# A transgenic mouse model ubiquitously overexpressing *Dnmt1s* mRNA lacks increased protein levels

Authors: Christian Grosser[1], Lothar Vassen[2], Bernhard Horsthemke[1], Nicholas Wagner[1]

Affiliations:

(1) Institute of Human Genetics, Faculty of Medicine, University of Duisburg-Essen, Hufelandstrasse 55, D-45122 Essen, Germany

(2) Institute for Cell Biology, Faculty of Medicine, University of Duisburg-Essen, Hufelandstrasse 55, D-45122 Essen, Germany

Corresponding author:

Nicholas Wagner

Institute of Human Genetics

Faculty of Medicine

University of Duisburg-Essen

Hufelandstrasse 55

D-45122 Essen

Germany

Email: nicholas.wagner@uni-due.de

Phone: +49 201 / 723-4653

Fax: +49 201 / 723-5900

## Abstract

Several types of cancer are characterized by global hypomethylation accompanied by regional hypermethylation and overexpression of DNA methyltransferase (cytosine-5) 1 (*Dnmt1*). In addition to the established role of Dnmt1 as maintenance methyltransferase, it has been suggested that Dnmt1 might also methylate certain target sites *de novo*. We created a transgenic mouse model to investigate whether the overexpression of the somatic form of *Dnmt1*, *Dnmt1s*, is sufficient to cause erroneous methylation and disease. Because ubiquitous Dnmt1 overexpression has been reported to be embryonic lethal, we designed a CAG promoter-driven Cre-*loxP* conditional transgene containing a floxed *EGFP* sequence followed by the *Dnmt1s* coding sequence. The *EGFP* sequence is excised and transgenic *Dnmt1s* expression is activated at specific time points or in specific tissues depending on the *Cre* deleter strain used for cross-ins. Pronucleus injections with the *Dnmt1s* transgene construct resulted in six founder lines as verified by PCR, Southern blot and EGFP fluorescence. We performed cross-ins with a CMV-*Cre* deleter strain to clarify if ubiquitous overexpression of *Dnmt1s* alone causes the previously observed embryonic lethality. Unexpectedly, these cross-ins yielded viable recombined offspring that ubiquitously overexpressed *Dnmt1* mRNA at tissue-dependent levels of up to 229-fold. However, the recombined offspring did not significantly overexpress Dnmt1 protein and showed no apparent signs of disease or pathological phenotype. Here, we describe the establishment of our *Dnmt1s*-transgenic mouse model and propose possibilities for the absence of transgenic protein.

## Introduction

DNA methylation is an important part of epigenetic signaling and is involved in differentiation, X-inactivation and genomic imprinting (Bird 2002). During DNA replication, DNA methyltransferase (cytosine-5) 1 (Dnmt1) methylates newly synthesized cytosines in a hemimethylated CpG context, a process referred to as maintenance methylation (Bestor & Ingram 1983; Gruenbaum et al. 1982). However, various *in vitro* studies have also attributed *de novo* methylation activity to Dnmt1 (Christman et al. 1995; Okano et al. 1998; Yoder et al. 1997).

The *Dnmt1* gene is located on chromosome 9 in the mouse and has three mRNA isoforms, known as *Dnmt1o*, *Dnmt1p* and *Dnmt1s*, that differ in alternative first exons. *Dnmt1o* and *Dnmt1p* are sex-specific isoforms that are only expressed in the oocyte/zygote up to the 8-cell stage and in the pachytene spermatocyte, respectively. *Dnmt1s* is the ubiquitously expressed somatic isoform, and the only isoform that is expressed from the preimplantation stage onward and throughout development in human and murine adult somatic cells (Cirio et al. 2008; Kurihara et al. 2008; Mertineit et al. 1998; Yoder et al. 1996). Dnmt1s contains several domains for activity regulation at the N-terminal end and a catalytic center at the C-terminal end (Chuang et al. 1997; Gruenbaum et al. 1983; Leonhardt et al. 1992). The Dnmt1 protein is target of many different post-translational modifications, which are established by numerous proteins and affect either the activity or the stability of Dnmt1 (Qin et al. 2011).

Tumor cells are characterized by global hypomethylation of the genome (Ehrlich 2002; Feinberg & Vogelstein 1983), accompanied by region-specific hypermethylation (Baylin et al. 1986; Robertson & Jones 2000). Region-specific hypermethylation is often closely associated with epigenetic gene silencing of tumor suppressor genes (Greger et al. 1994; Jones & Baylin 2002). Dnmt1 protein overexpression is a hallmark of many cancers

and likely to be a contributor to DNA hypermethylation. Alterations in DNA methylation patterns can play an important role in the development of tumors, although it is uncertain if erroneous DNA methylation alone is sufficient to trigger cancer development (Bestor 2003; Jones & Baylin 2007; Lyko 2005). Further studies have suggested that *de novo* methylation has a causal role in the development of tumors, and that hypermethylation is already detectable in the earliest stages of tumor development (Chan et al. 2006; Myohanen et al. 1998). One *in vitro* study addressing *Dnmt1* overexpression by stable transfection of cDNA in murine NIH-3T3 cells indicated that Dnmt1-induced hypermethylation may promote cell transformation (Wu et al. 1993).

The only *in vivo* study of *Dnmt1* overexpression that we are aware of is a mouse model by Biniszkiewicz *et al.* that was created using a BAC transgene (Biniszkiewicz et al. 2002). This mouse model was capable of expressing all three *Dnmt1* isoforms and it had an embryonic lethal phenotype. Therefore, we created a conditional transgenic mouse model for *Dnmt1s* overexpression using the Cre-*loxP* system to avoid embryonic lethality and study long-term effects. To clarify if the somatic isoform of Dnmt1 alone causes the embryonic lethality observed in the mouse model by Biniszkiewicz *et al.* (Biniszkiewicz et al. 2002), we triggered ubiquitous overexpression of *Dnmt1s*. Unexpectedly, the resulting offspring was viable with high levels of *Dnmt1* mRNA, but no increase in Dnmt1 protein levels. Transgenic animals showed no apparent signs of developmental defects or cancer. Further (conditional) cross-ins were not performed due to the lack of transgenic protein. In this study, we report on the establishment of our *Dnmt1s*-transgenic mouse model and provide hypotheses for the absence of transgenic protein.

## Materials and Methods

### Ethics statement

Adult mice (C57BL/6J, C3H) were obtained from the Central Animal Facility of the University Hospital Essen and were maintained in a 12-h light/dark cycle with access to food and water *ad libitum*. All animal experiments were approved by the institutional animal care committee of the German government (*Landesamt für Natur, Umwelt und Verbraucherschutz*; LANUV G034/06).

### Construction of the *Dnmt1s* transgene cassette

The pVL1393 plasmid containing the murine tagged *Dnmt1s* cDNA (NM_010066.4) was kindly provided by S. Pradhan from New England Biolabs (Ipswich, MA, USA). The tagged *Dnmt1s* cDNA sequence was excised from the pVL1393 vector and cloned into the pLCMV-*ECFP-loxP* plasmid after the second *loxP* site, resulting in pLCMV-*ECFP-loxP-Dnmt1s*. The CMV promoter of the pLCMV-*ECFP-loxP-Dnmt1s* plasmid was replaced with the CMV early enhancer/chicken β-actin (CAG) promoter to achieve a stronger ubiquitous expression of the transgene. Additionally, the *ECFP* sequence was replaced with an *EGFP* sequence resulting in the final version of the *Dnmt1s* transgene plasmid, which we called pLCAG-*EGFP-loxP-Dnmt1s* (Fig. 1).

## Cell culture

All cell culture media and supplements were obtained from Life Technologies (Darmstadt, Germany). NIH-3T3 murine fibroblast cells (ATCC 1658, adherent, DMEM) were cultivated in 75 $cm^2$ culture flasks with 12 ml culture medium at 37 °C in an incubator at 5% $CO_2$ and a relative humidity of 90%. All procedures were performed under a laminar flow hood. All media and reagents were autoclaved or sterile-filtered before use.

## Transient transfection of plasmid DNA into 3T3 murine fibroblasts

Roti-Fect reagent (Carl Roth, Karlsruhe, Germany) was used for transient transfection of NIH-3T3 cells with slight alterations to the supplier's protocol. 20 µg of plasmid DNA was transferred into a 1.5 ml reaction tube and filled to 400 µl with culture medium. In a separate 1.5 ml reaction tube 90 µl of Roti-Fect were mixed with 310 µl culture medium. For co-transfections, 20 µg of each plasmid DNA was used, and the volumes of Roti-Fect and culture medium were doubled. Both tubes were combined and incubated for 30 min at RT. During incubation, cells were washed twice with PBS and 8 ml culture medium was added. After incubation, the DNA-Roti-Fect mixture was added to the cultures and mixed by gently tilting the flask. Transfected cultures were incubated for 6 h at standard conditions, washed once with PBS, and incubated in culture medium (without transfection reagent) for a further 18 h before preparation of whole cell protein extracts. Successful transfection was checked by replacing the culture medium with PBS and observing EGFP fluorescence under the microscope, and by western blot analysis of whole cell lysates with a RGS-$His_6$-tag specific antibody.

## Luciferase Reporter Vectors

The HSV-TK promoter was cloned into a basic vector with no promoter (pGL4.10[*luc2*], Promega) to generate a HSV-TK-driven reporter construct that contained the GCCACC Kozak consensus sequence directly upstream of the luciferase start codon (pGL4.10[HSV-TK/GCCACC/*luc2*]) (Fig. S4B, top construct). Using the Geneart Site-Directed Mutagenesis System (Invitrogen) according to manufacturer's instructions, a second construct was created where the Kozak consensus sequence was changed to TTCCCC (pGL4.10[HSV-TK/TTCCCC/*luc2*]) (Fig. S4B, second from top), the same Kozak sequence as the one at the beginning of RGS-$His_6$-*Dnmt1s*. In addition, variants of the two constructs were created that were extended by the transgenic RGS-$His_6$ tag to encode N-terminally tagged luciferase, leading to pGL4.10[HSV-TK/GCCACC+tag/*luc2*] and pGL4.10[HSV-TK/TTCCCC+tag/*luc2*] (Fig. S4B, lower two constructs). The *Renilla* co-reporter plasmid pGL4.74[*hRluc*/TK] (Promega) was used for normalization.

## Luciferase Assays

Cells were seeded at $1x10^4$ cells per well in 96-well plate the day before transfection. Fifty nanograms of promoterless control plasmid (pGL4.10[*luc2*]) or one of the four pGL4.10[HSV-TK/*luc2*] target plasmids (GCCACC, TTCCCC, GCCACC+tag, TTCCCC+tag) were co-transfected with 2.5 ng of pGL4.74[*hRluc*/TK] co-reporter plasmid into NIH-3T3 cells using Lipofectamine 2000 (Invitrogen) for 24 h according to manufacturer's instructions. The luciferase assay was conducted using the Dual-Luciferase Reporter Assay System (Promega) and a Centro LB 960 luminometer (Berthold Technologies, Bad Wildbad, Germany) according to manufacturer's instructions. The firefly

luciferase readings were normalized to the *Renilla* luciferase readings. Results are representative of quintuplicate samples from three independent experiments and presented as mean ± standard deviation. Data groups were compared using the unpaired Student's *t*-test.

## Pronucleus injections

Pronucleus injections were performed by the Central Animal Facility of the University Hospital Essen. The pLCAG-*EGFP-loxP-Dnmt1s* plasmid was digested with NotI, and the transgene was separated from the vector backbone by agarose gel electrophoresis without ethidium bromide in gel or buffer. The fragment containing the *Dnmt1s* transgene construct was excised from the agarose gel and purified using the Qiaquick Gel Extraction Kit (Qiagen, Hilden, Germany), strictly following the manufacturer's protocol, except that the DNA was eluted in specialized pronucleus injection buffer provided by the Central Animal Facility of the University Hospital Essen. A total of 26 pronucleus injections in oocytes of C3H mice were performed, and the oocytes were transferred to C57BL/6J females. The pronucleus injections resulted in three female and three male transgenic mice. The six transgenic animals served as founder mice and were backcrossed into the C57BL/6J background. Mice used for analyses were in the F2 generation or later.

## Southern blot

4 µg of mouse tail DNA was digested with EcoRI, resolved and transferred to nylon membranes using standard methods. Membranes were incubated with a radioactively marked

probe that hybridized to a 1880 bp region spanning from bases 3 – 1882 of the *Dnmt1s* sequence.

Transgene sequencing

The whole transgene construct of an unrecombined mouse of line 4 as well as the CAG-*Dnmt1s* transition area of a recombined mouse of the same line were sequenced. Fragments were PCR-amplified, gel-purified (MinElute Gel Purification Kit, Qiagen) and Sanger sequenced.

PCR analysis

The PCRs for transgene insertion, recombination and *Cre* status, and the RT-PCR for verification of transgenic RNA expression were performed using the primers listed in Text S1.

Organ extraction from mice

Mice were euthanized using $CO_2$. Immediately after extraction, a small piece (50 – 100 mg) of each organ of was transferred to RNAlater stabilization reagent (Qiagen) for RNA preparations. The rest of the tissue was snap-frozen in liquid nitrogen for preparation of DNA or protein.

RNA extraction and cDNA preparation

RNA was prepared from RNAlater-stabilized (Qiagen) mouse tissue using the FastPrep FP120 Homogenizer (Thermo Fisher Scientific, Waltham, MA, USA) and Lysing Matrix D ceramic beads (MP Biomedicals, Eschwege, Germany). Depending on tissue type, either QIAzol lysis reagent followed by the RNA Cleanup protocol of the RNeasy Mini Kit was used for RNA isolation, or the protocol for purification of total RNA from animal tissues of the RNeasy Mini Kit including an additional homogenization step using QIAshredder columns (all Qiagen). An on-column DNase digestion was performed using the RNase-Free DNase Set (Qiagen) for all samples. RNA samples were treated with RQ1 RNase-Free DNase (Promega, Mannheim, Germany) and transcribed into cDNA using the ABI GeneAmp RNA PCR Kit with random hexamers (Life Technologies).

TaqMan quantitative real-time RT-PCR

TaqMan gene expression assays were performed to determine total *Dnmt1* transcript levels normalized to *Gapdh* expression (ABI TaqMan Gene Expression Assays: *Dnmt1* Mm01151063_m1, *Gapdh* Mm99999915_g1; Life Technologies). Depending on tissue type, 1.25 – 6.25 ng cDNA (converted from total RNA) were used in 10 µl reaction volumes (including ABI TaqMan Gene Expression Assay & Master Mix; Life Technologies). The 384-well plates were pipetted using an EPmotion pipetting robot (Qiagen) and run in an ABI 7900HT Fast Real-Time PCR System (Life Technologies) using following cycle program: 50 °C 2 min, 95 °C 10 min, 45 cycles (95 °C 15 s, 60 °C 1 min). The cycle threshold (Ct) values of each well were imported to Microsoft Excel for further analysis using the Delta-Delta Ct method.

240 Whole cell protein extracts from cultured cells and mouse organs

241

242 Pelleted cells were sonified at a concentration of 200 µg/ml in whole cell extract buffer (30
243 mM Tris-HCL, pH 8; 0.42 M NaCl; 0.5 mM EDTA; 20% Glycerol) containing Halt Protease
244 and Phosphatase Inhibitor (Thermo Fisher Scientific). Frozen (-80 °C) tissue was cut into
245 small pieces and sonified in whole cell extract buffer with inhibitors or RIPA buffer (50 mM
246 Hepes pH 7.4, 50 mM NaCl, 1% NP40, 0.5% Na-DOC, 0.1% SDS, 1% DTT, Halt Protease
247 and Phosphatase Inhibitor, Thermo Fisher Scientific). Samples were then centrifuged at
248 16,000 x g for 30 min at 4 °C. Protein quantification was performed using Coomassie Plus or
249 BCA Assay with iodoacetamide as compatibility reagent (both assays Thermo Fisher
250 Scientific).

251

252

253 Western blot

254

255 Western blotting was performed with crude protein extracts or with Ni-NTA-enriched protein
256 extracts (using Qiagen Ni-NTA Agarose). Gels were blotted onto nitrocellulose or PVDF
257 membranes with semi-dry or wet blot techniques. Antibodies are listed in Text S2.

## Results

### The conditional *Dnmt1s* transgene

We generated a conditional Cre-*loxP* transgene construct (pLCAG-*EGFP-loxP-Dnmt1s*) for the overexpression of *Dnmt1s* in mice. The transgene was constructed to initially express an EGFP marker protein ubiquitously driven by the CMV immediate early enhancer/chicken beta-actin promoter/rabbit beta-globin intron composite (CAG) promoter (Niwa et al. 1991). The *EGFP* sequence of the transgene is flanked by *loxP* sites (Fig. 1A). Upon recombination with Cre recombinase, the sequence between the *loxP* sites is looped out, removed and degraded by cellular mechanisms, preventing its re-insertion (Fig. 1B). Due to the excision of the *EGFP* and polyadenylation (pA) sequences, the *Dnmt1s* sequence moves up to the CAG promoter, resulting in the expression of *Dnmt1s* in the recombined version of the transgene (Fig. 1C).

### Cell culture testing of the pLCAG-*EGFP-loxP-Dnmt1s* transgene

The functionality of the transgene was tested in cell culture before pronucleus injections were carried out. Transient transfection of pLCAG-*EGFP-loxP-Dnmt1s* plasmid DNA into NIH-3T3 cells was performed to control for the expression of the EGFP marker protein from the unrecombined transgene by fluorescence microscopy. About 50% of cells were successfully transfected and expressed the EGFP marker protein (Fig. 2A), indicating functional promoter and *EGFP* sequences. No EGFP fluorescence could be observed in co-transfections with pCL-*Cre* plasmid DNA, indicating that the *EGFP* cassette was excised upon recombination (not shown).

To verify that RGS-$His_6$-tagged Dnmt1s was being expressed instead of the EGFP marker protein in the recombined cells, whole cell lysates were prepared from the *Dnmt1s/Cre* co-transfected culture dishes, separated by SDS-PAGE and probed with an RGS-$His_4$ antibody in a western blot. Equal loading was confirmed by a Ponceau-red stain of the blotted membrane prior to antibody incubation (not shown). A band of the expected size of Dnmt1s was present in *Dnmt1s/Cre* co-transfections (Fig. 2B, lane 1 and 2), showing that RGS-$His_6$-tagged Dnmt1s was expressed from the recombined transgene. No band was obtained for single transfections with pLCAG-*EGFP-loxP-Dnmt1s* plasmid DNA (Fig. 2B, lane 3), showing that the transgene does not express detectable amounts of Dnmt1s until recombination has taken place.

Founder mice express the EGFP marker protein

Pronucleus injections with the *Dnmt1s* transgene construct resulted in six transgenic founder lines with random integration of the unrecombined transgene, of which one line was discontinued due to breeding difficulties. Mice that carry the unrecombined version of the transgene should express EGFP protein ubiquitously, but not transgenic Dnmt1s. EGFP expression was controlled by screening the mouse tail snips gathered for genotyping with a fluorescence microscope before DNA isolation. Mouse tail snips from unrecombined transgenic mice showed strong EGFP expression (Fig. 3A, B). EGFP was also expressed in the paws (Fig. 3C, D), ears, nose, kidney, spleen and liver (data not shown) from mice carrying the unrecombined transgene, giving solid evidence that the transgene was successfully inserted and CAG-driven expression of the transgene was ubiquitous.

In addition to fluorescence microscopy, PCR and Southern blots were performed to verify transgene insertion. The PCR primers spanned a region from the *EGFP* sequence into

the *Dnmt1s* sequence, so that a 532 bp product indicated successful transgene insertion (Fig. 3E). The PCR reaction was multiplexed with primers spanning a region of the endogenous *Rag1* gene, so a 295 bp product indicated a successful PCR procedure. The Southern blot served as another control for transgene insertion. The transgenic status was indicated by one or more signals additional to those produced by the endogenous *Dnmt1* (Fig. 3F).

Ubiquitously recombined *Dnmt1s*-transgenic mice are viable

The five founder mice carrying the unrecombined version of the *Dnmt1s* transgene were crossed with a CMV-*Cre* deleter mouse strain, which ubiquitously expresses *Cre* under the control of the CMV promoter, to trigger transgene recombination in the zygote, leading to the ubiquitous overexpression of *Dnmt1s* in all tissues. To test for the expected recombination event, genomic DNA from tail snips of the offspring was analyzed by Southern blot and PCR using a forward primer located in the CAG promoter sequence and a reverse primer in the *Dnmt1s* sequence (Fig. 1A, C). This primer selection amplifies different product sizes from unrecombined and recombined versions of the transgene, being 1299 bp and 218 bp in size, respectively (Fig. 4A). Offspring from these cross-ins exhibited mendelian inheritance of both the *Dnmt1s* and the *Cre* transgenes. Unexpectedly, recombined offspring was viable and did not show any apparent signs of developmental defects or pathological phenotype. Dams showed no signs of carrying unborn dead embryos. Mice carrying recombined copies of the *Dnmt1s* transgene were then bred with C57BL/6J wild type mice, and only *Cre*-negative offspring with a recombined *Dnmt1s* transgene was chosen for further breeding to exclude influence of Cre recombinase on further studies.

336 Recombined mice overexpress *Dnmt1* mRNA in a tissue-dependent manner

337

338 Reverse transcriptase PCR (RT-PCR) analysis was performed to determine if the transgenic
339 mice express mRNA encoding RGS-$His_6$-tagged Dnmt1s. To specifically detect transgenic
340 RNA and avoid detection of endogenous *Dnmt1s* mRNA, the forward primer for the RT-PCR
341 was placed in the tag sequence, allowing only the specific amplification of a 209 bp
342 transgenic sequence. Because the *Dnmt1s* sequence in the transgene is intron-less, the cDNA
343 obtained from the RT reaction is identical to the genomic transgenic *Dnmt1s* sequence, so
344 both would yield a fragment in the RT-PCR. Therefore, RNA preparations were treated twice
345 with DNase (on column and in solution) to make sure that they were free of genomic DNA.
346 Transgenic *Dnmt1s* mRNA was expressed in recombined mice at different levels in a tissue-
347 dependent manner, with highest levels in heart and skeletal muscle (Fig. 4B). Wild type
348 siblings did not yield any RT-PCR products (Fig. S1).

349 To compare the expression levels of *Dnmt1* mRNA between transgenic mice and wild
350 type siblings, relative quantitative RT-PCR (qRT-PCR) was performed with ABI TaqMan
351 probes (Life Technologies, Darmstadt, Germany) that do not discriminate between transgenic
352 and endogenous *Dnmt1*. Transgenic line 4 showed the strongest overexpression of total
353 *Dnmt1* compared to sibling wild type controls, with tissue-dependent expression levels (Fig.
354 5). Overexpression in transgenic lines 2 and 3 was lower but followed a similar tissue-
355 dependent trend (see Fig. S2 for line 2 qRT-PCR). Transgenic mouse lines 2, 3, and 4 were
356 chosen for further breeding as they showed the highest *Dnmt1* expression levels, and lines 1
357 and 5 were discontinued. To check for point mutations and to verify correct recombination,
358 we sequenced the complete unrecombined transgene from genomic DNA of a line 4 mouse as
359 well as the CAG-*Dnmt1s* transition area after Cre-mediated recombination in a mouse of the
360 same line (data not shown). We found no point mutations, and after recombination the *EGFP*

sequence was successfully removed and the *Dnmt1s* sequence moved up to the CAG promoter.

Increased *Dnmt1* mRNA levels do not lead to Dnmt1 protein overexpression

We used western blotting techniques for the specific detection of tagged transgenic Dnmt1s protein or for the semi-quantitative detection of total Dnmt1 protein levels, including whole cell or Ni-NTA-enriched protein extracts, different blotting techniques (semi-dry blot, wet blot), different blotting membranes (nitrocellulose and PVDF), and different antibodies specific either for the His tag or for Dnmt1 (see Text S2 for a list of antibodies). Although transgenic mRNA was expressed in all tissues analyzed, we could not detect RGS-$His_6$-tagged transgenic Dnmt1s protein in any of the analyzed tissues from adult transgenic mouse lines 2, 3 and 4 or in transgenic mouse embryonic fibroblasts (MEFs) of 13.5 days *post coitum* embryos from line 4 (data not shown). Quantification of total Dnmt1 protein expression using the quantification software suites QuantityOne (Ver. 4.6.5, Bio-Rad, Munich, Germany) and MultiGauge (Ver. 3.0, Fujifilm, Düsseldorf, Germany) showed no significant changes between transgenic and wild type samples of lines 2 – 4 (for representative total Dnmt1 western blots from mouse tissue see Fig. 6 and from MEFs see Fig. S3). None of the western blots showed signs of degraded Dnmt1 protein.

The transgenic Kozak sequence and the RGS-$His_6$ tag lead to reduction of protein translation

We next investigated why our transgene expresses EGFP before recombination, but not Dnmt1s after recombination *in vivo*. Therefore, we performed luciferase assays to analyze more closely the effects of (i) the different Kozak sequences and (ii) the RGS-$His_6$ tag (Fig.

S4A) on translation. We cloned the different Kozak sequences used in our mouse model into a standard luciferase reporter plasmid and measured their influence on translation (Fig. S4B). The Kozak sequence used in our *Dnmt1s* transgene cassette (TTCCCC) reduced protein expression to about 70% compared to the Kozak sequence from the *EGFP* cassette, which had the optimal Kozak consensus (GCCACC) sequence (Fig. 7). To more closely resemble the situation in our mouse model, we tested the same two Kozak sequences using a luciferase reporter vector carrying an N-terminal RGS-$His_6$ tag identical to the one in our mouse model (Fig. S4B). The RGS-$His_6$ tag reduced the luciferase activity to about 30% in both cases, independent of the Kozak sequence (Fig. 7).

## Discussion

Global hypomethylation together with regional hypermethylation and overexpression of *Dnmt1* are characteristic for several types of cancer (Fernandez et al. 2012; Hervouet et al. 2010; Issa 2000). Despite the established role of Dnmt1 as maintenance methyltransferase, a number of publications on *Dnmt1*-overexpressing cells suggest that Dnmt1 might also have the ability to methylate certain target sites *de novo* (Feltus et al. 2003; Jair et al. 2006). We created a conditional mouse model to study the effects of Dnmt1s overexpression.

Studying the consequences of gene overexpression *in vivo* becomes problematic when the resulting changes are lethal at an early stage of development. This was the case in the *Dnmt1*-transgenic mouse model published by Biniszkiewicz *et al.* (Biniszkiewicz et al. 2002). They generated *Dnmt1*-overexpressing mice using a BAC clone that included all three isoforms of *Dnmt1*, leaving unclear which isoform(s) triggered the embryonic lethal phenotype. Therefore, we cloned a transgene using the murine cDNA of the somatic isoform, *Dnmt1s*, while leaving out the oocyte- and spermatocyte-specific isoforms, and generated a mouse model for its conditional overexpression to circumvent embryonic lethality and study long-term consequences.

We created a Cre-*loxP* transgene construct (Hoess & Abremski 1985) so that transgenic *Dnmt1s* expression can be activated in specific tissues or at specific time points in development, depending on the expression pattern of the promoters used to drive the transgene and *Cre* expression. The conditional transgene is not controlled by the endogenous *Dnmt1* promoter, but by the CAG promoter, which has been shown to be active ubiquitously from the preimplantation stage throughout development (Okabe et al. 1997; Sakai & Miyazaki 1997). Accordingly, EGFP was expressed ubiquitously from the unrecombined transgene in

our mouse model. This should provide the basis for activation of *Dnmt1s* overexpression in any tissue at any time point, as long as an appropriate *Cre* deleter strain is available.

We performed cross-ins with a CMV-*Cre* deleter strain that ubiquitously expresses Cre recombinase (Schwenk et al. 1995) to generate a transgenic mouse that resembles the mouse in the publication of Biniszkiewicz *et al.* (Biniszkiewicz et al. 2002) but is characterized by the ubiquitous overexpression of only the somatic isoform of *Dnmt1*. Surprisingly, these cross-ins yielded only viable offspring that showed no apparent signs of disease or pathological phenotype. Molecular biological analyses of recombined offspring showed that the *EGFP* sequence was excised and transgenic *Dnmt1s* mRNA was expressed in all tissues tested. The level of *Dnmt1* mRNA expression in recombined transgenic animals compared to wild type siblings was increased in a tissue-dependent manner (between 1.6- and 229-fold in line 4), with the highest expression levels in heart and skeletal muscle. The differential expression depending on tissue type most likely resulted from differences in the efficacy of the CAG promoter in distinct tissues, as previously shown by Ishikawa and colleagues, who also found the highest activity of their CAG-driven transgene in murine heart and muscle (Ishikawa & Herschman 2011).

Although transgenic mRNA was expressed in all tissues analyzed and qRT-PCR showed ubiquitous overexpression of *Dnmt1*, we could not detect tagged transgenic Dnmt1s protein or increased total Dnmt1 protein levels in various tissues from adult transgenic mice or in transgenic MEFs. Similar to our study, Choi *et al.* generated a human *DNMT1* transgene that was also N-terminally tagged (Choi et al. 2011). Stable integration into the genome of HEK293T cells led to a ~10-fold *DNMT1* mRNA overexpression, but only a 1.5-fold increase of total DNMT1 protein compared to control cells. *Dnmt1* mRNA expression in our mouse model was increased up to 229-fold depending on tissue type, but using semi-quantitative western blotting we could not detect any significant changes in total Dnmt1 protein levels. Both studies taken together suggest that high levels of *DNMT1/Dnmt1* mRNA overexpression

are not reflected on the protein level. In contrast to Choi and colleagues (Choi et al. 2011), we could not specifically detect tagged transgenic Dnmt1 protein. This might be caused by the second ATG translation initiation site close to the N-terminus (Fig. S4A, bottom construct) originating from the *Dnmt1s* cDNA sequence, which may result in an overexpressed protein lacking the N-terminal RGS-$His_6$ tag (Qiagen 2002). However, a significant overexpression of untagged transgenic protein appears rather unlikely as the second ATG site possesses a weaker Kozak consensus sequence, and we could not detect an increase in total Dnmt1 protein levels in transgenic mice or MEFs. Importantly, testing of our conditional transgene construct in murine NIH-3T3 fibroblasts prior to pronucleus injections demonstrated the full functionality of our transgene, and using western blots we could readily detect transgenic RGS-$His_6$-tagged Dnmt1s in extracts from *Dnmt1s/Cre* co-transfected cells. Our mouse model exhibits high levels of *Dnmt1* mRNA overexpression but no increased protein levels.

One possibility for the lack of Dnmt1s overexpression is that translation might have been impaired due to differences to the wild type *Dnmt1s* sequence. Our transgene lacks the complete endogenous 5'UTR and all introns, and the endogenous 3'UTR is present but shortened. The CAG promoter that drives our transgene carries its own 5'UTR and an intron from the rabbit beta-globin gene and has successfully been used for the expression of intron-less cDNA-based transgenes in mouse models (Ikawa et al. 1995; Kamei et al. 2006; Sakamoto et al. 2002). EGFP was expressed in our mouse model before recombination, but not Dnmt1s after recombination. As the two cassettes share the same promoter, but differ in their Kozak sequences, we compared the translation efficiency of those sequences using luciferase reporter assays. The translation with the *Dnmt1s* Kozak sequence had only about 70% efficiency compared to the *EGFP* Kozak sequence. The N-terminal RGS-$His_6$ tag caused a reduction of luciferase activity to about 30% regardless of Kozak sequence. It has been shown that an N-terminal $His_6$ tag does not significantly reduce the activity of luciferase (Min & Steghens 1999), so it is more likely that the tag impaired the translation of the protein.

Nevertheless, we would expect to detect RGS-$His_6$-tagged Dnmt1s and a rise in total Dnmt1 protein levels in our mouse model, at least in samples with the highest mRNA overexpression.

It has only recently been shown that the 3'UTR of *Dnmt1* mRNA contains a conserved block functional elements that are essential for efficient *Dnmt1* translation (Rutledge et al. 2014). The BAC transgenic mouse model by Biniszkiewicz *et al.* (Biniszkiewicz et al. 2002) contained the full-length UTRs and led to Dnmt1 overexpression. The transgenic cells by Choi *et al.* (Choi et al. 2011) showed only low levels of Dnmt1 overexpression, but the publication does not provide information on the UTRs used in the transgene construct. Our transgene contains an incomplete 3'UTR, which may have contributed to the reduced or abolished translation of *Dnmt1* observed in our model.

Another possibility for the lack of Dnmt1 protein overexpression could be that newly synthesized protein might have been degraded. Dnmt1 stability and activity are controlled by a complex interlinked network of numerous post-translational modifications (Qin et al. 2011), and Agoston and colleagues have shown that the DNMT1 protein contains an N-terminal destruction domain that controls its stability (Agoston et al. 2005). In addition to the reduced translation demonstrated in the luciferase assays, the RGS-$His_6$ tag in our transgene may have affected these and other pathways controlling Dnmt1s protein stability. However, we saw no indication of degraded Dnmt1 protein in our western blots. In the cell culture validation of our transgene construct, impaired translation or protein degradation may have been compensated because the transient transfection protocols we used resulted in high copy numbers per cell. In our transgenic mouse model, only one or a few copies remain in the genome after recombination (Lakso et al. 1992; Orban et al. 1992), so inefficient translation or protein degradation may have led to little or no transgenic protein.

In summary, we generated three independent *Dnmt1s*-transgenic mouse lines. Transgene testing in NIH-3T3 cells showed full functionality of the transgene construct in unrecombined and recombined states, and mice carrying the unrecombined transgene

498 ubiquitously expressed EGFP. However, we could not detect an increase of Dnmt1 protein in
499 recombined mice despite ubiquitous tissue-dependent *Dnmt1* mRNA overexpression. It
500 remains unclear if this is due to intrinsic features of the transgene or a stringent control of
501 Dnmt1 levels by the cell.

## Acknowledgements

We would like to thank S. Pradhan from New England Biolabs for providing the *Dnmt1s* cDNA and for helpful discussion. We would also like to thank Ralph Waldschütz and Wojciech Wegrzyn for pronucleus injections and *Cre* genotyping.

## Funding

This project is supported by the Deutsche Forschungsgemeinschaft (DFG) graduate research training program no. 1431: "Transcription, Chromatin Structure and DNA Repair in Development and Differentiation".

## References


Agoston AT, Argani P, Yegnasubramanian S, De Marzo AM, Ansari-Lari MA, Hicks JL, Davidson NE, and Nelson WG. 2005. Increased protein stability causes DNA methyltransferase 1 dysregulation in breast cancer. *The Journal of biological chemistry* 280:18302-18310.

Baylin SB, Hoppener JW, de Bustros A, Steenbergh PH, Lips CJ, and Nelkin BD. 1986. DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. *Cancer research* 46:2917-2922.

Bestor TH. 2003. Unanswered questions about the role of promoter methylation in carcinogenesis. *Annals of the New York Academy of Sciences* 983:22-27.

Bestor TH, and Ingram VM. 1983. Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proceedings of the National Academy of Sciences of the United States of America* 80:5559-5563.

Biniszkiewicz D, Gribnau J, Ramsahoye B, Gaudet F, Eggan K, Humpherys D, Mastrangelo MA, Jun Z, Walter J, and Jaenisch R. 2002. Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Molecular and cellular biology* 22:2124-2135.

Bird A. 2002. DNA methylation patterns and epigenetic memory. *Genes & development* 16:6-21.

Chan TL, Yuen ST, Kong CK, Chan YW, Chan AS, Ng WF, Tsui WY, Lo MW, Tam WY, Li VS, and Leung SY. 2006. Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nature genetics* 38:1178-1183.

Choi SH, Heo K, Byun HM, An W, Lu W, and Yang AS. 2011. Identification of preferential target sites for human DNA methyltransferases. *Nucleic acids research* 39:104-118.

Christman JK, Sheikhnejad G, Marasco CJ, and Sufrin JR. 1995. 5-Methyl-2'-deoxycytidine in single-stranded DNA can act in cis to signal de novo DNA methylation. *Proceedings of the National Academy of Sciences of the United States of America* 92:7347-7351.

Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, and Li BF. 1997. Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* 277:1996-2000.

Cirio MC, Ratnam S, Ding F, Reinhart B, Navara C, and Chaillet JR. 2008. Preimplantation expression of the somatic form of Dnmt1 suggests a role in the inheritance of genomic imprints. *BMC developmental biology* 8:9.

Ehrlich M. 2002. DNA methylation in cancer: too much, but also too little. *Oncogene* 21:5400-5413.

Feinberg AP, and Vogelstein B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 301:89-92.

Feltus FA, Lee EK, Costello JF, Plass C, and Vertino PM. 2003. Predicting aberrant CpG island methylation. *Proceedings of the National Academy of Sciences of the United States of America* 100:12253-12258.

Fernandez AF, Assenov Y, Martin-Subero JI, Balint B, Siebert R, Taniguchi H, Yamamoto H, Hidalgo M, Tan AC, Galm O, Ferrer I, Sanchez-Cespedes M, Villanueva A, Carmona J, Sanchez-Mut JV, Berdasco M, Moreno V, Capella G, Monk D, Ballestar E, Ropero S, Martinez R, Sanchez-Carbayo M, Prosper F, Agirre X, Fraga MF, Grana O, Perez-Jurado L, Mora J, Puig S, Prat J, Badimon L, Puca AA, Meltzer SJ, Lengauer T, Bridgewater J, Bock C, and Esteller M. 2012. A DNA methylation fingerprint of 1628 human samples. *Genome research* 22:407-419.

Greger V, Debus N, Lohmann D, Hopping W, Passarge E, and Horsthemke B. 1994. Frequency and parental origin of hypermethylated RB1 alleles in retinoblastoma. *Human genetics* 94:491-496.

Gruenbaum Y, Cedar H, and Razin A. 1982. Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* 295:620-622.

Gruenbaum Y, Szyf M, Cedar H, and Razin A. 1983. Methylation of replicating and post-replicated mouse L-cell DNA. *Proceedings of the National Academy of Sciences of the United States of America* 80:4919-4921.

Hervouet E, Vallette FM, and Cartron PF. 2010. Impact of the DNA methyltransferases expression on the methylation status of apoptosis-associated genes in glioblastoma multiforme. *Cell death & disease* 1:e8.

Hoess RH, and Abremski K. 1985. Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. *Journal of molecular biology* 181:351-362.
Ikawa M, Kominami K, Yoshimura Y, Tanaka K, Nishimune Y, and Okabe M. 1995. A rapid and non-invasive selection of transgenic embryos before implantation using green fluorescent protein (GFP). *FEBS letters* 375:125-128.
Ishikawa TO, and Herschman HR. 2011. Conditional bicistronic Cre reporter line expressing both firefly luciferase and beta-galactosidase. *Molecular imaging and biology : MIB : the official publication of the Academy of Molecular Imaging* 13:284-292.
Issa JP. 2000. The epigenetics of colorectal cancer. *Annals of the New York Academy of Sciences* 910:140-153; discussion 153-145.
Jair KW, Bachman KE, Suzuki H, Ting AH, Rhee I, Yen RW, Baylin SB, and Schuebel KE. 2006. De novo CpG island methylation in human cancer cells. *Cancer research* 66:682-692.
Jones PA, and Baylin SB. 2002. The fundamental role of epigenetic events in cancer. *Nature reviews Genetics* 3:415-428.
Jones PA, and Baylin SB. 2007. The epigenomics of cancer. *Cell* 128:683-692.
Kamei K, Ishikawa TO, and Herschman HR. 2006. Transgenic mouse for conditional, tissue-specific Cox-2 overexpression. *Genesis* 44:177-182.
Kurihara Y, Kawamura Y, Uchijima Y, Amamo T, Kobayashi H, Asano T, and Kurihara H. 2008. Maintenance of genomic methylation patterns during preimplantation development requires the somatic form of DNA methyltransferase 1. *Developmental biology* 313:335-346.
Lakso M, Sauer B, Mosinger B, Jr., Lee EJ, Manning RW, Yu SH, Mulder KL, and Westphal H. 1992. Targeted oncogene activation by site-specific recombination in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* 89:6232-6236.
Leonhardt H, Page AW, Weier HU, and Bestor TH. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71:865-873.
Lyko F. 2005. Novel methods for analysis of genomic DNA methylation. *Analytical and bioanalytical chemistry* 381:67-68.
Mertineit C, Yoder JA, Taketo T, Laird DW, Trasler JM, and Bestor TH. 1998. Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development* 125:889-897.
Min KL, and Steghens JP. 1999. The emitting species dissociated from the enzyme can emit the light in Photinus pyralis luciferase system. *Biochemical and biophysical research communications* 265:273-278.
Myohanen SK, Baylin SB, and Herman JG. 1998. Hypermethylation can selectively silence individual p16ink4A alleles in neoplasia. *Cancer research* 58:591-593.
Niwa H, Yamamura K, and Miyazaki J. 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* 108:193-199.
Okabe M, Ikawa M, Kominami K, Nakanishi T, and Nishimune Y. 1997. 'Green mice' as a source of ubiquitous green cells. *FEBS letters* 407:313-319.
Okano M, Xie S, and Li E. 1998. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature genetics* 19:219-220.
Orban PC, Chui D, and Marth JD. 1992. Tissue- and site-specific DNA recombination in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* 89:6861-6865.
Qiagen. 2002. *QIAexpress Detection and Assay Handbook*. Valencia, CA.
Qin W, Leonhardt H, and Pichler G. 2011. Regulation of DNA methyltransferase 1 by interactions and modifications. *Nucleus* 2:392-402.
Robertson KD, and Jones PA. 2000. DNA methylation: past, present and future directions. *Carcinogenesis* 21:461-467.
Rutledge CE, Lau HT, Mangan H, Hardy LL, Sunnotel O, Guo F, MacNicol AM, Walsh CP, and Lees-Murdock DJ. 2014. Efficient translation of Dnmt1 requires cytoplasmic polyadenylation and Musashi binding elements. *PloS one* 9:e88385.
Sakai K, and Miyazaki J. 1997. A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. *Biochemical and biophysical research communications* 237:318-324.
Sakamoto M, Yuasa K, Yoshimura M, Yokota T, Ikemoto T, Suzuki M, Dickson G, Miyagoe-Suzuki Y, and Takeda S. 2002. Micro-dystrophin cDNA ameliorates dystrophic phenotypes when

introduced into mdx mice as a transgene. *Biochemical and biophysical research communications* 293:1265-1272.

Schwenk F, Baron U, and Rajewsky K. 1995. A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. *Nucleic acids research* 23:5080-5081.

Wu J, Issa JP, Herman J, Bassett DE, Jr., Nelkin BD, and Baylin SB. 1993. Expression of an exogenous eukaryotic DNA methyltransferase gene induces transformation of NIH 3T3 cells. *Proceedings of the National Academy of Sciences of the United States of America* 90:8891-8895.

Yoder JA, Soman NS, Verdine GL, and Bestor TH. 1997. DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *Journal of molecular biology* 270:385-395.

Yoder JA, Yen RW, Vertino PM, Bestor TH, and Baylin SB. 1996. New 5' regions of the murine and human genes for DNA (cytosine-5)-methyltransferase. *The Journal of biological chemistry* 271:31092-31097.

## Figures

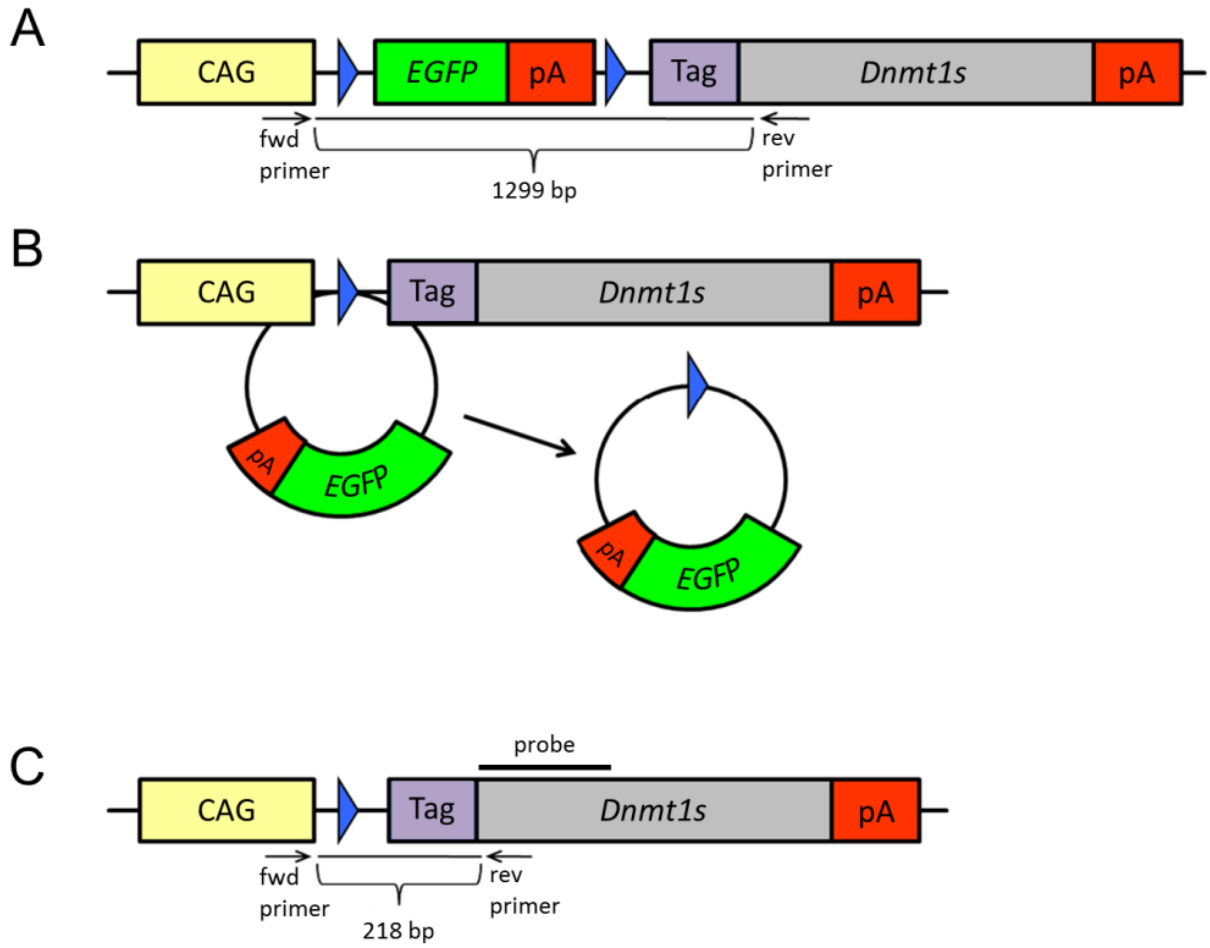



**Figure 1. The mechanism of the transgene.** (**A**) The CAG promoter initially drives *EGFP* expression. (**B**) Upon recombination with Cre recombinase, the *EGFP* sequence is looped out, removed and degraded. (**C**) After recombination, the CAG promoter drives the expression of RGS-$His_6$-tagged *Dnmt1s*. The black bar indicates the binding site of the Southern blot probe. Yellow box - CAG promoter; blue arrows - *loxP* sites; green box - *EGFP*; red boxes - poly A signal; purple box - RGS-$His_6$ tag; grey box - *Dnmt1s* (Figure not drawn to scale)

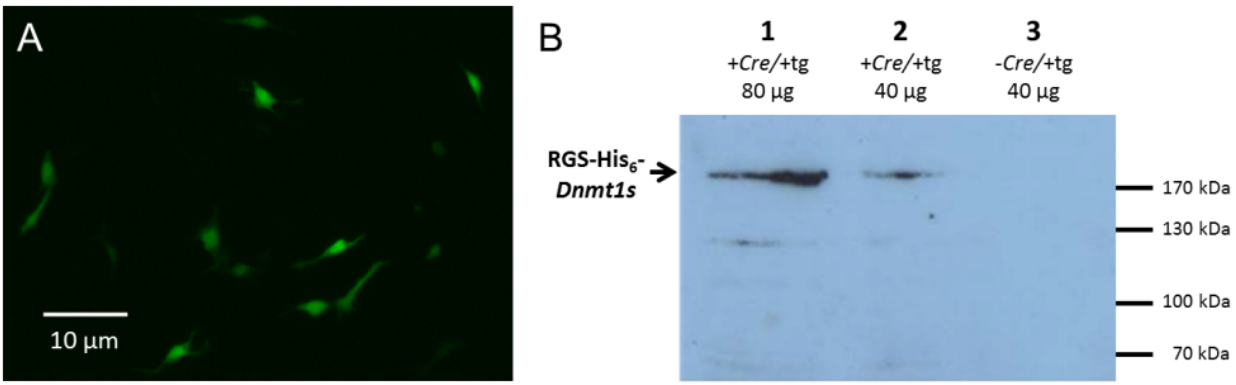



**Figure 2. Cell culture testing of the pLCAG-*EGFP-loxP-Dnmt1s* transgene.** (**A**) Transient transfection of NIH-3T3 cells with pLCAG-*EGFP-loxP-Dnmt1s* plasmid DNA showed EGFP fluorescence. (**B**) Western blot analysis using an RGS-$His_4$-tag antibody on whole cell lysates of NIH-3T3 cells (lane 1: 80 µg, lanes 2 and 3: 40 µg) Lane 1 and 2: Co-transfections with pLCAG-*EGFP-loxP-Dnmt1s* and pCL-*Cre* plasmid DNA. Bands of about 172 kDa indicate the expression of RGS-$His_6$-Dnmt1s protein. Lane 3: Transfection with pLCAG-*EGFP-loxP-Dnmt1s* plasmid DNA alone did not result in expression of RGS-$His_6$-tagged Dnmt1s protein, showing the requirement of recombination. tg: pLCAG-*EGFP-loxP-Dnmt1s* status; *Cre*: pCL-*Cre* status

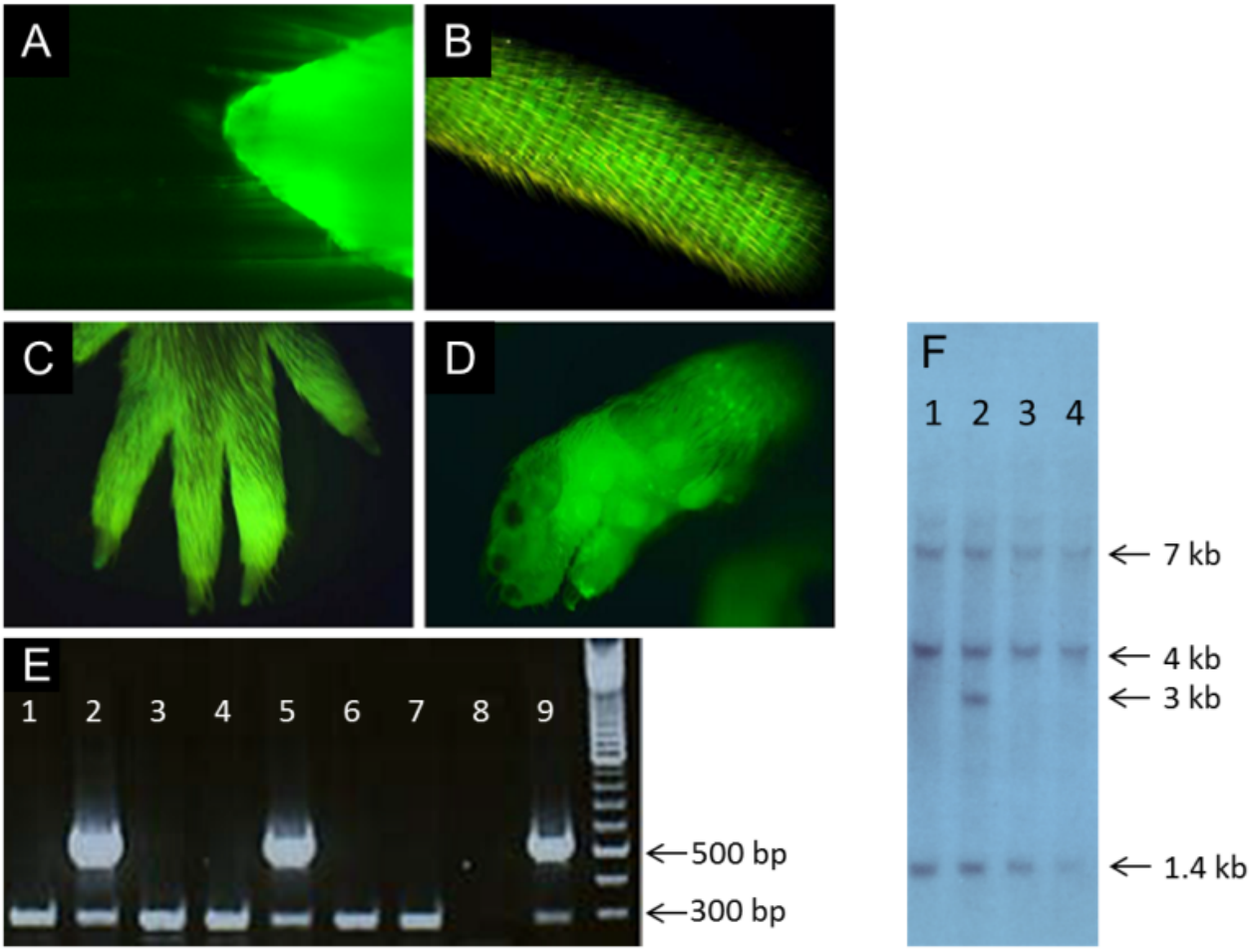
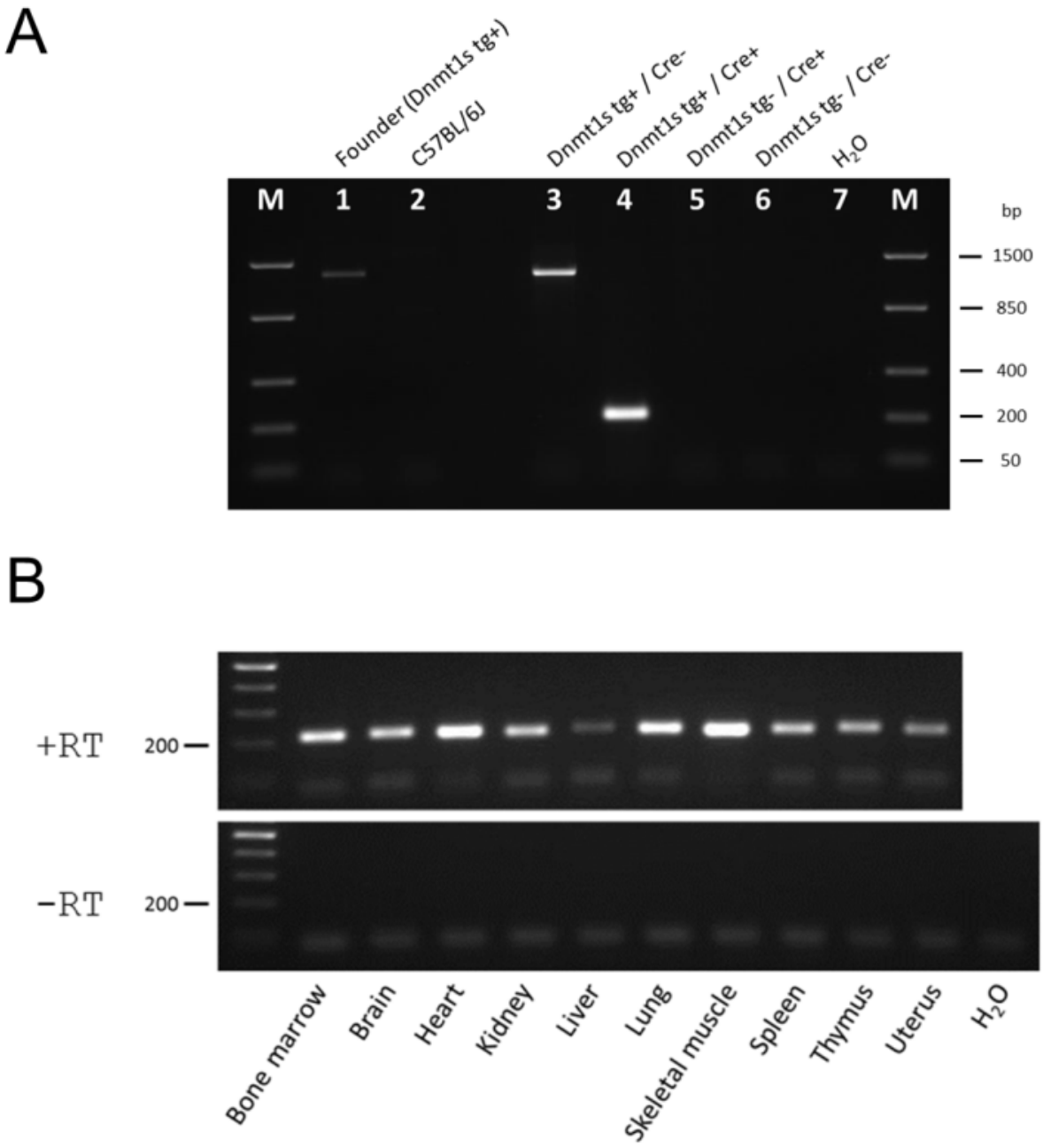



**Figure 3. EGFP is expressed *in vivo* from the unrecombined transgene.** (**A – D**) Fluorescence microscopy of mouse tail snips (**A** and **B**) and paws (**C** and **D**) of transgenic mice carrying the unrecombined transgene showed EGFP marker protein expression. (**E**) The multiplex PCR resulted in a transgene-specific 532 bp fragment for unrecombined transgenic mice (lanes 2, 5, and 9) and a 295 bp control fragment from the endogenous *Rag1* gene for all samples. PCR reactions without a 532 bp fragment (lanes 1, 3, 4, 6, 7) indicated wild type mice. PCR reactions lacking the control fragment (lane 8) indicated a failed PCR and were repeated. Far right lane: 100 bp DNA ladder. (**F**) Southern blots of wild type animals showed three bands produced by endogenous *Dnmt1* (1.4 kb, 4 kb, 7 kb). Transgenic offspring from pronucleus injections showed at least one additional signal. For founder line 4, this was one 3 kb band (lane 2).




670

671 **Figure 4. The *Dnmt1s* transgene recombines upon CMV-*Cre* cross-in and expresses**
672 **transgenic *Dnmt1s* mRNA ubiquitously.** (**A**) DNA from an unrecombined transgenic
673 founder mouse produced a 1299 bp PCR product (lane 1), whereas DNA from a wild type
674 mouse did not yield any product (lane 2). DNA from offspring of CMV-*Cre* cross-ins
675 produced a 1299 bp PCR product when Cre was absent (lane 3) or a 218 bp product in the
676 presence of Cre recombinase (lane 4), indicating successful transgene recombination. Mice
677 carrying no *Dnmt1s* transgene construct did not yield any product, regardless of presence

(lane 5) or absence (lane 6) of Cre. Lane 7 - $H_2O$ control; M - FastRuler Low Range DNA Ladder (Thermo Fisher Scientific) (**B**) mRNA encoding RGS-$His_6$-tagged *Dnmt1s* was specifically expressed in various tissues of recombined transgenic mice, indicated by a 209 bp band. As input amounts were equal for all samples, different band intensities indicate tissue-specific expression levels. +RT: complete RT reaction; -RT: control reaction without reverse transcriptase; Marker: FastRuler Low Range Ladder (Thermo Fisher Scientific)

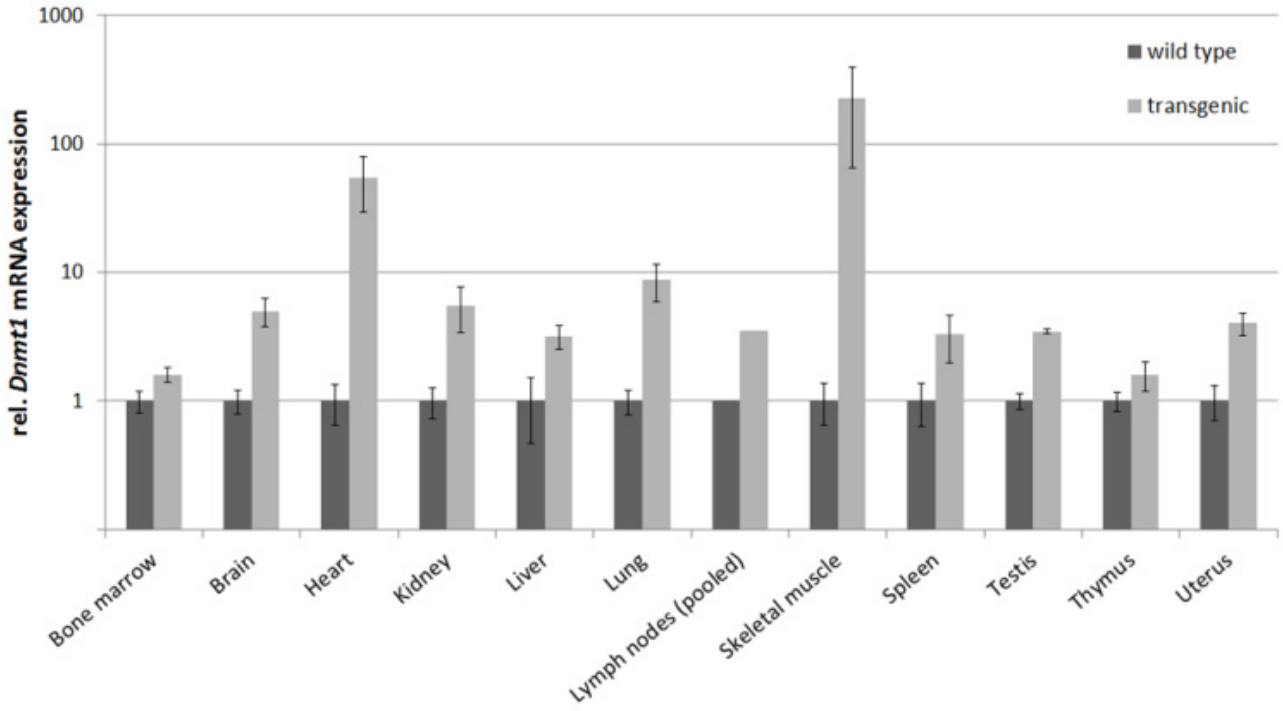



**Figure 5. Recombined mice of line 4 overexpress *Dnmt1* mRNA.** The relative expression of *Dnmt1* normalized to *Gapdh* in tissues from transgenic mice was significantly higher than in wild type siblings, and overexpression was tissue-dependent [Mean ± SD; n ≥ 5 (except for lymph nodes: 5 pooled samples); $P < 0.05$, *t*-test]. *Dnmt1* expression of wild type animals was set to 1 for each tissue. The highest *Dnmt1* mRNA expression was ~ 230-fold in muscle (228.5 ± 163.5) and ~ 55-fold heart (54.7 ± 25.6). *Dnmt1* overexpression was between 3- and 10-fold in brain (5.0 ± 1.3), kidney (5.5 ± 2.1), liver (3.2 ± 0.7), lung (8.7 ± 2.7), pooled lymph nodes (3.5), spleen (3.3 ± 1.3), testis (3.5 ± 0.2) and uterus (4.0 ± 0.8). The lowest overexpression was in bone marrow (1.6 ± 0.2) and thymus (1.6 ± 0.4)

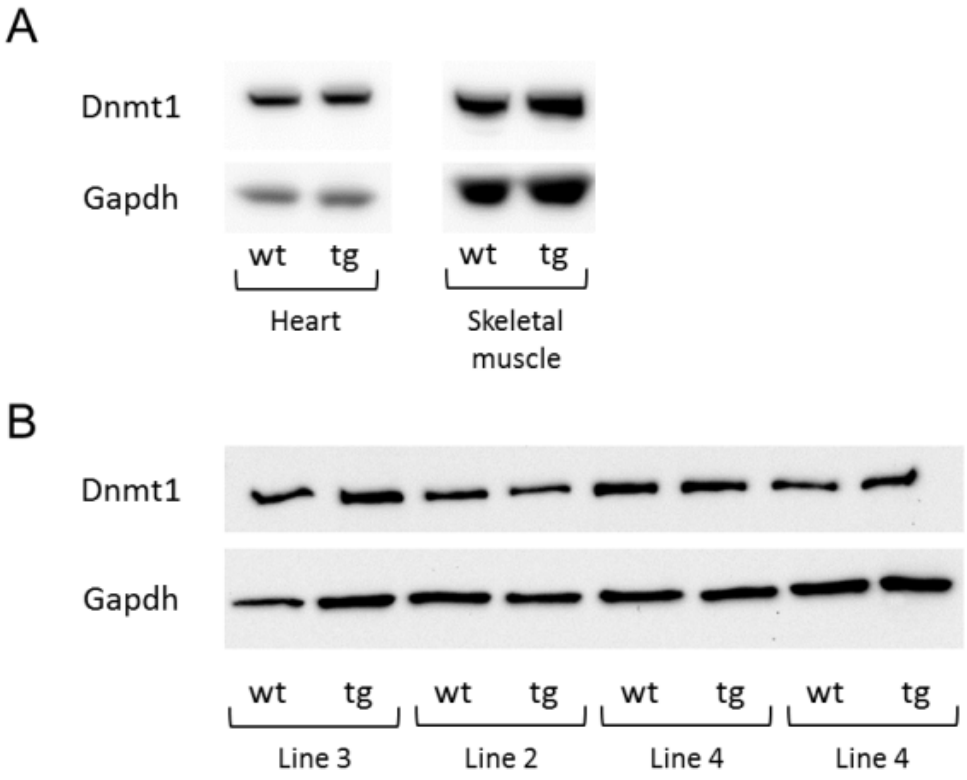



**Figure 6. Increased *Dnmt1* mRNA levels do not lead to Dnmt1 protein overexpression.** Detection of total (both endogenous and transgenic) Dnmt1 with a Dnmt1-specific antibody. Tissues with the strongest mRNA overexpression (heart and muscle) of line 4 (**A**) and brain tissue lysates from lines 2, 3 and 4 (**B**) showed similar Dnmt1 protein levels. wt - wild type; tg - transgenic

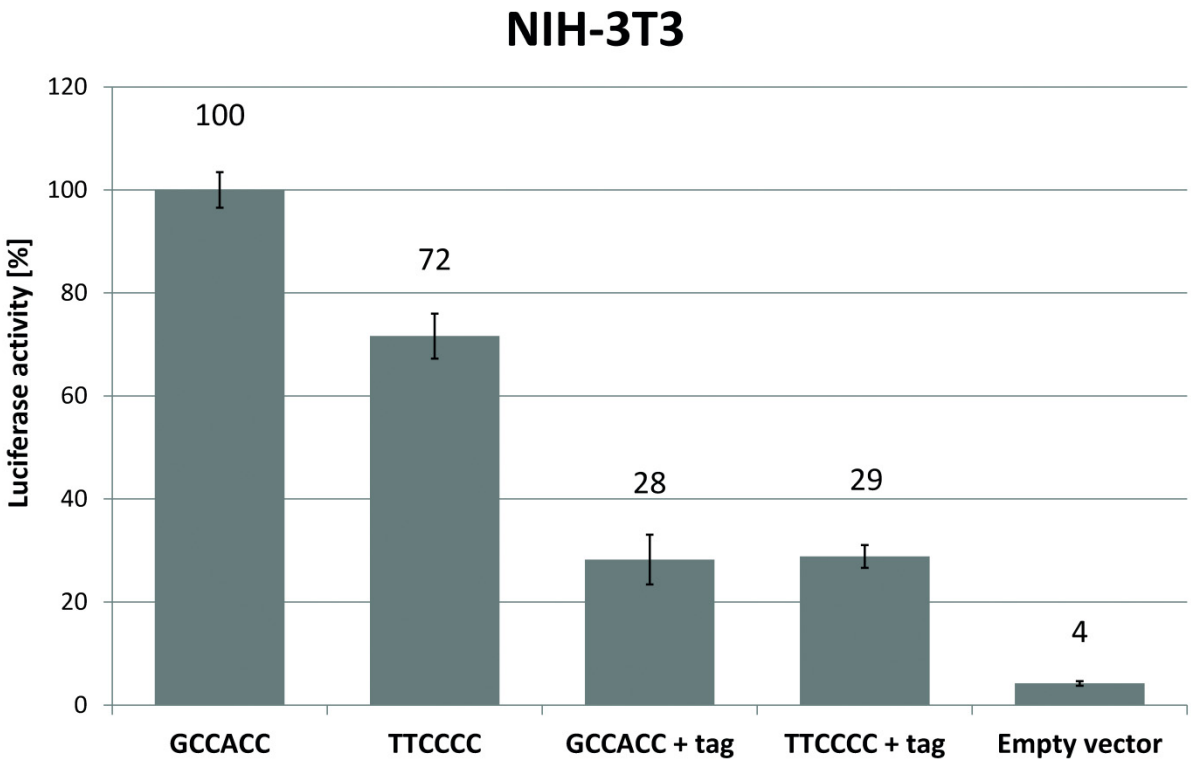



**Figure 7. The transgenic *Dnmt1s* Kozak and RGS-His$_6$ tag sequences do not prevent translation.** The *Dnmt1s* Kozak sequence (TTCCCC) from the recombined transgene reduces luciferase activity to 72% compared to the Kozak consensus sequence (GCCACC) of the *EGFP* cassette in the unrecombined transgene. Insertion of the RGS-His$_6$ tag in front of the luciferase start codon of both constructs leads to an even stronger reduction of luciferase activity to about 30% regardless of Kozak sequence. ($P < 0.01$ by unpaired Student's *t*-test for all samples compared to GCCACC, which was set to 100%; Empty vector: promoterless control plasmid pGL4.10[*luc2*]).

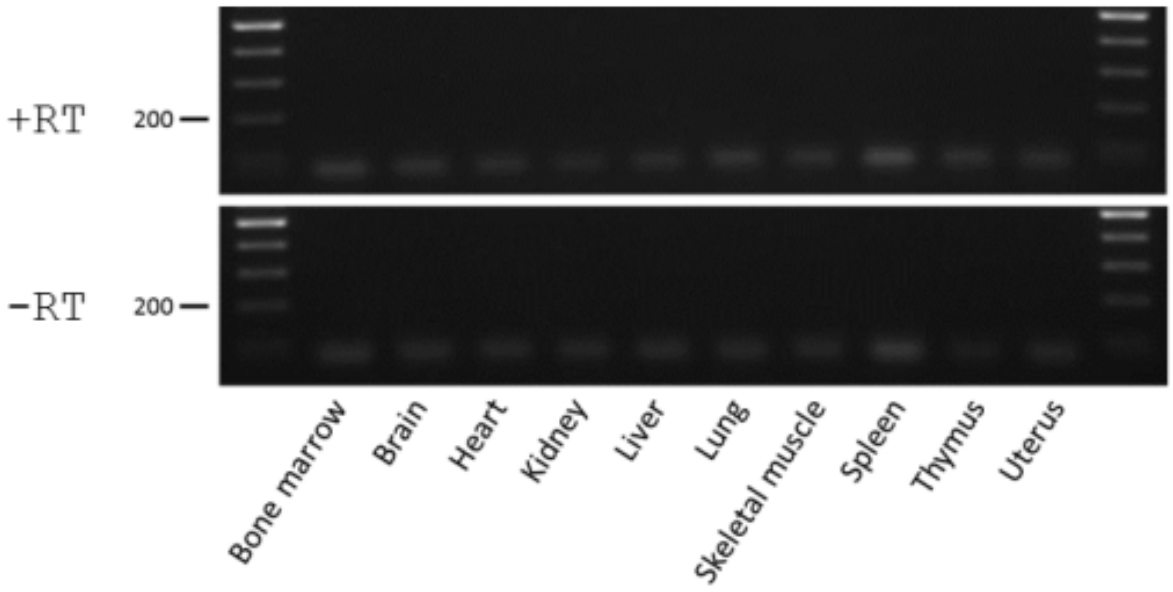



**Figure S1. cDNA from a wild type mouse does not yield a product in transgene-specific RT-PCR.** RT-PCR on cDNA from a wild type mouse related to Fig. 4B. Reactions were prepared using the same master mixes; for controls see Fig. 4B. +RT: complete RT reaction; -RT: control reaction without reverse transcriptase; Marker: FastRuler Low Range Ladder (Thermo Fisher Scientific)

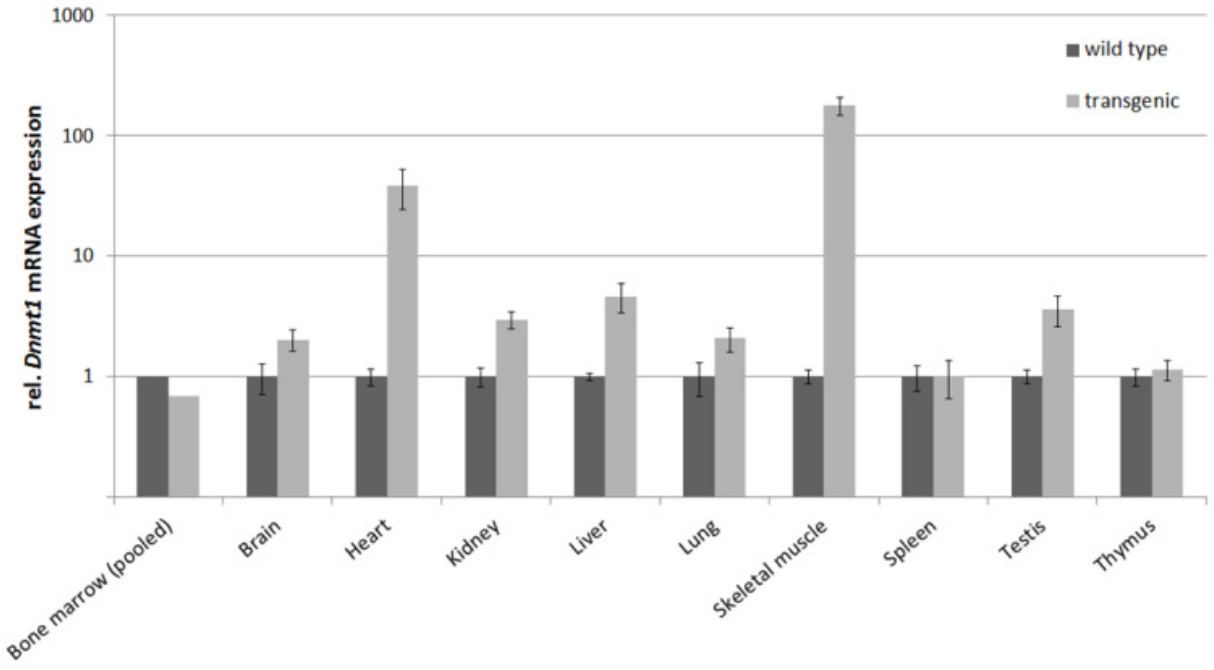



**Figure S2. Recombined mice of line 2 overexpress *Dnmt1* mRNA with a tissue-dependent trend similar to line 4.** The relative expression of *Dnmt1* normalized to *Gapdh* in tissues from transgenic line 2 was significantly higher ($P < 0.01$, *t*-test) than in wild type siblings in all tissues except spleen and thymus (Mean ± SD; $n \geq 3$, except for bone marrow: 5 pooled samples). The highest *Dnmt1* mRNA expression was ~180-fold in muscle (179.2 ± 30.7) and ~40-fold heart (38.6 ± 14.0). *Dnmt1* overexpression was between 2- and 10-fold in brain (2.0 ± 0.4), kidney (2.9 ± 0.5), liver (4.6 ± 1.3), lung (2.1 ± 0.5), testis (3.6 ± 1.0). No significant overexpression was found in thymus (1.1 ± 0.2), spleen (1.01 ± 0.4) and pooled bone marrow (0.7)

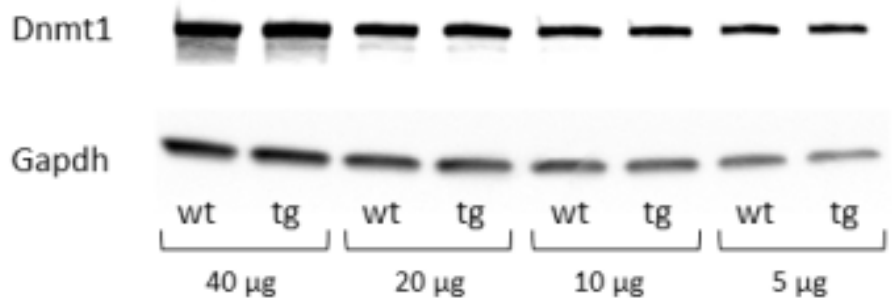



**Figure S3. Increased *Dnmt1* mRNA levels do not lead to Dnmt1 protein overexpression in MEFs.** We performed western blots with a Dnmt1-specific antibody on protein lysates of MEFs from day 13.5 recombined transgenic embryos which showed a 6.4-fold *Dnmt1* mRNA overexpression (data not shown). We used four different protein concentrations to minimize the chances that slight differences in total Dnmt1 expression are covered by too high overall amounts of Dnmt1 protein. However, even at 5 µg of protein per lane there was no measurable significant increase of total Dnmt1 levels in MEFs from recombined transgenic embryos. wt - wild type; tg - transgenic


**Figure S4. Construct design for functional testing of the Kozak sequences.** (**A**) Overview of the Kozak sequences of *EGFP* in the unrecombined version (top) and RGS-His$_6$-*Dnmt1s* in the recombined version (bottom) of the transgene. (**B**) Schematic of the luciferase reporter constructs to test the Kozak sequences of the *EGFP* cassette (top) and the RGS-His$_6$-*Dnmt1s* cassette (second from top) alone, and in combination with an RGS-His$_6$ tag (bottom two constructs). (CDS: coding sequence; Kozak_cons: Kozak consensus sequence; Kozak_tg: Kozak sequence of transgenic RGS-His$_6$-*Dnmt1s*)

**Text S1. Primers.**

All primers were created with the Primer3 software (Koressaar & Remm 2007; Untergasser et al. 2012) available at: http://primer3.wi.mit.edu

All sequences are given in 5' to 3' direction.

Transgenic status (Fig. 3e), 532 bp

| | |
|---|---|
| GFP_US1 | CCAACGAGAAGCGCGATCACATG |
| Dnmt1_rec_LS1 | CCGGGAGCGAGCCTGCCGGGGAG |

Rag1 control fragment (Fig. 3e), 295 bp

| | |
|---|---|
| mRag1.1 | GCTGATGGGAAGTCAAGCGAC |
| mRag1.3 | GGGAACTGCTGAACTTTCTGTG |

Southern blot probe (Fig. 3f), 1880 bp

| | |
|---|---|
| Dnmt1_g_US3 | GCCAGCGCGAACAGCTCCAGCCC |
| Dnmt1_g_LS3 | GATAGACCAGCTTGGTGGTGGTGGC |

Transgene recombination status (Fig. 4a), 1299 bp or 218 bp

| | |
|---|---|
| CAG_US1 | CTGGTTATTGTGCTGTCTCATC |
| Dnmt1_rec_LS1 | CCGGGAGCGAGCCTGCCGGGGAG |

Transgene-specific RT-PCR (Fig. 4b, Fig. S1), 209 bp

| | |
|---|---|
| HIS_RT_US2 | TCATCACCATCACCATATGCC |
| PRIMER3_RT_LS | CCAAGTCACACAACTGGCTTT |

**References for Text S1**

Koressaar T, and Remm M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289-1291.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, and Rozen SG. 2012. Primer3—new capabilities and interfaces. Nucleic acids research 40:e115.

781 **Text S2. Antibodies.**

782 Anti-Dnmt1 Antibody (#ab13537, Abcam, Cambridge, UK)

783 Anti-RGS-His Antibody, BSA-free (#34650, Qiagen, Hilden, Germany)

784 Anti-Penta-His Antibody, BSA-free (#34660, Qiagen)

785 Anti-Tetra-His Antibody, BSA-free (#34670, Qiagen)

786 Anti-His Antibody (#27-4710-01, GE Healthcare, Munich, Germany)

787 Anti-$(HIS)_6$-Tag Epitope Antibody Clone 13/45/31-2 (#DIA-900, Dianova, Hamburg,

788 Germany)

789 Anti-GAPDH Antibody (#2118, Cell Signaling Technology, Danvers, MA, USA)

790 Stabilized Goat Anti-Mouse HRP Conjugated (#32430, Thermo Fisher Scientific, Waltham,

791 MA, USA)

792 Stabilized Goat Anti-Rabbit HRP Conjugated (#32460, Thermo Fisher Scientific)