The role of DNA modifications in pluripotency and differentiation

Christine Silvia Schmidt

Dissertation
an der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

vorgelegt von
Christine Silvia Schmidt
aus Ingolstadt

München, den 09. August 2012

Erstgutachter: Prof. Dr. Heinrich Leonhardt

Zweitgutachter: PD. Dr. Anna Friedl

Tag der mündlichen Prüfung: 30.10.2012

Content

Summary

DNA methylation plays a crucial role in the epigenetic control of gene expression during mammalian development and differentiation. Whereas the *de novo* DNA methyltransferases (Dnmts), Dnmt3a and Dnmt3b, establish DNA methylation patterns during development; Dnmt1 stably maintains DNA methylation patterns during replication. DNA methylation patterns change dynamically during development and lineage specification, yet very little is known about how DNA methylation affects gene expression profiles upon differentiation. Therefore, we determined genome-wide expression profiles during differentiation of severely hypomethylated embryonic stem cells (ESCs) lacking either the maintenance enzyme Dnmt1 (*dnmt1*$^{-/-}$ ESCs) or all three major Dnmts (*dnmt1*$^{-/-}$*; dnmt3a*$^{-/-}$*, dnmt3b*$^{-/-}$ or TKO ESCs), resulting in complete loss of DNA methylation, and assayed their potential to transit in and out of the ESC state. Our results clearly demonstrate that upon initial differentiation to embryoid bodies (EBs), wild type, *dnmt1*$^{-/-}$ and TKO cells are able to activate differentiation processes. However, transcription profiles of *dnmt1*$^{-/-}$ and TKO EBs progressively diverge with prolonged EB culture, with *dnmt1*$^{-/-}$ EBs being more similar to wild type EBs, indicating a higher differentiation potential of *dnmt1*$^{-/-}$ EBs compared to TKO EBs. Remarkably though, after dissociation of late EBs and further cultivation under pluripotency promoting conditions, both *dnmt1*$^{-/-}$ and TKO but not wild type cells rapidly revert to expression profiles typical of undifferentiated ESCs. Thus, while DNA methylation is dispensable for the initial activation of differentiation programs, it seems to be crucial for permanently restricting the developmental fate during differentiation.

Based on the essential role of Uhrf1 in maintenance DNA methylation, we investigated the structurally highly similar second member of the Uhrf protein family, Uhrf2, whose function in maintenance methylation or other biological processes is completely unknown. Expression analysis of *uhrf1* and *uhrf2* in various cell lines and tissues revealed a time- and developmental switch in transcript levels of both genes with *uhrf1* being highly expressed in undifferentiated, proliferating cells and *uhrf2* being predominately expressed in differentiated, non-dividing cells. These opposite expression patterns together with no detectable effect on DNA methylation levels upon knock down of *uhrf2* suggests that Uhrf2 is rather involved in maintaining DNA methylation patterns in differentiated cells and points to non-redundant functions of both proteins.

The discovery of the “6th base” of the genome, 5-hydroxymethylcytosine (5hmC), resulting from the oxidation of 5mC by the family of Tet dioxygenases (Tet1-3), once again ignited the debate about how DNA methylation marks can be modified and removed. To gain insights into the biological function of this newly identified modification, we developed a sensitive enzymatic assay for quantification of global 5hmC levels in genomic DNA. Similar to 5mC levels, we found that also 5hmC levels dynamically change during differentiation of ESCs to EBs, which correlates with the differential expression of *tet1-3*. Furthermore, we characterized a novel endonuclease, PvuRts1I that selectively cleaves 5hmC containing DNA and show first data on its application as a tool to map and analyze 5hmC patterns in mammalian genomes.

Finally, we investigated designer transcription activator-like effector (dTALEs) proteins targeting the oct4 locus. Our results show that the epigenetic state of the target locus interferes with the ability of dTALEs to activate transcriptionally silent genes, which however can be overcome using dTALEs in combination with low doses of epigenetic inhibitors.

In conclusion, this work gives further insights into the biological roles of methylation mark writers (Dnmts), readers (Uhrfs) and modifiers (Tets) and advances our understanding on the function of DNA methylation in the epigenetic control of gene expression during development and cellular differentiation.

1. Introduction

1.1 Epigenetic gene regulation

"DNA is just a tape carrying Information and a tape is no good without a player. Epigenetics is about the tape player." Bryan Turner (Birmingham, UK).

The term "epigenetics" was introduced by Conrad Waddington in 1942 referring to the molecular mechanisms that translate the genetic information into an observable phenotype (Waddington, 1942). Nowadays, epigenetics is used to describe mechanisms that control gene function in a potentially heritable way without altering the DNA sequence and thus provide a mechanism to maintain cellular identity in a long-term way (reviewed in Bird, 2002).

During embryogenesis a single zygote gives rise to various cell types, which are genetically identically, but show structural and functional differences due to differential gene expression. Developmental and environmental signals synergistically activate complex transcription factor networks which together with epigenetic modifications induce differentiation programs. These epigenetic mechanisms include DNA methylation, histone modifications, histone variants and nucleosome remodelling and are crucial for the establishment of specific cell lineages and cell types during differentiation (Figure 1).

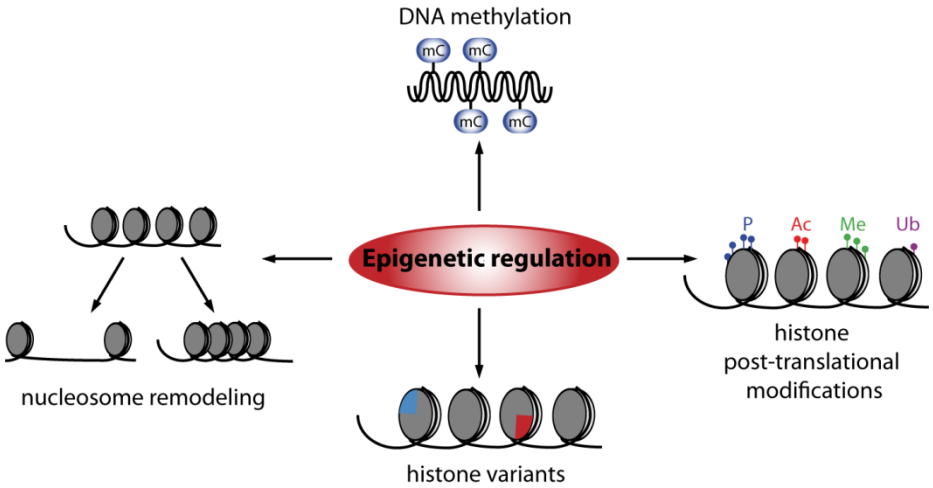


Figure 1: Overview of epigenetic mechanisms.
Methylation of DNA occurs at cytosines mostly within a CpG context and is generally associated with gene repression. Histones are subjected to various post-translational modifications including among others phosphorylation (P), acetylation (Ac), methylation (Me) and ubiquitination (Ub). The modified histones are recognized and bound by effector proteins, which translate the marks into specific molecular consequences like transcription, repair or condensation. Furthermore, the canonical histones can be replaced by histone variants containing different sequences which results in changes in the chromatin structure. In addition, nucleosome remodelling e.g. nucleosome eviction, sliding or insertion by ATP dependent chromatin remodelers leads to changes in chromatin structure and accessibility. Importantly, all of these processes are highly interconnected. Further factors that contribute to epigenetic regulation are non-coding RNAs which for instance play an important role during X Chromosome inactivation. Also the nuclear architecture, like the formation of distinct chromatin domains, including euchromatin and heterochromatin, is influenced by epigenetic mechanisms.

Despite the heritability of epigenetic marks, they can be reprogrammed under specific circumstances. On the one hand, reprogramming can be achieved experimentally by generating induced pluripotent stem cells (iPSCs) through ectopic expression of various transcription factors or on the other hand also naturally by the genome-wide DNA demethylation wave occurring during germ cell development. Thus, epigenetic marks fulfill dual roles in proliferating cells; they help to preserve cellular identity while concomitantly conferring cellular plasticity that is needed to adapt to environmental cues and differentiation signals. Moreover, epigenetic abnormalities contribute to the development of diseases like e.g. cancer, hence emphasizing the importance of proper epigenetic gene regulation to faithfully maintain genomic integrity and function.

1.1.1 DNA methylation

DNA methylation is the longest known and probably most studied epigenetic modification. In mammals, this post-replicative mark occurs exclusively at the C5 position of cytosine residues as 5-methylcytosine (5mC) mainly in the context of CpG dinucleotides and is generally associated with stable gene silencing.

Only around 1 % of total DNA bases consist of 5mC in human somatic cells, which accounts for approximately 60-80 % of all CpG dinucleotides being methylated in human and mouse (Ehrlich et al., 1982). Methylated sequences are quite diverse and include single copy genes, intergenic regions and repetitive sequences. Frequently, CpG dinucleotides are found within CpG islands, regions which encompass around 1000 bp with a G+C content of at least 50 % and a CpG frequency (observed versus expected) of at least 0.6 (Illingworth and Bird, 2009). Remarkably, about 60-70 % of human gene promoters are overlapping with CpG islands, which are usually unmethylated during development and in tissues and mostly are located within housekeeping genes (Ehrlich et al., 1982; Antequera and Bird, 1993). Only a small fraction (4-8 %) of CpG islands become tissue-specifically methylated, leading to stable silencing of the associated genes in somatic tissues (De Smet et al., 1999; Shen et al., 2007; Illingworth et al., 2008). Methylation of CpG islands plays a crucial role in genomic imprinting and X chromosome inactivation and has been linked to developmental diseases and cancer (Bird, 2002).

Interestingly, up to 70 % of tissue-specific DNA methylation has been detected at promoters and enhancers containing a low CpG density. Many of these low CpG density regions are located close (~2 kb) to CpG islands and hence are referred to as CpG island shores. A stronger correlation between gene expression and differentially methylated regions can be found at CpG island shores than at CpG islands (Doi et al., 2009; Irizarry et al., 2009; Ji et al., 2010). A comparison of DNA methylation between mouse and human revealed that

differentially methylated regions (DMR) – CpG islands which show tissue specific DNA methylation – are highly conserved across species and can even be used to distinguish various tissues irrespective of the origin of species (Irizarry et al., 2009).

DNA methylation occurs not only at promoters and enhancers, but it has also been detected at gene bodies, where it - in contrast to promoter methylation - positively correlates with transcription (Hellman and Chess, 2007; Rauch et al., 2009). Until today the function of gene body methylation is not well understood and several hypotheses have been suggested. On the one hand, it could help to dampen transcriptional noise by repressing expression of antisense transcripts (Peter A., 1999) or on the other hand, evidence from studies in plants indicate that it might influence elongation efficiency (Zilberman et al., 2007). Interestingly, increased gene body methylation was found on the active human X chromosome compared to the inactive X chromosome and therefore might be involved in the augmented expression of X-linked genes and dosis compensation in mammals (Hellman and Chess, 2007).

Another pivotal function of DNA methylation is the stable silencing of repetitive sequences including parasitic sequences like transposons and endogenous retroviruses in the genome (Yoder et al., 1997b). Direct evidence that these repetitive sequences are hypermethylated and therefore transcriptionally inactive in somatic cells comes from observations in embryos lacking Dnmt1, which heavily induce the expression of the transpositionally active family of intracisternal A particle (IAP) elements in the mouse genome (Walsh et al., 1998). In cancer cells, global hypomethylation at repetitive sequences is observed and it is believed that the reactivation of endoparastic sequences contributes to genomic instability (Gaudet et al., 2003). Therefore, methylation of repetitive sequences plays a crucial role in the preseveration of chromosomal integrity and stability as well as in the prevention of translocations and gene disruptions (Walsh et al., 1998; Gaudet et al., 2003; Esteller, 2007).

DNA methylation occurring in a non-CpG context was detected in embryonic stem cells (ESCs), where it is especially enriched in gene bodies and depleted in protein binding sites and enhancers (Ramsahoye et al., 2000; Lister et al., 2009; Laurent et al., 2010). Non-CpG context methylation can occur at CpA and CpT sites, however mCpA seems to be predominant form of non-CpG methylation (Ramsahoye et al., 2000; Laurent et al., 2010). As non-CpG methylation was found to decrease during differentiation, it has been suggested to be associated with pluripotency (Lister et al., 2009; Laurent et al., 2010).

1.1.2 Histone modifications

Within the cell, DNA is packed into chromatin. The first level of compaction is the nucleosome, which consists of a histone octamer of the four core histones H2.A, H2.B, H3 and H4 and of 146 bp DNA wrapped around in 1.65 turns. The fifth histone H1 is the linker

histone, which binds to DNA linking two nucleosomes and stabilizes higher order chromatin structure. The canonical histones consist of an unstructured N-terminal tail, a globular domain and a short C-terminus. The highly unstructured tails are targets of numerous post-translational modifications including acteylation, methylation, phosphorylation, ubiquitination, SUMOylation, ADP ribosylation, citrullination and glucosamine N-acetylation and proline isomerization (Kouzarides, 2007). These modifications are reversibly set by various writers and erasers including histone acetyltransferases and histone deacetylases, kinases and phosphatases, arginine methyltransferases as well as histone lysine methyltransferases and demethylases, respectively.

The modification of histone tails plays a crucial role in two basic processes; firstly, the establishment of global chromatin regions, like euchromatin and heterochromatin and secondly, the facilitation of DNA based functions like transcription, DNA replication and repair, chromosome condensation (Kouzarides, 2007) and alternative splicing (Luco et al., 2010). Posttranslationally modified histones alter the histone-DNA interactions either directly or via recruitment of specific chromatin associated proteins that recognize and translate the various histone modifications into specific biological consequences. Furthermore, the incorporation of histone variants through ATP dependent nucleosome-remodeling exchanger complexes can alter histone-DNA interactions which influence nucleosome positioning and gene expression (reviewed in (Portela and Esteller, 2010) and see also Figure 1 in Chapter 1.1).

Based on distinct modifications of histone tails, the genome can roughly be divided into two distinct chromatin regions; which reflect gene expression activity. The actively transcribed and accessible euchromatic regions contain high levels of acetylated histone tails as well as trimethylation on histone H3K4. The inactive and highly condensed heterochromatin is characterized by low levels of acetylation and high levels of trimethylation on histone H3K9, H3K27 and H4K20 (Li et al., 2007a). Moreover, bivalent chromatin domains have been found in embryonic stem cells (ESCs), where predominantly promoters involved in differentiation carry the active histone H3K4me3 mark simultaneously with the repressive histone H3K27me3 mark. It has been proposed that this bivalency keeps developmental genes in a silent state but poises them for rapid lineage-specific activation or repression (Bernstein et al., 2006; Mikkelsen et al., 2007; Ku et al., 2008). The enzymes responsible for setting these bivalent marks are two antagonistically acting epigenetic regulators of gene activity; Polycomb group (PcG) proteins act as transcriptional repressors which establish histone H3K27me3 marks and consist of two complexes, Polycomb repressor complex 1 and 2 (PRC1 and PRC2). More specifically, PRC1 is composed of the core subunits Ring1a and 1b together with various other proteins and catalyzes monoubiquitination of histone H2A. PRC2

consists of the three core proteins Enhancer of zeste 2 (Ezh2), Embryonic ectoderm development (Eed) and Suppressor of zeste 12 (Suz12) and are responsible for catalyzing di- and trimethylation of histone H3K27 via the SET domain of Ezh2 (Akasaka et al., 1996; Core et al., 1997; del Mar Lorente et al., 2000). Several studies suggest a synergistic action of PRC1 and PRC2 to mediate transcriptional repression. After trimethylation of H3K27, PRC1 binds to the repressive marks, catalyzes monoubiquitation of histone H2A which interferes with transcription. However, recent data point to more complex functions of PRC1 and PRC2 as both can be part of various heterogeneous complexes affecting many different target genes. By contrast, trithorax group (trxG) proteins are transcriptional activators catalyzing trimethlyation of histone H3K4 and are also part of multiprotein complexes including histone methyltransferases and nucleosome remodelers (reviewed in Orkin and Hochedlinger, 2011). The two protein groups, PcGs and TrxG proteins and their associated histone modifications play crucial roles in the plasticity of stem cell states, developmental transitions, maintenance of lineage-specific transcription programs, genomic imprinting, X inactivation and cancer. Importantly, both protein complexes are not involved in the initial regulation of expression but have a critical function in maintaining a transcriptionally active (trxG) or silent (PcG) state through many rounds of cell division and hence provide a system for cellular memory (Ringrose and Paro, 2004; Sparmann and van Lohuizen, 2006; Schuettengruber et al., 2007).

1.2 Regulation of DNA methylation in mammalian cells

1.2.1 Writers of DNA methylation marks – the family of DNA methyltransferases

DNA methylation marks are set and maintained by the family of mammalian DNA methyltransferases (Dnmts). The *de novo* Dnmts, Dnmt3a and Dnmt3b, together with their catalytically inactive cofactor Dnmt3L, are responsible for the establishment of methylation marks during differentiation, which are then progressed by the *maintenance* Dnmt, Dnmt1, throughout the cell cycle. All active DNA cytosine methyltransferases harbor a highly conserved C-terminal catalytical domain containing all 10 sequence motifs commonly found in bacterial cytosine methyltransferases (Goll and Bestor, 2005) (Figure 2).

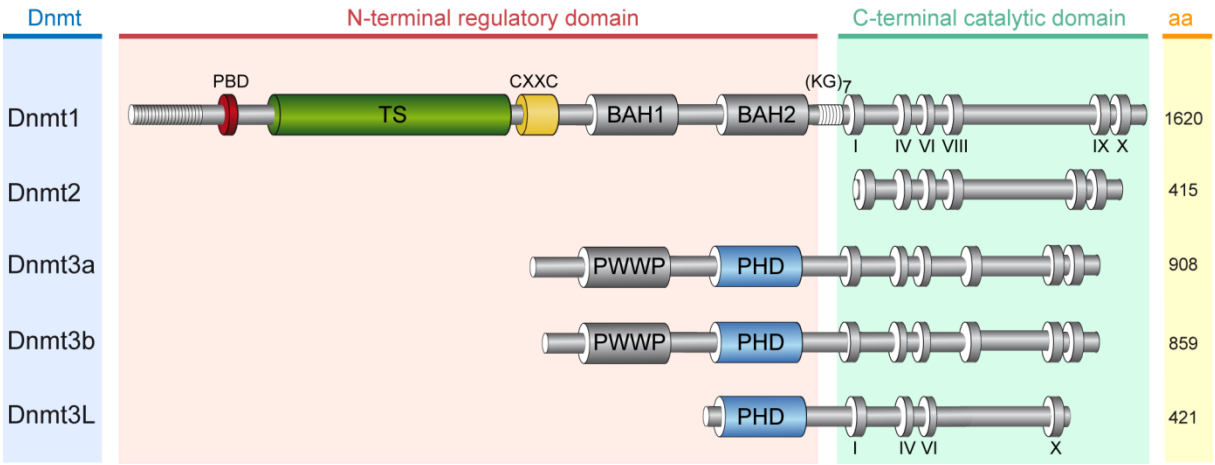


Figure 2. Domain structure of the mammalian Dnmt family.
Whereas the C-terminal catalytical domain is highly conserved among all members of the Dnmt family, the N-terminal regulatory domain shows some striking differences. Dnmt1 harbors the largest N-terminal region containing the PCNA binding domain (PBD), the targeting sequence (TS) responsible for pericentric heterochromatin localization, the CXXC domain followed by two bromo adjacent homology domains (BAH1 and 2). A linker consisting of 7 lysine - glyine repeats $(KG)_7$ connects the N- and C-terminal part of Dnmt1. The N-terminal region of Dnmt3a and 3b harbors only two distinct domains, the Pro-Trp-Trp-Pro (PWWP) motif containing domain and the plant – homeodomain (PHD). The catalytically inactive Dnmt3L only contains the PHD domain. Dnmt2 is the only member which does not contain a N-terminal regulatory domain. The length of the various Dnmts is indicated in amino acids (aa) (adapted from (Rottach et al., 2009).

Given the high conservation between prokaryotic and eukaryotic Dnmts, it is generally assumed that the same mechanism is applied by all enzymes. After substrate recognition and flipping of the target cytosine out of the DNA double helix, a covalent complex is formed with the C-6 position of the cytosine. Subsequently, a methyl group is transferred from the methyl group donor S-Adenosyl-L-Methionine (AdoMet) to the activated C-5 position and the enzyme is finally released by ß-elimination (Flynn et al., 1996; Flynn and Reich, 1998; Pradhan et al., 1999).

1.2.1.1 Dnmt1

The *maintenance* enzyme Dnmt1 was the first eukaryotic Dnmt to be discovered and cloned (Stein et al., 1982; Bestor et al., 1988). Experiments on cell lines and mice carrying homozygous mutations of *dnmt1* underline the importance of properly maintaining genomic DNA methylation patterns. Whereas complete genetic ablation of *dnmt1* leads to embryonic lethality at the midgastrula stage, embryos carrying a hypomorphic *dnmt1* allele with around 10% residual *dnmt1* expression are able to develop (Li et al., 1992; Lei et al., 1996). These mutant mice have a globally hypomethylated genome, display chromosomal instability and develop aggressive tumors several months after birth (Gaudet et al., 2003). Interestingly, *dnmt1*$^{-/-}$ embryos show biallelic expression of some imprinted genes and ectopic transient X Chromosome inactivation, emphasizing the crucial function of Dnmt1 in genomic imprinting and X Chromosome inactivation (Li et al., 1993; Beard et al., 1995). ESCs derived from *dnmt1*$^{-/-}$ blastocysts are viable and are not impaired in their ability to self-renew, although these cells contain low (around 20 %) but stable levels of DNA methylation mostly in repetitive sequences (Li et al., 1992; Lei et al., 1996). By contrast, somatic cells lacking Dnmt1 are impaired in their ability to proliferate and survive and show p53-dependent apoptosis, which can be overcome by additionally knocking out p53 in these cells (Jackson-Grusby et al., 2001; Chen et al., 2007; Spada et al., 2007).

Dnmt1 is the major and most abundant Dnmt in mammalian cells, as it is expressed in all proliferating cells. To provide a faithful mechanism for the stable inheritance of DNA methylation patterns, several factors act in combination. Firstly, Dnmt1 expression is cell cycle-dependent regulated, resulting in highest *dnmt1* transcription during S-G2 phase and minimal expression in quiescent (G0) cells (Robertson et al., 2000b; Tatematsu et al., 2000a). Consistently, *dnmt1* is down regulated in non-proliferating cells, with the exception of post-mitotic neurons, where Dnmt1 seems to be localized in the cytoplasma (Goto et al., 1994; Inano et al., 2000). Secondly, Dnmt1 is directly coupled to the replication machinery via direct interaction with the replication factor proliferating cell nuclear antigen (PCNA). Two different domains of the N-terminal part are involved in the distinct subnuclear localization of Dnmt1 during the cell cycle. Whereas the PCNA binding domain (PBD) is responsible for the recruitment of Dnmt1 to replication sites during early to mid S-phase (Leonhardt et al., 1992; Chuang et al., 1997), the protein is then associated with (peri-) centromeric heterochromatin via its Targeting sequence (TS) domain during late S to G2-phase (Easwaran et al., 2004). Thirdly, Dnmt1 has a preference for hemimethylated DNA (Yoder et al., 1997a; Frauer and Leonhardt, 2009), substrates which are generated during semi-conservative replication. Fourthly, the recently identified protein Uhrf1 has been identified as a crucial factor for maintenance methylation as this protein is proposed to recruit Dnmt1 to replication foci via its

binding to hemimethylated CpG sites (see also chapter 1.2.2) (Bostick et al., 2007; Papait et al., 2007; Sharif et al., 2007).

The association of Dnmt1 with the replication machinery suggests a combined replication of genetic and epigenetic information. Interestingly, the PBD has not only a crucial role in recruiting Dnmt1 to replication foci, but also to sites of DNA damage, indicating that Dnmt1 is also involved in restoring epigenetic information after DNA repair (Mortusewicz et al., 2005).

Several studies aiming to elucidate the basis of maintenance methylation - the preference of Dnmt1 for hemimethylated DNA - were performed over the last years. The first crystal structure of Dnmt1 in complex with unmethylated DNA was just recently solved and sheds some more light on the complex regulation and function of this enzyme. Whereas it has been known that the CXXC domain of Dnmt1 preferentially binds to unmethylated CpG sites (Fatemi et al., 2001; Pradhan et al., 2008; Frauer et al., 2011), structural insights now suggest a crucial inhibitory role of this binding in maintenance methylation. Upon binding of the CXXC domain to unmethylated DNA, an autoinhibitory linker connecting the CXXC domain with the first BAH domain is positioned in the active center of Dnmt1, thereby preventing the entering of (unmethylated) DNA in the catalytical pocket and thus aberrant DNA methylation (Song et al., 2011). The same group published a second structure of Dnmt1 covalently bound to a DNA substrate containing a hemimethylated CpG site. Comparison of this new structure to the inhibitory state shows that most conformational changes occur within the catalytical pocket. Interestingly, several amino acids in the catalytical center specifically target the DNA substrate to distinguish its methylation state and similarly to prokaryotic methyltransferases, the target cytosine is flipped out of the double helix (Song et al., 2012). Additionally, other crystal structures of Dnmt1 suggest an inhibitory role of the TS domain by binding directly to the DNA binding pocket within the catalytical center, which has to be displaced in order to allow the DNA methylation reaction (Syeda et al., 2011; Takeshita et al., 2011). Taken together, all these structures give first, but limited, insights into the complex process of maintenance methylation, which seems to involve a multi-step cascade of conformational changes.

The catalytic domain of Dnmt1 contains all conserved motifs described to be necessary in prokaryotic methyltransferases to catalyze the methylation reaction. Nonetheless, the catalytic domain of Dnmt1 alone is not sufficient for enzymatic activity but needs intramolecular interaction with the N-terminal regulatory domain for allosteric activation (Zimmermann et al., 1997; Margot et al., 2000; Fatemi et al., 2001). Besides this intramolecular interactions, also a range of intermolecular interactions with numerous chromatin-associated proteins were reported for Dnmt1, thereby linking the enzyme to diverse biological functions including cell cycle regulation, DNA repair, chromatin structure as

well as tumorigenesis. Important interaction partners are the H3K9 histone methyltransferases G9a (Estève et al., 2006), Suv39h1 (Fuks et al., 2003a) as well as components of the Polycomb group 2 complex (PRC2), Eed and Ezh2, involved in H3K27 methylation (Viré et al., 2006) and the histone deacetylases HDAC1/2(Fuks et al., 2000; Robertson et al., 2000a; Rountree et al., 2000b). Furthermore, the Heterochromatin Protein 1 (HP1) (Fuks et al., 2003a) and various chromatin remodelers including members of the SNF2 family of ATPases, like Smarca5 (Robertson et al., 2004) and related proteins like Lsh (Myant and Stancheva, 2008) as well as transcriptional regulators, among them the Dnmt1-associated protein 1 (Dmap1) (Rountree et al., 2000a), have been shown to interact with Dnmt1. These interactions emphasize the high interconnectivity of various epigenetic pathways which lead to the establishment and maintenance of transcriptionally inactive chromatin. Additionally, many interacting proteins have been demonstrated to subject Dnmt1 to numerous posttranslational modifications, which modulate Dnmt1 abundance, stability and activity. More specifically, recent reports suggest that the stability and abundance of Dnmt1 is controlled by acetylation and ubiquitination in a cell-cycle dependent manner. Whereas Dnmt1 is acetylated by the histone acetyltransferase (HAT) Kat5 and subsequently ubiquitinated by the E3 ubiquitin ligase Uhrf1 leading to its proteolytic digestion, Dnmt1 abundance increases in early to late S-phase via the concerted action of the deubiquitinase Usp7 and deacteylase HDAC1, thereby antagonizing degradation and enhancing the stability of Dnmt1 (Du et al., 2010; Qin et al., 2011). Furthermore, phosphorylation followed by methylation of Dnmt1 has been shown to function antagonistically to regulate Dnmt1 stability during the cell cycle (Estève et al., 2009, 2011). Another interesting example is the sumoylation of Dnmt1 which has been suggested to enhance its catalytic activity *in vitro* (Lee and Muller, 2009).

1.2.1.2 Dnmt2

The second member of the Dnmt family, Dnmt2, so far has not been shown to harbor Dnmt activity, but rather has RNA methyltransferase activity as it has been implicated in methylating aspartic acid transfer RNA ($tRNA^{Asp}$) (Goll et al., 2006). Although this protein is the most strongly conserved one, contradicting reports about the role of Dnmt2 in mammalian cells aside from its RNA methyltransferase activity exist. The first reports about the discovery of Dnmt2 describe no DNA methyltransferase activity although the enzyme contains all conserved methyltransferases motives (Okano et al., 1998b; Yoder and Bestor, 1998). However, DNA binding properties of human DNMT2 at least *in vitro* have been suggested by a crystal structure of human DNMT2 in complex with the demethylated cofactor S-adenosyl L-homocystein and a superimposition with bacterial restriction methyltransferase M.HhaI showed remarkably similar orientations of all important sequence motifs (Dong et al.,

2001). Furthermore, studies by Hermann *et al.* indicate a very weak DNA methyltransferase activity of DNMT2 *in vitro* (Hermann et al., 2003). Taken together, although sequence and structure comparison analyses suggest a Dnmt role for Dnmt2, so far most genetic and biochemical data failed to prove this function *in vitro* and *in vivo*.

1.2.1.3 The Dnmt3 family

During development and gametogenesis, Dnmt3a and Dnmt3b are responsible for the establishment of global DNA methylation patterns (Okano et al., 1999; Kaneda et al., 2004). Consistent with this, both proteins are highly expressed in embryonic stem cells (ESCs) and down regulated in differentiated somatic cells and tissues (Okano et al., 1998a). Mice lacking *dnmt3a* survive for 4 weeks after birth, suggesting that Dnmt3a is not crucial for early embryonic developmental processes but seems to play a pivotal role in the methylation of genes critical for neonatal viability (Okano et al., 1999). Furthermore, studies on conditional *dnmt3a* knockout mice revealed a crucial function of Dnmt3a in both, maternal and paternal imprinting, the mono-allelic expression of genes dependent on the origin of parent (Kaneda et al., 2004). In contrast, embryos deficient for *dnmt3b* show early embryonic lethality (E 9.5), indicating that Dnmt3b plays an important role during early developmentally processes. Dnmt3b was found to specifically methylate centromeric minor satellite repeats. This is consistent with a phenotyp described in patients suffering from the rare autosomal recessive human ICF (immunodeficiency, centromer instability and facial anomalies) syndrome, which is caused by point mutations in DNMT3B (Hansen et al., 1999; Xu et al., 1999). The specific loss of methylation at minor pericentric satellite DNA is thus assumed to be critical for maintaining chromosome stability. Moreover, loss of both *de novo* Dnmts leads to early embryonic lethality and *dnmt3a*$^{-/-}$*, 3b*$^{-/-}$ ESCs progressively become globally hypomethylated and fail to methylate newly integrated proviral sequences (Okano et al., 1999).

Interestingly, although Dnmt3a and Dnmt3b are closely related, they seem to have non-overlapping functions during development as suggested by the different phenotypes of their respective knockout mice. Both proteins consist of a regulatory N-terminal domain linked to a conserved C-terminal catalytical domain (Figure 2). The regulatory domain contains the PWWP domain shown to be involved in chromatin targeting of both enzymes. Furthermore, the PWWP domain was reported to bind to trimethylated histone H3K36, which seems to enhance DNA methylation activity of Dnmt3a and 3b (Ge et al., 2004; Dhayalan et al., 2010). The N-terminal domain also harbors a PHD which has been reported to specifically bind to unmethylated H3K4 and is responsible for multiple interactions with various chromatin proteins including HDACs, HP1 and the histone methyltransferase Suv39h1 (Fuks et al., 2003a).

The third homologue of the Dnmt3 family is Dnmt3L, which consists of a short N-terminal domain and the C-terminal catalytical domain, but has no methyltransferase activity (Figure 2). Nonetheless, Dnmt3L has been shown to function as an important cofactor for Dnmt3a and Dnmt3b as it interacts and colocalizes with both proteins during early embryonic development and was proposed to stimulate their activity (Hata et al., 2002; Xie et al., 2006). In germ cells, Dnmt3L together with Dnmt3a is responsible for establishing *de novo* DNA methylation at imprinted genes, presumably via its binding to unmethylated H3K4 and subsequent recruitment or activation of Dnmt3a2, a variant of Dnmt3a (Ooi et al., 2007). In line with this, *dnmt3l*$^{-/-}$ mice are viable but methylation of sequences that are normally maternally methylated is absent in oocytes. Furthermore, reactivation of retrotransposons and meiotic catastrophes is observed in spermatocytes from *dnmt3l*$^{-/-}$ mice. Consequently, the mice are sterile after birth, although surprisingly, global DNA methylation levels are not altered (Bourc'his et al., 2001; Bourc'his and Bestor, 2004).

It is still not well understood how sequence specific *de novo* DNA methylation patterns are established. One recent study identified small methylation-determining regions (MDRs) within proximal promoter regions that mediate both hypomethylation and *de novo* methylation, indicating that target specificity of DNA methylation patterns is conveyed by the local DNA sequence itself (Lienert et al., 2011). However, also DNA binding proteins and the local chromatin environment are likely to influence DNA methylation adding again to the complexity of how methylation patterns are established.

1.2.1.4 Cooperative functions of mammalian Dnmts

Although a non-overlapping function of the two types of enzymes, Dnmt1 as the *maintenance* and Dnmt3a and 3b as the *de novo* Dnmts, has been proposed, accumulating evidence suggests that a clear categorical distinction of *maintenance* and *de novo* methylation might not be possible. Firstly, a cooperative function of Dnmts has been proposed for the maintenance of methylation at repetitive sequences (Liang et al., 2002) and is underlined by data showing that Dnmt1 can interact with both *de novo* Dnmts (Fatemi et al., 2002; Kim et al., 2002). Secondly, the progressive loss of DNA methylation after inactivation of *dnmt3a* and *dnmt3b* in ESCs can be rescued by reintroduction of both proteins, indicating that they participate in the maintenance of global DNA methylation patterns (Chen et al., 2003). Thirdly, recent data revealed that Dnmt3a and Dnmt3b are selectively anchored to methylated nucleosomes, leading to their stabilization in the cells, whereas unbound Dnmt3a and Dnmt3b proteins are degraded by the proteosomal machinery. This compartmentalization has been suggested to abolish aberrant *de novo* methylation by Dnmt3a and Dnmt3b and by specifically binding to methylated sites, the two proteins would only target CpG sites that were missed by Dnmt1 after DNA replication. Altogether, this

implies that both *de novo* Dnmts work synergistically with Dnmt1 to stably propagate DNA methylation patterns (Jeong et al., 2009; Sharma et al., 2011).

1.2.2 Readers of DNA methylation marks – methylcytosine binding proteins

DNA methylation is considered to mediate transcriptional silencing for which two possible modes of repression have been described so far. The methyl group can directly interfere with the binding of transcription factors at their target sites (Becker et al., 1987). The second mechanism of DNA methylation mediated transcriptional repression involves direct and specific binding of the methyl group by methyl-CpG binding proteins (MBPs) which in turn recruit repressive chromatin modifiers. Until now, three different MBP families have been described as the readers of DNA methylation marks: the methyl-CpG binding domain (MBD) family, the Kaiso protein family, and the ubiquitin-like plant homeodomain and RING finger domain-containing (Uhrf) protein family.

1.2.2.1 The MBD protein family

The family of MBD proteins consists of five members (MBD1, MBD2, MBD3, MBD4 and MeCP2; Figure 3) and all proteins, except MBD3, preferentially bind methylated DNA via their MBD domain (Hendrich and Bird, 1998; Saito and Ishikawa, 2002). Besides their MBD domain, MBD1, MBD2 and MeCP2 contain a non-conserved transcription repressor domain (TRD) which in the case of MeCP2 mediates interaction with Dnmt1 (Kimura and Shiota, 2003). Also MBD2 and MBD3 were shown to form complexes with Dnmt1 and were suggested to be involved in maintaining DNA methylation during DNA replication (Tatematsu et al., 2000b).

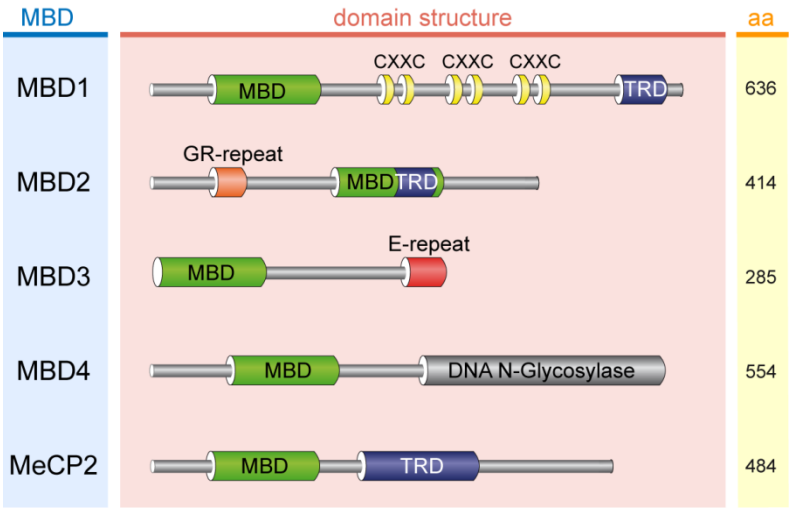


Figure 3. Domain structure of the MBD protein family.
Aside from the common MBD domain, MBD1 contains a cystein rich domain with 3 additional CXXC-type zinc finger motifs. MBD2 and MBD3 contain repeat sequences, Glycine arginine (GR) and glutamine (E) repeats, respectively. MBD4 is the only member containing a C-terminal glycosylase domain implicated in base excision repair. Numbers indicate the length of the proteins in amino acids (aa) (adapted from (Rottach et al., 2009).

All members of the MBD family mediate transcriptional repression in an HDAC-dependent manner and all except of MBD4 have been shown to interact with nucleosome remodeling complexes like NuRD, which establish a repressive chromatin environment (Jones et al., 1998; Hendrich and Tweedie, 2003; Kondo et al., 2005). Furthermore, MBD1 and MeCP2 have been found to interact with histone H3K9 methyltransferases as well as the Heterochromatin Protein 1 (HP1), thereby functioning as a link between DNA methylation and repressive histone modifications to stabilize transcriptional repression (Fujita et al., 2003; Fuks et al., 2003b; Sarraf and Stancheva, 2004; Agarwal et al., 2007). These two members contain dual binding sites and can in addition to methylated DNA also bind to unmethylated DNA. MBD1 and MeCP2 can also induce chromatin compaction in the absence of DNA methylation, suggesting that they facilitate transcriptional repression not only through the recruitment of histone deacteylases but also through their properties to generate highly condensed secondary and tertiary chromatin structures which constitute a physical barrier for the assembly of activating transcription complexes at these sites (Georgel et al., 2003; Jørgensen et al., 2004; Nikitina et al., 2007). Notably, MBD4 is the only member containing a thymine DNA glycosylase domain which has been described to be involved in the repair of TG mismatches generated by the deamination of 5-methylcytosine, implying a role of MBD4 in active DNA demethylation (Bellacosa et al., 1999; Hendrich et al., 1999); see also chapter 1.2.3).

Interestingly, the founding member of the MBD family, MeCP2 is ubiquitously expressed but the most abundant in brain tissue, indicating a functional role for MeCP2 in the nervous system (Shahbazian and Zoghbi, 2002). In line with this, mutations of the MeCP2 genes were identified as the primary cause of the rare neurodevelopmental RETT syndrome (RTT) in humans (Amir et al., 1999), implying that MeCP2 participates in the epigenetic regulation of neuronal function. Surprisingly though, genome-wide mapping of MeCP2 binding sites in human neurons revealed actively transcribed regions as the primary binding targets and only a minority of methylated CpGs sites were bound by MeCP2, leading to the assumption that transcriptional repression might not be the pivotal role of MeCP2 (Yasui et al., 2007). By contrast, recent genome-wide binding data in mouse neurons suggest that MeCP2 preferentially associates with methylated regions and that loss of MeCP2 leads to global changes in chromatin structures including increased histone acetylation levels and higher levels of the linker histone H1. These data indicate that MeCP2 globally functions as a transcriptional dampener and not as a gene-specific transcriptional repressor (Skene et al., 2010).

The unexpectedly very mild phenotype of mice lacking MBD proteins suggests functional redundancy among the proteins, especially in light of the dramatic phenotype of global loss

of DNA methylation and their early embryonic death (Chen et al., 2001; Guy et al., 2001; Hendrich et al., 2001; Zhao et al., 2003). However, MDB1, MBD2 and MeCP2 were shown to bind to different foci within a cell, arguing for dinstinct functions (Ballestar et al., 2003; Klose et al., 2005). Other explanations for the viability of knockout mice could be that either DNA methylation is sufficient to induce transcriptional silencing also in the absence of MBD proteins which then would function solely in maintaining the repressed state, or that other non-MBD proteins like the Kaiso and Uhrf family proteins can compensate the loss in these knock out mice (Sasai and Defossez, 2009).

1.2.2.2 The Kaiso protein family

All three members of the Kaiso protein family, Kaiso (ZBTB33) and the Kaiso-like proteins ZBTB4 and ZBTB38 harbor a three zinc-finger motif with which they preferentially bind methylated DNA (Prokhortchouk et al., 2001; Filion et al., 2006) (Figure 4).

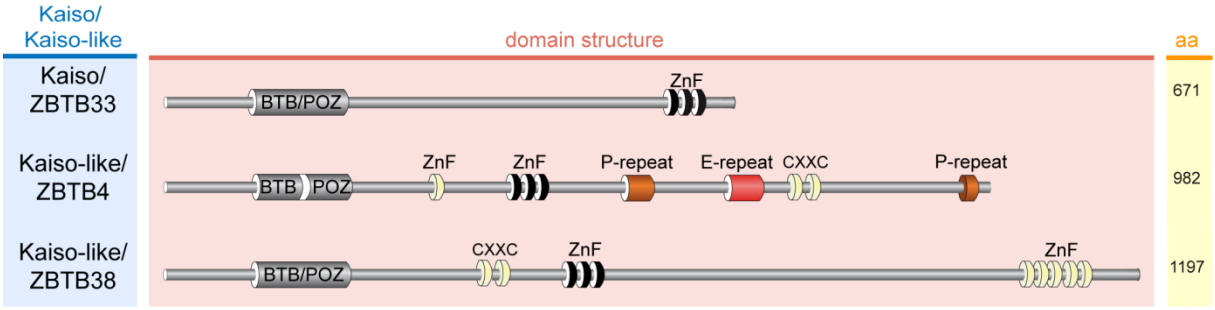


Figure 4. Schematic overview of the Kaiso protein family.
Binding to methylated DNA is mediated by their Krüppel-like, triple C2H2 ZnF domain marked in black. All members contain a broad complex, tramtrack and bric a brac/poxvirus and zinc finger domain (BTB/POZ) which in case of ZBTB4 harbors an insertion of 60 amino acids (aa) depicted in white. CXXC: CXXC type Zinc finger domain; ZnF: Zinc finger domain; P- and E-repeat: Proline and glutamine repeats are indicated (modified from Rottach et al., 2009).

Whereas Kaiso requires at least two symmetrical methylated CpG (mCpG) sites for efficient binding, Kaiso-like proteins are able to bind to single mCpGs. Similar to MBD proteins, Kaiso was shown to bind to and recruit nucleosome remodeling complexes to mediate methylation-dependent transcriptional repression via histone deaceylation and H3K9 methylation (Yoon et al., 2003). Interestingly, Kaiso proteins were shown to bind with even higher affinity to an unmethylated specific consensus sequences, termed Kaiso binding sequence (KBS) (Daniel et al., 2002). A KBS was identified at the promoter of the wnt/ß-catenin/TCF target promoter matrilysin, which is bound and repressed by Kaiso via recruitment of N-CoR and HDAC3 *in vivo*, demonstrating the bi-modal binding properties of Kaiso proteins (Spring et al., 2005). In contrast to Kaiso, transcriptional silencing mediated by ZBTB38 and ZBTB4 involves the recruitment of the corepressor protein CtBP and the Sin3A/HDAC repressor complex, respectively (Sasai et al., 2005; Weber et al., 2008). Until today, the targets of Kaiso and Kaiso-like proteins are not fully identified yet and only Kaiso has been knocked out in mice so far, however leading to no obvious phenotype (Prokhortchouk et al., 2001). This raises again

the question of redundancy of these proteins and given their overlapping gene expression patterns in some, but not all adult tissues (Daniel and Reynolds, 1999; Filion et al., 2006), these proteins might have context dependent unique and overlapping functions. Future studies targeting all members of the Kaiso protein family will clarify their role in transcriptional silencing.

1.2.2.3 The Uhrf protein family

The third family of methylcytosine binding proteins involves two members, Uhrf1 (also called Np95/ICBP90) and Uhrf2 (also called Np97/NIRF) (Figure 5). Both proteins consist of a multi-functional modular structure, containing a Set- and Ring associated (SRA) domain shown to recognize methylated DNA (Unoki et al., 2004).

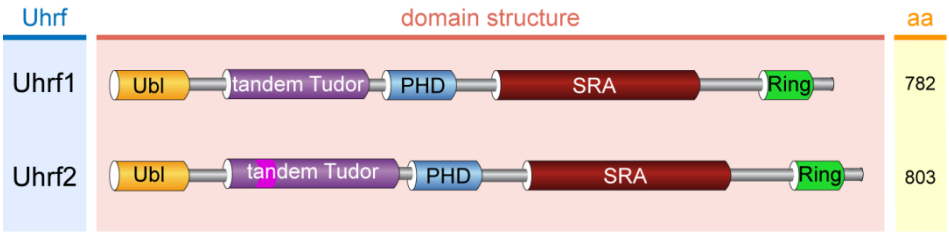


Figure 5. Domain structure of the Uhrf protein family.
Both members harbor a very similar domain structure containing a ubiquitin-like domain (Ubl), a tandem Tudor domain, followed by a plant homeo domain (PHD), a Set- and Ring associated (SRA) domain and a really interesting new gene (Ring) domain. Only the tandem Tudor domain of both proteins differs as Uhrf2 contains an insertion of 33 additional amino acids (aa) as depicted in brighter purple within the tandem Tudor domain. Numbers indicate the length of the protein in aa (modified from Rottach et al., 2009).

The founding member Uhrf1 was initially identified as a protein involved in cell cycle regulation and in the DNA damage response pathway (Bonapace et al., 2002; Muto et al., 2002). Further studies suggest an important role of Uhrf1 in cell proliferation as it has been implicated in silencing tumor suppressor genes in breast cancer cells possibly via the recruitment of the repressive chromatin modifying enzymes histone H3K9 methyltransferase G9a and histone deacetylase HDAC1 (Unoki et al., 2004; Kim et al., 2009). Uhrf1 localizes to replication foci during mid- to late S-phase and plays a crucial role in replicating heterochromatic regions (Uemura et al., 2000; Miura et al., 2001; Papait et al., 2007). Interestingly, the PHD finger of Uhrf1 has been shown to play an essential role in inducing large-scale reorganization of the pericentromeric heterochromatin, which might be critical for the replication of heterochromatic regions by permitting access of the replication machinery to densely packed structures (Papait et al., 2008). In addition, the PHD finger can bind to methylated histone H3K9 which seems to be crucial for the proper localization of Uhrf1 to heterochromatic regions (Karagianni et al., 2008). The RING domain of Uhrf1 harbors a E3 ubiquitin ligase activity and can ubiquitinate histone H3 *in vitro*, although the biological significance of Uhrf1-mediated histone ubiquitination remains to be determined (Citterio et al., 2004). Recent data demonstrated that the tandem Tudor domain of Uhrf1 recognizes and

selectively binds trimetylated histone H3K9 via an aromatic cage structure (Rottach et al., 2010). This structure shows striking similarity to the hydrophobic cage of the chromodomain of HP1, which is known to bind to repressive chromatin marks (H3K9me3) and to associate with pericentromeric heterochromatin (Jacobs and Khorasanizadeh, 2002).
Strikingly, recent findings suggest a crucial function of Uhrf1 in maintaining DNA methylation. Uhrf1 colocalizes with Dnmt1 throughout S-phase of the cell cycle and directly associates with Dnmt1 during replication (Bostick et al., 2007; Sharif et al., 2007). Embryos lacking *uhrf1* show a drastic phenotype remarkably similar to *dnmt1*$^{-/-}$ embryos, including global genomic DNA hypomethylation and early embryonic lethality. In line with this, decreased methylation at imprinted regions and major satellites as well as de-repression of endogenous retrotransposons were detected in *uhrf1*$^{-/-}$ ESCs and embryos (Bostick et al., 2007; Sharif et al., 2007) Furthermore, Uhrf1 binds with a slight, but significant preference, to hemimehylated substrates via its SRA domain and therefore was suggested to recruit Dnmt1 to hemimethylated CpG sites generated during DNA replication (Bostick et al., 2007; Sharif et al., 2007; Rottach et al., 2010). Crystal structures of Uhrf1 in complex with hemimethylated DNA revealed that the SRA domain recognizes its substrate by flipping the methylated base out of the DNA helix. Such a base-flipping mechanism has been demonstrated for bacterial Dnmts and some DNA repair enzymes which then further modify the flipped base. Until today, Uhrf1 has been the only non-enzymatic protein identified so far catalyzing this base-flapping mechanism (Arita et al., 2008; Avvakumov et al., 2008; Hashimoto et al., 2008).

In conclusion, by binding to hemi-methylated DNA and to repressive histone marks, together with the recruitment of repressive histone modifiers, Uhrf1 allows epigenetic crosstalk by linking the two major epigenetic pathways for gene silencing. Interestingly, the second member of the Uhrf protein family, Uhrf2, shows a similar modular structure to Uhrf1, raising the question whether Uhrf2 has a similar function. So far, first data point to a role of Uhrf2 in cell cycle regulation and possibly as a tumor suppressor protein and has been shown to harbor auto-ubiquitin ligase activity (Mori et al., 2002, 2004, 2011; Li et al., 2004). However its role in maintaining DNA methylation patterns still remains elusive.

1.2.3 Modifiers of DNA methylation marks – the family of Tet proteins

Since the discovery of DNA methylation, it was always considered to be a quite stable epigenetic mark. In 2009, however, Tahiliani and colleagues performed a computational homology search to the trypanosome thymine hydroxylases J base containing proteins JBP1 and JBP2 and discovered the ten-eleven translocation (Tet) protein family, which can convert 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC). By performing high resolution mass spectrometry, small amounts of 5hmC (0.03 % of all Cs) were detected in the genomic DNA of embryonic stem cells (ESCs) and expression analysis of *tet1* demonstrated the

presence of *tet1* transcripts in these cells (Tahiliani et al., 2009). The Tet protein family consists of three members, Tet1, Tet2 and Tet3 (Figure 6).

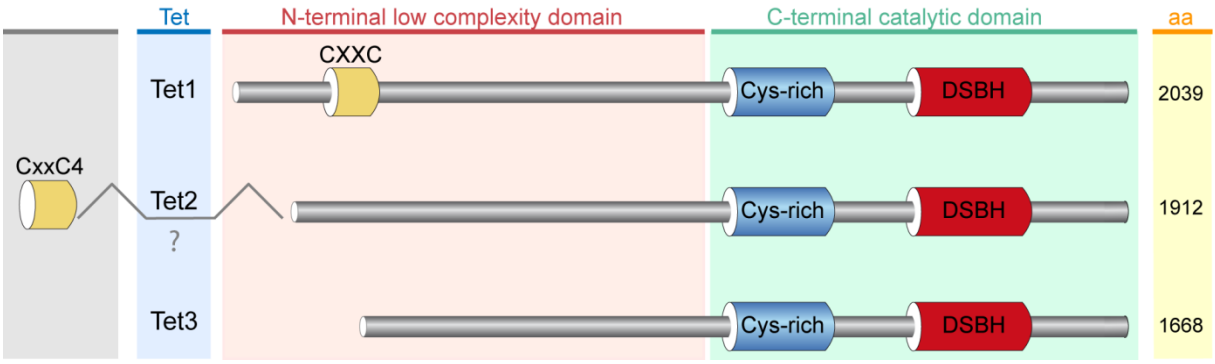


Figure 6. Schematic overview of the family of Tet proteins
The N-terminal, low complexity region is predicted to function as a protein interaction surface. As the CxxC4 protein was found in close chromosomal proximity it was suggested that a local chromosomal inversion removed the CXXC domain from the Tet2 gene. However, CxxC4 might interact with Tet2 and possibly is involved in targeting Tet2 to genomic DNA (Iyer et al., 2009). Contradicting reports about the presence of a CXXC domain in Tet3 exist and therefore a CXXC domain for Tet3 is not included in the overview. The length of the proteins is indicated in amino acids (aa).

All members contain a C-terminal 2-oxoglutarate (2OG)- and Fe(II) -dependent dioxygenase domain (DSBH), catalyzing the conversion of 5mC to 5hmC *in vitro* and *in vivo* (Tahiliani et al., 2009; Ito et al., 2010). Remarkably, recent data demonstrate that the Tet proteins can even further oxidize 5hmC to 5-formylcytosine (5fC) and 5 carboxylcytosine (5caC) and that these two cytosine derivates are detectable in mouse genomic DNA of ESCs and tissues (Ito et al., 2011). Adjacent to the DSBH, all Tet proteins harbor a Cystein-rich domain predicted to be involved in DNA binding. Interestingly, so far a CXXC domain was only identified in Tet1 which shows similarity to the CXXC domain in Dnmt1 and preferentially binds to CpG-rich DNA (Iyer et al., 2009).

1.2.3.1 Tet1

The founding member Tet1 was first described as a fusion partner of the histone H3K4 lysine methyltransferase MLL in a subgroup of patients suffering from acute myeloid leukemia harboring the translocation t(10;11)(q22;q23) (Lorsbach et al., 2003). The finding that *tet1* is highly expressed in ESCs, but decreases during differentiation, points to a role of Tet1 in regulating pluripotency and differentiation potential. Indeed, a knockdown of *tet1* in ESCs results not only in global reduction of 5hmC level, but also revealed a crucial role of Tet1 in maintaining ESCs in an undifferentiated state, possibly by keeping the promoter of the pluripotency associated gene *nanog* in a hypomethylated and therefore active state (Ito et al., 2010; Freudenberg et al., 2011). In line with this, *tet1* as well as *tet2* were shown to be regulated by Oct4 and Sox2 (Koh et al., 2011) and depletion of both proteins results in decreased expression of pluripotency-related genes with concordant increase in their promoter DNA methylation (Ficz et al., 2011; Koh et al., 2011). Furthermore, Tet1 seems to

be important for the specification of the inner cell mass, as reduced *tet1* levels in pre-implantation embryos results in a skewed differentiation towards trophoectodermal lineage (Ito et al., 2010; Koh et al., 2011). In stark contrast to the dramatic phenotypes upon *tet1* knock down, *tet1*$^{-/-}$ ESCs show only a partial reduction of global 5hmC level solely affecting the expression of few genes and no defect in maintaining pluripotency could be detected. Moreover, genetic ablation of *tet1* still gave viable and fertile mice, however with a slightly reduced body size compared to their littermates (Dawlaty et al., 2011). Hence, further studies including additional Tet knockout models are needed to clarify the function of Tet1 in pluripotency and during differentiation.

Genome-wide binding data of Tet1 shed some light on its biological function. Several studies showed an enrichment of Tet1 within gene bodies, preferentially at exons, as well as at transcriptional start sites and promoters (Ficz et al., 2011; Pastor et al., 2011; Williams et al., 2011; Wu et al., 2011b; Xu et al., 2011). Furthermore, Tet1 binding is especially observed at high CpG dense promoters (HCP) carrying histone H3K4me3 marks and thus negatively correlates with DNA methylation marks (Williams et al., 2011; Wu et al., 2011b; Xu et al., 2011). As a positive correlation between CpG content and Tet1 binding exists, it has been suggested that Tet1 binding to CG rich promoters keeps these sequences in a hypomethylated, activated state. Consistently, depletion of Tet1 leads to increased DNA methylation at CpG rich sequences (Wu et al., 2011a).

Surprisingly, these genome-wide binding profiles of Tet1 revealed a dual function of the protein in transcriptional regulation, as it can function not only as an activator but also as a repressor of target genes (Williams et al., 2011; Wu et al., 2011a; Xu et al., 2011). Active genes controlled by Tet1 are involved in pluripotency whereas genes repressed by Tet1 play a role in differentiation processes (Wu et al., 2011a). Many of the genes repressed by Tet1 are targets of Polycomp Repressive Complex 2 (PRC2) carrying bivalent histone modifications (see also chapter 1.1.2) and it has been shown that Tet1 is enriched on those genes (Williams et al., 2011; Wu et al., 2011a, 2011c). As loss of Tet1 abolishes binding of PRC2 to genes carrying bivalent marks and the fact that binding of PRC2 is blocked by 5mC, it has been suggested that PRC2 is indirectly recruited to bivalent genes by Tet1, which are kept in a hypomethylated and therefore accessible state for PCR2 (Wu et al., 2011a). Additional support for a role of Tet1 in transcriptional repression comes from the observation that Tet1 interacts and colocalizes with a substantial amount of target genes of the Sin3A co-repressor complex, implying an important function of this complex in Tet1-mediated repression. The Tet1- mediated repression by the Sin3A complex seems to be independent of its catalytical activity. In line with this, the same target genes are also up regulated in hypomethylated ESCs lacking Tet1, indicating that these genes are not directly controlled by

DNA hydroxymethylation *per se* but rather by the presence of Tet1 itself on these promoters (Williams et al., 2011).

1.2.3.2 Tet2

The second member of the Tet family, Tet2, is also expressed in ESCs but its depletion had only minor consequences in ESCs suggesting a possible function in other biological contexts (Ito et al., 2010; Koh et al., 2011). Indeed, human TET2 has been shown to be crucial for haematopoiesis as mutations in *TET2* have been frequently found in various human myeloid malignancies including myelodysplastic syndromes, myeloproliferative neoplasms, and chronic myelomonocytic leukemia (Langemeijer et al., 2009; Ko et al., 2010). Many of these mutations affect the catalytic activity of TET2 and it has been suggested that these mutations occur early during tumorigenesis (Ko et al., 2010). Studies on *tet2* knockout mice revealed a crucial function of Tet2 in self-renewal, proliferation and differentiation of hematopoietic stem cells (HSCs). Whereas *tet2*$^{-/-}$ mice are viable and appear phenotypically normal after birth, they seem to be more proned to develop hematopoietic malignancies already within one year after birth. Furthermore, loss of *tet2* in mice leads to a global decrease of 5hmC and an elevated number of HSCs. In agreement with this, *in vitro* experiments suggest that Tet2 functions as a tumor suppressor during hematopoietic cell homeostasis as cells depleted of *tet2* show an increased self-renewing, but decreased differentiation capacity (Ko et al., 2011; Li et al., 2011; Moran-Crusio et al., 2011; Quivoron et al., 2011). Interestingly, also mutations in the metabolic enzymes isocitrate dehydrogenase (*IDH) 1* and *IDH2* similarly impair the differentiation of HSC and are often found in patients suffering from acute myeloid leukemias. The mutant forms of the enzymes produce predominantly the metabolite 2-hydroxyglutarate, which inhibits the hydroxylation reaction by Tet2, probably by outcompeting the actual co-factor of Tet2 α-ketoglutarate (Figueroa et al., 2010; Konstandin et al., 2011).

1.2.3.3 Tet3

Probably the least investigated member of the Tet family so far, Tet3, has been reported to be only very low abundant in ESCs and therefore seems not to play a role in ESC biology. However, recent data suggest a fundamental role for Tet3 in the reprogramming of the paternal genome after fertilization (Gu et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011). More specifically, as levels of global 5mC drastically decline in the paternal pronucleus after fertilization, a concomitant increase of 5hmC was found in the male genome. Intriguingly, *tet3* was shown to be highly expressed in oocytes and zygotes and as the other two members of the family, *tet1* and *tet2* were nearly absent at these stages, it was suggested that Tet3 is the responsible enzyme for the conversion of 5mC to 5hmC in the paternal pronucleus (Iqbal et al., 2011; Wossidlo et al., 2011). The maternal pronucleus seems to be protected from Tet3

mediated hydroxylation by Stella/Dppa3 (Nakamura et al., 2006; Wossidlo et al., 2011). Further evidence for the involvement of Tet3 in the global reprogramming of the paternal genome comes from observations in *tet3*$^{-/-}$ zygotes, which fail to reduce global 5mC levels and show impaired demethylation of *oct4* and *nanog* in the male pronucleus. Interestingly, deletion of *tet3* in female germ lines leads to reduced fecundity of female mice and their heterozygous mutant offspring are more prone to developmental failures (Gu et al., 2011).

Based on the fact that 5hmC was specifically detected in the paternal pronucleus, it was speculated that Tet3-mediated oxidation of 5hmC is part of an active DNA demethylation process which would further involve the removal of 5hmC or its derivates by DNA repair enzymes of the base excision repair (BER) pathway like the tymine DNA glycosylase (Tdg). However, latest data on paternal and maternal chromosome spreads of pre-implantation embryos now suggest that 5hmC gets diluted in the paternal genome in a replication dependent, passive manner and not via the involvement of repair enzymes (Inoue and Zhang, 2011). Furthermore, the same authors show that concomitantly to 5hmC, also 5fC and 5caC can be detected in preimplantation embryos and also become replication-dependent diluted, indicating that these two newly identified modifications are quite stable and not just temporary intermediates, suggesting a possibly functional role of 5fC and 5caC during preimplantation development (Inoue et al., 2011).

1.2.3.4 DNA hydroxymethylation - 5hmC

Although the presence of a 6th DNA base, 5hmC, in mammalian genomic DNA was already reported in 1972 (Penn et al., 1972), appreciable attention to 5hmC started only in 2009 with the discovery of the Tet protein family (Tahiliani et al., 2009). 5hmC was not only detected in genomic DNA of ESCs, but also in Purkinje cells of the mouse cerebellum, 0.6 % of all Cs were found to be hydroxylated. Although this seems to be very low levels of genomic 5hmC, it translates to approximately 40 % of all 5mCs being hydroxylated (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009). The finding that 5hmC level are relatively high in ESCs but decrease upon ESC differentiation, suggests a crucial role of the newly discovered modification in pluripotency (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009). To advance understanding of the role of 5hmC, several groups performed genome-wide mapping of 5hmC in ESCs. Similar to the binding profile of Tet1, 5hmC is mainly found in gene bodies with a specific enrichment at exons and near transcriptional start sites (TSS). Furthermore, 5hmC can be detected on promoters with intermediate levels of CpG sites (ICPs) carrying predominantly bivalent chromatin marks and its enrichment negatively correlates with CG content (Ficz et al., 2011; Pastor et al., 2011; Wu et al., 2011b; Xu et al., 2011). Since 5hmC is derived from 5mC, both modifications show some co-existence in gene bodies. However, in contrast to the genomic distribution of 5mC, 5hmC is not enriched at heterochromatic

regions like repetitive sequences, pointing to different biological roles of the two modifications (Ficz et al., 2011; Pastor et al., 2011; Williams et al., 2011).

1.2.3.5 Possible mechanisms of DNA demethylation

The observation of dynamic reprogramming of DNA methylation patterns during development, the global demethylation wave in the paternal pronucleus shortly after fertilization (Mayer et al., 2000; Oswald et al., 2000) as well as in primordial germ cells in post-implantation embryos (Hajkova et al., 2002; Lee et al., 2002; Yamazaki et al., 2003), has raised the question of mechanisms for active DNA demethylation processes. Especially the fact that loss of DNA methylation in the paternal pronucleus occurs remarkably fast, being completed within 4-8 hours after fertilization even before the first cell division is finished, excludes the possibility of passive DNA demethylation and great effort has been undertaken in finding potential enzymes which could actively remove DNA methylation marks.

Two major mechanisms of active DNA demethylation have been proposed; the deamination pathway involving sequential deamination by AID/Apobec and subsequent excision by components of the BER machinery and the oxidation pathway including several oxidation steps of 5hmC to 5fC and 5caC followed by excision by Tdg (Figure 7).

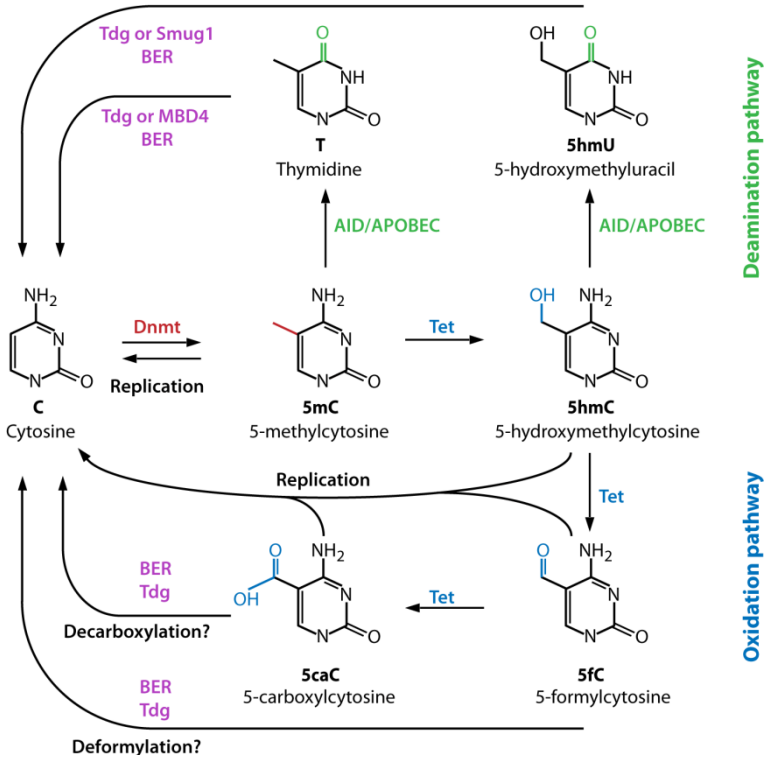


Figure 7. Overview of active and passive DNA demethylation pathways
Genomic 5mC can be removed by passive, replication-dependent dilution or actively involving several intermediate steps, among others, 5hmC as an intermediate product. In principal, two pathways for active DNA demethylation have been proposed, the deamination pathway including Tets, members of the family of AID deaminases and BER enzymes or the oxidation pathway including progressive oxidation of 5mC to 5cac by Tets and BER enzymes (adapted from Branco et al., 2012).

The deamination pathway is proposed to be a two-step process in which deamination is followed by base-excision repair. Crucial members of this suggested pathway involve the activation- induced cytidine deaminase (AID)/ apolipoprotein B mRNA editing enzyme, catalytic polypeptide (APOBEC) family, which have been shown to deaminate 5mC to thymine resulting in a G:T mismatch (Morgan et al., 2004). The mismatch is then recognized and excised by the DNA gylcosylases Tdg or MBD4. Evidence for this pathway comes from studies using zebrafish embryos showing that active demethylation involves deamination of AID and subsequent cleavage by MBD4 followed by replacement with cytosine via BER enzymes. Interestingly, this process was shown to be promoted by the damage response protein Gadd45a (Rai et al., 2008). Furthermore, mammalian PGCs deficient for AID display incomplete erasure of global DNA methylation marks (Popp et al., 2010). However, recent data in neural cells suggest that the same deaminases can also convert 5hmC to 5hmU, as overexpression of both, AID/ Apobec as well as Tet1, resulted in an elevated global accumulation of 5hmU. In line with this, Tdg and the single-strand-selective monofunctional uracil-DNA glycosylase 1 (SMUG1) have been reported to exhibit a strong activity towards 5hmU:G mismatches and hence were proposed to be involved in this pathway (Cortellino et al., 2011; Guo et al., 2011). The fact that Tdg was also found to interact with AID and Gadd45 further supports the existence of this proposed pathway (Cortellino et al., 2011). Importantly, by regulating endogenous levels of 5mC, the Tet1-oxidation/ deamination-mediated pathway has been implicated in neuronal activity- induced, region-specific demethylation of neurogenic niche factors in the dentate gyrus of adult mouse brain (Guo et al., 2011).

The initial step of the oxidation pathway for active DNA demethylation involves hydroxylation of 5mC to 5hmC and successive oxidation of 5hmC to 5fC and 5caC (Ito et al., 2010, 2011; Pfaffeneder et al., 2011). Both oxidation products, 5fC and 5caC were shown to be recognized and cleaved by Tdg, followed by repair via BER enzymes (He et al., 2011; Maiti and Drohat, 2011). However, also other strategies for the removal of 5fC and 5caC, possibly by directly deformylating 5fC or decarboxylating 5caC to generate C could exist and need further examination. Additionally, 5hmC as well as both cytosine derivates have also been suggested to become replication-dependent diluted in a passive demethylation pathway as has been reported in the paternal pronucleus in preimplantation embryos (Inoue and Zhang, 2011; Inoue et al., 2011). In support of this proposed passive mechanism, Dnmt1 binding to DNA was found to be impaired upon hydroxylation of 5mC to 5hmC (Valinluck and Sowers, 2007).

1.3 Embryonic stem cells as a model system for differentiation processes *in vitro*

Embryonic stem cells (ESCs) are derived from the explanted inner cell mass (ICM) of preimplantation blastocysts. Importantly, under appropriate culture conditions, ESCs are distinguished by two distinctive properties, unlimited capacity of self renewal and pluripotency which refers to the ability of ESCs to differentiate into all kinds of cell types including germ cells *in vitro* and *in vivo* (reviewed in Nichols and Smith, 2011). Therefore, ESCs are a powerful tool and attractive model to study the molecular mechanisms underlying pluripotency and cell fate choices during differentiation. Furthermore, the recent discovery that the expression of four transcription factors, Oct4, Sox2, Klf4 and c-Myc (Takahashi and Yamanaka, 2006), is sufficient for somatic cellular reprogramming - the conversion of a differentiated somatic cell to an ESC- like state, also referred to as induced pluripotent stem cells (iPS) - has focused attention to the molecular basis and regulatory mechanisms that establish and maintain pluripotency.

1.3.1 The pluripotency network

To keep ESCs in an undifferentiated, pluripotent and self- renewal state, a complex network of signal transduction pathways in combination with transcription factors needs to be maintained. Key signaling pathways include the leukemia inhibitory factor LIF/ JAK/ Stat3, PI3K/ Akt, BMP2/ Smad and wnt signaling in combination with the dual inhibition of the FGF/ Erk and GSK3 signaling pathways (Niwa et al., 1998, 2009; Matsuda et al., 1999; Sato et al., 2003; Ying et al., 2003a, 2008; Paling et al., 2004; Watanabe et al., 2006; Berge et al., 2011; Griffiths et al., 2011) (Figure 8). These conditions promote the expression of master transcriptional regulators of the pluripotency network, *oct4* (also called *pouf5l*), *nanog* and *sox2* (Okamoto et al., 1990; Nichols et al., 1998; Avilion et al., 2003; Chambers et al., 2003; Mitsui et al., 2003). Together with epigenetic modifiers, non- coding RNA and the c-Myc transcriptional network, these crucial regulators form a core regulatory transcriptional network that promotes expression of pluripotency associated genes and represses genes for lineage commitment and differentiation (Boyer et al., 2005; Loh et al., 2006; Kim et al., 2008a) and reviewed in Ng and Surani, 2011; Orkin and Hochedlinger, 2011).

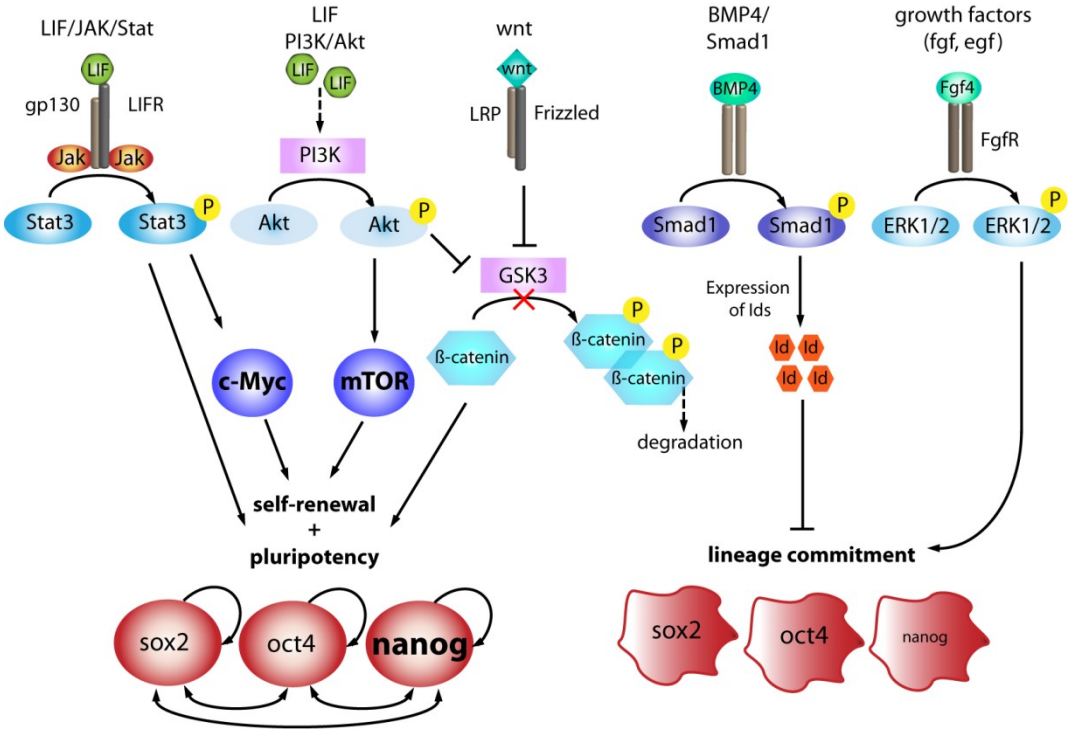


Figure 8. Overview of signaling pathways that maintain pluripotency and self-renewal in mouse ESCs. Several pathways synergistically act to keep ESCs in a pluripotent and self- renewal state. LIF can act via two different pathways: Binding to the gp130/ LIF receptor (LIFR) activates the Janus Kinase (Jak) which then phosphorylates signal transducer and activator of transcription 3 (Stat3). Activated Stat3 promotes the transcription of c-Myc and hence self- renewal and pluripotency. LIF can also activate the Phosphatidyl Inositol 3 Kinase (PI3K) which ultimately leads to activation of a serine/ threonine kinase (Akt), subsequent modulation of mTOR signaling and hence stimulation of proliferation and suppression of cell death. At the same time, Akt inhibits the glycogen synthase kinase 3ß (GSK3) normally involved in phosphorylation of ß-catenin followed by its degradation. By inhibiting GSK3, either by wnt signaling or Akt, ß-catenin can shuffle to the nucleus and stabilizes pluripotency by inhibiting Tcf3, a repressor of core pluripotency- associated transcription factors. In addition, the bone morphogenetic protein (BMP)/ Smad signaling contributes to pluripotency. Phosphorylated Smad1 triggers the expression of Inhibitor of differentiation (Id) proteins which block transcription factors involved in lineage commitment. Also ESCs themselves contribute to the metastable state of pluripotency by producing Fgf4 which drives ESCs to lineage commitment. In the pluripotent state, *nanog* is highly and - unlike *oct4* and *sox2* - heterogeneously expressed and counteracts Erk signaling. Based on the heterogeneous expression of *nanog*, cells with low *nanog* levels are more prone to Erk signaling and likely induce differentiation programs. (reviewed in (Okita and Yamanaka, 2006; Niwa, 2011; Welham et al., 2011).

A deeper understanding of how pluripotency is established and maintained by transcription factors and epigenetic modifications is of great interest not only for the facilitation of directed programming of ESCs to specific lineages but also for somatic cell reprogramming and hence holds great promises for the development of new therapies for diseases and regenerative medicine. Especially the master regulators Oct4, Sox2 and Nanog of the core transcriptional network have been extensively studied over the past years. Oct4, a POU domain-containing transcription factor encoded by *Pouf5I*, lies in the center of the network and is essential for the formation of a pluripotent founder cell population in the embryo (Nichols et al., 1998). The Oct4 levels within a cell need to be tightly controlled as already two fold changes in expression drastically affect the stem cell fate. Whereas an increase of *oct4* results in differentiation towards primitive endoderm and mesoderm, a decrease of *oct4*

leads to loss of pluripotency and differentiation to the trophoectodermal lineage (Niwa et al., 2000). Oct4 is one of the critical, so far irreplaceable factors of somatic cell reprogramming to iPS cells (Takahashi and Yamanaka, 2006; Nakagawa et al., 2007a).

Sox2 has been shown to function as a transcriptional partner of Oct4 (Avilion et al., 2003). More specifically, an enhancer highly active in ESCs contains binding motifs for Oct4 and Sox2 and regulates the expression of pluripotency- associated genes, including *nanog* and *fgf4*. It has been shown that Oct4 and Sox2 collaborate to synergistically activate the enhancer, thereby promoting the expression of genes important to maintain pluripotency (Yuan et al., 1995; Rodda et al., 2005; Masui et al., 2007). The Oct4- Sox2 enhancers also regulate the expression of *oct4* and *sox2* themselves by a positive- feedback loop (Tomioka et al., 2002; Chew et al., 2005; Okumura-Nakanishi et al., 2005). Sox2 is part of the SRY-related HMG box protein family and its genetic ablation in mice results in early embryonic lethality (Avilion et al., 2003) and deletion of *sox2* in ESCs leads to differentiation primarily into the trophoectodermal lineage (Masui et al., 2007).

The third member of the core transcriptional network, Nanog, is a homeodomain containing protein. Whereas lack of *nanog* in ESCs leads to loss of pluripotency and differentiation to the endodermal lineage, overexpression of *nanog* prevents the induction of differentiation and maintains pluripotency independently of the LIF/ Stat3 pathway (Chambers et al., 2003; Mitsui et al., 2003). Consequently, Nanog is often referred to as the gate keeper of pluripotency, as it needs to be tightly controlled to allow reprogramming and differentiation (Silva et al., 2009); see also Figure 8). N*anog* is known to show a very heterogeneous expression and latest data suggest that the heterogeneity of *nanog* expression is regulated on a chromosomal level. In early pre-implantation embryos, *nanog* is monoallelically expressed. However, its expression is gradually switched to biallelic expression as the inner cell mass (ICM) matures to the naïve epiblast and acquires ground- state pluripotency (Miyanari and Torres-Padilla, 2012).

Genome- wide comparative binding data of Oct4, Sox2 and Nanog revealed substantial overlapping binding sites on both, active and repressed promoters and enhancers in human ESCs. Actively transcribed genes occupied by one or two of the core regulators encode not only their own genes, but also transcription factors for components of signaling pathways including wnt and Tgf-ß pathways. Hence, the core factors promote not only their own expression by forming an interconnected autoregulatory loop but also promote the expression of genes encoding essential components of key signaling pathways as well as chromatin modifying proteins. Inactive genes co- bound by Oct4, Sox2 and Nanog are predominately genes involved in lineage commitment and differentiation. Taken together, the highly interconnected autoregulatory loop of the core transcription factors generates a

metastable state in ESCs: on the one hand, the positive feedback loop promotes pluripotency if all three core factors are expressed at appropriate levels; on the other hand, perturbation of one factor by e.g. low expression unbalances the positive loop and might thereby favor entrance into a differentiation program (Boyer et al., 2005; Loh et al., 2006).

Although the three transcriptional regulators, Oct4, Nanog and Sox2, build the core of the pluripotency network, proteomic studies using affinity purification and mass spectrometry identified a large protein- protein interaction network including many transcription factors and chromatin modifying complexes that contribute to the maintenance of ESC properties. Pluripotency related interacting proteins include Sal4, Rex1, Dax1, Klf4, Essrb, and Tcl1 as well as proteins linked to important signaling pathways to maintain pluripotency like Smad1, Stat3, and Tcf3 (Wang et al., 2006; Chen et al., 2008; Cole et al., 2008; Kim et al., 2008a, 2010; Mallanna et al., 2010; Pardo et al., 2010; van den Berg et al., 2010). Interacting proteins of chromatin modifying complexes involve components of the Swi/Snf (also called BAF) nucleosome remodeling complex, the NuRD/HDAC complex and polycomb complex 2 (Figure 9).

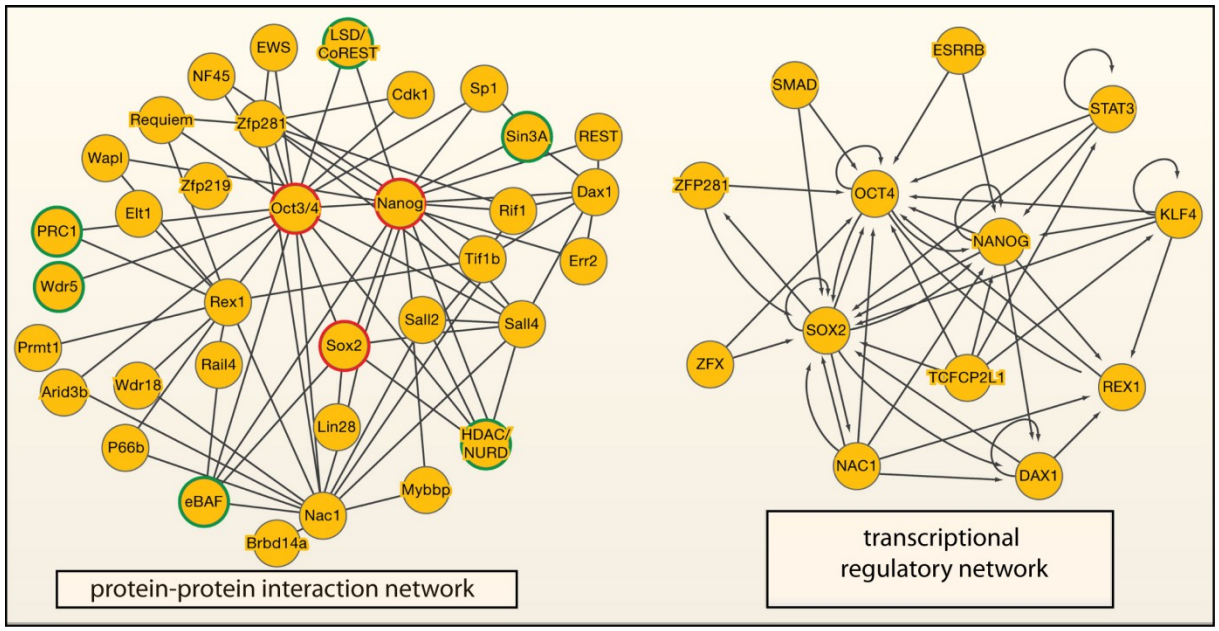


Figure 9. Overview of the *oct4* centric module in ESCs
The *oct4* centric module consists of a highly connected protein interaction network with Oct4, Nanog and Sox2 (circled in red in left picture) building important nodes within the network. The key regulators also directly or indirectly interact with various chromatin regulators (circled in green in left middle). Genome- wide binding data of the key regulators using chromatin immunoprecipitation (ChIP) followed by hybridization on a Chip or Sequencing revealed a complex transcriptional regulatory network (right picture). Many components of the network show autoregulatory but also interconnectivity mechanisms (adapted from Kim et al., 2010; Orkin and Hochedlinger, 2011).

Further studies investigating the genome- wide binding behavior of several key transcription factors in ESCs revealed a complex transcriptional regulatory network with many autoregulatory loops. Additionally, many transcription factors co- occupy target genes important for pluripotency, leading to a high interconnectivity among the components of the

network (Figure 9). Consistent with this, Oct4 was shown to function as an anchor protein for the assembly of the multiprotein complexes on target genes (Chen et al., 2008; van den Berg et al., 2010).

By combining protein- protein and protein- DNA interaction studies, the various transcription factors active in ESCs cluster into two different modules: the Oct4- centric module, and the Myc- centric module (Chen et al., 2008; Kim et al., 2008a, 2010). The Oct4- centric module is also referred to as the core transcriptional network which consists aside from the master transcriptional regulators Oct4, Nanog and Sox2 of other transcription factors important for pluripotency (see also Figure 9). The Myc- centric module consists mainly of proteins associated with cell cycle regulation and metabolism, including c-Myc, n-Myc, Zfx, E2F1 and E2F4. Interestingly, components of the Myc regulated network include also various chromatin modifying enzymes like the HATs GCN5, p300 and Tip60-p400 complex (Kim et al., 2008a, 2010; Lin et al., 2009).

Interestingly, recent data suggest that the key regulators Oct4, Nanog and Sox2 also play crucial roles in cell fate choice and initiation of developmental processes. By integrating external differentiation signals which modulate Oct4 and Sox2 protein level and change their genome wide binding properties, lineage selection is induced without prior activation of any lineage specific markers (Boiani and Scholer, 2005; Thomson et al., 2011)

1.3.2 The epigenetic landscape in embryonic stem cells

ESCs harbor a unique epigenetic landscape characterized by a relatively less condensed, open chromatin structure compared to the more restricted, closed chromatin conformation observed in differentiated cells (Efroni et al., 2008). Possibly as a consequence of the open chromatin structure, global transcriptional hyperactivity is observed in ESCs together with an enrichment of activating histone marks including histone H3 trimethylation at K4, K36 and K79 and acetylation of H4K16. In line with this, repressive histone modifications like trimethylation of histone H3 at K9 and K27 are underrepresented in ESCs (Gaspar-Maia et al., 2010). Furthermore, many chromatin associated proteins like Histone H1 and Hp1 show hyperdynamic binding behaviors in ESCs and heterochromatic regions are more dispersed and less abundant compared to differentiated cells (Meshorer et al., 2006). Recent data demonstrate a similar open chromatin conformation in cells of the ICM of mouse blastocysts at day 3.5, confirming that the chromatin structure of ESCs indeed resembles the *in vivo* structure in the embryo (Ahmed et al., 2010).

In addition to the open chromatin structure, ESCs also harbor distinct epigenetic modifications that contribute to the maintenance of pluripotency and self renewal. Genome-wide detection of trimethylation of H3K4 and H3K27 in ESCs revealed a unique distribution

of these marks that reflect the cellular state. Whereas actively transcribed genes are marked by trimethylation of H3K4, stably repressed genes in ESCs are enriched in trimethylation of H3K27 (Guenther et al., 2007; Mikkelsen et al., 2007; Pan et al., 2007; Zhao et al., 2007). Repressed genes which are crucial for differentiation- related processes carry bivalent chromatin domains leaving them poised for rapid activation or silencing during lineage commitment. The occurrence of bivalency in ESCs has been suggested to be integral for pluripotency as it is used as a mechanism to transiently repress lineage marker genes without permanently silencing them (see also chapter 1.1.2) (Bernstein et al., 2006). Interestingly, genes which carry neither of the two histone marks have been shown to be targets of DNA methylation and are almost exclusively classified as intermediate and low CpG containing promoters. More specifically, around 30 % of all genes in ESCs are controlled by DNA methylation and are predominately involved in differentiation and reproduction and therefore mainly repressed in ESCs. Consequently, housekeeping genes and genes crucial for pluripotency are unmethylated in the undifferentiated state and are enriched for trimethylation of H3K4 marks. Consistent with this, comparison of binding data for Oct4, Nanog and proteins of the PcG complex with mapping data of histone H3K4/ K27 trimethylation and DNA methylation marks revealed hardly any overlap between the different components and chromatin marks that contribute to the maintenance of pluripotency and self- renewal. Hence, it has been suggested that DNA methylation functions as an additional, parallel layer of epigenetic control to suppress differentiation and keep the undifferentiated state together with the pluripotency core network and actions of PcG associated proteins (Fouse et al., 2008).

Taken together, all these distinct epigenetic modifications and chromatin features ensure a highly dynamic state that confers the necessary plasticity to ESCs to maintain pluripotency but also to integrate intrinsic and extrinsic signals to permit progression towards differentiation programs. Whereas epigenetic factors and chromatin remodelers participate to stabilize the pluripotent state, they are usually not strictly necessary for the establishment and maintenance of pluripotency. In support of this, loss of components of PRC1, PRC2 or both does not impair the self renewal capacity of ESCs (O'Carroll et al., 2001; Wang et al., 2002; Voncken et al., 2003; Leeb et al., 2010). However, catalytical inactivation of both, PRC1 and PRC2, results in unstable ESCs with more spontaneous differentiation in culture (Leeb et al., 2010). Consistently, ESCs lacking Dnmts are viable and show no defect in proliferation (Tsumura et al., 2006). Therefore, it has been suggested that the pluripotent state of ESCs represents a ground state which is relatively independent from epigenetic regulation. However, as most ESCs lacking key epigenetic factors are impaired in their ability to differentiate it is postulated that epigenetic regulation plays especially a crucial role during cell fate choice and lineage commitment (Silva and Smith, 2008; Leeb et al., 2010).

1.3.3 Differentiation of embryonic stem cells *in vitro* recapitulates early developmental processes

The early mouse development is characterized by two different stages of the embryo, the pre- implantation and post- implantation embryo and reflects the gradual loss of cell potency and concomitant progression of differentiation. During pre- implantation development, the zygote develops to the early blastocyst (day 3.5), which consists of two different cell lineages; the inner cell mass (ICM) and the trophoectoderm (TE) (Figure 10).

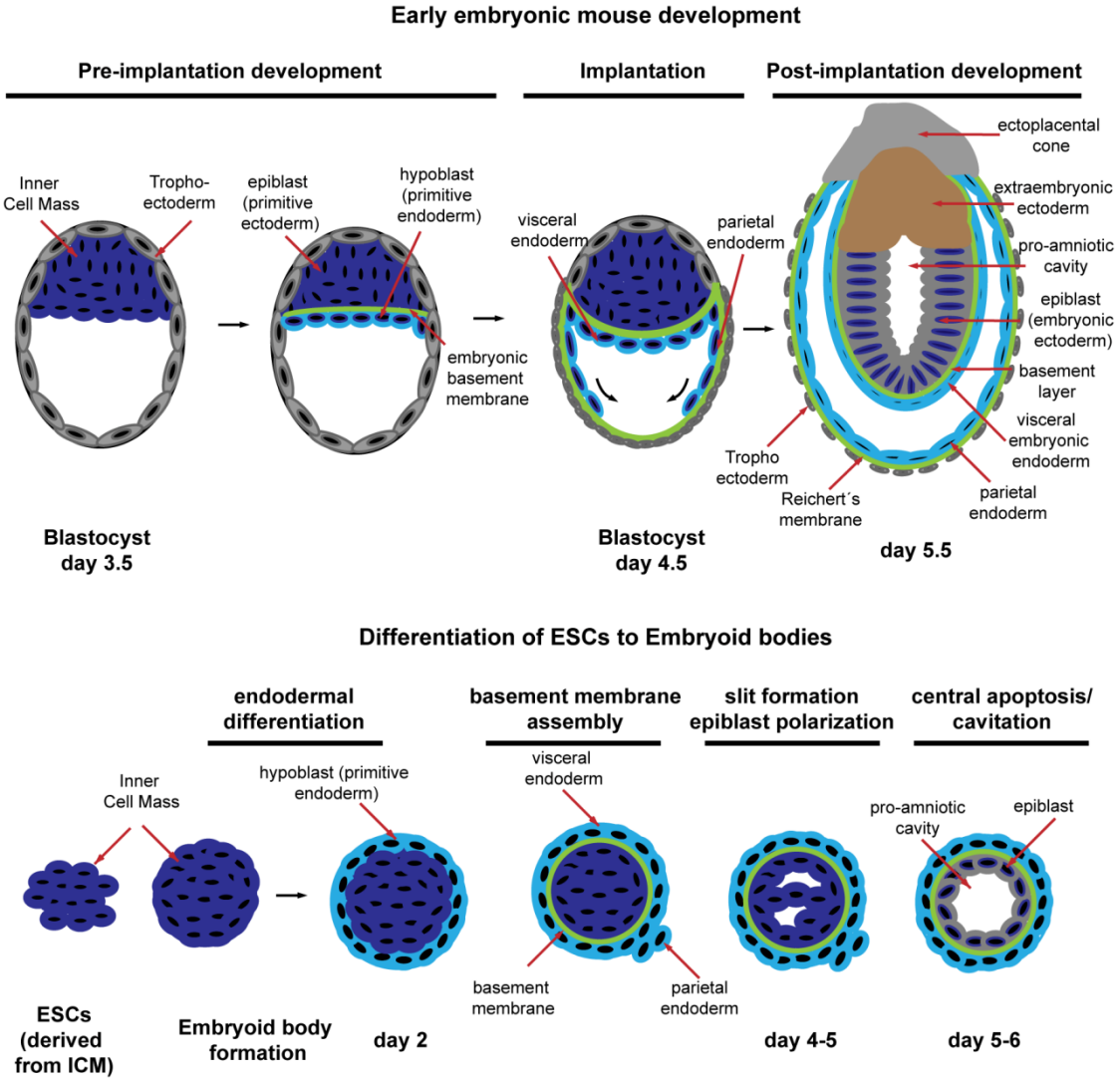


Figure 10. Comparison of mouse embryonic and embryoid body development
The differentiation of ESCs to embryoid bodies resembles early mouse developmental processes as they also initiate endodermal differentiation, followed by basement membrane assembly and epiplast polarization. After 5-6 days, a pro- amniotic cavity is built by apoptosis of the central cells in the epiblast (summarized in and adapted from (Li and Yurchenco, 2006).

Reciprocal repression of Oct4 in the ICM and caudal related homeobox2 (Cdx2) in TE seem to initiate the segregation of the two lineages in the pre- implantation embryo. More specifically, the ICM is characterized by high levels of Oct4 which inhibit Cdx2 expression in this lineage, whereas Cdx2 is highly expressed in the TE and negatively regulates Oct4

expression in the TE (Niwa et al., 2005). The TE cells give rise to extraembryonic tissues important for the implantation of the embryo. In contrast, the ICM further develops into the epiblast (primitive ectoderm) and the hypoblast (primitive endoderm) by day 4.5 (late blastocyst stage) which are separated by an embryonic basement layer. After implantation, the hypoblast gives rise to parietal and visceral primitive endoderm and together with the trophoblast contributes to extraembryonic tissues including the placenta. In contrast, the central cells of the epiblast undergo apoptosis, leading to the formation of the pro-amniotic cavity. Afterwards, during gastrulation, the epiblast gives rise to the three embryonic germ layers ectoderm, mesoderm and endoderm, which differentiate into all cell types of an organism. Notably, cells from the epiblast also contribute to the extraembryonic mesoderm of the yolk sac and additionally form primordial germ cells (PGCs) around day 7.

ESCs derived from the ICM can be differentiated in culture to further investigate processes of lineage commitment and cell fate specification. In principle, there are three different approaches to differentiate ESCs: i) monolayer culture of ESCs on extracellular matrix proteins ii) direct culture of ESCs on supportive stromal layers and iii) culture of ESCs as spherical cell aggregates called Embryoid Bodies (EBs). Among the different methods, the formation of EBs provides an undirected way to induce differentiation which recapitulates early developmental processes including temporal and spatial expression patterns during early embryogenesis. More specifically, the formation of EBs in culture resembles the transition from the ICM to a structure containing two germ layers which is similar to the early egg cylinder stage at the beginning of gastrula. Hence, EBs consist of an outer secretory endodermal layer and an inner epiblast layer, separated by a basement layer and produce a central cavity with progressive differentiation. Within the EB, progenitors and even terminally differentiated cells of all three germ layers can be generated including among others cardiogenic, myogenic, neurogenic and hematopoietic lineages (reviewed in (Guan et al., 1999). Several protocols exist to induce EB formation in culture; for example ESCs can be cultured in suspension in non- adherent dishes generating EBs with very heterogeneous size and shape. To form more homogenous EBs, the hanging drop method is usually used where a defined, equal number of cells is plated in a drop on inverted petri dishes, which allows the formation of EBs at the bottom of the drop. After 3-4 days, the EBs are then cultured in suspension to induce later differentiation stages (summarized in Desbaillets et al., 2000).

1.3.4 The epigenetic landscape changes dynamically during differentiation

As the *in vitro* differentiation of ESCs recapitulates early embryonic developmental processes it is widely used as a powerful tool to study epigenetic processes during early embryogenesis and lineage commitment. Already shortly after fertilization, the developing zygote undergoes major epigenetic reprogramming events. The male nucleus is extensively remodeled,

involving replacement of protamines with histones and global paternal- specific active DNA demethylation except for imprinted regions and repetitive sequences (see also chapter 1.1.3). Shortly afterwards, the maternal nucleus becomes passively demethylated as a consequence of the active exclusion of Dnmt1 from the nucleus (Carlson et al., 1992; Cardoso and Leonhardt, 1999). Around this time, the expression of pluripotency associated genes begins which are enriched in histone H3K4 trimethylation whereas developmental genes are kept silenced by the PcG repressive system. During differentiation, pluripotency genes are then silenced and at the same time, the expression of developmental genes is induced together with an increase of trimethylation of histone H3K4. Remarkably, a second wave of global DNA demethylation occurs around day 7 with the establishment of PGCs which includes reactivation of the X- chromosome and erasure of imprinted genes. Concordantly, somatic differentiation programs are repressed and the pluripotency potential is reacquired. During this time, pluripotent embryonic germ cells (EGCs) can be derived from PGCs under appropriate culture conditions (summarized in Orkin and Hochedlinger, 2011).

1.3.4.1 Role of DNA methylation during development

DNA methylation has been shown to play an important role in silencing pluripotency genes. Especially the multistep cascade for silencing of *oct4* during differentiation is well characterized and involves loss of the nucleosome- depleted regions at the distal enhancer, setting of repressive histone marks H3K9me3 by the histone methyltransferase G9a and subsequent recruitment of Hp1 which then is followed by *de novo* methylation of the promoter via Dnmt3a and Dnmt3b (Feldman et al., 2006; Li et al., 2007b; You et al., 2011). Besides pluripotency genes, also germ line specific genes and clusters like the protocadherin gene family as well as the reproductive homeobox X- linked (rhox) gene cluster have been identified as targets of DNA methylation (Oda et al., 2006; Illingworth et al., 2008; Borgel et al., 2010). Furthermore, DNA methylation seems to play a crucial role during lineage commitment as hypomethlyated embryos die early during development (Li et al., 1992, see also chapter 1.2.1). How *de novo* DNA methylation contributes to lineage commitment, is not yet well understood and several controversial studies about the developmental potential of hypomethylated cells exist. Results by Panning and Jaenisch suggest that differentiation of *dnmt1*$^{-/-}$ ESCs leads to apoptosis whereas Jackson and colleagues show that *dnmt1*$^{-/-}$ ESCs stay viable upon LIF removal, but fail to initiate differentiation (Panning and Jaenisch, 1996; Jackson et al., 2004). Furthermore, *dnmt1*$^{-/-}$ EBs were reported to express high levels of trophoblast markers and in contrast to wild type (wt) EBs can be differentiated into trophoblast stem cells under appropriate culture conditions (Ng et al., 2008). Similarly, it has been shown that TKO ESCs exhibit a growth defect and increased apoptosis upon EB differentiation. In addition, chimeric mouse embryos derived by aggregation of wt and TKO

embryos show that TKO derived cells mostly contributed to extraembryonic tissues but also to a small extent to the embryo proper at least detectable until an early postgastrulation stage (day 8.5) (Tsumura et al., 2006; Sakaue et al., 2010).

1.4 Aims of the work

In mammals, DNA methylation plays important roles in the epigenetic control of gene expression during development and differentiation and it is crucial for maintaining genomic stability. The functional significance of DNA methylation has been mainly inferred from genomic methylcytosine profiles in a limited selection of cell types and developmental stages and very little is known about how DNA methylation and Dnmts actually affect transcription programs during cellular differentiation. Based on the controversial reports about the developmental potential of hypomethylated cells (see previous chapter), it is still not clear to what extent hypomethylated cells are able to commit to and progress along embryonic lineages.

One of the main objectives of this thesis was to elucidate the role of DNA methylation and Dnmts during differentiation. To this aim I used wild type (wt) ESCs, ESCs lacking the maintenance Dnmt (*dnmt1*$^{-/-}$ ESCs) as well as ESCs lacking all three major Dnmts (*dnmt1*$^{-/-}$; *dnmt3a*$^{-/-}$; *dnmt3b*$^{-/-}$, or TKO ESCs) and generated their corresponding Embryoid Bodies (EBs). Using this differentiation system, I analyzed the potential of hypomethylated ESCs to silence pluripotency genes, studied their developmental potential and investigated whether differentiated EBs from all three cell lines can revert back to the undifferentiated state under appropriate culture conditions (chapter 3.1).

Recent reports identified the multi-domain protein Uhrf1 as an essential co- factor for maintenance DNA methylation. The second member of the Uhrf family, Uhrf2, is structurally very similar, but very little is known about its biological function(s). To gain first insights into the function of Uhrf2, I analyzed expression patterns of *uhrf1* and *uhrf2* in various cellular contexts and investigated whether Uhrf2 plays a role in maintenance DNA methylation (chapter 3.2).

The discovery of the "6th base" of the genome, 5hmC, a potential intermediate in DNA demethylation, and the Tet1-3 enzymes, which have been shown to be responsible for oxidation of 5mC to 5hmC, raised fundamental questions about the biological relevance of this newly identified modification. To advance understanding of the functions(s) of 5hmC and Tets, I analyzed expression levels of *tet1-3* in ESCs, during differentiation and in different tissues and used a newly identified, 5hmC specific endonuclease to map 5hmC levels in genomic DNA (chapter 3.3).

Finally, I aimed at analyzing whether designer transcription activator- like effectors (dTALEs) can be used for activation of target promoters, whether the epigenetic state of a target promoter sequence interferes with the action of dTALEs and how this interference can be overcome (chapter 3.4).

2. Materials and Methods

2.1 Materials

2.1.1 Technical devices

Devices	Type	Supplier
Agarosegel system	Mupid-Ex	Advance co
Bacterial incubator	UL 40	Memmert GmbH
Bacterial shaker	Certomat H+R	B.Braun
Cell culture microscope	EVOS xl	AMG
Centrifuge	Avanti J30I	Beckman Coulter GmbH
CO_2 incubator	Binder CB150	BINDER Inc.
Epiflourescence microscope	Axiophot 2	Carl Zeiss MicroImaging GmbH
FACS Aria II instrument	Sorp	Becton Dickinson
Fixed angle rotor	fixed angle 1720	Hettich Zentrifugen
Fixed angle rotor	JA-14	Beckman Coulter GmbH
Freezer (-20°C)	Comfort	neoLab Migge Laborbedarf GmbH
Freezer (-80°C)	MDF-594	SANYO GmbH
Fridge	Premium	Liebherr
Geldocumentation system	UV System	INTAS
Imaging system	Typhoon Trio Variable Mode Imager	GE Healthcare
Liquid Scintillation Analyzer	Tri-Carb 2100TR	Packard
Magnetic Stirrer	Yellowline MSH-basic	IKA GmbH & CO. KG
Microwave		Siemens AG
Photometer	NanoVue	GE Healthcare
Pipettes	Eppendorf Research)	Eppendorf AG
Pipettor (Mobile)	PIPETBOY acu	INTEGRA Biosciences GmbH
Power Supply unit	Bio-Rad PowerPac 300	Bio-Rad Laboratories GmbH
Real-Time PCR System	7500 Fast	Applied Biosystems
Roller mixer	RM5	CAT
SDS PAGE system	Mini-Protean Tetra	Bio-Rad Laboratories GmbH
Shaker	DOS-10L	neoLab Migge Laborbedarf GmbH
Sonifier	Branson Digital Sonifier 450D	G. Heinemann Ultraschall- und Labortechnick
Sterile Bench	Herasafe KS, Class II	Fisher Scientific GmbH
Table centrifuge	Mikro 22R	Hettich Zentrifugen
Table top centrifuge	Centrifuge 5454	Eppendorf AG
Table top centrifuge	Rotina 38R	Hettich Zentrifugen
Thermo shaker	Thermomixer comfort	Eppendorf AG
TECAN infinite M1000 plate reader		Tecan
Vortex mixer	Neolab 7-2020	neoLab Migge laborbedarf GmbH
Water bath	Type 1013	GFL
Waving platform shaker	Polymax 1040	Heidolph Instruments GmbH&Co

2.1.2 Consumables

Consumables	Supplier
Cell culture plates & flasks	Falcon, Becton Dickinson GmbH
Centrifuge bottle (polypropylene, 250ml)	Beckman Coulter GmbH
Erlenmeyer flasks	SCHOTT AG
Falcon Tubes (15ml, 50ml)	Becton Dickinson GmbH
Glass Pasteur pipettes (230mm)	Brand GmbH +Co KG
Laboratory bottle (100ml, 250ml, 500ml, 1L)	SCHOTT AG
Latex exam gloves “Satin PLUS”	Kimberely-Clarke Europe
Microscope coverslips (Ø20mm)	Menzel GmbH + Co KG
Nitrile laboratory gloves	SLG Süd-Laborbedarf GmbH
Parafilm M sealing film	neoLab Migge Laborbedarf GmbH
Pipette tips (10µl, 300µl, 1000µl)	BrandTech Scientific
Pipette tips with filter (10µl, 200µl, 100µl)	SLG Süd-Laborbedarf GmbH
Reaction tubes (1.5ml)	Eppendorf AG
Reaction tubes (0.2ml, 0.5ml)	Eppendorf AG
Serological pipettes	Carl Roth
Soft wipes (KIMTECH science)	Kimberely-Clark Europe
Snaptwist vials	Simport
Whatman filter paper	Whatmann GmbH

2.1.3 Reagents and consumables

Reagents & Kits	Supplier
5-aza-2'-deoxycytidine	Sigma-Aldrich
7-amino-actinomycin D	Invitrogen
5-hydroxymethyl-dCTP	Bioline GmbH
5-methyl-dCTP	Jena Bioscience GmbH
Accutase	PAA Laboratories GmbH
Alexa Flour® 647 Mouse anti-Oct3/4	BD Biosciences
Alexa Flour® 647 Mouse IgG1 K isotype control	BD Biosciences
Ammoniumperoxodisulfate (APS)	Carl Roth GmbH
Anti-oct4 mouse (goat)	Santa Cruz
Anti-goat-Alexa Fluor 647 (rabbit)	Molecular Probes
ß-mercaptoethanol	Invitrogen GmbH
B27 (50x)	Invitrogen GmbH
Betaine	Sigma
BSA	PAA Laboratories GmbH
Bromphenol blue sodium salt	AppliChem GmbH
Cell Dissociation Buffer	Invitrogen GmbH
Chloroform	Roth GmbH
Cytofix	BD Biosciences
4’,6-Diamidino-2-phenylindol (DAPI)	Roche Diagnostics
DMEM/F12	Invitrogen GmbH
DMEM, high glucose with L-glutamine	PAA Laboratories GmbH
Dimethylsulfoxide (DMSO)	AppliChem GmbH
Disodiumhydrogenphosphate (Na_2HPO_4)	Carl Roth GmbH
dNTPs	PeqLab

Eco Plus scintillation liquid	Carl Roth GmbH
EDTA-dihydrate	AppliChem GmbH
EGF	Peprotech
Ethanol (98%)	AppliChem GmbH
Ethidiumbromide	AppliChem GmbH
Euromed-N	Euroclone
EZ DNA Methylation-Gold Kit	Zymo research
Fetal bovine serum (FBS)	PAA Laboratories GmbH
FGF2	Peprotech
Formaldehyde (36,5%)	Sigma-Aldrich
Fugene HD	Roche Diagnostics
Gelatine	Sigma
GeneChip Mouse Gene 1.0 ST microarrays	Affymetrix
Gentamicin (50mg/ml)	PAA Laboratories GmbH
GlutaMax I (200mM)	Invitrogen GmbH
Glycerol	Carl Roth GmbH
Glycin	Carl Roth GmbH
HEPES	PAA Laboratories GmbH
High-Capacity cDNA Reverse Transcription Kit with RNase Inhibitor	Applied Biosystems
Hydrochloric acid (HCl)	Carl Roth GmbH
Hoechst 33258	Invitrogen GmbH
Isopropanole	Carl Roth GmbH
Isopropyl β-d-thiogalactopyranoside (IPTG)	AppliChem GmbH
Knock-out DMEM	Gibco
LIF (ESGRO)	Millipore
L-glutamine (200mM)	PAA Laboratories GmbH
Lipofectamine 2000	Invitrogen GmbH
LR recombination	Invitrogen GmbH
Magnesium chloride ($MgCl_2$)	Sigma-Aldrich
Mammalian Protease Inhibitor	Calbiochem Merck
MEM Non-essential Amnio Acid Solution	PAA Laboratories GmbH
N2 (100x)	Invitrogen GmbH
Neurobasal medium	Invitrogen GmbH
Nickel-nitrilotriacetic acid column	QIAGEN
Nonylphenoxypolyethoxylethanol (NP-40)	Sigma
Nucleospin	Macherey-Nagel
NucleoSpin Triprep Kit	Macherey-Nagel
OptiMEM	Invitrogen
PBS (phosphate buffer saline)	PAA Laboratories GmbH
Penincillin/Strepomycin	PAA Laboratories GmbH
Perm/Wash Buffer	BD Biosciences
pET28b vector	Novagen
pENTR-D-TOPO	Invitrogen GmbH
pGL-3 basic vector	Promega
Phusion HF DNA Polymerase	Finnzymes
Polyethylenimine (PEI)	Sigma
Potassium chloride	Carl Roth GmbH

Potassiumdihydrogenphosphate (KH_2PO_4)	Merck
Power SYBR Green PCR Master Mix	Applied Biosystems
Propidium Iodide	Sigma
Qiagen Hot Start Polymerase	QIAGEN
QIAmp DNA Mini Kit	QIAGEN
ResourceQ anion exchange column	GE Healthcare
Rotiphorese Gel 30 (Acrylamide)	Carl Roth GmbH
RNAse A	AppliChem
RNase-free DNase I	Roche Diagnostics
RNeasy kit	QIAGEN
siRNA, ON-TARGET plus non-targeting pool, mouse	Dharmacon, Thermo Scientific
siRNA, ON-TARGET plus Set of 4, uhrf2 mouse	Dharmacon, Thermo Scientific
Smart Ladder	Eurogentec Deutschland GmbH
Sodium chloride (NaCl)	Carl Roth GmbH
Sodium hydroxide (NaOH)	Carl Roth GmbH
Sodium sulfate (Na_2SO_4)	Sigma-Aldrich
StrataClone™ PCR Cloning Kit	Agilent Technologies
StrataCone™ SoloPack® Competent Cells	Agilent Technologies
Streptomycin	PAA Laboratories GmbH
Superdex S-200 preparative gel filtration column	GE Healthcare
S-SYBR green I nucleic acid gel stain (10.000x)	Invitrogen GmbH
T4 DNA Ligase	New England Biolabs
TaqMan probes	Applied Biosystems
TaqMan Gene expression Master Mix	Applied Biosystems
TEMED	Merck
Tetramethylammonium-chloride (TMAC)	Sigma-Aldrich
Trichloroacetic acid (TCA)	Sigma-Aldrich
Trichostatin A	Sigma-Aldrich
Tris	Carl Roth GmbH
Triton X-100	Carl Roth GmbH
Trizol	Invitrogen GmbH
Trypsin/EDTA	PAA Laboratories GmbH
Tween	Carl Roth GmbH
Universal agarose	Bio&SELL e.K.
UDP-glucose	Sigma-Aldrich
UDP-[^{3}H]glucose	Hartmann Analytic GmbH
Valproic acid sodium salt	Sigma-Aldrich
Vectashield	Vector Laboratories
WT Expression Kit	Ambion
WT Terminal Labeling and Controls Kit	Affymetrix
Zero Blunt® PCR Cloning Kit	Invitrogen

2.1.4 Cell lines

ESCs	Parental cell line	Genetic background	Introduced modification
J1 wt	J1	129/sv	Wild type
dnmt1$^{c/c}$ J1	J1	129/sv	*dnmt1 c* allele homozygous null
TKO	J1	129/sv	*dnmt1, dnmt3a and 3b* triple homozygote null
E14 wt	E14	129/ola	Wild type
uhrf1$^{-/-}$ E14	E14	129/ola	homozygote *uhrf1* null
JM8A3.N1	JM8	C57Bl/6N	Wild type
uhrf2$^{-/+}$ JM8A3.N1 EPD0373-2C02	JM8	C57Bl/6N	heterozygote *uhrf2* null, clone EPD0373-2C02
ogESCs	wt J1	129/sv	stable cell line expressing eGFP driven by the oct4 promoter (pc2055)
ogNSCs	wt J1	129/sv	stable cell line expressing eGFP driven by the oct4 promoter (pc2055), differentiated to neural stem cells (NSCs)

somatic cell	description
C2C12	Mouse myoblasts
NIH3T3	Mouse fibroblasts
HEK293T	Human embryonic kidney

2.1.5 Primer sequences

2.1.5.1 TaqMan Assay ID numbers for relative quantification using qPCR

All PCRs were carried out using the standard program (see 2.2.2.3) and Annealing temperature of 60°C.

Gene Name	Assay ID	Amplicon length
dnmt3a	Mm00432884_m1	80 bp
dnmt3b	Mm01240113_m1	83 bp
gapdh	Mm99999915_g1	107 bp
oct4	Mm00658129_gH	100 bp
nanog	Mm01617761_g1	100 bp
uhrf1	Mm00477865_m1	83 bp
uhrf2	Mm00520043_m1	62 bp

2.1.5.2 SYBR Green Primer sequences for relative quantification using qPCR

All PCRs were conducted using standard conditions (see 2.2.2.3).

Gene	Forward Primer (5´-3´)	Reverse Primer (5´-3´)	Amplicon length
brachyury	CTC CAA CCT ATG CGG ACA ATT C	ATG ACT CAC AGG CAG CAT GCT	110 bp
dnmt1	GGC GGA AAT CAA AGG AGG AT	CCT GGG TCT GGA ACT TCT TTT ATC	101 bp
eomes	ACC GGC ACC AAA CTG AGA TGA	GGG GTT GAG TCC GTT TAT GTT GAA	85 bp
fgf5	GAT CTA CCC GGA TGG CAA AG	TGC TGA AAA CTC CTC GTA TTC CT	110 bp
gapdh	CAT GGC CTT CCG TGT TCC TA	CTT CAC CAC CTT CTT GAT GTC ATC	100bp
gata6	CAA AAG CTT GCT CCG GTA ACA	GGT CGC TTG TGT AGA AGG AGA AG	110 bp
hnf4a	CAA GAG GTC CAT GGT GTT TAA GG	CGG CTC ATC TCC GCT AGC T	90 bp
nestin	ACT CTG CTG GAG GCT GAG AAC T	CAA GGA AAT GCA GCT TCA GCT T	100 bp
pcna	GAC TTA GAT GTG GAG CAA CTT GGA	GGC TAA GGT CTC GGC ATA TAC G	100 bp
sox1	GCT TCG GAG GAC AAA AGA CAA	AAG AGC TGG CGG GAA GTA AAC	101 bp
tet1	CCA GGA AGA GGC GAC TAC GTT	TTA GTG TTG TGT GAA CCT GAT TTA TTG T	100 bp
tet 2	ACA AAG CTG ATG GAA AAT GCA A	GGT GCC TCT GGA GTG TTG GT	100 bp
tet3	GAG CAC GCC AGA GAA GAT CAA	CAG GCT TTG CTG GGA CAA TC	100 bp

2.1.5.3 SYBR Green Primer sequences for quantification of PvuRts1I digested products using qPCR

PCRs were performed using standard conditions (see 2.2.2.3).

Name	Forward Primer (5´-3´)	Reverse Primer (5´-3´)
nanog primer 1	GAT TTC CCC AGG TTT CCC AAT	GAG TCA GAC CTT GCT GCC AAA
nanog primer 2	TGC CTA ATG ACA AGA ATC ACA TCA	GCC CCT AAG TAG AAA TCA TAG AAC A
oct4 primer	GAG AAG GAT GTG AGT GCC AAG AT	GGA ATG GGA ACA GGG AAA CA

2.1.5.4 Primer sequences for bisulfite sequencing

All PCR reactions were carried out as semi-nested PCRs.

Gene	Forward Primer (5´-3´)	Reverse Primer (5´-3´)
brachyury (PCR 1)	GGT GGG AGT TAG TGG TAG TTT AT	CAA AAC CCT AAC TCC TAA AAC CA
brachyury (PCR 2)	TTT AAA GTT GTT ATA TTT GGG GAG GT	= Reverse from PCR1

fgf5 (PCR1)	TAG GGT GGT TTT TTA GTG GAG AAA T	ATT ATC AAA AAC CAC CCA ATC ACC
fgf5 (PCR2)	ATG GTA GGG GTT AGT AAT TTG GAA	= Reverse from PCR1
nestin (PCR1)	GGA AGA ATT TTT TTA GAT GTG GGA G	CAA CCT AAA TAC TCA ACC ACC C
nestin (PCR2)	GAG GAG TAG AAT TAG TTG TTT AGT	= Reverse from PCR1
sox1 (PCR1)	GTT AGT TTA GGT TGG GTT TTA TGA AAT	AAT CCC TAT CTC AAA ACC TAC TAC
sox1 (PCR2)	= Forward from PCR1	CCT AAA CCT ATC AAT ATA AAC CCT AT
tet1 (PCR1)	TTTTTAGGATGTTATTTGGATGATT	CACAACCTTTACTAAACCCTATACC
tet1 (PCR2)	= Forward from PCR1	TATCTCCCCAATACAAACCTC

2.1.5.5 Primer sequences for Pyrosequencing

Gene	Forward Primer (5´-3´)	Reverse Primer (5´-3´)
oct4 (PCR 1)	ATG GGT TGA AAT ATT GGG TTT ATT TA	ACC CTC TAA CCT TAA CCT CTA AC
oct4 (PCR 2)	GTA AGA ATT GAG GAG TGG TTT TAG	= Reverse from PCR1 with 5´biotinylated
nanog (PCR1)	TAG TTT GGG TTA TTT TAT AGT TTT TTT TG	CCA AAA AAA CCC ACA CTC ATA TC
nanog (PCR2)	AAT GTT TAT GGT GGA TTT TGT AGG T	= Reverse from PCR1 with 5´biotinylated
major satellite	AAAATGAGAAACATCCACTTG	CCATGATTTTCAGTTTTCTT
h19 promoter	ATAGTTATTGTTTATAGTTT	AGGAATATGTTATATTTAT

2.2 Methods

2.2.1 Methods of Cell Biology

2.2.1.1 Cultivation of ESCs and somatic cells

The undifferentiated mouse ESCs derived from the genetic background 129/ sv and 129/ ola (see 2.1.4) were maintained on gelatin-coated dishes in DMEM supplemented with 16 % FBS , 0,1 mM ß- mercaptoethanol, 2 mM L- glutamine, 1x MEM Non- essential Amnio Acid Solution, 100 U /ml pencillin, 100 µg/ ml streptomycin and 1000 U/ ml recombinant mouse. Mouse ESCs derived from C57Bl6/ N strain were cultured in the same medium except Knock- out DMEM was supplemented with 10 % FBS. The somatic cell lines, C2C12 myoblasts, NIH3T3 fibroblasts and HEK293T kidney cells were cultured in DMEM supplemented with 20 % or 10 % FBS, respectively and 50 µg/ ml gentamycin. All cell lines were cultured in a 37°C incubator with a humified atmosphere of 5 % CO2. Cells were spilt every two to three days in a ratio ranging from 1:8 - 1:10 for ESCs and HEK293T, 1:20 – 1:40 for C2C12 and 1:4 – 1:6 for NIH3T3 cells.

2.2.1.2 Generation of transgenic cell lines

The generation of the transgenic cell line was done by Sebastian Bultmann. J1 wt ESCs were transfected with p*oct4*-eGFP promoter construct using Fugene according to the manufacturer´s instructions and cells with stable integrated reporter contruct were identified by repeated FACS sorting for eGFP expression using the FACS Aria II instrument. Finally, to obtain a clonal transgenic cell line (ogESCs), single cell sorting was performed.

2.2.1.3 Differentiation of ESCs to neural stem cells (NSCs)

Differentiation of ogESCs to ogNSCs was conducted by Sebastian Bultmann. In brief, $3{,}5x10^5$ cells were plated in a 25 cm² culture flask with N2B27 medium containing 1000 U/ ml of LIF (Ying et al., 2003b). To induce neural differentiation, medium was replaced with N2B27 without LIF the next day. After 7 days cells were plated in Euromed-N supplemented with 20 ng/ ml EGF and FGF2. After additional 5 days, neurospheres were collected and plated in gelatin coated flasks in the same medium to allow outgrowth of NSCs. NSCs were then maintained in N2B27 medium containing 20 ng/ ml EGF and FGF2.

2.2.1.4 Differentiation of ESCs to EBs

To induce EB formation, ESCs were resuspended in the same medium as above but without LIF and cultured in hanging drops (600 cells/ 20 µl drop) for four days. Culture dishes were filled with PBS to avoid drying of the hanging drops. After four days, EBs were cultured in

bacterial culture dishes to avoid attachment of the EBs and the medium was replaced every four days.

2.2.1.5 Replating of EBs

For replating of cells from late EBs, EBs were washed three times with PBS and dissociated to single cells by repeated cycles of Accutase treatment at 37°C. The reaction was stopped by adding fresh EB medium (ES cell medium without LIF) to the dissociated cells. After centrifugation, dissociated cells were resuspended in fresh EB medium and equal numbers of cells were seeded in duplicate on gelatin- coated plates and LIF was added to one of the duplicate plates to a final concentration of 1000 U/ ml.

2.2.1.6 Transfection of plasmids

For transient reporter assays of oct4 dTALEs, HEK293T cells were transfected with polyethylenimine. Transfections of ogESCs and ogNSCs were carried out using Lipofectamine 2000 according to the manufacturer´s instructions.

2.2.1.7 Transient knock- down using small interfering RNAs (siRNAs)

For silencing of *uhrf2* in ESCs, two different siRNAs targeting *uhrf2* as well as a pool of non-targeting siRNA were used to control for off- target effects. Cells were plated at a density of 6-7*10^4 cells/ cm^2 in normal ESC medium. After approximately one hour, siRNA diluted in OptiMEM with a final concentration of 50 nM was transfected using Lipofectamine 2000 according to the manufacturer´s conditions. Medium was changed after 24 hours and cells were split and transfected every second day for a total of 8-10 days. To monitor for knock-down efficiency, residual cells were harvested and lysed in RP1 buffer of the Triprep kit every second day. Lysates were stored at -80°C until the end of the experiments, so that all samples could be processed simultaneously.

2.2.1.8 Treatment of ogNSCs with epigenetic inhibitors

Valproic acid sodium salt (VPA) and 5-aza-2'-deoxycytidine (5-azadC) was dissolved in PBS at a concentration of 250 mM and 30 mM, respectively and sterile filtered. Trichostatin A (TSA) was dissolved in DMSO at a concentration of 5 mM. Prior to treatment with epigenetic inhibitors, ogNSCs were transfected with the T-83 construct (for details see 2.2.1.6) and after 12 hours, medium was exchanged with medium containing dilutions of the respective inhibitor or a combination thereof and cells were cultured for additional 36 hours. Final concentrations of the inhibitors were for TSA 30 nM, VPA 620 µM, 5-azadC 10 nM and for the combination of VPA 310 µM + 5-azadC 5 nM.

2.2.1.9 Intracellular protein staining using FACS

To stain intracellular Oct4 protein levels during differentiation, EBs were washed twice in PBS and incubated in Cell Dissociation Buffer for 45 min at 37°C. An equal volume of EB medium was added and EBs were dissociated to single cells by vortexing. To distinguish dead from living cells, 7- amino- actinomycin D (7- AAD) was added at a concentration of 5 µg/ ml for 20 min at 4°C. After multiple washing steps in PBS, cells were fixed in Cytofix for 20 min at 4°C and washed twice in Perm/ Wash solution. Antibody staining was performed in Perm/ Wash buffer for 30 min at 4°C using Alexa Flour® 647 Mouse anti- Oct3/4 and Alexa Flour® 647 Mouse IgG1 K isotype control. After washing twice with Perm/ Wash buffer, cells were resuspended in Annexin buffer (10 mM HEPES, 140 mM NaCl, 2.5 mM $CaCl_2$, pH 7.4) and analyzed with a FACS Aria II instrument. Data analysis was performed using FlowJo version 7.2.5.

2.2.1.10 Serum starvation of fibroblasts

To analyse whether *uhrf2* shows an expression pattern that is dependent on the proliferative state of the cells, NIH3T3 cells were plated to about 60 % confluency one day before starvation. After washing twice with PBS, cells were cultured in starvation medium containing 0.5 % FBS and 1 % Gentamycin for 48 hours. Cells were released from starvation by adding fresh medium and samples were taken every 3, 6, 9, 12 and 24 hours after serum addition.

2.2.1.11 Analysis of cell cycle profile using FACS

After trypsinization and resuspension of the cell pellet in PBS, cells were fixed by slowly adding 100 % Ethanol while simultaneously vortexing to ensure homogenous fixation of the cells. Cells were incubated on ice for 15 min, washed twice in PBS and incubated in PBS containing 50 µg/ ml Propidium Iodide (PI), 0.1 mg/ ml RNAse A and 0.05 % Triton X-100 for 30 min at 37°C. Cells were then analyzed using a FACS Aria II instrument and data analysis was performed using FlowJo version 7.2.5.

2.2.1.12 Immunofluorescence staining

Immunostaining of Oct4 protein in ogNSCs was performed by Sebastian Bultmann. In brief, ogNSCs were grown on cover slips, transiently transfected with the oct4- dTALE T-83, fixed with 2 % paraformaldehyde in PBS and permeabilized in PBS containing 0.2 % Triton X-100. Oct4 was stained using a primary goat- anti- Oct4 antibody (dilution 1:1000) and a secondary rabbit- anti- goat antibody coupled to Alexa Fluor 647 (dilution 1:2000). The antibodies were diluted in PBS containing 0.02 % Tween 20 and 2 % BSA. Cells were counterstained with DAPI and mounted in Vectashield (Vector Laboratories) Images were acquired with a Zeiss Axioplan 2 fluorescence microscope equipped with a Plan-NEOFLUAR 40x/1,3 oil objective.

2.2.2 Methods of Molecular Biology

2.2.2.1 RNA extraction

Total RNA from cultured cells, EBs and tissue samples from 6 weeks old 129sv mice were isolated with TRIzol reagent or the NucleoSpin Triprep Kit according to the manufacturer´s instructions. To avoid genomic DNA contamination, isolated RNA was digested with recombinant RNase- free DNase I and further purified with the QIAGEN RNeasy kit. RNA concentration and purity was determined using a NanoVue spectrophotometer. Quality of RNA was assessed by agarose gel electrophoresis and RNA was stored at -20°C - -80C°.

2.2.2.2 cDNA synthesis

500 ng total RNA were used for cDNA synthesis with the High- Capacity cDNA Reverse Transcription Kit with RNase Inhibitor according to the manufacturer´s instructions and using random hexamer. Samples without reverse transcriptase were prepared in parallel to control for genomic DNA contamination. cDNA was stored at -20°C.

2.2.2.3 Quantification of mRNA levels using Real- time PCR

Transcript levels of genes of interest were quantified relative to transcript levels of stably expressed housekeeping genes with quantitative PCR using a 7500 Fast Real- Time PCR cycler with a 96 well block. For quantification, either TaqMan probes in combination with TaqMan Gene expression Master Mix or Primers designed with Primer Express were used together with the Power SYBR Green PCR Master Mix. TaqMan gene expression assay IDs and Primer sequences are listed in 2.1.5.1 and 2.1.5.2, respectively. Each sample was analyzed in technical triplicates for each primer set with a total volume of 10 µl. In every reaction, equal amounts of cDNA were diluted 1:25 – 1:50 and final concentrations for each SYBR green primer pair were 500 nM. TaqMan assays were used as recommended by the manufacturer. All primer pairs were tested for specificity and efficiency by using non template controls, melting curve analysis (in the case of SYBR green primers after each run), agarose gel electrophoresis and dilution series of cDNAs. The following program was used for amplification:

Pre-incubation	2min	50°C
Enzyme activation	10min	95°C
	15sec	95°C
40x cycles of amplification	1min	60°C
	5sec	95°C
Melting curve	1min	60°C
	2.5°C/ min	95°C

Gene expression levels were normalized to *gapdh* and calculated with the sequence detection system software v1.3 using the comparative CT method (ΔΔCT method).

2.2.2.4 Genome- wide expression profiling using Microarrays

Microarray hybridization was carried out by our collaborators Dietmar Martin and Kerstin Maier at the Gene Center Affymetrix Microarray Platform (Laboratory of Patrick Cramer, LMU Munich). Sample preparation for microarray analysis was performed using the WT Expression Kit and WT Terminal Labeling and Controls Kit with 300 ng input RNA. Samples were hybridized to GeneChip Mouse Gene 1.0 ST microarrays according to the manufacturer's instructions. Bioinformatical analysis including Principal Component Analysis (PCA) was performed by out collaborators Benedikt Zacher and Achim Tresch (Laboratory of Achim Tresch, LMU Munich). Quality control, normalization and further statistical analyses were carried out using the R/Bioconductor programming environment. Linear Models for Microarray Data (limma) was used to compute fold changes and p-values (Smyth, 2005). Genes with a fold change of 2 and a false discovery rate below 0.05 were considered differentially expressed. These genes are listed in the appendix (chapter 6). A one-sided Kolmogorov-Smirnov test was used to compare gene expression changes in wt and knockout EBs globally as well as for bivalent and non-bivalent genes between day 0 and 4 of differentiation. I performed Gene ontology (GO), tissue expression and chromosomal location analysis using DAVID (Huang et al., 2008, 2009). For GO enrichment analysis the lowest level of GO categories for biological process (GO_BP_ALL) was used. To identify significantly enriched GO categories, the corrected p-value calculated according to the Benjamini- Hochberg method was used for multiple testing corrections. Microarray data have been deposited into the GEO database under the accession number GSE36679.

2.2.2.5 DNA extraction

Genomic DNA from ESCs, EBs and tissue samples from 6 weeks old 129sv mice were extracted using either the NucleoSpin Triprep Kit or the QIAmp DNA Mini Kit according to the manufacturer´s instructions. Shearing of fragments and concentration measurements with Hoechst was done by Sebastian Bultmann. Isolated genomic DNA samples were sheared to 500 - 1500 bp fragments by sonication to reduce the viscosity and improve homogeneity and the concentration was measured by fluorometry. 50 µl of diluted sample were mixed with 50 µl of 2x TNE (Tris 20 mM, pH 7.4; NaCl 400 mM and EDTA 2mM) containing 200 ng/ µl of Hoechst 33258 and fluorescence was measured in a TECAN infinite M1000 plate reader (Ex: 350/10; Em: 455/10). Serial dilutions ranging from 20 to 2000 ng/ ml of the 5hmC containing reference DNA fragment (see below) were used as standard for quantification.

2.2.2.6 Preparation of reference DNA fragments for 5hmC glucosylation assay

Reference DNA fragments (1139 bp) containing 0 % and 100 % 5hmC were prepared by PCR, using dCTP and 5-hydroxymethyl- dCTP, respectively. As a template, T4 phage DNA

was used with the following primers: 5'-TGG AGA AGG AGA ATG AAG AAT AAT-3' and 5'-GTG AAG TAA GTA ATA AAT GGA TTG-3', Phusion HF DNA Polymerase and the following cycling profile: one cycle of 98°C, 2 min and 35 cycles of 98°C, 10s; 58°C, 10s; 72°C, 30s. Note that the Primer sequences do not contain any cytosine residues. PCR products were purified by gel electrophoresis and silica column purification.

2.2.2.7 Quantitative 5hmC glucosylation assay

Assays for quantification of 5hmC glucosylation were performed by Alexandra Szwagierczak and Sebastian Bultmann. Reactions contained 150 mM NaCl, 20 mM Tris, pH 8.0, 25 mM $CaCl_2$, 1 mM DTT, 2.8 µM "cold" UDP-glucose, 0.86 nM UDP- [^{3}H]glucose (glucose-6-^{3}H; 60 Ci / mmol;), 1 µg of DNA substrate and 36 nM recombinant β- gt in a total volume of 50 µl. After incubation for 20 min at RT, reactions were terminated by heating at 65°C for 10 min. From each reaction, 20 µl were spotted in duplicate on paper filters and DNA was precipitated by incubation in 5 % TCA for 15 min at RT. Afterwards, filters were washed twice with 5 % TCA and once with 70 % ethanol. Remaining radioactivity was measured using a Liquid Scintillation Analyzer Tri- Carb 2100TR with quench indicating parameter set on tSIE/ AEC (transformed spectral index of the external standard / automatic efficiency control) in 4 ml of Rotiszint Eco Plus scintillation liquid in snaptwist vials. Samples were measured for 30 min or until the 2 σ value reached 2 %. To calculate the percentage of 5hmC per total cytosine from the incorporation of [^{3}H]glucose, a calibration curve was used which was measured with the reference fragments for every experiment. The percentage of 5hmC was then corrected for the difference in C abundance between reference fragment (35 %) and mouse genome (42 %).

2.2.2.8 Preparation of reference DNA fragments for testing PvuRts1I specificity

Reference DNA fragments containing exclusively 5hmC, 5mC or unmodified cytosine residues were prepared by PCR using 5-hydroxymethyl- dCTP, 5-methyl- and dCTP, respectively. As a template for all reference fragments, T4 phage DNA was used for PCR amplification by Phusion HF DNA Polymerase and primer 5'-GTG AAG TAA GTA ATA AAT GGA TTG-3', which does not contain cytosine residues. To generate the reference 1139 bp fragment with 100 % 5hmC for restriction with PvuRts1I the second primer was 5'-TGG AGA AGG AGA ATG AAG AAT AAT- 3', which also does not contain cytosine residues. To generate the 800 and 500 bp control substrates containing only 5mC and only unmodified cytosine the second primer was 5'-GCC ATA TTG ATA ATG AAA TTA AAT GTA-3' and 5'-TCA GCA ATT TTA ATA TTT CCA TCT TC-3', respectively. After PCR amplification, products were purified by gel electrophoresis followed by silica column purification.

2.2.2.9 Preparation of DNA substrates for Linker – PvuRTS1I – analysis

To prepare substrates containing different 5hmC levels (0 %, 1 %, 2,5 %, 5 %, 10 %), genomic DNA from wt JM8A3.N1 was used as a template to amplify the Region III of the *nanog* promoter (Hattori et al., 2007) by PCR using Phusion HF DNA Polymerase (Finnzymes) and the following primers: For 5´ TCA GGA GTT TGG GAC CAG CTA 3´ and Rev 5´CCC CCC TCA AGC CTC CTA 3´, resulting in a 867 bp fragment. To generate *nanog* promoter substrates containing increasing 5hmC levels, 5-hydroxymethyl- dCTP and dCTP were added to the PCR reactions at appropriate ratios and PCR products were purified by silica column purification. Successful incorporation of 5hmC levels was confirmed by the quantificative 5hmC glucosylation assay. The radioactive assay was conducted by Sebastian Bultmann.

2.2.2.10 Preparation of linker

A linker containing a random 3´overhang was generated by annealing the following primer: For 5´CTC GTA GAC TGC GTA CCA TG NN 3´ and Rev 5´CA TGG TAC GCA GTC TAC CAG 3´. Successful annealing of oligos was checked by loading 100 ng of each oligo (annealed and unannealed) together with the loading dye xylenol mixed with bromphenol blue (ratio 1:1) on a 15 % non- denaturating PAGE. Ultra low range marker was used to control for size, gel was stained with SYBR green (1:1000) for 20 mins and scanned on a Typhoon imaging system using filter sets for excitation at 488 nm and emission of 520 nm.

2.2.2.11 Digestion with PvuRts1I

Sheared genomic DNA of wt and TKO ESCs or 250 ng of each fragment with increasing 5hmC content were digested with 2 U of PvuRts1I in reaction buffer containing 150 mM NaCl, 20 mM Tris, pH 8.0, 5 mM $MgCl_2$, 1 mM DTT for 15 min at 22°C followed by a heat inactivation of 20 min at 60°C. To control for successful digestion, reference substrates (see 2.2.2.8) were digested using the same conditions and analyzed on a 1 % agarose gel electrophoresis.

2.2.2.12 Detection of PvuRts1I digested fragments using qPCR

10 ng of digested or mock treated genomic DNA was amplified by qPCR using two different primer pairs specific for the *nanog* upstream regulatory region at standard cycling conditions (see 2.1.5.3). To ensure that equal amounts of gDNA were added to each well, a primer pair specifically amplifying the *oct4* locus was used to correct for pipetting mistakes. Each sample was analyzed in technical triplates and amplification of single products was tested by melting curve analysis after each run.

2.2.2.12 Detection of PvuRts1I digested fragments using linker - PCR strategy

25 ng of each digested fragment was ligated to a linker containing the random 3´ overhang. The ligation reaction was carried out using T4 DNA Ligase overnight at 16°C. As a control for the specificity of the ligation, each fragment was Mock- ligated, where no linker was added to the reaction. To selectively amplify fragments cut by PvuRts1I, the ligated products were amplified by PCR with Phusion HF DNA Polymerase using a Linker- specific Forward Primer (For 5´CTC GTA GAC TGC GTA CCA TG 3´) and gene- specific Revers Primers (PP1 Rev: 5´ GAG TCA GAC CTT GCT GCC AAA 3´, and PP2 Rev 5´ GCC GTC TAA GCA ATG GAA GAA 3´), resulting in different fragment sizes depending on the cutting site. The PCR pattern was analyzed on a 2 % agarose gel and a library of the 10 % 5hmC fragments was generated using the Zero Blunt® PCR Cloning Kit. Randomly selected clones were sequenced and analyzed for potential cutting sites.

2.2.2.13 DNA methylation analysis

Bisulfite conversion of genomic DNA (0.5 - 1.5 µg) was carried out using the EZ DNA Methylation-Gold Kit according to the manufacturer´s instructions. For amplification we used Qiagen Hot Start Polymerase in 1x Qiagen Hot Start Polymerase buffer supplemented with 0.2 mM dNTPs, 0.2 µM forward primer, 0.2 µM reverse primer, 1.3 mM betaine and 60 mM tetramethylammonium-chloride. Primer sequences (listed in 2.1.5.4 and 2.1.5.5) were designed using MethPrimer (Li and Dahiya, 2002) and CpG islands were identified with the CpG Island Searcher using default settings (Takai and Jones, 2002). DNA methylation analysis by pyrosequencing was carried out by Varionostic GmbH (Ulm, Germany). For bisulfite sequencing, PCR products were subcloned using the StrataClone™ PCR Cloning Kit and StrataCone™ SoloPack® Competent Cells according to the manufacturer´s instructions. Sequences were analyzed with BISMA software using default settings (Rohde et al., 2010).

2.2.2.14 Cloning of oct4 dTALEs

Cloning of oct4 dTALEs was performed by our collaboration partner Robert Morbitzer (group of Thomas Lahaye, LMU Munich). A Gateway cassette from pGWB5 (Nakagawa et al., 2007b) was amplified (Forward primer 5´GGG GCG ATC GCA CAA GTT TGT ACA AAA AAG CTG AAC GAG 3´; reverse primer 5´ GGG GCG GCC GCA ACC ACT TTG TAC AAG AAA GCT GAA CG 3´), thereby adding AsiSI and NotI restriction sites and subcloned into pCAG_mCH (Niwa et al., 1991) using AsiSI and NotI, generating pCAG_mCH_GW. The VP16AD was amplified from RSV E2F1-VP16 using the following primers: Forward primer 5´ GGG GGT CTC TCA CCA TGG ATC CTG CCC CCC CGA CCG ATG TCA GC 3, Reverse primer 5´ GGG GGT CTC CCT TCT ACC CAC CGT ACT CGT CAA TTC CAA GG 3´ (Johnson et al., 1994), thereby adding a BamH1 restriction site to the 5' end and

cloned into pENTR-D-TOPO generating pENTR-D-BamHI_VP16AD. TALE repeat arrays were generated via multi-fragment cut-ligation using golden gate cloning (Engler et al., 2009) and ligated either into pENTR-D-TALE-∆rep-BpiI-AC or pENTR-D-TALE-∆rep-BpiI-AC-VP16AD. All entry clones were transferred by LR recombination into the expression vector pCAG_mCh_GW.

2.2.2.15 Cloning of oct4 reporter contruct

Cloning of the oct4 reporter construct was performed by Sebastian Bultmann. The vector GOF-18 (Yeom et al., 1996) was cut by XhoI/ AvrII, resulting in the basepairs -1 to -4716 upstream of the transcriptional start site of *oct4*, which was ligated into pGL-3 basic together with a linker oligo (5´ CCT AGG TGA GCC GTC TTT CCA CCA GGC CCC CGG CTC GGG GTG CGA TCG CCG CCC ATG G 3´) using XhoI/ NcoI. Subsequently, the luciferase ORF was removed by with KasI/ FseI and replaced by the eGFP ORF (Forward primer 5´ AAA GGC GCC AGT GAG CAA GGG CG 3´, reverse primer 5´. AAA GGC CGG CCT TAC TTG TAC AGC TCG TCC 3´).

2.2.3 Methods of Biochemistry

2.2.3.1 Protein expression and purification

Cloning of both sequences was performed by Sebastian Bultmann and Alexandra Szwagierczak. The protein expression and purification was done by Alexandra Szwagierczak. The sequence encoding bacteriophage T4 β- gt and PvuRts1I was synthesized at Mr. Gene GmbH (Regensburg) and cloned into pET28b vector. BL21(DE3) *E. coli* cells carrying the expression construct were grown at 37 °C until A_{600} = 0.6–0.7 and induced with 1 mM isopropyl β-d-thiogalactopyranoside for 16 h at 20 °C and 18 °C, respectively. Lysates from ß- gt and PvuRts1I expressing bacteria were prepared by sonication in 300 mM NaCl, 50 mM Na_2HPO_4, pH 8.0, 10 mM imidazole, 1 mM β-mercaptoethanol and for PvuRts1I, additionally 10 % glycerol was added. After clearing lysates by centrifugation, they were applied to nickel-nitrilotriacetic acid column pre-equilibrated with lysis buffer. Washing and elution were performed with lysis buffer containing 20 and 250 mM imidazole, respectively. Eluted proteins were applied to a Superdex S-200 preparative gel filtration column in 150 mM NaCl, 20 mM Tris, pH 8.0, 1 mM DTT and in the case of PvuRts1I, 10 % glyercol was again added. Fractions containing the β- gt or PvuRts1I peak were pooled and applied to a ResourceQ anion exchange column in order to eliminate residual contaminants, resulting in pure β- gt and PvuRts1I in the flowthrough. The stability of PvuRts1I upon storage was improved by supplementation with 10 % glycerol.

2.2.3.2 Reporter gene assay for dTALE activity

The transient reporter gene assay was conducted by Sebastian Bultmann. HEK293T cells were plated in 6 well plates, grown to 70

% confluence and co-transfected with the reporter plasmid and the respective dTALE construct. 48 h after transfection cells were lysed in PBS containing 0,5 % NP40 and mammalian protease inhibitors. The lysate was cleared by centrifugation and eGFP and mCherry fluorescence was measured with a Tecan Infinite M1000 plate reader.

3. Results

3.1 Reversion of differentiation programs in globally hypomethylated embryonic stem cells

To elucidate the role of DNA methylation during ESC differentiation, we chose a system, the differentiation of wt, *dnmt1*$^{-/-}$ and TKO cells to Embryoid Bodies (EBs), which represents an undirected way of differentiating cells, allowing for a less biased analysis of the differentiation process. EB formation was performed using the hanging drop method, where a definite number of cells per drop (600 cells/ 20 µl) are spotted on the lid of a petri dish (Figure 11). Inversion of the lid causes the cells to accumulate at the bottom of the drop resulting in the formation of EBs. After 4 days, the EBs are harvested from the hanging drops and cultured in suspension to promote further differentiation and allow exchange of medium.

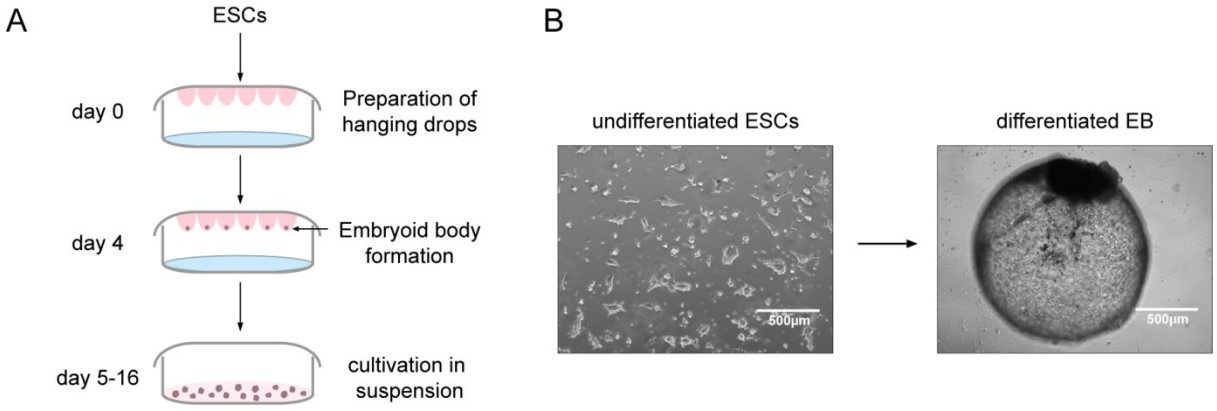


Figure 11. Overview of hanging drop method to induce EB formation.
(A) Schematic representation of EB formation by the hanging drop method. (B) Phase contrast pictures of undifferentiated ESCs (left) grown on gelatinized dishes in standard ESC medium and a differentiated EB grown in suspension in EB medium at day 12.

3.1.1 Incomplete silencing of pluripotency genes during differentiation in globally hypomethylated cells

To gain insights into the specific role of DNA methylation in silencing pluripotency associated genes during differentiation, we generated EBs from wt, *dnmt1*$^{-/-}$ and TKO ESCs and analyzed expression as well as promoter DNA methylation of the pluripotency genes *oct4* and *nanog* (Fig. 12). Interestingly, wt and *dnmt1*$^{-/-}$ EBs already showed a drastic down regulation of both transcripts at day 4, whereas DNA methylation of the respective promoters could be detected only by day 8. Consistent with previously reported data (Sakaue et al., 2010), TKO EBs exhibited similar initial silencing kinetics for nanog transcription, however, a drastic delay in the down regulation of *oct4* mRNA levels was observed. Strikingly, in both Dnmt knock out EBs, silencing of *oct4* and *nanog* transcripts was partial as compared to wt EBs and significant residual levels of both transcripts were detectable even in late EBs. It is interesting to note that although both knock out EBs are severely hypomethylated, *dnmt1*$^{-/-}$

EBs seem less affected by the loss of DNA methylation compared to TKO EBs as both pluripotency genes are more efficiently silenced.

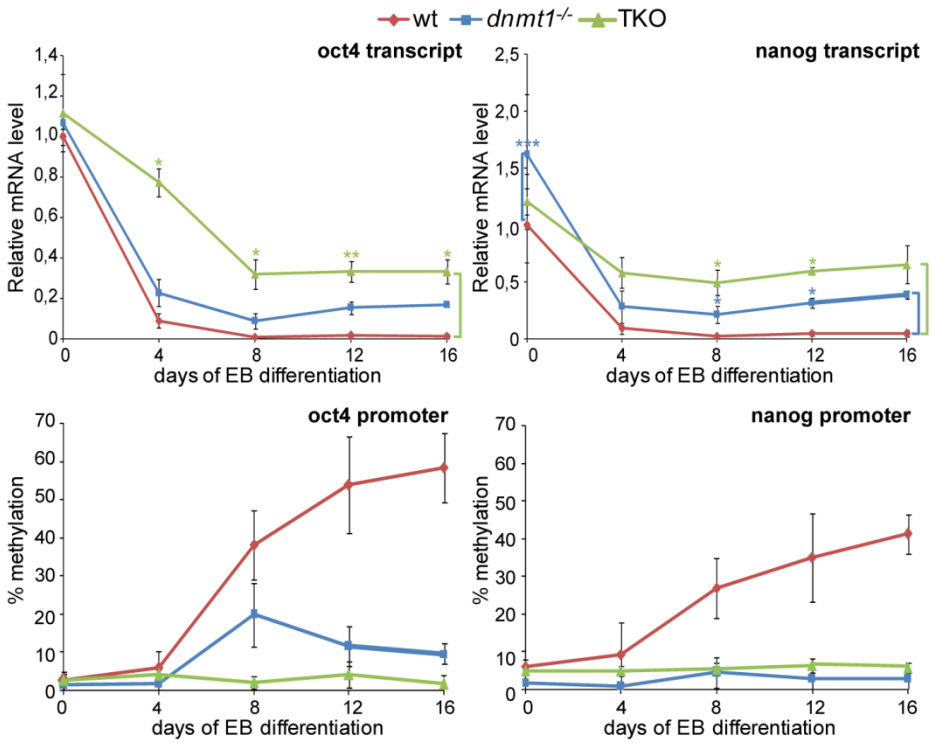


Figure 12. Pluripotency genes are not completely silenced in hypomethylated cells during EB formation. *oct4* and *nanog* transcript levels (upper panels) und respective promoter DNA methylation (lower panels) during differentiation of wt, *dnmt1*$^{-/-}$ and TKO ESCs to EBs. Expression levels were determined by qPCR and are displayed relative to day 0 in wt ESCs and standard errors are the mean of 3 biological replicates. DNA methylation analysis was performed using bisulfite conversion followed by PCR and pyrosequencing. Five and four CpG sites within the proximal promoter regions of *oct4* and *nanog*, respectively were measured and averaged values of two biological replicates and the standard deviation are displayed. At every time point, statistical analysis was performed between wt and *dnmt1*$^{-/-}$ as well as between wt and TKO ESCs and EBs, respectively. Significant differences between wt and knock out ESCs and EBs are marked by an asterisk; significance level: * $p<0.05$; ** $p<0.001$; *** $p< 0.0001$ (Student t- test). Pyrosequencing data for *dnmt1*$^{-/-}$ EBs duplicates as well as data from one wtJ1 EB unicate was provided by Daniela Meilinger.

Taken together, our experiments reveal that promoter DNA methylation is dispensable for the initial silencing of *oct4* and *nanog* but necessary for their complete shutdown during differentiation.

3.1.2 Complete and uniform downregulation of Oct4 protein level in hypomethylated EBs

We next asked whether the partial down regulation of pluripotency transcripts also results in an incomplete reduction of pluripotency markers at the protein level and whether the incomplete silencing occurs homogenously in all cells or it is due to the appearance of heterogeneous subpopulations during differentiation. To address these questions, we studied Oct4 in greater detail and followed its protein level during EB formation by intracellular FACS staining using commercial antibodies (Fig. 13). The FACS machine was operated with the help of Sebastian Bultmann.

As expected, all cell lines displayed similar Oct4 level in the pluripotent state. After 4 days of differentiation, less than 2 % cells of wt EBs are Oct4 positive, but a significantly higher proportion of positive cells could be detected in both knock out EBs. As observed at the transcript level (Fig. 12) a dramatic delay in the down regulation of Oct4 protein was also observed in TKO EBs as more than 13 % of cells were still Oct4 positive at day 4. Furthermore, we again found that *dnmt1*$^{-/-}$ EBs showed an intermediate phenotype as only 4.75 % of all cells still contained detectable Oct4 levels. Surprisingly, at later time points, less than 1 % of all cells were Oct4 positive, independent of the methylation status. As only a single Oct4 positive population was detectable at all analyzed time points, we conclude that the down regulation of Oct4 during differentiation occurs homogenously not only in wt but also in hypomethylated cells.

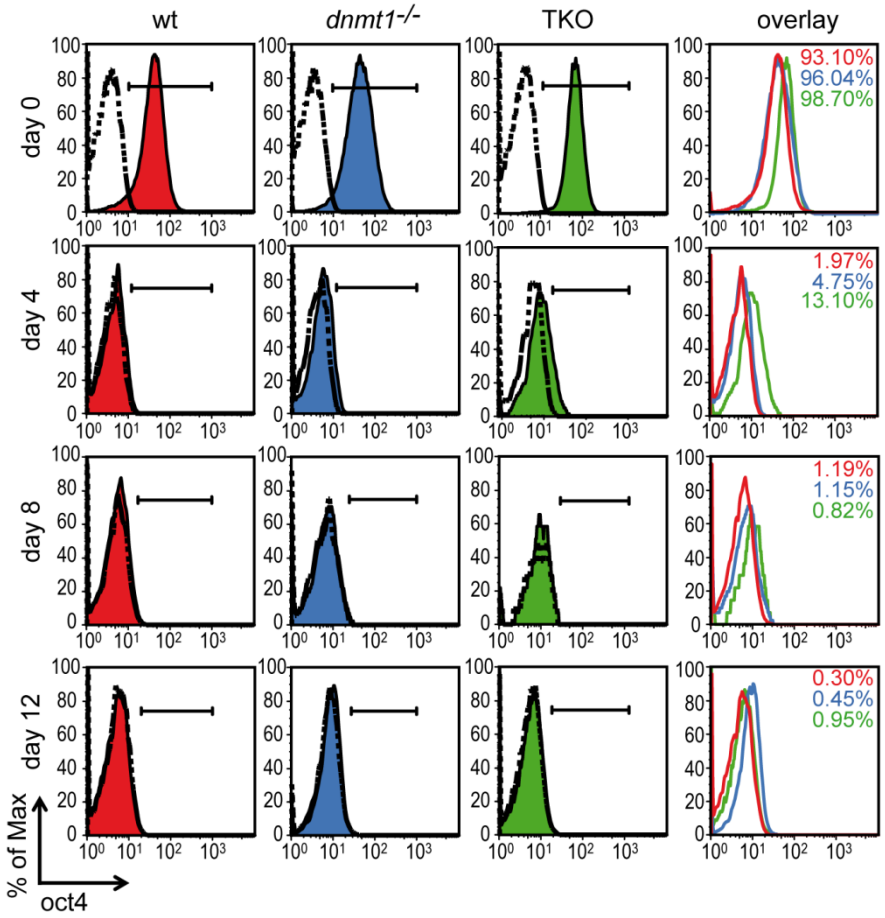


Figure 13. Homogenous downregulation of Oct4 protein levels during differentiation.
Shaded histograms depict Oct4 intracellular staining by FACS analysis. Dashed lines indicate isotype control staining and horizontal bars represent individual gates set for Oct4 positive cells based on the isotype control. The overlay shows Oct4 expression for each cell line at each day. Numbers in red, blue and green indicate the percentages of Oct4 positive cells in wt, *dnmt1*$^{-/-}$ and TKO ESCs/EBs, respectively, after subtraction of background signal from isotype controls.

3.1.3 DNA methylation is dispensable for the initiation of differentiation programs

As we found that hypomethylated cells are capable of partially down regulating pluripotency associated genes, we next asked whether these cells are able to initiate differentiation

programs. To address this question, we performed genome- wide expression analysis of wt, *dnmt1*$^{-/-}$ and TKO cells at three different time points: i) in the undifferentiated state (day 0), ii) during early differentiation (day 4) and iii) during a late differentiation (day 16) stage in duplicate experiments. The Microarray hybridizations were performed by our collaborators Kerstin Maier and Dietmar Martin (Laboratory of Patrick Cramer) and the initial processing of the microarray data was done by our collaborators Benedikt Zacher and Achim Tresch (Laboratory of Achim Tresch). They performed the Principal Component Analysis (Fig. 14), identified differentially regulated genes, compared differentially expressed genes in undifferentiated TKO ESCs identified in our study to published data (Fig. 16B) and did the analysis of bivalent genes and all gene changes between d0-4 (Fig. 21, 23). I performed all other analyses, including GO, Kegg pathway, chromosomal location and tissue expression analysis using DAVID software (Huang et al., 2008, 2009; for details see chapter 2.2.2.4). Differentially regulated genes identified at the various time points are listed in chapter 6.

The Principal Component Analysis (PCA) revealed that the microarray data can be separated according to day of differentiation and genotype (Fig. 14).

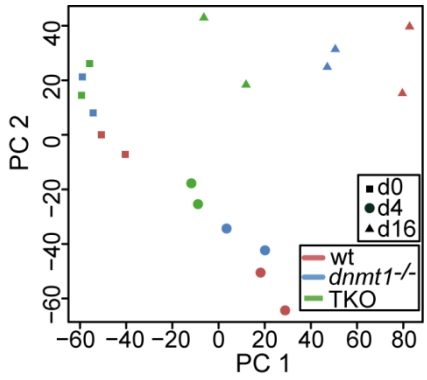


Figure 14. Principal Component Analysis of genome- wide expression data from wt, *dnmt1*$^{-/-}$ and TKO ESCs during differentiation

RNA samples from undifferentiated ESCs and respective EBs at day 4 and 16 of differentiation were subjected to microarray expression analysis. Data from independent biological duplicates per cell line and per time point were processed by two dimensional principal component analysis. Two principal components (PC), PC1 and PC2, were identified (see also text for details).

In the undifferentiated state, all cell lines were closely clustered, suggesting that loss of DNA methylation in the undifferentiated state affects the expression of relatively few genes. However, principal component 1 provides a clear differentiation signature, as it increases continuously with culture time. This increase was less pronounced among the knockout genotypes, reflecting a crucial role of DNA methylation during differentiation. Intriguingly, at both differentiation time points, *dnmt1*$^{-/-}$ EBs took an intermediate position on the median line between TKO and wt EBs, again emphasizing the less severe phenotype of *dnmt1*$^{-/-}$ EBs compared to TKO EBs.

Using a two- fold cutoff and false discovery rate below 5 %, we calculated gene expression changes in undifferentiated *dnmt1*$^{-/-}$ and TKO ESCs relative to wt ESC lines (Fig. 15A). Furthermore, we determined gene expression changes for each genotype between the undifferentiated ESC state and day 4 EBs (Fig. 15B), as well as between day 4 and day 16 EBs (Fig. 15C). As suggested by PCA, the expression of only few genes was altered in undifferentiated hypomethylated ESCs (54 genes in *dnmt1*$^{-/-}$ ESCs, 82 genes in TKO ESCs). As DNA methylation is mostly considered a repressive mark involved in silencing of genes (Siegfried et al., 1999), we reasoned that genes being down regulated in the knock outs are most likely the results of indirect effects and not due to the lack of DNA methylation. Therefore we focused most of the analysis on the set of up regulated genes in the knock out ESCs and EBs.

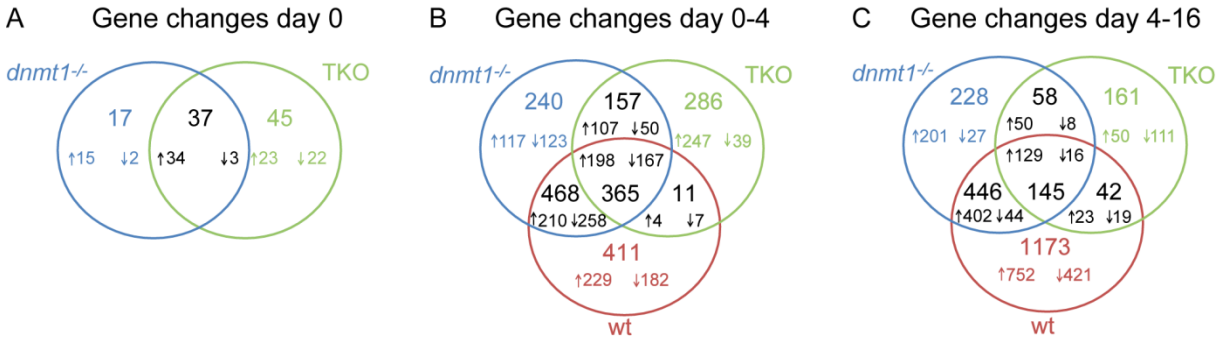


Figure 15. Analysis of genome- wide expression data from wt, *dnmt1*$^{-/-}$ and TKO ESCs during differentiation.

A) Venn diagram of differentially expressed genes in *dnmt1*$^{-/-}$ (blue circles) and TKO (green circles) ESCs compared to wt ESCs. (B, C) Venn diagram of gene expression changes occurring between day 0-4 (C) and day 4-16 (C) in knockout and wt (red circles) EBs. Expression changes occurring in wt EBs between day 0-4 and day 4-16 were identified and compared to expression patterns in *dnmt1*$^{-/-}$ and/ or TKO EBs. Larger numbers indicate total numbers of gene changes, whereas smaller numbers refer to up- (↑) and down- (↓) regulated transcripts in the respective sectors. Data were averaged from independent biological duplicates.

Gene ontology (GO) enrichment analysis of the 57 genes up regulated in TKO ESCs revealed that the most enriched categories are involved in reproductive processes like gamete generation, oogenesis and spermatogenesis (Fig. 16A).

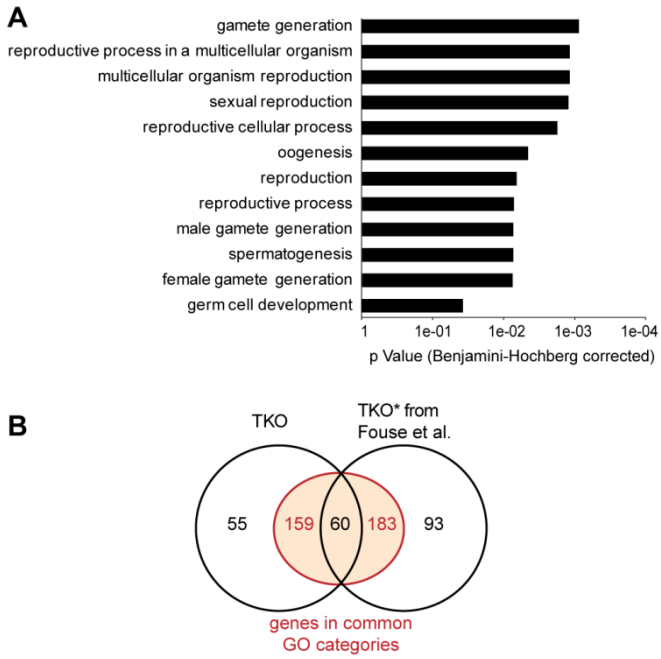


Figure 16. Gene ontology enrichment of transcriptome analysis of TKO ESCs compared to wt ESCs
(A) 57 genes were up regulated in TKO ESCs compared to wt ESCs (Fig. 15) and can be grouped in 12 GO categories. (B) Venn diagram showing the comparison of differentially expressed genes in TKO ESCs identified in the current study and deregulated genes identified by Fouse *et al*. (Fouse et al., 2008) in compound *dnmt3a/3b*$^{-/-}$ ESCs with stable Dnmt1 knockdown (TKO*; black circles). To directly compare differentially expressed genes in the two studies, the same cut off criteria based solely on a fold change >2 were used. These less stringent cut off criteria lead to a significant higher number of deregulated genes in our TKO ESCs (274 genes compared to 82 with additional cut off criteria of $p<0.05$). However, only 60 commonly differentially regulated genes between our TKO ESCs and the 336 genes identified by Fouse *et al*. can be identified. Those genes can also be grouped within the same GO categories comparable to the ones in (A). A comparison of the GO terms associated with the individual genes in the lists from the two studies shows a significant functional overlap (red numbers in red shaded circle). 159 genes identified from our study are associated with the same GO terms as genes from Fouse *et al*. and vice verse, 183 genes from their data are linked to the same GO terms as genes from our data set.

In contrast to the data obtained from TKO ESCs, no enriched categories could be identified for the 49 up regulated genes in *dnmt1*$^{-/-}$ ESCs. However, in both mutant cell lines, deregulated genes were mainly located on the X chromosome (*dnmt1*$^{-/-}$ 13.1 %; TKO 40.4 %). In this context, it is important to note that both wt and knock out ESCs are derived from male (XY) mice and hence lack X chromosome inactivation. Therefore, the relatively high amount of X- linked, upregulated genes in both hypomethylated ESCs is not linked to a deregulated X inactivation process. Furthermore, in *dnmt1*$^{-/-}$ ESCs many deregulated genes are also located on the Y chromosome (13.3 %). The up regulated transcripts of both cell lines were mostly genes known to exhibit testis-specific expression (*dnmt1*$^{-/-}$ 57.7 %; TKO 61.5 %). These data are consistent with previous results from Fouse and colleagues (Fouse et al., 2008), who performed genome-wide expression analysis on a *dnmt3* double knock ESC line (*dnmt3a*$^{-/-}$, *dnmt3b*$^{-/-}$) with constitutive knockdown of *dnmt1* (TKO*). Surprisingly though, using the same cut off criteria (fold change >2) as in Fouse *et al*., only 60 out of 274 differentially expressed genes identified in our study for TKO were also found to be deregulated in the previous study (Fig. 16B). This may at least in part be due to the less

stringent cut off criteria used in Fouse et al. However, when we compare the GO terms associated with the differentially regulated genes, we found a significant functional overlap between these two studies.

To investigate the role of DNA methylation during differentiation, we first determined the gene changes occurring in wt EBs between day 0-4 and day 4-16 and used this as a baseline of expression changes normally occurring during ESC differentiation. We then compared these changes to the gene changes occurring in the knock outs at the respective days. This strategy allowed us to identify concordant gene changes of wt and knock out EBs which are regulated independently of DNA methylation (Fig. 15B, C). In the first differentiation period (d0-4), approximately the same number of genes in wt (1255) and *dnmt1*$^{-/-}$ EBs (1230) were altered and two thirds (833) of the genes concordantly changed in the two EB populations. In comparison, 819 genes changed their transcript level in TKO EBs, which translates to about two thirds of gene changes compared to those in wt EBs. However, only one third of the genes (376) changed concertedly in wt and TKO EBs (Fig. 15B). Strikingly, in the second differentiation period (d4-16) total transcript level changes in *dnmt1*$^{-/-}$ (879) and TKO EBs (406) were roughly half and one fifth of those in wt EBs (1808), respectively (Fig. 16C). Furthermore, one third (593) of the genes changing in *dnmt1*$^{-/-}$ but only one tenth (187) of transcripts altered in TKO are concordantly regulated compared to wt EBs. The observation of a high proportion of concerted gene expression changes in mutant and wt EBs during the early differentiation stage, suggests that hypomethylated cells are able to induce differentiation programs to some extent. However, the decreased amount of concordant transcription level changes between day 4-16 implies that the progression of differentiation programs in hypomethylated EBs is impaired, in particular in TKO EBs, indicating that the presence of Dnmt3 proteins could enable hypomethylated *dnmt1*$^{-/-}$ EBs the execution of differentiation programs to similar extents as in wt EBs.

Intriguingly, we detected 365 concerted gene changes in EBs of all three genotypes during the first four days of differentiation (Fig. 15B). GO analysis of the 198 conjointly up regulated genes yielded categories involved in developmental processes like anatomical structure development, system and organ development and cell differentiation being the most enriched (Fig. 17A).

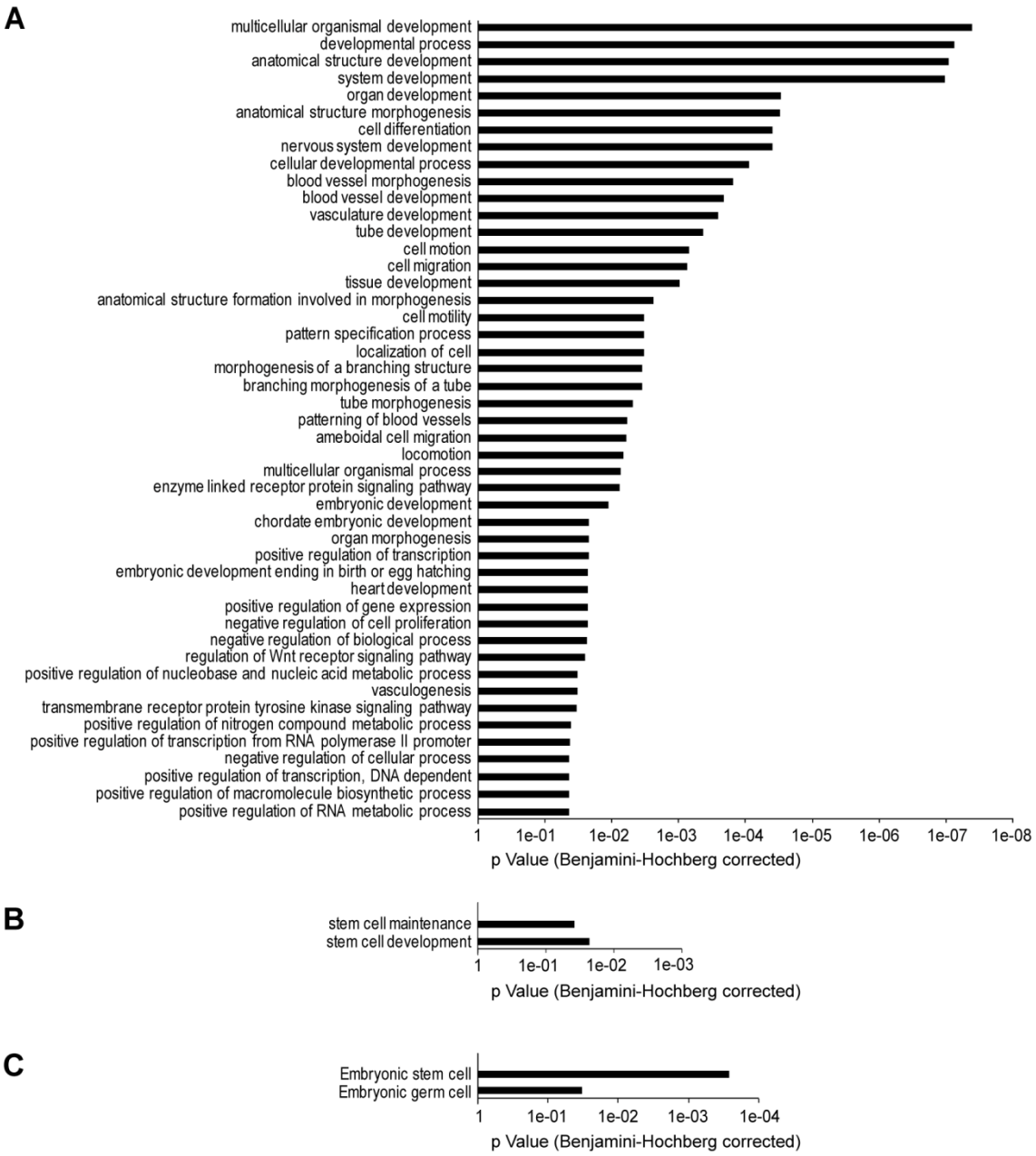


Figure 17. Gene ontology enrichment and cell type specific expression of concomitantly regulated genes in wt, *dnmt1*$^{-/-}$ and TKO EBs during day 0-4 of differentiation

From the 365 concordant genes changes (Fig. 15), 198 genes were upregulated and are associated with 47 GO terms (A) involved in various developmental processes. On the contrary, 167 genes were concordantly down regulated and are connected to stem cell maintenance and development GO categories (B). The latter set of genes is specifically expressed in ESCs and EGCs specifically (C).

Furthermore, genes playing roles in transcription, metabolism and signaling pathways like the wnt receptor and tyrosine kinase pathways were identified by Kegg pathway analysis (not shown). By contrast, the expression of 167 genes was down regulated in all three cell lines and they could be categorized into genes involved in stem cell maintenance and development and are known to be mainly expressed in ESCs and embryonic germ cells (EGCs) (Fig. 17B, C).

These data clearly demonstrate that all cell lines independent of their genotype were able to activate differentiation programs by down regulating genes associated with stem cell fate and up regulating genes required for developmental processes.

3.1.4 Cells lacking Dnmt1 possess a greater differentiation potential than TKO cells

During the later differentiation time point (day 4-16), only 145 commonly expressed genes in wt, *dnmt1*$^{-/-}$ and TKO EBs could be identified using our strategy described above (Fig. 15C). Enriched GO categories for the 129 concerted up regulated genes in all three cell lines were involved in lipid metabolic processes and were mainly expressed in extraembryonic tissues (Fig. 18).

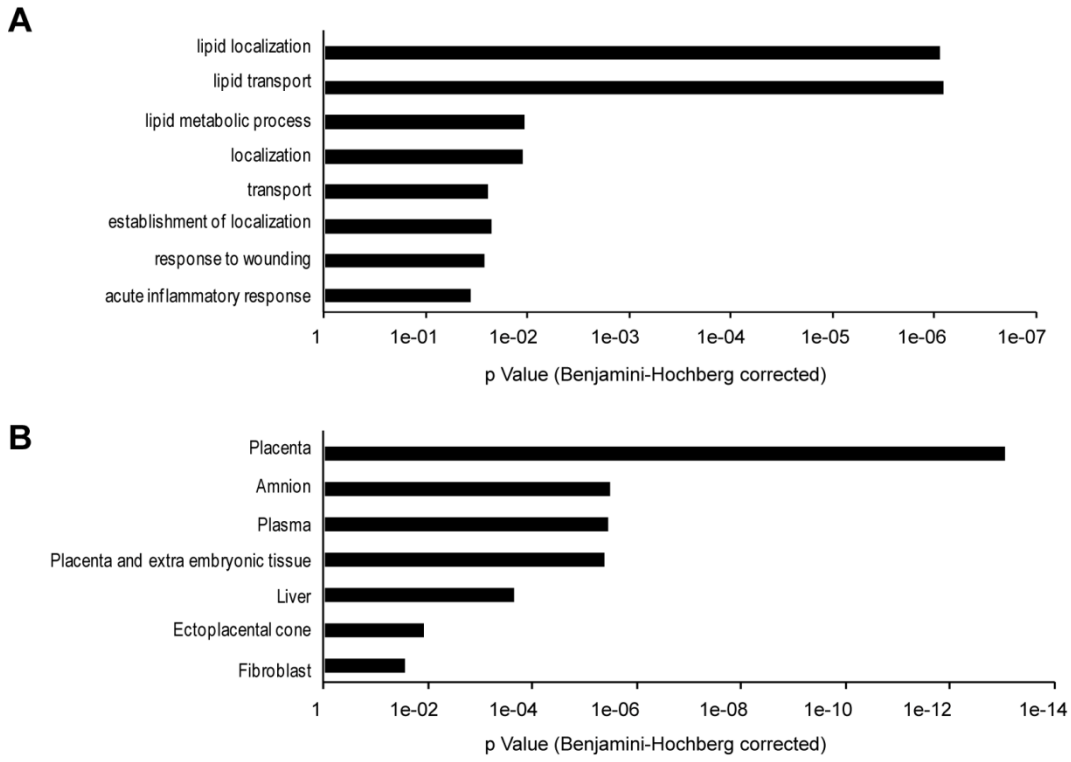


Fig. 18. Gene ontology enrichment and cell type specific expression of concomitantly regulated genes in wt, *dnmt1*$^{-/-}$ and TKO EBs during day 4-16 of differentiation

From the 145 concordant genes changes (Fig. 15C), almost 90 % of these genes were upregulated and were associated with 8 GO terms (A) mainly involved in lipid metabolism. The up regulated genes are mostly expressed in extraembryonic tissues like placenta and amnion (B). Only 16 genes were concordantly down regulated.

Intriguingly, substantially more genes were concertedly regulated in wt and *dnmt1*$^{-/-}$ at both time points of differentiation. In addition to the 365 genes concordantly altered in all three cell lines, 468 more genes were conjointly regulated in wt and *dnmt1*$^{-/-}$ EBs at d0-4 (Fig. 15B). Highly enriched GO terms of the 210 up regulated genes were associated with developmental processes including cell differentiation and proliferation as well as tissue and organ development (Fig. 19). Kegg pathway analysis indicate that these genes are, among others, associated with the wnt and TGF-ß signaling pathway (not shown). Furthermore, the

258 genes conjointly down regulated in wt and *dnmt1*$^{-/-}$ EBs yielded 35 enriched GO categories involved in the regulation of expression, metabolism and developmental processes (not shown).

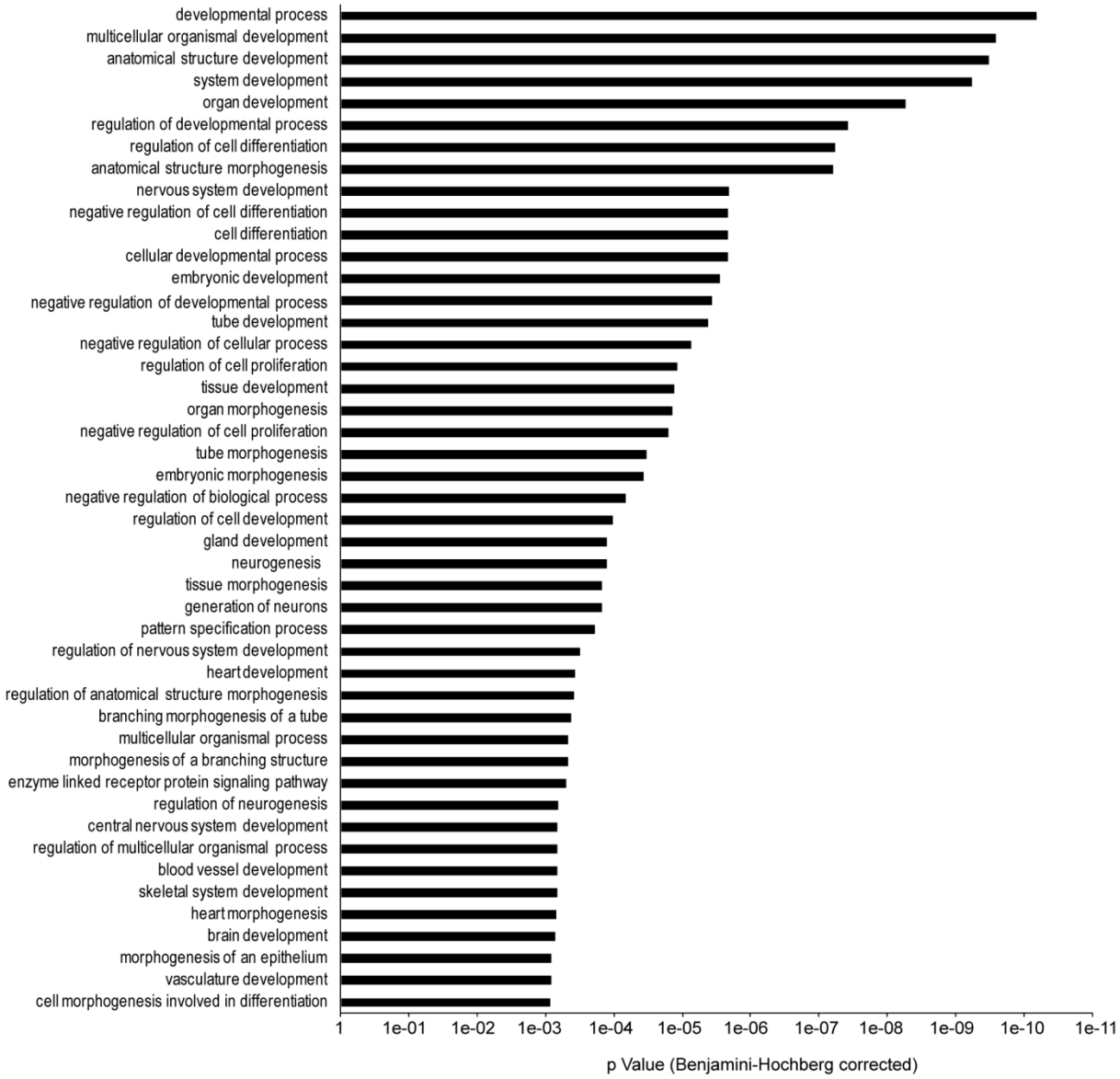


Figure 19. Enriched GO categories of conjointly up regulated genes in wt and *dnmt1*$^{-/-}$ EBs after 4 days of differentiation

A total of 468 genes could be identified as commonly regulated genes in wt and *dnmt1*$^{-/-}$ EBs after 4 days of differentiation (Fig. 16C). Among these genes, 210 are up- and 258 are downregulated. Shown here are the 46 enriched GO categories for the up regulated genes with a p-value < $1x10^{-3}$ (1e-03). Using a cut off p-value < 0.05 corrected by Benjamini-Hochberg increases the number of enriched categories to 144. In addition, these 210 genes, among others, play a role in the wnt and TGF-ß signaling pathways according to Kegg pathway analysis. For the concordantly down regulated genes, using a p-value < 0.05 we also found 35 enriched GO categories which include genes involved in the regulation of expression, metabolism and developmental processes. However, the p-values for these categories are above $1x10^{-3}$ and therefore are not displayed here for consistency with the upregulated genes.

Importantly, also at a later differentiation time point (day 4-16), 446 genes were commonly regulated in wt and *dnmt1*$^{-/-}$ EBs, and more than 90 % of those genes (402) were up regulated (Fig. 20).

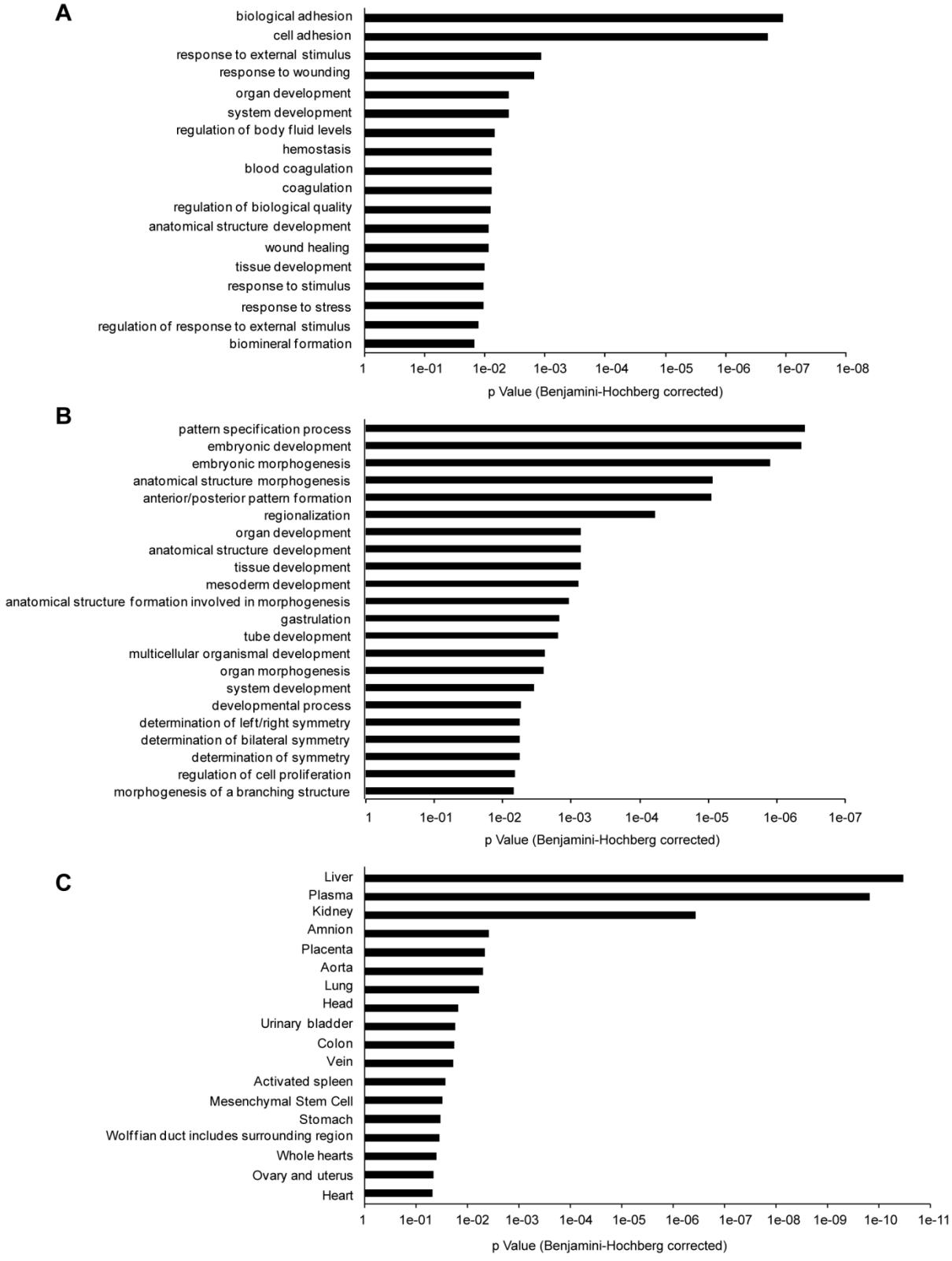


Figure 20. Enriched GO categories of concordantly up regulated genes in wt and *dnmt1*$^{-/-}$ EBs during day 4-16 of differentiation.

Later during differentiation, 446 genes were still commonly regulated in wt and *dnmt1*$^{-/-}$ EBs. More than 90 % of those genes (402) were upregulated and only 44 were downregulated. The most enriched GO terms with a p-value below $1x10^{-2}$ (1e-02) of the up regulated (18 GO terms) and down regulated (17 GO terms) genes are displayed in A and B, respectively. Using a cut off p-value < 0.05 corrected by Benjamini-Hochberg increases the number of enriched categories to 20 for up regulated genes and 38 for down regulated genes. The up regulated genes are expressed in various tissues (C), the highest enriched in liver, plasma and kidney (p-value <.1e-06).

Enriched GO terms for the up regulated genes were mainly associated to cell adhesion, hemostasis, organ development and metabolism (Fig. 20A) and are annotated as being expressed in various tissues like liver, plasma and kidney (Fig. 20C). In contrast, the few genes (44) concertedly down regulated in wt and *dnmt1*$^{-/-}$ late EBs showed enriched GO categories of pattern specification and embryonic development (Fig. 20B). Taken together, these data reveal a previously unappreciated progression and execution of developmental programs in *dnmt1*$^{-/-}$ EBs and indicate that *dnmt1*$^{-/-}$ cells possess a greater differentiation potential compared to TKO cells.

To analyze the different developmental potentials of *dnmt1*$^{-/-}$ and TKO EBs in greater detail we focused on gene changes occurring at the first differentiation period (day 0-4) as TKO EBs seemed unable to progress to later differentiation stages. We compared the total expression changes in TKO or *dnmt1*$^{-/-}$ EBs relative to wt EBs for all genes (25528) (Fig. 21).

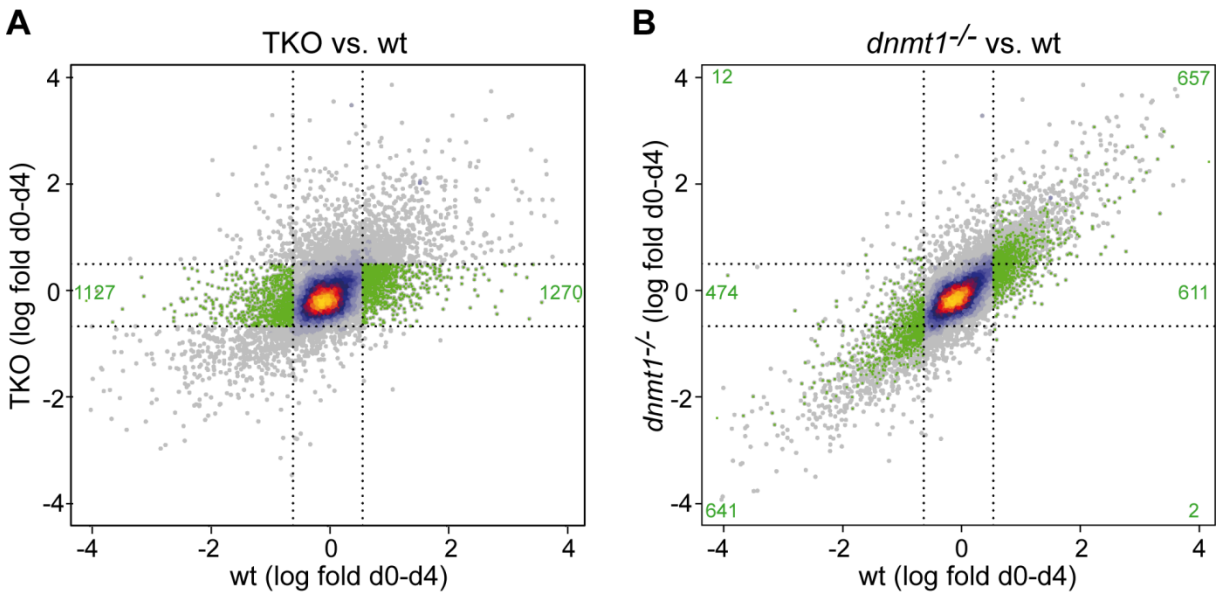


Figure 21. Genome- wide expression changes of TKO and *dnmt1*$^{-/-}$ EBs compared to wt EBs at day 0-4 of differentiation

Scatter plots of genome- wide expression changes between day 0 and 4 in TKO (A) and *dnmt1*$^{-/-}$ EBs (B) relative to wt EBs. Data points are heat- colored according to their local point density with yellow representing high gene density and blue low gene density. From a total of 25528 genes, around 17 % of all genes are changing in wt during differentiation as indicated by the vertical corridor between the dotted lines. Genes whose expression levels increased and decreased in wt but not in TKO EBs ("TKO non- responders"), are marked in green in both plots. These genes are located within the horizontal dotted lines to the left and right of the vertical dotted lines in (A). Many of the "TKO non- responding" (green marked) genes change similarly in *dnmt1*$^{-/-}$ and wt EBs (B). The number of "TKO non– responders" are reported in green in the respective sectors.

In this analysis we used a two- fold cutoff as well, but did not filter for false discovery rate, resulting in higher numbers of gene expression changes than in the case of Fig. 15B. Interestingly, data points clustered much closer to the diagonal line in the plot showing expression changes in *dnmt1*$^{-/-}$ relative to wt EBs (Fig. 21B), confirming the higher concordance of transcript level changes between these EB populations. In TKO EBs, 2397

genes did not respond to the differentiation signals (“non-responders”; green dots in Fig. 21A). To investigate, whether those “non-responders” were also deregulated in *dnmt1*$^{-/-}$ EBs, we compared the expression profile of *dnmt1*$^{-/-}$ versus wt EBs and again highlighted the non-responding genes identified in TKO EBs (green dots in Fig. 21B). Intriguingly, more than half of these genes (1298 out of 2397) did change similarly in *dnmt1*$^{-/-}$ and wt EBs and that in total 70 % of gene expression changes were concordant between et EBs and *dnmt1*$^{-/-}$ EBs.

The further progression of differentiation programs in *dnmt1*$^{-/-}$ compared to TKO EBs suggests that the expression of the *de novo* Dnmts, *dnmt3a* and *dnmt3b* confers *dnmt1*$^{-/-}$ cells with greater developmental potential. Therefore, we analyzed *dnmt3a* and *dnmt3b* transcript levels in wt and *dnmt1*$^{-/-}$ EBs after 4 and 16 days of differentiation (Fig. 22).

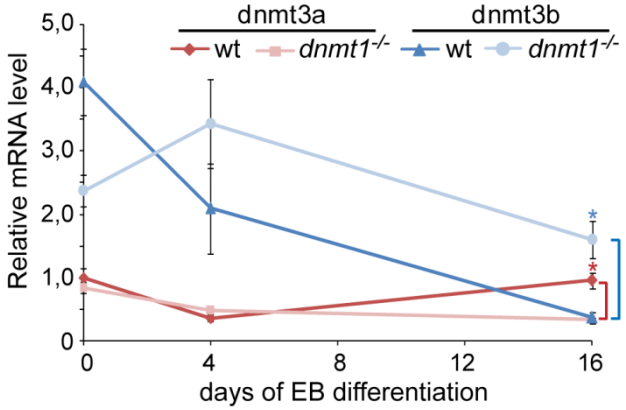


Figure 22: Expression of *dnmt3a* and *dnmt3b* in wt and *dnmt1*$^{-/-}$ cell lines during EB formation.
Expression analysis was performed by qPCR and is displayed relative to *dnmt3a* in wt ESCs at day 0, so that transcript levels of *dnmt3a* and *dnmt3b* are directly comparable to each other. Mean values and standard errors from three independent experiments are shown. At each analyzed time point, statistical analysis was performed between wt and *dnmt1*$^{-/-}$ ESCs and EBs and significant differences are marked by an asterisk. Significance level $p<0.05$ as determined by Student t-test is displayed.

In wt EBs, *dnmt3b* was highly expressed in the undifferentiated state and during early differentiation stages (day 4), but was drastically down regulated during later differentiation stages. In contrast, *dnmt3a* was highly expressed in late EBs, indicating a time- and developmental stage- dependent switch between *dnmt3a* and *dnmt3b* expression. This observation confirms previously published immunostaining data in ESCs, reporting lower protein levels of Dnmt3a compared to Dnmt3b as well as the switch of *de novo dnmt* expression during differentiation (Watanabe et al., 2002). Surprisingly, *dnmt1*$^{-/-}$ EBs showed a completely opposite expression pattern of both *de novo* Dnmts. Whereas *dnmt3a* mRNA levels progressively decreased during EB formation, *dnmt3b* is highly up regulated at all time points during differentiation. The high *dnmt3b* transcript levels might contribute to the further progression of differentiation programs in *dnmt1*$^{-/-}$ versus TKO EBs.

3.1.5 Most bivalent genes are silenced independently of *de novo* methylation or Dnmt proteins during early EB differentiation

In mouse ESCs, around 2978 genes carry bivalent (H3K4me3/H3K27me3) chromatin domains and it has been proposed that this bivalency keeps developmental genes in a silent state but poises them for rapid activation during differentiation (Bernstein et al., 2006; Ku et al., 2008). Interestingly, Mohn *et al*. showed that 93 % of all bivalent promoters contain CpG islands and suggest a switch from the bivalent state to DNA methylation during differentiation of ESCs to neural progenitors, indicating a lineage- specific crosstalk between the two repressive modifications. To test whether a similar transition occurs more generally in undirected EB differentiation conditions, we repeated the same analysis of gene expression changes as above, but analyzed bivalent and non- bivalent gene sets separately (Figure 23).

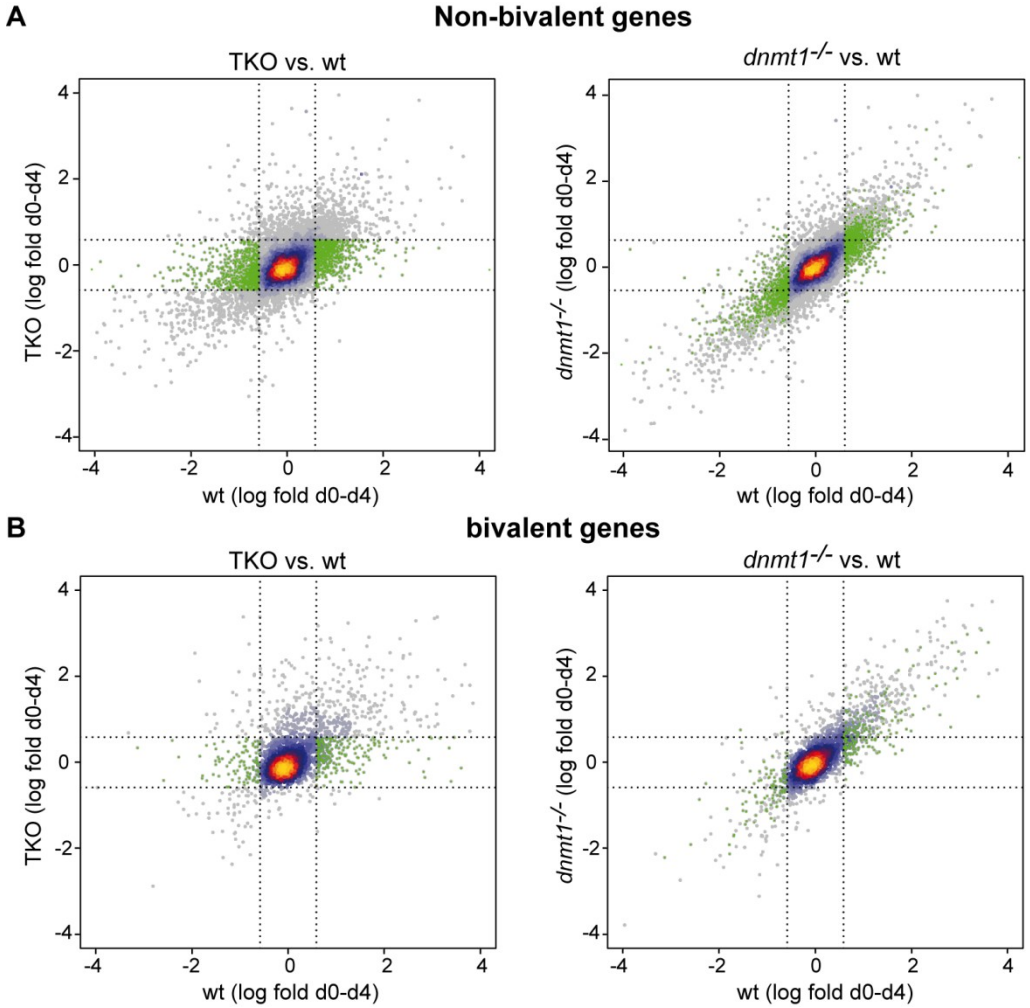


Figure 23. Expression changes of non-bivalent and bivalent genes in TKO and *dnmt1*$^{-/-}$ EBs relative to wt EBs at day 0-4 of differentiation

Scatter plots of expression changes for all 22550 non- bivalent genes (A) and all 2978 bivalent genes (B) in TKO (y-axis; left plots) and *dnmt1*$^{-/-}$EBs (y-axis; right plots) relative to wt EBs (x-axis). Genes whose expression levels alter in wt but not in TKO EBs ("TKO non- responders", horizontal corridor between the dotted lines of A and B) are marked in green in both panels of A and B. The other "responding" genes are heat- colored according to their local point density. An increased density of data points in the central doted squares of plots in (B) relative to (A) could be observed, indicating that the majority of genes carrying bivalent chromatin domains in the ESCs do not change their expression level between d0-4 in any of the genotypes with respect to the non- bivalent genes.

Despite the different number of genes, this analysis revealed similar distributions for the separate gene sets as for all genes together. Therefore, no general trend for genes which are bivalent in the undifferentiated state and then switch to repression by DNA methylation or Dnmt protein- dependent mechanisms during early differentiation could be observed. Instead, bivalent genes do not tend to change their expression levels under EB differentiation conditions since these genes cluster more at the center of the plots (Fig. 23). This is expected as EBs are composed of heterogeneous cell populations and bivalent genes are only activated in specific cell types and kept silenced in all other lineages.

3.1.6 Dnmts are required for silencing selected bivalent genes during differentiation

We also analyzed genes specifically upregulated in *dnmt1*$^{-/-}$ and/or TKO EBs to identify genes controlled by DNA methylation and/or Dnmt3 proteins. We detected 107 genes upregulated in both hypomethylated EBs during the first four days of differentiation (d0-4). These genes are enriched in GO terms for developmental processes like system, organ and anatomical structure development (Fig. 24A).

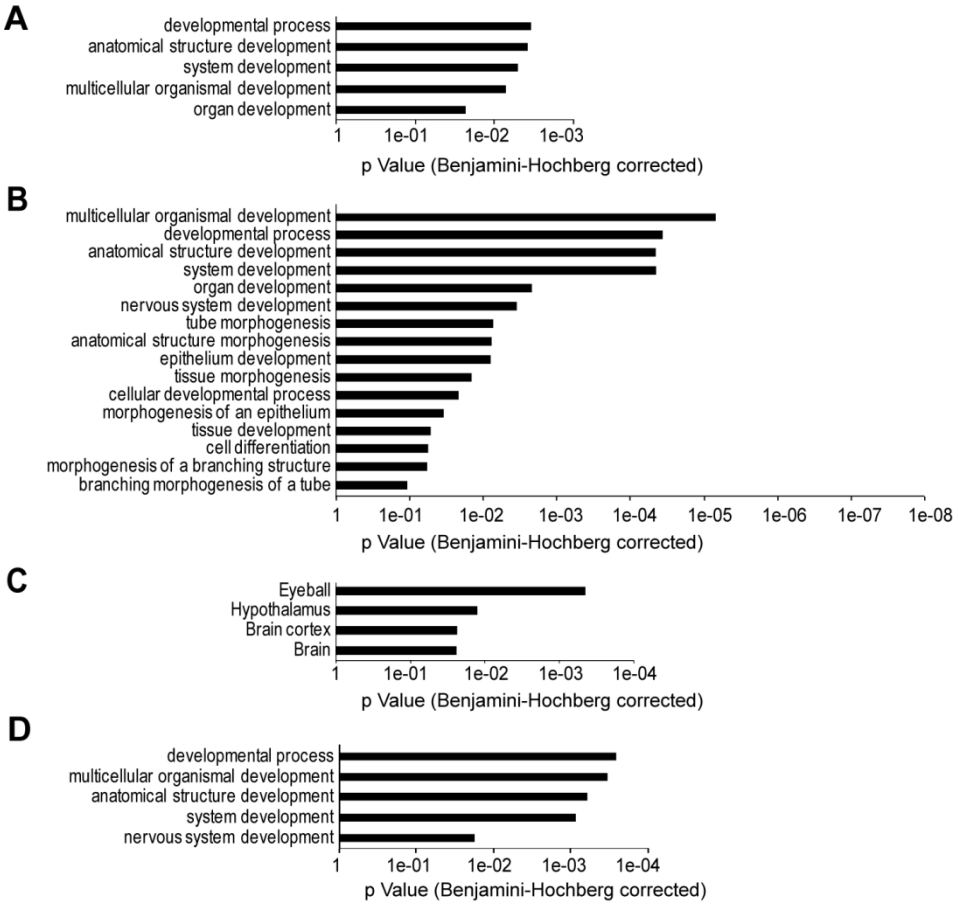


Figure 24. GO and tissue expression enrichment for genes concertedly upregulated in *dnmt1*$^{-/-}$ and TKO EBs as well as solely upregulated in *dnmt1*$^{-/-}$ or TKO EBs after the first 4 days of differentiation.
Enriched GO categories for genes commonly upregulated in *dnmt1*$^{-/-}$ and TKO EBs (A) and for genes solely upregulated in TKO (B) and *dnmt1*$^{-/-}$ EBs (D). Enriched tissues identified from genes upregulated in TKO EBs are displayed in (C).

In TKO EBs, more genes (353) were upregulated in comparison to the 224 genes upregulated in *dnmt1*$^{-/-}$ EBs (Fig. 15). Intriguingly, genes exclusively upregulated in either *dnmt1*$^{-/-}$ and/ or TKO EBs have been shown to play a role in the development of the nervous system (Fig. 24B, D) and in the case of TKO EBs, many exhibit a brain– specific expression (Fig. 24C). Up regulated genes in TKO EBs involved in neural differentiation include *nestin*, *hes6*, *fapb7/blbp*, *otx2* and *fzd3*, suggesting that DNA methylation is crucial for the repression of early neural lineage specification. In line with this, some of the genes were shown to be involved in the *in vitro* differentiation of ESCs to progenitors of neural rosettes (Abranches et al., 2009).

Interestingly, *nestin* and *otx2* carry bivalent domains in the ESC state and may have been missed in our global analysis of bivalent gene expression in hypomethylated EBs (Fig. 23). Therefore, we analyzed whether the expression of selected bivalent genes, including genes encoding early neural markers, are controlled by DNA methylation and/ or Dnmt proteins. To address this question, we quantified mRNA levels in wt, *dnmt1*$^{-/-}$ and TKO ESCs and EBs and analyzed CpG island methylation for *nestin* (all ectodermal progenitors) and *sox1* (early neuroepithelial progenitors). We extended the same analysis to *fgf5* (primitive ectoderm) and *brachury* (early mesoderm) as representatives of non- neural bivalent genes (Fig. 25). As a control we also analyzed a non- bivalent gene, *tet1*, which encodes the founding member of the Tet hydroxylase family and is highly expressed in ESCs, but silenced upon EB differentiation (Szwagierczak et al., 2010). As expected, transcript level analysis of *nestin*, *sox1, fgf5* and *brachury* revealed altered expression profiles in mutant ESCs/EBs, although to variable extents and at different time points during differentiation. Interestingly, out of all analyzed bivalent genes, an increase in CpG island methylation was only detectable for *nestin* and *sox1* (Fig. 25B, D). However, it is important to note that methylated sites were not found in all clones analyzed by bisulfite sequencing, indicating that the increased methylation of these genes only occurs in selected lineages. Furthermore, the observation that some genes showed relatively high global transcript levels is consistent with the presence of several clones which do not acquire any or very little DNA methylation at all stages. On the contrary, in all analyzed clones of *tet1* an increase in methylation was detectable in wt EBs, although surprisingly, *tet1* mRNA levels were not altered in mutant EBs (Fig. 25I, J). These data point towards a role of DNA methylation in mediating the suppression of bivalent genes involved in early neural differentiation. However, the proper repression of other selected bivalent genes like *fgf5* and *brachyury* possibly requires the presence of Dnmt proteins, but not the DNA methylation mark *per se* (Fig. 25F, H).

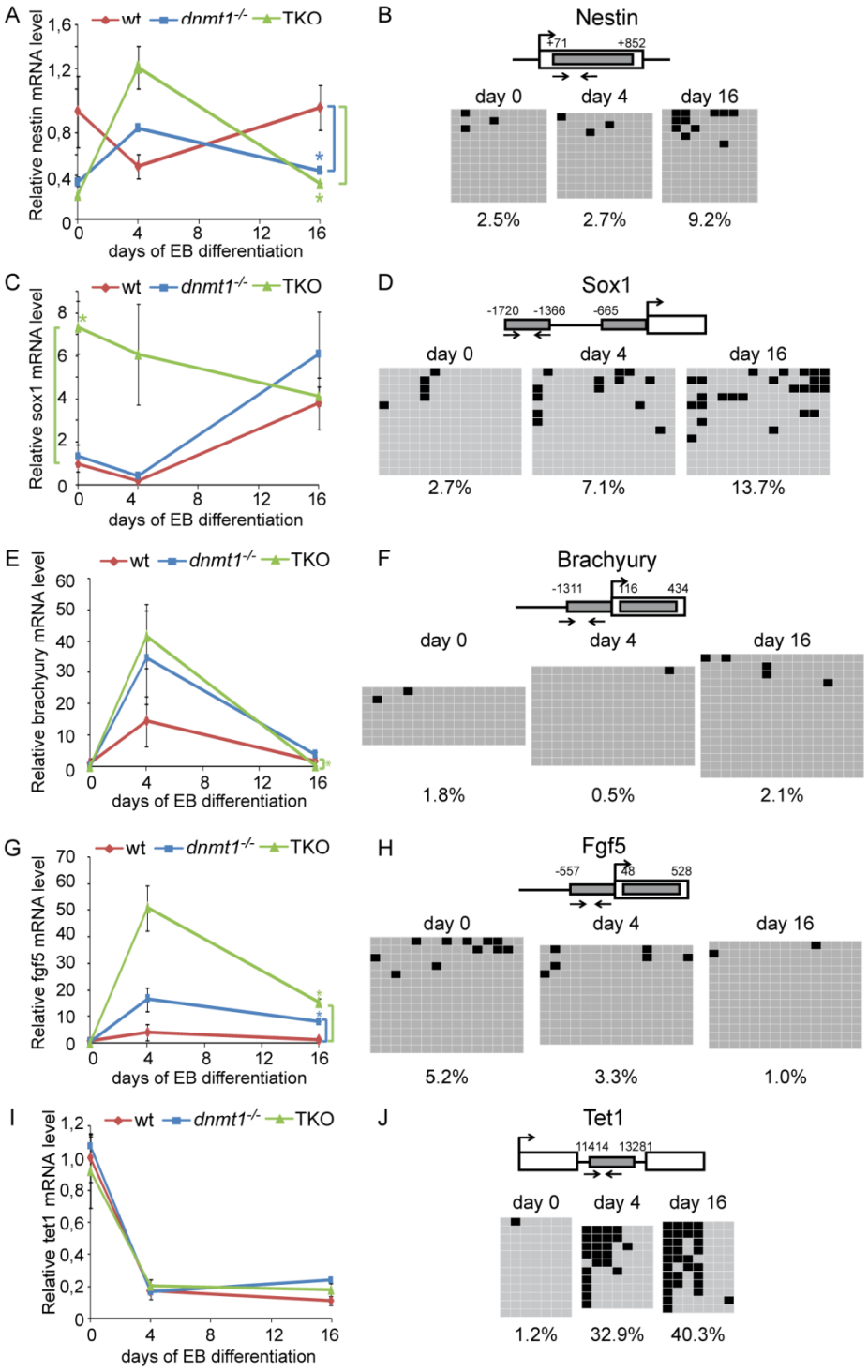


Figure 25. CpG islands of selected bivalent genes important for early neurodevelopmental processes are deregulated in mutant ESCs/EBs and methylated in subpopulations of differentiating wt cells.
Expression levels and CpG island methylation of bivalent genes *nestin* (A,B), *sox1* (C,D), *brachyury* (E,F) and *fgf5* (G,H) as well as non- bivalent control gene *tet1* (I,J) in undifferentiated ESCs (day 0) and at day 4 and 16 of EB differentiation. (A,C,E,G,I) Transcript levels were determined by qPCR in wt, *dnmt1*$^{-/-}$ and TKO ESCs/EBs as indicated and mean values and standard errors from biological triplicates are displayed. To directly compare the cell lines to each other, all values were normalized to undifferentiated wt ESCs. Asterisks indicate the significance level $p < 0.05$ (Student t-test). (B,D,F,H,J) DNA methylation analysis was performed using bisulfite sequencing in wt ESCs/EBs. In the gene cartoons large arrows indicate transcriptional start sites (TSS), open rectangles represent exons, grey shaded rectangles represent CpG islands and numbers depict boarders of CpG islands with respect to the TSS. Small arrows indicate the analyzed regions. In the panels grey and black squares indicate unmethylated and methylated CpG sites, respectively. Percentages of total CpG methylation within the analyzed regions/clones are indicated at the bottom of each panel.

3.1.7 Hypomethylated cells from late EBs revert to the undifferentiated state

Our data clearly demonstrated that hypomethylated EBs downregulated Oct4 protein levels, although silencing of *oct4* and *nanog* transcription was only partial relative to wt EBs. At the same time, both mutant EBs were able to activate differentiation programs, but the progression of differentiation programs was changed to an extent that depended on the presence of Dnmt3 proteins as *dnmt1*$^{-/-}$ cells containing both *de novo* Dnmts, possessed a greater differentiation potential compared to TKO cells. Thus, we wondered whether cells derived from late *dnmt1*$^{-/-}$ and TKO EBs could revert to the undifferentiated state in the presence of pluripotency promoting conditions. To address this question, we dissociated 12 days old wt, *dnmt1*$^{-/-}$ and TKO EBs into single cells and plated equal numbers of cells in the presence or absence of LIF (Fig. 26).

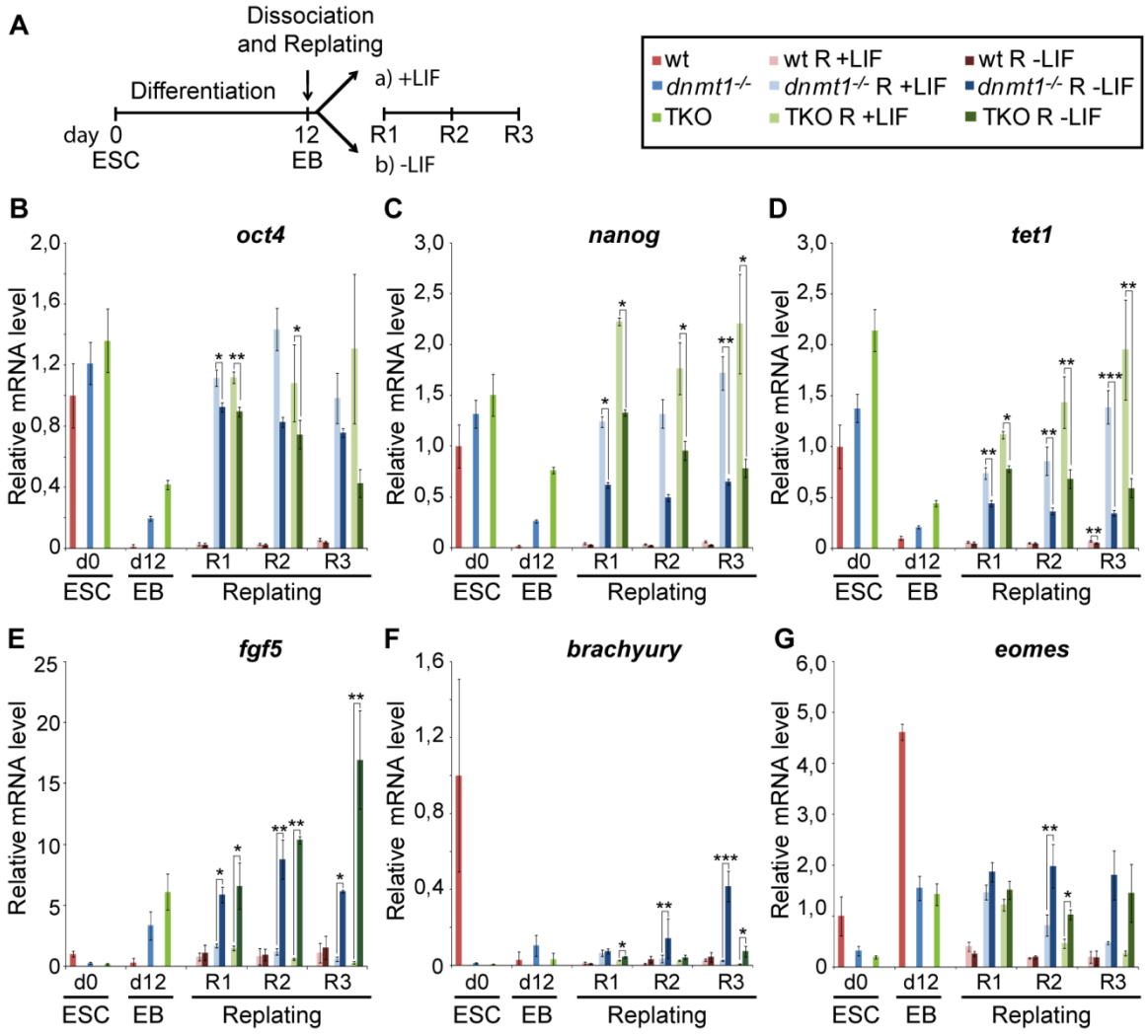


Figure 26. Reversible exit from the undifferentiated ESC state in cells with severe DNA hypomethylation
(A) Overview of experimental set up. EBs were cultured for 12 days, dissociated and their cells plated and cultured for three days (R1-3) in the presence or absence of LIF. Expression analysis of *oct4* (B), *nanog* (C), *tet1* (D), *fgf5* (E), *brachyury* (F) and *eomes* (G) was performed by qPCR in ESCs (day 0), EBs (day 12) and 1, 2 or 3 days after replating (R1-3). Wt samples are represented in shades of red, *dnmt1*$^{-/-}$ in shades of blue and TKO in shades of green colors; as indicated in the box at the upper right corner. Transcript levels are relative to day 0 in

wt ESCs and standard errors are the mean of 3 biological replicates. Asterisks indicate significance levels according to a student t-test * p<0.05; ** p<0.001; *** p< 0.0001.

We then analyzed the expression of pluripotency associated genes as well as markers of differentiation after one (R1), two (R2) or three (R3) days of replating. We chose this time point (day 12) as previous results clearly demonstrated that Oct4 protein levels were homogenously downregulated in all cell lines (Fig. 13). As expected, mRNA levels of pluripotency associated factors as well as differentiation markers did not show any response to the presence or absence of LIF in wt cells upon replating. By contrast, *dnmt1*$^{-/-}$ and TKO cells upregulated the expression of pluripotency associated factors within 24 h after replating. This effect was particularly obvious in the presence of LIF and cells eventually restored the transcription to levels originally measured in the respective undifferentiated ESC cultures. Furthermore, depending on the presence or absence of LIF, the expression of differentiation markers decreased or increased, respectively. Again, the addition of LIF resulted in expression levels similar to those originally determined in the undifferentiated state, demonstrating that cells from 12 days old *dnmt1*$^{-/-}$ and TKO EBs are still responsive to pluripotency promoting conditions. In the absence of LIF, the same replated cells from hypomethylated EBs kept or even augmented the expression of differentiation markers.

These data clearly demonstrate that even though hypomethylated EBs can activate and can progress in their differentiation programs (in the case of *dnmt1*$^{-/-}$), cells isolated from these mutant EBs can fully revert to the undifferentiated ESC state under appropriate culture conditions. This underlines the fundamental role of DNA methylation in establishing a barrier for the reactivation of pluripotency genes and in stabilizing transcriptionally silent states during differentiation.

3.2 Distinct functions of the two members of the Uhrf protein family, Uhrf1 and Uhrf2

Recently, Uhrf1 was identified as a crucial factor in the maintenance of DNA methylation patterns by recruiting Dnmt1 to hemi- methylated CpG sites (Bostick et al., 2007; Sharif et al., 2007). By binding to hemi- methylated CpG sites and to repressive histone marks, Uhrf1 connects the two major silencing pathways (see also Chapter 1.2.2). Interestingly, the second member of the Uhrf protein family, Uhrf2, contains a highly similar domain structure as Uhrf1 (see also Figure 5); however its functional role in maintenance DNA methylation or other biological processes remains elusive. Data published by Pichler *et al*. revealed that Uhrf2 is also able to bind to hemi- methylated CpG sites via its SRA domain, as well as to repressive histone marks via its tandem tudor domain. Furthermore, co- immunoprecipitation studies showed that Uhrf2 interacts with all Dnmts (Dnmt1, Dnmt3a and 3b), suggesting a potential role of Uhrf2 in the regulation of DNA methylation (Pichler et al., 2011).

3.2.1 Uhrf1 and Uhrf2 are differentially expressed in ESCs, various adult tissues, during differentiation to EBs and quiescence (serum starvation)

To elucidate the function of Uhrf2, we first analyzed the expression of *uhrf1* and *uhrf2*, in various ESCs, somatic cell lines, during Embryoid Body (EB) differentiation and in adult mouse tissues (Fig. 27).

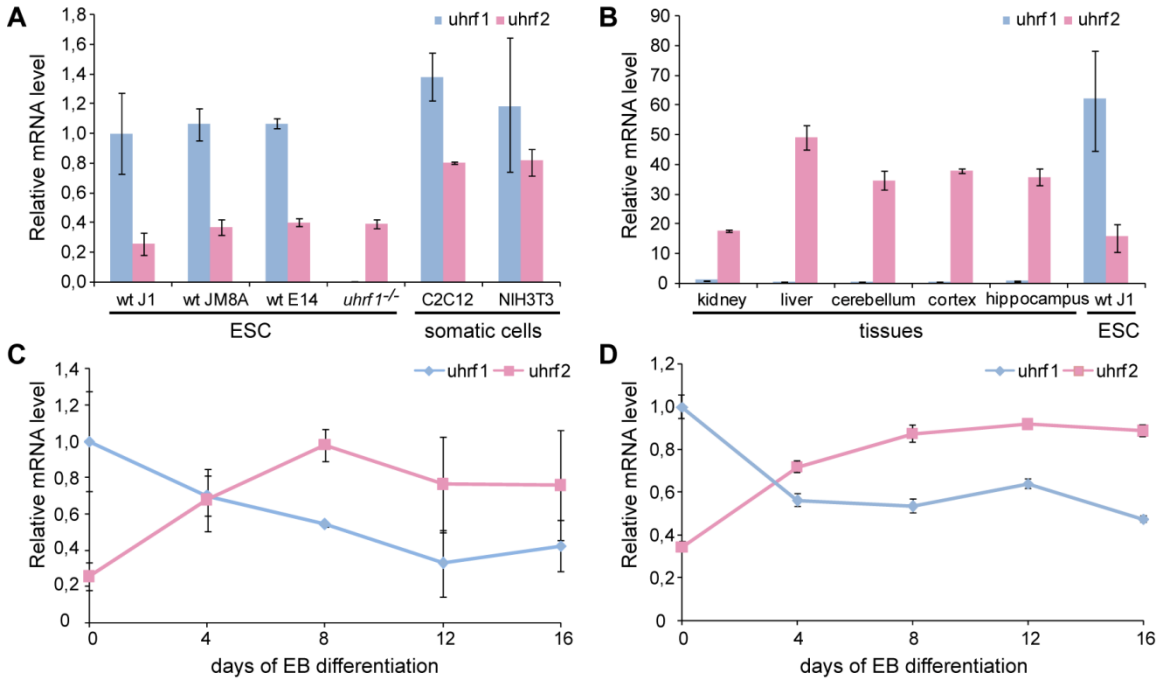


Figure 27. Opposite expression pattern of *uhrf1* and *uhrf2* in ESCs, during differentiation and in tissues
Transcript levels of *uhrf1* and *uhrf2* were measured by qPCR in different wildtype ESCs (wt J1, wt, JM8A, wt E14), *uhrf1*$^{-/-}$ ESCs and somatic cells (A), in various adult mouse tissues (B) and during differentiation of wt J1 (C) and wt E14 (D) ESCs to embryoid bodies (EBs). Expression levels are displayed relative to *uhrf1* in wtJ1 (A), day 0 of differentiation (C and D) and to kidney (B) (*uhrf1* set to 1). Error bars represent standard error of at least two independent experiments, with the exception of (D), where the technical error of one experiment is displayed and each sample was measured in triplicates. Part of these data were published in (Pichler et al., 2011).

The analysis revealed that the expression profile of both genes switches during differentiation of ESCs. *Uhrf1* is highly expressed in all analyzed wild type ESCs (wt J1, wt JM8A, wt E14) and progressively down regulated during differentiation of ESCs to EBs. These data are consistent with previous reports demonstrating that *uhrf1* is highly expressed in proliferating cells (Muto et al., 1995; Fujimori et al., 1998). In contrast, *uhrf2* expression is higher in somatic cells when compared to the transcript levels detected in ESCs and its expression is up regulated during EB differentiation. Furthermore, in differentiated adult mouse tissues like kidney, liver and several brain tissues, *uhrf1* mRNA levels are mostly absent and *uhrf2* is the predominately expressed member of the Uhrf protein family. The opposite expression pattern of *uhrf1* and *uhrf2* argues against a functional redundancy of both proteins. Consistent with this, we could not detect a compensatory up- regulation of *uhrf2* in *uhrf1*$^{-/-}$ ESCs. Hence the switch in transcription levels of both members of the Uhrf protein family indicates different functional roles of *uhrf1* and *uhrf2* during differentiation.

As the expression of *uhrf1* is known to be regulated in a proliferation- dependent manner (Uemura et al., 2000; Miura et al., 2001), we wondered whether *uhrf2* shows a similar regulation dependency on the progression of the cell cycle. To address this questions we serum- deprived NIH3T3 mouse fibroblasts for 48h which caused the majority of the cells to accumulate in a quiescent state (G1/G0 phase of the cell cycle) (Fig. 28).

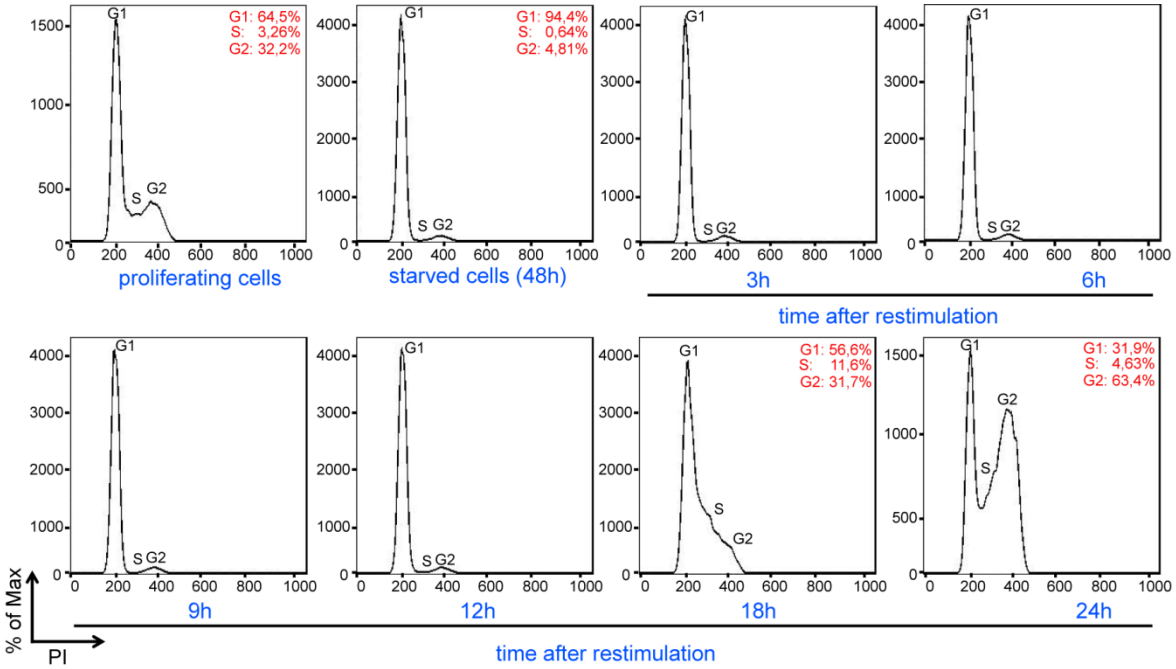


Figure 28. Serum- starvation of NIH3T3 fibroblasts results in the arrest of the majority of the cells in G1/G0 phase of the cell cycle.

To measure the cell cycle profile of proliferating, serum- deprived and restimulated NIH3T3 fibroblasts, cells were fixed, stained with Propidium Iodide (PI) and analyzed by FACS. The starvation of cells was very effective as less than 1 % of the cells were in S-phase and almost 95 % of all cells were arrested in G1/G0- phase as compared to the proliferative state, where only two thirds of the cells were in G1- phase and more than 3 % in S- phase. A change in the cell cycle profiles could be detected 18hours after restimulation by the addition of serum. Numbers in red describe the percentage of cells in the different phases of the cell cycle (G1, S, and G2). The analysis was performed with the help of my colleague Sebastian Bultmann.

By adding serum we then restimulated proliferation of the cells and took samples for expression analysis after 3, 6, 9, 12 and 24h after serum addition. As a control, we followed the expression of *dnmt1* and *pcna* during normal growth, starvation and restimulation, as the expression of both genes is known to be dependent of the proliferation of the cells. Consistent with previous data (Uemura et al., 2000; Miura et al., 2001), all three genes show a proliferation- dependent expression (Fig. 29). During serum starvation transcript levels of *uhrf1*, *dnmt1* and *pcna* were significantly reduced, but reached initial expression level about 12 hours after serum addition with respect to normally proliferating cells. In contrast, *uhrf2* expression was highly up- regulated in the quiescent state and mRNA levels were drastically down regulated upon restimulation. Taken together, these data argue for distinct roles of *uhrf1* and *uhrf2*.

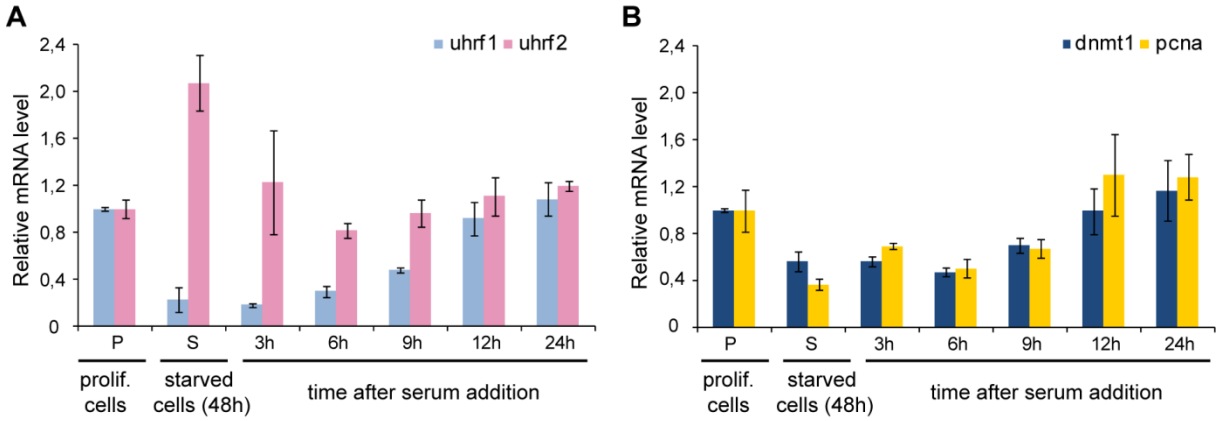


Figure 29. *Uhrf1*, but not *uhrf2* shows proliferation dependent expression.
Expression analysis of *uhrf1*, *uhrf2*, *dnmt1* and *pcna* in proliferating (P), serum- starved (S) and restimulated NIH3T3 fibroblasts after 3, 6, 9, 12 and 24h of serum addition. Transcript levels were determined by qPCR and are displayed relative to the proliferative state of each gene (P is set to 1). Error bars represent standard errors of three biological replicates, each measured in triplicates.

3.2.2 Uhrf2 does not play a role in maintenance DNA methylation in proliferating cells

To gain more insights into the role of Uhrf2 in regulating DNA methylation we performed transient knock downs of *uhrf2* in wt and *uhrf1*$^{-/-}$ ESCs with two different siRNAs. Scrambled siRNA (control siRNA) was used to control for possible off- target effects. As currently no specific antibody against murine Uhrf2 is available, we monitored the knock down efficiency by qPCR (Fig. 30A, B). After two days of treatment, *uhrf2* expression was already very efficiently down regulated by both specific siRNAs and transcript levels stayed low during consecutive rounds of siRNA treatment. As previous studies on RNAi- mediated knock down of *dnmt1* in HCT116 cells have demonstrated that a dramatic effect on genomic DNA methylation could only be observed after at least 8 days of siRNA treatment (Spada et al., 2007), we chose to analyze DNA methylation levels after 10 days of repeated siRNA transfections. This prolonged treatment should also ensure a good knock down efficiency of *uhrf2* and accumulation of potential errors in DNA methylation patterns. Hence, after 10 days

of siRNA treatment, we isolated genomic DNA and analyzed the DNA methylation status at repetitive sequences like major satellites using bisulfite treatment followed by pyrosequencing (Fig. 30C, D). As expected, *uhrf1*$^{-/-}$ ESCs show a drastic reduction in DNA methylation at repetitive sequences in comparison to wt ESCs. However, DNA methylation levels were neither affected in wt nor in *uhrf1*$^{-/-}$ ESCs by the reduced *uhrf2* mRNA level, suggesting that Uhrf2 might not play a role in the regulation of maintenance DNA methylation in ESCs.

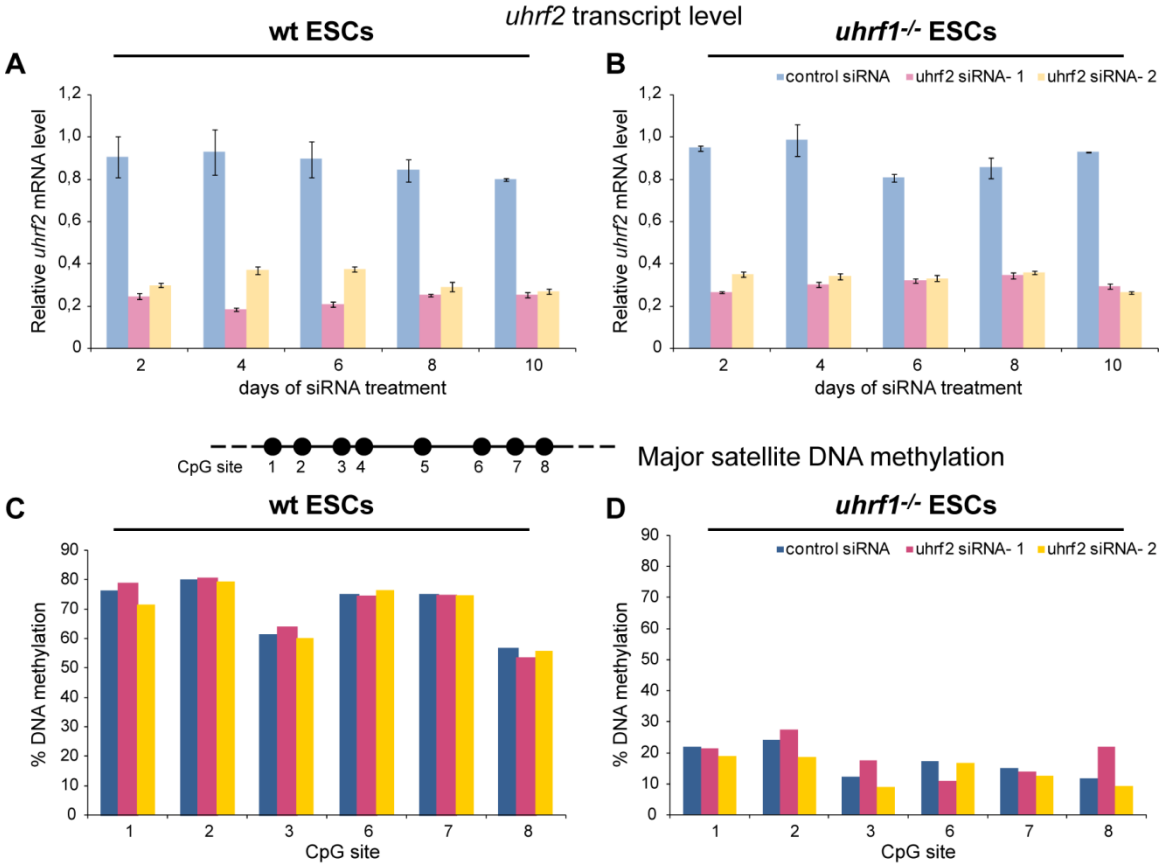


Figure 30. Transient knock down of *uhrf2* in ESCs does not affect DNA methylation levels at repetitive sequences.
(A and B) Wildtype (left side) and *uhrf1*$^{-/-}$ (right side) ESCs were treated with two different siRNAs targeting *uhrf2* (siRNA- 1 and siRNA- 2). Cells were transfected with siRNA every second day for a total of 10 days and samples were taken for qPCR analysis to investigate the success of *uhrf2* downregulation. Control siRNA was used to monitor possible off- targets caused by repeated transfections with siRNA. Untreated cells (not shown) were used for normalization and set to 1. Shown is the standard error of technical replicates of one biological replicate. (C and D). DNA methylation levels were analyzed at Major satellite sequences in ESCs treated with specific or control siRNA for 10 days using bisulfite treatment followed by pyrosequencing.

Since our first experiments of silencing *uhrf2* expression by RNAi showed no effect on maintenance DNA methylation, we wondered whether the residual *uhrf2* mRNA levels (about 20 % with respect to control treated cells) were sufficient to maintain relative high Uhrf2 protein levels in the cells. As we were unable to monitor the remaining levels of Uhrf2 protein after RNAi treatment due to the lack of a Uhrf2 specific antibody we acquired *uhrf2*$^{-/+}$ ESCs, which only express about 50 % of *uhrf2* with respect to their wt cells (Figure 31A, B).

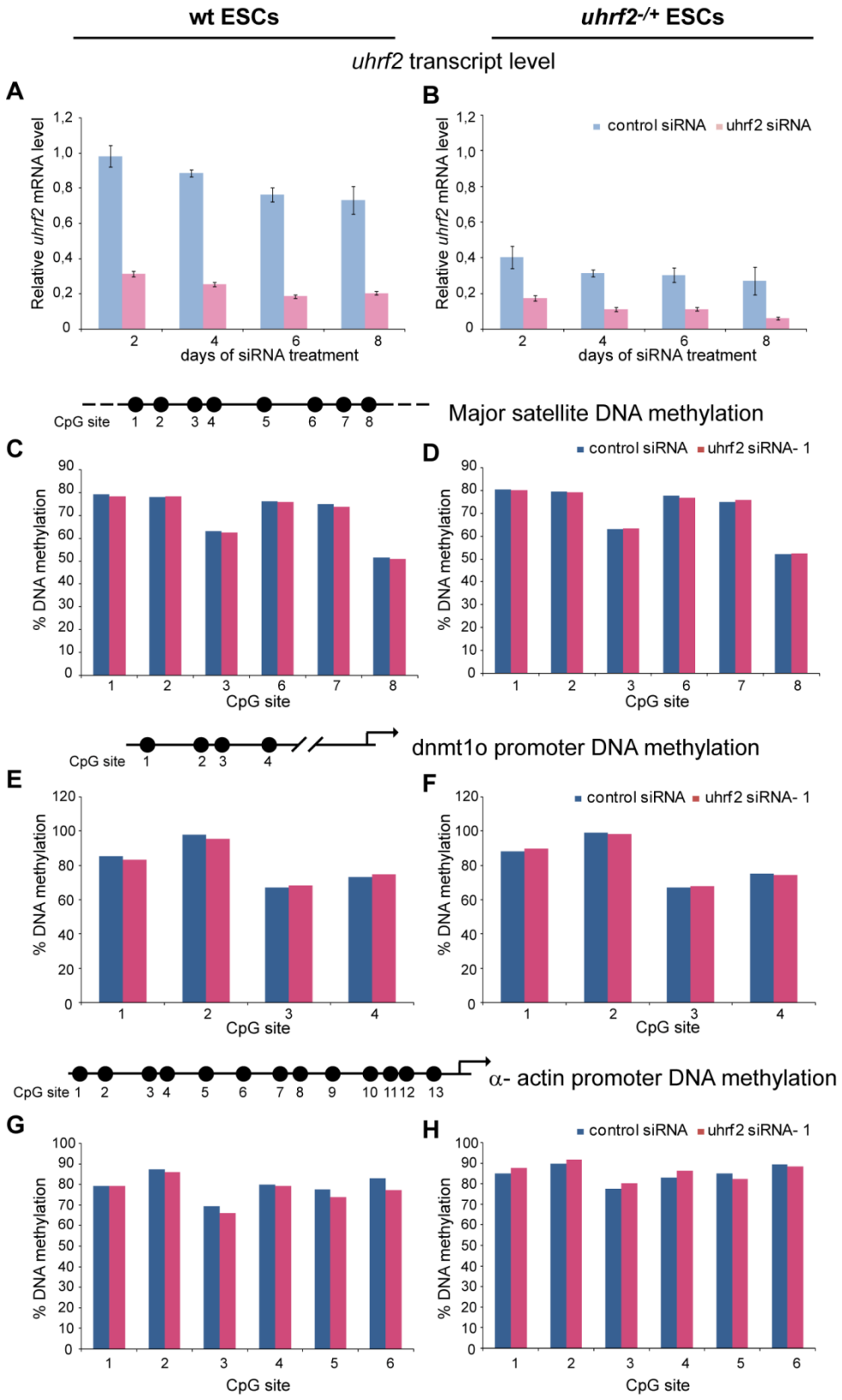


Figure 31. RNAi- mediated silencing of *uhrf2* in ESCs does not affect maintenance DNA methylation.
(A and B) Wildtype (left side) and *uhrf2*$^{+/-}$ (right side) ESCs were either transfected with control or specific siRNA (siRNA- 1) every second day and knock down efficiency was monitored by qPCR. Untreated wt cells (not shown) were used as reference and set to 1 so that levels in wt and *uhrf2*$^{-/+}$ ESCs are directly comparable. Shown is the standard error of technical replicates of one biological replicate. (C–G). Pyrosequencing data at major satellite sequences (C, D), *dnmt1o* promoter (E, F) and *skeletal α-actin* promoter (G, H) in wt (left) and *uhrf2*$^{-/+}$ ESCs treated with specific or control siRNA for 8 days. For the skeletal α-actin promoter, only the results from the first 6 CpG sites are displayed.

We hypothesized that RNAi treatment of these cells should achieve a higher knock down efficiency. Hence, we treated wt and *uhrf2*$^{-/+}$ with control or specific siRNA for 8 days and monitored efficiency of *uhrf2* silencing by qPCR. In wt ESCs, we obtained similar down regulation of *uhrf2* to about 80 % compared to control treated cells. As expected, in ESCs heterozygous for *uhrf2*, mRNA levels of *uhrf2* were reduced to about 95 % with respective to wt cells, demonstrating that the lower starting transcript levels indeed improved the total knock down efficiency. After 8 days, we then analyzed DNA methylation levels at repetitive sequences as well as single- copy genes in both cell lines (Fig. 31C-G). Although *uhrf2* expression was drastically reduced after 8 days of siRNA treatment, neither an effect on DNA methylation level at repetitive sequences (i.e. major satellites) nor at single copy genes (i.e. promoter of *dnmt1o* or *skeletal α- actin*) was detected. These data, together with the previous RNAi mediated knock down experiment suggest that Uhrf2 does not play a role in maintenance DNA methylation in undifferentiated ESCs.

The observation that *uhrf2* transcript levels increase during differentiation could imply a role of Uhrf2 in maintenance methylation during development. As *uhrf2*$^{-/+}$ ESCs show reduced mRNA levels of *uhrf2,* we wondered whether these cells exhibit any defects in maintaining DNA methylation patterns during differentiation. Therefore, we differentiated wt and *uhrf2*$^{-/+}$ ESCs to EBs and analyzed the expression of *uhrf1* and *uhrf2* in the undifferentiated state (day 0) as well as 4 and 16 days after differentiation (Fig. 32A,B). Whereas *uhrf1* transcript levels were similarly down regulated independently of the genotype of the cell lines, *uhrf2* expression in *uhrf2*$^{-/+}$ ESCs was increased in both EBs during differentiation. However, in heterozygous EBs, *uhrf2* mRNA levels reached on 50% of the transcript levels measured in wt EBs. We then analyzed the level of DNA methylation in the ESCs state (d0) and at both time points during differentiation (d4 and d16) by bisulfite treatment followed by pyrosequencing (Fig. 32C-F). For both analyzed sequences, no difference in DNA methylation was observed in *uhrf2*$^{-/+}$ ESCs and during EB differentiation, possibly because of the remaining Uhrf2 level in the heterozygous cells.

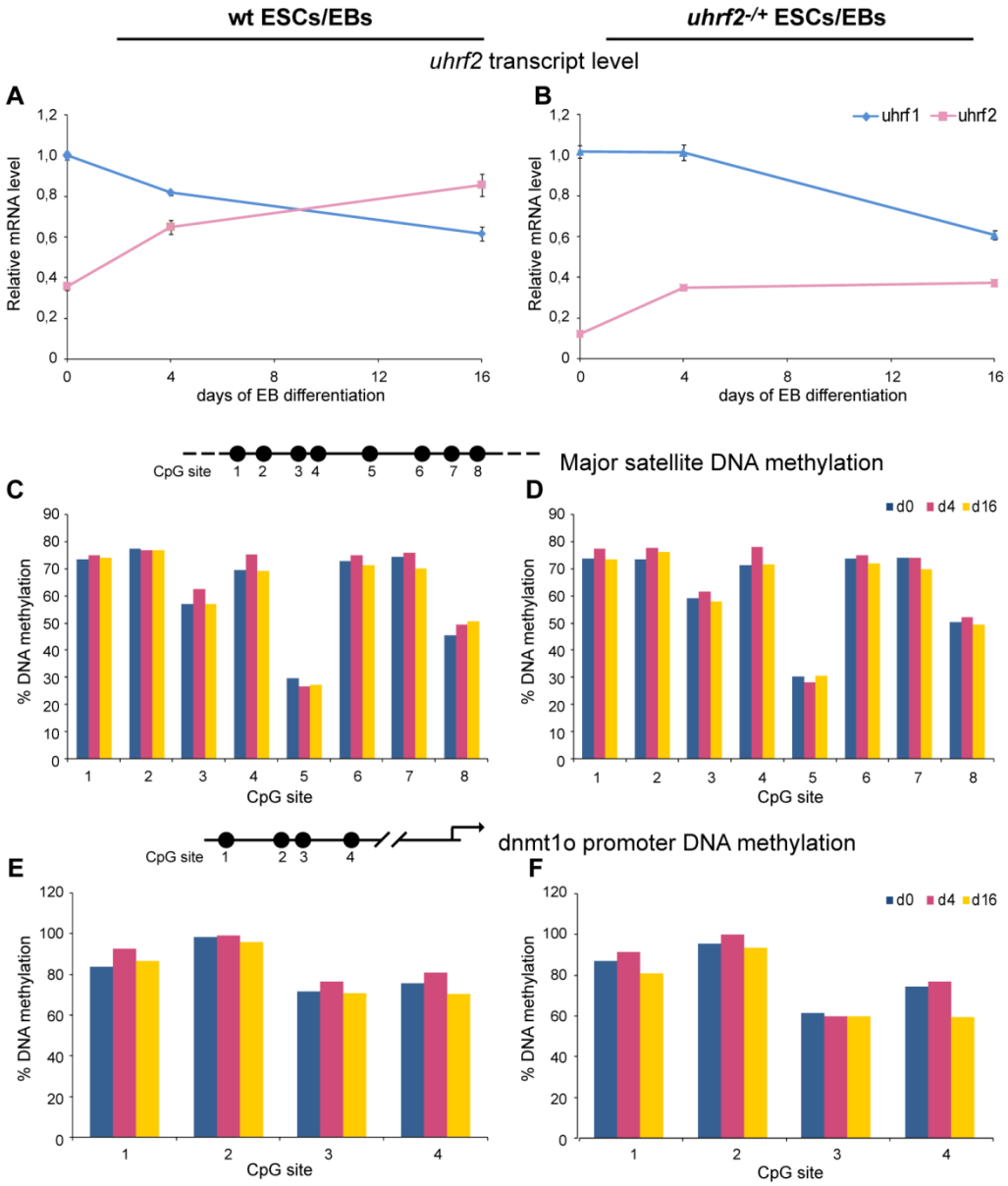


Figure 32. Reduced uhrf2 transcript levels do not alter DNA methylation during differentiation.
Wildtype (left) and *uhrf2*$^{-/+}$ (right) ESCs were differentiated to EBs and samples for measuring expression levels as well as DNA methylation levels were taken and analyzed in the pluripotent state (d0) as well as 4 and 16 days after differentiation. (A and B) Expression analysis of *uhrf1* and *uhrf2* in wt and *uhrf2*$^{-/+}$ ESCs and EBs by qPCR. Transcript levels of *uhrf1* in undifferentiated wt cells were used as a reference (set to 1), so that mRNA levels between the two cell lines (wt and *uhrf2*$^{-/+}$) are directly comparable. Error bars represent the standard error of three technical replicates. (C–F). DNA methylation of ESCs (d0) and 4 and 16 days old EBs at major satellite sequences (C, D) and the *dnmt1o* promoter (E, F) was measured by pyrosequencing.

In general, the residual levels of *uhrf2* mRNA in *uhrf2*$^{-/+}$ ESCs and during RNAi- mediated knock down could mask the real phenotype of Uhrf2 and hence might not be the best approach to analyze the function of Uhrf2, especially in light of a missing antibody to control for knock down efficiency at the protein level. Therefore, the best solution to analyze a possible function of Uhrf2 in maintenance DNA methylation, would be to generate *uhrf2*$^{-/-}$ ESCs and then analyze DNA methylation pattern in the undifferentiated state as well as during differentiation and mouse development.

3.3 Novel methods to quantify and map 5 hmC in genomic DNA

The discovery of the 6th base 5 hydroxymethylcytosine (5hmC) in the genome of ESCs and in several tissues with particularly high levels of 5hmC in the central nervous system (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009), raised interest in the functional role of this newly identified modification. Several hypotheses about the possible function of 5hmC have been formulated, including roles as an epigenetic mark and/ or an intermediate in the active demethylation pathway (see also chapter 1.2.3). Furthermore, mutations in *tet2* have been linked to various myelodysplastic syndromes including myeloid leukemia, suggesting that aberrant global 5hmC patterns might contribute to the development of myeloid malignancies (Langemeijer et al., 2009; Ko et al., 2010).

Two properties of the newly identified modification make it technically challenging to quantify and selectively detect 5hmC: its low abundance and its structural similarity to the more abundant 5mC. Moreover, the classical approach to determine genomic 5mC levels, bisulfite treatment, cannot be used to distinguish 5hmC from 5mC. Therefore, several tools have been adapted or newly developed to quantify and map 5hmC. These methods include among others thin layer chromatography (TLC), liquid chromatography in combination with mass spectrometry (LS- MS) and the use of antibodies for detection. However, these techniques have several drawbacks as in the case of TLC accuracy of 5hmC detection proves to be difficult or the use of antibodies for detection bears the risk of unspecific binding (reviewed in (Branco et al., 2012).

3.3.1 Sensitive enzymatic quantification of global hmC levels

Given all the limitations and drawbacks of the methods used so far to quantify global 5 hmC levels, we developed a sensitive, enzymatic assay for accurate quantification of genomic 5hmC levels. The establishment of the glucosylation assay and the quantification of global 5hmC levels in various tissues, ESCs and during EB formation was carried out by my colleagues Aleksandra Szwagierczak and Sebastian Bultmann. I contributed material from ESCs and EBs, participated in the isolation of genomic DNA from tissues, isolated RNA from all samples and analyzed the expression of *tet1-3* transcript levels by qPCR.

The basis of the newly developed method is the use of glucosyltransferases of T- even bacteriophages, which have been shown to specifically transfer glucose from an uridine 5´-diphosphate (UDP)- glucose donor to the hydroxymethyl group of 5hmC. Interestingly, the DNA from T4 bacteriophages is devoid of cytosine residues and instead contains 5hmC residues, which are modified by α- and ß- glucosyltransferases (α- and ß- gt). As previous studies demonstrated that ß- gt is more efficient for *in vitro* glucosylation assays, we focused on ß- gt rather than α- gt (Kornberg et al., 1961; Georgopoulos and Revel, 1971). Using a

standard curve with known 5hmC content, we showed that the incorporation of isotopically labeled glucose ([^{3}H]glucose) in DNA can be reliably used to measure the abundance of 5hmC. Hence, with this ß- gt glucosylation assay, genomic 5hmC can be specifically labeled and accurately quantified by comparison to a standard curve. We applied this assay to various adult mouse tissues as well as to two different mouse strains of undifferentiated ESCs and their corresponding differentiated EBs (wt J1 and wt E14) and correlated the hmC levels to the relative transcript levels of *tet1-3* (Figure 33).

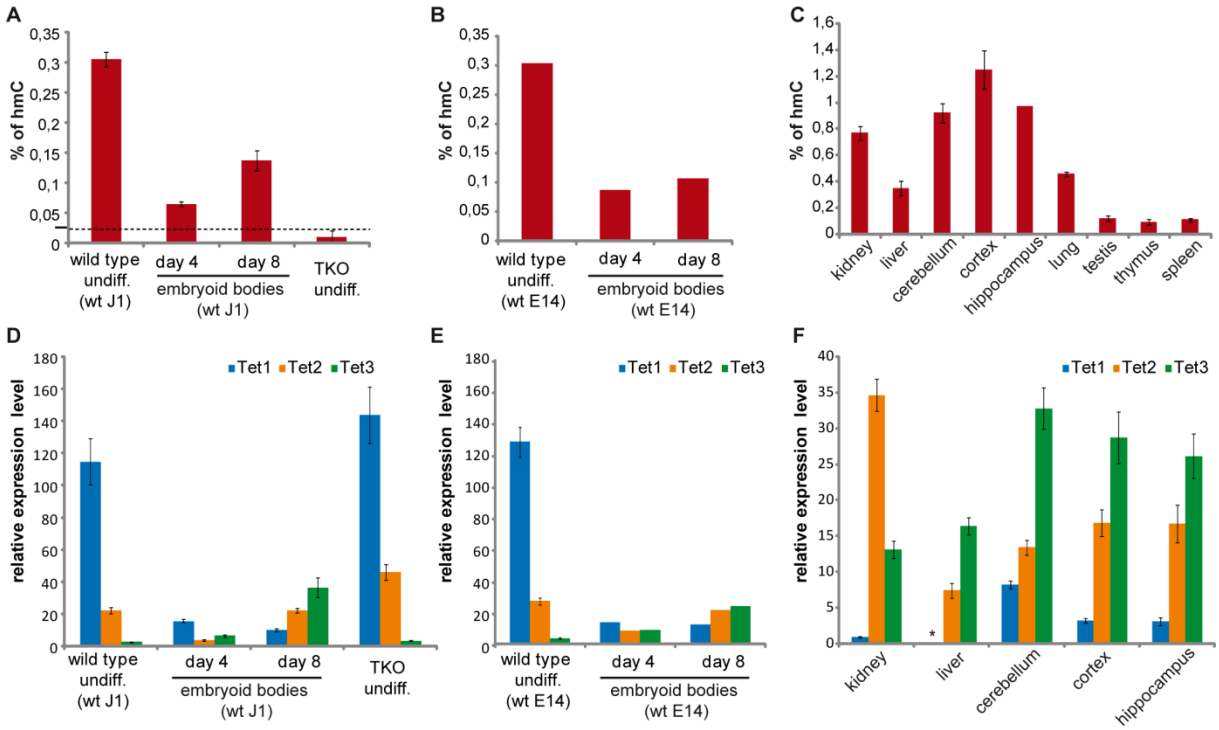


Figure 33. Quantification of genomic 5hmC and *tet* mRNA levels in ESCs, EBs and tissues.
(A-C) Measurements of global 5hmC levels using the 5hmC glucosylation assays. A calibration curve was used as a reference to calculate the percentage of 5hmC per total cytosine. Error bars in (A) and (C) represent standard deviation of two (A) or one (C) biological replicate, each measured in two independent assays, whereas samples in (B) and the hippocampus were only measured once. Every sample was measured twice in every assay. The dashed line in (A) represents the calculated limit of detection (0.025 %). (D-F) qPCR analysis of *tet* expression. Transcript levels were calculated relative to *tet1* in kidney (set to 1), so that values in D, E and F are directly comparable. Error bars represent standard deviation from two (D, E) and one (F) biological replicate, each determined from two independent reactions of cDNA synthesis, with the exception of wt E14 EBs after 4 and 8 days, which were only measured from one cDNA synthesis and therefore lack error bars. Each sample was measured in triplicates in the qPCR reaction. Asterisk depicts that obtained values where at the limit of detection and therefore could not be quantified accurately. Note that genomic DNA and RNA used in A/B/C and D/E/F, respectively, were isolated from the same cell and tissue sample. Data were published in (Szwagierczak et al., 2010).

Our measurements using the glucosylation assay revealed that both wild type ESCs contained 0.3 % 5hmC relative to total cytosine (Fig. 33A and B). As a control, we also measured genomic 5hmC level in Dnmt1, Dnmt3a/3b triple knockout (TKO) ESCs, which are devoid of any genomic cytosine methylation and hence should also contain very little, if any, hydroxymethylation. Indeed, we could only measure background levels of 5hmC in TKO ESCs below the calculated detection limit of 0.025 % of our assay. Expression analysis by qPCR showed that *tet1-3* transcript levels were comparable in both wild type as well as TKO

ESCs, with *tet1* transcripts being the most abundant and *tet3* the least expressed in undifferentiated ESCs (Fig. 33D and E). Previous studies revealed that differentiation of ESCs by withdrawal of LIF from monolayer cultures for 5 days leads to decreased genomic 5hmC with a concomitant reduction of *tet1* transcript levels (Tahiliani et al., 2009). We differentiated both wild type ESCs to EBs and followed 5hmC levels and *tet1-3* transcript dynamics. Interestingly, we found that 5hmC levels initially decrease during differentiation of ESCs (day 4 EBs), but increased again after 4 additional days of EB culture (day 8 EBs). Furthermore, we could observe distinct transcription dynamics of *tet* genes during EB differentiation. *Tet1* was predominately expressed in the undifferentiated state but drastically down regulated already in the first 4 days and further decreased with differentiation. Similarly, *tet2* transcripts were down regulated at day 4 of differentiation, but were fully restored in EBs at day 8. By contrast, *tet3* transcript levels drastically increased during the course of differentiation. After the first 4 days, *tet3* mRNA levels doubled and increased by up to 20 fold in prolonged EB cultures as compared to the levels in the undifferentiated state. Therefore, the high levels of *tet1*, together with the lower *tet2* transcripts in the undifferentiated state correlate with the relatively high 5hmC levels. The partial recovery of 5hmC after 8 days of EB culture correlates with increased expression of *tet2* and *tet3.*

We then measured 5hmC as well as *tet1-3* mRNA levels in several adult tissues (Fig. 33C and F). Consistent with previous reports (Kriaucionis and Heintz, 2009), we found the highest levels of genomic 5hmC in brain tissues which correlated with high levels of *tet3* and to a lower extent *tet2* transcripts. In general, all analyzed adult tissues were characterized by high levels of *tet3* and low levels of *tet1*, whereas undifferentiated ESCs show the reversed expression pattern. Strikingly, we detect a relatively high amount of 5hmC in kidney with concomitant high levels of *tet2* in this tissue. The predominate expression of *tet2* in kidney is consistent with the observation, that one of the described phenotype of *tet2*$^{-/-}$ mice is a cellular defect in proximal convoluted tubules of the kidney (Tang et al., 2008).

Taken together, our analysis demonstrates a correlation of the different amounts of genomic 5hmC in various adult tissues, ESCs and during differentiation with the differential expression of *tet* genes.

3.3.2 The 5hmC specific endonuclease *Pvu*RtsI1 as a tool to profile genomic 5hmC patterns

To gain insights into the functional role(s) of the newly discovered "6th base", it will be necessary to determine genomic 5hmC patterns. As described earlier, the genome of T4 bacteriophages contains exclusively 5hmC instead of cytosine residues, which are modified by α- and ß- glucosyltransferases. The switch of cytosine to 5hmC in the T4 genome is thought to have evolved as a protection system against the restriction enzymes of the host

bacteria after infection. As a strategy to counter the phage`s measures, bacteria have evolved a system of restriction enzymes that specifically recognize modified cytosines. One of the bacteria enzymes, the endonuclease PvuRts1I has been demonstrated to cleave glucosylated 5hmC and its restriction activity was shown to be modulated by glucosylation in a complex way. PvuRts1I is encoded by a single gene found on the kanamycin resistance plasmid Rts1 which was originally isolated from *Proteus vulgaris* (Janosi et al., 1994). Interestingly, the growth of 5hmC containing T- even bacteriophages, but not that of T-odd phages containing 5mC or λ- phages devoid of any modified base is restricted by bacteria carrying the Rts1 plasmid (Janosi et al., 1994). This suggests that PvuRts1I could be a useful tool to discriminate 5hmC from 5mC or unmodified cytosine.

To address this question, we purified recombinant PvuRts1I and showed that it selectively cleaves non- glucosylated 5hmC containing DNA with even higher efficiency than α- or ß-glucosylated DNA. We then determined the cleavage pattern of PvuRts1I by generating libraries of restriction fragments of the whole non- glucosylated T4 genome or a reference fragment produced from the same genome containing exclusively hdroxymethylated cytosine. Random sequencing of more than 100 clones from each library revealed a consensus sequence of $^{hm}CN_{11\text{-}12}/N_{9\text{-}10}G$ with a 2 nucleotide 3´- overhang. Furthermore, by comparing DNA substrates containing one single PvuRts1I consensus site ($^{hm}CN_{12}/N_{10}G$) with either hmC or mC in symmetrical or asymmetrical configuration or unmodified C, we found that sites with symmetric hmC are the preferred substrates of PvuRts1I. These experiments were conducted by my colleague Aleksandra Szwagierczak and are published in (Szwagierczak et al., 2011).

All the experiments performed so far used T4 genome or artificial DNA substrates as templates. As a next step, we wanted to investigate whether PvuRts1I could be used as a tool to map 5hmC patterns in mammalian genomic DNA. All the following experiments were conducted by me, except for the radioactive measurements of 5hmC levels in DNA substrates, which were conducted by my colleague Aleksandra Szwagierczak.

To analyze the efficiency of PvuRts1I digestion for mammalian genomic DNA, we chose the upstream regulatory region III of the mouse *nanog* gene (Hattori et al., 2007). We selected this region because very recent data indicate that Tet1 binds to this region and keeps the *nanog* promoter in a hypomethylated, active state. Further evidence for this hypothesis comes from the observation that this region acquires CpG methylation upon knock down of Tet1 in ESCs (Ito et al., 2010). Hence, the upstream regulatory region of *nanog* represents a potential 5hmC containing sequence in ESCs.

Firstly, we digested or mock- treated genomic DNA from wild type and TKO ESCs with PvuRts1I and used two different primer pairs (Primer 1 and Primer 2) to analyze the decrease in product after digestion compared to mock digested samples (Figure 34). The primers were chosen in *nanog* regions which according to Ito *et al*. show increased DNA methylation upon tet1 knock down in ESCs (Ito et al., 2010).

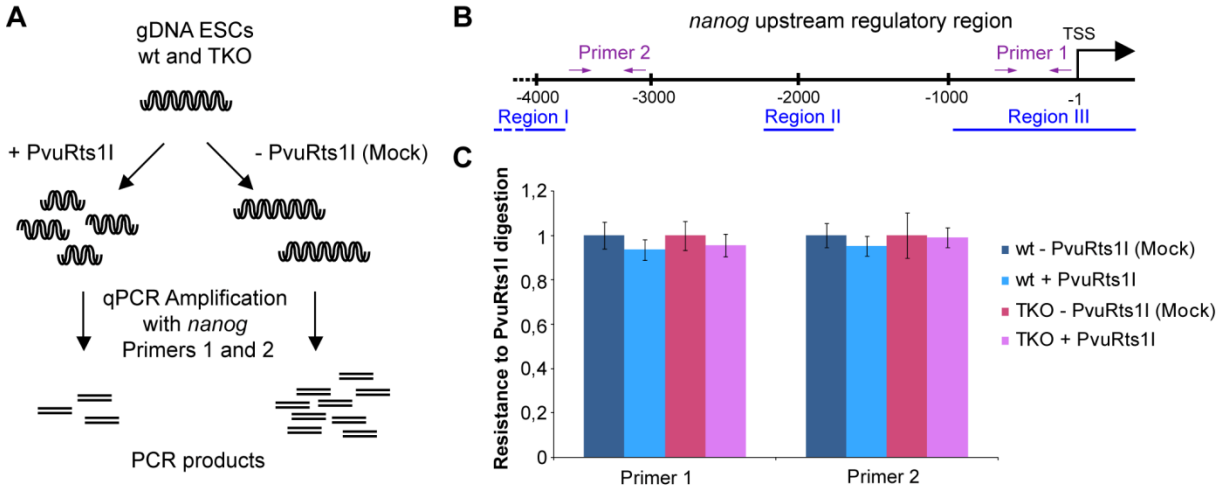


Figure 34. Amplification of PvuRts1I digested fragments does not reduce amount of PCR products. Genomic DNA from wt or TKO ESCs was digested with PvuRts1I or mock treated and amplified with primers specific for the *nanog* locus by qPCR. A locus containing 5hmC and digested with PvuRts1I should result in fewer template for the qPCR. As TKO ESCs are devoid of any DNA methylation, genomic DNA from TKO serves as a negative control for the specificity of the digestion reaction. A) shows an outline of the strategy used to detect 5hmC in mammalian genomic DNA whereas B) describes the upstream regulatory region of the *nanog* locus and also marks the location of the primers used in the qPCR reaction. The results of the amplification of fragments treated or not treated are depicted in C). Shown is the technical error of one representative experiment and each sample was measured in triplicates in the reaction. Each mock treated sample for each analyzed region and cell line was set to 1 to calculate the change in product after digestion.

However, we could not detect a decreased resistance to PvuRts1I digestion in the two *nanog* promoter regions as qPCR amplification with both primer pairs did not lead to a reduction of products in the digested samples. This could be due to the low abundance of 5hmC in genomic DNA of ESCs which makes it technically challenging to detect slight differences in decreased amounts of PCR products. We then thought of a strategy to positively identify rare PvuRts1I digestion products. As digestion with PvuRts1I results in fragments with a two nucleotide 3´ overhang, we generated a linker with a random two nucleotide 3´overhang (Fig. 35A), which we ligated to the digested products. We then used *nanog* specific primers paired with a linker specific primer to amplify ligation products. Unfortunately even using this adapted protocol we were unable to detect any amplification products (data not shown). The lack of amplification products may be explained by an extremely rare occurrence of 5hmC at PvuRts1I cleavage sites of this locus, especially as PvuRts1I preferentially cleaves sites with symmetrical 5hmC configuration. In addition, it could be due to inefficient digestion with PvuRts1I or a combination of both explanations. In this respect it is important to note that a positive identification of 5hmC in the upstream regulatory region of the *nanog* locus has not

been conclusively demonstrated for ESCs (Ito et al., 2010) and is still highly controversial as another group could not confirm the reduced *nanog* expression upon *tet1* knock down in ESCs (Koh et al., 2011). Taken together, it is uncertain whether the *nanog* locus is modified by hydroxymethylation in ESCs and is therefore not suitable to establish the cut- ligation-amplification strategy to detect PvuRts1I digestion products.

Consequently, we decided to generate substrates with defined amounts of 5hmC to validate the PvuRts1I cut- ligation amplification protocol for the identification of 5hmC sites. We used primers specific for the region III of the nanog promoter to amplify fragments by PCR in the presence of increasing amounts of 5-hydroxymethyl- dCTP (Figure 35). The successful incorporation of proportional levels of 5hmC was confirmed by the ß- glucosylation assay.

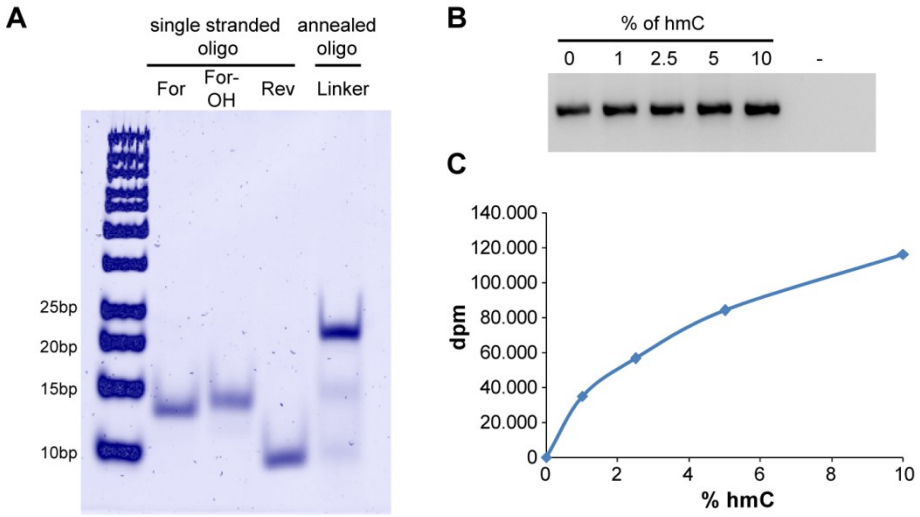
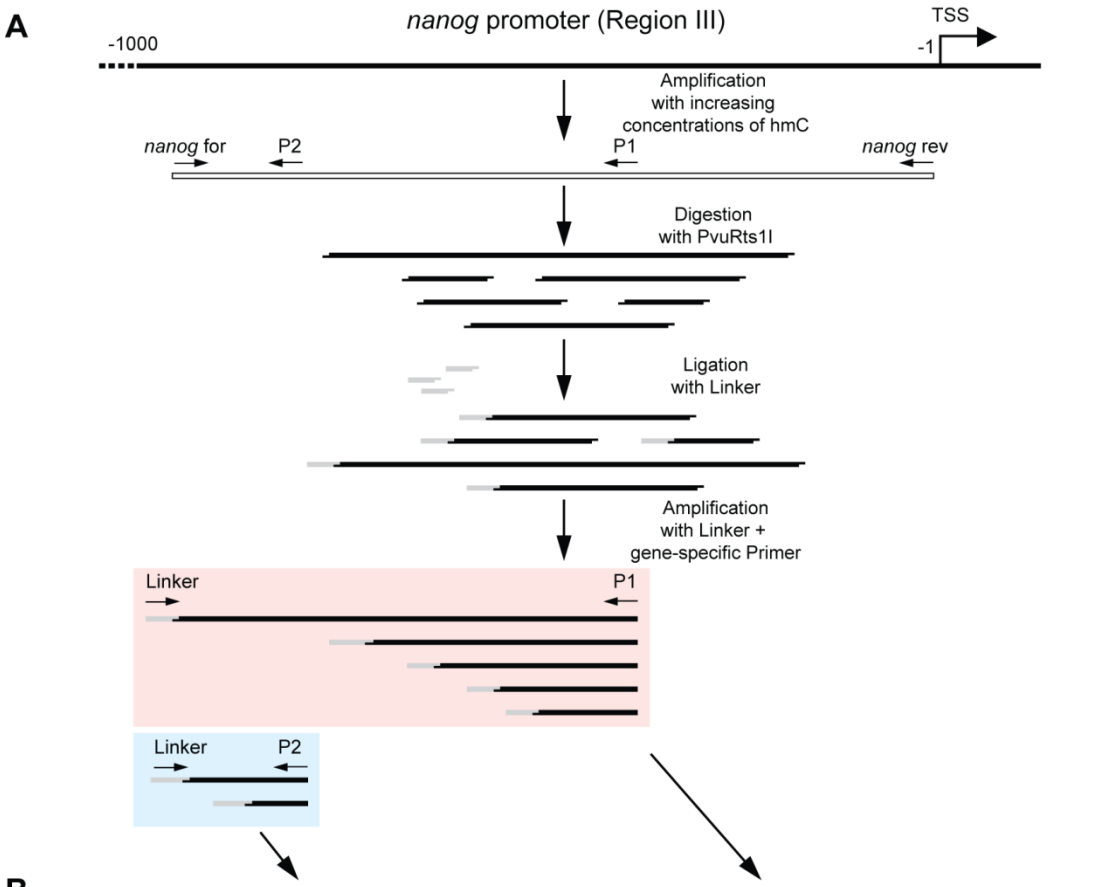


Figure 35. Preparation of linker oligos and substrates with increasing 5hmC concentration.
A) Equal amounts of single stranded forward primer (For), forward primer containing a random two nucleotide 3´overhang (For-OH), reverse primer (Rev) and annealed oligo (Linker) were loaded on a 15 % non-denaturating polyacrylamide gel (PAGE) and stained with SYBR-Green. B) PCR products (867 bp) containing increasing 5hmC levels were generated by amplification of the proximal upstream regulatory region of the *nanog* locus (Region III) and the addition of 5-hydroxymethyl- dCTP and dCTP at appropriate ratios. Minus indicates negative control of the PCR reaction. (C) The random incorporation of 5hmC into the PCR fragments was confirmed by ß-glucosylation assay.

The PCR products with increasing, randomly distributed 5hmC sites were then digested with PvuRts1I, ligated to a linker with random two nucleotide overhangs to match PvuRts1I and ligation products were detected by PCR amplification using two distinct *nanog* specific primers (nanog P1 and P2) each paired with a linker specific primer. PCR products were analyzed on an agarose gel (Figure 35B) and randomly cloned and sequenced. Indeed, we could detect fragments with ends corresponding to the PvuRts1I cleavage pattern however only in products from high 5hmC content (10 %). While fragments containing 1 % 5hmC, which is the highest global 5hmC content to be reported in certain brain tissues, only show background signal (Fig. 36). Our results using the linker/ amplification strategy clearly suggest that a high local concentration of 5hmC facilitates the detection of digestion products by PvuRts1I.



C

Primer/PCR product	Clones analyzed	Consensus sequence		
		$^{hm}CN_{12}/N_{10}G$	$^{hm}CN_{12}/N_9G$	$^{hm}CN_{11}/N_9G$
P1, 1% ^{hm}C	7	0	0	0
P1, 10% ^{hm}C	16	7	3	1*
P2, 10% ^{hm}C	20	14	4	0

Figure 36. Identification of PvuRts1I digestion fragments of substrates with increasing 5hmC level.
A) Outline of cut-ligation-amplification strategy to identify PvuRts1I cleavage sites. After generation of PCR fragments with increasing 5hmC amount, fragments were digested with PvuRts1I and ligated to a linker. Two different nanog specific primers (P1 and P2) were used in combination with the linker specific primer to positively amplify PvuRts1I cleaved sites. B) Agarose gels of obtained PCR fragments indicate the presence of several products after PvuRts1I digestion. The percentage of 5hmC contained in the original substrates and the presence of the linker in the ligation reaction are depicted. NTC: no template control in the PCR reaction. C) Products from the PCR reaction (B) were randomly cloned and sequenced. In the table, the numbers of sequences containing ends corresponding to one of the PvuRts1I cleavage site and the site subtype are summarized. The asterisk indicates a sequence which is reported under two categories because it could not be unambiguously assigned to the consenus site hmCN12/N9G or hmCN11/N9G due to occurrence of consecutive C residues. Data were published in (Szwagierczak et al., 2011).

3.4 Targeted transcriptional activation of silent *oct4* pluripotency genes by combining designer TALEs and inhibition of epigenetic modifiers

The ability to study and engineer biological processes by specifically manipulating the genetic information within a cell has been a long- sought goal for scientists. Some natural occurring DNA- binding proteins including zinc fingers and meganucleases have been engineered to achieve site- specific manipulation of the genome. Especially Zinc fingers fused to transcriptional activator or repressor proteins have been successfully used to carry out site- specific modifications nearby their binding sites (Beerli et al., 2000; Blancafort et al., 2003). However, the design and development of new DNA binding proteins that recognize user- defined target sequences is often difficult and expensive. Intriguingly, recent studies showed that transcription activator- like effector proteins (TALEs) from the plant pathogenic bacteria *Xanthomonas* harbor a DNA binding domain which can be tailored to specifically target new sequences (Boch et al., 2009; Bogdanove and Voytas, 2011). These natural effector proteins are injected into plant cells where the TALE modulates the gene expression of the host genome to contribute to bacterial colonization and survival (Kay et al., 2007; Römer et al., 2007). The central DNA binding domain of the modular TALE proteins is composed of a variable number of tandem repeats of a 34-35 amino acid- sequence motif. Each repeat monomer binds to one base and the various repeats differ from each other only at position 12 and 13. These two residues are also known as the repeat variable diresidues (RVDs) which confer base preference as different RVDs specifically recognize different DNA base pairs (Boch et al., 2009; Moscou and Bogdanove, 2009). The decryption of this TALE code enables the assembly of TALE repeat arrays that target any user-defined DNA sequence (Bogdanove et al., 2010). Indeed, recent data using a variety of designer TALEs (dTALEs) targeting several genes demonstrated that they can specifically modulate transcription from the genome in human cells (Zhang et al., 2011a). However, genes like the pluripotency gene *oct4* could not be activated in the study, raising the question if and how epigenetic modifications might affect the performance of dTALEs.

To shed some light on how the epigenetic environment might affect the ability of TALEs to activate oct4, we generated five different dTALEs each targeting 19 bp sequences within 100 bp upstream of the transcriptional start site of the murine *oct4* promoter. Furthermore, the activation domain (AD) of the Xanthomonas TALE was replaced with the VP-16AD of the herpes simplex. The generation of the dTALEs was carried out by our collaborator Robert Morbitzer from the group of Thomas Lahaye. Reporter assays, Immunflourescence and most of the DNA methylation analysis were done by Sebastian Bultmann. I contributed all expression data and participated in collecting and analyzing DNA methylation patterns.

Using a transient gene reporter assay, we tested the activity of the five dTALEs by co-transfecting an *oct4*- promoter driven eGFP (p*oct4*-eGFP) construct and a constitutively expressed dTALE plasmid. Notably, we found that the efficiency to act as transcriptional activators greatly varied - up to 25 fold difference - between the tested dTALEs on the p*oct4*-eGFP reporter. Next we analyzed the efficiency of dTALE mediated transcriptional activation on methylated p*oct4*-GFP plasmid to investigate whether the epigenetic state of the promoter does influence the efficiency of activation. Interestingly, all tested dTALEs were able to induce eGFP expression albeit to a lower extent when compared to the unmethylated reporter construct. We then used the dTALE (T-83) which achieved the strongest transcriptional activation in both transient reporter assays to activate endogenous oct4 expression in ESCs. To test the activation potential of T-83, we generated ESC lines carrying a stable integrated *oct4* promoter construct regulating the expression of eGFP (ogESCs) and transfected the reporter cell with the dTALE T-83 fused to mCherry. Transfected cells were identified and selected by FACS, the intensity of eGFP fluorescence was measured and endogenous *oct4* mRNA level were analyzed by qPCR (Figure 37).

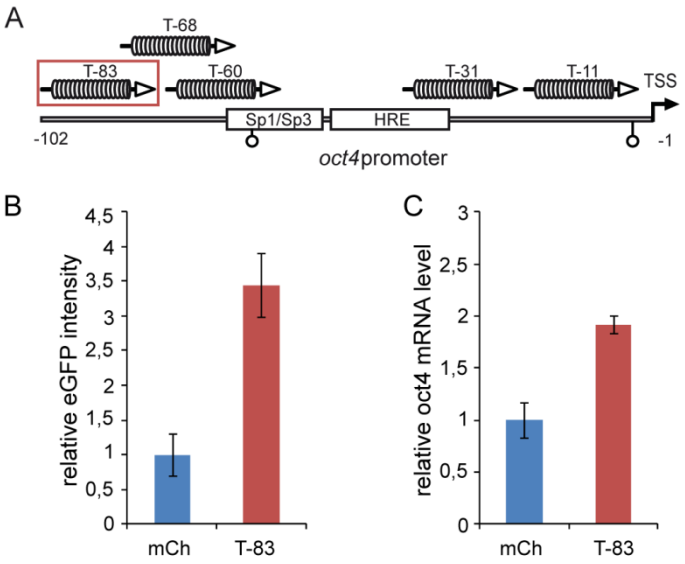


Figure 37. Hyperactivation of endogenous oct4 expression in ESCs by dTALEs.
(A) Schematic overview of the 102 bp long region upstream of the transcriptional start site (TSS) of the *oct4* locus. The binding site of the Sp1/Sp3 transcription factors, the hormone responsive element (HRE) and two CpG sites (open circles) are depicted. The location of the target sequence of the oct4 specific dTALEs is shown and the dTALEs are named according to the distance between the 5´end of their binding sequence and the TSS. The most activating dTALE according to the transient reporter assays, T-83 (red box), was used to hyperactive the endogenous *oct4* locus in ESCs carrying a stable integrated plasmid where expression of eGFP is under control of the upstream regulatory region of *oct4* (ogESCs). After transfection with control (mCherry) vector or T-83 plasmid fused to mCherry, cells were analyzed by FACS and relative eGFP intensity of mCherry positive ogESCs was measured (B). Transfected, mCherry positive cells were single sorted by FACS and endogenous *oct4* transcript levels were analyzed by qPCR (C). Shown are average values and standard deviation from three biological replicates. Data were published in (Bultmann et al., 2012).

Cells transfected with dTALE T-83, but not with control plasmid, showed a 3-4 fold higher mean eGFP fluorescence intensity compared to control treated cells. Furthermore, transcript

levels of the endogenous *oct4* locus were about twofold higher than cells transfected with control plasmid, demonstrating that the dTALEs can hyperactivate endogenous *oct4* expression in ogESCs. The relatively low activation rate of *oct4* is probably due to the high basal *oct4* expression level in ESCs and to the regulation of *oct4* by negative feedback loops on its own promoter (Pan et al., 2006).

To investigate whether dTALEs are able to activate a transcriptionally silent endogenous *oct4* locus, we differentiated ogESCs into neural stem cells (ogNSCs) since the *oct4* promoter is epigenetically silenced during differentiation and *oct4* expression is therefore shut down (Kim et al., 2008b). Consistently, we could neither detect any p*oct4*-eGFP fluorescence nor *oct4* transcripts in untransfected ogNSCs (Fig. 38A,B). Transfection of control or dTALE T-83 could however neither activated the endogenous locus nor the *oct4*eGFP locus since neither *oct4* mRNA nor eGFP fluorescence was measured. Given that the *oct4* promoter acquires repressive histone modifications and DNA methylation upon differentiation, we wondered whether the different epigenetic states of the *oct4* locus in ogESCs and ogNSCs could be the reason for the lack of activation upon dTALE transfection. In ogESCs, *oct4* is actively transcribed and therefore the promoter region might be more accessible to dTALEs, whereas in ogNSCs, *oct4* is silenced and the locus could be in a more condensed and inaccessible conformation for dTALE binding and activation. Hence, we wondered whether inhibition of repressive epigenetic modifications could overcome the barrier for dTALE mediated activation of the transcriptionally silent *oct4* promoter. To test this hypothesis, we applied histone deacteylase (HDAC) inhibitors Trichostatin A (TSA) (Yoshida et al., 1990) or valproic acid (VPA) (Göttlicher et al., 2001) as well as the Dnmt inhibitor 5-aza- 2´- deoxycytidine (5-azadC) (Santi et al., 1983) which interfere with the two major epigenetic silencing mechanisms in mammals (Fig. 38). Twelve hours after transfection with empty or T-83 plasmid, ogNSCs were treated with respective inhibitor or a combination thereof (VPA + 5-azadC) for additional 36 hours. In line with our hypothesis, the treatment of cells with VPA, 5-azadC or a combination of both significantly increased eGFP fluorescence intensity only in cells transfected with dTALE and not with control vector. In line with this, endogenous transcript levels of *oct4* were induced up to 60 % with respect to the levels in ogESCs. In contrast, addition of TSA to transfected cells (control or dTALE) neither enhanced eGFP fluorescence nor endogenous *oct4* mRNA level. A combination of 5-azadC and VPA did not show any additive nor synergistic effects but led to similar activation as the addition of the single epigenetic inhibitors.

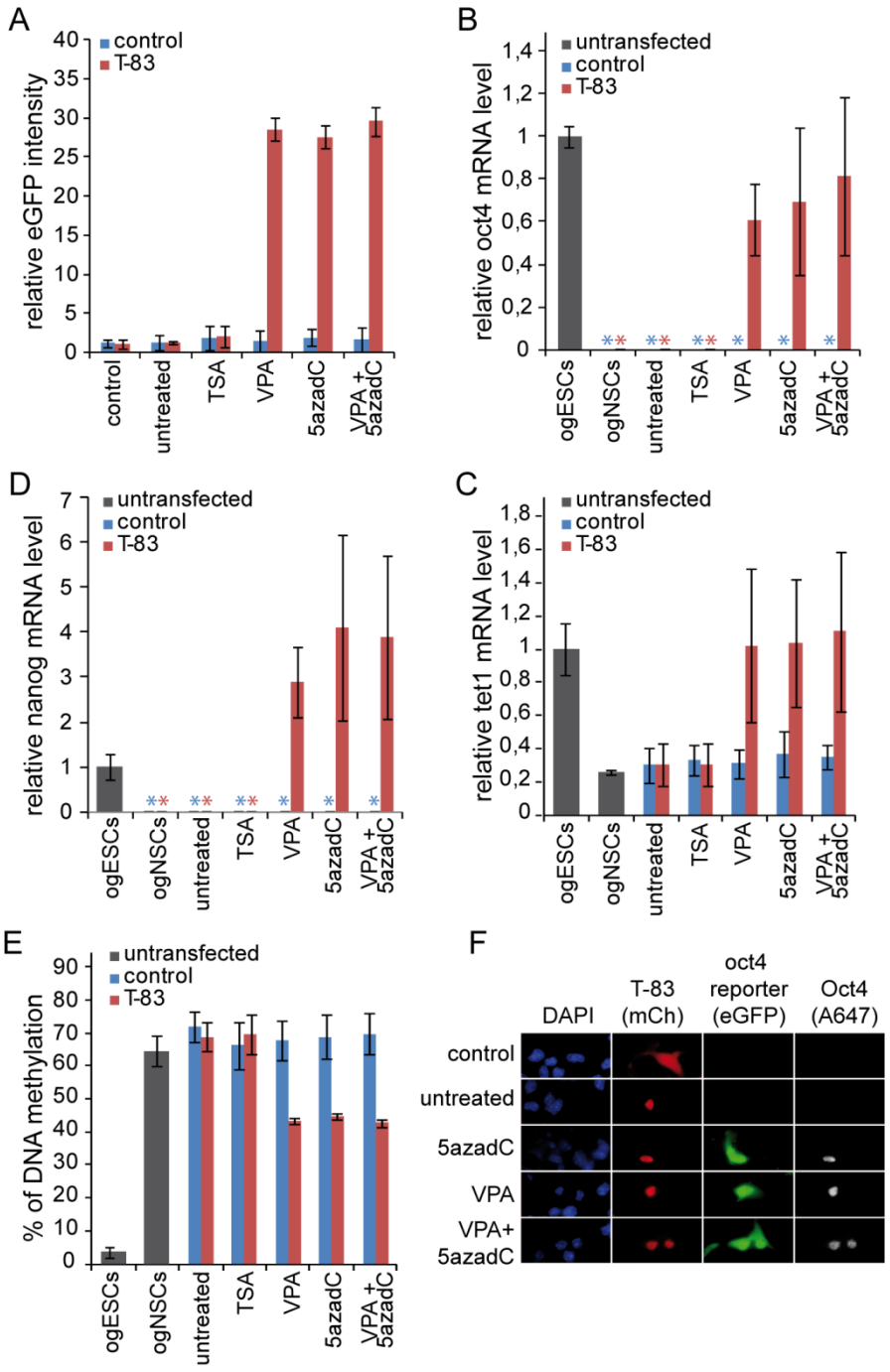


Figure 38. Activation of the silent *oct4* locus in NSCs requires inhibition of repressive epigenetic mechanisms in combination with dTALEs.

NSCs carrying a stable integrated *oct4* promoter plasmid controlling eGFP expression were derived from ESCs, transfected with either control (blue) or T-83 (red) plasmid and treated with various epigenetic inhibitors: TSA (30 nM), VPA (620 µM), 5- azadC (10 nM) or a combination of VPA (310 µM) and 5- azadC (5 nM). (A) Relative eGFP intensities of transfected and treated cells were measured by flow cytometry. After RNA isolation of positively transfected and sorted cells, mRNA levels of *oct4* (B), *nanog* (C) and *tet1* (D) were quantified by qPCR. Untransfected ogESCs and ogNSCs were used as controls and all expression levels were calculated relative to ogESCs (ogESCs set to 1). In the very same samples, DNA methylation levels at the *oct4* promoter were analyzed by bisulfite-treatment followed by pyrosequencing and ogESCs and ogNSCs served as reference (E). Shown is the percentage of DNA methylation measured over 5 CpG sites in the proximal part of the *oct4* promoter. Oct4 protein level was confirmed by antibody staining (Alexa-647) and fluorescence microscopy (F). ogNSCs were transfected with T-83 dTALE and either untreated or treated with 5-azadC (10 nM) and samples were counterstained with DAPI. Transfected cells with T-83 or control plasmid are displayed in the mCherry channel, whereas eGFP fluorescence reflects expression of oct4 reporter construct. Error bars show standard

deviation from two to three biological replicates. Samples where no transcript could be detected by qPCR are marked with asterisks. Data were published in (Bultmann et al., 2012).

Importantly, the fact that the treatment with inhibitors alone did neither induce reporter nor endogenous *oct4* expression, demonstrates that the observed transcriptional activation was the result of the synergistic action of both, dTALE and inhibitors on the *oct4* locus. Furthermore, ogNSCs transfected with dTALE and treated with VPA, 5-azadc or combination of both did not only upregulate *oct4* transcript levels but also Oct4 protein levels (Figure 38F). In addition, also downstream target genes of Oct4, like *nanog* and *tet1*, were specifically activated in these cells, suggesting that the pluripotency network could at least be partially activated via the activation of Oct4 by dTALEs in combination with epigenetic inhibitors (Fig. 38D, C). In contrast, the expression of genes which are not part of the Oct4 regulatory network were not altered upon treatment with epigenetic inhibitors and/ or transfection with dTALE (Figure 39).

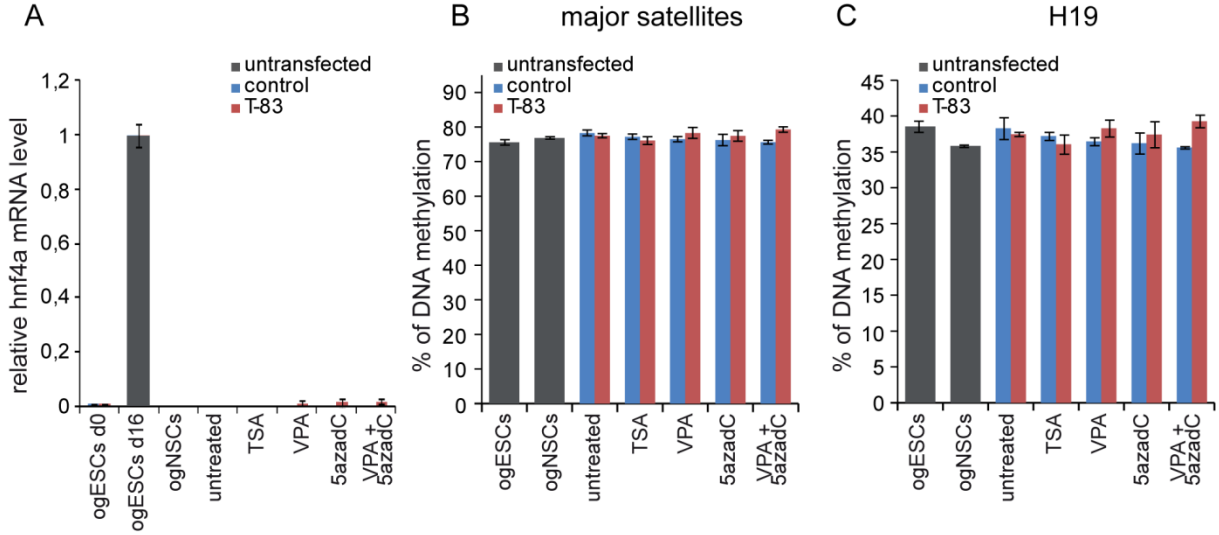


Figure 39. Transcript and DNA methylation levels of control genes and loci, respectively, are not affected by dTALE mediated activation of oct4 promoter or addition of epigenetic inhibitor.
Cells were treated as described in Figure 37. (A) Relative expression of the late endodermal marker gene hnf4a was measured by qPCR in untransfected, control or T-83 transfected cells. As a positive control, 16 day old EBs derived from ogESCs (ogEBs) were analyzed and set as reference (set to 1). DNA methylation levels at major satellite repeats (B) and the imprinted locus h19 (C) were analyzed by bisulfite treatment coupled with pyrosequencing. Average values of 8 (B) and 6 CpG (C) sites are displayed. Error bars represent standard deviation from two to three biological replicates. Data were published in (Bultmann et al., 2012).

As the application of both, 5-azadC and VPA, has been shown to induce DNA demethylation (Santi et al., 1983; Dong et al., 2010), we analyzed the effect of both inhibitors on the DNA methylation level at the *oct4* locus. Strikingly, in samples treated with inhibitors alone and/or transfected with control vector, DNA methylation remained unchanged. In contrast, cells transfected with dTALE T-83 and/ or treatment with VPA and/ or 5-azadC resulted in a remarkable reduction of around 30 % of the original methylation observed at the *oct4* promoter (Figure 38E). It is important to note that reduced DNA methylation levels were not observed on the imprinted *h19* locus or at major satellite repeats upon treatment with

inhibitors in combination with dTALE T-83 transfection (Fig. 39). This clearly demonstrates that the observed effect occurs specifically at the *oct4* promoter and suggests that dTALEs in combination with chemical manipulation of epigenetic modifiers facilitate targeted transcriptional activation of epigenetically silenced target genes.

4. Discussion

4.1 Global DNA hypomethylation prevents consolidation of differentiation programs and allows reversion to the ESC state

4.1.1 DNA methylation is not required for the initial down regulation of pluripotency genes

Numerous studies have analyzed DNA methylation profiles in several cell lines, including ESCs, during ESC differentiation and various developmental stages. However, most studies concentrated on mapping DNA methylation patterns without taking into account that not only DNA methylation as a mark *per se* but also the presence of the Dnmts themselves independent of their catalytical activity might be a crucial factor for the initiation and execution of differentiation programs. Especially expression data from *in vitro* differentiated progeny of hypomethylated ESCs lacking specific Dnmts are very seldom. This is probably due to previous studies reporting very limited survival and/ or proliferation of *dnmt1*$^{-/-}$ and TKO ESCs upon differentiation (Lei et al., 1996; Panning and Jaenisch, 1996; Jackson et al., 2004; Sakaue et al., 2010). To elucidate the role of Dnmts and DNA methylation during differentiation we generated Embryoid Bodies (EBs) from wt, *dnmt1*$^{-/-}$ and TKO ESCs and analyzed their differentiation potential (Fig. 40).

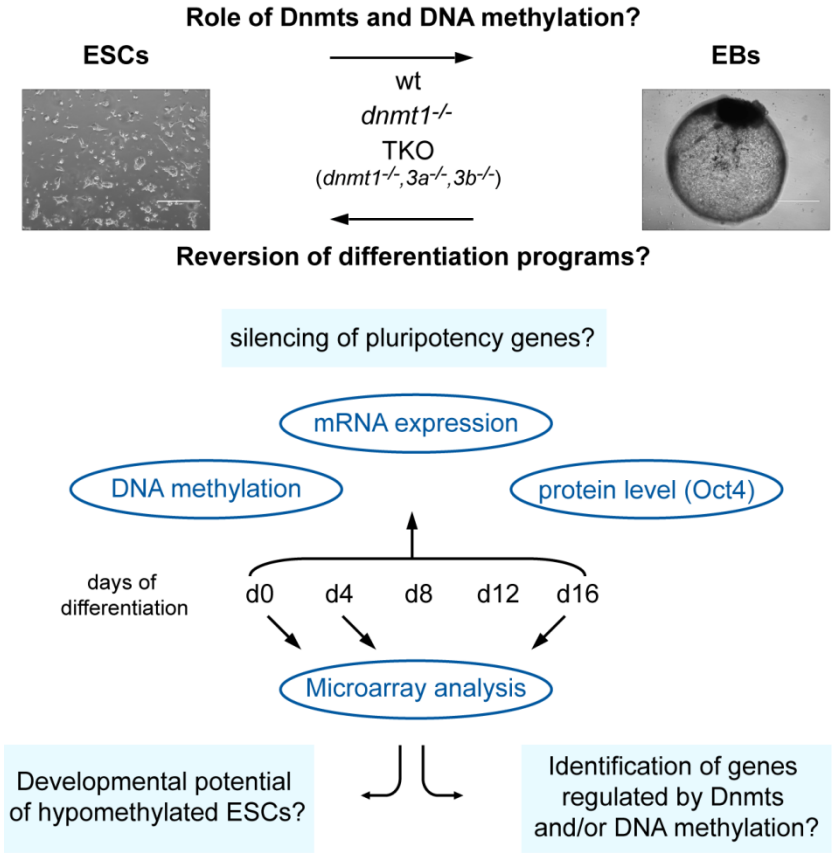


Figure 40. The experimental strategy and key questions of the project are outlined.

The formation of EBs provides an unbiased and undirected differentiation model leading to the generation of a broad range of cell fates and therefore is commonly used to evaluate the differentiation potential of cells. We investigated transcript and DNA methylation levels of selected genes involved in pluripotency and differentiation and also measured protein levels of the master regulator Oct4 during the course of differentiation. Furthermore, we analyzed genome- wide expression changes at three different time points during differentiation to gain insights into the developmental potential of globally hypomethylated cells.

In our study, we find that the expression of the pluripotency master genes*, oct4* and *nanog*, is down regulated in *dnmt1*$^{-/-}$ and TKO EBs, but residual mRNA levels of both genes can even be detected in 16 days old EBs (Fig. 12). However, intracellular protein staining by FACS revealed that all cells uniformly down regulated Oct4 protein level to the same basal levels present in wt EBs after 8 to 12 days of EB culture, indicating that also globally hypomethylated cells did homogenously exit from the ESC state (Fig. 13). This observation is supported by the relatively high concordance of genome- wide expression changes in mutant relative to wt EBs after 4 days of differentiation. In all three cell lines, transcript levels of genes involved in stem cell maintenance and development were down regulated, whereas the expression of genes involved in differentiation processes were up regulated (Fig. 17). Our results also show that the initial down regulation of *oct4* and *nanog* transcript levels occurs independently of DNA methylation as an increase in promoter methylation in wt EBs is only detectable after 6 to 8 days of differentiation (Fig. 12). This is in line with previous studies demonstrating that DNA methylation is not required for initiation of silencing, but necessary to maintain long- term repression (Feldman et al., 2006; Sato et al., 2006; Athanasiadou et al., 2010).

Furthermore, our experiments revealed that silencing of *oct4* occurs in a timely delayed manner exclusively in TKO, but not in *dnmt1*$^{-/-}$ EBs, after 4 days of differentiation and this delay is detectable at transcript as well as protein level. Both mutant EBs are globally hypomethylated, but the major difference between these two knock out EBs is that TKO cells additionally lack both *de novo* Dnmts. The notion, that the delay in *oct4* silencing selectively occurs in TKO EBs very early during differentiation, suggests that the presence of *de novo* Dnmts independent of their catalytical activity plays a role in the initiation of *oct4* down regulation. The process of *oct4* silencing during differentiation has been shown to consist of a multistep cascade including the loss of the nucleosome- depleted regions at the distal enhancer of *oct4*, binding of the transcriptional repressor GCNF, transfer of the repressive histone mark H3K9me3 by the histone methyltransferase G9a and subsequent recruitment of Hp1 which is then followed by *de novo* methylation of the promoter via Dnmt3a and Dnmt3b (Feldman et al., 2006; Sato et al., 2006; Li et al., 2007b; Epsztejn-Litman et al., 2008; You et

al., 2011). Several studies indicated that Dnmt3a and 3b also interact with various epigenetic factors like GCNF and G9a which supposedly leads to the recruitment of the *de novo* Dnmts to the *oct4* promoter (Sato et al., 2006; Epsztejn-Litman et al., 2008). Although DNA methylation seems to function as a secondary epigenetic event in *oct4* silencing, our data indicate that the presence of Dnmt3a and 3b proteins at the *oct4* locus independent from their catalytic activity might contribute to the efficient initiation of *oct4* silencing possibly by supporting the recruitment of repressive histone modifying enzymes. It would be interested to analyze whether a delay in *oct4* down regulation can also be observed in ESCs lacking Dnmt3a and/or Dnm3b and to perform genetic complementation assays of TKO ESCs with catalytical active or inactive Dnmt3a and/ or Dnmt3b proteins. Furthermore, it would be important to examine at what time point during differentiation Dnmt3 proteins bind to the *oct4* locus and to identify epigenetic factors which are possibly recruited by Dnmt3 proteins very early during differentiation before any promoter methylation is detectable at the locus.

4.1.2 *dnmt1*$^{-/-}$ and TKO ESCs show differences in their developmental potential

To analyze the differentiation potential of wt, *dnmt1*$^{-/-}$ and TKO ESCs, we performed global expression analysis in the pluripotent state (d0) and at two time points during differentiation (d4 and d16), which revealed several important results. Firstly, hypomethylated cells are able to initiate differentiation processes as many concordant transcript changes between mutant and wt EBs were detected after 4 days of differentiation. In line with this, the expression of genes associated with pluripotency was down regulated and transcripts involved in lineage selection and developmental processes were up regulated in all three cell lines independently of the genotype. These results indicate that DNA methylation is dispensable for the activation of differentiation programs. Secondly, although both mutant cells are globally hypomethylated, *dnmt1*$^{-/-}$ and TKO EBs show significant differences in their ability to execute differentiation programs. This idea is supported by the observation that the expression profiles of TKO EBs show a high degree of divergence from those in wt EBs after 16 days of EB culture and the few genes could only be grouped into GO categories involved in metabolic processes. In contrast, we still detect many concordant transcript changes in *dnmt1*$^{-/-}$ and wt EBs at day 16 of differentiation. Most of these commonly expressed genes were related to developmental processes including cell differentiation and proliferation as well as organ development. These data clearly point to a previously unappreciated progression of transcription programs in differentiated *dnmt1*$^{-/-}$ cells.

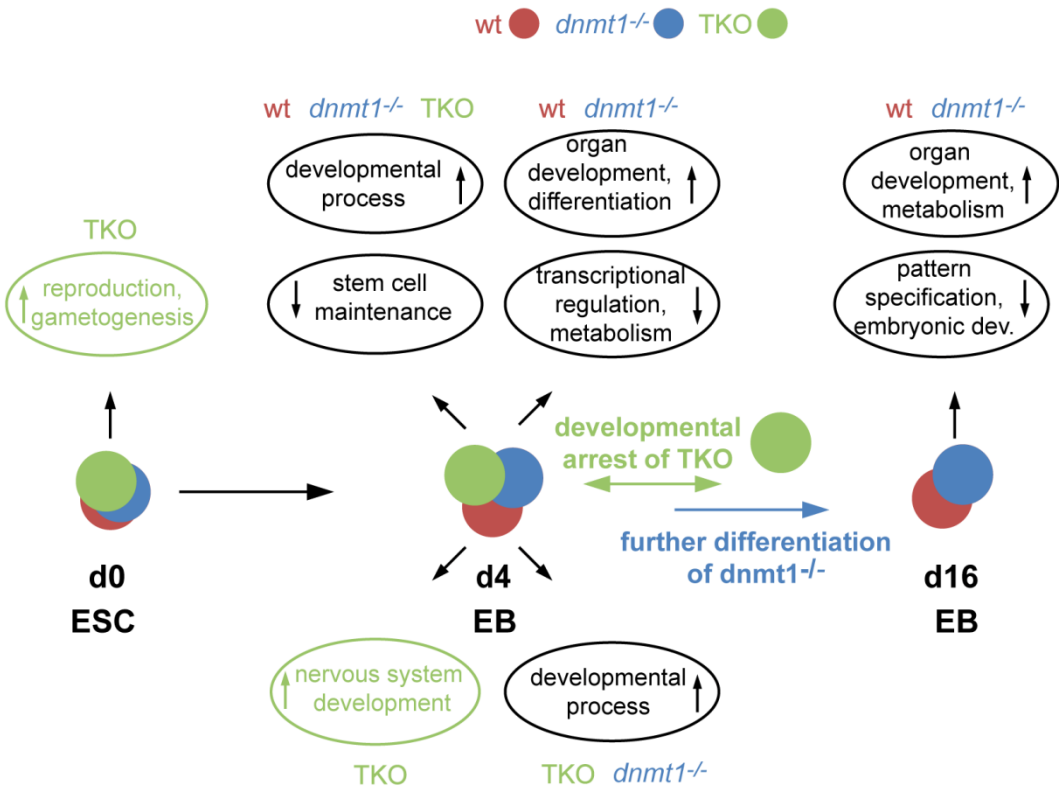


Figure 41. Summary of gene changes and corresponding GO categories occurring in wt (red), *dnmt1*$^{-/-}$ (blue) and TKO (green) ESCs and EBs identified by Microarray analysis.
In the undifferentiated state, all cell lines show very similar expression profiles, however, during differentiation, expression patterns progressively diverge, especially in the case of TKO EBs.

However, our analysis also implies a substantial limited differentiation potential in TKO and, to a lower extent, in *dnmt1*$^{-/-}$ cells. In this regard it is important to note that while TKO ESCs are virtual devoid of DNA methylation, *dnmt1*$^{-/-}$ ESCs contain about 20 % residual genomic methylation, although the methylation is mainly restricted to repetitive sequences (Lei et al., 1996; Liang et al., 2002; Tsumura et al., 2006). Nonetheless, examination of the expression profiles of both hypomethylated cells clearly shows a different response of TKO and *dnmt1*$^{-/-}$ cells to differentiation conditions, suggesting that the presence of Dnmt3 proteins can partially compensate for the loss of Dnmt1. In line with this we detected higher levels of *dnmt3b*, but slightly lower levels of *dnmt3a* transcripts in *dnmt1*$^{-/-}$ EBs compared to wt EBs (Fig. 22). Interestingly, it has been shown that the activity of the *dnmt3b* promoter is regulated by DNA methylation (Nimura et al., 2006). Hence, the reduced global DNA methylation levels in *dnmt1*$^{-/-}$ EBs could contribute to the higher expression of *dnmt3b* and possibly compensates for the lower *dnmt3a* transcript levels observed in mutant EBs.

Besides a possible compensatory role of Dnmt3 proteins in *dnmt1*$^{-/-}$ EBs, it could also be that the *de novo* Dnmts fulfill functions in transcriptional regulation independent of their catalytical activity like we have seen e.g. in the silencing of the bivalent genes *fgf5* and *brachyury* (Fig. 25), although here it will be important to confirm binding of Dnmt3 proteins at those promoters. The notion that Dnmts can mediate transcriptional repression independent of their

catalytical domain has been demonstrated for Dnmt1. Among the numerous Dnmt1-interacting proteins, also several histone deacteylases have been identified (see also chapter 1.2.1) and it is believed that Dnmt1 recruits these repressive chromatin modifying enzymes to mediate gene silencing independently of DNA methylation (Fuks et al., 2000; Robertson et al., 2000a; Rountree et al., 2000b). A recent report suggests that also Dnmt3b could fulfill regulatory functions aside from its DNA methylation activity. More specifically, the study by Martins-Taylor *et al*. shows that knock down of *dnmt3b* during neural differentiation alters the timing of differentiation and lineage choice. Interestingly, affected genes were not targets of DNA methylation but the proximal promoters of these lineage genes were directly bound by Dnmt3b. The finding that upon *dnmt3b* knock down, also the repressive histone mark H3K37me3 and binding of EZH2, the H3K27 methyltransferase component of the Polycomb repressive complex 2 (PRC2), at deregulated genes was reduced compared to control treated cells, implies that Dnmt3b might play an important role in the recruitment of EZH2 and/ or in maintaining EZH2 binding at these promoters (Martins-Taylor et al., 2012). In this context it is interesting to note that up regulated genes in TKO EBs were predominately involved in neural fate specification, implying that DNA methylation and/ or Dnmts play a crucial role in neural development. In general, further studies comparing global binding patterns of Dnmts to DNA methylation profiles would be necessary to shed more light on the role of DNA methylation and Dnmts in controlling transcription programs during development. The analysis of global binding profiles of Dnmt3s and DNA methylation maps could be used to identify targets of Dnmts independent of their catalytical activity. Conversely, the knockout cell lines could be complemented with catalytical mutants of Dnmts to distinguish catalytical dependent and independent targets of the methyltransferases. In addition, as our results suggest that the presence of Dnmt3s in *dnmt1*$^{-/-}$ could contribute to their milder phenotype, it would be crucial to analyze the transcription profiles and developmental potential of ESCs lacking Dnmt3a and/or Dnmt3b. A comparison of these results with the data from *dnmt1*$^{-/-}$ and TKO ESCs and EBs would shed light on how the various Dnmt proteins contribute to transcriptional control during differentiation.

Key transcription factor genes for cell fate choice and lineage commitment are known to carry bivalent chromatin domains, which are mainly resolved after differentiation initiation, either by loss of H3K27me3 for transcriptional activation or loss of H3K4me3 or both H3 marks for gene silencing (Bernstein et al., 2006; Mikkelsen et al., 2007). In the latter case, the loss of these histone marks is believed to be accompanied by gain of DNA methylation for permanently sealing transcription, as has been proposed for the differentiation into the neural lineage (Mohn et al., 2008). We hypothesized that DNA methylation might represent a general mechanism for the final silencing of bivalent genes and expected that most bivalent genes would be deregulated in *dnmt1*$^{-/-}$ and/ or TKO EBs. However, our analysis revealed

that only selected bivalent genes involved in early neuroectodermal differentiation like *nestin* and *sox1* did gain DNA methylation in a subset of differentiated wt cells and were deregulated in the absence of Dnmts. Hence, our results support the idea that *de novo* methylation represents the long- term silencing mechanism for selected bivalent genes in specific lineages and does not function as a general mechanism for the repression of bivalent genes during differentiation.

4.1.3 Parallels and crosstalk between the two major repressive pathways– DNA methylation and Polycomb repressive system

Intriguingly, a comparison of the phenotypes observed in ESCs devoid of functional Polycomb repressive complexes (*ring1b*$^{-/-}$ *eed*$^{-/-}$ double knock out (DKO) ESCs) show striking similarities to the phenotypes detected in ESCs lacking all three major Dnmts. Analogous to Dnmt TKO ESCs, Polycomb DKO ESCs are able to self- renew in culture, but are severely impaired in their ability to differentiate. More specifically, *in vitro* differentiation to EBs revealed that the Polycomb double mutant cells formed EBs, although with reduced size compared to wt or single mutant EBs and up regulated differentiation markers and down regulated the expression of pluripotency genes. Furthermore, Polycomb DKO cells injected into blastocysts were able to contribute to the inner cell mass (ICM), but failed to contribute to lineages at later developmental stages (E10.5) in chimera embryos (Leeb et al., 2010). Hence, similar to Dnmt TKO ESCs, Polycomb DKO ESCs are able to activate differentiation programs but fail to proceed with differentiation (Leeb et al., 2010; Sakaue et al., 2010). ESCs lacking either PRC1 or PRC2 show no defect in differentiation and can contribute to cells of all three lineages (Leeb et al., 2010). In line with this, we detect a milder phenotype of ESCs lacking solely Dnmt1 since also at later differentiation stages, many concordant expression changes between wt and *dnmt1*$^{-/-}$ EBs could be detected.

Another parallel between the two repressive systems is the higher number of derepressed genes upon simultaneous knock out of PRC1/PRC2 ESCs and TKO ESCs compared to cells lacking only one component. In both cases, complete ablation of all functional components results in approximately twice as many deregulated genes as in single knockouts of ESCs and in the case of the Polycomb DKO, also genomic repeats like endogenous retroviral elements (ERVs) are derepressed (Leeb et al., 2010). The finding that the majority of Polycomb targets are repetitive sequences suggests a potential role of repeat sequences in Polycomb mediated gene silencing, probably by serving as a binding platform for PcG proteins. Indeed, several studies indicate that genomic repeats play central roles in regulating transcription often via the transcription of long non- coding RNAs or miRNAs from these repetitive sequences, which subsequently interfere with expression (Faulkner et al., 2009; Kaneko et al., 2011; Cabianca et al., 2012). For instance, recent data demonstrate that

a long noncoding RNA generated from the repeat element D4Z4 of patients suffering from a specific form of muscular dystrophy (Facioscapulohumeral muscular dystrophy, FSMD) is involved in the progression of the disease. In healthy humans, D4Z4 repeats are silenced by Polycomb group proteins, however with the onset of the disease, these repeats are progressively lost, leading to reduced Polycomb binding. The insufficient Polycomb protein binding allows the production of a long non- coding RNA at these repeats which recruits an activating Trithorax complex that promotes further derepression of the repetitive arrays (Cabianca et al., 2012). Interestingly, in our study we find that *dnmt1*$^{-/-}$ ESCs possess a greater differentiation potential compared to TKO ESCs, although both cells are severely globally hypomethylated. However, *dnmt1*$^{-/-}$ ESCs (and possibly EBs thereof) harbor a substantial residual methylation of repetitive sequences (Lei et al., 1996; Biniszkiewicz et al., 2002; Chen et al., 2003) which could - similar to PcG mediated repression - serve as a platform for gene silencing during differentiation. It will be crucial to determine methylation levels at repetitive sequences in *dnmt1*$^{-/-}$ EBs to exclude the possibility that the milder phenotype observed in these cells is due to higher methylation at single copy genes and/ or repetitive elements. However, this seems unlikely as our methylation analysis at e.g. the *oct4* promoter in *dnmt1*$^{-/-}$ EBs during differentiation shows an inital slight increase in DNA methylation, but the levels drop dramatically during further differentiation (Fig. 12). In addition, no remarkable increase in promoter methylation at the *nanog* locus in *dnmt1*$^{-/-}$ EBs could be detected at any time point during EB formation. Therefore, it is unlikely that a transient increase in DNA methylation at selected genes could lead to the high number of concordant expression changes between *dnmt1*$^{-/-}$ and wt EBs after 16 days of EB culture. In conclusion, the residual methylation at repetitive sequences and/or the presence of Dnmt3 proteins could account for the greater differentiation potential of *dnmt1*$^{-/-}$ ESCs compared to TKO ESCs.

Taken together, there seem to be striking parallels between the roles and possible modes of action of the two major repressive systems, DNA methylation and Polycomb system, in contributing to cell fate choice and developmental potential. Moreover, there also seems to be a specific crosstalk between both systems since e.g. ESCs with simultaneous ablation of PRC1 and PRC2 show not only reduced levels in H3K27me3, but also partial loss of DNA methylation at repetitive sequences (Leeb et al., 2010). This is consistent with a previous study reporting that components of PRC2 are involved in the recruitment of Dnmts and subsequent methylation of Polycomb target genes (Viré et al., 2006). Additionally, it has been shown that 93 % of all bivalent promoters in ESCs contain CpG islands and that genes repressed by PcG frequently become targets of DNA methylation during neuronal differentiation, indicating a lineage- specific crosstalk between the two repressive modifications (Mohn et al., 2008). In conclusion, both repressive pathways do not only show

parallels in their functions and actions, but also specific crosstalk between DNA methylation and Polycomb repressive systems can be observed dependent on the genomic context as well as cell lineage and developmental stage.

4.1. 4 Improved reprogramming by transient, simultaneous inactivation of Dnmt1 and p53?

Our experiments reveal that globally hypomethylated cells isolated from 12 days old EBs can fully and rapidly revert to the undifferentiated state when cultured in pluripotency promoting conditions. This is especially surprising for replated *dnmt1*$^{-/-}$ cells as these cells show a higher degree of differentiation compared to TKO cells. The observed reversion to the undifferentiated ESC state is unlikely the result of a subpopulation of undifferentiated cells within *dnmt1*$^{-/-}$ and TKO EBs, as their cells homogenously express the same basal levels of Oct4 protein as differentiated cells from wt EBs, which do not show any response to replating in medium containing LIF. Additionally, the fast kinetics of re- increasing the expression of pluripotency genes and concomitant reduction of lineage marker transcripts argues against a non- differentiating subpopulation as the reversion occurs within three days after replating exclusively in hypomethylated cells but not in wt cells in the presence of LIF. However, the observation that *oct4* and *nanog* mRNA level are not completely silencing in both hypomethylated cells, underscores that DNA methylation is crucial for complete and permanent repression of their transcription and thus enforces canalization of cell fate choice and lineage commitment upon differentiation.

The fact that cells lacking DNA methylation are not able to stably maintain cell fate is in line with reports showing that global inhibition of Dnmt activity e.g. by the addition of the Dnmt inhibitor 5-aza-deoxycytidine, 5-azadC, accelerates the rate of reprogramming of differentiated somatic cells to pluripotency (Mikkelsen et al., 2008; Shi et al., 2008). This facilitation of somatic cellular reprogramming is likely due to the more efficient and complete demethylation of pluripotency genes. However, the use of small molecules like 5-azadC for cellular reprogramming also harbors the risk of forming covalent and potentially mutagenic Dnmt- DNA adducts. Our finding that cells lacking Dnmt1 can efficiently revert to the ESC state, suggests that transient and specific inhibition of Dnmt1 activity in combination with pluripotency promoting conditions might be sufficient to facilitate reprogramming to the pluripotent state. Based on the recently described crystal structures of Dnmt1 (Song et al., 2011, 2012; Syeda et al., 2011; Takeshita et al., 2011), it will be easier to design more specific and less toxic compounds for Dnmt1 inhibition which could be used instead of unspecific inhibitors like 5-azadC.

Transient inhibition of Dnmt1 activity could be combined with functional inactivation of p53, which has been shown to increase the efficiency of iPS derivation by overcoming proliferative senescence of somatic cells (Hong et al., 2009; Kawamura et al., 2009; Li et al., 2009; Marion et al., 2009; Utikal et al., 2009). However, as p53 is an important guardian of chromosome stability and integrity, transient inactivation of p53 by either RNAi or a p53-peptide inhibitor, which forms a hetero- tetramer with the endogenous p53 and leads to its reduced transcriptional activity (Wada et al., 2012), would be safer ways to overcome the replicative senescence. In addition to the latter effect, transient p53 inactivation would also support a rapid passive demethylation by high proliferation rates. Furthermore, the silencing of p53 could prevent the death of not yet dedifferentiated cells which would probably die upon Dnmt1 inhibition and subsequent global demethylation, phenotypes which are observed upon genetic ablation of Dnmt1 in fibroblasts (Jackson-Grusby et al., 2001) .

4.2 Uhrf proteins link the two major repressive epigenetic pathways

It is now well accepted that both DNA methylation and histone modifications are involved in the epigenetic control of gene expression during development. Although both modifications are established by a completely different set of enzymes, increasing evidence suggests a high interconnectivity between these two epigenetic systems (see also chapter 4.1.3). In addition to Polycomb- mediated H3K27me3, also the repressive histone mark H3K9me3 has been linked to DNA methylation, since e.g. the loss of the histone methyltransferase G9a not only leads to reduced H3K9me3 levels, but also decreased DNA methylation levels (Feldman et al., 2006). However how this epigenetic crosstalk is mediated and translated into defined chromatin states within the cell is still poorly understood. The discovery of the multi- domain Uhrf protein family (see chapter 1.2.2.), has shed some light on how the two repressive epigenetic pathways are connected with each other. The two members of the Uhrf protein family harbor a conserved multi- domain structure and were shown to behave biochemically very similar. Both proteins can bind to hemi- methylated DNA as well as to methylated H3K9 and co- immunoprecipitation studies revealed that Uhrf1 and Uhrf2 interact with Dnmt1, Dnmt3a, Dnmt3b and G9a, suggesting functional redundancy of both proteins (Bostick et al., 2007; Sharif et al., 2007; Rottach et al., 2010; Pichler et al., 2011; Zhang et al., 2011b). Whereas Uhrf1 has been shown to function as an essential co- factor in maintaining DNA methylation patterns (Bostick et al., 2007; Sharif et al., 2007), it remains elusive whether Uhrf2 also plays a role in DNA methylation and/ or has additional functions in other biological contexts. So far, Uhrf2 has been implicated in cell cycle regulation (Mori et al., 2002, 2004; Li et al., 2004) and in the intranuclear degradation of polyglutamine aggregates via its E3 ubiquitin ligase activity (Iwata et al., 2009).

4.2.1 Uhrf1 and Uhrf2 show no functional redundancy

To gain more insights into the function of Uhrf2, we analyzed transcript levels of *uhrf1* and *uhrf2* in ESCs, ESC differentiation to Embryoid Bodies (EBs), somatic cells and adult mouse tissues. Interestingly, we found that the expression profile of *uhrf1* and *uhrf2* show striking differences. In undifferentiated ESCs, *uhrf1* is predominately expressed, but the transcript levels progressively decrease during differentiation of ESCs to EBs (Fig. 27), confirming previous data showing that *uhrf1* is highly expressed in proliferating cells (Muto et al., 1995; Fujimori et al., 1998). On the contrary, *uhrf2* mRNA level increase with prolonged EB culture, indicating a time- and developmental switch in *uhrf1* and *uhrf2* expression. In line with this, *uhrf2* transcript levels were prevalent in differentiated adult mouse tissues and very little if any expression of *uhrf1* could be detected. Furthermore, we found that both proteins behaved differently during serum- starvation of fibroblasts. Whereas *uhrf1*, like *dnmt1,* is proliferation- dependent regulated, *uhrf2* is up regulated in serum- starved, quiescent

fibroblasts (Fig. 29). Interestingly, this opposite expression pattern of *uhrf1* and *uhrf2* dependent on the proliferative state of the cell is in line with previous studies analyzing the expression pattern of both genes in primordial germ cells (PGCs). Shortly after PGC specification around E.6.5 days, these cells arrest in the G2 phase of the cell cycle and become quiescent. During this time, *uhrf1*, *dnmt1* as well as *dnmt3b* transcription dramatically decreases, whereas *uhrf2* has been shown to be specifically upregulated, at least in *in vitro* PGC- like cells. However, this high expression of *uhrf2* is maintained despite re- entry of the cells in an active proliferative state during PGC migration, together with a re-increase of *dnmt1* expression (Kurimoto et al., 2008). Since migrating PGCs still contain a relatively high DNA methylation level although the essential co- factor Uhrf1 is absent, it has been speculated that Uhrf2 might compensate for the loss of Uhrf1 and contributes to the maintenance of DNA methylation until the PGCs enter the genital ride around E.10.5 days, where they become fully reprogrammed, including genome- wide erasure of DNA methylation patterns (reviewed in Hackett et al., 2012).

However, it is still unclear whether Uhrf2 plays a role in maintenance methylation. Our knock-down studies of *uhrf2* in wildtype, *uhrf1*$^{-/-}$ as well as heterozygous *uhrf2*$^{-/+}$ ESCs did not reveal any difference in DNA methylation levels on repetitive sequences and single copy genes upon loss of *uhrf2* (Fig. 30 and 31). Consistent with this, a recently published study on human UHRF2 could also not detect an effect on global DNA methylation levels upon knock down of *UHRF2* (Zhang et al., 2011b). Furthermore, genetic complementation of *uhrf1*$^{-/-}$ ESCs with Uhrf2 did not lead to restoration of DNA methylation levels, indicating that Uhrf1 and Uhrf2 are not functional redundant (Pichler et al., 2011; Zhang et al., 2011b). In line with this, it has been shown that, unlike Uhrf1, Uhrf2 fails to recruit Dnmt1 to replication foci during the S- phase of the cell cycle, hence providing a possible explanation for the inability of Uhrf2 to rescue DNA methylation patterns in genetic complementation assays. Additionally, this result underscores the S- phase dependent interaction between Dnmt1 and Uhrf1 as a crucial regulatory mechanism for maintenance methylation (Zhang et al., 2011b).

4.2.2 What is the function of Uhrf2?

Nonetheless, it has been shown that Uhrf2, like Uhrf1, can interact with Dnmts at least *in vitro*. The functional consequences of these interactions are still unknown. One possibility is that Uhrf2 needs additional proteins that mediate the interaction to Dnmts and which are not present in proliferating cells, where high levels of Uhrf1 are sufficient to promote the maintenance of DNA methylation. This would lead to the hypothesis that Uhrf2 might only be involved in maintenance methylation when Uhrf1 is absent e.g. in non- proliferating, differentiated cells or in specialized cells like PGCs. In support of this, it has been shown that

Uhrf2, in contrast to Uhrf1, only preferentially binds to hemi-methylated DNA, the substrate of maintenance methylation, when simultaneously bound to the repressive heterochromatin mark H3K9me3 (Pichler et al., 2011). This cooperative binding of Uhrf2 might play a role in the tighter control of gene silencing in differentiated cells, whereas Uhrf1 might mediate a less stringent control in undifferentiated cells to allow the cellular plasticity and open chromatin conformation of ESCs. To understand the functional role of Uhrf2 it will be necessary to generate *uhrf2*$^{-/-}$ knock-out mice and ESCs and analyze DNA methylation patterns in the undifferentiated state as well as during ESC differentiation. Furthermore, the recently established culture system which reconstitutes the PGC specification pathway (Hayashi et al., 2011) would be an attractive model to analyze the role of Uhrf2 in maintaining DNA methylation patterns during germ cell development.

4.3 Role of 5hmC and Tets during development

4.3.1 Novel methods to quantify and map 5hmC

DNA methylation has always been considered a quite stable epigenetic modification, which once it is established, cannot be actively removed. The discovery of the "6th base" of the genome, 5- hydroxymethylcytosine (5hmC), together with the identification of the family of Tet proteins, changed this long believed paradigm and scientists started to focus on the biological function of this newly identified modification. To gain insights into the functional role of 5hmC, a first challenge is to develop new methods which can selectively detect 5hmC and discriminate it from the more abundant and structurally very similar 5mC (see also chapter 3.3) and hence allow the measurement of 5hmC levels in various cell lines and tissues.

To this aim we developed a novel method to quantify global 5hmC levels in genomic DNA. We sought to exploit the ß- glucosyltransferase of T4 bacteriophages, an enzyme known to modify 5hmC and that evolved as a defense mechanism in the struggle between prokaryotes and their viruses. Using radiolabeled glucose, we showed that 5hmC can be specifically labeled and by generating reference fragments with known, but varying 5hmC content, we verified that the incorporation of isotopically labeled glucose in genomic DNA occurs linear within a range of 0.25 %- 2 % 5hmC content. Thus, global 5hmC level can be specifically labeled and accurately quantified by comparison to a standard curve, which was measured in each assay (Szwagierczak et al., 2010).

First, we applied this assay to two different wildtype ESCs as well as during their *in vitro* differentiation to Embryoid Bodies (EBs) and also measured transcript levels of *tet1-3* in the very same samples (Figure 33). We found that ESCs contain relatively high 5hmC levels (0.3 % 5hmC relative to total cytosine), which drastically drop during differentiation to EBs. Interestingly, *tet1* transcript levels were prevalent in the undifferentiated state, but decreased during EB formation, suggesting that Tet1 is the main enzyme responsible for the generation of 5hmC in ESCs (Szwagierczak et al., 2010). These data are consistent with previous publications showing that *tet1* is predominately expressed in ESCs, but declines during monolayer differentiation of ESCs upon removal of LIF (Tahiliani et al., 2009). Furthermore, increasing evidence points to a role of Tet1 and 5hmC in the regulation of pluripotency and developmental potential. Knock-down of *tet1* in ESCs results not only in reduced 5hmC level with a concomitant increase in 5mC at selected loci, but also in the deregulation of pluripotency associated genes (Ito et al., 2010; Freudenberg et al., 2011). In line with this, it has been shown that *tet1* depletion in pre- implantation embryos and ESCs leads to loss of pluripotency and skewed differentiation towards the trophoectodermal lineage (Ito et al.,

2010; Koh et al., 2011). The finding that *tet1* is directly regulated by the master regulator Oct4 also integrates Tet1 and 5hmC in the pluripotency network (Koh et al., 2011). However, several studies could not confirm the deregulation of pluripotency genes upon *tet1* knock down in ESCs and found only modest reduction of global 5hmC level with a minor increase of 5mC (Koh et al., 2011; Williams et al., 2011). Additionally, genetic ablation of *tet1* revealed viable and fertile mice (Dawlaty et al., 2011), raising uncertainty about the importance of Tet1 in ESC maintenance. These contradictory results emphasize the need for further experiments including the generation of inducible, conditional knock out mice, which would allow a more specific approach to investigate the function of Tet1 e.g. in certain tissues or at defined time points during development.

In contrast to the expression profile of *tet1*, we found that *tet3* mRNA levels were very low in ESCs, but increased with differentiation and prolonged EB culture (Fig. 33). *Tet2* transcript levels dropped during the first 4 days of EB differentiation, but the mRNA levels recovered to the levels initially found in ESCs after 4 more days of EB culture. Remarkably, the initial reduction of 5hmC after 4 days of differentiation was followed by a re- increase of global 5hmC level in 8 days old EBs. Therefore, our results suggest that the relatively high abundance of 5hmC in undifferentiated ESCs correlates with high expression levels of *tet1* and to a lower extent, *tet2*. The partial recovery of genomic 5hmC in 8 days old EBs correlates with higher *tet2* and *tet3* transcript levels (Szwagierczak et al., 2010). Interestingly, the distinct expression profiles of *tet1-3* are in line with a study showing that during reprogramming of fibroblasts, the initial high levels of *tet3* transcripts substantially decrease during this process, whereas *tet1* and *tet2* transcript levels as well as global 5hmC content concomitantly increase (Koh et al., 2011).

We next analyzed genomic 5hmC as well as *tet1-3* transcript levels in several adult mouse tissues (Fig. 33). In line with previous reports (Kriaucionis and Heintz, 2009), 5hmC was the most abundant in brain tissues which correlated with high levels of *tet3* and lower levels of *tet2*. In general, we found that all analyzed tissues typically contained high levels of *tet3* but low levels of *tet1*, whereas undifferentiated ESCs are characterized by the exactly opposite expression pattern. However, kidney seems to be an exemption as we measured relatively high level of genomic 5hmC together with high *tet2* levels (Szwagierczak et al., 2010). The prevalent expression of *tet2* in kidney is consistent with reports showing that one of the phenotypes described in *tet2*$^{-/-}$ mice is a cellular defect in proximal convoluted tubules of the kidney (Tang et al., 2008).

In conclusion, our analysis revealed that genomic 5hmC can be detected not only in various brain regions, but also in other analyzed tissues like kidney and liver. Furthermore, we found

that different amounts of global 5hmC level correlate with differential expression of *tet1-3* genes (Figure 42).

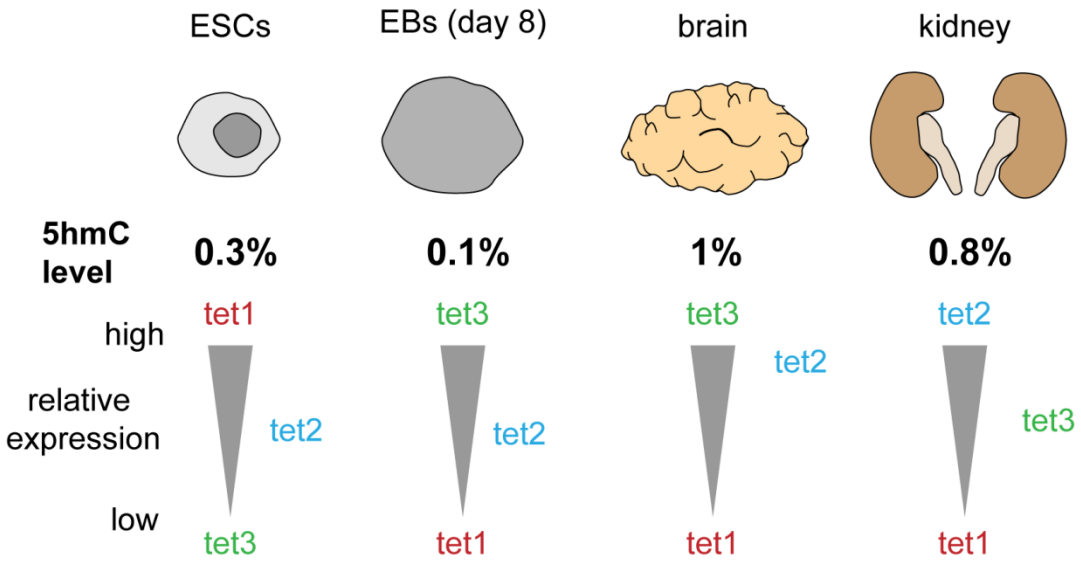


Figure 42. Abundance of 5hmC in ESCs, EBs and various tissues correlates with the differential expression of *tet1-3* genes.

To elucidate the biological significance of 5hmC in mammalian genomes, it is crucial to determine the distribution of the novel modification in genomic DNA: For this purpose we again exploited an enzyme from bacteria that was evolved as a strategy to counter the phage´s measures. The endonuclease PvuRts1I has been shown to cleave glucosylated 5hmC and *in vivo* studies found that T- even phages containing exclusively genomic 5hmC, but not T- odd phages which contain 5mC or unmodified C, were selectively limited in their growth by the presence of a plasmid encoding PvuRts1I (see also chapter 3.3.2). We wondered whether PvuRts1I could be used as a tool to discriminate 5hmC from 5mC and unmodified C and therefore purified the enzyme and tested its activity *in vitro* (Szwagierczak et al., 2011).

Our analysis demonstrates that PvuRtsI1 selectively cleaves 5hmC containing DNA and revealed the consensus sequence $^{hm}CN_{11\text{-}12}/N_{9\text{-}10}G$ with a 2 nucleotide 3´-overhang as the cleavage site of PvuRtsI1. We then wanted to use PvuRtsI1 to map 5hmC pattern in genomic DNA and, based on previous report, chose the upstream regulatory region of *nanog* as a potential region containing 5hmC (Ito et al., 2010). However, our attempt to measure the decrease of product after PvuRtsI1 digestion compared to mock digested samples did not reveal any difference in products between the two different samples (Fig. 34). Also, our devised strategy to positively identify rare digestion of PvuRtsI1 products by ligating linkers with random 2 nucleotide overhang in combination with PCR amplification using a linker specific primer paired with a *nanog* specific primer did not reveal any amplification products. In this context it is important to note that until today no positive identification of 5hmC at the upstream regulatory region of *nanog* has been demonstrated, raising uncertainty about the

suitability of this locus to establish the linker- amplification strategy also because it is still highly debated whether *nanog* is a target of Tet1 and hydroxylation at all (Koh et al., 2011). Nonetheless, the use of substrates with defined 5hmC amount showed that the cut and ligation strategy could in principal be used to map 5hmC patterns, although a high local concentration of 5hmC clearly facilitates the detection of digestion products by PvuRtsI1 (Fig. 36) (Szwagierczak et al., 2011). The presence of high local 5hmC concentrations does not seem very unlikely given that in the case of cerebellum, the measured 5hmC levels translate to approximately 40 % of all 5mCs being hydroxylated (Kriaucionis and Heintz, 2009). Furthermore, several studies performed global mapping of 5hmC in ESCs and revealed a non- linear distribution of this modification with specific enrichment of 5hmC within gene bodies (specifically at exons) and at transcriptional start sites and promoters (Ficz et al., 2011; Pastor et al., 2011; Williams et al., 2011; Wu et al., 2011a; Xu et al., 2011). Hence, the enrichment of the digested fragments using our cut/ligation strategy could be applied to generate libraries for massive parallel sequencing and/or microarray hybridizations for genome- wide mapping of 5hmC.

4.3.2 5hmC- an intermediate of demethylation or a stable epigenetic modification?

The discovery of 5hmC in mammalian genomes led to the formulation of two principal hypotheses about the biological role of the 6th base. As a potentially stable base, 5hmC itself might represent a novel epigenetic modification which alters chromatin structure and possibly influences the local transcriptional state. Alternatively, it has been suggested that 5hmC serves as an intermediate stage in the DNA demethylation pathway, although it is still debated whether the demethylation occurs actively and/or passively. Active DNA demethylation has been proposed to involve specific DNA repair mechanisms such as deamination by the cytidine deaminases AID/APOBEC leading to the conversion of 5hmC to 5 hydroxymethyluracil (5hmU), which would then be removed by enzymes of the BER pathway like Tdg or MBD4 (see also chapter 1.2.3 and Fig. 7) (Cortellino et al., 2011; Guo et al., 2011). Further evidence for 5hmC as an intermediate in active DNA demethylation came from studies showing that Tet enzymes can even further oxidize 5hmC to 5 formylcytosine (5fC) and 5 carboxylcytosine (5caC) (Ito et al., 2011). Interestingly, these two oxidation products are also recognized and cleaved by Tdg, offering another mechanism of active DNA demethylation (He et al., 2011; Maiti and Drohat, 2011). Alternatively, 5caC could be decarboxylated to unmodified C by a yet to be identified decarboxylase, which would offer a demethylation pathway without the involvement of the DNA repair machinery. However, it has also been suggested that 5hmC as well as both cytosine derivates are part of a passive demethylation pathway. In line with this, it has been shown that 5hmC, 5fC and 5caC become replication- dependent diluted in the paternal pronucleus in preimplantation embryos

(Inoue and Zhang, 2011; Inoue et al., 2011). In support of a passive demethylation mechanism, it has been shown that 5hmC containing DNA cannot be methylated by Dnmt1 (Valinluck and Sowers, 2007). Based on these results, it is now widely accepted that 5hmC plays a role in DNA demethylation, however, additional function(s) of 5hmC as a stable epigenetic mark are discussed. Especially the high abundance of 5hmC in post- mitotic neurons suggests a function as a epigenetic mark, possibly by changing the local chromatin environment via the recruitment or displacement of proteins (Kriaucionis and Heintz, 2009). Evidence strengthening this hypothesis comes from the finding that the methylcytosine binding protein MeCP2, which is highly abundant in brain tissues, does not recognize 5hmC and therefore might prevent the establishment of repressive chromatin structures (Frauer et al., 2011). Conversely, MBD3 has been suggested as a first possible effector protein which selectively recognizes 5hmC. MBD3 recruitment was shown to be dependent on Tet1-catalyzed hydroxymethylation and suggests a mechanism of how possible effects of 5hmC could be translated within the cell. However, what biological consequences such a possible 5hmC signaling could have is still completely unknown (Yildirim et al., 2011).

4.4 designer TALEs– novel tools for genome editing

The recent discovery of the TALE DNA binding code has enabled the engineering of specific TALEs that bind to user- defined target sequences in a number of different cell types and organisms including plant and mammalian cells (Boch et al., 2009; Moscou and Bogdanove, 2009; Zhang et al., 2011a). These designer TALEs (dTALEs) can be tethered to a variety of effector domains, e.g. activating or repressing, to modulate the transcription of target genes *in vivo*. A recent study synthesized a large number of dTALEs fused to activation domains and targeted the promoters of the "Yamanaka- factors" (oct4, sox2, Klf4, c-myc), factors known to be essential for efficient reprogramming of fibroblasts to induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006). Two genes, *sox2* and *klf4*, were successfully targeted by specific dTALEs and endogenous genes could be activated. However, dTALEs targeting the epigenetically regulated pluripotency gene *oct4* failed to activate the endogenous gene (Zhang et al., 2011a). This raises the question whether epigenetic modifications like DNA methylation interfere with dTALE- mediated transcriptional activation. Our study analyzing 5 different dTALEs fused to the activating domain VP-16 of the herpes simplex, and targeting distinct sites in the *oct4* promoter revealed differences in the efficiency to activate *oct4* in transient gene reporter assays (see also chapter 3.4). Furthermore, we found that methylation of the *oct4* promoter construct did not interfere with dTALE- mediated activation, but reduced its efficiency. We then chose the dTALE with the highest activation efficiency as measured in our transient reporter assays and used this dTALE to hyperactivate endogenous *oct4* expression in ESCs (Fig. 37). However, the same locus could not be activated by dTALEs in neural stem cells (NSCs), where *oct4* is transcriptionally silent, suggesting that repressive epigenetic mechanisms restrict dTALE activity. Therefore, we targeted epigenetic modulators and inhibited the activity of histone deacetylases by valproic acid (VPA) or DNA methyltransferases by 5-aza- deoxycytidine (5-azadC), which facilitated dTALE- mediated transcriptional activation of the epigenetically silent *oct4* locus in NSCs (Fig. 38) (Bultmann et al., 2012). It is well established that silencing of the *oct4* promoter during differentiation consists of a multi-step cascade involving histone H3K9 methylation as well as DNA methylation. This tight epigenetic control of *oct4* seems to constitute a barrier for the inappropriate reactivation of *oct4*, thereby preventing uncontrolled proliferation and cancer (Gidekel et al., 2003; Looijenga et al., 2003; Feldman et al., 2006). Hence, it is conceivable that the repressive epigenetic environment might restrict the access of dTALEs to their target site. In line with this, crystal structures of two TALEs showed that the protein forms a right- handed superhelix which wraps around the DNA helix along the major groove, thereby directly competing with nucleosome positioning (Deng et al., 2012; Mak et al., 2012).

Interestingly, we found that addition of the HDAC inhibitor VPA, but not Trichostatin A (TSA), allowed dTALE- mediated activation of *oct4* (Bultmann et al., 2012). In line with this, it has been shown that both inhibitors have different efficacies for cellular reprogramming and *oct4* promoter activation, implying different target specificities of both inhibitors (Kim and Bang, 2006; Huangfu et al., 2008). Whereas it has been reported that high concentrations of VPA and/or 5-azadC lead to demethylation and reactivation of silent genes (Santi et al., 1983; Dong et al., 2010), we could only detect reduced DNA methylation levels when the inhibitors were applied in combination with the dTALEs, implying a synergistic effect. One possible explanation for this observation could be that dTALE binding to the *oct4* locus might interfere with Dnmt binding and maintenance methylation. The synergistic effect of low concentrations of epigenetic inhibitors and dTALEs indicates that inactive target genes could be activated without global DNA demethylation and hence would avoid unwanted side effects. Indeed, we found that DNA methylation levels were only reduced at the *oct4* promoter, but not at repetitive sequences like major satellites or the imprinted *h19* locus (Fig. 39). Furthermore, we found that dTALE mediated transcriptional activation of *oct4* also reactivated Oct4 target genes of the pluripotency network e.g. *nanog* and *tet1* (Fig. 38). This shows that the reactivated Oct4 also directly affected its downstream targets, leading to the reactivation of several members of the core pluripotency network. Hence, dTALEs could possibly also be used for reprogramming of somatic cells to iPSCs.

In general, dTALEs can not only be applied to activate any desired gene within a cell, but also repressor domains could be tethered to dTALE to allow silencing of specific genes at a user- defined time point. Such repressor domains could be chromatin-modifying enzymes, like histone deacetylases or DNA methyltransferases. Furthermore, by coupling dTALEs to nucleases like *FokI*, site- specific DNA double strand breaks can be generated that enable targeted DNA cleavage for gene knockouts and genome editing (summarized in Bogdanove and Voytas, 2011) (Figure 43).

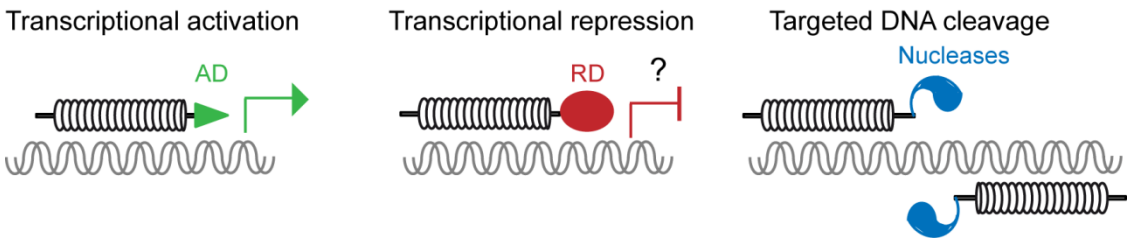


Figure 43. Overview of possibilities for genome editing and manipulation using engineered dTALEs.
TALE proteins fused to activation domains (AD) or repression domains (RD) could be used for gene regulation in vivo, including transcriptional activation and repression of user-defined target genes. Moreover, specific gene knock outs can be generated by tethering a nuclease to TALE proteins. Note that FokI, a nuclease mainly used for this purpose only functions as a heterodimer, thereby reducing off-target effects (reviewed in Bogdanove and Voytas, 2011).

5. Annex

5.1 References

Abranches, E., Silva, M., Pradier, L., Schulz, H., Hummel, O., Henrique, D., and Bekman, E. (2009). Neural Differentiation of Embryonic Stem Cells In Vitro: A Road Map to Neurogenesis in the Embryo. PLoS ONE *4*, e6286.

Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U., Becker, A., Leonhardt, H., and Cardoso, M.C. (2007). MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. Nucl. Acids Res. *35*, 5402–5408.

Ahmed, K., Dehghani, H., Rugg-Gunn, P., Fussner, E., Rossant, J., and Bazett-Jones, D.P. (2010). Global Chromatin Architecture Reflects Pluripotency and Lineage Commitment in the Early Mouse Embryo. PLoS ONE *5*, e10531.

Akasaka, T., Kanno, M., Balling, R., Mieza, M.A., Taniguchi, M., and Koseki, H. (1996). A role for mel-18, a Polycomb group-related vertebrate gene, during theanteroposterior specification of the axial skeleton. Development *122*, 1513–1522.

Amir, R.E., Veyver, I.B.V. den, Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nature Genetics *23*, 185–188.

Antequera, F., and Bird, A. (1993). Number of CpG islands and genes in human and mouse. Proceedings of the National Academy of Sciences *90*, 11995–11999.

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., and Shirakawa, M. (2008). Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature *455*, 818–821.

Athanasiadou, R., de Sousa, D., Myant, K., Merusi, C., Stancheva, I., and Bird, A. (2010). Targeting of De Novo DNA Methylation Throughout the Oct-4 Gene Regulatory Region in Differentiating Embryonic Stem Cells. PLoS ONE *5*, e9937.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes & Development *17*, 126 –140.

Avvakumov, G.V., Walker, J.R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C.H., and Dhe-Paganon, S. (2008). Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. Nature *455*, 822–825.

Ballestar, E., Paz, M.F., Valle, L., Wei, S., Fraga, M.F., Espada, J., Cruz Cigudosa, J., Huang, T.H.-M., and Esteller, M. (2003). Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. EMBO J *22*, 6335–6345.

Beard, C., Li, E., and Jaenisch, R. (1995). Loss of methylation activates Xist in somatic but not in embryonic cells. Genes Dev. *9*, 2325–2334.

Becker, P.B., Ruppert, S., and Schütz, G. (1987). Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. Cell *51*, 435–443.

Beerli, R.R., Dreier, B., and Barbas, C.F. (2000). Positive and Negative Regulation of Endogenous Genes by Designed Transcription Factors. PNAS *97*, 1495–1500.

Bellacosa, A., Cicchillitti, L., Schepis, F., Riccio, A., Yeung, A.T., Matsumoto, Y., Golemis, E.A., Genuardi, M., and Neri, G. (1999). MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. PNAS *96*, 3969–3974.

van den Berg, D.L.C., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2010). An Oct4-Centered Protein Interaction Network in Embryonic Stem Cells. Cell Stem Cell *6*, 369–381.

Berge, D. ten, Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nat Cell Biol *13*, 1070–1075.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell *125*, 315–326.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells: The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. Journal of Molecular Biology *203*, 971–983.

Biniszkiewicz, D., Gribnau, J., Ramsahoye, B., Gaudet, F., Eggan, K., Humpherys, D., Mastrangelo, M.-A., Jun, Z., Walter, J., and Jaenisch, R. (2002). Dnmt1 Overexpression Causes Genomic Hypermethylation, Loss of Imprinting, and Embryonic Lethality. Mol. Cell. Biol. *22*, 2124–2135.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes & Development *16*, 6 –21.

Blancafort, P., Magnenat, L., and Barbas, C.F. (2003). Scanning the human genome with combinatorial transcription factor libraries. Nature Biotechnology *21*, 269–274.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. Science *326*, 1509–1512.

Bogdanove, A.J., Schornack, S., and Lahaye, T. (2010). TAL effectors: finding plant genes for disease and defense. Current Opinion in Plant Biology *13*, 394–401.

Bogdanove, A.J., and Voytas, D.F. (2011). TAL Effectors: Customizable Proteins for DNA Targeting. Science *333*, 1843–1846.

Boiani, M., and Scholer, H.R. (2005). Regulatory networks in embryo-derived pluripotent stem cells. Nat Rev Mol Cell Biol *6*, 872–881.

Bonapace, I.M., Latella, L., Papait, R., Nicassio, F., Sacco, A., Muto, M., Crescenzi, M., and Di Fiore, P.P. (2002). Np95 is regulated by E1A during mitotic reactivation of terminally differentiated cells and is essential for S phase entry. J Cell Biol *157*, 909–914.

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schubeler, D., Sasaki, H., Forne, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. Nat Genet *42*, 1093–1100.

Bostick, M., Kim, J.K., Estève, P.-O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. Science *317*, 1760–1764.

Bourc'his, D., and Bestor, T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature *431*, 96–99.

Bourc'his, D., Xu, G.-L., Lin, C.-S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the Establishment of Maternal Genomic Imprints. Science *294*, 2536 –2539.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. Cell *122*, 947–956.

Branco, M.R., Ficz, G., and Reik, W. (2012). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nat Rev Genet *13*, 7–13.

Bultmann, S., Morbitzer, R., Schmidt, C.S., Thanisch, K., Spada, F., Elsaesser, J., Lahaye, T., and Leonhardt, H. (2012). Targeted Transcriptional Activation of Silent Oct4 Pluripotency Gene by Combining Designer TALEs and Inhibition of Epigenetic Modifiers. Nucl. Acids Res.

Cabianca, D.S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y., and Gabellini, D. (2012). A Long ncRNA Links Copy Number Variation to a Polycomb/Trithorax Epigenetic Switch in FSHD Muscular Dystrophy. Cell *149*, 819–831.

Cardoso, M.C., and Leonhardt, H. (1999). DNA Methyltransferase Is Actively Retained in the Cytoplasm During Early Development. J Cell Biol *147*, 25–32.

Carlson, L.L., Page, A.W., and Bestor, T.H. (1992). Properties and Localization of DNA Methyltransferase in Preimplantation Mouse Embryos: Implications for Genomic Imprinting. Genes Dev. *6*, 2536–2541.

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

Chen, R.Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nature Genetics *27*, 327–331.

Chen, T., Hevi, S., Gay, F., Tsujimoto, N., He, T., Zhang, B., Ueda, Y., and Li, E. (2007). Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. Nat Genet *39*, 391–396.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and Maintenance of Genomic Methylation Patterns in Mouse Embryonic Stem Cells by Dnmt3a and Dnmt3b. Mol. Cell. Biol. *23*, 5594–5605.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. Cell *133*, 1106–1117.

Chew, J.-L., Loh, Y.-H., Zhang, W., Chen, X., Tam, W.-L., Yeap, L.-S., Li, P., Ang, Y.-S., Lim, B., Robson, P., et al. (2005). Reciprocal Transcriptional Regulation of Pou5f1 and Sox2 Via the Oct4/Sox2 Complex in Embryonic Stem Cells. Mol. Cell. Biol. *25*, 6031–6046.

Chuang, L.S.-H., Ian, H.-I., Koh, T.-W., Ng, H.-H., Xu, G., and Li, B.F.L. (1997). Human DNA-(Cytosine-5) Methyltransferase-PCNA Complex as a Target for p21WAF1. Science *277*, 1996–2000.

Citterio, E., Papait, R., Nicassio, F., Vecchi, M., Gomiero, P., Mantovani, R., Di Fiore, P.P., and Bonapace, I.M. (2004). Np95 Is a Histone-Binding Protein Endowed with Ubiquitin Ligase Activity. Mol. Cell. Biol. *24*, 2526–2535.

Cole, M.F., Johnstone, S.E., Newman, J.J., Kagey, M.H., and Young, R.A. (2008). Tcf3 Is an Integral Component of the Core Regulatory Circuitry of Embryonic Stem Cells. Genes Dev. *22*, 746–755.

Core, N., Bel, S., Gaunt, S.J., Aurrand-Lions, M., Pearce, J., Fisher, A., and Djabali, M. (1997). Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice. Development *124*, 721–729.

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.

Daniel, J.M., and Reynolds, A.B. (1999). The Catenin p120 ctn Interacts with Kaiso, a Novel BTB/POZ Domain Zinc Finger Transcription Factor. Mol. Cell. Biol. *19*, 3614–3623.

Daniel, J.M., Spring, C.M., Crawford, H.C., Reynolds, A.B., and Baig, A. (2002). The p120ctn-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. Nucl. Acids Res. *30*, 2911–2919.

Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.-C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.-W., Page, D.C., 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.

Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.-K., Shi, Y., and Yan, N. (2012). Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. Science *335*, 720–723.

Desbaillets, I., Ziegler, U., Groscurth, P., and Gassmann, M. (2000). Embryoid Bodies: An in Vitro Model of Mouse Embryogenesis. Exp Physiol *85*, 645–651.

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2010). The Dnmt3a PWWP Domain Reads Histone 3 Lysine 36 Trimethylation and Guides DNA Methylation. J. Biol. Chem. *285*, 26114–26120.

Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., et al. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat Genet *41*, 1350–1353.

Dong, A., Yoder, J.A., Zhang, X., Zhou, L., Bestor, T.H., and Cheng, X. (2001). Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. Nucleic Acids Research *29*, 439–448.

Dong, E., Chen, Y., Gavin, D.P., Grayson, D.R., and Guidotti, A. (2010). Valproate induces DNA demethylation in nuclear extracts from adult mouse brain. Epigenetics *5*, 730–735.

Du, Z., Song, J., Wang, Y., Zhao, Y., Guda, K., Yang, S., Kao, H.-Y., Xu, Y., Willis, J., Markowitz, S.D., et al. (2010). DNMT1 Stability Is Regulated by Proteins Coordinating Deubiquitination and Acetylation-Driven Ubiquitination. Sci. Signal. *3*, ra80.

Easwaran, H.P., Schermelleh, L., Leonhardt, H., and Cardoso, M.C. (2004). Replication-independent chromatin loading of Dnmt1 during G2 and M phases. EMBO Rep *5*, 1181–1186.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D.G., Buetow, K.H., et al. (2008). Global transcription in pluripotent embryonic stem cells. Cell Stem Cell *2*, 437–447.

Ehrlich, M., Gama-Sosa, M.A., Huang, L.-H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. Nucleic Acids Research *10*, 2709–2721.

Engler, C., Gruetzner, R., Kandzia, R., and Marillonnet, S. (2009). Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. PLoS ONE *4*, e5553.

Epsztejn-Litman, S., Feldman, N., Abu-Remaileh, M., Shufaro, Y., Gerson, A., Ueda, J., Deplus, R., Fuks, F., Shinkai, Y., Cedar, H., et al. (2008). De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. Nat Struct Mol Biol *15*, 1176–1183.

Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet *8*, 286–298.

Estève, P.-O., Chang, Y., Samaranayake, M., Upadhyay, A.K., Horton, J.R., Feehery, G.R., Cheng, X., and Pradhan, S. (2011). A methylation and phosphorylation switch between an adjacent lysine and serine determines human DNMT1 stability. Nat Struct Mol Biol *18*, 42–48.

Estève, P.-O., Chin, H.G., Benner, J., Feehery, G.R., Samaranayake, M., Horwitz, G.A., Jacobsen, S.E., and Pradhan, S. (2009). Regulation of DNMT1 stability through SET7-mediated lysine methylation in mammalian cells. PNAS *106*, 5076–5081.

Estève, P.-O., Chin, H.G., Smallwood, A., Feehery, G.R., Gangisetty, O., Karpf, A.R., Carey, M.F., and Pradhan, S. (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. Genes Dev. *20*, 3089–3103.

Fatemi, M., Hermann, A., Gowher, H., and Jeltsch, A. (2002). Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. European Journal of Biochemistry *269*, 4981–4984.

Fatemi, M., Hermann, A., Pradhan, S., and 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. Journal of Molecular Biology *309*, 1189–1199.

Faulkner, G.J., Kimura, Y., Daub, C.O., Wani, S., Plessy, C., Irvine, K.M., Schroder, K., Cloonan, N., Steptoe, A.L., Lassmann, T., et al. (2009). The regulated retrotransposon transcriptome of mammalian cells. Nature Genetics *41*, 563–571.

Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol *8*, 188–194.

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature *advance online publication*,.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F., et al. (2010). Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation. Cancer Cell *18*, 553–567.

Filion, G.J.P., Zhenilo, S., Salozhin, S., Yamada, D., Prokhortchouk, E., and Defossez, P.-A. (2006). A Family of Human Zinc Finger Proteins That Bind Methylated DNA and Repress Transcription. Mol Cell Biol *26*, 169–181.

Flynn, J., Glickman, J.F., and Reich, N.O. (1996). Murine DNA Cytosine-C5 Methyltransferase: Pre-Steady- and Steady-State Kinetic Analysis with Regulatory DNA Sequences†. Biochemistry *35*, 7308–7315.

Flynn, J., and Reich, N. (1998). Murine DNA (Cytosine-5-)-methyltransferase: Steady-State and Substrate Trapping Analyses of the Kinetic Mechanism. Biochemistry *37*, 15162–15169.

Fouse, S.D., Shen, Y., Pellegrini, M., Cole, S., Meissner, A., Van Neste, L., Jaenisch, R., and 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.

Frauer, C., and Leonhardt, H. (2009). A versatile non-radioactive assay for DNA methyltransferase activity and DNA binding. Nucleic Acids Research *37*, e22–e22.

Frauer, C., Rottach, A., Meilinger, D., Bultmann, S., Fellinger, K., Hasenöder, S., Wang, M., Qin, W., Söding, J., Spada, F., et al. (2011). Different Binding Properties and Function of CXXC Zinc Finger Domains in Dnmt1 and Tet1. PLoS ONE *6*, e16627.

Freudenberg, J.M., Ghosh, S., Lackford, B.L., Yellaboina, S., Zheng, X., Li, R., Cuddapah, S., Wade, P.A., Hu, G., and Jothi, R. (2011). Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. Nucleic Acids Research.

Fujimori, A., Matsuda, Y., Takemoto, Y., Hashimoto, Y., Kubo, E., Araki, R., Fukumura, R., Mita, K., Tatsumi, K., and Muto, M. (1998). Cloning and mapping of Np95 gene which encodes a novel nuclear protein associated with cell proliferation. Mamm. Genome *9*, 1032–1035.

Fujita, N., Watanabe, S., Ichimura, T., Tsuruzoe, S., Shinkai, Y., Tachibana, M., Chiba, T., and Nakao, M. (2003). Methyl-CpG Binding Domain 1 (MBD1) Interacts with the Suv39h1-HP1 Heterochromatic Complex for DNA Methylation-based Transcriptional Repression. J. Biol. Chem. *278*, 24132–24138.

Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L., and Kouzarides, T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. Nature Genetics *24*, 88–91.

Fuks, F., Hurd, P.J., Deplus, R., and Kouzarides, T. (2003a). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucl. Acids Res. *31*, 2305–2312.

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T. (2003b). The Methyl-CpG-binding Protein MeCP2 Links DNA Methylation to Histone Methylation. J. Biol. Chem. *278*, 4035–4040.

Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2010). Open chromatin in pluripotency and reprogramming. Nature Reviews Molecular Cell Biology *12*, 36–47.

Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., and Jaenisch, R. (2003). Induction of tumors in mice by genomic hypomethylation. Science *300*, 489–492.

Ge, Y.-Z., Pu, M.-T., Gowher, H., Wu, H.-P., Ding, J.-P., Jeltsch, A., and Xu, G.-L. (2004). Chromatin Targeting of de Novo DNA Methyltransferases by the PWWP Domain. J. Biol. Chem. *279*, 25447–25454.

Georgel, P.T., Horowitz-Scherer, R.A., Adkins, N., Woodcock, C.L., Wade, P.A., and Hansen, J.C. (2003). Chromatin Compaction by Human MeCP2. J. Biol. Chem. *278*, 32181–32188.

Georgopoulos, C.P., and Revel, H.R. (1971). Studies with glucosyl transferase mutants of the T-even bacteriophages. Virology *44*, 271–285.

Gidekel, S., Pizov, G., Bergman, Y., and Pikarsky, E. (2003). Oct-3/4 is a dose-dependent oncogenic fate determinant. Cancer Cell *4*, 361–370.

Goll, M.G., and Bestor, T.H. (2005). EUKARYOTIC CYTOSINE METHYLTRANSFERASES. Annual Review of Biochemistry *74*, 481–514.

Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.-L., Zhang, X., Golic, K.G., Jacobsen, S.E., and Bestor, T.H. (2006). Methylation of tRNAAsp by the DNA Methyltransferase Homolog Dnmt2. Science *311*, 395–398.

Goto, K., Numata, M., Komura, J.-I., Ono, T., Bestor, T.H., and Kondo, H. (1994). Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice. Differentiation *56*, 39–44.

Göttlicher, M., Minucci, S., Zhu, P., Krämer, O.H., Schimpf, A., Giavara, S., Sleeman, J.P., Lo Coco, F., Nervi, C., Pelicci, P.G., et al. (2001). Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. EMBO J *20*, 6969–6978.

Griffiths, D.S., Li, J., Dawson, M.A., Trotter, M.W.B., Cheng, Y.-H., Smith, A.M., Mansfield, W., Liu, P., Kouzarides, T., Nichols, J., et al. (2011). LIF independent JAK signalling to chromatin in embryonic stem cells uncovered from an adult stem cell disease. Nat Cell Biol *13*, 13–21.

Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-F., Liu, W., Xie, Z.-G., Shi, L., He, X., Jin, S., et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *advance online publication*,.

Guan, K., Rohwedel, J., and Wobus, A. (1999). Embryonic stem cell differentiation models: cardiogenesis, myogenesis, neurogenesis, epithelial and vascular smooth muscle cell differentiation in vitro. Cytotechnology *30*, 211–226.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. Cell *130*, 77–88.

Guo, J.U., Su, Y., Zhong, C., Ming, G., and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423–434.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E., and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nature Genetics *27*, 322–326.

Hackett, J.A., Zylicz, J.J., and Surani, M.A. (2012). Parallel mechanisms of epigenetic reprogramming in the germline. Trends in Genetics *28*, 164–174.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. Mechanisms of Development *117*, 15–23.

Hansen, R.S., Wijmenga, C., Luo, P., Stanek, A.M., Canfield, T.K., Weemaes, C.M.R., and Gartler, S.M. (1999). The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. PNAS *96*, 14412–14417.

Hashimoto, H., Horton, J.R., Zhang, X., Bostick, M., Jacobsen, S.E., and Cheng, X. (2008). The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. Nature *455*, 826–829.

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

Hattori, N., Imao, Y., Nishino, K., Hattori, N., Ohgane, J., Yagi, S., Tanaka, S., and Shiota, K. (2007). Epigenetic regulation of Nanog gene in embryonic stem and trophoblast stem cells. Genes to Cells *12*, 387–396.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells. Cell *In Press, Corrected Proof*,.

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.

Hellman, A., and Chess, A. (2007). Gene Body-Specific Methylation on the Active X Chromosome. Science *315*, 1141–1143.

Hendrich, B., and Bird, A. (1998). Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins. Molecular and Cellular Biology *18*, 6538 –6547.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V.A., and Bird, A. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev. *15*, 710–723.

Hendrich, B., Hardeland, U., Ng, H.-H., Jiricny, J., and Bird, A. (1999). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. Nature *401*, 301–304.

Hendrich, B., and Tweedie, S. (2003). The methyl-CpG binding domain and the evolving role of DNA methylation in animals. Trends in Genetics *19*, 269–277.

Hermann, A., Schmitt, S., and Jeltsch, A. (2003). The Human Dnmt2 Has Residual DNA-(Cytosine-C5) Methyltransferase Activity. Journal of Biological Chemistry *278*, 31717–31721.

Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature *460*, 1132–1135.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protocols *4*, 44–57.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research *37*, 1 –13.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nature Biotechnology *26*, 795–797.

Illingworth, R., Kerr, A., DeSousa, D., Jørgensen, H., Ellis, P., Stalker, J., Jackson, D., Clee, C., Plumb, R., Rogers, J., et al. (2008). A Novel CpG Island Set Identifies Tissue-Specific Methylation at Developmental Gene Loci. PLoS Biol *6*, e22.

Illingworth, R.S., and Bird, A.P. (2009). CpG islands – "A rough guide."FEBS Letters *583*, 1713–1720.

Inano, K., Suetake, I., Ueda, T., Miyake, Y., Nakamura, M., Okada, M., and Tajima, S. (2000). Maintenance-Type DNA Methyltransferase Is Highly Expressed in Post-Mitotic Neurons and Localized in the Cytoplasmic Compartment. Journal of Biochemistry *128*, 315–321.

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

Inoue, A., and Zhang, Y. (2011). Replication-Dependent Loss of 5-Hydroxymethylcytosine in Mouse Preimplantation Embryos. Science.

Iqbal, K., Jin, S.-G., Pfeifer, G.P., and Szabó, P.E. (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.

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet *41*, 178–186.

Ito, S., D/'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and 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, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. Science.

Iwata, A., Nagashima, Y., Matsumoto, L., Suzuki, T., Yamanaka, T., Date, H., Deoka, K., Nukina, N., and Tsuji, S. (2009). Intranuclear Degradation of Polyglutamine Aggregates by the Ubiquitin-Proteasome System. J. Biol. Chem. *284*, 9796–9803.

Iyer, L.M., Tahiliani, M., Rao, A., and Aravind, L. (2009). Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle *8*, 1698–1710.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). Severe Global DNA Hypomethylation Blocks Differentiation and Induces Histone Hyperacetylation in Embryonic Stem Cells. Mol. Cell. Biol. *24*, 8862–8871.

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E., et al. (2001). Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat Genet *27*, 31–39.

Jacobs, S.A., and Khorasanizadeh, S. (2002). Structure of HP1 Chromodomain Bound to a Lysine 9-Methylated Histone H3 Tail. Science *295*, 2080–2083.

Jaenisch, R., and Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet.

Janosi, L., Yonemitsu, H., Hong, H., and Kaji, A. (1994). Molecular Cloning and Expression of a Novel Hydroxymethylcytosine-specific Restriction Enzyme (PvuRts1I) Modulated by Glucosylation of DNA. Journal of Molecular Biology *242*, 45–61.

Jeong, S., Liang, G., Sharma, S., Lin, J.C., Choi, S.H., Han, H., Yoo, C.B., Egger, G., Yang, A.S., and Jones, P.A. (2009). Selective Anchoring of DNA Methyltransferases 3A and 3B to Nucleosomes Containing Methylated DNA. Mol Cell Biol *29*, 5366–5376.

Ji, H., Ehrlich, L.I.R., Seita, J., Murakami, P., Doi, A., Lindau, P., Lee, H., Aryee, M.J., Irizarry, R.A., Kim, K., et al. (2010). Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature *467*, 338–342.

Johnson, D.G., Cress, W.D., Jakoi, L., and Nevins, J.R. (1994). Oncogenic capacity of the E2F1 gene. Proc. Natl. Acad. Sci. U.S.A. *91*, 12823–12827.

Jones, P.L., Veenstra, G.C.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genetics *19*, 187–191.

Jørgensen, H.F., Ben-Porath, I., and Bird, A.P. (2004). Mbd1 Is Recruited to both Methylated and Nonmethylated CpGs via Distinct DNA Binding Domains. Mol. Cell. Biol. *24*, 3387–3395.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900–903.

Kaneko, H., Dridi, S., Tarallo, V., Gelfand, B.D., Fowler, B.J., Cho, W.G., Kleinman, M.E., Ponicsan, S.L., Hauswirth, W.W., Chiodo, V.A., et al. (2011). DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. Nature *471*, 325–330.

Karagianni, P., Amazit, L., Qin, J., and Wong, J. (2008). ICBP90, a Novel Methyl K9 H3 Binding Protein Linking Protein Ubiquitination with Heterochromatin Formation. Mol. Cell. Biol. *28*, 705–717.

Kawamura, T., Suzuki, J., Wang, Y.V., Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., and Belmonte, J.C.I. (2009). Linking the p53 tumour suppressor pathway to somatic cell reprogramming. Nature *460*, 1140–1144.

Kay, S., Hahn, S., Marois, E., Hause, G., and Bonas, U. (2007). A Bacterial Effector Acts as a Plant Transcription Factor and Induces a Cell Size Regulator. Science *318*, 648–651.

Kim, G.-D., Ni, J., Kelesoglu, N., Roberts, R.J., and Pradhan, S. (2002). Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. EMBO J *21*, 4183–4195.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008a). An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. Cell *132*, 1049–1061.

Kim, J., Woo, A.J., Chu, J., Snow, J.W., Fujiwara, Y., Kim, C.G., Cantor, A.B., and Orkin, S.H. (2010). A Myc Network Accounts for Similarities between Embryonic Stem and Cancer Cell Transcription Programs. Cell *143*, 313–324.

Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Ara|[uacute]|zo-Bravo, M.J., Ruau, D., Han, D.W., Zenke, M., et al. (2008b). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. Nature *454*, 646–650.

Kim, J.K., Estève, P.-O., Jacobsen, S.E., and Pradhan, S. (2009). UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells. Nucleic Acids Res *37*, 493–505.

Kim, T.-Y., and Bang, Y.-J. (2006). Histone Deacetylase Inhibitors for Cancer Therapy. Epigenetics *1*, 14–23.

Kimura, H., and Shiota, K. (2003). Methyl-CpG-binding Protein, MeCP2, Is a Target Molecule for Maintenance DNA Methyltransferase, Dnmt1. J. Biol. Chem. *278*, 4806–4812.

Klose, R.J., Sarraf, S.A., Schmiedeberg, L., McDermott, S.M., Stancheva, I., and Bird, A.P. (2005). DNA Binding Selectivity of MeCP2 Due to a Requirement for A/T Sequences Adjacent to Methyl-CpG. Molecular Cell *19*, 667–678.

Ko, M., Bandukwala, H.S., An, J., Lamperti, E.D., Thompson, E.C., Hastie, R., Tsangaratou, A., Rajewsky, K., Koralov, S.B., and Rao, A. (2011). Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. Proc Natl Acad Sci U S A *108*, 14566–14571.

Ko, M., Huang, Y., Jankowska, A.M., Pape, U.J., Tahiliani, M., Bandukwala, H.S., An, J., Lamperti, E.D., Koh, K.P., Ganetzky, R., et al. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature *468*, 839–843.

Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C.A., 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.

Kondo, E., Gu, Z., Horii, A., and Fukushige, S. (2005). The Thymine DNA Glycosylase MBD4 Represses Transcription and Is Associated with Methylated p16INK4a and hMLH1 Genes. Mol. Cell. Biol. *25*, 4388–4396.

Konstandin, N., Bultmann, S., Szwagierczak, A., Dufour, A., Ksienzyk, B., Schneider, F., Herold, T., Mulaw, M., Kakadia, P.M., Schneider, S., et al. (2011). Genomic 5-

hydroxymethylcytosine levels correlate with TET2 mutations and a distinct global gene expression pattern in secondary acute myeloid leukemia. Leukemia *25*, 1649–1652.

Kornberg, S.R., Zimmerman, S.B., and Kornberg, A. (1961). Glucosylation of Deoxyribonucleic Acid by Enzymes from Bacteriophage-Infected Escherichia Coli. J. Biol. Chem. *236*, 1487–1493.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693–705.

Kriaucionis, S., and Heintz, N. (2009). The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. Science *324*, 929–930.

Ku, M., Koche, R.P., Rheinbay, E., Mendenhall, E.M., Endoh, M., Mikkelsen, T.S., Presser, A., Nusbaum, C., Xie, X., Chi, A.S., et al. (2008). Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains. PLoS Genet *4*, e1000242.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K., and Saitou, M. (2008). Complex Genome-Wide Transcription Dynamics Orchestrated by Blimp1 for the Specification of the Germ Cell Lineage in Mice. Genes Dev. *22*, 1617–1635.

Langemeijer, S.M.C., Kuiper, R.P., Berends, M., Knops, R., Aslanyan, M.G., Massop, M., Stevens-Linders, E., Hoogen, P. van, Kessel, A.G. van, Raymakers, R.A.P., et al. (2009). Acquired mutations in TET2 are common in myelodysplastic syndromes. Nature Genetics *41*, 838–842.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Kin Sung, K.W., Rigoutsos, I., Loring, J., et al. (2010). Dynamic changes in the human methylome during differentiation. Genome Research *20*, 320–331.

Lee, B., and Muller, M.T. (2009). SUMOylation enhances DNA methyltransferase 1 activity. Biochemical Journal *421*, 449–461.

Lee, J., Inoue, K., Ono, R., Ogonuki, N., Kohda, T., Kaneko-Ishino, T., Ogura, A., and Ishino, F. (2002). Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. Development *129*, 1807–1817.

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., and Wutz, A. (2010). Polycomb Complexes Act Redundantly to Repress Genomic Repeats and Genes. Genes Dev. *24*, 265–276.

Lei, H., Oh, S. p., Okano, M., Juttermann, R., Goss, K. a., Jaenisch, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195 –3205.

Leonhardt, H., Page, A.W., Weier, H.-U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell *71*, 865–873.

Li, B., Carey, M., and Workman, J.L. (2007a). The Role of Chromatin during Transcription. Cell *128*, 707–719.

Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. , Published Online: 25 December 1993; | Doi:10.1038/366362a0 *366*, 362–365.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915–926.

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Canamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature *460*, 1136–1139.

Li, J.-Y., Pu, M.-T., Hirasawa, R., Li, B.-Z., Huang, Y.-N., Zeng, R., Jing, N.-H., Chen, T., Li, E., Sasaki, H., et al. (2007b). Synergistic Function of DNA Methyltransferases Dnmt3a and Dnmt3b in the Methylation of Oct4 and Nanog. Molecular and Cellular Biology *27*, 8748 – 8759.

Li, L.-C., and Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. Bioinformatics *18*, 1427 –1431.

Li, Y., Mori, T., Hata, H., Homma, Y., and Kochi, H. (2004). NIRF induces G1 arrest and associates with Cdk2. Biochemical and Biophysical Research Communications *319*, 464–468.

Li, Z., Cai, X., Cai, C.-L., Wang, J., Zhang, W., Petersen, B.E., Yang, F.-C., and Xu, M. (2011). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood *118*, 4509–4518.

Liang, G., Chan, M.F., Tomigahara, Y., Tsai, Y.C., Gonzales, F.A., Li, E., Laird, P.W., and Jones, P.A. (2002). Cooperativity between DNA Methyltransferases in the Maintenance Methylation of Repetitive Elements. Mol. Cell. Biol. *22*, 480–491.

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

Lin, C.-H., Lin, C., Tanaka, H., Fero, M.L., and Eisenman, R.N. (2009). Gene Regulation and Epigenetic Remodeling in Murine Embryonic Stem Cells by c-Myc. PLoS One *4*,.

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

Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet *38*, 431–440.

Looijenga, L.H.J., Stoop, H., De Leeuw, H.P.J.C., De Gouveia Brazao, C.A., Gillis, A.J.M., Van Roozendaal, K.E.P., Van Zoelen, E.J.J., Weber, R.F.A., Wolffenbuttel, K.P., Van Dekken, H., et al. (2003). POU5F1 (OCT3/4) Identifies Cells with Pluripotent Potential in Human Germ Cell Tumors. Cancer Res *63*, 2244–2250.

Lorsbach, R.B., Moore, J., Mathew, S., Raimondi, S.C., Mukatira, S.T., and Downing, J.R. (2003). TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia *17*, 637–641.

Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of Alternative Splicing by Histone Modifications. Science *327*, 996–1000.

Maiti, A., and Drohat, A.C. (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.

Mak, A.N.-S., Bradley, P., Cernadas, R.A., Bogdanove, A.J., and Stoddard, B.L. (2012). The Crystal Structure of TAL Effector PthXo1 Bound to Its DNA Target. Science *335*, 716–719.

Mallanna, S.K., Ormsbee, B.D., Iacovino, M., Gilmore, J.M., Cox, J.L., Kyba, M., Washburn, M.P., and Rizzino, A. (2010). Proteomic Analysis of Sox2-associated Proteins During Early Stages of Mouse Embryonic Stem Cell Differentiation Identifies Sox21 as a Novel Regulator of Stem Cell Fate. Stem Cells *28*, 1715–1727.

del Mar Lorente, M., Marcos-Gutierrez, C., Perez, C., Schoorlemmer, J., Ramirez, A., Magin, T., and Vidal, M. (2000). Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. Development *127*, 5093–5100.

Margot, J.B., Aguirre-Arteta, A.M., Di Giacco, B.V., Pradhan, S., Roberts, R.J., Cardoso, M.C., and Leonhardt, H. (2000). Structure and function of the mouse DNA methyltransferase gene: Dnmt1 shows a tripartite structure. Journal of Molecular Biology *297*, 293–300.

Marion, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature *460*, 1149–1153.

Martins-Taylor, K., Schroeder, D.I., Lasalle, J.M., Lalande, M., and Xu, R.-H. (2012). Role of DNMT3B in the regulation of early neural and neural crest specifiers. Epigenetics: Official Journal of the DNA Methylation Society *7*,.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., et al. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat. Cell Biol. *9*, 625–635.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J *18*, 4261–4269.

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

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells. Developmental Cell *10*, 105–116.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49–55.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells. Cell *113*, 631–642.

Miura, M., Watanabe, H., Sasaki, T., Tatsumi, K., and Muto, M. (2001). Dynamic Changes in Subnuclear NP95 Location during the Cell Cycle and Its Spatial Relationship with DNA Replication Foci. Experimental Cell Research *263*, 202–208.

Miyanari, Y., and Torres-Padilla, M.-E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. Nature *advance online publication*,.

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and SchÃ¼beler, D. (2008). Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. Molecular Cell *30*, 755–766.

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M.E., Vasanthakumar, A., Patel, J., Zhao, X., et al. (2011). Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell *20*, 11–24.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (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.

Mori, T., Ikeda, D.D., Fukushima, T., Takenoshita, S., and Kochi, H. (2011). NIRF constitutes a nodal point in the cell cycle network and is a candidate tumor suppressor. Cell Cycle *10*, 3284–3299.

Mori, T., Li, Y., Hata, H., and Kochi, H. (2004). NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a PEST-containing nuclear protein. FEBS Letters *557*, 209–214.

Mori, T., Li, Y., Hata, H., Ono, K., and Kochi, H. (2002). NIRF, a novel RING finger protein, is involved in cell-cycle regulation. Biochemical and Biophysical Research Communications *296*, 530–536.

Mortusewicz, O., Schermelleh, L., Walter, J., Cardoso, M.C., and Leonhardt, H. (2005). Recruitment of DNA methyltransferase I to DNA repair sites. Proceedings of the National Academy of Sciences of the United States of America *102*, 8905–8909.

Moscou, M.J., and Bogdanove, A.J. (2009). A Simple Cipher Governs DNA Recognition by TAL Effectors. Science *326*, 1501–1501.

Muto, M., Kanari, Y., Kubo, E., Takabe, T., Kurihara, T., Fujimori, A., and Tatsumi, K. (2002). Targeted Disruption of Np95 Gene Renders Murine Embryonic Stem Cells Hypersensitive to DNA Damaging Agents and DNA Replication Blocks. J. Biol. Chem. *277*, 34549–34555.

Muto, M., Utsuyama, M., Horiguchi, T., Kubo, E., Sado, T., and Hirokawa, K. (1995). The characterization of the monoclonal antibody Th-10a, specific for a nuclear protein appearing in the S phase of the cell cycle in normal thymocytes and its unregulated expression in lymphoma cell lines. Cell Proliferation *28*, 645–657.

Myant, K., and Stancheva, I. (2008). LSH Cooperates with DNA Methyltransferases To Repress Transcription. Mol. Cell. Biol. *28*, 215–226.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2007a). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nature Biotechnology *26*, 101–106.

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T., and Kimura, T. (2007b). Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. Journal of Bioscience and Bioengineering *104*, 34–41.

Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M., et al. (2006). PGC7/Stella protects against DNA demethylation in early embryogenesis. Nature Cell Biology *9*, 64–71.

Ng, H.-H., and Surani, M.A. (2011). The transcriptional and signalling networks of pluripotency. Nature Cell Biology *13*, 490–496.

Ng, R.K., Dean, W., Dawson, C., Lucifero, D., Madeja, Z., Reik, W., and Hemberger, M. (2008). Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. Nature Cell Biology *10*, 1280–1290.

Nichols, J., and Smith, A. (2011). The origin and identity of embryonic stem cells. Development *138*, 3 –8.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. Cell *95*, 379–391.

Nikitina, T., Shi, X., Ghosh, R.P., Horowitz-Scherer, R.A., Hansen, J.C., and Woodcock, C.L. (2007). Multiple Modes of Interaction between the Methylated DNA Binding Protein MeCP2 and Chromatin. Mol. Cell. Biol. *27*, 864–877.

Nimura, K., Ishida, C., Koriyama, H., Hata, K., Yamanaka, S., Li, E., Ura, K., and Kaneda, Y. (2006). Dnmt3a2 targets endogenous Dnmt3L to ES cell chromatin and induces regional DNA methylation. Genes to Cells *11*, 1225–1237.

Niwa, H. (2011). Wnt: What's Needed To maintain pluripotency? Nat Cell Biol *13*, 1024–1026.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-Renewal of Pluripotent Embryonic Stem Cells Is Mediated Via Activation of STAT3. Genes Dev. *12*, 2048–2060.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat. Genet. *24*, 372–376.

Niwa, H., Ogawa, K., Shimosato, D., and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature *460*, 118–122.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation. Cell *123*, 917–929.

Niwa, H., Yamamura, K., and Miyazaki, J. (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene *108*, 193–199.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The Polycomb-Group Gene Ezh2 Is Required for Early Mouse Development. Mol Cell Biol *21*, 4330–4336.

Oda, M., Yamagiwa, A., Yamamoto, S., Nakayama, T., Tsumura, A., Sasaki, H., Nakao, K., Li, E., and Okano, M. (2006). DNA methylation regulates long-range gene silencing of an X-linked homeobox gene cluster in a lineage-specific manner. Genes & Development *20*, 3382 –3394.

Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M., and Hamada, H. (1990). A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell *60*, 461–472.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. Cell *99*, 247–257.

Okano, M., Xie, S., and Li, E. (1998a). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nature Genetics *19*, 219–220.

Okano, M., Xie, S., and Li, E. (1998b). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Research *26*, 2536–2540.

Okita, K., and Yamanaka, S. (2006). Intracellular signaling pathways regulating pluripotency of embryonic stem cells. Curr Stem Cell Res Ther *1*, 103–111.

Okumura-Nakanishi, S., Saito, M., Niwa, H., and Ishikawa, F. (2005). Oct-3/4 and Sox2 Regulate Oct-3/4 Gene in Embryonic Stem Cells. J. Biol. Chem. *280*, 5307–5317.

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

Orkin, S.H., and Hochedlinger, K. (2011). Chromatin Connections to Pluripotency and Cellular Reprogramming. Cell *145*, 835–850.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Current Biology *10*, 475–478.

Paling, N.R.D., Wheadon, H., Bone, H.K., and Welham, M.J. (2004). Regulation of Embryonic Stem Cell Self-Renewal by Phosphoinositide 3-Kinase-Dependent Signaling. J. Biol. Chem. *279*, 48063–48070.

Pan, G., Li, J., Zhou, Y., Zheng, H., and Pei, D. (2006). A Negative Feedback Loop of Transcription Factors That Controls Stem Cell Pluripotency and Self-Renewal. FASEB J *20*, 1730–1732.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and Thomson, J.A. (2007). Whole-Genome Analysis of Histone H3 Lysine 4 and Lysine 27 Methylation in Human Embryonic Stem Cells. Cell Stem Cell *1*, 299–312.

Panning, B., and Jaenisch, R. (1996). DNA hypomethylation can activate Xist expression and silence X-linked genes. Genes & Development *10*, 1991 –2002.

Papait, R., Pistore, C., Grazini, U., Babbio, F., Cogliati, S., Pecoraro, D., Brino, L., Morand, A.-L., Dechampesme, A.-M., Spada, F., et al. (2008). The PHD Domain of Np95 (mUHRF1) Is Involved in Large-Scale Reorganization of Pericentromeric Heterochromatin. Mol. Biol. Cell *19*, 3554–3563.

Papait, R., Pistore, C., Negri, D., Pecoraro, D., Cantarini, L., and Bonapace, I.M. (2007). Np95 Is Implicated in Pericentromeric Heterochromatin Replication and in Major Satellite Silencing. Molecular Biology of the Cell *18*, 1098–1106.

Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M.M., and Choudhary, J. (2010). An Expanded Oct4 Interaction Network: Implications for Stem Cell Biology, Development, and Disease. Cell Stem Cell *6*, 382–395.

Pastor, W.A., Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., et al. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature *473*, 394–397.

Penn, N.W., Suwalski, R., O'Riley, C., Bojanowski, K., and Yura, R. (1972). The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. Biochem J *126*, 781–790.

Peter A., J. (1999). The DNA methylation paradox. Trends in Genetics *15*, 34–37.

Pfaffeneder, T., Hackner, B., Truß, M., Münzel, M., Müller, M., Deiml, C.A., Hagemeier, C., and Carell, T. (2011). The Discovery of 5-Formylcytosine in Embryonic Stem Cell DNA. Angewandte Chemie International Edition *50*, 7008–7012.

Pichler, G., Wolf, P., Schmidt, C.S., Meilinger, D., Schneider, K., Frauer, C., Fellinger, K., Rottach, A., and Leonhardt, H. (2011). Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways. J Cell Biochem *112*, 1585–2593.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by Aid deficiency. Nature *463*, 1101–1105.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nat Biotech *28*, 1057–1068.

, P.-O., Chin, H.G., Samaranayake, M., Kim, G.-D., and Pradhan, S. (2008). CXXC Domain of Human DNMT1 Is Essential for Enzymatic Activity. Biochemistry *47*, 10000–10009.

Pradhan, S., Bacolla, A., Wells, R.D., and Roberts, R.J. (1999). Recombinant Human DNA (Cytosine-5) Methyltransferase. Journal of Biological Chemistry *274*, 33002–33010.

Prokhortchouk, A., Hendrich, B., Jørgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., and Prokhortchouk, E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev. *15*, 1613–1618.

Qin, W., Leonhardt, H., and Spada, F. (2011). Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1. Journal of Cellular Biochemistry *112*, 439–444.

Quivoron, C., Couronné, L., Della Valle, V., Lopez, C.K., Plo, I., Wagner-Ballon, O., Do Cruzeiro, M., Delhommeau, F., Arnulf, B., Stern, M.-H., et al. (2011). TET2 Inactivation Results in Pleiotropic Hematopoietic Abnormalities in Mouse and Is a Recurrent Event during Human Lymphomagenesis. Cancer Cell *20*, 25–38.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA Demethylation in Zebrafish Involves the Coupling of a Deaminase, a Glycosylase, and Gadd45. Cell *135*, 1201–1212.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proceedings of the National Academy of Sciences *97*, 5237–5242.

Rauch, T.A., Wu, X., Zhong, X., Riggs, A.D., and Pfeifer, G.P. (2009). A human B cell methylome at 100−base pair resolution. Proceedings of the National Academy of Sciences *106*, 671–678.

Ringrose, L., and Paro, R. (2004). EPIGENETIC REGULATION OF CELLULAR MEMORY BY THE POLYCOMB AND TRITHORAX GROUP PROTEINS. Annual Review of Genetics *38*, 413–443.

Robertson, A.K., Geiman, T.M., Sankpal, U.T., Hager, G.L., and Robertson, K.D. (2004). Effects of chromatin structure on the enzymatic and DNA binding functions of DNA methyltransferases DNMT1 and Dnmt3a in vitro. Biochemical and Biophysical Research Communications *322*, 110–118.

Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L., and Wolffe, A.P. (2000a). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nature Genetics *25*, 338–342.

Robertson, K.D., Keyomarsi, K., Gonzales, F.A., Velicescu, M., and Jones, P.A. (2000b). Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G0/G1 to S phase transition in normal and tumor cells. Nucleic Acids Research *28*, 2108–2113.

Rodda, D.J., Chew, J.-L., Lim, L.-H., Loh, Y.-H., Wang, B., Ng, H.-H., and Robson, P. (2005). Transcriptional Regulation of Nanog by OCT4 and SOX2. J. Biol. Chem. *280*, 24731–24737.

Rohde, C., Zhang, Y., Reinhardt, R., and Jeltsch, A. (2010). BISMA--fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. BMC Bioinformatics *11*, 230.

Römer, P., Hahn, S., Jordan, T., Strauß, T., Bonas, U., and Lahaye, T. (2007). Plant Pathogen Recognition Mediated by Promoter Activation of the Pepper Bs3 Resistance Gene. Science *318*, 645–648.

Rottach, A., Frauer, C., Pichler, G., Bonapace, I.M., Spada, F., and Leonhardt, H. (2010). The multi-domain protein Np95 connects DNA methylation and histone modification. Nucl. Acids Res. *38*, 1796–1804.

Rottach, A., Leonhardt, H., and Spada, F. (2009). DNA methylation-mediated epigenetic control. Journal of Cellular Biochemistry *108*, 43–51.

Rountree, M.R., Bachman, K.E., and Baylin, S.B. (2000b). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat Genet *25*, 269–277.

Rountree, M.R., Bachman, K.E., and Baylin, S.B. (2000a). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nature Genetics *25*, 269–277.

Saito, M., and Ishikawa, F. (2002). The mCpG-binding Domain of Human MBD3 Does Not Bind to mCpG but Interacts with NuRD/Mi2 Components HDAC1 and MTA2. J. Biol. Chem. *277*, 35434–35439.

Sakaue, M., Ohta, H., Kumaki, Y., Oda, M., Sakaide, Y., Matsuoka, C., Yamagiwa, A., Niwa, H., Wakayama, T., and Okano, M. (2010). DNA Methylation Is Dispensable for the Growth and Survival of the Extraembryonic Lineages. Current Biology *20*, 1452–1457.

Santi, D.V., Garrett, C.E., and Barr, P.J. (1983). On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs. Cell *33*, 9–10.

Sarraf, S.A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol. Cell *15*, 595–605.

Sasai, N., and Defossez, P.-A. (2009). Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. Int. J. Dev. Biol. *53*, 323–334.

Sasai, N., Matsuda, E., Sarashina, E., Ishida, Y., and Kawaichi, M. (2005). Identification of a novel BTB-zinc finger transcriptional repressor, CIBZ, that interacts with CtBP corepressor. Genes to Cells *10*, 871–885.

Sato, N., Kondo, M., and Arai, K. (2006). The orphan nuclear receptor GCNF recruits DNA methyltransferase for Oct-3/4 silencing. Biochemical and Biophysical Research Communications *344*, 845–851.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2003). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nature Medicine *10*, 55–63.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome Regulation by Polycomb and Trithorax Proteins. Cell *128*, 735–745.

Shahbazian, M.D., and Zoghbi, H.Y. (2002). Rett Syndrome and MeCP2: Linking Epigenetics and Neuronal Function. The American Journal of Human Genetics *71*, 1259–1272.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., et al. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908–912.

Sharma, S., De Carvalho, D.D., Jeong, S., Jones, P.A., and Liang, G. (2011). Nucleosomes Containing Methylated DNA Stabilize DNA Methyltransferases 3A/3B and Ensure Faithful Epigenetic Inheritance. PLoS Genet *7*, e1001286.

Shen, L., Kondo, Y., Guo, Y., Zhang, J., Zhang, L., Ahmed, S., Shu, J., Chen, X., Waterland, R.A., and Issa, J.-P.J. (2007). Genome-Wide Profiling of DNA Methylation Reveals a Class of Normally Methylated CpG Island Promoters. PLoS Genet *3*, e181.

Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Schöler, H.R., and Ding, S. (2008). Induction of Pluripotent Stem Cells from Mouse Embryonic Fibroblasts by Oct4 and Klf4 with Small-Molecule Compounds. Cell Stem Cell *3*, 568–574.

Siegfried, Z., Eden, S., Mendelsohn, M., Feng, X., Tsuberi, B.-Z., and Cedar, H. (1999). DNA methylation represses transcription in vivo. Nat Genet *22*, 203–206.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog Is the Gateway to the Pluripotent Ground State. Cell *138*, 722–737.

Silva, J., and Smith, A. (2008). Capturing Pluripotency. Cell *132*, 532–536.

Skene, P.J., Illingworth, R.S., Webb, S., Kerr, A.R.W., James, K.D., Turner, D.J., Andrews, R., and Bird, A.P. (2010). Neuronal MeCP2 Is Expressed at Near Histone-Octamer Levels and Globally Alters the Chromatin State. Molecular Cell *37*, 457–468.

De Smet, C., Lurquin, C., Lethé, B., Martelange, V., and Boon, T. (1999). DNA Methylation Is the Primary Silencing Mechanism for a Set of Germ Line- and Tumor-Specific Genes with a CpG-Rich Promoter. Molecular and Cellular Biology *19*, 7327–7335.

Smyth, G.K. (2005). Limma: linear models for microarray data. In Bioinformatics and Computational Biology Solutions Using R and Bioconductor, R. Gentleman, V.J. Carey, W. Huber, R.A. Irizarry, and S. Dudoit, eds. (New York: Springer), pp. 397–420.

Song, J., Rechkoblit, O., Bestor, T.H., and Patel, D.J. (2011). Structure of DNMT1-DNA Complex Reveals a Role for Autoinhibition in Maintenance DNA Methylation. Science *331*, 1036 –1040.

Song, J., Teplova, M., Ishibe-Murakami, S., and Patel, D.J. (2012). Structure-Based Mechanistic Insights into DNMT1-Mediated Maintenance DNA Methylation. Science *335*, 709 –712.

Spada, F., Haemmer, A., Kuch, D., Rothbauer, U., Schermelleh, L., Kremmer, E., Carell, T., Längst, G., and Leonhardt, H. (2007). DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. The Journal of Cell Biology *176*, 565 –571.

Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer *6*, 846–856.

Spring, C.M., Kelly, K.F., O'Kelly, I., Graham, M., Crawford, H.C., and Daniel, J.M. (2005). The catenin p120ctn inhibits Kaiso-mediated transcriptional repression of the β-catenin/TCF target gene matrilysin. Experimental Cell Research *305*, 253–265.

Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A., and Cedar, H. (1982). Clonal inheritance of the pattern of DNA methylation in mouse cells. Proceedings of the National Academy of Sciences *79*, 61–65.

Syeda, F., Fagan, R.L., Wean, M., Avvakumov, G.V., Walker, J.R., Xue, S., Dhe-Paganon, S., and Brenner, C. (2011). The Replication Focus Targeting Sequence (RFTS) Domain Is a DNA-competitive Inhibitor of Dnmt1. J. Biol. Chem. *286*, 15344–15351.

Szwagierczak, A., Brachmann, A., Schmidt, C.S., Bultmann, S., Leonhardt, H., and Spada, F. (2011). Characterization of PvuRts1I endonuclease as a tool to investigate genomic 5–hydroxymethylcytosine. Nucleic Acids Research.

Szwagierczak, A., Bultmann, S., Schmidt, C.S., Spada, F., and Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Research *38*, e181.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., 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., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Takai, D., and Jones, P.A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc. Natl. Acad. Sci. U.S.A. *99*, 3740–3745.

Takeshita, K., Suetake, I., Yamashita, E., Suga, M., Narita, H., Nakagawa, A., and Tajima, S. (2011). Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). Proceedings of the National Academy of Sciences *108*, 9055–9059.

Tang, H., Araki, K., Li, Z., and Yamamura, K. (2008). Characterization of Ayu17-449 gene expression and resultant kidney pathology in a knockout mouse model. Transgenic Research *17*, 599–608.

Tatematsu, K., Yamazaki, T., and Ishikawa, F. (2000a). MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase. Genes to Cells *5*, 677–688.

Tatematsu, K., Yamazaki, T., and Ishikawa, F. (2000b). MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase. Genes to Cells *5*, 677–688.

Thomson, M., Liu, S.J., Zou, L.-N., Smith, Z., Meissner, A., and Ramanathan, S. (2011). Pluripotency Factors in Embryonic Stem Cells Regulate Differentiation into Germ Layers. Cell *145*, 875–889.

Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M., and Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4–Sox-2 complex. Nucleic Acids Res *30*, 3202–3213.

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

Uemura, T., Kubo, E., Kanari, Y., Ikemura, T., Tatsumi, K., and Muto, M. (2000). Temporal and Spatial Localization of Novel Nuclear Protein NP95 in Mitotic and Meiotic Cells. Cell Structure and Function *25*, 149–159.

Unoki, M., Nishidate, T., and Nakamura, Y. (2004). ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. Oncogene *23*, 7601–7610.

Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. (2009). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature *460*, 1145–1148.

Valinluck, V., and Sowers, L.C. (2007). Endogenous Cytosine Damage Products Alter the Site Selectivity of Human DNA Maintenance Methyltransferase DNMT1. Cancer Res *67*, 946–950.

Vir|[eacute]|, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Eynde, A.V., Bernard, D., Vanderwinden, J.-M., et al. (2005). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871–874.

Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.-M., et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871–874.

Voncken, J.W., Roelen, B.A.J., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., and van Lohuizen, M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proc Natl Acad Sci U S A *100*, 2468–2473.

Wada, J., Kamada, R., Imagawa, T., Chuman, Y., and Sakaguchi, K. (2012). Inhibition of tumor suppressor protein p53-dependent transcription by a tetramerization domain peptide via hetero-oligomerization. Bioorganic & Medicinal Chemistry Letters *22*, 2780–2783.

Waddington, C.H. (1942). The Epigenotype. International Journal of Epidemiology.

Walsh, C.P., Chaillet, J.R., and Bestor, T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet *20*, 116–117.

Wang, J., Mager, J., Schnedier, E., and Magnuson, T. (2002). The mouse PcG gene eed is required for Hox gene repression and extraembryonic development. Mamm. Genome *13*, 493–503.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364–368.

Watanabe, D., Suetake, I., Tada, T., and Tajima, S. (2002). Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. Mechanisms of Development *118*, 187–190.

Watanabe, S., Umehara, H., Murayama, K., Okabe, M., Kimura, T., and Nakano, T. (2006). Activation of Akt signaling is sufficient to maintain pluripotency in mouse and primate embryonic stem cells. Oncogene *25*, 2697–2707.

Weber, A., Marquardt, J., Elzi, D., Forster, N., Starke, S., Glaum, A., Yamada, D., Defossez, P.-A., Delrow, J., Eisenman, R.N., et al. (2008). Zbtb4 represses transcription of P21CIP1 and controls the cellular response to p53 activation. The EMBO Journal *27*, 1563–1574.

Welham, M.J., Kingham, E., Sanchez‐Ripoll, Y., Kumpfmueller, B., Storm, M., and Bone, H. (2011). Controlling embryonic stem cell proliferation and pluripotency: the role of PI3K- and GSK-3-dependent signalling. Biochemical Society Transactions *39*, 674–678.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A.C., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *advance online publication*,.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nature Communications *2*, 241.

Wu, H., D/'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Eve Sun, Y., and Zhang, Y. (2011a). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature *advance online publication*,.

Wu, H., D'Alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., and Zhang, Y. (2011b). Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev. *25*, 679–684.

Xie, Z.-H., Huang, Y.-N., Chen, Z.-X., Riggs, A.D., Ding, J.-P., Gowher, H., Jeltsch, A., Sasaki, H., Hata, K., and Xu, G.-L. (2006). Mutations in DNA methyltransferase DNMT3B in ICF syndrome affect its regulation by DNMT3L. Hum. Mol. Genet. *15*, 1375–1385.

Xu, G.-L., Bestor, T.H., Bourc'his, D., Hsieh, C.-L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and Viegas-P|[eacute]|quignot, E. (1999). Chromosome instability and

immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature *402*, 187–191.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., et al. (2011). Genome-wide Regulation of 5hmC, 5mC, and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells. Molecular Cell *42*, 451–464.

Yamazaki, Y., Mann, M.R.W., Lee, S.S., Marh, J., McCarrey, J.R., Yanagimachi, R., and Bartolomei, M.S. (2003). Reprogramming of primordial germ cells begins before migration into the genital ridge, making these cells inadequate donors for reproductive cloning. PNAS *100*, 12207–12212.

Yasui, D.H., Peddada, S., Bieda, M.C., Vallero, R.O., Hogart, A., Nagarajan, R.P., Thatcher, K.N., Farnham, P.J., and LaSalle, J.M. (2007). Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. PNAS *104*, 19416–19421.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., and Scholer, H.R. (1996). Germline Regulatory Element of Oct-4 Specific for the Totipotent Cycle of Embryonal Cells. Development *122*, 881–894.

Yildirim, O., Li, R., Hung, J.-H., Chen, P.B., Dong, X., Ee, L.-S., Weng, Z., Rando, O.J., and Fazzio, T.G. (2011). Mbd3/NURD Complex Regulates Expression of 5-Hydroxymethylcytosine Marked Genes in Embryonic Stem Cells. Cell *147*, 1498–1510.

Ying, Q.-L., Nichols, J., Chambers, I., and Smith, A. (2003a). 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., and Smith, A. (2003b). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nature Biotechnology *21*, 183–186.

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

Yoder, J.A., and Bestor, T.H. (1998). A Candidate Mammalian DNA Methyltransferase Related to pmt1p of Fission Yeast. Human Molecular Genetics *7*, 279–284.

Yoder, J.A., Soman, N.S., Verdine, G.L., and Bestor, T.H. (1997a). DNA (cytosine-5)-methyltransferases in mouse cells and tissues. studies with a mechanism-based probe. Journal of Molecular Biology *270*, 385–395.

Yoder, J.A., Walsh, C.P., and Bestor, T.H. (1997b). Cytosine methylation and the ecology of intragenomic parasites. Trends in Genetics *13*, 335–340.

Yoon, H.-G., Chan, D.W., Reynolds, A.B., Qin, J., and Wong, J. (2003). N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. Mol. Cell *12*, 723–734.

Yoshida, M., Kijima, M., Akita, M., and Beppu, T. (1990). Potent and Specific Inhibition of Mammalian Histone Deacetylase Both in Vivo and in Vitro by Trichostatin A. J. Biol. Chem. *265*, 17174–17179.

You, J.S., Kelly, T.K., De Carvalho, D.D., Taberlay, P.C., Liang, G., and Jones, P.A. (2011). OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers

of epigenetic regulation of its target genes. Proceedings of the National Academy of Sciences *108*, 14497 –14502.

Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995). Developmental-Specific Activity of the FGF-4 Enhancer Requires the Synergistic Action of Sox2 and Oct-3. Genes Dev. *9*, 2635–2645.

Li and Yurchenco (2006). Methods in Molecular Biology #329: Embryonic Stem Cell Protocols: Volume I: Isolation and Characterization (Springer).

Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M., and Arlotta, P. (2011a). Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nature Biotechnology *29*, 149–153.

Zhang, J., Gao, Q., Li, P., Liu, X., Jia, Y., Wu, W., Li, J., Dong, S., Koseki, H., and Wong, J. (2011b). S phase-dependent interaction with DNMT1 dictates the role of UHRF1 but not UHRF2 in DNA methylation maintenance. Cell Res *21*, 1723–1739.

Zhao, X., Ueba, T., Christie, B.R., Barkho, B., McConnell, M.J., Nakashima, K., Lein, E.S., Eadie, B.D., Willhoite, A.R., Muotri, A.R., et al. (2003). Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. PNAS *100*, 6777–6782.

Zhao, X.D., Han, X., Chew, J.L., Liu, J., Chiu, K.P., Choo, A., Orlov, Y.L., Sung, W.-K., Shahab, A., Kuznetsov, V.A., et al. (2007). Whole-Genome Mapping of Histone H3 Lys4 and 27 Trimethylations Reveals Distinct Genomic Compartments in Human Embryonic Stem Cells. Cell Stem Cell *1*, 286–298.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet *39*, 61–69.

Zimmermann, C., Guhl, E., and Graessmann, A. (1997). Mouse DNA methyltransferase (MTase) deletion mutants that retain the catalytic domain display neither de novo nor maintenance methylation activity in vivo. Biol. Chem. *378*, 393–405.

5.2 Abbreviations

5-azadC: 5-aza cytidine

5cac: 5-carboxylcytosine

5fc: 5-formylcytosine

5mC: 5-methylcytosine

5hmC: 5-hydroxymethylcytosine

aa: amino acid

Ac: acetylation

AID: activation-induced cytidine deaminase

Apobec: apolipoprotein B mRNA editing enzyme, catalytic polypeptide

ATP: adenosine-5´-triphosphate

ß-gt: beta-glucosyltransferase

BAH: Bromo Adjacent Homology domain

BER: base excision repair

bp: base pair

CpG: cytosine-phosphatidyl-guanine

Da: dalton

DMAP: Dnmt1-associated protein

DMR: differentially methylated regions

DNA: deoxyribonucleic acid

Dnmt: DNA methyltransferase

dTALE: designer transcription activator-like effector proteins

EB: embryoid body

Eed: embryonic ectodermal development

EGC: embryonic germ cell

EGF: epidermal growth factor

EGFP: enhanced green fluorescent protein

ESC: embryonic stem cell

Ezh2: enhancer of zeste homolog 2

FACS: fluorescence-activated cell sorting

FGF: fibroblast growth factor

GO: gene ontology

G phase: Gap phase

HAT: histone acetyltransferase

HDAC: histone deactelyase

Hp1: heterochromatin binding protein 1

HSC: hematopoetic stem cell

IAP: intracisternal A particle
ICBP90: inverted CCAAT box binding protein of 90 kDa
ICF: immunodeficiency, centromer instability and facial anomalies
ICM: inner cell mass
IDH: isocitrate dehydrogenase
iPSC: induced pluripotent stem cells
KBS: kaiso binding sequence
LIF: leukemia inhibitory factor
MBD: methyl-CpG binding domain
MBP: methyl-CpG binding protein
MDR: methylation-determining regions
Me: methylation
MeCP2: methyl-CpG binding protein 2
mRNA: messenger RNA
Np95/97: nueclar protein of 95/97 kDa
NIRF: Np95/ICBP90-like RING finger protein
NSC: neural stem cell
NuRD: nucleosome remodeling complex
Oct4: octamer binding transcription factor 4
p: phosphorylation
PBD: PCNA binding domain
PCA: Principal component analysis
PcG: Polycomb group
PCNA: proliferating cell nuclear antigen
PHD: Plant Homeo Domain
PGC: Primordial germ cells
PI: Propidium iodide
PRC: Polycomb repressive complex
qPCR: quantitative PCR
RING: really interesting new gene
RNA: ribonucleic acid
S-Adomet: S-adenosyl-L-methionine
S phase: synthesis phase
SRA: Set- and Ring-associated
Suv39h1: suppressor of variegation 3-9 homolog 1
Tdg: tymidine glycosylase
TE: trophoectoderm

Tet: Ten-eleven translocation

TKO: triple knock out of *dnmt1*$^{-/-}$*; dnm3a*$^{-/-}$*; dnmt3b*$^{-/-}$ ESCs

TRD: transcription repressor domain

trxG: trithorax group

TS: Targeting Sequence

TSA: Trichostatin A

TSS: transcriptional start site

Ub: ubiquitin

Uhrf: ubiquitin-like containing PhD and Ring finger domain proteins

VPA: valproic acid

wt: wild type

5.3 Declaration

Declaration according to the “Promotionsordnung der LMU München für die Fakultät Biologie“.

Betreuung: Hiermit erkläre ich, dass die vorgelegte Arbeit an der LMU von Herrn Prof. Dr. Heinrich Leonhardt betreut wurde.

Anfertigung: Ich versichere hiermit an Eides stattes, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfsmittel angefertigt wurde.

Prüfung: Hiermit erkläre ich, dass die Dissertation weder als ganzes noch in Teilen an einem anderen Ort einer Prüfungskommission vorgelegt wurde. Weiterhin habe ich weder an einem anderen Ort eine Promotion angestrebt noch angemeldet noch versucht eine Doktorprüfung abzulegen.

München, den 9. August 2012

_____

Christine Schmidt

5.4 Acknowledgements

First of all, I would like to thank my doctor father and direct supervisor **Prof. Dr. Heinrich Leonhardt** for giving me the unique opportunity to conduct my PhD thesis in your group. Thank you for all your support, your brilliant ideas, the stimulating and fruitful discussions and your interest in my work. Also, I would like to thank you for letting me develop my ideas. The time in your lab not only sharpened my scientific thinking and widened my technical skill set, but also greatly influenced me personally.

I thank **Dr. Fabio Spada** for his motivating discussions and brilliant ideas. Thank you for sharing your wide scientific knowledge with me, for always taking time to answer my questions and supporting and appreciating my work.

I would like to thank the members of my thesis advisory committee, **Dr. Sandra Hake** and **Dr. Zuzana Storchova**, for the interest in my work, the motivating discussions and valuable feedback during my progress reports.

I am grateful to **Dr. Achim Tresch** and **Benedikt Zacher** for the successful cooperation and the analysis of my Microarray data. Furthermore, I would like to thank **Prof. Dr. Dietmar** and **Kerstin Maier** for conducting the Microarray hybridizations.

I thank **Dr. Hans-Jörg Schäffer**, **Maxi Reif** and **Dr. Ingrid Wolf** from the IMPRS coordination office for the financial support and the great time I had during all the IMPRS seminars, soft skill workshops and retreats. You guys really do I great job and I am grateful to be part of such a wonderful PhD program. Thanks also to all my IMPRS colleagues from Bambodia! ☺

I am grateful to the IDK-NBT for the financial support and the organization of the interesting lectures and workshops. Special thanks to **Marilena Pinto**, **Dr. Marie-Christine Blüm** and **Dr. Susanne Henning** from the coordination office.

A very big thank you goes to **all members of the Leonhardt group** and the **Chromobody team** for creating a nice and supportive working atmosphere. **Anja** and **Susanne**, thank you for taking care of all the organization of our daily lab work! Especially I want to thank **Anja** for all the support and for being the good soul of the lab! Thanks to **Dani** and **Andrea** for taking care of all the teaching and bringing so many nice and talented students to our lab!

I would like to thank **Sebastian** for all his support and the great team work, which resulted in so many successful projects. Thank you for all the scientific and personal discussions, the fun we had during our countless coffee breaks and for always having an open ear for me! A very special thank you goes to **Carina**, who helped me so much during my PhD thesis, not

only scientifically, but also personally. Thank you for all the great lunch and coffee breaks, your friendship and for always being there for me! I really miss you a lot!! Also, a special thank you goes to **Garwin**, for all his support during my PhD thesis, the great discussions during our coffee breaks and for just being you! I wish you all the best in the States – and I will miss you a lot!! I thank **Anja, Dani, Sebastian, Fabio** and **Udo** for the great office atmosphere we had and have! Thanks also to **Katharina**, **Katrin** and **Patricia** for the nice lunch times and scientific and non-scientific discussions. **Katrin**, thank you for always giving me a ride when I needed one! **Katharina**, thank you for bringing so much more organization to our lab and for your good-natured spirit!

I would also like to thank all other current and former members of the Leonhardt group, **Weihua, Wen, Nan, Congdi, Boris, Irina, Jürgen, Andreas, Mengxi, Jonas, Kamila, Bijan, Andrea, Ola, Tina, Kathrin, Olli, Kourosh, Uli, Jacqueline, Daniel** and **Lothar** for being great colleagues. Special thanks also go to **Tina, Martin, Nils** and **Carina** for all the great card game nights and fun we had and have together!

I thank all my closest friends, **Babsi, Sanni, Michelle, Hannah, Micha** und **Chrissi** for the great friendship and all the fun times we had and have together! You guys are just awesome!

A very special thank you goes to my sister **Claudia**, for all her support, believe in me, encouraging words and most importantly, all the fun we have together! Danke dass du immer für mich da bist! Ich bin so froh dich als große Schwester zu haben! The C-Sisters rock! ☺

Also, I am immensely grateful to my parents! Liebe **Mama** und **Papa**, vielen Dank dass ihr mich immer unterstützt und immer für mich da ward, seid und sein werdet! Ihr seid die besten Eltern der Welt! Ohne Euch wäre ich heute nicht da wo ich jetzt stehe!

Last but not least, I would like to thank my boyfriend **Christian** for his unconditional love and support during all the up and downs of my PhD thesis. Thank you for always believing in me, your encouranging words and for making me happy. I can´t wait until we finally can be together again and start a new chapter in our life together.

To all those wonderful people, I dedicate this thesis

6. Appendix

6.1 Differentially expressed genes in the pluripotent state

Genes upregulated in dnmt1-/- ESCs compared to wt ESCs			
Gene symbol	**Entrez Gene ID**	**Fold change to wt ESCs**	**Gene name**
1700013H16Rik	75514	1.67	RIKEN cDNA 1700013H16 gene
AV320801	331531	2.85	expressed sequence AV320801
Dazl	13164	2.01	deleted in azoospermia-like
Dcdc2a	195208	1.34	doublecortin domain containing 2a
Dnajc5g	231098	1.36	DnaJ (Hsp40) homolog, subfamily C, member 5 gamma
Efhc2	74405	3.44	EF-hand domain (C-terminal) containing 2
Fkbp6	94244	1.81	FK506 binding protein 6
Fthl17	83457	1.65	ferritin, heavy polypeptide-like 17
Gm13154	433804	1.21	predicted gene 13154
Gm13498	227885	1.83	predicted gene 13498
Gm2889	100040658	2.1	hypothetical protein LOC100041609; hypothetical protein LOC100044795; predicted gene 3395; similar to gag polyprotein; hypothetical protein LOC100047557; hypothetical protein LOC100040347; hypothetical protein LOC100044384; hypothetical protein LOC100045342; hypothetical protein LOC100038979; predicted gene 2889
Gm4638	100043775	2.78	predicted gene 4638
Gm5128	331529	3.08	predicted gene 7903; predicted gene 5128
Gm5635	434729	1.18	predicted gene 5635
Gpat2	215456	1.45	RIKEN cDNA A530057A03 gene
Gpx6	75512	1.23	glutathione peroxidase 6
Gtsf1	74174	1.67	gametocyte specific factor 1
Hormad1	67981	2.42	HORMA domain containing 1; predicted gene 7167
Kdm5d	20592	1.04	lysine (K)-specific demethylase 5D
LOC280487	280487	2.78	pol polyprotein
Magea1	17137	1.42	melanoma antigen, family A, 1
Nckap1l	105855	1.54	NCK associated protein 1 like
Nlrc4	268973	1.48	NLR family, CARD domain containing 4
Nlrp4c	83564	1.61	NLR family, pyrin domain containing 4C
Nxf3	245610	1.72	nuclear RNA export factor 3
Olfr307	258610	2.51	olfactory receptor 307
Pnma5	385377	1.58	paraneoplastic antigen family 5
Pramel3	83565	2.96	preferentially expressed antigen in melanoma-like 3
Rbmy1a1	19657	1.58	RNA binding motif protein, Y chromosome, family 1, member A1
Rnf17	30054	1.28	ring finger protein 17
Rpl39l	68172	2.41	ribosomal protein L39-like
Scml2	107815	1.64	similar to sex comb on midleg-like 2 (Drosophila); sex comb on midleg-like 2 (Drosophila)
Serpina3m	20717	1.45	serine (or cysteine) peptidase inhibitor, clade A, member 3M
Slc5a4b	64454	1.92	solute carrier family 5 (neutral amino acid transporters, system A), member 4b
Smc1b	140557	2.27	structural maintenance of chromosomes 1B
Sohlh2	74434	1.4	spermatogenesis and oogenesis specific basic helix-loop-helix 2
Stk31	77485	1.69	serine threonine kinase 31
Ube1y1	22202	1.35	similar to ubiquitin activating enzyme E1; ubiquitin-activating enzyme E1, Chr Y 1
Usp9y	107868	1.39	ubiquitin specific peptidase 9, Y chromosome
Vmn2r-ps104	100041915	1.05	vomeronasal 2, receptor, pseudogene 104
Wfdc15a	68221	1.85	WAP four-disulfide core domain 15A
Xlr3a	22445	3.35	hypothetical protein LOC100044094; similar to X-linked lymphocyte-regulated protein 3A; X-linked lymphocyte-regulated 3E, pseudogene; hypothetical protein LOC100044314; similar to X-LINKED LYMPHOCYTE-REGULATED PROTEIN 3A (XLR RELATED PROTEIN A12); X-linked lymphocyte-regulated 3C; X-linked lymphocyte-regulated 3A; X-linked lymphocyte-regulated 3B; X-linked lymphocyte-regulated 3D, pseudogene
Xlr3b	574437	3.48	hypothetical protein LOC100044094; similar to X-linked lymphocyte-regulated protein 3A; X-linked lymphocyte-regulated 3E, pseudogene; hypothetical protein LOC100044314; similar to X-LINKED LYMPHOCYTE-REGULATED PROTEIN 3A (XLR RELATED PROTEIN A12); X-linked lymphocyte-regulated 3C; X-linked lymphocyte-regulated 3A; X-linked lymphocyte-regulated 3B; X-linked lymphocyte-regulated 3D, pseudogene

Xlr3c	22446	3.54	hypothetical protein LOC100044094; similar to X-linked lymphocyte-regulated protein 3A; X-linked lymphocyte-regulated 3E, pseudogene; hypothetical protein LOC100044314; similar to X-LINKED LYMPHOCYTE-REGULATED PROTEIN 3A (XLR RELATED PROTEIN A12); X-linked lymphocyte-regulated 3C; X-linked lymphocyte-regulated 3A; X-linked lymphocyte-regulated 3B; X-linked lymphocyte-regulated 3D, pseudogene
Xlr4b	27083	3.07	X-linked lymphocyte-regulated 4D; X-linked lymphocyte-regulated 4E, pseudogene; X-linked lymphocyte-regulated 4B; X-linked lymphocyte-regulated 4C; hypothetical protein LOC100044049
Xlr4c	72891	3.36	X-linked lymphocyte-regulated 4D; X-linked lymphocyte-regulated 4E, pseudogene; X-linked lymphocyte-regulated 4B; X-linked lymphocyte-regulated 4C; hypothetical protein LOC100044049
Xlr5a	574438	4.48	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Xlr5b	627081	5.05	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Zfy1	22767	1.94	zinc finger protein 1, Y linked

Genes down regulated in dnmt1-/- ESCs compared to wt ESCs			
Gene symbol	**Entrez Gene ID**	**Fold change to wt ESCs**	**Gene name**
Dnmt1	13433	-1.02	DNA methyltransferase (cytosine-5) 1
Eif4g3	230861	-1.03	eukaryotic translation initiation factor 4 gamma, 3; similar to Eukaryotic translation initiation factor 4 gamma 3 (eIF-4-gamma 3) (eIF-4G 3) (eIF4G 3) (eIF-4-gamma II) (eIF4GII)
Frmd4b	232288	-1.13	FERM domain containing 4B
Lphn2	99633	-1.53	latrophilin 2
Sema6a	20358	-2.12	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A

Genes upregulated in TKO ESCs compared to wt ESCs			
Gene symbol	**Entrez Gene ID**	**Fold change to wt ESCs**	**Gene name**
1700013H16Rik	75514	1.81	RIKEN cDNA 1700013H16 gene
1700080O16Rik	74279	1.14	RIKEN cDNA 1700080O16 gene
4930591A17Rik	68175	1.51	RIKEN cDNA 4930591A17 gene
9030617O03Rik	217830	1.04	RIKEN cDNA 9030617O03 gene
Asz1	74068	1.49	ankyrin repeat, SAM and basic leucine zipper domain containing 1
AV320801	331531	3.61	expressed sequence AV320801
Camk1d	227541	1.14	calcium/calmodulin-dependent protein kinase ID
Ctcfl	664799	1.99	CCCTC-binding factor (zinc finger protein)-like
Dazl	13164	2.39	deleted in azoospermia-like
Dcdc2a	195208	1.52	doublecortin domain containing 2a
Dnajc5g	231098	1.71	DnaJ (Hsp40) homolog, subfamily C, member 5 gamma
Efhc2	74405	3.58	EF-hand domain (C-terminal) containing 2
Fkbp6	94244	2.16	FK506 binding protein 6
Fthl17	83457	1.5	ferritin, heavy polypeptide-like 17
Gm13498	227885	2.38	predicted gene 13498
Gm2889	100040658	2.35	hypothetical protein LOC100041609; hypothetical protein LOC100044795; predicted gene 3395; similar to gag polyprotein; hypothetical protein LOC100047557; hypothetical protein LOC100040347; hypothetical protein LOC100044384; hypothetical protein LOC100045342; hypothetical protein LOC100038979; predicted gene 2889
Gm5128	331529	3.91	predicted gene 7903; predicted gene 5128
Gpa33	59290	1.15	glycoprotein A33 (transmembrane)
Gpat2	215456	1.34	RIKEN cDNA A530057A03 gene
Gpx6	75512	1.43	glutathione peroxidase 6
Gtsf1	74174	1.59	gametocyte specific factor 1
Hormad1	67981	2.71	HORMA domain containing 1; predicted gene 7167
Magea1	17137	2.42	melanoma antigen, family A, 1
Magea2	17138	2.01	melanoma antigen, family A, 2
Magea4	17140	1.86	melanoma antigen, family A, 4
Magea5	17141	1.87	melanoma antigen, family A, 5
Magea6	17142	1.58	melanoma antigen, family A, 6
Mov10l1	83456	1.46	Moloney leukemia virus 10-like 1
Nckap1l	105855	1.75	NCK associated protein 1 like
Nlrp4c	83564	1.58	NLR family, pyrin domain containing 4C
Nxf3	245610	1.61	nuclear RNA export factor 3
Olfr307	258610	2.17	olfactory receptor 307
Pde8a	18584	1.13	phosphodiesterase 8A; similar to cAMP-specific cyclic nucleotide phosphodiesterase PDE8; MMPDE8

Piwil2	57746	1.3	piwi-like homolog 2 (Drosophila)
Pnma5	385377	1.62	paraneoplastic antigen family 5
Ppp3ca	19055	1.05	protein phosphatase 3, catalytic subunit, alpha isoform
Pramel3	83565	3.75	preferentially expressed antigen in melanoma-like 3
Rhox13	73614	1.25	reproductive homeobox 13
Rhox2a	75199	1.27	reproductive homeobox 2C; reproductive homeobox 2B; reproductive homeobox 2A; reproductive homeobox 2E
Rhox2e	100040016	1.26	reproductive homeobox 2C; reproductive homeobox 2B; reproductive homeobox 2A; reproductive homeobox 2E
Rnf17	30054	1.06	ring finger protein 17
Scml2	107815	2.16	similar to sex comb on midleg-like 2 (Drosophila); sex comb on midleg-like 2 (Drosophila)
Smc1b	140557	2.47	structural maintenance of chromosomes 1B
Snx25	102141	1.42	sorting nexin 25
Sohlh2	74434	1.53	spermatogenesis and oogenesis specific basic helix-loop-helix 2
Spink10	328971	1.12	serine peptidase inhibitor, Kazal type 10
Stra8	20899	1.12	stimulated by retinoic acid gene 8
Taf9b	407786	1.46	TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor
Tuba3b	22147	1.55	predicted gene 5366; tubulin, alpha 3B; tubulin, alpha 3A
Wfdc15a	68221	1.82	WAP four-disulfide core domain 15A
Xlr3a	22445	3.16	hypothetical protein LOC100044094; similar to X-linked lymphocyte-regulated protein 3A; X-linked lymphocyte-regulated 3E, pseudogene; hypothetical protein LOC100044314; similar to X-LINKED LYMPHOCYTE-REGULATED PROTEIN 3A (XLR RELATED PROTEIN A12); X-linked lymphocyte-regulated 3C; X-linked lymphocyte-regulated 3A; X-linked lymphocyte-regulated 3B; X-linked lymphocyte-regulated 3D, pseudogene
Xlr3b	574437	3.32	hypothetical protein LOC100044094; similar to X-linked lymphocyte-regulated protein 3A; X-linked lymphocyte-regulated 3E, pseudogene; hypothetical protein LOC100044314; similar to X-LINKED LYMPHOCYTE-REGULATED PROTEIN 3A (XLR RELATED PROTEIN A12); X-linked lymphocyte-regulated 3C; X-linked lymphocyte-regulated 3A; X-linked lymphocyte-regulated 3B; X-linked lymphocyte-regulated 3D, pseudogene
Xlr3c	22446	3.36	hypothetical protein LOC100044094; similar to X-linked lymphocyte-regulated protein 3A; X-linked lymphocyte-regulated 3E, pseudogene; hypothetical protein LOC100044314; similar to X-LINKED LYMPHOCYTE-REGULATED PROTEIN 3A (XLR RELATED PROTEIN A12); X-linked lymphocyte-regulated 3C; X-linked lymphocyte-regulated 3A; X-linked lymphocyte-regulated 3B; X-linked lymphocyte-regulated 3D, pseudogene
Xlr4b	27083	3.09	X-linked lymphocyte-regulated 4D; X-linked lymphocyte-regulated 4E, pseudogene; X-linked lymphocyte-regulated 4B; X-linked lymphocyte-regulated 4C; hypothetical protein LOC100044049
Xlr4c	72891	3.34	X-linked lymphocyte-regulated 4D; X-linked lymphocyte-regulated 4E, pseudogene; X-linked lymphocyte-regulated 4B; X-linked lymphocyte-regulated 4C; hypothetical protein LOC100044049
Xlr5a	574438	4.63	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Xlr5b	627081	5.21	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719

Genes downregulated in TKO ESCs compared to wt ESCs			
Gene symbol	**Entrez Gene ID**	**Fold change to wt ESCs**	**Gene name**
Atxn1l	52335	-1.01	ataxin 1-like
Clcn5	12728	-1.12	chloride channel 5
Cyr61	16007	-2.14	cysteine rich protein 61
Dnmt1	13433	-1.27	DNA methyltransferase (cytosine-5) 1
Dnmt3a	13435	-1.62	DNA methyltransferase 3A
Dnmt3b	13436	-2.52	DNA methyltransferase 3B
Dpysl2	12934	-1.23	dihydropyrimidinase-like 2
Fkbp10	14230	-1.02	FK506 binding protein 10
Fndc3c1	333564	-2.86	fibronectin type III domain containing 3C1
Gja1	14609	-1.18	gap junction protein, alpha 1
Krt18	16668	-2.46	keratin 18
Krt19	16669	-2.9	keratin 19
Lpar6	67168	-1.66	purinergic receptor P2Y, G-protein coupled, 5
Lphn2	99633	-1.86	latrophilin 2
Lrrc8c	100604	-1.44	leucine rich repeat containing 8 family, member C
Mphosph6	68533	-1.03	M phase phosphoprotein 6; predicted gene 11448
Nes	18008	-1.9	nestin
Nnat	18111	-2.13	neuronatin
Ogfr	72075	-1.05	opioid growth factor receptor
Sema6a	20358	-2.38	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A
Sh3bgrl	56726	-1.02	SH3-binding domain glutamic acid-rich protein like

Sox4	20677	-1	SRY-box containing gene 19; SRY-box containing gene 4
Tagln	21345	-2.29	transgelin
Thbs1	21825	-2.61	thrombospondin 1; similar to thrombospondin 1
Tlcd1	68385	-1.09	TLC domain containing 1

Concordant upregulated genes in dnmt1-/- and TKO ESCs compared to wt ESCs				
Gene symbol	**Entrez Gene ID**	**Fold change dnmt1-/- to wt ESCs**	**Fold change TKO to wt ESCs**	**Gene name**
1700013H16Rik	75514	1.67	1.81	RIKEN cDNA 1700013H16 gene
AV320801	331531	2.85	3.61	expressed sequence AV320801
Dazl	13164	2.01	2.39	deleted in azoospermia-like
Dcdc2a	195208	1.34	1.52	doublecortin domain containing 2a
Dnajc5g	231098	1.36	1.71	DnaJ (Hsp40) homolog, subfamily C, member 5 gamma
Efhc2	74405	3.44	3.58	EF-hand domain (C-terminal) containing 2
Fkbp6	94244	1.81	2.16	FK506 binding protein 6
Fthl17	83457	1.65	1.5	ferritin, heavy polypeptide-like 17
Gm13498	227885	1.83	2.38	predicted gene 13498
Gm2889	100040658	2.1	2.35	hypothetical protein LOC100041609; hypothetical protein LOC100044795; predicted gene 3395; similar to gag polyprotein; hypothetical protein LOC100047557; hypothetical protein LOC100040347; hypothetical protein LOC100044384; hypothetical protein LOC100045342; hypothetical protein LOC100038979; predicted gene 2889
Gm5128	331529	3.08	3.91	predicted gene 7903; predicted gene 5128
Gpat2	215456	1.45	1.34	RIKEN cDNA A530057A03 gene
Gpx6	75512	1.23	1.43	glutathione peroxidase 6
Gtsf1	74174	1.67	1.59	gametocyte specific factor 1
Hormad1	67981	2.42	2.71	HORMA domain containing 1; predicted gene 7167
Magea1	17137	1.42	2.42	melanoma antigen, family A, 1
Nckap1l	105855	1.54	1.75	NCK associated protein 1 like
Nlrp4c	83564	1.61	1.58	NLR family, pyrin domain containing 4C
Nxf3	245610	1.72	1.61	nuclear RNA export factor 3
Olfr307	258610	2.51	2.17	olfactory receptor 307
Pnma5	385377	1.58	1.62	paraneoplastic antigen family 5
Pramel3	83565	2.96	3.75	preferentially expressed antigen in melanoma-like 3
Rnf17	30054	1.28	1.06	ring finger protein 17
Scml2	107815	1.64	2.16	similar to sex comb on midleg-like 2 (Drosophila); sex comb on midleg-like 2 (Drosophila)
Smc1b	140557	1.4	2.47	structural maintenance of chromosomes 1B
Sohlh2	74434	1.05	1.53	spermatogenesis and oogenesis specific basic helix-loop-helix 2
Wfdc15a	68221	1.85	1.82	WAP four-disulfide core domain 15A
Xlr3a	22445	3.35	3.16	hypothetical protein LOC100044094; similar to X-linked lymphocyte-regulated protein 3A; X-linked lymphocyte-regulated 3E, pseudogene; hypothetical protein LOC100044314; similar to X-LINKED LYMPHOCYTE-REGULATED PROTEIN 3A (XLR RELATED PROTEIN A12); X-linked lymphocyte-regulated 3C; X-linked lymphocyte-regulated 3A; X-linked lymphocyte-regulated 3B; X-linked lymphocyte-regulated 3D, pseudogene
Xlr3b	574437	3.48	3.32	hypothetical protein LOC100044094; similar to X-linked lymphocyte-regulated protein 3A; X-linked lymphocyte-regulated 3E, pseudogene; hypothetical protein LOC100044314; similar to X-LINKED LYMPHOCYTE-REGULATED PROTEIN 3A (XLR RELATED PROTEIN A12); X-linked lymphocyte-regulated 3C; X-linked lymphocyte-regulated 3A; X-linked lymphocyte-regulated 3B; X-linked lymphocyte-regulated 3D, pseudogene
Xlr3c	22446	3.54	3.36	hypothetical protein LOC100044094; similar to X-linked lymphocyte-regulated protein 3A; X-linked lymphocyte-regulated 3E, pseudogene; hypothetical protein LOC100044314; similar to X-LINKED LYMPHOCYTE-REGULATED PROTEIN 3A (XLR RELATED PROTEIN A12); X-linked lymphocyte-regulated 3C; X-linked lymphocyte-regulated 3A; X-linked lymphocyte-regulated 3B; X-linked lymphocyte-regulated 3D, pseudogene
Xlr4b	27083	3.07	3.09	X-linked lymphocyte-regulated 4D; X-linked lymphocyte-regulated 4E, pseudogene; X-linked lymphocyte-regulated 4B; X-linked lymphocyte-regulated 4C; hypothetical protein LOC100044049
Xlr4c	72891	3.36	3.34	X-linked lymphocyte-regulated 4D; X-linked lymphocyte-regulated 4E, pseudogene; X-linked lymphocyte-regulated 4B; X-linked lymphocyte-regulated 4C; hypothetical protein LOC100044049
Xlr5a	574438	4.48	4.63	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Xlr5b	627081	5.05	5.21	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719

Concordant downregulated genes in dnmt1-/- and TKO ESCs compared to wt ESCs				
Gene symbol	**Entrez Gene ID**	**Fold change dnmt1-/- to wt ESCs**	**Fold change TKO to wt ESCs**	**Gene name**
Dnmt1	13433	-1.02	-1.27	DNA methyltransferase (cytosine-5) 1
Lphn2	99633	-1.53	-1.86	latrophilin 2
Sema6a	20358	-2.12	-2.38	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A

6.2 Differentially regulated genes during the first differentiation period (day 0 -4)

Unique upregulated genes in wt EBs d0-4			
Gene symbol	**Entrez Gene ID**	**Fold change between d0_4**	**Gene Name**
1110012D08Rik	73827	1.15	RIKEN cDNA 1110012D08 gene
1500011H22Rik	68948	1.1	RIKEN cDNA 1500011H22 gene
2010011I20Rik	67017	1.17	RIKEN cDNA 2010011I20 gene
2410018L13Rik	69732	1.15	RIKEN cDNA 2410018L13 gene
3110062M04Rik	78412	1.25	RIKEN cDNA 3110062M04 gene
3632451O06Rik	67419	1.87	RIKEN cDNA 3632451O06 gene
4933433P14Rik	66787	1	RIKEN cDNA 4933433P14 gene
9430020K01Rik	240185	1.18	RIKEN cDNA 9430020K01 gene
Aacs	78894	1.28	acetoacetyl-CoA synthetase
Abcd2	26874	1.16	ATP-binding cassette, sub-family D (ALD), member 2
Abtb2	99382	1	ankyrin repeat and BTB (POZ) domain containing 2
Adam10	11487	1.4	a disintegrin and metallopeptidase domain 10
Adamts10	224697	1.59	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 10
Add3	27360	1.05	adducin 3 (gamma)
Adrbk2	320129	1.05	adrenergic receptor kinase, beta 2
Akr1c19	432720	1.1	aldo-keto reductase family 1, member C19
Akr1e1	56043	1.28	aldo-keto reductase family 1, member E1
Aldh1l1	107747	1.06	similar to Aldehyde dehydrogenase 1 family, member L1; aldehyde dehydrogenase 1 family, member L1
Amfr	23802	1.02	autocrine motility factor receptor
Angptl2	26360	1.03	angiopoietin-like 2
Ano1	101772	1.39	anoctamin 1, calcium activated chloride channel
Apbb1	11785	1.19	amyloid beta (A4) precursor protein-binding, family B, member 1
Aph1b	208117	1.19	anterior pharynx defective 1b homolog (C. elegans)
Arhgap29	214137	1.46	Rho GTPase activating protein 29
Arid3a	13496	1.29	AT rich interactive domain 3A (BRIGHT-like)
Asb12	70392	2.4	ankyrin repeat and SOCS box-containing 12
Atp2b4	381290	1.06	ATPase, Ca++ transporting, plasma membrane 4
AV249152	216560	1.01	expressed sequence AV249152
AW551984	244810	1.83	expressed sequence AW551984
B930095G15Rik	320268	1.29	RIKEN cDNA B930095G15 gene
Bag2	213539	1.08	BCL2-associated athanogene 2
BC005764	216152	1.22	cDNA sequence BC005764
BC051142	407788	1.15	cDNA sequence BC051142; testis specific basic protein
Bmf	171543	1.01	BCL2 modifying factor
Bmp4	12159	1.88	bone morphogenetic protein 4
Bmp7	12162	1.5	bone morphogenetic protein 7
Ccdc136	232664	1.28	coiled-coil domain containing 136
Ccdc46	76380	1.15	coiled-coil domain containing 46
Ccdc8	434130	1.1	coiled-coil domain containing 8
Ccl25	20300	1.01	chemokine (C-C motif) ligand 25
Cd55	13136	2	CD55 antigen
Cdc42ep5	58804	1.03	CDC42 effector protein (Rho GTPase binding) 5
Cdkn1b	12576	1.15	cyclin-dependent kinase inhibitor 1B

Chst2	54371	1.12	carbohydrate sulfotransferase 2
Col4a1	12826	1.33	collagen, type IV, alpha 1
Col4a2	12827	1.14	collagen, type IV, alpha 2
Col4a6	94216	1.36	collagen, type IV, alpha 6
Copz2	56358	1.56	coatomer protein complex, subunit zeta 2
Cpm	70574	2.51	carboxypeptidase M
Creb3l2	208647	1.02	cAMP responsive element binding protein 3-like 2
Ctsh	13036	2.59	cathepsin H
Ctso	229445	1.53	cathepsin O
Ctxn1	330695	1.02	cortexin 1
Cyp51	13121	1.49	cytochrome P450, family 51
D10Ertd610e	52666	1.19	DNA segment, Chr 10, ERATO Doi 610, expressed
D130062J21Rik	100038651	1.5	RIKEN cDNA D130062J21 gene
D430042O09Rik	233865	1.04	RIKEN cDNA D430042O09 gene
Dapk2	13143	1.53	death-associated protein kinase 2
Ddah2	51793	1.32	dimethylarginine dimethylaminohydrolase 2
Dhrs7	66375	1.31	dehydrogenase/reductase (SDR family) member 7
Dnajc12	30045	1.06	DnaJ (Hsp40) homolog, subfamily C, member 12
Dynlt3	67117	1.09	dynein light chain Tctex-type 3
E130112L23Rik	268739	1.05	RIKEN cDNA E130112L23 gene
Edil3	13612	1.8	EGF-like repeats and discoidin I-like domains 3
Egflam	268780	1.06	EGF-like, fibronectin type III and laminin G domains
Egln3	112407	1.89	EGL nine homolog 3 (C. elegans)
Elovl1	54325	1.37	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 1
Emp2	13731	1.39	epithelial membrane protein 2
Emp3	13732	1.03	epithelial membrane protein 3
Erbb2	13866	1.14	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
Fam114a1	68303	1.38	family with sequence similarity 114, member A1
Fam125b	72543	1.38	family with sequence similarity 125, member B; similar to RIKEN cDNA 2610528K11 gene
Fam126a	84652	1.2	family with sequence similarity 126, member A
Fam57a	116972	1.37	family with sequence similarity 57, member A
Fam65a	75687	1.05	family with sequence similarity 65, member A
Fbln1	14114	1.38	fibulin 1
Fbxl20	72194	1.05	F-box and leucine-rich repeat protein 20
Fgf15	14170	1.25	fibroblast growth factor 15
Fgfr2	14183	1.53	fibroblast growth factor receptor 2
Fhdc1	229474	1.26	FH2 domain containing 1
Flrt3	71436	1.44	fibronectin leucine rich transmembrane protein 3; similar to fibronectin leucine rich transmembrane protein 3
Fut8	53618	1.03	fucosyltransferase 8
Fzd4	14366	1.77	frizzled homolog 4 (Drosophila)
Gabre	14404	1.73	similar to gamma-aminobutyric acid (GABA-A) receptor, subunit epsilon; gamma-aminobutyric acid (GABA) A receptor, subunit epsilon
Gapdhs	14447	1.06	glyceraldehyde-3-phosphate dehydrogenase, spermatogenic
Gm10554	100038541	1.19	predicted gene 10554
Gm10561	628004	1.28	predicted gene 10561
Gm11818	208820	1.02	predicted gene 11818
Gm4983	245297	1.17	predicted gene 4983
Gng12	14701	1.18	guanine nucleotide binding protein (G protein), gamma 12
Golim4	73124	1.04	golgi integral membrane protein 4
Gpr124	78560	2.07	G protein-coupled receptor 124
Grik5	14809	1.04	glutamate receptor, ionotropic, kainate 5 (gamma 2)
Grina	66168	1.28	glutamate receptor, ionotropic, N-methyl D-aspartate-associated protein 1 (glutamate binding)
Gucy1b3	54195	1.71	guanylate cyclase 1, soluble, beta 3
H2-Ab1	14961	1.57	histocompatibility 2, class II antigen A, beta 1; response to metastatic cancers 2; similar to H-2 class II histocompatibility antigen, A-D beta chain precursor
Herc1	235439	1.16	hect (homologous to the E6-AP (UBE3A) carboxyl terminus) domain and RCC1 (CHC1)-like domain (RLD) 1
Hey2	15214	2.47	hairy/enhancer-of-split related with YRPW motif 2
Hgsnat	52120	1.12	heparan-alpha-glucosaminide N-acetyltransferase
Hmgcs1	208715	1.28	similar to Hmgcs1 protein; 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1

Hook3	320191	1.06	hook homolog 3 (Drosophila)
Hoxd1	15429	1.33	homeo box D1
Hpcal1	53602	1.01	hippocalcin-like 1
Hs3st3b1	54710	1.12	heparan sulfate (glucosamine) 3-O-sulfotransferase 3B1
Hsd17b7	15490	1.08	hydroxysteroid (17-beta) dehydrogenase 7
Idh1	15926	1.11	isocitrate dehydrogenase 1 (NADP+), soluble
Igfbp5	16011	1.67	insulin-like growth factor binding protein 5
Inpp4a	269180	1.07	inositol polyphosphate-4-phosphatase, type I
Inpp5f	101490	1.1	inositol polyphosphate-5-phosphatase F
Ipp	16351	1.1	IAP promoted placental gene
Itga4	16401	1.63	integrin alpha 4
Itgb5	16419	1.05	integrin beta 5
Itm2a	16431	1.37	integral membrane protein 2A
Katnal2	71206	1.03	katanin p60 subunit A-like 2
Kcnq5	226922	1.28	potassium voltage-gated channel, subfamily Q, member 5
Kdelr3	105785	1.12	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3
Kitl	17311	1.39	kit ligand
Klk1b22	13646	1.46	kallikrein 1-related peptidase b22
L1cam	16728	1.06	L1 cell adhesion molecule
Large	16795	1.29	like-glycosyltransferase
Lass4	67260	1.5	LAG1 homolog, ceramide synthase 4
Letmd1	68614	1.17	LETM1 domain containing 1
Lipa	16889	1.11	lysosomal acid lipase A
LOC677548	677548	1.09	similar to Hippocalcin-like protein 1 (Visinin-like protein 3) (VILIP-3) (Neural visinin-like protein 3) (NVL-3) (NVP-3)
Lrig2	269473	1.02	leucine-rich repeats and immunoglobulin-like domains 2
Lrp4	228357	1.04	low density lipoprotein receptor-related protein 4
Lss	16987	1.02	lanosterol synthase
Maged1	94275	1.17	melanoma antigen, family D, 1
Man2a1	17158	1.12	mannosidase 2, alpha 1
Map2k6	26399	1.11	mitogen-activated protein kinase kinase 6
Med12	59024	1.09	mediator of RNA polymerase II transcription, subunit 12 homolog (yeast)
Mex3b	108797	1.25	mex3 homolog B (C. elegans)
Mllt3	70122	1.23	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 3
Mmp15	17388	1.79	matrix metallopeptidase 15
Mpp2	50997	1.08	membrane protein, palmitoylated 2 (MAGUK p55 subfamily member 2)
Mvd	192156	1.18	mevalonate (diphospho) decarboxylase
Mxra8	74761	1.69	matrix-remodelling associated 8
Myo1b	17912	1.17	myosin IB
Narf	67608	1.33	nuclear prelamin A recognition factor
Ncrna00086	320237	1.06	non-protein coding RNA 86
Ndrg1	17988	1.28	N-myc downstream regulated gene 1
Neil3	234258	1.54	nei like 3 (E. coli)
Nelf	56876	1.13	nasal embryonic LHRH factor
Nfatc4	73181	1.05	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 4
Npc2	67963	1.14	Niemann Pick type C2
Npr2	230103	1.05	natriuretic peptide receptor 2
Nsmaf	18201	1.06	neutral sphingomyelinase (N-SMase) activation associated factor
P4ha1	18451	1.85	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha 1 polypeptide
P4ha2	18452	1.72	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha II polypeptide
Pbx2	18515	1.42	pre B-cell leukemia transcription factor 2
Pde3a	54611	1.21	phosphodiesterase 3A, cGMP inhibited
Pdlim4	30794	1.37	PDZ and LIM domain 4
Phc3	241915	1.09	polyhomeotic-like 3 (Drosophila)
Pik3ip1	216505	1.14	phosphoinositide-3-kinase interacting protein 1
Pja1	18744	1.07	praja1, RING-H2 motif containing
Pja2	224938	1.18	praja 2, RING-H2 motif containing
Plekha6	240753	1.09	pleckstrin homology domain containing, family A member 6
Ppap2b	67916	1.03	phosphatidic acid phosphatase type 2B
Ppil6	73075	1.22	peptidylprolyl isomerase (cyclophilin)-like 6

Ppp1r3c	53412	1.23	protein phosphatase 1, regulatory (inhibitor) subunit 3C
Ppp3ca	19055	1.69	protein phosphatase 3, catalytic subunit, alpha isoform
Prex1	277360	1.09	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 1
Prmt2	15468	1.03	protein arginine N-methyltransferase 2
Pros1	19128	1.07	protein S (alpha)
Prrg1	546336	1.1	proline rich Gla (G-carboxyglutamic acid) 1
Purg	75029	1.05	purine-rich element binding protein G
Rabgap1l	29809	1.17	RAB GTPase activating protein 1-like
Rapgef5	217944	1.05	Rap guanine nucleotide exchange factor (GEF) 5
Rdh11	17252	1.15	retinol dehydrogenase 11
Reln	19699	1.64	reelin
Ripk2	192656	1.19	receptor (TNFRSF)-interacting serine-threonine kinase 2
Rnpep	215615	1.06	arginyl aminopeptidase (aminopeptidase B)
Rras	20130	1.07	Harvey rat sarcoma oncogene, subgroup R
S100a1	20193	1.01	S100 calcium binding protein A1
Samd14	217125	1.02	sterile alpha motif domain containing 14
Sc4mol	66234	1.15	sterol-C4-methyl oxidase-like
Sc5d	235293	1.43	sterol-C5-desaturase (fungal ERG3, delta-5-desaturase) homolog (S. cerevisae)
Scarna17	100217466	1.03	small Cajal body-specific RNA 17
Scd1	20249	1.12	stearoyl-Coenzyme A desaturase 1
Scn7a	20272	1.14	sodium channel, voltage-gated, type VII, alpha
Sept10	103080	1.28	septin 10
Sertad4	214791	1.27	SERTA domain containing 4
Sesn1	140742	1.11	sestrin 1
Sh3yl1	24057	1.16	Sh3 domain YSC-like 1
Sirpa	19261	1.26	signal-regulatory protein alpha
Slc25a10	27376	1.01	solute carrier family 25 (mitochondrial carrier, dicarboxylate transporter), member 10
Slc38a4	69354	1.26	solute carrier family 38, member 4
Slc5a7	63993	1.43	solute carrier family 5 (choline transporter), member 7
Slit2	20563	1.72	slit homolog 2 (Drosophila)
Smpd1	20597	1.05	sphingomyelin phosphodiesterase 1, acid lysosomal
Snai2	20583	2.82	snail homolog 2 (Drosophila)
Snx25	102141	1.11	sorting nexin 25
Spata6	67946	1.14	spermatogenesis associated 6
Spo11	26972	1.1	sporulation protein, meiosis-specific, SPO11 homolog (S. cerevisiae)
Sprr1a	20753	1.28	small proline-rich protein 1A
St3gal5	20454	1.6	ST3 beta-galactoside alpha-2,3-sialyltransferase 5
St6galnac3	20447	1.03	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3
St6galnac4	20448	1.1	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4
Stard4	170459	1.37	StAR-related lipid transfer (START) domain containing 4
Stx2	13852	1.22	syntaxin 2
Taz	66826	1.14	tafazzin
Tbx2	21385	1.31	T-box 2
Tdo2	56720	4.21	tryptophan 2,3-dioxygenase
Tgfbr3	21814	1.11	transforming growth factor, beta receptor III
Thra	21833	1.1	thyroid hormone receptor alpha; similar to thyroid hormone receptor
Tiam2	24001	1.04	T-cell lymphoma invasion and metastasis 2
Tm7sf2	73166	1.42	transmembrane 7 superfamily member 2
Tmem119	231633	1.42	transmembrane protein 119
Tmem176a	66058	1.06	transmembrane protein 176A
Tmx1	72736	1.24	thioredoxin-related transmembrane protein 1
Tnnc1	21924	1.53	troponin C, cardiac/slow skeletal
Ttc12	235330	1.01	tetratricopeptide repeat domain 12; similar to tetratricopeptide repeat domain 12
Ttc3	22129	1.27	tetratricopeptide repeat domain 3
Ttc8	76260	1.08	tetratricopeptide repeat domain 8
Ttr	22139	2.15	transthyretin
Twist1	22160	1.88	twist homolog 1 (Drosophila)
Twist2	13345	1.41	twist homolog 2 (Drosophila)
Vegfa	22339	1.29	vascular endothelial growth factor A

Vim	22352	1.48	vimentin
Wbp1	22377	1.05	WW domain binding protein 1
Wdr19	213081	1	WD repeat domain 19
Whsc1	107823	1.08	Wolf-Hirschhorn syndrome candidate 1 (human)
Zc3hav1l	209032	1.2	zinc finger CCCH-type, antiviral 1-like
Zcchc24	71918	1.5	zinc finger, CCHC domain containing 24
Zeb1	21417	1.1	zinc finger E-box binding homeobox 1
Zfp395	380912	1.04	zinc finger protein 395
Zfp516	329003	1.99	zinc finger protein 516
Zhx1	22770	1.06	zinc fingers and homeoboxes 1
Zim1	22776	1.28	zinc finger, imprinted 1

Unique downregulated genes in wt EBs d0-4			
Gene symbol	**Entrez Gene ID**	**Fold change between d0_4**	**Gene Name**
1110012J17Rik	68617	-1.03	RIKEN cDNA 1110012J17 gene
1190005I06Rik	68918	-1.15	RIKEN cDNA 1190005I06 gene
1700020N01Rik	67692	-1.12	RIKEN cDNA 1700020N01 gene
1700024J04Rik	71848	-1.09	RIKEN cDNA 1700024J04 gene
2810474O19Rik	67246	-1.05	RIKEN cDNA 2810474O19 gene
4930572J05Rik	223626	-1.51	RIKEN cDNA 4930572J05 gene
4933437F05Rik	71275	-1.32	RIKEN cDNA 4933437F05 gene
Abca1	11303	-1.51	ATP-binding cassette, sub-family A (ABC1), member 1
Abcb1b	18669	-1.44	ATP-binding cassette, sub-family B (MDR/TAP), member 1B
AI747699	381236	-1.2	predicted gene 8981; similar to lipase-like, ab-hydrolase domain containing 2; expressed sequence AI747699
Aim2	383619	-2.36	absent in melanoma 2
Akap1	11640	-1.01	A kinase (PRKA) anchor protein 1
Angptl4	57875	-1.03	angiopoietin-like 4
Anxa1	16952	-2.34	annexin A1
Arl4a	11861	-1.06	ADP-ribosylation factor-like 4A
As3mt	57344	-1.15	arsenic (+3 oxidation state) methyltransferase
Atp10a	11982	-1.48	ATPase, class V, type 10A
Atp1b1	11931	-1.77	ATPase, Na+/K+ transporting, beta 1 polypeptide
B3gnt2	53625	-1.34	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2
BB287469	544881	-2.75	predicted gene 6804; predicted gene 2046; predicted gene 8607; expressed sequence BB287469; predicted gene 2075; predicted gene 2022; predicted gene 4027
Bcam	57278	-1.59	basal cell adhesion molecule
Bcat1	12035	-1.44	branched chain aminotransferase 1, cytosolic; similar to branched chain aminotransferase 1, cytosolic
Bdh2	69772	-1.56	3-hydroxybutyrate dehydrogenase, type 2
Bend3	331623	-1.08	BEN domain containing 3
Bmp8b	12164	-1.02	bone morphogenetic protein 8b
Bspry	192120	-1.07	B-box and SPRY domain containing
C230052I12Rik	101831	-1.26	RIKEN cDNA C230052I12 gene
Cacna1a	12286	-1.05	calcium channel, voltage-dependent, P/Q type, alpha 1A subunit
Capsl	75568	-1.29	calcyphosine-like
Cbr3	109857	-1.84	carbonyl reductase 3
Cd34	12490	-1	CD34 antigen
Cdh1	12550	-1.37	cadherin 1
Cenpv	73139	-1.32	centromere protein V
Chek2	50883	-1.05	CHK2 checkpoint homolog (S. pombe)
Cldn4	12740	-1.27	claudin 4
Cngb1	333329	-1.04	cyclic nucleotide gated channel beta 1
Col18a1	12822	-1.28	collagen, type XVIII, alpha 1
Ctgf	14219	-1.81	connective tissue growth factor
Cth	107869	-1.31	cystathionase (cystathionine gamma-lyase)
Cxcl16	66102	-1.53	chemokine (C-X-C motif) ligand 16
Cyp2j13	230459	-1.09	cytochrome P450, family 2, subfamily j, polypeptide 13

D6Mm5e	110958	-1.02	DNA segment, Chr 6, Miriam Meisler 5, expressed
Dbx1	13172	-1.56	developing brain homeobox 1
Ddx4	13206	-2.28	DEAD (Asp-Glu-Ala-Asp) box polypeptide 4
Depdc6	97998	-1.07	DEP domain containing 6
Depdc7	211896	-1.75	DEP domain containing 7
Dnahc8	13417	-1.33	similar to axonemal dynein heavy chain 8 long form; similar to dynein, axonemal, heavy chain 8; dynein, axonemal, heavy chain 8
Dnmt3a	13435	-1.02	DNA methyltransferase 3A
Dtx3l	209200	-1.25	deltex 3-like (Drosophila)
Dusp14	56405	-1.09	dual specificity phosphatase 14
Dyrk3	226419	-1.03	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 3
E130012A19Rik	103551	-1.18	RIKEN cDNA E130012A19 gene
E2f8	108961	-1.22	E2F transcription factor 8
Edn1	13614	-1.38	endothelin 1
EG665955	665955	-1.03	predicted gene, EG665955
Elavl2	15569	-1.08	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2 (Hu antigen B)
Emp1	13730	-1.81	epithelial membrane protein 1
Eps8	13860	-1.01	epidermal growth factor receptor pathway substrate 8
F3	14066	-1.27	coagulation factor III
Fam111a	107373	-1.95	RIKEN cDNA 4632417K18 gene
Fancl	67030	-1.14	similar to Fanconi anemia, complementation group L; Fanconi anemia, complementation group L
Fetub	59083	-2.53	fetuin beta
Fgd1	14163	-1.14	FYVE, RhoGEF and PH domain containing 1
Fgf18	14172	-1.06	fibroblast growth factor 18
Fgfbp1	14181	-1.45	fibroblast growth factor binding protein 1
Fmr1nb	207854	-1.56	fragile X mental retardation 1 neighbor
Foxd3	15221	-1.13	forkhead box D3
Foxi3	232077	-1.12	forkhead box I3
Foxo1	56458	-1.6	forkhead box O1
Foxp1	108655	-1.45	forkhead box P1
Gadd45a	13197	-1.72	growth arrest and DNA-damage-inducible 45 alpha
Gadd45b	17873	-1.13	growth arrest and DNA-damage-inducible 45 beta
Gclm	14630	-1.19	glutamate-cysteine ligase, modifier subunit
Gm10785	100038392	-1.15	predicted gene 10785
Gm16368	100039226	-1.98	predicted gene 2056; predicted gene 5442; predicted gene 8300; predicted gene 5788; predicted gene, 100039042; similar to X-linked eukaryotic translation initiation factor 1A; predicted gene 8332; predicted gene 2035; predicted gene 16368
Gm4340	100043292	-3.38	predicted gene 4340
Gm5712	435755	-1.07	predicted gene 5712
Gm6723	626952	-1.1	L antigen family, member 3; predicted gene 6723; L antigen family, member 3 pseudogene
Gm7682	665551	-1.9625	predicted gene 7682
Gm9350	668774	-1.56	predicted gene 9350
Gm9359	668786	-1.44	tripartite motif-containing 13; predicted gene 9359
Gm98	225908	-1.25	predicted gene 98
Gpc4	14735	-1.35	glypican 4; similar to Glypican 4
Grtp1	66790	-1.34	GH regulated TBC protein 1
Gstm3	14864	-1.15	glutathione S-transferase, mu 3
Hist1h2aa	319163	-2.57	histone cluster 1, H2aa
Hmgxb4	70823	-1.01	similar to Hmgb2l1 protein; HMG box domain containing 4
Hook1	77963	-1.16	hook homolog 1 (Drosophila)
Hs6st1	50785	-1.22	heparan sulfate 6-O-sulfotransferase 1
Igfbp2	16008	-1.18	insulin-like growth factor binding protein 2
Il33	77125	-1.12	interleukin 33
Itga3	16400	-2.24	integrin alpha 3
Jup	16480	-1.08	junction plakoglobin
Krtdap	64661	-1.27	keratinocyte differentiation associated protein
Laptm5	16792	-2.01	lysosomal-associated protein transmembrane 5
Lefty2	320202	-3.32	left-right determination factor 2
Lgals9	16859	-1.13	lectin, galactose binding, soluble 9
Lgr4	107515	-1.27	leucine-rich repeat-containing G protein-coupled receptor 4
LOC624931	624931	-1.955	similar to D5Ertd577e protein

Lrp2	14725	-2.41	low density lipoprotein receptor-related protein 2
Mal2	105853	-1.15	mal, T-cell differentiation protein 2
Mfap5	50530	-1.15	microfibrillar associated protein 5
Mgat4a	269181	-1.26	mannoside acetylglucosaminyltransferase 4, isoenzyme A
Mlkl	74568	-1.47	mixed lineage kinase domain-like
Mov10l1	83456	-1.11	Moloney leukemia virus 10-like 1
Mpzl2	14012	-1.72	myelin protein zero-like 2
Mt2	17750	-1.2	metallothionein 2
Mtap7d2	78283	-2	MAP7 domain containing 2
Mtap7d3	320923	-1.21	MAP7 domain containing 3
Nckap1l	105855	-1.6	NCK associated protein 1 like
Ngfr	18053	-1.31	nerve growth factor receptor (TNFR superfamily, member 16)
Nhedc1	74446	-1.05	Na+/H+ exchanger domain containing 1
Nlrp14	76858	-1.52	NLR family, pyrin domain containing 14
Nol6	230082	-1.14	nucleolar protein family 6 (RNA-associated)
Nt5e	23959	-1.17	5' nucleotidase, ecto
Ntn1	18208	-1.52	similar to Netrin-1 precursor; netrin 1
Nuak2	74137	-1.16	NUAK family, SNF1-like kinase, 2
Olfr161	258859	-1.17	olfactory receptor 161
Phyhipl	70911	-1.9	phytanoyl-CoA hydroxylase interacting protein-like
Plaur	18793	-1.03	plasminogen activator, urokinase receptor
Ppap2a	19012	-1.42	phosphatidic acid phosphatase type 2A
Ppif	105675	-1.01	peptidylprolyl isomerase F (cyclophilin F)
Pramef12	77632	-1.01	PRAME family member 12
Prex2	109294	-2.37	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 2
Prps1	19139	-1.42	mirror-image polydactyly gene 1 homolog (human); phosphoribosyl pyrophosphate synthetase 1; phosphoribosyl pyrophosphate synthetase 1-like 1
Ptgs2	19225	-1.37	prostaglandin-endoperoxide synthase 2
Ptpn3	545622	-1.37	protein tyrosine phosphatase, non-receptor type 3
Rasd2	75141	-1.191111111	RASD family, member 2
Rasgrp1	19419	-1.02	RAS guanyl releasing protein 1
Rnd1	223881	-1	Rho family GTPase 1
Rnf138	56515	-1.05	ring finger protein 138
Rps4y2	66184	-1.305	predicted gene 6816; ribosomal protein S4, Y-linked 2
Rrp9	27966	-1.14	RRP9, small subunit (SSU) processome component, homolog (yeast)
Scand3	71970	-1	SCAN domain containing 3
Sec24d	69608	-1.65	Sec24 related gene family, member D (S. cerevisiae)
Sema3e	20349	-1.07	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E; hypothetical protein LOC100044162
Serpine1	18787	-1.17	serine (or cysteine) peptidase inhibitor, clade E, member 1
Serpini1	20713	-1.17	serine (or cysteine) peptidase inhibitor, clade I, member 1
Sesn2	230784	-1.18	sestrin 2
Set	56086	-1.35	predicted gene, EG625349; predicted gene 5789; predicted gene 7085; predicted gene 5708; predicted gene 6847; SET translocation; cDNA sequence BC085271; predicted gene 7239; similar to protein phosphatase 2A inhibitor-2 I-2PP2A; predicted gene 9531
Sfmbt2	353282	-1.03	Scm-like with four mbt domains 2
Sfn	55948	-2.305	predicted gene 5279; stratifin; similar to 14-3-3 protein sigma (Stratifin); predicted gene 7850
Sgk1	20393	-1.34	serum/glucocorticoid regulated kinase 1
Six6os1	75801	-1.03	Six6 opposite strand transcript 1
Slc25a31	73333	-1.53	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 31
Slc35f2	72022	-1.94	solute carrier family 35, member F2
Slc7a5	20539	-1.06	similar to solute carrier family 7 (cationic amino acid transporter, y+ system), member 5; similar to Solute carrier family 7 (cationic amino acid transporter, y+ system), member 5; solute carrier family 7 (cationic amino acid transporter, y+ system), member 5
Sms	20603	-1.235	predicted gene 7270; predicted gene 14680; spermine synthase
Spna1	20739	-1.85	spectrin alpha 1
Ston2	108800	-1.1	stonin 2
Sycp1	20957	-1.21	synaptonemal complex protein 1; similar to testicular protein
Sycp3	20962	-1.79	synaptonemal complex protein 3
Syngr1	20972	-1.35	synaptogyrin 1
Syt4	20983	-1.21	synaptotagmin IV

Taf4a	228980	-1.07	TAF4A RNA polymerase II, TATA box binding protein (TBP)-associated factor; similar to TAF4A RNA polymerase II, TATA box binding protein (TBP)-associated factor
Tagln	21345	-1.46	transgelin
Tcf3	21423	-1.04	transcription factor E2a
Tcof1	21453	-1.1	Treacher Collins Franceschetti syndrome 1, homolog
Thbs1	21825	-1.85	thrombospondin 1; similar to thrombospondin 1
Tmem106a	217203	-1.05	transmembrane protein 106A
Tmem30b	238257	-1.02	transmembrane protein 30B
Tmem54	66260	-1.62	transmembrane protein 54
Tmem79	71913	-1.02	transmembrane protein 79
Trh	22044	-1.15	thyrotropin releasing hormone
Ube1y1	22202	-1.69	similar to ubiquitin activating enzyme E1; ubiquitin-activating enzyme E1, Chr Y 1
Ubl3	24109	-1.09	ubiquitin-like 3
Ugt8a	22239	-1.69	UDP galactosyltransferase 8A
Usp44	327799	-1.04	ubiquitin specific peptidase 44
Utp14a	72554	-1.03	UTP14, U3 small nucleolar ribonucleoprotein, homolog A (yeast)
Wsb2	59043	-1.36	WD repeat and SOCS box-containing 2
X99384	27355	-1.04	cDNA sequence X99384
Xrcc5	22596	-1.21	X-ray repair complementing defective repair in Chinese hamster cells 5
Zdhhc23	332175	-1.24	zinc finger, DHHC domain containing 23
Zfp345	545471	-1.37	zinc finger protein 345
Zfp36	22695	-1.07	zinc finger protein 36
Zfp518	72672	-1.03	zinc finger protein 518
Zfp97	22759	-1.06	zinc finger protein 97; cDNA sequence BC018101
Zfy1	22767	-1.36	zinc finger protein 1, Y linked
Zic5	65100	-1.93	similar to zinc finger protein of the cerebellum 5; predicted gene 12241; zinc finger protein of the cerebellum 5
Zmiz2	52915	-1.08	zinc finger, MIZ-type containing 2
Zscan4c	245109	-3.723333333	zinc finger and SCAN domain containing 4F; zinc finger and SCAN domain containing 4E; zinc finger and SCAN domain containing 4D; expresssed sequence BQ559217; zinc finger and SCAN domain containing 4C; zinc finger and SCAN domain containing 4B; zinc finger and SCAN domain containing 4, pseudogene 1; predicted gene 4186; zinc finger and SCAN domain containing 4, pseudogene 2; zinc finger and SCAN domain containing 4, pseudogene 3; similar to Gene model 397, (NCBI)

Upregulated genes in dnmt1-/- EBs compared to wt EBs between d0-4			
Gene symbol	**Entrez Gene ID**	**Fold change compared to wt**	**Gene name**
1700048O20Rik	69430	1.06	RIKEN cDNA 1700048O20 gene
1810011O10Rik	69068	1.47	RIKEN cDNA 1810011O10 gene
2310014D11Rik	69633	1.23	RIKEN cDNA 2310014D11 gene
2610034B18Rik	70420	1.4	RIKEN cDNA 2610034B18 gene
2700046A07Rik	78449	1.42	RIKEN cDNA 2700046A07 gene
3110007F17Rik	73061	1.57	predicted gene 5945; RIKEN cDNA 3110007F17 gene; predicted gene 2411; predicted gene 5167; predicted gene 6604; predicted gene 14957
3830403N18Rik	70691	2.6	RIKEN cDNA 3830403N18 gene
5930434B04Rik	381356	1.06	RIKEN cDNA 5930434B04 gene; hypothetical protein LOC100047034
A130022J15Rik	101351	1.52	RIKEN cDNA A130022J15 gene
Abca5	217265	1.07	ATP-binding cassette, sub-family A (ABC1), member 5
Acp2	11432	1.01	acid phosphatase 2, lysosomal
Acss3	380660	1.04	acyl-CoA synthetase short-chain family member 3
Acvr1b	11479	1	activin A receptor, type 1B
Adcy2	210044	1.02	adenylate cyclase 2
Adra2b	11552	2.11	adrenergic receptor, alpha 2b
Agpat6	102247	1.09	1-acylglycerol-3-phosphate O-acyltransferase 6 (lysophosphatidic acid acyltransferase, zeta)
Antxr1	69538	1.12	anthrax toxin receptor 1
Arg2	11847	1.2	arginase type II
Arhgef9	236915	1.15	CDC42 guanine nucleotide exchange factor (GEF) 9
Asz1	74068	1.45	ankyrin repeat, SAM and basic leucine zipper domain containing 1
Bach1	12013	1.05	BTB and CNC homology 1
Baiap2l1	66898	1	BAI1-associated protein 2-like 1
Baz2b	407823	1.01	bromodomain adjacent to zinc finger domain, 2B

Calcrl	54598	2.04	calcitonin receptor-like
Cald1	109624	1.37	caldesmon 1
Camk2d	108058	1.18	calcium/calmodulin-dependent protein kinase II, delta
Capn2	12334	1.72	calpain 2
Car13	71934	1.42	carbonic anhydrase 13
Car3	12350	1.2	carbonic anhydrase 3
Cdh5	12562	1.56	cadherin 5
Cited2	17684	1.36	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2
Clcn5	12728	1.55	chloride channel 5
Cmtm8	70031	1.26	CKLF-like MARVEL transmembrane domain containing 8
Cnrip1	380686	1.65	cannabinoid receptor interacting protein 1
Cpd	12874	1.44	carboxypeptidase D; similar to carboxypeptidase D
Cpe	12876	1.93	carboxypeptidase E; similar to carboxypeptidase E
Cpxm2	55987	1.56	carboxypeptidase X 2 (M14 family)
Ctsc	13032	2.71	cathepsin C
Ctse	13034	1.58	cathepsin E
Cxadr	13052	1.13	coxsackie virus and adenovirus receptor
Cxcr7	12778	1	chemokine (C-X-C motif) receptor 7
Cxx1b	553127	1.395	CAAX box 1 homolog A (human); CAAX box 1 homolog B (human); similar to mammalian retrotransposon derived 8b
Cxx1c	72865	1.87	CAAX box 1 homolog C (human)
Cyp2j6	13110	1.25	cytochrome P450, family 2, subfamily j, polypeptide 6
D830030K20Rik	320333	1.135	predicted gene 2971; predicted gene 3532; predicted gene 3033; RIKEN cDNA D830030K20 gene; predicted gene 3278; predicted gene 8271; predicted gene 2244; predicted gene 3043; predicted gene 3537; predicted gene 10408; predicted gene 8050; predicted gene 3755
Dact1	59036	1.09	dapper homolog 1, antagonist of beta-catenin (xenopus)
Dclk1	13175	1.7	doublecortin-like kinase 1
Dock11	75974	1.17	dedicator of cytokinesis 11
Dpysl5	65254	1.16	dihydropyrimidinase-like 5
Dub1	13531	1.18	deubiquitinating enzyme 1; similar to DUB-1
Dusp22	105352	1.01	dual specificity phosphatase 22
Dusp4	319520	2.2	dual specificity phosphatase 4
Efna5	13640	1.1	ephrin A5
Efnb2	13642	1.52	ephrin B2
Eif4g3	230861	1.11	eukaryotic translation initiation factor 4 gamma, 3; similar to Eukaryotic translation initiation factor 4 gamma 3 (eIF-4-gamma 3) (eIF-4G 3) (eIF4G 3) (eIF-4-gamma II) (eIF4GII)
Elmod2	244548	1.03	ELMO domain containing 2
Eml3	225898	1.06	echinoderm microtubule associated protein like 3
ENSMUSG00000079376	545001	1.17	predicted gene, ENSMUSG00000079376
Epb4.1l3	13823	1.81	erythrocyte protein band 4.1-like 3
Erbb2ip	59079	1.01	Erbb2 interacting protein
F2r	14062	1.97	coagulation factor II (thrombin) receptor
Fabp7	12140	1.88	fatty acid binding protein 7, brain
Fam135a	68187	1.12	family with sequence similarity 135, member A
Fam149a	212326	1.04	family with sequence similarity 149, member A
Fam163a	329274	1.32	family with sequence similarity 163, member A
Fam38b	667742	1.32	family with sequence similarity 38, member B2
Fam64a	109212	1.14	RIKEN cDNA 6720460F02 gene
Fgf5	14176	1.86	fibroblast growth factor 5
Flnc	68794	1.49	filamin C, gamma
Fn1	14268	1.04	fibronectin 1
Frk	14302	1.34	fyn-related kinase
Frmd4b	232288	1.58	FERM domain containing 4B
Fst	14313	2.24	follistatin
Fuca2	66848	1.17	fucosidase, alpha-L- 2, plasma
Furin	18550	1.41	furin (paired basic amino acid cleaving enzyme)
Fzd2	57265	1.16	frizzled homolog 2 (Drosophila)
Fzd3	14365	1.05	frizzled homolog 3 (Drosophila)
Gabra3	14396	1.04	gamma-aminobutyric acid (GABA) A receptor, subunit alpha 3
Gdpd3	68616	3.66	glycerophosphodiester phosphodiesterase domain containing 3

Gfra2	14586	1.2	glial cell line derived neurotrophic factor family receptor alpha 2
Gfra3	14587	1.23	glial cell line derived neurotrophic factor family receptor alpha 3
Ghr	14600	1.79	growth hormone receptor
Gja1	14609	1.09	gap junction protein, alpha 1
Gm10021	622931	1.04	predicted gene 10021
Gm12387	621880	1.2	predicted gene 12387
Gm3264	100041306	1.17	predicted gene 3494; predicted gene 3099; predicted gene 6676; predicted gene 3518; alpha7-takusan; RIKEN cDNA 4930555G01 gene; predicted gene, 100039441; predicted gene 3715; RIKEN cDNA B930046C15 gene;
Gm3696	100042149	3.36	predicted gene 3494; predicted gene 3099; predicted gene 6676; predicted gene 3518; alpha7-takusan; RIKEN cDNA 4930555G01 gene; predicted gene, 100039441;
Gm5458	432825	1.135	predicted gene 3494; predicted gene 3099; predicted gene 6676; predicted gene 3518; alpha7-takusan; RIKEN cDNA 4930555G01 gene; predicted gene, 100039441;
Gm5666	435373	1.7	predicted gene 5666
Gm773	331416	2.07	predicted gene 773
Gm9943	100036531	1.5	predicted gene 9943
Gpm6b	14758	1.16	glycoprotein m6b
Gpr116	224792	1.65	G protein-coupled receptor 116
Gpr165	76206	1.27	G protein-coupled receptor 165
Gpx8	69590	1.76	glutathione peroxidase 8 (putative)
Gria3	53623	1.96	glutamate receptor, ionotropic, AMPA3 (alpha 3)
Gstm5	14866	2.36	glutathione S-transferase, mu 5
Gypc	71683	1.72	glycophorin C
H2afy2	404634	1.18	H2A histone family, member Y2
H2-gs10	436493	1.01	MHC class I like protein GS10
H2-T22	15039	1.47	histocompatibility 2, T region locus 9; hypothetical protein LOC100044191; histocompatibility 2, T region locus 10; hypothetical protein LOC100044190; histocompatibility 2, T region locus 22
Hapln1	12950	3.94	hyaluronan and proteoglycan link protein 1
Hbb-y	15135	4.06	similar to beta-globin; hemoglobin Y, beta-like embryonic chain
Hdx	245596	1.37	highly divergent homeobox
Heph	15203	1.34	hephaestin
Hn1l	52009	1.15	hematological and neurological expressed 1-like
Igdcc3	19289	1.23	immunoglobulin superfamily, DCC subclass, member 3
Igf1r	16001	1.22	insulin-like growth factor I receptor
Ilk	16202	1.17	integrin linked kinase; predicted gene 6263
Inadl	12695	1.36	InaD-like (Drosophila)
Irs1	16367	1.3	insulin receptor substrate 1
Irs2	384783	1.2	insulin receptor substrate 2
Itga8	241226	1.68	integrin alpha 8
Kcne3	57442	1.29	potassium voltage-gated channel, Isk-related subfamily, gene 3; hypothetical protein LOC100044693
Kcnj5	16521	1.56	potassium inwardly-rectifying channel, subfamily J, member 5
Kcnmb1	16533	1.04	potassium large conductance calcium-activated channel, subfamily M, beta member 1
Kdm5d	20592	1.01	lysine (K)-specific demethylase 5D
Kif3a	16568	1.02	kinesin family member 3A
Klf6	23849	1.32	Kruppel-like factor 6
Krt19	16669	2.18	keratin 19
Lamp2	16784	1.13	lysosomal-associated membrane protein 2
Lcp1	18826	1.38	lymphocyte cytosolic protein 1
Ldb2	16826	1.8	LIM domain binding 2
Lgals3bp	19039	1	lectin, galactoside-binding, soluble, 3 binding protein
Lhfp	108927	1.17	lipoma HMGIC fusion partner
Limd1	29806	1.1	LIM domains containing 1
Lmbrd1	68421	1.17	LMBR1 domain containing 1
LOC280487	280487	3.36	pol polyprotein
Lphn2	99633	1.98125	latrophilin 2
Lrrc8c	100604	1.08	leucine rich repeat containing 8 family, member C
Map4k5	399510	1.09	mitogen-activated protein kinase kinase kinase kinase 5
Mapkapk2	17164	1.36	MAP kinase-activated protein kinase 2
Mcc	328949	1.23	mutated in colorectal cancers
Mesp1	17292	2.21	mesoderm posterior 1
Mmd	67468	1.73	monocyte to macrophage differentiation-associated; similar to monocyte to macrophage differentiation-associated

Mmp9	17395	1.36	matrix metallopeptidase 9
Mpp5	56217	1.3	membrane protein, palmitoylated 5 (MAGUK p55 subfamily member 5)
Msl1	74026	1.04	similar to RIKEN cDNA 4121402D02 gene; male-specific lethal 1 homolog (Drosophila)
Msl3	17692	1.15	male-specific lethal 3 homolog (Drosophila)
Mthfd2l	665563	1.145	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2-like
Myh10	77579	1.04	myosin, heavy polypeptide 10, non-muscle
Naalad2	72560	1.22	N-acetylated alpha-linked acidic dipeptidase 2
Nme5	75533	1.19	non-metastatic cells 5, protein expressed in (nucleoside-diphosphate kinase)
Nnat	18111	1.75	neuronatin
Notch2	18129	1.14	Notch gene homolog 2 (Drosophila)
Ogfr	72075	1.2	opioid growth factor receptor
Olfm1	56177	1.19	olfactomedin 1
Olfr893	258333	1.41	olfactory receptor 893
Ophn1	94190	1.21	oligophrenin 1
Padi3	18601	1.74	peptidyl arginine deiminase, type III
Parp8	52552	1.55	poly (ADP-ribose) polymerase family, member 8
Pbx3	18516	1.49	similar to PBX3a; pre B-cell leukemia transcription factor 3
Pctp	18559	1.12	phosphatidylcholine transfer protein
Pdgfrl	68797	1.34	platelet-derived growth factor receptor-like
Penk	18619	1.81	preproenkephalin
Phldb2	208177	2.29	pleckstrin homology-like domain, family B, member 2
Pla2g12b	69836	1.47	phospholipase A2, group XIIB
Plin2	11520	1.5	adipose differentiation related protein
Plxna4	243743	1.01	plexin A4
Pnpt1	71701	1.14	similar to polynucleotide phosphorylase-like protein; polyribonucleotide nucleotidyltransferase 1
Ppfibp1	67533	1.24	PTPRF interacting protein, binding protein 1 (liprin beta 1); similar to mKIAA1230 protein
Ppt2	54397	1.25	palmitoyl-protein thioesterase 2
Prkce	18754	1.06	RIKEN cDNA 9630025F12 gene; protein kinase C, epsilon
Prom1	19126	1.18	prominin 1
Pxmp3	19302	1.55	peroxisomal biogenesis factor 5-like; peroxisomal membrane protein 3
Pygl	110095	1.52	liver glycogen phosphorylase
Rap1a	109905	1.09	predicted gene 9392; similar to Raichu404X; RAS-related protein-1a
Rap2c	72065	1.34	similar to RAP2C, member of RAS oncogene family; RAP2C, member of RAS oncogene family
Rbm24	666794	1.09	RNA binding motif protein 24
Rbmy1a1	19657	1.893333333	RNA binding motif protein, Y chromosome, family 1, member A1
Rbp1	19659	1.53	retinol binding protein 1, cellular
Rfx3	19726	1.05	regulatory factor X, 3 (influences HLA class II expression); similar to Regulatory factor X, 3 (influences HLA class II expression)
Rgs17	56533	1.58	regulator of G-protein signaling 17
Rhob	11852	1.17	ras homolog gene family, member B
Rnf128	66889	1.78	ring finger protein 128
Samd3	268288	1.34	sterile alpha motif domain containing 3
Samd5	320825	1.07	sterile alpha motif domain containing 5
Sema6a	20358	1.72	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A
Sept8	20362	1.85	septin 8
Serpinb9	20723	1.39	serine (or cysteine) peptidase inhibitor, clade B, member 9
Sfxn4	94281	1.03	sideroflexin 4
Sgk3	170755	1.06	serum/glucocorticoid regulated kinase 3
Slc16a10	72472	1.02	solute carrier family 16 (monocarboxylic acid transporters), member 10
Slc18a2	214084	1.89	solute carrier family 18 (vesicular monoamine), member 2
Slc38a5	209837	1.04	solute carrier family 38, member 5
Slc39a8	67547	2.86	solute carrier family 39 (metal ion transporter), member 8
Slc43a3	58207	1.27	solute carrier family 43, member 3
Slc6a6	21366	1.09	solute carrier family 6 (neurotransmitter transporter, taurine), member 6
Sox4	20677	1.01	SRY-box containing gene 19; SRY-box containing gene 4
Sparcl1	13602	1.44	SPARC-like 1
Spats2l	67198	1.29	RIKEN cDNA 2810022L02 gene
St3gal1	20442	1.04	ST3 beta-galactoside alpha-2,3-sialyltransferase 1
Stag2	20843	1.01	stromal antigen 2

Stxbp6	217517	1.28	syntaxin binding protein 6 (amisyn)
Syne2	319565	1.156666667	synaptic nuclear envelope 2
T	20997	5.02	brachyury
Tank	21353	1.19	TRAF family member-associated Nf-kappa B activator
Tcf4	21413	1.04	transcription factor 4
Tex13	83555	1.39	testis expressed gene 13
Thoc7	66231	1.06	THO complex 7 homolog (Drosophila); similar to Thoc7 protein
Tmem128	66309	1.13	transmembrane protein 128
Tmem144	70652	1.13	transmembrane protein 144
Tmem47	192216	1.29	transmembrane protein 47
Tmem80	71448	1.22	transmembrane protein 80
Tpm1	22003	1.57	tropomyosin 1, alpha
Tsc22d3	14605	1.1	TSC22 domain family, member 3
Tspan6	56496	2.59	tetraspanin 6
Tspan7	21912	1.14	tetraspanin 7
Ttc18	76670	1.66	tetratricopeptide repeat domain 18
Tuft1	22156	1.07	similar to tuftelin; tuftelin 1
Usp18	24110	1.69	ubiquitin specific peptidase 18; similar to ubiquitin specific protease UBP43
Usp25	30940	1.26	ubiquitin specific peptidase 25
Usp9y	107868	1.29	ubiquitin specific peptidase 9, Y chromosome
Vmn2r50	434117	1.01	vomeronasal 2, receptor 50
Vwa5a	67776	1.16	von Willebrand factor A domain containing 5A
Wnt3	22415	2.14	wingless-related MMTV integration site 3
Zc3h12b	547176	1.28	zinc finger CCCH-type containing 12B
Zfp362	230761	1.03	zinc finger protein 362
Zfp608	269023	1.54	zinc finger protein 608
Zfp772	232855	1.01	cDNA sequence BC023179
Zhx2	387609	1.15	zinc fingers and homeoboxes 2

Downregulated genes in dnmt1-/- EBs compared to wt EBs between d0-4			
Gene symbol	**Entrez Gene ID**	**Fold change compared to wt**	**Gene name**
1700007K13Rik	69327	-1.14	RIKEN cDNA 1700007K13 gene
2200001I15Rik	69134	-1.06	RIKEN cDNA 2200001I15 gene
2310003C23Rik	76425	-1.01	predicted gene 5206; RIKEN cDNA 2310003C23 gene
2410017P07Rik	103268	-1.04	RIKEN cDNA 2410017P07 gene
2700023E23Rik	70036	-1.08	RIKEN cDNA 2700023E23 gene
4930429B21Rik	67576	-1.14	RIKEN cDNA 4930429B21 gene
4930502E18Rik	75013	-1.05	RIKEN cDNA 4930502E18 gene
4930591A17Rik	68175	-1.54	RIKEN cDNA 4930591A17 gene
5730590G19Rik	77011	-1.15	RIKEN cDNA 5730590G19 gene; similar to RIKEN cDNA 5730590G19-like
A2m	232345	-1.3	alpha-2-macroglobulin
Aass	30956	-1.15	aminoadipate-semialdehyde synthase
Abcg1	11307	-1.12	ATP-binding cassette, sub-family G (WHITE), member 1
Adamts14	237360	-1	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 14
Adap2	216991	-1.41	ArfGAP with dual PH domains 2
Aes	14797	-1.4	amino-terminal enhancer of split
Aire	11634	-1.53	autoimmune regulator (autoimmune polyendocrinopathy candidiasis ectodermal dystrophy)
Aldh1b1	72535	-1.27	aldehyde dehydrogenase 1 family, member B1
Aplp1	11803	-1.47	amyloid beta (A4) precursor-like protein 1
Apobec2	11811	-1.32	apolipoprotein B mRNA editing enzyme, catalytic polypeptide 2; similar to APOBEC-2 protein
Arrdc3	105171	-1.02	arrestin domain containing 3
C130073F10Rik	242574	-1.18	RIKEN cDNA C130073F10 gene
Camk2b	12323	-1.25	calcium/calmodulin-dependent protein kinase II, beta
Ccdc3	74186	-1.17	coiled-coil domain containing 3
Cd1d1	12479	-1.04	CD1d1 antigen; CD1d2 antigen
Cd37	12493	-1.34	CD37 antigen

Cd80	12519	-1.57	CD80 antigen
Cdkl4	381113	-1.2	cyclin-dependent kinase-like 4
Cenpm	66570	-1.48	centromere protein M
Cpn1	93721	-1.18	carboxypeptidase N, polypeptide 1
Cpne5	240058	-1.69	copine V; similar to Copine V
Crmp1	12933	-1.03	collapsin response mediator protein 1
Crtap	56693	-1.13	cartilage associated protein
Ctcfl	664799	-1.24	CCCTC-binding factor (zinc finger protein)-like
Ctsw	13041	-1.16	cathepsin W
Cyb5r1	72017	-1.01	cytochrome b5 reductase 1
Cyct	13067	-1.01	cytochrome c, testis
D1Pas1	110957	-1.11	DNA segment, Chr 1, Pasteur Institute 1
D630013G24Rik	319825	-1.23	RIKEN cDNA D630013G24 gene
D630023F18Rik	98303	-1.22	RIKEN cDNA D630023F18 gene
Dcdc2a	195208	-1.59	doublecortin domain containing 2a
Ddr2	18214	-1.37	discoidin domain receptor family, member 2
Derl3	70377	-1.12	Der1-like domain family, member 3
Dgka	13139	-1.2	diacylglycerol kinase, alpha
Dgke	56077	-1.12	diacylglycerol kinase, epsilon
Dmc1	13404	-1.04	DMC1 dosage suppressor of mck1 homolog, meiosis-specific homologous recombination (yeast)
Dnajc5g	231098	-1.92	DnaJ (Hsp40) homolog, subfamily C, member 5 gamma
Dpp4	13482	-1.49	dipeptidylpeptidase 4
Efhc2	74405	-2.27	EF-hand domain (C-terminal) containing 2
Egr1	13653	-1.01	early growth response 1
Epas1	13819	-1.56	endothelial PAS domain protein 1; similar to Endothelial PAS domain protein 1
Fam63a	75007	-1.24	family with sequence similarity 63, member A
Fkbp6	94244	-1.58	FK506 binding protein 6
Foxr1	382074	-1.32	forkhead box R1
Fry	320365	-1.29	furry homolog (Drosophila)
Galnt6	207839	-1.5	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 6
Gjb3	14620	-1.64	gap junction protein, beta 3
Glb1	12091	-1.24	galactosidase, beta 1
Gm10047	791327	-1.16	predicted gene 10047
Gm10139	628676	-1.05	predicted gene 10139
Gm13498	227885	-1.76	predicted gene 13498
Gm1564	268491	-1.74	predicted gene 1564
Gm15698	217066	-1.27	predicted gene 15698
Gm281	238939	-1.05	predicted gene 281
Gm2889	100040658	-1.982	hypothetical protein LOC100041609; hypothetical protein LOC100044795; predicted gene 3395; similar to gag polyprotein; hypothetical protein LOC100047557; hypothetical protein LOC100040347; hypothetical protein LOC100044384; hypothetical protein LOC100045342; hypothetical protein LOC100038979; predicted gene 2889
Gm340	381224	-1.16	predicted gene 340
Gm5077	317677	-1.16	predicted gene 5077
Gm5488	433036	-1.04	predicted gene 5488
Gm7325	653016	-1.08	predicted gene 7325
Gm7455	665033	-1.07	predicted gene 7455
Gm949	381142	-1.37	predicted gene 949
Gm9886	791286	-1.38	predicted gene 9886
Gpx2	14776	-1.59	glutathione peroxidase 2
Gpx4	625249	-1.025	heterogeneous nuclear ribonucleoprotein L-like; glutathione peroxidase 4
Grhl3	230824	-1.28	grainyhead-like 3 (Drosophila)
Hvcn1	74096	-1	hydrogen voltage-gated channel 1
Icosl	50723	-1.09	icos ligand
Ildr1	106347	-1.05	immunoglobulin-like domain containing receptor 1
Jag2	16450	-1.31	jagged 2
Jak3	16453	-1.07	Janus kinase 3
Kank3	80880	-1.11	KN motif and ankyrin repeat domains 3
Kdm3a	104263	-1.25	lysine (K)-specific demethylase 3A
Lama1	16772	-1.72	laminin, alpha 1

Lamc1	226519	-1.13	laminin, gamma 1
Macf1	11426	-1.12	microtubule-actin crosslinking factor 1
Magea1	17137	-1.27	melanoma antigen, family A, 1
Magea2	17138	-1.18	melanoma antigen, family A, 2
Mc5r	17203	-1.2	melanocortin 5 receptor
Mcam	84004	-1.36	melanoma cell adhesion molecule
Mfng	17305	-1.13	MFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase
Mia1	12587	-1.04	melanoma inhibitory activity 1
Mlana	77836	-1.21	melan-A
Mmp11	17385	-1.02	matrix metallopeptidase 11
Mmrn2	105450	-1.63	multimerin 2
Mras	17532	-1.31	muscle and microspikes RAS
Mylpf	17907	-1.67	myosin light chain, phosphorylatable, fast skeletal muscle
Myo1g	246177	-1	myosin IG
Naaa	67111	-1.3	N-acylethanolamine acid amidase
Naprt1	223646	-1.09	nicotinate phosphoribosyltransferase domain containing 1
Ncoa1	17977	-1.12	similar to Nuclear receptor coactivator 1 (NCoA-1) (Steroid receptor coactivator 1) (SRC-1) (Nuclear receptor coactivator protein 1) (mNRC-1); nuclear receptor coactivator 1
Necab1	69352	-1.59	N-terminal EF-hand calcium binding protein 1
Nedd4l	83814	-1.2	neural precursor cell expressed, developmentally down-regulated gene 4-like
Nfu1	56748	-1.34	NFU1 iron-sulfur cluster scaffold homolog (S. cerevisiae); predicted gene 7859
Nlrc4	268973	-2.24	NLR family, CARD domain containing 4
Nlrp4f	97895	-1.3	NLR family, pyrin domain containing 4F
Nlrp9b	243874	-1.04	NLR family, pyrin domain containing 9B
Nos1	18125	-1.31	nitric oxide synthase 1, neuronal
Nrg4	83961	-1.42	neuregulin 4
Olfr307	258610	-2.59	olfactory receptor 307
P4htm	74443	-1.12	prolyl 4-hydroxylase, transmembrane (endoplasmic reticulum)
Padi4	18602	-1.07	peptidyl arginine deiminase, type IV
Pcyt1b	236899	-1.19	phosphate cytidylyltransferase 1, choline, beta isoform
Pfkp	56421	-1.45	phosphofructokinase, platelet
Pigl	327942	-1.09	phosphatidylinositol glycan anchor biosynthesis, class L
Pik3cd	18707	-1.08	phosphatidylinositol 3-kinase catalytic delta polypeptide; RIKEN cDNA 2610208K16 gene
Pitpnc1	71795	-1.11	phosphatidylinositol transfer protein, cytoplasmic 1
Piwil4	330890	-1.29	piwi-like homolog 4 (Drosophila)
Pla2g7	27226	-1.07	phospholipase A2, group VII (platelet-activating factor acetylhydrolase, plasma)
Plac9	211623	-1.013333333	placenta specific 9; predicted gene 10393; predicted gene 9780
Plekhg5	269608	-1.27	pleckstrin homology domain containing, family G (with RhoGef domain) member 5
Plk5	216166	-1.06	polo-like kinase 5 (Drosophila)
Plxdc1	72324	-1.11	plexin domain containing 1
Pole4	66979	-1.31	polymerase (DNA-directed), epsilon 4 (p12 subunit)
Porcn	53627	-1.21	porcupine homolog (Drosophila)
Ppp2r2c	269643	-1.41	protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), gamma isoform
Ppp2r5c	26931	-1.14	protein phosphatase 2, regulatory subunit B (B56), gamma isoform
Pramel6	347711	-1.33	preferentially expressed antigen in melanoma like 6
Prmt8	381813	-1.39	protein arginine N-methyltransferase 8
Prune	229589	-1.04	predicted gene 5217; prune homolog (Drosophila)
Prune2	353211	-1.09	RIKEN cDNA A230083H22 gene
Ptchd3	74675	-1.12	patched domain containing 3
Ptk2b	19229	-1.3	PTK2 protein tyrosine kinase 2 beta
Ptp4a3	19245	-1.1	protein tyrosine phosphatase 4a3
Pycard	66824	-1.13	PYD and CARD domain containing
R3hdml	100043899	-1.08	R3H domain containing-like
Rab27a	11891	-1.55	RAB27A, member RAS oncogene family
Rad9b	231724	-1.09	RAD9 homolog B (S. cerevisiae)
Rdm1	66599	-1.23	RAD52 motif 1
Rlbp1	19771	-1.03	retinaldehyde binding protein 1
Skil	20482	-1.32	SKI-like
Slc1a1	20510	-1.78	solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1

Slc29a1	63959	-1.09	solute carrier family 29 (nucleoside transporters), member 1
Slc29a4	243328	-1.04	solute carrier family 29 (nucleoside transporters), member 4; similar to Solute carrier family 29 (nucleoside transporters), member 4
Slc5a1	20537	-1.52	solute carrier family 5 (sodium/glucose cotransporter), member 1
Slco4c1	227394	-1.4	solute carrier organic anion transporter family, member 4C1
Slfn9	237886	-1.05	similar to putative protein; schlafen 9; similar to schlafen 9
Smc1b	140557	-2.39	structural maintenance of chromosomes 1B
Sntb2	20650	-1.04	similar to beta-2-syntrophin; syntrophin, basic 2
Sox15	20670	-1.03	SRY-box containing gene 16; SRY-box containing gene 15
Sp110	109032	-1.03	predicted gene 15753; Sp110 nuclear body protein
Spint2	20733	-1.01	serine protease inhibitor, Kunitz type 2
Stag3	50878	-1.43	stromal antigen 3
Steap3	68428	-1.06	STEAP family member 3
Stk30	26448	-1.45	renal tumor antigen
Stra8	20899	-1.26	stimulated by retinoic acid gene 8
Syngr3	20974	-1.03	synaptogyrin 3
Taf9b	407786	-1.24	TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor
Tcfeb	21425	-1.3	transcription factor EB
Tmem64	100201	-1.19	transmembrane protein 64
Tmprss12	75002	-1.2	transmembrane protease, serine 12
Tnfsf12-tnfsf13	619441	-1.13	tumor necrosis factor (ligand) superfamily, member 12; tumor necrosis factor (ligand) superfamily, member 12-member 13; tumor necrosis factor (ligand) superfamily, member 13
Trib3	228775	-1.19	tribbles homolog 3 (Drosophila)
Tuba3b	22147	-1.28	predicted gene 5366; tubulin, alpha 3B; tubulin, alpha 3A
Tubb2b	73710	-1.1	tubulin, beta 2a, pseudogene 2; tubulin, beta 2B
Ubr5	70790	-1.14	ubiquitin protein ligase E3 component n-recognin 5
Ung	22256	-1.11	uracil DNA glycosylase
Upk1a	109637	-1.11	uroplakin 1A
Usp50	75083	-1.26	ubiquitin specific peptidase 50
Wdtc1	230796	-1.19	WD and tetratricopeptide repeats 1; similar to WD and tetratricopeptide repeats 1
Xlr5a	574438	-1.22	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Xlr5b	627081	-2.63	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Zcwpw1	381678	-2.7	paired immunoglobin-like type 2 receptor beta 2; zinc finger, CW type with PWWP domain 1
Zcwpw1	545812	-1.54	paired immunoglobin-like type 2 receptor beta 2; zinc finger, CW type with PWWP domain 1
Zfp710	209225	-1.17	zinc finger protein 710

Upregulated genes in TKO EBs compared to wt EBs between d0-4			
Gene symbol	**Entrez Gene ID**	**Fold change compared to wt**	**Gene name**
544988	100042149	1.05	predicted gene, 544988
1300003B13Rik	74149	1.13	RIKEN cDNA 1300003B13 gene; hypothetical protein LOC100044281
1300014I06Rik	66895	2.37	RIKEN cDNA 1300014I06 gene
1700048O20Rik	69430	1.42	RIKEN cDNA 1700048O20 gene
1810011O10Rik	69068	1.37	RIKEN cDNA 1810011O10 gene
2610018G03Rik	70415	1.3	RIKEN cDNA 2610018G03 gene
4631416L12Rik	622434	1.33	RIKEN cDNA 4631416L12 gene
4930550L24Rik	75352	1.33	RIKEN cDNA 4930550L24 gene
5930434B04Rik	381356	1.02	RIKEN cDNA 5930434B04 gene; hypothetical protein LOC100047034
6430411K18Rik	76880	1.06	retrotransposon-like 1; RIKEN cDNA 6430411K18 gene
9230105E10Rik	319236	1.526666667	similar to tripartite motif protein TRIM5; RIKEN cDNA 9230105E10 gene
A130022J15Rik	101351	1.69	RIKEN cDNA A130022J15 gene
A2bp1	268859	1.06	ataxin 2 binding protein 1
Aadat	23923	1	aminoadipate aminotransferase
Acadl	11363	1.31	acyl-Coenzyme A dehydrogenase, long-chain
Acsl4	50790	1.04	acyl-CoA synthetase long-chain family member 4
Adamts1	11504	1	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 1
Adamts15	235130	1.78	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 15

Aff3	16764	1.5	AF4/FMR2 family, member 3; similar to AF4/FMR2 family member 3 (LAF-4 protein) (Lymphoid nuclear protein related to AF4)
Afp	11576	1.42	alpha fetoprotein
Agpat3	28169	1.15	1-acylglycerol-3-phosphate O-acyltransferase 3
AI314180	230249	1.01	expressed sequence AI314180
Ano5	233246	1.04	anoctamin 5
Ap1m2	11768	1.29	adaptor protein complex AP-1, mu 2 subunit
Apba2	11784	1.22	amyloid beta (A4) precursor protein-binding, family A, member 2
Ar	11835	1.49	androgen receptor
Arhgap18	73910	1.88	Rho GTPase activating protein 18
Arhgef5	54324	1.285	Rho guanine nucleotide exchange factor (GEF) 5
Arhgef9	236915	1.18	CDC42 guanine nucleotide exchange factor (GEF) 9
Arid2	77044	1.04	AT rich interactive domain 2 (ARID, RFX-like); RIKEN cDNA 1700124K17 gene
Arl4c	320982	1.28	similar to ADP-ribosylation factor-like protein 7; ADP-ribosylation factor-like 4C
Arrb1	109689	1.14	arrestin, beta 1
Bach1	12013	1.25	BTB and CNC homology 1
Bbx	70508	1.23	bobby sox homolog (Drosophila)
Bcl9	77578	1.02	B-cell CLL/lymphoma 9
Cables1	63955	1.25	CDK5 and Abl enzyme substrate 1
Cacnb2	12296	1.04	calcium channel, voltage-dependent, beta 2 subunit
Calcr	12311	2.3	calcitonin receptor
Calcrl	54598	2.16	calcitonin receptor-like
Camk2d	108058	1.65	calcium/calmodulin-dependent protein kinase II, delta
Car13	71934	1.67	carbonic anhydrase 13
Car2	12349	2.86	carbonic anhydrase 2
Ccdc85a	216613	1.08	coiled-coil domain containing 85A
Ccl20	20297	1.01	chemokine (C-C motif) ligand 20
Cda	72269	1.09	cytidine deaminase
Cdc14a	229776	1.12	similar to Dual specificity protein phosphatase CDC14A (CDC14 cell division cycle 14 homolog A); CDC14 cell division cycle 14 homolog A (S. cerevisiae)
Cdcp1	109332	1.12	CUB domain containing protein 1
Cdh10	320873	1.47	cadherin 10
Cdh5	12562	1.7	cadherin 5
Cdh9	12565	1	cadherin 9
Cdkn1a	12575	1.23	cyclin-dependent kinase inhibitor 1A (P21)
Chst15	77590	2.09	carbohydrate (N-acetylgalactosamine 4-sulfate 6-O) sulfotransferase 15
Clcn5	12728	1.54	chloride channel 5
Cldn6	54419	1.81	claudin 6
Cldn7	53624	1.61	claudin 7
Cmtm8	70031	1.26	CKLF-like MARVEL transmembrane domain containing 8
Cnksr3	215748	1.13	Cnksr family member 3
Cpe	12876	2.88	carboxypeptidase E; similar to carboxypeptidase E
Cpt1b	12895	1.02	carnitine palmitoyltransferase 1b, muscle
Csrp1	13007	1.11	cysteine and glycine-rich protein 1
Ctnnd2	18163	1.2	catenin (cadherin associated protein), delta 2
Ctsc	13032	2.94	cathepsin C
Cxadr	13052	1.36	coxsackie virus and adenovirus receptor
Cxcl12	20315	1.43	chemokine (C-X-C motif) ligand 12
Cxx1b	553127	1.545	CAAX box 1 homolog A (human); CAAX box 1 homolog B (human); similar to mammalian retrotransposon derived 8b
Cxx1c	72865	2.1	CAAX box 1 homolog C (human)
Cyp2j6	13110	1.01	cytochrome P450, family 2, subfamily j, polypeptide 6
Cyp4a12b	13118	1.88	cytochrome P450, family 4, subfamily a, polypeptide 12B
Cyr61	16007	2.24	cysteine rich protein 61
Cysltr1	58861	1.66	cysteinyl leukotriene receptor 1
D830030K20Rik	320333	1	predicted gene 2971; predicted gene 3532; predicted gene 3033; RIKEN cDNA D830030K20 gene; predicted gene 3278; predicted gene 8271; predicted gene 2244; predicted gene 3043; predicted gene 3537; predicted gene 10408; predicted gene 8050; predicted gene 3755
Dapk1	69635	1.22	death associated protein kinase 1
Dbc1	56710	1.08	deleted in bladder cancer 1 (human)
Dclk1	13175	1.77	doublecortin-like kinase 1

Dennd5b	320560	1.02	DENN/MADD domain containing 5B
Dhrs3	20148	1.63	dehydrogenase/reductase (SDR family) member 3
Diap2	54004	1	diaphanous homolog 2 (Drosophila)
Dnmt3b	13436	1.35	DNA methyltransferase 3B
Dock11	75974	1.38	dedicator of cytokinesis 11
Dpysl5	65254	1.3	dihydropyrimidinase-like 5
Dusp4	319520	2.79	dual specificity phosphatase 4
Dusp6	67603	2.21	dual specificity phosphatase 6
Ebf1	13591	1.895	early B-cell factor 1
Efna5	13640	1.97	ephrin A5
Efnb2	13642	1.32	ephrin B2
Eif4g3	230861	1.14	eukaryotic translation initiation factor 4 gamma, 3; similar to Eukaryotic translation initiation factor 4 gamma 3 (eIF-4-gamma 3) (eIF-4G 3) (eIF4G 3) (eIF-4-gamma II) (eIF4GII)
Eltd1	170757	1.69	EGF, latrophilin seven transmembrane domain containing 1
Enpp2	18606	2.42	ectonucleotide pyrophosphatase/phosphodiesterase 2
ENSMUSG00000068790	100042149	1.41	predicted gene 3494; predicted gene 3099; predicted gene 6676; predicted gene 3518; alpha7-takusan; RIKEN cDNA 4930555G01 gene; predicted gene, 100039441; predicted gene 3715; RIKEN cDNA B930046C15 gene;
Epb4.1l3	13823	1.85	erythrocyte protein band 4.1-like 3
Epha1	13835	2.57	Eph receptor A1
Epha7	13841	1.85	Eph receptor A7
Erap1	80898	2.16	endoplasmic reticulum aminopeptidase 1
Erbb2ip	59079	1.16	Erbb2 interacting protein
Erg	13876	1.18	avian erythroblastosis virus E-26 (v-ets) oncogene related
Errfi1	74155	1	ERBB receptor feedback inhibitor 1
Esm1	71690	1.07	endothelial cell-specific molecule 1
Exoc3l	277978	1.82	exocyst complex component 3-like
F2r	14062	2.21	coagulation factor II (thrombin) receptor
Fabp7	12140	1.61	fatty acid binding protein 7, brain
Fads2	56473	1.12	fatty acid desaturase 2
Fam131b	76156	1.08	family with sequence similarity 131, member B
Fam171b	241520	1.76	family with sequence similarity 171, member B
Fam38b	667742	1.36	family with sequence similarity 38, member B2
Fam55c	385658	1.21	family with sequence similarity 55, member C
Fbxo32	67731	1	F-box protein 32
Fermt1	241639	1.07	fermitin family homolog 1 (Drosophila)
Fgf5	14176	4.03	fibroblast growth factor 5
Fgf8	14179	1.42	fibroblast growth factor 8
Fhl3	14201	1.03	four and a half LIM domains 3
Fras1	231470	1.32	Fraser syndrome 1 homolog (human)
Frem2	242022	1.35	Fras1 related extracellular matrix protein 2
Frk	14302	1.01	fyn-related kinase
Frmd4b	232288	1.57	FERM domain containing 4B
Fst	14313	3.17	follistatin
Fuca2	66848	1.23	fucosidase, alpha-L- 2, plasma
Fv1	14349	1.04	Friend virus susceptibility 1
Fzd2	57265	1.15	frizzled homolog 2 (Drosophila)
Fzd3	14365	1.2	frizzled homolog 3 (Drosophila)
Fzd7	14369	1.15	frizzled homolog 7 (Drosophila)
Gabra3	14396	1.67	gamma-aminobutyric acid (GABA) A receptor, subunit alpha 3
Gabra4	14397	1.24	gamma-aminobutyric acid (GABA) A receptor, subunit alpha 4
Galnt12	230145	1.26	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 12
Galnt3	14425	2.13	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3
Gatm	67092	1.63	glycine amidinotransferase (L-arginine:glycine amidinotransferase)
Gatsl2	80909	1.04	GATS protein-like 2
Gbp3	55932	2.28	guanylate binding protein 3
Gbp4	17472	1.15	guanylate binding protein 4
Gcom1	102371	1.02	GRINL1A complex locus
Gdpd3	68616	3.94	glycerophosphodiester phosphodiesterase domain containing 3
Ghr	14600	1.92	growth hormone receptor
Gja1	14609	1.74	gap junction protein, alpha 1

Glcci1	170772	1.25	similar to glucocorticoid induced transcript 1; predicted gene 5815; glucocorticoid induced transcript 1
Gli3	14634	1.41	GLI-Kruppel family member GLI3
Glrx	93692	1.85	glutaredoxin
Gm10021	622931	1.21	predicted gene 10021
Gm4638	100043775	3.56	predicted gene 4638
Gpm6a	234267	2.15	glycoprotein m6a
Gpm6b	14758	1.05	glycoprotein m6b
Gpx8	69590	1.49	glutathione peroxidase 8 (putative)
Gria3	53623	1.56	glutamate receptor, ionotropic, AMPA3 (alpha 3)
Grik3	14807	1.84	glutamate receptor, ionotropic, kainate 3
Grip1	74053	1	glutamate receptor interacting protein 1
Gstm1	14862	1.12	similar to Glutathione S-transferase Mu 1 (GST class-mu 1) (Glutathione S-transferase GT8.7) (pmGT10) (GST 1-1); predicted gene 5562; glutathione S-transferase, mu 1
Gstm5	14866	1.86	glutathione S-transferase, mu 5
H2-gs10	436493	1.78	MHC class I like protein GS10
H2-Q6	110557	1.17	histocompatibility 2, Q region locus 1; histocompatibility 2, Q region locus 9; similar to H-2 class I histocompatibility antigen, L-D alpha chain precursor; histocompatibility 2, Q region locus 8; histocompatibility 2, Q region locus 2; similar to MHC class Ib antigen; histocompatibility 2, Q region locus 7; histocompatibility 2, Q region locus 6; hypothetical protein LOC100044307; similar to H-2 class I histocompatibility antigen, Q7 alpha chain precursor (QA-2 antigen); RIKEN cDNA 0610037M15 gene
H2-T10	15024	1.17	histocompatibility 2, T region locus 9; hypothetical protein LOC100044191; histocompatibility 2, T region locus 10; hypothetical protein LOC100044190; histocompatibility 2, T region locus 22
H2-T22	15039	1.98	histocompatibility 2, T region locus 9; hypothetical protein LOC100044191; histocompatibility 2, T region locus 10; hypothetical protein LOC100044190; histocompatibility 2, T region locus 22
Hapln1	12950	3.37	hyaluronan and proteoglycan link protein 1
Hdx	245596	1.12	highly divergent homeobox
Heph	15203	1.79	hephaestin
Hes6	55927	1.09	hairy and enhancer of split 6 (Drosophila)
Hivep2	15273	1.35	human immunodeficiency virus type I enhancer binding protein 2
Hn1l	52009	1.06	hematological and neurological expressed 1-like
Hspa4l	18415	1	heat shock protein 4 like
Hunk	26559	1.28	similar to putative serine/threonine protein kinase MAK-V; similar to hormonally upregulated Neu-associated kinase; hormonally upregulated Neu-associated kinase
Id4	15904	2.03	inhibitor of DNA binding 4
Igdcc3	19289	1.26	immunoglobulin superfamily, DCC subclass, member 3
Igf1r	16001	1.51	insulin-like growth factor I receptor
Igfbp3	16009	2.74	insulin-like growth factor binding protein 3
Il17rd	171463	1.655	interleukin 17 receptor D
Ilk	16202	1.23	integrin linked kinase; predicted gene 6263
Inadl	12695	1.57	InaD-like (Drosophila)
Irgm1	15944	3.63	immunity-related GTPase family M member 1
Irs1	16367	1.71	insulin receptor substrate 1
Itga1	109700	2.32	integrin alpha 1
Itga8	241226	1.71	integrin alpha 8
Itm2c	64294	1.07	integral membrane protein 2C
Itpkb	320404	1.37	inositol 1,4,5-trisphosphate 3-kinase B
Jakmip2	76217	1.04	janus kinase and microtubule interacting protein 2
Kcng3	225030	1.76	potassium voltage-gated channel, subfamily G, member 3
Kcnh8	211468	1.13	potassium voltage-gated channel, subfamily H (eag-related), member 8
Kcnj3	16519	2.1	potassium inwardly-rectifying channel, subfamily J, member 3
Kcnk1	16525	1.1	potassium channel, subfamily K, member 1
Kctd12b	207474	1.18	potassium channel tetramerisation domain containing 12b
Kdm5d	20592	1	lysine (K)-specific demethylase 5D
Kif1a	16560	2.42	kinesin family member 1A
Kif21a	16564	1.5	kinesin family member 21A
Klf6	23849	1.26	Kruppel-like factor 6
Krt19	16669	1.9	keratin 19
Lcp1	18826	1.74	lymphocyte cytosolic protein 1
Ldb2	16826	1.77	LIM domain binding 2
Ldhb	16832	1.08	lactate dehydrogenase B; predicted gene 5514
Limd1	29806	1.22	LIM domains containing 1

Lmo4	16911	1.67	LIM domain only 4
LOC280487	280487	3.56	pol polyprotein
Lpar4	78134	2.34	lysophosphatidic acid receptor 4
Lpcat4	99010	1.14	lysophosphatidylcholine acyltransferase 4
Lphn2	99633	2.715	latrophilin 2
Lrrc1	214345	1.21	leucine rich repeat containing 1
Lrrc8c	100604	2.06	leucine rich repeat containing 8 family, member C
Lrrn1	16979	1.57	leucine rich repeat protein 1, neuronal
Lsr	54135	1.01	lipolysis stimulated lipoprotein receptor
Lyn	17096	1.41	Yamaguchi sarcoma viral (v-yes-1) oncogene homolog
Lypd6	320343	1.13	LY6/PLAUR domain containing 6
Mab21l2	23937	1.3	mab-21-like 2 (C. elegans)
Magel2	27385	1.78	melanoma antigen, family L, 2
Magi2	50791	1.09	membrane associated guanylate kinase, WW and PDZ domain containing 2
Mapk12	29857	1.05	mitogen-activated protein kinase 12
Mapk4	225724	1.33	mitogen-activated protein kinase 4
Mapre2	212307	1.19	microtubule-associated protein, RP/EB family, member 2
Marveld2	218518	1.03	MARVEL (membrane-associating) domain containing 2
Mcc	328949	1.47	mutated in colorectal cancers
Mcoln3	171166	1.41	mucolipin 3
Mecom	14013	1.57	ecotropic viral integration site 1
Med14	26896	1.05	mediator complex subunit 14
Met	17295	1.52	met proto-oncogene
Mfap3l	71306	1.06	microfibrillar-associated protein 3-like
Mkx	210719	1.12	mohawk homeobox
Mmp14	17387	1.13	matrix metallopeptidase 14 (membrane-inserted)
Mmp25	240047	2.17	matrix metallopeptidase 25
Morf4l2	56397	1.07	predicted gene 5521; similar to mortality factor 4 like 2; mortality factor 4 like 2
Moxd1	59012	1.24	monooxygenase, DBH-like 1
Mthfd2l	665563	1.435	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2-like
Naalad2	72560	1.31	N-acetylated alpha-linked acidic dipeptidase 2
Nefl	18039	2.53	neurofilament, light polypeptide
Nefm	18040	3.37	neurofilament, medium polypeptide
Nes	18008	2.02	nestin
Nhsl1	215819	1.15	NHS-like 1
Nnat	18111	2.29	neuronatin
Notch2	18129	1.21	Notch gene homolog 2 (Drosophila)
Npy1r	18166	1.33	neuropeptide Y receptor Y1
Nr6a1	14536	1.01	nuclear receptor subfamily 6, group A, member 1
Nrcam	319504	1.48	neuron-glia-CAM-related cell adhesion molecule
Ogfr	72075	1.35	opioid growth factor receptor
Ophn1	94190	1.29	oligophrenin 1
Otx2	18424	1.68	orthodenticle homolog 2 (Drosophila)
Parp10	671535	1.18	similar to Plec1 protein
Parp8	52552	2.84	poly (ADP-ribose) polymerase family, member 8
Pcdh1	75599	1.23	protocadherin 1
Pde2a	207728	1.27	phosphodiesterase 2A, cGMP-stimulated
Pdpn	14726	1.01	podoplanin
Penk	18619	1.83	preproenkephalin
Pgap1	241062	1.15	post-GPI attachment to proteins 1
Phldb2	208177	2.04	pleckstrin homology-like domain, family B, member 2
Pik3r1	18708	1.09	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1 (p85 alpha)
Pim2	18715	2.77	proviral integration site 2
Pitx2	18741	1.1	paired-like homeodomain transcription factor 2
Plagl1	22634	1.67	pleiomorphic adenoma gene-like 1
Plch1	269437	1.18	phospholipase C, eta 1
Plekha1	101476	1.1	pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 1
Plekhg2	101497	1.31	pleckstrin homology domain containing, family G (with RhoGef domain) member 2
Plin2	11520	1.47	adipose differentiation related protein

Plk2	20620	1.21	polo-like kinase 2 (Drosophila)
Plxdc2	67448	1.35	plexin domain containing 2
Plxnb1	235611	1.06	plexin B1
Plxnc1	54712	1.5	plexin C1; similar to plexin C1
Pphln1	223828	1.58	periphilin 1
Ppp4r4	74521	3.28	protein phosphatase 4, regulatory subunit 4
Pramel3	83565	1.02	preferentially expressed antigen in melanoma-like 3
Prickle1	106042	1.27	prickle like 1 (Drosophila)
Prkch	18755	1.06	protein kinase C, eta
Prom1	19126	2.43	prominin 1
Prss8	76560	1.19	protease, serine, 8 (prostasin)
Psme1	19186	1.1	predicted gene 7776; proteasome (prosome, macropain) 28 subunit, alpha
Ptbp2	56195	1.07	polypyrimidine tract binding protein 2
Ptk7	71461	1.04	PTK7 protein tyrosine kinase 7
Ptprm	19274	1.13	protein tyrosine phosphatase, receptor type, M
Rab25	53868	1.83	RAB25, member RAS oncogene family
Rab38	72433	1.16	RAB38, member of RAS oncogene family
Rasef	242505	1.99	RAS and EF hand domain containing
Rdh10	98711	1.19	retinol dehydrogenase 10 (all-trans)
Rem2	140743	1.17	rad and gem related GTP binding protein 2
Rerg	232441	1.43	RAS-like, estrogen-regulated, growth-inhibitor
Rfx3	19726	1.06	regulatory factor X, 3 (influences HLA class II expression); similar to Regulatory factor X, 3 (influences HLA class II expression)
Rfx6	320995	1.33	regulatory factor X, 6
Rgs17	56533	1.6	regulator of G-protein signaling 17
Rgs2	19735	2.18	regulator of G-protein signaling 2
Rhob	11852	1.11	ras homolog gene family, member B
Rnf122	68867	1.13	ring finger protein 122
Rnf128	66889	1.86	ring finger protein 128
Rnpepl1	108657	1.06	arginyl aminopeptidase (aminopeptidase B)-like 1
Rragd	52187	1.02	Ras-related GTP binding D
Rtl1	353326	1.325	retrotransposon-like 1; RIKEN cDNA 6430411K18 gene
S1pr1	13609	1.24	sphingosine-1-phosphate receptor 1
Sall2	50524	1.68	sal-like 2 (Drosophila)
Sat1	20229	1.04	similar to spermidine/spermine N1-acetyltransferase; predicted gene 5552; spermidine/spermine N1-acetyl transferase 1
Satb1	20230	1.18	special AT-rich sequence binding protein 1
Scamp1	107767	1.36	secretory carrier membrane protein 1
Sema6a	20358	3.37	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A
Sepp1	20363	1.33	selenoprotein P, plasma, 1
Serinc5	218442	1.91	serine incorporator 5
Serpinb9	20723	1.71	serine (or cysteine) peptidase inhibitor, clade B, member 9
Sesn3	75747	1.03	sestrin 3
Setbp1	240427	1.45	SET binding protein 1
Setd7	73251	1.18	SET domain containing (lysine methyltransferase) 7
Sfrp2	20319	1.39	secreted frizzled-related protein 2
Sgk3	170755	1.39	serum/glucocorticoid regulated kinase 3
Shc4	271849	1.31	SHC (Src homology 2 domain containing) family, member 4
Slc16a2	20502	1.12	similar to X-linked PEST-containing transporter; solute carrier family 16 (monocarboxylic acid transporters), member 2
Slc35d3	76157	1.18	solute carrier family 35, member D3
Slc35f1	215085	1.22	solute carrier family 35, member F1
Slc39a8	67547	2.61	solute carrier family 39 (metal ion transporter), member 8
Slc5a5	114479	1.33	solute carrier family 5 (sodium iodide symporter), member 5
Slc9a3r1	26941	1.2	solute carrier family 9 (sodium/hydrogen exchanger), member 3 regulator 1
Slc9a7	236727	1.84	solute carrier family 9 (sodium/hydrogen exchanger), member 7
Slco5a1	240726	1.06	solute carrier organic anion transporter family, member 5A1
Socs2	216233	1.12	suppressor of cytokine signaling 2; predicted gene 8000
Sox18	20672	1.13	SRY-box containing gene 18
Sox4	20677	1.6	SRY-box containing gene 19; SRY-box containing gene 4
Sp8	320145	1.14	trans-acting transcription factor 8

Spats2l	67198	1.51	RIKEN cDNA 2810022L02 gene
Spon1	233744	1.51	spondin 1, (f-spondin) extracellular matrix protein
Spry2	24064	1.08	sprouty homolog 2 (Drosophila)
Ssbp2	66970	1.99	single-stranded DNA binding protein 2; predicted gene 12470
St6gal2	240119	2.04	beta galactoside alpha 2,6 sialyltransferase 2
Stag2	20843	1.13	stromal antigen 2
Stard10	56018	1.06	START domain containing 10
Stau2	29819	1.14	staufen (RNA binding protein) homolog 2 (Drosophila)
Stox2	71069	1.32	storkhead box 2
Stx3	20908	1.44	syntaxin 3
Stxbp6	217517	1.42	syntaxin binding protein 6 (amisyn)
Syne2	319565	1.736666667	synaptic nuclear envelope 2
Syngap1	240057	1.06	synaptic Ras GTPase activating protein 1 homolog (rat)
T	20997	5.67	Brachyury
Tapbp	21356	1.22	TAP binding protein
Tbc1d5	72238	1.1	TBC1 domain family, member 5
Tc2n	74413	1.31	tandem C2 domains, nuclear
Tceal8	66684	1.28	transcription elongation factor A (SII)-like 8; similar to transcription elongation factor A (SII)-like 8
Tcf4	21413	1.08	transcription factor 4
Tcirg1	27060	1.02	T-cell, immune regulator 1, ATPase, H+ transporting, lysosomal V0 protein A3
Tec	21682	1.02	tec protein tyrosine kinase
Tex13	83555	2.02	testis expressed gene 13
Tex15	104271	1.35	testis expressed gene 15
Tie1	21846	1.25	tyrosine kinase with immunoglobulin-like and EGF-like domains 1
Tmc4	353499	1.18	transmembrane channel-like gene family 4
Tmem173	72512	1.1	transmembrane protein 173
Tmem189	407243	1.16	predicted gene 6194; transmembrane protein 189
Tmem200a	77220	1.82	RIKEN cDNA C030003D03 gene
Tmem47	192216	1.2	transmembrane protein 47
Tnfaip8	106869	1.06	tumor necrosis factor, alpha-induced protein 8
Tpbg	21983	1.76	trophoblast glycoprotein
Tpm1	22003	1.69	tropomyosin 1, alpha
Trib2	217410	1.18	tribbles homolog 2 (Drosophila)
Trpa1	277328	1.04	transient receptor potential cation channel, subfamily A, member 1
Tspan5	56224	1.12	tetraspanin 5
Tspan6	56496	2.03	tetraspanin 6
Tspan7	21912	1.49	tetraspanin 7
Tspan8	216350	1.6	tetraspanin 8
Ttc18	76670	1.21	tetratricopeptide repeat domain 18
Uap1l1	227620	1.03	UDP-N-acteylglucosamine pyrophosphorylase 1-like 1
Usp25	30940	1.31	ubiquitin specific peptidase 25
Utrn	22288	1.05	utrophin
Vaultrc5	378472	1.22	vault RNA component 5
Vcl	22330	1.17	vinculin
Vwa5a	67776	1.65	von Willebrand factor A domain containing 5A
Wnt8a	20890	2.39	wingless-related MMTV integration site 8A
Zfp36l2	12193	1.53	zinc finger protein 36, C3H type-like 2
Zfp608	269023	2.27	zinc finger protein 608
Zfp869	66869	1.2	zinc finger protein 869
Zmynd8	228880	1	zinc finger, MYND-type containing 8
Zyx	22793	1.04	zyxin

Downpregulated genes in TKO EBs compared to wt EBs between d0-4			
Gene symbol	**Entrez Gene ID**	**Fold change compared to wt**	**Gene name**
2410017P07Rik	103268	-1.15	RIKEN cDNA 2410017P07 gene
4930528F23Rik	75178	-1.97	RIKEN cDNA 4930528F23 gene
4930591A17Rik	68175	-1.59	RIKEN cDNA 4930591A17 gene

9030617O03Rik	217830	-1.57	RIKEN cDNA 9030617O03 gene
A2m	232345	-1.6	alpha-2-macroglobulin
Acp6	66659	-1.04	acid phosphatase 6, lysophosphatidic
Aire	11634	-1.51	autoimmune regulator (autoimmune polyendocrinopathy candidiasis ectodermal dystrophy)
Akap2	11641	-1.06	A kinase (PRKA) anchor protein 2; paralemmin 2
Alpk3	116904	-1.8	alpha-kinase 3
Apod	11815	-1.04	apolipoprotein D
Arrdc4	66412	-1.1	arrestin domain containing 4
Ascl2	17173	-1.07	achaete-scute complex homolog 2 (Drosophila)
Atp9a	11981	-1.21	ATPase, class II, type 9A
Bcap29	12033	-1.53	B-cell receptor-associated protein 29
Btnl7	195349	-1.42	butyrophilin-like 7
C330024D21Rik	320479	-1.04	RIKEN cDNA C330024D21 gene
Cd37	12493	-1.23	CD37 antigen
Cenpm	66570	-1.26	centromere protein M
Chac1	69065	-1.3	ChaC, cation transport regulator-like 1 (E. coli)
Cmklr1	14747	-1.12	chemokine-like receptor 1
Cpn1	93721	-1.28	carboxypeptidase N, polypeptide 1
Crtap	56693	-1.02	cartilage associated protein
Ctcfl	664799	-1.82	CCCTC-binding factor (zinc finger protein)-like
D1Pas1	110957	-1.13	DNA segment, Chr 1, Pasteur Institute 1
Dcdc2a	195208	-1.21	doublecortin domain containing 2a
Dgke	56077	-1.08	diacylglycerol kinase, epsilon
Dnajc5g	231098	-1.9	DnaJ (Hsp40) homolog, subfamily C, member 5 gamma
Ech1	51798	-1.2	enoyl coenzyme A hydratase 1, peroxisomal
Efhc2	74405	-2.6	EF-hand domain (C-terminal) containing 2
Galnt6	207839	-1.5	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 6
Gfpt2	14584	-1.23	glutamine fructose-6-phosphate transaminase 2
Gjb3	14620	-1.55	gap junction protein, beta 3
Gjb5	14622	-1.43	gap junction protein, beta 5
Gm10863	100041655	-1.36	hypothetical protein LOC100041655
Gm13498	227885	-2.7	predicted gene 13498
Gm15698	217066	-1.22	predicted gene 15698
Gm216	241112	-1.08	predicted gene 216
Gm2889	100040658	-2.348	hypothetical protein LOC100041609; hypothetical protein LOC100044795; predicted gene 3395; similar to gag polyprotein; hypothetical protein LOC100047557; hypothetical protein LOC100040347; hypothetical protein LOC100044384; hypothetical protein LOC100045342; hypothetical protein LOC100038979; predicted gene 2889
Gpx2	14776	-1.41	glutathione peroxidase 2
Grb10	14783	-1.61	growth factor receptor bound protein 10
Hmmr	15366	-1.1	hyaluronan mediated motility receptor (RHAMM)
Magea1	17137	-1.71	melanoma antigen, family A, 1
Magea2	17138	-1.53	melanoma antigen, family A, 2
Magea4	17140	-1.04	melanoma antigen, family A, 4
Magea5	17141	-1.38	melanoma antigen, family A, 5
Magea6	17142	-1.29	melanoma antigen, family A, 6
Mapt	17762	-1.1	microtubule-associated protein tau
Mia1	12587	-1.31	melanoma inhibitory activity 1
Mras	17532	-1.42	muscle and microspikes RAS
Msln	56047	-1.14	mesothelin
Naprt1	223646	-1.01	nicotinate phosphoribosyltransferase domain containing 1
Necab1	69352	-1.13	N-terminal EF-hand calcium binding protein 1
Nfu1	56748	-1.16	NFU1 iron-sulfur cluster scaffold homolog (S. cerevisiae); predicted gene 7859
Nlrp4f	97895	-1.23	NLR family, pyrin domain containing 4F
Nos1	18125	-1.19	nitric oxide synthase 1, neuronal
Nupr1	56312	-1.04	nuclear protein 1
Olfr307	258610	-1.36	olfactory receptor 307
P4htm	74443	-1.01	prolyl 4-hydroxylase, transmembrane (endoplasmic reticulum)
Pcyt1b	236899	-1.15	phosphate cytidylyltransferase 1, choline, beta isoform
Phlda2	22113	-1.34	pleckstrin homology-like domain, family A, member 2
Pigl	327942	-1.02	phosphatidylinositol glycan anchor biosynthesis, class L

Piwil4	330890	-1.51	piwi-like homolog 4 (Drosophila)
Pkd2l1	329064	-1.04	polycystic kidney disease 2-like 1
Prmt8	381813	-1.26	protein arginine N-methyltransferase 8
R3hdml	100043899	-1.03	R3H domain containing-like
Rhox10	434769	-1.3	reproductive homeobox 10
Rlbp1	19771	-1.06	retinaldehyde binding protein 1
Rtp3	235636	-1.08	receptor transporter protein 3
Slc11a1	18173	-1.02	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1
Slc1a1	20510	-1.47	solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1
Slc5a1	20537	-1.14	solute carrier family 5 (sodium/glucose cotransporter), member 1
Smc1b	140557	-2.15	structural maintenance of chromosomes 1B
Spink10	328971	-1.18	serine peptidase inhibitor, Kazal type 10
Stag3	50878	-1.09	stromal antigen 3
Stk30	26448	-1.03	renal tumor antigen
Stra8	20899	-1.47	stimulated by retinoic acid gene 8
Tcfeb	21425	-1.25	transcription factor EB
Tmem40	94346	-1.12	transmembrane protein 40
Trib3	228775	-1.43	tribbles homolog 3 (Drosophila)
Tsx	22127	-1.04	testis specific X-linked gene
Txnrd3	232223	-1.05	thioredoxin reductase 3
Ubash3a	328795	-1.1	ubiquitin associated and SH3 domain containing, A
Ubxn11	67586	-1.03	UBX domain protein 11
Usp50	75083	-1.08	ubiquitin specific peptidase 50
Xlr5a	574438	-2.68	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Xlr5b	627081	-2.8	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Yars	107271	-1.05	tyrosyl-tRNA synthetase
Zcwpw1	381678	-1.15	paired immunoglobin-like type 2 receptor beta 2; zinc finger, CW type with PWWP domain 1

Genes concordantly upregulated in wt, dnmt1-/- and TKO Ebs between d0-4					
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change dnmt1-/-**	**Fold change TKO**	**Gene name**
2310021P13Rik	268721	1.01	1.23	1.12	RIKEN cDNA 2310021P13 gene
2810432L12Rik	67063	1.05	1.26	1.09	RIKEN cDNA 2810432L12 gene
5730469M10Rik	70564	1.4	1.75	1.12	RIKEN cDNA 5730469M10 gene
7420416P09Rik	432677	2.15	2	2.14	RIKEN cDNA 7420416P09 gene
Adamts3	330119	1.718666667	1.555454545	1.05	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 3
Adamts6	108154	1.37	1.37	2.71	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 6
Adamts9	101401	1.63	1.02	1.215	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 9
Alox15	11687	1.9	2.71	2.2	arachidonate 15-lipoxygenase
Amot	27494	4.31	4.18	3.84	angiomotin
Ano10	102566	2.03	2.36	1.52	anoctamin 10
Antxr2	71914	1.59	1.99	1.99	anthrax toxin receptor 2
Anxa5	11747	1.42	2.4	2.48	annexin A5
Aplnr	23796	3.64	3.86	1.49	apelin receptor
App	11820	1.98	2.16	2.15	amyloid beta (A4) precursor protein
Armcx3	71703	1.936666667	1.773333333	2.133333333	armadillo repeat containing, X-linked 3; hypothetical protein LOC100044266; predicted gene 9299
Armcx3	100044266	3.67	5.58	2.52	armadillo repeat containing, X-linked 3; hypothetical protein LOC100044266; predicted gene 9299
Asb4	65255	1.59	2.05	1.64	ankyrin repeat and SOCS box-containing 4
Atrnl1	226255	1.13	1.83	1.17	attractin like 1
Axin2	12006	1.02	1.83	1.65	axin2
B2m	12010	2.04	2.06	1.07	beta-2 microglobulin
BC023829	236848	1.5	1.38	1.45	cDNA sequence BC023829
Bin1	30948	1.19	1.32	1.54	bridging integrator 1
C530008M17Rik	320827	1.68	2.08	2.9	RIKEN cDNA C530008M17 gene

Cachd1	320508	2.12	3.04	2.54	cache domain containing 1; similar to Cache domain containing 1
Cap2	67252	2.51	2.34	1.97	CAP, adenylate cyclase-associated protein, 2 (yeast)
Car14	23831	3.04	2.51	3.34	carbonic anhydrase 14
Car4	12351	1.66	1.37	1.46	carbonic anhydrase 4
Casp6	12368	1.11	1.74	1.72	caspase 6
Casp8	12370	1.27	1.1	1.62	caspase 8
Ccnjl	380694	5.18	4.06	2.11	cyclin J-like
Cdh11	12552	2.4	2.81	2.47	cadherin 11
Cdh2	12558	1.19	1.51	1.26	cadherin 2; similar to N-cadherin
Cgrrf1	68755	1.11	1.13	1.18	cell growth regulator with ring finger domain 1
Clcn4-2	12727	1.58	1.7	1.78	chloride channel 4-2
Clic1	114584	1.85	1.68	1.16	chloride intracellular channel 1
Cmtm3	68119	1.36	1.59	1.18	CKLF-like MARVEL transmembrane domain containing 3
Cobll1	319876	1.07	1.41	1.64	Cobl-like 1
Colec12	140792	1.69	1.69	1.4	collectin sub-family member 12
Commd3	12238	1.42	1.48	1.85	COMM domain containing 3
Crispld1	83691	4.28	4.11	1.37	cysteine-rich secretory protein LCCL domain containing 1
Cxcr4	12767	3.68	3.73	2.27	chemokine (C-X-C motif) receptor 4
Cyp26a1	13082	1.01	1.5	1.47	cytochrome P450, family 26, subfamily a, polypeptide 1
Cyp39a1	56050	1.39	1.86	1.54	cytochrome P450, family 39, subfamily a, polypeptide 1
Dab2	13132	1.4	1.41	1.53	disabled homolog 2 (Drosophila)
Dbn1	56320	1.71	1.35	1.34	drebrin 1
Ddx26b	236790	1.69	1.69	1.67	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 26B
Dgkk	331374	1.12	1.3	1.02	diacylglycerol kinase kappa
Dock7	67299	1.51	1.3	1.47	dedicator of cytokinesis 7
Dpysl2	12934	1.14	1.44	1.76	dihydropyrimidinase-like 2
Dsc2	13506	1.53	1.99	1.38	desmocollin 2
Dsp	109620	2.75	3.75	4.27	desmoplakin
Eomes	13813	1.93	2.1	1.07	eomesodermin homolog (Xenopus laevis)
Ets1	23871	1.16	1.84	2.15	E26 avian leukemia oncogene 1, 5' domain
Etv2	14008	2.14	2.13	1.66	similar to ETS related protein 71; ets variant gene 2
Fam115a	77574	1.37	1.02	1.31	family with sequence similarity 115, member A
Farp1	223254	1.3	1.2075	1.42	FERM, RhoGEF (Arhgef) and pleckstrin domain protein 1 (chondrocyte-derived); similar to FERMRhoGEF (Arhgef) and pleckstrin domain protein 1
Fat3	270120	2.63	2.54	1.14	FAT tumor suppressor homolog 3 (Drosophila)
Fgf10	14165	1.94	1.25	1.32	fibroblast growth factor 10
Fli1	14247	1.69	1.92	2.38	Friend leukemia integration 1
Fndc3c1	333564	1.53	3.19	4.37	fibronectin type III domain containing 3C1
Foxh1	14106	1.45	1.31	1	forkhead box H1
Frzb	20378	2.53	3.26	2.51	frizzled-related protein
Gata3	14462	1.7	1.4	1.02	GATA binding protein 3
Gata6	14465	2.65	2.7	1.16	GATA binding protein 6
Gpc6	23888	1.91	1.46	1.19	predicted gene 4672; glypican 6; similar to Glypican 6
Gprc5c	70355	1.615	1.745	1.84	G protein-coupled receptor, family C, group 5, member C; similar to G protein-coupled receptor, family C, group 5, member C
Gria4	14802	1.32	1.05	1.53	glutamate receptor, ionotropic, AMPA4 (alpha 4); hypothetical protein LOC100044208
H19	14955	3.58	3.15	2.825	H19 fetal liver mRNA
Has2	15117	3.09	3.74	4.44	hyaluronan synthase 2
Hdac7	56233	1.25	1.21	1.06	histone deacetylase 7; similar to histone deacetylase 7A
Hey1	15213	1.5	1.46	1.94	hairy/enhancer-of-split related with YRPW motif 1
Hmcn1	545370	1.312121212	1.378	1.5806	hemicentin 1
Hmgb3	15354	1.855	1.815	1.39	predicted gene 11805; predicted gene 8850; high mobility group box 3; similar to High mobility group protein 4 (HMG-4) (High mobility group protein 2a) (HMG-2a)
Id3	15903	1.56	1.94	2.14	inhibitor of DNA binding 3
Ier5	15939	1.25	1.65	1.12	immediate early response 5
Igfbp4	16010	2.39	2.9	2.41	insulin-like growth factor binding protein 4
Insr	16337	1.245	1.41	1.25	insulin receptor
Jak2	16452	1.43	1.32	1.31	Janus kinase 2

Jam3	83964	1.16	1.64	1.25	junction adhesion molecule 3
Jun	16476	1.25	1.52	1.44	Jun oncogene
Kdr	16542	4.12	4.24	3.41	kinase insert domain protein receptor
Kif1c	16562	1.43	1.38	1.24	kinesin family member 1C
Klhdc2	69554	1.13	1.35	1.18	kelch domain containing 2
Klhl6	239743	2.75	4.92	3.82	kelch-like 6 (Drosophila)
Krt18	16668	1.64	2.82	2.69	keratin 18
Krt8	16691	1.65	2.85	1.92	predicted gene 5604; keratin 8
L3mbtl3	237339	1.59	1.6	1.54	l(3)mbt-like 3 (Drosophila)
Leprel1	210530	2.15	2.6	1.55	leprecan-like 1
Lmo2	16909	2.03	1.69	1.05	LIM domain only 2
Lpar6	67168	1.05	1.03	2.57	purinergic receptor P2Y, G-protein coupled, 5
Lrrn4	320974	1.19	1.06	1.32	leucine rich repeat neuronal 4
Maged2	80884	1.3	1.45	1.18	similar to melanoma antigen family D, 2; melanoma antigen, family D, 2
Magi3	99470	1.86	1.18	1.15	membrane associated guanylate kinase, WW and PDZ domain containing 3
Marcks	17118	1.96	2.06	1.79	myristoylated alanine rich protein kinase C substrate
Mef2a	17258	1.08	1.28	1.21	similar to Myocyte enhancer factor 2A; myocyte enhancer factor 2A
Mest	17294	1.31	1.29	1.05	mesoderm specific transcript
Mixl1	27217	2.66	3.34	3.28	Mix1 homeobox-like 1 (Xenopus laevis)
Mmp2	17390	1.9	2.1	1.46	matrix metallopeptidase 2
Mogat2	233549	3.17	3.31	1.99	monoacylglycerol O-acyltransferase 2
Mpdz	17475	1.29	1.55	1.66	multiple PDZ domain protein
Mrps6	121022	1.19	1.46	1.16	mitochondrial ribosomal protein S6
Msx1	17701	2.41	2.32	1.33	homeobox, msh-like 1
Mum1l1	245631	2.49	2.71	1.81	melanoma associated antigen (mutated) 1-like 1
Ndn	17984	1.43	1.62	1.29	necdin
Nrarp	67122	1.11	1.25	1.6	Notch-regulated ankyrin repeat protein
Oat	18242	1.14	1.54	1.52	ornithine aminotransferase
Paip1	218693	1.27	1.25	1.02	polyadenylate binding protein-interacting protein 1; similar to poly(A) binding protein interacting protein 1
Pam	18484	1.69	1.86	1.92	peptidylglycine alpha-amidating monooxygenase
Papss1	23971	1.18	1.31	1.25	3'-phosphoadenosine 5'-phosphosulfate synthase 1
Pbx1	18514	1.97	2.55	2.05	pre B-cell leukemia transcription factor 1; region containing RIKEN cDNA 2310056B04 gene; pre B-cell leukemia transcription factor 1
Pcdh18	73173	2.69	3.16	1.98	protocadherin 18
Pcdh7	54216	3.09	2.53	1.35	protocadherin 7
Pde5a	242202	1.31	2.03	2.27	phosphodiesterase 5A, cGMP-specific
Pdgfra	18595	3.78	4.24	1.79	platelet derived growth factor receptor, alpha polypeptide
Pdlim3	53318	1.39	1.47	1.78	PDZ and LIM domain 3
Pdzrn3	55983	1.44	1.78	2.09	PDZ domain containing RING finger 3
Pea15a	18611	1.03	1.17	1.39	phosphoprotein enriched in astrocytes 15A
Peg10	170676	1.32	2.1	1.45	paternally expressed 10
Peg3	18616	3.21	2.32	1.78	paternally expressed 3; antisense transcript gene of Peg3
Pik3r3	18710	1.43	1.26	1.28	phosphatidylinositol 3 kinase, regulatory subunit, polypeptide 3 (p55)
Pip4k2a	18718	1.34	1.31	1.22	phosphatidylinositol-5-phosphate 4-kinase, type II, alpha
Pkdcc	106522	1.13	1.15	1.31	protein kinase domain containing, cytoplasmic
Plcl2	224860	1.15	1.53	1.19	phospholipase C-like 2
Plekho1	67220	1.54	1.56	1.54	pleckstrin homology domain containing, family O member 1
Pls3	102866	1.63	1.65	2.12	plastin 3 (T-isoform)
Plxnd1	67784	1.36	1.35	1.46	plexin D1
Podxl	27205	1.99	1.77	1.45	podocalyxin-like
Prkd1	18760	2.39	1.79	1.43	protein kinase D1
Prtg	235472	2.5	3.32	2.4	protogenin homolog (Gallus gallus)
Ptpn13	19249	1.92	1.83	2.03	protein tyrosine phosphatase, non-receptor type 13
Ptpn9	56294	1.45	1.95	1.78	protein tyrosine phosphatase, non-receptor type 9
Pxdn	69675	1.53	1.76	1.77	peroxidasin homolog (Drosophila)
Ramp2	54409	1.81	1.59	1.56	receptor (calcitonin) activity modifying protein 2
Rasgrp3	240168	2.05	3.2	2.9	RAS, guanyl releasing protein 3

Rbms1	56878	1.685	1.875	1.29	RNA binding motif, single stranded interacting protein 1
Rcbtb2	105670	1.49	1.9	2	regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing protein 2
Reep5	13476	1.12	1.26	1.07	receptor accessory protein 5
Rgl1	19731	1.78	1.56	1.31	ral guanine nucleotide dissociation stimulator,-like 1
Rhobtb3	73296	1.31	1.23	1.57	Rho-related BTB domain containing 3
Rhoj	80837	1.25	1.51	1.07	ras homolog gene family, member J
Rnf103	22644	1.45	1.33	1.24	ring finger protein 103
Rnf217	268291	1.9	1.89	1.5	ring finger protein 217
Rspo3	72780	3.43	3.11	1.54	R-spondin 3 homolog (Xenopus laevis)
Sema4c	20353	2.06	1.84	1.64	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4C
Sema5a	20356	1.85	1.36	1.31	sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A
Sepn1	74777	1.04	1.24	1.02	selenoprotein N, 1
Sgce	20392	1.78	2.34	1.3	sarcoglycan, epsilon
Sh3bgrl	56726	1.26	2	1.99	SH3-binding domain glutamic acid-rich protein like
Slc15a2	57738	1.6	1.27	1.42	solute carrier family 15 (H+/peptide transporter), member 2
Slc22a23	73102	1.58	1.54	1.59	solute carrier family 22, member 23
Slc25a24	229731	1.44	2.68	2.13	solute carrier family 25 (mitochondrial carrier, phosphate carrier), member 24
Slc40a1	53945	1.25	1.73	3.23	solute carrier family 40 (iron-regulated transporter), member 1
Slc4a4	54403	1.02	1.3	1.04	solute carrier family 4 (anion exchanger), member 4
Slc5a3	53881	2.78	2.69	1.49	solute carrier family 5 (inositol transporters), member 3
Slc7a8	50934	1.45	1.33	1.62	solute carrier family 7 (cationic amino acid transporter, y+ system), member 8
Slit3	20564	1.31	1.15	1.7	slit homolog 3 (Drosophila)
Smad1	17125	1.61	1.91	1.36	MAD homolog 1 (Drosophila)
Smo	319757	1.25	1.4	1.2	predicted gene 4066; smoothened homolog (Drosophila)
Snai1	20613	1.19	1.96	1.09	snail homolog 1 (Drosophila)
Snx33	235406	1.35	1.43	1.11	sorting nexin 33
Soat1	20652	1.43	2.27	2.21	sterol O-acyltransferase 1
Sox11	20666	1.26	1.31	1.6	SRY-box containing gene 11
St3gal6	54613	2.55	2.84	1.76	ST3 beta-galactoside alpha-2,3-sialyltransferase 6
St6gal1	20440	1.4	1.39	1.61	beta galactoside alpha 2,6 sialyltransferase 1
Stk39	53416	2.33	2.27	2.87	serine/threonine kinase 39, STE20/SPS1 homolog (yeast)
Stxbp4	20913	1.59	1.415	1.08	syntaxin binding protein 4
Stxbp5	78808	1.55	1.39	1.17	syntaxin binding protein 5 (tomosyn)
Suv39h1	20937	2.17	2.26	1.46	suppressor of variegation 3-9 homolog 1 (Drosophila)
Tax1bp3	76281	1.06	1.32	1.415	Tax1 (human T-cell leukemia virus type I) binding protein 3; predicted gene 13597; similar to Tax1 (human T-cell leukemia virus type I) binding protein 3
Tbx20	57246	3.26	3.64	1.92	T-box 20
Tcf12	21406	1.64	1.45	1.47	transcription factor 12
Tcn2	21452	1.33	1.19	1.04	transcobalamin 2
Tead2	21677	1.2	1.23	1.24	TEA domain family member 2
Tes	21753	1.25	1.6	2.25	testis derived transcript
Tmc7	209760	2	1.73	1.61	transmembrane channel-like gene family 7; similar to Tmc7 protein
Tmem164	209497	1.6	1.6	1.47	transmembrane protein 164
Tmem176b	65963	2.01	1.82	1.04	transmembrane protein 176B
Tnfrsf19	29820	1.54	1.32	1.07	tumor necrosis factor receptor superfamily, member 19
Tpm4	326618	1.08	1.24	1.185	tropomyosin 4; predicted gene 7809
Ttc21b	73668	1.13	1.36	1.4	tetratricopeptide repeat domain 21B
Uggt2	66435	1.16	1.4	1.2	UDP-glucose ceramide glucosyltransferase-like 2
Unc5b	107449	1.05	1.63	2.2	unc-5 homolog B (C. elegans)
Vamp8	22320	1.08	1.15	1.37	vesicle-associated membrane protein 8
Vav3	57257	1.7	1.37	1.23	vav 3 oncogene
Vcan	13003	2.14	1.97	1.78	versican
Vopp1	232023	1.66	2.38	2.1	expressed sequence AW146242
Yaf2	67057	1.24	1.4	1.17	YY1 associated factor 2
Zdhhc2	70546	1.63	1.91	2.19	zinc finger, DHHC domain containing 2

Zeb2	24136	2.14	2.34	1.27	zinc finger E-box binding homeobox 2
Zfp358	140482	1.38	1.14	1.02	zinc finger protein 358
Zfp521	225207	1.96	2.12	1.97	zinc finger protein 521
Zfp711	245595	3.61	2.06	1.79	zinc finger protein 711

Concordantly downregulated genes in wt, dnmt1-/- and TKO EBs between d0-4					
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change dnmt1-/-**	**Fold change TKO**	**Gene name**
1190003J15Rik	76974	-1.38	-1.53	-1.06	RIKEN cDNA 1190003J15 gene
1700001L05Rik	69291	-2.13	-2.48	-1.31	RIKEN cDNA 1700001L05 gene
1700019N12Rik	67077	-1.12	-1.04	-1.04	RIKEN cDNA 1700019N12 gene
1700029P11Rik	66346	-1.63	-1.92	-1.87	RIKEN cDNA 1700029P11 gene; predicted gene 9050
1700061G19Rik	78625	-1.5	-2.04	-1.93	RIKEN cDNA 1700061G19 gene
1700072H12Rik	73493	-1.17	-1.2	-1.17	RIKEN cDNA 1700072H12 gene
2410004A20Rik	66991	-1.09	-2.08	-1.79	RIKEN cDNA 2410004A20 gene
2410137M14Rik	76797	-1.65	-1.88	-1.61	RIKEN cDNA 2410137M14 gene
2410141K09Rik	76803	-3.22	-2.04	-1.28	RIKEN cDNA 2410141K09 gene; predicted gene 7771; predicted gene 7896
8430410A17Rik	232210	-1.19	-1.05	-1.17	RIKEN cDNA 8430410A17 gene
Abcb1a	18671	-1.56	-1.43	-1.63	ATP-binding cassette, sub-family B (MDR/TAP), member 1A
Accsl	381411	-2.42	-2.23	-2.46	1-aminocyclopropane-1-carboxylate synthase homolog (Arabidopsis)(non-functional)-like
Actn3	11474	-1.52	-1.23	-1.52	actinin alpha 3
Adap1	231821	-1.82	-2.01	-2.06	ArfGAP with dual PH domains 1
Adrb3	11556	-1.12	-1.72	-1.22	adrenergic receptor, beta 3
Aoah	27052	-1.22	-1.64	-1.03	acyloxyacyl hydrolase
Arid5b	71371	-1.65	-1.96	-1.19	similar to modulator recognition factor 2; AT rich interactive domain 5B (MRF1-like)
Atcay	16467	-1.2	-1.1	-1.1	ataxia, cerebellar, Cayman type homolog (human)
Atp1a3	232975	-1.01	-1.72	-1.16	ATPase, Na+/K+ transporting, alpha 3 polypeptide
AU018091	245128	-4.02	-3.4	-2.03	expressed sequence AU018091
BC021614	225884	-1.09	-1.14	-1.28	cDNA sequence BC021614
Bcl3	12051	-1.64	-1.79	-1.67	B-cell leukemia/lymphoma 3
Bnc2	242509	-2.4	-2.31	-1.06	basonuclin 2
Btbd11	74007	-1.24	-1.23	-1.19	BTB (POZ) domain containing 11
C330019L16Rik	208111	-1.91	-1.38	-1.38	RIKEN cDNA C330019L16 gene
C330022B21Rik	78699	-2.65	-1.98	-1.36	RIKEN cDNA C330022B21 gene
Calml4	75600	-1.34	-1.97	-1.21	calmodulin-like 4
Camk1d	227541	-1.12	-1.7	-1.18	calcium/calmodulin-dependent protein kinase ID
Cd68	12514	-1.02	-1.57	-1.37	CD68 antigen
Cd97	26364	-1.3	-1.8	-1.2	CD97 antigen
Cdyl2	75796	-2.16	-2.93	-2.06	chromodomain protein, Y chromosome-like 2
Cep55	74107	-1.47	-1.65	-1.04	centrosomal protein 55
Chrna9	231252	-1.31	-1.27	-1.55	cholinergic receptor, nicotinic, alpha polypeptide 9
Clgn	12745	-2.4	-2.3	-1.66	calmegin
Cltb	74325	-2.88	-2.68	-1.96	clathrin, light polypeptide (Lcb)
Cnpy1	269637	-2.42	-1.97	-1.1	canopy 1 homolog (zebrafish)
Cobl	12808	-1.5	-1.5	-1.31	cordon-bleu
Cphx	105594	-1.0433333 33	-1.286666667	-1.02	predicted gene 2135; predicted gene 2104; cytoplasmic polyadenylated homeobox
Cpsf4l	52670	-1.71	-2.06	-1.47	cleavage and polyadenylation specific factor 4-like
Ctnnal1	54366	-1.49	-1.53	-1	catenin (cadherin associated protein), alpha-like 1
D14Ertd668e	219132	-2.1	-2.69	-1.59	predicted gene 6907; predicted gene 6904; predicted gene 4902; DNA segment, Chr 14, ERATO Doi 668, expressed; PHD finger protein 11
Ddc	13195	-1.02	-1.13	-1.12	dopa decarboxylase
Dnmt3l	54427	-4.6	-3.53	-2.09	similar to DNA cytosine-5 methyltransferase 3-like protein; DNA (cytosine-5-)-methyltransferase 3-like
Dppa2	73703	-3.98	-2.98	-1.28	similar to developmental pluripotency-associated 2; predicted gene 9298; developmental pluripotency associated 2
Dusp27	240892	-2.41	-3.41	-2.73	dual specificity phosphatase 27 (putative)
Enox1	239188	-1.88	-2.07	-1.17	ecto-NOX disulfide-thiol exchanger 1
Ephx1	13849	-1.27	-2.17	-1.89	epoxide hydrolase 1, microsomal

Esrrb	26380	-4.11	-3.58	-1.58	estrogen related receptor, beta
Fblim1	74202	-2.05	-2.52	-1.87	filamin binding LIM protein 1
Fbxo15	50764	-5.41	-3.95	-2.25	F-box protein 15
Fcgrt	14132	-1.02	-1.55	-1.21	Fc receptor, IgG, alpha chain transporter
Fgf17	14171	-2.4	-1.93	-1.54	fibroblast growth factor 17
Fgf4	14175	-1.91	-2.19	-1.4	fibroblast growth factor 4
Folr1	14275	-1.38	-1.48	-1.09	folate receptor 1 (adult)
Gatsl3	71962	-1.25	-1.37	-1.07	GATS protein-like 3
Gbx2	14472	-4.36	-3.2	-1.33	gastrulation brain homeobox 2
Gdf3	14562	-1.42	-1.48	-1.15	growth differentiation factor 3
Gli1	14632	-1.99	-1.73	-1.56	GLI-Kruppel family member GLI1
Glod5	69824	-4.52	-3.07	-1.54	glyoxalase domain containing 5
Gm10324	628709	-1.855	-1.74	-1.35	predicted gene 10324
Gm10351	100127434	-1.11	-1.65	-1.12	predicted gene 10351
Gm10522	100038407	-2.29	-1.6	-1.03	predicted gene 10522
Gm13051	626316	-2.02	-2	-1.05	predicted gene 13051; predicted gene 8935; predicted gene 13151
Gm13152	195531	-1.76	-1.703333333	-1.163333333	predicted gene 13152
Gm13154	433804	-3.77	-2.25	-1.8	predicted gene 13154
Gm13242	100041379	-4.73	-3.34	-1.74	predicted gene 13242
Gm13251	433791	-2.115	-1.525	-1.12	predicted gene 13251; predicted gene, OTTMUSG00000010657; RIKEN cDNA 1700029I01 gene
Gm15514	434225	-1.42	-1.7	-1.52	predicted gene 15514
Gm4782	213320	-1.92	-2.2	-1.34	predicted gene 4782; predicted gene 5989
Gm5301	384356	-1.886666667	-2.153333333	-1.243333333	predicted gene 5301
Gpa33	59290	-1.96	-2.84	-1.92	glycoprotein A33 (transmembrane)
Gpx6	75512	-1.74	-2.45	-1.28	glutathione peroxidase 6
Gtsf1l	68236	-2.04	-2.45	-1.37	zinc finger protein 850; gametocyte specific factor 1-like
H2-M5	240095	-1.6	-1.62	-1.07	histocompatibility 2, M region locus 5
Hck	15162	-2.81	-2.74	-2.88	hemopoietic cell kinase
Hormad1	67981	-1.17	-2.34	-2.025	HORMA domain containing 1; predicted gene 7167
Hormad2	75828	-2.29	-1.11	-1.17	HORMA domain containing 2
Hsd17b14	66065	-2.96	-2.89	-2.02	hydroxysteroid (17-beta) dehydrogenase 14
Hsf2bp	74377	-3.66	-3.18	-1.78	heat shock transcription factor 2 binding protein
Icam1	15894	-1.46	-1.96	-1.39	intercellular adhesion molecule 1
Inhbb	16324	-1.11	-1.11	-1.66	inhibin beta-B
Inpp5d	16331	-2.4	-1.87	-1.35	inositol polyphosphate-5-phosphatase D
Itgb7	16421	-1.42	-1.68	-1.35	integrin beta 7
Jam2	67374	-3.98	-3.84	-2.15	junction adhesion molecule 2
Klf2	16598	-2.95	-2.18	-1.2	Kruppel-like factor 2 (lung)
Klf4	16600	-1.02	-1.84	-1.96	Kruppel-like factor 4 (gut)
Klhl13	67455	-2.26	-1.74	-1.35	kelch-like 13 (Drosophila)
Krt17	16667	-1.8	-1.97	-1.45	keratin 17
Krt42	68239	-1.62	-1.51	-1.13	keratin 42
Lrrc2	74249	-2.57	-3.29	-2.21	leucine rich repeat containing 2; similar to Leucine-rich repeat-containing protein 2
Mael	98558	-2.15	-2.67	-1.51	maelstrom homolog (Drosophila)
Mageb16	71967	-4.295	-1.785	-1.555	predicted gene 15072; RIKEN cDNA 2410003J06 gene
Mlh3	217716	-1.31	-1.45	-1.15	mutL homolog 3 (E coli)
Morc1	17450	-3.38	-3.67	-2.53	microrchidia 1
Mreg	381269	-2.62	-2.26	-1.2	melanoregulin
Ms4a14	383435	-1.55	-1.76	-1.15	membrane-spanning 4-domains, subfamily A, member 14
Myof	226101	-1.34	-1.71	-1.16	myoferlin
Myst4	54169	-2.28	-2.21	-1.64	MYST histone acetyltransferase monocytic leukemia 4
Nefh	380684	-2.4	-2.48	-1.39	similar to neurofilament protein; neurofilament, heavy polypeptide
Neurod1	18012	-1.61	-1.55	-1.18	neurogenic differentiation 1; neurogenic differentiation 5
Nfatc2ip	18020	-1.37	-1.38	-1.01	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2 interacting protein
Nphs1	54631	-1.49	-1.94	-1.04	nephrosis 1 homolog, nephrin (human)
Nqo1	18104	-2.51	-2.33	-1.39	NAD(P)H dehydrogenase, quinone 1

Nr5a2	26424	-2.45	-2.43	-1.12	nuclear receptor subfamily 5, group A, member 2
Nup210l	77595	-1.09	-1.236666667	-1.11	nucleoporin 210-like
Obox6	252830	-1	-1.76	-1.38	predicted gene 5496; oocyte specific homeobox 6
Ooep	67968	-2.19	-2.5	-1.54	oocyte expressed protein homolog (dog)
Pcolce	18542	-2.06	-1.96	-2.34	procollagen C-endopeptidase enhancer protein
Pdgfc	54635	-1.16	-2.61	-1.35	platelet-derived growth factor, C polypeptide
Pim1	18712	-2.11	-1.88	-1.06	proviral integration site 1
Piwil2	57746	-1.37	-1.7	-1.61	piwi-like homolog 2 (Drosophila)
Pla2g10	26565	-1.55	-2.04	-1.12	phospholipase A2, group X
Pnma5	385377	-2.16	-2.17	-2.08	paraneoplastic antigen family 5
Psma8	73677	-2.95	-2.6	-1.78	proteasome (prosome, macropain) subunit, alpha type, 8
Ptch2	19207	-1.86	-1.91	-1.1	patched homolog 2
Reep6	70335	-1.37	-1.61	-1.3	receptor accessory protein 6
Rex2	19715	-4.02	-2.47	-1.3	hypothetical protein LOC100048814; predicted gene 13138; reduced expression 2; predicted gene 13164
Ripply1	622473	-1.65	-1.3	-1.41	ripply1 homolog (zebrafish)
Rmnd5b	66089	-1.39	-1.41	-1.21	required for meiotic nuclear division 5 homolog B (S. cerevisiae)
Rnf125	67664	-2.54	-2.51	-1.37	ring finger protein 125
Rnf17	30054	-1.91	-1.51	-1.49	ring finger protein 17
Rpl10l	238217	-1.94	-1.27	-1.15	predicted gene 14460; predicted gene 13891; predicted gene 2387; ribosomal protein 10; predicted gene 7476; predicted gene 4167; predicted gene 5621; predicted gene 3379; similar to QM protein; predicted gene 11450; predicted gene 6564; predicted gene 3405; predicted gene 10041; predicted gene 4892; ribosomal protein L10-like
Sept1	54204	-1.36	-2.43	-1.58	septin 1
Serpinb6c	97848	-1.55	-1.57	-1.64	serine (or cysteine) peptidase inhibitor, clade B, member 6c
Sfrp4	20379	-2.07	-2.28	-1.24	secreted frizzled-related protein 4
Slc15a1	56643	-1.1	-1.8	-1.13	solute carrier family 15 (oligopeptide transporter), member 1
Slc25a12	78830	-2.33	-2.29	-1.83	solute carrier family 25 (mitochondrial carrier, Aralar), member 12
Slc28a1	434203	-1.74	-1.48	-1.23	solute carrier family 28 (sodium-coupled nucleoside transporter), member 1
Slc47a1	67473	-1.04	-1.52	-1.25	solute carrier family 47, member 1
Slc5a11	233836	-1.28	-1.55	-1.58	solute carrier family 5 (sodium/glucose cotransporter), member 11
Slc5a4b	64454	-1.11	-2.49	-1.57	solute carrier family 5 (neutral amino acid transporters, system A), member 4b
Slc7a1	11987	-1.35	-1.68	-2.22	solute carrier family 7 (cationic amino acid transporter, y+ system), member 1
Socs3	12702	-1.51	-1.85	-1.87	suppressor of cytokine signaling 3
Sod2	20656	-1.06	-1.43	-1.32	superoxide dismutase 2, mitochondrial
Spats1	71020	-1.01	-1.39	-1.29	spermatogenesis associated, serine-rich 1
Spp1	20750	-4.35	-4.19	-2.05	secreted phosphoprotein 1
Stk31	77485	-1.91	-2.44	-1.2	serine threonine kinase 31
Stmn2	20257	-1.48	-1.85	-1.29	stathmin-like 2
Susd2	71733	-1.41	-1.78	-1.78	sushi domain containing 2
Syce1	74075	-2.35	-2.09	-1.69	synaptonemal complex central element protein 1
Tcea3	21401	-2.01	-2.2	-1.29	transcription elongation factor A (SII), 3
Tcl1	21432	-3.06	-3.31	-2.37	T-cell lymphoma breakpoint 1
Tdrd12	71981	-4.27	-3.25	-2.23	tudor domain containing 12
Tet1	52463	-2.326666667	-2.63	-1.686666667	tet oncogene 1
Tet2	214133	-1.89	-2.34	-2.22	tet oncogene family member 2
Tex14	83560	-2.01	-2.64	-2.26	testis expressed gene 14
Tgm1	21816	-2.21	-1.31	-2.05	transglutaminase 1, K polypeptide
Timp1	21857	-1.9	-1.72	-1.18	tissue inhibitor of metalloproteinase 1
Tm7sf3	67623	-1.52	-1.63	-1.46	transmembrane 7 superfamily member 3
Trap1a	22037	-2.92	-1.45	-1.3	tumor rejection antigen P1A
Trim25	217069	-1.31	-1.23	-1.08	tripartite motif-containing 25
Triml1	244448	-2.25	-2.45	-2.07	tripartite motif family-like 1
Triml2	622117	-3.5	-2.85	-1.93	tripartite motif family-like 2
Ttc39b	69863	-2.25	-2.09	-1.145	tetratricopeptide repeat domain 39B
Ttll6	237930	-1.06	-1.15	-1.18	tubulin tyrosine ligase-like family, member 6

Ttpa	50500	-1.24	-1.43	-1.16	tocopherol (alpha) transfer protein
Tubb3	22152	-1.82	-2.02	-1.24	tubulin, beta 3; tubulin, beta 3, pseudogene 1
Uchl1	22223	-3.04	-2.64	-1.11	ubiquitin carboxy-terminal hydrolase L1
Vmn2r-ps104	100041915	-1.84	-1.76	-1.09	vomeronasal 2, receptor, pseudogene 104
Zbtb32	58206	-1.21	-1.62	-1.27	zinc finger and BTB domain containing 32
Zfp229	381067	-1.4	-1.82	-1.34	zinc finger protein
Zfp296	63872	-2.15	-1.79	-1.15	zinc finger protein 296
Zfp42	22702	-5.06	-3.96	-1.85	zinc finger protein 42
Zfp57	22715	-1.88	-2.12	-1.11	zinc finger protein 57
Zfp640	386626	-4.806	-2.92	-1.805	zinc finger protein 640
Zp3	22788	-1.7	-1.82	-1.84	zona pellucida glycoprotein 3
Zswim1	71971	-1.74	-1.58	-1.08	zinc finger, SWIM domain containing 1

Concordantly upregulated genes in wt and dnmt1-/- EBs between d0-4				
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change dnmt1-/-**	**Gene name**
1110002B05Rik	104725	1.16	1.17	RIKEN cDNA 1110002B05 gene
1110032A03Rik	68721	1.37	1.28	RIKEN cDNA 1110032A03 gene; hypothetical protein LOC100048251
1110038D17Rik	68778	1.15	1.06	RIKEN cDNA 1110038D17 gene
1110051M20Rik	228356	1.15	1.21	RIKEN cDNA 1110051M20 gene
1700025G04Rik	69399	1.345	1.02	RIKEN cDNA 1700025G04 gene
2010106G01Rik	66552	1.08	1.11	RIKEN cDNA 2010106G01 gene
2610008E11Rik	72128	1.37	1.04	RIKEN cDNA 2610008E11 gene
6230427J02Rik	68176	1.96	2.03	RIKEN cDNA 6230427J02 gene
6330403K07Rik	103712	2.09	1.88	RIKEN cDNA 6330403K07 gene
9230110C19Rik	234912	1.42	1.36	RIKEN cDNA 9230110C19 gene
Ablim1	226251	1.1	1.02	actin-binding LIM protein 1
Adcy6	11512	1.29	1.55	adenylate cyclase 6
Adssl1	11565	2.03	1.1	adenylosuccinate synthetase like 1
Akt3	23797	1.39	1.53	thymoma viral proto-oncogene 3
Alg8	381903	1.23	1.06	asparagine-linked glycosylation 8 homolog (yeast, alpha-1,3-glucosyltransferase)
Amn1	232566	1.16	1.01	antagonist of mitotic exit network 1 homolog (S. cerevisiae)
Anxa4	11746	1.17	1.26	annexin A4
Apcdd1	494504	2.49	2.46	adenomatosis polyposis coli down-regulated 1
Arsa	11883	1.74	1.8	arylsulfatase A
B130016D09Rik	436015	2.28	2.09	RIKEN cDNA B130016D09 cDNA
Bach2	12014	1.1	1.08	BTB and CNC homology 2
Bambi	68010	2.32	1.49	BMP and activin membrane-bound inhibitor, homolog (Xenopus laevis)
Bmp2	12156	2.98	2.57	bone morphogenetic protein 2
Bmp5	12160	3.24	3.61	bone morphogenetic protein 5
Bmper	73230	3.16	2.97	BMP-binding endothelial regulator
Camk1g	215303	1.05	1.26	calcium/calmodulin-dependent protein kinase I gamma
Capn6	12338	4.56	3.61	calpain 6
Casd1	213819	1.41	1.02	CAS1 domain containing 1; similar to O-acetyltransferase
Ccng2	12452	2.44	1.35	cyclin G2
Cd24a	12484	1.68	1.54	CD24a antigen
Cd44	12505	1.8	1.17	CD44 antigen
Cd99l2	171486	1.42	1.07	CD99 antigen-like 2
Cdx2	12591	2.26	1.24	caudal type homeo box 2
Cer1	12622	2.28	3.15	cerberus 1 homolog (Xenopus laevis)
Cfc1	12627	2.84	2.77	cripto, FRL-1, cryptic family 1
Cgnl1	68178	1.79	1.58	cingulin-like 1
Chst14	72136	1.22	1.47	carbohydrate (N-acetylgalactosamine 4-0) sulfotransferase 14
Clic6	209195	2.63	2.21	chloride intracellular channel 6
Cnksr2	245684	2.36	1.29	connector enhancer of kinase suppressor of Ras 2
Col1a1	12842	1.22	1.43	collagen, type I, alpha 1
Crb2	241324	2.01	1.64	crumbs homolog 2 (Drosophila)
Cttnbp2	30785	1.08	1.14	cortactin binding protein 2

Cyp4f15	106648	1.4	1.19	cytochrome P450, family 4, subfamily f, polypeptide 15
Dkk1	13380	2.15	2.27	dickkopf homolog 1 (Xenopus laevis)
Dlg5	71228	1.24	1.07	discs, large homolog 5 (Drosophila)
Dok4	114255	2.95	1.95	docking protein 4
E330009J07Rik	243780	1.02	1.19	RIKEN cDNA E330009J07 gene
E330027M22Rik	100038419	2.46	1.67	RIKEN cDNA gene, E330027M22Rik
Ednra	13617	1.83	1.38	endothelin receptor type A
Efemp2	58859	1.07	1.04	epidermal growth factor-containing fibulin-like extracellular matrix protein 2
Efha1	68514	1.23	1.38	EF hand domain family A1
Efna1	13636	2.26	1.51	ephrin A1
Efnb1	13641	1.4	1.2	ephrin B1
Eid1	58521	1.49	1.47	EP300 interacting inhibitor of differentiation 1
Emb	13723	1.39	1.62	embigin
Extl2	58193	1.09	1.2	exostoses (multiple)-like 2
Fam55d	244853	1	1.04	family with sequence similarity 55, member D
Fbln2	14115	1.44	1.58	fibulin 2
Fbn1	14118	1.68	1.68	fibrillin 1
Fbn2	14119	2.44	1.7	similar to fibrillin 2; fibrillin 2
Fgf3	14174	2.52	1.77	fibroblast growth factor 3
Fkbp14	231997	1.34	1.02	FK506 binding protein 14
Foxf1a	15227	4.015	3.255	forkhead box F1a
Foxo4	54601	1.65	1.35	forkhead box O4
Frat1	14296	1.13	1.14	frequently rearranged in advanced T-cell lymphomas
Frem1	329872	1.66	1.72	Fras1 related extracellular matrix protein 1
Frmd6	319710	1.22	1.33	predicted gene 5780; FERM domain containing 6
Fstl1	14314	1.19	1.3	follistatin-like 1
Fut10	171167	1.14	1.32	fucosyltransferase 10
Fut11	73068	1.01	1.17	fucosyltransferase 11
Galnt1	14423	1.42	1.57	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 1
Galnt7	108150	1.67	1.36	UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase 7
Gata2	14461	1.43	1.64	GATA binding protein 2
Gata4	14463	2.19	2.67	GATA binding protein 4
Gem	14579	1.66	1.39	GTP binding protein (gene overexpressed in skeletal muscle)
Gm10397	100038401	1.5	1.58	predicted gene 10397
Gm9958	791294	1.17	1.03	predicted gene 9958
Gnas	14683	1.14	1.535	GNAS (guanine nucleotide binding protein, alpha stimulating) complex locus
Gnb4	14696	1.44	1.05	guanine nucleotide binding protein (G protein), beta 4
Gpc3	14734	2.89	2.12	glypican 3
Greb1	268527	2.28	2.21	gene regulated by estrogen in breast cancer protein
Gsbs	19051	1.45	1.28	G substrate
Gulp1	70676	1.71	1.6	GULP, engulfment adaptor PTB domain containing 1
H1f0	14958	1.88	1.69	H1 histone family, member 0
Hand1	15110	2.92	2.18	heart and neural crest derivatives expressed transcript 1
Hes1	15205	1.09	1.02	hairy and enhancer of split 1 (Drosophila)
Hmga2	15364	1.73	2.13	predicted gene 7996; high mobility group AT-hook 2
Hoxb2	103889	1.94	1.43	homeo box B2
Ifngr2	15980	1.28	1.05	interferon gamma receptor 2
Igf2	16002	3.39	1.53	insulin-like growth factor 2
Il10rb	16155	1.23	1.39	interleukin 10 receptor, beta
Ip6k2	76500	1.1	1	hypothetical protein LOC100044300; inositol hexaphosphate kinase 2
Isl1	16392	1.99	1.07	ISL1 transcription factor, LIM/homeodomain
Kcnmb4	58802	1.13	1.14	potassium large conductance calcium-activated channel, subfamily M, beta member 4
Kif21b	16565	1.67	1.05	kinesin family member 21B
Klhl23	277396	1	1.2	kelch-like 23 (Drosophila)
Lama4	16775	3.06	2.34	laminin, alpha 4
Lbh	77889	1.11	1.14	limb-bud and heart
Lgr5	14160	2.92	2.79	leucine rich repeat containing G protein coupled receptor 5
Lhfpl2	218454	2.42	2.26	lipoma HMGIC fusion partner-like 2

Lhx1	16869	2.38	3.01	LIM homeobox protein 1
Lix1l	280411	1.51	1.18	Lix1-like
Lrp1	16971	1.95	1.19	low density lipoprotein receptor-related protein 1
Lsp1	16985	1.22	1.23	lymphocyte specific 1
Maml3	433586	1.25	1.22	mastermind like 3 (Drosophila)
Man1c1	230815	1.37	1.03	mannosidase, alpha, class 1C, member 1
Man2a2	140481	1	1.01	mannosidase 2, alpha 2
Meis2	17536	1.45	1.52	Meis homeobox 2
Mfap2	17150	1.19	1.06	microfibrillar-associated protein 2
Morc4	75746	2.75	2.66	microrchidia 4
Mospd1	70380	1.86	1.24	predicted gene 2147; motile sperm domain containing 1
Ms4a4d	66607	2.67	2.26	membrane-spanning 4-domains, subfamily A, member 4D
Msx2	17702	3.35	2.55	similar to homeobox protein; homeobox, msh-like 2
Nfxl1	100978	2.18	1.91	nuclear transcription factor, X-box binding-like 1
Nkd1	93960	2.04	2.24	naked cuticle 1 homolog (Drosophila); similar to naked cuticle 1 homolog
Nrk	27206	1.73	1.44	Nik related kinase
Nrp1	18186	5.31	4.62	neuropilin 1
Nxf7	170722	2.08	2.85	nuclear RNA export factor 7
Opn3	13603	1.81	1.71	opsin 3
Pcmtd2	245867	1.55	1.02	protein-L-isoaspartate (D-aspartate) O-methyltransferase domain containing 2
Pdgfrb	18596	2.25	1.28	platelet derived growth factor receptor, beta polypeptide
Pigp	56176	1.64	1.22	phosphatidylinositol glycan anchor biosynthesis, class P; predicted gene 9001
Pknox2	208076	1.07	1.12	Pbx/knotted 1 homeobox 2
Plac1	56096	3.16	2.3	placental specific protein 1
Plxna2	18845	1.67	1.16	plexin A2
Pmp22	18858	3.84	4.33	peripheral myelin protein 22
Polg	18975	1.16	1.04	polymerase (DNA directed), gamma
Ppic	19038	2.48	2.46	peptidylprolyl isomerase C
Prdm6	225518	3.45	3.07	PR domain containing 6
Prkar2b	19088	2.71	1.82	protein kinase, cAMP dependent regulatory, type II beta
Prox1	19130	1.93	1.56	prospero-related homeobox 1
Prss12	19142	1.11	1.59	protease, serine, 12 neurotrypsin (motopsin)
Pth1r	19228	1.13	1.24	parathyroid hormone 1 receptor
Ptpla	30963	1.38	1.26	protein tyrosine phosphatase-like (proline instead of catalytic arginine), member a
Rcn1	19672	1.5	1.27	reticulocalbin 1
Reck	53614	1.7	1.06	reversion-inducing-cysteine-rich protein with kazal motifs; similar to Reversion-inducing cysteine-rich protein with Kazal motifs precursor (mRECK)
Rftn1	76438	2.13	2.26	raftlin lipid raft linker 1
Rgs5	19737	2.98	1.91	regulator of G-protein signaling 5
Rhou	69581	1.13	1.48	ras homolog gene family, member U
Rnf130	59044	1.14	1.19	ring finger protein 130; similar to Ring finger protein 130
Rnf19a	30945	1.5	1.02	ring finger protein 19A
Rora	19883	1.82	1.2	RAR-related orphan receptor alpha
Rps6ka6	67071	1.58	1.3	ribosomal protein S6 kinase polypeptide 6
Runx1	12394	1.9	2.61	runt related transcription factor 1
Sccpdh	109232	1.8	1.51	similar to Saccharopine dehydrogenase (putative); saccharopine dehydrogenase (putative)
Sdc2	15529	1.2	1.22	syndecan 2
Sdc3	20970	1.08	1.16	syndecan 3
Sdcbp	53378	1.15	1.27	similar to syntenin; syndecan binding protein
Sema3a	20346	1.76	1.63	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A; hypothetical protein LOC100044161
Sema6d	214968	1.49	1.71	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6D
Senp7	66315	1.63	1.28	SUMO1/sentrin specific peptidase 7
Serpinh1	12406	1.3	1.35	serine (or cysteine) peptidase inhibitor, clade H, member 1
Sgcb	24051	1.5	1.56	sarcoglycan, beta (dystrophin-associated glycoprotein)
Sh3kbp1	58194	1.11	1.04	SH3-domain kinase binding protein 1
Sh3pxd2b	268396	1.22	1.14	SH3 and PX domains 2B
Slc22a21	56517	1.3	1.51	solute carrier family 22 (organic cation transporter), member 21
Slc36a4	234967	1.25	1.15	solute carrier family 36 (proton/amino acid symporter), member 4

Slc44a5	242259	1.91	2.58	solute carrier family 44, member 5
Slc9a6	236794	1.31	1.02	solute carrier family 9 (sodium/hydrogen exchanger), member 6
Smad3	17127	1.43	1.19	MAD homolog 3 (Drosophila)
Smad6	17130	2.49	2.06	MAD homolog 6 (Drosophila)
Smarca2	67155	1.79	1.8	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2
Snx1	56440	1.15	1.11	sorting nexin 1
Snx6	72183	1.12	1.11	similar to sorting nexin 6; sorting nexin 6
Sox17	20671	1.41	1.01	SRY-box containing gene 17
Spin2	278240	2.7	1.55	spindlin family, member 2
Stard13	243362	1.14	1.28	StAR-related lipid transfer (START) domain containing 13; similar to serologically defined colon cancer antigen 13
Stard8	236920	3.11	2.15	START domain containing 8
Stat5b	20851	1.27	1.21	signal transducer and activator of transcription 5B
Stx18	71116	1.06	1.17	syntaxin 18
Tbc1d12	209478	1.12	1.22	TBC1D12: TBC1 domain family, member 12
Tbx3	21386	1.89	1.11	T-box 3
Tceal1	237052	1.5	1.52	transcription elongation factor A (SII)-like 1
Tgfb1	21803	1.24	1.09	transforming growth factor, beta 1
Tgfb2	21808	1.57	1.4	transforming growth factor, beta 2
Timp2	21858	1.28	1.3	tissue inhibitor of metalloproteinase 2
Tlcd2	380712	1.03	1.06	TLC domain containing 2
Tm6sf1	107769	1.1	1.6	transmembrane 6 superfamily member 1
Tmem106b	71900	1.01	1.03	transmembrane protein 106B
Tmem132a	98170	1.01	1.11	transmembrane protein 132A
Tmem22	245020	2.57	2.99	transmembrane protein 22
Tmem88	67020	1.31	1.17	transmembrane protein 88
Tmem9b	56786	1.27	1.08	TMEM9 domain family, member B
Tmprss2	50528	2.36	1.96	transmembrane protease, serine 2
Tnfrsf1a	21937	1.5	1.08	tumor necrosis factor receptor superfamily, member 1a
Top2b	21974	1.92	1.69	topoisomerase (DNA) II beta
Trpc3	22065	3.12	3.12	transient receptor potential cation channel, subfamily C, member 3
Tspan12	269831	1.16	1.1	tetraspanin 12
Tspan2	70747	1.65	1.29	tetraspanin 2
Ube2h	22214	1.9	1.26	predicted gene 2058; similar to Ubiquitin-conjugating enzyme UbcH2; ubiquitin-conjugating enzyme E2H
Ugdh	22235	1.17	1.11	UDP-glucose dehydrogenase
Uhrf2	109113	1.89	1.8	ubiquitin-like, containing PHD and RING finger domains 2
Ulk2	29869	3.78	2.14	Unc-51 like kinase 2 (C. elegans)
Unc5c	22253	1.93	1.41	unc-5 homolog C (C. elegans)
Usp3	235441	1.11	1.18	ubiquitin specific peptidase 3
Usp47	74996	1.39	1.14	ubiquitin specific peptidase 47
Usp51	635253	1.35	1.46	ubiquitin specific protease 51
Vamp4	53330	1.8	1.03	vesicle-associated membrane protein 4
Vasn	246154	1.39	1.79	vasorin
Vgll4	232334	2.76	2.19	vestigial like 4 (Drosophila)
Vldlr	22359	2.12	2.02	very low density lipoprotein receptor
Vstm2b	58188	1.46	1.63	hypothetical protein LOC100045106; V-set and transmembrane domain containing 2B
Wnt2	22413	3.08	3.17	wingless-related MMTV integration site 2
Wnt5a	22418	1.42	1.62	wingless-related MMTV integration site 5A
Zc3hav1	78781	1.39	1.31	zinc finger CCCH type, antiviral 1
Zcchc11	230594	1.02	1.12	zinc finger, CCHC domain containing 11
Zcchc12	72693	1.21	1.25	zinc finger, CCHC domain containing 12
Zfp334	228876	1.05	1.02	zinc finger protein 334
Zfpm1	22761	1.38	1.55	zinc finger protein, multitype 1; similar to FOG
Zkscan3	72739	1.09	1.11	zinc finger with KRAB and SCAN domains 3

Concordantly downregulated genes in wt and dnmt1-/- EBs between d0-4				
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change dnmt1-/-**	**Gene name**

1700019D03Rik	67080	-2.18	-2.1	RIKEN cDNA 1700019D03 gene
1700029I01Rik	70005	-1.84	-1.33	predicted gene 13251; predicted gene, OTTMUSG00000010657; RIKEN cDNA 1700029I01 gene
2410012M07Rik	71979	-1.41	-1.68	RIKEN cDNA 2410012M07 gene
2610305D13Rik	112422	-3.47	-2.19	RIKEN cDNA 2610305D13 gene
2610528J11Rik	66451	-1.29	-1.06	RIKEN cDNA 2610528J11 gene
2900011O08Rik	67254	-1.37	-1.2	RIKEN cDNA 2900011O08 gene
4930519F16Rik	75106	-1.05	-1.24	RIKEN cDNA 4930519F16 gene
4932441K18Rik	353170	-1.02	-1.07	predicted gene 8258; similar to factor inhibiting activating transcription factor 4 (ATF4)-mediated transcription; RIKEN cDNA 4932441K18 gene
9630033F20Rik	319801	-1.24	-1.18	RIKEN cDNA 9630033F20 gene; similar to RIKEN cDNA 9630033F20 gene
A830080D01Rik	382252	-1.25	-1.09	RIKEN cDNA A830080D01 gene
Abcc4	239273	-2.28	-1.6	ATP-binding cassette, sub-family C (CFTR/MRP), member 4
Acsl1	14081	-1.14	-1.14	acyl-CoA synthetase long-chain family member 1
Actr3b	242894	-1.41	-1.32	ARP3 actin-related protein 3 homolog B (yeast)
Adam23	23792	-2.14	-1.28	a disintegrin and metallopeptidase domain 23; similar to ADAM23
Agtrap	11610	-1.18	-1.29	angiotensin II, type I receptor-associated protein
Alg13	67574	-1.88	-1.73	asparagine-linked glycosylation 13 homolog (S. cerevisiae)
Ankrd45	73844	-1.26	-1.41	ankyrin repeat domain 45
Ano9	71345	-1.39	-2.06	anoctamin 9
Anxa11	11744	-1.79	-1.17	annexin A11; predicted gene 2260; predicted gene 2274
Apobec1	11810	-1.39	-1.32	apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1
Apobec3	80287	-1.91	-2.08	apolipoprotein B mRNA editing enzyme, catalytic polypeptide 3
Aqp3	11828	-1.47	-1.98	aquaporin 3
Arhgap30	226652	-1.24	-1.41	Rho GTPase activating protein 30
Arhgef3	71704	-1.77	-1.55	Rho guanine nucleotide exchange factor (GEF) 3
Ass1	11898	-1.575	-1.84	argininosuccinate synthetase 1
Avpi1	69534	-1.1	-1.12	arginine vasopressin-induced 1
B3gnt7	227327	-1.31	-1	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7
B4galnt3	330406	-1.04	-1.52	beta-1,4-N-acetyl-galactosaminyl transferase 3
BC032203	210982	-1.51	-1.4	cDNA sequence BC032203
Bcat2	12036	-1.18	-1.18	branched chain aminotransferase 2, mitochondrial
Bnipl	171388	-1.13	-1.39	BCL2/adenovirus E1B 19kD interacting protein like
C330016O10Rik	212706	-1.76	-1.45	RIKEN cDNA C330016O10 gene
C77370	245555	-2.17	-1.62	expressed sequence C77370
Calcoco2	76815	-2.865	-2.745	predicted gene 11701; calcium binding and coiled-coil domain 2
Cbfa2t2	12396	-1.07	-1.06	core-binding factor, runt domain, alpha subunit 2, translocated to, 2 (human)
Cbx7	52609	-1.63	-1.37	chromobox homolog 7
Ccnb1ip1	239083	-1.27	-1.04	cyclin B1 interacting protein 1
Ccne1	12447	-1.92	-1.45	cyclin E1
Ccrn4l	12457	-1.49	-1.29	similar to carbon catabolite repression 4 protein homolog; CCR4 carbon catabolite repression 4-like (S. cerevisiae)
Cdh3	12560	-1.27	-1.39	cadherin 3
Cdv3	321022	-2.32	-1.59	carnitine deficiency-associated gene expressed in ventricle 3; predicted gene 7236
Ceacam1	26365	-1.41	-1.74	carcinoembryonic antigen-related cell adhesion molecule 1; carcinoembryonic antigen-related cell adhesion molecule 2
Chchd10	103172	-2.88	-2.53	coiled-coil-helix-coiled-coil-helix domain containing 10
Chd9	109151	-2.1	-1.29	chromodomain helicase DNA binding protein 9
Ckb	12709	-2.83	-1.6	similar to creatine kinase, brain; predicted gene 12892; creatine kinase, brain
Cnnm2	94219	-1.23	-1.23	cyclin M2
Cpt1a	12894	-1.93	-1.48	carnitine palmitoyltransferase 1a, liver
Csrnp1	215418	-1.02	-1.16	cysteine-serine-rich nuclear protein 1
Ctbp2	13017	-1.5	-1.27	C-terminal binding protein 2
Dapp1	26377	-1.89	-1.29	dual adaptor for phosphotyrosine and 3-phosphoinositides 1
Dazl	13164	-1.22	-1.48	deleted in azoospermia-like
Dclk2	70762	-1.07	-1.62	doublecortin-like kinase 2
Ddx58	230073	-1.9	-1.75	DEAD (Asp-Glu-Ala-Asp) box polypeptide 58
Dennd2c	329727	-1.45	-1.48	DENN/MADD domain containing 2C
Diap1	13367	-1.21	-1.05	diaphanous homolog 1 (Drosophila)
Dmrt1	50796	-1.42	-1.54	doublesex and mab-3 related transcription factor 1
Dmrtc1c	71083	-1.53	-1.08	DMRT-like family C1c2; DMRT-like family C1c

Dnajc21	78244	-1.44	-1.61	DnaJ (Hsp40) homolog, subfamily C, member 21
Dnajc6	72685	-2.09	-1.88	DnaJ (Hsp40) homolog, subfamily C, member 6
Dock6	319899	-1.09	-1.178571429	dedicator of cytokinesis 6
Dppa3	73708	-1.625	-1.655	developmental pluripotency-associated 3; predicted gene 6269
Dppa4	73693	-3.89	-2.66	predicted gene 5501; developmental pluripotency associated 4
Dppa5a	434423	-3.62	-2.27	developmental pluripotency associated 5A; similar to developmental pluripotency associated 5; hypothetical protein LOC674522
Duxbl	278672	-1.066666667	-1.16	predicted gene 10394; predicted gene 10391; double homeobox B-like
Efhd1	98363	-1.37	-1.34	EF hand domain containing 1
Elavl3	15571	-1.18	-1.1	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 3 (Hu antigen C)
En2	13799	-1.05	-1.61	engrailed 2
Enah	13800	-1.28	-1.34	enabled homolog (Drosophila)
Enpp3	209558	-4.06	-2.3	ectonucleotide pyrophosphatase/phosphodiesterase 3
Epb4.1l4a	13824	-1.53	-1.28	erythrocyte protein band 4.1-like 4a
Epb4.9	13829	-1.38	-1.21	erythrocyte protein band 4.9
Epha2	13836	-2.47	-1.74	Eph receptor A2
Ephx2	13850	-1.6	-2.19	epoxide hydrolase 2, cytoplasmic
Eras	353283	-1.43	-1.3	ES cell-expressed Ras
Etv5	104156	-1.94	-1.715	ets variant gene 5
Ezr	22350	-1.19	-1.02	ezrin; hypothetical protein LOC100044177
Fabp3	14077	-1.495	-1.42	fatty acid binding protein 3, muscle and heart; similar to mammary-derived growth inhibitor
Fam169a	320557	-1.42	-1.33	family with sequence similarity 169, member A
Fam178b	381337	-1.67	-1.39	family with sequence similarity 178, member B
Fgd4	224014	-1.27	-1.51	FYVE, RhoGEF and PH domain containing 4
Fkbp5	14229	-1.4	-1.32	FK506 binding protein 5
Frrs1	20321	-1.88	-1.69	ferric-chelate reductase 1
Fut9	14348	-1.96	-1.56	fucosyltransferase 9
Fzd5	14367	-1.63	-1.44	frizzled homolog 5 (Drosophila)
Gab1	14388	-2.29	-1.8	growth factor receptor bound protein 2-associated protein 1
Gap43	14432	-2	-1.12	growth associated protein 43
Gca	227960	-1.68	-1.09	grancalcin
Gcnt2	14538	-1.89	-1.5	glucosaminyl (N-acetyl) transferase 2, I-branching enzyme
Gldc	104174	-1.62	-1.08	glycine decarboxylase
Gm10451	100041694	-2.05	-1.2	hypothetical protein LOC100041694
Gm12569	622699	-2	-1.6	predicted gene 12569
Gm13225	622846	-3.81	-2.51	predicted gene 13225
Gm13235	385211	-3.81	-2.51	predicted gene 13235
Gm13939	100038454	-2.38	-2.33	predicted gene 13939
Gm6762	627480	-1.47	-1.16	cDNA sequence BC004728; predicted gene 6762
Gm6792	627821	-2.52	-2.06	predicted gene 6792
Gng3	14704	-2.26	-1.81	guanine nucleotide binding protein (G protein), gamma 3
Grhl1	195733	-1.17	-1.18	grainyhead-like 1 (Drosophila); similar to Grhl1 protein
Grhl2	252973	-1.46	-1.44	grainyhead-like 2 (Drosophila)
Gsta4	14860	-1.48	-1.32	glutathione S-transferase, alpha 4
Hap1	15114	-1.06	-1.08	huntingtin-associated protein 1
Helb	117599	-1.07	-1.33	helicase (DNA) B
Hexb	15212	-1.34	-1.46	hexosaminidase B
Hhip	15245	-1.65	-1.59	Hedgehog-interacting protein
Hpdl	242642	-1.28	-1.02	4-hydroxyphenylpyruvate dioxygenase-like
Hspb1	15507	-1.73	-1.585	heat shock protein 1
Hspbap1	66667	-1.2	-1.05	Hspb associated protein 1
Ifitm3	66141	-1.78	-1.16	interferon induced transmembrane protein 3
Impa2	114663	-1.15	-1.11	inositol (myo)-1(or 4)-monophosphatase 2
Ina	226180	-2.54	-1.905	internexin neuronal intermediate filament protein, alpha
Ino80	68142	-1.13	-1.53	INO80 homolog (S. cerevisiae); similar to yeast INO80-like protein
Irak3	73914	-1.98	-2.17	interleukin-1 receptor-associated kinase 3
Itga6	16403	-1.87	-1.32	integrin alpha 6
Itpk1	217837	-1.49	-1.49	inositol 1,3,4-triphosphate 5/6 kinase

Jarid2	16468	-1.6	-1.8	jumonji, AT rich interactive domain 2
Jub	16475	-1.63	-1.21	ajuba
Kat2b	18519	-2.08	-1.94	K(lysine) acetyltransferase 2B; predicted gene 5109
Kbtbd11	74901	-1.62	-1.65	kelch repeat and BTB (POZ) domain containing 11
Kcnk5	16529	-1.52	-1.8	potassium channel, subfamily K, member 5
Kdm4c	76804	-1.21	-1.17	lysine (K)-specific demethylase 4C
Kirrel2	243911	-1.27	-1.79	kin of IRRE like 2 (Drosophila)
Klf3	16599	-1.51	-1.16	Kruppel-like factor 3 (basic); similar to BKLF
Klf5	12224	-2.02	-1.53	Kruppel-like factor 5
Klf9	16601	-1.86	-1.86	Kruppel-like factor 9
L1td1	381591	-3.56	-2.46	LINE-1 type transposase domain containing 1
L3mbtl2	214669	-1.29	-1.51	l(3)mbt-like 2 (Drosophila)
Lck	16818	-1.41	-1.26	lymphocyte protein tyrosine kinase
Lefty1	13590	-3.97	-3.78	left right determination factor 1
Lima1	65970	-1.69	-1.4	LIM domain and actin binding 1
Liph	239759	-3.44	-3.68	lipase, member H
Lrrc34	71827	-2.24	-1.7	leucine rich repeat containing 34
Lrrc8d	231549	-1.11	-1.44	leucine rich repeat containing 8D
Ltbp4	108075	-1.39	-1.21	latent transforming growth factor beta binding protein 4
Manba	110173	-1.32	-1.52	mannosidase, beta A, lysosomal
Mcf2	109904	-2.7	-2.31	mcf.2 transforming sequence
Mep1b	17288	-1.54	-1.1	meprin 1 beta
Mkrn1	54484	-2.22	-1.84	makorin, ring finger protein, 1
Mkrn1-ps1	353327	-2.04	-1.74	makorin, ring finger protein 1, pseudogene 1
Mllt6	246198	-1.02	-1.01	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 6
Mov10	17454	-1.8	-1.81	Moloney leukemia virus 10; predicted gene 7357
Msh6	17688	-1.36	-1.29	mutS homolog 6 (E. coli)
Mtap7	17761	-1.64	-1.49	microtubule-associated protein 7
Mtmr7	54384	-1.16	-1.26	myotubularin related protein 7
Muc3	666339	-1.22	-1.33	mucin 3, intestinal
Mvp	78388	-1.08	-1.25	major vault protein
Mybl2	17865	-1.97	-1.83	myeloblastosis oncogene-like 2
Mycn	18109	-2.01	-1.24	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)
Myo10	17909	-1.04	-1.01	myosin X
Nanog	71950	-2.59	-2.04	similar to Nanog homeobox; Nanog homeobox
Nav2	78286	-1.78	-1.23	neuron navigator 2
Nid2	18074	-1.5	-1.36	nidogen 2
Nodal	18119	-2.21	-1.82	nodal
Nom1	433864	-3.27	-1.92	nucleolar protein with MIF4G domain 1
Notum	77583	-1.16	-1.86	notum pectinacetylesterase homolog (Drosophila)
Nr0b1	11614	-5.6	-4.25	nuclear receptor subfamily 0, group B, member 1
Nup210	54563	-1.17	-1.02	nucleoporin 210
Olfr985	258854	-1.27	-1.84	olfactory receptor 985
Pax6	18508	-1.78	-2.13	paired box gene 6
Pcolce2	76477	-2.17	-1.67	procollagen C-endopeptidase enhancer 2
Pcsk6	18553	-1.08	-1.1	proprotein convertase subtilisin/kexin type 6
Pde1b	18574	-1.79	-1.92	phosphodiesterase 1B, Ca2+-calmodulin dependent
Pdzd4	245469	-2.44	-1.36	PDZ domain containing 4
Phc1	13619	-2.08	-1.58	polyhomeotic-like 1 (Drosophila)
Phf17	269424	-2.49	-1.75	PHD finger protein 17
Pipox	19193	-3.23	-2.8	pipecolic acid oxidase
Pla2g1b	18778	-3.79	-3.11	phospholipase A2, group IB, pancreas
Plcb4	18798	-1.23	-1.12	phospholipase C, beta 4
Plcg2	234779	-1.44	-1.14	phospholipase C, gamma 2
Plekha7	233765	-1.04	-1.1	pleckstrin homology domain containing, family A member 7
Plk3	12795	-1.88	-2.08	polo-like kinase 3 (Drosophila)
Pls1	102502	-1.05	-1.14	plastin 1 (I-isoform)
Pltp	18830	-1.32	-1.24	phospholipid transfer protein
Pml	18854	-2.06	-1.89	promyelocytic leukemia

Pmm1	29858	-1.27	-1.02	phosphomannomutase 1
Pnldc1	240023	-2.02	-1.49	poly(A)-specific ribonuclease (PARN)-like domain containing 1
Pou5f1	18999	-2.64	-1.66	POU domain, class 5, transcription factor 1
Ppm1j	71887	-1.06	-1.24	protein phosphatase 1J
Prdm14	383491	-1.12	-2.16	PR domain containing 14
Prrg4	228413	-2.02	-1.18	proline rich Gla (G-carboxyglutamic acid) 4 (transmembrane); similar to Transmembrane gamma-carboxyglutamic acid protein 4 precursor (Proline-rich Gla protein 4) (Proline-rich gamma-carboxyglutamic acid protein 4)
Pycr2	69051	-1.2	-1.21	pyrroline-5-carboxylate reductase family, member 2
Ramp3	56089	-1.7	-1.32	receptor (calcitonin) activity modifying protein 3; similar to receptor activity modifying protein 3
Ranbp17	66011	-1.56	-1.31	RAN binding protein 17
Rarg	19411	-1.34	-1.46	retinoic acid receptor, gamma
Rasgrp2	19395	-1.43	-1.37	RAS, guanyl releasing protein 2
Rbpj	19664	-2.13	-2.005	recombination signal binding protein for immunoglobulin kappa J region
Rest	19712	-1.75	-1.72	RE1-silencing transcription factor
Rif1	51869	-1.18	-1.2	Rap1 interacting factor 1 homolog (yeast)
Rnf165	225743	-1.34	-1.3	ring finger protein 165
Rpl39l	68172	-2.26	-1.79	ribosomal protein L39-like
Rpp25	102614	-2.06	-2.1	ribonuclease P 25 subunit (human)
Scrn1	69938	-1.26	-1.21	secernin 1
Sema4b	20352	-1.18	-1.1	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4B
Sema4d	20354	-1.69	-1.69	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4D
Sh3gl2	20404	-1.62	-1.4	SH3-domain GRB2-like 2
Si	20431	-1.35	-1.45	silver
Six1	20471	-1.7	-1.71	sine oculis-related homeobox 1 homolog (Drosophila)
Slc12a6	107723	-1.675	-1.415	solute carrier family 12, member 6
Slc12a8	171286	-1.48	-1.67	solute carrier family 12 (potassium/chloride transporters), member 8
Slc17a9	228993	-1.4	-1.54	solute carrier family 17, member 9
Slc23a1	20522	-1.71	-1.23	solute carrier family 23 (nucleobase transporters), member 1
Slc25a40	319653	-1.37	-1.03	solute carrier family 25, member 40
Slc27a2	26458	-3.13	-2.21	solute carrier family 27 (fatty acid transporter), member 2
Slc37a1	224674	-1.81	-1.22	solute carrier family 37 (glycerol-3-phosphate transporter), member 1
Slc39a14	213053	-1.16	-1.4	solute carrier family 39 (zinc transporter), member 14
Slc4a11	269356	-1.03	-1.12	solute carrier family 4, sodium bicarbonate transporter-like, member 11
Slc6a15	103098	-1.5	-1.12	solute carrier family 6 (neurotransmitter transporter), member 15
Slc7a3	11989	-3.09	-2.43	solute carrier family 7 (cationic amino acid transporter, y+ system), member 3
Slc7a7	20540	-2.19	-1.61	solute carrier family 7 (cationic amino acid transporter, y+ system), member 7
Smoc1	64075	-1.39	-1.07	SPARC related modular calcium binding 1
Smpdl3b	100340	-1.64	-1.56	sphingomyelin phosphodiesterase, acid-like 3B
Sohlh2	74434	-1.77	-1.36	spermatogenesis and oogenesis specific basic helix-loop-helix 2
Sox2	20674	-3.59	-2.65	SRY-box containing gene 2
Spata22	380709	-1.04	-1.03	spermatogenesis associated 22; olfactory receptor 20
Spic	20728	-1.5	-1.77	Spi-C transcription factor (Spi-1/PU.1 related)
Spry4	24066	-2.25	-2.12	sprouty homolog 4 (Drosophila)
Stat4	20849	-1.97	-1.33	signal transducer and activator of transcription 4
Sult6b1	73671	-1.3	-1.58	sulfotransferase family, cytosolic, 6B, member 1
Syt11	229521	-1.55	-1.5	synaptotagmin XI; similar to synaptotagmin XI
Syt9	60510	-1.39	-1.05	synaptotagmin IX
Tcf15	21407	-2.29	-1.82	transcription factor 15
Tcfcp2l1	81879	-2.96	-2.47	transcription factor CP2-like 1
Tcfe3	209446	-1.22	-1.07	transcription factor E3
Tdh	433463	-4.71	-2.735	L-threonine dehydrogenase; predicted gene 13929
Tesk2	230661	-1.25	-1.3	testis-specific kinase 2
Tex11	83558	-2.17	-1.28	testis expressed gene 11
Tex19.1	73679	-2.32	-1.48	testis expressed gene 19.1
Tjp2	21873	-1.33	-1.18	tight junction protein 2
Tle4	21888	-1.95	-1.32	transducin-like enhancer of split 4, homolog of Drosophila E(spl)
Tm4sf5	75604	-1.25	-1.01	transmembrane 4 superfamily member 5
Tmem63a	208795	-1.01	-1.13	transmembrane protein 63a

Tmem8	60455	-1.24	-1.03	transmembrane protein 8 (five membrane-spanning domains)
Tns3	319939	-2.01	-1.95	tensin 3
Tpd52	21985	-1.175	-1.1	similar to Tpd52 protein; tumor protein D52
Trim13	66597	-1.82	-1.34	tripartite motif-containing 13; predicted gene 9359
Trim2	80890	-1.52	-1.6	tripartite motif-containing 2
Trim6	94088	-1.03	-1.46	tripartite motif-containing 6; similar to Tripartite motif protein 6
Tst	22117	-1.13	-1.18	thiosulfate sulfurtransferase, mitochondrial
Tuba3a	22144	-1.02	-1.09	predicted gene 5366; tubulin, alpha 3B; tubulin, alpha 3A
Tuba4a	22145	-1.44	-1.43	tubulin, alpha 4A
Twf2	23999	-1.26	-1.47	twinfilin, actin-binding protein, homolog 2 (Drosophila)
Ubxn2a	217379	-1.18	-1.12	UBX domain protein 2A; predicted gene 6245
Ulk1	22241	-1.03	-1.2	Unc-51 like kinase 1 (C. elegans)
Usp26	83563	-3.45	-1.9	ubiquitin specific peptidase 26
Usp28	235323	-1.92	-1.83	ubiquitin specific peptidase 28
Utf1	22286	-4.39	-3.28	undifferentiated embryonic cell transcription factor 1
Wdr31	71354	-1.29	-1.23	WD repeat domain 31
Wfdc15a	68221	-1.04	-1.29	WAP four-disulfide core domain 15A
Zbtb8a	73680	-2.44	-2.23	zinc finger and BTB domain containing 8a
Zfp219	69890	-1.05	-1.49	zinc finger protein 219
Zfp428	232969	-1.29	-1.41	zinc finger protein 428
Zfp459	328274	-2.43	-1.65	zinc finger protein 459
Zfp462	242466	-1.31	-1.58	zinc finger protein 462
Zfp532	328977	-1.37	-1.66	zinc finger protein 532
Zfp809	235047	-1.12	-1.09	zinc finger protein 809
Zfp819	74400	-2.15	-2.25	zinc finger protein 819
Zic3	22773	-1.61	-1.25	zinc finger protein of the cerebellum 3
Zscan10	332221	-1.24	-1.17	zinc finger and SCAN domain containing 10
Zyg11a	230590	-1.19	-1.94	zyg-11 homolog A (C. elegans)

Concordantly upregulated genes in wt and TKO EBs between d0-4				
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change TKO**	**Gene name**
Itm2a	16431	1.37	1.77	integral membrane protein 2A
Slit2	20563	1.72	1.24	slit homolog 2 (Drosophila)
Itgb5	16419	1.05	1.07	integrin beta 5
Cpm	70574	2.51	1.76	carboxypeptidase M

Concordantly downregulated genes in wt and TKO EBs between d0-4				
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change TKO**	**Gene name**
4933437F05Rik	71275	-1.32	-1.05	RIKEN cDNA 4933437F05 gene
Emp1	13730	-1.81	-2.04	epithelial membrane protein 1
Mal2	105853	-1.15	1.62	mal, T-cell differentiation protein 2
Sema3e	20349	-1.07	1.56	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E; hypothetical protein LOC100044162
Slc25a31	73333	-1.53	1.12	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 31
Trh	22044	-1.15	1.19	thyrotropin releasing hormone
Xrcc5	22596	-1.21	-1.17	X-ray repair complementing defective repair in Chinese hamster cells 5

6.3 Differentially regulated genes during the second differentiation period (day 4 - 16)

Unique upregulated genes in wt EBs d4-16			
Gene symbol	**Entrez Gene ID**	**Fold change between d4_16**	**Gene Name**
0610040J01Rik	76261	1	RIKEN cDNA 0610040J01 gene
1110032A04Rik	66183	2.3	RIKEN cDNA 1110032A04 gene
1200009O22Rik	66873	1.04	RIKEN cDNA 1200009O22 gene

1600012P17Rik	72025	2	RIKEN cDNA 1600012P17 gene
1810010H24Rik	69066	1.52	RIKEN cDNA 1810010H24 gene
2210407C18Rik	78354	2.06	RIKEN cDNA 2210407C18 gene
2210415F13Rik	70163	3.28	RIKEN cDNA 2210415F13 gene
2310043J07Rik	69665	2.16	RIKEN cDNA 2310043J07 gene
2310067E19Rik	76455	1.39	RIKEN cDNA 2310067E19 gene
2610203C20Rik	100042464	1.45	hypothetical protein LOC100042464
2810055G20Rik	77994	1.485	RIKEN cDNA 2810055G20 gene
3830417A13Rik	70696	2	RIKEN cDNA 3830417A13 gene
4921506M07Rik	70846	1.36	RIKEN cDNA 4921506M07 gene
4930402H24Rik	228602	1.31	RIKEN cDNA 4930402H24 gene
4932442L08Rik	631145	1.15	similar to Putative uncharacterized protein CXorf58
5330426P16Rik	68190	1.2	RIKEN cDNA 5330426P16 gene
5430435G22Rik	226421	1.23	RIKEN cDNA 5430435G22 gene
A630033H20Rik	213438	1.21	RIKEN cDNA A630033H20 gene
A730069N07Rik	244425	1.14	RIKEN cDNA A730069N07 gene
AA986860	212439	1.27	expressed sequence AA986860
Aadac	67758	1.12	arylacetamide deacetylase (esterase)
Abca5	217265	1.33	ATP-binding cassette, sub-family A (ABC1), member 5
Abca8b	27404	1.44	ATP-binding cassette, sub-family A (ABC1), member 8b
Abcc3	76408	1.38	ATP-binding cassette, sub-family C (CFTR/MRP), member 3
Abcd1	11666	1.18	ATP-binding cassette, sub-family D (ALD), member 1
Abcd3	19299	1.13	ATP-binding cassette, sub-family D (ALD), member 3
Abhd2	54608	1.12	abhydrolase domain containing 2
Abi3bp	320712	1.03	ABI gene family, member 3 (NESH) binding protein
Acaa2	52538	1.24	acetyl-Coenzyme A acyltransferase 2 (mitochondrial 3-oxoacyl-Coenzyme A thiolase)
Acadl	11363	1.07	acyl-Coenzyme A dehydrogenase, long-chain
Acat1	110446	1.16	acetyl-Coenzyme A acetyltransferase 1
Ace2	70008	2.79	angiotensin I converting enzyme (peptidyl-dipeptidase A) 2
Acnat1	230161	1.38	novel protein similar to bile acid-Coenzyme A: amino acid N-acyltransferase Baat; expressed sequence AI132189
Acot2	171210	1.01	acyl-CoA thioesterase 2
Acsl5	433256	1.4	acyl-CoA synthetase long-chain family member 5
Acsm1	117147	1.04	acyl-CoA synthetase medium-chain family member 1
Acsm3	20216	1.58	acyl-CoA synthetase medium-chain family member 3
Acss1	68738	1.13	acyl-CoA synthetase short-chain family member 1
Actg2	11468	2.38	actin, gamma 2, smooth muscle, enteric
Actn2	11472	1.4	actinin alpha 2
Adam10	11487	1.21	a disintegrin and metallopeptidase domain 10
Adam12	11489	1.11	a disintegrin and metallopeptidase domain 12 (meltrin alpha)
Adamdec1	58860	1.94	ADAM-like, decysin 1
Adamts12	239337	1.3	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 12
Adamts5	23794	1.16	similar to a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 5 (aggrecanase-2); a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 5 (aggrecanase-2)
Adamtsl1	77739	1.2	ADAMTS-like 1
Adcy7	11513	1.14	adenylate cyclase 7
Aes	14797	1.85	amino-terminal enhancer of split
Aff2	14266	1.04	AF4/FMR2 family, member 2
Agt	11606	3.01	angiotensinogen (serpin peptidase inhibitor, clade A, member 8)
Ahr	11622	1.65	aryl-hydrocarbon receptor
AI607873	226691	1.05	expressed sequence AI607873
AI747448	99709	2.2	expressed sequence AI747448
Aifm2	71361	1.11	apoptosis-inducing factor, mitochondrion-associated 2
Aim2	383619	1.19	absent in melanoma 2
Ak5	229949	1.2	adenylate kinase 5
Akr1c13	27384	1.36	aldo-keto reductase family 1, member C13
Akr1c14	105387	2.95	aldo-keto reductase family 1, member C14
Akt1	11651	1.14	thymoma viral proto-oncogene 1; similar to serine/threonine protein kinase
Alcam	11658	3.64	activated leukocyte cell adhesion molecule
Aldh1a1	11668	2.8	aldehyde dehydrogenase family 1, subfamily A1

Aldh1a2	19378	1.59	aldehyde dehydrogenase family 1, subfamily A2
Aldh1a7	26358	4.23	aldehyde dehydrogenase family 1, subfamily A7
Aldh1b1	72535	1.86	aldehyde dehydrogenase 1 family, member B1
Aldh1l2	216188	1.85	aldehyde dehydrogenase 1 family, member L2
Alox12	11684	1.02	arachidonate 12-lipoxygenase
Anpep	16790	1.15	alanyl (membrane) aminopeptidase
Antxr2	71914	1.2	anthrax toxin receptor 2
Anxa10	26359	3.57	annexin A10
Anxa11	11744	1	annexin A11; predicted gene 2260; predicted gene 2274
Anxa13	69787	1.33	annexin A13
Anxa3	11745	2.29	similar to Anxa3; annexin A3
Anxa5	11747	1.76	annexin A5
Anxa6	11749	1.02	annexin A6
Anxa8	11752	1.1	annexin A8
Apcs	20219	1.05	serum amyloid P-component
Apobec1	11810	2.06	apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1
Apoc3	11814	1.34	apolipoprotein C-III
Apoh	11818	2.54	apolipoprotein H
Arap1	69710	1.08	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1
Arfgap3	66251	1.57	ADP-ribosylation factor GTPase activating protein 3
Arg1	11846	1.46	arginase, liver
Arhgap24	231532	1.12	Rho GTPase activating protein 24
Arhgap31	12549	1.46	Rho GTPase activating protein 31; Synonyms CdGAP, mKIAA1204
Arhgdib	11857	1.44	Rho, GDP dissociation inhibitor (GDI) beta
Arhgef10	234094	1.15	Rho guanine nucleotide exchange factor (GEF) 10
Arrdc4	66412	1.455	arrestin domain containing 4
Arsb	11881	1.05	arylsulfatase B
Asb2	65256	1.22	ankyrin repeat and SOCS box-containing 2
Aspa	11484	2.83	aspartoacylase
Aspn	66695	1.35	asporin
Atp2c2	69047	1.29	ATPase, Ca++ transporting, type 2C, member 2
Atp6v0a4	140494	1.15	ATPase, H+ transporting, lysosomal V0 subunit A4
Atp6v0d2	242341	1.21	ATPase, H+ transporting, lysosomal V0 subunit D2
Atp9a	11981	1.37	ATPase, class II, type 9A
AU018091	245128	1.93	expressed sequence AU018091
Axl	26362	2.4	AXL receptor tyrosine kinase
B130024G19Rik	434198	1.76	RIKEN cDNA B130024G19 gene
B3gnt9	97440	1.67	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 9
BB287469	544881	2.49	predicted gene 6804; predicted gene 2046; predicted gene 8607; expressed sequence BB287469; predicted gene 2075; predicted gene 2022; predicted gene 4027
Bcam	57278	1.42	basal cell adhesion molecule
Bcl11a	14025	1.55	B-cell CLL/lymphoma 11A (zinc finger protein)
Blvra	109778	1.5	biliverdin reductase A
C1qa	12259	2.78	complement component 1, q subcomponent, alpha polypeptide
C1qb	12260	3.13	complement component 1, q subcomponent, beta polypeptide
C1qc	12262	3.31	complement component 1, q subcomponent, C chain
C1qtnf6	72709	1.16	C1q and tumor necrosis factor related protein 6
C1rb	667277	1.62	complement component 1, r subcomponent; predicted gene 8551
C2	12263	1.1	complement component 2 (within H-2S)
C630004H02Rik	217310	1.48	hypothetical protein LOC100043986; RIKEN cDNA C630004H02 gene
Camk2n1	66259	1.15	calcium/calmodulin-dependent protein kinase II inhibitor 1
Car8	12319	1.88	carbonic anhydrase 8; similar to Carbonic anhydrase-related protein (CARP) (CA-VIII)
Casp12	12364	1.95	caspase 12; hypothetical protein LOC100044205
Cav1	12389	1.43	caveolin 1, caveolae protein
Cav2	12390	1.17	caveolin 2
Cbr2	12409	1.41	carbonyl reductase 2
Cbx4	12418	1.24	chromobox homolog 4 (Drosophila Pc class)
Ccdc80	67896	2.6	coiled-coil domain containing 80
Ccl12	20293	1.55	chemokine (C-C motif) ligand 12; similar to monocyte chemoattractant protein-5
Ccl3	20302	4.87	chemokine (C-C motif) ligand 3

Ccl6	20305	1.04	chemokine (C-C motif) ligand 6
Ccnd2	12444	1.34	cyclin D2
Ccr5	12774	1.88	chemokine (C-C motif) receptor 5
Cd14	12475	2.23	CD14 antigen
Cd34	12490	2.95	CD34 antigen
Cd36	12491	1.84	CD36 antigen
Cd38	12494	1.11	CD38 antigen
Cd47	16423	1.29	CD47 antigen (Rh-related antigen, integrin-associated signal transducer)
Cd52	23833	1.69	CD52 antigen
Cd53	12508	3.93	CD53 antigen
Cd63	12512	1.08	CD63 antigen
Cd72	12517	1.4	CD72 antigen
Cd93	17064	2	CD93 antigen
Cd97	26364	1.41	CD97 antigen
Cdc42ep2	104252	1.18	CDC42 effector protein (Rho GTPase binding) 2
Cdh1	12550	1.19	cadherin 1
Cdh17	12557	1.08	cadherin 17
Cdhr5	72040	2.56	cadherin-related family member 5; Synonyms 1810074H01Rik, Mucdhl, Mupcdh
Cdkn1a	12575	1.89	cyclin-dependent kinase inhibitor 1A (P21)
Cdkn2c	12580	1.04	cyclin-dependent kinase inhibitor 2C (p18, inhibits CDK4)
Cds1	74596	1.06	CDP-diacylglycerol synthase 1
Ceacam3	384557	1.33	carcinoembryonic antigen-related cell adhesion molecule 3; similar to Carcinoembryonic antigen-related cell adhesion molecule 3 precursor (Carcinoembryonic antigen CGM1)
Cebpa	12606	1.11	CCAAT/enhancer binding protein (C/EBP), alpha
Ces3	104158	1.1	carboxylesterase 3
Cfh	12628	2.74	complement component factor h; similar to complement component factor H
Cftr	12638	1.64	cystic fibrosis transmembrane conductance regulator homolog
Cgn	70737	1.04	cingulin; cDNA sequence BC021767
Chchd10	103172	1.27	coiled-coil-helix-coiled-coil-helix domain containing 10
Chrm2	243764	1.63	cholinergic receptor, muscarinic 2, cardiac
Ckb	12709	1.61	similar to creatine kinase, brain; predicted gene 12892; creatine kinase, brain
Ckmt1	12716	1.94	creatine kinase, mitochondrial 1, ubiquitous
Clca1	12722	1.57	chloride channel calcium activated 1
Clca2	80797	1.1	chloride channel calcium activated 2
Clca3	23844	1.09	chloride channel calcium activated 3
Cldn1	12737	1.46	claudin 1
Cldn23	71908	1.3	claudin 23
Cldn3	12739	1.53	claudin 3
Cldn8	54420	1.5	claudin 8
Clec2d	93694	1.67	C-type lectin domain family 2, member d
Clec2h	94071	1.33	C-type lectin domain family 2, member h
Clec4d	17474	1.85	C-type lectin domain family 4, member d
Clec4n	56620	2.73	C-type lectin domain family 4, member n
Clmn	94040	1.7	calmin
Clrn3	212070	1.18	clarin 3
Cltb	74325	1.72	clathrin, light polypeptide (Lcb)
Cmtm8	70031	1.3	CKLF-like MARVEL transmembrane domain containing 8
Col6a2	12834	1.88	collagen, type VI, alpha 2
Colec10	239447	1.15	collectin sub-family member 10
Copz2	56358	1.56	coatomer protein complex, subunit zeta 2
Coro2a	107684	1.06	coronin, actin binding protein 2A
Cps1	227231	2.02	carbamoyl-phosphate synthetase 1
Cpt1a	12894	3.04	carnitine palmitoyltransferase 1a, liver
Creb3	12913	1.08	cAMP responsive element binding protein 3
Crp	12944	2.85	C-reactive protein, pentraxin-related
Csf2rb	12983	1.02	colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage)
Csrp3	13009	1.55	cysteine and glycine-rich protein 3
Ctgf	14219	2.13	connective tissue growth factor
Ctsd	13033	1.16	cathepsin D
Ctse	13034	1.91	cathepsin E

Ctsr	56835	1.48	cathepsin R
Ctss	13040	4.28	cathepsin S
Cx3cr1	13051	1.49	similar to chemokine receptor CX3CR1; chemokine (C-X3-C) receptor 1
Cxcl16	66102	1.24	chemokine (C-X-C motif) ligand 16
Cyba	13057	1.01	cytochrome b-245, alpha polypeptide
Cybb	13058	2.71	cytochrome b-245, beta polypeptide
Cym	229697	6.57	similar to prochymosin; chymosin
Cyp2c40	13099	1.53	cytochrome P450, family 2, subfamily c, polypeptide 40; similar to RIKEN cDNA C730004C24 gene; cytochrome P450, family 2, subfamily c, polypeptide 69; cytochrome P450, family 2, subfamily c, polypeptide 67
Cyp2c55	72082	2.2	cytochrome P450, family 2, subfamily c, polypeptide 55
Cyp2c67	545288	1.09	cytochrome P450, family 2, subfamily c, polypeptide 40; similar to RIKEN cDNA C730004C24 gene; cytochrome P450, family 2, subfamily c, polypeptide 69; cytochrome P450, family 2, subfamily c, polypeptide 67
Cyp2c68	433247	1.24	cytochrome P450, family 2, subfamily c, polypeptide 68
Cyp2c70	226105	1.79	cytochrome P450, family 2, subfamily c, polypeptide 70
Cyp2d26	76279	1.41	cytochrome P450, family 2, subfamily d, polypeptide 26
Cyp2f2	13107	3.23	cytochrome P450, family 2, subfamily f, polypeptide 2
Cyp2s1	74134	1.39	cytochrome P450, family 2, subfamily s, polypeptide 1
Cyp4a12b	13118	1.11	cytochrome P450, family 4, subfamily a, polypeptide 12B
Cyp4v3	102294	1.01	cytochrome P450, family 4, subfamily v, polypeptide 3
Cyth4	72318	1.05	cytohesin 4
D4Bwg0951e	52829	2.99	DNA segment, Chr 4, Brigham & Women's Genetics 0951 expressed
Dapp1	26377	1.31	dual adaptor for phosphotyrosine and 3-phosphoinositides 1
Dazap2	23994	1.07	similar to DAZ associated protein 2; predicted gene 2444; DAZ associated protein 2
Dcdc2a	195208	1.07	doublecortin domain containing 2a
Ddx60	234311	2.33	DEAD (Asp-Glu-Ala-Asp) box polypeptide 60
Degs2	70059	1.2	degenerative spermatocyte homolog 2 (Drosophila), lipid desaturase
Des	13346	1.31	desmin
Dkk2	56811	2.78	dickkopf homolog 2 (Xenopus laevis)
Dmbt1	12945	1.53	deleted in malignant brain tumors 1
Dmrtc1c	71083	1.11	DMRT-like family C1c2; DMRT-like family C1c
Dnajb4	67035	1.25	DnaJ (Hsp40) homolog, subfamily B, member 4
Dnm1	13429	1.01	dynamin 1
Dock8	76088	1.23	dedicator of cytokinesis 8
Doxl2	243376	1.09	diamine oxidase-like protein 2
Dpt	56429	1.08	dermatopontin
Dsg2	13511	1.29	desmoglein 2; similar to Dsg2 protein
Dtx4	207521	1.31	deltex 4 homolog (Drosophila)
Dusp1	19252	1.03	dual specificity phosphatase 1
Dynlt3	67117	1.08	dynein light chain Tctex-type 3
Ech1	51798	1.09	enoyl coenzyme A hydratase 1, peroxisomal
Efhd2	27984	1.24	similar to EF hand domain containing 2; EF hand domain containing 2
Egfr	13649	1.93	epidermal growth factor receptor
Ehd2	259300	1.41	EH-domain containing 2
Ehf	13661	1.96	ets homologous factor
Ehhadh	74147	1.19	enoyl-Coenzyme A, hydratase/3-hydroxyacyl Coenzyme A dehydrogenase
Eid1	58521	1.1	EP300 interacting inhibitor of differentiation 1
Eif2ak3	13666	1.02	eukaryotic translation initiation factor 2 alpha kinase 3
Elf3	13710	1.34	E74-like factor 3
Elmo3	234683	1.09	engulfment and cell motility 3, ced-12 homolog (C. elegans)
Eltd1	170757	1.98	EGF, latrophilin seven transmembrane domain containing 1
Emcn	59308	2.59	endomucin
Emilin1	100952	1.8	elastin microfibril interfacer 1
Emp3	13732	1.41	epithelial membrane protein 3
Emr1	13733	2.11	EGF-like module containing, mucin-like, hormone receptor-like sequence 1
Enpp4	224794	1.03	ectonucleotide pyrophosphatase/phosphodiesterase 4
Entpd4	67464	1.03	ectonucleoside triphosphate diphosphohydrolase 4
Epha3	13837	2.99	Eph receptor A3
Epha7	13841	1.95	Eph receptor A7
Epn3	71889	1.45	epsin 3

Erbb2ip	59079	1.12	Erbb2 interacting protein
Erbb3	13867	1.6	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)
Esm1	71690	1.29	endothelial cell-specific molecule 1
Esrp2	77411	2.03	epithelial splicing regulatory protein 2
Exph5	320051	1.17	exophilin 5
F10	14058	1.24	coagulation factor X
F13a1	74145	4.54	coagulation factor XIII, A1 subunit
Fabp2	14079	1.31	fatty acid binding protein 2, intestinal
Fabp4	11770	1.51	fatty acid binding protein 4, adipocyte
Fam101b	76566	1.34	family with sequence similarity 101, member B
Fam174b	100038347	1.57	family with sequence similarity 174, member B
Fam20c	80752	1.1	family with sequence similarity 20, member C
Fam70a	245386	1.98	family with sequence similarity 70, member A
Fam84a	105005	1.24	family with sequence similarity 84, member A
Fas	14102	1.37	Fas (TNF receptor superfamily member 6)
Fat4	329628	1.89	FAT tumor suppressor homolog 4 (Drosophila)
Fbp1	14121	1.94	fructose bisphosphatase 1
Fbxo32	67731	1.56	F-box protein 32
Fcer1g	14127	2.59	Fc receptor, IgE, high affinity I, gamma polypeptide
Fcgr2b	14130	1	Fc receptor, IgG, low affinity IIb
Fdx1	14148	1.05	ferredoxin 1
Fermt1	241639	1.32	fermitin family homolog 1 (Drosophila)
Fetub	59083	1.44	fetuin beta
Fgd3	30938	1.38	FYVE, RhoGEF and PH domain containing 3
Fgfr3	14184	1.75	fibroblast growth factor receptor 3
Fgl1	234199	1.46	fibrinogen-like protein 1
Figf	14205	1.52	c-fos induced growth factor
Fkbp7	14231	1.43	FK506 binding protein 7
Fosl2	14284	1.14	similar to fos-like antigen 2; fos-like antigen 2
Foxa1	15375	1.62	forkhead box A1; similar to Hepatocyte nuclear factor 3-alpha (HNF-3A) (Forkhead box protein A1)
Foxo1	56458	1.07	forkhead box O1
Foxq1	15220	1.37	forkhead box Q1
Fras1	231470	1.72	Fraser syndrome 1 homolog (human)
Frem1	329872	1.78	Fras1 related extracellular matrix protein 1
Frem2	242022	1.1	Fras1 related extracellular matrix protein 2
Frk	14302	2.03	fyn-related kinase
Fry	320365	1.19	furry homolog (Drosophila)
Fstl1	14314	1.22	follistatin-like 1
Fxyd5	18301	2.17	FXYD domain-containing ion transport regulator 5
Fzd1	14362	1.45	frizzled homolog 1 (Drosophila)
G0s2	14373	1.26	G0/G1 switch gene 2
Gab2	14389	1.02	growth factor receptor bound protein 2-associated protein 2
Gabrp	216643	2.44	gamma-aminobutyric acid (GABA) A receptor, pi
Gadd45b	17873	2.07	growth arrest and DNA-damage-inducible 45 beta
Galnt4	14426	1.12	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 4
Gap43	14432	1.09	growth associated protein 43
Gatm	67092	1.12	glycine amidinotransferase (L-arginine:glycine amidinotransferase)
Gca	227960	1.1	grancalcin
Gcnt3	72077	1.17	glucosaminyl (N-acetyl) transferase 3, mucin type
Gda	14544	1.14	guanine deaminase
Gfra1	14585	1.27	glial cell line derived neurotrophic factor family receptor alpha 1
Ghr	14600	2.01	growth hormone receptor
Gjb1	14618	1.44	gap junction protein, beta 1
Glcci1	170772	1.08	similar to glucocorticoid induced transcript 1; predicted gene 5815; glucocorticoid induced transcript 1
Glrx	93692	1.45	glutaredoxin
Gm10639	100042314	2.99	predicted gene 10639
Gm4340	100043292	2.42	predicted gene 4340
Gm52	214292	1.72	predicted gene 52
Gm648	270599	1.56	predicted gene 648

Gm885	380732	2.25	predicted gene 885
Gp49a	14727	3.65	glycoprotein 49 A; leukocyte immunoglobulin-like receptor, subfamily B, member 4
Gpbar1	227289	1.21	G protein-coupled bile acid receptor 1
Gpm6b	14758	1.2	glycoprotein m6b
Gpr97	54672	1.16	G protein-coupled receptor 97
Gpx2	14776	1.61	glutathione peroxidase 2
Gria3	53623	1.65	glutamate receptor, ionotropic, AMPA3 (alpha 3)
Grtp1	66790	1.04	GH regulated TBC protein 1
Gsdmc2	331063	2.45	gasdermin C2
Gsdmc3	270328	2.61	gasdermin C3
Gsta2	14858	3.04	glutathione S-transferase, alpha 2 (Yc2)
Gsta4	14860	1.33	glutathione S-transferase, alpha 4
Gstm2	14863	1.36	predicted gene 6665; glutathione S-transferase, mu 2
Gsto1	14873	1.47	glutathione S-transferase omega 1
Gucy1a3	60596	1.74	guanylate cyclase 1, soluble, alpha 3
Gucy1b3	54195	1.57	guanylate cyclase 1, soluble, beta 3
H2-D1	14964	1.15	histocompatibility 2, D region; histocompatibility 2, D region locus 1
Hc	15139	1	hemolytic complement
Hcfc2	67933	1.37	host cell factor C2
Hcls1	15163	1.26	hematopoietic cell specific Lyn substrate 1
Heg1	77446	1.12	HEG homolog 1 (zebrafish)
Heph	15203	1.06	hephaestin
Herpud1	64209	1.36	homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1
Hgd	15233	1.04	homogentisate 1, 2-dioxygenase
Hic1	15248	1.11	hypermethylated in cancer 1
Hif3a	53417	1.27	hypoxia inducible factor 3, alpha subunit
Hipk3	15259	1.23	homeodomain interacting protein kinase 3; similar to homeodomain interacting protein kinase 3
Hist1h1c	50708	1.31	histone cluster 1, H1c
Hist1h2bc	68024	2.06	histone cluster 1, H2bg; histone cluster 1, H2be; histone cluster 2, H2bb; histone cluster 1, H2bc
Hk1	15275	1.18	hexokinase 1
Hmgcs2	15360	3.55	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2
Hpgds	54486	1.48	prostaglandin D2 synthase 2, hematopoietic
Hpse	15442	1.7	heparanase
Hs6st1	50785	1.18	heparan sulfate 6-O-sulfotransferase 1
Hsd17b13	243168	2.79	hydroxysteroid (17-beta) dehydrogenase 13
Hsd3b2	15493	1.37	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 2
Hspa1a	193740	1.115	heat shock protein 1B; heat shock protein 1A; heat shock protein 1-like
Hspb1	15507	1.435	heat shock protein 1
Hspb8	80888	1.08	heat shock protein 8
Htra1	56213	1.57	HtrA serine peptidase 1
Idh2	269951	1.12	isocitrate dehydrogenase 2 (NADP+), mitochondrial
Ifi202b	26388	2.16	interferon activated gene 202B
Ifi204	15951	1.76	interferon activated gene 204
Ifi27l2a	76933	1.05	interferon, alpha-inducible protein 27 like 2A
Ifi47	15953	1.12	interferon gamma inducible protein 47
Ifih1	71586	1.46	interferon induced with helicase C domain 1
Ifit2	15958	1.31	interferon-induced protein with tetratricopeptide repeats 2
Ifitm3	66141	1.85	interferon induced transmembrane protein 3
Ifngr1	15979	1.4	interferon gamma receptor 1
Ifngr2	15980	1.08	interferon gamma receptor 2
Igf1	16000	2.42	insulin-like growth factor 1
Igfbp2	16008	1.29	insulin-like growth factor binding protein 2
Igsf5	72058	1.16	Purkinje cell protein 4; immunoglobulin superfamily, member 5
Iigp1	60440	1.38	interferon inducible GTPase 1; interferon-inducible GTPase-like
Il13ra1	16164	2.15	interleukin 13 receptor, alpha 1
Il2rg	16186	1.6	predicted gene 614; interleukin 2 receptor, gamma chain
Il33	77125	1.37	interleukin 33
Il4ra	16190	1.4	interleukin 4 receptor, alpha

Inhba	16323	1.43	inhibin beta-A
Irf2bp2	270110	1.37	interferon regulatory factor 2 binding protein 2
Irf8	15900	1.42	interferon regulatory factor 8
Isg20	57444	1.27	interferon-stimulated protein
Itga1	109700	1.43	integrin alpha 1
Itga8	241226	1.18	integrin alpha 8
Itgam	16409	1.53	integrin alpha M
Itgb2	16414	1.13	integrin beta 2
Itgb4	192897	1.24	integrin beta 4
Itgb6	16420	2.03	integrin beta 6
Itgb8	320910	1.07	integrin beta 8
Itih3	16426	1.72	inter-alpha trypsin inhibitor, heavy chain 3
Itih4	16427	2.05	inter alpha-trypsin inhibitor, heavy chain 4
Itm2a	16431	2.55	integral membrane protein 2A
Itpkb	320404	1.2	inositol 1,4,5-trisphosphate 3-kinase B
Jag1	16449	1	jagged 1
Jund	16478	1.16	Jun proto-oncogene related gene d
Jup	16480	1.27	junction plakoglobin
Kcne3	57442	1.46	potassium voltage-gated channel, Isk-related subfamily, gene 3; hypothetical protein LOC100044693
Kcne4	57814	1.59	potassium voltage-gated channel, Isk-related subfamily, gene 4
Kcnj2	16518	2.08	potassium inwardly-rectifying channel, subfamily J, member 2
Klf3	16599	1.15	Kruppel-like factor 3 (basic); similar to BKLF
Klf5	12224	2.07	Kruppel-like factor 5
Klhl13	67455	1.7	kelch-like 13 (Drosophila)
Kng2	385643	2.14	kininogen 2
Krt15	16665	4.19	keratin 15
Krt4	16682	3.5	keratin 4
Krt5	110308	2.86	keratin 5
Kynu	70789	1.94	kynureninase (L-kynurenine hydrolase)
Lama2	16773	1.1	laminin, alpha 2
Lama5	16776	1.38	laminin, alpha 5
Lamb3	16780	1.12	laminin, beta 3
Laptm5	16792	2.45	lysosomal-associated protein transmembrane 5
Lass3	545975	2.06	LAG1 homolog, ceramide synthase 3
Lbh	77889	1.88	limb-bud and heart
Ldb2	16826	1.17	LIM domain binding 2
Lepr	16847	1.17	leptin receptor
Lgals4	16855	3.9	lectin, galactose binding, soluble 6; hypothetical protein LOC100044254; lectin, galactose binding, soluble 4
Lgals8	56048	1.36	lectin, galactose binding, soluble 8
Lgi2	246316	1.46	leucine-rich repeat LGI family, member 2
Lgr4	107515	1.06	leucine-rich repeat-containing G protein-coupled receptor 4
Lilrb4	14728	2.86	glycoprotein 49 A; leukocyte immunoglobulin-like receptor, subfamily B, member 4
Limk2	16886	1.06	LIM motif-containing protein kinase 2
Lims2	225341	1.47	LIM and senescent cell antigen like domains 2
Lmna	16905	1.11	lamin A
Lmo7	380928	1.35	LIM domain only 7
Lonrf3	74365	1.7	LON peptidase N-terminal domain and ring finger 3
Loxl1	16949	1.52	lysyl oxidase-like 1
Lphn3	319387	1.13	latrophilin 3
Lpin2	64898	1.12	lipin 2
Lrig1	16206	1.36	leucine-rich repeats and immunoglobulin-like domains 1
Lrrc17	74511	2.23	leucine rich repeat containing 17
Lrrc32	434215	1.03	leucine rich repeat containing 32
Lsp1	16985	1.28	lymphocyte specific 1
Ltbp4	108075	1.29	latent transforming growth factor beta binding protein 4
Ly86	17084	1.6	lymphocyte antigen 86
Ly96	17087	1.43	lymphocyte antigen 96
Lyn	17096	1.55	Yamaguchi sarcoma viral (v-yes-1) oncogene homolog
Lypd6b	71897	1.66	LY6/PLAUR domain containing 6B

Lyve1	114332	1.39	lymphatic vessel endothelial hyaluronan receptor 1
Macc1	238455	1.71	metastasis associated in colon cancer 1
Mageb16	71967	1.3	predicted gene 15072; RIKEN cDNA 2410003J06 gene
Mal	17153	2.37	myelin and lymphocyte protein, T-cell differentiation protein
Mal2	105853	3.45	mal, T-cell differentiation protein 2
Mamdc4	381352	1.49	MAM domain containing 4
Mapk13	26415	1.21	mitogen-activated protein kinase 13
Masp1	17174	1.35	mannan-binding lectin serine peptidase 1
Mat1a	11720	1.76	methionine adenosyltransferase I, alpha
Mbnl1	56758	2.5	muscleblind-like 1 (Drosophila)
Mbnl3	171170	2.34	muscleblind-like 3 (Drosophila)
Mcam	84004	1.38	melanoma cell adhesion molecule
Mecom	14013	2.76	ecotropic viral integration site 1
Mef2c	17260	1.69	myocyte enhancer factor 2C
Meg3	17263	1.26	maternally expressed 3
Meis1	17268	2.9	Meis homeobox 1
Meis2	17536	2.12	Meis homeobox 2
Met	17295	1.42	met proto-oncogene
Mfap2	17150	1.58	microfibrillar-associated protein 2
Mfap4	76293	1.58	microfibrillar-associated protein 4
Mfi2	30060	1.02	antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5
Mfsd6	98682	1.2	major facilitator superfamily domain containing 6
Mgll	23945	1.89	monoglyceride lipase
Mmaa	109136	1.05	methylmalonic aciduria (cobalamin deficiency) type A
Mmp12	17381	3.41	matrix metallopeptidase 12
Mmp14	17387	1.1	matrix metallopeptidase 14 (membrane-inserted)
Mmp16	17389	1.25	matrix metallopeptidase 16
Mmp19	58223	1.07	matrix metallopeptidase 19
Mmp2	17390	1.11	matrix metallopeptidase 2
Mpeg1	17476	3.68	macrophage expressed gene 1
Mpp1	17524	1.14	membrane protein, palmitoylated
Mr1	15064	2.17	major histocompatibility complex, class I-related
Mrc1	17533	1.83	mannose receptor, C type 1
Ms4a4a	666907	2.07	membrane-spanning 4-domains, subfamily A, member 4A
Ms4a6b	69774	1.54	membrane-spanning 4-domains, subfamily A, member 6B
Ms4a6c	73656	1.71	membrane-spanning 4-domains, subfamily A, member 6C
Ms4a6d	68774	2.04	membrane-spanning 4-domains, subfamily A, member 6D
Msln	56047	2.51	mesothelin
Msr1	20288	2.1	macrophage scavenger receptor 1
Mt1	17748	2.27	metallothionein 1
Mt2	17750	2.19	metallothionein 2
Muc4	140474	1.02	mucin 4
Myh14	71960	1.03	myosin, heavy polypeptide 14
Myl2	17906	1.64	myosin, light polypeptide 2, regulatory, cardiac, slow
Myl9	98932	2.26	myosin, light polypeptide 9, regulatory
Myo15b	217328	1.63	myosin XVB
Myo16	244281	1.96	myosin XVI
Myo1c	17913	1.07	similar to nuclear myosin I beta; myosin IC
Myo1f	17916	1.07	myosin IF
Naaladl2	635702	1.43	N-acetylated alpha-linked acidic dipeptidase-like 2
Naga	17939	1.13	N-acetyl galactosaminidase, alpha
Naglu	27419	1.07	alpha-N-acetylglucosaminidase (Sanfilippo disease IIIB)
Ncam1	17967	1.52	neural cell adhesion molecule 1
Ncf1	17969	1.19	neutrophil cytosolic factor 1
Nckap1l	105855	2.63	NCK associated protein 1 like
Ndrg2	29811	1.58	N-myc downstream regulated gene 2
Nell1	338352	1.41	NEL-like 1 (chicken)
Neurl3	214854	1.13	neuralized homolog 3 homolog (Drosophila)
Nexn	68810	1.32	nexilin

Nfia	18027	1.875	nuclear factor I/A
Nfix	18032	1.89	nuclear factor I/X
Nfkbiz	80859	1.32	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta
Ngly1	59007	1.01	N-glycanase 1
Nipal2	223473	1.14	NIPA-like domain containing 2
Nnat	18111	1.1	neuronatin
Nos2	18126	1.14	nitric oxide synthase 2, inducible
Nov	18133	1.02	nephroblastoma overexpressed gene
Npnt	114249	1.45	nephronectin
Nppa	230899	1.07	natriuretic peptide precursor type A
Npy1r	18166	1.02	neuropeptide Y receptor Y1
Nr1h4	20186	1.05	nuclear receptor subfamily 1, group H, member 4
Nr2f1	13865	1.45	nuclear receptor subfamily 2, group F, member 1
Nr3c1	14815	1.7	nuclear receptor subfamily 3, group C, member 1
Nupr1	56312	1.55	nuclear protein 1
Oas1g	23960	1.81	2'-5' oligoadenylate synthetase 1G
Oasl2	23962	1.47	2'-5' oligoadenylate synthetase-like 2
Ociad2	433904	1.07	OCIA domain containing 2
Ogn	18295	3.22	osteoglycin
Olfm4	380924	1.85	olfactomedin 4
Olr1	108078	1.24	oxidized low density lipoprotein (lectin-like) receptor 1
Onecut1	15379	1.915	one cut domain, family member 1
Os9	216440	1.3	amplified in osteosarcoma
Osr1	23967	1.08	odd-skipped related 1 (Drosophila)
Ostf1	20409	1.12	osteoclast stimulating factor 1
Pah	18478	1.34	phenylalanine hydroxylase
Pappa	18491	1.36	pregnancy-associated plasma protein A
Pappa2	23850	1.793333333	pappalysin 2
Paqr9	75552	1.22	progestin and adipoQ receptor family member IX
Parp14	547253	1.47	poly (ADP-ribose) polymerase family, member 14
Parva	57342	1.88	parvin, alpha
Pcdh11x	245578	1.52	protocadherin 11 X-linked
Pcdh18	73173	1.28	protocadherin 18
Pcdh9	211712	1.56	protocadherin 9
Pcdhb11	93882	1.6	protocadherin beta 11
Pcdhb16	93887	2	protocadherin beta 16
Pcdhb17	93888	1.42	protocadherin beta 17
Pcdhb18	93889	1.16	protocadherin beta 18
Pcdhb19	93890	1.41	protocadherin beta 19
Pcdhb20	93891	1.06	protocadherin beta 20
Pcolce	18542	3.1	procollagen C-endopeptidase enhancer protein
Pde1a	18573	1.76	phosphodiesterase 1A, calmodulin-dependent
Pdgfc	54635	1.88	platelet-derived growth factor, C polypeptide
Pdk2	18604	1.11	pyruvate dehydrogenase kinase, isoenzyme 2
Pdzk1ip1	67182	1.32	PDZK1 interacting protein 1
Pgf	18654	1.95	placental growth factor
Phactr2	215789	1.12	phosphatase and actin regulator 2
Phyhd1	227696	1.02	phytanoyl-CoA dioxygenase domain containing 1
Pik3ap1	83490	1.33	phosphoinositide-3-kinase adaptor protein 1
Pik3r1	18708	1.38	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1 (p85 alpha)
Pkhd1	241035	1.27	polycystic kidney and hepatic disease 1
Pla2g7	27226	1.02	phospholipase A2, group VII (platelet-activating factor acetylhydrolase, plasma)
Plaur	18793	1.89	plasminogen activator, urokinase receptor
Plbd1	66857	1.46	phospholipase B domain containing 1; similar to RIKEN cDNA 1100001H23 gene
Plcd1	18799	1.16	phospholipase C, delta 1
Plk2	20620	1.3	polo-like kinase 2 (Drosophila)
Pllp	67801	1.02	plasma membrane proteolipid
Plod1	18822	1.08	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1
Pltp	18830	1.53	phospholipid transfer protein

Plxnb2	140570	1.33	plexin B2
Plxnc1	54712	1.16	plexin C1; similar to plexin C1
Pmepa1	65112	1.63	prostate transmembrane protein, androgen induced 1; similar to Nedd4 WW binding protein 4
Pof1b	69693	1.79	premature ovarian failure 1B
Pold4	69745	1.12	polymerase (DNA-directed), delta 4
Pon2	330260	1.09	paraoxonase 2
Ppap2a	19012	1.47	phosphatidic acid phosphatase type 2A
Ppargc1a	19017	1.78	peroxisome proliferative activated receptor, gamma, coactivator 1 alpha
Ppl	19041	1.1	periplakin
Ppm1k	243382	1.27	protein phosphatase 1K (PP2C domain containing)
Ppp2r2b	72930	1.87	protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), beta isoform
Ppp2r3a	235542	1.21	protein phosphatase 2 (formerly 2A), regulatory subunit B", alpha; RIKEN cDNA 3222402P14 gene
Pramef12	77632	2.48	PRAME family member 12
Prdm16	70673	1.09	PR domain containing 16
Prelp	116847	1.19	proline arginine-rich end leucine-rich repeat
Prkcdbp	109042	1.52	protein kinase C, delta binding protein
Prkch	18755	1.7	protein kinase C, eta
Prkg1	19091	1.49	protein kinase, cGMP-dependent, type I
Prnp	19122	1.81	prion protein
Prrx1	18933	1.02	paired related homeobox 1
Prss23	76453	1.51	protease, serine, 23
Psca	72373	2.23	prostate stem cell antigen
Psd3	234353	1.17	pleckstrin and Sec7 domain containing 3
Psg27	545925	1.08	pregnancy-specific glycoprotein 27
Psg28	114871	1.17	pregnancy-specific glycoprotein 28
Ptger3	19218	1.22	prostaglandin E receptor 3 (subtype EP3)
Ptgs1	19224	1.62	prostaglandin-endoperoxide synthase 1
Ptk2b	19229	1.12	PTK2 protein tyrosine kinase 2 beta
Ptp4a3	19245	1.11	protein tyrosine phosphatase 4a3
Ptprc	19264	1.51	protein tyrosine phosphatase, receptor type, C
Ptprj	19271	1.055	protein tyrosine phosphatase, receptor type, J; predicted gene 13768; predicted gene 13767
Ptrf	19285	1.65	polymerase I and transcript release factor
Pvrl4	71740	1.1	poliovirus receptor-related 4
Pygb	110078	1.23	brain glycogen phosphorylase
Pygl	110095	1.24	liver glycogen phosphorylase
Pzp	11287	3.28	pregnancy zone protein
Rab25	53868	1.49	RAB25, member RAS oncogene family
Rab27b	80718	1.97	RAB27b, member RAS oncogene family
Rab3d	19340	1.02	RAB3D, member RAS oncogene family
Rarb	218772	2.76	retinoic acid receptor, beta
Rarres2	71660	1.51	retinoic acid receptor responder (tazarotene induced) 2
Rasd1	19416	1.01	RAS, dexamethasone-induced 1
Rasgef1b	320292	1.38	RasGEF domain family, member 1B; hypothetical protein LOC100044232
Rasl11b	68939	1.51	RAS-like, family 11, member B
Rassf8	71323	1.02	Ras association (RalGDS/AF-6) domain family (N-terminal) member 8
Rassf9	237504	2.52	Ras association (RalGDS/AF-6) domain family (N-terminal) member 9
Rbp1	19659	1.61	retinol binding protein 1, cellular
Rbp2	19660	1.68	retinol binding protein 2, cellular
Rcn3	52377	1.15	reticulocalbin 3, EF-hand calcium binding domain
Rerg	232441	1.42	RAS-like, estrogen-regulated, growth-inhibitor
Rgn	19733	1.18	regucalcin
Rgs2	19735	1.41	regulator of G-protein signaling 2
Rhox12	382282	1.15	reproductive homeobox 12
Rnf125	67664	1.27	ring finger protein 125
Rnf213	672511	1.445	ring finger protein 213
Rtl1	353326	1.23	retrotransposon-like 1; RIKEN cDNA 6430411K18 gene
S100a11	20195	1.34	predicted gene 7665; S100 calcium binding protein A11 (calgizzarin); predicted gene 5068
S100a14	66166	1.46	S100 calcium binding protein A14
S100a16	67860	1.46	S100 calcium binding protein A16

S100a6	20200	2.64	S100 calcium binding protein A6 (calcyclin)
S1pr1	13609	1.04	sphingosine-1-phosphate receptor 1
Samd5	320825	1.1	sterile alpha motif domain containing 5
Sardh	192166	1.07	sarcosine dehydrogenase
Sat1	20229	1.09	similar to spermidine/spermine N1-acetyltransferase; predicted gene 5552; spermidine/spermine N1-acetyl transferase 1
Scarb2	12492	1.28	scavenger receptor class B, member 2
Scel	64929	1.45	sciellin
Scrn1	69938	1.13	secernin 1
Selenbp2	20342	1.18	selenium binding protein 2
Sema3a	20346	1.99	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A
Sema3c	20348	4.86	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3C
Sema3d	108151	1.12	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3D; hypothetical protein LOC100044160
Sema4d	20354	1.19	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4D
Sept4	18952	1.35	septin 4
Serinc2	230779	1.52	serine incorporator 2; hypothetical protein LOC100044221
Serpina10	217847	1.15	serine (or cysteine) peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 10
Serpina1c	20702	3.95	serine (or cysteine) peptidase inhibitor, clade A, member 1C
Serpinb1a	66222	1.56	serine (or cysteine) peptidase inhibitor, clade B, member 1a
Serpinb6b	20708	2.47	serine (or cysteine) peptidase inhibitor, clade B, member 6b
Serpinb8	20725	1.04	serine (or cysteine) peptidase inhibitor, clade B, member 8
Serpinb9b	20706	2.32	serine (or cysteine) peptidase inhibitor, clade B, member 9b
Serpinc1	11905	1.18	serine (or cysteine) peptidase inhibitor, clade C (antithrombin), member 1
Serpine2	20720	1.43	serine (or cysteine) peptidase inhibitor, clade E, member 2
Serpinf1	20317	2.14	serine (or cysteine) peptidase inhibitor, clade F, member 1
Sfrp1	20377	2.74	secreted frizzled-related protein 1
Sfrp2	20319	1.56	secreted frizzled-related protein 2
Sgce	20392	1.13	sarcoglycan, epsilon
Shisa3	330096	1.12	shisa homolog 3 (Xenopus laevis)
Shisa5	66940	1.31	shisa homolog 5 (Xenopus laevis)
Slc14a1	108052	1.24	solute carrier family 14 (urea transporter), member 1
Slc15a3	65221	1.94	solute carrier family 15, member 3
Slc16a7	20503	1.53	solute carrier family 16 (monocarboxylic acid transporters), member 7
Slc17a5	235504	1.03	solute carrier family 17 (anion/sugar transporter), member 5
Slc25a12	78830	1.37	solute carrier family 25 (mitochondrial carrier, Aralar), member 12
Slc26a7	208890	1	solute carrier family 26, member 7
Slc2a2	20526	1.75	solute carrier family 2 (facilitated glucose transporter), member 2
Slc35f5	74150	1.36	solute carrier family 35, member F5
Slc36a4	234967	1.18	solute carrier family 36 (proton/amino acid symporter), member 4
Slc37a2	56857	1.37	solute carrier family 37 (glycerol-3-phosphate transporter), member 2
Slc6a14	56774	1.15	solute carrier family 6 (neurotransmitter transporter), member 14
Slc6a2	20538	1.42	solute carrier family 6 (neurotransmitter transporter, noradrenalin), member 2
Slc7a11	26570	1.42	solute carrier family 7 (cationic amino acid transporter, y+ system), member 11
Slc7a2	11988	1.05	solute carrier family 7 (cationic amino acid transporter, y+ system), member 2
Slc9a9	331004	1.27	solute carrier family 9 (sodium/hydrogen exchanger), member 9
Slco4a1	108115	1.65	solute carrier organic anion transporter family, member 4a1
Slfn2	20556	1.95	schlafen 2
Slit2	20563	1.41	slit homolog 2 (Drosophila)
Slitrk6	239250	2.18	SLIT and NTRK-like family, member 6
Smpd3	58994	1.18	sphingomyelin phosphodiesterase 3, neutral
Smyd1	12180	1.12	SET and MYND domain containing 1
Sncaip	67847	1.03	synuclein, alpha interacting protein (synphilin)
Sostdc1	66042	1.15	sclerostin domain containing 1
Sox2	20674	1.48	SRY-box containing gene 2
Sox5	20678	1.13	SRY-box containing gene 5
Spon1	233744	1.41	spondin 1, (f-spondin) extracellular matrix protein
Sprr3	20766	1.29	small proline-rich protein 3
Sqrdl	59010	1.63	sulfide quinone reductase-like (yeast)
Sri	109552	1.01	sorcin

Srpx	51795	1	sushi-repeat-containing protein; retinitis pigmentosa GTPase regulator
Srpx2	68792	2.6	sushi-repeat-containing protein, X-linked 2
Srr	27364	1.23	serine racemase
Ssbp2	66970	1.3	single-stranded DNA binding protein 2; predicted gene 12470
St3gal4	20443	1.2	ST3 beta-galactoside alpha-2,3-sialyltransferase 4
St5	76954	1.06	suppression of tumorigenicity 5
St8sia4	20452	1.3	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4
Stab1	192187	1.34	stabilin 1
Stambp	70527	1.14	STAM binding protein
Stat1	20846	1.66	signal transducer and activator of transcription 1
Steap2	74051	1.07	six transmembrane epithelial antigen of prostate 2
Stom	13830	1.82	stomatin
Styk1	243659	1.04	serine/threonine/tyrosine kinase 1
Sult1d1	53315	1.87	sulfotransferase family 1D, member 1
Syne1	64009	1.15	synaptic nuclear envelope 1
Syngr1	20972	1.61	synaptogyrin 1
Sytl2	83671	2.33	synaptotagmin-like 2
Tac2	21334	1.03	tachykinin 2
Tagln	21345	1.54	transgelin
Tbc1d8b	245638	1.87	TBC1 domain family, member 8B
Tcp11l2	216198	1.43	t-complex 11 (mouse) like 2
Tecrl	243078	1.34	steroid 5 alpha-reductase 2-like 2
Tet2	214133	1.41	tet oncogene family member 2
Tff1	21784	1.45	trefoil factor 1; predicted gene 3090
Tgfb3	21809	2.2	transforming growth factor, beta 3
Tgm1	21816	1.17	transglutaminase 1, K polypeptide
Thbs1	21825	3.07	thrombospondin 1; similar to thrombospondin 1
Tifa	211550	1.43	TRAF-interacting protein with forkhead-associated domain; similar to Traf2 binding protein
Timp1	21857	1.27	tissue inhibitor of metalloproteinase 1
Timp2	21858	1.6	tissue inhibitor of metalloproteinase 2
Tll1	21892	1.1	tolloid-like
Tlr13	279572	1.12	toll-like receptor 13
Tlr7	170743	1.31	toll-like receptor 7
Tm4sf1	17112	1.34	transmembrane 4 superfamily member 1
Tm4sf4	229302	2.47	transmembrane 4 superfamily member 4
Tmem100	67888	1.45	transmembrane protein 100
Tmem125	230678	1.06	transmembrane protein 125
Tmem173	72512	1.37	transmembrane protein 173
Tmem184a	231832	1.01	transmembrane protein 184a
Tmem195	319660	1.23	transmembrane protein 195
Tmem26	327766	1.81	transmembrane protein 26
Tmem27	57394	2.68	transmembrane protein 27
Tmem45a	56277	3.45	transmembrane protein 45a
Tmem63a	208795	1.41	transmembrane protein 63a
Tmem86a	67893	1.37	transmembrane protein 86A
Tmem87b	72477	1.08	transmembrane protein 87B
Tmem98	103743	1.18	transmembrane protein 98
Tmprss11b	319875	1.36	transmembrane protease, serine 11b N terminal like
Tmsb4x	19241	2.39	thymosin, beta 4, X chromosome; similar to thymosin beta-4
Tmtc2	278279	1.39	transmembrane and tetratricopeptide repeat containing 2
Tmx4	52837	1.02	thioredoxin-related transmembrane protein 4
Tnfrsf9	21942	2.23	tumor necrosis factor receptor superfamily, member 9
Tns1	21961	1.535	tensin 1
Tns3	319939	1.27	tensin 3
Tns4	217169	1.76	tensin 4
Tob1	22057	1.6	transducer of ErbB-2.1
Tpbpb	116913	1.93	trophoblast specific protein beta
Trim29	72169	1.52	tripartite motif-containing 29
Trim30	20128	1.84	tripartite motif-containing 30

Trim55	381485	1.03	tripartite motif-containing 55
Trmt2b	215201	1.15	TRM2 tRNA methyltransferase 2 homolog B (S. cerevisiae)
Trpm4	68667	1.03	transient receptor potential cation channel, subfamily M, member 4
Trps1	83925	1.1	trichorhinophalangeal syndrome I (human); similar to Trps1 protein
Tshz1	110796	1.37	teashirt zinc finger family member 1
Tshz2	228911	1.85	teashirt zinc finger family member 2
Tspan1	66805	1.4	tetraspanin 1
Tspan15	70423	1.32	tetraspanin 15
Tyrobp	22177	2.29	TYRO protein tyrosine kinase binding protein
U46068	228801	4.68	cDNA sequence U46068
Ugt1a9	394434	1.61	similar to UDP glycosyltransferase 1 family polypeptide A13; similar to UGT1.6; UDP glucuronosyltransferase 1 family, polypeptide A1; UDP glucuronosyltransferase 1 family, polypeptide A2; UDP glycosyltransferase 1 family, polypeptide A10; UDP glucuronosyltransferase 1 family, polypeptide A5; UDP glycosyltransferase 1 family, polypeptide A cluster; UDP glucuronosyltransferase 1 family, polypeptide A9; UDP glucuronosyltransferase 1 family, polypeptide A8; similar to UDP glycosyltransferase 1 family, polypeptide A8; UDP glucuronosyltransferase 1 family, polypeptide A7C; UDP glucuronosyltransferase 1 family, polypeptide A6A; similar to UDP glucuronosyltransferase 1 family, polypeptide A6B; UDP glucuronosyltransferase 1 family, polypeptide A6B
Ugt2b35	243085	1.8	UDP glucuronosyltransferase 2 family, polypeptide B35
Ugt2b36	231396	1.19	UDP glucuronosyltransferase 2 family, polypeptide B36
Ugt8a	22239	1.82	UDP galactosyltransferase 8A
Unc93b1	54445	1.41	unc-93 homolog B1 (C. elegans)
Upk1a	109637	1.22	uroplakin 1A
Usp18	24110	1.98	ubiquitin specific peptidase 18; similar to ubiquitin specific protease UBP43
Usp46	69727	1.16	ubiquitin specific peptidase 46
Vcan	13003	1.21	versican
Vgll3	73569	1.01	vestigial like 3 (Drosophila)
Vsig1	78789	2.08	V-set and immunoglobulin domain containing 1
Vtcn1	242122	1.44	V-set domain containing T cell activation inhibitor 1
Wfdc1	67866	1.62	WAP four-disulfide core domain 1
Wfdc2	67701	2.42	WAP four-disulfide core domain 2
Wisp1	22402	1.67	WNT1 inducible signaling pathway protein 1
Wsb2	59043	1.555	WD repeat and SOCS box-containing 2
Wwc1	211652	1.17	WW, C2 and coiled-coil domain containing 1
Xaf1	327959	1.1	XIAP associated factor 1
Xbp1	22433	1.07	X-box binding protein 1
Ypel5	383295	1.01	yippee-like 5 (Drosophila)
Zadh2	225791	1.66	zinc binding alcohol dehydrogenase, domain containing 2
Zbtb20	56490	1.305	zinc finger and BTB domain containing 20
Zc3h12c	244871	1	zinc finger CCCH type containing 12C
Zfhx4	80892	1.1	zinc finger homeodomain 4
Zfp827	622675	1.37	zinc finger protein 827
Zscan4c	245109	2.226666667	zinc finger and SCAN domain containing 4C; pseudogene 3; similar to Gene model 397, (NCBI)

Unique downpregulated genes in wt EBs d4-16			
Gene symbol	**Entrez Gene ID**	**Fold change between d4_16**	**Gene Name**
1810032O08Rik	66293	-1.11	RIKEN cDNA 1810032O08 gene
2410127L17Rik	67383	-1.09	RIKEN cDNA 2410127L17 gene; predicted gene 7622
2610002D18Rik	69885	-1.29	RIKEN cDNA 2610002D18 gene
2610024G14Rik	56412	-1.05	RIKEN cDNA 2610024G14 gene
2610039C10Rik	66578	-1.11	RIKEN cDNA 2610039C10 gene
2610318N02Rik	70458	-1.485	RIKEN cDNA 2610318N02 gene; predicted gene 6960
2610528E23Rik	66497	-1.18	RIKEN cDNA 2610528E23 gene
4732471D19Rik	319719	-1.05	RIKEN cDNA 4732471D19 gene
4930422G04Rik	71643	-1.04	RIKEN cDNA 4930422G04 gene
4930503L19Rik	269033	-1.14	RIKEN cDNA 4930503L19 gene
Aacs	78894	-1.03	acetoacetyl-CoA synthetase
Abcd4	19300	-1.16	ATP-binding cassette, sub-family D (ALD), member 4
Abtb2	99382	-1.17	ankyrin repeat and BTB (POZ) domain containing 2

Acaca	107476	-1.01	predicted gene 5182; acetyl-Coenzyme A carboxylase alpha
Acat2	110460	-1.03	acetyl-Coenzyme A acetyltransferase 2
Acsl3	74205	-1.02	acyl-CoA synthetase long-chain family member 3
Acsl6	216739	-1.09	acyl-CoA synthetase long-chain family member 6
Adamts3	330119	-1.195	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 3
Adcyap1r1	11517	-1.19	adenylate cyclase activating polypeptide 1 receptor 1
AK129341	234915	-1.09	similar to CDNA sequence AK129341; cDNA sequence AK129341
Akap12	83397	-1.19	A kinase (PRKA) anchor protein (gravin) 12
Akr1b3	11677	-1.4525	aldo-keto reductase family 1, member B3 (aldose reductase)
Amhr2	110542	-1.22	anti-Mullerian hormone type 2 receptor
Amph	218038	-1.22	amphiphysin
Arhgap19	71085	-1.22	Rho GTPase activating protein 19
Arid3b	56380	-1.74	AT rich interactive domain 3B (BRIGHT-like)
Atf7ip2	75329	-1.205	activating transcription factor 7 interacting protein 2
Atp10d	231287	-1.06	ATPase, class V, type 10D
B3gnt5	108105	-1.66	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 5
Bard1	12021	-1.08	BRCA1 associated RING domain 1
BC030867	217216	-1.05	cDNA sequence BC030867
BC051142	407788	-1.09	cDNA sequence BC051142; testis specific basic protein
Bccip	66165	-1.05	BRCA2 and CDKN1A interacting protein
Blm	12144	-1.37	Bloom syndrome homolog (human)
Bop1	12181	-1.23	block of proliferation 1
Brca1	12189	-1.03	breast cancer 1
Brca2	12190	-1.05	breast cancer 2
Bub1	12235	-1.39	budding uninhibited by benzimidazoles 1 homolog (S. cerevisiae)
C330019L16Rik	208111	-1.48	RIKEN cDNA C330019L16 gene
C330024D21Rik	320479	-1.87	RIKEN cDNA C330024D21 gene
C330027C09Rik	224171	-1.22	RIKEN cDNA C330027C09 gene
C530008M17Rik	320827	-1.04	RIKEN cDNA C530008M17 gene
C79407	217653	-1.22	expressed sequence C79407
Cad	69719	-1.28	carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase
Calcr	12311	-1.09	calcitonin receptor
Camkv	235604	-1.29	CaM kinase-like vesicle-associated
Car14	23831	-1.62	carbonic anhydrase 14
Casp8ap2	26885	-1.12	caspase 8 associated protein 2
Ccdc136	232664	-1.05	coiled-coil domain containing 136
Ccdc18	73254	-1.41	coiled-coil domain containing 18
Ccdc21	70012	-1.09	coiled-coil domain containing 21
Ccnb1ip1	239083	-2.15	cyclin B1 interacting protein 1
Ccnd1	12443	-1.13	cyclin D1
Ccne2	12448	-1.21	cyclin E2
Cdc20	107995	-1.03	cell division cycle 20 homolog (S. cerevisiae)
Cdc6	23834	-1.23	cell division cycle 6 homolog (S. cerevisiae); predicted gene 9430; similar to cell division cycle 6 homolog
Cdca5	67849	-1.25	cell division cycle associated 5
Cdca7	66953	-1.08	cell division cycle associated 7
Cdca8	52276	-1.14	cell division cycle associated 8
Cdh2	12558	-1.36	cadherin 2; similar to N-cadherin
Cdx2	12591	-1.71	caudal type homeo box 2
Cebpz	12607	-1.02	CCAAT/enhancer binding protein zeta
Cecr2	330409	-1.79	cat eye syndrome chromosome region, candidate 2 homolog (human)
Cenph	26886	-1.18	centromere protein H
Cenpk	60411	-1.33	centromere protein K
Cenpn	72155	-1.05	centromere protein N
Cenpp	66336	-1.16	centromere protein P
Cep76	225659	-1.05	centrosomal protein 76
Cep78	208518	-1.17	centrosomal protein 78
Chd7	320790	-1.042	chromodomain helicase DNA binding protein 7
Chek1	12649	-1.16	checkpoint kinase 1 homolog (S. pombe)
Chn1	108699	-1.04	chimerin (chimaerin) 1

Cnksr2	245684	-1.26	connector enhancer of kinase suppressor of Ras 2
Crb2	241324	-1.77	crumbs homolog 2 (Drosophila)
Crispld1	83691	-1.43	cysteine-rich secretory protein LCCL domain containing 1
Crlf3	54394	-1.01	cytokine receptor-like factor 3
Crmp1	12933	-1.13	collapsin response mediator protein 1
Cstf3	228410	-1	cleavage stimulation factor, 3' pre-RNA, subunit 3
Cyp26a1	13082	-2.52	cytochrome P450, family 26, subfamily a, polypeptide 1
Cyp51	13121	-1.22	cytochrome P450, family 51
D10Wsu102e	28109	-1.1	DNA segment, Chr 10, Wayne State University 102, expressed
Dapk2	13143	-1.15	death-associated protein kinase 2
Ddx11	320209	-1.02	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homolog, S. cerevisiae)
Ddx21	56200	-1.21	DEAD (Asp-Glu-Ala-Asp) box polypeptide 21
Ddx39	68278	-1.05	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39
Depdc1a	76131	-1.045	DEP domain containing 1a
Depdc1b	218581	-1.27	DEP domain containing 1B
Dhfr	13361	-1.43	dihydrofolate reductase
Dimt1	66254	-1.1	DIM1 dimethyladenosine transferase 1-like (S. cerevisiae)
Dkc1	245474	-1.24	dyskeratosis congenita 1, dyskerin homolog (human)
Dkk1	13380	-1.27	dickkopf homolog 1 (Xenopus laevis)
Dna2	327762	-1.02	DNA replication helicase 2 homolog (yeast)
Dnmt3b	13436	-2.09	DNA methyltransferase 3B
Dok4	114255	-1.81	docking protein 4
Donson	60364	-1.04	downstream neighbor of SON
Dscc1	72107	-1.33	defective in sister chromatid cohesion 1 homolog (S. cerevisiae)
Dus4l	71916	-1.16	dihydrouridine synthase 4-like (S. cerevisiae)
Dusp4	319520	-1.6	dual specificity phosphatase 4
Dut	110074	-1.08	deoxyuridine triphosphatase
E2f5	13559	-1.24	E2F transcription factor 5
E330027M22Rik	100038419	-2.25	RIKEN cDNA gene, E330027M22Rik
Efna3	13638	-1.28	ephrin A3; similar to Ephrin A3
Eif2b3	108067	-1.04	eukaryotic translation initiation factor 2B, subunit 3
Elavl2	15569	-1.12	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2 (Hu antigen B)
Emb	13723	-1.94	embigin
Eme1	268465	-1.1	essential meiotic endonuclease 1 homolog 1 (S. pombe)
Ercc6l	236930	-1.12	excision repair cross-complementing rodent repair deficiency complementation group 6 - like
Esco2	71988	-1.02	establishment of cohesion 1 homolog 2 (S. cerevisiae)
Espl1	105988	-1.04	extra spindle poles-like 1 (S. cerevisiae)
Exo1	26909	-1.46	exonuclease 1
Exosc8	69639	-1.13	exosome component 8
F630043A04Rik	219114	-1.15	RIKEN cDNA F630043A04 gene
F730047E07Rik	212377	-1.04	RIKEN cDNA F730047E07 gene
Fam123c	211383	-1.27	family with sequence similarity 123, member C
Fam136a	66488	-1.01	RIKEN cDNA 2010309E21 gene; similar to CG5323-PA; predicted gene 6396; predicted gene 6624
Fam169a	320557	-1.66	family with sequence similarity 169, member A
Fam54a	71804	-1.07	similar to DUF729 domain containing 1; family with sequence similarity 54, member A
Fam64a	109212	-1.03	RIKEN cDNA 6720460F02 gene
Fancd2	211651	-1.25	Fanconi anemia, complementation group D2
Fancm	104806	-1.11	Fanconi anemia, complementation group M
Fbl	14113	-1.07	similar to Fibrillarin; fibrillarin
Fbxo5	67141	-1.05	F-box protein 5
Fen1	14156	-1.18	flap structure specific endonuclease 1
Fndc3c1	333564	-1.81	fibronectin type III domain containing 3C1
Fscn1	14086	-1.01	fascin homolog 1, actin bundling protein (Strongylocentrotus purpuratus)
Gemin4	276919	-1.04	gem (nuclear organelle) associated protein 4
Gemin6	67242	-1.295	predicted gene 6253; gem (nuclear organelle) associated protein 6; similar to gem (nuclear organelle) associated protein 6
Gen1	209334	-1.36	Gen homolog 1, endonuclease (Drosophila)
Gfpt2	14584	-1.73	glutamine fructose-6-phosphate transaminase 2
Gins1	69270	-1.38	GINS complex subunit 1 (Psf1 homolog)

Gls2	216456	-1.07	glutaminase 2 (liver, mitochondrial)
Glt1d1	319804	-1.18	glycosyltransferase 1 domain containing 1
Gm10661	100038435	-1.21	predicted gene 10661
Gm10664	100038489	-1.52	predicted gene 10664
Gm15542	100043546	-1.15	predicted gene 15542; predicted gene 7816; similar to DnaJ-like protein; predicted gene 6335; DnaJ (Hsp40) homolog, subfamily A, member 1, pseudogene; DnaJ (Hsp40) homolog, subfamily A, member 1
Gm4893	235279	-1.03	predicted gene 11221; similar to Protein C14orf111 homolog; predicted gene 4893; FCF1 small subunit (SSU) processome component homolog (S. cerevisiae)
Gm5595	434179	-1.26	predicted gene 5595; predicted gene 2381
Gm7239	638487	-1.31	predicted gene, EG625349; predicted gene 5789; predicted gene 7085; predicted gene 5708; predicted gene 6847; SET translocation; cDNA sequence BC085271; predicted gene 7239; similar to protein phosphatase 2A inhibitor-2 I-2PP2A; predicted gene 9531
Gmnn	57441	-1.225	geminin
Gng3	14704	-1.17	guanine nucleotide binding protein (G protein), gamma 3
Gnl3	30877	-1.03	guanine nucleotide binding protein-like 3 (nucleolar)
Gpatch4	66614	-1.1	G patch domain containing 4
Greb1	268527	-1.56	gene regulated by estrogen in breast cancer protein
Grik3	14807	-1.29	glutamate receptor, ionotropic, kainate 3
Grwd1	101612	-1.18	glutamate-rich WD repeat containing 1
H2-Oa	15001	-1.11	histocompatibility 2, O region alpha locus
Hand1	15110	-2.19	heart and neural crest derivatives expressed transcript 1
Has2	15117	-1.69	hyaluronan synthase 2
Haus6	230376	-1.08	HAUS augmin-like complex, subunit 6
Heatr1	217995	-1.19	HEAT repeat containing 1
Hells	15201	-1.2	helicase, lymphoid specific
Herc1	235439	-1.06	hect (homologous to the E6-AP (UBE3A) carboxyl terminus) domain and RCC1 (CHC1)-like domain (RLD) 1
Hey2	15214	-1.68	hairy/enhancer-of-split related with YRPW motif 2
Hmga1	15361	-1.34	high mobility group AT-hook I, related sequence 1; high mobility group AT-hook 1
Hmgb3	15354	-1.115	predicted gene 11805; predicted gene 8850; high mobility group box 3; similar to High mobility group protein 4 (HMG-4) (High mobility group protein 2a) (HMG-2a)
Hmmr	15366	-1.04	hyaluronan mediated motility receptor (RHAMM)
Hoxb6	15414	-1.03	homeo box B6
Hoxd1	15429	-1.47	homeo box D1
Hspbap1	66667	-1.09	Hspb associated protein 1
Hsph1	15505	-1.05	heat shock 105kDa/110kDa protein 1
Id3	15903	-1.32	inhibitor of DNA binding 3
Idi1	319554	-1.655	isopentenyl-diphosphate delta isomerase; similar to Isopentenyl-diphosphate delta isomerase; predicted gene 7655
Ifrd2	15983	-1.05	interferon-related developmental regulator 2
Igf2bp1	140486	-1.39	insulin-like growth factor 2 mRNA binding protein 1
Il17rd	171463	-1.38	interleukin 17 receptor D
Isy1	57905	-1.04	ISY1 splicing factor homolog (S. cerevisiae)
Jph1	57339	-1.185	junctophilin 1
Kbtbd8	243574	-1.56	kelch repeat and BTB (POZ) domain containing 8
Kif11	16551	-1.06	kinesin family member 11
Kif14	381293	-1.07	kinesin family member 14
Kif15	209737	-1.38	kinesin family member 15
Kif18a	228421	-1.04	kinesin family member 18A
Kif1a	16560	-1.27	kinesin family member 1A
Kif20a	19348	-1.1	kinesin family member 20A
Kif20b	240641	-1.07	kinesin family member 20B
Kif22	110033	-1.16	kinesin family member 22
Kif4	16571	-1.01	kinesin family member 4
Klf8	245671	-1.78	Kruppel-like factor 8
Klhl12	240756	-1.03	kelch-like 12 (Drosophila)
Kntc1	208628	-1.05	kinetochore associated 1
Lbr	98386	-1.26	lamin B receptor
Lin28a	83557	-1.79	lin-28 homolog A (C. elegans)
Lin28b	380669	-1.54	lin-28 homolog B (C. elegans)
Lrp11	237253	-1.06	low density lipoprotein receptor-related protein 11
Lrrn4	320974	-2.04	leucine rich repeat neuronal 4
Lss	16987	-1.12	lanosterol synthase

Lyar	17089	-1.31	Ly1 antibody reactive clone
Mak16	67920	-1.05	MAK16 homolog (S. cerevisiae)
Mapk8ip2	60597	-1.23	mitogen-activated protein kinase 8 interacting protein 2
Mastl	67121	-1.11	microtubule associated serine/threonine kinase-like
Mcm10	70024	-1.37	minichromosome maintenance deficient 10 (S. cerevisiae)
Mcm2	17216	-1.05	minichromosome maintenance deficient 2 mitotin (S. cerevisiae)
Mcm4	17217	-1.16	minichromosome maintenance deficient 4 homolog (S. cerevisiae)
Mcm5	17218	-1.08	minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)
Mcm7	17220	-1.07	minichromosome maintenance deficient 7 (S. cerevisiae)
Mcm8	66634	-1.06	minichromosome maintenance deficient 8 (S. cerevisiae)
Mdm1	17245	-1.12	transformed mouse 3T3 cell double minute 1
Mdn1	100019	-1.2	midasin homolog (yeast)
Melk	17279	-1.19	maternal embryonic leucine zipper kinase
Mirhg1	75957	-1.24	microRNA host gene 1 (non-protein coding)
Mki67ip	67949	-1.1	Mki67 (FHA domain) interacting nucleolar phosphoprotein
Mlf1ip	71876	-1.16	myeloid leukemia factor 1 interacting protein
Mlh1	17350	-1.36	mutL homolog 1 (E. coli)
Mnd1	76915	-1.84	meiotic nuclear divisions 1 homolog (S. cerevisiae); predicted gene 3833; similar to Meiotic nuclear divisions 1 homolog (S. cerevisiae)
Mphosph10	67973	-1.11	M-phase phosphoprotein 10 (U3 small nucleolar ribonucleoprotein)
Mpp6	56524	-1.1	membrane protein, palmitoylated 6 (MAGUK p55 subfamily member 6)
Mrto4	69902	-1.245	predicted gene 9178; MRT4, mRNA turnover 4, homolog (S. cerevisiae); predicted gene 5633
Msto1	229524	-1	misato homolog 1 (Drosophila)
Mtap7d3	320923	-1.29	MAP7 domain containing 3
Mthfd1	108156	-1.03	methylenetetrahydrofolate dehydrogenase (NADP+ dependent), methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthase
Mum1	68114	-1	melanoma associated antigen (mutated) 1
Mum1l1	245631	-1.82	melanoma associated antigen (mutated) 1-like 1
Mvd	192156	-1.41	mevalonate (diphospho) decarboxylase
Myb	17863	-1.33	myeloblastosis oncogene
Naf1	234344	-1.34	nuclear assembly factor 1 homolog (S. cerevisiae)
Nasp	50927	-1.47	nuclear autoantigenic sperm protein (histone-binding); similar to nuclear autoantigenic sperm protein; NASP
Nat10	98956	-1.04	N-acetyltransferase 10
Ncapg	54392	-1.31	non-SMC condensin I complex, subunit G
Ncapg2	76044	-1.04	non-SMC condensin II complex, subunit G2
Ncl	17975	-1.42	nucleolin
Ncrna00086	320237	-1.39	non-protein coding RNA 86
Ndc80	67052	-1.04	NDC80 homolog, kinetochore complex component (S. cerevisiae)
Nfkbil2	72749	-1.06	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 2
Nfxl1	100978	-1.616666667	nuclear transcription factor, X-box binding-like 1
Nhej1	75570	-1.06	nonhomologous end-joining factor 1
Ninl	78177	-1.16	ninein-like
Nkx1-2	20231	-1.07	NK1 transcription factor related, locus 2 (Drosophila)
Nmt2	18108	-1.05	N-myristoyltransferase 2
Noc3l	57753	-1.01	nucleolar complex associated 3 homolog (S. cerevisiae)
Nolc1	70769	-1.14	nucleolar and coiled-body phosphoprotein 1
Nop58	55989	-1	NOP58 ribonucleoprotein homolog (yeast)
Npm3	18150	-1.24	nucleoplasmin 3; nucleoplasmin 3, pseudogene 1
Nr6a1	14536	-1.08	nuclear receptor subfamily 6, group A, member 1
Nsdhl	18194	-1.02	NAD(P) dependent steroid dehydrogenase-like
Nudt11	58242	-1.26	nudix (nucleoside diphosphate linked moiety X)-type motif 11; nudix (nucleoside diphosphate linked moiety X)-type motif 10
Nudt15	214254	-1.17	nudix (nucleoside diphosphate linked moiety X)-type motif 15
Nup107	103468	-1.11	nucleoporin 107
Nup35	69482	-1.225	predicted gene 4353; nucleoporin 35
Nup37	69736	-1.12	similar to nucleoporin 37; nucleoporin 37
Nup43	69912	-1.3	nucleoporin 43
Nutf2	68051	-1.16	similar to Chain A, D92n,D94n Double Point Mutant Of Human Nuclear Transport Factor 2 (Ntf2); predicted gene 10349; predicted gene 4682; nuclear transport factor 2; similar to nuclear transport factor 2; predicted gene 9386; predicted gene 10333
Nxf3	245610	-1.03	nuclear RNA export factor 3

Oip5	70645	-1.25	Opa interacting protein 5
Orc1l	18392	-1.6	origin recognition complex, subunit 1-like (S.cereviaiae)
Orc2l	18393	-1.05	origin recognition complex, subunit 2-like (S. cerevisiae)
Osgepl1	72085	-1.03	O-sialoglycoprotein endopeptidase-like 1
Osgin2	209212	-1.36	oxidative stress induced growth inhibitor family member 2
Pcbp4	59092	-1.21	poly(rC) binding protein 4
Pcyt1b	236899	-1.4	phosphate cytidylyltransferase 1, choline, beta isoform
Pdcd11	18572	-1	programmed cell death 11
Pdlim4	30794	-1.04	PDZ and LIM domain 4
Pdss1	56075	-1.07	prenyl (solanesyl) diphosphate synthase, subunit 1
Pfas	237823	-1	phosphoribosylformylglycinamidine synthase (FGAR amidotransferase)
Pfkfb1	18639	-1.21	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 1
Plk1	18817	-1.02	polo-like kinase 1 (Drosophila)
Plk4	20873	-1.1	polo-like kinase 4 (Drosophila)
Pm20d2	242377	-1.07	peptidase M20 domain containing 2
Pmvk	68603	-1.1	phosphomevalonate kinase
Pold3	67967	-1.03	polymerase (DNA-directed), delta 3, accessory subunit
Pole	18973	-1.23	polymerase (DNA directed), epsilon
Pole2	18974	-1.17	polymerase (DNA directed), epsilon 2 (p59 subunit)
Polg	18975	-1.4	polymerase (DNA directed), gamma
Polr1e	64424	-1.16	polymerase (RNA) I polypeptide E
Polr3g	67486	-1.74	polymerase (RNA) III (DNA directed) polypeptide G
Pou5f1	18999	-1.68	POU domain, class 5, transcription factor 1
Ppa1	67895	-1.01	pyrophosphatase (inorganic) 1
Ppat	231327	-1.28	phosphoribosyl pyrophosphate amidotransferase
Ppih	66101	-1.105	similar to peptidyl prolyl isomerase H; predicted gene 7879; predicted gene 9088; predicted gene 8719; predicted gene 11585; similar to Peptidyl-prolyl cis-trans isomerase H (PPIase H) (Rotamase H); peptidyl prolyl isomerase H
Ppil1	100047806	-1.21	similar to peptidylprolyl isomerase-like 1; peptidylprolyl isomerase (cyclophilin)-like 1
Ppil1	68816	-1.85	similar to peptidylprolyl isomerase-like 1; peptidylprolyl isomerase (cyclophilin)-like 1
Prdm6	225518	-1.02	PR domain containing 6
Prim1	19075	-1.13	DNA primase, p49 subunit
Prim2	19076	-1.73	DNA primase, p58 subunit
Prkd1	18760	-1.02	protein kinase D1
Prmt3	71974	-1.49	protein arginine N-methyltransferase 3
Prph	19132	-2.4	peripherin
Prtg	235472	-1.04	protogenin homolog (Gallus gallus)
Ptcd3	69956	-1.07	pentatricopeptide repeat domain 3
Pus7	78697	-1.17	pseudouridylate synthase 7 homolog (S. cerevisiae)
Rad51	19361	-1.13	RAD51 homolog (S. cerevisiae)
Rad54l	19366	-1.16	RAD54 like (S. cerevisiae)
Rangrf	57785	-1.04	predicted gene 7791; similar to Ran-interacting protein MOG1; predicted gene 15711; RAN guanine nucleotide release factor; predicted gene 4535; predicted gene 8572
Rbm19	74111	-1.76	RNA binding motif protein 19
Rbpms2	71973	-1.22	predicted gene 3470; RNA binding protein with multiple splicing 2
Rcl1	59028	-1.25	RNA terminal phosphate cyclase-like 1
Rcor2	104383	-1.23	REST corepressor 2
Rdh11	17252	-1.01	retinol dehydrogenase 11
Rfc3	69263	-1.03	replication factor C (activator 1) 3
Rhebl1	69159	-1.62	Ras homolog enriched in brain like 1
Ripk2	192656	-1.53	receptor (TNFRSF)-interacting serine-threonine kinase 2
Ripk3	56532	-1.13	receptor-interacting serine-threonine kinase 3
Rnu73a	19870	-1.12	U73B small nuclear RNA; U73A small nuclear RNA
Rnu73b	19871	-1.34	U73B small nuclear RNA; U73A small nuclear RNA
Rock2	19878	-1.29	Rho-associated coiled-coil containing protein kinase 2
Rpl12	269261	-1.15	similar to ribosomal protein L12; small nucleolar RNA, H/ACA box 65;similar to 60S ribosomal protein L12; ribosomal protein L12;
Rpl13	270106	-1.53	similar to ribosomal protein L13; similar to 60S ribosomal protein L13; ribosomal protein L13
Rpl23a	268449	-1.28	; ribosomal protein L23a; similar to 60S ribosomal protein L23a;
Rpp30	54364	-1.03	ribonuclease P/MRP 30 subunit (human)
Rpp40	208366	-1.18	ribonuclease P 40 subunit (human)
Rps12	20042	-1.23	ribosomal protein S12;

Rps13	68052	-1.26	similar to ribosomal protein S13; predicted gene 12270; predicted gene 6834; predicted gene 15483; predicted gene 6573; ribosomal protein S13; predicted gene 10159
Rps6ka6	67071	-2.24	ribosomal protein S6 kinase polypeptide 6
Rragb	245670	-1.35	Ras-related GTP binding B
Rrm2	20135	-1.03	ribonucleotide reductase M2
Rrp12	107094	-1.01	ribosomal RNA processing 12 homolog (S. cerevisiae)
Rrp1b	72462	-1.2	ribosomal RNA processing 1 homolog B (S. cerevisiae)
Ruvbl1	56505	-1.02	RuvB-like protein 1
Sall1	58198	-1.33	sal-like 1 (Drosophila)
Sall4	99377	-2.07	sal-like 4 (Drosophila)
Scarna8	100217448	-1.02	small Cajal body-specific RNA 8
Scd1	20249	-1.44	stearoyl-Coenzyme A desaturase 1
Scyl3	240880	-1.11	SCY1-like 3 (S. cerevisiae)
Sgk1	20393	-1.07	serum/glucocorticoid regulated kinase 1
Sgol1	72415	-1.09	shugoshin-like 1 (S. pombe)
Shcbp1	20419	-1.09	Shc SH2-domain binding protein 1
Siah1b	20438	-1.14	seven in absentia 1B
Sirt1	93759	-1.05	sirtuin 1 (silent mating type information regulation 2, homolog) 1 (S. cerevisiae)
Six6os1	75801	-1.53	Six6 opposite strand transcript 1
Slc43a1	72401	-1.04	solute carrier family 43, member 1
Slc4a5	232156	-1.22	solute carrier family 4, sodium bicarbonate cotransporter, member 5
Smad6	17130	-1.48	MAD homolog 6 (Drosophila)
Smarcd1	83797	-1.12	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 1
Smn1	20595	-1.04	survival motor neuron 1
Smpd4	77626	-1.07	sphingomyelin phosphodiesterase 4
Smyd5	232187	-1.01	SET and MYND domain containing 5
Snhg1	83673	-1.7125	small nucleolar RNA host gene (non-protein coding) 1
Snora61	100217440	-1.15	small nucleolar RNA, H/ACA box 61
Snora62	104433	-1.01	small nucleolar RNA, H/ACA box 62
Snora65	104367	-1.19	similar to ribosomal protein L12; predicted gene 7117; small nucleolar RNA, H/ACA box 65; predicted gene 11425; predicted gene 6285; predicted gene 5962; predicted gene 9396; similar to 60S ribosomal protein L12; ribosomal protein L12; predicted gene 6336
Snora69	104369	-1.67	small nucleolar RNA, H/ACA box 69
Snora7a	100217451	-1.05	small nucleolar RNA, H/ACA box 7A
Snord1c	100216536	-1.4	small nucleolar RNA, C/D box 1C
Snord34	27210	-1.38	small nucleolar RNA, C/D box 34
Snord37	100217454	-1.25	small nucleolar RNA, C/D box 37
Snord38a	100217424	-1.29	small nucleolar RNA, C/D box 38A
Snord49a	100217455	-1.29	small nucleolar RNA, C/D box 49A
Snord58b	100217457	-1.3	small nucleolar RNA, C/D box 58B
Snord82	80828	-1.33	small nucleolar RNA, C/D box 82
Snord87	266793	-1.52	small nucleolar RNA, C/D box 87
Snrnp25	78372	-1.22	small nuclear ribonucleoprotein 25 (U11/U12)
Snrpg	68011	-1.01	predicted gene 8186; small nuclear ribonucleoprotein polypeptide G
Snrpn	20646	-1.16	small nuclear ribonucleoprotein N; SNRPN upstream reading frame; predicted gene 5802; similar to SNRPN upstream reading frame protein
Spag5	54141	-1.11	sperm associated antigen 5
Spata13	219140	-1.17	spermatogenesis associated 13
Spata5	57815	-1.01	spermatogenesis associated 5
Spin2	278240	-2.1	spindlin family, member 2
Spo11	26972	-1.06	sporulation protein, meiosis-specific, SPO11 homolog (S. cerevisiae)
Stard8	236920	-2.07	START domain containing 8
Stmn2	20257	-2.09	stathmin-like 2
Sult4a1	29859	-2.1	sulfotransferase family 4A, member 1
Suv39h1	20937	-1.52	suppressor of variegation 3-9 homolog 1 (Drosophila)
Suv39h2	64707	-1.44	suppressor of variegation 3-9 homolog 2 (Drosophila)
Taf4b	72504	-1.36	TAF4B RNA polymerase II, TATA box binding protein (TBP)-associated factor
Tbx20	57246	-1.43	T-box 20
Tbx3	21386	-1.82	T-box 3
Tcf7	21414	-1.9	transcription factor 7, T-cell specific
Tcfl5	277353	-2.17	transcription factor-like 5 (basic helix-loop-helix)

Tdrd5	214575	-1.06	tudor domain containing 5
Terf1	21749	-1.02	telomeric repeat binding factor 1
Thop1	50492	-1.03	thimet oligopeptidase 1
Tiam2	24001	-1.11	T-cell lymphoma invasion and metastasis 2
Timm8a1	30058	-1.035	predicted gene 9797; translocase of inner mitochondrial membrane 8 homolog a1 (yeast); similar to small zinc finger-like protein
Tipin	66131	-1.15	timeless interacting protein
Tmem28	620592	-1.18	transmembrane protein 28
Tmem88	67020	-1.82	transmembrane protein 88
Tnfrsf19	29820	-1.07	tumor necrosis factor receptor superfamily, member 19
Tnfrsf22	79202	-1.2	tumor necrosis factor receptor superfamily, member 22
Top2a	21973	-1.19	topoisomerase (DNA) II alpha
Trim36	28105	-1.06	tripartite motif-containing 36
Trim59	66949	-1.24	similar to mouse RING finger 1; tripartite motif-containing 59
Trim6	94088	-1.01	tripartite motif-containing 6; similar to Tripartite motif protein 6
Trim71	636931	-1.93	similar to LIN41; tripartite motif-containing 71
Trip13	69716	-1.48	thyroid hormone receptor interactor 13
Trmt11	73681	-1.07	tRNA methyltransferase 11 homolog (S. cerevisiae)
Tsen2	381802	-1	tRNA splicing endonuclease 2 homolog (S. cerevisiae)
Tspan2	70747	-1.11	tetraspanin 2
Ttc27	74196	-1.39	tetratricopeptide repeat domain 27
Ttk	22137	-1.24	Ttk protein kinase
Ube2cbp	70348	-1.09	ubiquitin-conjugating enzyme E2C binding protein
Uchl5	56207	-1.06	ubiquitin carboxyl-terminal esterase L5
Umps	22247	-1.02	uridine monophosphate synthetase
Urb1	207932	-1.08	URB1 ribosome biogenesis 1 homolog (S. cerevisiae)
Urod	22275	-1.02	uroporphyrinogen decarboxylase
Usp44	327799	-1.46	ubiquitin specific peptidase 44
Utp14a	72554	-1.29	UTP14, U3 small nucleolar ribonucleoprotein, homolog A (yeast)
Utp18	217109	-1.12	UTP18, small subunit (SSU) processome component, homolog (yeast)
Utp20	70683	-1.05	UTP20, small subunit (SSU) processome component, homolog (yeast)
Vrk1	22367	-1.08	vaccinia related kinase 1
Wdhd1	218973	-1.33	WD repeat and HMG-box DNA binding protein 1
Wdr12	57750	-1.265	WD repeat domain 12; predicted gene 4879
Wdr3	269470	-1	WD repeat domain 3
Wdr46	57315	-1.03	WD repeat domain 46
Wdr54	75659	-1.37	WD repeat domain 54
Wdr55	67936	-1.22	WD repeat domain 55
Wdr75	73674	-1.02	WD repeat domain 75
Wnt5a	22418	-1.43	wingless-related MMTV integration site 5A
Xpo4	57258	-1	exportin 4
Yars2	70120	-1.07	tyrosyl-tRNA synthetase 2 (mitochondrial)
Zc3hav1	78781	-1.09	zinc finger CCCH type, antiviral 1
Zdbf2	73884	-1.33	zinc finger, DBF-type containing 2
Zfp105	22646	-1.13	zinc finger protein 105
Zfp280c	208968	-1.21	zinc finger protein 280C
Zfp345	545471	-1.41	zinc finger protein 345
Zfp365	216049	-1.04	zinc finger protein 365
Zfp7	223669	-1.08	zinc finger protein 7
Zfp711	245595	-2.19	zinc finger protein 711
Zic2	22772	-1.85	zinc finger protein of the cerebellum 2
Zic3	22773	-1.39	zinc finger protein of the cerebellum 3
Zic5	65100	-1.42	similar to zinc finger protein of the cerebellum 5; predicted gene 12241; zinc finger protein of the cerebellum 5
Znhit3	448850	-1.11	zinc finger, HIT type 3
Zscan10	332221	-2.02	zinc finger and SCAN domain containing 10
Zwilch	68014	-1.13	Zwilch, kinetochore associated, homolog (Drosophila)

Upregulated genes in dnmt1-/- EBs compared to wt EBs between d4-16			
Gene symbol	**Entrez Gene ID**	**Fold change compared to wt**	**Gene name**
0610031J06Rik	56700	1.05	RIKEN cDNA 0610031J06 gene
1700001L05Rik	69291	1.42	RIKEN cDNA 1700001L05 gene
1700013H16Rik	75514	2.2	RIKEN cDNA 1700013H16 gene
2210010C17Rik	70080	1.85	RIKEN cDNA 2210010C17 gene
2510049J12Rik	70291	1.36	RIKEN cDNA 2510049J12 gene
2610019F03Rik	72148	1.13	RIKEN cDNA 2610019F03 gene
3830431G21Rik	217682	1.16	RIKEN cDNA 3830431G21 gene
4632428N05Rik	74048	1.28	RIKEN cDNA 4632428N05 gene
4933426M11Rik	217684	1.02	RIKEN cDNA 4933426M11 gene
5730469M10Rik	70564	1.02	RIKEN cDNA 5730469M10 gene
9130017N09Rik	78906	1.27	RIKEN cDNA 9130017N09 gene
9930013L23Rik	80982	1.75	RIKEN cDNA 9930013L23 gene
A4galt	239559	1.19	alpha 1,4-galactosyltransferase
Aass	30956	1.42	aminoadipate-semialdehyde synthase
Abcb11	27413	1.01	ATP-binding cassette, sub-family B (MDR/TAP), member 11
Abcc4	239273	1.46	ATP-binding cassette, sub-family C (CFTR/MRP), member 4
Acat3	224530	1.25	acetyl-Coenzyme A acetyltransferase 3
Acp5	11433	1.08	acid phosphatase 5, tartrate resistant
Adamts9	101401	1.148	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 9
Adap2	216991	1.16	ArfGAP with dual PH domains 2
Amn	93835	1.74	amnionless
Angpt2	11601	1.13	angiopoietin 2
Apoc1	11812	1.72	apolipoprotein C-I
Apoe	11816	1.53	apolipoprotein E
Arhgap28	268970	1.44	Rho GTPase activating protein 28
Arhgap8	73167	1.21	Rho GTPase activating protein 8; proline rich 5 (renal)
Arhgef16	230972	1.05	Rho guanine nucleotide exchange factor (GEF) 16
As3mt	57344	1.31	arsenic (+3 oxidation state) methyltransferase
Atp6v0a1	11975	1.24	ATPase, H+ transporting, lysosomal V0 subunit A1
Atp7b	11979	1.85	ATPase, Cu++ transporting, beta polypeptide
Atp8a2	50769	1.81	ATPase, aminophospholipid transporter-like, class I, type 8A, member 2
Atxn1	20238	1.16	ataxin 1
Bace2	56175	1.61	beta-site APP-cleaving enzyme 2
Bcmo1	63857	1.88	beta-carotene 15,15'-monooxygenase
Bend5	67621	1.5	BEN domain containing 5
Car7	12354	3.21	carbonic anhydrase 7
Casq1	12372	1.18	calsequestrin 1
Cck	12424	1.14	cholecystokinin
Ccl2	20296	1.17	chemokine (C-C motif) ligand 2
Ceacam15	101434	2.73	carcinoembryonic antigen-related cell adhesion molecule 15
Ceacam9	26368	2.63	carcinoembryonic antigen-related cell adhesion molecule 9
Chrnb1	11443	1.08	cholinergic receptor, nicotinic, beta polypeptide 1 (muscle)
Cited1	12705	2.15	Cbp/p300-interacting transactivator with Glu/Asp-rich carboxy-terminal domain 1
Cldn6	54419	1.07	claudin 6
Clic6	209195	2.69	chloride intracellular channel 6
Cln6	76524	1.35	ceroid-lipofuscinosis, neuronal 6
Clps	109791	2.15	colipase, pancreatic
Cndp2	66054	1.56	CNDP dipeptidase 2 (metallopeptidase M20 family)
Cntn5	244682	1.57	contactin 5
Cpeb2	231207	1.51	cytoplasmic polyadenylation element binding protein 2
Cpn1	93721	2.14	carboxypeptidase N, polypeptide 1
Crabp2	12904	1.11	cellular retinoic acid binding protein II
Creb3l1	26427	1.48	cAMP responsive element binding protein 3-like 1
Creb3l3	208677	1.73	cAMP responsive element binding protein 3-like 3
Cst3	13010	1.11	cystatin C
Ctdspl	69274	1.16	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase-like

Cts7	56092	1.97	cathepsin 7
Ctsh	13036	3.91	cathepsin H
D14Ertd668e	219132	2.09	predicted gene 6907; predicted gene 6904; predicted gene 4902; DNA segment, Chr 14, ERATO Doi 668, expressed; PHD finger protein 11
Dazl	13164	2.61	deleted in azoospermia-like
Ddah1	69219	1.72	dimethylarginine dimethylaminohydrolase 1
Depdc7	211896	1.46	DEP domain containing 7
Dkkl1	50722	1.59	dickkopf-like 1
Dnajb9	27362	1.04	predicted gene 6568; DnaJ (Hsp40) homolog, subfamily B, member 9
Dnajc22	72778	1.27	DnaJ (Hsp40) homolog, subfamily C, member 22
Dpcr1	268949	2.8	diffuse panbronchiolitis critical region 1 (human)
Dpep1	13479	1.02	dipeptidase 1 (renal)
Dpep3	71854	1.14	dipeptidase 3
Dpp4	13482	3.8	dipeptidylpeptidase 4
dram2	67171	1.2	DNA-damage regulated autophagy modulator 2; oldy symbol Tmem77
Eif2ak2	19106	1.16	eukaryotic translation initiation factor 2-alpha kinase 2
Entpd2	12496	1.28	ectonucleoside triphosphate diphosphohydrolase 2
Epas1	13819	2.73	endothelial PAS domain protein 1; similar to Endothelial PAS domain protein 1
Eps8	13860	1.08	epidermal growth factor receptor pathway substrate 8
Eps8l3	99662	1.57	EPS8-like 3
Ero1l	50527	1.79	ERO1-like (S. cerevisiae)
Exoc6	107371	1.07	exocyst complex component 6
Fgfr4	14186	1.22	fibroblast growth factor receptor 4
Fhdc1	229474	1.1	FH2 domain containing 1
Flrt3	71436	1.29	fibronectin leucine rich transmembrane protein 3; similar to fibronectin leucine rich transmembrane protein 3
Fmo4	226564	1.9	flavin containing monooxygenase 4
Fndc3a	319448	1.24	fibronectin type III domain containing 3A
Foxa3	15377	1.12	forkhead box A3
Fuca2	66848	1.22	fucosidase, alpha-L- 2, plasma
Gab1	14388	1.58	growth factor receptor bound protein 2-associated protein 1
Galnt6	207839	1.71	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 6
Gcnt4	218476	1.14	predicted gene 73
Gdpd5	233552	1.11	glycerophosphodiester phosphodiesterase domain containing 5
Gkn1	66283	2.91	gastrokine 1
Gm12569	622699	1.1	predicted gene 12569
Gm1527	385263	1.09	predicted gene 1527; predicted gene 6558
Gm2889	100040658	1.91	hypothetical protein LOC100041609; hypothetical protein LOC100044795; predicted gene 3395; similar to gag polyprotein; hypothetical protein LOC100047557; hypothetical protein LOC100040347; hypothetical protein LOC100044384; hypothetical protein LOC100045342; hypothetical protein LOC100038979; predicted gene 2889
Gm2889	100041537	1.265	hypothetical protein LOC100041609; hypothetical protein LOC100044795; predicted gene 3395; similar to gag polyprotein; hypothetical protein LOC100047557; hypothetical protein LOC100040347; hypothetical protein LOC100044384; hypothetical protein LOC100045342; hypothetical protein LOC100038979; predicted gene 2889
Gm371	236914	1.13	predicted gene 371
Gm4779	212753	1.17	predicted gene 4779
Gm5077	317677	1.09	predicted gene 5077
Gm7455	665033	1.44	predicted gene 7455
Gm9	194854	1.07	predicted gene 9
Gns	75612	1.17	glucosamine (N-acetyl)-6-sulfatase
Gpam	14732	1.07	glycerol-3-phosphate acyltransferase, mitochondrial
Gpx3	14778	2.19	glutathione peroxidase 3
Gramd1b	235283	2.37	GRAM domain containing 1B
Gramd1c	207798	1.53	GRAM domain containing 1C
Grb10	14783	1.66	growth factor receptor bound protein 10
Hbegf	15200	1.18	heparin-binding EGF-like growth factor
Hnf1b	21410	1.92	HNF1 homeobox B
Hs3st1	15476	1.86	heparan sulfate (glucosamine) 3-O-sulfotransferase 1
Hsf2bp	74377	1.36	heat shock transcription factor 2 binding protein
Ifi27l2b	217845	1.08	interferon, alpha-inducible protein 27 like 2B
Igf2r	16004	1.56	insulin-like growth factor 2 receptor
Igfbp6	16012	1.09	insulin-like growth factor binding protein 6

Il1rl2	107527	1.13	interleukin 1 receptor-like 2
Il22ra1	230828	2.16	interleukin 22 receptor, alpha 1
Insig1	231070	1.17	insulin induced gene 1
Irf7	54123	1.73	interferon regulatory factor 7
Kit	16590	1.57	kit oncogene
Klf4	16600	1.08	Kruppel-like factor 4 (gut)
Klhl2	77113	1.2	kelch-like 2, Mayven (Drosophila)
Kmo	98256	2.22	kynurenine 3-monooxygenase (kynurenine 3-hydroxylase)
Lamb2	16779	1.13	laminin, beta 2
Lcat	16816	1.08	lecithin cholesterol acyltransferase
Lgi1	56839	1.22	leucine-rich repeat LGI family, member 1; predicted gene 3888
Lrpap1	16976	1.66	low density lipoprotein receptor-related protein associated protein 1
Lrrc8b	433926	1.07	leucine rich repeat containing 8 family, member B
Lrrc8d	231549	1.05	leucine rich repeat containing 8D
Mael	98558	1.8	maelstrom homolog (Drosophila)
Magea1	17137	1.12	melanoma antigen, family A, 1
Man2b2	17160	1.09	mannosidase 2, alpha B2
Marveld2	218518	1.16	MARVEL (membrane-associating) domain containing 2
Mcoln2	68279	1.6	mucolipin 2
Mcoln3	171166	2.2	mucolipin 3
Me1	17436	1.48	predicted gene 7049; similar to NADP-dependent malic enzyme (NADP-ME) (Malic enzyme 1); malic enzyme 1, NADP(+)-dependent, cytosolic
Mmp3	17392	1.12	matrix metallopeptidase 3
Mobkl2b	214944	1	MOB1, Mps One Binder kinase activator-like 2B (yeast)
Morc1	17450	1.31	microrchidia 1
Mosc1	66112	1.74	similar to MOSC domain-containing protein 1, mitochondrial; MOCO sulphurase C-terminal domain containing 1
Mtmr7	54384	1.35	myotubularin related protein 7
Naaa	67111	1.15	N-acylethanolamine acid amidase
Neu1	18010	1.77	neuraminidase 1
Npc1l1	237636	1.12	NPC1-like 1
Nuak2	74137	1.08	NUAK family, SNF1-like kinase, 2
Oas1c	114643	1.19	2'-5' oligoadenylate synthetase 1C
Olfr1054	259021	1.61	olfactory receptor 1054
Osbpl6	99031	1.28	oxysterol binding protein-like 6
P4ha1	18451	1.54	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha 1 polypeptide
Pdia5	72599	1.25	protein disulfide isomerase associated 5
Pdpn	14726	2.07	podoplanin
Pdzk1	59020	1.53	PDZ domain containing 1
Pgam2	56012	1.17	phosphoglycerate mutase 2
Pgc	109820	3.81	progastricsin (pepsinogen C)
Phf16	382207	1.25	PHD finger protein 16
Phf17	269424	1.51	PHD finger protein 17
Pik3cb	74769	1.21	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Pilra	231805	1.09	paired immunoglobin-like type 2 receptor alpha
Plekhb2	226971	1.41	pleckstrin homology domain containing, family B (evectins) member 2
Plekhh1	211945	1.3	pleckstrin homology domain containing, family H (with MyTH4 domain) member 1
Plk3	12795	1.08	polo-like kinase 3 (Drosophila)
Plscr1	22038	1	phospholipid scramblase 1; predicted gene 5530; hypothetical protein LOC677340
Podxl	27205	1.47	podocalyxin-like
Prl2c1	666317	4.36	Prolactin family 2, subfamily c, member 1
Prl3d1	18775	4.8	prolactin family 3, subfamily d, member 1
Prl3d3	215029	3.28	prolactin family 3, subfamily d, member 3
Prl5a1	28078	2.65	prolactin family 5, subfamily a, member 1
Prl7a1	19113	4.32	prolactin family 7, subfamily a, member 1
Prl7a2	19114	1.07	prolactin family 7, subfamily a, member 2
Prlr	19116	1.71	prolactin receptor
Procr	19124	1.5	protein C receptor, endothelial
Pros1	19128	1.04	protein S (alpha)
Psg29	114872	2.99	pregnancy-specific glycoprotein 29

Pth1r	19228	1.48	parathyroid hormone 1 receptor
Ptk6	20459	1.84	PTK6 protein tyrosine kinase 6
Pyy	217212	2.37	peptide YY
Rab20	19332	1.04	RAB20, member RAS oncogene family
Rab30	75985	1.32	RAB30, member RAS oncogene family
Ralgapa2	241694	1.2	Ral GTPase activating protein, alpha subunit 2 (catalytic)
Rbmy1a1	19657	2.063333333	RNA binding motif protein, Y chromosome, family 1, member A1
Rfx6	320995	1.57	regulatory factor X, 6
Rgs4	19736	1.62	regulator of G-protein signaling 4
Rhobtb1	69288	1.19	Rho-related BTB domain containing 1
Rhpn2	52428	1.05	rhophilin, Rho GTPase binding protein 2
Rnasel	24014	1.33	ribonuclease L (2', 5'-oligoisoadenylate synthetase-dependent)
S100a1	20193	1.27	S100 calcium binding protein A1
Scml2	107815	1.19	similar to sex comb on midleg-like 2 (Drosophila); sex comb on midleg-like 2 (Drosophila)
Scpep1	74617	1.31	serine carboxypeptidase 1
Serpinb6c	97848	1.85	serine (or cysteine) peptidase inhibitor, clade B, member 6c
Sfmbt2	353282	1.47	Scm-like with four mbt domains 2
Sgpp2	433323	1.05	sphingosine-1-phosphate phosphotase 2
Sh3d19	27059	1.25	SH3 domain protein D19
Shroom3	27428	1.03	shroom family member 3
Siae	22619	1.23	sialic acid acetylesterase
Slc13a3	114644	1.64	solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 3
Slc13a4	243755	1.97	solute carrier family 13 (sodium/sulfate symporters), member 4
Slc16a9	66859	1.14	solute carrier family 16 (monocarboxylic acid transporters), member 9
Slc23a1	20522	1.68	solute carrier family 23 (nucleobase transporters), member 1
Slc23a3	22626	1.96	solute carrier family 23 (nucleobase transporters), member 3
Slc28a2	269346	1.27	solute carrier family 28 (sodium-coupled nucleoside transporter), member 2
Slc30a1	22782	1.27	solute carrier family 30 (zinc transporter), member 1
Slc47a1	67473	1.21	solute carrier family 47, member 1
Slc6a4	15567	1.45	solute carrier family 6 (neurotransmitter transporter, serotonin), member 4
Slc9a6	236794	1.42	solute carrier family 9 (sodium/hydrogen exchanger), member 6
Slco4c1	227394	1.95	solute carrier organic anion transporter family, member 4C1
Smc1b	140557	2.07	structural maintenance of chromosomes 1B
Smoc1	64075	1.22	SPARC related modular calcium binding 1
Sntb1	20649	1.14	syntrophin, basic 1
Snx9	66616	1.21	similar to Sorting nexin 9; sorting nexin 9
Socs3	12702	1.28	suppressor of cytokine signaling 3
Sohlh2	74434	2.65	spermatogenesis and oogenesis specific basic helix-loop-helix 2
Spns2	216892	1.23	spinster homolog 2 (Drosophila)
Spock1	20745	1.39	sparc/osteonectin, cwcv and kazal-like domains proteoglycan 1
St3gal1	20442	1.5	ST3 beta-galactoside alpha-2,3-sialyltransferase 1
Stra8	20899	1.41	stimulated by retinoic acid gene 8
Stx7	53331	1.19	syntaxin 7
Sycp1	20957	1.45	synaptonemal complex protein 1; similar to testicular protein
Taf7l	74469	2.17	TAF7-like RNA polymerase II, TATA box binding protein (TBP)-associated factor
Tcf7l2	21416	1.1	transcription factor 7-like 2, T-cell specific, HMG-box
Tcfec	21426	1.32	transcription factor EC
Tcn2	21452	1.37	transcobalamin 2
Tdh	58865	1.76	L-threonine dehydrogenase; predicted gene 13929
Tex11	83558	1.48	testis expressed gene 11
Tex14	83560	1.16	testis expressed gene 14
Thbd	21824	1.06	thrombomodulin
Tm6sf2	107770	1.17	transmembrane 6 superfamily member 2
Tmem116	77462	1.06	transmembrane protein 116
Tmem130	243339	1.04	transmembrane protein 130
Tmem144	70652	1.01	transmembrane protein 144
Tmprss2	50528	2.03	transmembrane protease, serine 2
Tnfrsf12a	27279	1.22	tumor necrosis factor receptor superfamily, member 12a
Tor1aip2	240832	1.01	torsin A interacting protein 2

Tram2	170829	1.04	translocating chain-associating membrane protein 2
Trap1a	22037	1.36	tumor rejection antigen P1A
Trim25	217069	1.04	tripartite motif-containing 25
Triml2	622117	2	tripartite motif family-like 2
Trpm6	225997	1.2	transient receptor potential cation channel, subfamily M, member 6
Tspan9	109246	1.36	tetraspanin 9
Ttc18	76670	1.6	tetratricopeptide repeat domain 18
Tuba3b	22147	2.3	predicted gene 5366; tubulin, alpha 3B; tubulin, alpha 3A
Ubash3b	72828	1.14	ubiquitin associated and SH3 domain containing, B
Upk3b	100647	1.59	uroplakin 3B
Usp26	83563	1.36	ubiquitin specific peptidase 26
Utf1	22286	1.45	undifferentiated embryonic cell transcription factor 1
Vamp8	22320	1.26	vesicle-associated membrane protein 8
Vkorc1	27973	1.06	vitamin K epoxide reductase complex, subunit 1
Wnk2	75607	1.06	WNK lysine deficient protein kinase 2
Xist	213742	4.99	inactive X specific transcripts
Xlr5a	574438	2.08	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Xlr5b	627081	2.16	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Ypel2	77864	1.56	yippee-like 2 (Drosophila)
Zbtb16	235320	1.79	zinc finger and BTB domain containing 16
Zfp42	22702	1.57	zinc finger protein 42

Downregulated genes in dnmt1-/- EBs compared to wt EBs between d4-16			
Gene symbol	**Entrez Gene ID**	**Fold change compared to wt**	**Gene name**
1700008I05Rik	71841	-1.45	RIKEN cDNA 1700008I05 gene
3110007F17Rik	73061	-1.06	predicted gene 5945; RIKEN cDNA 3110007F17 gene; predicted gene 2411; predicted gene 5167; predicted gene 6604; predicted gene 14957
Arhgap36	75404	-1.72	Rho GTPase activating protein 36; Synonym 1100001E04Rik
Car3	12350	-1.77	carbonic anhydrase 3
Cth	107869	-1.28	cystathionase (cystathionine gamma-lyase)
Cxxc1	72865	-1.13	CAAX box 1 homolog C (human)
Dclk1	13175	-1.46	doublecortin-like kinase 1
Dmrta1	242523	-1.04	doublesex and mab-3 related transcription factor like family A1
Dub1	13531	-1.29	deubiquitinating enzyme 1; similar to DUB-1
Efna2	13637	-1.1	ephrin A2
Epor	13857	-1.07	erythropoietin receptor
Exoc3l	277978	-1.82	exocyst complex component 3-like
Frat1	14296	-1.17	frequently rearranged in advanced T-cell lymphomas
Gfra3	14587	-1.29	glial cell line derived neurotrophic factor family receptor alpha 3
Gm12387	621880	-1.06	predicted gene 12387
Gpr165	76206	-1.14	G protein-coupled receptor 165
Gria4	14802	-1.14	glutamate receptor, ionotropic, AMPA4 (alpha 4)
Khk	16548	-1.02	ketohexokinase
L1cam	16728	-1.02	L1 cell adhesion molecule
Lmo2	16909	-1.13	LIM domain only 2
Mesp1	17292	-2.3	mesoderm posterior 1
Mesp2	17293	-1.4	mesoderm posterior 2
Ncs1	14299	-1.31	neuronal calcium sensor 1, Freq
Nkd1	93960	-1.36	naked cuticle 1 homolog (Drosophila); similar to naked cuticle 1 homolog
Olfm1	56177	-1.04	olfactomedin 1
Olfr893	258333	-1.2	olfactory receptor 893
Pgm2l1	70974	-1.14	phosphoglucomutase 2-like 1
Rasgrp3	240168	-1.55	RAS, guanyl releasing protein 3
Samd3	268288	-1.17	sterile alpha motif domain containing 3
Slc38a5	209837	-1.08	solute carrier family 38, member 5
Slc39a10	227059	-1.07	solute carrier family 39 (zinc transporter), member 10

Tmem47	192216	-1.07	transmembrane protein 47
Wdr16	71860	-1.15	WD repeat domain 16

Upregulated genes in TKO EBs compared to wt EBs between d4-16			
Gene symbol	**Entrez Gene ID**	**Fold change compared to wt**	**Gene name**
1700013H16Rik	75514	2.57	RIKEN cDNA 1700013H16 gene
1700080O16Rik	74279	1.12	RIKEN cDNA 1700080O16 gene
2410004A20Rik	66991	1.26	RIKEN cDNA 2410004A20 gene
4632428N05Rik	74048	1.05	RIKEN cDNA 4632428N05 gene
Agpat9	231510	1.37	1-acylglycerol-3-phosphate O-acyltransferase 9
Aire	11634	1.47	autoimmune regulator (autoimmune polyendocrinopathy candidiasis ectodermal dystrophy)
Alpk3	116904	1.06	alpha-kinase 3
Anxa2	12306	2.5	similar to Annexin A2 (Annexin II) (Lipocortin II) (Calpactin I heavy chain) (Chromobindin-8) (p36) (Protein I) (Placental anticoagulant protein IV) (PAP-IV); annexin A2
As3mt	57344	1.1	arsenic (+3 oxidation state) methyltransferase
Ascl2	17173	1.28	achaete-scute complex homolog 2 (Drosophila)
Btnl7	195349	1.18	butyrophilin-like 7
Car7	12354	1.64	carbonic anhydrase 7
Cdyl2	75796	1.36	chromodomain protein, Y chromosome-like 2
Ceacam15	101434	2.53	carcinoembryonic antigen-related cell adhesion molecule 15
Ceacam9	26368	2.58	carcinoembryonic antigen-related cell adhesion molecule 9
Clic6	209195	2.17	chloride intracellular channel 6
Cln6	76524	1.28	ceroid-lipofuscinosis, neuronal 6
Cntn5	244682	1.65	contactin 5
Cpn1	93721	1.26	carboxypeptidase N, polypeptide 1
Crygd	12967	1.49	crystallin, gamma D
Ctsh	13036	2.26	cathepsin H
D14Ertd668e	219132	1.58	predicted gene 6907; predicted gene 6904; predicted gene 4902; DNA segment, Chr 14, ERATO Doi 668, expressed; PHD finger protein 11
Dazl	13164	2.57	deleted in azoospermia-like
Dkkl1	50722	1.63	dickkopf-like 1
Dmrt1	50796	1.2	doublesex and mab-3 related transcription factor 1
Dpep3	71854	1.22	dipeptidase 3
Dpp4	13482	1.97	dipeptidylpeptidase 4
Dppa3	73708	1.37	developmental pluripotency-associated 3; predicted gene 6269
Dysf	26903	1.1	dysferlin
Efnb1	13641	1.07	ephrin B1
Gata3	14462	1.51	GATA binding protein 3
Gjb3	14620	1.87	gap junction protein, beta 3
Gjb5	14622	1.51	gap junction protein, beta 5
Gm10439	382243	2.02	predicted gene, OTTMUSG00000018964; predicted gene, OTTMUSG00000019138; predicted gene 15097; predicted gene 15085; predicted gene 15093; predicted gene 15109; predicted gene 15114; predicted gene 15128; ovary testis transcribed; similar to ovary testis transcribed; novel protein similar to ovary testis transcribed (Ott)
Gm13498	227885	1.62	predicted gene 13498
Gm2889	100040658	1.46	hypothetical protein LOC100041609; hypothetical protein LOC100044795; predicted gene 3395; similar to gag polyprotein; hypothetical protein LOC100047557; hypothetical protein LOC100040347; hypothetical protein LOC100044384; hypothetical protein LOC100045342; hypothetical protein LOC100038979; predicted gene 2889
Gm773	331416	2.48	predicted gene 773
Gm9	194854	1.5	predicted gene 9
Got1l1	76615	1.09	glutamic-oxaloacetic transaminase 1-like 1
Gramd1c	207798	1.5	GRAM domain containing 1C
Grin1	14810	1.63	glutamate receptor, ionotropic, NMDA1 (zeta 1)
Hormad2	75828	1.34	HORMA domain containing 2
Htatip2	53415	1.04	HIV-1 tat interactive protein 2, homolog (human)
Igfbp6	16012	1.22	insulin-like growth factor binding protein 6
Inpp5d	16331	1.28	inositol polyphosphate-5-phosphatase D
Klf4	16600	1.04	Kruppel-like factor 4 (gut)
Krt8	16691	1.97	predicted gene 5604; keratin 8
Luzp4	434865	2.34	predicted gene, OTTMUSG00000019001; leucine zipper protein 4

Ly6a	110454	2.42	lymphocyte antigen 6 complex, locus A
Ly75	17076	1.25	lymphocyte antigen 75
Magea1	17137	2.01	melanoma antigen, family A, 1
Magea2	17138	1.8	melanoma antigen, family A, 2
Magea6	17142	1.62	melanoma antigen, family A, 6
Me1	17436	1.03	predicted gene 7049; similar to NADP-dependent malic enzyme (NADP-ME) (Malic enzyme 1); malic enzyme 1, NADP(+)-dependent, cytosolic
Morc1	17450	1.31	microrchidia 1
Nxf7	170722	2.21	nuclear RNA export factor 7
Olfr1054	259021	1.06	olfactory receptor 1054
Ott	18422	2.211111111	predicted gene, OTTMUSG00000018964; predicted gene, OTTMUSG00000019138; predicted gene 15097; predicted gene 15085; predicted gene 15093; predicted gene 15109; predicted gene 15114; predicted gene 15128; ovary testis transcribed; similar to ovary testis transcribed; novel protein similar to ovary testis transcribed (Ott)
Pgc	109820	2.22	progastricsin (pepsinogen C)
Phf16	382207	1.11	PHD finger protein 16
Plac1	56096	2.9	placental specific protein 1
Pnma5	385377	1.28	paraneoplastic antigen family 5
Prl2c1	666317	3.14	Prolactin family 2, subfamily c, member 1
Prl3d1	18775	3.77	prolactin family 3, subfamily d, member 1
Prl5a1	28078	2.81	prolactin family 5, subfamily a, member 1
Prl7a1	19113	3.2	prolactin family 7, subfamily a, member 1
Prss42	235628	1.1	protease serine 42; old symbol Tessp2
Rbm44	329207	1.41	RNA binding motif protein 44
Rbmy1a1	19657	1.723333333	RNA binding motif protein, Y chromosome, family 1, member A1
Sbsn	282619	1.13	suprabasin
Scml2	107815	1.44	similar to sex comb on midleg-like 2 (Drosophila); sex comb on midleg-like 2 (Drosophila)
Serpinb1c	380839	1.33	serine (or cysteine) peptidase inhibitor, clade B, member 1c
Sfmbt2	353282	1.88	Scm-like with four mbt domains 2
Shroom1	71774	1.08	shroom family member 1
Slc13a4	243755	1.99	solute carrier family 13 (sodium/sulfate symporters), member 4
Slc25a20	57279	1.15	solute carrier family 25 (mitochondrial carnitine/acylcarnitine translocase), member 20
Slc6a4	15567	1.21	solute carrier family 6 (neurotransmitter transporter, serotonin), member 4
Smc1b	140557	2.65	structural maintenance of chromosomes 1B
Smtnl2	276829	1.24	smoothelin-like 2
Sohlh2	74434	1.94	spermatogenesis and oogenesis specific basic helix-loop-helix 2
Sox15	20670	1.13	SRY-box containing gene 16; SRY-box containing gene 15
Stag3	50878	1.33	stromal antigen 3
Stra8	20899	1.82	stimulated by retinoic acid gene 8
Tcfap2c	21420	1.77	transcription factor AP-2, gamma
Tex14	83560	1.5	testis expressed gene 14
Tmem231	234740	1.05	transmembrane protein 231; synonym 4932417I16Rik
Tmprss2	50528	1.84	transmembrane protease, serine 2
Ttc18	76670	2.45	tetratricopeptide repeat domain 18
Tuba3b	22147	1.12	predicted gene 5366; tubulin, alpha 3B; tubulin, alpha 3A
Usp26	83563	2.19	ubiquitin specific peptidase 26
Xist	213742	2.78	inactive X specific transcripts
Xlr5a	574438	2.31	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Xlr5b	627081	2.42	X-linked lymphocyte-regulated 5B; X-linked lymphocyte-regulated 5D, pseudogene; X-linked lymphocyte-regulated 5A; X-linked lymphocyte-regulated 5E, pseudogene; predicted gene, EG667719
Zbtb10	229055	1.41	zinc finger and BTB domain containing 10
Zbtb16	235320	1.13	zinc finger and BTB domain containing 16
Zc3h6	78751	1.19	zinc finger CCCH type containing 6

Downregulated genes in TKO EBs compared to wt EBs between d4-16			
Gene symbol	**Entrez Gene ID**	**Fold change compared to wt**	**Gene name**
A130022J15Rik	101351	-1.15	RIKEN cDNA A130022J15 gene
Aadat	23923	-1.01	aminoadipate aminotransferase

Adamts15	235130	-1.35	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 15
Adamts6	108154	-1.5	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 6
Aff3	16764	-1.15	AF4/FMR2 family, member 3; similar to AF4/FMR2 family member 3 (LAF-4 protein) (Lymphoid nuclear protein related to AF4)
AI504432	229694	-1.17	expressed sequence AI504432
Antxr1	69538	-1.04	anthrax toxin receptor 1
Apba2	11784	-1.07	amyloid beta (A4) precursor protein-binding, family A, member 2
Arrb1	109689	-1.1	arrestin, beta 1
Bend3	331623	-1.03	BEN domain containing 3
Cachd1	320508	-1.55	cache domain containing 1; similar to Cache domain containing 1
Car3	12350	-1.64	carbonic anhydrase 3
Cask	12361	-1.29	calcium/calmodulin-dependent serine protein kinase (MAGUK family)
Casp3	12367	-1.09	caspase 3
Ccdc85a	216613	-1.08	coiled-coil domain containing 85A
Ccl20	20297	-1.03	chemokine (C-C motif) ligand 20
Cdh10	320873	-1.07	cadherin 10
Cdh11	12552	-1.22	cadherin 11
Cdh9	12565	-1.14	cadherin 9
Coro1a	12721	-1.27	coronin, actin binding protein 1A
Cpe	12876	-1.54	carboxypeptidase E; similar to carboxypeptidase E
Cxx1b	553127	-1.19	CAAX box 1 homolog A (human); CAAX box 1 homolog B (human); similar to mammalian retrotransposon derived 8b
Cxx1c	72865	-1.62	CAAX box 1 homolog C (human)
Cyp2j6	13110	-1.12	cytochrome P450, family 2, subfamily j, polypeptide 6
Cysltr1	58861	-1.29	cysteinyl leukotriene receptor 1
Cyyr1	224405	-1.19	cysteine and tyrosine-rich protein 1
Dab1	13131	-1.19	disabled homolog 1 (Drosophila)
Dapk1	69635	-1.01	death associated protein kinase 1
Dbn1	56320	-1.11	drebrin 1
Dclk1	13175	-1.55	doublecortin-like kinase 1
Ebf1	13591	-1.915	early B-cell factor 1
Elovl4	83603	-1	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 4
Enpp2	18606	-2.1	ectonucleotide pyrophosphatase/phosphodiesterase 2
Etv1	14009	-1.725	ets variant gene 1
Etv6	14011	-1.18	ets variant gene 6 (TEL oncogene)
Exoc3l	277978	-2.3	exocyst complex component 3-like
Fabp7	12140	-1.78	fatty acid binding protein 7, brain
Fam115c	232748	-1.16	family with sequence similarity 115, member C
Fam171b	241520	-1.85	family with sequence similarity 171, member B
Fam19a4	320701	-1.42	family with sequence similarity 19, member A4
Fam38b	667742	-1.29	family with sequence similarity 38, member B2
Fat3	270120	-1.2525	FAT tumor suppressor homolog 3 (Drosophila)
Foxi3	232077	-1.23	forkhead box I3
Fst	14313	-3.1	follistatin
Fzd3	14365	-1.06	frizzled homolog 3 (Drosophila)
Gabra3	14396	-1.5	gamma-aminobutyric acid (GABA) A receptor, subunit alpha 3
Gria4	14802	-1.47	glutamate receptor, ionotropic, AMPA4 (alpha 4); hypothetical protein LOC100044208
Gtf3c6	67371	-1.12	general transcription factor IIIC, polypeptide 6, alpha
Gucy1a2	234889	-1.25	hypothetical protein LOC100044212; guanylate cyclase 1, soluble, alpha 2
Hmcn1	545370	-1.351666667	hemicentin 1
Hunk	26559	-1.13	similar to putative serine/threonine protein kinase MAK-V; similar to hormonally upregulated Neu-associated kinase; hormonally upregulated Neu-associated kinase
Ids	15931	-1	iduronate 2-sulfatase
Jakmip2	76217	-1.64	janus kinase and microtubule interacting protein 2
Jam3	83964	-1.08	junction adhesion molecule 3
Kcng3	225030	-1	potassium voltage-gated channel, subfamily G, member 3
Kcnh8	211468	-1.15	potassium voltage-gated channel, subfamily H (eag-related), member 8
Kif3c	16570	-1.1	kinesin family member 3C
Lpar4	78134	-2.26	lysophosphatidic acid receptor 4
Lphn2	99633	-1.64	latrophilin 2

Magel2	27385	-1.81	melanoma antigen, family L, 2
Mapk4	225724	-1.04	mitogen-activated protein kinase 4
Mapre2	212307	-1.24	microtubule-associated protein, RP/EB family, member 2
Mcc	328949	-1.33	mutated in colorectal cancers
Mkx	210719	-1.05	mohawk homeobox
Mmp25	240047	-1.81	matrix metallopeptidase 25
Naalad2	72560	-1.37	N-acetylated alpha-linked acidic dipeptidase 2
Nefl	18039	-2.71	neurofilament, light polypeptide
Neto2	74513	-2.1	neuropilin (NRP) and tolloid (TLL)-like 2
Nova1	664883	-1.19	neuro-oncological ventral antigen 1
Nrarp	67122	-1.1	Notch-regulated ankyrin repeat protein
Nudt19	110959	-1.03	nudix (nucleoside diphosphate linked moiety X)-type motif 19
Parp8	52552	-1.61	poly (ADP-ribose) polymerase family, member 8
Pbx1	18514	-1.05	pre B-cell leukemia transcription factor 1; region containing RIKEN cDNA 2310056B04 gene; pre B-cell leukemia transcription factor 1
Pde2a	207728	-1.2	phosphodiesterase 2A, cGMP-stimulated
Pde4d	238871	-1.14	phosphodiesterase 4D, cAMP specific
Pde5a	242202	-1.21	phosphodiesterase 5A, cGMP-specific
Pgap1	241062	-1.34	post-GPI attachment to proteins 1
Plch1	269437	-1.06	phospholipase C, eta 1
Plxdc2	67448	-1.24	plexin domain containing 2
Ppp4r4	74521	-2.66	protein phosphatase 4, regulatory subunit 4
Prom1	19126	-1.2	prominin 1
Prps2	110639	-1.06	phosphoribosyl pyrophosphate synthetase 2
Ptk7	71461	-1.14	PTK7 protein tyrosine kinase 7
Rab38	72433	-1.13	RAB38, member of RAS oncogene family
Rasgrp3	240168	-1.48	RAS, guanyl releasing protein 3
Rragd	52187	-1.09	Ras-related GTP binding D
Rxrg	20183	-1.44	retinoid X receptor gamma
Sall2	50524	-1.22	sal-like 2 (Drosophila)
Sall3	20689	-1.02	sal-like 3 (Drosophila)
Sema6a	20358	-1.84	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A
Shc4	271849	-1.31	SHC (Src homology 2 domain containing) family, member 4
Slc27a2	26458	-2.02	solute carrier family 27 (fatty acid transporter), member 2
Slc35f1	215085	-1.11	solute carrier family 35, member F1
Slco5a1	240726	-1.04	solute carrier organic anion transporter family, member 5A1
Smarca1	93761	-1.43	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1
Socs2	216233	-1.34	suppressor of cytokine signaling 2; predicted gene 8000
Sox4	20677	-1	SRY-box containing gene 19; SRY-box containing gene 4
Sp8	320145	-1.07	trans-acting transcription factor 8
Spry2	24064	-1.43	sprouty homolog 2 (Drosophila)
Sssca1	56390	-1.155	Sjogren's syndrome/scleroderma autoantigen 1 homolog (human)
St6gal2	240119	-2.08	beta galactoside alpha 2,6 sialyltransferase 2
Stox2	71069	-1.21	storkhead box 2
Sulf1	240725	-1.66	sulfatase 1
Syngap1	240057	-1.25	synaptic Ras GTPase activating protein 1 homolog (rat)
Syt11	229521	-1.3	synaptotagmin XI; similar to synaptotagmin XI
T	20997	-5.13	brachyruy
Thbs3	21827	-1.18	thrombospondin 3
Tmem47	192216	-1.04	transmembrane protein 47
Tmem90b	433485	-1.41	transmembrane protein 90B; old symbol Gm14134
Trpa1	277328	-1.09	transient receptor potential cation channel, subfamily A, member 1
Tspan6	56496	-1.74	tetraspanin 6
Wnt8a	20890	-2.3	wingless-related MMTV integration site 8A
Zdhhc15	108672	-1.06	zinc finger, DHHC domain containing 15
Zfp521	225207	-1.04	zinc finger protein 521
Zfp608	269023	-1.19	zinc finger protein 608
Zfp946	74149	-1.18	RIKEN cDNA 1300003B13 gene; hypothetical protein LOC100044281, old symbol 1300003B13Rik
Zfp882	382019	-1.48	zinc finger protein 882

Concordantly upregulated genes in wt, dnmt1-/- and TKO Ebs between d4-16					
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change dnmt1-/-**	**Fold change TKO**	**Gene name**
1110003E01Rik	482771	1.44	1.25	1.06	RIKEN cDNA 1110003E01 gene
1600029D21Rik	425870	2.68	2.04	2.27	RIKEN cDNA 1600029D21 gene
A2m	445751	2.26	2.19	2.1	alpha-2-macroglobulin
Abcc2	451709	2.68	2.91	2.08	ATP-binding cassette, sub-family C (CFTR/MRP), member 2
Afp	452310	7.45	6.91	4.42	alpha fetoprotein
Akap2	424354	1.2	1.085	1.02	A kinase (PRKA) anchor protein 2; paralemmin 2
Aldh3b2	482013	1.68	1.06	1.5	aldehyde dehydrogenase 3 family, member B2; RIKEN cDNA 1700055N04 gene
Anxa1	425879	4.82	2.15	1.28	annexin A1
Apoa1	431055	5.81	5.68	4.13	apolipoprotein A-I
Apoa2	483037	3.72	5.14	2.44	apolipoprotein A-II
Apob	447688	4.82	5.1	3.61	apolipoprotein B
Apoc2	461471	3.26	4.65	3.16	apolipoprotein C-II
Apom	460325	2.39	5.32	3.67	apolipoprotein M
Aqp3	476854	1.98	2.56	1.83	aquaporin 3
Aqp8	425939	3.37	4.15	2.57	aquaporin 8
Arhgef6	421496	4.29	4.06	3.54	Rac/Cdc42 guanine nucleotide exchange factor (GEF) 6
Atp11a	433815	2.35	1.66	1.2	ATPase, class VI, type 11A
B4galnt2	430216	2.26	4.28	2.66	beta-1,4-N-acetyl-galactosaminyl transferase 2
Cd68	462582	2.76	1.92	1.83	CD68 antigen
Cd82	451736	3.05	2.69	2.42	CD82 antigen
Cdh5	446981	2.85	2.29	1.36	cadherin 5
Cdhr2	268663	2.6	3.25	1.51	cadherin-related family member 2
Cdkn1c	475371	1.62	2.8	1.94	cyclin-dependent kinase inhibitor 1C (P57)
Cfb	425397	3.17	3.12	1.06	complement factor B
Cfi	432302	2.38	2.87	1.66	complement component factor i
Clu	443598	3.67	2.84	1.3	similar to clusterin; clusterin
Csf1r	430511	3.41	1.95	1.27	colony stimulating factor 1 receptor
Ctsb	477096	2.47	2.01	1.25	cathepsin B
Ctsj	468291	3.51	3.89	2.43	cathepsin J
Ctsz	441359	1.08	1.85	1.09	cathepsin Z
Cubn	474472	3.76	5.01	3.16	cubilin (intrinsic factor-cobalamin receptor)
Dab2	448546	2.22	3.59	2.36	disabled homolog 2 (Drosophila)
Dmrtc1b	432258	1.13	1.24	1.22	DMRT-like family C1b
Dram1	426028	1.12	2.28	1.16	RIKEN cDNA 1200002N14 gene
Enpep	444098	2.86	3.34	1.18	glutamyl aminopeptidase
Entpd1	452537	1.67	1.85	1.99	ectonucleoside triphosphate diphosphohydrolase 1
Ephx1	460565	2.97	1.41	1.07	epoxide hydrolase 1, microsomal
Eps8l2	425066	1.77	1.66	1.13	EPS8-like 2
Ets2	428496	1.59	1.39	1.34	E26 avian leukemia oncogene 2, 3' domain
F2	448405	2.9	3.71	1.45	coagulation factor II
Fabp3	422118	1.71	2.66	1.815	fatty acid binding protein 3, muscle and heart; similar to mammary-derived growth inhibitor
Fgb	430358	5.52	5.53	3.04	fibrinogen beta chain
Fndc3b	456688	1.65	1.75	1.68	fibronectin type III domain containing 3B
Folr1	430285	1.82	2.9	1.29	folate receptor 1 (adult)
Fstl3	425869	1.59	1.48	1.69	follistatin-like 3
Gc	449785	4.62	3.49	1.3	group specific component
Gjb2	432541	2.9	2.45	2.67	gap junction protein, beta 2
Gm2a	439285	2.43	2.2	1.89	GM2 ganglioside activator protein
Gpr116	446182	2.93	2.18	1.95	G protein-coupled receptor 116
Gpr126	462385	2.83	1.69	1.28	G protein-coupled receptor 126
Gsn	438270	2.13	2.27	2	gelsolin
Gstm3	434159	4.325	2.37	1.48	glutathione S-transferase, mu 3
Habp2	448617	2.3	3.45	1.41	hyaluronic acid binding protein 2

Hkdc1	471239	1.86	2.96	1.61	hexokinase domain containing 1
Hnf4a	442159	1.8	2.25	1.28	hepatic nuclear factor 4, alpha
Hpgd	456111	2.82	1.45	1.21	hydroxyprostaglandin dehydrogenase 15 (NAD)
Hsd17b2	427962	4.01	3.33	2.31	hydroxysteroid (17-beta) dehydrogenase 2
Igfbp1	467817	5.1	5.04	1.31	insulin-like growth factor binding protein 1
Itgb3	461496	1.41	1.26	2.12	integrin beta 3
Kdelr3	441810	2.33	2.48	1.47	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3
Klb	426337	1.21	2.89	1.15	klotho beta
Krt14	481757	1.94	1.77	2.15	keratin 14
Krt19	448028	2.23	1.77	2.48	keratin 19
Krt7	433153	4.01	2.25	2.68	keratin 7
Lgals2	429829	2.26	4.99	2.82	lectin, galactose-binding, soluble 2
Lgals3	434018	2.36	1.77	1.42	lectin, galactose binding, soluble 3
Lgals9	452854	4.72	3.44	2.65	lectin, galactose binding, soluble 9
Lgmn	446275	1.58	2.55	1.93	legumain
Mbnl2	457193	1.8	1.92	1.34	muscleblind-like 2
Mfge8	455998	1.24	1.56	1.51	milk fat globule-EGF factor 8 protein
Mia2	433125	2.6	3.91	2.75	melanoma inhibitory activity 2
Mttp	433105	1.63	3.67	2.34	microsomal triglyceride transfer protein
Muc13	455791	3.29	3.07	1.6	mucin 13, epithelial transmembrane
Mvp	429449	1.73	1.85	1.34	major vault protein
Myl4	447204	3.21	2.86	1.62	myosin, light polypeptide 4
Myof	478519	2.59	2.31	1.3	myoferlin
Ndrg1	465817	2.29	1.68	1.98	N-myc downstream regulated gene 1
Nid1	478950	2.25	1.86	1.35	similar to Nidogen precursor (Entactin); nidogen 1; similar to Nid1 protein
Nostrin	464009	1.45	3.15	2.1	nitric oxide synthase trafficker
Nqo1	481143	3.8	2.75	1.56	NAD(P)H dehydrogenase, quinone 1
Nr2f2	434891	4.37	1.98	1.24	similar to COUP-TFI; nuclear receptor subfamily 2, group F, member 2
Nrk	479394	2.8	3.04	2.98	Nik related kinase
P2rx4	463310	1.2	2.55	1.23	purinergic receptor P2X, ligand-gated ion channel 4
Pcbd1	436208	1.35	2.41	1.6	pterin 4 alpha carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha (TCF1) 1
Pde8a	435616	1.18	1.42	1.07	phosphodiesterase 8A; similar to cAMP-specific cyclic nucleotide phosphodiesterase PDE8; MMPDE8
Pla2g12b	466028	1.49	2.35	1.56	phospholipase A2, group XIIB
Pparg	459227	2.03	1.15	1.55	peroxisome proliferator activated receptor gamma
Prl2a1	458150	5.01	4.87	3.54	prolactin family 2, subfamily a, member 1
Prl2b1	437269	3.44	2.71	1.76	prolactin family 2, subfamily b, member 1
Prl2c3	425724	2.11	4.69	2.36	prolactin family 2, subfamily c, member 3; prolactin family 2, subfamily c, member 4
Prl2c5	437114	6.04	6.99	5.09	prolactin family 2, subfamily c, member 5
Prl4a1	453799	1.08	3.38	2.85	prolactin family 4, subfamily a, member 1
Prl7b1	459814	3.58	4.4	2.95	prolactin family 7, subfamily b, member 1
Prl7d1	464374	4.2	5.17	4	prolactin family 7, subfamily d, member 1
Prl8a9	428579	4.58	3.61	2.03	prolactin family8, subfamily a, member 9
Prr13	424501	1.3	1.55	1.1	proline rich 13
Rassf4	423866	1.28	1.44	1.01	Ras association (RalGDS/AF-6) domain family member 4
Rbp4	433458	4.6	6.09	4.52	retinol binding protein 4, plasma
Reep6	469675	1.48	1.79	1.05	receptor accessory protein 6
Rhox6	429981	2.98	2.13	2.38	reproductive homeobox 6
Rhox9	445775	3.57	2.68	2.55	reproductive homeobox 9
Sct	423804	2.28	2.59	2.25	secretin
Serpinb9d	478140	1.7	2.57	2.12	serine (or cysteine) peptidase inhibitor, clade B, member 9d
Serpinb9e	424972	5.61	4.96	4.24	serine (or cysteine) peptidase inhibitor, clade B, member 9e
Serpinb9f	425674	4.36	4.19	3.53	serine (or cysteine) peptidase inhibitor, clade B, member 9f
Serpinb9g	442453	4.3	4.12	3.415	serine (or cysteine) peptidase inhibitor, clade B, member 9g
Serpine1	422178	2.53	3.24	2.85	serine (or cysteine) peptidase inhibitor, clade E, member 1
Serpinf2	479980	3.82	3.39	1.58	serine (or cysteine) peptidase inhibitor, clade F, member 2
Slc1a1	451162	1.24	1.46	1.46	solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1
Slc34a2	440546	2.25	3.23	2.24	solute carrier family 34 (sodium phosphate), member 2

Slc44a4	457306	1.62	1.87	1.03	solute carrier family 44, member 4
Slc7a1	475083	1.15	1.14	1.34	solute carrier family 7 (cationic amino acid transporter, y+ system), member 1
Slc7a9	456728	1.55	3.25	1.54	solute carrier family 7 (cationic amino acid transporter, y+ system), member 9
Slco2a1	475464	2.99	3.62	1.9	solute carrier organic anion transporter family, member 2a1
Soat2	464985	1.29	2.34	1.61	sterol O-acyltransferase 2
Tfpi	464234	1.98	1.73	1.21	tissue factor pathway inhibitor
Tgm2	437315	3.44	3.05	2.5	transglutaminase 2, C polypeptide
Tinagl1	480993	3.27	3.03	2.24	tubulointerstitial nephritis antigen-like 1
Tmem106a	433088	1.35	1.66	1.04	transmembrane protein 106A
Tnnc1	435600	1.4	1.99	1.08	troponin C, cardiac/slow skeletal
Tnnt2	454486	3.42	2.26	1.08	troponin T2, cardiac
Trf	433235	6.81	6.09	4.07	transferrin
Tspan8	475891	3.96	3.07	1.41	tetraspanin 8
Ttr	439049	4.83	5.14	3.68	transthyretin
Txnip	468587	1.57	1.13	1.22	thioredoxin interacting protein
Vegfa	480917	1.86	1.59	1.52	vascular endothelial growth factor A
Vil1	444016	1.82	2.2	1.22	villin 1
Vill	464133	1.64	1.17	1.25	villin-like
Vnn1	436866	4.12	3.62	2.3	vanin 1

Concordantly downregulated genes in wt, dnmt1-/- and TKO Ebs between d4-16					
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change dnmt1-/-**	**Fold change TKO**	**Gene name**
7420416P09Rik	432677	-3.31	-1.09	-2.02	RIKEN cDNA 7420416P09 gene
Alox15	11687	-2.77	-2.95	-2.62	arachidonate 15-lipoxygenase
Amot	27494	-1.89	-1.29	-1.17	angiomotin
Asb4	65255	-1.41	-2.74	-1.47	ankyrin repeat and SOCS box-containing 4
Axin2	12006	-1.1	-1.02	-1.36	axin2
Dgkk	331374	-1.94	-1.07	-1.18	diacylglycerol kinase kappa
Dpysl5	65254	-1.42	-1.16	-1.04	dihydropyrimidinase-like 5
Eomes	13813	-3.34	-2.45	-2.75	eomesodermin homolog (Xenopus laevis)
Etv2	14008	-1.44	-1.91	-1.78	similar to ETS related protein 71; ets variant gene 2
Fgf15	14170	-2.48	-1.21	-1.37	fibroblast growth factor 15
Hdx	245596	-1.32	-1.26	-1.45	highly divergent homeobox
Ina	226180	-1.29	-1.09	-1.47	internexin neuronal intermediate filament protein, alpha
Mixl1	27217	-2.48	-2.58	-3.04	Mix1 homeobox-like 1 (Xenopus laevis)
Phf6	70998	-1.11	-1.17	-1.2	PHD finger protein 6
Trh	22044	-2.22	-1.15	-2.06	thyrotropin releasing hormone

Concordantly upregulated genes in wt and dnmt1-/- EBs between d4-16				
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change dnmt1-/-**	**Gene name**
0610010O12Rik	66060	1.08	1.54	RIKEN cDNA 0610010O12 gene
1110036O03Rik	66180	1.38	1.17	RIKEN cDNA 1110036O03 gene
1200009I06Rik	74190	1.08	1.27	RIKEN cDNA 1200009I06 gene
1300017J02Rik	71775	1.65	1.72	RIKEN cDNA 1300017J02 gene
1810011O10Rik	69068	2.99	1.48	RIKEN cDNA 1810011O10 gene
2610528J11Rik	66451	1.95	2.43	RIKEN cDNA 2610528J11 gene
2810459M11Rik	72792	1.04	1.59	RIKEN cDNA 2810459M11 gene
4930506M07Rik	71653	1.42	1.09	RIKEN cDNA 4930506M07 gene
8430408G22Rik	213393	2.75	1.72	RIKEN cDNA 8430408G22 gene
9530068E07Rik	213673	1.62	1.06	RIKEN cDNA 9530068E07 gene
A1cf	69865	1.2	1.24	APOBEC1 complementation factor
A430107O13Rik	214642	2.46	1.22	RIKEN cDNA A430107O13 gene
Abca1	11303	2.87	1.82	ATP-binding cassette, sub-family A (ABC1), member 1
Abca3	27410	1.03	1.28	ATP-binding cassette, sub-family A (ABC1), member 3
Acox1	11430	1.52	1.25	acyl-Coenzyme A oxidase 1, palmitoyl

Acsl1	14081	1.42	1.14	acyl-CoA synthetase long-chain family member 1
Acta2	11475	3.85	2.56	actin, alpha 2, smooth muscle, aorta
Actc1	11464	5.01	3.88	actin, alpha, cardiac muscle 1; similar to alpha-actin (AA 27-375)
Adamts1	11504	2.79	1.67	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 1
Adh1	11522	3.7	1.35	alcohol dehydrogenase 1 (class I)
Adm	11535	2.25	1.74	adrenomedullin
Afap1l2	226250	1.41	1.38	actin filament associated protein 1-like 2
Afm	280662	2.33	1.56	afamin
Aga	11593	1.44	1.46	aspartylglucosaminidase
Ahnak2	100041194	2.02	1.27	AHNAK nucleoprotein 2; similar to Unknown (protein for IMAGE:3599271)
Ahsg	11625	6.56	2.53	alpha-2-HS-glycoprotein
Akr1c12	622402	1.48	1.01	aldo-keto reductase family 1, member C12
Akr1c19	432720	3.01	1.48	aldo-keto reductase family 1, member C19
Alb	11657	7.84	4.96	albumin
Aldh3a2	11671	1.33	1.15	aldehyde dehydrogenase family 3, subfamily A2
Aldob	230163	3.11	1.8	aldolase B, fructose-bisphosphate
Ambp	11699	5.34	3.21	alpha 1 microglobulin/bikunin
Ang	11727	1.43	1.9	angiogenin, ribonuclease, RNase A family, 5
Angpt1	11600	1.87	1.38	angiopoietin 1
Ank	11732	1.34	1.07	progressive ankylosis
Ankrd1	107765	3.1	2.24	ankyrin repeat domain 1 (cardiac muscle)
Ano1	101772	1.95	1.19	anoctamin 1, calcium activated chloride channel
Anxa4	11746	1.35	1.27	annexin A4
Apoa4	11808	2.75	3.2	apolipoprotein A-IV
App	11820	1.2	1.11	amyloid beta (A4) precursor protein
Arhgap10	78514	1.2	1.16	Rho GTPase activating protein 10
Arhgap18	73910	1.59	1.36	Rho GTPase activating protein 18
Arhgef3	71704	1.57	1.36	Rho guanine nucleotide exchange factor (GEF) 3
Arid5b	71371	1.37	1.72	similar to modulator recognition factor 2; AT rich interactive domain 5B (MRF1-like)
Asah2	54447	1.04	1.16	N-acylsphingosine amidohydrolase 2
Ass1	11898	1.16	1.02	argininosuccinate synthetase 1
Atf3	11910	1.78	1.12	activating transcription factor 3
Atp1b1	11931	3.57	2.44	ATPase, Na+/K+ transporting, beta 1 polypeptide
Atp8b1	54670	3.21	1.07	ATPase, class I, type 8B, member 1
B2m	12010	1.93	1.61	beta-2 microglobulin
Baiap2l1	66898	1.13	1.16	BAI1-associated protein 2-like 1
BC025446	223631	1.5	2.3	cDNA sequence BC025446
BC037703	242125	1.71	1.32	cDNA sequence BC037703
Bgn	12111	4.74	2.99	biglycan
Bicc1	83675	2.75	1.23	bicaudal C homolog 1 (Drosophila)
Bmp2k	140780	1.21	1.03	predicted gene 4521; BMP2 inducible kinase
Bnc2	242509	1.85	1.08	basonuclin 2
Bst1	12182	1.39	1.97	bone marrow stromal cell antigen 1
Btg1	12226	1.49	1.2	B-cell translocation gene 1, anti-proliferative; similar to myocardial vascular inhibition factor
C1s	50908	2.33	2.09	similar to Complement component 1, s subcomponent; complement component 1, s subcomponent
C3	12266	3.34	1.55	complement component 3; similar to complement component C3 prepropeptide, last
C3ar1	12267	4.61	1.55	complement component 3a receptor 1
Calml4	75600	2.17	1.63	calmodulin-like 4
Ccdc3	74186	2.46	1.09	coiled-coil domain containing 3
Ccdc68	381175	1.8575	1.37	coiled-coil domain containing 68
Ccdc141	545428	2.31	2.35	coiled-coil domain containing 141
Ccl9	20308	1.71	1.02	chemokine (C-C motif) ligand 9
Cd9	12527	1.14	1.4	CD9 antigen
Cdcp1	109332	1.13	1.11	CUB domain containing protein 1
Cdh3	12560	2.09	1.96	cadherin 3
Cdh6	12563	1.25	1.29	cadherin 6
Cdo1	12583	1.71	1.34	cysteine dioxygenase 1, cytosolic

Chst15	77590	1.28	1.16	carbohydrate (N-acetylgalactosamine 4-sulfate 6-O) sulfotransferase 15
Cldn18	56492	2.76	1.89	claudin 18
Cldn2	12738	1.49	2.61	claudin 2
Cldn4	12740	2.4	1.22	claudin 4
Cldn7	53624	1.15	1.3	claudin 7
Clic5	224796	1.34	1.17	chloride intracellular channel 5
Cmbl	69574	1.43	1.21	carboxymethylenebutenolidase-like (Pseudomonas)
Col12a1	12816	2.89	1.3	collagen, type XII, alpha 1
Col14a1	12818	3.11	1.15	collagen, type XIV, alpha 1
Col1a1	12842	3.21	2.21	collagen, type I, alpha 1
Col1a2	12843	5.13	2.86	collagen, type I, alpha 2
Col3a1	12825	5.26	3	collagen, type III, alpha 1
Col4a1	12826	1.31	2.55	collagen, type IV, alpha 1
Col4a2	12827	1.47	2.76	collagen, type IV, alpha 2
Col5a2	12832	3.22	1.07	collagen, type V, alpha 2
Col6a1	12833	2.86	1.48	collagen, type VI, alpha 1
Col6a3	12835	2.73	1.25	collagen, type VI, alpha 3
Cp	12870	2.7	1.01	ceruloplasmin
Cpb2	56373	1.68	1.28	carboxypeptidase B2 (plasma)
Cpeb4	67579	1.34	1.05	cytoplasmic polyadenylation element binding protein 4
Cpm	70574	1.71	2.12	carboxypeptidase M
Creb3l2	208647	1.21	1.35	cAMP responsive element binding protein 3-like 2
Crim1	50766	2.37	2.68	cysteine rich transmembrane BMP regulator 1 (chordin like)
Crispld2	78892	1.89	1.01	cysteine-rich secretory protein LCCL domain containing 2
Cryab	12955	2.23	1.34	crystallin, alpha B
Ctsa	19025	1.31	1.47	cathepsin A
Ctsl	13039	1.08	1.63	cathepsin L
Ctsq	104002	3.02	2.93	cathepsin Q
Cxcl10	15945	2.37	1	chemokine (C-X-C motif) ligand 10; similar to Small inducible cytokine B10 precursor (CXCL10) (Interferon-gamma-induced protein CRG-2) (Gamma-IP10) (IP-10) (C7)
Cxcl12	20315	2.58	1.19	chemokine (C-X-C motif) ligand 12
Cybrd1	73649	1.61	1.85	cytochrome b reductase 1
Cyp11a1	13070	2.83	2.02	cytochrome P450, family 11, subfamily a, polypeptide 1
Cyp1b1	13078	2.92	1.62	cytochrome P450, family 1, subfamily b, polypeptide 1
Cyp2c65	72303	2.25	1.64	cytochrome P450, family 2, subfamily c, polypeptide 65
Cyp3a13	13113	1.48	1.51	cytochrome P450, family 3, subfamily a, polypeptide 13
D0H4S114	27528	2.85	1.04	DNA segment, human D4S114
Dap	223453	2.05	1.87	death-associated protein
Dcn	13179	6.97	4.9	decorin
Ddc	13195	2.44	1.59	dopa decarboxylase
Ddr2	18214	2.13	1.89	discoidin domain receptor family, member 2
Ddx58	230073	2.24	1.92	DEAD (Asp-Glu-Ala-Asp) box polypeptide 58
Dgat2	67800	1.01	1.36	diacylglycerol O-acyltransferase 2
Dhrs3	20148	3.59	1.27	dehydrogenase/reductase (SDR family) member 3
Dhrs9	241452	1.2	1.78	dehydrogenase/reductase (SDR family) member 9
Dio3	107585	1	1.68	deiodinase, iodothyronine type III
Dkk3	50781	1.67	1.06	dickkopf homolog 3 (Xenopus laevis)
Dlk1	13386	4.04	3.71	delta-like 1 homolog (Drosophila)
Dnm3os	474332	2.43	1.13	dynamin 3, opposite strand
Dsc2	13506	1.88	1.5	desmocollin 2
E130203B14Rik	320736	2.38	1.36	RIKEN cDNA E130203B14 gene
Ecm1	13601	1.05	1.22	extracellular matrix protein 1
Ednra	13617	1.82	1.42	endothelin receptor type A
Ednrb	13618	2.45	1.51	endothelin receptor type B
Elf1	13709	1.36	1.23	E74-like factor 1
Elovl7	74559	1.19	1.16	ELOVL family member 7, elongation of long chain fatty acids (yeast)
Emp1	13730	4.17	1.83	epithelial membrane protein 1
Emp2	13731	1.23	1.29	epithelial membrane protein 2
Enpp1	18605	1.61	1.38	ectonucleotide pyrophosphatase/phosphodiesterase 1

Enpp3	209558	1.19	1.66	ectonucleotide pyrophosphatase/phosphodiesterase 3
Enpp5	83965	1.53	1.74	ectonucleotide pyrophosphatase/phosphodiesterase 5
Entpd5	12499	1.66	1.3	ectonucleoside triphosphate diphosphohydrolase 5
Errfi1	74155	1.34	1.14	ERBB receptor feedback inhibitor 1
Esam	69524	1.43	1.19	endothelial cell-specific adhesion molecule
Ezr	22350	1.33	1.28	ezrin; hypothetical protein LOC100044177
F3	14066	2.07	3.17	coagulation factor III
Fabp1	14080	4.32	1.88	fatty acid binding protein 1, liver
Fam114a1	68303	2.07	1.67	family with sequence similarity 114, member A1
Fam129a	63913	1.03	1.22	family with sequence similarity 129, member A
Fam134b	66270	1.58	2.67	family with sequence similarity 134, member B
Fam189a2	381217	1.1	1.56	predicted gene 967
Fam46a	212943	1.13	1.32	family with sequence similarity 46, member A
Fam83b	208994	1.53	1.65	family with sequence similarity 83, member B
Fbp2	14120	1.46	1.01	fructose bisphosphatase 2
Fbxo15	50764	1.81	2.37	F-box protein 15
Fcgrt	14132	3.38	2.32	Fc receptor, IgG, alpha chain transporter
Fga	14161	2.64	2.155	fibrinogen alpha chain
Fgd4	224014	1.15	1.17	FYVE, RhoGEF and PH domain containing 4
Fgg	99571	4.73	3.47	fibrinogen gamma chain
Flrt2	399558	2.23	1.08	fibronectin leucine rich transmembrane protein 2
Fmo1	14261	2.17	2.96	flavin containing monooxygenase 1
Fnd3c2	331491	2.71	1.84	fibronectin type III domain containing 3C2
Fndc1	68655	1.36	1.04	fibronectin type III domain containing 1; similar to fibronectin type III domain containing 1
Gas6	14456	2.03	1.37	growth arrest specific 6
Gbp6	229900	1.51	1.37	guanylate binding protein 6
Gcnt2	14538	1.35	1.57	glucosaminyl (N-acetyl) transferase 2, I-branching enzyme
Gipc2	54120	2.5	2.32	GIPC PDZ domain containing family, member 2
Gkn2	66284	1.98	4.64	gastrokine 2
Glipr1	73690	1.64	2.28	GLI pathogenesis-related 1 (glioma)
Gm10768	100038628	2.4	3.04	predicted gene 10768
Gm10786	100038539	1.81	2.48	predicted gene 10786
Gm15070	100038527	1.95	2.06	predicted gene 15070
Golt1a	68338	1.02	1.84	golgi transport 1 homolog A (S. cerevisiae)
Gpnmb	93695	3.41	1.41	glycoprotein (transmembrane) nmb
Gpr39	71111	1.45	1.77	G protein-coupled receptor 39
Gpr50	14765	3.5	1.24	G-protein-coupled receptor 50
Gpr56	14766	2.21	1.05	G protein-coupled receptor 56
Gprc5a	232431	1.94	1.5	G protein-coupled receptor, family C, group 5, member A
Grina	66168	1.41	1.14	glutamate receptor, ionotropic, N-methyl D-aspartate-associated protein 1 (glutamate binding)
Gsta1	14857	4.57	1.615	glutathione S-transferase, alpha 1 (Ya)
Gsta3	14859	4.02	1.36	glutathione S-transferase, alpha 3
Gstm1	14862	3.22	2.5	similar to Glutathione S-transferase Mu 1 (GST class-mu 1) (Glutathione S-transferase GT8.7) (pmGT10) (GST 1-1); predicted gene 5562; glutathione S-transferase, mu 1
Gstm4	14865	1.29	1.35	glutathione S-transferase, mu 4
Gucy2c	14917	1.27	1.04	guanylate cyclase 2c
Hexb	15212	2.31	1.98	hexosaminidase B
Hgf	15234	2.5	1.08	hepatocyte growth factor
Hhip	15245	2.34	1.88	Hedgehog-interacting protein
Hpx	15458	3.33	3.24	hemopexin
Hsd11b2	15484	1.09	1.6	hydroxysteroid 11-beta dehydrogenase 2
Hsd3b7	101502	1.11	1.04	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7
Icam1	15894	1.08	1.34	intercellular adhesion molecule 1
Ifi27l1	52668	2.21	1.73	interferon, alpha-inducible protein 27 like 1
Ifi30	65972	1.58	2.39	interferon gamma inducible protein 30
Ifi35	70110	1.08	1.03	interferon-induced protein 35
Ifi44	99899	2.6	1.79	interferon-induced protein 44
Ifit1	15957	2.39	1.83	interferon-induced protein with tetratricopeptide repeats 1
Ifit3	15959	1.91	1.82	interferon-induced protein with tetratricopeptide repeats 3

Igf2	16002	1.27	2.72	insulin-like growth factor 2
Igfbp5	16011	3.86	2.94	insulin-like growth factor binding protein 5
Igfbp7	29817	4.29	2.95	insulin-like growth factor binding protein 7
Il10rb	16155	1.79	1.22	interleukin 10 receptor, beta
Il1r1	16177	2.44	2.21	interleukin 1 receptor, type I
Irak2	108960	1.14	1.08	interleukin-1 receptor-associated kinase 2
Irf6	54139	1.31	1.34	interferon regulatory factor 6
Irf9	16391	1.27	1.32	interferon regulatory factor 9
Islr	26968	2.71	1.17	immunoglobulin superfamily containing leucine-rich repeat
Itga3	16400	2.17	1.73	integrin alpha 3
Itga6	16403	1.63	1.41	integrin alpha 6
Itgav	16410	1.15	1.09	integrin alpha V
Itih2	16425	3.82	1.48	inter-alpha trypsin inhibitor, heavy chain 2
Itm2b	16432	1.76	1.2	integral membrane protein 2B
Kng1	16644	5.39	3.72	kininogen 1
Krt13	16663	1.17	1.01	keratin 13
Krt20	66809	2	1.8	keratin 20
Lama1	16772	2.19	4.4	laminin, alpha 1
Lamb1-1	16777	1.27	1.96	laminin B1 subunit 1
Lamc1	226519	1.81	2.41	laminin, gamma 1
Lamc2	16782	2.03	1.34	laminin, gamma 2
Lamp2	16784	1.71	1.23	lysosomal-associated membrane protein 2
Lbp	16803	2.05	1.43	lipopolysaccharide binding protein
Lcn2	16819	2.28	2.44	lipocalin 2
Lcor	212391	1.39	1.29	ligand dependent nuclear receptor corepressor
Lcp1	18826	2.51	2.54	lymphocyte cytosolic protein 1
Lgals3bp	19039	3.56	2.15	lectin, galactoside-binding, soluble, 3 binding protein
Lima1	65970	1.73	1.59	LIM domain and actin binding 1
Lipa	16889	1.47	1.31	lysosomal acid lipase A
Liph	239759	2.58	2.14	lipase, member H
Lox	16948	3.65	1.7	lysyl oxidase
Loxl2	94352	1.6	2.83	lysyl oxidase-like 2
Lpl	16956	2.34	1.73	lipoprotein lipase; similar to Lipoprotein lipase precursor (LPL)
Lpp	210126	1.135	1.15	LIM domain containing preferred translocation partner in lipoma
Lrp2	14725	1.39	2.16	low density lipoprotein receptor-related protein 2
Lum	17022	7.07	3.8	lumican
Lyz2	17105	5.1	2.01	lysozyme 2
Mab21l2	23937	2.74	1.64	mab-21-like 2 (C. elegans)
Maf	17132	1.3	1.01	similar to c-Maf long form; avian musculoaponeurotic fibrosarcoma (v-maf) AS42 oncogene homolog
Man1a	17155	1.78	1.32	mannosidase 1, alpha
Manba	110173	1.33	1.49	mannosidase, beta A, lysosomal
Maob	109731	3.66	3.23	monoamine oxidase B
Mboat1	218121	2.2	1.28	membrane bound O-acyltransferase domain containing 1
Megf9	230316	1.17	1.28	multiple EGF-like-domains 9
Mertk	17289	1.24	1.18	c-mer proto-oncogene tyrosine kinase
Mfsd7c	217721	1.77	2.27	major facilitator superfamily domain containing 7C
Mgat4a	269181	2.85	1.3	mannoside acetylglucosaminyltransferase 4, isoenzyme A
Mgst1	56615	2.74	1.61	microsomal glutathione S-transferase 1
Mgst2	211666	2.18	2.09	microsomal glutathione S-transferase 2
Mical2	320878	1.61	1.45	microtubule associated monoxygenase, calponin and LIM domain containing 2
Mov10	17454	1.14	1.33	Moloney leukemia virus 10; predicted gene 7357
Mpzl2	14012	1.94	1.17	myelin protein zero-like 2
Mreg	381269	2.01	1.76	melanoregulin
Mtmr11	194126	1.03	1.87	myotubularin related protein 11
Mtus1	102103	1.62	1.9	mitochondrial tumor suppressor 1
Muc1	17829	3.32	1.54	mucin 1, transmembrane
Mybpc3	17868	1.53	2.61	myosin binding protein C, cardiac
Myh6	17888	3.32	2.53	myosin, heavy polypeptide 6, cardiac muscle, alpha
Myh7	140781	2.4	1.48	myosin, heavy polypeptide 7, cardiac muscle, beta

Myl3	17897	1.53	2.03	myosin, light polypeptide 3
Myl7	17898	3.45	1.75	myosin, light polypeptide 7, regulatory
Mylip	218203	1.13	1.37	myosin regulatory light chain interacting protein
Myo1d	338367	2.39	1.52	myosin ID
Myo5b	17919	1.23	1.17	myosin VB
Myo6	17920	1.94	2.55	myosin VI
Myom1	17929	1.38	1.05	myomesin 1
Nebl	74103	1.263333333	1.42	nebulette
Nedd4l	83814	1.6	1.51	neural precursor cell expressed, developmentally down-regulated gene 4-like
Nedd9	18003	1.67	1.31	neural precursor cell expressed, developmentally down-regulated gene 9
Nek6	59126	1.61	1.12	NIMA (never in mitosis gene a)-related expressed kinase 6
Nepn	66650	1.04	2.11	nephrocan
Nfib	18028	3.685	1.925	nuclear factor I/B
Nhlrc3	212114	1.35	1.71	NHL repeat containing 3
Nkain4	58237	1.35	1.03	Na+/K+ transporting ATPase interacting 4
Npas2	18143	1.07	1.41	neuronal PAS domain protein 2
Npl	74091	1.1	2.23	N-acetylneuraminate pyruvate lyase
Npr3	18162	3.15	1.23	natriuretic peptide receptor 3
Nrn1	68404	2.95	1.85	neuritin 1
Oaf	102644	1.82	1.17	OAF homolog (Drosophila)
Ocln	18260	2.05	1.88	occludin
Olfml1	244198	2.25	1.49	olfactomedin-like 1
Olfml3	99543	2.69	1.39	olfactomedin-like 3
Osmr	18414	2.11	1.98	oncostatin M receptor
Palld	72333	1.69	1.07	palladin, cytoskeletal associated protein
Pamr1	210622	4.25	2.26	peptidase domain containing associated with muscle regeneration 1
Pde3a	54611	1.41	1.82	phosphodiesterase 3A, cGMP inhibited
Pdlim1	54132	1.4	1.49	PDZ and LIM domain 1 (elfin); predicted gene 5864; similar to PDZ and LIM domain protein 1 (LIM domain protein CLP-36) (C-terminal LIM domain protein 1) (Elfin)
Pf4	56744	3.33	2.01	platelet factor 4
Pga5	58803	3.48	4.2	pepsinogen 5, group I
Pgcp	54381	1.57	1.01	plasma glutamate carboxypeptidase
Pgm5	226041	2.35	1.37	phosphoglucomutase 5
Pi15	94227	2.85	1.34	peptidase inhibitor 15
Pip5k1b	18719	1.11	1.29	phosphatidylinositol-4-phosphate 5-kinase, type 1 beta; similar to phosphatidylinositol 4-phosphate 5-kinase type I-alpha
Pklr	18770	1.26	1.16	pyruvate kinase liver and red blood cell
Pla1a	85031	1.17	2.03	phospholipase A1 member A
Pla2g4a	18783	3.56	1.2	phospholipase A2, group IVA (cytosolic, calcium-dependent)
Plac8	231507	2.46	2.21	placenta-specific 8
Plagl1	22634	5.16	2.42	pleiomorphic adenoma gene-like 1
Plat	18791	2.09	2.21	plasminogen activator, tissue
Plau	18792	2.1	1.22	plasminogen activator, urokinase
Plek	56193	3.01	1.59	pleckstrin
Plg	18815	1.89	1.23	plasminogen
Plin2	11520	1.59	1.17	adipose differentiation related protein
Pln	18821	3.03	2.17	phospholamban
Plod2	26432	1.14	2.04	procollagen lysine, 2-oxoglutarate 5-dioxygenase 2
Pls1	102502	1.67	2.24	plastin 1 (I-isoform)
Popdc2	64082	1.46	1.04	popeye domain containing 2
Postn	50706	5.21	4.05	periostin, osteoblast specific factor
Ppbp	57349	1.58	1.32	pro-platelet basic protein
Ppfibp2	19024	1.28	1.63	protein tyrosine phosphatase, receptor-type, F interacting protein, binding protein 2
Ppp1r3c	53412	1.36	1.25	protein phosphatase 1, regulatory (inhibitor) subunit 3C
Prkg2	19092	1.24	1.47	protein kinase, cGMP-dependent, type II
Prl3b1	18776	3.38	2.75	prolactin family 3, subfamily b, member 1
Prl6a1	19111	1.26	1.27	prolactin family 6, subfamily a, member 1
Proc	19123	1.09	1.47	protein C
Prss35	244954	2.51	2.4	protease, serine, 35

Prss8	76560	1.13	1.73	protease, serine, 8 (prostasin)
Psap	19156	1.47	1.03	prosaposin
Ptgs2	19225	2.97	2.55	prostaglandin-endoperoxide synthase 2
Ptn	19242	3.16	1.92	pleiotrophin
Pttg1ip	108705	1.04	1.16	pituitary tumor-transforming 1 interacting protein
Pvr	52118	1.12	1.78	poliovirus receptor
Qsox1	104009	2.23	1.7	quiescin Q6 sulfhydryl oxidase 1
Rab11fip5	52055	1.02	1.31	RAB11 family interacting protein 5 (class I)
Rab17	19329	1.15	1.04	RAB17, member RAS oncogene family; similar to RAB17, member RAS oncogene family
Rap1gap	110351	1.23	1.14	Rap1 GTPase-activating protein
Rassf6	73246	1.27	1.67	Ras association (RalGDS/AF-6) domain family member 6
Rbl2	19651	1.35	1.09	retinoblastoma-like 2
Rbm46	633285	1.05	1.14	RNA binding motif protein 46
Rbm47	245945	1.71	2.17	RNA binding motif protein 47
Rbms3	207181	1.78	1.07	RNA binding motif, single stranded interacting protein
Rdh10	98711	1.38	1.13	retinol dehydrogenase 10 (all-trans)
Rhoc	11853	1.23	1.13	ras homolog gene family, member C
Rnf144b	218215	1.09	1.16	ring finger protein 144B
Robo2	268902	1.77	1.43	roundabout homolog 2 (Drosophila)
Rsad2	58185	1.96	1.56	radical S-adenosyl methionine domain containing 2
Rtp4	67775	2.86	2.22	receptor transporter protein 4
S100g	12309	3.56	5.25	S100 calcium binding protein G
S1pr3	13610	2.28	1	sphingosine-1-phosphate receptor 3
Saa3	20210	1.78	1.79	serum amyloid A 3
Samd9l	209086	1.84	1.27	sterile alpha motif domain containing 9-like
Sdc4	20971	1.38	1.52	syndecan 4
Sdpr	20324	2.75	1.39	serum deprivation response
Sec24d	69608	2.36	1.52	Sec24 related gene family, member D (S. cerevisiae)
Sel1/3	231238	1.15	1.97	sel-1 suppressor of lin-12-like 3 (C. elegans)
Selenbp1	20341	1.75	1.09	selenium binding protein 1; hypothetical protein LOC100044204
Sema4g	26456	1.16	1.09	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4G
Sepp1	20363	3.44	3.08	selenoprotein P, plasma, 1
Serpina1a	20700	4.62	1.12	serine (or cysteine) peptidase inhibitor, clade A, member 1A
Serpina1b	20701	4.63	1.4	serine (or cysteine) preptidase inhibitor, clade A, member 1B
Serpina1e	20704	4.72	1.13	predicted gene 8893; serine (or cysteine) peptidase inhibitor, clade A, member 1E
Serpina3n	20716	4.51	2.26	serine (or cysteine) peptidase inhibitor, clade A, member 3N
Serpina6	12401	4.9	1.25	serine (or cysteine) peptidase inhibitor, clade A, member 6
Serpind1	15160	3.36	2.04	serine (or cysteine) peptidase inhibitor, clade D, member 1
Serping1	12258	2.7	1.98	serine (or cysteine) peptidase inhibitor, clade G, member 1
Sgms2	74442	1.79	1.23	sphingomyelin synthase 2
Sh2d4a	72281	2.53	1.44	SH2 domain containing 4A
Slc16a12	240638	3.24	1.02	solute carrier family 16 (monocarboxylic acid transporters), member 12
Slc39a5	72002	1.12	1.75	solute carrier family 39 (metal ion transporter), member 5
Slc40a1	53945	2.47	2.04	solute carrier family 40 (iron-regulated transporter), member 1
Slc41a2	338365	1.21	1.49	solute carrier family 41, member 2
Slc44a3	213603	2.43	1.72	solute carrier family 44, member 3
Slc8a1	20541	1.795	1.42	solute carrier family 8 (sodium/calcium exchanger), member 1
Sorbs2	234214	3.47	1.23	sorbin and SH3 domain containing 2
Sord	20322	1.41	1.035	sorbitol dehydrogenase
Sparc	20692	1.51	1.75	secreted acidic cysteine rich glycoprotein; similar to Secreted acidic cysteine rich glycoprotein
Sparcl1	13602	3.73	1.28	SPARC-like 1
Sphkap	77629	1.16	1.11	SPHK1 interactor, AKAP domain containing
Spint2	20733	1.24	1.14	serine protease inhibitor, Kunitz type 2
Spon2	100689	2.11	1.74	spondin 2, extracellular matrix protein
Spp1	20750	5.26	2.21	secreted phosphoprotein 1
Spp2	75396	4.25	4.17	secreted phosphoprotein 2
Sprr1a	20753	3.8	1.42	small proline-rich protein 1A
Srgn	19073	4.22	4.62	serglycin

Stard10	56018	1.65	1	START domain containing 10
Steap4	117167	1.31	1.39	STEAP family member 4
Sult1e1	20860	1.02	2.12	sulfotransferase family 1E, member 1
Susd2	71733	1.33	2.17	sushi domain containing 2
Taf9b	407786	1.1	1.7	TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor
Tcf21	21412	2.72	2	transcription factor 21
Tdrd7	100121	1.58	1.13	tudor domain containing 7
Tec	21682	1.39	1.58	tec protein tyrosine kinase
Tff3	21786	2.45	2.87	trefoil factor 3, intestinal
Tgfbi	21810	2.91	2.04	transforming growth factor, beta induced
Tgfbr2	21813	1.71	1.43	transforming growth factor, beta receptor II
Tgfbr3	21814	1.72	1.83	transforming growth factor, beta receptor III
Thbs2	21826	3.04	1.43	thrombospondin 2
Timd2	171284	1.63	1.72	T-cell immunoglobulin and mucin domain containing 2
Timp3	21859	2.01	2.09	tissue inhibitor of metalloproteinase 3
Tm4sf5	75604	1.62	2.36	transmembrane 4 superfamily member 5
Tmbim1	69660	1.6	1.31	transmembrane BAX inhibitor motif containing 1
Tmc4	353499	1.92	1.23	transmembrane channel-like gene family 4
Tmem140	68487	2.44	1.96	transmembrane protein 140
Tmem150a	232086	1.18	1.59	transmembrane protein 150A
Tmem176a	66058	2.33	1.27	transmembrane protein 176A
Tmem176b	65963	2.77	1.3	transmembrane protein 176B
Tmem30b	238257	1.83	1.48	transmembrane protein 30B
Tmem37	170706	1.48	1.52	transmembrane protein 37
Tnc	21923	3.66	1.74	tenascin C
Tnni1	21952	1.5	1.47	troponin I, skeletal, slow 1
Tpbpa	21984	5.23	3.31	trophoblast specific protein alpha
Tpd52	21985	1.36	1.235	similar to Tpd52 protein; tumor protein D52
Ttn	22138	1.39	1.31	titin
Tuft1	22156	1.44	1.11	similar to tuftelin; tuftelin 1
Ugt2b34	100727	6.59	4.34	UDP glucuronosyltransferase 2 family, polypeptide B34
Upk1b	22268	4.01	1.84	uroplakin 1B
Vaultrc5	378472	1.76	1.2	vault RNA component 5
Vcam1	22329	3.3	1.62	vascular cell adhesion molecule 1
Vsig2	57276	1.97	1.6	V-set and immunoglobulin domain containing 2
Vtn	22370	4.98	2.26	vitronectin
Wipi1	52639	1.57	1.26	WD repeat domain, phosphoinositide interacting 1
Zfp36	22695	1.7	1.04	zinc finger protein 36
Zfp9	22750	1.33	1.12	zinc finger protein 9
Zim1	22776	2.51	2.22	zinc finger, imprinted 1

Concordantly downregulated genes in wt and dnmt1-/- EBs between d4-16				
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change dnmt1-/-**	**Gene name**
6030429G01Rik	436022	-1.75	-1.82	RIKEN cDNA 6030429G01 gene
Apcdd1	494504	-2.2	-2.07	adenomatosis polyposis coli down-regulated 1
Aplnr	23796	-1.1	-1.35	apelin receptor
Atp11c	320940	-2.49	-1.68	ATPase, class VI, type 11C
B130016D09Rik	436015	-1.56	-1.5	RIKEN cDNA B130016D09 cDNA
BC023829	236848	-3.31	-2.04	cDNA sequence BC023829
Bmper	73230	-1.67	-1.02	BMP-binding endothelial regulator
Cer1	12622	-1.86	-2.55	cerberus 1 homolog (Xenopus laevis)
Cfc1	12627	-2.62	-1.75	cripto, FRL-1, cryptic family 1
Coch	12810	-2.3	-1.56	coagulation factor C homolog (Limulus polyphemus)
Cpxm2	55987	-1.34	-1.47	carboxypeptidase X 2 (M14 family)
Cxcr4	12767	-1.54	-2.45	chemokine (C-X-C motif) receptor 4
Cyp4f15	106648	-1.11	-1.42	cytochrome P450, family 4, subfamily f, polypeptide 15
Dll3	13389	-1.91	-2.19	delta-like 3 (Drosophila)
Dock11	75974	-1.09	-1.41	dedicator of cytokinesis 11

Fgf10	14165	-2.3	-1.87	fibroblast growth factor 10
Fgf3	14174	-2.26	-1.95	fibroblast growth factor 3
Foxf1a	15227	-1.455	-1.455	forkhead box F1a
Foxh1	14106	-3.27	-2.1	forkhead box H1
Gas5	14455	-1.523333333	-1.05	growth arrest specific 5
Ifitm1	68713	-1.04	-1.16	interferon induced transmembrane protein 1
Kdr	16542	-1.81	-1.92	kinase insert domain protein receptor
Lef1	16842	-1.68	-1.53	lymphoid enhancer binding factor 1
Lgr5	14160	-2.46	-2.29	leucine rich repeat containing G protein coupled receptor 5
Lhx1	16869	-2.53	-2.52	LIM homeobox protein 1
Mns1	17427	-1.73	-1.61	meiosis-specific nuclear structural protein 1
Mogat2	233549	-2.04	-1.12	monoacylglycerol O-acyltransferase 2
Ms4a4d	66607	-2.54	-1.31	membrane-spanning 4-domains, subfamily A, member 4D
Msx1	17701	-2.35	-1.61	homeobox, msh-like 1
Msx2	17702	-3	-1.49	similar to homeobox protein; homeobox, msh-like 2
Myc	17869	-1.59	-1.07	myelocytomatosis oncogene
Nkrf	77286	-1.34	-1.01	NF-kappaB repressing factor
Padi3	18601	-1.5	-1.49	peptidyl arginine deiminase, type III
Pop1	67724	-1.39	-1.07	processing of precursor 1, ribonuclease P/MRP family, (S. cerevisiae)
Rftn1	76438	-1.83	-1.99	raftlin lipid raft linker 1
Rspo3	72780	-1.96	-1.99	R-spondin 3 homolog (Xenopus laevis)
Slc44a5	242259	-1.57	-2.09	solute carrier family 44, member 5
Snord52	100217427	-1.73	-1.2	small nucleolar RNA, C/D box 52
St6galnac4	20448	-1.08	-1.31	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4
Taf1d	75316	-1.52	-1.173333333	TATA box binding protein (Tbp)-associated factor, RNA polymerase I, D; predicted gene 13487
Trpc3	22065	-1.94	-1.63	transient receptor potential cation channel, subfamily C, member 3
Vstm2b	58188	-1.95	-1.59	hypothetical protein LOC100045106; V-set and transmembrane domain containing 2B
Wnt3	22415	-1.66	-1.82	wingless-related MMTV integration site 3
Zcchc12	72693	-1.03	-1.85	zinc finger, CCHC domain containing 12

Concordantly upregulated genes in wt and TKO EBs between d4-16 (Note that genes marked in red show opposite expression pattern in wt and TKO Ebs)				
Gene Symbol	**Entrez Gene ID**	**fold change wt**	**fold change TKO**	**Gene Name**
Antxr2	71914	1.2	-1.15	liver glycogen phosphorylase
Atp9a	11981	1.37	1.15	Ras association (RalGDS/AF-6) domain family (N-terminal) member 9
Car8	12319	1.88	1.16	epsin 3
Copz2	56358	1.56	1.94	DMRT-like family C1c2; DMRT-like family C1c
Cyp2c55	72082	2.2	1.42	Eph receptor A7
Cyp2s1	74134	1.39	1.19	regulator of G-protein signaling 2
Cyp4a12b	13118	1.11	-1.47	septin 4
Dmrtc1c	71083	1.11	1.77	carbonic anhydrase 8; similar to Carbonic anhydrase-related protein (CARP) (CA-VIII)
Epha7	13841	1.95	-1.15	heat shock protein 8
Epn3	71889	1.45	1.2	sarcoglycan, epsilon
Gadd45b	17873	2.07	1.56	LIM domain binding 2
Gpr97	54672	1.16	1.3	cytochrome P450, family 2, subfamily s, polypeptide 1
Hspb8	80888	1.08	1.45	serine (or cysteine) peptidase inhibitor, clade B, member 1a
Ldb2	16826	1.17	-1.24	coatomer protein complex, subunit zeta 2
Mecom	14013	2.76	-1.14	growth arrest and DNA-damage-inducible 45 beta
Nnat	18111	1.1	-1.39	anthrax toxin receptor 2
Npy1r	18166	1.02	-1.28	cytochrome P450, family 2, subfamily c, polypeptide 55
Pygl	110095	1.24	1.44	ecotropic viral integration site 1
Rassf9	237504	2.52	-1.02	cytochrome P450, family 4, subfamily a, polypeptide 12B
Rgs2	19735	1.41	-1.53	ATPase, class II, type 9A
Sept4	18952	1.35	1.52	G protein-coupled receptor 97
Serpinb1a	66222	1.56	1.42	neuronatin
Sgce	20392	1.13	1.01	neuropeptide Y receptor Y1

Concordantly downregulated genes in wt and TKO EBs between d4-16				
Gene symbol	**Entrez Gene ID**	**Fold change wt**	**Fold change TKO**	**Gene name**
Akr1b3	11677	-1.4525	-1.01	aldo-keto reductase family 1, member B3 (aldose reductase)
Calcr	12311	-1.09	-1.25	calcitonin receptor
Crispld1	83691	-1.43	-1.46	cysteine-rich secretory protein LCCL domain containing 1
Crmp1	12933	-1.13	-1.19	collapsin response mediator protein 1
Dnmt3b	13436	-2.09	-1.12	DNA methyltransferase 3B
Dut	110074	-1.08	-1.02	deoxyuridine triphosphatase
Gng3	14704	-1.17	-1.8	guanine nucleotide binding protein (G protein), gamma 3
Grik3	14807	-1.29	-1.55	glutamate receptor, ionotropic, kainate 3
Hand1	15110	-2.19	1.16	heart and neural crest derivatives expressed transcript 1
Il17rd	171463	-1.38	-1.39	interleukin 17 receptor D
Kif1a	16560	-1.27	-1.09	kinesin family member 1A
Mtap7d3	320923	-1.29	-2.12	MAP7 domain containing 3
Rragb	245670	-1.35	-1.05	Ras-related GTP binding B
Tnfrsf19	29820	-1.07	-1.13	tumor necrosis factor receptor superfamily, member 19
Usp44	327799	-1.46	-1.05	ubiquitin specific peptidase 44
Zc3hav1	78781	-1.09	1.48	zinc finger CCCH type, antiviral 1
Zdbf2	73884	-1.33	-1.14	zinc finger, DBF-type containing 2
Zfp280c	208968	-1.21	-1.24	zinc finger protein 280C
Zfp365	216049	-1.04	-1.1	zinc finger protein 365

7. Publications

Schmidt CS, Bultmann S, Meilinger D, Zacher B, Tresch A, Maier K, Peter C, Martin DE, Leonhardt H, Spada F: Global DNA hypomethylation prevents consolidation of differentiation programs and allows reversion to the embryonic stem cell state. **PLoS One**. 2012; 7(12) : e52629

Solovei I, Wang AS, Thanisch K, **Schmidt CS**, Krebs S, Zwerger M, Devys D, Peichl L, Herrmann H, Blum H, Engelkamp D, Stewart CL, Leonhardt H, Joffe B: LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation.
Manuscript under revision at Cell

Bultmann S, Morbitzer R, **Schmidt CS**, Thanisch K, Spada F, Elsaesser J, Lahaye T, Leonhardt H (2012). Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers.
Nucleic Acids Res., 1–10, (2012)

Pichler G, Wolf P, **Schmidt CS**, Meilinger D, Schneider K, Frauer C, Fellinger K, Rottach A, Leonhardt H (2011). Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways. **J Cell Biochem**, 112(9):2585-93

Szwagierczak A, Brachmann A, **Schmidt CS**, Bultmann S, Leonhardt H, Spada F (2011). Characterization of PvuRts1I endonuclease as a tool to investigate genomic 5-hydroxymethylcytosine. ***Nucleic Acids Res***, 39(12):5149-56

Szwagierczak A, Bultmann S, **Schmidt CS**, Spada F, Leonhardt H (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA.
Nucleic Acids Res.;38(19):e181

Patent application

„Novel methods for detecting hydroxymethylcytosine” (B71043) at the European patent office, The Hague, ***pending***