/or the fitness of HSCs, at least when subjected to the transplantation procedure and/or when they are under competitive pressure *in vivo*.

Next, we examined the methylation status of hematopoietic precursor cells. Given the pervasive effects of the H1511D mutation, we compared the extent of methylation in Lin⁻ cells of the bone marrow. Significantly reduced methylation levels of CpG dinucleotides were detected ([Figure 7A](#)), clustered within 38,000 differentially methylated regions (DMRs), spread across the entire genome ([Table S3](#)). We then determined gene expression profiles of HSCs, and the MPP1 and MPP4 precursor populations between wild-type and Dnmt1^{H1511D} mutant mice by RNA-seq ([Table S4](#)). The data indicated surprisingly few significant changes of generally modest degree; however, overexpression of *Fos* and *FosB* genes was consistently found in mutant progenitor subsets ([Figure 7B](#); [Table S4](#)). The higher expression levels of these two genes were associated with hypomethylation of several locus-associated DMRs in Lin⁻ cells ([Figures 7C, S6A and S6B](#)). The DMRs in the *Fos* and *FosB* genes coincide neither with active chromatin marks nor with signatures of open chromatin that are found in HSCs, MPPs, and CLP, respectively ([Sun et al., 2014a](#); [Lara-Astiaso et al., 2014](#); [Yoshida et al., 2019](#)) (see [Figure S6](#) for H3K4me1), indicating a

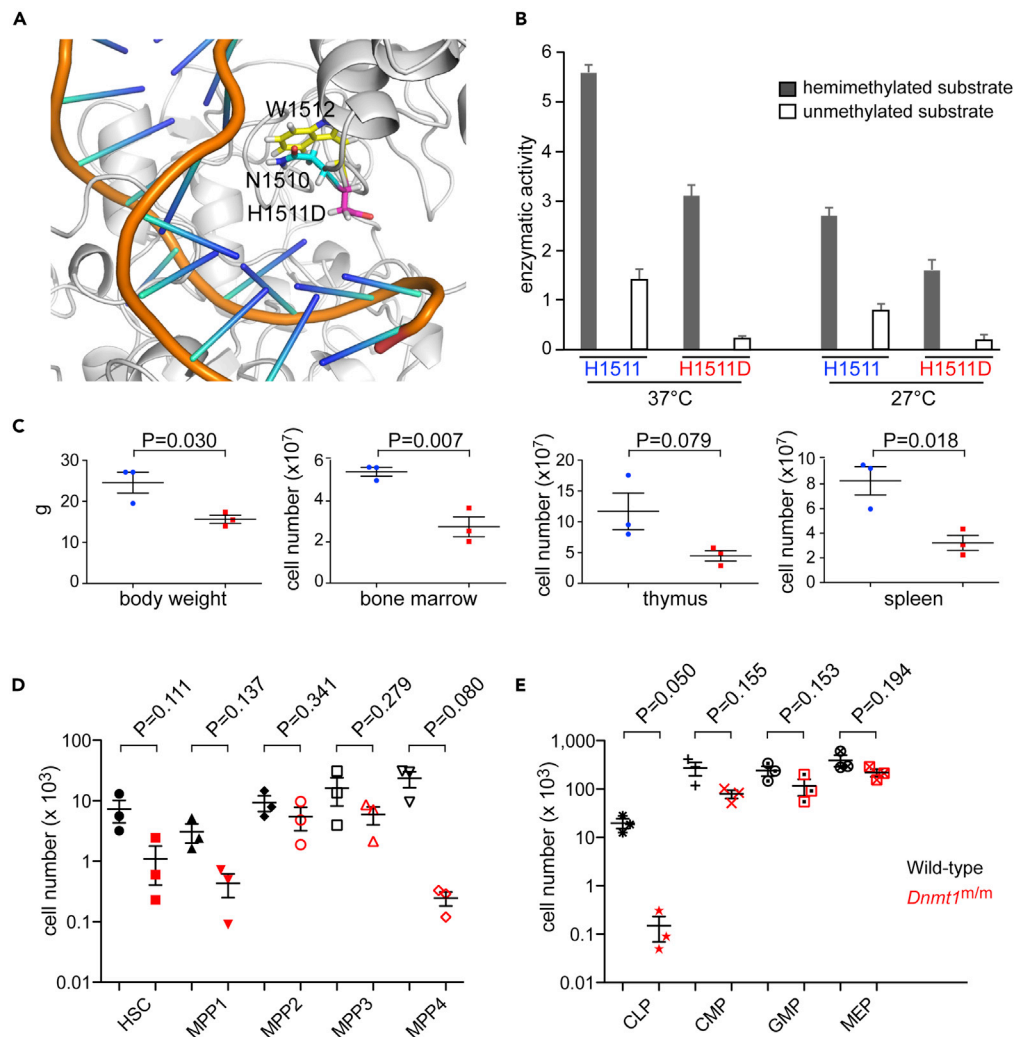


Figure 5. A Mouse Model of Impaired Dnmt1 Activity

(A) Structure of the distorted catalytic pocket of the mouse H1511D Dnmt1 variant. The hemimethylated DNA substrate is indicated in orange.

(B) Enzymatic activities (mol CH₃ incorporated/h/mol of Dnmt1) of mouse wild-type and H1511D variant Dnmt1 proteins as a function of temperature in degree Celsius.

(C) Phenotypic features of H1511D homozygous mutant in 6-week-old mice. Body weight (left panel); absolute cell numbers of bone marrow, thymus, and spleen.

(D) Absolute cell numbers of HSC and MPP precursor populations in bone marrow.

(E) Absolute cell numbers of CLP, CMP, GMP, and MEP precursor populations in bone marrow. Note the dramatic loss of the common lymphoid progenitor (CLP).

(C–E) Wild-type (blue data points) and *Dnmt1*-mutant (red data points). See also [Figures S4, 5, and 7](#).

complex relationship between hypomethylation and overexpression of these genes in mutant progenitor cells. Overexpression of *Fos* is known to selectively induce apoptosis in B cell precursors ([Imoto et al., 1996](#); [Hu et al., 1996](#)); in the H1511D mutant, we found that pre-B cells of mutant mice exhibit moderately increased death rates ([Figure 7D](#)). This effect may contribute to reduced B cell numbers, because Lin⁻ cells in the bone marrow proliferate normally ([Figures S7A and S7B](#)). By contrast, overexpression of *Fos* was shown to have no effect on T cell differentiation ([Ruther et al., 1988](#); [Fridkis-Hareli et al., 1992](#)); indeed, despite higher expression levels of *Fos* and *Fosb* in thymocytes ([Figure 7C](#)), no changes in proliferation ([Figure S7D](#)) and levels of apoptosis were observed ([Figure S7E](#)). Collectively, these findings suggest that the drastic reduction of CLPs is the main reason for the severe cytopenia in the downstream lymphocyte differentiation pathways in mutant mice; in addition, the detrimental effect of the mutation in the B cell lineage is magnified by increased B cell apoptosis.

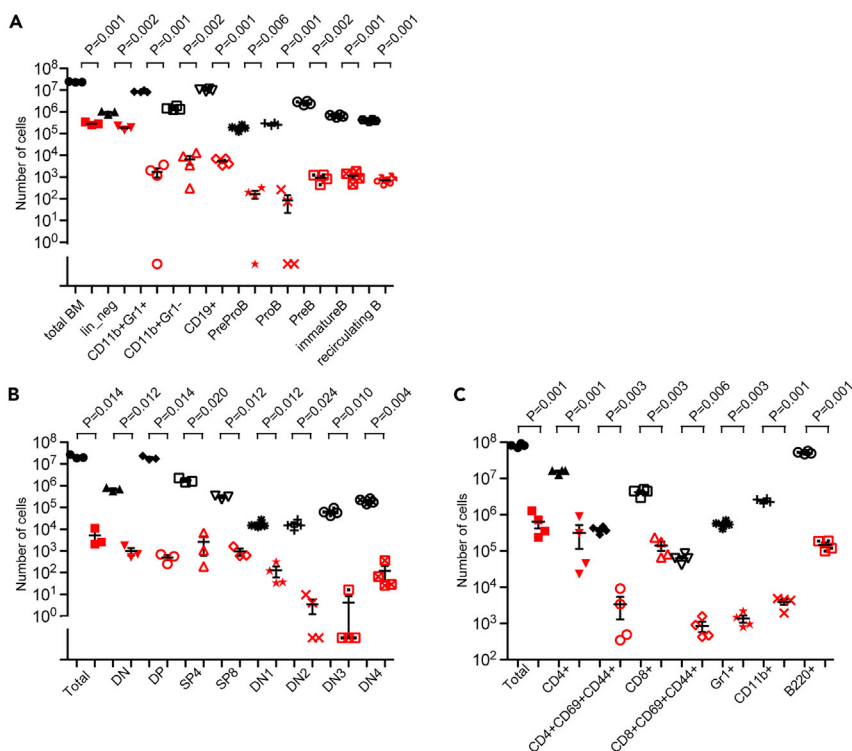


Figure 6. Impaired Fitness of Hematopoietic Precursor Populations of H1511D Homozygous Mice

The outcome of competitive transfer of allotype-tagged precursor populations was examined after 12 weeks. Results for wild-type donor cells are shown in black symbols and those for H1511D mutant hematopoietic cells in red symbols.

(A) Numbers of different cell types in the bone marrow.

(B) Numbers of different cell types in the thymus.

(C) Numbers of different cell types in the spleen.

See also [Figures S4, 5, and 7](#).

DISCUSSION

After cell division, multiple mechanisms ensure the faithful transfer of epigenetic information from the parental to the daughter cell. The DNA maintenance methylase, DNMT1, reestablishes the pattern of cytosine methylation on the newly synthesized strand after DNA replication and repair synthesis ([Bergman and Cedar, 2013](#); [Smith and Meissner, 2013](#)). Hence, loss of DNMT1 leads to epigenetic aberrations culminating in cellular lethality. From an experimental perspective, the absolute requirement of DNMT1 for cell viability hampers studies on the role of DNA maintenance methylation during development and cell differentiation and thus precludes an assessment of tissue function in adults. This is well illustrated by the early lethality associated with deletion of DNMT1 in mice ([Li et al., 1992](#)) and apparent null alleles of *dnmt1* in zebrafish ([Anderson et al., 2009](#)). In the latter case, maternal contributions of the enzyme to the developing embryo and larva can buffer the lack of zygotic activity only up to a certain developmental stage, highlighting the continuous requirement of (at least some level of) enzyme activity for development and differentiation. Harnessing the power of forward genetic screens, we identified two viable hypomorphic recessive alleles of DNMT1 in teleost fish. Orthologous amino acid replacements at the two critical positions in the catalytic site of mouse Dnmt1 protein conferred temperature sensitivity to the enzyme *in vitro*; this observation suggests an explanation for why the corresponding mouse alleles exhibited embryonic lethality. However, because these studies had identified the physiologically relevant region in the catalytic pocket, it became possible by structure-guided mutation of the adjacent histidine residue to generate a novel mouse *Dnmt1* allele (H1511D) with properties amenable to organismic studies. In humans, dominant missense mutations in *DNMT1* result in premature protein degradation and are associated with neuronal disorders ([Klein et al., 2011](#); [Winkelman et al., 2012](#); [Sun et al., 2014b](#)). Our zebrafish mutant exhibits mild haploinsufficiency with respect to body mass, but this effect is not seen in the *Dnmt1*^{+H1511D} mice; moreover, no gross behavioral abnormalities were detectable in our mutant animals. In mice, low levels of wild-type Dnmt1 protein are also associated with defects in lymphoid differentiation ([Broske et al., 2009](#)), a phenotype that partially overlaps with the

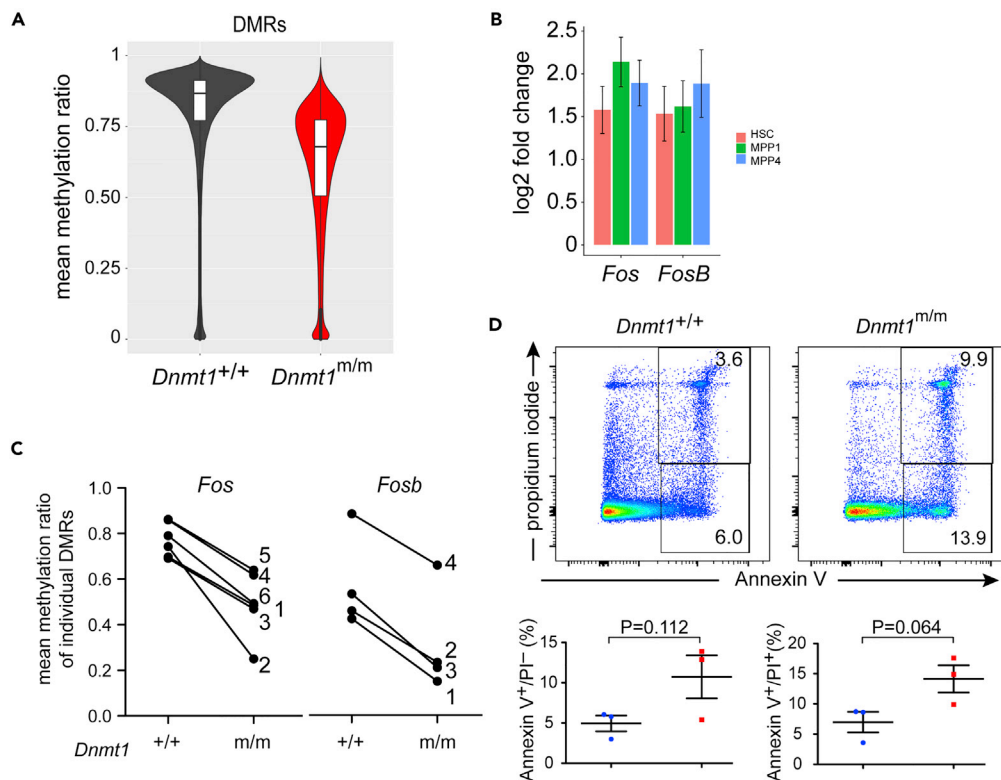


Figure 7. Characterization of the Hematopoietic Compartment of H1511D Mutant Mice

(A) Extent of methylation in DMRs in lineage-negative bone marrow precursor cells.

(B) Expression levels of *Fos* and *FosB* genes in hematopoietic precursor populations of H1511D mutants relative to controls.

(C) Mean methylation ratios of gene-associated DMRs in Lin⁻ bone marrow cells of the indicated genotypes. The numbers refer to the map positions in Figure S6.

(D) Elevated levels of apoptosis in pro-B cells of *Dnmt1*-mutant animals. Flow cytometric analysis (top panels); AnnexinV⁺/PI⁻ cells are in early apoptosis, whereas AnnexinV⁺/PI⁺ cells are in late stages of apoptosis. Quantification of results (bottom panels); wild-type (blue data points) and *Dnmt1*-mutant (red data points).

See also Figure S6.

one observed here. However, an important distinction is the fact that the low levels of *Dnmt1* lead to leukemic transformation and death at 12 weeks of age (Broske et al., 2009; Gaudet et al., 2003). Neither in our zebrafish and medaka mutants nor in the H1511D mutant mice, leukemia development has yet been detected. This feature represents an important distinction to the previous model, because it enables the analysis of hematopoietic abnormalities without the complication of concomitant malignant transformation. Collectively, the outcomes observed with the different *DNMT1* alleles support the notion of a complex genotype-phenotype relationship, encouraging further site-specific mutagenesis studies.

The identification of viable alleles in two fish species, and the subsequent structure-guided development of a viable mouse allele, allowed us to examine the physiological consequences of impaired DNA methylation during the development of species that are separated by several hundred million years of independent evolution. It is remarkable that, among the myriad cellular differentiation events underlying the development and reproduction of the vertebrate body, and the many evolutionary innovations in organismal physiology distinguishing fish and mammals, lymphoid differentiation appears to be a common target of aberrations in DNA methylation. Although this general conclusion is well supported by our present study, there are certain notable differences among the three species. For instance, with respect to adult lymphopoiesis, we find that the phenotype in zebrafish is more severe than that in medaka; however, at present, we cannot distinguish between different degrees of functional impairment of the enzyme and divergent biological features of the two species. Nonetheless, inter-individual differences in the severity of lymphopenia in adult zebrafish are a sign of (albeit insufficient) recovery

of lymphopoiesis, reminiscent of the more complete restoration in medaka mutants. It is known that larval and adult phases of zebrafish T cell development are genetically separable (Schorpp et al., 2006; Hess and Boehm, 2012; Tian et al., 2017); hence, it is possible that adult lymphoid progenitors undergo at least partial restoration of aberrant DNA methylation patterns. In some cases, this recovery process may be sufficient to rescue a certain degree of adult lymphopoiesis. On the other hand, the compensatory process often fails, such that a considerable fraction of mutant zebrafish accepts allogeneic transplants without the requirement of prior conditioning. Hence, from a practical perspective, the *dnmt1* mutants can be added to the growing list of immunodeficient fish lines, such as those carrying mutations in *c-myb*, *rag2*, *prkdc*, and *zap70* (Hess and Boehm, 2016; Hess et al., 2013; Moore et al., 2016a; Moore et al., 2016b) that have been used for transplantation experiments in diverse settings (Yan et al., 2019). Mice homozygous for the H1511D mutation are phenotypically more similar to zebrafish than to medaka, as indicated by the reduced lymphoid progenitor cells in the bone marrow and the thymus and the sustained peripheral cytopenia. Although our mutants are viable and grossly normal except for a smaller body size, it is well possible that they exhibit subtle abnormalities other than those of the hematopoietic system studied here; however, whatever their nature may be, they do not interfere with viability and fertility.

With respect to the molecular mechanism underlying the pathology in *dnmt1* mutants, we propose that subtle changes in the interaction of the DNMT1 enzyme with its hemimethylated target sites underlie the observed reduced levels of DNA cytosine methylation, as implied by the crystal structure of the mouse DNMT1 protein (Song et al., 2012; Takeshita et al., 2011) and the molecular dynamics simulations reported herein. It is conceivable that the alterations in the catalytic site observed in the mutants change the mode and/or the efficiency with which the enzyme engages certain CpG dinucleotides. One obvious confounding factor is the sequence context of the CpG dinucleotide; other variables include chromatin context and altered interactions with accessory proteins in the methylase complex. Additional contributions to the altered methylation pattern could arise from compensatory *de novo* methylation processes. The structure-function relationships reported here should guide the generation of further mutants that may exhibit a different phenotypic spectrum than those uncovered in the present study.

In conclusion, our study establishes a set of three animal models with which to examine the physiological consequences of impaired maintenance methylation throughout life. The opportunity to study this in the organismal context is unprecedented and should therefore be applicable to many biological questions. Indeed, although we have focused here on hematological features of the mutant phenotype, in-depth analyses of other organ systems and physiological responses may reveal other so far undetected abnormalities. When viewed from an evolutionary angle, our study suggests that DNA methylation is required to protect the development of lymphocytes within the hematopoietic system of teleosts and mammals, which have independently evolved for several hundred millions years. Hence, it appears that an ancient epigenetic mechanism was repurposed to enable the establishment of an evolutionarily new cell type in vertebrates, underlying their adaptive immune systems.

Limitations of the Study

Our work demonstrates that the lymphoid lineage is consistently affected in three animal models homozygous for hypomorphic mutations in *DNMT1*. However, although viability is not compromised, the mutation also affects other organ systems, such as growth rate etc. Further studies are required to examine these additional features of the phenotypes. With respect to the mouse model described here, hematopoietic precursors should be examined in more detail, in particular with respect to possible signs of haploinsufficiency. Moreover, it will be important to determine whether the consistent upregulation of *Fos* and *FosB* genes is causal for the observed lymphoid defects or whether it is an indirect consequence of DNA hypomethylation in this lineage.

Resource Availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Thomas Boehm (boehm@ie-freiburg.mpg.de).

Materials Availability

All unique reagents generated in this study are available from the lead Contact with a completed Materials Transfer Agreement.

Data and Code Availability

The R code necessary to reproduce statistical analyses and results presented here is reported in Supplemental Code available at https://github.com/katsikora/lwanami2019_SupplementaryCodeAndData_B. The methylation and RNA_seq data have been deposited in the GEO database under accession number GSE98648.

METHODS

All methods can be found in the accompanying [Transparent Methods](#) supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.101260>.

ACKNOWLEDGMENTS

We are grateful to NBRP Medaka (<https://shigen.nig.ac.jp/medaka/>) for providing Kaga and *gyokuro* medaka strains and T. Sasado for help with their characterization. We thank U. Bönisch, S. Diehl, and A. Richter for help with next-generation sequencing, Sagar for help with RNA isolation, Benoît Kanzler for help with generating the mouse *Dnmt1* mutants, and T. Clapes and E. Trompouki for help with competitive bone marrow transplantations. The members of the Boehm laboratory are gratefully acknowledged for help during various stages of the project. This project has received funding from the Max Planck Society, the Deutsche Forschungsgemeinschaft, and the European Research Council under the European Union's Seventh Framework Program (FP7/2007–2013) ERC grant agreement no. 323126. This research was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP19am0101070 (support number 1563). This work was performed in part under the International Collaborative Research Program of Institute for Protein Research, Osaka University, ICR-16-17.

AUTHOR CONTRIBUTIONS

All authors designed experiments and analyzed data. N.I., D.-F.L., I.T., C.O'M., and I.S. performed experiments. Y. T., M.F.-S, and H.K. developed the medaka mutants; K.S. carried out the bioinformatic analyses; K.T. and Y.Y. carried out protein structure analysis; I.S. and S.T. carried out *in vitro* functional analysis; T.C. and E.T. established bone marrow chimeras; N.I. and T.B. directed the study and wrote the paper with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 19, 2020

Revised: April 28, 2020

Accepted: June 8, 2020

Published: July 24, 2020

REFERENCES

- Anderson, R.M., Bosch, J.A., Goll, M.G., Hesselson, D., Dong, P.D., Shin, D., Chi, N.C., Shin, C.H., Schlegel, A., Halpern, M., et al. (2009). Loss of *Dnmt1* catalytic activity reveals multiple roles for DNA methylation during pancreas development and regeneration. *Dev. Biol.* 334, 213–223.
- Bajoghli, B., Aghaallaei, N., Hess, I., Rode, I., Netuschil, N., Tay, B.-H., Venkatesh, B., Yu, J.-K., Kaltenbach, S.L., Holland, N.D., et al. (2009). Evolution of genetic networks underlying the emergence of thymopoiesis in vertebrates. *Cell* 138, 186–197.
- Bergman, Y., and Cedar, H. (2013). DNA methylation dynamics in health and disease. *Nat. Struct. Mol. Biol.* 20, 274–281.
- Boehm, T., Bleul, C.C., and Schorpp, M. (2003). Genetic dissection of thymus development in mouse and zebrafish. *Immunol. Rev.* 195, 15–27.
- Boehm, T., and Swann, J.B. (2014). Origin and evolution of adaptive immunity. *Annu. Rev. Anim. Biosci.* 2, 259–283.
- Broske, A.M., Vockentanz, L., Kharazi, S., Huska, M.R., Mancini, E., Scheller, M., Kuhl, C., Enns, A., Prinz, M., Jaenisch, R., et al. (2009). DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat. Genet.* 41, 1207–1215.
- Cedar, H., and Bergman, Y. (2011). Epigenetics of haematopoietic cell development. *Nat. Rev. Immunol.* 11, 478–488.
- Cheng, J., Yang, H., Fang, J., Ma, L., Gong, R., Wang, P., Li, Z., and Xu, Y. (2015). Molecular mechanism for USP7-mediated DNMT1 stabilization by acetylation. *Nat. Commun.* 6, 7023.
- Clark, M.R., Mandal, M., Ochiai, K., and Singh, H. (2014). Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nat. Rev. Immunol.* 14, 69–80.
- Cumano, A., Berthault, C., Ramond, C., Petit, M., Golub, R., Bandeira, A., and Pereira, P. (2019). New molecular insights into immune cell development. *Annu. Rev. Immunol.* 37, 497–519.
- Fridkis-Hareli, M., Abel, L., Eisenbach, L., and Globerson, A. (1992). Differentiation patterns of CD4/CD8 thymocyte subsets in cocultures of fetal

- thymus and lymphohemopoietic cells from c-fos transgenic and normal mice. *Cell Immunol.* 141, 279–292.
- Furutani-Seiki, M., Sasado, T., Morinaga, C., Suwa, H., Niwa, K., Yoda, H., Deguchi, T., Hirose, Y., Yasuoka, A., Henrich, T., et al. (2004). A systematic genome-wide screen for mutations affecting organogenesis in Medaka, *Oryzias latipes*. *Mech. Dev.* 121, 647–658.
- Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., and Jaenisch, R. (2003). Induction of tumors in mice by genomic hypomethylation. *Science* 300, 489–492.
- Guillamot, M., Cimmino, L., and Aifantis, I. (2016). The impact of DNA methylation in hematopoietic malignancies. *Trends Cancer* 2, 70–83.
- Hess, I., and Boehm, T. (2012). Intravital imaging of thymopoiesis reveals dynamic lympho-epithelial interactions. *Immunity* 36, 298–309.
- Hess, I., and Boehm, T. (2016). Stable multilineage xenogeneic replacement of definitive hematopoiesis in adult zebrafish. *Sci. Rep.* 6, 19634.
- Hess, I., Iwanami, N., Schorpp, M., and Boehm, T. (2013). Zebrafish model for allogeneic hematopoietic cell transplantation not requiring preconditioning. *Proc. Natl. Acad. Sci. U S A* 110, 4327–4332.
- Hu, L., Hatano, M., Ruther, U., and Tokuhisa, T. (1996). Overexpression of c-Fos induces apoptosis of CD43+ pro-B cells. *J. Immunol.* 157, 3804–3811.
- Imoto, S., Hu, L., Tomita, Y., Phuchareon, J., Ruther, U., and Tokuhisa, T. (1996). A regulatory role of c-Fos in the development of precursor B lymphocytes mediated by interleukin-7. *Cell Immunol.* 169, 67–74.
- Iwamoto, T. (2004). Stages of normal development in the medaka *Oryzias latipes*. *Mech. Dev.* 121, 605–618.
- Iwanami, N., Takahama, Y., Kunimatsu, S., Li, J., Takei, R., Ishikura, Y., Suwa, H., Niwa, K., Sasado, T., Morinaga, C., et al. (2004). Mutations affecting thymus organogenesis in Medaka, *Oryzias latipes*. *Mech. Dev.* 121, 779–789.
- Iwanami, N., Higuchi, T., Sasano, Y., Fujiwara, T., Hoa, V.Q., Okada, M., Talukder, S.R., Kunimatsu, S., Li, J., Saito, F., et al. (2008). WDR55 is a nucleolar modulator of ribosomal RNA synthesis, cell cycle progression, and teleost organ development. *PLoS Genet.* 4, e1000171.
- Iwanami, N., Okada, M., Hoa, V.Q., Seo, Y., Mitani, H., Sasaki, T., Shimizu, N., Kondoh, H., Furutani-Seiki, M., and Takahama, Y. (2009). Ethylnitrosourea-induced thymus-defective mutants identify roles of KIAA1440, TRRAP, and SKIV2L2 in teleost organ development. *Eur. J. Immunol.* 39, 2606–2616.
- Iwanami, N., Sikora, K., Richter, A.S., Mönnich, M., Guerri, L., Soza-Ried, C., Lawir, D.-F., Mateos, F., Hess, I., O'Meara, C.P., et al. (2016). Forward genetic screens in zebrafish identify pre-mRNA-processing pathways regulating early T cell development. *Cell Rep.* 17, 2259–2270.
- Jeong, M., and Goodell, M.A. (2014). New answers to old questions from genome-wide maps of DNA methylation in hematopoietic cells. *Exp. Hematol.* 42, 609–617.
- Klein, C.J., Botuyan, M.V., Wu, Y., Ward, C.J., Nicholson, G.A., Hammans, S., Hojo, K., Yamanishi, H., Karpf, A.R., Wallace, D.C., et al. (2011). Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss. *Nat. Genet.* 43, 595–600.
- Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretsky, I., Jaitin, D.A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S., et al. (2014). Immunogenetics. Chromatin state dynamics during blood formation. *Science* 345, 943–949.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915–926.
- Li, J., Iwanami, N., Hoa, V.Q., Furutani-Seiki, M., and Takahama, Y. (2007). Noninvasive intravital imaging of thymocyte dynamics in medaka. *J. Immunol.* 179, 1605–1615.
- Medvedovic, J., Ebert, A., Tagoh, H., and Busslinger, M. (2011). Pax5: a master regulator of B cell development and leukemogenesis. *Adv. Immunol.* 111, 179–206.
- Moore, J.C., Mulligan, T.S., Yordan, N.T., Castranova, D., Pham, V.N., Tang, Q., Lobbardi, R., Anselmo, A., Liwski, R.S., Berman, J.N., et al. (2016a). T cell immune deficiency in zap70 mutant zebrafish. *Mol. Cell. Biol.* 36, 2868–2876.
- Moore, J.C., Tang, Q., Yordan, N.T., Moore, F.E., Garcia, E.G., Lobbardi, R., Ramakrishnan, A., Marvin, D.L., Anselmo, A., Sadreyev, R.I., et al. (2016b). Single-cell imaging of normal and malignant cell engraftment into optically clear prkdc-null SCID zebrafish. *J. Exp. Med.* 213, 2575–2589.
- Potok, M.E., Nix, D.A., Parnell, T.J., and Cairns, B.R. (2013). Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* 153, 759–772.
- Rothenberg, E.V., Ungerback, J., and Champhekar, A. (2016). Forging T-lymphocyte identity: intersecting networks of transcriptional control. *Adv. Immunol.* 129, 109–174.
- Ruther, U., Muller, W., Sumida, T., Tokuhisa, T., Rajewsky, K., and Wagner, E.F. (1988). c-fos expression interferes with thymus development in transgenic mice. *Cell* 53, 847–856.
- Schorpp, M., Leicht, M., Nold, E., Hammerschmidt, M., Haas-Assenbaum, A., Wiest, W., and Boehm, T. (2002). A zebrafish orthologue (whnb) of the mouse nude gene is expressed in the epithelial compartment of the embryonic thymic rudiment. *Mech. Dev.* 118, 179–185.
- Schorpp, M., Bialecki, M., Diekhoff, D., Walderich, B., Odenthal, J., Maischein, H.-M., Zapata, A.G., Tübingen 2000 Screen Consortium, Freiburg Screening Group, and Boehm, T. (2006). Conserved functions of Ikaros in vertebrate lymphocyte development: genetic evidence for distinct larval and adult phases of T cell development and two lineages of B cells in zebrafish. *J. Immunol.* 177, 2463–2476.
- Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* 14, 204–220.
- Song, J., Teplova, M., Ishibe-Murakami, S., and Patel, D.J. (2012). Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. *Science* 335, 709–712.
- Sun, D., Luo, M., Jeong, M., Rodriguez, B., Xia, Z., Hannah, R., Wang, H., Le, T., Faull, K.F., Chen, R., et al. (2014a). Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell* 14, 673–688.
- Sun, Z., Wu, Y., Ordog, T., Baheti, S., Nie, J., Duan, X., Hojo, K., Kocher, J.P., Dyck, P.J., and Klein, C.J. (2014b). Aberrant signature methylome by DNMT1 hot spot mutation in hereditary sensory and autonomic neuropathy 1E. *Epigenetics* 9, 1184–1193.
- Takeshita, K., Suetake, I., Yamashita, E., Suga, M., Narita, H., Nakagawa, A., and Tajima, S. (2011). Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). *Proc. Natl. Acad. Sci. U S A* 108, 9055–9059.
- Tian, Y., Xu, J., Feng, S., He, S., Zhao, S., Zhu, L., Jin, W., Dai, Y., Luo, L., Qu, J.Y., et al. (2017). The first wave of T lymphopoiesis in zebrafish arises from aorta endothelium independent of hematopoietic stem cells. *J. Exp. Med.* 214, 3347–3360.
- Winkelmann, J., Lin, L., Schormair, B., Kornum, B.R., Faraco, J., Plazzi, G., Melberg, A., Cornelio, F., Urban, A.E., Pizza, F., et al. (2012). Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum. Mol. Genet.* 21, 2205–2210.
- Yan, C., Brunson, D.C., Tang, Q., Do, D., Iftimia, N.A., Moore, J.C., Hayes, M.N., Welker, A.M., Garcia, E.G., Dubash, T.D., et al. (2019). Visualizing engrafted human cancer and therapy responses in immunodeficient zebrafish. *Cell* 177, 1903–1914.e14.
- Yoshida, H., Lareau, C.A., Ramirez, R.N., Rose, S.A., Maier, B., Wroblewska, A., Desland, F., Chudnovskiy, A., Mortha, A., Dominguez, C., et al. (2019). The cis-regulatory atlas of the mouse immune system. *Cell* 176, 897–912.
- Zhang, Z.M., Liu, S., Lin, K., Luo, Y., Perry, J.J., Wang, Y., and Song, J. (2015). Crystal structure of human DNA methyltransferase 1. *J. Mol. Biol.* 427, 2520–2531.

Supplemental Information

Epigenetic Protection of Vertebrate

Lymphoid Progenitor Cells by *Dnmt1*

Norimasa Iwanami, Kohei Takeshita, Divine-Fondzenyuy Lawir, Isao Suetake, Shoji Tajima, Katarzyna Sikora, Inês Trancoso, Connor ÓMeara, Iliana Siamishi, Yousuke Takahama, Makoto Furutani-Seiki, Hisato Kondoh, Yasushige Yonezawa, Michael Schorpp, and Thomas Boehm

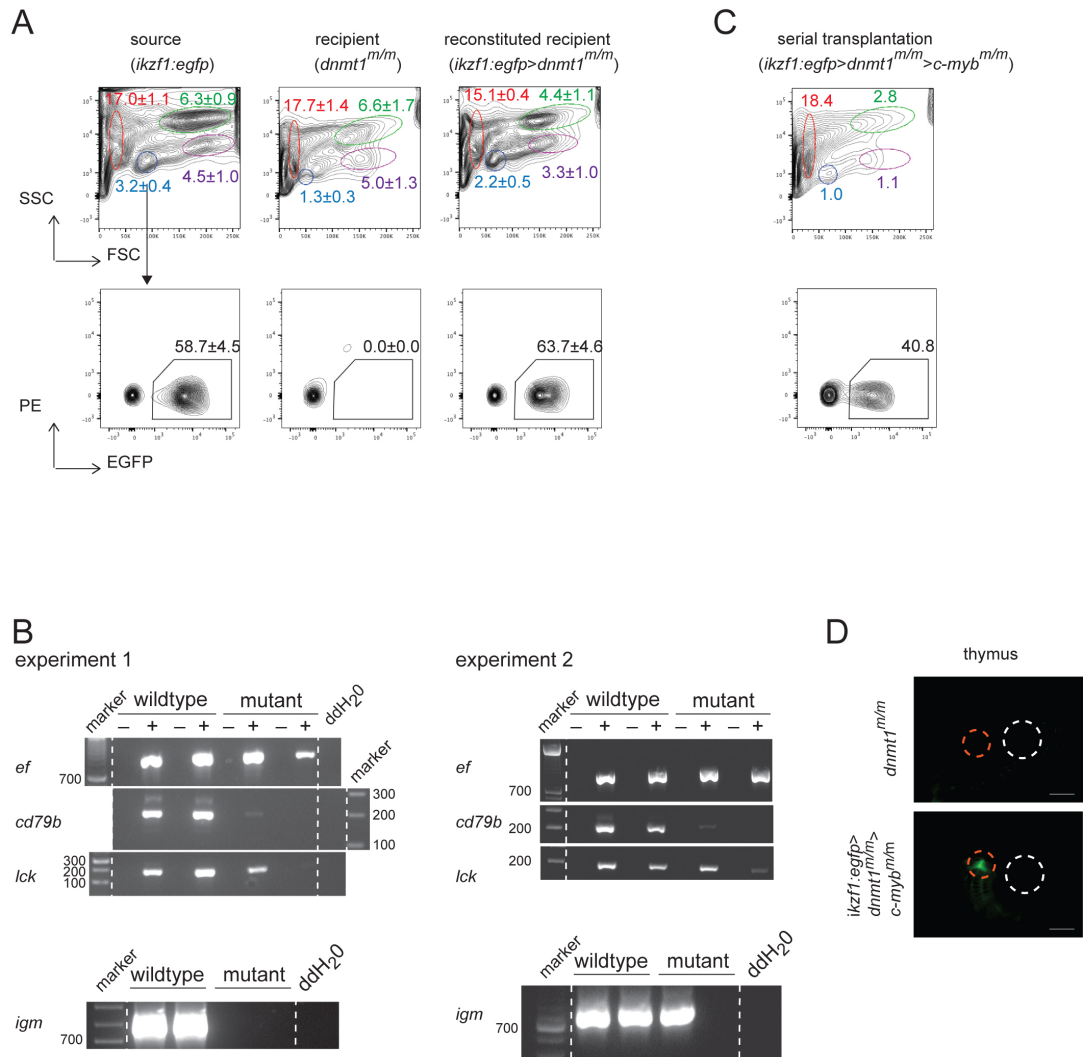


Figure S1. Long-term reconstitution of transplanted kidney cells in *dnmt1* mutant zebrafish. Related to Figures 1, 2.

(A) The top panels exemplify light scatter profiles of whole kidney marrow cells isolated from fish of the indicated genotypes and transplantation histories; the percentages of erythroid (red), lymphoid (light blue), myeloid (green), and precursor (purple) cell populations are indicated as mean±s.e.m.; the bottom panels exemplify EGFP expression patterns of cells in lymphoid populations, mean±s.e.m. is shown; n=7, 4, 6, respectively.

(B) Gel images of RT-PCR amplicons using kidney-derived RNA at 7 wpf. The results of 2 independent experiments are shown. Two *dnmt1^{+/+}* (WT) fish and 2 *dnmt1^{m/m}* (MT) fish were used for each experiment. The symbols – and + indicate cDNA templates generated without and with addition of reverse transcriptase to the reaction, respectively. *ef*, amplicon specific for elongation factor as a control for successful synthesis of cDNA; *cd79b*, amplicon specific for the B-cell marker CD79B factor; *lck*, amplicon specific for the T cell marker LCK; *igm*; amplicon specific for the B-cell marker immunoglobulin IGM; ddH₂O denotes PCR reactions without template. Size markers are indicated in bp.

(C) Flow cytometric profiles of whole kidney marrow cells isolated from a *c-myb* mutant after receiving cells from a *dnmt1* mutant fish previously transplanted with wild-type whole kidney marrow cells; n=1.

(D) Fluorescent microscopic images of *dnmt1* mutant and the secondary *c-myb* mutant recipient shown in (B), 37 days after the secondary transplantation. Red circles indicate thymic region, and white circles indicate the location of the eye. Scale bars, 1mm.

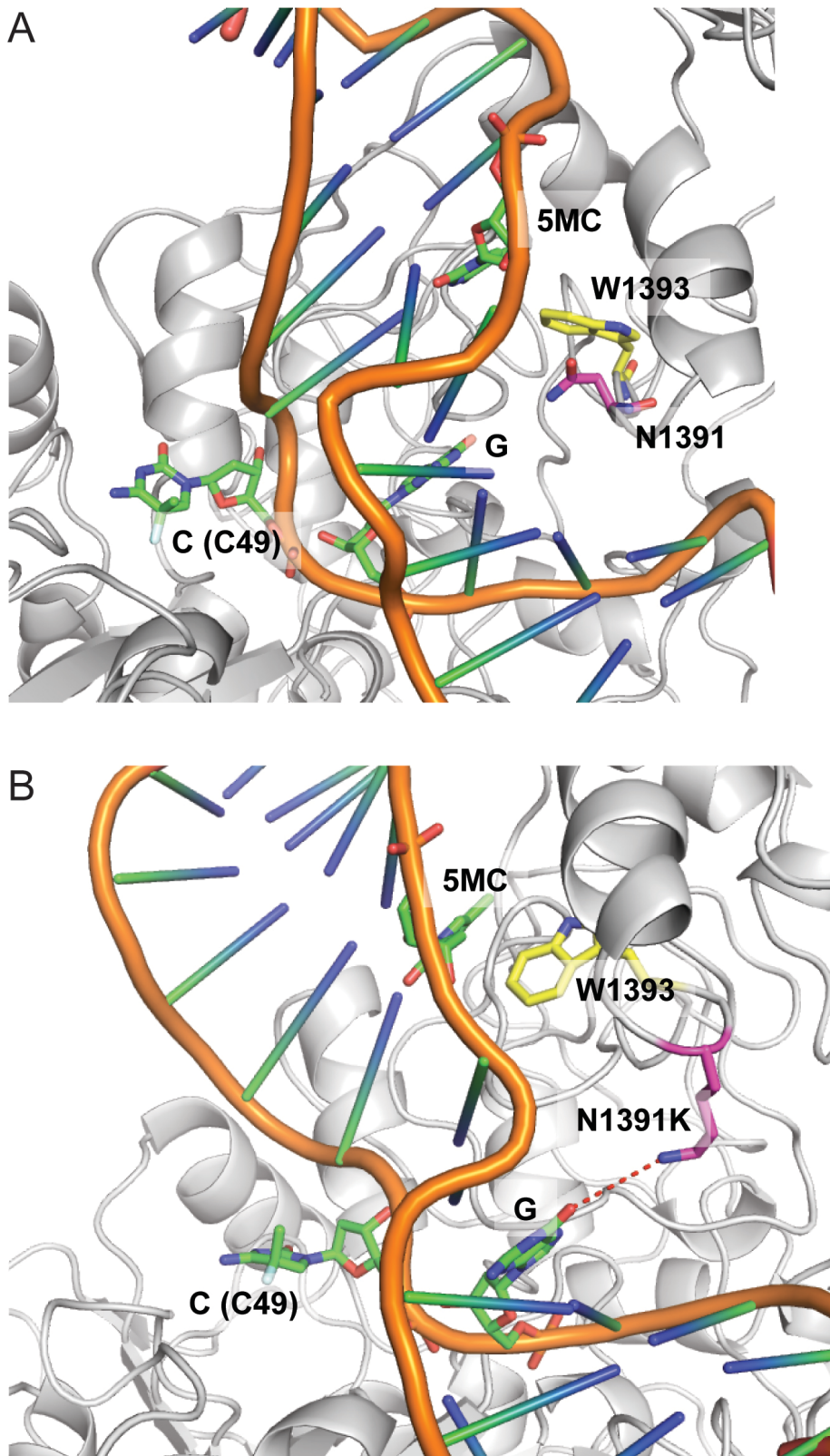


Figure S2. Structures of catalytic centers of zebrafish wild-type and mutant dnmt1 proteins as determined by molecular dynamics simulation. Related to Figures 1, 2, 4.
 (A) Wild-type structure.
 (B) N1391K mutant structure. Note the hydrogen bonds between the mutant lysine residue and the guanine base.

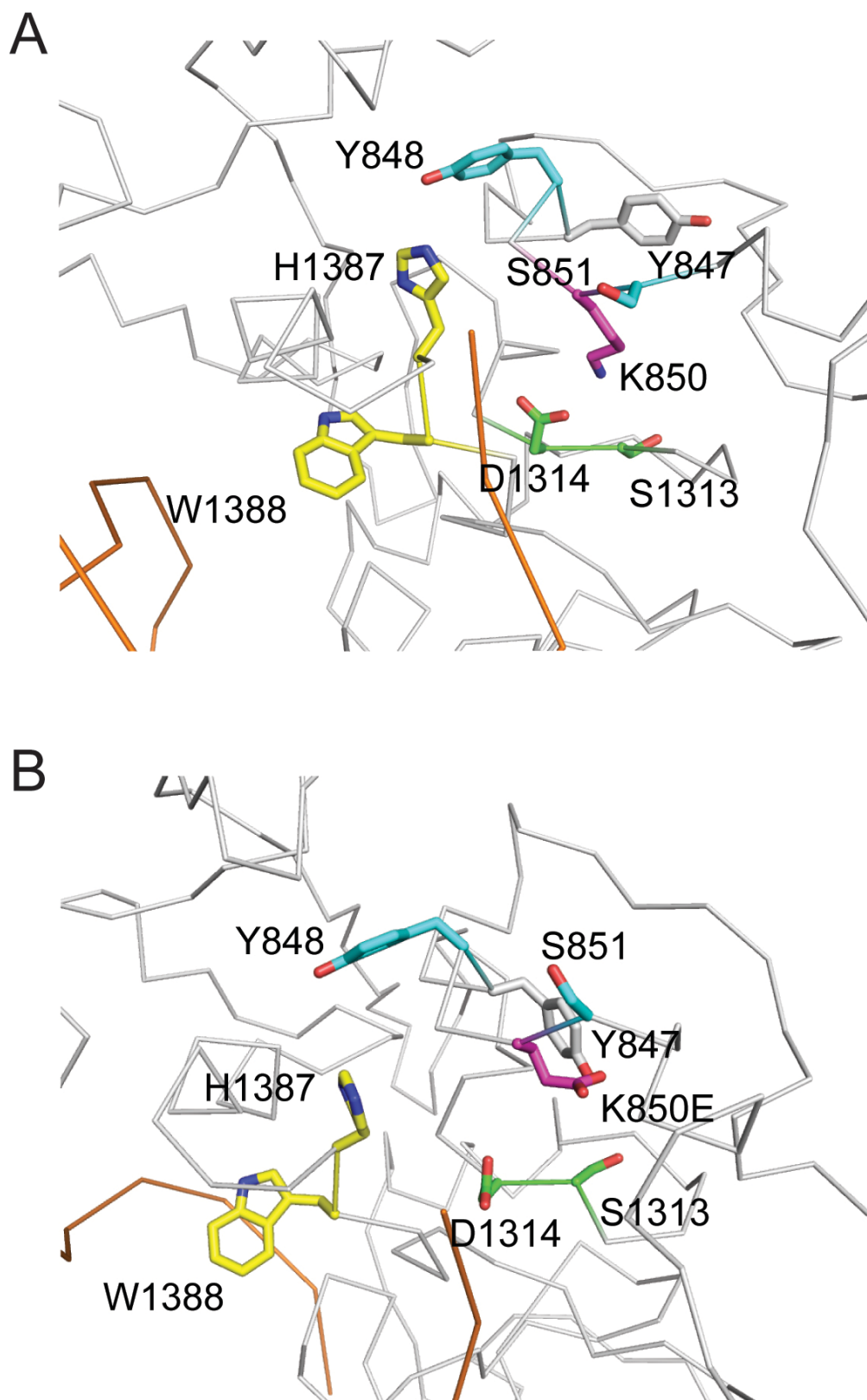


Figure S3. Structures of catalytic centers of medaka wild-type and mutant dnmt1 proteins as determined by molecular dynamics simulation. Related to Figures 1, 3, 4.

(A) Wild-type structure.

(B) K850E mutant structure. Note the distorted structure of the catalytic center.

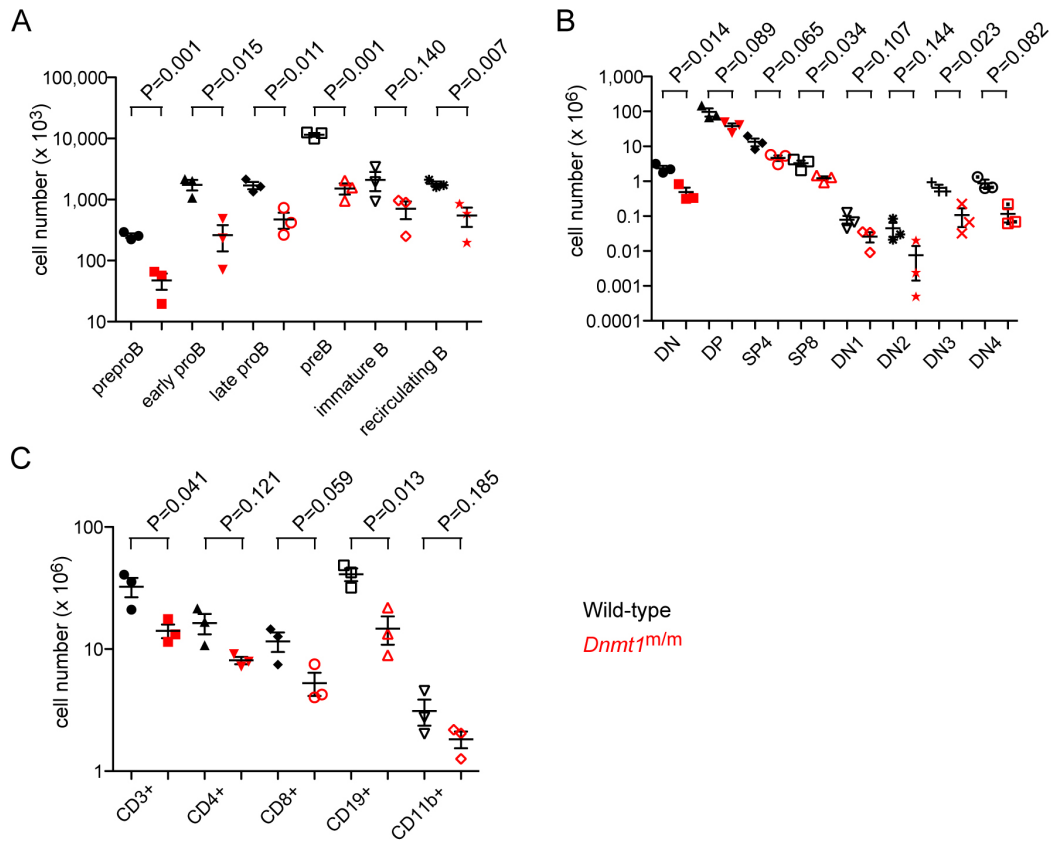


Figure S4. Impaired hematopoiesis in H1511D homozygous mice. Related to Figures 5, 6.

(A) B cell types in bone marrow.

(B) Cell types in thymus.

(C) Cell types in spleen.

Cell populations were measured by quantitative flow cytometry.

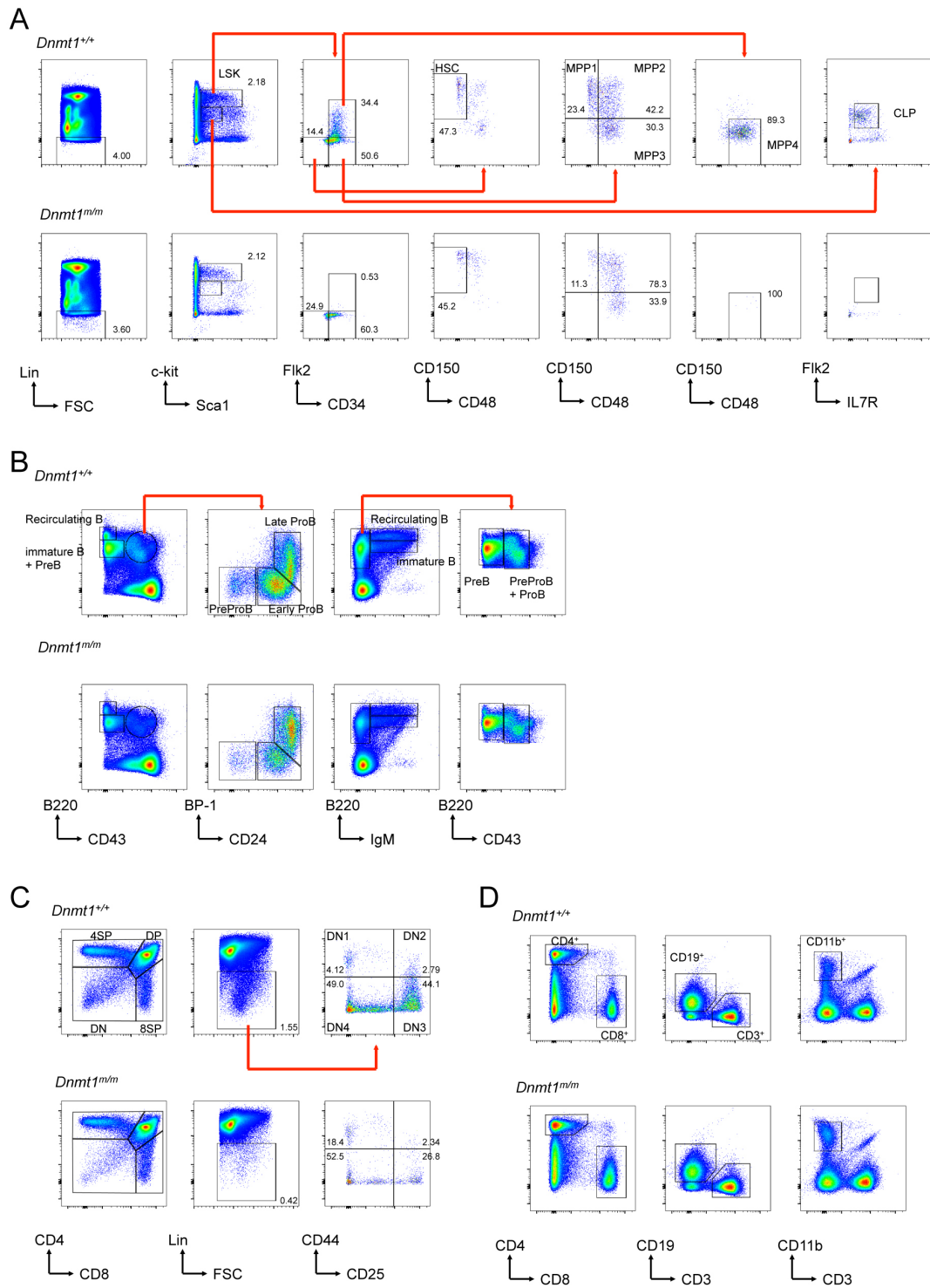


Figure S5. Gating strategies used for the characterization of the hematopoietic compartment of wildtype and H1511D mutant mice. Related to Figures 5, 6.

(A) Characterization of precursor populations in the bone marrow.

(B) Characterization of B cell precursor cells in the bone marrow.

(C) Characterization of T cell precursor cells in the thymus.

(D) Characterization of splenocytes.

Cell populations were measured by quantitative flow cytometry.

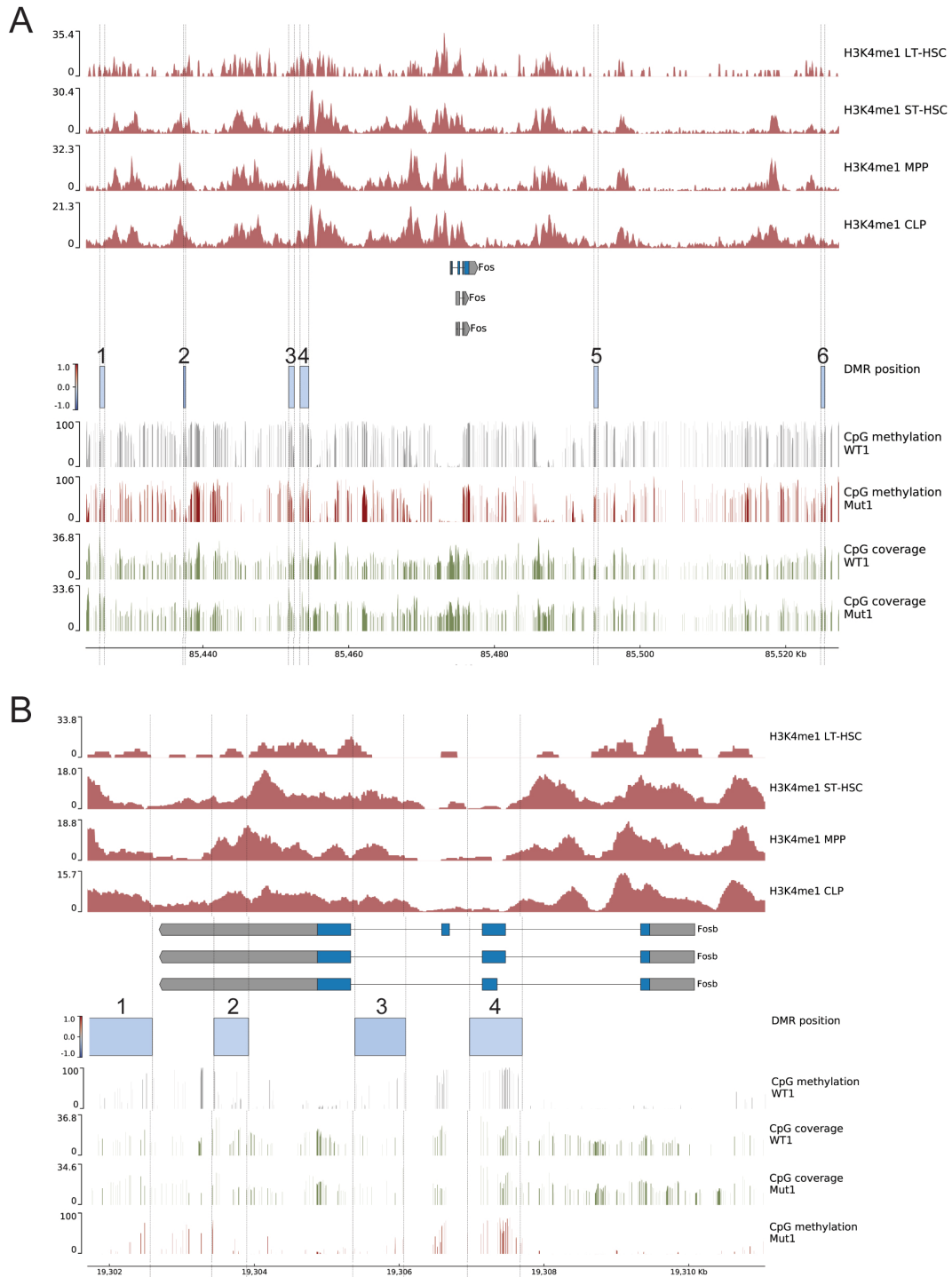


Figure S6. DMRs in the vicinity of *Fos* and *Fosb* genes. Related to Figure 7.

(A) Localization of 6 DMRs in the *Fos* locus; see Figure 7C.

(B) Localization of 4 DMRs in the *Fosb* locus; see Figure 7C.

Dashed vertical lines highlight the position of differentially methylated DMRs (hypomethylated) in wild-type and mutant mice.

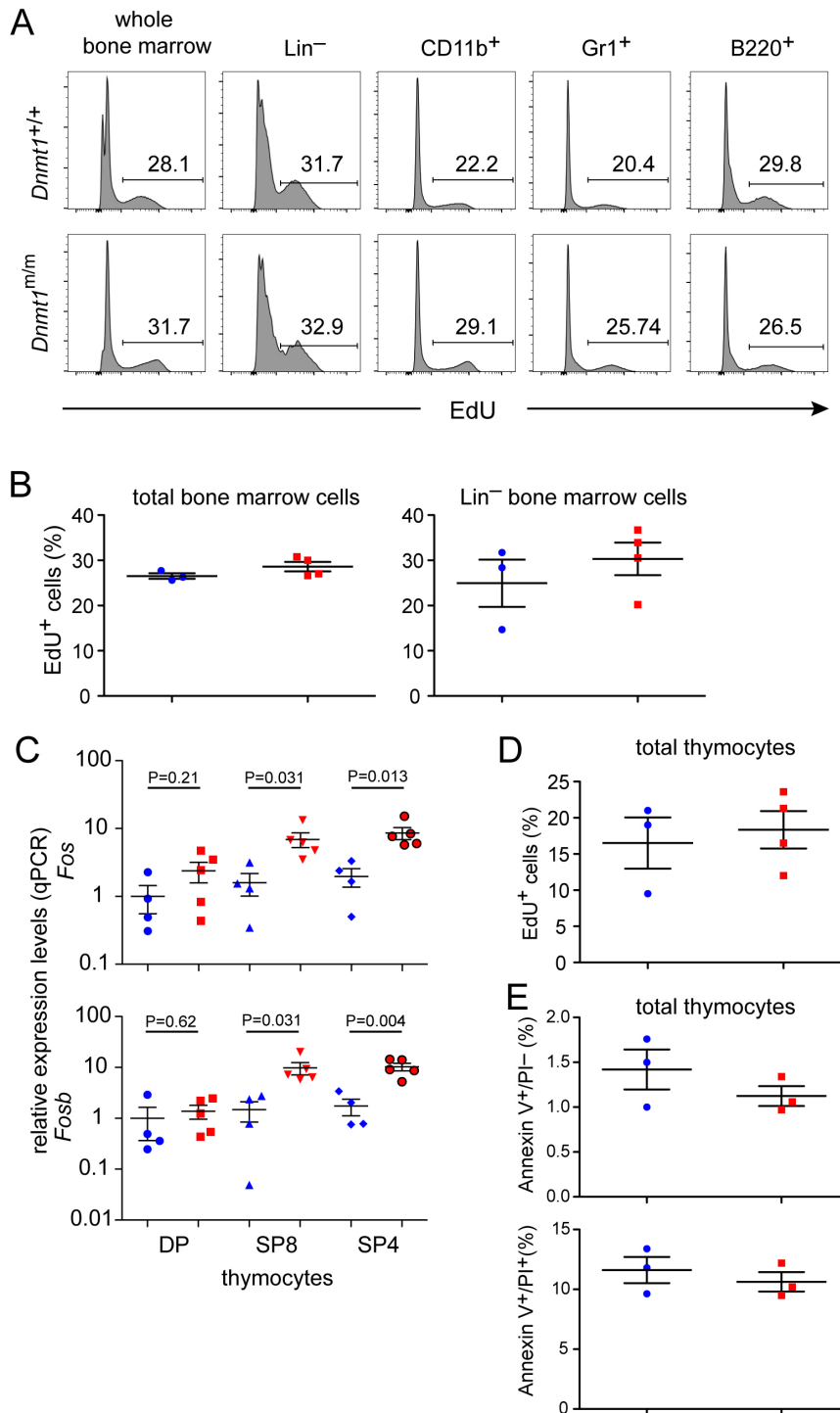


Figure S7. Characterization of the hematopoietic compartment of H1511D mutant mice. Related to Figures 5, 6.

(A) Representative histograms showing EdU incorporation for the indicated bone marrow cell populations. Note the unperturbed proliferative potential of bone marrow cell populations in *Dnmt1*-mutants. (B) Unperturbed proliferative potential of bone marrow lin-negative precursor cell populations in *Dnmt1*-mutants.

(C) Elevated expression levels of *Fos* and *FosB* genes in thymocytes of *Dnmt1*-mutants as determined by qPCR.

(D) Unperturbed proliferative potential of thymocytes of and *Dnmt1*-mutant animals.

(E) Lack of increased apoptosis in thymocytes of *Dnmt1*-mutant animals.

B-E, wild-type (blue data points) and *Dnmt1*-mutant (red data points).

Table S1. Sequence polymorphisms in the critical genomic interval of the gkr medaka mutant. Related to Figure 3.

Row.names	POS	EFFECT	GENE	GENEID	HGVS_C	HGVS_P	Comments			
8:8390707_G/A	8390707	missense_variant	GB-ALPHA3	ENSORLG00000005267	c.208G>A	p.Ala70Thr	a) b)H,M			f)
8:8404841_G/A	8404841	missense_variant	zgc:92880	ENSORLG00000005283	c.373C>T	p.Leu125Phe			d)	f)
8:8420687_C/T	8420687	missense_variant	KIAA0172	ENSORLG00000005329	c.73G>A	p.Ala25Thr			d)	
8:8437467_G/A	8437467	missense_variant	DOCK6	ENSORLG00000005365	c.3554C>T	p.Pro1185Leu	a)		d)	
8:8499625_G/A	8499625	missense_variant	pde4a	ENSORLG00000005439	c.214C>T	p.Leu72Phe	a)	c)		f)
8:8538307_G/A	8538307	missense_variant	raver1	ENSORLG00000005458	c.1357C>T	p.Pro453Ser		c)		
8:8539847_T/A	8539847	missense_variant	raver1	ENSORLG00000005458	c.793A>T	p.Thr265Ser			d)	
8:8622951_A/C	8622951	missense_variant	dnmt1	ENSORLG00000005582	c.1743A>C	p.Glu581Asp			d)	
8:8624154_G/G	8624154	missense_variant	dnmt1	ENSORLG00000005582	c.2533A>G	p.Lys845Glu				
8:8633660_G/T	8633660	missense_variant	p2ry11	ENSORLG00000005606	c.406C>A	p.Leu136Ile	a) b)M	c)		
8:8634159_A/T	8634159	stop_gained	p2ry11	ENSORLG00000005606	c.14T>A	p.Leu5*	a) b)M		d)	e)
8:8638287_T/G	8638287	missense_variant	ppan	ENSORLG00000005634	c.575A>C	p.His192Pro			d)	
8:8641575_G/A	8641575	missense_variant	angptl6	ENSORLG00000005649	c.113G>A	p.Arg38His			d)	
8:8642057_G/T	8642057	missense_variant	angptl6	ENSORLG00000005649	c.515G>T	p.Gly172Val			d)	
8:8648875_G/A	8648875	missense_variant	ENSORLG00000005655	ENSORLG00000005655	c.297G>A	p.Met99Ile	a) b)H,M			f)
8:8649032_T/C	8649032	missense_variant	ENSORLG00000005655	ENSORLG00000005655	c.454T>C	p.Phe152Leu	a) b)H,M			f)
8:8649110_C/T	8649110	missense_variant	ENSORLG00000005655	ENSORLG00000005655	c.532C>T	p.Arg178Cys	a) b)H,M			f)
8:8666931_A/G	8666931	missense_variant	TRIP10__1_of_2_	ENSORLG00000005760	c.1616A>G	p.Asn539Ser			d)	f)
8:8776418_G/T	8776418	missense_variant	col5a3a	ENSORLG00000005829	c.3922C>A	p.Pro1308Thr	a)		d)	f)
8:8783104_G/T	8783104	missense_variant	col5a3a	ENSORLG00000005829	c.3259C>A	p.Leu1087Met	a)		d)	f)
8:8795210_C/A	8795210	missense_variant	col5a3a	ENSORLG00000005829	c.1278G>T	p.Met426Ile	a)		d)	f)
8:8804389_C/T	8804389	missense_variant	col5a3a	ENSORLG00000005829	c.481G>A	p.Val161Ile	a)		d)	f)
8:8831897_G/A	8831897	missense_variant	rdh8a	ENSORLG00000005853	c.920G>A	p.Arg307Gln		c)		f)
8:8905744_C/T	8905744	missense_variant	CAMSAP3__1_of_2_	ENSORLG00000006040	c.1081G>A	p.Ala361Thr	a)	c)		
8:8908131_G/C	8908131	missense_variant	CAMSAP3__1_of_2_	ENSORLG00000006040	c.437C>G	p.Ala146Gly	a)		d)	
8:8913101_T/A	8913101	missense_variant	CAMSAP3__1_of_2_	ENSORLG00000006040	c.299A>T	p.Lys100Ile	a)		d)	
8:9049797_G/T	9049797	missense_variant	MYH13__6_of_11_	ENSORLG00000006359	c.1903G>T	p.Ala635Ser			d)	f)
8:9104419_G/T	9104419	missense_variant	brd4	ENSORLG00000006490	c.2353G>T	p.Ala785Ser		c)		f)
8:9217392_C/G	9217392	missense_variant	ENSORLG00000006644	ENSORLG00000006644	c.98G>C	p.Arg33Thr		b)H,M,Z		
8:9217417_C/T	9217417	missense_variant	ENSORLG00000006644	ENSORLG00000006644	c.73G>A	p.Gly25Ser		b)H,M,Z		
8:9224420_C/T	9224420	missense_variant	trim35-13	ENSORLG00000006660	c.520C>T	p.His174Tyr	a)		b)H,M	
8:9224496_C/G	9224496	missense_variant	trim35-13	ENSORLG00000006660	c.596C>G	p.Ser199Cys	a)		b)H,M	
8:9229906_A/G	9229906	missense_variant	slc27a1a	ENSORLG00000006698	c.455T>C	p.Val152Ala			d)	f)
8:9233995_T/C	9233995	missense_variant	ptger1c	ENSORLG00000006707	c.25A>G	p.Lys9Glu		b)H,M		f)
8:9234003_C/T	9234003	missense_variant	ptger1c	ENSORLG00000006707	c.17G>A	p.Ser6Asn		b)H,M		f)
8:9460925_G/C	9460925	missense_variant	dnm2a	ENSORLG00000006841	c.2427C>G	p.Phe809Leu				f)

- a) Ensembl transcript does not contain start codon.
- b) No ortholog in human (H), mouse (M), or zebrafish (Z).
- c) Amino acid residues are not conserved among 4 species at around mutation. Alignment is difficult.
- d) Mutation residue is not conserved among 4 species.
- e) One of two variants (exon1) contains the mutation nucleotide. Neither exon1 prediction might be wrong.
- f) At least one more paralogue in medaka.

Table S2. Primers and oligonucleotides used in this study. Related to Figures 1 and 3.

Primer name	Primer sequence	Note
medaka mapping primers		
1345-F	TGCAATGACCGAGGATCTGTAAAG	
1345-R	CTTCTGGGCTAAAATCTCAGGCAA	
Scaf31_1470_F	CACAATCGTCCGCCCTCTTC	
Scaf31_1470_R	CCTCGCATCTGAAGACATTG	
medaka dnmt1 genotyping		
ol_dnmt1_L1	TTCTGTATGCGCTGGCTTGC	
ol_dnmt1_R1	CAGTTTGACGTCTGCAGTGTC	ol_dnmt1_R1 also for sequencing
medaka bisulfite primers		
ol_cpn1_BSF1	TGTTYGGAATGTAAGAGATTAGATAGAGATTG	Annealing temperature: 56°C
ol_cpn1_BSR1	TCRCTAACTCCAACACATAAAAAATAAC	
P5_ol_cpn1_F2	ACACTCTTCCCTACACGACGCTCTCCGATCTTTTTAGTTTTYGTGTGTTTATTTGAGTTGAG	Annealing temperature: 55°C
P7_ol_cpn1_R1	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCTTCRCTAACTCCAACACATAAAAAATAAC	
ol_irx3a_BSF1	GATTTTTTAAATGGGATAAGAAGAGATGAGTTGTGGG	Annealing temperature: 58°C
ol_irx3a_BSR1	TTAAACRACAAAATAAECTACATTCAAAACAAC	
P5_ol_irx3a_F2	ACACTCTTCCCTACACGACGCTCTCCGATCTGGGGTYGTGGTTGTTTTGTTAAAG	Annealing temperature: 55°C
P7_ol_irx3a_R2	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCTAACRACTATAAAAAATTTACAAAAACCAAAAC	
ol_krt4_BSF1	ATGTGGGTGATGAGATAATATGGTTTATAG	Annealing temperature: 55°C
ol_krt4_BSR1	CAAATTTTTTAAAAAATCCTCCCTCTTCTAAAC	
P5_ol_krt4_F2	ACACTCTTCCCTACACGACGCTCTCCGATCTAAAGYGGTTGTTGAGGGATTGATTTG	Annealing temperature: 57°C
P7_ol_krt4_R2	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCTTCTATACCATAACCAATCAACTCTCCAC	
ol_krt8_BSF1	GGYGGTATGGGGATTTGGGAGATG	Annealing temperature: 63°C
ol_krt8_BSR1	AACCCTCCCTCRCTACTACTTTAAAACTCCTCC	
P5_ol_krt8_F2	ACACTCTTCCCTACACGACGCTCTCCGATCTAAGTTTTYGAATTTTTATGGAAGAGTTAAG	Annealing temperature: 55°C
P7_ol_krt8_R2	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCTAACRAACCAATTACTCTAACCAACCCAAAAATCC	
ol_ntla_BSF1	AAYGTTTTGAGTTGAATTTGATTTAGTGTGAAAAAG	Annealing temperature: 58°C
ol_ntla_BSR1	AACCAACRCCTAAAAACATCTCCTAAAC	
P5_ol_ntla_F1	ACACTCTTCCCTACACGACGCTCTCCGATCTAAYGTTTTGAGTTGAATTTGATTAGTGTGAAAAAG	Annealing temperature: 57°C
P7_ol_ntla_R2	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCTAAAAACRAAATCTATAAAACAAATTTAAAAAATTAACC	
ol_rarga_BSF1	AGTTAGGTAGGAATTAGATTGTATTTAGTTG	Annealing temperature: 55°C
ol_rarga_BSR1	TACTACTCTACAATCAATAACAAAAACCCCTTAAC	
P5_ol_rarga_F2	ACACTCTTCCCTACACGACGCTCTCCGATCTTATTAATTTAGATAGTTGTTGTTGTTGTTGATTAG	Annealing temperature: 57°C
P7_ol_rarga_R2	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCTAATATATCRACAAAAAAAACAAACAAAACTCAC	
ol_pou5f1_BSF1	ATTTTAGGTYGTAGGGTAGAAGATTTTATGGGGTATGG	Annealing temperature: 56°C
ol_pou5f1_BSR1	CTTCAACCRRAACTTACAACCAAAATTTAAATC	
P5_ol_pou5f1_F2	ACACTCTTCCCTACACGACGCTCTCCGATCTGATTGTTTTYGGGTTTGGGTTATTGGGGGGTTG	Annealing temperature: 59°C
P7_ol_pou5f1_R2	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCTAACCTACTTACAACAAACCACTTATAAC	
ol_rgc101640_BSF1	AGGAATTTGTGATAAATTGAAAGTAGTGGAGTG	Annealing temperature: 55°C
ol_rgc101640_BSR1	ACCTCTCCACAATCAATAACATACATATAC	
P5_ol_rgc101640F2	ACACTCTTCCCTACACGACGCTCTCCGATCTTTAYGGTAGTTAAGGAAGGAAGAAGATAG	Annealing temperature: 57°C
P7_ol_rgc101640R2	GTGACTGGAGTTCAGACGTGTGCTCTCCGATCTTATCCCAACTATCTTACTTACATAATAAACTCTCTC	

Table S4. Differentially expressed genes in mouse hematopoietic progenitor populations.
Related to Figure 7.

HSCs

	baseMean	log2FoldChar	lfcSE	stat	pvalue	padj
Fos__chr12	224.086878	1.57818887	0.27528112	5.73300799	9.87E-09	6.04E-05
Slc25a31__chr3	13.6636173	2.16486176	0.37323785	5.80022027	6.62E-09	6.04E-05
Hspa1b__chr17	15.7211853	2.00249014	0.37231931	5.3784214	7.51E-08	0.00030668
n-R5-8s1__chr18	33.9826963	1.7617384	0.34726974	5.07311232	3.91E-07	0.00119796
3830403N18Rik__chrX	10.6482505	1.85223151	0.37108556	4.99138666	5.99E-07	0.00146799
Jun__chr4	376.401288	1.30609389	0.2673006	4.88623635	1.03E-06	0.00209743
Fosb__chr7	115.444473	1.53445223	0.31932398	4.80531477	1.55E-06	0.00236475
Xlr3a__chrX	11.7733374	1.80175627	0.37298699	4.8306143	1.36E-06	0.00236475

MPP1

	baseMean	log2FoldChar	lfcSE	stat	pvalue	padj
Fos__chr12	181.652816	2.13959281	0.28994635	7.3792713	1.59E-13	3.24E-09
Gm10801__chr2	12712.2057	1.60943243	0.23855998	6.74644774	1.52E-11	1.54E-07
Gm26870__chr9	3632.56799	1.50105305	0.24793635	6.05418714	1.41E-09	9.58E-06
Gm17535__chr9	416.695834	1.62142188	0.29697577	5.45977835	4.77E-08	0.00024283
Fosb__chr7	122.053456	1.61859893	0.30050091	5.38633628	7.19E-08	0.00029303
Slc17a8__chr10	21.0513276	1.63427191	0.37906329	4.31134318	1.62E-05	0.04723095

MPP4

	baseMean	log2FoldChar	lfcSE	stat	pvalue	padj
Fos__chr12	218.987584	1.8925346	0.26778623	7.06733351	1.58E-12	2.31E-08
Slc25a31__chr3	11.7190425	2.42489572	0.41258584	5.87731203	4.17E-09	3.05E-05
3830403N18Rik__chrX	12.7715231	2.37805736	0.41040171	5.79446266	6.85E-09	3.34E-05
Hspa1a__chr17	17.5242296	2.27939166	0.41160531	5.53780923	3.06E-08	0.00011203
Tifab__chr13	55.6836923	1.67305569	0.31646211	5.28674882	1.25E-07	0.00036434
Cryab__chr9	10.250812	1.98444916	0.41161383	4.82114301	1.43E-06	0.00348067
Fosb__chr7	82.3970523	1.88705981	0.39593392	4.76609784	1.88E-06	0.00392588
Jun__chr4	283.010291	1.44605207	0.30679563	4.71340496	2.44E-06	0.00445535
Egr1__chr18	76.7558932	1.71870469	0.38567116	4.45639929	8.33E-06	0.01354956
Ube2c__chr2	219.965014	1.24885469	0.28639192	4.36064915	1.30E-05	0.01897306
Clec10a__chr11	7.4482292	1.75459929	0.40887703	4.29126401	1.78E-05	0.02363025
Stag3__chr5	27.1886181	1.43762347	0.35052816	4.10130666	4.11E-05	0.04995012
Xlr4a__chrX	7.67843443	1.68392317	0.41238281	4.08339809	4.44E-05	0.04995012

Table S5. Antibodies used in this study. Related to Figures 5, 6, S4, S5, and S7.

Antigen/Staining Reagent	Clone	Conjugate	Source	Identifier
Thymocyte stage analysis				
CD4	GK1.5	APC Cy7	Biolegend	100414
CD8a	53-6.7	PE	eBioscience	12-0081-85
CD44	IM7	APC	eBioscience	17-0441-81
CD25	PC61	PE Cy7	BD Biosciences	552880
B220 (CD45R)	RA3-6B2	FITC	Biolegend	103206
TCR $\gamma\delta$	eBioGL3	FITC	eBioscience	11-5711-82
NK1.1	PK136	FITC	Biolegend	108706
CD11c	HL3	FITC	BD Biosciences	557400
CD11b (Mac1)	M1/70	FITC	BD Biosciences	557396
Annexin V detection in thymocytes				
CD4	GK1.5	PE Dazzle	Biolegend	100455
CD8a	53-6.7	BV421	Biolegend	100738
CD44	IM7	APC	eBioscience	17-0441-81
CD25	PC61	PE Cy7	BD Biosciences	552880
TCR $\gamma\delta$	GL3	PerCP Cy5.5	Biolegend	118118
B220 (CD45R)	RA3-6B2	PE	Biolegend	103208
NK1.1	PK136	PE	eBioscience	12-5941-83
CD11c	N418	PE	eBioscience	12-0114-82
CD11b (Mac1)	M1/70	PE	BD Biosciences	553311
Thymocyte stages TCR-β				
CD4	GK1.5	APC Cy7	Biolegend	100414
CD8a	53-6.7	BV421	Biolegend	100738
CD44	IM7	APC	eBioscience	17-0441-81
CD25	PC61	PE Cy7	BD Biosciences	552880
B220 (CD45R)	RA3-6B2	FITC	Biolegend	103206
NK1.1	PK136	FITC	Biolegend	108706
CD11c	HL3	FITC	BD Biosciences	557400
CD11b (Mac1)	M1/70	FITC	BD Biosciences	557396
TCR $\gamma\delta$	GL3	PerCP Cy5.5	Biolegend	118118
TCR- β	H57-597	PE	eBioscience	12-5961-83
Thymocyte stages sort				
CD4	GK1.5	APC Cy7	Biolegend	100414
CD8a	53-6.7	BV421	Biolegend	100738
CD44	IM7	PE	BD Biosciences	553134
CD25	PC61	Alexa Fluor 647	Biolegend	102020
B220 (CD45R)	RA3-6B2	FITC	Biolegend	103206
TCR $\gamma\delta$	eBioGL3	FITC	eBioscience	11-5711-82
NK1.1	PK136	FITC	Biolegend	108706
CD11c	HL3	FITC	BD Biosciences	557400
CD11b (Mac1)	M1/70	FITC	BD Biosciences	557396
B cell stage analysis				
IgM	II/41	FITC	BD Biosciences	553437
BP-1 (Ly51)	6C3	PE	Thermo Fischer	12-5891-83
CD43	S7	APC	BD Biosciences	560663
B220 (CD45R)	RA3-6B2	PE Cy7	eBioscience	25-0452-82
CD24	M1/69	eF450	eBioscience	48-0242-82
Annexin V detection in B cells in the bone marrow				
IgM	II/41	PE	eBioscience	12-5790-81
CD43	S7	APC	BD Biosciences	560663
B220 (CD45R)	RA3-6B2	PE Cy7	eBioscience	25-0452-82

CD24	M1/69	eF450	eBioscience	48-0242-82
CD3	145-2C11	PerCP Cy5.5	Biolegend	100328
Peripheral lymphocyte analysis				
CD4	GK1.5	APC Cy7	Biolegend	100414
CD8a	53-6.7	BV421	Biolegend	100738
TCR $\gamma\delta$	GL3	PerCP Cy5.5	Biolegend	118118
B220 (CD45R)	RA3-6B2	FITC	Biolegend	103206
CD11c	HL3	FITC	BD Biosciences	557400
CD11b (Mac1)	M1/70	FITC	BD Biosciences	557396
CD3e	145-2C11	APC	Thermo Fischer	17-0031-82
α -GalCer Loaded CD1 tetramer		PE	ProImmune	E001-2X
EdU incorporation in Thymus				
CD4	GK1.5	APC Cy7	Biolegend	100414
CD8a	53-6.7	BV421	Biolegend	100738
CD44	IM7	APC	eBioscience	17-0441-81
CD25	PC61	BV605	Biolegend	102035
Sorting of CFSE labelled thymocytes				
CD8a	53-6.7	PE Cy7	eBioscience	25-0081-82
CD4	GK1.5	PE Dazzle	Biolegend	100455
TCR β	H57-597	BV421	Biolegend	109229
CD62L	MEL-14	Alexa Fluor 700	Biolegend	104426
CD44	IM7	APC	eBioscience	17-0441-81
CD19	eBio1D3	PE	eBioscience	12-0193-83
HSC-MPP phenotyping				
CD3e	145-2C11	FITC	Biolegend	100306
B220 (CD45R)	RA3-6B2	FITC	Biolegend	103206
CD11b (Mac1)	M1/70	FITC	BD Biosciences	557396
Gr1 (Ly-6G/Ly-6C)	RB6-8C5	FITC	Biolegend	108406
TER-119	TER-119	FITC	eBioscience	11-5921-81
Sca1 (Ly-6A/E)	D7	APC	eBioscience	17-5981-81
Ckit (CD117)	2B8	BV421	Biolegend	105827
Flk2 (CD135)	A2F10.1	PE	BD Biosciences	561068
CD34	RAM34	Alexa Fluor 700	eBioscience	56-0341-82
CD150 (SLAMF7)	TC15-12F12.2	BV605	Biolegend	115927
CD48	HM48-1	APC Cy7	Biolegend	103431
CLP phenotyping				
CD3e	145-2C11	FITC	Biolegend	100306
B220 (CD45R)	RA3-6B2	FITC	Biolegend	103206
CD11b (Mac1)	M1/70	FITC	BD Biosciences	557396
Gr1 (Ly-6G/Ly-6C)	RB6-8C5	FITC	Biolegend	108406
TER-119	TER-119	FITC	eBioscience	11-5921-81
Sca1 (Ly-6A/E)	D7	APC	eBioscience	17-5981-81
Ckit (CD117)	2B8	BV421	Biolegend	105827
IL-7Ra	A7R34	PE Dazzle	Biolegend	135031
Flk2 (CD135)	A2F10	PerCP efluor710	eBioscience	46-1351-80
Peripheral blood reconstitution				
CD45.1	A20	PE	eBioscience	12-0453-83
CD45.2	104	BV421	Biolegend	109831
CD11b (Mac1)	M1/70	FITC	BD Biosciences	557396
Gr1 (Ly-6G/Ly-6C)	RB6-8C5	FITC	Biolegend	108406
CD3e	145-2C11	APC	Thermo Fischer	17-0031-82
B220 (CD45R)	RA3-6B2	PE Cy7	eBioscience	25-0452-82
HSC engraftment analysis				
CD3e	145-2C11	FITC	Biolegend	100306
B220 (CD45R)	RA3-6B2	FITC	Biolegend	103206

CD11b (Mac1)	M1/70	FITC	BD Biosciences	557396
Gr1 (Ly-6G/Ly-6C)	RB6-8C5	FITC	Biolegend	108406
Ter119	TER-119	FITC	eBioscience	11-5921-81
Sca1 (Ly-6A/E)	D7	PE Cy7	Biolegend	108113
CD48	HM48-1	PerCP efluor710	eBioscience	46-0481-82
CD150 (SLAMF6)	TC15-12F12.2	PE Dazzle	Biolegend	115935
Flk2 (CD135)	A2F10.1	PE	BD Biosciences	561068
CD45.2	104	Alexa Fluor 700	Biolegend	109822
CD45.1	A20	APC Cy7	Biolegend	110716

Thymocyte engraftment analysis

CD4	GK1.5	PE Cy7	Biolegend	100422
CD8	53-6.7	BV421	Biolegend	100738
CD44	IM7	PE	BD Biosciences	553134
CD25	PC61	BV605	Biolegend	102035
TCR $\gamma\delta$	GL3	PerCP Cy5.5	Biolegend	118118
B220 (CD45R)	RA3-6B2	FITC	Biolegend	103206
NK1.1	PK136	FITC	Biolegend	108706
CD11c	HL3	FITC	BD Biosciences	557400
CD11b (Mac1)	M1/70	FITC	BD Biosciences	557396
CD45.2	104	Alexa Fluor 700	Biolegend	109822
CD45.1	A20	APC Cy7	Biolegend	110716

Analysis of B cell engraftment in the bone marrow

IgM	II/41	FITC	BD Biosciences	553437
CD43	S7	PE	BD Biosciences	553271
B220 (CD45R)	RA3-6B2	PE Cy7	eBioscience	25-0452-82
CD24	M1/69	eF450	eBioscience	48-0242-82
CD45.2	104	Alexa Fluor 700	Biolegend	109822
CD45.1	A20	APC Cy7	Biolegend	110716

Analysis of engraftment of peripheral leukocytes

CD4	GK1.5	PE Cy7	Biolegend	100422
CD8	53-6.7	BV421	Biolegend	100738
B220 (CD45R)	RA3-6B2	FITC	Biolegend	103206
TCR $\gamma\delta$	GL3	PerCP Cy5.5	Biolegend	118118
CD11b (Mac1)	M1/70	PE	BD Biosciences	553311
Gr1 (Ly-6G/Ly-6C)	RB6-8C5	PE	Biolegend	108408
CD45.2	104	Alexa Fluor 700	Biolegend	109822
CD45.1	A20	APC Cy7	Biolegend	110716

Transparent Methods

Fish lines. The zebrafish (*D. rerio*) wild-type strain TLEK (Tüpfel long fin/Ekkwill), the *ikaros:eGFP* transgenic zebrafish (Bajoghli et al., 2009), the medaka (*O. latipes*) strain cab and the *rag1:eGFP* transgenic medaka (Li et al., 2007) are maintained in the animal facility of the Max Planck Institute of Immunobiology and Epigenetics. The medaka Kaga (strain ID: IB833) and the *gyokuro* (strain ID: MT369) strains (Furutani-Seiki et al., 2004; Iwanami et al., 2008; Iwanami et al., 2004) were supplied by the National Bioresource Project (NBRP) Medaka (<https://shigen.nig.ac.jp/medaka/>). All animal experiments involving medaka and zebrafish were approved by the institute's review committee and conducted under licenses from the local government (Regierungspräsidium Freiburg; license 35-9185.81/G-15/115). The developmental stages of medaka were designated as described (Iwamatsu, 2004).

ENU mutagenesis of zebrafish and recovery of *dnmt1* mutant. The IY071 (allele t^{25501}) mutant (Iwanami et al., 2016) was established in collaboration with Tübingen 2000 Screen consortium (Boehm et al., 2003).

ENU mutagenesis of medaka and recovery of *dnmt1* mutant. For meiotic recombination mapping of the responsible mutation in the *gyokuro* line (*j48-12B*), the panels of PCR length polymorphism (PLP) markers in M-marker 2009 (Kimura and Naruse, 2010) were used for bulk segment analysis using DNA isolated from 21 *gkr* embryos and 38 wild-type siblings derived from an *gkr/cab* x kaga mapping cross. 103 *gkr* embryos were analyzed using markers described in MLBase (http://mbase.nig.ac.jp/mbase/ml_base.html) and custom-made PLP primers on chromosome 8. Primers used are listed in Table S2. In order to identify the mutation, genomic DNA was extracted from 87 *gkr* and 82 siblings derived from a *gkr/cab* x kaga mapping cross at stage 35 after *in situ* hybridization using a *rag1*-specific probe. After whole genome sequencing (Iwanami et al., 2016), sequencing reads were mapped to the medaka reference genome (MEDAKA1; Ensembl release 72) and any polymorphisms in the coding regions of the genes in the critical chromosomal interval between *gkr* mutants and wild-type siblings were recorded. Of the identified variants, those found in mutant DNA but identical to the reference sequence were not considered further, because the reference HdrR strain is closely related to the cab strain. To safeguard against the possibility of misphenotyped embryos included in the pools used for preparation of DNA, more than 80% of reads from the mutant pool and less than 50% of reads from wild-type pool needed to differ from the reference sequence, with an additional requirement that the coverage was at least 10 reads from each pool. No mutation was found in splice donor and acceptor sites; however, one nonsense and 36 missense mutations were identified (Table S1). The derived protein sequences of human, mouse, and zebrafish orthologs of candidate genes were aligned with medaka sequences to examine the degree of conservation of the mutated residues among these species (Table S1). Using this criterion, only two missense mutations (K850E in *dnmt1* and F809L in *dnm2a*) were deemed to be interesting candidates. Since there is a paralog of *dnm2a*, *dnm2*, located on chromosome 1 with possible associated functional redundancy, we focused on the *dnmt1* K850E missense mutation as the one responsible for the *gkr* phenotype.

RNA *in situ* hybridization. For hybridization using whole embryos or tissue slides, the relevant zebrafish probes were taken from (Schorpp et al., 2006; Soza-Ried et al., 2010). For the medaka *mpx* probe, nucleotides 645-1453 in Genbank accession number XM_004074804 sequence were cloned into pGEM-T easy vector (Promega). After linearization of the plasmid with *SacI*, *in vitro* transcription was carried out with T7 polymerase using DIG RNA labeling kit (Roche). The other medaka probes were taken from (Bajoghli et al., 2009; Li et al., 2007).

Histological analysis. Histological analysis of fish specimens followed (Schorpp et al., 2006).

RNA extraction and cDNA synthesis. Total RNA was extracted using TRI Reagent (Sigma) following the manufacturer's instructions. After treatment with DNaseI (Promega), RNA extraction using TRI Reagent was repeated. Superscript III Reverse Transcriptase (Invitrogen) and random hexamers were used for cDNA synthesis from total RNA.

RT-PCR. For RT-PCR to examine expression of zebrafish genes, primer sequences were taken from (Hess and Boehm, 2016; Hess et al., 2013; Schorpp et al., 2006). Cycle numbers were as follows; *ef* (*elongation factor*): 30 cycles, *cd79b* and *lck*: 35 cycles, *igm* (igV_{H1}-Cm): 30 cycles plus nested 30 cycles.

Quantitative PCR. qPCR was carried out (Boehm et al., 2003); (Iwanami et al., 2011); (Iwanami et al., 2016) using SYBR Premix Ex Taq (Takara) and 7500 fast real-time PCR system (Applied Biosystems). Zebrafish *actb1* and mouse *Hprt* were used as reference genes. The primer sets for zebrafish and mouse genes were purchased from BioRad.

Expression of Dnmt1 protein variants. The cloning and expression of Dnmt1 variants (aa 291-1620) followed (Vilkaitis et al., 2005; Berkyurek et al., 2014).

***in vitro* methylation assay.** The enzymatic activities of Dnmt1 variants were assayed *in vitro* following (Sugiyama et al., 2010; Garvilles et al., 2016), using unmethylated and hemi-methylated substrates (Suetake et al., 2006). Briefly, an annealed oligonucleotide (100 nM) was methylated with recombinant Dnmt1 (2 nM) in the presence of 2.2 μM [3H]-S-adenosyl-methionine (PerkinElmer Life Sciences) in 25 μl of buffer (5 mM EDTA, 20% glycerol, 0.2 mM DTT, 0.2 mM phenylmethylsulfonyl fluoride, and 20 mM Tris-HCl, 7.4) at 37°C for 1 h. The radioactivity incorporated into the DNA was measured using a liquid scintillation counter.

Flow cytometry of zebrafish cells. Flow cytometric analysis of light-scatter characteristics of WKM cells followed (Traver et al., 2003); staining with hydroxystilbamidine (Enzo Life Sciences; final concentration 1 μg/mL) was used to exclude dead cells.

Zebrafish transplantation experiments. Transplantation of whole kidney marrow cells of zebrafish followed (Hess et al., 2013).

Imaging of zebrafish and medaka specimens. Embryos and larvae were anesthetized and immobilized in 3% methylcellulose. Fluorescence microscopy was performed with Zeiss Imager.Z1.

Mouse lines. *Dnmt1* mutants (*Dnmt1* K979E; *Dnmt1* N1510K; *Dnmt1* H1511D) were generated by CRISPR-Cas9 technology and locus-specific sgRNAs (Table S2). Sequence-specific single-stranded repair oligonucleotides (Table S2) contained additional mismatches to avoid cleavage by the pre-assembled sgRNA/Cas9 RNP. The targeting sequences for guide RNAs were designed according to (Hwang et al., 2013) and cloned into the pDR274 vector (Addgene plasmid #42250). After digestion with *DraI* restriction enzyme (New England Biolabs), sgRNA was generated by *in vitro* transcription using MAXIscript T7 Transcription Kit (Thermo). For injection into fertilized eggs, two sgRNAs per target site were combined (final concentration 25 ng/μL), recombinant Cas9 protein from *Streptococcus pyogenes* (PNA Bio; final concentration 50ng/μL), and repair oligonucleotide (final concentration 5 μM) were mixed on ice in 10mM Tris, pH 7.5; 0.15mM EDTA; approximately 1–2 nL of the solution were injected per fertilized egg. When introduced into Balb/c, C57BL/6 and FVB genetic backgrounds, homozygosity of K979E or N1510K led to embryonic lethality. For the K979E mutant, the following results were obtained. Balb/c background: (a) Newborn mice: 19 mice of +/+ genotype; 17 mice of +/- genotype; 0 mice of m/m genotype. (b) E12.5 mouse embryos: 10 mice of +/+ genotype; 9 mice of +/- genotype; 0 mice of m/m genotype. C57BL/6 background: (a) Newborn mice: 18 mice of +/+ genotype; 25 mice of +/- genotype; 0 mice of m/m genotype. FVB background: No m/m mice were ever observed, but precise numbers of mice of +/+ genotype and +/- genotype were not determined. For the N1510K mutant, the following results were obtained. Balb/c background: (a) E8.5 mouse embryos: 2 mice of +/+ genotype; 13 mice of +/- genotype; 2 mice of m/m genotype. (b) E10.5 mouse embryos: 5 mice of +/+ genotype; 9 mice of +/- genotype; 0 mice of m/m genotype. C57BL/6 background: (a) E10.5 mouse embryos: 2 mice of +/+ genotype; 9 mice of +/- genotype; 2 mice of m/m genotype. (b) E13.5-E15.5 mouse embryos: 8 mice of +/+ genotype; 24 mice of +/- genotype; 0 mice of m/m genotype. Balb/c, C57BL/6, and FVB background: Newborn mice: No m/m mice were ever observed, but precise numbers of mice of +/+ genotype and +/- genotype were not determined. By contrast, the Balb/c background supported the survival of mice homozygous for the H1511D mutation. All mice were kept in the animal facility of the Max Planck Institute of Immunobiology and Epigenetics under specific pathogen-free conditions. All animal experiments were performed in

accordance with the relevant guidelines and regulations, approved by the review committee of the Max Planck Institute of Immunobiology and Epigenetics and the Regierungspräsidium Freiburg, Germany (license AZ 35-9185.81/G-15/35).

Flow cytometry and cell sorting of mouse cells. Phenotyping of lymphocyte populations was performed by flow cytometry after preparation of single cell suspensions from lymphoid organs and staining using antibodies listed in Table S5. Single cell suspension of thymus and spleen were prepared in FACS buffer (2% FBS, 1 mM ethylenediaminetetraacetic acid (EDTA), 1% penicillin- streptomycin, in PBS) by tissue homogenization with a syringe plunger against a 40 μm cell strainer. For preparation of cell suspensions from bone marrow femur, tibia and pelvis were flushed with FACS buffer using a 10 ml syringe and a 26 gauge-needle and then passed through 40 μm cell strainer to obtain single cell suspensions. For red blood cell lysis, the cell suspensions were treated with ACK lysis buffer (0.15M NH_4Cl , 10mM KHCO_3 , 0.1mM EDTA in H_2O , pH 7.2-7.4), washed and resuspended in FACS buffer. Data were collected on a LSRFortessa and/or LSRII apparatus (BD Biosciences) and were analysed with FlowJo software version 10; cell sorting was done using a FACSAria instrument (BD Biosciences). In some experiments, the MojoSort mouse hematopoietic progenitor cell isolation kit (BioLegend) was used in order to analyze or sort lineage-negative bone marrow cells.

EdU staining and cell cycle analysis of mouse cells. For *in vivo* cell cycle analysis of thymocytes, mice received a single intra-peritoneal injection of EdU (5-ethynyl-2'-deoxyuridine) diluted in PBS at a dose of 50 mg/kg body weight. After 16h of EdU exposure, the mice were sacrificed and single cell suspensions of bone marrow and thymus were prepared. Single cell suspensions of EdU labeled cells were processed using the Click-iT EdU Flow Cytometry Assay (Thermo Fisher) according to the manufacturers protocol. Briefly, the cells were washed with 1% BSA in PBS and stained with surface antibodies (listed in Table S5) at 4 °C. The cells were washed again and fixed in Click-iT fixative for 15 min, washed once and then permeabilized in saponin-based permeabilization and wash reagent for 15 min. The Click-iT reaction cocktail was added to the cells for additional 30 min incubation. The cells were finally washed and analyzed by flow cytometry. All steps after the cell surface antibody staining were carried out at room temperature.

Annexin V apoptosis detection in developing mouse lymphocytes. The Annexin V (FITC) apoptosis detection Kit (eBioscience) was used for detection of apoptotic cells during development of T and B lymphocytes. Cell suspensions prepared from thymus and bone marrow were stained with surface antibodies (listed in Table S5) in FACS buffer (2% FBS, 1 mM EDTA, 1% penicillin- streptomycin, in PBS). After washing with 1 X diluted binding buffer (in dH_2O) they were resuspended again in 1X binding buffer and incubated with FITC-conjugated Annexin V for 15 minutes at room temperature. The cells were finally washed with 1X binding buffer, resuspended in 1X binding buffer containing propidium iodide, and analyzed by flow cytometry.

Competitive bone marrow reconstitution in mice. Bone marrow (BM) cell suspensions from CD45.2 *Dnmt1^{m/m}* and CD45.1/CD45.2 wild-type mice on the Balb/c background (7-12 weeks of age) were prepared and an aliquot was stained with surface antibodies for flow cytometric determination of HSC numbers. A bone marrow cell suspension of the desired donors, each containing 100 HSCs, were combined from the required sources before transplantation (wild-type donor [Ly5.1/Ly5.2]; 1.4×10^6 cells; H1511D/H1511D mutant donor [Ly5.2/Ly5.2]; 2.7×10^6 cells). The mixtures of CD45.1/CD45.2 wild-type and CD45.2 *Dnmt1^{m/m}* bone marrow cells were resuspended in 100 μL PBS and transplanted into lethally irradiated (a total of 9.5 Gy delivered in two doses of 5 Gy and 4.5 Gy, separated by a 3h interval) CD45.2 recipient mice (7-12 weeks of age) by tail vein injection. Peripheral blood was obtained from the recipient mice at 4 and 8 weeks after transplantation. Following red blood cell lysis, the contributions of CD45.2⁺ donor derived cells to the T cell (CD3⁺), B cell (B220⁺) and myeloid (CD11b/Gr1⁺) lineages were assessed by flow cytometry. At 12 weeks after transplantation, the mice were sacrificed and BM, thymus and spleen were collected for flow cytometric analysis using the antibodies listed in Table S5.

RNA_seq analysis of lineage-negative mouse bone marrow cells. For RNA-seq analyses, HSCs (n=3 biological replicates), MPP1 (n=2), and MPP4 (n=4) cells were isolated from bone marrow cells of *Dnmt1^{+/+}* and *Dnmt1^{m/m}* animals. Raw sequencing reads were processed with the snakePipes (Bhardwaj et al., 2019) version 0.5 scRNAseq-mapcount workflow to produce a count table per sample. Read trimming

was enabled, and annotation was filtered to remove entries with keyword "decay" or "pseudogene" in biotype description. Postprocessing and differential gene expression analysis were performed in R version 3.4.0. Counts were normalized by downscaling to 200,000 transcripts per cell with RaceID3 (Herman et al., 2018) and rounded to integer counts. Outlier samples were removed. Differentially expressed genes were called between *Dnmt1* mutant and wildtype replicates in each cell population separately, using R packages zingeR [<https://github.com/statOmics/zingeR>] version 0.1.0 and DESeq2 (Love et al., 2014) version 1.18.1. Sequencing batch was not included in the model. Gene lists were filtered for FDR<5% and absolute log₂ fold change of at least 1.

Whole genome bisulfite sequencing. Genomic DNA was extracted from three *dnmt1*^{+/+} and three *dnmt1*^{m/m} zebrafish at 18 dpf using the DNeasy blood and tissue kit (Qiagen). 1µg and 0.5µg of DNA was used for bisulfite reactions and library construction using the TruSeq DNA PCR-free library preparation kit (Illumina) and the EpiGnome Methyl-Seq kit (Epicentre), respectively. The fragments were sequenced in paired-end 100bp mode on 1 lane of Illumina HiSeq 2500 instrument.

Zebrafish whole genome methylation analysis. Raw sequencing reads were trimmed with cutadapt version 1.9.1 (Martin, 2011) as follows: First, 2 (TruSeq) or 6 (EpiGnome) 5'-most nucleotides were hard-trimmed and Illumina adapter sequences removed. Bisulfite-specific operations on reads and reference genome were performed with methylCtools version 0.9.4 (Hovestadt et al., 2014). Bisulfite-converted reads were mapped to bisulfite-converted GRCz10 zebrafish genome with bwa-mem version 0.7.12 separately for the two library types. Back-converted bam files were sorted with samtools version 1.3.1, PCR duplicates removed and read group information added with Picard tools v1.136. The two resulting bam files per sample were merged with samtools and methylation bias profiled with MethylDackel v0.1.7 [<https://github.com/dpryan79/MethylDackel>]. Methylated and unmethylated read counts per CpG position were extracted with methylCtools v0.9.4 with mapping quality threshold of 10, SNP detection, counting only 1 of two overlapping paired end reads, skipping 5 nucleotides from each read length and zero-padding of uncovered positions.

Medaka amplicon methylation analysis. The medaka CpG islands orthologous to the eight zebrafish CpG islands described by Potok et al. (Potok et al., 2013) were identified as described below. Primers for medaka amplicons were designed using Bisulfite Primer Seeker 12S (Zymo Research) and are listed in Supplementary Table 2. Genomic DNA was extracted from three *dnmt1*^{+/+} and three *dnmt1*^{m/m} medaka at 18 dpf using the DNeasy blood and tissue kit (Qiagen). 150ng of DNA was used for the bisulfite reaction using the EZ DNA methylation kit (Zymo Research) and the bisulfite-treated DNA was recovered in 15µl of elution buffer. 1µl of bisulfite-treated DNA solution was used for 30 cycles of PCR amplification with the indicated bisulfite primers in a reaction volume of 20µl; 1µl of the first PCR reaction was used for a nested PCR reaction with P5 and P7 adaptor-attached bisulfite primers. After the nested PCR reactions, amplicons were purified using PCR purification kit (Qiagen) and qualified and quantified by gel electrophoresis, the High Sensitivity DNA kit (Agilent) for the Agilent Bioanalyzer and the Qubit dsDNA HS assay kit (Invitrogen). The eight amplicons derived from the same original source were mixed at 1ng/µl each and 15µl of the mixture were amplified using P5 and P7 primers attached to barcodes using NEBNext DNA library prep kit (New England Biolabs), following the manufacturer's protocol. The amplicons were purified using AMPure XP (Beckman Coulter) and qualified and quantified using Bioanalyzer and Qubit dsDNA HS assay kit. Finally, all amplicons were mixed at 1ng/µl each and 5 µl of mixed amplicons were used for sequencing in paired-end 150bp mode on 1 lane of MiSeq Sequencing System (Illumina). Raw sequencing reads were trimmed with cutadapt version 1.9.1 (Martin, 2011), removing the first two 5'-most nucleotides and Illumina adapter sequences. Base quality trimming was performed using prinseq lite version 0.20.4. Bisulfite-specific mapping to bisulfite-converted MEDAKA1.72 genome was performed with Bismark v0.14.6 and Bowtie2 v2.2.8 in non-directional mode allowing for dovetail alignments. Samtools sorted and indexed bams were used for methylation bias calculation (MethylDackel v0.1.7) and read count extraction in target regions with methylCtools version 0.9.4 (Hovestadt et al., 2014), using MAPQ threshold 10, SNP detection, counting one of two overlapping mates, skipping 5 nucleotides from read ends, and zero-padding.

Post-processing and statistical analysis of methylation data in selected target regions of fish genomes.

All postprocessing and statistical analysis were performed in R version 3.2.3. Read counts for CpG positions with illegal nucleotide frequency of at least 0.25 were set to NA. Detailed methylation analysis was performed on CpG sites from main chromosomes only. Methylated as well as unmethylated read counts from the two strands were summed per CpG position; coverage was calculated as sum of methylated and unmethylated reads per position on both strands. Methylation ratio (beta value) was calculated as ratio of methylated reads to coverage per site. Beta values for sites with coverage less than 10 reads were set to NA. A methylation matrix was obtained by merging methylation values for single samples with CpG positions in rows and samples in columns.

Zebrafish CpG methylation values obtained through whole genome bisulfite sequencing were intersected using bedtools version 2.23.0 with genomic intervals selected from Supplementary Table S4 of Potok et al. (Potok et al., 2013) lifted over to GRCz10 on the UCSC webpage. Single CpG methylation values were aggregated to mean values per interval allowing for maximum 1 NA in the total of six samples, and at least 20% of non-NA CpGs per interval. Mean beta values per interval were plotted for every replicate in each genotype group with ggplot2 version 2.1.0.

Medaka CpG methylation values were obtained through amplicon sequencing of the selected zebrafish genomic regions indicated above lifted over to MEDAKA1 using UCSC tools web service and corrected for primer placement. Single CpG methylation values were aggregated to mean values per interval allowing for a maximum of 1 NA in a total of 6 samples, and at least 20% of non-NA CpGs per interval. Mean beta values per interval were plotted for every replicate in each genotype group with ggplot2 version 2.1.0.

Between-group statistical testing of methylation ratio differences in zebrafish and medaka was performed on logit-transformed interval mean beta values (adjust offset 0.025, R package cars version 2.1-1). P values from two-sided two-sample t-test (with 3 replicates per genotype group) are reported.

DNA methylation analysis of mouse Lin⁻ cells. Raw sequencing reads were processed with the snakePipes (Bhardwaj et al., 2019) version 1.2.0 WGBS workflow providing mm10 genome to produce bam files filtered for PCR duplicates. This was done separately for reads obtained with each library preparation kit. The resulting bam files were merged per sample. Merged bams from the previous step were processed with the snakePipes (Bhardwaj et al., 2019) version 1.2.0 WGBS workflow using the `-fromBam` argument and providing a sample sheet. Gencode version m9 gene models were used. This produced a.o. a list of differentially methylated regions between the mutant and the wildtype sample groups filtered for FDR<2% and absolute methylation difference between groups of at least 20%. Bigwig tracks with WGBS methylation and coverage were generated with snakePipes (Bhardwaj et al., 2019).

Reanalysis of publicly available mouse hematopoietic progenitor cell ChIP-seq data. ChIP reads for mouse HSC H3k27me3 (Sun et al., 2014) were downloaded from ENA (GSE47765), as were ChIP reads for mouse LT/ST-HSC and MPP H3K4me3, H3K4me1 and H3K27ac (Lara-Astiaso et al, 2014; GSE59636), as well as ATAC-seq reads for mouse LT/ST-HSC, MPP3/4 and proB.CLP (Yoshida et al., 2019; GSE100738). Fastq files were merged for multiple runs of the same sample.

Reads were mapped to mouse genome GRCm38 and processed with snakePipes (Bhardwaj et al., 2019) version 2.1.1 DNA-mapping and ATAC-seq workflows to produce bigwig files with normalized coverage. Genomic tracks plots for specific genes were obtained with pyGenomeTracks (Ramirez et al., 2018) version 3.2. To align CHIP and ATAC_seq reads to DMRs, normalized coverage bigwig files were obtained for H3K27me3, H3K4me3, H3K4me1 and H3K27ac marks and for open chromatin, as described above, gencode version m9 gene model gtf for GRCm38, bed files with DMR positions, bigwig files with filtered CpG coverage for WT and Mut replicates, bigwig files with methylation value (0-100%) for wild-type and mutant replicates.

The plot for *Fos* was generated for genomic interval chr12:85423890-85527273 (gene locus + 50kb flanks); for *Fosb*, chr7:19301696-19311051 (1 kb flanks); for *ikzf1*, chr11:11684003-11773926 (1kb flanks); for *Ebfl*, chr11:44616317-45009091 (1kb flanks).

Statistical analysis of animal phenotypes. No randomization of animals was done in the present studies; phenotypes were recorded by a blinded observer before genotyping. In zebrafish and medaka, the sex of animals can only be determined in late adult stage; therefore, in the present studies, no consideration was given to sex. No animals were excluded from analyses. Samples size was estimated from the degrees of variability in previous analyses (Boehm et al., 2003); (Iwanami et al., 2016); (Schorpp et al., 2006) in order

to be able to detect biologically meaningful differences in examined parameters, usually 20% difference from control values. t-tests were performed for samples with equal variance; otherwise, F-tests were used.

Molecular dynamics simulation. Homology models were constructed using the structure of the complex of mouse Dnmt1 with a hemi-methylated CpG DNA substrate (PDB ID: 4da4) (Song et al., 2012) using SWISS-MODEL (Waterhouse et al., 2018). Missing residues in the protein were modeled using the Modellar program (Sali and Blundell, 1993), and missing hydrogen atoms were added under conditions of pH 7. For stabilization and equilibration of homology models, structures were set at the center of a rectangular box and were solvated in a rectangular box filled with TIP3P (Jorgensen et al., 1983) model waters; appropriate numbers of sodium and chloride ions were added to achieve charge neutral systems with an ionic concentration of 150 mM. The number of water molecules was defined to be about 17,500. The closest distance between the protein and the rectangular box was set to 17 Å. The force field of proteins and ions were calculated using amber03 force field (Duan et al., 2003); the water molecules were rigidified using the SETTLE algorithm (Miyamoto and Kollman, 1992); electrostatic interactions were calculated using the particle-mesh-Ewald method (Essmann et al., 1995) using a cutoff radius of van der Waals interactions of 8 Å. Adequate restrained potentials were allied to the heavy atoms at the original positions. The total potential energy was minimized using the steepest decent method. The system was then simulated using a constant number of atoms at a constant temperature (310K) with the Berendsen algorithm (Berendsen et al., 1984). Chemical bonds in the protein were treated as rigid using the linear constraint solver (LINCS) algorithm (Hess, 2008). The time step was set to 2 fs, and the simulation with a constant number of atoms, constant pressure (1 atm), and constant temperature (310K) was conducted using the Parrinello-Rahman barostat (Parrinello and Rahman, 1980) and Bussi–Donadio–Parrinello thermostat (Bussi et al., 2007). The restrained potential was then released, and a production run was initiated. Coordinates of the whole system were recorded every 2 ps for the analysis; however, the first 5-ns trajectories were used to equilibrate the system and were then discarded. Molecular dynamic simulations were performed using the Gromacs-4.5.5 program package (Pronk et al., 2013). For the modeling of the medaka protein, the dnmt1 variant X1 was used (accession number XM_023957476.1; XP_023813244.1) comprising a total of 1,496 amino acids; mouse residue K979 corresponds to medaka residue K850, and zebrafish N1391 corresponds to medaka N1386.

Data availability. The R code necessary to reproduce statistical analyses and results is reported in Supplemental Code available at https://github.com/katsikora/Iwanami2019_SupplementaryCodeAndData_B. The methylation and RNA_seq data have been deposited in the GEO database under accession number GSE98648.

Supplemental references

Berendsen, H.J.C., Postma, J.P.M., Gunsteren van, W.F., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* *81*, 3684-3690.

Berkyurek, A.C., Suetake, I., Arita, K., Takeshita, K., Nakagawa, A., Shirakawa, M., and Tajima, S. (2014) The DNA Methyltransferase Dnmt1 Directly Interacts with the SET and RING Finger-associated (SRA) Domain of the Multifunctional Protein Uhrf1 to Facilitate Accession of the Catalytic Center to Hemi-methylated DNA. *J Biol Chem* *289*, 379-386.

Bhardwaj, V., Heyne, S., Sikora, K., Rabbani, L., Rauer, M., Kilpert, F., Richter, A.S., Ryan, D.P., and Manke, T. (2019). snakePipes: facilitating flexible, scalable and integrative epigenomic analysis. *Bioinformatics* *35*, 4757-4759.

Boehm, T., Bleul, C.C., and Schorpp, M. (2003). Genetic dissection of thymus development in mouse and zebrafish. *Immunol Rev* *195*, 15-27.

Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *J Chem Phys* *126*, 014101.

Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., *et al.* (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* *24*, 1999-2012.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen L. G. (1995) A smooth particle mesh Ewald method. *J Chem Phys* *101*, 8577-8593.

Garvilles, R.G., Hasegawa, T., Kimura, H., Sharif, J., Muto, M., Koseki, H., Takahashi, S., Suetake, I., and Tajima, S. (2015) Dual Functions of the RFTS Domain of Dnmt1 in Replication-Coupled DNA Methylation and in Protection of the Genome from Aberrant Methylation. *PLoS One* *10*, e0137509.

Herman, J.S., Sagar, and Grun, D. (2018). FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat Methods* *15*, 379-386.

Hess, B. (2008). P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J Chem Theory Comput* *4*, 116-122.

Hovestadt, V., Jones, D.T.W., Picelli, S., Wang, W., Kool, M., Northcott, P.A., Sultan, M., Stachurski, K., Ryzhova, M., Warnatz, H.-J., *et al.* (2014). Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature* *510*, 537-541.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* *31*, 227-229.

Iwanami, N., Mateos, F., Hess, I., Riffel, N., Soza-Ried, C., Schorpp, M., and Boehm, T. (2011). Genetic evidence for an evolutionarily conserved role of IL-7 signaling in T cell development of zebrafish. *J. Immunol.* *186*, 7060-7066.

Jorgensen, W.L., Chandrasekhar, J., and Madura, J.D. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* *79*, 926-935.

Kimura, T., and Naruse, K. (2010). M-marker 2009, a marker set for mapping medaka mutants using PCR length polymorphisms with an automated microchip gel electrophoresis system. *Biotechniques* *49*, 582-583.

Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretzky, I. *et al.* (2014) Chromatin state dynamics

during blood formation. *Science* 345, 943-949.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10-12.

Miyamoto, S., and Kollman, P.A. (1992). SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *Journal of Computational Chemistry* 13, 952-962.

Parrinello, M., and Rahman, A. (1980). Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Physical Review Letters* 45, 1196-1199.

Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M.R., Smith, J.C., Kasson, P.M., van der Spoel, D., *et al.* (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29, 845-854.

Ramírez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., *et al.* (2018) High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun* 15, 189. doi:10.1038/s41467-017-02525-w

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234, 779-815.

Soza-Ried, C., Hess, I., Netuschil, N., Schorpp, M., and Boehm, T. (2010). Essential role of *c-myb* in definitive hematopoiesis is evolutionarily conserved. *Proc. Natl. Acad. Sci. USA* 107, 17304-17308.

Suetake, I., Hayata, D., and Tajima, S. (2006). The amino-terminus of mouse DNA methyltransferase 1 forms an independent domain and binds to DNA with the sequence involving PCNA binding motif. *J Biochem* 140, 763-776.

Sugiyama, Y., Hatano, N., Sueyoshi, N., Suetake, I., Tajima, S., Kinoshita, E., Kinoshita-Kikuta, E., Koike, T., and Kameshita, I. (2010). The DNA-binding activity of mouse DNA methyltransferase 1 is regulated by phosphorylation with casein kinase 1delta/epsilon. *Biochem J* 427, 489-497.

Sun, D., Luo, M., Jeong, M., Rodriguez, B. *et al.* (2014) Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell* 14, 673-688.

Traver, D., Paw, B.H., Poss, K.D., Penberthy, W.T., Lin, S., and Zon, L.I. (2003). Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat Immunol* 4, 1238-1246.

Vilkaitis, G., Suetake, I., Klimasauskas, S., and Tajima, S. (2005). Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *J Biol Chem* 280, 64-72.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., *et al.* (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46, W296-W303.

Yoshida, H., Lareau, C.A., Ramirez, R.N., Rose, S.A. *et al.* (2019) The cis-Regulatory Atlas of the Mouse Immune System. *Cell* 176, 897-912.