The Role of GADD45 Proteins and NEIL DNA Glycosylases in Mouse Embryonic Stem Cells

Dissertation

Zur Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie

Der Johannes Gutenberg-Universität Mainz

Katrin Mercedes Schüle

geb. am 10.08.1990 in La Jolla

Mainz, 2019

Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung:

Table of Contents

1 Summary

The establishment of DNA methylation patterns is key to ensure the epigenetic and transcriptional changes required for accurate progression during early mouse development. These DNA methylation patterns are, in contrast to longstanding presumption, dynamically regulated by a passive and an active process of DNA demethylation. The identification of TET proteins and TDG as mediators of the active DNA demethylation pathway has raised great interest in understanding their involvement in early mouse development. Our laboratory and others discovered players in active DNA demethylation. First, the GADD45 (Growth arrest and DNA damage-inducible) protein family, fulfilling a dual function by enhancing TET and TDG activity. Second, the nei endonuclease VIII-like family of DNA glycosylases (NEIL1 and NEIL2) capable to promote the substrate turnover of TDG. The precise role of GADD45 and NEIL proteins in promoting DNA demethylation during development is not well understood, and addressed in this study in two separate chapters using mouse embryonic stem cells as versatile *in vitro* system resembling early mouse development.

Mouse embryonic stem cells (mESC) deficient for all three *Gadd45* genes showed deregulation and hypermethylation of a specific subset of genes attributed to the totipotent 2-cell (2C) stage. 2C-reporter analysis corroborated GADD45 proteins as novel regulators of the 2C-like state, a mESC state with expanded fate potential comparable to the 2-cell embryo. *Gadd45a/b/g* deficient mESCs revealed reduced cycling into the 2C-like state resulting in reduced fate potential to transdifferentiate into trophoblast stem cells. Gene expression analysis of *Gadd45a/b* deficient 2-cell mouse embryos showed transcriptional changes for zygotic genome activation (ZGA) specific genes supporting the hypothesis that GADD45 proteins are involved in zygotic genome activation *in vivo*.

Neil1 or *Neil2* deficient mESCs failed to differentiate into the neuronal and neural crest lineage. Notably, this defect was not associated with their implication in active DNA demethylation, but rather with their function in base excision repair in mitochondria. Elevated mitochondrial DNA damage in *Neil1* and *Neil2* deficient mESCs activated an intrinsic mitochondrial p53 response impairing neuronal specification *in vitro*.

In summary, these results highlight the importance of the proteins GADD45α,β,γ and NEIL1,2 at distinct stages of early mouse development, and attribute their causal involvement in DNA demethylation and DNA repair, respectively.

2 Zusammenfassung

Die Etablierung des DNA-Methylierungsmusters ist entscheidend für epigenetische und transkriptionelle Veränderungen während der frühen Mausentwicklung. Entgegen der langläufigen Annahme handelt es sich bei DNA Demethylierung um einen dynamisch regulierten Prozess. Die Entdeckung der aktiven DNA-Demethylierung durch TET Proteine und TDG weckte großes Interesse an der Analyse der Funktion dieser Proteine während der frühen Mausentwicklung. Dieses und andere Forschungslabore konnten zwei Akteure der aktiven DNA Demethylierung identifizieren. Erstens, die GADD45 (Growth arrest and DNA damage-inducible) Proteinfamilie, welche eine Doppelfunktion der TET- und TDG-Aktivitätenerhöhung erfüllt. Zweitens, die nei endonuclease VIII-ähnliche Familie der DNA Glykosylasen (NEIL1 und NEIL2), welche den Substratumsatz von TDG fördern. Die genaue Rolle dieser Proteine in DNA-Demethylierung während der Embryonalentwicklung ist nicht eindeutig geklärt und wurde in der vorliegenden Studie in zwei separaten Kapiteln mittels muriner embryonaler Stammzellen (mESCs) als vielseitiges *In-vitro*-Sytem der frühen Mausentwicklung analysiert.

Embryonale Mausstammzellen (mESC), die für alle drei *Gadd45* Gene defizient sind, zeigten eine Deregulierung und Hypermethylierung einer bestimmten Teilmenge von Genen des totipotenten 2-Zellen-(2C)-Stadium. 2C-Reporteranalysen bestätigten GADD45-Proteine als neuartige Regulatoren des 2C-ähnlichen Zustands, eines mESCs-Zustands mit erweitertem Differenzierungspotenzial, vergleichbar mit dem 2-Zell-Embryo. *Gadd45a/b/g* defiziente mESCs wiesen ein reduziertes Umschalten in das 2-Zell-Stadium auf, welches ein reduziertes Differenzierungspotenzial in Trophoblaststammzellen verursachte. Die Expressionsanalyse von *Gadd45a/b*-defizienten 2-Zell-Mausembryonen zeigte transkriptionelle Veränderungen für ZGA-spezifische Gene, die die Hypothese stützen, dass GADD45 Proteine an der zygotischen Genomaktivierung *in vivo* beteiligt sind.

Neil1 oder *Neil2* defiziente mESCs wiesen eine verminderte Differenzierungskapazität zu neuronalen und Neuralleisten-Vorläuferzellen auf. Dieser Defekt wurde nicht über die Funktion beider Proteine in der aktiven DNA-Demethylierung ausgelöst, sondern über deren Funktion in der Basenexzisionsreparatur an mitochondrialer DNA (mtDNA). Mechanistisch lösten erhöhte mtDNA Schäden in *Neil1* und *Neil2* defizienten mESCs eine intrinsische, mitochondriale p53-vermittelte DNA Schadensantwort aus, die die neuronale Spezifikation *in vitro* verhinderte.

Diese Arbeit hebt die Bedeutung der Proteine GADD45α,β,γ und NEIL1,2 in verschiedenen Phasen der frühen Mausentwicklung hervor und ordnet deren Beteiligung an der DNA-Demethylierung beziehungsweise DNA-Reparatur als kausal ein.

3 Introduction

3.1 DNA methylation and demethylation

3.1.1 DNA methylation

DNA methylation is one of the key ‘silencing‘ epigenetic modifications playing fundamental roles in embryonic development (Smith and Meissner 2013). Main functions of DNA methylation are X-chromosome inactivation, repression of transposable elements, genomic imprinting and regulation of transcription (Smith and Meissner 2013). DNA methylation patterns are established by the *de novo* DNA methyltransferase 3a (DNMT3a) and DNA methyltransferase 3b (DNMT3b) depositing a methyl group to carbon 5 of cytosines generating 5-methylcytosine (5mC) within a CpG dinucleotide (Cedar et al. 1979; Okano et al. 1999). These DNA methylation patterns are mitotically inherited by the action of DNA methyltransferase 1 (DNMT1) (Leonhardt et al. 1992). The inheritance of full methylation of the nascent strand during replication is ensured by two mechanisms. First, hemi-methylated DNA is the preferential substrate of DNMT1 (Fatemi et al. 2001; Hermann et al. 2004). Second, UHRF1 recognizes hemi-methylated DNA and recruits DNMT1 to the sites of action (Bostick et al. 2007). In 2000, DNMT3L was identified as a cofactor for DNMT3a/b (Aapola et al. 2000). DNMT3L is structurally related to DNMT3a/b but possesses no enzymatic activity (Hata et al. 2002). However, it is required for mediating methylation of maternally imprinted genes and spermatogenesis (Bourc'his et al. 2001; Hata et al. 2002).

DNA methylation in vertebrates is distributed throughout the genome (Suzuki and Bird 2008). Of note, it almost exclusively occurs in the context of symmetric CpG-sequences (Law and Jacobsen 2010). From a global perspective, methylation patterns are bimodaly distributed (Meissner et al. 2008). 60-80% of CpGs in somatic cells are fully methylated, whereas CpG-rich regions, termed CpG-islands (CGI), are unmethylated (Deaton and Bird 2011). Fully methylated regions (FMR) are established to maintain genomic integrity by repressing transcription of repetitive elements or transposable elements (Gopalakrishnan et al. 2009; Smith and Meissner 2013; Walsh et al. 1998). Moreover, this genome-wide repression ensures tissue specific expression (Wu and Zhang 2014). Unmethylated, and thereby non-repressed regions, mostly correlate with developmental regulatory genes (tissue-dependent) and transcription start sites of housekeeping genes (tissue-independent). Methylation of promoter CGIs is rare, however it plays a crucial role in long-term silencing of imprinted genes, the X-chromosome in females and genes that are exclusively expressed in germ cells (Jones 2012). Fully methylated and unmethylated regions are both considered to be stably (non-) methylated. In contrast, lowly methylated regions (LMRs), which show an intermediate to low level of methylation, are suggested to fine tune gene regulation (Stadler et al. 2011). LMRs appear in few CpGs and correlate with CpG-distal regulatory elements. LMRs overlap with enhancer marks and show dynamic changes during differentiation. Non-CpG methylation is to date an underexplored field (Jang et al. 2017; Lister et al. 2009).

DNA methylation achieves its repressive function by various processes. First, the mark itself impairs binding of methylation sensitive transcription factors (Watt and Molloy 1988). Second, 5mC can be bound by methyl-binding proteins further recruiting epigenetic repressors to manifest silencing of these regions, i.e. histone deacetylases (Jones et al. 1998; Nan et al. 1998). Third, DNA methylation can influence nucleosome positioning (Chodavarapu et al. 2010).

Unmethylated regions, like CpG-islands, are protected against methylation. Histone variants like H2A.Z and trimethylation of histone 3 at lysine 4 (H3K4me3) as well as transcription factor binding itself have been shown to possess DNMT-repulsive functions (Ooi et al. 2007; Zilberman et al. 2008).

3.1.2 DNA demethylation

Methylation of cytosine was long time considered to be enzymatically non-reversible since a direct removal of the methyl group from 5mC by a carbon-carbon bond cleavage is thermodynamically unfavorable. However, DNA methylation can be passively removed by preventing the action of DNMT1 during replication, thereby inhibiting the maintenance of DNA methylation patterns (Bird 2002; Dean 2008). This passive dilution can occur either genome-wide i.e. by downregulation of *Dnmt1* expression or nuclear exclusion of DNMT1, or site-specifically by DNMT-repulsive functions. Importantly, this passive dilution mechanism requires the replication process and is therefore restricted to proliferative cells. Intriguingly, during the last decade, researchers uncovered a replication-independent, enzymatic removal of 5mC from genomic DNA, the active DNA demethylation pathway. The mechanism involves the enzymatic function of the Ten-eleven translocation (TET) family proteins (Kohli and Zhang 2013). The family of TET proteins belongs to the Fe^{2+} and α-ketoglutarate dependent family of dioxygenases and consists of three members, TET1, TET2 and TET3. TET proteins iteratively oxidize 5mC to 5-hydroxymethylcytosine (5hmC) (Kriaucionis and Heintz 2009; Tahiliani et al. 2009), 5-formylcytosine (5fC) (Ito et al. 2011; Pfaffeneder et al. 2011) and 5-carboxylcytosine (5caC) (He et al. 2011; Ito et al. 2011). Subsequently, the bases 5fC and 5caC are excised by the thymine DNA glycosylase (TDG) creating an apurinic/apyrimidinic (AP) site and initiating a base excision repair (BER) pathway (He et al. 2011; Maiti and Drohat 2011). BER includes the incorporation of an unmodified cytosine restoring the unmodified DNA structure. This process involves active modification of 5mC and active removal of the oxidative derivatives (AM-AR) (Figure 3.1). Of note, oxidized derivatives of 5mC are also passively diluted upon replication since these modifications do not constitute a substrate for DNMT1-mediated maintenance methylation (Figure 3.1, AM-PD, active modification – passive dilution) (Akalin et al. 2012; Hashimoto et al. 2012). The contribution of other proteins to this pathway is a crucial topic and under high investigation revealing the importance of TET regulation by i.e. posttranslational modifications or TET recruitment to specific sites (Arab et al. 2019; Bauer et al. 2015; Nakagawa et al. 2015; Schäfer et al. 2018).

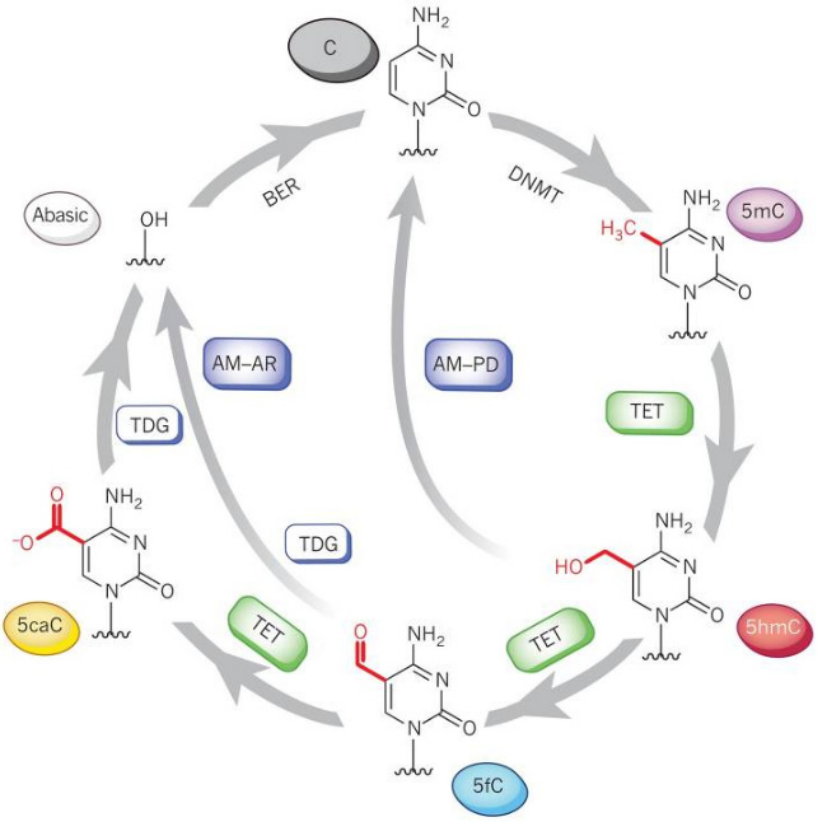

Figure 3.1: TET-mediated pathway of active DNA demethylation

Cytosine is methylated by DNA-methyltransferases (DNMT) generating 5-methylcytosine (5mC). 5mC is iteratively oxidized by the ten-eleven-translocation (TET) enzymes to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). 5hmC, 5fC and 5caC can be passively diluted during replication (AM-PD, active modification - passive dilution). In the AM-AR (active modification – active removal) pathway, 5fC and 5caC can be excised by the thymine DNA glycosylase (TDG) creating an apurinic/apyrimidinic (AP) site including downstream base excision repair (BER) to restore unmodified cytosine. Figure from (Kohli and Zhang 2013).

Noteworthy, 5hmC, 5fC and 5caC may not only act as a demethylation intermediate but also possess their own function as an epigenetic mark. A mass-spectrometry based approach identified specific reader proteins for all oxidized derivatives implying individual functions for each derivative (Spruijt et al. 2013). Moreover, distribution of 5hmC is highly tissue-specific with neurons showing highest abundancy indicating a tissue-specific function of 5hmC (Kinney et al. 2011; Kriaucionis and Heintz 2009). 5fC and 5caC have been shown to interfere with transcriptional elongation by RNA polymerase II (Wang et al. 2015). Furthermore, 5fC is suspected to affect the structure of the DNA double helix due to its chemical properties (Raiber et al. 2015).

TET proteins are expressed in various cell types including the zygote (Wossidlo et al. 2011), primordial germ cells (Hajkova et al. 2010), embryonic stem cells (Ito et al. 2010; Koh et al. 2011) and neurons (Kriaucionis and Heintz 2009), which makes these cell types of great interest to study the biological importance of TET-mediated active DNA demethylation. Interestingly, several studies indicate that

5hmC is enriched at promoters and enhancers of developmental regulatory genes with low-to-intermediate CpG density (Wu et al. 2018). It is important to note that TET proteins also possess a non-catalytic repressive function (Williams et al. 2011).

Besides the biochemically and fully *in vivo* validated TET-TDG pathway, other DNA demethylation processes were proposed, but their existence is unconfirmed. MBD2 was suggested to act as a direct 5mC demethylase (Bhattacharya et al. 1999). Moreover, the direct excision of 5mC by TDG or MBD4 and the deamination of 5mC by AID/APOBEC or DNMT3a/b coupled to BER have been proposed to restore unmodified cytosine (Bhutani et al. 2010; Métivier et al. 2008; Morgan et al. 2004; Zhu, Zheng, Angliker et al. 2000; Zhu, Zheng, Hess et al. 2000).

3.2 Early mouse development

3.2.1 Mouse preimplantation and early phases of postimplantation development

Every animal body is a remarkably complex interplay between different cell types and cell associations fulfilling highly context-specific functions. However, development of animal life starts with a single totipotent cell, the zygote, which is formed by the fusion of two gametes (the oocyte and the sperm) (Chazaud and Yamanaka 2016). Post-fertilization development in the zygote is divided in five pronuclear phases (PN0-5). PN0-2 embryos are in G1-phase, PN3-4 embryos are in the replicating S-phase and PN5 embryos are largely in G2-phase (Santos et al. 2013). One day after zygote formation the first cell division occurs (2-cell stage), which is followed by two subsequent symmetric mitotic divisions giving rise to the 8-cell stage embryo within the zona pellucida (Figure 3.2). The zona pellucida is a specialized extracellular matrix protecting the embryo and inhibiting the growth of the overall cytoplasmic volume. As a consequence, each division of the cells (also termed blastomeres) during this cleavage phase results in cell size reduction (Aiken et al. 2004). At the late 8-cell stage, blastomeres undergo compaction and polarization forming the early morula. Up to this stage, each blastomere of the embryo retains the full totipotent capacity to develop into any cell type of the embryonic or extraembryonic proper (Kelly 1977). At the morula stage, two asymmetric cell divisions segregate the first two lineages: the trophectoderm (TE) giving rise to the placenta, and the inner cell mass (ICM). Transforming a blastocyst with implant abilities requires a process called cavitation. During this phase, the ICM differentiates into the epiblast (EPI) giving rise to the fetal cells, and the primitive endoderm (PE) assigned to the extra-embryonic yolk sac. Subsequently, the blastocyst is comprised of a fluid-filled cavity, the epiblast and the primitive endoderm surrounded by the trophectoderm. The blastocyst hatches the zona pellucida, now capable of uterine implantation.

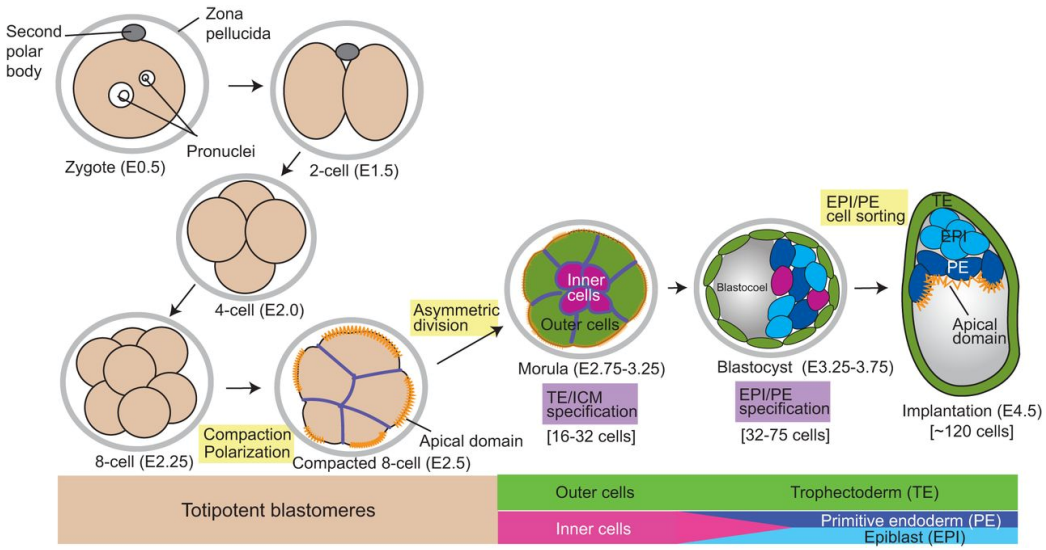

Figure 3.2: Mouse preimplantation development

Schematic illustrating mouse preimplantation development. The zygote divides three times in a symmetric manner forming the early 8-cell stage. Subsequent compaction, polarization and asymmetric division result in the morula stage that defines the first linear segregation between trophectoderm (TE) and inner cell mass (ICM). Further development requires differentiation of the ICM into the epiblast (EPI) and the primitive endoderm (PE) and a process called cavitation leading to a blastocyst capable of implantation. Green, trophectoderm; light blue, epiblast; dark blue, primitive endoderm; yellow, morphogenetic changes; purple, lineage specification steps; orange, apical domains of cell. Figure from (Chazaud and Yamanaka 2016).

The postimplantation phase starts with the process of gastrulation. Due to the high complexity of the postimplantion development, I will describe only the early phases below. During gastrulation the three primary germ layers mesoderm, endoderm and ectoderm evolve (Tam and Loebel 2007). This lineage segregation is highly dependent on cell movements and surrounding signals. At the onset of gastrulation, the primitive streak forms in the posterior region of the embryo. Cells that migrate through the primitive streak differentiate into the mesoderm and the endoderm. At this stage, the time of migration determines the future lineage specification of the cells (i.e. cells ingressing first will differentiate into the extra-embryonic mesoderm). Cells that do not ingress the primitive streak comprise the ectoderm. Generalized, the mesoderm gives rise to skeletal and cardiac muscle, endothelial cells and red blood cells. The endoderm contributes to lung, thyroid and digestive cells, whereas the ectoderm yields the nervous system and the skin.

Interestingly, recent studies suggest the occurrence of asymmetric cell division before the morula stage as molecular heterogeneity seems to be apparent already at the 4- or even 2-cell stage (Torres-Padilla et al. 2007; Wang et al. 2018). Overexpression experiments in single blastomeres suggest that asymmetric expression at these early stages predispose the fate of the single blastomere to the ICM or the TE. Hence, in a strict view, only the zygote and the 2-cell stage embryos are totipotent.

3.2.1.1 Zygotic genome activation

A key event during mouse preimplantation development is the zygotic genome activation (ZGA) occurring at the 2-cell stage. Prior to ZGA, the zygote is fully dependent on maternally provided proteins and mRNAs. In a process called maternal-to-zygotic switch, the maternal RNAs and proteins are depleted and concurrently the transcriptional expression from the zygotic genome is activated. To ensure this process, the cell cycles of the first two cell divisions are elongated as compared to the subsequent ones (20 hours versus 12 hours) (Artus and Cohen-Tannoudji 2008).

ZGA occurs in two phases: a first minor wave, mostly attributed from the male pronucleus (Wiekowski et al. 1993), and a second major wave at the 2-cell stage (Aoki et al. 1997; Hamatani et al. 2004). Understanding the molecular mechanisms of ZGA is of high interest albeit it is technically challenging to assign the contribution of maternally provided or zygotic transcripts to the observed outcome. However, recent studies identified the general modes of activation. Besides the importance of transcription factors, cell cycle regulation and chromatin reorganization have been shown to play a fundamental role in ZGA (Lee, M. T. et al. 2014). Upon fertilization, two compacted chromatins that originated from two distinct cell types need to be unified to create an entire new genome. Consequently, one barrier to ZGA is chromatin decompaction. At ZGA the chromatin is highly accessible (CHO et al. 2002) partially caused by intensive chromatin remodeling through histone exchange (Aiken et al. 2004; Chang et al. 2005; Fu et al. 2003; Nonchev and Tsanev 1990). Furthermore, the erasure of repressive marks contributes to ZGA, i.e. loss of H3K4me3 and increase in H3K27ac (Dahl et al. 2016). In addition, another study showed a novel role for retrotransposons as key regulators of ZGA (Macfarlan et al. 2012). Recent studies identified a zygotic pioneering transcription factor, double homeobox protein (DUX) (Hendrickson et al. 2017; De Iaco et al. 2017). DUX binds to a consensus motif and initiates transcription of key zygotic genes and retroviral elements (Hendrickson et al. 2017). Expression of *Dux* during the minor wave of ZGA is suggested to be induced by the maternally provided DPPA2 and DPPA4 (Eckersley-Maslin et al. 2019).

3.2.2 DNA (de)methylation processes during early mouse development

DNA methylation as well as DNA demethylation are implicated in important regulatory mechanisms during early mouse development. In general, de novo DNA methylation correlates with lineage commitment, whereas erasure of DNA methylation marks is linked to expanded developmental potential. To implement totipotency, epigenetic reprogramming and therefore dramatic methylation losses occur in the early embryos and in primordial germ cells (PGCs). The sperm and the oocyte display specific methylation patterns that need to be erased to reacquire developmental potency. Upon fertilization, the male (~ 90% methylated CpGs) and the female pronuclei (~40% methylated CpGs) undergo global demethylation (Kobayashi et al. 2012; Lee, H. J. et al. 2014; Messerschmidt et al. 2014). Interestingly, DNA demethylation of maternal- and paternal DNA occurs in two different fashions (Eckersley-Maslin et al. 2018). The loss of methylation on maternal DNA is mostly attributed to a

passive manner by nuclear exclusion of DNMT1 (Carlson et al. 1992). Although active DNA demethylation of the maternal DNA has been reported, its contribution is probably rather low (Guo et al. 2014; Shen et al. 2014). In contrast, the highly-methylated paternal DNA is apparently demethylated before the onset of replication, hence, by an active demethylation process (Mayer et al. 2000). Indeed, TET3, unlike TET1 and TET2, is expressed in the zygote and iteratively oxidizes 5mC specifically on the paternal pronucleus (Gu et al. 2011; Iqbal et al. 2011; Wossidlo et al. 2011). Subsequently, the oxidized derivatives of 5mC are out diluted by DNA replication and seem not to be subject to a base excision activity (Inoue et al. 2011; Inoue and Zhang 2011). In line, *Tdg* is expressed at very low levels until the blastocyst stage excluding a major contribution of TDG to DNA demethylation of the paternal genome. (Jessop et al. 2018). Interestingly, the maternal genome shields 5mC from TET3-dependent oxidation by recruiting STELLA to H3K9me2, a mark that is not found in male pronucleus, and thereby altering chromatin configuration (Nakamura et al. 2012). However, recent studies challenge the role of TET3 as the exclusive factor being responsible for the initial loss of 5mC in the male pronucleus. Detailed time-resolving measurements of the loss of 5mC and the occurrence of its oxidation products indicated an initial 5mC drop in early PN3, which was prior to 5hmC and 5fC detection at late PN3 (Amouroux et al. 2016; Santos et al. 2013; Zhu et al. 2017). Moreover, maternal TET3 knockout zygotes still displayed partial DNA demethylation (Guo et al. 2014; Peat et al. 2014; Shen et al. 2014). Alternative pathways to TET3-mediated 5mC loss have been suggested like DNA strand break induced BER or deamination involved DNA repair-based mechanisms (Hajkova et al. 2010; Santos et al. 2013; Wossidlo et al. 2010). Albeit the global demethylation precedes zygotic genome activation (Figure 3.3), the impact of DNA demethylation on ZGA is still controversial (Shen et al. 2014). However, the globally occurring DNA demethylation (excluding imprinted control regions (ICR)) is of high importance to reacquire totipotency.

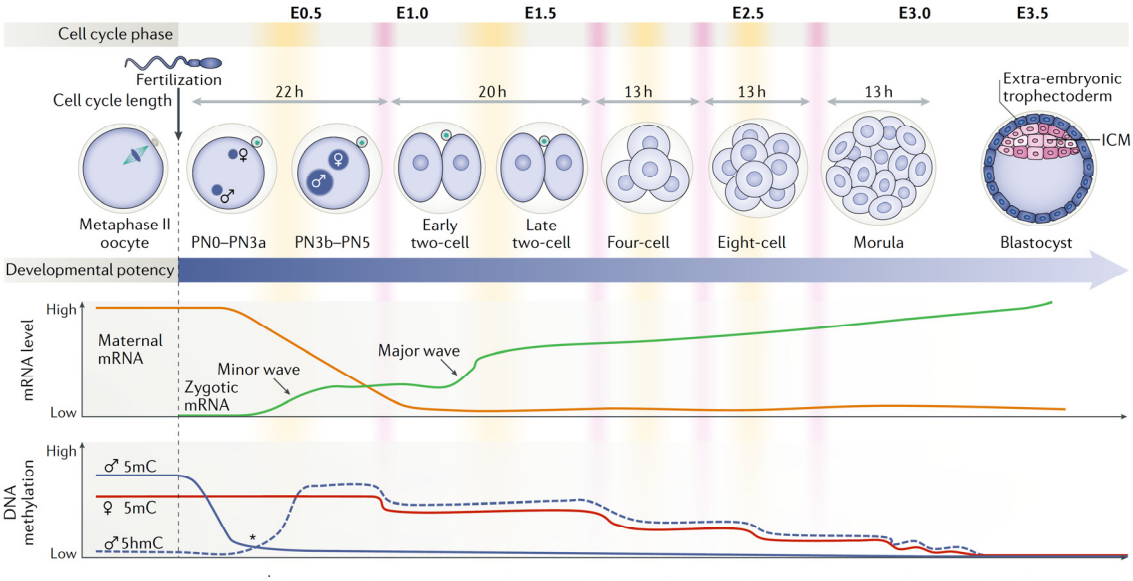

Figure 3.3: Schematic of global transcription and DNA methylation changes during preimplantation development.

Following fertilization the zygotic genome is activated in two phases, the minor and the major wave (green). This is accompanied by the decline of maternal mRNAs (orange). 5-methylcytoysine (5mC, blue line) levels in the paternal genome are rapidly removed in the early pronuclear (PN) phases with concomitant increase in 5-hydroxymethylcytosine (5hmC, blue dashed line) levels. Maternal DNA methylation is mostly lost passively upon replication. * indicates concurrent DNA methylation and hydroxymethylation. Yellow bars, S-phase; red bars, M-phase; E, embryonic day; ICM, inner cell mass. Figure from (Eckersley-Maslin et al. 2018).

Subsequent to implantation, DNA methylation patterns need to be re-established to define lineage commitment and to regulate cell-type specific expression. The *de novo* DNA methyltransferases DNMT3a and DNMT3b establish this methylation pattern as determined by the DNA sequence and CpG density (Krebs et al. 2014; Okano et al. 1999). Consequently, mutations of the *de novo* DNA methylation machinery result in postimplantation defects and *Dnmt3a/b* double knockout mice die around E11.5 (Okano et al. 1999). Interestingly, the TET family has also been implicated in lineage commitment. *Tet1/2/3* triple knockout embryos develop morphologically normal until gastrulation. Upon gastrulation, impaired DNA demethylation results in misregulated Lefty-Nodal signaling affecting early body plan formation (Dai et al. 2016). Unlike for other DNA glycosylases, *Tdg* deficiency in mice results in embryonic lethality suggestive of an implication of TDG in methylation-regulated gene expression (Cortázar et al. 2011). However, it remains unresolved if the function of TDG in active DNA demethylation or in DNA damage repair is responsible for this severe phenotype.

Primordial germ cell (PGCs) specification is a further example of methylation reprogramming during early mouse development. PGCs differentiate from epiblast cells with high methylation levels comparable to somatic cells. While migrating to the genital ridges, PGCs lose 90% of their methylation levels in two phases. Beginning from E8.0, the first wave of DNA demethylation occurs in a passive manner by i.e. downregulation of *Uhrf1* and nuclear exclusion of UHRF1 (Kagiwada et al. 2013; Magnúsdóttir et al. 2013; Seisenberger et al. 2012). Specific sequences, like ICRs or germline-specific

genes remain methylated until the second wave of DNA demethylation, which seems to involve 5mC oxidation by TET1 and TET2 (Hackett et al. 2013; Vincent et al. 2013). The precise mode of action, however, is not fully understood.

3.3 Mouse embryonic stem cells – *in vitro* models of early mouse development

Research of early mouse development is challenging due to the limited amount of material available from early embryos. Although embryos can be pooled and single-cell analysis techniques are emerging *in vitro* models are experimentally highly valuable. In 1981 Evans and Martins isolated mouse embryonic stem cells (mESCs) from blastocyst stage embryos (Evans and Kaufman 1981; Martin 1981). Since then mESCs and its utilization have become a versatile tool for the understanding of the molecular mechanisms of early mouse development.

3.3.1 Mouse embryonic stem cells and in vitro differentiation

Mouse ESCs are isolated from the ICM of preimplantation embryos featuring pluripotency and immortal self-renewal potential *in vitro* (Evans and Kaufman 1981; Martin 1981). On a transcriptional level mESCs resemble the early naïve epiblast consequently with the ability to differentiate into all three germ layers (mesoderm, endoderm, ectoderm) (Boroviak et al. 2014). The key regulators of pluripotency are the transcription factors OCT4, SOX2 and NANOG (Avilion et al. 2003; Chambers et al. 2003; Niwa et al. 2000). The interplay of these transcription factors ensures the expression of core pluripotency markers and the repression of lineage commitment factors in part by cross-regulating each other (Young 2011). Importantly, mESCs are capable to contribute to the germ line upon blastocyst injection into a host animal (Bradley et al. 1984). This facilitates transgenic mouse generation by preselecting successful genetic modifications *in vitro* (Bradley et al. 1992).

To maintain pluripotency mESCs require specific culture conditions. Initially, mESCs were cultured on mitotically inactivated mouse embryonic fibroblast (MEF) feeders in serum-supplemented or conditioned medium to provide necessary factors for pluripotency (Evans and Kaufman 1981; Martin 1981). Later, LIF has been shown to be capable to maintain pluripotency in the absence of MEF feeders through activation of the JAK/STAT signaling pathway (Niwa et al. 2000; Smith et al. 1988; Williams et al. 1988). Although mESCs cultured on MEF feeders with serum and LIF is still widespread, some mESC lines were adapted to serum-supplemented LIF conditions on gelatin-coated surfaces (Smith 1991). Serum-supplementation fulfills two functions. First, it suppresses autocrine FGF signaling, which induces neuronal development in an unrepressed state (Kunath et al. 2007; Ying, Stavridis et al. 2003). Second, the serum protein BMP maintains pluripotency, but only in combination with LIF (Ying, Nichols et al. 2003). Under these serum-supplemented culture conditions, mESCs resemble a heterogeneous cell population with a diverse developmental potential (Hayashi et al. 2008; Macfarlan et al. 2012; Toyooka et al. 2008). To obtain a more homogenous population and to evolve ground-state pluripotency, mESCs are cultured in serum-free medium termed ‘2i‘ (Ying et al. 2008). 2i medium contains two inhibitors (i), a MEK/ERK and a GSK3 inhibitor, in the presence of LIF. MEK/ERKi and

GSK3i repress FGF-mediated differentiation and sustain self-renewal capacity, respectively. Latter effect is caused by activation of the canonical Wnt-signaling pathway, which has been proposed to be important for pluripotent maintenance (Wray et al. 2010).

In regard to methylation patterns, mESCs cultured in serum-containing medium predominantly resemble the postimplantation stage whereas mESCs cultured in 2i are globally hypomethylated mirroring the preimplantation development (Habibi et al. 2013). Global hypomethylation in 2i-medium cultured mESCs is caused by repression of *Dnmt3a/b* on mRNA and protein level and by promoting TET1 activity (Ficz et al. 2013; Leitch et al. 2013; Sim et al. 2017).

Reversing lineage commitment of somatic cells and regaining pluripotency has been of great interest in the past decade. Work by Takahashi and Yamanaka identified a mechanism to reprogram mouse embryonic or adult fibroblasts to induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka 2006). Overexpression of *Oct4*, *Sox2*, *c-myc* and *Klf4* resulted in expression of stem cell marker genes, display of mESC growth behavior and morphology and chimeric contribution when injected into a host blastocyst.

Apart from understanding the molecular mechanisms of pluripotency, mESCs can be utilized to investigate lineage commitment by differentiating them into various cell types upon specific culture conditions. In the absence of LIF, mESCs differentiate in an undirected manner and express mesodermal and endodermal marker (Smith 2001). A method to induce defined differentiation of mESCs is cellular aggregation in cell suspension, also termed embryoid body (EB) differentiation (Keller 2005). EBs in part represent the developmental program of a postimplantation blastocyst since they self-organize and give rise to derivatives of all three germ layers. In the presence of certain growth factors, these precursor cell types can further differentiate along a defined lineage developing to i.e. cardiomyocytes (Doetschman et al. 1985), hematopoetic precursors (Wiles and Keller 1991), adipocytes (Dani et al. 1997) or smooth muscle cells (Yamashita et al. 2000). Addition of retinoic acid biases cell fate to the neuronal lineage (Bibel et al. 2007).

Mouse ESCs are isolated from the ICM and therefore lost their potential to differentiate into the trophoblast lineage. However, in chimera assays a minor fraction of mESCs contributes also to the trophoblast lineage (Beddington and Robertson 1989). Of note, mESCs can sporadically transdifferentiate into mouse trophoblast stem cells (mTSCs) *in vitro* (Hayashi et al. 2010). This transdifferentiation can be enhanced by BMP4.

3.3.1.1 DNA (de)methylation in embryonic stem cells and in vitro differentiation

Mouse ESCs are a valuable model system to investigate the interplay between DNA (de)methylation and lineage commitment. Knockout (KO) mESC studies suggest that regulation of DNA methylation patterns is dispensable for sustaining pluripotency. *Dnmt1/3a/3b* triple knockout (TKO) mESCs are pluripotent and show normal growth behavior with only minor transcriptional changes (Tsumura et al.

2006). However, the specific regulation of DNA methylation is required upon lineage commitment. While initial differentiation of DNMT TKO mESCs is inducible, later stage embryoid bodies show incomplete silencing of the pluripotency machinery (Schmidt et al. 2012). This scenario accurately reflects the *in vivo* situation where *Dnmt* TKO embryos develop postimplantation but are impaired in regulating lineage commitment. *Tet* TKO mESCs express pluripotency marker, but only poorly contribute to chimeric embryos. In line with this, *Tet* TKO mESCs have poor differentiation potential *in vitro* (Dawlaty et al. 2014). Noteworthy, *Tet1* and *Tet2* but not *Tet3* are expressed in undifferentiated mESCs (Koh et al. 2011). *Tet3* expression is induced during differentiation. Similar to TETs, TDG is dispensable for mESC pluripotency and self-renewal but *Tdg* KO mESCs are unable to differentiate into neuronal progenitor cells (Cortázar et al. 2011).

DNA methylation patterns seem to build major barriers for reprogramming as demonstrated by the two following examples. First, iPSC reprogramming is enhanced in presence of DNA methylation blocking agents (Wernig et al. 2008). Second, mESC to mTSC transdifferentiation involves demethylation of the promoter of the mTSC-specific transcription factor *Elf5* (Ng et al. 2008). Consequently, mESCs devoid of DNA methylation are capable of differentiating into trophoblast derivatives.

3.3.2 2C-like cells

Mouse ESCs cultured in serum condition display a heterogeneous mixture of subpopulations owing different developmental potential (Chambers et al. 2007; Hayashi et al. 2008; Singh et al. 2007; Toyooka et al. 2008; Zalzman et al. 2010). In 2012, Macfarlan and colleagues identified an additional subpopulation in mESCs revolutionizing the study of ZGA *in vitro* (Macfarlan et al. 2012). Expression analysis of this subpopulation revealed the transcription profile of 2-cell stage embryos including ZGA-specific genes. This cell fraction, termed 2C-like cells, represent 1-5% of the mESC population and express, similar to 2-cell embryos, high levels of the murine endogenous retrovirus with leucine tRNA primer (*MERVL*) retrotransposon (Macfarlan et al. 2012). Each cell fluctuates through this state at least once within nine passages presumably to sustain the pluripotent potential (Amano et al. 2013; Zalzman et al. 2010). Importantly, upon injection into preimplantation embryos, 2C-like cells contribute to the embryonic as well as extra-embryonic tissue indicating expanded fate potential (Macfarlan et al. 2012). Moreover, the 2C-like state is activated upon iPSC reprogramming indicating its ability to recapture developmental potential (Eckersley-Maslin et al. 2016). 2C-like cells not only share the transcriptional profile with 2-cell embryos but also manifest the 2-cell-specific epigenetic state i.e. high levels of H3K4me2 and histone acetylation and increased chromatin accessibility (Eckersley-Maslin et al. 2016; Hendrickson et al. 2017; Macfarlan et al. 2012). Two markers are characteristic for the 2C-like state, the highly active *MERVL* (Macfarlan et al. 2012). and the transcription factor zinc finger and SCAN domain-containing protein (ZSCAN4), which is exclusively expressed in 2-cell stage embryos during preimplantation development (Zalzman et al. 2010). Reporter systems using fragments of the *MERVL* sequence or the *Zscan4* promoter provided insight in the molecular mechanism regulating 2C-like

cycling. Till date, several 2C-like repressors were identified, the histone modifying enzymes KAP1, LSD1 and G9a (Macfarlan et al. 2012), the histone chaperone CAF-1 (Ishiuchi et al. 2015) and the LINE1-nucleolin complex (Percharde et al. 2018) among others. In contrast, positive regulators remained elusive until the discovery of the transcription factor DUX as the first described 2C-like promoter (Hendrickson et al. 2017; De Iaco et al. 2017). DUX binds to a consensus motif and initiates transcription of key zygotic genes and retroviral elements (Hendrickson et al. 2017). *Dux* expression itself is suggested to be regulated by DPPA2 and DPPA4, which are maternally provided in the mouse zygote (Eckersley-Maslin et al. 2019). Other potential zygotic (transcription) factors responsible for the activation of the key event during preimplantation development remain to be discovered.

The role of DNA (de)methylation in 2C-like regulation is contradictory. 2C-like cells display global DNA demethylation upon 2C-like cycling (Eckersley-Maslin et al. 2016). However, induction of global DNA demethylation using a *Dnmt1* inducible-knockout system does not affect occurrence of 2C-like cells. This might indicate that DNA demethylation is important for 2C-like maintenance but not for the establishment of this state. Inconsistently, 2i medium, which induces global DNA demethylation, was shown to diminish the 2C-like population (Eckersley-Maslin et al. 2016; Macfarlan et al. 2012). *Tet* KO studies provoked even more discrepancies. *Tet1/2/3* TKO mESCs express higher levels of 2C-like genes and harbor an increase in 2C-like cell population as defined by ZSCAN4 protein level (Lu et al. 2014). Interestingly, the *MERVL*- reporter was stably integrated in the *Tet* TKO mESCs, but solely used to study the maintenance of *MERVL*$^{+}$ cells after sorting. In contrast, *Tet1/2/3*-deficient epiblast-derived stem cells show a decrease in *Zscan4* expression (Khoueiry et al. 2017).

3.4 GADD45 protein family

The GADD45 (Growth arrest and DNA damage-inducible) family is comprised of three small (18 kDa), evolutionary conserved, highly acidic proteins, which show a similarity of around 55% in amino acid sequence (Liebermann and Hoffman 2008; Zhan et al. 1994). GADD45α, GADD45β (alias MYD118) and GADD45γ (alias CR6) belong to a superfamily of RNA-binding proteins named L7Ae/L30e/S12 (Pfam entry: PF01248). Albeit GADD45 proteins do not show any enzymatic activity they fulfill various and frequently redundant functions via interaction with effector proteins (Akalin et al. 2012; Liebermann and Hoffman 2008; Ma, Guo et al. 2009).

3.4.1 General functions of GADD45 proteins

GADD45 proteins are well established sensors of cellular stress upon genotoxic or differentiation signals (Liebermann and Hoffman 2008). GADD45 proteins accomplish their functions by interacting with various proteins to orchestrate the stress response. PCNA, p21, cdc2/cyclin B1 complex, p38 and MEKK4 are just a few examples belonging to the diversified network of GADD45 protein-protein interactions. GADD45 proteins possess three major functions upon cellular stress. First, they regulate cell cycle progression by interacting with CRIF1, p21 and the cdc2/cyclin B1 complex resulting in G1/S and G2/M cell cycle arrest, respectively (Chung et al. 2003; Liebermann and Hoffman 2008; Vairapandi

et al. 2002). Second, GADD45 proteins are implicated in nucleotide excision repair (NER) and BER due to their interaction with PCNA and their chromatin relaxation, facilitating DNA accessibility (Carrier et al. 1999; Smith et al. 1994). Third, GADD45 proteins can cause apoptosis/cell survival/differentiation mainly by interaction and activation of MAPK-signaling members (i.e. MEKK4, p38) (Bulavin et al. 2003; Takekawa and Saito 1998). Latest findings revealed a novel role for GADD45α/ING1 in aging due to regulation of enhancer methylation (see below and (Schäfer et al. 2018)).

Induction as well as the response pathway of GADD45 proteins upon genotoxic stress is highly cell-type specific. The versatile and cell-type specific functions of GADD45 proteins are displayed by the pleiotropic phenotypes observed in *Gadd45* single mutant mice. *Gadd45* single mutant mice are viable and fertile (Hollander et al. 1999; Lu et al. 2001; Lu et al. 2004). However, described phenotypes comprise among others increased genomic instability and tumorigenesis, immune deficiencies, exencephaly, impaired long-term memory and sex reversal (Gierl et al. 2012; Hildesheim et al. 2002; Hollander et al. 1999; Leach et al. 2012; Liu et al. 2005; Lu et al. 2001; Sultan et al. 2012). Additional phenotypes might be suppressed due to compensatory redundancy within the GADD45 family of proteins. Although challenging, *Gadd45* triple knockout mice have been generated and they are viable and fertile (Cai et al. 2006). However, statistical analysis of phenotypes is lacking due to low offspring frequency and therefore remain elusive.

3.4.2 DNA demethylation and GADD45 proteins

GADD45 proteins are implicated in diverse repair-mediated DNA demethylation pathways. Initially, Gadd45a was found as a hit in a *Xenopus* cDNA screen for factors that reactivate expression on a methylation silenced reporter plasmid (Barreto et al. 2007). Gadd45a binds XPG and was postulated to recruit the NER machinery for repair-mediated DNA demethylation. Further studies supported the promoting function of GADD45 proteins in site-specific DNA demethylation in mammals (Jin et al. 2008; Rajput et al. 2016; Schäfer et al. 2010; Schmitz et al. 2009). In addition, Gadd45a promotes demethylation by coupled 5mC deamination and Mbd4-mediated T:G mismatch repair in zebrafish (Rai et al. 2008). Moreover, human GADD45α has been shown to impair DNMT1 activity during homology-directed repair (Lee et al. 2012). Recent investigations discovered that GADD45 proteins are also involved in TET-TDG-mediated DNA demethylation. In this regard, GADD45 fulfills a dual function by enhancing TET1 activity as well as 5fC and 5caC removal by TDG (Kienhöfer et al. 2015; Li et al. 2015).

GADD45 proteins are RNA binders, which opens up the intriguing possibility that they promote demethylation via RNA-guides recruiting the demethylation machinery to specific genomic loci. Indeed, GADD45α binds the long non-coding RNA *TARID* to facilitate demethylation of the *Tcf21* locus involving DNA-RNA complementary hybrid structures (Arab et al. 2014). Moreover, GADD45α is an R-loop binder and recruits TET to numerous promoter CGIs (Arab et al. 2019). In addition, GADD45

proteins also use protein factors to recruit the DNA demethylation machinery to sites of action (e.g. TAF12 and ING1) (Schäfer et al. 2013; Schäfer et al. 2018; Schmitz et al. 2009). Interestingly, GADD45α has been shown to bind hemi-methylated DNA which might regulate the maintenance methylation (Lee et al. 2012). The importance of GADD45 proteins promoting DNA demethylation in early mouse development remains elusive. *Dnmt*, *Tet* and *Tdg* knockout studies demonstrated the significance of DNA methylation regulators in early mouse development therefore evolving great interest in the involvement of GADD45 proteins in this process.

3.5 DNA glycosylases

The DNA of a cell is continuously exposed to potential threats resulting in i.e. oxidation, alkylation or deamination of the base (Lindahl and Wood 1999). Cells respond to this damage by either inducing apoptosis or, if the damage is repaired, cell survival including activation of checkpoint pathways (Fleck and Nielsen 2004). If unrepaired, these damaged bases cause mutations or stall replication contributing to genomic instability (Kunz et al. 2009; Lindahl and Wood 1999). The genetic lesions can cause neurodegenerative or cardiovascular diseases, premature aging and carcinogenesis (Chen et al. 2012; Dizdaroglu 2015; Elnakish et al. 2013; Sohal and Orr 2012). DNA glycosylases counteract genetic lesions and fulfill the initial steps in BER, the recognition and the excision of a damaged base (Jacobs and Schär 2012). DNA glycosylases have the same mode of action in recognizing damaged bases and possess in part overlapping substrate spectra (Jacobs and Schär 2012). Mechanistically, DNA glycosylases cleave the N-glycosidic bond between the damaged base and the ribose moiety creating abasic sites that are further processed by AP endonucleases, DNA polymerases and ligases to complete BER. Besides the well-studied function of DNA glycosylases in base damage removal, they are implicated in the process of active DNA demethylation. (He et al. 2011; Jost et al. 1995; Maiti and Drohat 2011; Schomacher et al. 2016; Zhu, Zheng, Angliker et al. 2000).

3.5.1 NEIL DNA glycosylases

The nei endonuclease VIII-like family of DNA glycosylases constitute three (NEIL1, NEIL2 and NEIL3) of the eleven known DNA glycosylases identified in mammals (Jacobs and Schär 2012). NEIL DNA glycosylases preferentially process oxidized base lesions like thymine glycol (Tg), spiroiminodihydantoin (Sp), guanidinohydantoin (Gh), 2,6-Diamino-4-hydroxy-5-formamidopyrimidine (FapyG), 4,6-diamino-5-formamidopyrimidine (FapyA) (Schomacher and Niehrs 2017). NEIL1 and NEIL2 are bifunctional enzymes acting not only as DNA glycosylases in base removal but also as AP lyases. Both enzymes process abasic sites via β, δ-elimination to create a 1 nucleotide gap with 3‘- and 5‘-phosphate residues (Schomacher and Niehrs 2017). NEIL3 differs from NEIL1 and NEIL2 as it functions mainly as a single-strand specific and monofunctional DNA glycosylase. In addition, NEIL3 is capable to unhook psoralen-induced DNA interstrand cross-links (Krokeide et al. 2013; Semlow et al. 2016).

Single *Neil*-deficient mice are viable, fertile and bear rather mild phenotypes very similar to other glycosylase-deficient mouse mutants. Mice deficient for *Neil1* display metabolic syndrome, increased brain damage and impaired short-term spatial memory retention (Canugovi et al. 2012; Vartanian et al. 2006). The latter two only occurring upon ischemic stroke. Moreover, *Neil1*-deficient mice exhibit an increased frequency of mitochondrial DNA damage (Vartanian et al. 2006). *Neil2* knockout mice accumulate oxidized DNA bases with increasing age and are sensitive to inflammatory agents (Chakraborty et al. 2015). *Neil3* deficient mice show a reduced number in microglia, loss of proliferative neural progenitors upon ischemic stroke and impaired adult neurogenesis (Regnell et al. 2012; Sejersted et al. 2011). In contrast, deficiency in *Tdg* or downstream BER factors cause embryonic lethality (Cortázar et al. 2011; Friedberg and Meira 2006).

The rather mild phenotypes of single *Neil*-deficient mice may be due to functional redundancy of the residual family members. Surprisingly, *Neil* mutant mice deficient for all three family members do not show any overt and additive phenotype nor a cancer predisposition (Rolseth et al. 2017). Thus, NEIL DNA glycosylases might not compensate for each other but rather have acquired specialized functions. Of note, *Neil1/2/3* triple knockout mice were characterized under physiological conditions. Environmental stress might trigger phenotypical abnormalities. In *Xenopus*, knockdown of *neil2* causes a severe neural crest differentiation defect (Schomacher et al. 2016).

3.5.2 NEIL DNA glycosylases and DNA demethylation

The involvement of DNA glycosylases in DNA demethylation was first proposed for TDG supposed to excise 5mC directly (Zhu, Zheng, Hess et al. 2000). However, subsequent studies convincingly demonstrated that TDG removes 5fC and 5caC rather than 5mC or 5hmC (Cortázar et al. 2011; He et al. 2011; Maiti and Drohat 2011). In search of additional DNA glycosylases involved in 5fC and 5caC excision, our laboratory performed a glycosylase knockdown screen in HeLa cells (Schomacher et al. 2016). Interestingly, knockdown of *NEIL1* and *NEIL2*, but not *NEIL3* impaired the excision of 5fC and 5caC *in vitro*. In line, NEIL DNA glycosylases bind oxidized derivatives of 5mC and overexpression of *Neil1, 2* and *3* in mESCs could compensate for a *Tdg* knockout (Müller et al. 2014; Spruijt et al. 2013). Unexpectedly, NEIL1 and NEIL2 do not directly excise 5fC or 5caC but rather promote the substrate turnover of TDG (Schomacher et al. 2016). TDG as a monofunctional DNA glycosylase is product inhibited upon excision of 5fC or 5caC and, hence, stalled at the AP-site. NEIL1 and NEIL2 release stalled TDG by competing for AP site binding, thereby enhancing 5fC and 5caC turnover by TDG. In *Xenopus*, Neil2 is required for 5fC/5caC removal *in vivo*. The function of NEIL1 and NEIL2 in promoting 5fC and 5caC removal in mouse development remains elusive.

3.6 Aim of the thesis

Previous studies in our laboratory identified GADD45 proteins, NEIL1 and NEIL2 as regulators of active DNA demethylation. Dynamic regulation of DNA methylation is a crucial process during mouse development. However, the physiological importance of GADD45 or NEIL1 and NEIL2 in DNA demethylation in early mouse development remains elusive. The aim of the thesis is to address this open question using mESCs as *in vitro* system. The results are represented in two chapters.

In chapter 1, I investigated the role of GADD45 proteins in 2C regulation. To verify GADD45 proteins as novel regulators of the 2C-like state, a 2C-specific reporter had to be introduced in control and *Gadd45* TKO mESCs. Gene expression and methylation analysis in corroboration with overexpression of specific GADD45 interactors in 2C-positive or negative *Gadd45* TKO and control mESC should give insight in the regulatory mechanism of GADD45 proteins. Furthermore, RNA-sequencing analysis of *Gadd45* deficient 2-cell mouse embryos should reveal the *in vivo* function of GADD45 proteins. This work has been published in *Genes and Development* (Schüle et al. 2019). The publication is attached as submitted to the journal and references from chapter 1 are listed separately. I carried out 2C-related experiments.

In chapter 2, I interrogated the cause and the consequences of the neuronal/neural crest differentiation defect in *Neil1* or *Neil2* deficient mESCs. Embryoid body differentiation assays with addition of potential rescuing agents should elucidate the origin of the differentiation defect. Transciptomic and proteomic analysis should identify the potential downstream effects of *Neil1* and *Neil2* deficiency.

4 Results

4.1 Chapter 1: GADD45 promotes locus specific DNA demethylation and 2C cycling in embryonic stem cells (Schüle et al., *Genes Dev* 33, 782-798, 2019)

GADD45 promotes locus specific DNA demethylation and 2C cycling in embryonic stem cells

Katrin M. Schüle[1,3], Manuel Leichsenring[1,3], Tommaso Andreani[1], Viviana Vastolo[1], Medhavi Mallick[1], Michael U. Musheev[1], Emil Karaulanov[1], and Christof Niehrs[1,2,4]

Affiliations:

[1] Institute of Molecular Biology (IMB), 55128 Mainz, Germany

[2] German Cancer Research Center (DKFZ), Division of Molecular Embryology, 69120 Heidelberg, Germany

[3]Equal contribution

[4]Corresponding author: c.niehrs@imb-mainz.de

Shortened title: GADD45 in DNA demethylation & 2C cycling

Keywords: demethylation, GADD45, 5-hydroxymethylcytosine, TET, 2-cell embryo, ZGA, ZSCAN4.

Mouse embryonic stem cell (ESC) cultures contain a rare cell population of '2C-like' cells resembling 2 cell embryos, the key stage of zygotic genome activation (ZGA). Little is known about positive regulators of the 2C-like state and 2-cell stage embryos. Here we show that GADD45 (Growth arrest and DNA damage 45) proteins, regulators of TET mediated DNA demethylation, promote both states. Methylome analysis of *Gadd45a,b,g* triple-knockout (TKO) ESCs revealed locus-specific DNA hypermethylation of ~7,000 sites, which are enriched for enhancers and loci undergoing TET-TDG mediated demethylation. Gene expression is misregulated in TKOs, notably upon differentiation, and displays signatures of DNMT and TET targets. TKOs manifest impaired transition into the 2C-like state and exhibit DNA hypermethylation and downregulation of 2C-like state specific genes. *Gadd45a,b* double mutant mouse embryos display embryonic sublethality, show deregulated ZGA gene expression, and developmental arrest. Our study reveals an unexpected role of GADD45 proteins in embryonic 2 cell stage regulation.

Introduction

Mouse embryonic stem cells (ESCs) are a model for the inner cell mass around implantation stage. ESCs are heterogeneous and contain subpopulations with different properties (Hayashi et al. 2008; Macfarlan et al. 2012; Toyooka et al. 2008; Zalzman et al. 2010). One of these subpopulations (1-5%) is transcriptionally and epigenetically similar to the 2-cell stage embryo and hence referred to as '2C-like' (Macfarlan et al. 2012; Zalzman et al. 2010). The embryonic 2-cell stage is a key phase of mouse development during which the major wave of zygotic genome activation occurs (ZGA; reviewed in (Eckersley-Maslin et al. 2018; Jukam et al. 2017; Svoboda 2017). During this period, the bulk of the genome becomes transcriptional active, which is accompanied by extensive chromatin modification. 2C-like ESCs exhibit unique molecular features of totipotent cleavage-stage cells and in chimeras they can contribute to both embryonic and extraembryonic derivatives, including trophoblast (Choi et al. 2017; De Iaco et al. 2017; Ishiuchi et al. 2015; Macfarlan et al. 2012; Rodriguez-Terrones et al. 2018; Whiddon et al. 2017). ESCs cycle in and out of this transient 2C-like state at least once within nine passages (Zalzman et al. 2010). Characteristic markers for the 2C-like state are murine endogenous retrovirus with leucine tRNA primer (*MERVL*) retrotransposon and Zinc finger and SCAN domain-containing protein 4 (*Zscan4*; Macfarlan et al. 2012; Zalzman et al. 2010). Thus, 2C-like ESCs model essential aspects of the 2-cell stage embryo and ZGA (reviewed in (Eckersley-Maslin et al. 2018; Ishiuchi and Torres-Padilla 2013)). Few positive regulators of the 2C-like state or ZGA are known, including the transcriptional regulators ZSCAN4 (Amano et al. 2013; Falco et al. 2007; Hirata et al. 2012; Zalzman et al. 2010), DUX (Hendrickson et al. 2017; De Iaco et al. 2017), STELLA (Huang et al. 2017), TBX3 (Dan et al. 2013), as well as DPPA2 and DPPA4 (Eckersley-Maslin et al. 2019).

Here we report a role for the small gene family *Gadd45* (Growth arrest and DNA damage protein 45) *a,-b* and *-g*, in regulation of the 2C-like state. GADD45α is a stress response protein, which interacts with the key enzymes of the DNA demethylation machinery, TET1 (TET methylcytosine dioxygenase

1) and TDG (Thymine DNA glycosylase; Barreto et al. 2007; Cortellino et al. 2011; Kienhöfer et al. 2015; Li et al. 2015). TET enzymes convert 5-methylcytosine (5mC) sequentially to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC; Guo et al. 2011; He et al. 2011; Ito et al. 2011; Kriaucionis and Heintz 2009; Tahiliani et al. 2009). DNA repair via TDG removes 5fC and 5caC to restore unmethylated cytosine (Cortázar et al. 2011; Cortellino et al. 2011; Shen et al. 2013). GADD45α is an adapter protein that tethers TET/TDG to sites of DNA demethylation, which functions in locus-specific DNA demethylation (Arab et al. 2014; Barreto et al. 2007; Cortellino et al. 2011; Li et al. 2010; Sabag et al. 2014; Zhang, et al. 2011a). GADD45α recruits TET/TDG to specific sites in the genome via additional cofactors (Arab et al. 2014; Arab et al. 2019; Schäfer et al. 2013; Schäfer et al. 2018).

Since not only *Gadd45a* but also *Gadd45b* and *Gadd45g* promote DNA demethylation (Gavin et al. 2015; Jarome et al. 2015; Ma et al. 2009; Rai et al. 2008; Sen et al. 2010) and since single mouse mutants are viable (Hollander et al. 1999; Lu et al. 2001; Lu et al. 2004), this raises the question of whether the genes have overlapping roles in development and differentiation. To address this question we have generated and characterized *Gadd45a,b,g* triple–knockout mouse embryonic stem cells (TKO-ESCs). We find that GADD45 proteins are dispensable for maintaining pluripotency and self-renewal. However, methylome analysis indicates that GADD45 proteins are required for DNA demethylation of specific loci and normal gene expression. Moreover, GADD45 proteins promote the 2C-like state and *Gadd45a,b* double mutant mouse embryos show partial deregulation of ZGA genes at the 2-cell stage and developmental arrest. Collectively, the results indicate that GADD45 proteins act redundantly to promote locus-specific demethylation as well as embryonic 2-cell stage.

Results

Gadd45 **TKO-ESCs are pluripotent and self-renew**

We generated homozygous deletions in *Gadd45a*, *Gadd45b* and *Gadd45g* in ESCs using the CRISPR/Cas9 system (Jinek et al. 2012; Le Cong et al. 2013; Mali et al. 2013). Six gRNAs were cotransfected, two for each *Gadd45* gene, to create 300-700 bp deletions between the 5'-UTR and the second intron, covering the start codon (Fig. 1A). Out of 276 colonies obtained after selection, three independent *Gadd45* triple-knockout (TKO) ESC clones were obtained (Fig. S1A). Sequencing confirmed deletion of the respective genomic regions in the TKO-ESCs (Fig. S1B) and Western blot and mass-spectrometry analysis showed that both GADD45α and GADD45β were undetectable in TKO-ESCs (Fig. S1C-D). GADD45γ was undetectable in both wild type and mutant ESCs (Fig. S1E) and even if truncated GADD45γ protein was expressed, it would be non-functional since deletion of exons 1 and 2 include the dimerization domains (aa43-86) required for GADD45γ function (Zhang, et al. 2011b). To generate three independent wild type ESC control (Co) clones, ESCs were transfected with Cas9 and the selection marker, but without specific gRNAs.

The *Gadd45* TKO-ESCs showed no apparent loss of pluripotency, *Oct4*, *Nanog* and *Sox2* expression was not reduced (Fig. 1B), and their morphology as well as growth rate was normal (Fig. S2A-C). In teratoma assays, TKO-ESCs gave rise to derivatives of all three germ layers (Fig. 1C).

Given the previously described functions of GADD45 proteins in active DNA demethylation, we analyzed global 5mC, 5hmC, 5fC, and 5caC levels by quantitative mass spectrometry. Global levels of 5mC and its oxidative derivatives were only mildly affected in *Gadd45* TKO-ESCs (Fig. 1D-G). There was a slight increase of 5mC with concomitant slight decrease of oxidized cytosine derivatives in TKO-ESCs compared to Co-ESCs. This result is consistent with GADD45 proteins acting not in global- but in locus-specific demethylation.

Mouse ESCs are a model for the inner cell mass (ICM) around implantation stage with relatively high methylation at promoters, enhancers and bivalent loci (Habibi et al. 2013). However, ESCs can be reverted to a hypomethylated ground state more similar to pre-implantation embryos using small-molecule MEK and GSK3β inhibitors ('2i'; Ficz et al. 2013; Leitch et al. 2013), as well by vitamin C (Blaschke et al. 2013). Hence, we induced global DNA demethylation via vitamin C or 2i treatment, however global levels of cytosine modifications in *Gadd45* TKO-ESCs were changing similarly to Co-ESCs (Fig. S2D – S2E).

We conclude that *Gadd45* genes are dispensable for ESC maintenance, as is the case for *Tet* and *Dnmt* genes (Dawlaty et al. 2011; Dawlaty et al. 2014; Lei et al. 1996; Okano et al. 1999; Tsumura et al. 2006).

Loci undergoing TET-dependent oxidation are hypermethylated in TKO-ESCs

To unravel DNA methylation changes we performed whole-genome bisulfite sequencing (WGBS) of Co- and TKO-ESCs to obtain base-pair-resolution methylomes. Co-ESCs showed the characteristic bimodal distribution of CpG methylation, but the distribution in TKO-ESCs was skewed towards higher methylation (Fig. 2A). To call differentially methylated CpG-regions (DMRs) in TKO-ESCs, we used a stringent cutoff of 5% FDR and > 30% methylation difference on at least two CpGs. WGBS does not discriminate 5hmC from 5mC (Booth et al. 2012), hence we may underestimate the number of GADD45- dependent TET target sites. DMRs were broadly distributed on all 19 autosomes (data not shown). The DMR analysis identified 6,904 hypermethylated, but only 34 hypomethylated regions in TKO-ESCs (Fig. S3A). Thus, although global 5mC levels were only slightly increased by LC-MS/MS, the skewed bimodal methylation pattern and the ~200-fold bias towards hypermethylated DMRs indicate a locus-specific DNA hypermethylation in TKO-ESCs.

We therefore focused on the hypermethylated DMRs (hyper-DMRs). The majority overlapped with intronic and intergenic regions (Fig. S3B). There was around 2.5-fold enrichment for enhancers and coding exons (Fig. 2B). Moreover, hyper-DMRs were enriched at sites marked by 5fC or 5caC after *Tdg* knockdown (Shen et al. 2013) and to a lesser extent at sites marked by 5hmC (Kong et al. 2016; Shen et al. 2013). Hyper-DMRs were also enriched for TET-dependent hyper-DMRs (Lu et al. 2014).

Positional correlation analysis centered on hyper-DMRs revealed prominent overlap with sites marked by 5hmC (Fig. 2C). Hyper-DMRs also overlapped with 5fC/5caC peaks accumulating in *Tdg* knockdown ESCs (Shen et al. 2013), indicating that hyper-DMRs are targets of TET/TDG-mediated cytosine oxidation and excision in ESCs. In contrast, there was little overlap with 5mC sites, indicating that the association of hyper-DMRs occurred specifically with oxidized cytosines.

In another positional correlation analysis, we plotted the average levels of methylation change between TKO- and Co-ESCs against the center of genomic features derived from a wide panel of published genome-wide mapping data sets in ESCs, including 5mC oxidative derivatives, DNA-binding factors, and major histone modifications. Moreover, we divided the analysis between proximal and distal elements with regard to gene transcription start sites (TSS). This analysis corroborated that the main co-occurrence of hypermethylation in TKO-ESCs was with sites marked by 5fC and 5caC, both in proximal and distal sites (Fig. 2D, top three rows). Hypermethylation accumulated also at the center of 5hmC peaks, but to a lower level and restricted to distal loci.

Conversely, CpG islands (CGIs), which typically occur at proximal sites, showed the lowest levels of methylation difference, consistent with the fact that they are constitutively unmethylated (Deaton and Bird 2011). Other genomic features correlated with promoter CGIs, such as Pol2, TBP, and transcription elongation factors (Nelfa, Spt5) followed this trend. In general, this methylation signature parallels the signature in *Tet1,2,3* TKO-ESCs (Lu et al. 2014), whereby TET mediated DNA demethylation i) mainly occurs at distal regulatory elements and ii) affects sites marked in control cells by 5fC and 5caC more pronounced than those marked by 5hmC. We conclude that the hypermethylation signature in TKO-ESCs closely correlates with loci processed by TET/TDG, consistent with GADD45 proteins acting in locus-specific DNA demethylation.

We segmented hyper-DMRs and carried out transcription factor (TF) binding motif analysis using HOMER (Fig. 3A). The most prominent hit in all DMRs was Klf5, a possible reader of methylated DNA (Liu et al. 2014; Spruijt et al. 2013), which promotes the pluripotent ESC state and is required for trophectoderm development (Ema et al. 2008; Parisi et al. 2008). Other prominent hits were Ets-like TF binding elements (Ehf, Etv1, Fli1, Elk4/1). Interestingly, hyper-DMRs overlapping with enhancers harboring 5hmC were enriched for motifs of Zscan4, a key regulator of the 2C-like state.

Methylation regulated genes are downregulated in *Gadd45* TKO-ESCs

To identify genes differentially expressed upon *Gadd45* deficiency RNA-seq analysis was carried out under three culture conditions: normal serum culture, and two conditions inducing global demethylation, vitamin C and 2i treatment (Blaschke et al. 2013; Ficz et al. 2013; Leitch et al. 2013). A total of 135 genes were differentially expressed in *Gadd45* TKO-ESCs versus Co-ESCs during normal serum culture (FDR 10%). This number decreased sharply upon vitamin C- or 2i treatment (Fig. 3B), supporting that gene deregulation in TKO-ESCs is due, directly or indirectly, to DNA hypermethylation. Around 25%

of deregulated genes overlapped with hyper-DMR associated genes (Fig. S3C) further indicating that only a moderate fraction of GADD45-dependent genes are methylation-regulated. This is in line with generally modest correlation between gene expression and DNA methylation in ESCs (Karimi et al. 2011; Lu et al. 2014) Genes downregulated in *Gadd45* TKO-ESCs overlapped significantly with genes upregulated in *Dnmt1*$^{-/-}$/*Dnmt3a*$^{-/-}$/*Dnmt3b*$^{-/-}$ TKO-ESCs (Fig. 3C; Karimi et al. 2011). Consequently, localization of genes downregulated in *Gadd45* TKO-ESCs is enriched on the X-chromosome (Benjamini p= 2.2x10-6), as has been observed for upregulated genes in *Dnmt* TKO-ESCs (Fouse et al. 2008). These genes tend to be involved in germ cell regulation (Wang et al. 2001). Moreover, genes deregulated in *Gadd45* TKO-ESCs correlated with genes deregulated upon *Tet1* knockdown in ESCs (Fig. 3D; Huang et al. 2014).

Previous studies showed that DNA methylation and demethylation play a more important role during differentiation than during pluripotency (Dawlaty et al. 2014; Jackson et al. 2004; Lei et al. 1996; Okano et al. 1999; Sakaue et al. 2010). We therefore subjected TKO-ESCs to three differentiation protocols and analyzed transcriptome changes by RNA-seq. First, ESCs were differentiated for 8 days as embryoid bodies (EBs). Second, we differentiated ESCs for 6 days in serum-free monolayer culture. Third, ESCs were differentiated for 4 days as EBs and then treated for 4 more days with retinoic acid (RA). While the latter two protocols favor neuronal differentiation, the unguided EB culture allows differentiation into all three germ layers (Bibel et al. 2007; Ying et al. 2003). *Gadd45a, b* and *g* were all expressed at varying levels under these differentiation regimes (Fig. S3D).

The number of differentially expressed genes in TKO cells versus Co cells was 907 upon monolayer differentiation and 659 in EBs (FDR 10%, Fig. 3E), and thus gene deregulation in *Gadd45* TKO cells is indeed increased upon differentiation (compare Fig. 3B, E). Only 22 genes were differentially expressed in RA-treated EBs, suggesting that the particular neural lineages induced by RA are less sensitive to *Gadd45* deficiency. Only 48 genes were commonly deregulated in TKO cells between EB and monolayer differentiation (not shown), indicating that the function of the *Gadd45* genes is highly context dependent. Interestingly, genes differentially expressed in *Gadd45* TKO EBs were more than twice as likely to be marked by 5fC in their promoter regions compared to unaffected genes (Fig. S3E), supporting that GADD45-dependent genes are prone to undergo active DNA demethylation during EB differentiation.

GO term analysis in TKO EBs showed that downregulated genes were highly enriched for developmental terms such as system development, cell fate determination, cell migration, and axon guidance (Fig. S4A). Less pronounced GO term enrichment was found for genes upregulated in TKOs (Fig. S4B). For genes downregulated in monolayer differentiated TKO cells, GO term enrichment was also found for cell motility related terms (Fig. S4C), whereas genes upregulated were enriched for developmental and neuronal functions (Fig. S4D). Despite sharing similar GO ontologies, the genes affected in *Gadd45* TKO EBs were largely distinct from those affected in *Gadd45* TKO monolayer cells

(not shown). We conclude that in differentiating ESCs, GADD45 proteins regulate genes related to developmental, neuronal, and cell motility function, consistent with the involvment of GADD45 proteins in the regulation of neural development (Huang et al. 2010; Kaufmann and Niehrs 2011; Ma et al. 2009).

We validated selected genes commonly deregulated in *Gadd45* TKO-, *Dnmt* TKO- and *Tet1* knockdown ESCs by qPCR. Commonly deregulated genes did not significantly cluster in terms of gene ontology (GO) enrichment (not shown), but included various germline specific (e.g. *Rhox2a* and *Asz1*) and pluripotency related (e.g. *Pramel6* and *Pramel7*) genes (Fig. 3F and Fig. S5A).

To analyze the *Gadd45* gene redundancy in differential gene expression, we conducted rescue experiments. Transient combined overexpression of *Gadd45a*, *Gadd45b* and *Gadd45g* rescued downregulation of a panel of misregulated genes in *Gadd45* TKO-ESCs and further increased their expression levels in Co-ESCs (Fig. S5A). Not only combined, but also individual *Gadd45a*, *Gadd45b* or *Gadd45g* overexpression was effective in these rescue experiments, indicating that *Gadd45* genes in ESCs function redundantly (Fig. S5B-C). Genes downregulated in *Gadd45* TKO-ESCs were also induced by 5'-deoxyazacytidine treatment (Fig. 3G), further supporting that gene downregulation in TKO-ESCs involved DNA hypermethylation.

To test directly for DNA hypermethylation in TKO-ESCs, we analyzed the methylation status of regulatory elements in the vicinity of selected GADD45-dependent genes, which are shared with DNMT- and TET1-dependent genes. The majority of CpG dinucleotides in the *Rhox2a* promoter (Fig. 3H), the *Pramel6* promoter (Fig. S6A), and the *Gm364* promoter (Fig. S6B) were hypermethylated in *Gadd45* TKO-ESCs. In contrast, in two control gene promoters displaying high and low methylation levels, respectively (*Sry, Rerg*), no methylation changes were observed in TKO-ESCs (Fig. S6C-D).

We conclude that GADD45 proteins act redundantly to maintain normal expression and methylation levels of selected genes in ESCs, and that this list of genes overlaps with TET1 and DNMT target genes.

***Gadd45* TKO-ESCs show impaired 2C-like state**

We used the ESCAPE database (Xu et al. 2014), which integrates high-content data from embryonic stem cells, to conduct enrichment analysis of genes downregulated in undifferentiated *Gadd45* TKO-ESCs (Fig. 4A). The top hit returned was a set of genes reported upregulated upon *Gadd45a* overexpression (Nishiyama et al. 2009), corroborating the validity of the ESCAPE analysis. Among the other hits with similar significance were genes upregulated upon misexpression of *Zscan4* (zinc finger and SCAN domain-containing protein 4; (Nishiyama et al. 2009), a transcription factor whose recognition motif was enriched in hyper-DMRs (Fig. 3A). *Zscan4* is a marker and regulator of the 2-cell embryo (Falco et al. 2007). Consistent with the overlap between GADD45- and ZSCAN4-regulated genes, we found that genes upregulated in 2C-like cells tend to be downregulated in *Gadd45* TKO-ESCs, suggesting a requirement of *Gadd45* genes in regulating the 2C-like state (Fig. 4B). Indeed,

although only 97 genes in *Gadd45* TKO-ESCs were downregulated (> 0.5-fold, 10% FDR), these were enriched in genes specifically expressed in 2C-like cells (Fig. S7A; Macfarlan et al. 2012). In contrast, not a single gene up-regulated in *Gadd45* TKO-ESCs overlapped with the 2C gene set. 2C-specific genes were only modestly enriched in the vicinity of hyper-DMRs (Fig. 4C), which is expectable since the 2C-like cells only represent a small fraction of the ESC population. Of these 178 hyper-DMR- and 2C-like associated genes, six were also downregulated in the *Gadd45* TKO-ESCs (*Igfbp2, Inpp4b, Pramel6, Pramel7, Snhg11* and *Tmem92*). Other hypermethylated 2C genes may be downregulated only upon 2C cycling and hence escape detection.

To confirm these results, we monitored the expression of prominent 2C-associated genes in Co- and TKO-ESCs after combined overexpression of *Gadd45a*, *-b* and *-g* or GFP (Fig. 4D). Overexpression of Gadd45 genes not only rescued downregulation of the majority of tested 2C-associated genes in TKO-ESCs, but also increased expression levels even in Co-ESCs. In contrast, expression of retroviral elements unrelated to the 2C status (intracisternal A-particle, IAP) was unchanged.

To investigate the role of GADD45 in regulating the 2C-like state, we stably introduced a 2C-reporter, Zscan4c::eGFP (Zalzman et al. 2010), in Co- and TKO-ESCs (Fig. 4E). Flow cytometry revealed a significant reduction of Zscan4$^+$ cells in TKO ESCs (Fig. 4F). Combined overexpression of the *Gadd45* genes rescued the reduction of Zscan4$^+$ cells in TKO-ESCs to control levels (Fig. 4G). Likewise, individual overexpression of *Gadd45a, -b* or *-g* increased the percentage of Zscan4$^+$ cells, as well as of cells harboring the mERVL 2C-reporter (Macfarlan et al. 2012), suggesting redundant function of the GADD45 proteins in 2C regulation (Fig. 4H and Fig S7B). Interestingly, long term 2i treatment (seven passages), which induces hypomethylation (Ficz et al. 2013; Leitch et al. 2013), abolished the observed difference in 2C-like population (Fig. S7C). In contrast, overexpression of *Tet1*, *Tet2* or *Tdg* did not affect the frequency of Zscan4$^+$ cells in TKO- nor in Co-ESCs (Fig. S7D-F).

Zscan4$^+$ sorted cells show a global demethylation relative to unsorted ESCs (Eckersley-Maslin et al. 2016). We confirmed by mass-spectrometry that Co Zscan4$^+$ cells show reduced 5mC levels (compare Fig. 4I Co 'unsorted' to 'Zscan4$^+$') while Zscan4$^+$ TKO cells showed hypermethylation. No difference was found for 5hmC (Fig. 4J).

Gene expression analysis of TKO-ESCs (Fig. 5A) revealed reduced RNA levels of 2C-specific genes not only in unsorted- but also in Zscan4$^+$ TKO-ESCs (e.g. *Zscan4, Sp110, Tcstv1*). This reduction was not due to deficient maintenance of 2C-like state but reflected deficient entry, since cycling-out of the Zscan4+ state occurred with the same kinetics in Co- and TKO-ESCs (Fig. 5B).

Among the downregulated 2C genes was *Dux*, a key regulator of ESCs cycling into the 2C-like state (Hendrickson et al. 2017; De Iaco et al. 2017). This suggested that GADD45 may act upstream of DUX to promote the 2C-like state. Concordantly, overexpression of *Dux* using a doxycycline-inducible plasmid restored the reduced number of 2C-like cells in TKO-ESCs (Fig. 5C). The expression level of

repressors of the 2C-like state (*Trim28, Lsd1, G9a, Chaf1a*) was not altered in unsorted, Zscan4^{+} or Zscan4^{-} TKO-ESCs (Fig. S7G).

Finally, we explored the consequences of 2C misregulation in ESC transdifferentiation. While ESCs normally do not give rise to trophoblast, they do sporadically transdifferentiate into this lineage, which can be further enhanced by BMP4 treatment (Beddington and Robertson 1989; Hayashi et al. 2010). Trophoblast transdifferentiation involves 2C cycling, since 2C-like cells are not lineage restricted (Macfarlan et al. 2012) and cycling through the 2C-state is critical to restore the full developmental capacity in ESCs (Amano et al. 2013). We treated ESCs with BMP4, which induced 2C-associated genes (Fig. 5D) as well as the trophoblast stem cells markers *Cdx2* and *Elf5* (Fig. 5E) supporting that ESCs employ the 2C-like state during transdifferentiation. Induction of both 2C- and trophoblast markers was greatly reduced in TKO-ESCs. Moreover, expression of placental markers induced by BMP4 (e.g. *Serpin*, *Psg* and *Prl* gene families) was almost abolished in TKOs (Fig. 5F). The reduced transdifferentiation potential of *Gadd45* TKO-ESCs is consistent with impaired 2C-like cycling. Interestingly, *Dnmt1* deficiency yields the opposite phenotype to *Gadd45* deficiency, i.e. activation of trophoblast lineage markers (Cambuli et al. 2014). Our data are therefore in line with the notion that DNA methylation is a barrier to ESC- to trophoblast transdifferentiation (Ng et al. 2008).

(*Gadd45a/Gadd45b*)$^{-/-}$ mice are sublethal and show partially impaired ZGA gene expression

Interrogating a database of early mouse transcriptome (Park et al. 2015) revealed that *Gadd45a* and *Gadd45b* belong to a "2-Cell Transient" cluster, showing a peak of expression specifically in 2-cell embryos (Fig. 6A). *Gadd45g* belongs to the "Major ZGA cluster" but shows low expression during cleavage stages. This raised the possibility that the impairment of 2C cycling in ESCs, in fact, reflects a role for GADD45 in the embryonic 2-cell stage, coinciding with the major phase of zygotic genome activation (Eckersley-Maslin et al. 2018). Single *Gadd45a, -b*, or *-g* mutants are viable and fertile (Hollander et al. 1999; Lu et al. 2001; Lu et al. 2004) but our results in ESCs indicate that they may compensate for each other in 2C-state regulation. As generation of triple mutants is challenging, we generated *Gadd45a,b* double mutant (DKO) mice by intercrossing double heterozygous animals. All nine possible genotypes were obtained at expected Mendelian ratios, except for the homozygous *Gadd45a,b* double mutants, whose frequency was 50% reduced (Fig. 6B, arrow). Surviving DKO mice showed normal body size, but displayed phenotypic abnormalities characteristic of neural tube closure defects (NTDs) such as curly tail and spina bifida. Litter size of DKO intercrossing was 50% reduced compared to double heterozygous crosses (Fig. 6C). At embryonic stage E13.5, 50% (n=24) of DKO embryos resulting from double heterozygous breeding and 80% (n=37) from homozygous DKO breeding, showed, beyond curly tail and spina bifida, also defects like exencephaly and cranial hemorrhage (Fig. 6D-E, S7H). Sublethality, exencephaly, and cranial hemorrhaging are also observed in *Tet1, 2* DKO mice (Dawlaty et al. 2013). The increased phenotypical abnormalities of DKO embryos

resulting from breeding homozygous DKOs compared to double heterozygous mice hints at a requirement of GADD45 in the germ line, as is the case for TET1 (Yamaguchi et al. 2013).

To investigate the impact of *Gadd45a,b* deficiency on gene expression we performed transcriptome analysis of 2-cell stage embryos from *Gadd45a,b* DKO crosses. Transposable elements (TE) were hardly affected in DKO embryos. There were no downregulated- and few upregulated TE transcripts, including LINE1 elements (L1MC, Lx2B; Fig. 6F). Also the number of deregulated non-repetitive genes in DKOs was limited (n= 104, 10% FDR, Fig. 6G), in accord with subviability of DKO mice. However, misregulated genes showed a clear signature of impaired 2-cell stage entry: First, genes downregulated in DKOs corresponded to genes upregulated in 2-cell stage embryos compared to oocytes (Macfarlan et al. 2012), while genes upregulated in DKOs overlapped with genes upregulated in oocytes (Fig. 6H-I). Second, ZGA genes associated specifically with the 2C-like state (Macfarlan et al. 2012) overlapped with none of the up- but seven downregulated genes in DKOs (Fig. 6J). The enrichment of maternal- and depletion of zygotic transcripts in DKO embryos indicates that GADD45α,β promote the embryonic 2-cell stage. Hence, we monitored the development of wild type (WT) and DKO preimplantation embryos in vitro. No differences were observed for development until the 2-cell stage between WT and DKO embryos (data not shown). However, only ~40% of the DKO embryos reached 8-cell stage versus ~80% of the wild type embryos (Fig. 6K). The affected DKO embryos remained at 2- or 4-cell stage, or died. Incomplete penetrance of preimplantation defects are also observed in other mutant mice (Narducci et al. 2002). However, the impaired in vitro pre-implantation development supports our findings in mouse embryonic stem cells showing that GADD45 proteins are involved in 2-cell stage regulation.

Discussion

GADD45α,β,γ are adaptors for TET/TDG mediated DNA demethylation but in which physiological processes and at which genomic loci they mediate demethylation remains poorly understood. In addition, they show overlapping expression and hence may act functionally redundant, complicating their analysis. Here we present analyses of triple-knockout ESCs and *Gadd45a,b* DKO mice showing that i) GADD45α,β,γ are not required to maintain pluripotency and self-renewal in ESCs. Yet, ii) GADD45 proteins are required for locus-specific demethylation of ~7,000 sites, notably on enhancers and at sites harboring oxidized 5mC; iii) *Gadd45*-mutant ESCs display methylation-related gene misexpression; iv) GADD45 proteins promote the 2C-like ESC state and 2-cell embryo stage, regulating a subset of ZGA-specific genes.

GADD45 proteins are not required for pluripotency but for differentiation of ESCs

We find that *Gadd45* TKO-ESCs remain pluripotent and self-renew, as has been found for other DNA demethylation deficient ESCs, such as *Tet1,2,3* and *Tdg* knockout ESCs (Cortázar et al. 2011; Dawlaty et al. 2014). A previous study (Li et al. 2015) reported that also *Gadd45a,b* double knockout ESCs remain pluripotent. In agreement with this study, overall oxidized cytosine levels were mostly

unchanged in *Gadd45* TKO-ESCs. This is expected, because GADD45α functions in locus-specific, rather than global demethylation (Arab et al. 2014; Arab et al. 2019; Schäfer et al. 2013; Schmitz et al. 2009). Our rescue experiments indicate that all three GADD45 proteins can compensate for the loss of all three genes, supporting their functional redundancy.

While GADD45 proteins were dispensable for overall ESC maintenance, a subset of genes was downregulated and hypermethylated in TKO-ESCs. Moreover, DNA-hypomethylating 2i, vitamin C, and 5 azadeoxycytidine treatment rescued this downregulation, suggesting that it was the direct or indirect consequence of DNA hypermethylation. Similarly, previous reduced representation bisulfite sequencing identified 68 hypermethylated but no hypomethylated loci in *Gadd45a,b* DKO-ESCs (Li et al. 2015). Most of these sites overlapped with 5hmC- and 5fC-enriched regions, corroborating a role in DNA demethylation. Taken together with the overlap between the genes misregulated in *Gadd45* and *Tet* mutants, this supports the conclusion that these two protein families cooperate in enzymatic DNA demethylation (Arab et al. 2014; Arab et al. 2019; Kienhöfer et al. 2015; Li et al. 2015).

Even though GADD45 proteins act locally and DNMTs globally, there was a significant overlap between genes misregulated in *Gadd45* TKO-ESCs and *Dnmt* TKO-ESCs. However, despite millions of genomic sites that become unmethylated in *Dnmt* TKO-ESCs, including at least 6,100 promoters, only a few hundred genes are actually de-repressed (Karimi et al. 2011). These 'hotspot' genes misregulated in *Dnmt* TKO-ESCs are enriched for a role in germ cell development and localize on the X-chromosome. Likewise, genes in the *MageA* and *Rhox* clusters that are prominently reactivated in the DNMT TKO-ESCs, are downregulated in *Gadd45* TKO-ESCs. Thus, even though DNMTs act globally, respective gene expression changes in ESCs only occur on hotspots, which are prone to react to DNA methylation changes, such as germ cell-specific genes. Indeed, TET1 is a prominent regulator of gene expression in germ cells (Hill et al. 2018; Yamaguchi et al. 2012), suggesting that ESCs recapitulate aspects of germ cell gene regulation via DNMT-TET mediated methylation-demethylation.

In contrast to a limited role of GADD45 proteins during ESC pluripotency, gene deregulation in TKOs was greatly increased upon EB- and monolayer differentiation. Neuronal genes were particularly affected, corroborating a role for *Gadd45* genes during neural differentiation (Huang et al. 2010; Kaufmann and Niehrs 2011; Ma et al. 2009). Indeed, while *Gadd45a,b* DKO mice were viable, they were sublethal, with embryos showing gross abnormalities including defects in neural tube closure and brain hemorrhage. This phenocopies *Tet1,2* DKOs, which are also sublethal, with affected embryos showing exencephaly and cranial hemorrhaging, and where the defects were attributed to reduced 5hmC, increased 5mC, and aberrant imprinting (Dawlaty et al. 2013).

TET/TDG processed sites are main targets of GADD45 mediated DNA demethylation

Our methylome analysis of TKO-ESCs revealed ~7,000 hyper-DMRs, the greatest number of locus-specific methylation changes so far reported to be associated with GADD45 function. Since only 34

hypo-DMRs were detected (arguing against generalized methylation-misregulation), and since the most significant association of the hyper-DMRs was with oxidized 5mC, we conclude that the hyper-DMRs arose as the result of impaired enzymatic DNA demethylation. This is consistent with direct interaction of GADD45α with both TET1 and TDG (Barreto et al. 2007; Cortellino et al. 2011; Kienhöfer et al. 2015; Li et al. 2015). Moreover, GADD45-dependent hyper-DMRs are enriched at TET-dependent hyper-DMRs (Lu et al. 2014) with the caveat that the number of called DMRs is not directly comparable, due to differences in ESC lines and in DMR-calling algorithms used. Consistently, we reach similar conclusions for *Gadd45* TKO-ESCs as have been reported for *Tet1,2,3* TKO-ESCs (Lu et al. 2014). First, the most important concordance of hyper-DMRs was with sites marked by oxidized 5mC derivatives. Enrichment of hyper-DMRs at TDG 5fC- and 5caC target sites concurs with the conclusion of a dual role of GADD45α, to stimulate TET1 chemical processivity for iterative 5mC oxidation, and to enhance TDG processing of 5fC/5caC (Kienhöfer et al. 2015; Li et al. 2015). Second, enhancers are the most enriched genomic target of both GADD45 and TET-mediated DNA demethylation. Enhancers in ESCs tend to be hypomethylated (Kieffer-Kwon et al. 2013; Stadler et al. 2011; Ziller et al. 2013) and thus GADD45 plays a significant role in this phenomenon. Enhancers are also the main target of GADD45α mediated DNA demethylation in mouse embryonic fibroblasts (Schäfer et al. 2018).

Both in *Tet*- and *Gadd45*-TKO-ESCs, enhancer hypermethylation affected expression of relatively few genes, consistent with the minor role of 5mC in gene repression of early embryos and ESCs (Bogdanovic et al. 2011; Fouse et al. 2008). The regulatory role of DNA demethylation for *Gadd45* may manifest more upon later cell differentiation, as e.g. mutation of *Dnmt* and *Tet* in ESCs leads to differentiation defects (Dawlaty et al. 2014; Jackson et al. 2004) and to lethality in whole embryos (Dai et al. 2016; Okano et al. 1999). Hence, the dynamics of DNA methylation in ESCs may reflect a role as molecular memory for subsequent enhancer regulation in differentiating cells, rather than in regulation of acute gene expression (Kim et al. 2018).

GADD45 proteins promote a 2C-like state and ZGA-specific gene expression

A main finding of this study is that *Gadd45* TKO-ESCs displayed impaired cycling into the 2C-like state and *Gadd45a,b* DKOs showed partial gene misregulation at the 2-cell stage, as well as embryonic sublethality with developmental stage arrest. This suggests that 2C-cycling defects in TKO-ESCs mirror a role of GADD45 during 2-cell stage development and ZGA. The 2C-like state of ESCs recapitulates key aspects of the 2-cell stage mouse embryo, both phenotypically and molecularly (reviewed in (Eckersley-Maslin et al. 2018; Ishiuchi and Torres-Padilla 2013)). Concordantly, *Gadd45* expression peaks in 2-cell stage mouse embryos and while *Gadd45a,b* DKOs display only moderate gene misregulation, our results may underestimate their role, as *Gadd45g* could partially compensate for *Gadd45a,b* deficiency in DKOs (Bogdanovic et al. 2011).

Entry into the 2C-like state is accompanied by genome-wide DNA demethylation (Dan et al. 2017; Eckersley-Maslin et al. 2016) However, ESCs cultured in serum are in a distinct, potentially artefactual

epigenetic state: The de novo DNA methylation machinery, only transiently active in embryos during implantation, is instead continuously operating in ESCs (Brandeis et al. 1994; Lienert et al. 2011). Hence, downregulation of 2C genes in TKO-ESCs might rather reflect a role of GADD45 proteins during early 2-cell stage in protecting against de novo methylation. This is because the genome of 2-cell embryos is already hypomethylated due to global demethylation occurring in zygotes (reviewed in (Eckersley-Maslin et al. 2018; Lee et al. 2014; Messerschmidt et al. 2014)) and requires TET3 to protect against de novo DNA methylation (Amouroux et al. 2016). It is this protection against de novo DNA methylation that may require GADD45.

In apparent contradiction to these findings, transition to the 2C-like state in Tet TKO-ESCs is actually enhanced, as TET proteins repress type III ERVs and 2C-specific genes (Lu et al. 2014). Moreover, *Tet* TKO mouse embryos develop past implantation stage to gastrulae, albeit with altered expression of a few hundred genes at the blastocyst stage (Dai et al. 2016). If GADD45 act via TETs to affect 2C-like state and 2-cell embryo development, how can this discrepancy be explained? First, *Tet* TKO mouse embryos were generated by crossing mice with TET-deficient germ cells (Dai et al. 2016), and hence with zygotes that had adapted to a state completely devoid of TET enzymatic activity. Furthermore, close inspection of *Tet1,3* DKO mice showed unexpectedly variable expression of ~150 genes in 8-cell stage embryos, among them the *Zscan4* cluster (Kang et al. 2015). In fact, acute *Tet1,2,3* knockdown in oocytes leads to developmental arrest at 2-cell stage with severe ZGA gene misregulation (M. Wossidlo, pers. commun.). Second, TETs have a dual function, involving both their catalytic- as well as non-catalytic, gene repressive function, the latter being the dominant gene-regulatory mode of TET1 in ESCs (Williams et al. 2011). The repressive role involves recruitment of KAP1/TRIM28 by TETs, a negative 2C-like state regulator (Lu et al. 2014). Consistent with a dual role, *Tet*-deficient ESCs and epiblast-derived stem cells can display either up- or downregulation of *Zscan4* cluster expression, depending on culture conditions (Khoueiry et al. 2017; Yang et al. 2016). Similarly, for LINE1 regulation TETs play a dual role in ESCs, activating and silencing expression via LINE1 promoter demethylation and via recruiting the SIN3A repressor, respectively (La Rica et al. 2016).

In analogy, TET enzymes may play a dual role in the regulation of the key developmental transition at the 2-cell stage: First, a repressive one, whereby possibly TET represses LINE1 elements, whose transcripts silence *Dux* expression and thereby the 2-cell state (Percharde et al. 2018). Second, a promoting one, as supported by our study, whereby GADD45 targets TETs to promote the 2-cell stage via demethylation and protection against de-novo methylation.

Material & Methods

Statistics

Unless otherwise indicated, statistical significance was tested with unpaired, two-tailed Student's t-test with three biological replicates. Unless otherwise indicated, data are presented as mean ± standard deviation from three independent clones. *p<0.05, **p<0.005, ***p<0.0005.

Gene ontology (GO) enrichment analysis for the ontology "biological process" was carried out using GOrilla (Eden et al. 2009), using the parameter "Two unranked lists of genes" and the whole mouse transcriptome as a background list.

Commonly deregulated genes were identified using publicly available datasets. Overlaps were identified using ENSEMBL or RefSeq IDs if available, and gene symbols if not. Statistical significance of overlaps was determined using hypergeometric test. Only genes detectable by both platforms employed were considered. For scatterplots, only genes which were significantly deregulated in both respective studies are shown.

Quantitative PCR

Total RNA was isolated using a Qiagen RNeasy mini kit with on-column DNase digest (Qiagen). First strand cDNA was generated using SuperScriptII reverse transcriptase (Invitrogen). Real-time PCR was performed in technical duplicates using Roche LightCycler480 probes master and primers in combination with predesigned monocolor hydrolysis probes of the Roche Universal Probe Library (UPL). For quantification, Roche LC480 quantification software module was used. Expression levels were normalized to Gapdh or Tbp expression first, followed by normalization to control conditions specific to the individual experiment, as indicated.

Quantitative mass spectrometry of DNA modifications

Genomic DNA sample preparation and quantification of 5mC and its oxidative derivatives was carried out as described before (Schomacher et al. 2016).

ESC culture, treatment and differentiation

ESCs were cultured on gelatinized cell culture vessels in LIF-conditioned DMEM supplemented with 15% ESC-grade FBS, 2mM L-Glutamine, 50U/ml Penicillin/Streptomycin, 1xNEAA, 1mM sodium pyruvate, 100µM β-mercaptoethanol, at 37°C, 5% CO2 and 21% O2, with daily medium changes. Cells were passaged every 2 days once reaching approximately 60% confluency with a ratio of 1:8. For ESC culture on mouse embryonic fibroblast (MEF) feeder, ESCs were plated in a 12-well preplated with CF-1 MEF feeder (ATCC). For growth curve analysis cells were counted in duplicates using a TC10 automated cell counter (Biorad).

For EB differentiation, 3.5x106 mESCs were plated on non-adherent 10cm bacterial dishes (Greiner) in 15ml CA medium (Bibel et al. 2007). CA medium was changed every other day. For retinoic acid induced EB differentiation, CA medium was supplemented with retinoic acid (final concentration 5µM) at days 4 and 6 of differentiation. EBs were harvested after 8 days of differentiation.

For monolayer differentiation, approximately 5.000 ESCs per cm^2 were plated on gelatinized cell culture plates in regular ESC medium on the evening before differentiation. On the next morning, cells were washed twice with PBS and medium was changed to N2B27 medium (50% Advanced DMEM/F12, 50% Neurobasal, 2mM L-Glutamine, 50U/ml Penicillin/Streptomycin, 50µg/ml BSA, 0.5x N2-supplement (Invitrogen), 0.5x B27-supplement (Invitrogen). Medium was changed on days 3 and 5 of differentiation. Cells were harvested after 6 days of differentiation.

For 2i treatment, 230.000 ESCs were plated per 6-well in regular ESC medium. On the next day, cells were washed twice with PBS and cultivated up to 72h in 2i medium (N2B27 medium, 1µM PD0325901, 3µM CHIR99021, 4% LIF-supernatant).

For vitamin C treatment, 230.000 ESCs per well were plated in 6-well-plates in regular ESC medium. The next day, vitamin C (L-ascorbic acid 2-phosphate sesqui-magnesium salt) was added to a final concentration of 100µg/ml. Cells were incubated in vitamin C containing medium for up to 72h and passaged once during that time.

For 5-azacytidine treatment, ESCs were incubated in 10µM 5-deoxyazacytidine in ESC medium for 48h.

For transdifferentiation, ESCs were resuspended in DMEM supplemented with 15% Knockout Serum Replacement (Invitrogen), 2mM L-Glutamine, 50U/ml Penicillin/Streptomycin, 1xNEAA, 1mM Na-pyruvate, 100µM β-mercaptethanol medium and 10ng/ml recombinant hBMP-4 (R&D Systems) after passaging and plated at a density of 104 cells per cm^2 on gelatinized cell culture vessels. Medium was changed on days 2, 4, 6 and 7 of transdifferentiation. Cells were harvested after 8 days.

For transient overexpression, preplated ESCs were transfected with Lipofectamine 2000 (Thermo-Fischer) according to the manufacturer´s recommendations. Medium was renewed on the next day and cells were harvested 48h post transfection.

CRISPR/Cas9 mediated knockout and introduction of stable expressing Zscan4c::eGFP

$1.45x10^6$ ESCs were seeded on 10cm dishes each and transfected on the next day with either i) 13.5µg empty px330 vector and 1.5µg pPuro, ii) 2.25µg of each of the six guide RNA containing plasmids (see Supplemental Material and Methods) and 1.5µg pPuro using 45µl Lipofectamine 2000 (Thermo-Fischer) in a volume of 1.5ml OptiMEM, but otherwise as decribed above. ESCs were passaged from 1x10cm to 2x15cm dishes the following day. Cells were selected with 2µg/ml puromycin from 48h post-transfection and kept under selection until the day of freezing. Forming ESC colonies were transferred

to gelatinized 48-well plates 6 days after passaging. On the next day, colonies were washed once with 500µl PBS, dissociated with 100µl 0.25% trypsin for 90s at 37°C, quenched with 800µl of ESC medium and plated on a new 48-well plate. Three days later, cells were passaged and partially used for genotyping PCR (Primers see Supplemental Material and Methods). For transfection of Zscan4c::eGFP $1x10^6$ ESCs were seeded on gelatinized 10cm dishes and transfected 4 hours after seeding using Lipofectamine 2000 (Thermo-Fischer). Cells were selected with 5µg/ml blasticidin 48h post-transfection for eleven days, expanded and frozen for further analysis. For transfection of 2C::tdTomato reporter (Addgene #40281) 1x106 ESCs were seeded on gelatinized 10cm dishes and transfected 4 hours after seeding using Lipofectamine 2000 (Thermo-Fischer). Cells were selected with 150µg/ml hygromycin 48h post-transfection for seven days, expanded and frozen for further analysis.

Plasmids

pZscan4-Emerald was a kind gift from M. Ku (Zalzman et al. 2010). pCW57.1_Luciferase and pCW57.1_mDuxCA-3xHA were a kind gift from B.Cairns (Hendrickson et al. 2017). The expression constructs in this study were pCS2+FLAG_hTDG (Schomacher et al. 2016), pCS2+-GFP (Barreto et al. 2007), pCS2+myc-MmGadd45a (M. Gierl, unpubl.), pCS2+myc-MmGadd45b(M. Gierl, unpubl.), pCS2+myc-MmGadd45g (Gierl et al. 2012), pRKW2-mTet2 (Ko et al. 2010). The catalytic domain of mouse Tet1 was inserted into pCS2+ as N-terminal HA-tag expression construct (A. Ernst, unpubl.).

Teratoma assays

ESCs were sent cryo-conserved to EPO Berlin GmbH, where they were thawed and passaged twice before transplantation. Cells were resuspended in PBS, mixed with Matrigel and transplanted into the flanks of three NSG mice per mESC clone. Tumor weight and size were measured twice per week. Animals injected with the same ESC clone were sacrificed once the average tumor size reach approximately 1.0 cm^3. Tumors were excised, weighed and cut. 1/3rd was shock frozen and 2/3rds were fixed in formalin. Formalin fixed tumors were paraffin embedded, sectioned, and stained with hematoxylin and eosin.

Flow cytometry analysis and fluorescence activated cell sorting

Cells were detached using 0.25% trypsin and resuspended in PBS containing 2.5% ESC grade and 5mM EDTA pH 8. Suspended cells were then analysed by the BD LSRFortessaSORP flow cytometry system using DiVa software. For downstream analysis on sorted ESCs, cells were sorted according to the fluorescence intensity of eGFP into PBS using BD FACSAria III SORP with a 85 µm nozzle. Data analysis was performed with FlowJo software (version 10.5.3).

Whole-genome bisulfite sequencing

Genomic DNA from Control and Gadd45 TKO-ESCs was purified using a QIAamp DNA Mini kit (Qiagen) according to the manufacturer´s recommendations. An additional RNAseA treatment

(10mg/ml, Qiagen) was done after cell lysis. WGBS library preparation was carried out using the TruSeq PCR-Free Library Prep Kit (LT) and the Epitect Kit (Qiagen) for bisulfite conversion. Sequencing was performed on Illumina HiSeq X in paired-end mode (PE150).

RNA sequencing

ESC samples

Total RNA was isolated using the RNeasy mini kit (Qiagen) with on-column DNase digest. NGS library preparation was performed using Illumina's TruSeq Stranded mRNA HT Sample Prep Kit with dual-indexing following the standard protocol (Illumina Part # 15031047 Rev. D). Libraries were profiled with a DNA 1000 chip on an Agilent 2100 Bioanalyzer and quantified using the Qubit dsDNA HS Assay Kit on a Qubit 2.0 Fluorometer (Life Technologies). All 36 samples were pooled in equimolar ratio and sequenced on 8 HiSeq 2000 lanes for 35 cycles plus additional 16 cycles for the i7 and i5 index reads.

Isolation and culture of preimplantation embryos

For RNA-sequencing, 2-cell stage embryos were collected from 3-week old wild type (n=4) and (*Gadd45a/Gadd45b*)$^{-/-}$ (n=4) superovulated females 20 h after the appearance of the vaginal plug. 2-cell stage embryos coming from the same litter were pooled and considered as one biological sample. Embryos were collected in M2 medium (Sigma) supplemented with 0.3 mg/ml hyaluronidase (Sigma), washed twice with PBS and directly transferred into lysis buffer (Smart Seq v4 Ultra Low Input RNA Kit for Sequencing, Takara). Sample preparation was done according to the manufacturer's recommendations. cDNA was amplified using 11 PCR cycles. NGS library preparation was performed using NuGEN´s Ovation Ultralow System V2 1-96 (2014). Libraries were prepared with a starting amount of 2,36ng of fragmented cDNA and were amplified in 11 PCR cycles. Libraries were profiled in a High Sensitivity DNA on a 2100 Bioanalyzer (Agilent technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2.0 Fluorometer (Life technologies). All 8 samples were pooled in equimolar ratio and sequenced on one NextSeq 500 Highoutput FC, SR for 85 cycles plus 7 cycles for the index read.

For the in vitro development assay, mouse preimplantation embryos were collected as described above and then cultured in EmbryoMax® Human Tubal Fluid medium (Millipore) at 37 °C, 5% CO_2 for 48 h. The experiment was perfomed in biological triplicates using three independent breeding for wild type and (*Gadd45a/ Gadd45b*) $^{-/-}$ animals.

Animal experiments

Gadd45a and *Gadd45b* knock out mice were kindly provided by M. C. Hollander (Gupta et al. 2005; Hollander et al. 1999). Both strains were backcrossed several generations into the C57BL/6N background and interbred to obtain *Gadd45a* $^{+/-}$/ *Gadd45b*$^{+/-}$ (DHet) mice, which were further intercrossed to generate wild type, *Gadd45a* $^{+/-}$/ *Gadd45b*$^{+/-}$ and *Gadd45a*$^{-/-}$/ *Gadd45b* $^{-/-}$ (DKO) animals

from homogeneous genetic background. Mice were housed under 12:12 light/dark cycles and provided with ad libitum food and water, in accordance with national and European guidelines. For embryo isolation at stage E13.5, timed matings were set up between DHet mice and DKO animals. All procedures were performed with the approval of the ethical committees on animal care and use of the federal states of Rheinland-Pfalz, Germany.

Data availability

All NGS data have been deposited in the NCBI's Gene Expression Omnibus (GEO) under superseries accession number GSE127720.

Acknowledgements

We thank C. Christopoulou for assistance with the rescue experiments. We thank M. C. Hollander for *Gadd45a* and *Gadd45b* mutant mice. B. Cairns, M. Ku, A. Rao and H. Richly kindly provided reagents. We thank C. Scholz for technical support. Contributions by the IMB Core Facilities Flow Cytometry, Genomics, Proteomics, Microscopy and DKFZ High Throughput Sequencing Unit are gratefully acknowledged. M.U.M was supported by Natural Sciences and Engineering Research Council of Canada Postdoctoral Fellowship (NSERC-PDF). This work was supported by an ERC advanced grant ("Demethylase").

Author contributions

K.M.S. conceived and conducted most 2C related experiments. M.L. generated and characterized the TKO ESCs. V.V. carried out all knock-out mouse analyses. T.A., M.M. and E.K. carried out bioinformatics analyses. M.U.M. performed LC-MS/MS measurements. All authors analyzed and discussed the data. C.N. conceived and coordinated the study and wrote the paper with contribution from K.M.S and M.L.

References

Amano T, Hirata T, Falco G, Monti M, Sharova LV, Amano M, Sheer S, Hoang HG, Piao Y, Stagg CA, et al. 2013. Zscan4 restores the developmental potency of embryonic stem cells, Nat Commun 4: 1966.

Amouroux R, Nashun B, Shirane K, Nakagawa S, Hill PW, D'Souza Z, Nakayama M, Matsuda M, Turp A, Ndjetehe E, et al. 2016. De novo DNA methylation drives 5hmC accumulation in mouse zygotes, Nat Cell Biol 18: 225–233.

Arab K, Karaulanov E, Musheev M, Trnka P, Schäfer A, Grummt I, Niehrs C. 2019. GADD45A binds R-loops and recruits TET1 to CpG island promoters, Nat Genet.

Arab K, Park YJ, Lindroth AM, Schäfer A, Oakes C, Weichenhan D, Lukanova A, Lundin E, Risch A, Meister M, et al. 2014. Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A, Mol Cell 55: 604–614.

Barreto G, Schäfer A, Marhold J, Stach D, Swaminathan SK, Handa V, Döderlein G, Maltry N, Wu W, Lyko F, et al. 2007. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation, Nature 445: 671–675.

Beddington RS, Robertson EJ. 1989. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo, Development 105: 733–737.

Bibel M, Richter J, Lacroix E, Barde Y-A. 2007. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells, Nat Protoc 2: 1034–1043.

Blaschke K, Ebata KT, Karimi MM, Zepeda-Martínez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, et al. 2013. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells, Nature 500: 222–226.

Bogdanovic O, Long SW, van Heeringen SJ, Brinkman AB, Gómez-Skarmeta JL, Stunnenberg HG, Jones PL, Veenstra GJC. 2011. Temporal uncoupling of the DNA methylome and transcriptional repression during embryogenesis, Genome Res 21: 1313–1327.

Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, Balasubramanian S. 2012. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution, Science 336: 934–937.

Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, Temper V, Razin A, Cedar H. 1994. Sp1 elements protect a CpG island from de novo methylation, Nature 371: 435–438.

Cambuli F, Murray A, Dean W, Dudzinska D, Krueger F, Andrews S, Senner CE, Cook SJ, Hemberger M. 2014. Epigenetic memory of the first cell fate decision prevents complete ES cell reprogramming into trophoblast, Nat Commun 5: 5538.

Choi YJ, Lin C-P, Risso D, Chen S, Kim TA, Tan MH, Li JB, Wu Y, Chen C, Xuan Z, et al. 2017. Deficiency of microRNA miR-34a expands cell fate potential in pluripotent stem cells, Science 355.

Cortázar D, Kunz C, Selfridge J, Lettieri T, Saito Y, MacDougall E, Wirz A, Schuermann D, Jacobs AL, Siegrist F, et al. 2011. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability, Nature 470: 419–423.

Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, et al. 2011. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair, Cell 146: 67–79.

Dai H-Q, Wang B-A, Yang L, Chen J-J, Zhu G-C, Sun M-L, Ge H, Wang R, Chapman DL, Tang F, et al. 2016. TET-mediated DNA demethylation controls gastrulation by regulating Lefty-Nodal signalling, Nature 538: 528–532.

Dan J, Li M, Yang J, Li J, Okuka M, Ye X, Liu L. 2013. Roles for Tbx3 in regulation of two-cell state and telomere elongation in mouse ES cells, Sci Rep 3: 3492.

Dan J, Rousseau P, Hardikar S, Veland N, Wong J, Autexier C, Chen T. 2017. Zscan4 Inhibits Maintenance DNA Methylation to Facilitate Telomere Elongation in Mouse Embryonic Stem Cells, Cell Rep 20: 1936–1949.

Dawlaty MM, Breiling A, Le T, Barrasa MI, Raddatz G, Gao Q, Powell BE, Cheng AW, Faull KF, Lyko F, et al. 2014. Loss of Tet enzymes compromises proper differentiation of embryonic stem cells, Dev Cell 29: 102–111.

Dawlaty MM, Breiling A, Le T, Raddatz G, Barrasa MI, Cheng AW, Gao Q, Powell BE, Li Z, Xu M, et al. 2013. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development, Dev Cell 24: 310–323.

Dawlaty MM, Ganz K, Powell BE, Hu Y-C, Markoulaki S, Cheng AW, Gao Q, Kim J, Choi S-W, Page DC, et al. 2011. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development, Cell Stem Cell 9: 166–175.

Deaton AM, Bird A. 2011. CpG islands and the regulation of transcription, Genes Dev 25: 1010–1022.

Eckersley-Maslin M, Alda-Catalinas C, Blotenburg M, Kreibich E, Krueger C, Reik W. 2019. Dppa2 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program, Genes Dev.

Eckersley-Maslin MA, Alda-Catalinas C, Reik W. 2018. Dynamics of the epigenetic landscape during the maternal-to-zygotic transition, Nat Rev Mol Cell Biol.

Eckersley-Maslin MA, Svensson V, Krueger C, Stubbs TM, Giehr P, Krueger F, Miragaia RJ, Kyriakopoulos C, Berrens RV, Milagre I, et al. 2016. MERVL/Zscan4 Network Activation Results in Transient Genome-wide DNA Demethylation of mESCs, Cell Rep 17: 179–192.

Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. 2009. GOrilla. A tool for discovery and visualization of enriched GO terms in ranked gene lists, BMC Bioinformatics 10: 48.

Ema M, Mori D, Niwa H, Hasegawa Y, Yamanaka Y, Hitoshi S, Mimura J, Kawabe Y-i, Hosoya T, Morita M, et al. 2008. Krüppel-like factor 5 is essential for blastocyst development and the normal self-renewal of mouse ESCs, Cell Stem Cell 3: 555–567.

Falco G, Lee S-L, Stanghellini I, Bassey UC, Hamatani T, Ko MSH. 2007. Zscan4. A novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells, Dev Biol 307: 539–550.

Ficz G, Hore TA, Santos F, Lee HJ, Dean W, Arand J, Krueger F, Oxley D, Paul Y-L, Walter J, et al. 2013. FGF Signaling Inhibition in ESCs Drives Rapid Genome-wide Demethylation to the Epigenetic Ground State of Pluripotency, Cell Stem Cell 13: 351–359.

Fouse SD, Shen Y, Pellegrini M, Cole S, Meissner A, van Neste L, Jaenisch R, Fan G. 2008. Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation, Cell Stem Cell 2: 160–169.

Gavin DP, Kusumo H, Sharma RP, Guizzetti M, Guidotti A, Pandey SC. 2015. Gadd45b and N-methyl-D-aspartate induced DNA demethylation in postmitotic neurons, Epigenomics 7: 567–579.

Gierl MS, Gruhn WH, Seggern A von, Maltry N, Niehrs C. 2012. GADD45G functions in male sex determination by promoting p38 signaling and Sry expression, Dev Cell 23: 1032–1042.

Guo JU, Su Y, Zhong C, Ming G-l, Song H. 2011. Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond, Cell Cycle 10: 2662–2668.

Gupta M, Gupta SK, Balliet AG, Hollander MC, Fornace AJ, Hoffman B, Liebermann DA. 2005. Hematopoietic cells from Gadd45a- and Gadd45b-deficient mice are sensitized to genotoxic-stress-induced apoptosis, Oncogene 24: 7170–7179.

Habibi E, Brinkman AB, Arand J, Kroeze LI, Kerstens HHD, Matarese F, Lepikhov K, Gut M, Brun-Heath I, Hubner NC, et al. 2013. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells, Cell Stem Cell 13: 360–369.

Hayashi K, Sousa Lopes SMC de, Tang F, Lao K, Surani MA. 2008. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states, Cell Stem Cell 3: 391–401.

Hayashi Y, Furue MK, Tanaka S, Hirose M, Wakisaka N, Danno H, Ohnuma K, Oeda S, Aihara Y, Shiota K, et al. 2010. BMP4 induction of trophoblast from mouse embryonic stem cells in defined culture conditions on laminin, In Vitro Cell Dev Biol Anim 46: 416–430.

He Y-F, Li B-Z, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, et al. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA, Science 333: 1303–1307.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities, Mol Cell 38: 576–589.

Hendrickson PG, Doráis JA, Grow EJ, Whiddon JL, Lim J-W, Wike CL, Weaver BD, Pflueger C, Emery BR, Wilcox AL, et al. 2017. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons, Nat Genet 49: 925–934.

Hill PWS, Leitch HG, Requena CE, Sun Z, Amouroux R, Roman-Trufero M, Borkowska M, Terragni J, Vaisvila R, Linnett S, et al. 2018. Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte, Nature 555: 392–396.

Hirata T, Amano T, Nakatake Y, Amano M, Piao Y, Hoang HG, Ko MSH. 2012. Zscan4 transiently reactivates early embryonic genes during the generation of induced pluripotent stem cells, Sci Rep 2: 208.

Hollander MC, Sheikh MS, Bulavin DV, Lundgren K, Augeri-Henmueller L, Shehee R, Molinaro TA, Kim KE, Tolosa E, Ashwell JD, et al. 1999. Genomic instability in Gadd45a-deficient mice, Nat Genet 23: 176–184.

Huang HS, Kubish GM, Redmond TM, Turner DL, Thompson RC, Murphy GG, Uhler MD. 2010. Direct transcriptional induction of Gadd45gamma by Ascl1 during neuronal differentiation, Mol Cell Neurosci 44: 282–296.

Huang Y, Chavez L, Chang X, Wang X, Pastor WA, Kang J, Zepeda-Martínez JA, Pape UJ, Jacobsen SE, Peters B, et al. 2014. Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells, Proc Natl Acad Sci U S A 111: 1361–1366.

Huang Y, Kim JK, Do DV, Lee C, Penfold CA, Zylicz JJ, Marioni JC, Hackett JA, Surani MA. 2017. Stella modulates transcriptional and endogenous retrovirus programs during maternal-to-zygotic transition, Elife 6.

Iaco A de, Planet E, Coluccio A, Verp S, Duc J, Trono D. 2017. DUX-family transcription factors regulate zygotic genome activation in placental mammals, Nat Genet 49: 941–945.

Ishiuchi T, Enriquez-Gasca R, Mizutani E, Bošković A, Ziegler-Birling C, Rodriguez-Terrones D, Wakayama T, Vaquerizas JM, Torres-Padilla M-E. 2015. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly, Nat Struct Mol Biol 22: 662–671.

Ishiuchi T, Torres-Padilla M-E. 2013. Towards an understanding of the regulatory mechanisms of totipotency, Curr Opin Genet Dev 23: 512–518.

Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine, Science 333: 1300–1303.

Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J, Ramsahoye B. 2004. Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells, Mol Cell Biol 24: 8862–8871.

Jarome TJ, Butler AA, Nichols JN, Pacheco NL, Lubin FD. 2015. NF-κB mediates Gadd45β expression and DNA demethylation in the hippocampus during fear memory formation, Front. Mol. Neurosci. 8: 539.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity, Science 337: 816–821.

Jukam D, Shariati SAM, Skotheim JM. 2017. Zygotic Genome Activation in Vertebrates, Dev Cell 42: 316–332.

Kang J, Lienhard M, Pastor WA, Chawla A, Novotny M, Tsagaratou A, Lasken RS, Thompson EC, Surani MA, Koralov SB, et al. 2015. Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis, Proc Natl Acad Sci U S A 112: E4236-45.

Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, Shinkai Y, Mager DL, Jones S, Hirst M, et al. 2011. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs, Cell Stem Cell 8: 676–687.

Kaufmann LT, Niehrs C. 2011. Gadd45a and Gadd45g regulate neural development and exit from pluripotency in Xenopus, Mech Dev 128: 401–411.

Khoueiry R, Sohni A, Thienpont B, Luo X, Velde JV, Bartoccetti M, Boeckx B, Zwijsen A, Rao A, Lambrechts D, et al. 2017. Lineage-specific functions of TET1 in the postimplantation mouse embryo, Nat Genet 49: 1061–1072.

Kieffer-Kwon K-R, Tang Z, Mathe E, Qian J, Sung M-H, Li G, Resch W, Baek S, Pruett N, Grøntved L, et al. 2013. Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation, Cell 155: 1507–1520.

Kienhöfer S, Musheev MU, Stapf U, Helm M, Schomacher L, Niehrs C, Schäfer A. 2015. GADD45a physically and functionally interacts with TET1, Differentiation 90: 59–68.

Kim HS, Tan Y, Ma W, Merkurjev D, Destici E, Ma Q, Suter T, Ohgi K, Friedman M, Skowronska-Krawczyk D, et al. 2018. Pluripotency factors functionally premark cell-type-restricted enhancers in ES cells, Nature 556: 510–514.

Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, An J, Lamperti ED, Koh KP, Ganetzky R, et al. 2010. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2, Nature 468: 839–843.

Kong L, Tan L, Lv R, Shi Z, Xiong L, Wu F, Rabidou K, Smith M, He C, Zhang L, et al. 2016. A primary role of TET proteins in establishment and maintenance of De Novo bivalency at CpG islands, Nucleic Acids Res 44: 8682–8692.

Kriaucionis S, Heintz N. 2009. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain, Science 324: 929–930.

La Rica L de, Deniz Ö, Cheng KCL, Todd CD, Cruz C, Houseley J, Branco MR. 2016. TET-dependent regulation of retrotransposable elements in mouse embryonic stem cells, Genome Biol 17: 234.

Le Cong, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. 2013. Multiplex genome engineering using CRISPR/Cas systems, Science 339: 819–823.

Lee HJ, Hore TA, Reik W. 2014. Reprogramming the methylome: erasing memory and creating diversity, Cell Stem Cell 14: 710–719.

Lei H, Oh SP, Okano M, Jüttermann R, Goss KA, Jaenisch R, Li E. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells, Development 122: 3195–3205.

Leitch HG, McEwen KR, Turp A, Encheva V, Carroll T, Grabole N, Mansfield W, Nashun B, Knezovich JG, Smith A, et al. 2013. Naive pluripotency is associated with global DNA hypomethylation, Nat Struct Mol Biol 20: 311–316.

Li Y, Zhao M, Yin H, Gao F, Wu X, Luo Y, Zhao S, Zhang X, Su Y, Hu N, et al. 2010. Overexpression of the growth arrest and DNA damage-induced 45alpha gene contributes to autoimmunity by promoting DNA demethylation in lupus T cells, Arthritis Rheum 62: 1438–1447.

Li Z, Gu T-P, Weber AR, Shen J-Z, Li B-Z, Xie Z-G, Yin R, Guo F, Liu X, Tang F, et al. 2015. Gadd45a promotes DNA demethylation through TDG, Nucleic Acids Res 43: 3986–3997.

Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schübeler D. 2011. Identification of genetic elements that autonomously determine DNA methylation states, Nat Genet 43: 1091–1097.

Liu Y, Olanrewaju YO, Zheng Y, Hashimoto H, Blumenthal RM, Zhang X, Cheng X. 2014. Structural basis for Klf4 recognition of methylated DNA, Nucleic Acids Res 42: 4859–4867.

Lu B, Ferrandino AF, Flavell RA. 2004. Gadd45beta is important for perpetuating cognate and inflammatory signals in T cells, Nat Immunol 5: 38–44.

Lu B, Yu H, Chow C, Li B, Zheng W, Davis RJ, Flavell RA. 2001. GADD45gamma mediates the activation of the p38 and JNK MAP kinase pathways and cytokine production in effector TH1 cells, Immunity 14: 583–590.

Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. 2014. Role of Tet proteins in enhancer activity and telomere elongation, Genes Dev 28: 2103–2119.

Ma DK, Jang M-H, Guo JU, Kitabatake Y, Chang M-l, Pow-anpongkul N, Flavell RA, Lu B, Ming G-l, Song H. 2009. Neuronal Activity-Induced Gadd45b Promotes Epigenetic DNA Demethylation and Adult Neurogenesis, Science 323: 1074–1077.

Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity, Nature 487: 57–63.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9, Science 339: 823–826.

Messerschmidt DM, Knowles BB, Solter D. 2014. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos, Genes Dev 28: 812–828.

Narducci MG, Fiorenza MT, Kang S-M, Bevilacqua A, Di Giacomo M, Remotti D, Picchio MC, Fidanza V, Cooper MD, Croce CM, et al. 2002. TCL1 participates in early embryonic development and is overexpressed in human seminomas, Proc Natl Acad Sci U S A 99: 11712–11717.

Ng RK, Dean W, Dawson C, Lucifero D, Madeja Z, Reik W, Hemberger M. 2008. Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5, Nat Cell Biol 10: 1280–1290.

Nishiyama A, Xin L, Sharov AA, Thomas M, Mowrer G, Meyers E, Piao Y, Mehta S, Yee S, Nakatake Y, et al. 2009. Uncovering early response of gene regulatory networks in ESCs by systematic induction of transcription factors, Cell Stem Cell 5: 420–433.

Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development, Cell 99: 247–257.

Parisi S, Passaro F, Aloia L, Manabe I, Nagai R, Pastore L, Russo T. 2008. Klf5 is involved in self-renewal of mouse embryonic stem cells, J Cell Sci 121: 2629–2634.

Park S-J, Shirahige K, Ohsugi M, Nakai K. 2015. DBTMEE. A database of transcriptome in mouse early embryos, Nucleic Acids Res 43: D771-6.

Percharde M, Lin C-J, Yin Y, Guan J, Peixoto GA, Bulut-Karslioglu A, Biechele S, Huang B, Shen X, Ramalho-Santos M. 2018. A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity, Cell 174: 391-405.e19.

Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR. 2008. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45, Cell 135: 1201–1212.

Rodriguez-Terrones D, Gaume X, Ishiuchi T, Weiss A, Kopp A, Kruse K, Penning A, Vaquerizas JM, Brino L, Torres-Padilla M-E. 2018. A molecular roadmap for the emergence of early-embryonic-like cells in culture, Nat Genet 50: 106–119.

Sabag O, Zamir A, Keshet I, Hecht M, Ludwig G, Tabib A, Moss J, Cedar H. 2014. Establishment of methylation patterns in ES cells, Nat Struct Mol Biol 21: 110–112.

Sakaue M, Ohta H, Kumaki Y, Oda M, Sakaide Y, Matsuoka C, Yamagiwa A, Niwa H, Wakayama T, Okano M. 2010. DNA methylation is dispensable for the growth and survival of the extraembryonic lineages, Curr Biol 20: 1452–1457.

Schäfer A, Karaulanov E, Stapf U, Döderlein G, Niehrs C. 2013. Ing1 functions in DNA demethylation by directing Gadd45a to H3K4me3, Genes Dev 27: 261–273.

Schäfer A, Mekker B, Mallick M, Vastolo V, Karaulanov E, Sebastian D, Lippen C von der, Epe B, Downes DJ, Scholz C, et al. 2018. Impaired DNA demethylation of C/EBP sites causes premature aging, Genes Dev 32: 742–762.

Schmitz K-M, Schmitt N, Hoffmann-Rohrer U, Schäfer A, Grummt I, Mayer C. 2009. TAF12 Recruits Gadd45a and the Nucleotide Excision Repair Complex to the Promoter of rRNA Genes Leading to Active DNA Demethylation, Mol Cell 33: 344–353.

Schomacher L, Han D, Musheev MU, Arab K, Kienhöfer S, Seggern A von, Niehrs C. 2016. Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation, Nat Struct Mol Biol 23: 116–124.

Sen GL, Reuter JA, Webster DE, Zhu L, Khavari PA. 2010. DNMT1 maintains progenitor function in self-renewing somatic tissue, Nature 463: 563–567.

Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung H-L, Zhang K, Zhang Y. 2013. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics, Cell 153: 692–706.

Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PWTC, Bauer C, Münzel M, Wagner M, Müller M, Khan F, et al. 2013. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives, Cell 152: 1146–1159.

Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, et al. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions, Nature 480: 490–495.

Svoboda P. 2017. Mammalian zygotic genome activation, Semin Cell Dev Biol.

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, et al. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, Science 324: 930–935.

Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H. 2008. Identification and characterization of subpopulations in undifferentiated ES cell culture, Development 135: 909–918.

Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S-i, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, et al. 2006. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b, Genes Cells 11: 805–814.

Wang PJ, McCarrey JR, Yang F, Page DC. 2001. An abundance of X-linked genes expressed in spermatogonia, Nat Genet 27: 422–426.

Whiddon JL, Langford AT, Wong C-J, Zhong JW, Tapscott SJ. 2017. Conservation and innovation in the DUX4-family gene network, Nat Genet 49: 935–940.

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PAC, Rappsilber J, Helin K. 2011. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity, Nature 473: 343–348.

Xu H, Ang Y-S, Sevilla A, Lemischka IR, Ma'ayan A. 2014. Construction and validation of a regulatory network for pluripotency and self-renewal of mouse embryonic stem cells, PLoS Comput Biol 10: e1003777.

Yamaguchi S, Hong K, Liu R, Shen L, Inoue A, Diep D, Zhang K, Zhang Y. 2012. Tet1 controls meiosis by regulating meiotic gene expression, Nature 492: 443–447.

Yamaguchi S, Shen L, Liu Y, Sendler D, Zhang Y. 2013. Role of Tet1 in erasure of genomic imprinting, Nature 504: 460–464.

Yang J, Guo R, Wang H, Ye X, Zhou Z, Dan J, Wang H, Gong P, Deng W, Yin Y, et al. 2016. Tet Enzymes Regulate Telomere Maintenance and Chromosomal Stability of Mouse ESCs, Cell Rep 15: 1809–1821.

Ying Q-L, Stavridis M, Griffiths D, Li M, Smith A. 2003. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture, Nat Biotechnol 21: 183–186.

Zalzman M, Falco G, Sharova LV, Nishiyama A, Thomas M, Lee S-L, Stagg CA, Hoang HG, Yang H-T, Indig FE, et al. 2010. Zscan4 regulates telomere elongation and genomic stability in ES cells, Nature 464: 858–863.

Zhang R-p, Shao J-z, Xiang L-x. 2011. GADD45A protein plays an essential role in active DNA demethylation during terminal osteogenic differentiation of adipose-derived mesenchymal stem cells, J Biol Chem 286: 41083–41094.

Zhang W, Fu S, Liu X, Zhao X, Zhang W, Peng W, Wu C, Li Y, Li X, Bartlam M, et al. 2011. Crystal structure of human Gadd45γ corrected reveals an active dimer, Protein Cell 2: 814–826.

Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LT-Y, Kohlbacher O, Jager PL de, Rosen ED, Bennett DA, Bernstein BE, et al. 2013. Charting a dynamic DNA methylation landscape of the human genome, Nature 500: 477–481.

Figures

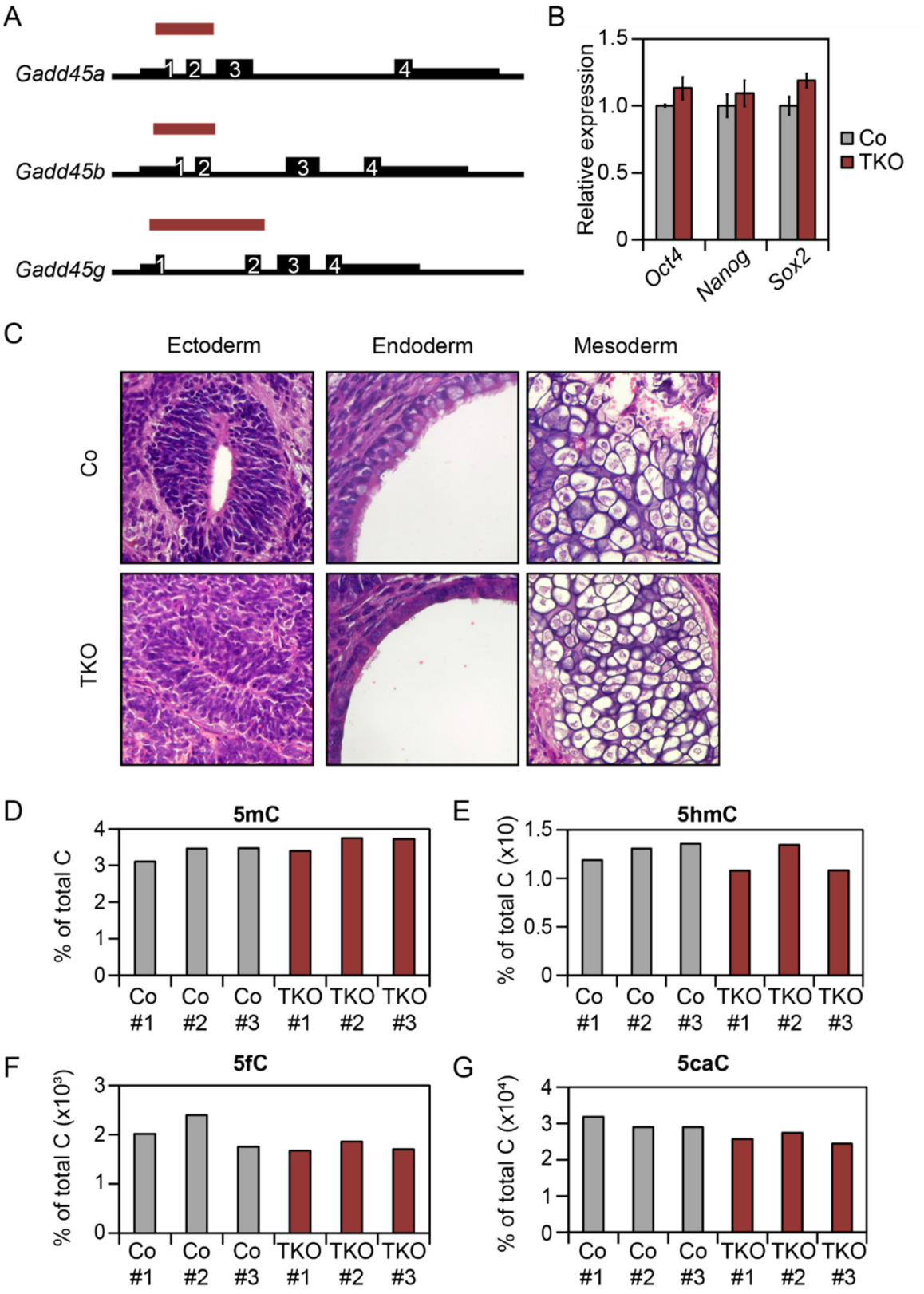
A
Gadd45a
Gadd45b
Gadd45g
1
2
3
4
B
Relative expression
1.5
1.0
0.5
0
Co
TKO
Oct4
Nanog
Sox2
C
Ectoderm
Endoderm
Mesoderm
Co
TKO
D
5mC
% of total C
4
3
2
1
0
Co #1
Co #2
Co #3
TKO #1
TKO #2
TKO #3
E
5hmC
% of total C (x10)
1.5
1.0
0.5
0
F
5fC
% of total C (x10³)
G
5caC
% of total C (x10⁴)

Figure 1: *Gadd45* TKO-ESCs are pluripotent and show normal global levels of DNA modifications

(A) Scheme of the CRISPR/Cas9-mediated *Gadd45* knockout strategy. Numbers, exons; red bars, location of deletion.

(B) Relative expression of representative pluripotency markers in Control- (Co) and *Gadd45* TKO-ESCs measured by qPCR. Expression values are relative to the average expression in Co-ESCs.

(C) Hematoxylin/Eosin staining of paraffin Co and *Gadd45* TKO teratoma sections. Representative examples for ectoderm, endoderm and mesoderm derivatives are shown.

(D-G) 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) levels in individual Co- and *Gadd45* TKO-ESC clones determined by LC-MS/MS. Values are given as % of total cytosine (C).

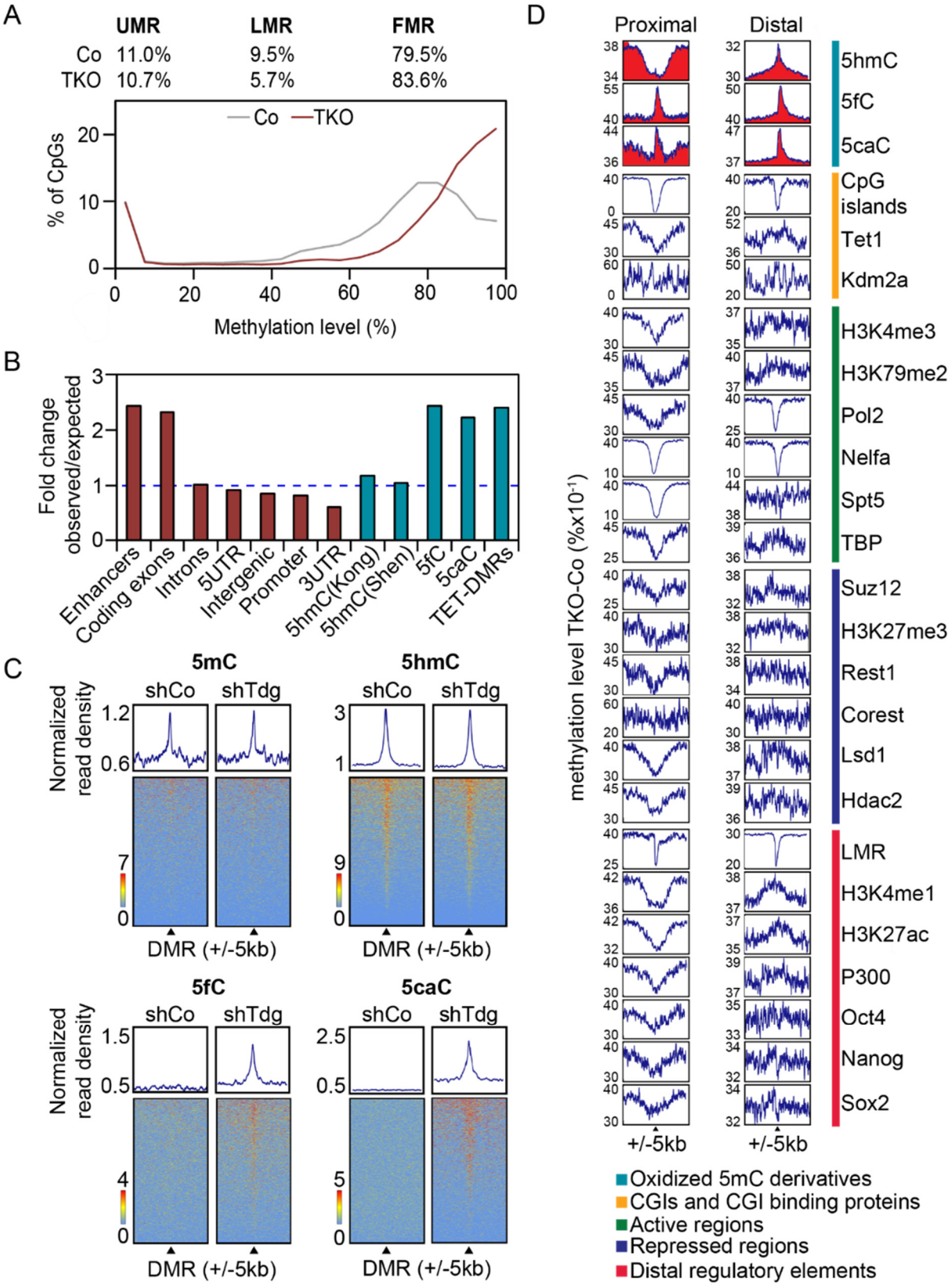
A
UMR
LMR
FMR
Co 11.0% 9.5% 79.5%
TKO 10.7% 5.7% 83.6%
Co
TKO
% of CpGs
Methylation level (%)
B
Fold change observed/expected
Enhancers
Coding exons
Introns
5UTR
Intergenic
Promoter
3UTR
5hmC(Kong)
5hmC(Shen)
5fC
5caC
TET-DMRs
C
5mC
5hmC
5fC
5caC
shCo
shTdg
Normalized read density
DMR (+/-5kb)
D
Proximal
Distal
methylation level TKO-Co (%x10-1)
5hmC
5fC
5caC
CpG islands
Tet1
Kdm2a
H3K4me3
H3K79me2
Pol2
Nelfa
Spt5
TBP
Suz12
H3K27me3
Rest1
Corest
Lsd1
Hdac2
LMR
H3K4me1
H3K27ac
P300
Oct4
Nanog
Sox2
+/-5kb
Oxidized 5mC derivatives
CGIs and CGI binding proteins
Active regions
Repressed regions
Distal regulatory elements

Figure 2: Loci undergoing TET-dependent oxidation are hypermethylated in *Gadd45* TKO-ESCs

(A) The bimodal methylation pattern of individual CpG sites in ESCs is skewed to higher methylation level in *Gadd45* TKO-ESCs. Methylation levels are shown as the average of two biological replicates. UMR, unmethylated regions; LMR, lowly methylated regions; FMR, fully methylated regions as defined by (Stadler et al. 2011).

(B) Relative enrichment of *Gadd45* TKO hypermethylated DMRs (hyper-DMRs) at various genomic elements (red), oxidative 5mC derivatives and *Tet*-TKO hyper-DMRs (blue).

(C) Heat maps of depicted DNA modifications (Shen et al. 2013) centered (+/- 5kb) on *Gadd45* TKO hyper-DMRs (black triangles) in untreated and shTdg-treated ESCs.

(D) Average methylation differences between *Gadd45* TKO- and Co-ESCs around centers (+/- 5kb) of annotated genomic features. Methylation differences are shown for proximal- (within 1 kb of a gene TSS) and distal features. Red areas highlight oxidized 5mC derivatives.

A

Hyper-DMR	Name	Motif	p-value
All	Klf5		$1e^{-23}$
	Klf6		$1e^{-10}$
	Ehf		$1e^{-8}$
Promoters	n.s.	n.s.	n.s.
Enhancers	n.s.	n.s.	n.s.
5hmC	Klf5		$1e^{-8}$
	Etv1		$1e^{-7}$
	Fli1		$1e^{-6}$
5fC	Ventx		$1e^{-10}$
	Elk4		$1e^{-7}$
	Elk1		$1e^{-6}$
Promoters 5hmC	n.s.	n.s.	n.s.
Promoters 5fC	Smad2		$1e^{-8}$
	Nfic		$1e^{-6}$
Enhancers 5hmC	Zscan4		$1e^{-7}$
Enhancers 5fC	n.s.	n.s.	n.s.

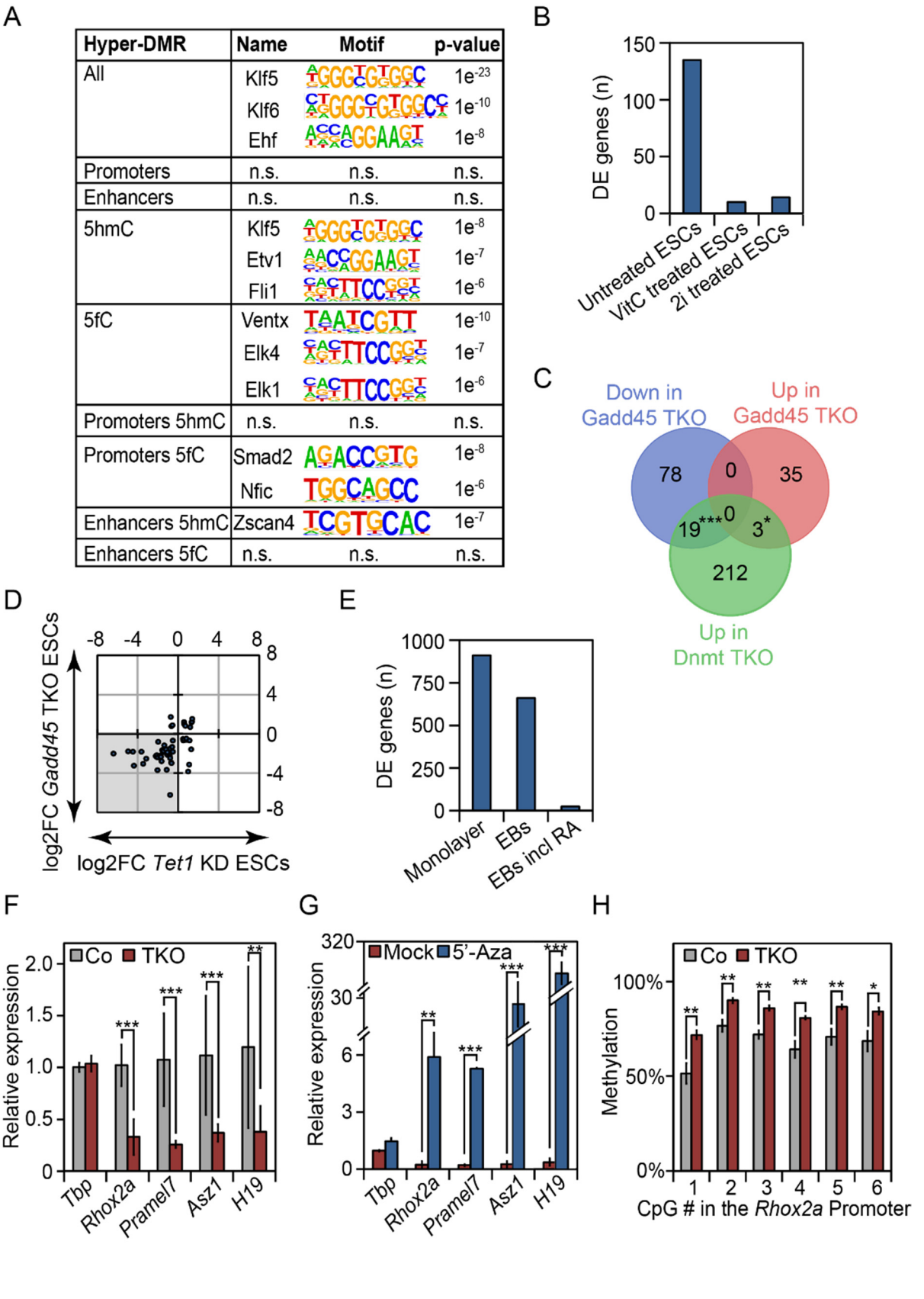

Figure 3: Methylation-regulated genes are downregulated in *Gadd45* TKO-ESCs

(A) Motif analysis of hyper-DMRs in *Gadd45* TKO-ESCs using HOMER (Heinz et al. 2010).

(B) Differentially expressed (DE) genes (FDR 10%) identified by RNA-Seq in *Gadd45* TKO-ESCs versus Control- (Co) ESCs, which were untreated, 72h vitamin C treated (VitC), or 72h 2i treated.

(C) Overlap of genes down- and upregulated in untreated *Gadd45* TKO-ESCs and genes upregulated in *Dnmt1,2,3* TKO-ESCs (Karimi et al. 2011).

(D) Scatterplot of the common deregulated genes (grey) in *Gadd45* TKO and *Tet1* knockdown ESCs (Log2FC, Log2 of Fold Change versus control ESCs; Huang et al. 2014).

(E) Differentially expressed (DE) genes (FDR 10%) identified by RNA-Seq in *Gadd45* TKO cells versus Co cells upon monolayer differentiation (Monolayer), embryoid body differentiation (EBs) or retinoic acid (RA) stimulation during EB differentiation.

(F) Expression of selected GADD45-dependent genes in Co- and *Gadd45* TKO-ESCs measured by qPCR. Expression is relative to Co-ESCs. Data are presented as mean ± SD from n=3 independent clones and n=3 independent experiments.

(G) GADD45-regulated genes are DNA methylation-sensitive. Relative expression levels of selected GADD45-dependent genes in *Gadd45* TKO-ESCs upon 48h of DMSO (Mock) or 5'-azadeoxycytidine (5'-Aza) treatment measured by qPCR. Expression is relative to DMSO treated Co-ESCs.

(H) Hypermethylation of GADD45-dependent genes. DNA methylation of indicated CpGs in the *Rhox2a* promoter in Co- and *Gadd45* TKO-ESCs monitored by site-specific bisulfite sequencing.

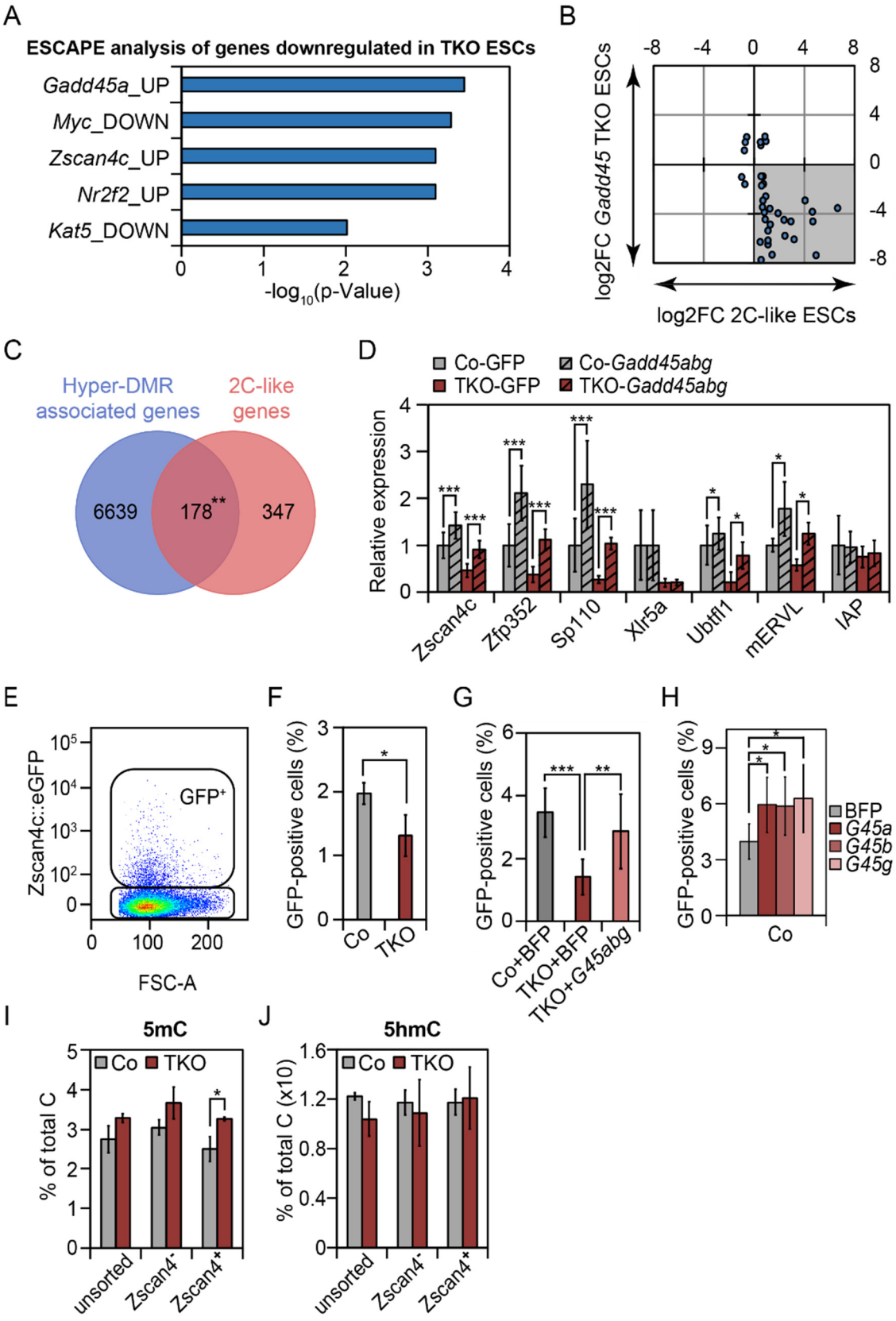
A
ESCAPE analysis of genes downregulated in TKO ESCs
Gadd45a_UP
Myc_DOWN
Zscan4c_UP
Nr2f2_UP
Kat5_DOWN
-log10(p-Value)
B
log2FC Gadd45 TKO ESCs
log2FC 2C-like ESCs
C
Hyper-DMR associated genes
2C-like genes
6639
178**
347
D
Co-GFP
Co-Gadd45abg
TKO-GFP
TKO-Gadd45abg
Relative expression
Zscan4c
Zfp352
Sp110
Xlr5a
Ubtfl1
mERVL
IAP
E
Zscan4c::eGFP
GFP+
FSC-A
F
GFP-positive cells (%)
Co
TKO
G
GFP-positive cells (%)
Co+BFP
TKO+BFP
TKO+G45abg
H
GFP-positive cells (%)
BFP
G45a
G45b
G45g
Co
I
5mC
Co
TKO
% of total C
unsorted
Zscan4-
Zscan4+
J
5hmC
Co
TKO
% of total C (x10)
unsorted
Zscan4-
Zscan4+

Figure 4: *Gadd45* TKO-ESCs show impairment of the 2C-like state

(A) Enrichment analysis of genes downregulated in *Gadd45* TKO-ESCs using ESCAPE database (Xu et al. 2014). Enrichment analysis shows a significant overlap with genes deregulated (UP or DOWN) upon overexpression of e.g. *Gadd45a* or *Zscan4c*.

(B) Scatterplot of the common deregulated genes (grey) in *Gadd45* TKO-ESCs and upregulated in the 2C-like state (Log2FC, Log2 of Fold Change versus control ESCs; Macfarlan et al. 2012)

(C) Overlap between 2C- (Macfarlan et al. 2012) and hyper-DMR-associated genes.

(D) qPCR expression analysis of selected 2C-associated genes and retroviral elements in Control- (Co) and *Gadd45* TKO-ESCs, 48h after transfection with the indicated genes. Expression is relative to GFP transfected Co-ESCs. Data are presented as mean ± SD from n = 3 independent clones and n=2 independent experiments. Statistical significance was tested with two-tailed, paired Student's t-test.

(E) Scatterplot showing Zscan4c::eGFP positive cells in bulk Co-ESCs measured by flow cytometry analysis. Representative gates used for bulk analysis are boxed.

(F) Flow cytometry analysis of Zscan4c::eGFP positive cells in Co- and *Gadd45* TKO-ESCs.

(G) Flow cytometry analysis of Zscan4c::eGFP positive cells in Co- or *Gadd45*-TKO-ESC, 48h after transfection with the indicated genes. Data are presented as means ± SD from n = 3 independent clones and n=2 independent experiments. Statistical significance was tested with two-tailed, unpaired (TKO versus Co) or paired (BFP versus *Gadd45* overexpression) Student's t-test.

(H) Flow cytometry analysis of Zscan4::eGFP positive cells in Co-ESCs, 48h after transfection with the indicated genes. Statistical significance was tested with two-tailed, paired Student's t-test.

(I-J) 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) levels in unsorted, Zscan4^{-} or Zscan4^{+} sorted Co- and *Gadd45* TKO-ESC clones determined by LC-MS/MS. Values are % of total cytosine (C).

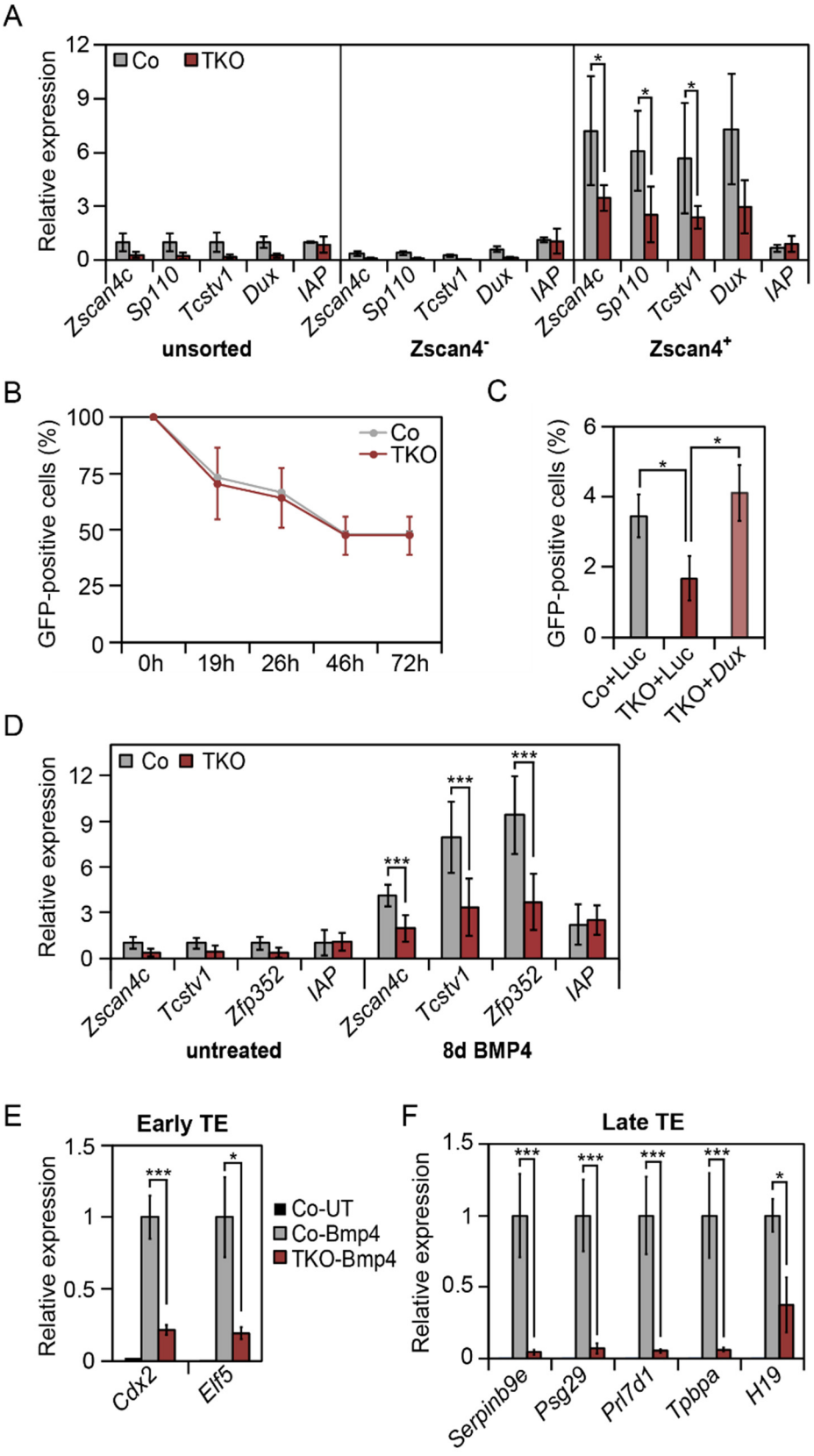
A
Relative expression
Co
TKO
Zscan4c
Sp110
Tcstv1
Dux
IAP
unsorted
Zscan4-
Zscan4+
B
GFP-positive cells (%)
0h
19h
26h
46h
72h
C
GFP-positive cells (%)
Co+Luc
TKO+Luc
TKO+Dux
D
untreated
8d BMP4
Zfp352
E
Early TE
Co-UT
Co-Bmp4
TKO-Bmp4
Cdx2
Elf5
F
Late TE
Serpinb9e
Psg29
Prl7d1
Tpbpa
H19

Figure 5: *Gadd45* TKO-ESCs show reduced 2C-like gene expression and transdifferentiation potential

(A) qPCR expression analysis of selected 2C-associated genes in Control- (Co) and *Gadd45* TKO-ESC in unsorted, Zscan4^{-} or Zscan4^{+} FACS-sorted cells. Expression is relative to the average expression in unsorted Co-ESCs. IAP, intracisternal A-particle

(B) Reanalysis of Zcan4^{+} sorted Co- and TKO-ESCs at the indicated time points using flow cytometry.

(C) *Dux* overexpression restores the reduced number of 2C-like cells in *Gadd45* TKO-ESCs. Flow cytometry analysis of Zscan4c::eGFP positive cells in Co- and *Gadd45* TKO-ESC, 48h after transfection with the indicated genes and 24 hours after doxycycline addition. Statistical significance was tested with two-tailed, unpaired (TKO versus control) or paired (*Luciferase* versus *Dux* overexpression) Student's t-test.

(D) qPCR expression analysis of selected 2C-associated genes in untreated or BMP4 treated (8 days) Co- and *Gadd45* TKO-ESC. Expression is relative to untreated Co-ESCs. Data are presented as mean ± SD from n=3 independent clones and n=3 independent experiments.

(E-F) Impairment of trophectoderm transdifferentiation in *Gadd45* TKO-ESCs. *Gadd45* TKO and Co-ESCs were treated for 8 days with BMP4. Induction of early (E) and late (F) trophectoderm (TE) marker genes is relative to BMP4-treated Co-ESCs. UT, untreated

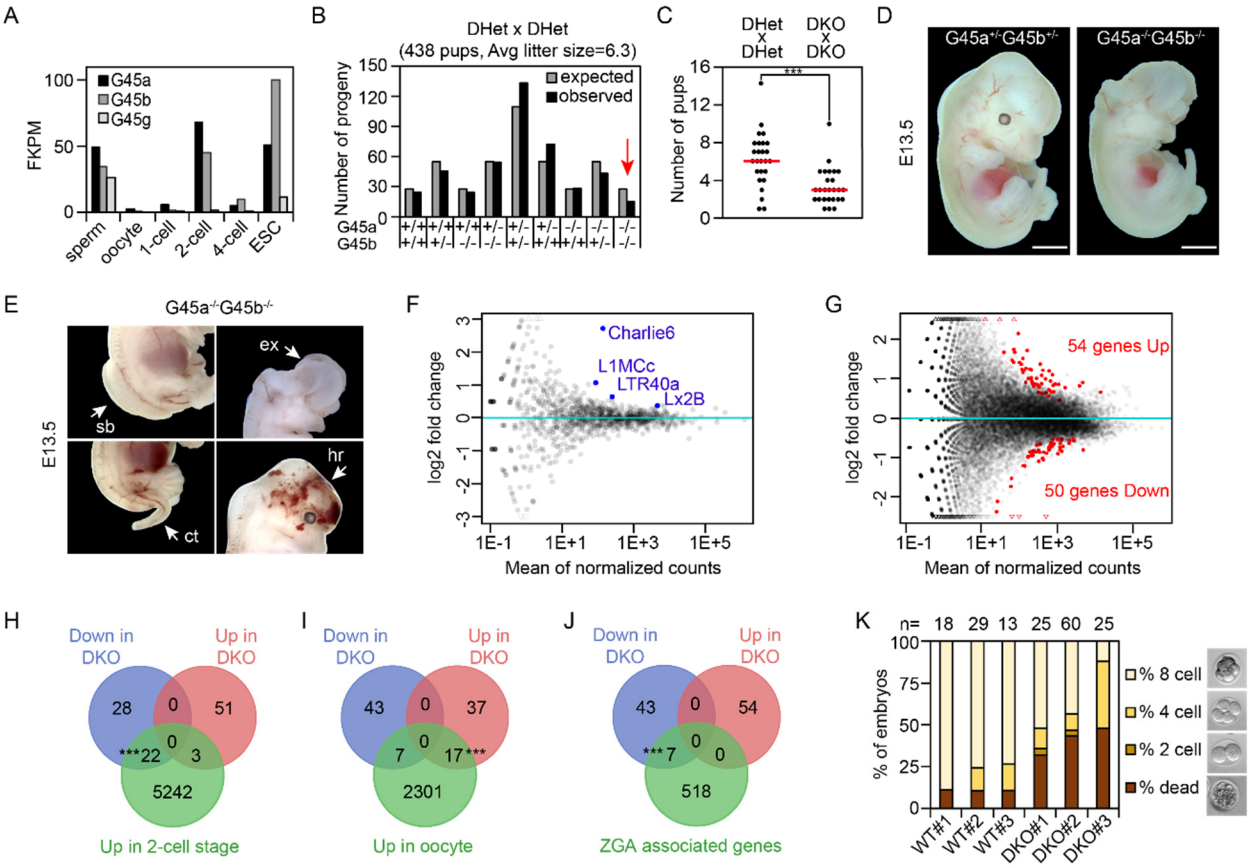

Figure 6: (*Gadd45a/Gadd45b*)$^{-/-}$ mice show embryonic sublethality

(A) Expression analysis of *Gadd45a, -b, -g* in preimplantation embryos and ESCs from DBTMEE database (Park et al. 2015). FPKM, Fragments Per Kilobase of exon model per Million reads mapped.

(B) Genotypic analysis of progenies from *Gadd45a,b* double heterozygous (DHet) mice showing expected and observed Mendelian ratios.

(C) (*Gadd45a/Gadd45b*)$^{-/-}$ (DKO) intercrossing show reduced liter size compared to intercrossed double heterozygous animals. Data points indicate the number of pups per litter. Red lines indicate average litter size. Statistical significance was tested with two-tailed, paired Student's t-test.

(D) Images of E13.5 heterozygous and homozygous *Gadd45a/Gadd45b* mouse embryos. Scale bar, 2 mm.

(E) Developmental abnormalities in DKO embryos showing curly tail (ct), exencephaly (ex), cranial hemorrhage (he), and spina bifida (sb).

(F-G) RNA-seq differential expression analysis of repeats (F) and genes (G) in 2-cell stage DKO embryos. Significantly deregulated (FDR 10%) repeats and genes are highlighted in blue and red respectively.

(H-J) Overlap of genes up- and downregulated in *Gadd45a,b* DKO 2-cell stage embryos and genes normally upregulated in (H) 2-cell stage embryos, (I) oocytes, or (J) during ZGA (Macfarlan et al. 2012).

(K) In vitro development of wild type (WT) and *Gadd45a,b* DKO preimplantation embryos isolated from three independent breeding. Development was scored 24 hours after isolation. Representative microscopic images of the scored embryonic stages are shown.

Supplemental information

GADD45 promotes locus specific DNA demethylation and 2C cycling in embryonic stem cells

Inventory

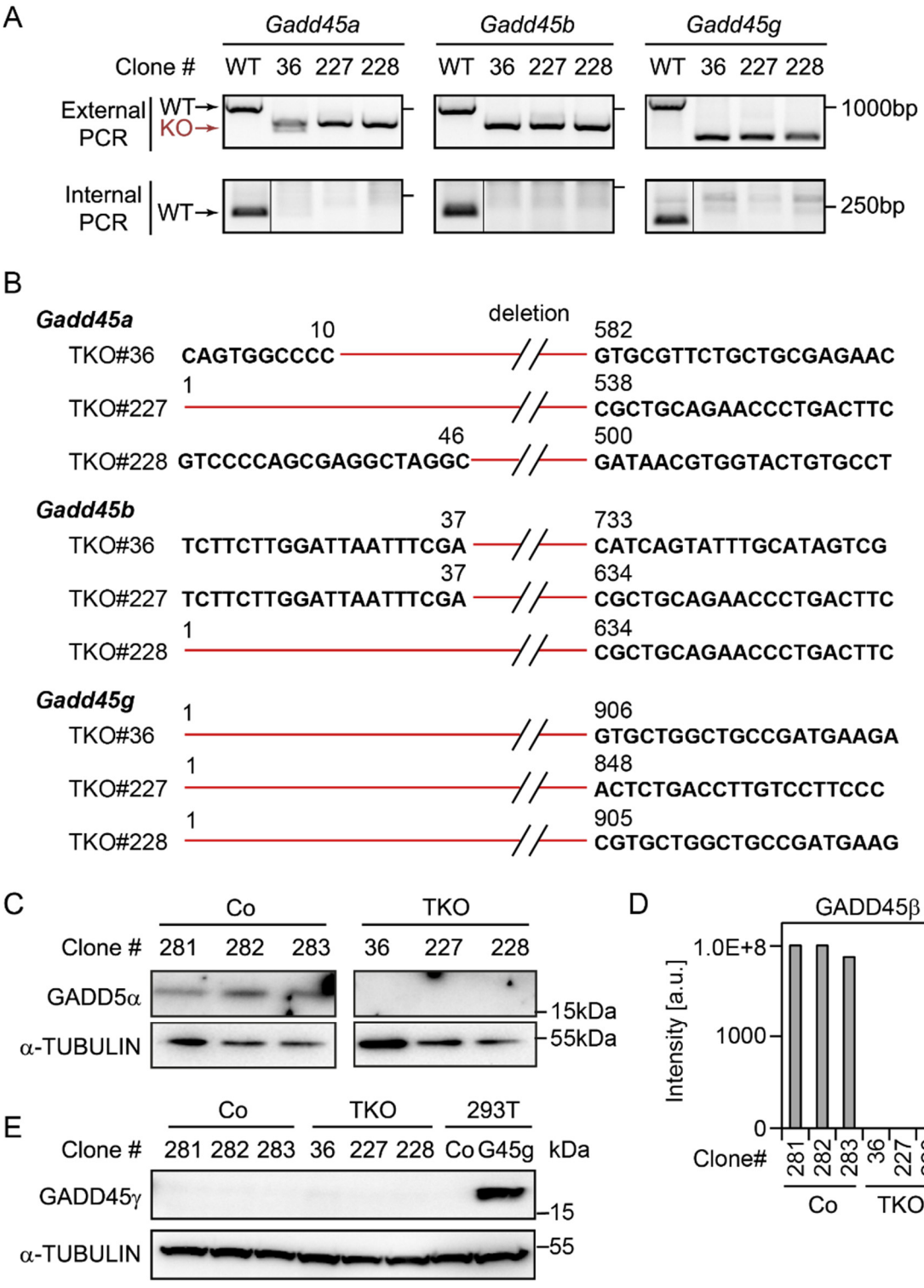
A
Gadd45a
Gadd45b
Gadd45g
Clone # WT 36 227 228 WT 36 227 228 WT 36 227 228
External PCR
WT
KO
1000bp
Internal PCR
WT
250bp
B
Gadd45a
deletion
TKO#36 10 CAGTGGCCCC 582 GTGCGTTCTGCTGCGAGAAC
TKO#227 1 538 CGCTGCAGAACCCTGACTTC
TKO#228 46 GTCCCCAGCGAGGCTAGGC 500 GATAACGTGGTACTGTGCCT
Gadd45b
TKO#36 37 TCTTCTTGGATTAATTTCGA 733 CATCAGTATTTGCATAGTCG
TKO#227 37 TCTTCTTGGATTAATTTCGA 634 CGCTGCAGAACCCTGACTTC
TKO#228 1 634 CGCTGCAGAACCCTGACTTC
Gadd45g
TKO#36 1 906 GTGCTGGCTGCCGATGAAGA
TKO#227 1 848 ACTCTGACCTTGTCCTTCCC
TKO#228 1 905 CGTGCTGGCTGCCGATGAAG
C
Co TKO
Clone # 281 282 283 36 227 228
GADD5α
α-TUBULIN
15kDa
55kDa
D
GADD45β
1.0E+8
1000
0
Intensity [a.u.]
Clone# 281 282 283 36 227 228
Co TKO
E
Co TKO 293T
Clone # 281 282 283 36 227 228 Co G45g kDa
GADD45γ
α-TUBULIN
15
55

Figure S1: Confirmation of successful CRISPR/Cas9 mediated *Gadd45* triple knockout in ESCs

(*A*) Genotyping PCRs using primer pairs external and internal to the desired deletion. External primer pairs are expected to produce a smaller PCR product in the knockout (KO, *Gadd45a* = 619bp, *Gadd45b* = 594bp, *Gadd45g* = 368bp) compared to the wildtype (WT, *Gadd45a* = 972bp, *Gadd45b* = 975bp, *Gadd45g* = 1089bp) genotype. Internal primers are expected to only produce a PCR product in the WT genotype. Red arrow, expected KO PCR product size; Black arrow, expected WT PCR product size.

(*B*) DNA-sequences (partial) of the *Gadd45a*,*-b*,*-g* gene of three independent *Gadd45* TKO-ESC clones. Red lines indicate deletions.

(*C*) Western Blot analysis of three independent Control- (Co) and *Gadd45* TKO-ESC clones demonstrate depletion of GADD45α protein.

(*D*) GADD45β protein is depleted in all three independent *Gadd45* TKO-ESCs. Immunoprecipitation coupled to mass spectrometry analysis of Co- and *Gadd45* TKO-ESC clones is shown.

(*E*) GADD45γ protein is not detectable in ESC whole cell lysates. Western blot analysis of three independent Co- and *Gadd45* TKO-ESC clones. Empty vector and *Gadd45g* transfected HEK293T cell lysates were used as controls.

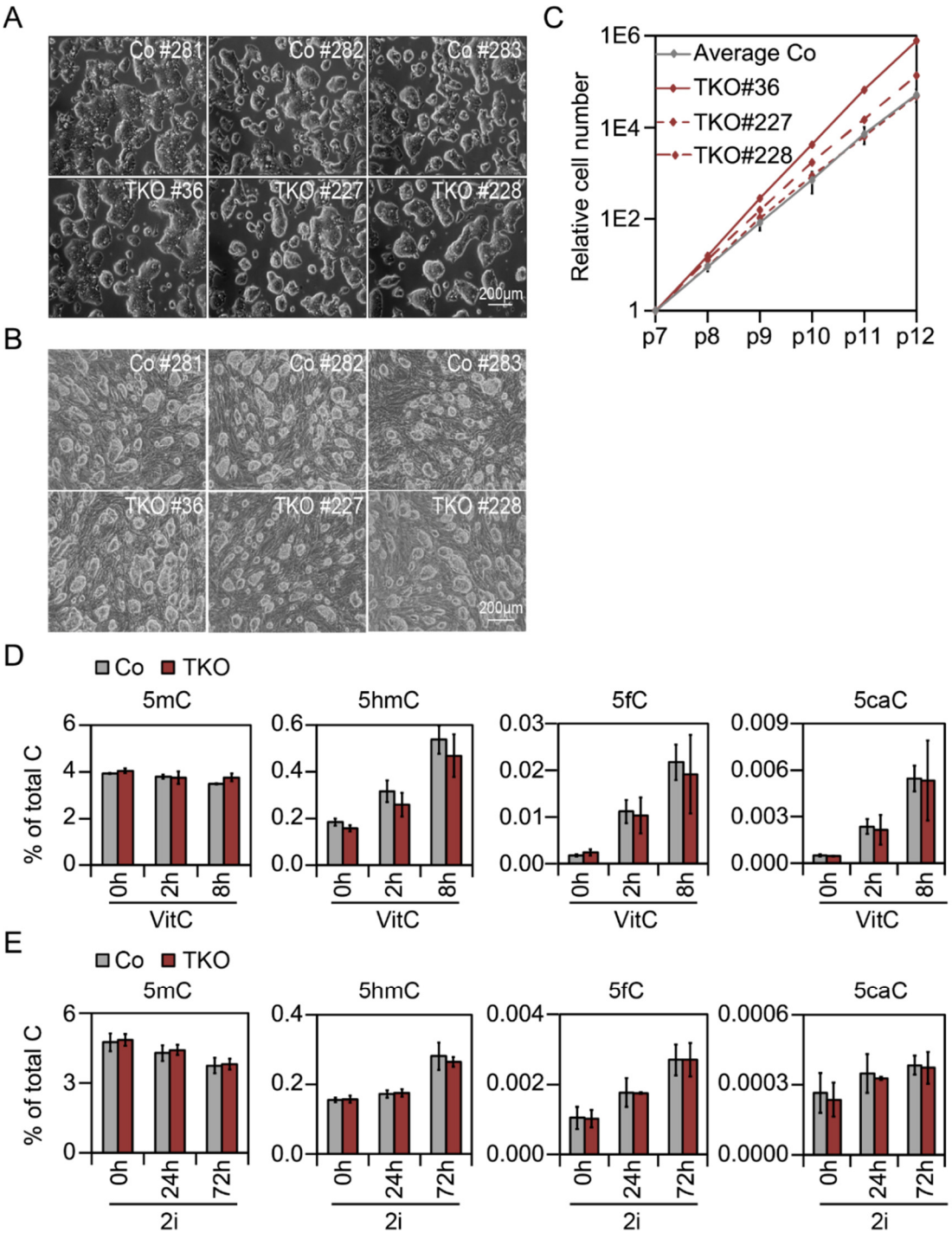
A
Co #281
Co #282
Co #283
TKO #36
TKO #227
TKO #228
200µm
B
Co #281
Co #282
Co #283
TKO #36
TKO #227
TKO #228
200µm
C
Relative cell number
1E6
1E4
1E2
1
Average Co
TKO#36
TKO#227
TKO#228
p7
p8
p9
p10
p11
p12
D
Co
TKO
5mC
5hmC
5fC
5caC
% of total C
0h
2h
8h
VitC
E
Co
TKO
5mC
5hmC
5fC
5caC
% of total C
0h
24h
72h
2i

Figure S2: *Gadd45* TKO-ESC retain normal morphology and undergo normal global DNA demethylation upon shift to ground state pluripotency

(*A*) Representative light microscopy images showing the morphology of three independent Control- (Co) and *Gadd45* TKO-ESC.

(*B*) Representative light microscopy images showing Co- and *Gadd45* TKO-ESC morphology grown on MEF feeder cells.

(*C*) Growth curves of Co- and *Gadd45* TKO-ESCs showing normal growth behavior. Co- ESCs are represented as means ± SD from n = 3 independent clones.

(*D-E*) 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) levels in Co- and *Gadd45* TKO-ESCs after vitamin C (*D*) and 2i treatment (*E*), determined by LC-MS/MS. Values are % of total cytosine (C).

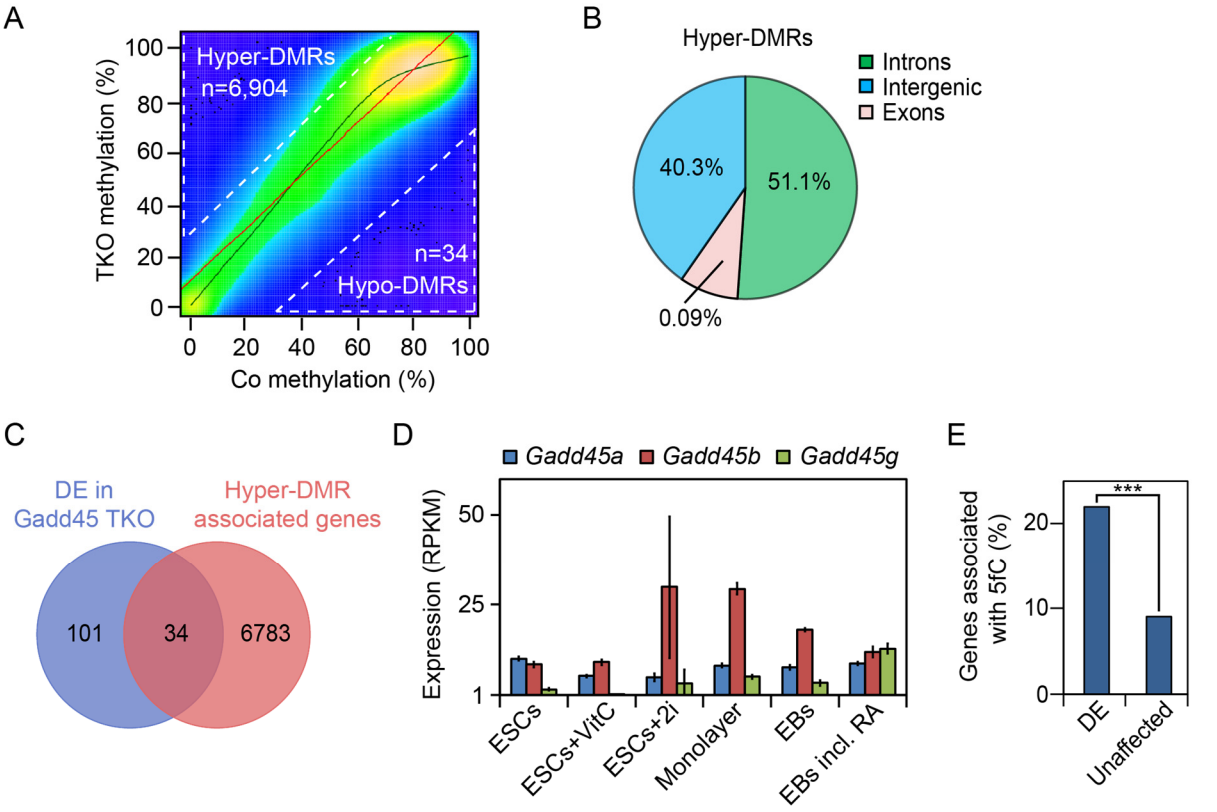

Figure S3: *Gadd45* TKO-ESCs show hypermethylation and hyper-DMRs overlap with introns and intergenic regions

(*A*) Global methylation comparison between Control- (Co) and *Gadd45* TKO-ESCs. Scatter plot shows hyper- and hypomethylated differentially methylated regions (DMRs).

(*B*) Distribution of hyper-DMRs at introns, intergenic regions and exons. Hyper-DMRs overlapping with two features were counted twice.

(*C*) Overlap of deregulated (DE) and hyper-DMR associated genes in *Gadd45* TKO-ESCs.

(*D*) Expression level of *Gadd45a*, *Gadd45b* and *Gadd45g* in Co cells measured by RNA-seq under the indicated conditions: Untreated ESCs, 72h vitamin C treated ESCs (ESCs + VitC), 72h 2i treated ESCs (ESCs + 2i), 6 days monolayer differentiated cells (Monolayer), 8 days differentiated embryoid bodies (EBs), 8 days differentiated embryoid bodies treated for 4 days with retinoic acid (EBs incl. RA). Expression values are represented as reads per kilobase pair per million reads (RPKM).

(*E*) Enrichment of differentially expressed (DE) genes in *Gadd45* TKO EBs with 5-formylcytosine marks in their promoter region (Song et al. 2013). ***p<0.0005; Chi-square test with Yates' correction.

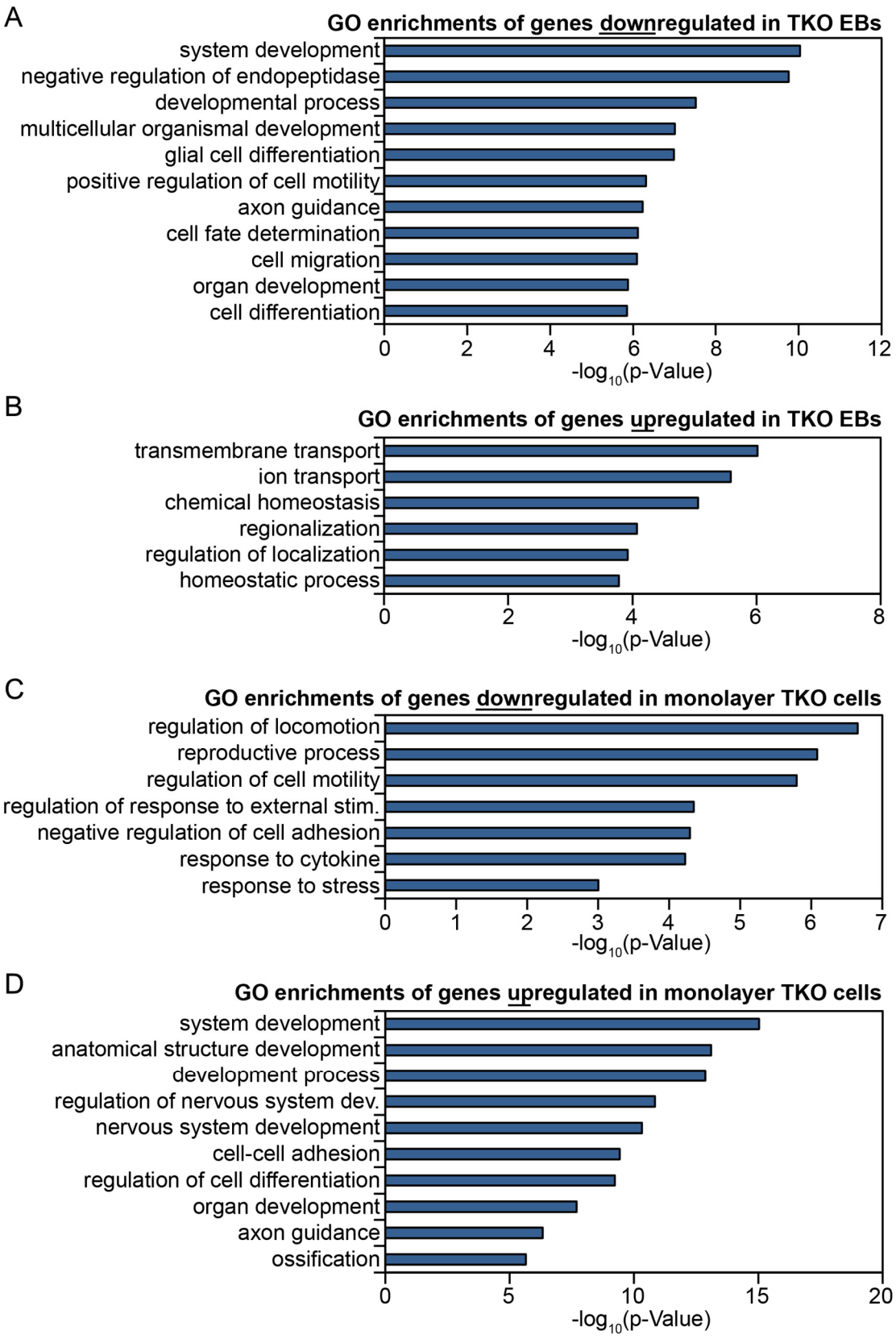

Figure S4: GO term enrichment in *Gadd45*-dependent genes during differentiation

(*A-D*) Gene ontology (GO) biological process terms enriched in (*A*) genes downregulated in *Gadd45* TKO EBs, (B) upregulated in *Gadd45* TKO EBs, (*C*,*D*) down- or upregulated in *Gadd45* TKO monolayer differentiated cells.

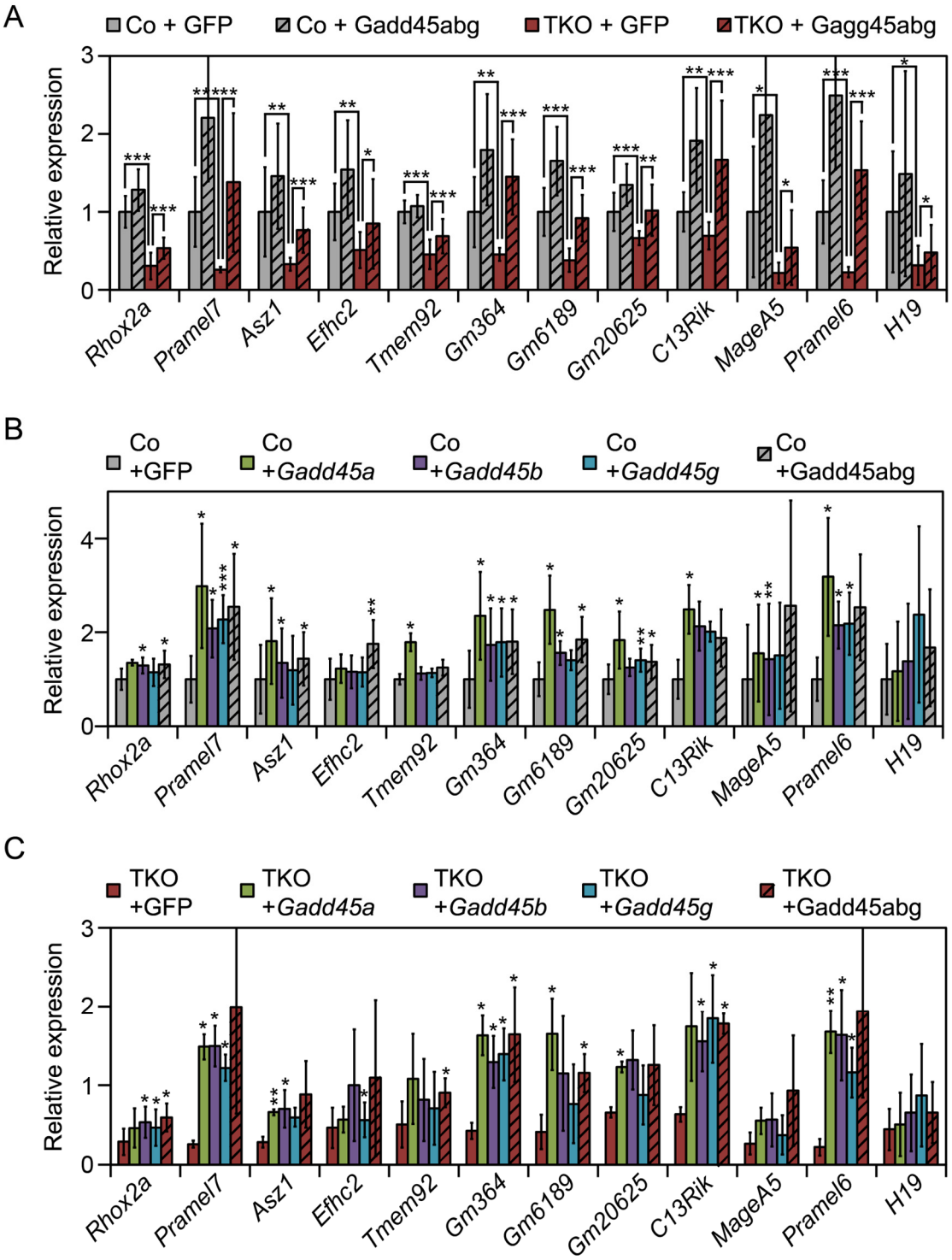

Figure S5: Transient *Gadd45* overexpression rescues gene-downregulation in *Gadd45* TKO-ESCs

(*A-C*) qPCR expression analysis of selected GADD45-dependent genes in Control- (Co) and *Gadd45* TKO-ESCs, after 48h transfection with combined (A) or individual *Gadd45a*, *Gadd45b* and *Gadd45g* (B-C). GFP is used as a control. Expression is relative to GFP transfected Co-ESCs. Data are presented as means ± SD from n = 3 independent clones and n=3 (A) or n=1 (B-C) independent experiments. Statistical significance was tested with two-tailed, unpaired (TKO versus control) or paired (GFP versus *Gadd45* overexpression) Student's t-test.

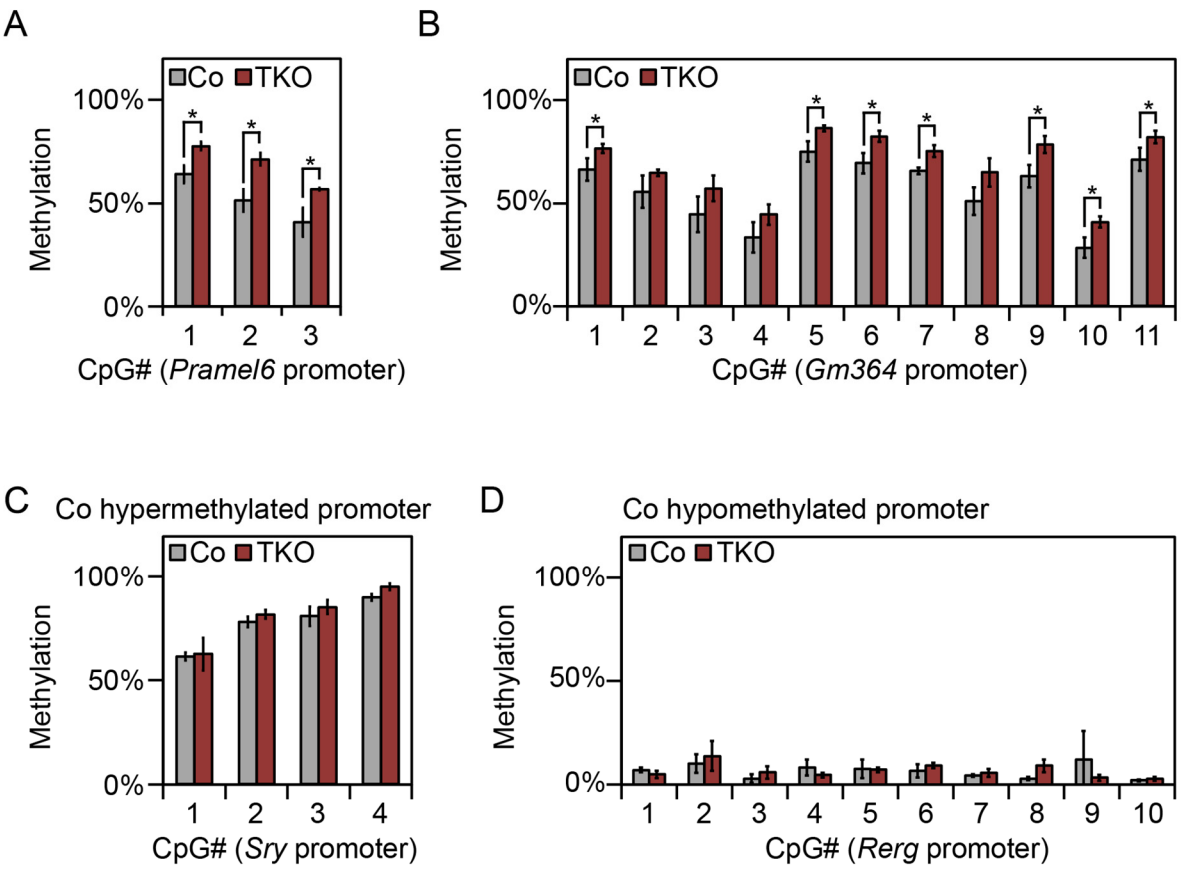

Figure S6: DNA methylation of additional genes in Co and *Gadd45* TKO-ESCs

(*A-B*) Hypermethylation of GADD45-dependent genes. DNA methylation of indicated CpGs in the *Pramel6* promoter (*A*) and the *Gm364* promoter (*B*) in Control- (Co) and *Gadd45* TKO-ESCs measured by site specific bisulfite sequencing.

(*C-D*) Examples of methylation-unaffected genes. GADD45-independent DNA-methylation of indicated CpGs in the hypermethylated *Sry* promoter (*C*) and the hypomethylated *Rerg* promoter (*D*) in Co- and *Gadd45* TKO-ESCs measured by site specific bisulfite sequencing.

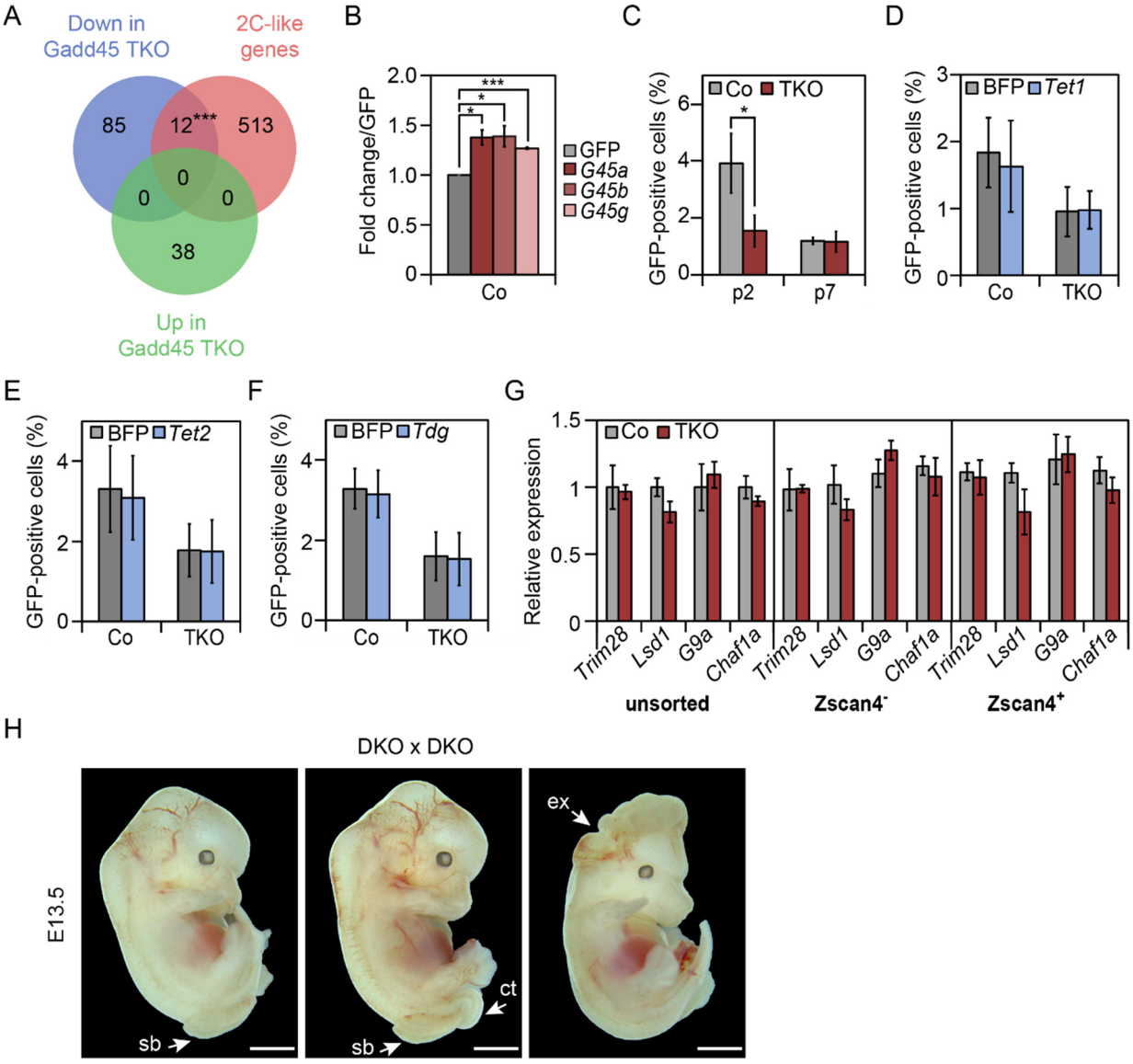

Figure S7:

(*A*) Overlap of genes down- and upregulated in *Gadd45* TKO-ESCs and 2C-like genes (Macfarlan et al. 2012).

(*B*) Flow cytometry analysis of 2C::tdTomato positive cells in Co-ESCs, 48h after transfection with the indicated genes. 2C::tdTomato^{+} cells upon *Gadd45a*, *Gadd45b* or *Gadd45g* are represented relative to 2C::tdTomato^{+} cells upon GFP transfection. Statistical significance was tested with two-tailed, paired Student's t-test.

(*C*) Flow cytometry analysis of Zscan4c::eGFP positive cells in Control- (Co) and *Gadd45* TKO-ESCs treated with 2i-medium at the indicated passage number.

(*D-F*) Flow cytometry analysis of Zscan4c::eGFP positive cells in Co- and *Gadd45* TKO-ESCs, 48h after transfection with the indicated genes.

(*G*) qPCR expression analysis of selected 2C-repressor genes in Co- and *Gadd45* TKO-ESCs in unsorted, Zscan4^{-} or Zscan4^{+} FACS-sorted cells. Expression is relative to the average expression in unsorted Co- ESCs.

(*H*) Images of developmental abnormalities in E13.5 DKO embryos from double knockout crosses showing curly tail (ct), exencephaly (ex) and spina bifida (sb). Scale bar, 2 mm.

Supplemental Material and Methods

RNA-Seq bioinformatic analyses

RNAseq of ESC clones

Sample demultiplexing and FastQ file generation was performed using Illumina's bcl2fastq Conversion Software v1.8.4. The raw sequence reads were assessed with the quality control tool FastQC v.0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and then aligned to the mouse reference genome (NCBIM37/mm9 build from Illumina iGenomes) using TopHat v.1.4.1 (Trapnell et al. 2009) with a GTF transcript annotation file from Illumina iGenomes, Ensembl release 66. The mapped data were quality assessed with RSeQC v.2.3.2 (Wang et al. 2012), sorted with SamTools v.0.1.19 (Li et al. 2009) and summarised on the gene level using the tool htseq-count from the HTSeq package v.0.6.1 (Anders et al. 2015) with stranded option "-s reverse".

Pairwise differential expression analyses were performed using the Bioconductor package edgeR v.3.4.2 (Robinson et al. 2010) following the classic (exactTest) analysis workflow and a 10% FDR (false discovery rate) cutoff. For browser visualisation, the mapped data in BAM format were converted into BedGraph files using BEDTools v.2.16.2 (Quinlan and Hall 2010) and subsequently to reads-per-million-normalized BigWig coverage tracks using the UCSC utility bedGraphToBigWig. Browser tracks were converted between genome assemblies using CrossMap (Zhao et al. 2014).

RNAseq of 2-cell stage embryos

Sample demultiplexing and FastQ file generation was performed using Illumina bcl2fastq v.2.19.1 and overall sequence quality was assessed with FastQC v.0.11.5. The reads were aligned to the mouse reference genome NCBIM37/mm9 with annotation from Gencode release M1 using STAR v.2.5.2b (Dobin et al. 2013) default parameters except '--outFilterMismatchNoverLmax 0.04'. RNA-seq duplication rate analysis using dupRadar (Sayols et al. 2016) revealed very good library complexity for all samples. Mapped read summarization on the gene level was performed using Subread featureCounts v.1.5.1 (Liao et al. 2014) with default parameters. For transposon expression analysis, the raw RNA-seq reads were aligned using Bowtie v.1.2.2 (Langmead et al. 2009) with parameters ‘-n 3 -l 60 -e 150 -M 1 --best --strata’ and reads overlapping with annotated gene exons were removed using BEDTools v.2.27.1. Read summarization on repeat elements (repName, n=1217) was performed using Subread featureCounts v.1.5.1 with parameters ‘-F SAF -M --largestOverlap’ using RepeatMasker mm9 annotation from the UCSC Table Browser (https://genome.ucsc.edu/cgi-bin/hgTables?hgsid=711598791_ANdA08B09Q7lJRUvekj8SPnWmx7f). Differential expression analysis of genes or repeats was performed using the Bioconductor package DESeq2 v.1.18.1 with Independent Hypothesis Weighting (IHW) multiple testing procedure at 10% FDR (Ignatiadis et al. 2016; Love et al. 2014).

Genome-wide methylome analysis

WGBS data sets contained an average of 490 and 479 million paired end reads (151 bp) for two replicates (= two independent clones) of Gadd45 TKO and Co mouse ESCs respectively. Low-quality bases and adapter-containing reads were trimmed from raw data using Trim Galore v0.4.4 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with default parameters. Reads were aligned over the reference mouse genome (NCBI38/mm10) using Bismark v.0.18.0 (Krueger and Andrews 2011) with parameters, '-n 1 -I 0 -X 1000 --score_min L,0,-0.6' resulting in unique mapping rate of 72% on average. Methylation calling for individual cytosines was performed using bismark_methylation_extractor with parameters '-p --ignore 5 --ignore_r2 5 --ample_memory –bedGraph --counts'. Differentially methylated regions (DMRs) between TKO- and Co- ESCs were identified using methylKit 1.3.3 (Akalin et al. 2012) with parameters: (1) Minimum read coverage: 10x per strand, (2) Minimum methylation difference: 30%, (3) Minimum number of consecutive CpGs affected: 2, (4) using a fixed window of 100 bp with (5) false discovery rate <= 5%. Consecutive DMRs were merged. The TKO vs Co scatterplot was generated using methylKit. Gene regulatory regions (i.e. exons, introns, 5UTR, promoter, intergenic, 3UTR) were downloaded from UCSC table browser (https://genome.ucsc.edu/cgi-bin/hgTables?hgsid=690092087_JliNVUb0ssiztYi4Ees7TTAZA3PF).
Fold change enrichment at each regulatory feature was computed as the ratio between the number of hyper-DMRs and the number of random genomic regions. The random set of genomic regions was generated using 'shuffleBed –incl' from BEDTools (Quinlan and Hall 2010). Heatmaps and frequency plots were generated using deepTools v.3.0.1 (Ramírez et al. 2014) with the command 'computeMatrix reference-point -a 5000 -b 5000 --missingDataAsZero –skipZeros'. Methylation differences were calculated by subtracting the average methylation level in 100bp bins between TKO- and Co-ESCs from the center of each feature in +/- 5kb. In this, features within +/- 1 kb of transcription start sites (TSS) were considered as proximal elements and rest as distal elements. We used the following published data sets: 5hmC (Kong et al. 2016), 5hmC, 5fC and 5caC (Shen et al. 2013); TET-dependent DMRs (Lu et al. 2014); Tet1 (Wu et al. 2011); Kdm2aA (Blackledge et al. 2010); Oct4, Sox2, Nanog, p300, Lsd1, RNAPol2, H3K4me1, H4K4me3, H3K79me2, H3K27ac, Hdac2, Rest1, Corest, H3K27me3, Suz12 (Whyte et al. 2012), Nelfa, Spt5 (Rahl et al. 2010); TBP (Kagey et al. 2010); LMR (Stadler et al. 2011). mm9 genome features were converted to mm10 using Batch Coordinate Conversion (liftOver) tool from UCSC Genome Browser Utilities (https://genome.ucsc.edu/cgi-bin/hgLiftOver). Motif analysis was performed using HOMER v3.12 (Heinz et al. 2010) by segmenting the hyper-DMRs at enhancers, promoters and 5mC oxidized products using parameters '-size 25 –len 8'.

Hyper-DMRs were associated with near-by genes using GREAT_v0.3.0 (McLean et al. 2010) with default parameters. All genes that were expressed in at least one sample of ESCs (RNA-seq, this manuscript) were used as background list.

Site specific bisulfite sequencing

Genomic DNA was extracted using the DNeasy Blood & Tissue kit (Qiagen) and bisulfite converted using the Epitect kit (Qiagen) according to the manufacturer's instruction. Regions of interest were amplified by PCR using the AccuPrime kit (Thermo Fisher). An aliquot of each PCR reaction was run on a 1% agarose gel and band intensity was quantified with the ImageLab 3.0 software (BioRad). For each biological replicate amplicons were mixed in equimolar amounts to yield 2µg total DNA and purified using the PCR purification kit (Qiagen) according to the manufacturer's instructions. Amplicon libraries were prepared from 100ng input DNA, using NuGEN's Amplicon Sequencing System (NuGEN, 9092-256), with primers for dual indexing, from the Ovation Low Complexity kit (NuGEN, 9092-16). Libraries were pooled together with a genomic DNA library (to increase pool complexity for sequencing), in equimolar ratio. 10pmol of the resulting pool was spiked in with phiX control library (10%, Illumina) and sequenced on a MiSeq (Illumina), paired-end, for 2x 251 cycles plus 2x 8 cycles for index reads. Read quality was evaluated using FastQC and bases with Phred scores <30 were trimmed using Trim Galore with standard parameters. Reads were aligned to the reference amplicons using the mapping algorithm of Bismark. The methylation extractor and bedGraph2cytosine algorithms of Bismark were then used to calculate methylation levels for each CpG site.

Immunoblotting and immunoprecipitation

Cells were lysed (150mM NaCl, 20mM Tris pH 7.5, 2mM EDTA, 10% glycerol, 1% triton-x and Complete Mini Protease Inhibitor Cocktail (Roche)) and disrupted by sonication for 20 cycles 15 sec on/off, high setting, in a Bioruptor (Diagenode). For immunoblotting, lysates were cleared by centrifugation and protein concentrations were estimated by bicinchoninic acid (BCA) assay using bovine serum albumine samples with known concentrations as standard. 50µg of cell lysate were adjusted to equal volumes, mixed with NuPAGE® LDS Sample Buffer (4X) (Novex®) and boiled for 10min at 70°C. Protein samples were separated on a polyacrylamid gel and transferred to a polyvinylidene difluoride (PVDF) membrane. Membrane was blocked with 5% skimmed milk in Tris buffered saline with 0.1% Tween (TBS-T) for 1 hour and incubated with primary antibody diluted in blocking buffer at 4°C overnight. Membrane was washed three times with TBS-T before incubated with the corresponding secondary antibodys diluted in blocking solution for 1 hour at room temperature. Signals were developed with SuperSignal West Pico or Femto Chemiluminescent Substrate (Thermo Scientific) and analyzed using a ChemiDoc (BioRad) with ImageLab software.

For immunoprecipitation, cells were lysed using the same buffer described above. Cell disruption was done by sonication for 28 cycles 15 sec on/off, high setting in a Bioruptor (Diagenode). Cell lysates were cleared using centrifugation and incubated on a rotating wheel overnight at 4°C with specific antibody. Agarose G beads (Roche) were added and incubated for 2h at 4°C on a rotating wheel. Beads were washed 3 times in lysis buffer and protein was eluted with NuPAGE® LDS Sample Buffer (1X) (Novex®) with freshly added DTT (100Mm final concentration).

Antibodies

Primary antibodies used in this study were rabbit anti GADD45α (Santa Cruz, sc-797), mouse anti α Tubulin (Sigma, T5186), goat anti GADD45β (Santa Cruz, sc-8776), guniea pig anti GADD45γ (homemade), mouse anti c-Myc (Sigma, M4439). Secondary antibodies were HRP-coupled goat anti mouse IgG (Dianova, 115-035-146), HRP-coupled goat anti rabbit IgG (Dianova, 111-035-144), HRP-coupled goat anti guinea pig IgG (Dianova, 106-035-003), HRP-coupled donkey anti goat IgG (Dianova, 705-035-003).

Mass spectrometry sample preparation

Samples were boiled in 1x NuPAGE LDS Sample Buffer (Life technologies) containing 100mM DTT at 70°C for 10 minutes. Samples were separated on a 10% NuPAGE Bis-Tris gel (Life technologies) for 10 at 180V in MES running buffer (Life technologies) and fixed in 7% acetic acid containing 40% methanol and subsequently stained for 30 minutes using Colloidal Blue staining kit (Life technologies). Protein lane was excised from the gel, chopped and destained (50% ethanol in 25 mM ammmonium bicarbonate) at room temperature for 15 min and dehydrated (100% acetonitrile) for 10 minutes on a rotating wheel. Samples were vacuum dried and subsequently rehydrated and reduced for 60 minutes in reduction buffer (10mM DTT in 50mM NH4HCO3 pH 8.0) at 56°C. Samples were then alkylated in 50 mM iodoacetamide in 50mM ammmonium bicarbonate pH 8.0 at room temperature for 45 minutes in the dark. Dehydrated and vacuum dried samples were trypsin digested (1µg trypsin/sample in 50mM Triethylammonium bicarbonate buffer pH 8.0) at 37°C overnight. This was followed by stepwise peptide extraction: two times extraction solution (30% acetonitrile) and 100% acetonitrile at 25°C shaking at 1400 rpm for 15 minutes. Peptides were purified and desalted using C18 stage tips (Rappsilber et al. 2007). 3.5 µL peptides were loaded and separated on C18 column (New Objective) with 75 µm inner diameter self-packed with 1.9µm Reprosil beads (Dr. Maisch) which was mounted to an EasyLC1000 HPLC (Thermo).

Mass spectrometry measurement and data analysis

Peptides were separatd using reversed-phase chromatography gradient (Buffer A: 0.1% formic acid, Buffer B: 80% acetonitrile and 0.1% formic acid, Gradient: 0-67 min 0-22% Buffer B, 67-88 min 22-40% Buffer B, 89-92 min 40-95% Buffer B). Eluted peptides were directly sprayed into a Q Exactive Plus mass spectrometer from Thermo operating in positive scan mode with a full scan resolution of 70,000; AGC target 3x10^6; max IT = 20ms; Scan range 300 - 1650 m/z and a Top10 MSMS method. Normalized collision energy was set to 25 and MSMS scan mode operated with resolution of 17,000; AGC target 1x10^5; max IT = 120 ms.

Database search was performed using MaxQuant Version 1.5.2.8 (Cox and Mann 2008) against Mus Musculus Ensembl database (downloaded 23th January 2017, 89063 entries), with Trypsin/P as digestion enzyme allowing 2 missed cleavages. Following settings were applied: variable modification:

Acetyl (Protein N-term); Oxidation (M), fixed modifications: Carbamidomethyl (C), FDR of 1% on peptide and protein level was applied.

Proteins with at least two peptides (one of them unique) were considered as identified. Proteins matching reverse database or common contamination list as well as proteins with peptides only identified by peptides with modification were filtered out. Further bioinformatics analysis was conducted in R (https://www.R-project.org/) using existing libraries (knitr, imbproteomicsr, plyr, reshape, ggplot2 and psych).

qPCR (expression) primer

Gene / Repeats	Forward primer	Reverse primer	UPL Probe
0610005C13Rik	ctatgacaccgcttgacacc	caaaatcgctgtggattgg	13
Asz1	agttgcctgtaaattctcacaagat	ccagttcttcacaaactgaagtaaaat	7
Cdx2	caccatcaggaggaaaagtga	ctgcggttctgaaaccaaat	34
Chaf1a	catgtccaaacctagcaacg	tcactcatgcggaaatgct	75
Dux	ccactggattggggtagaaa	ggtatccatgcccgaattta	69
Efhc2	gtttgagcctatagagaataattcagg	caggctttttcacacgacttc	31
Elf5	gactgtcacagccgaacaag	ccaggatgccacagttctct	56
G9a	cccctgatgagccagaac	aagcaactgggccaggata	79
Gadd45a	gctgccaagctgctcaac	tcgtcgtcttcgtcagca	98
Gadd45b	cggccaaactgatgaatgt	atctgcagagcgatatcatcc	79
Gadd45g	gtccgccaaagtcctgaat	gctatgtcgccctcatcttc	71
Gapdh	agcttgtcatcaacgggaag	tttgatgttagtggggtctcg	9
Gm20625	cacagctgcgactgaacaat	gatagggcctcagcacctg	66
Gm364	tgctatgctattttcggctgt	gtcaaaaaggctctccacca	73
Gm6189	cagctgggaaaccctgaa	tcctggttgtagtgcagacg	67
H19	cggtgtgatggagaggaca	agacggcttctacgacaagg	42
IAP	tcaaggacagggtattgttg	tcgggtgagtctttctggtac	-
Lsd1	ttctggagggtatggagacg	gcttcctgagaggtcattcg	60
MageA5	tgctgtaccctcattggaca	cccaaacaaagcgaaaattg	25
mERVL	tttctcaaggcccaccaatagt	gacacctttttaactatgcgagc	-
Nanog	gcctccagcagatgcaag	ggttttgaaaccaggtcttaacc	25
Oct4	aatgccgtgaagttggagaa	ccttctgcagggctttcat	95
Pramel6	ccaactttggcagtctccag	ccgtaattgacaaggtcttcaaa	27
Pramel7	gatcatttctctgttctgcaaactc	aatggggcctccagacac	94
Prl7d1	tgatccaaccgtgctcct	ctggcatttatgggtgcag	7

Psg29	gcagggggttctactcacag	caatggtgactttggaagtgg	62
Rhox2a	gggagtgagtgaagccacag	gcctttacagcctcttataactcct	91
Serpinb9e	caccaaggaagtgcccttta	ttcacataggcatggaaaaatg	4
Sox2	gacgtcgtagcggtgcat	acggcagctacagcatga	68
Sp110	ccgggacaattccttcatc	attgtgcaccactttggaca	93
Tbp	ccaggaaataattctggctca	ggggagctgtgatgtgaagt	97
Tcstv1	gaaccatccatcctcaggaac	ccctgaaggtaaatcctccac	55
Tmem92	gagagcatttgggtcacctc	tgtgggaggtgtcatctgaa	77
Tpbpa	tgaagagctgaaccactgga	caggcataggatgactaggaagat	107
Trim28	caaccagcggaaatgtgag	ggtaccaccaggctgctc	2
Ubtfl1	agctcctcctgcacaatttc	caaggcaagcagttgaagaa	42
Xlr5a	tgggatcttcggaggtttag	gaaccttcctgggtcagaca	80
Zfp352	acaagactactgaacttttcatcttca	gggaaggctgtgtattgcag	20
Zscan4c	gactgaactatctaacatcctcagca	ttgcaacattcttctctctttga	41

PCR (genotyping) primer

Genotyping PCR	Forward primer	Reverse primer
CRISPR *Gadd45a* external	cgctccgtaccctatcacaa	aactctgccttgctttggtg
CRISPR *Gadd45b* external	catctccagccaatctcagc	caagacccatcgtgcatcag
CRISPR *Gadd45g* external	aaatctgcaggctccagtct	cacgcggggcctttctac
CRISPR *Gadd45a* internal	ggccacttacacgttgagc	ggtaagtgtgcctgccga
CRISPR *Gadd45b* internal	tcatgaccctggaagagctg	acattcatcagtttggccgc
CRISPR *Gadd45g* internal	cgccagatcatgactagggt	cagaagttcgtgcagtgctt

PCR (bisulfite converted template amplification) primer

Associated gene	Genomic element	Forward primer	Reverse primer
Gm364	Promoter	tttaaaagaaagattttatgaggttgaa	caaataaaccaccaaaaaataaaaca
Pramel6	Promoter	tttaagaatttggattaaaggtgtg	tttaaacactaacacaaaaaaaacc
Rerg	Promoter	agaaaggatttagttgaagtttgtaggt	aactaaaaaactctcccaccctaac
Rhox2a	Promoter	aattaaataaggtaggaggatttag	atcccacatcaaacttataattaac
Sry	Promoter	tttttgtttgtttgtttgtttttgt	accactcctctataacactttaaccc

gRNA sequences

Target	Reverse off-target frequency	Guide RNA Sequence	PAM Motif
Gadd45a Exon1	99	GCCCCCACAGTGCGCGGCGC	TGG
Gadd45a Intron 2	91	GGTGAGTTGCATACGCTCAG	GGG
Gadd45b Exon 1	90	GAAGTCCCACCGCCTCCGGA	AGG
Gadd45b Intron 2	86	GCACCCCCCTTTCTCGGGCG	TGG
Gadd45g Exon 1	86	GTTATCCAAAAGAGTGCGGA	GGG
Gadd45g Intron 2	92	GTAGAAAGGCCCCGCGTGCA	GGG

Supplemental References

Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. 2012. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles, Genome Biol 13: R87.

Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data, Bioinformatics 31: 166–169.

Blackledge NP, Zhou JC, Tolstorukov MY, Farcas AM, Park PJ, Klose RJ. 2010. CpG islands recruit a histone H3 lysine 36 demethylase, Mol Cell 38: 179–190.

Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, Nat Biotechnol 26: 1367–1372.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner, Bioinformatics 29: 15–21.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities, Mol Cell 38: 576–589.

Ignatiadis N, Klaus B, Zaugg JB, Huber W. 2016. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing, Nat Methods 13: 577–580.

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture, Nature 467: 430–435.

Kong L, Tan L, Lv R, Shi Z, Xiong L, Wu F, Rabidou K, Smith M, He C, Zhang L, et al. 2016. A primary role of TET proteins in establishment and maintenance of De Novo bivalency at CpG islands, Nucleic Acids Res 44: 8682–8692.

Krueger F, Andrews SR. 2011. Bismark. A flexible aligner and methylation caller for Bisulfite-Seq applications, Bioinformatics 27: 1571–1572.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, Genome Biol 10: R25.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools, Bioinformatics 25: 2078–2079.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features, Bioinformatics 30: 923–930.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol 15: 550.

Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. 2014. Role of Tet proteins in enhancer activity and telomere elongation, Genes Dev 28: 2103–2119.

Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity, Nature 487: 57–63.

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions, Nat Biotechnol 28: 495–501.

Quinlan AR, Hall IM. 2010. BEDTools. A flexible suite of utilities for comparing genomic features, Bioinformatics 26: 841–842.

Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA. 2010. c-Myc regulates transcriptional pause release, Cell 141: 432–445.

Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data, Nucleic Acids Res 42: W187-91.

Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips, Nat Protoc 2: 1896–1906.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics 26: 139–140.

Sayols S, Scherzinger D, Klein H. 2016. dupRadar: a Bioconductor package for the assessment of PCR artifacts in RNA-Seq data, BMC Bioinformatics 17: 428.

Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung H-L, Zhang K, Zhang Y. 2013. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics, Cell 153: 692–706.

Song C-X, Szulwach KE, Dai Q, Fu Y, Mao S-Q, Lin L, Street C, Li Y, Poidevin M, Wu H, et al. 2013. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming, Cell 153: 678–691.

Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, et al. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions, Nature 480: 490–495.

Trapnell C, Pachter L, Salzberg SL. 2009. TopHat. Discovering splice junctions with RNA-Seq, Bioinformatics 25: 1105–1111.

Wang L, Wang S, Li W. 2012. RSeQC. Quality control of RNA-seq experiments, Bioinformatics 28: 2184–2185.

Whyte WA, Bilodeau S, Orlando DA, Hoke HA, Frampton GM, Foster CT, Cowley SM, Young RA. 2012. Enhancer decommissioning by LSD1 during embryonic stem cell differentiation, Nature 482: 221–225.

Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y. 2011. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells, Nature 473: 389–393.

Zhao H, Sun Z, Wang J, Huang H, Kocher J-P, Wang L. 2014. CrossMap: a versatile tool for coordinate conversion between genome assemblies, Bioinformatics 30: 1006–1007.

4.2 Chapter 2: Role of NEIL1 and NEIL2 DNA glycosylases during differentiation

4.2.1 Introduction

DNA methylation is a key epigenetic silencing mark modulating lineage commitment during early mouse development (Kim and Costello 2017). Upon implantation, pluripotency genes require silencing whereas lineage specific expression is activated. DNA methylating and demethylating factors can fulfill both functions by establishing and removing methylation marks. *Dnmt* and *Tet* knockout studies demonstrated the importance of dynamic rather than stable methylation levels. *Dnmt3a/b* and *Tet1/2/3* deficient mice bear postimplantation defects causing lethality at E11.5 in *Dnmt3a/b* double knockout mice (Dai et al. 2016; Okano et al. 1999). Do other factors involved in DNA (de)methylation reveal these phenotypes as well? The deficiency of the base excision repair factor *Tdg*, which has been involved in the active DNA demethylation pathway, shows a drastic embryonic lethal phenotype (Cortázar et al. 2011). Since this lethality is not consistent with phenotypes of mice deficient of any other DNA glycosylase family member, the authors suggested that it is caused by methylation misregulated gene-expression. However, it is not clear to date whether the involvement of TDG in BER or in DNA demethylation ultimately causes this striking phenotype.

The base excision repair DNA glycosylases NEIL1 and NEIL2 are newly discovered regulators of TDG in active DNA demethylation (Schomacher et al. 2016). TDG recognizes and excises 5fC and 5caC generating an AP site. TDG stalls at the AP site and is product-inhibited. NEIL1 and NEIL2 possess a higher affinity towards AP-sites, displace TDG and thereby promote substrate turnover of TDG. Although NEIL1 and NEIL2 are hereby implicated in DNA demethylation, knockout studies in mice do not show any severe developmental defects as observed for other DNA (de)methylating factors (Canugovi et al. 2012; Chakraborty et al. 2015; Vartanian et al. 2006). In contrast, knockdown of *neil2* in *Xenopus* causes a severe neural crest defect. In line with this, *in vitro* differentiated mESCs deficient for *Neil1* or *Neil2* show a neuronal/neural crest defect (unpublished data). How can we explain such discrepancy between *in vitro* and *in vivo* differentiation? Can the cause of the differentiation defect be attributed to their role in DNA demethylation or in base excision repair? *In vivo*, phenotypes of *Neil*-deficient mice are similar to the mild phenotypes caused by other DNA glycosylases suggesting the biological importance of their role in base excision repair. However, the differentiation defects of mESCs *in vitro* suggest for an involvement in DNA demethylation.

4.2.2 Results

4.2.2.1 Neil1 and Neil2 deficient mESCs show a neuronal/neural crest differentiation defect

To investigate the role of the novel DNA demethylation factors NEIL1 and NEIL2 in pluripotency and differentiation, *Neil1* and *Neil2* single knockout (SKO) mESCs were generated using the CRISPR/Cas9 system ([redacted], unpublished data). Three clones were generated for each deletion covering the coding region of the catalytic domain. To obtain three independent Control (Co) clones, mESCs were subjected to the CRISPR/Cas9 process without specific gRNAs. *Neil1* and *Neil2* deficient mESCs showed normal growth behavior (Figure 4.1A) and morphology (data not shown) when compared to Co mESCs. Expression of pluripotency markers (*Oct4*, *Nanog* and *Sox2*) was not altered (Figure 4.1B). These results showed that maintenance of pluripotency was not affected by loss of *Neil1* or *Neil2*, which is in line with sustained pluripotent capacity in *Dnmt*, *Tet* and *Tdg* deficient mESCs (Cortázar et al. 2011; Dawlaty et al. 2014; Tsumura et al. 2006).

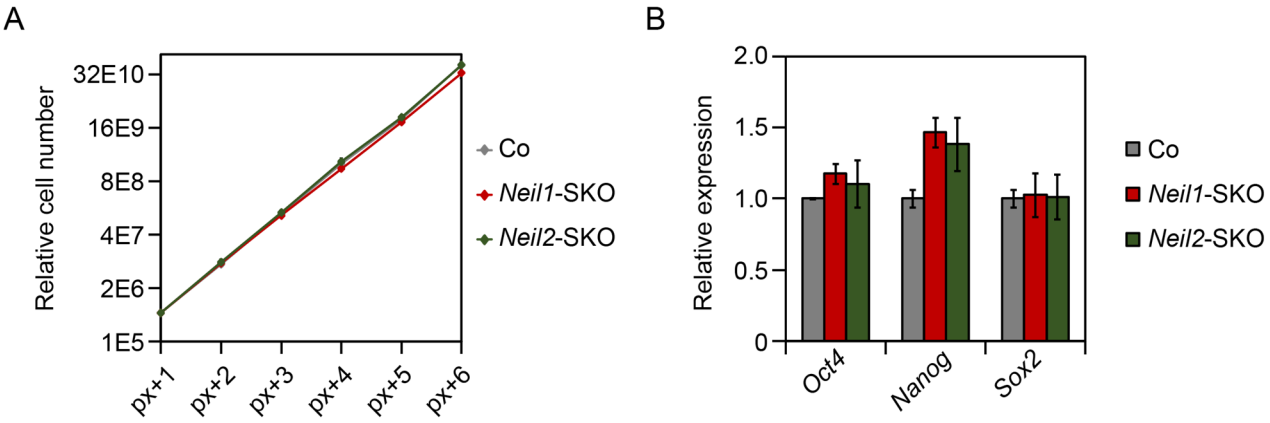

Figure 4.1: *Neil1* and *Neil2* deficient mESCs maintain pluripotency

(A) Growth curve of *Neil1* and *Neil2* deficient (*Neil1- and Neil2*-SKO) and control (Co) mESCs counted for six passages (p). (B) Relative expression of pluripotency markers *Oct4*, *Nanog* and *Sox2* in *Neil1* and *Neil2* deficient and Co mESCs measured by qPCR. Expression levels are relative to average expression in Co mESCs.

Although an aberrant methylation landscape does not affect pluripotency, mESCs deficient for DNA (de)methylating factors show impaired differentiation potential (Dawlaty et al. 2014; Okano et al. 1999). We therefore sought to investigate the importance of NEIL1 and NEIL2 in defining lineage commitment using an *in vitro* differentiation assay. *Neil1* and *Neil2* deficient mESCs were subjected to embryoid body (EB) differentiation for 8 days. EBs were differentiated in the absence and presence of retinoic acid (RA) beginning from day 4 on, the latter one favoring neural lineage development (Bibel et al. 2007). Surprisingly, *Neil1* and *Neil2* deficient mESCs showed a striking and significant neuronal/neural crest defect upon retinoic acid addition (*Neurog1*, *Pax6* and *Pax3*), while mesodermal differentiation (*Eomes*) was negligible affected (Figure 4.2). Endodermal marker expression (*Gata6*) was reduced in *Neil1* and *Neil2* deficient EBs. However, this reduction was quite inconsistent between different EB differentiations probably caused by variations in the undirected EB protocol. These results confirm the neuronal/neural crest defect as observed by [redacted].

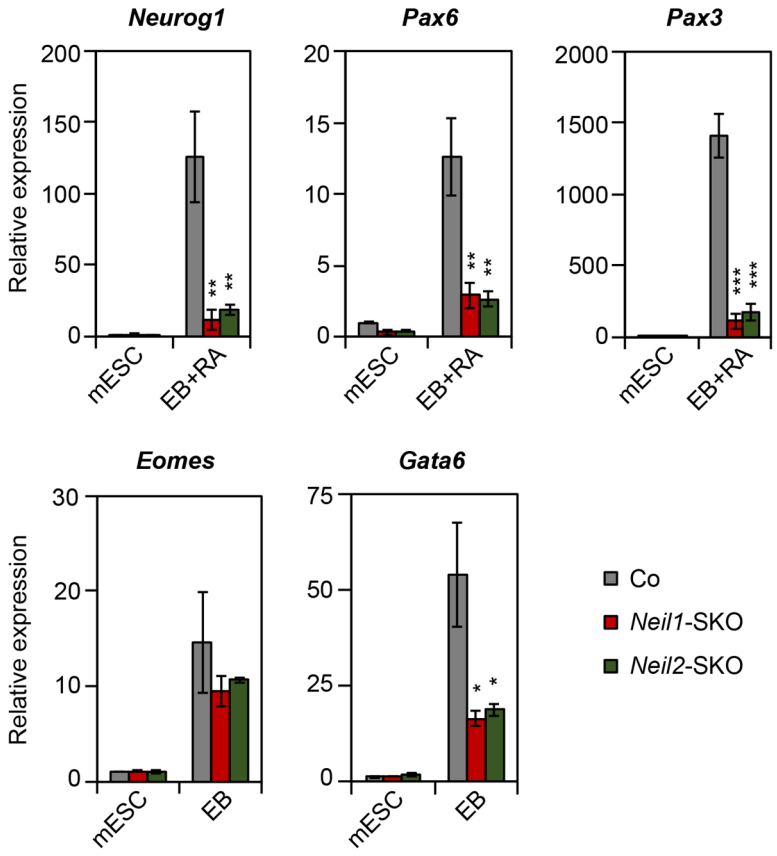

Figure 4.2: *Neil1* and *Neil2* deficient mESCs show a neuronal/neural crest differentiation defect.

Expression analysis of neuronal/neural crest (*Neurog1*, *Pax6, Pax3*), mesodermal (*Eomes*) and endodermal (*Gata6*) marker in *Neil1* and *Neil2* deficient (*Neil1- and Neil2*-SKO) and control (Co) mESCs measured by qPCR. Expression levels are relative to average expression in Co mESCs.

This result is in line with the finding that *Tdg* deficient mESCs are impaired in terminal neuronal differentiation (Cortázar et al. 2011). It is important to note that *Tdg* KO mESCs differentiate normal in EB (+RA) conditions ([redacted], unpublished data). Therefore, *Neil1* and *Neil2* deficient mESCs show a more severe phenotype *in vitro* than *Tdg* deficient mESCs. The severity of the phenotype is reflected in *Xenopus* upon *neil2* knockdown, but here restricted to neural crest (Schomacher et al. 2016). Interestingly, 5fC and 5caC accumulate upon knockdown of *neil2* in *Xenopus* indicating the physiological importance of Neil2 promoting the substrate turnover of TDG. Global 5fC and 5caC levels were measured in undifferentiated and EB (+RA) differentiated *Neil1* and *Neil2* deficient and Co mESCs using liquid chromatography-tandem mass spectrometry (LC-MS/MS) ([redacted] and [redacted], unpublished data). Notably, there was no accumulation of 5fC or 5caC suggesting that if anything, NEIL1 and NEIL2 promoting TDG´s turnover has a local-specific rather than global effect in mESCs.

4.2.2.2 The cause of the neuronal/neural crest differentiation defect – rescue experiments

Neil1 and *Neil2* deficient mESCs showed a drastic neuronal/neural crest differentiation defect upon EB+RA differentiation. To understand the molecular mechanism of this differentiation defect, *Neil1* and *Neil2* deficient mESCs were subjected to RNA-sequencing (RNA-seq) analysis in undifferentiated, EB and EB+RA differentiated state ([redacted] and [redacted], unpublished data). The number of differentially expressed genes increased remarkably upon differentiation. Pathway analysis of down- and upregulated genes did not show any enrichment hits in undifferentiated conditions, but neural crest was the top hit for downregulated genes in *Neil1* and *Neil2* deficient EBs and EBs+RA confirming our qPCR results. Moreover, WNT-signaling appeared as one of the top hits upon differentiation. Tightly regulated WNT-signaling is key to maintain pluripotency and to mediate differentiation (Sokol 2011). I performed various rescue experiment manipulating WNT-signaling (i.e. addition of WNT3a-conditioned medium, GSK3β-inhibition or AXIN stabilization) throughout or in specific time-intervals during differentiation, but none of these conditions could rescue the neuronal/neural crest defect (data not shown). This suggests that deregulation of factors of the highly controlled WNT-signaling upon differentiation is most likely a consequence and not the cause of the differentiation defect.

LC-MS/MS analysis of differentiated *Neil1* and *Neil2* deficient mESCs revealed that there is no global effect of NEIL1 and NEIL2 in promoting substrate turnover of TDG upon mESC differentiation. However, *Neil1* and *Neil2* deficiency might have a site-specific effect thereby causing the differentiation defect. To test this, I cultured *Neil1* and *Neil2* deficient and Co mESCs in 2i medium to induce global DNA hypomethylation and thereby preventing the accumulation of 5fC and 5caC (Ficz et al. 2013; Leitch et al. 2013). If NEIL1 and NEIL2 are required to mediate substrate turnover of 5fC and 5caC and are therefore required for DNA demethylation at specific sites, these sites become already demethylated upon 2i-treatment. 2i medium contains two inhibitors, one of them enhancing WNT-signaling via GSK3β inhibition. To evaluate the attribution of the hypomethylating properties of 2i medium, I additionally treated mESCs with 1i medium only containing GSK3βi. If both media compositions rescue the phenotype, it is not attributable to hypomethylation.

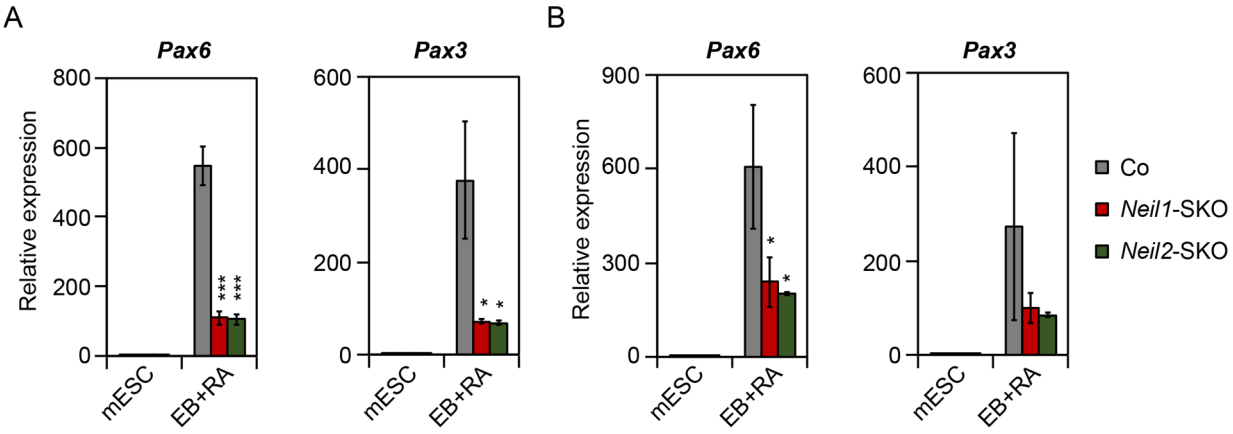

Figure 4.3: Global hypomethylation does not rescue the differentiation defect

(A-B) Expression analysis of neuronal (*Pax6*) and neural crest (*Pax3*) marker in *Neil1* and *Neil2* deficient (*Neil1*- and *Neil2*-SKO) and control (Co) mESCs and upon differentiation measured by qPCR. Expression levels are relative to average expression in Co mESCs. mESCs were cultured in 1i (A) or 2i (B) medium before *in vitro* differentiation.

Pre-treatment of *Neil1* and *Neil2* deficient mESCs with 2i or 1i medium did not increase the differentiation potential (Figure 4.3). Although the differentiation defect was less significant in 2i compared to 1i pre-treated *Neil1* and *Neil2* deficient mESCs, the effects were minor. This suggests that accumulation of 5fC and 5caC has no effect on neuronal/neural crest marker expression at this state of differentiation.

DNA glycosylases are involved in DNA repair mechanisms to secure genome stability (Jacobs and Schär 2012). It is essential to ensure the functionality of DNA repair processes during differentiation. Therefore, irrespective of promoting DNA demethylation, NEIL DNA glycosylases might be essential for the repair of oxidative lesions during differentiation. Standard embryonic stem cell cultures are performed at 20% O_2 (equivalent to 142 mmHg partial pressure of oxygen (pO_2)), whereas the developing embryo *in vivo* encounters oxygen levels between 0-30 mmHg (Millman et al. 2009; Powers et al. 2008). Consequently, mESCs differentiated *in vitro* experience more oxidative stress due to reactive oxygen species (ROS) generation at high O_2 levels (Jagannathan et al. 2016). In order to reduce the DNA damage potentially induced by high oxygen levels, EB (+RA) differentiation was performed in low oxygen conditions (5% O_2, Figure 4.4). However, low oxygen conditions did not rescue the neuronal/neural crest defect in *Neil1* and *Neil2* deficient mESCs. In contrary, low oxygen conditions impaired the differentiation potential dramatically even in Co mESCs, except for endodermal marker expression, which showed a ~10-fold increase (compare Figure 4.4 with Figure 4.2).

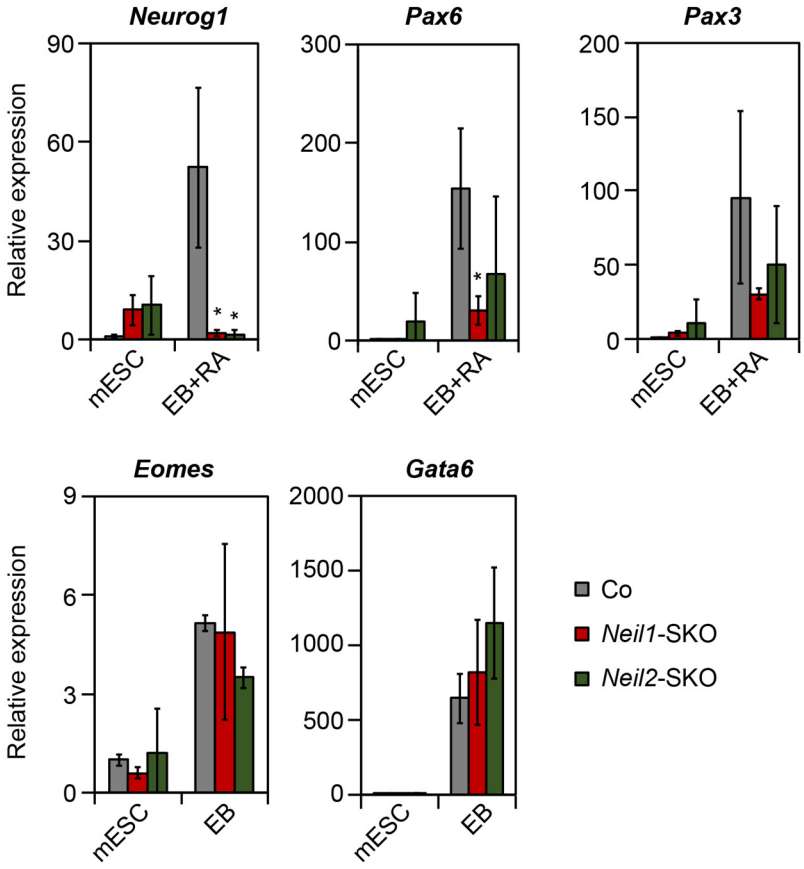

Figure 4.4: Reduced oxygen conditions do not rescue differentiation defect

Expression analysis of neuronal/neural crest (*Neurog1*, *Pax6, Pax3*), mesodermal (*Eomes*) and endodermal (*Gata6*) marker in *Neil1* and *Neil2* deficient (*Neil1*- and *Neil2*-SKO) and control (Co) mESCs and upon differentiation in reduced oxygen conditions (5% O_2) measured by qPCR. Expression levels are relative to average expression in Co mESCs.

In addition to high pO_2, standard mESC cultures utilize high glucose containing medium. High glucose conditions have been shown to increase ROS (Aguiari et al. 2008; Subramaniyan and Natarajan 2017) and might therefore trigger the neuronal/neural crest differentiation defect. Nevertheless, low glucose culture conditions had no positive impact on the differentiation ability of *Neil1* and *Neil2* deficient mESCs (data not shown).

Upon differentiation ROS levels increase and have been shown to be an important cell signaling molecule (Tormos et al. 2011). However, excess of ROS upon differentiation can cause DNA lesions, which might require NEIL1 and NEIL2 for repair and genome stability. Interestingly, neuronal development in particular is associated with an increase in mitochondrial number and therefore increase in oxidative stress (Khacho and Slack 2018). Consequently, I attempted to rescue the neuronal/neural crest defect by differentiating *Neil1* and *Neil2* deficient mESCs in presence of various antioxidants (Vitamine E, C, D and n-acetylcysteine (NAC) (Clarkson and Thompson 2000; Kerksick and Willoughby 2005)). However, none of the applied antioxidants could rescue the differentiation defect, although Vitamin C and Vitamin D addition had a promoting effect on neuronal/neural crest marker

expression in Co mESCs (summarized in Table 1). Surprisingly, NAC addition mimicked the *Neil1* and *Neil2* deficiency.

Treatment	Neuronal/ Neural crest	Mesodermal	Endodermal
Vitamine E	-	-	-
Vitamine C	+	-	-
Vitamine D	+	-	-
NAC	X	-	-

Table 1: Summary rescue experiments affecting ROS levels

Different conditions either had no effect (-), inhibiting effects even in control mESCs (X) or enhancing effects only in control mESCs (+). NAC, n-acetylcysteine.

4.2.2.3 Neil1 and Neil2 deficient mESCs show an activated p53 response upon neuronal differentiation

Rescue experiments were not sufficient to trace back the cause of the differentiation defect in *Neil1* or *Neil2* deficient mESCs. Interestingly, knockdown of *neil2* in *Xenopus* shows a similar phenotype, a specific neural crest defect. RNA-seq analysis of *Xenopus* stage 23 embryos upon *neil2* knockdown revealed Tp53 pathway as top hit for upregulated genes in a pathway enrichment analysis ([redacted] and [redacted], unpublished data). Expression analysis of selected p53 target genes (Fischer 2017) in undifferentiated and differentiated *Neil1* and *Neil2* deficient mESCs showed a significant increase in p53 targets including regulators of cell cycle (*p21*), feedback (*Mdm2*, *Ccng1*), translation (*Sesn2*) or apopotosis (*Perp1*) upon differentiation, most strikingly upon neuronal differentiation (EB+RA) (Figure 4.5). The result indicates the occurrence of a p53 response in *Neil1* and *Neil2* deficient mESCs as observed in *Xenopus* Neil2 morphants. In line, p53 inhibition upon differentiation rescued the neuronal/neural crest phenotype ([redacted], unpublished data).

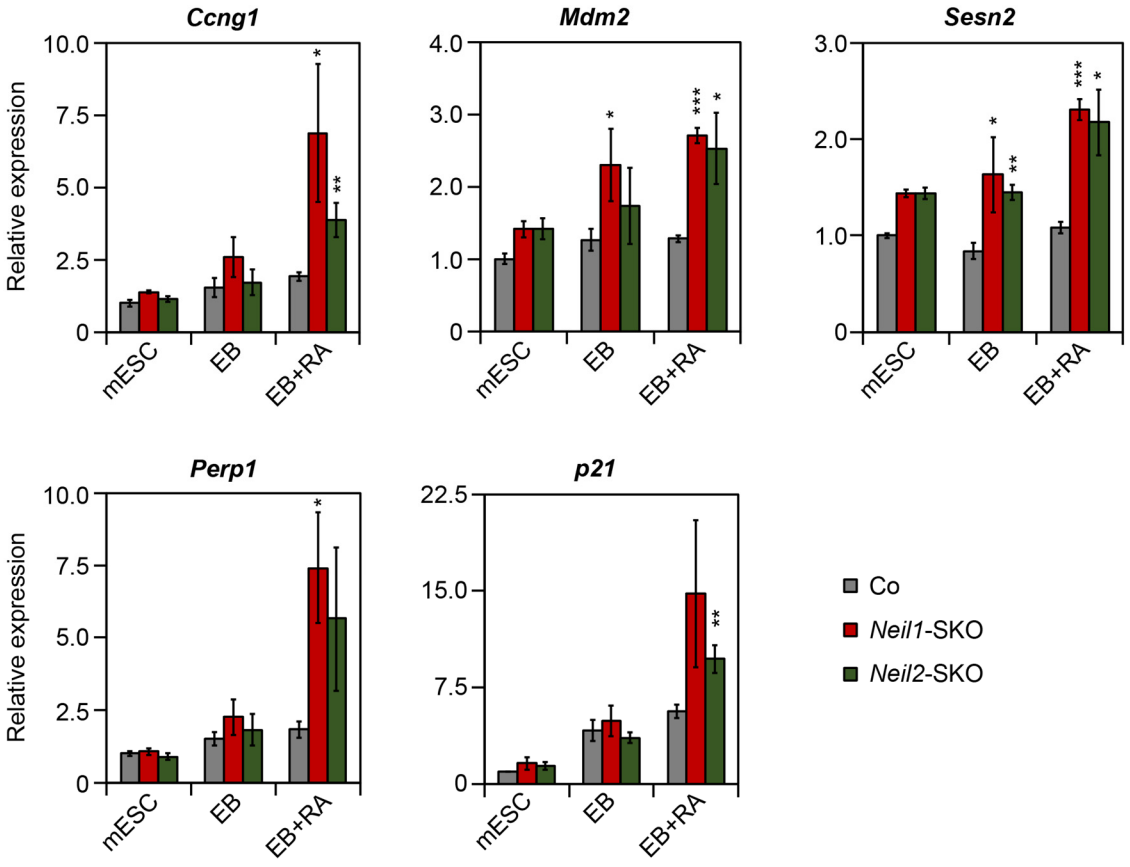

Figure 4.5: p53 target genes are upregulated in *Neil1* and *Neil2* deficient mESCs upon RA-mediated differentiation

Expression analysis of selected p53 target genes in *Neil1* and *Neil2* deficient (*Neil1*- and *Neil2*-SKO) and control (Co) mESCs and upon differentiation measured by qPCR. Expression levels are relative to average expression in Co mESCs.

Although *Trp53* is expressed in mouse embryonic stem cells and upon differentiation (Jain and Barton 2018), I was unable to detect p53 by Western blot analysis in mESCs or EBs (+RA) using different antibodies (data not shown). I therefore investigated the downstream effector pathways of activated p53, in particular cell cycle arrest and induction of apoptosis. CYCLIN D1 and cleaved CASPASE-3 levels were analyzed for p53-mediated cell cycle arrest or induction of apoptosis, respectively. CYCLIN D1 modulates the cell cycle progression of cells through G1- and S-phase (Liu, L. et al. 2017). Reduced CYCLIN D1 levels can cause G1-arrest and p53 has been shown to repress *Cyclin D1* (Liu, L. et al. 2017; Rocha et al. 2003). Induction of apoptosis by p53 is mediated by activation, hence cleavage of caspases (Schuler et al. 2000). Western Blot analysis of *Neil1* and *Neil2* deficient EBs and EBs+RA (Figure 4.6) showed a striking reduction in CYCLIN D1 and CASPASE-3 (uncleaved and cleaved) levels specifically upon neuronal differentiation (Figure 4.6B, EB+RA). This suggests that both downstream pathways of p53 are affected in *Neil1* and *Neil2* deficient EBs+RA. However, there is no evidence for induction of apoptosis since CASPASE-3 levels were decreased and not increased. Of note, it has been shown that CASPASE-3 in mESCs is induced upon, and actually compulsory for RA-stimulated differentiation (Fujita et al. 2008). Hence, CASPASE-3 levels were reduced in *Neil1* and

Neil2 deficient EBs+RA due to impaired neuronal differentiation irrespective of its implication in apoptosis. EB differentiation showed no difference in CYCLIN D1 or CASPASE-3 levels between the genotypes (Figure 4.6A, EB).

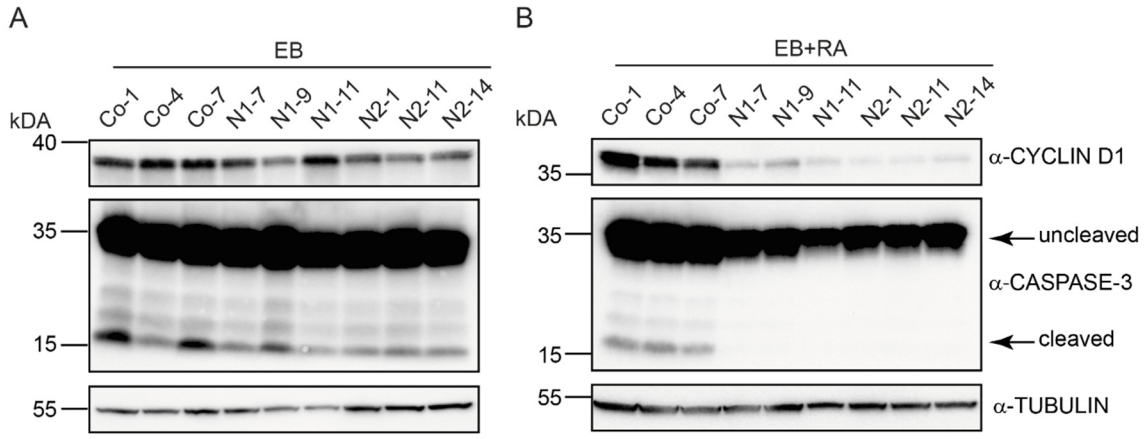

Figure 4.6: Reduced CYCLIN D1 and CASPASE-3 levels in *Neil1* and *Neil2* deficient mESCs upon RA-stimulated differentiation.

(A-B) Western blot analysis for CYCLIN D1 and CASPASE-3 in three independent control (Co, lane 1-3) *Neil1* (N1, lane 4-6) and *Neil2* deficient (N2, lane 7-9) EBs (A) and EBs+RA (B). Uncleaved and cleaved CASPASE-3 are indicated. TUBULIN served as loading control.

To investigate the effects of *Neil1* and *Neil2* deficiency on cell cycle in more detail, I performed cell cycle analysis of undifferentiated, EB and EB+RA differentiated *Neil1* and *Neil2* deficient and Co mESCs using flow cytometry (Figure 4.7). Embryonic stem cells possess an unique cell cycle profile with a characteristic short G1-phase and a high proportion of S-phase cells (White and Dalton 2005). Upon differentiation, the cell cycle of mESCs modulates to less cells in S-phase and more cells in G1-phase. Indeed, and as expected, cell cycle profiles of undifferentiated mESCs were clearly distinguishable from differentiated cells (Figure 4.7 compare panel A with panel B+C). However, I could not observe clear-cut differences in cell cycle profiles between the genotypes in any of the tested conditions. The result was unexpected given the obvious changes in CYCLIN D1 levels upon neuronal differentiation (Figure 4.6B).

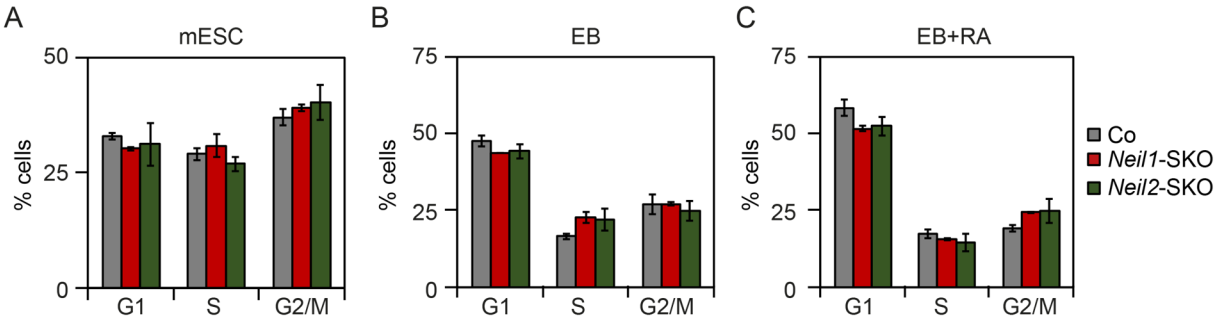

Figure 4.7: Cell cycle profiles are not affected in *Neil1* or *Neil2* deficient mESCs, EBs or EBs+RA.

(A-C) Cell cycle analysis of control (Co), *Neil1* or *Neil2* deficient (*Neil1*- and *Neil2*-SKO) mESCs (A), EBs (B) or EBs+RA (C) using flow cytometry.

Since cleavage of CASPASE-3 was not induced upon EB+RA differentiation in *Neil1* and *Neil2* deficient mESCs, I tested the cleavage of another potential effector caspase, CASPASE-7 (Figure 4.8). However, I could not observe any changes in activated (cleaved) CASPASE-7 levels in neither of the two differentiation protocols (EB and EB+RA) nor between the genotypes. The results suggests that the p53 response does not cause a caspase-induced apoptosis at day 8 of differentiation.

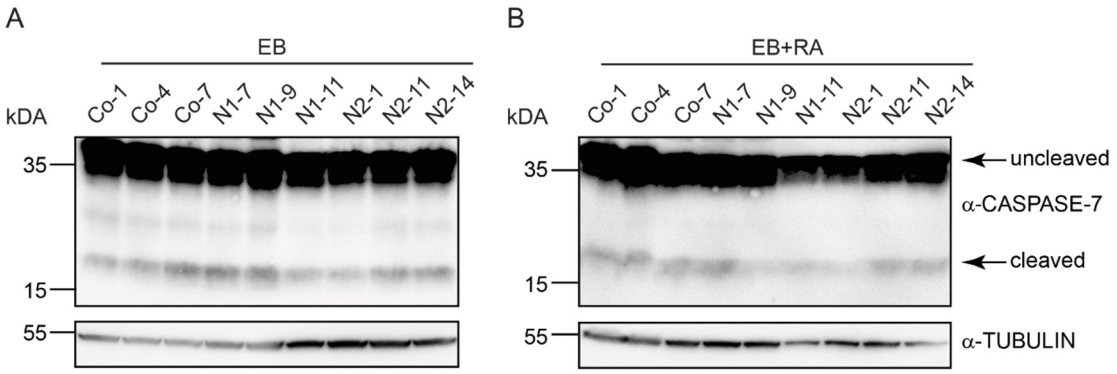

Figure 4.8: *Neil1* or *Neil2* deficient EBs (+RA) do not show altered (activated) CASPASE-7 levels

(A-B) Western blot analysis for CASPASE-7 in three independent control (Co, lane 1-3), *Neil1* (N1, lane 4-6) and *Neil2* (N2, lane 7-9) deficient EBs (A) and EBs+RA (B). Uncleaved and cleaved CASPASE-7 are indicated. TUBULIN served as loading control.

Accumulation of DNA lesions during RA-stimulated differentiation and induction of a p53 response might lead to detectable cell cycle or apoptosis effects already before the manifestation of the neuronal/neural crest defects seen at day 8 of differentiation. Therefore, I tested CYCLIN D1 levels at day 6 of differentiation without or with RA addition (Figure 4.9 A+C). Interestingly, *Neil2* deficient cells showed decreased CYCLIN D1 levels in both conditions, whereas *Neil1* deficiency showed normal levels compared to Co cells. The minor changes for CYCLIN D1 levels in *Neil2* deficient EBs (+RA) were not reflected in flow cytometry analysis, where the distribution of cells in G1-, S- or G2/M-phase was equal between all genotypes in both differentiation conditions (Figure 4.9 B+D).

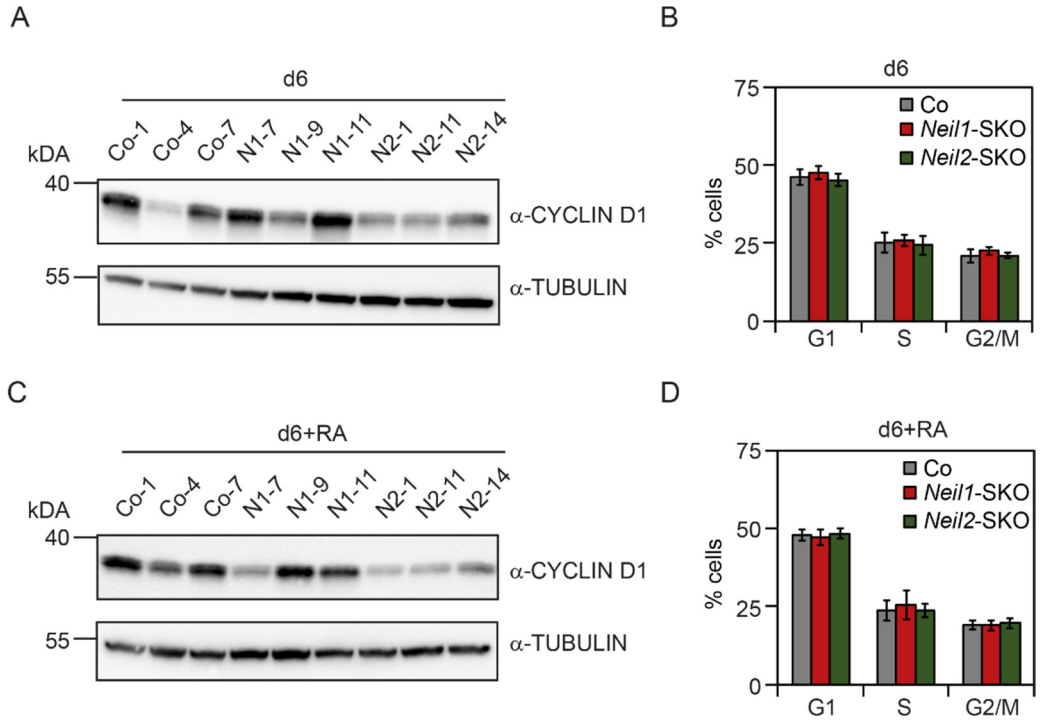

Figure 4.9: Cell cycle profiles are not affected in *Neil1* or *Neil2* deficient EBs (+RA) at d6 of differentiation

(A+C) Western blot analysis for CYCLIN D1 in three independent control (Co, lane 1-3), *Neil1* (N1, lane 4-6) and *Neil2* deficient (N2, lane 7-9) EBs (A, d6) and EBs+RA (C, d6+RA) at day (d) 6 of differentiation. TUBULIN served as loading control. (B+D) Cell cycle analysis of control (Co), *Neil1* or *Neil2* deficient (*Neil1*- and *Neil2*-SKO) EBs (B, d6) or EBs+RA (D, d6+RA) using flow cytometry.

To investigate the potential induction of apoptosis at earlier time points of differentiation, protein levels of inactive (uncleaved) and active (cleaved) effector caspases were analyzed (Figure 4.10). Uncleaved CASPASE-3 levels were not altered between the genotypes at day 6 of differentiation in EBs and EBs+RA, indicating the occurrence of the neuronal/neural crest defect at later stages of differentiation. I could not detect any increase in cleaved CASPASE-3 levels as indicator of apoptosis. Similarly, uncleaved and cleaved CASPASE-7 levels were not changed between genotypes nor differentiation regimes.

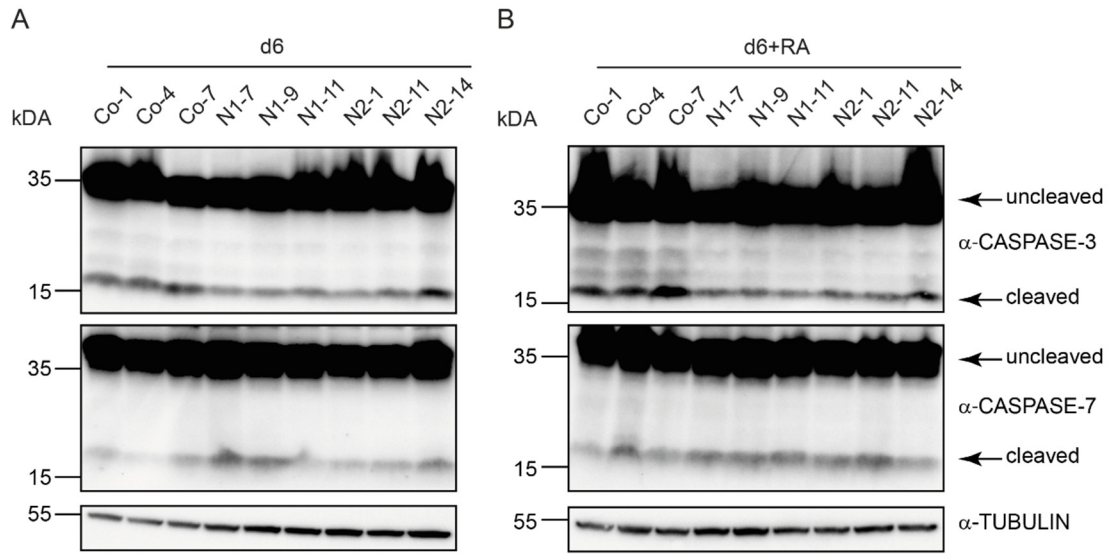

Figure 4.10: (Activated) CASPASE-3 or -7 levels are unaffected in *Neil1* or *Neil2* deficient EBs (+RA) at d6 of differentiation

(A-B) Western blot analysis for CASPASE-3 and CASPASE-7 in three independent control (Co, lane 1-3), *Neil1* (N1, lane 4-6) and *Neil2* deficient (N2, lane 7-9) EBs (A) and EBs+RA (B). Uncleaved and cleaved CASPASE-3/-7 are indicated. TUBULIN served as loading control.

4.2.2.4 *Mitochondrial p53 repsonse in Neil1 and Neil2 deficient mESCs upon neuronal differentiation*

The observed neuronal/neural crest differentiation defect is presumably caused by an activated p53 response. But why do *Neil1* and *Neil2* deficiencies trigger a p53 response? Interestingly, *Neil3* deficient mESCs do not show this phenotype ([redacted], unpublished data). Like *Tdg* deficient mESCs, they differentiate normal until day 8 of EB (+RA) differentiation. What are the differences between NEIL1, NEIL2 and NEIL3? First, NEIL3 mainly acts as monofunctional glycosylase whereas NEIL1 and NEIL2 fulfill bifunctional DNA glycosylase activity. Second, NEIL1 and NEIL2 are localized in the nucleus as well as the mitochondria mediating BER in both organelles. Therefore, oxidative damage was measured in genomic and mitochondrial DNA of Co, *Neil1* and *Neil2* deficient mESCs, EBs and EBs+RA [redacted] and [redacted], unpublished data). Interestingly, oxidative damage increased only in mtDNA in *Neil1* and *Neil2* deficient cells and specifically upon neuronal differentiation, which coincides with the differentiation defect. Consequently, I analyzed *Neil1* and *Neil2* deficient cells for a mitochondrial p53 response. The intrinsic mitochondrial apoptotic pathway is mediated by the BCL-2 superfamily (Tsujimoto 1998). The BCL-2 superfamily comprises anti- and pro-apoptotic family members like BCL-2 and BAX or BAK, respectively. Together these proteins balance the response to cellular stress until the pro-apoptotic signals exceed a threshold (Hemann and Lowe 2006). *Bcl-2*, *Bax* and *Bak* have been shown to be direct p53 targets (Fischer 2017). Consequently, I investigated the expression of these mitochondrial p53 targets in undifferentiated, EB and EB+RA differentiated *Neil1* and *Neil2* deficient and Co mESCs (Figure 4.11). Interestingly, the anti-apoptotic factor *Bcl-2* was downregulated whereas the pro-apoptotic factor *Bax* was upregulated in *Neil1* and *Neil2* deficient EBs (+RA). The pro-apoptotic factor *Bak* was only upregulated upon neuronal

differentiation (EB+RA). The results indicate an imbalance between anti- and pro-apoptotic factors favoring induction of apoptosis.

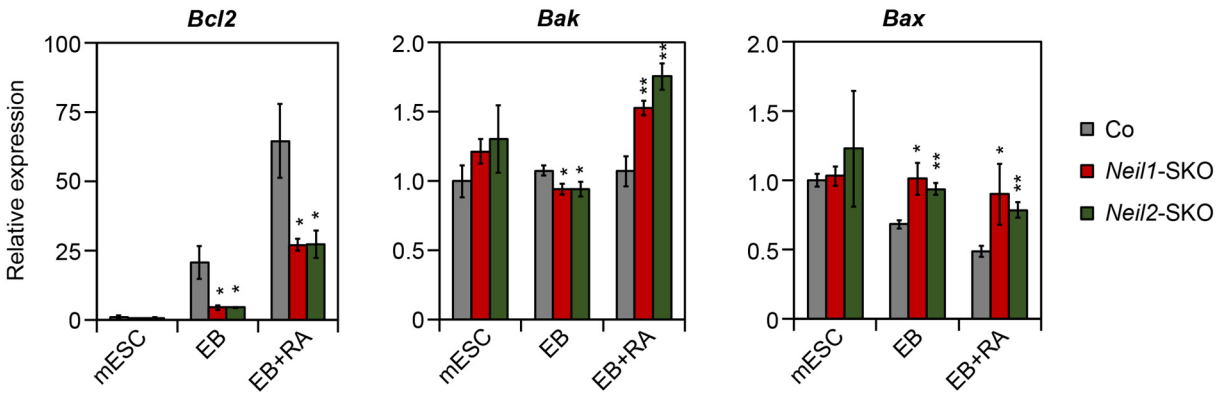

Figure 4.11: Mitochondrial p53-pathway is activated in *Neil1* and *Neil2* deficient mESCs upon differentiation

Expression analysis of selected mitochondrial p53 target genes in *Neil1* and *Neil2* deficient (*Neil1*- and *Neil2*-SKO) or control (Co) mESCs and upon differentiation (EB, EB+RA) measured by qPCR. Expression levels are relative to average expression in Co mESCs.

Upon exceeding a threshold of pro-apoptotic signals, the mitochondria release apoptogenic factors like CYTOCHROME-C and apoptosis-inducing factor (AIF) (Tsujimoto 1998). After release into the cytoplasm, CYTOCHROME-C and AIF activate caspases, which in turn cause apoptosis. Interestingly, AIF is only released upon mitochondrial dysfunction and activates apoptosis also in a caspase-independent manner, which may be in line with the apparent lack of CASPASE-3 and -7 activation upon EB+RA differentiation in *Neil1* and *Neil2* knockouts. Cytosolic, nuclear and mitochondrial fractions of RA-differentiated *Neil1* and *Neil2* deficient and Co mESCs were isolated to trace the CYTOCHROME-C and AIF release (Figure 4.12). As expected, the cytosolic marker TUBULIN was enriched in cytosolic, AIF in mitochondrial and RAD21 in nuclear fractions.

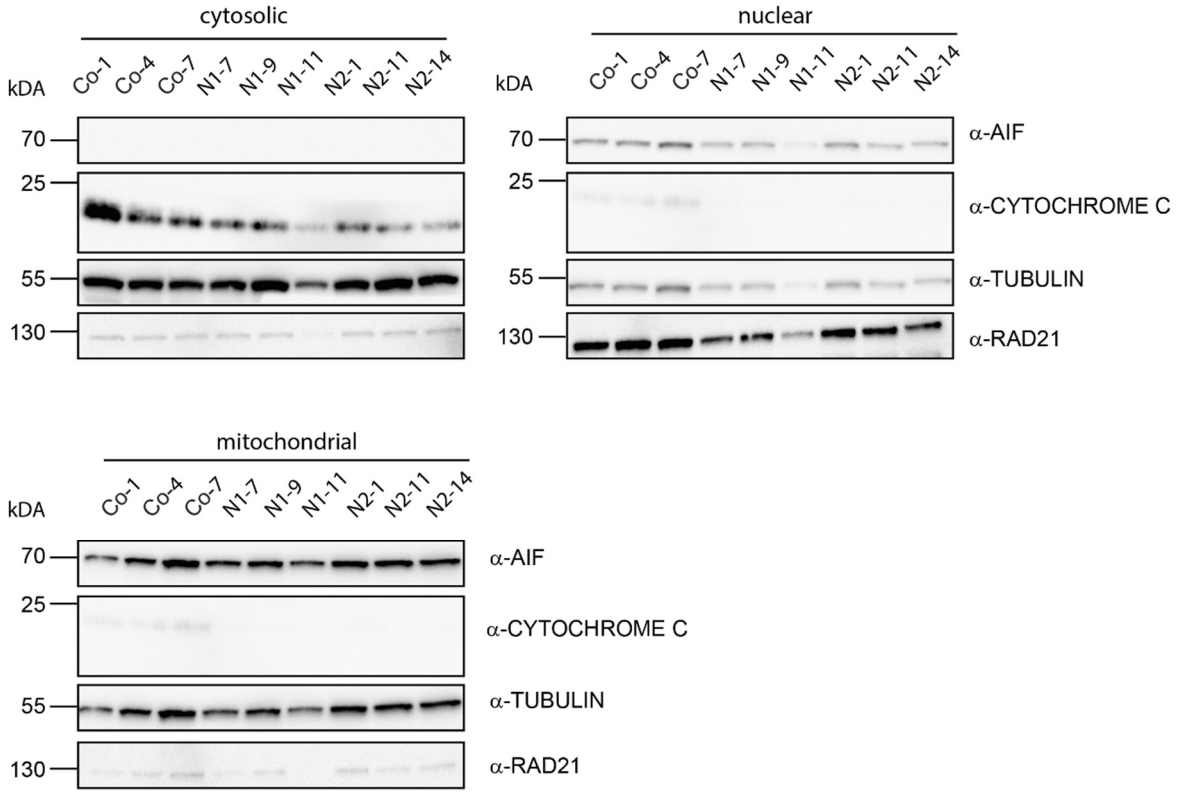

Figure 4.12: No AIF-release upon neuronal differentiation in *Neil1* and *Neil2* deficient mESCs.

Western Blot analysis for AIF and CYTOCHROME C in cytosolic, nuclear and mitochondrial fractions of control (Co, lane 1-3), *Neil1* (N1, lane 4-6) or *Neil2* (N2, lane 7-9) deficient EBs+RA. TUBULIN and RAD21 served as loading control.

The apoptogenic factor CYTOCHROME-C was exclusively detectable in the cytosolic fraction showing that CYTOCHROME-C release occurred upon neuronal differentiation. However, I could not detect an obvious difference between the genotypes and if anything, levels were reduced in *Neil1* and *Neil2* deficient EBs+RA. The result supports the suggestion that the mitochondrial caspase-induced apoptotic pathway was not specifically activated in .*Neil1* and *Neil2* deficient EBs+RA. AIF protein was highly detectable in the mitochondrial fraction, but also lowly detectable in the nuclear fraction indicating that AIF was released and transferred to the nucleus presumably mediating caspase-independent apoptosis. However, no significant changes in AIF levels were observed between the genotypes.

4.2.3 Discussion

The DNA repair proteins NEIL1 and NEIL2 have been identified as players of the active DNA demethylation pathway. In this respect, NEIL1 and NEIL2 promote the substrate turnover of TDG during 5fC and 5caC excision. Mice deficient for either *Neil1* or *Neil2* do not show severe developmental defects, which is in contrast to phenotypes of mice deficient for *Tet* or *Tdg* and to the severe neural crest phenotype observed in *neil2* knockdown *Xenopus* embryos (Cortázar et al. 2011; Dai et al. 2016; Schomacher et al. 2016). The requirement of NEIL1 and NEIL2 in promoting DNA demethylation during mouse embryonic development had not been investigated to date. Hence, we analyzed the role of NEIL1 and NEIL2 in pluripotency and lineage commitment using mouse embryonic stem cells as model system. The analysis of CRISPR/Cas9 generated *Neil1* and *Neil2* deficient mESCs resulted in two main findings. First, NEIL1 and NEIL2 are not required to maintain pluripotency of mESCs, but are essential for neuronal/neural crest differentiation *in vitro*. Second, this specific neuronal/neural crest defect is presumably caused by a mitochondrial DNA damage-induced p53-response and not by attenuated 5fC and 5caC removal.

We found that NEIL1 and NEIL2 are not required to maintain pluripotency. This finding is in line with sustained pluripotency in absence of other DNA (de)methylation factors (Cortázar et al. 2011; Dawlaty et al. 2014; Tsumura et al. 2006). Similarly, NEIL1 and NEIL2 are required for proper differentiation like DNMTs, TETs and TDG (Cortázar et al. 2011; Dawlaty et al. 2014; Schmidt et al. 2012) mirroring the characteristic effects of a DNA demethylating factor in mESCs. However, deletion of none of these factors showed a comparably strong specificity for a neuronal/neural crest defect at this stage of differentiation. Although *Tdg*-deficient mESCs are impaired in terminal neuronal differentiation *in vitro* (Cortázar et al. 2011), NEIL1 and NEIL2 appear to be essential at earlier stages of differentiation. Hence, it is likely, that there is no requirement of NEIL1 and NEIL2 in promoting substrate turnover at this stage. Consistently, there is no accumulation of 5fC or 5caC upon differentiation. Instead, there is a significant increase in mitochondrial DNA (mtDNA) damage upon RA-mediated differentiation, highly suggesting that this accumulation of mtDNA damage causes this serve phenotype *in vitro*. TDG solely localizes to the nucleus (Prakash and Doublié 2015) clarifying the discrepancy between the phenotypes of proteins promoting the same process in the active DNA demethylation pathway.

Our study revealed an induction of the p53 pathway in *Neil1* and *Neil2* deficient mESCs upon neuronal differentiation *in vitro*. The function of p53 in the DNA damage response pathway is well studied and established. However, its role in embryogenesis is not clear to date. Activation of p53 has been shown to be important for differentiation (Lin et al. 2005). Upon DNA damage, activated p53 suppresses expression of the pluripotency marker *Nanog* and causes differentiation. This novel mechanism ensures genome stability of the pluripotent pool by initiating differentiation of damaged mESCs. Although this implies a beneficial effect of activated p53 in differentiation, the function of p53 in neural crest development seems to be opposing. p53 activation impairs cranial neural crest cell differentiation as

highlighted by craniofacial defects upon p53 stabilization in chick embryos (Rinon et al. 2011). Moreover, deficiency of *Tcof1*, whose mutations leads to the craniofacial disorder Treacher Collins Syndrom (TCS), results in p53 activation (Jones et al. 2008). Inhibition of p53 rescues this craniofacial defect by presumably preventing apoptosis in neural crest cells. Another study showed that nucleolar dysfunction and ribosomal DNA (rDNA) damage cause p53 activation in TCS (Calo et al. 2018). Furthermore, aberrant activation of p53 causes phenotypes characteristic of CHARGE syndrome, a developmental disorder of defective neural crest-derived tissues. In 70-90% of the cases, CHARGE syndrome results from mutations in *CHD7* and loss of *CHD7* in neural crest cells has been associated with p53 activation (van Nostrand et al. 2014). Our results support the proposition that aberrant p53 activation causes defects in neural crest cell development.

Why is there no p53 response in *Neil1* and *Neil2* deficient undifferentiated mESCs? And why is the differentiation defect specifically occurring in neuronal/neural crest progenitors? Mouse ESCs have a high demand of safeguarding genome stability, as they are the only renewable source of all cells of an embryo. In contrast, some of the critical DNA damage responses do not exist in mESCs showing that the highly proliferative mESC developed their own unique DNA damage response to cope with constitutive cellular stress, which is characterized by high sensitivity and robustness, mediated by ESC-specific factors (i.e. FILIA or SALL4) (Fu et al. 2017; Tichy 2011). Hypothetically, NEIL1 and NEIL2 may simply not be required for this specialized DNA damage response. Moreover, mESC express higher levels of DNA repair factors when compared to differentiated cells (Tichy 2011) raising the possibility that other DNA glycosylases involved in repair of oxidative lesions might compensate for *Neil1* and *Neil2* deficiency (i.e. *Nthl1*). Why is the differentiation defect specifically occurring in neuronal/neural crest progenitors? Embryonic stem cells rely on glycolysis as energy source since they possess only few and immature mitochondria (Facucho-Oliveira et al. 2007). Consequently, mESCs encounter less mitochondrial ROS than somatic cells. Upon differentiation mitochondria mature and mESCs switch their energy production from glycolysis to oxidative phosphorylation causing high ROS levels (Xu et al. 2013). Although ROS is an essential signaling molecule in cells, high levels of ROS can have deleterious effects like DNA damage (Sena and Chandel 2012). Neurons in particular have a high demand of energy, pinpointing the importance of mitochondrial dynamics in neurogenesis (Khacho and Slack 2018). High ROS levels cause damage of nuclear DNA, but given the spatial origin of ROS as a byproduct of oxidative phosphorylation, the primary target for ROS damage is mitochondrial DNA. If unrepaired, i.e. in absence of NEIL1 and NEIL2, this damage might lead to mitochondrial dysfunction specifically during neuronal differentiation. Interestingly, addition of the antioxidant NAC revealed an unexpected effect by mimicking *Neil1* and *Neil2* deficiency upon neuronal differentiation. NAC has been shown to induce mitochondrial mediated apoptosis (Liu, Y. et al. 2017) supporting our proposition of the importance of mtDNA integrity upon neuronal differentiation.

One shortcoming of this study is the missing detection of downstream effectors of p53. Although I could detect reduced CYCLIN D1 levels in *Neil1* and *Neil2* deficient mESCs specifically upon neuronal

differentiation, flow cytometry analysis did not reflect cell cycle deregulation. CYCLIN D1 is required for neurogenesis in chick embryos in a cell cycle independent manner (Lukaszewicz and Anderson 2011) and in human embryonic stem cells CYCLIN D1 recruits coactivators to neuroectoderm genes (Pauklin et al. 2016). However, whether CYCLIN D1 is the cause or consequence of impaired neuronal differentiation remains open. Hence, the reduction in CYCLIN D1 levels in *Neil1* and *Neil2* deficient EBs+RA might rather reflect the neuronal/neural crest defect and not an effective sign of a p53 response. Further investigations are required to understand the causality between CYCLIN D1 protein levels and neuronal development. Alternatively, cell cycle effects might be masked since RA-stimulated differentiation give rises to a heterogeneous cell population.

This study reveals the importance of NEIL1 and NEIL2 in protecting mitochondrial DNA ensuring functional neuronal differentiation. Of note, this does not exclude the involvement of NEIL1 and NEIL2 in promoting 5fC and 5caC removal during development or in somatic tissues. Although global effects are not expected, site-specific functions of NEIL1 and NEIL2 could regulate important biological processes.

Mice deficient for *Neil1* or *Neil2* do not show any neuronal defects during development. These findings are clearly in contrast to our *in vitro* results in mESCs. How can we explain this discrepancy? Mouse embryonic stem cells cultured *in vitro* experience appreciable more environmental stress than *in vivo* regarding i.e. high oxygen partial pressure and high glucose compositions in the medium (Millman et al. 2009; Powers et al. 2008). On top of the intracellular stress occurring upon neuronal differentiation, these additional stressors might trigger an intrinsic p53 response. However, it is puzzling that addition of antioxidants, low oxygen or low glucose conditions were unable to rescue the neuronal/neural crest phenotype. Most likely, it is not sufficient to change one parameter of environmental stress to rescue this phenotype, or mitochondrial targeted antioxidants are required to reduce ROS levels specifically in the mitochondria. Hypothetically, culture conditions fully resembling the *in vivo* situation might prevent the neuronal/neural crest defect in *Neil1* and *Neil2* deficient cells. Given the complexity of embryonic development involving signaling cascades even from surrounding tissues (Bénazéraf and Pourquié 2013), this is quite unrealistic to accomplish. However, this might explain the rather mild phenotypes *in vivo*. Neuronal development is a crucial step in embryogenesis, which requires tight regulation of various processes. It is likely that compensatory mechanisms have evolved from the surrounding tissues to ensure functionality of this key process.

5 General Discussion

The precise role of 5mC deposition and removal during early mouse development is not fully understood. However, several studies have established DNA methylation as a hallmark of lineage commitment. Mice deficient for factors regulating DNA (de)methylation show severe developmental defects, which can be recapitulated *in vitro* (Cortázar et al. 2011; Dai et al. 2016; Dawlaty et al. 2014; Okano et al. 1999; Schmidt et al. 2012). We used mouse embryonic stem cells as an *in vitro* model to investigate the function of two DNA repair protein families (GADD45 and NEIL) recently identified as regulators of active DNA demethylation (Arab et al. 2019; Barreto et al. 2007; Kienhöfer et al. 2015; Schomacher et al. 2016). These studies shed light on the unexpected role of GADD45 proteins as 2-cell regulators and on the requirement of NEIL1 and NEIL2 in neuronal/neural crest differentiation. We hypothesize that the function of GADD45 proteins in 2-cell regulation is linked to their involvement in DNA demethylation. For NEIL1 and NEIL2, we propose a model, where these two proteins protect against a mitochondrial DNA damage response upon neuronal differentiation, although we cannot fully exclude the association of NEIL1 and NEIL2 in promoting 5fC and 5caC removal. Overall, these studies elucidate the importance of accurate investigation of multifunctional proteins during distinct stages of development.

5.1 *In vitro* models of early mouse development

Our *in vitro* analysis using mESCs proposed a requirement of GADD45 proteins for embryoid body and monolayer differentiation as well as of NEIL1 and NEIL2 for RA-mediated EB differentiation. Genes that were affected in *Gadd45a/b/g* deficient mESCs were mostly attributed to neuronal functions. This is in line with GADD45 proteins being involved in neural differentiation *in vivo* (Huang et al. 2010; Kaufmann and Niehrs 2011; Ma, Jang et al. 2009) and *Gadd45a/b* DKO embryos showing neural tube closure defects and brain hemorrhage (Schüle et al. 2019). This close correlation underlines the applicability of mESC differentiation as *in vitro* model for early mouse postimplantation development. However, we did not see this agreement between *in vitro* and *in vivo* data in our studies on NEIL proteins. The striking neuronal/neural crest defect of *Neil1* and *Neil2* deficient mESCs has not been recapitulated *in vivo*, since mice deficient for single *Neil* or all three *Neil* genes are viable without any severe developmental defects reported (Canugovi et al. 2012; Chakraborty et al. 2015; Regnell et al. 2012; Rolseth et al. 2017; Sejersted et al. 2011; Vartanian et al. 2006). This suggests, that defects exclusively occurring *in vitro*, as observed for *Neil1* and *Neil2* deficient mESCs, are triggered by additional environmental stressors mediated by culture conditions (i.e. high pO_2, high glucose concentration). Consequently, these phenotypes appear milder *in vivo*. *Neil1* or *Neil2* deficient mice challenged by high fat diet causing high ROS levels might show the severe phenotype *in vivo*. A rather mild phenotype in unchallenged conditions indicates a diversity of compensatory mechanisms that have evolved to protect against environmental stress during development. Various signals from the multifaceted environment and crosstalk between different cell types are essential for fate decisions

specifically in the early fate (Bénazéraf and Pourquié 2013), making these steps more susceptible to compensatory mechanisms.

Our study revealed an unexpected role for GADD45 proteins regulating the 2C-like state. Consequently, transferring this *in vitro* function to the *in vivo* development, embryos deficient for *Gadd45* genes should arrest at 2-cell stage. Since *Gadd45a* and *Gadd45b* are the major isoforms expressed at 2-cell stage, mice deficient for these genes are proposed to show this early preimplantation defect. *Gadd45a* or *Gadd45b* single knockout mice have been characterized and have not been reported to display any preimplantation defects (Hollander et al. 1999; Lu et al. 2004). However, our study in DKO mice revealed partial misregulation of 2-cell associated genes, embryonic sublethality and, importantly, partial developmental arrest in an *in vitro* preimplantation assay. Why do, on average, only 50% show the developmental arrest? First, partial defects might result from incomplete penetrance of preimplantation defects as observed for other mutants (Narducci et al. 2002). Second, other 2-cell regulators and the redundant family member GADD45γ might compensate for the loss of *Gadd45a/b*. It is likely that nature evolved a complex backup mechanism to secure developmental integrity, since all cells of the embryonic and extraembryonic tissues derive from these two cells. The effectiveness of these compensatory mechanisms *in vivo* is exemplified by the fact that *Dux* KO mice do not show any preimplantation defect (Chen and Zhang 2019). Contradictory, DUX has been established as a leading promoter of the 2C-like state and acute CRISPR/Cas9 mediated deletion of *Dux* in 2-cell stage embryos results in a striking preimplantation developmental arrest (Hendrickson et al. 2017; De Iaco et al. 2017).

Overall, both studies show the potential of mESCs and their versatile application as *in vitro* model for early mouse development. However, the presented data also underline the necessity to recapitulate *in vitro* findings *in vivo.* Further studies are fundamental to acquire culture conditions resembling the *in vivo* environment more accurately and to assign individual steps of the *in vitro* differentiation to mouse developmental stages.

5.2 DNA methylation and gene expression

Global levels of 5mC oxidative derivatives were altered neither in *Gadd45* TKO nor in *Neil1* and *Neil2* deficient mESCs suggesting that both protein families exert a locus-specific and not a global function in active DNA demethylation. Consequently, no striking changes in gene expression were expected in *Gadd45* TKO or *Neil1* and *Neil2* deficient mESCs. Although this was true for undifferentiated KO mESCs, both KO mESCs (*Gadd45* TKO and *Neil1* and *Neil2* SKO) showed massive gene deregulation upon differentiation. Can this deregulation be attributed to DNA demethylation? Previous studies have shown that the position of 5mC and the cellular context are key determinants for the transcriptional outcome, i.e. promoter demethylation directly results in gene activation. In contrast, 5hmC at enhancers mediates a molecular memory for differentiation (Kim et al. 2018; Stadler et al. 2011). Consequently, locus-specific imbalanced methylation at enhancers is not reflected in misregulated gene expression in undifferentiated, but rather in differentiated mESCs, like observed in *Gadd45a/b/g* and *Neil1* and *Neil2*

deficient mESCs. Methylome and gene expression analysis support this concept of memory function in *Gadd45* TKO mESCs. We identified ~7000 hyper-DMRs that were enriched at enhancers, but only weakly correlating with deregulated genes in undifferentiated *Gadd45a/b/g* deficient mESCs. Misregulation of gene expression dramatically increased upon differentiation presumably caused by imbalanced methylation at enhancers in mESCs and thereby loss of memory function.

These results support the importance of GADD45 proteins in gene regulation via its function in active DNA demethylation. Of note, this study lacks evidence for the direct effect of the functionally pleiotropic GADD45 proteins in mediating active DNA demethylation of specific sites. A shortcoming of my study is that direct binding of GADD45 proteins to hyper-DMR sites was not shown. Hence, one cannot exclude that hypermethylation in *Gadd45* TKO mESCs is caused by indirect effects.

Changes in methylation might not systematically affect gene expression in a direct way. DNA demethylation intermediates interact with specific binders, thereby suggesting an epigenetic function of each intermediate itself (Spruijt et al. 2013). Thus, imbalance of the DNA demethylation pathway might affect gene regulation by erroneous recruitment of specific factors. TDG is product-inhibited and stalled at AP-sites subsequent to 5fC and 5caC removal (Schomacher et al. 2016). This has been considered as a protective mechanism of the vulnerable AP-site to prevent genomic instability. In contrast, attenuated turnover might facilitate recruitment of additional co-activators via TDG scaffolding (i.e. CBP/p300 (Tini et al. 2002)) to these specific sites thereby manifesting an activated chromatin status. In this scenario, NEIL1 and NEIL2 promoting TDG´s substrate turnover would have a deleterious effect on gene expression. Importantly, our study does not support the requirement of NEIL1 and NEIL2 in 5fC and 5caC removal upon differentiation. We propose a model whereby NEIL1 and NEIL2 protect against an mtDNA damage-induced neuronal/neural crest defect. To test for locus-specific action of NEIL1 and NEIL2 in promoting active DNA demethylation in mESCs, methylase-assisted bisulfite (MAB)-sequencing should be performed to map 5fC and 5caC at base resolution.

5.3 Crosstalk between DNA methylation and DNA repair

GADD45α/β/γ and NEIL1 and NEIL2 are proteins that exhibit a dual function in active DNA demethylation and DNA damage repair, exemplifying the tight interplay between these two processes. The direct removal of the methyl moiety of cytosines is thermodynamically unfavorable. Therefore, 5mC is oxidized, excised and replaced with cytosine in a consecutive action of enzymes involved in active DNA demethylation and DNA repair. Is this interplay between DNA demethylation and DNA repair only a coincidence? Might DNA repair also depend on DNA demethylation? DNA demethylation intermediates, generated by TET-mediated oxidation, can be passively diluted during replication but they can also be recognized and excised by the BER enzyme TDG (Bird 2002; Dean 2008; He et al. 2011; Maiti and Drohat 2011). TDG itself, but also each intermediate of the active DNA demethylation pathway, may recruit additional factors, i.e. BER proteins (Spruijt et al. 2013; Tini et al. 2002) therefore promoting the DNA repair process. Moreover, the chemical properties of oxidized 5mC derivatives

cause an open chromatin formation. This facilitates access of BER proteins showing that the interplay between DNA demethylation and BER is tightly organized. Proteins involved in active DNA demethylation have a well-established role in DNA repair (i.e. TDG, NEIL and GADD45 proteins) with the exception of TET proteins. However, recent studies propose the involvement of TET proteins in DDR. TET-mediated 5hmC has been shown to colocalize with the DNA damage markers 53BP1 and γH2AX, with the notion to facilitate DNA repair (An et al. 2015; Kafer et al. 2016; Nakatani et al. 2015). In this context 5hmC can recruit DDR proteins to facilitate DNA repair in line with a proposed function of 5hmC as an epigenetic mark of DNA damage (Kafer et al. 2016). These studies open a new avenue for GADD45 and NEIL proteins in DNA repair. On top of their evolved role, these proteins could facilitate the establishment or the removal of 5hmC as repair mark. Diversely, TET3 has been shown to be a direct target of ATR upon the DNA damage response (Jiang et al. 2017). TET3-mediated DNA demethylation might affect gene expression of DNA repair proteins (Cimmino et al. 2015) thereby creating a novel model of DNA damage response. GADD45 and NEIL proteins may be involved in this gene activation. Integrity of DNA demethylation and DNA repair seem to be crucial for accurate neuronal development. As exemplified by our studies, neuronal defects caused by dual-functional proteins demand detailed analysis to attribute their function to any of the specific processes.

6 Material and Methods

6.1 Material

6.1.1 Equipment

-150°C freezer (Sanyo); -80°C freezer (Sanyo); balances (Sartorius, Kern); BD LSR-FortessaSORP (BD Biosciences); Trans blot turbo (BioRad); cell counter (BioRad); cell culture incubators (Thermo Scientific); centrifuges (Heraeus); Cryo-Safe Cooler (Belart); E-Gel electrophoresis system (Invitrogen); heating blocks (Eppendorf); laminar flow hoods (Dometric); LightCycler 480 (Roche); magnetic stirrer (Heidolph); microcentrifuges (Heraeus); microscope (Leica); multidispenser pipette (Eppendorf); Nanodrop 2000 spectrophotometer (Thermo Scientific); PAGE minigel chambers (BioRad); PCR thermocyclers (Biometra); pH meter (Mettler Toledo); pipet boy (Integra); pipettes (Eppendorf); power supplies (BioRad); stereomicroscope (Leica); transferpipette (Neolab); ultrapure water purification system (Millipore); UV photodocumentation system (BioRad); vortexer (Scientific industries); waterbaths (Neolab)

6.1.2 Chemicals and pre-made buffers

0.1% gelatine in ultrapure water (Millipore); 100x L-Glutamin (Thermo Fisher Scientific); 100x Non-essential amino acids (NEAA) (Thermo Fisher Scientific); 100x Penicillin/Streptomycin (Thermo Fisher Scientific); 100x sodium pyruvate (Thermo Fisher Scientific); 5xFS buffer (Invitrogen); 12% CriterionTM Bis-Tris Protein Gel (Biorad); α-tocopherol (Sigma); β-mercaptoethanol (Sigma-Aldrich); bicinchoninic acid (Sigma-Aldrich); bovine serum albumin (Sigma-Aldrich); bromophenol blue (Sigma-Aldrich); CHIR99021 (Biocat); cholecalciferol (Sigma-Aldrich); complete™ protease inhibitor cocktail tablets (Roche); dimethylsulfoxide (Sigma-Aldrich); dithiothreitol (Sigma-Aldrich); DMEM with glucose (Thermo Fisher Scientific); DMEM low glucose (Thermo Fisher Scientific); dNTPs (Thermo Fisher Scientific); EDTA (Sigma-Aldrich); ES-grade FBS (PAN Biotech); ethanol (Sigma-Aldrich); glycerol (Sigma-Aldrich); hydrochloric acid (Sigma-Aldrich); isopropanol (Sigma-Aldrich); L-ascorbic acid 2-phosphate sesquimagnesium salt (Sigma-Aldrich); magnesium chloride (Sigma-Aldrich); methanol (Sigma-Aldrich); N-acteyl-L-cysteine (Sigma-Aldrich); N-Ethylmaleimide (Sigma-Aldrich); Nonidet P-40 (Sigma-Aldrich); nuclease-free water (Qiagen); PD0325901 (Sigma-Aldrich); DPBS (Thermo Fisher Scientific); PhosStop™ tablets (Roche); potassium chloride (Sigma-Aldrich); Random primers (Invitrogen); retinoic acid (Sigma-Aldrich); restore™ Western Blot Stripping buffer (Thermo Fisher Scientific); skim milk powder (Sigma-Aldrich); sodium chloride (Sigma-Aldrich); sodium deoxycholate (Sigma-Aldrich); sodium dodecyl sulfate (Sigma-Aldrich); Tris base (Sigma-Aldrich); Tris-HCl (Sigma-Aldrich); Trypsin, 0.25% (Thermo Fisher Scientific); Tween-20 (Sigma-Aldrich).

6.1.3 Kits and enzymes

DNAse I (Qiagen); LIF (homemade supernatant IMB); miRNeasy mini kit (Qiagen); Probes Master (Roche); Ribolock (Thermo Fisher Scientific); RNase A (Qiagen); Superscript II Reverse Transcriptase (Invitrogen); SuperSignal West Pico/Femto (Thermo Scientific); Trans-Blot Turbo™ Transfer Pack (Biorad).

6.1.4 Cell culture media and buffers

6.1.4.1 Cell culture media

ESC medium	DMEM, 15% ESC-grade FBS, 2mM L-Glutamine, 50U/ml Penicilin/Streptomycin, 1mM sodium pyruvate, 1x NEAA, 100µM β-mercaptoethanol, 4% LIF-supernatant
2i medium	ESC medium, 1µM PD0325901, 3µM CHIR99021
1i medium	ESC medium, 3µM CHIR99021
CA medium	DMEM, 10% ESC-grade FBS, 2mM L-Glutamine, 50U/ml Penicilin/Streptomycin, 1x NEAA, 100µM β-mercaptoethanol

6.1.4.2 General molecular biology buffers

dNTP mix	5mM dATP, 5mM dCTP, 5mM dGTP, 5mM dTTP
10x HG buffer	400mM Tris-HCl pH 7.8, 250mM NaCl, 50mM $MgCl_2$
4x Lämmli	60mM Tris-Hcl pH 6.8, 2% sodium dodecyl sulfate, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue
NOP buffer	2% Nonidet P-40, 20mM Tris-HCl pH 7.5, 150mM NaCl. Stored at 4°C
NOP+ buffer	NOP buffer, 1x complete™ protease inhibitor, 1x PhosStop™, 2mM DTT, 5mM N-Ethylenmaleimide, prepare freshly
RIPA buffer	50mM Tris-Hc pH 8.0, 150mM NaCl, 1% (vol/vol) NP-40, 0.5% (wt/vol) sodium deoxycholate, 0.1% (wt/vol) SDS, 2mM EDTA
20x TBS	3M NaCl, 5.36 mmM KCl, 1M Tris-HCl pH 7.4, autoclaved

6.1.5 Antibodies

Antibody	Host species	Company	Catalogue #	Dilution
AIF	rabbit	Abcam	ab32516	1:1000
alpha-tubulin	mouse	Sigma	T5168	1:10000
Caspase-3	rabbit	Cell Signaling	9662	1:1000
Caspase-7	rabbit	Cell Signaling	9492	1:1000
Cyclin D1	rabbit	Abcam	ab134175	1:2000
Cytochrome c	rabbit	Abcam	ab133504	1:1000
Rad21	mouse	Merck Millipore	05-905	1:2000
anti-rabbit HRP	goat	Dianova	111-035-144	1:10000
anti-mouse HRP	goat	Dianova	115-035-146	1:10000

6.1.6 Primer sequences

Oligonucleotides were synthesized by Sigma Aldrich or IDT.

6.1.6.1 qPCR expression primers

Gene	Forward primer	Reverse primer	Probe
mmBak1	ggaatgcctacgaactcttca	ccagctgatgccactcttaaa	19
mmBax	gtgagcggctgcttgtct	gtgggggtcccgaagtag	83
mmBcl-2	agtacctgaaccggcatctg	ggggccatatagttccacaaa	75
mmCcng1	ttagtaggcctgtcggatcg	agcagtttctgagagtcagttgtc	108
mmEomes	accggcaccaaactgaga	aagctcaagaaaggaaacatgc	9
mmGata6	ggtctctacagcaagatgaatgg	tggcacaggacagtccaag	40
mmMdm2	tgtttggagtcccgagtttc	agccactaaatttctgtagatcattg	99
mmNanog	gcctccagcagatgcaag	ggttttgaaaccaggtcttaacc	25
mmNeurog1	gacctgtccagcttcctcac	tggaggctaggggctgtag	101
mmP21	aacatctcagggccgaaa	tgcgcttggagtgatagaaa	16
mmPax3	gcccacgtctattccacaa	gaatagtgctttggtgtacagtgc	69
mmPax6	gttccctgtcctgtggactc	accgcccttggttaaagtct	78
mmPerp	gaccccagatgcttgttttc	accagggagatgatctggaa	71
mmOct4	aatgccgtgaagttggagaa	ccttctgcagggctttcat	95
mmSesn2	acatccactgcgtctttgg	cgtcttgatatagattttgaggttcc	17
mmSox2	gacgtcgtagcggtgcat	acggcagctacagcatga	68
mmTbp	ggggagctgtgatgtgaagt	ccaggaaataattctggctca	97

6.2 Methods

6.2.1 Cell culture

6.2.1.1 Mouse embryonic stem cell culture

ESCs were cultured on pre-coated gelatinized (0.1% Gelatin, Millipore) plates in regular serum-supplemented ESC medium. Cells were passaged every other day by washing cells with pre-warmed PBS subsequent to detaching cells with pre-warmed 0.25% trypsin. Trypsinization was quenched by addition of serum-supplemented ESC medium. Cells were spun down for 3min, at 300g. Cells were resuspended in ESC medium and $1x10^6$ cells or 1/8th was plated on a new pre-coated gelatinized plate. For growth curve analysis, ESCs were counted every second day mixing trypan blue and cell suspension in a 1:1 ratio and using a TC10 automated cell counter.

For 2i and 1i treatment, $1x10^6$ cells were plated in serum-supplemented ESC medium containing a final concentration of 1 µM PD0325901 and 3 µM CHIR99021 and 3 µM CHIR99021, respectively. ESCs were treated with 2i and 1i medium for 8 days and split every other day, followed by regular EB differentiation.

6.2.1.2 In vitro embryoid body (EB) differentiation

For EB differentiation $3.5x10^6$ ($6x10^5$) cells were plated on non-adherent 10 cm (3.5 cm) bacterial dishes (Greiner) in 15 ml (2.5 ml) CA medium. Every other day, EBs were collected in a 50 ml (15 ml) falcon and settled down by gravity. EBs were resuspended in fresh CA medium and transferred on a new non-adherent 10 cm (3.5 cm) bacterial dish. For retinoic acid mediated EB differentiation, 5µM retinoic acid (final concentration) was added to CA medium on day 4 and day 6 of differentiation.

6.2.1.3 Antioxidant/low glucose treatments

For antioxidant treatments, regular EB differentiation were performed as described in 6.2.1.2. Antioxidants were added to CA medium with following final concentrations: 50 and 150µM vitamin E (α-tocopherol), 20 and 100 µg/ml vitamin C (L-ascorbic acid 2-phosphate sesquimagnesium salt hydrate), 5 and 20 µM vitamin D_3 (cholecalciferol), 2 and 5 mM N-acetyl-L-cysteine. As negative control, CA medium was supplemented with equal amounts of DMSO (vitamin E+D_3).

6.2.2 General molecular biology

6.2.2.1 Immunoblotting

Cells were resuspended in NOP+ buffer and lysed for 20 min on ice. Cell lysate was cleared by 15 minutes centrifugation at maximum speed and 4°C. Protein concentrations were estimated by BCA assay using BSA as standard. Equal protein amounts were mixed with 4x Lämmli buffer (1x final) and incubated for 5min at 95°C. Protein lysates were separated on a precast 12% Criterion™ XT Bis-Tris protein gel and transferred to a ployvinylidene difluoride (PVDF) membrane. SDS-PAGE, transfer to PVDF membrane and Western blot analysis was done according to standard protocols. Signals were developed with SuperSignal West Pico or Femto Chemiluminescent Substrate and analyzed using a ChemiDoc (Biorad) with ImageLab software.

For mitochondrial, cytosolic and nuclear fractionation, cells were resuspended in 500 µl ice-cold 1x HG buffer. Cell suspension was incubated for 10 min on ice. Subsequently, cells were disrupted with 30 strokes using a glass-glass homogenizer. Cell disruption was monitored with trypan blue under a light microscope. To stop the hypotonic lysis, 62.5 µl of 10x HG buffer were added. Unbroken cells were removed by two centrifugation steps at 1000x g, 3min and 4°C. The first pellet contained the nuclear fraction, supernatant contained cytosolic and mitochondrial fractions. The mitochondrial fraction was separated by centrifugation at 17,000x g for 10 min and 4°C. Supernatant was kept as cytosolic fraction. Pelleted mitochondrial fraction was washed twice with 1x HG buffer and lysed with 1x RIPA buffer. Nuclear fractions were resuspended in 497.5 µl 1x HG buffer plus 2.5 µl of 10% Nonidet P-40 and incubated for 10 min on ice to lyse unbroken cells. Nuclear fractions were separated by centrifugation at 1,200x g for 3 min and 4°C. Nuclear pellet was washed twice with 1x HG buffer and lysed with RIPA buffer. All buffers contained complete 1x complete™ protease inhibitor and 1x PhosStop™.

6.2.2.2 RNA isolation

RNA was isolated using the RNeasy mini kit (Qiagen) with a DNAse I on-colum digest according to the manufacturer`s recommendations. RNA was eluted in 30 or 50µl ddH_2O. Concentrations were measured on a Nanodrop 2000.

6.2.2.3 RT-qPCR

Equal amounts of RNA (1 µg) were used for cDNA synthesis using Superscript II reverse transcriptase according to the manufacturer`s recommendations. 1µg RNA was mixed with 1µl of 5mM dNTPs and 120 ng random primers in a final volume of 12 µl and incubated at 65°C for 5 min. The mixture was cooled to 4°C before 5x FS buffer, 2 µl 0.1 M DTT, 1 µl Ribolock and 1 µl Superscript II Polymerase were added. Samples were incubated at 25°C for 10 min, at 42°C for 90 min and at 72°C for 5 min. Prior qPCR analysis, cDNA was diluted 1:6 in ddH_2O.

For each qPCR reaction, 0.055 µl 100 µM forward primer, 0.055 µl 100 µM reverse primer, 0.11 µl UPL probe and 0.28 µl ddH_2O were mixed. Quantitative real time PCR was performed in technical duplicates in a 384-well format on a LightCycler 480. Following PCR program was used:

	PCR step	Temperature	Time	Ramp rate
1	Denaturation	95°C	10 min	4.8°C/sec
2	Denaturation	95°C	10 s	4.8°C/sec
3	Annealing	60°C	20 s	2.5°C/sec
4	Elongation + Signal acquisition	72°C	1 s	4.8°C/sec
5	Go to step [2], 49x times	-	-	-
6	Cooling	4°C	1 s	2.5°C/sec

Quantitative analysis was carried out with the Roche Lightcycler software. Cp values were determined and calculated according to the ddCp method (Livak and Schmittgen 2001) normalizing to *Tbp* as housekeeping gene.

6.2.3 Flow cytometry analysis

For cell cycle analysis, cells were washed with 1x PBS and resuspended in 0.25% trypsin to obtain single cell suspensions. Trypsinization was quenched by addition of serum-supplemented ESC medium. Cells were spun down for 3min at 300g, washed once with 1x PBS. Cell count was equalized to $2x10^7$ cells. Cells were fixed by adding 3 ml ice-cold ethanol drop-wise to the cell pellet while vortexing. Cell-ethanol mixture was vortexed for additional 15 s and incubated at -20°C for 30 min or stored. For propidium iodide staining, cells were washed twice with 1x PBS containing 100 µl/ml RNAse A and 0.1% ESC grade. Each centrifugation step was done at 4°C and 500x g for 5 min. Cell pellet was resuspended in 1 ml PBS containing 50 µg/ml propidium iodide and 1 mM EDTA and incubated at room temperature for 10 min in the dark. Cell suspension was transferred to flow cytometry tubes and

analysed by the BD LSR-FortessaSORP flow cytometry system using DiVa software. Data analysis was carried out with FlowJo software (version 10.5.3).

6.2.4 Statistics

Statistical significance was tested by the two-tailed unpaired Student´s t-test using the TTEST() function in Excel. Data are represented as mean ± standard deviation from three independent clones. Significances are displayed as $*P < 0.05$, $**P < 0.005$, $P*** < 0.0005$.

7 References

Aapola U, Kawasaki K, Scott HS, Ollila J, Vihinen M, Heino M, Shintani A, Minoshima S, Krohn K, Antonarakis SE, et al. 2000. Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family, *Genomics* **65:** 293–298.

Aguiari P, Leo S, Zavan B, Vindigni V, Rimessi A, Bianchi K, Franzin C, Cortivo R, Rossato M, Vettor R, et al. 2008. High glucose induces adipogenic differentiation of muscle-derived stem cells, *Proc Natl Acad Sci U S A* **105:** 1226–1231.

Aiken CEM, Swoboda PPL, Skepper JN, Johnson MH. 2004. The direct measurement of embryogenic volume and nucleo-cytoplasmic ratio during mouse pre-implantation development, *Reproduction* **128:** 527–535.

Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. 2012. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles, *Genome Biol* **13:** R87.

Amano T, Hirata T, Falco G, Monti M, Sharova LV, Amano M, Sheer S, Hoang HG, Piao Y, Stagg CA, et al. 2013. Zscan4 restores the developmental potency of embryonic stem cells, *Nat Commun* **4:** 1966.

Amouroux R, Nashun B, Shirane K, Nakagawa S, Hill PW, D'Souza Z, Nakayama M, Matsuda M, Turp A, Ndjetehe E, et al. 2016. De novo DNA methylation drives 5hmC accumulation in mouse zygotes, *Nat Cell Biol* **18:** 225–233.

An J, González-Avalos E, Chawla A, Jeong M, López-Moyado IF, Li W, Goodell MA, Chavez L, Ko M, Rao A. 2015. Acute loss of TET function results in aggressive myeloid cancer in mice, *Nat Commun* **6:** 10071.

Aoki F, Worrad DM, Schultz RM. 1997. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo, *Dev Biol* **181:** 296–307.

Arab K, Karaulanov E, Musheev M, Trnka P, Schäfer A, Grummt I, Niehrs C. 2019. GADD45A binds R-loops and recruits TET1 to CpG island promoters, *Nat Genet*.

Arab K, Park YJ, Lindroth AM, Schäfer A, Oakes C, Weichenhan D, Lukanova A, Lundin E, Risch A, Meister M, et al. 2014. Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A, *Mol Cell* **55:** 604–614.

Artus J, Cohen-Tannoudji M. 2008. Cell cycle regulation during early mouse embryogenesis, *Mol Cell Endocrinol* **282:** 78–86.

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function, *Genes Dev* **17:** 126–140.

Barreto G, Schäfer A, Marhold J, Stach D, Swaminathan SK, Handa V, Döderlein G, Maltry N, Wu W, Lyko F, et al. 2007. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation, *Nature* **445:** 671–675.

Bauer C, Göbel K, Nagaraj N, Colantuoni C, Wang M, Müller U, Kremmer E, Rottach A, Leonhardt H. 2015. Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT), *J Biol Chem* **290:** 4801–4812.

Beddington RS, Robertson EJ. 1989. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo, *Development* **105:** 733–737.

Bénazéraf B, Pourquié O. 2013. Formation and segmentation of the vertebrate body axis, *Annu Rev Cell Dev Biol* **29:** 1–26.

Bhattacharya SK, Ramchandani S, Cervoni N, Szyf M. 1999. A mammalian protein with specific demethylase activity for mCpG DNA, *Nature* **397:** 579–583.

Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. 2010. Reprogramming towards pluripotency requires AID-dependent DNA demethylation, *Nature* **463:** 1042–1047.

Bibel M, Richter J, Lacroix E, Barde Y-A. 2007. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells, *Nat Protoc* **2:** 1034–1043.

Bird A. 2002. DNA methylation patterns and epigenetic memory, *Genes Dev* **16:** 6–21.

Boroviak T, Loos R, Bertone P, Smith A, Nichols J. 2014. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification, *Nat Cell Biol* **16:** 516–528.

Bostick M, Kim JK, Estève P-O, Clark A, Pradhan S, Jacobsen SE. 2007. UHRF1 plays a role in maintaining DNA methylation in mammalian cells, *Science* **317:** 1760–1764.

Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. 2001. Dnmt3L and the establishment of maternal genomic imprints, *Science* **294:** 2536–2539.

Bradley A, Evans M, Kaufman MH, Robertson E. 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines, *Nature* **309:** 255–256.

Bradley A, Hasty P, Davis A, Ramirez-Solis R. 1992. Modifying the mouse: design and desire, *Biotechnology (N Y)* **10:** 534–539.

Bulavin DV, Kovalsky O, Hollander MC, Fornace AJ. 2003. Loss of oncogenic H-ras-induced cell cycle arrest and p38 mitogen-activated protein kinase activation by disruption of Gadd45a, *Mol Cell Biol* **23:** 3859–3871.

Cai Q, Dmitrieva NI, Ferraris JD, Michea LF, Salvador JM, Hollander MC, Fornace AJ, Fenton RA, Burg MB. 2006. Effects of expression of p53 and Gadd45 on osmotic tolerance of renal inner medullary cells, *Am J Physiol Renal Physiol* **291:** F341-9.

Calo E, Gu B, Bowen ME, Aryan F, Zalc A, Liang J, Flynn RA, Swigut T, Chang HY, Attardi LD, et al. 2018. Tissue-selective effects of nucleolar stress and rDNA damage in developmental disorders, *Nature* **554:** 112–117.

Canugovi C, Yoon JS, Feldman NH, Croteau DL, Mattson MP, Bohr VA. 2012. Endonuclease VIII-like 1 (NEIL1) promotes short-term spatial memory retention and protects from ischemic stroke-induced brain dysfunction and death in mice, *Proc Natl Acad Sci U S A* **109:** 14948–14953.

Carlson LL, Page AW, Bestor TH. 1992. Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting, *Genes Dev* **6:** 2536–2541.

Carrier F, Georgel PT, Pourquier P, Blake M, Kontny HU, Antinore MJ, Gariboldi M, Myers TG, Weinstein JN, Pommier Y, et al. 1999. Gadd45, a p53-responsive stress protein, modifies DNA accessibility on damaged chromatin, *Mol Cell Biol* **19:** 1673–1685.

Cedar H, Solage A, Glaser G, Razin A. 1979. Direct detection of methylated cytosine in DNA by use of the restriction enzyme MspI, *Nucleic Acids Res* **6:** 2125–2132.

Chakraborty A, Wakamiya M, Venkova-Canova T, Pandita RK, Aguilera-Aguirre L, Sarker AH, Singh DK, Hosoki K, Wood TG, Sharma G, et al. 2015. Neil2-null Mice Accumulate Oxidized DNA Bases in the Transcriptionally Active Sequences of the Genome and Are Susceptible to Innate Inflammation, *J Biol Chem* **290:** 24636–24648.

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. 2003. Functional Expression Cloning of Nanog, a Pluripotency Sustaining Factor in Embryonic Stem Cells, *Cell* **113:** 643–655.

Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A. 2007. Nanog safeguards pluripotency and mediates germline development, *Nature* **450:** 1230–1234.

Chang C-C, Ma Y, Jacobs S, Tian XC, Yang X, Rasmussen TP. 2005. A maternal store of macroH2A is removed from pronuclei prior to onset of somatic macroH2A expression in preimplantation embryos, *Dev Biol* **278:** 367–380.

Chazaud C, Yamanaka Y. 2016. Lineage specification in the mouse preimplantation embryo, *Development* **143:** 1063–1074.

Chen X, Guo C, Kong J. 2012. Oxidative stress in neurodegenerative diseases, *Neural Regen Res* **7:** 376–385.

Chen Z, Zhang Y. 2019. Loss of DUX causes minor defects in zygotic genome activation and is compatible with mouse development, *Nat Genet* **51:** 947–951.

CHO T, SAKAI S, NAGATA M, AOKI F. 2002. Involvement of chromatin structure in the regulation of mouse zygotic gene activation, *Animal Science Journal* **73:** 113–122.

Chodavarapu RK, Feng S, Bernatavichute YV, Chen P-Y, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, et al. 2010. Relationship between nucleosome positioning and DNA methylation, *Nature* **466:** 388–392.

Chung HK, Yi Y-W, Jung N-C, Kim D, Suh JM, Kim H, Park KC, Song JH, Kim DW, Hwang ES, et al. 2003. CR6-interacting factor 1 interacts with Gadd45 family proteins and modulates the cell cycle, *J Biol Chem* **278:** 28079–28088.

Cimmino L, Dawlaty MM, Ndiaye-Lobry D, Yap YS, Bakogianni S, Yu Y, Bhattacharyya S, Shaknovich R, Geng H, Lobry C, et al. 2015. TET1 is a tumor suppressor of hematopoietic malignancy, *Nat Immunol* **16:** 653–662.

Clarkson PM, Thompson HS. 2000. Antioxidants: what role do they play in physical activity and health?, *Am J Clin Nutr* **72:** 637S-46S.

Cortázar D, Kunz C, Selfridge J, Lettieri T, Saito Y, MacDougall E, Wirz A, Schuermann D, Jacobs AL, Siegrist F, et al. 2011. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability, *Nature* **470:** 419–423.

Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, Li G, Kuan S, Li B, Lee AY, et al. 2016. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition, *Nature* **537:** 548–552.

Dai H-Q, Wang B-A, Yang L, Chen J-J, Zhu G-C, Sun M-L, Ge H, Wang R, Chapman DL, Tang F, et al. 2016. TET-mediated DNA demethylation controls gastrulation by regulating Lefty-Nodal signalling, *Nature* **538:** 528–532.

Dani C, Smith AG, Dessolin S, Leroy P, Staccini L, Villageois P, Darimont C, Ailhaud G. 1997. Differentiation of embryonic stem cells into adipocytes in vitro, *J Cell Sci* **110 (Pt 11):** 1279–1285.

Dawlaty MM, Breiling A, Le T, Barrasa MI, Raddatz G, Gao Q, Powell BE, Cheng AW, Faull KF, Lyko F, et al. 2014. Loss of Tet enzymes compromises proper differentiation of embryonic stem cells, *Dev Cell* **29:** 102–111.

Dean W. 2008. The elusive Dnmt1 and its role during early development, *Epigenetics* **3:** 175–178.

Deaton AM, Bird A. 2011. CpG islands and the regulation of transcription, *Genes Dev* **25:** 1010–1022.

Dizdaroglu M. 2015. Oxidatively induced DNA damage and its repair in cancer, *Mutat Res Rev Mutat Res* **763:** 212–245.

Doetschman TC, Eistetter H, Katz M, Schmidt W, Kemler R. 1985. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium, *J Embryol Exp Morphol* **87:** 27–45.

Eckersley-Maslin M, Alda-Catalinas C, Blotenburg M, Kreibich E, Krueger C, Reik W. 2019. Dppa2 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program, *Genes Dev*.

Eckersley-Maslin MA, Alda-Catalinas C, Reik W. 2018. Dynamics of the epigenetic landscape during the maternal-to-zygotic transition, *Nat Rev Mol Cell Biol*.

Eckersley-Maslin MA, Svensson V, Krueger C, Stubbs TM, Giehr P, Krueger F, Miragaia RJ, Kyriakopoulos C, Berrens RV, Milagre I, et al. 2016. MERVL/Zscan4 Network Activation Results in Transient Genome-wide DNA Demethylation of mESCs, *Cell Rep* **17:** 179–192.

Elnakish MT, Hassanain HH, Janssen PM, Angelos MG, Khan M. 2013. Emerging role of oxidative stress in metabolic syndrome and cardiovascular diseases: important role of Rac/NADPH oxidase, *J Pathol* **231:** 290–300.

Evans MJ, Kaufman MH. 1981. Establishment in culture of pluripotential cells from mouse embryos, *Nature* **292:** 154–156.

Facucho-Oliveira JM, Alderson J, Spikings EC, Egginton S, St John JC. 2007. Mitochondrial DNA replication during differentiation of murine embryonic stem cells, *J Cell Sci* **120:** 4025–4034.

Fatemi M, Hermann A, Pradhan S, Jeltsch A. 2001. The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA, *J Mol Biol* **309:** 1189–1199.

Ficz G, Hore TA, Santos F, Lee HJ, Dean W, Arand J, Krueger F, Oxley D, Paul Y-L, Walter J, et al. 2013. FGF Signaling Inhibition in ESCs Drives Rapid Genome-wide Demethylation to the Epigenetic Ground State of Pluripotency, *Cell Stem Cell* **13:** 351–359.

Fischer M. 2017. Census and evaluation of p53 target genes, *Oncogene* **36:** 3943–3956.

Fleck O, Nielsen O. 2004. DNA repair, *J Cell Sci* **117:** 515–517.

Friedberg EC, Meira LB. 2006. Database of mouse strains carrying targeted mutations in genes affecting biological responses to DNA damage Version 7, *DNA Repair (Amst)* **5:** 189–209.

Fu G, Ghadam P, Sirotkin A, Khochbin S, Skoultchi AI, Clarke HJ. 2003. Mouse oocytes and early embryos express multiple histone H1 subtypes, *Biol Reprod* **68:** 1569–1576.

Fu X, Cui K, Yi Q, Yu L, Xu Y. 2017. DNA repair mechanisms in embryonic stem cells, *Cell Mol Life Sci* **74:** 487–493.

Fujita J, Crane AM, Souza MK, Dejosez M, Kyba M, Flavell RA, Thomson JA, Zwaka TP. 2008. Caspase activity mediates the differentiation of embryonic stem cells, *Cell Stem Cell* **2:** 595–601.

Gierl MS, Gruhn WH, Seggern A von, Maltry N, Niehrs C. 2012. GADD45G functions in male sex determination by promoting p38 signaling and Sry expression, *Dev Cell* **23:** 1032–1042.

Gopalakrishnan S, Sullivan BA, Trazzi S, Della Valle G, Robertson KD. 2009. DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions, *Hum Mol Genet* **18:** 3178–3193.

Gu T-P, Guo F, Yang H, Wu H-P, Xu G-F, Liu W, Xie Z-G, Shi L, He X, Jin S-G, et al. 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes, *Nature* **477:** 606–610.

Guo F, Li X, Liang D, Li T, Zhu P, Guo H, Wu X, Wen L, Gu T-P, Hu B, et al. 2014. Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote, *Cell Stem Cell* **15:** 447–459.

Habibi E, Brinkman AB, Arand J, Kroeze LI, Kerstens HHD, Matarese F, Lepikhov K, Gut M, Brun-Heath I, Hubner NC, et al. 2013. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells, *Cell Stem Cell* **13:** 360–369.

Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, Surani MA. 2013. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine, *Science* **339:** 448–452.

Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA. 2010. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway, *Science* **329:** 78–82.

Hamatani T, Carter MG, Sharov AA, Ko MSH. 2004. Dynamics of global gene expression changes during mouse preimplantation development, *Dev Cell* **6:** 117–131.

Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, Zhang X, Cheng X. 2012. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation, *Nucleic Acids Res* **40:** 4841–4849.

Hata K, Okano M, Lei H, Li E. 2002. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice, *Development* **129:** 1983–1993.

Hayashi K, Sousa Lopes SMC de, Tang F, Lao K, Surani MA. 2008. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states, *Cell Stem Cell* **3:** 391–401.

Hayashi Y, Furue MK, Tanaka S, Hirose M, Wakisaka N, Danno H, Ohnuma K, Oeda S, Aihara Y, Shiota K, et al. 2010. BMP4 induction of trophoblast from mouse embryonic stem cells in defined culture conditions on laminin, *In Vitro Cell Dev Biol Anim* **46:** 416–430.

He Y-F, Li B-Z, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, et al. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA, *Science* **333:** 1303–1307.

Hemann MT, Lowe SW. 2006. The p53-Bcl-2 connection, *Cell Death Differ* **13:** 1256–1259.

Hendrickson PG, Doráis JA, Grow EJ, Whiddon JL, Lim J-W, Wike CL, Weaver BD, Pflueger C, Emery BR, Wilcox AL, et al. 2017. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons, *Nat Genet* **49:** 925–934.

Hermann A, Goyal R, Jeltsch A. 2004. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites, *J Biol Chem* **279:** 48350–48359.

Hildesheim J, Bulavin DV, Anver MR, Alvord WG, Hollander MC, Vardanian L, Fornace AJ. 2002. Gadd45a protects against UV irradiation-induced skin tumors, and promotes apoptosis and stress signaling via MAPK and p53, *Cancer Res* **62:** 7305–7315.

Hollander MC, Sheikh MS, Bulavin DV, Lundgren K, Augeri-Henmueller L, Shehee R, Molinaro TA, Kim KE, Tolosa E, Ashwell JD, et al. 1999. Genomic instability in Gadd45a-deficient mice, *Nat Genet* **23:** 176–184.

Huang HS, Kubish GM, Redmond TM, Turner DL, Thompson RC, Murphy GG, Uhler MD. 2010. Direct transcriptional induction of Gadd45gamma by Ascl1 during neuronal differentiation, *Mol Cell Neurosci* **44:** 282–296.

Iaco A de, Planet E, Coluccio A, Verp S, Duc J, Trono D. 2017. DUX-family transcription factors regulate zygotic genome activation in placental mammals, *Nat Genet* **49:** 941–945.

Inoue A, Shen L, Dai Q, He C, Zhang Y. 2011. Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development, *Cell Res* **21:** 1670–1676.

Inoue A, Zhang Y. 2011. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos, *Science* **334:** 194.

Iqbal K, Jin S-G, Pfeifer GP, Szabó PE. 2011. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine, *Proc Natl Acad Sci U S A* **108:** 3642–3647.

Ishiuchi T, Enriquez-Gasca R, Mizutani E, Bošković A, Ziegler-Birling C, Rodriguez-Terrones D, Wakayama T, Vaquerizas JM, Torres-Padilla M-E. 2015. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly, *Nat Struct Mol Biol* **22:** 662–671.

Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification, *Nature* **466:** 1129–1133.

Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine, *Science* **333:** 1300–1303.

Jacobs AL, Schär P. 2012. DNA glycosylases: in DNA repair and beyond, *Chromosoma* **121:** 1–20.

Jagannathan L, Cuddapah S, Costa M. 2016. Oxidative stress under ambient and physiological oxygen tension in tissue culture, *Curr Pharmacol Rep* **2:** 64–72.

Jain AK, Barton MC. 2018. p53: emerging roles in stem cells, development and beyond, *Development* **145**.

Jang HS, Shin WJ, Lee JE, Do JT. 2017. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function, *Genes (Basel)* **8**.

Jessop P, Ruzov A, Gering M. 2018. Developmental Functions of the Dynamic DNA Methylome and Hydroxymethylome in the Mouse and Zebrafish: Similarities and Differences, *Front Cell Dev Biol* **6:** 27.

Jiang D, Wei S, Chen F, Zhang Y, Li J. 2017. TET3-mediated DNA oxidation promotes ATR-dependent DNA damage response, *EMBO Rep* **18:** 781–796.

Jin S-G, Guo C, Pfeifer GP. 2008. GADD45A does not promote DNA demethylation, *PLoS Genet* **4:** e1000013.

Jones NC, Lynn ML, Gaudenz K, Sakai D, Aoto K, Rey J-P, Glynn EF, Ellington L, Du C, Dixon J, et al. 2008. Prevention of the neurocristopathy Treacher Collins syndrome through inhibition of p53 function, *Nat Med* **14:** 125–133.

Jones PA. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond, *Nat Rev Genet* **13:** 484–492.

Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription, *Nat Genet* **19:** 187–191.

Jost JP, Siegmann M, Sun L, Leung R. 1995. Mechanisms of DNA demethylation in chicken embryos. Purification and properties of a 5-methylcytosine-DNA glycosylase, *J Biol Chem* **270:** 9734–9739.

Kafer GR, Li X, Horii T, Suetake I, Tajima S, Hatada I, Carlton PM. 2016. 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability, *Cell Rep* **14:** 1283–1292.

Kagiwada S, Kurimoto K, Hirota T, Yamaji M, Saitou M. 2013. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice, *EMBO J* **32:** 340–353.

Kaufmann LT, Niehrs C. 2011. Gadd45a and Gadd45g regulate neural development and exit from pluripotency in Xenopus, *Mech Dev* **128:** 401–411.

Keller G. 2005. Embryonic stem cell differentiation. Emergence of a new era in biology and medicine, *Genes Dev* **19:** 1129–1155.

Kelly SJ. 1977. Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres, *J Exp Zool* **200:** 365–376.

Kerksick C, Willoughby D. 2005. The antioxidant role of glutathione and N-acetyl-cysteine supplements and exercise-induced oxidative stress, *J Int Soc Sports Nutr* **2:** 38–44.

Khacho M, Slack RS. 2018. Mitochondrial dynamics in the regulation of neurogenesis: From development to the adult brain, *Dev Dyn* **247:** 47–53.

Khoueiry R, Sohni A, Thienpont B, Luo X, Velde JV, Bartoccetti M, Boeckx B, Zwijsen A, Rao A, Lambrechts D, et al. 2017. Lineage-specific functions of TET1 in the postimplantation mouse embryo, *Nat Genet* **49:** 1061–1072.

Kienhöfer S, Musheev MU, Stapf U, Helm M, Schomacher L, Niehrs C, Schäfer A. 2015. GADD45a physically and functionally interacts with TET1, *Differentiation* **90:** 59–68.

Kim HS, Tan Y, Ma W, Merkurjev D, Destici E, Ma Q, Suter T, Ohgi K, Friedman M, Skowronska-Krawczyk D, et al. 2018. Pluripotency factors functionally premark cell-type-restricted enhancers in ES cells, *Nature* **556:** 510–514.

Kim M, Costello J. 2017. DNA methylation. An epigenetic mark of cellular memory, *Exp Mol Med* **49:** e322.

Kinney SM, Chin HG, Vaisvila R, Bitinaite J, Zheng Y, Estève P-O, Feng S, Stroud H, Jacobsen SE, Pradhan S. 2011. Tissue-specific distribution and dynamic changes of 5-hydroxymethylcytosine in mammalian genomes, *J Biol Chem* **286:** 24685–24693.

Kobayashi H, Sakurai T, Imai M, Takahashi N, Fukuda A, Yayoi O, Sato S, Nakabayashi K, Hata K, Sotomaru Y, et al. 2012. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks, *PLoS Genet* **8:** e1002440.

Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G, et al. 2011. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells, *Cell Stem Cell* **8:** 200–213.

Kohli RM, Zhang Y. 2013. TET enzymes, TDG and the dynamics of DNA demethylation, *Nature* **502:** 472–479.

Krebs AR, Dessus-Babus S, Burger L, Schübeler D. 2014. High-throughput engineering of a mammalian genome reveals building principles of methylation states at CG rich regions, *Elife* **3:** e04094.

Kriaucionis S, Heintz N. 2009. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain, *Science* **324:** 929–930.

Krokeide SZ, Laerdahl JK, Salah M, Luna L, Cederkvist FH, Fleming AM, Burrows CJ, Dalhus B, Bjørås M. 2013. Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroimindiohydantoin and guanidinohydantoin, *DNA Repair (Amst)* **12:** 1159–1164.

Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith A. 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment, *Development* **134:** 2895–2902.

Kunz C, Saito Y, Schär P. 2009. DNA Repair in mammalian cells: Mismatched repair: variations on a theme, *Cell Mol Life Sci* **66:** 1021–1038.

Law JA, Jacobsen SE. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals, *Nat Rev Genet* **11:** 204–220.

Leach PT, Poplawski SG, Kenney JW, Hoffman B, Liebermann DA, Abel T, Gould TJ. 2012. Gadd45b knockout mice exhibit selective deficits in hippocampus-dependent long-term memory, *Learn Mem* **19:** 319–324.

Lee B, Morano A, Porcellini A, Muller MT. 2012. GADD45α inhibition of DNMT1 dependent DNA methylation during homology directed DNA repair, *Nucleic Acids Res* **40:** 2481–2493.

Lee HJ, Hore TA, Reik W. 2014. Reprogramming the methylome: erasing memory and creating diversity, *Cell Stem Cell* **14:** 710–719.

Lee MT, Bonneau AR, Giraldez AJ. 2014. Zygotic genome activation during the maternal-to-zygotic transition, *Annu Rev Cell Dev Biol* **30:** 581–613.

Leitch HG, McEwen KR, Turp A, Encheva V, Carroll T, Grabole N, Mansfield W, Nashun B, Knezovich JG, Smith A, et al. 2013. Naive pluripotency is associated with global DNA hypomethylation, *Nat Struct Mol Biol* **20:** 311–316.

Leonhardt H, Page AW, Weier HU, Bestor TH. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei, *Cell* **71:** 865–873.

Li Z, Gu T-P, Weber AR, Shen J-Z, Li B-Z, Xie Z-G, Yin R, Guo F, Liu X, Tang F, et al. 2015. Gadd45a promotes DNA demethylation through TDG, *Nucleic Acids Res* **43:** 3986–3997.

Liebermann DA, Hoffman B. 2008. Gadd45 in stress signaling, *J Mol Signal* **3:** 15.

Lin T, Chao C, Saito S'i, Mazur SJ, Murphy ME, Appella E, Xu Y. 2005. p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression, *Nat Cell Biol* **7:** 165–171.

Lindahl T, Wood RD. 1999. Quality control by DNA repair, *Science* **286:** 1897–1905.

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo Q-M, et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences, *Nature* **462:** 315–322.

Liu L, Michowski W, Inuzuka H, Shimizu K, Nihira NT, Chick JM, Li N, Geng Y, Meng AY, Ordureau A, et al. 2017. G1 cyclins link proliferation, pluripotency and differentiation of embryonic stem cells, *Nat Cell Biol* **19:** 177–188.

Liu L, Tran E, Zhao Y, Huang Y, Flavell R, Lu B. 2005. Gadd45 beta and Gadd45 gamma are critical for regulating autoimmunity, *J Exp Med* **202:** 1341–1347.

Liu Y, Liu K, Wang N, Zhang H. 2017. N-acetylcysteine induces apoptosis via the mitochondria-dependent pathway but not via endoplasmic reticulum stress in H9c2 cells, *Mol Med Rep* **16:** 6626–6633.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, *Methods* **25:** 402–408.

Lu B, Ferrandino AF, Flavell RA. 2004. Gadd45beta is important for perpetuating cognate and inflammatory signals in T cells, *Nat Immunol* **5:** 38–44.

Lu B, Yu H, Chow C, Li B, Zheng W, Davis RJ, Flavell RA. 2001. GADD45gamma mediates the activation of the p38 and JNK MAP kinase pathways and cytokine production in effector TH1 cells, *Immunity* **14:** 583–590.

Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. 2014. Role of Tet proteins in enhancer activity and telomere elongation, *Genes Dev* **28:** 2103–2119.

Lukaszewicz AI, Anderson DJ. 2011. Cyclin D1 promotes neurogenesis in the developing spinal cord in a cell cycle-independent manner, *Proc Natl Acad Sci U S A* **108:** 11632–11637.

Ma DK, Guo JU, Ming G-l, Song H. 2009. DNA excision repair proteins and Gadd45 as molecular players for active DNA demethylation, *Cell Cycle* **8:** 1526–1531.

Ma DK, Jang M-H, Guo JU, Kitabatake Y, Chang M-l, Pow-anpongkul N, Flavell RA, Lu B, Ming G-l, Song H. 2009. Neuronal Activity-Induced Gadd45b Promotes Epigenetic DNA Demethylation and Adult Neurogenesis, *Science* **323:** 1074–1077.

Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. 2012a. Embryonic stem cell potency fluctuates with endogenous retrovirus activity, *Nature* **487:** 57–63.

Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. 2012b. Embryonic stem cell potency fluctuates with endogenous retrovirus activity, *Nature* **487:** 57–63.

Magnúsdóttir E, Dietmann S, Murakami K, Günesdogan U, Tang F, Bao S, Diamanti E, Lao K, Gottgens B, Azim Surani M. 2013. A tripartite transcription factor network regulates primordial germ cell specification in mice, *Nat Cell Biol* **15:** 905–915.

Maiti A, Drohat AC. 2011. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites, *J Biol Chem* **286:** 35334–35338.

Martin GR. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells, *Proc Natl Acad Sci U S A* **78:** 7634–7638.

Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. 2000. Demethylation of the zygotic paternal genome, *Nature* **403:** 501–502.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells, *Nature* **454:** 766–770.

Messerschmidt DM, Knowles BB, Solter D. 2014. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos, *Genes Dev* **28:** 812–828.

Métivier R, Gallais R, Tiffoche C, Le Péron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G, et al. 2008. Cyclical DNA methylation of a transcriptionally active promoter, *Nature* **452:** 45–50.

Millman JR, Tan JH, Colton CK. 2009. The effects of low oxygen on self-renewal and differentiation of embryonic stem cells, *Curr Opin Organ Transplant* **14:** 694–700.

Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. 2004. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming, *J Biol Chem* **279:** 52353–52360.

Müller U, Bauer C, Siegl M, Rottach A, Leonhardt H. 2014. TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation, *Nucleic Acids Res* **42:** 8592–8604.

Nakagawa T, Lv L, Nakagawa M, Yu Y, Yu C, D'Alessio AC, Nakayama K, Fan H-Y, Chen X, Xiong Y. 2015. CRL4(VprBP) E3 ligase promotes monoubiquitylation and chromatin binding of TET dioxygenases, *Mol Cell* **57:** 247–260.

Nakamura T, Liu Y-J, Nakashima H, Umehara H, Inoue K, Matoba S, Tachibana M, Ogura A, Shinkai Y, Nakano T. 2012. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos, *Nature* **486:** 415–419.

Nakatani T, Yamagata K, Kimura T, Oda M, Nakashima H, Hori M, Sekita Y, Arakawa T, Nakamura T, Nakano T. 2015. Stella preserves maternal chromosome integrity by inhibiting 5hmC-induced γH2AX accumulation, *EMBO Rep* **16:** 582–589.

Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex, *Nature* **393:** 386–389.

Narducci MG, Fiorenza MT, Kang S-M, Bevilacqua A, Di Giacomo M, Remotti D, Picchio MC, Fidanza V, Cooper MD, Croce CM, et al. 2002. TCL1 participates in early embryonic development and is overexpressed in human seminomas, *Proc Natl Acad Sci U S A* **99:** 11712–11717.

Ng RK, Dean W, Dawson C, Lucifero D, Madeja Z, Reik W, Hemberger M. 2008. Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5, *Nat Cell Biol* **10:** 1280–1290.

Niwa H, Miyazaki J-i, Smith AG. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells, *Nat Genet* **24:** 372–376.

Nonchev S, Tsanev R. 1990. Protamine-histone replacement and DNA replication in the male mouse pronucleus, *Mol Reprod Dev* **25:** 72–76.

Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development, *Cell* **99:** 247–257.

Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin S-P, Allis CD, et al. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA, *Nature* **448:** 714–717.

Pauklin S, Madrigal P, Bertero A, Vallier L. 2016. Initiation of stem cell differentiation involves cell cycle-dependent regulation of developmental genes by Cyclin D, *Genes Dev* **30:** 421–433.

Peat JR, Dean W, Clark SJ, Krueger F, Smallwood SA, Ficz G, Kim JK, Marioni JC, Hore TA, Reik W. 2014. Genome-wide bisulfite sequencing in zygotes identifies demethylation targets and maps the contribution of TET3 oxidation, *Cell Rep* **9:** 1990–2000.

Percharde M, Lin C-J, Yin Y, Guan J, Peixoto GA, Bulut-Karslioglu A, Biechele S, Huang B, Shen X, Ramalho-Santos M. 2018. A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity, *Cell* **174:** 391-405.e19.

Pfaffeneder T, Hackner B, Truss M, Münzel M, Müller M, Deiml CA, Hagemeier C, Carell T. 2011. The discovery of 5-formylcytosine in embryonic stem cell DNA, *Angew Chem Int Ed Engl* **50:** 7008–7012.

Powers DE, Millman JR, Huang RB, Colton CK. 2008. Effects of oxygen on mouse embryonic stem cell growth, phenotype retention, and cellular energetics, *Biotechnol Bioeng* **101:** 241–254.

Prakash A, Doublié S. 2015. Base Excision Repair in the Mitochondria, *J Cell Biochem* **116:** 1490–1499.

Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR. 2008. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45, *Cell* **135:** 1201–1212.

Raiber E-A, Murat P, Chirgadze DY, Beraldi D, Luisi BF, Balasubramanian S. 2015. 5-Formylcytosine alters the structure of the DNA double helix, *Nat Struct Mol Biol* **22:** 44–49.

Rajput P, Pandey V, Kumar V. 2016. Stimulation of ribosomal RNA gene promoter by transcription factor Sp1 involves active DNA demethylation by Gadd45-NER pathway, *Biochim Biophys Acta* **1859:** 953–963.

Regnell CE, Hildrestrand GA, Sejersted Y, Medin T, Moldestad O, Rolseth V, Krokeide SZ, Suganthan R, Luna L, Bjørås M, et al. 2012. Hippocampal adult neurogenesis is maintained by Neil3-dependent repair of oxidative DNA lesions in neural progenitor cells, *Cell Rep* **2:** 503–510.

Rinon A, Molchadsky A, Nathan E, Yovel G, Rotter V, Sarig R, Tzahor E. 2011. p53 coordinates cranial neural crest cell growth and epithelial-mesenchymal transition/delamination processes, *Development* **138:** 1827–1838.

Rocha S, Martin AM, Meek DW, Perkins ND. 2003. p53 represses cyclin D1 transcription through down regulation of Bcl-3 and inducing increased association of the p52 NF-kappaB subunit with histone deacetylase 1, *Mol Cell Biol* **23:** 4713–4727.

Rolseth V, Luna L, Olsen AK, Suganthan R, Scheffler K, Neurauter CG, Esbensen Y, Kuśnierczyk A, Hildrestrand GA, Graupner A, et al. 2017. No cancer predisposition or increased spontaneous mutation frequencies in NEIL DNA glycosylases-deficient mice, *Sci Rep* **7:** 4384.

Santos F, Peat J, Burgess H, Rada C, Reik W, Dean W. 2013. Active demethylation in mouse zygotes involves cytosine deamination and base excision repair, *Epigenetics Chromatin* **6:** 39.

Schäfer A, Karaulanov E, Stapf U, Döderlein G, Niehrs C. 2013. Ing1 functions in DNA demethylation by directing Gadd45a to H3K4me3, *Genes Dev* **27:** 261–273.

Schäfer A, Mekker B, Mallick M, Vastolo V, Karaulanov E, Sebastian D, Lippen C von der, Epe B, Downes DJ, Scholz C, et al. 2018. Impaired DNA demethylation of C/EBP sites causes premature aging, *Genes Dev* **32:** 742–762.

Schäfer A, Schomacher L, Barreto G, Döderlein G, Niehrs C. 2010. Gemcitabine functions epigenetically by inhibiting repair mediated DNA demethylation, *PLoS ONE* **5:** e14060.

Schmidt CS, Bultmann S, Meilinger D, Zacher B, Tresch A, Maier KC, Peter C, Martin DE, Leonhardt H, Spada F. 2012. Global DNA hypomethylation prevents consolidation of differentiation programs and allows reversion to the embryonic stem cell state, *PLoS ONE* **7:** e52629.

Schmitz K-M, Schmitt N, Hoffmann-Rohrer U, Schäfer A, Grummt I, Mayer C. 2009. TAF12 Recruits Gadd45a and the Nucleotide Excision Repair Complex to the Promoter of rRNA Genes Leading to Active DNA Demethylation, *Mol Cell* **33:** 344–353.

Schomacher L, Han D, Musheev MU, Arab K, Kienhöfer S, Seggern A von, Niehrs C. 2016. Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation, *Nat Struct Mol Biol* **23:** 116–124.

Schomacher L, Niehrs C. 2017. DNA repair and erasure of 5-methylcytosine in vertebrates, *Bioessays* **39**.

Schüle KM, Leichsenring M, Andreani T, Vastolo V, Mallick M, Musheev MU, Karaulanov E, Niehrs C. 2019. GADD45 promotes locus-specific DNA demethylation and 2C cycling in embryonic stem cells, *Genes Dev* **33:** 782–798.

Schuler M, Bossy-Wetzel E, Goldstein JC, Fitzgerald P, Green DR. 2000. p53 induces apoptosis by caspase activation through mitochondrial cytochrome c release, *J Biol Chem* **275:** 7337–7342.

Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W. 2012. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells, *Mol Cell* **48:** 849–862.

Sejersted Y, Hildrestrand GA, Kunke D, Rolseth V, Krokeide SZ, Neurauter CG, Suganthan R, Atneosen-Åsegg M, Fleming AM, Saugstad OD, et al. 2011. Endonuclease VIII-like 3 (Neil3) DNA glycosylase promotes neurogenesis induced by hypoxia-ischemia, *Proc Natl Acad Sci U S A* **108:** 18802–18807.

Semlow DR, Zhang J, Budzowska M, Drohat AC, Walter JC. 2016. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase, *Cell* **167:** 498-511.e14.

Sena LA, Chandel NS. 2012. Physiological roles of mitochondrial reactive oxygen species, *Mol Cell* **48:** 158–167.

Shen L, Inoue A, He J, Liu Y, Lu F, Zhang Y. 2014. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes, *Cell Stem Cell* **15:** 459–471.

Sim Y-J, Kim M-S, Nayfeh A, Yun Y-J, Kim S-J, Park K-T, Kim C-H, Kim K-S. 2017. 2i Maintains a Naive Ground State in ESCs through Two Distinct Epigenetic Mechanisms, *Stem Cell Reports* **8:** 1312–1328.

Singh AM, Hamazaki T, Hankowski KE, Terada N. 2007. A heterogeneous expression pattern for Nanog in embryonic stem cells, *Stem Cells* **25:** 2534–2542.

Smith AG. 1991. Culture and differentiation of embryonic stem cells, *Journal of Tissue Culture Methods* **13:** 89–94.

Smith AG. 2001. Embryo-derived stem cells: of mice and men, *Annu Rev Cell Dev Biol* **17:** 435–462.

Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides, *Nature* **336:** 688–690.

Smith ML, Chen IT, Zhan Q, Bae I, Chen CY, Gilmer TM, Kastan MB, O'Connor PM, Fornace AJ. 1994. Interaction of the p53-regulated protein Gadd45 with proliferating cell nuclear antigen, *Science* **266:** 1376–1380.

Smith ZD, Meissner A. 2013. DNA methylation. Roles in mammalian development, *Nat Rev Genet* **14:** 204–220.

Sohal RS, Orr WC. 2012. The redox stress hypothesis of aging, *Free Radic Biol Med* **52:** 539–555.

Sokol SY. 2011. Maintaining embryonic stem cell pluripotency with Wnt signaling, *Development* **138:** 4341–4350.

Song C-X, Szulwach KE, Dai Q, Fu Y, Mao S-Q, Lin L, Street C, Li Y, Poidevin M, Wu H, et al. 2013. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming, *Cell* **153:** 678–691.

Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PWTC, Bauer C, Münzel M, Wagner M, Müller M, Khan F, et al. 2013. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives, *Cell* **152:** 1146–1159.

Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, et al. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions, *Nature* **480:** 490–495.

Subramaniyan SD, Natarajan AK. 2017. Citral, A Monoterpene Protect Against High Glucose Induced Oxidative Injury in HepG2 Cell In Vitro-An Experimental Study, *J Clin Diagn Res* **11:** BC10-BC15.

Sultan FA, Wang J, Tront J, Liebermann DA, Sweatt JD. 2012. Genetic deletion of Gadd45b, a regulator of active DNA demethylation, enhances long-term memory and synaptic plasticity, *J Neurosci* **32:** 17059–17066.

Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics, *Nat Rev Genet* **9:** 465–476.

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, et al. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, *Science* **324:** 930–935.

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors, *Cell* **126:** 663–676.

Takekawa M, Saito H. 1998. A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK, *Cell* **95:** 521–530.

Tam PPL, Loebel DAF. 2007. Gene function in mouse embryogenesis: get set for gastrulation, *Nat Rev Genet* **8:** 368–381.

Tichy ED. 2011. Mechanisms maintaining genomic integrity in embryonic stem cells and induced pluripotent stem cells, *Exp Biol Med (Maywood)* **236:** 987–996.

Tini M, Benecke A, Um S-J, Torchia J, Evans RM, Chambon P. 2002. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription, *Mol Cell* **9:** 265–277.

Tormos KV, Anso E, Hamanaka RB, Eisenbart J, Joseph J, Kalyanaraman B, Chandel NS. 2011. Mitochondrial complex III ROS regulate adipocyte differentiation, *Cell Metab* **14:** 537–544.

Torres-Padilla M-E, Parfitt D-E, Kouzarides T, Zernicka-Goetz M. 2007. Histone arginine methylation regulates pluripotency in the early mouse embryo, *Nature* **445:** 214–218.

Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H. 2008. Identification and characterization of subpopulations in undifferentiated ES cell culture, *Development* **135:** 909–918.

Tsujimoto Y. 1998. Role of Bcl-2 family proteins in apoptosis: apoptosomes or mitochondria?, *Genes Cells* **3:** 697–707.

Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S-i, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, et al. 2006. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b, *Genes Cells* **11:** 805–814.

Vairapandi M, Balliet AG, Hoffman B, Liebermann DA. 2002. GADD45b and GADD45g are cdc2/cyclinB1 kinase inhibitors with a role in S and G2/M cell cycle checkpoints induced by genotoxic stress, *J Cell Physiol* **192:** 327–338.

van Nostrand JL, Brady CA, Jung H, Fuentes DR, Kozak MM, Johnson TM, Lin C-Y, Lin C-J, Swiderski DL, Vogel H, et al. 2014. Inappropriate p53 activation during development induces features of CHARGE syndrome, *Nature* **514:** 228–232.

Vartanian V, Lowell B, Minko IG, Wood TG, Ceci JD, George S, Ballinger SW, Corless CL, McCullough AK, Lloyd RS. 2006. The metabolic syndrome resulting from a knockout of the NEIL1 DNA glycosylase, *Proc Natl Acad Sci U S A* **103:** 1864–1869.

Vincent JJ, Huang Y, Chen P-Y, Feng S, Calvopiña JH, Nee K, Lee SA, Le T, Yoon AJ, Faull K, et al. 2013. Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells, *Cell Stem Cell* **12:** 470–478.

Walsh CP, Chaillet JR, Bestor TH. 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation, *Nat Genet* **20:** 116–117.

Wang J, Wang L, Feng G, Wang Y, Li Y, Li X, Liu C, Jiao G, Huang C, Shi J, et al. 2018. Asymmetric Expression of LincGET Biases Cell Fate in Two-Cell Mouse Embryos, *Cell* **175:** 1887-1901.e18.

Wang L, Zhou Y, Xu L, Xiao R, Lu X, Chen L, Chong J, Li H, He C, Fu X-D, et al. 2015. Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex, *Nature* **523:** 621–625.

Watt F, Molloy PL. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter, *Genes Dev* **2:** 1136–1143.

Wernig M, Lengner CJ, Hanna J, Lodato MA, Steine E, Foreman R, Staerk J, Markoulaki S, Jaenisch R. 2008. A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types, *Nat Biotechnol* **26:** 916–924.

White J, Dalton S. 2005. Cell cycle control of embryonic stem cells, *Stem Cell Rev* **1:** 131–138.

Wiekowski M, Miranda M, DePamphilis ML. 1993. Requirements for promoter activity in mouse oocytes and embryos distinguish paternal pronuclei from maternal and zygotic nuclei, *Dev Biol* **159:** 366–378.

Wiles MV, Keller G. 1991. Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture, *Development* **111:** 259–267.

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PAC, Rappsilber J, Helin K. 2011. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity, *Nature* **473:** 343–348.

Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells, *Nature* **336:** 684–687.

Wossidlo M, Arand J, Sebastiano V, Lepikhov K, Boiani M, Reinhardt R, Schöler H, Walter J. 2010. Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes, *EMBO J* **29:** 1877–1888.

Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W, Walter J. 2011. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming, *Nat Commun* **2:** 241.

Wray J, Kalkan T, Smith AG. 2010. The ground state of pluripotency, *Biochem Soc Trans* **38:** 1027–1032.

Wu H, Zhang Y. 2014. Reversing DNA methylation. Mechanisms, genomics, and biological functions, *Cell* **156:** 45–68.

Wu X, Li G, Xie R. 2018. Decoding the role of TET family dioxygenases in lineage specification, *Epigenetics Chromatin* **11:** 58.

Xu X, Duan S, Yi F, Ocampo A, Liu G-H, Izpisua Belmonte JC. 2013. Mitochondrial regulation in pluripotent stem cells, *Cell Metab* **18:** 325–332.

Yamashita J, Itoh H, Hirashima M, Ogawa M, Nishikawa S, Yurugi T, Naito M, Nakao K. 2000. Flk1-positive cells derived from embryonic stem cells serve as vascular progenitors, *Nature* **408:** 92–96.

Ying QL, Nichols J, Chambers I, Smith A. 2003. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3, *Cell* **115:** 281–292.

Ying Q-L, Stavridis M, Griffiths D, Li M, Smith A. 2003. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture, *Nat Biotechnol* **21:** 183–186.

Ying Q-L, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A. 2008. The ground state of embryonic stem cell self-renewal, *Nature* **453:** 519–523.

Young RA. 2011. Control of the embryonic stem cell state, *Cell* **144:** 940–954.

Zalzman M, Falco G, Sharova LV, Nishiyama A, Thomas M, Lee S-L, Stagg CA, Hoang HG, Yang H-T, Indig FE, et al. 2010. Zscan4 regulates telomere elongation and genomic stability in ES cells, *Nature* **464:** 858–863.

Zhan Q, Lord KA, Alamo I, Hollander MC, Carrier F, Ron D, Kohn KW, Hoffman B, Liebermann DA, Fornace AJ. 1994. The gadd and MyD genes define a novel set of mammalian genes encoding acidic proteins that synergistically suppress cell growth, *Mol Cell Biol* **14:** 2361–2371.

Zhu B, Zheng Y, Angliker H, Schwarz S, Thiry S, Siegmann M, Jost JP. 2000. 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence, *Nucleic Acids Res* **28:** 4157–4165.

Zhu B, Zheng Y, Hess D, Angliker H, Schwarz S, Siegmann M, Thiry S, Jost JP. 2000. 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex, *Proc Natl Acad Sci U S A* **97:** 5135–5139.

Zhu C, Gao Y, Guo H, Xia B, Song J, Wu X, Zeng H, Kee K, Tang F, Yi C. 2017. Single-Cell 5-Formylcytosine Landscapes of Mammalian Early Embryos and ESCs at Single-Base Resolution, *Cell Stem Cell* **20:** 720-731.e5.

Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. 2008. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks, *Nature* **456:** 125–129.

8 List of abbreviations

2C	Two cell
5´aza	5´-azadeoxycytidine
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
aa	Amino acid
AID	Activation induced deaminase
AIF	Apoptosis inducing factor
AM-AR	Active modification – active removal
AM-PD	Active modification – passive dilution
APOBEC	Apolipoprotein B mRNA-editing enzyme complex
AP	Apurinic/apyrimidinic
APEX1	AP endonuclease
Asz1	Ankyrin repeat, SAM and basic leucine zipper domain containing protein 1
Bak	Bcl-2 homologous antagonist/killer
Bax	BCL2 associated X, apoptosis regulator
BCA	Bicinchoninic acid
Bcl-2	BCL2 apoptosis regulator
BER	Base excision repair
BFP	Blue fluorescent protein
BMP4	Bone morphogenetic protein 4
Bp	base pair
BSA	Bovine serum albumin
CA	Cellular aggregate
CBP	CREB binding protein
Ccng1	Cyclin G1
Cdc2	Cyclin dependent kinase 1
cDNA	Copy DNA
Cdx2	Caudal-type homeobox protein 2
CGI	Cytosine-Guanine dinucleotide islands
Chaf1a	Chromatin assembly factor I P150 subunit

CHARGE	Coloboma, heart defects, atresia choanae, growth retardation, genital abnormalities, ear abnormalities
Chd7	Chromodomain helicase DNA binding protein 7
Co	Control
CpG	Cytosine-Guanine dinucleotide
Crif1	CR6-interacting factor 1
CRISPR	Clustered regularly interspaced short palindromic repeats
DBTMEE	Database of Transcriptome in Mouse Early Embryos
DDR	DNA damage response
DE	Differentially expressed
DHet	Double heterozygous
DKO	Double knockout
DMEM	Dulbecco´s modified Eagle´s medium
DMR	Differentially methylated region
DMSO	Diethylsulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
dNTP	Nucleoside triphosphate
DPPA2	Developmental pluripotency-associated protein 2
DPPA4	Developmental pluripotency-associated protein 4
DTT	Dithiothreitol
DUX	Double homeobox protein
E	Embryonic day
EB	Embryoid body
EDTA	Ethylendiamintetraacetat
Ehf	ETS homologous factor
Elf5	E74 like ETS transcription factor 5
Elk1/4	ETS domain-containing protein 1/4
Epi	Epiblast
Etv1	ETS translocation variant 1
Eomes	Eomesodermin
ERK	Extracellular-signal regulated kinases
ESCAPE	Embryonic Stem Cell Atlas from Pluripotency Evidence
FACS	Fluorescence activated cell sorting
FapyA	4,6-diamino-5-formamidopyrimidine

FapyG	2,6-diamino-4-hydroxy-5-formamidopyrimidine
FBS	Fetal bovine serum
FC	Fold change
FDR	Fold discovery rate
FGF	Fibroblast growth factor
Fli1	Friend leukemia integration 1 transcription factor
FMR	Fully methylated region
FPKM	Fragments per kilobase of exon per million reads mapped
G1	Gap 1
G2/M	Gap2/Mitose
G9a	Euchromatic histone-lysine N-methyltransferase 2
Gadd45a	Growth arrest and DNA-damage inducible protein 45 alpha
Gadd45b	Growth arrest and DNA-damage inducible protein 45 beta
Gadd45g	Growth arrest and DNA-damage inducible protein 45 gamma
Gh	Guanidinohydantoin
GFP	Green fluorescent protein
gRNA	Guide ribonucleic acid
GO	Gene ontology
GSK3	Glycogen synthase kinase 3
H3K4me2/3	Histone H3 lysine 4 dimethylation/trimethylation
H3K9me2	Histone H3 lysine 9 dimethylation
H3K27ac	Histone H3 lysine 27 acetylation
HDAC	Histone deacetylase
IAP	Intracisternal A particle
ICM	Inner cell mass
ICR	Imprinted control region
ING1	Inhibitor of growth protein 1
iPSC	Induced pluripotent stem cell
JAK/STAT	Januskinase/Signal transducers and activators of transcription
JNK	c-Jun N-terminal kinases
KAP1/TRIM28	Kruppel-associated box domain-associated protein 1/Tripartite motif containing 28
Klf4/5	Kruppel like factor 4/5
KO	Knockout
LC-MS/MS	Liquid chromatography-tandem mass spectrometry

LIF	Leukemia inhibitory factor
LINE1	Long interspersed element-1
LMR	Lowly methylated region
Lsd1	Lysine-specific histone demethylase 1
MAB-seq	Methylation-assisted bisulfite sequencing
MageA	Melanoma-associated antigen
MAPK	Mitogen-activated protein kinase
MBD2	Methyl-CpG-binding domain protein 2
MBD4	Methyl-CpG-binding domain protein 4
Mdm2	Mouse double minute protein 2
MEF	Mouse embryonic fibroblast
MEK	Mitogen-activated protein kinase kinase
MEKK4	Mitogen-activated protein kinase kinase kinase 4
MERVL	murine endogenous retrovirus with leucine tRNA primer
mESC	Mouse embryonic stem cells
min	Minute
mRNA	Messenger RNA
mtDNA	Mitochondrial DNA
mTSC	Mouse trophoblast stem cells
NEIL	Nei endonuclease VIII-like
NAC	N-acetyl-L-cysteine
NEAA	Non-essential amino acid solution
Nelfa	Negative elongation factor A
NER	Nucleotide excision repair
NGS	Next generation sequencing
NP-40	Nonidet P40
NTD	Neural tube closure defect
Nthl1	Nth like DNA glycosylase 1
Oct4	Octamer-binding protein 4
PAGE	Polyacrylamide gel electrophoresis
Pax3	Paired box 3
Pax6	Paired box 6
PE	Primitive endoderm
Perp1	Plasma cell-induced resident endoplasmic reticulum protein

PBS	Phosphate buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PGC	Primordial germ cells
PN	Pronuclear phases
PNKP	Polynucleotide kinase 3'-phosphatase
pO_2	Partial pressure of oxygen
Pramel6/7	Preferentially expressed antigen in melanoma-like 6/7
Prl	Prolactin
Psg	Pregnancy-specific glycoprotein
PVDF	Polyvinylidendifluorid
qPCR	Quantitative polymerase chain reaction
RA	Retinoic acid
rDNA	Ribosomal deoxyribonucleic acid
Rerg	RAS like estrogen regulated growth inhibitor
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RPKM	Reads per kilobase pair per million reads
rpm	Rounds per minute
S	Synthesis
Sall4	Spalt like transcription factor 4
SD	Standard deviation
SDS	Sodium dodecy sulfate
Serpin	Serine proteinase inhibitor
SKO	Single knockout
Sox2	SRY-box 2
Sp	Spiroiminodihydantoin
Sp110	Speckled 110 kDa
Spt	Serine palmitoyltransferase 2
Sry	Sex-determining region Y protein
TAF12	TATA-box binding protein associated factor 12
TARID	TCF21 antisense RNA inducing promoter demethylation
TBP	TATA binding protein

TBS	Tris buffered saline
TBX3	T-box 3
Tcof1	Treacle ribosome biogenesis factor 1
TCS	Treacher-Collins-Syndrome
Tcstv1/3	Two-cell stage variable transcript 1/3
TF	Transcription factor
TDG	Thymine DNA glycosylase
Tg	Thymine glycol
TE	Transposable element
TE	Trophectoderm
TET	Ten-eleven translocation methylcytosine dioxygenase
TKO	Triple knockout
TSS	Transcription start site
UHRF1	Ubiquitin like with PHD and ring finger domains 1
UMR	Unmethylated region
UPL	Universal probe library
UT	Untreated
UTR	Untranslated region
UV	Ultraviolet
WGBS	Whole-genome bisulfite sequencing
Wnt	Wingless-type MMTV integration site family member
WT	Wildtype
XPG	Xeroderma pigmentosum group G-complementing protein
XRCC1	X-Ray repair cross complementing 1
ZGA	Zygotic genome activation
Zscan4	Zinc finger and SCAN domain containing 4

9 Acknowledgements

10 Lebenslauf