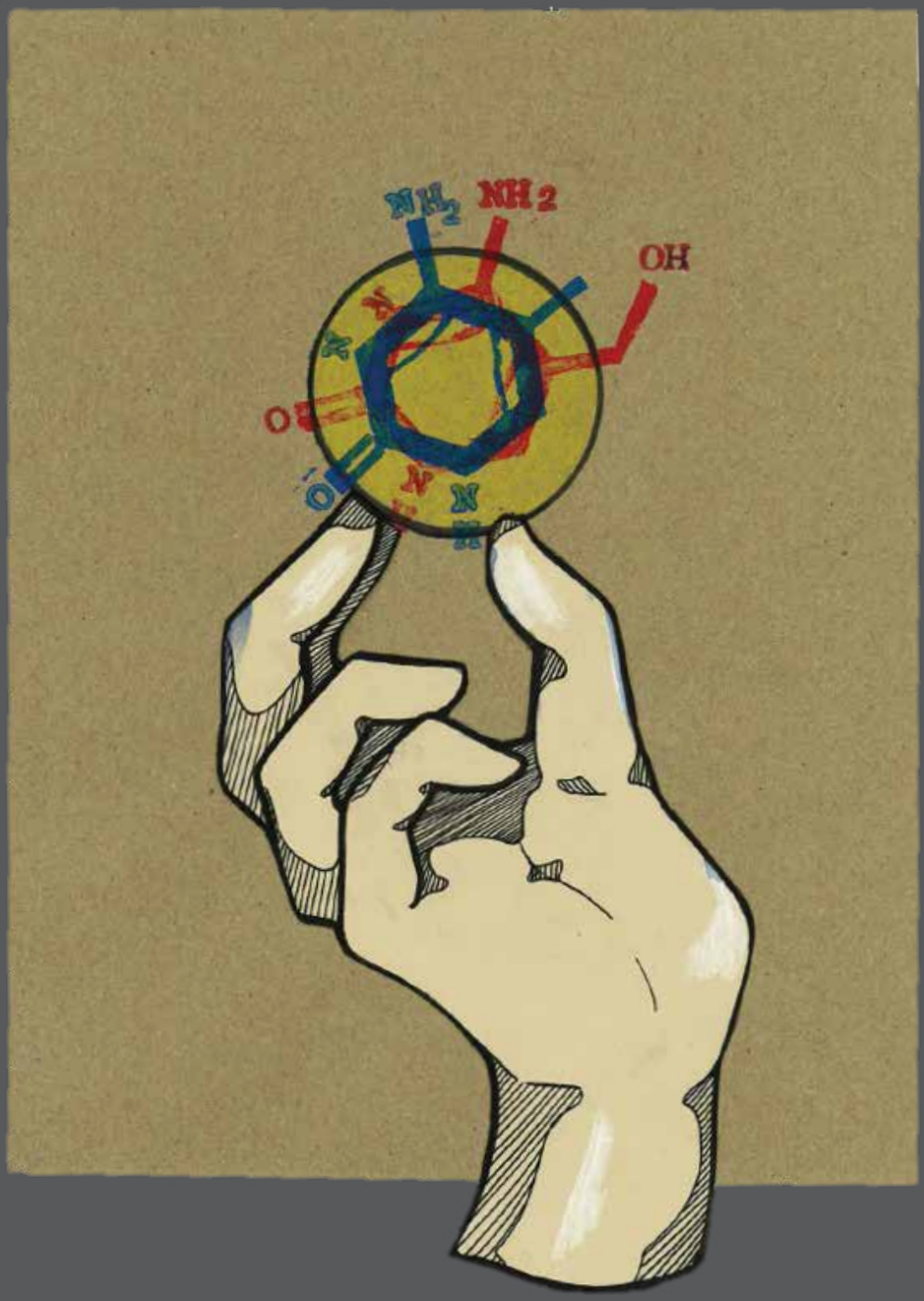


Characterization of the cancer-associated short isoform of Tet1 in DNA replication

by Florian D. Hastert

Characterization of the cancer-associated short isoform of Tet1 in DNA replication

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Dem Fachbereich Biologie der Technischen Universität Darmstadt
zur Erlangung des akademischen Grades eines
Doctor rerum naturalium
vorgelete Dissertation

von

Florian Dieter Hastert, M.Sc. Biologie
geboren am 10.06.1986 in Eberbach.

Referentin: Prof. Dr. M. Cristina Cardoso
Korreferent: Prof. Dr. Bodo Laube
Tag der Einreichung: 28.11.2018

Darmstadt 2019

Und ich weiß, was für ein Wetter über der Stadt lag,
in der Nacht, in der wir nicht wussten,
woher wir kommen, wohin wir gehen.

Contents

List of Figures

List of Tables

1 Summary/Zusammenfassung

1.1 Summary

Almost all cells of a metazoan organism share the same genetic information but can differ highly in their function and phenotype. The transition from the same genotype to different phenotypes is mostly achieved by epigenetic mechanisms, including DNA base modifications. The most prominent and best studied DNA modification is 5-methylcytosine (5mC), which was implicated in transcriptional repression. The levels of 5mC are maintained on the newly synthesized DNA strand, to ensure its faithful inheritance during mitotic cell division. Oxidation of the 5mC by Tet dioxygenases, to 5-hydroxymethylcytosine (5hmC) and even further, can drastically change the epigenetic information transported by this modification and hence modulate gene expression and the overall phenotype of a cell.
However, the exact regulation of Tet proteins and the putative cell cycle dependent inheritance of their oxidative products still is under investigation. In the present work, I therefore characterized the subnuclear distribution of different Tet proteins throughout the cell cycle and especially during S-phase and potential implications for cellular homoeostasis. I identified the recruitment of the short isoform of Tet1 (Tet1s), which is found to be overexpressed in different cancers, to sites of ongoing DNA replication in pericentric heterochromatin, which is rich in 5mC. I furthermore found that Tet1s localization during S-phase and also its catalytic activity is highly dependent on a conserved lysine residue. Moreover, I could show that Tet1 physically interacts with the E3 ubiquitin-ligase Uhrf1 and that the observed localization of Tet1s to sites of ongoing DNA replication in pericentric heterochromatin requires the presence of Uhrf1. Finally, I found that a breast cancer cell line that was shown to overexpress Tet1s, exhibits significantly changed 5mC and 5hmC levels, globally as well as *in situ*, in comparison to non-transformed breast epithelial cells.
In summary, my findings contribute to the understanding of how the epigenetic information of every cell can be diversified by the regulation of Tet protein localization and their catalytic activity.

1.2 Zusammenfassung

Nahezu alle Zellen eines metazoen Organismus besitzen die gleiche genetische Information, können sich jedoch in ihrer Funktion und ihrem Phänotyp stark unterscheiden. Der Übergang vom gleichen Genotyp zu verschiedenen Phänotypen wird hauptsächlich durch epigenetische Mechanismen, einschließlich verschiedener DNA-Basenmodifikationen erreicht. Die bekannteste und am besten untersuchte DNA-Modifikation ist 5-Methylcytosin (5mC), das an der Transkriptionsrepression beteiligt ist. Das 5mC-Muster wird während der S-Phase auf den neu synthetisierten DNA-Strang kopiert, um sicherzustellen, dass es während der mitotischen Zellteilung korrekt vererbt wid. Die Oxidation der 5mC-Methylgruppe durch Tet-Dioxygenasen zu 5-Hydroxymethylcytosin (5hmC) und darüber hinaus, kann die durch diese Modifikation transportierten epigenetischen Informationen drastisch verändern und somit die Genexpression und den Gesamtmetabolismus einer Zelle modulieren. Die genaue Regulation von Tet-Proteinen und die Zellzyklus-abhängige Aufrechterhaltung ihrer oxidativen Produkte wirft jedoch noch viele Fragen auf.
In der vorliegenden Arbeit habe ich daher die subnukleäre Verteilung verschiedener Tet-Proteine während des Zellzyklus und insbesondere während der S-Phase, sowie potenzielle Implikationen für die zelluläre Homöostase charakterisiert. Ich identifizierte die Rekrutierung der kurzen Isoform von Tet1 (Tet1s), die in verschiedenen Krebsarten überexprimiert wird, während der laufenden DNA-Replikation in perizentrischem Heterochromatin. Weiterhin fand ich, dass die Lokalisierung von Tet1 während der S-Phase und auch seine gerichtete katalytische Aktivität stark von einem konservierten Lysinrest abhängen. Außerdem konnte ich zeigen, dass Tet1 physisch mit der E3-Ubiquitin-Ligase Uhrf1 interagiert und dass die beobachtete Lokalisierung von Tet1s während der laufenden DNA-Replikation in perizentrischem Heterochromatin, die Anwesenheit von Uhrf1 erfordert. Schließlich konnte ich zeigen, dass eine Brustkrebszelllinie, die Tet1s überexprimiert, signifikant veränderte Level von 5mC und 5hmC sowohl global als auch textit in situ, im Vergleich zu nicht transformierten Brust-Epithelzellen, aufweist.
Zusammenfassend tragen meine Erkenntnisse zum generellen Verständnis bei, wie die epigenetischen Information jeder Zelle durch die Regulation der Tet-Proteinlokalisierung und ihrer katalytischen Aktivität diversifiziert werden kann.

2 Introduction

2.1 Information content of DNA and the impact of epigenetics

2.1.1 DNA – A linguistic perspective

Being presented with a naked stretch of eukaryotic, genomic DNA is similar to being presented with a piece of text that lacks all basic syntactic features like the simple separation of words by spaces. To simplify this analogy, transcriptional activity outside of genic regions is neglected and only a short genic stretch of DNA is observed in close detail.
A first layer of syntax is the correct order of cis-acting regulatory elements within any given genic region, like the promoter or the TATA-box and the open reading frame, associated with these elements. The distinction of the different, specialized components of this region is mainly achieved by defined sequence features. Consequently, all constituents of a genic region harbour conserved consensus sequences that clearly define them as promoter and transcription start sites or flank intronic and exonic regions, respectively.
If the regulatory elements and the promoter are bound by the transcription initiation complex and a RNA polymerase starts transcribing, a semantic meaning arises from this stretch of DNA. This eventually results in a protein product or a non-coding RNA and is comparable to reading out loud a text without paying attention to the actual combined meaning of words as long as basic grammatical rules are followed. If not only grammatical rules but also the meaningful combination of words is considered, pragmatic information is transported to the listener. (**Figure 1**)
In metazoans, almost all cells of an organism are genetically identical but still give rise to different cell types and consequently various, specialized tissues. Hence, a pragmatic mode of regulation for the information content of DNA is needed to control the precise, contextual correct activation of transcription in individual cells.
This pragmatic layer of DNA information interpretation is mainly achieved by epigenetic mechanisms and chromatin modifications, which finetune the interplay of syntactic and semantic DNA information content.

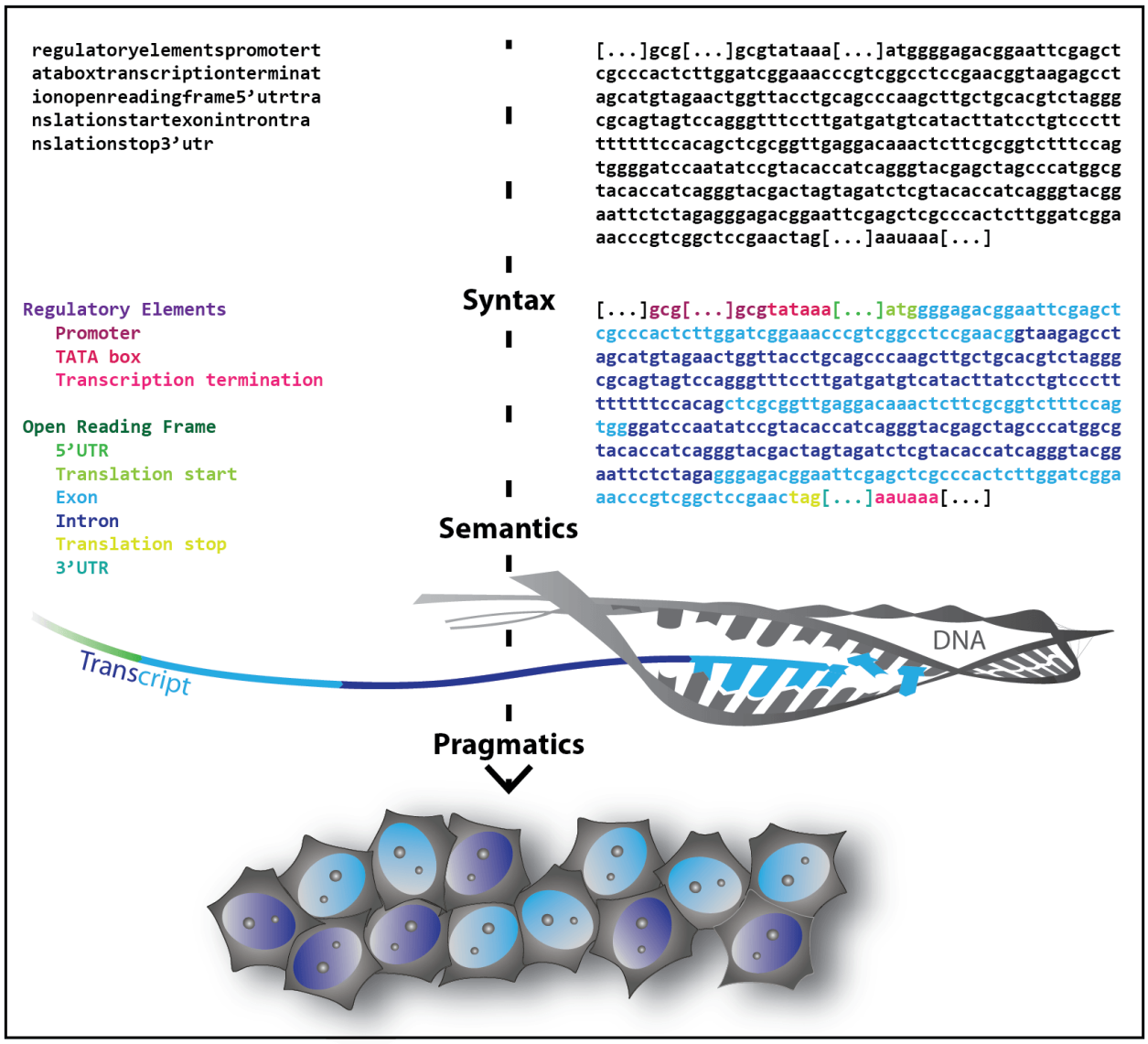


Figure 1: Illustration of the linguistic view of a genic DNA region. A non-structured DNA sequence that stretches across a genic region can be subdivided into regulatory elements and an open reading frame, according to the respective sequence features and is thereby given a logical syntax (on the righthand side). This is similar to the descriptive list on the lefthand side, whose syntax arises from the addition of spaces and bullet points. The combined interpretation of the regulatory elements and the open reading frame results in the transcription of this piece of DNA and is thereby given a semantic relevance in the form of (pre)mRNA and its product. The levels of a newly translated protein or a resulting non-coding RNA can differ highly between different cells, even within the same tissue. This heterogeneity shapes the individual phenotype of every cell and gives the initial DNA sequence a pragmatic meaning.

2.1.2 Revising Waddington

The term epigenetics was originally introduced by Conrad Hal Waddington to describe how phenotypical differences arise from genetic mutations. He used it, to bring together the fields of genetics and organismic development, then called epigenesis (Waddington, 1942). This definition would start to change with a discovery, made shortly after Waddington discussed his ideas on epigenetics and already five years before the initial description of the DNA double-helix structure (Watson and Crick, 1953). In 1948, Rollin Hotchkiss described a fifth peak on a paper chromatogram from thymus deoxyribonucleic acids, in addition to the four canonical DNA bases. He termed this novel mammalian DNA base "epicytosine", unknowingly anticipating the modern definition of epigenetics and additional layer of information that epigenetic chromatin modifications add to the genetic information of every eukaryotic cell (Hotchkiss, 1948). A sixth DNA base, named 5-hydroxymethylcytosine (5hmC), was identified in different vertebrate tissues like frog and mouse brain already in 1972 (Penn *et al.*, 1972). However, epicytosine or 5-methylcytosine (5mC), how it was later known, remained the only non-canonical mammalian DNA base that was studied to a larger extent for a long time.

5mC helped explaining how different cellular phenotypes can arise from the same nucleotide sequence. Consistently, it has meanwhile been established that the complex interplay of DNA modifications, different histone variants and their modifications as well as non-coding RNAs and nucleosomal arrangements and the thereby modulated binding of transcription factors serves as a hereditary transgenerational mechanism to regulate a plethora of cellular processes in higher eukaryotic organisms (**Figure 2**). The combinatorial effect of different epigenetic modifications furthermore defines the two main chromatin types, euchromatin and heterochromatin. Euchromatin is comparatively less condensed, gene-rich and characterized by high transcriptional activity. Heterochromatin, on the other hand, is generally considered to be transcriptionally inactive and can be further subdivided into two different classes, constitutive and facultative heterochromatin. The latter can become transcriptionally active in response to certain cellular signals. Constitutive heterochromatin, in contrast, is mainly composed of repetitive DNA sequences and stays transcriptionally inert in normal somatic cells (Joffe, Leonhardt, and Solovei, 2010).

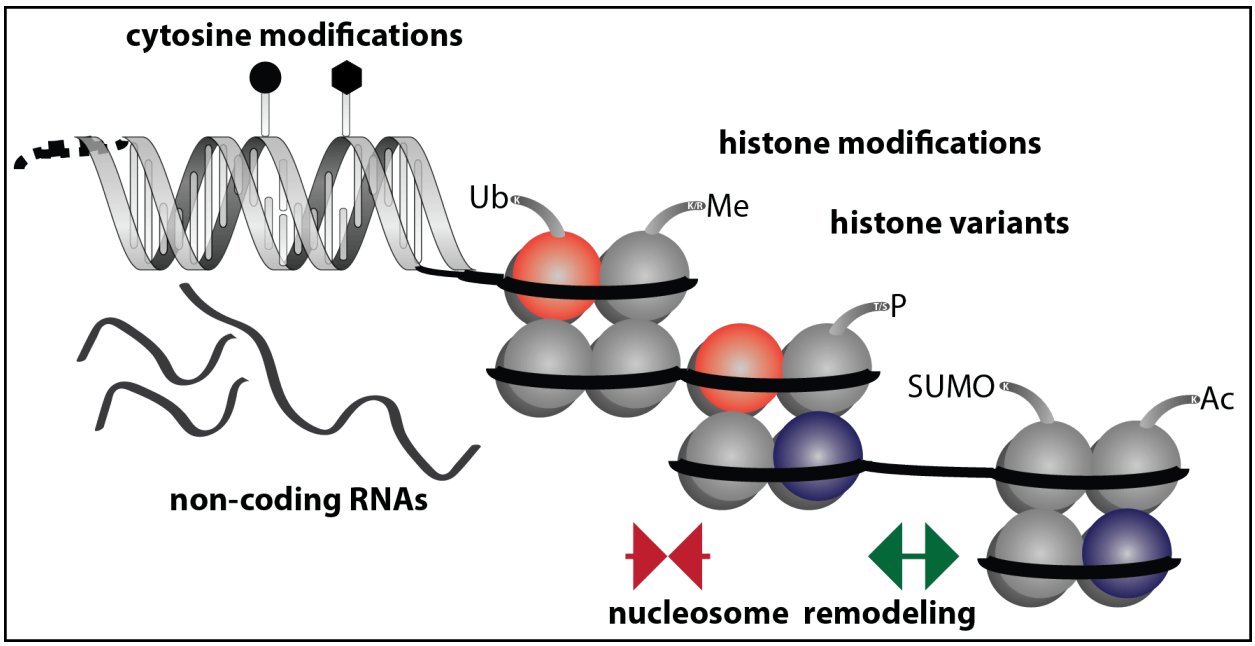


Figure 2: Overview over the different epigenetic mechanisms that regulate DNA metabolism. The epigenetic make-up of higher eukaryotic cells comprises many different chromatin modifications. Amongst them are covalent DNA- and histone modifications, but also various histone variants as well different non-coding RNAs and also nucleosomal arrangements. Mammalian DNA modifications comprise predominantly cytosine methylation and different oxidised forms thereof. Many post-translational histone tail modifications, like ubiquitination (Ub), SUMOylation (SUMO) and acetylation (Ac) are found exclusively on lysine residues. Methylation of histones (Me), however, is found on lysines and also on arginines and histone phosphorylation (P) is observed on threonines and serines. Besides the histone modifications, different histone variants that can vary between celltypes and even within the cell cycle, have been described to modulate important cellular processes. Non-coding RNAs act in diverse function to fine-tune transcriptional activity, for example by coating, to be silenced, DNA. These processes concomitantly affect nucleosome remodelling and thereby affect chromatin structure and consequently DNA metabolism.

2.1.3 Epigenetic mechanisms & chromatin modifications

Chromatin topology is highly dynamic and undergoes constant changes that correlate with its transcriptional activity, which in turn is highly dependent on epigenetic mechanisms and chromatin modifications (Hübner and Spector, 2010). The basic unit of chromatin organization is an approximately 146 base pair long stretch of DNA wrapped around a histone octamer, which consists of two copies of each of the core histones H2A, H2B, H3 and H4 that together form the so-called nucleosome (Luger *et al.*, 1997; Luger and Richmond, 1998). The drum-shaped nucleosome particles are further compacted via the binding of the linker histone H1, which sits between neighbouring nucleosomes and stabilizes the DNA that is wrapped around them (Hergeth and Schneider, 2015). The amino-terminal histone tails, which protrude from the core, are essential for the biological function of the nucleosome, as many post-translational modifications (PTMs) are deposited

here (Kouzarides, 2007). The most abundant histone modifications comprise lysine and arginine methylation, lysine acetylation, phosphorylation of serines or threonines, as well as ubiquitination and SUMOylation of lysines within the histone tails. Of the different groups of post-translational histone modifications, all of them have been implicated in transcriptional regulation, while some were shown to also regulate DNA repair or DNA replication.

Histone tail modifications can modulate chromatin structure, for example, by neutralizing the positive charge of lysine via the addition of acetyl groups, thereby opening up the chromatin. Histone hyperacetylation is therefore a characteristic feature of transcriptionally active euchromatic domains and mostly absent in heterochromatin. Methylation of histones, on the other hand, is found in transcriptionally active as well as silenced chromatin regions. While trimethylation of lysines 9, 20 and 27 in histone H3 (H3K9/20/27me3) is associated with heterochromatin, trimethylation of lysines 4, 36 and 79 (H3K4/36/79me3) is predominantly found in euchromatic domains (Karlic *et al.*, 2010). Each single histone tail modification can serve as a recruitment platform for further chromatin factors and hence downstream epigenetic processes (Bannister and Kouzarides, 2011). Although, conserved nucleosome modification patterns have been described in the context of promoters and different upstream regulatory elements (Wang *et al.*, 2008), the exact modification pattern of neighbouring nucleosomes can be highly diverse and is not mutually exclusive, resulting in the so-called histone code (Strahl and Allis, 2000; Jenuwein and Allis, 2001).

Besides their modification status, variants of the four core histones play an important role in various cellular processes and can be highly dynamic throughout the cell cycle (Boyarchuk, Montes de Oca, and Almouzni, 2011). Histone H2A and its variant H2A.Z, for example, are deposited in pericentric heterochromatin in discrete waves throughout the cell cycle and thereby contribute to the faithful progression of S-phase and G1, respectively (Boyarchuk *et al.*, 2014). Different histone variants and their modifications also serve as recognition sites as well as substrates for chromatin remodelling complexes like the SWI/SNF (SWItch/Sucrose Non-Fermentable) or the Mi-2/NuRD (Nucleosome Remodeling Deacetylase) complexes. These chromatin remodelling complexes consist of multiple different proteins, amongst them ATPases, which modulate nucleosomal repositioning, but also a plethora of epigenetic reader and writer proteins (Clapier *et al.*, 2017). The

Mi-2/NuRD-complex, for example, comprises histone deacetylases HDAC1 and HDAC2 as well as methyl-CpG binding domain protein (Mbd) family members Mbd2 and Mbd3 (Wade *et al.*, 1999), which together enable it to read chromatin signatures and finetune transcriptional activity. Besides this, the NuRD complex was also shown to bind the cell cycle independent histone variant H3.3, which is enriched in actively transcribed chromatin, thereby promoting gene activity (Kraushaar *et al.*, 2018) and to associate with the Tet1 dioxygenase (Shi *et al.*, 2013). While being crucial regulators of chromatin structure and function, many chromatin remodelling complexes were found to be themselves regulated by non-coding RNAs (ncRNAs) that often arise from non-protein coding DNA sequences and are therefore not translated (Palazzo and Lee, 2015). Accordingly, it could be shown that the SWI/SNF member Brg1/BAF can be removed from its chromatin binding sites by the long-ncRNA *Mhrt* to prevent nucleosome remodelling (Han *et al.*, 2014). Other ncRNAs, like *PARTICLE* or *HOTAIR*, have furthermore been implicated in regulating post-translational histone modifications, like lysine methylation, by acting as scaffolds for the assembly of distinct protein complexes for subsequent H3K27 methylation and H3K4 demethylation (Tsai *et al.*, 2010; O'Leary *et al.*, 2017). Finally, similar to acting as a scaffold for histone modifying enzymes, ncRNAs like *PARTICLE* were shown to modulate the activity of the maintenance DNA methyltransferase Dnmt1 and thereby affect global DNA methylation levels (O'Leary *et al.*, 2017).

2.2 DNA methylation and the Dnmt protein family

Methylation of the fifth carbon atom of cytosine is found in the DNA of many prokaryotic and with only a few exceptions eukaryotic organisms. It represents one of the most abundant and most conserved epigenetic modifications and regulates gene expression by silencing promoter regions and facilitating the compaction of chromatin (Lande-Diner *et al.*, 2007; Smith and Meissner, 2013). Mammalian DNA methylation occurs predominantly in the context of symmetric CpG dinucleotides and in the human genome between 60-80% of the about 28 million CpG sites, hence around 5% of all cytosines are methylated (Smith and Meissner, 2013). However, up to 200 base-pair long stretches of these dinucleotides, so-called CpG-islands (CGI), are often found in the context of promoters and transcription start sites (TSS), where they are sparsely methylated (Bird, 1986; Smith

and Meissner, 2013). The remaining fraction of long CpG stretches in the genome is usually hypermethylated in healthy, somatic cells and plays an important role in processes like X chromosome inactivation and genomic imprinting but also in cancerogenesis and neurodegeneration, where aberrant DNA methylation patterns are frequently observed (Bird, 2002; Klutstein *et al.*, 2016; Sanchez-Mut *et al.*, 2016). Consequently, repetitive and transposable elements within the human genome, like retroviruses or LINE1 and *Alu* elements are normally highly methylated to prevent their reactivation or in general their transcription (Jones, 2012). Moreover, DNA methylation contributes to the integrity of heterochromatin by serving as a binding platform for various proteins that can recognize and bind 5mC, like the members of the Mbd (methyl-CpG binding domain)-protein family (Ludwig, Zhang, and Cardoso, 2016). By modulating chromatin structure, 5mC has a direct effect on DNA transcription fidelity and can thereby affect alternative splicing. Consistently, DNA methylation and accompanying nucleosome occupancy in exons results in decreased binding of the architectural protein CTCF (CCCTC-binding factor), which in turn leads to exclusion of weak upstream exons (Shukla *et al.*, 2011; Marina *et al.*, 2015).
DNA cytosine methylation is catalysed by members of the DNA (cytosine-5)-methyltransferase (Dnmt) protein family that all share a highly conserved C-terminal catalytic domain. Despite this homology, of the six protein family members, only Dnmt1, Dnmt3A, Dnmt3B and Dnmt3C were shown to catalytically act on DNA (Jurkowska, Jurkowski, and Jeltsch, 2011; Barau *et al.*, 2016) (**Figure 3A**). While the Dnmt3-like protein Dnmt3L serves as a catalytically inactive regulatory factor (Chedin, Lieber, and Hsieh, 2002), Dnmt2 or Trdmt1 targets RNA and methylates different tRNAs, thereby protecting them from ribonuclease cleavage (Goll *et al.*, 2006; Schaefer *et al.*, 2010).
All catalytically active Dnmts belong to the class I of methyltransferases and act on non-methylated cytosines in DNA or RNA by flipping its substrate out of the sugar-phosphate backbone (Song *et al.*, 2012) (**Figure 3B**). This enables them, to transfer a methyl group from the methyl donor molecule S-adenosyl-L-methionine (SAM) to the cytosine residue, creating 5mC and S-adenosyl-L-homocysteine (SAH) (Jurkowska, Jurkowski, and Jeltsch, 2011) (**Figure 3C**).

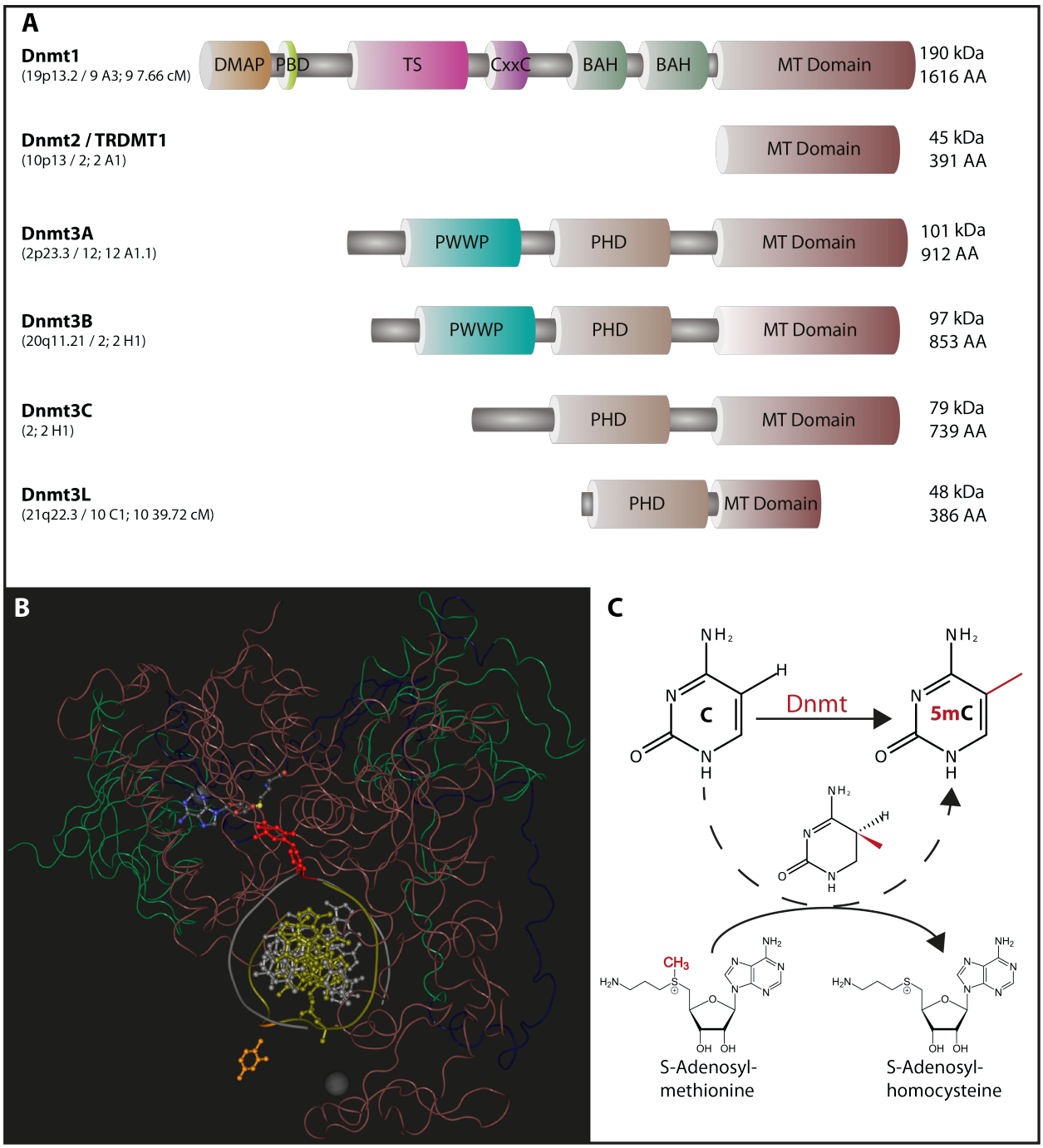


Figure 3: The Dnmt Methyltransferase protein family. (**A**) Domain organizations of human Dnmt1, Dnmt2, Dnmt3A, Dnmt3B, Dnmt3C and Dnmt3L including the DMAP (DMAP1 interacting site), PBD (PCNA binding domain), TS (Replication Foci Targeting Sequence), CxxC (CxxC Zinc-finger domain), BAH (Bromeo-adjacent homology), MT (methyltransferase domain). (**B**) Crystal structure of mouse DNMT1 (731-1602) bound to hemimethylated CpG DNA in complex with 5-methylcytosine (red), flipped out of the double helix and a 12 bp long stretch of double stranded DNA (golden and silver) at a resolution of 2.6 Å (PDB: 4DA4) (Song *et al.*, 2012). Zinc atom is displayed as grey sphere. The structural model was prepared with Cn3D (Wang *et al.*, 2000). (**C**) Illustration of the methylation reaction of Dnmt family proteins. The catalytic center of Dnmt proteins covalently bind cytosine and the methylgroup is transferred from the cofactor S-adenosyl-methionine to the C5 carbon atom of cytosine, leaving 5mC and demethylated S-adenosyl-homocysteine.

2.2.1 Dnmt3 *de novo* DNA methyltransferases

DNA cytosine methylation is initially established during embryogenesis and gametogenesis by the *de novo* DNA methyltransferases Dnmt3A and Dnmt3B, which both show a strong affinity for unmethylated DNA (Jurkowska, Jurkowski, and Jeltsch, 2011). In accordance with this, these two *de novo* DNA methyltransferases are highly expressed in embryonic stem cells and germ cells, but are down-regulated upon differentiation and in general in somatic tissues. However, Dnmt3A and Dnmt3B are not redundant in their biological function and relevance. While depletion of Dnmt3A results in embryonic lethality, Dnmt3B deficient mice die a couple of weeks after birth (Okano *et al.*, 1999; Ueda, 2006).
A third protein that resembles Dnmt3A and Dnmt3B structurally but is catalytically inert, Dnmt3L, was found to stimulate the catalytic activity of Dnmt3A and Dnmt3B (Chedin, Lieber, and Hsieh, 2002), for example by interacting with non-modified lysine 4 in the histone H3 tail (H3K4) and subsequently recruiting Dnmt3a for *de novo* methylation (Ooi *et al.*, 2007). The modification status of H3K4 furthermore regulates an autoinhibitory conformational change of Dnmt3A, which is further regulated by its interplay with the essential 5mC binding protein MeCP2 (Rajavelu *et al.*, 2018). In rodents, a third catalytically active *de novo* DNA methyltransferase, Dnmt3C, was described recently and shown to protect genome integrity of male germ cells by targeting the promoters of transposable elements and thereby silencing them (Barau *et al.*, 2016).

2.2.2 Dnmt1 maintenance methyltransferase

After DNA methylation patterns have been established by Dnmt3A and Dnmt3B they need to be maintained during the cell cycle to conserve the thereby transported epigenetic information. The mitotic inheritance of DNA methylation patterns is facilitated by the maintenance DNA methyltransferase Dnmt1. Its activity during DNA replication in S-phase prevents the enrichment of hemimethylated CpG sites and eventual loss of DNA methylation, due to the semi-conservative nature of DNA replication. In contrast to the comparatively short N-terminal domain of Dnmt3A and Dnmt3B, Dnmt1 harbours a long N-terminus with a multitude of additional regulatory domains. The importance of the N-terminal domains of Dnmt1 is underlined by the impaired catalytic activity of its C-terminal catalytic domain alone (Margot *et al.*, 2000). Furthermore its high preference

for hemimethylated DNA (Song *et al.*, 2011) and the localization to sites of ongoing DNA replication (Leonhardt *et al.*, 1992) is important for the role of Dnmt1 in DNA methylation maintenance during mitotic cell division. While Dnmt1 shows only a diffuse nuclear localization pattern during G1 and most parts of G2, a clear colocalization with the DNA replication machinery is observed throughout S-phase and in the transition from S-phase to G2 (Easwaran *et al.*, 2004). Furthermore, in contrast to Dnmt3A or Dnmt3B, Dnmt1 was shown to be recruited to DNA damage sites and hence plays a crucial role in maintaining epigenetic information not only during DNA replication, but also on newly synthesized DNA after successful repair (Mortusewicz *et al.*, 2005).

The catalytic activity of Dnmt1 is regulated by the concerted action of interacting proteins and the thereby modulated regulation of autoinhibitory conformational changes. Structural data and mutagenesis studies hint to a blocking of the catalytic pocket by the N-terminal CxxC- and TS-domains (Song *et al.*, 2011). Moreover, protein stability and therefore catalytic activity are regulated by ubiquitination and deubiquitination via Uhrf1 and Usp7, respectively. Besides regulating Dnmt1 protein stability, Uhrf1 plays a crucial role for Dnmt1 mediated DNA methylation maintenance, by stimulating its catalytic activity through direct protein-protein interactions and by binding and modifying the chromatin that surrounds hemimodified CpG sites (Bashtrykov *et al.*, 2013; Qin *et al.*, 2015).

2.3 The Uhrf protein family

Uhrf1 is the founding member of the Uhrf protein family and essential for the correct function of Dnmt1 during DNA methylation maintenance by multiple, different interactions with Dnmt1, methylated cytosine as well as histones in the vicinity of hemimethylated DNA (Arita *et al.*, 2008; Avvakumov *et al.*, 2008; Hashimoto *et al.*, 2008; Qin *et al.*, 2015). Uhrf1 was first discovered by two independent studies that set out to identify proteins that are recruited during S-phase or harbour a CCAAT-box binding motif, respectively (Fujimori *et al.*, 1998; Hopfner *et al.*, 2000). The human proteins Uhrf1 or ICBP90 (inverted CCAAT Box-binding Protein of 90 kDa) and the later identified Uhrf2 or NIRF (Np95/ICBP90 Ring Finger) show high sequence homology to their respective mouse orthologs Np95 (Nuclear protein of 95 kDa; mUhrf1) and Np97 (Nuclear protein of 97 kDa; mUhrf2), as they all share the same functional domains. At the very N-terminus resides the UBL (Ubiquitin-like)

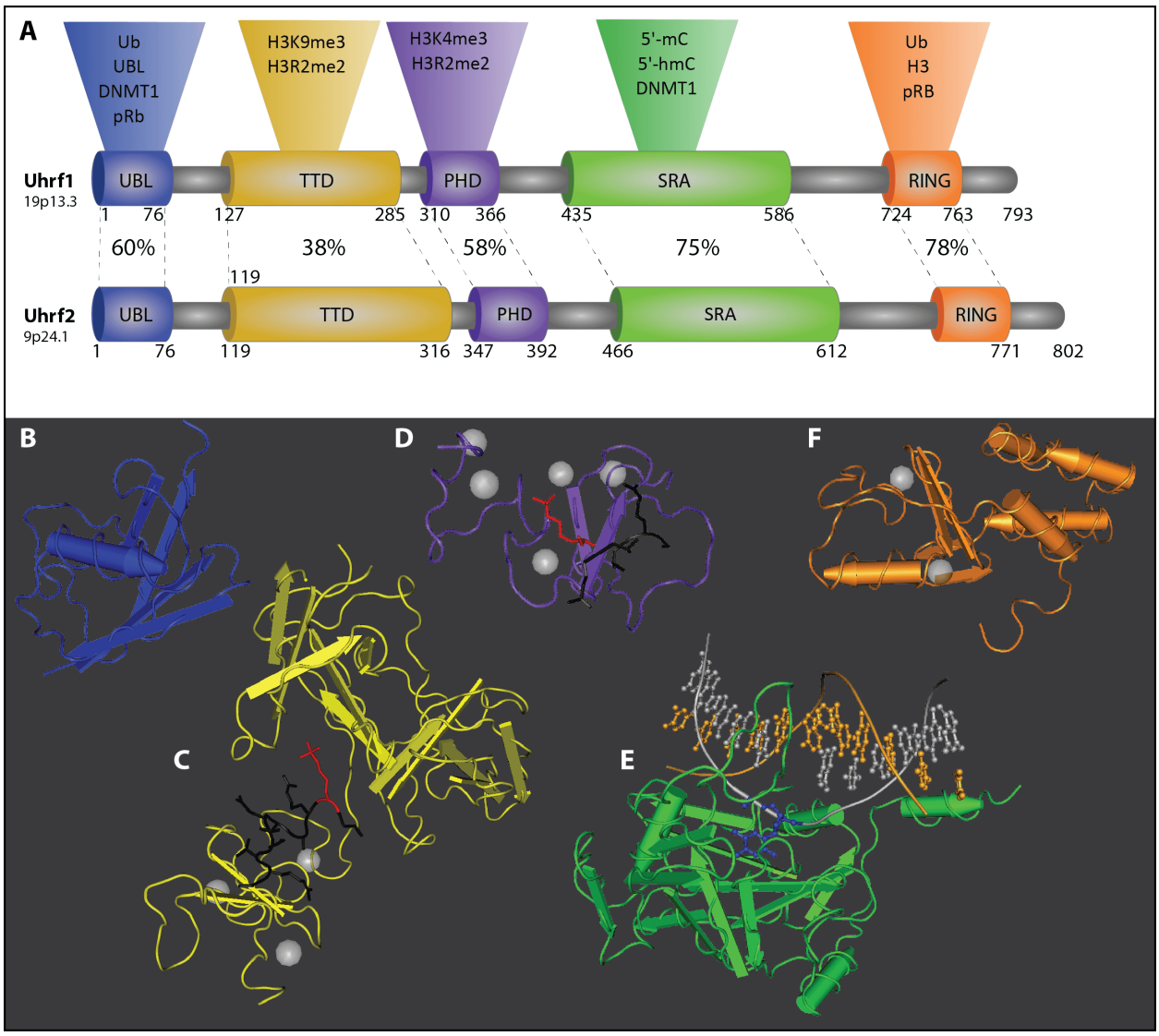


Figure 4: Uhrf1 and Uhrf2 are multidomain proteins with modules for DNA and histone interactions. (**A**) Domain organizations of human Uhrf1 and Uhrf2, including the ubiquitin-like (UBL) domain, histone modifications binding tandem tudor domain (TTD) and plant homeodomain (PHD), as well as the DNA interacting SET and RING associated (SRA) domain and the really interesting new gene (RING) domain, which harbours an E3-ligase activity. (**B**) crystal structure of the human Uhrf1 UBL domain at 2 Å (PDB: 2FAZ) (**C**) Crystal structure of the human Uhrf1 TTD domain in complex with trimethylated lysine 9 (red) of the histone H3 tail peptide at a resolution of 2.96 Å (PDB: 4GY5_A) (Cheng *et al.*, 2013) (**D**) Crystal structure of the human Uhrf1 PHD domain in complex with trimethylated lysine 4 (red) of the histone H3 tail petide (black) at a resolution of 1.9501 Å (PDB: 3SOW_A) (Rajakumara *et al.*, 2011). (**E**) Crystal structure of the SRA domain of human Uhrf1 in complex with 5-methylcytosine (red), flipped out of the double helix and a 12 bp long stretch of double stranded DNA (golden and silver) at a resolution of 2.2 Å (PDB: 3CLZ_A) (Avvakumov *et al.*, 2008) (**F**) Crystal structure of the RING domain of human Uhrf1 at a resolution of 1.74 Å (PDB: 3FL2_A); Zinc atoms are displayed as light spheres, alpha helices are displayed as barrel shaped arrows, beta sheets are displayed as flattened arrows. All structural models were prepared with Cn3D (Wang *et al.*, 2000).

domain followed by the histone binding tandem tudor domain (TTD) and plant homeodomain (PHD), the DNA-interacting SRA (SET and RING Associated) domain and the

E3-ligase activity mediating RING (Really Interesting New Gene) domain (**Figure 4A**).
The UBL domain shows 35% sequence homology with ubiquitin, as well as the typical ubiquitin alpha/beta fold and was proposed to interact with the proteasome and target ubiquitinated proteins for degradation. Recently, the UBL domain was shown to stimulate the activity of the C-terminal, E3-ligase activity mediating, RING domain (Foster *et al.*, 2018). In line with these findings, it was demonstrated that hUhrf2 rescues primary neurons from cytotoxic polyglutamine aggregates by targeting them for proteasomal degradation (Iwata *et al.*, 2009).
As previously mentioned, the regulation of Dnmt1 activity, is one of the most important biological functions of Uhrf1. Consistently, it was shown that the UBL domain of Uhrf1 directly interacts with the TS domain of Dnmt1, not only by targeting it for methylation maintenance but also by allosterically stimulating its catalytic activity (Li *et al.*, 2018).
Furthermore, it was shown that overexpression of Uhrf1 enhances ubiquitination levels of Dnmt1 *in vivo* (Qin, Leonhardt, and Spada, 2011) and also of histone H3, with a high preference over the other histones (Citterio, Papait, and Nicassio, 2004; Karagianni *et al.*, 2008). This dual ubiquitination was meanwhile linked to the regulation of Dnmt1 in DNA methylation maintenance during DNA replication (Qin *et al.*, 2015). In addition, Uhrf proteins were shown to be involved in cell cycle control by ubiquitinating PCNP *in vivo* and *in vitro* (Mori *et al.*, 2004).
The PHD domain is thought to be involved in gene activation and silencing via binding to different histone modifications. Consequently it has been shown that H3K4me3, a hallmark of transcriptionally active chromatin, is recognized and bound by the PHD domain with high specificity and can even be distinguished from the dimethylated variant of H3K4 (Wysocka *et al.*, 2006; Li *et al.*, 2006; Pena *et al.*, 2009). It was furthermore shown that the PHD domain can bind H3R2me2, a characteristic modification of active promoters (Rajakumara et al., 2011; Yuan et al., 2012). In this regard, the TTD domain was found to recognize the methylation state of H3K9 and H3R2 and facilitate binding to this residues together with the PHD domain (Arita *et al.*, 2012; Cheng *et al.*, 2013). The TTD domain was furthermore shown to bind H3K9me3 but not histone H3 tail acetylations (Rottach *et al.*, 2010). Recently, the TTD of Uhrf1 was found to recognize a H3K9me3-mimic peptide in ligase1, which targets it for methylation maintenance during DNA replication of pericentric

heterochromatin (Ferry *et al.*, 2017). The SRA (SET and RING associated) domain has in mammals only been described in Uhrf1 and Uhrf2. In plants, however, it is found in a variety of histone methyltransferases involved in epigenetic silencing (Baumbusch *et al.*, 2001). The SRA domain was shown to bind 5mC and also 5hmC (Frauer *et al.*, 2011b; Unoki, Nishidate, and Nakamura, 2004). It is able to flip methylated cytosine out of the DNA helix at hemimethylated sites and shows a up to 10 fold higher affinity for hemimethylated DNA compared to fully methylated DNA (Sharif *et al.*, 2007; Bostick *et al.*, 2007; Hashimoto *et al.*, 2008). Furthermore, the SRA domain of Uhrf1 was shown to stimulate the methylation activity of Dnmt1, most likely by allosterically removing the TS domain from the catalytic pocket and thereby allowing for DNA-binding and consequently, catalytic activity (Berkyurek *et al.*, 2014).

Albeit the highly similar domain structures of Uhrf1 and Uhrf2, their respective expression patterns and biological functions are far from redundant (Bronner *et al.*, 2007). While Uhrf1 was found do bind to 5mC and 5hmC with similar affinities (Frauer *et al.*, 2011b), Uhrf2 was implicated in only recognizing and binding to 5hmC (Spruijt *et al.*, 2013a). Although, Uhrf1 and also Uhrf2 were found to interact with Dnmt1, Dnmt3A and Dnmt3B and also the protein methyltransferase G9a in general, a S-phase dependent Dnmt1 interaction was only shown for Uhrf1 (Zhang *et al.*, 2011). Accordingly, one of the most important functions of Uhrf1 in cellular homeostasis, the targeting of Dnmt1 for methylation maintenance, can not be rescued by Uhrf2 in an Uhrf1 depleted background (Pichler *et al.*, 2011). This was explained with conformational constraints of Uhrf2 that arise from a different chromatin modification readout, compared to Uhrf1, and subsequent failure to ubiquitinate lysines in the histone H3 tail (Vaughan *et al.*, 2018). In stark contrast to the important role of Uhrf1 in DNA methylation maintenance, it was found that cancer-associated overexpression of Uhrf1 and also Uhrf2 is correlated with ubiquitination and subsequent proteasomal degradation of Dnmt3A and consequently global DNA hypomethylation (Jia *et al.*, 2016).

2.4 5mC oxidation and the Tet protein family

The already in 1972 described sixth mammalian DNA base 5-hydroxymethylcytosine was rediscovered in 2009, when it was found to be abundant in mouse brain and there especially in Purkinje neurons and granule cells (Kriaucionis and Heintz, 2009). A simultaneously

published study investigated homologous proteins of the trypanosomal thymine dioxygenases JBP1 and JBP2 (J-binding protein 1/2). Here, three mammalian JBP paralogs, Tet1, Tet2 and Tet3 were identified and Tet1 was shown to convert 5mC to 5hmC in an Fe(II)- and 2-oxoglutarate (2OG) dependent manner (Tahiliani *et al.*, 2009). The name Tet1 (Ten-eleven-translocation) was derived from its role as a t(10;11)(q22;q23) chromosomal translocation partner of the histone methyltransferase MLL/KMT2A (Mixed lineage leukemia/lysine methyltransferase 2A) in acute myeloid leukemia. This translocation results in a fusion protein that harbours the N-terminus of MLL and a C-terminal part of Tet1 (Ono, Taki, and Taketani, 2002; Lorsbach *et al.*, 2003).

The picture of Tet catalytic activity was completed by studies that found 5hmC to be further oxidised to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) in a Tet-dependent manner (Ito *et al.*, 2011; Pfaffeneder *et al.*, 2011). The oxidised forms of 5mC can serve as substrates for different proteins that are involved in base excision repair (BER), like Tdg or the Neil glycosylases (Maiti and Drohat, 2011; Müller *et al.*, 2014). Hence, it was initially assumed that 5hmC, 5fC and 5caC are primarily intermediates in a base excision repair mediated 5mC removal and accompanied loss of DNA methylation. In this line, the finding that Tet proteins can also oxidise thymine to 5-hydroxymethyluracil (5hmU) was in part attributed to mismatch induced DNA repair and an accompanied increase in 5mC removal (Pfaffeneder *et al.*, 2014) (**Figure 5C**). However, 5hmC and 5fC have meanwhile been identified to be stable DNA base modifications, whose levels fluctuate throughout the cell cycle and peak in late S-phase and G2, similar to the dynamics that are observed for 5mC in normal proliferating cells (Bachman *et al.*, 2014; Bachman *et al.*, 2015). In addition, 5caC was assigned an important signalling function in adult neurons where it regulates neurodegeneration (Jin *et al.*, 2016). Besides their signalling functions, the 5mC oxidation products have been implicated in altering the mechanical properties of the DNA double-helix. While 5mC stiffens the DNA strand, its oxidised forms, especially 5hmC and 5fC, can markedly increase DNA flexibility (Ngo *et al.*, 2016).

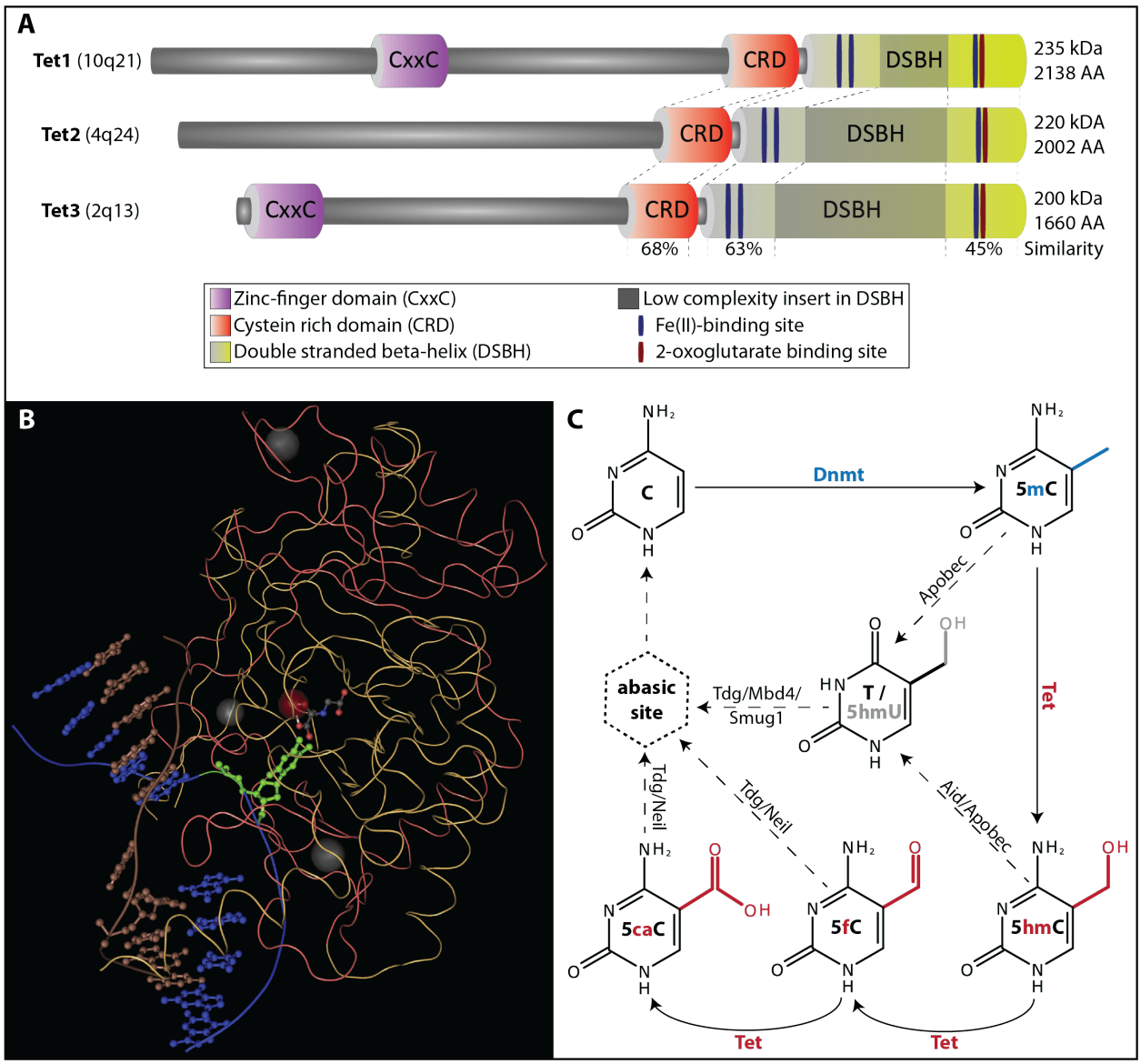


Figure 5: The Tet dioxygenase protein family. (**A**) Domain organizations of the long isoforms of human Tet1, Tet2 and Tet3. Shown are the N-terminal zinc-finger domain (CxxC), the cystein-rich domain (CRD) and the double stranded beta-helix (DSBH) which is interrupted by a low complexity insert and harbours the Fe(II)- and 2-oxoglutarate binding sites. (**B**) Crystal structure of human Tet2 bound to a 12 base pair long stretch of methylated DNA at a resolution of 2.03 Å (PDB 4NM6) (Hu *et al.*, 2013). The CRD is shown as faint red tube worms and the DSBH as faint yellow tube worms. The bound DNA is depicted as blue and brown stick representation helix with the methylated cytosine flipped out of the double helix shown in green. Zinc and iron atoms are shown as grey and red spheres, respectively. The 2-oxoglutarate analog, N-oxalylglycine is shown in stick representation. The structural model was visualized with Cn3D (Wang *et al.*, 2000). (**C**) Overview of the different cytosine modifications and the involved proteins. Cytosine is initially methylated by DNA methyltransferases (Dnmt) and is then oxidised by Tet proteins to 5hmC, 5fC and 5caC in an iterative manner, 2-OG and Fe(II)-dependent manner. 5mC or 5hmC can be deaminated by Apobec (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like)deaminases and 5hmC additionally by Aid (Activation-induced deaminase), creating thymidine (T) or 5-hydroxymethyluracil (5hmU), respectively. 5fC and 5caC as well as 5hmU and mismatched T are recognized and excised by different glycosylases involved in the base excision repair (BER) pathway, like Mbd4 (Methyl-CpG binding domain protein 4), Neil glycosylases, Smug1 (Single-strand selective monofunctional uracil DNA glycosylase) or Tdg (Thymine DNA glycosylase), eventually creating non-modified cytosine. The oxidised forms of 5mC can however transport epigenetic information and are not necessarily removed immediately.

2.4.1 The Tet dioxygenase family

All three members of the Tet protein family share high sequence similarities in their C-terminal catalytic domain, which is sufficient for their respective catalytic activity. This catalytic domain comprises the cysteine-rich domain (CRD) and the split double-stranded beta helix domain (DSBH) which harbours the Fe(II) and 2-oxoglutarate cofactor binding sites and is interrupted by a low-complexity insert (**Figure 5A**). Structural studies have shown that the CRD and the DSBH form a compact catalytic core, in which a conserved sequence motif within the CRD stabilizes the substrate-containing DNA double helix, while 5mC or its oxidised derivatives are flipped out of the DNA double helix (Hu *et al.*, 2013) (**Figure 5B**). Besides this, two recent studies identified monoubiquitination of a conserved lysine residue within the CRD of all three Tet proteins by the CRL4(VprBP) E3-ligase complex to modulate their catalytic activity and chromatin association (Yu *et al.*, 2013; Nakagawa *et al.*, 2015).

A structural feature that separates the three Tet proteins from one another is their N-terminal CxxC domain. Although the zinc-finger domains of Tet1 and Tet3 as well as the Tet2 associated IDAX/CXXC4, belong to the same subgroup of CxxC-domains (Xu *et al.*, 2018), they differ highly in their respective binding affinities and substrates. Tet2 lost its zinc-finger during evolution in a chromosomal inversion and it is now encoded by the genomically adjacent IDAX/CXXC4, which negatively regulates Tet2 expression and modulates its catalytic activity (Ko *et al.*, 2013). Tet1 and Tet3, in contrast, both kept their respective zinc-finger domains and the CxxC domain of Tet3 was shown to preferentially bind 5caC whereby it regulates the expression of lysosomal proteins in postmitotic neurons (Jin *et al.*, 2016). However, Tet1 CxxC domain was found to mostly bind non-modified DNA (Frauer *et al.*, 2011a), which was implicated in the prevention of DNA methylation spreading into unmethylated, euchromatic regions (Jin *et al.*, 2014). Besides this, the Tet1 CxxC was furthermore implicated in transcriptional fine tuning by facilitating Tet1 association with transcriptional start sites within CpG rich promoters and genic sequences. In line with these findings that assign Tet1 a role in transcriptional regulation, it was also found to interact with the Sin3a co-repressor complex and the polycomb repressive complex 2 (PRC2) in embryonic stem cells (Williams *et al.*, 2011; Neri *et al.*, 2013).

While three different isoforms of Tet2 and also Tet3 have been characterized to date

(Langemeijer *et al.*, 2009; Jin *et al.*, 2016), a N-terminally truncated Tet1 isoform, which also lacks the CxxC domain, was described just recently and attributed a role in mouse development, but also in cancerogenesis (Zhang *et al.*, 2016b; Good *et al.*, 2017).
Despite their high similarities in catalytic activity, major differences between the three Tet protein family members are observed in their respective expression levels throughout different tissues and developmental stages and hence their physiological roles. While Tet3 is predominantly expressed during early embryogenesis but also in post-mitotic neurons, Tet1 and Tet2 are found to be more ubiquitously expressed across different tissues from embryonic stem cells to somatic cells (Santiago *et al.*, 2014).
Although Tet proteins were found to be degraded in an ubiquitination-independent pathway via calpain proteases (Wang and Zhang, 2014), it was recently discovered that the cellular presence of viral protein R (Vpr) after HIV infection modulates the ubiquitination pattern of a conserved lysine residue within the CRD of Tet2. While the monoubiquitination of said lysine positively stimulates the chromatin association of Tet2 and also the activity of Tet1 and Tet3 (Nakagawa *et al.*, 2015), Vpr presence results in a polyubiquitination of said lysine and consequently proteasomal degradation (Lv *et al.*, 2018). This leads to increased HIV-1 replication through de-repression of interleukin 6 that is targetd by HDAC1/2 via Tet2 recruitment in non-infected cells.
Additionally, the activity of Tet proteins is not restricted to DNA. While already the first study to describe 5hmC in mammalian DNA hinted to its occurrence in RNA from crude rat tissue preparations (Penn *et al.*, 1972), further evidence for this was found recently. Accordingly, mammalian Tet proteins were shown to act not only on DNA but also to oxidise 5mC in RNA and it was demonstrated that ectopic overexpression of either of the three Tet proteins leads to a 5hmC increase in RNA (Fu *et al.*, 2014). It was furthermore found that ascorbic acid, which is known to enhance Tet activity, can increase the levels of 5hmC and also 5fC not only in DNA but also in RNA (Zhang *et al.*, 2016a). In line with these findings, recent studies identified a Tet homologue in *Drosophila melanogaster*, called dTet, whose knockout resulted in impaired brain development and late pupal lethality (Delatte *et al.*, 2016; Wang *et al.*, 2018). Despite the almost complete absence of 5mC or 5hmC in *Drosophila* DNA (Capuano *et al.*, 2014), a 5hmC enrichment in polyadenylated mRNA and accompanied positive effect on translation was demonstrated (Delatte *et al.*, 2016).

2.5 Cytosine modifications in development and diseases

The loss of *de novo* DNA methyltransferases Dnmt3A and Dnmt3B or the maintenance methyltransferase Dnmt1 and the accompanying embryonic and early postnatal lethality in mice highlights the importance of cytosine methylation for normal development in mammals (Li, Bestor, and Jaenisch, 1992; Okano *et al.*, 1999; Ueda, 2006). However, not only DNA methylation but also its removal in the right context is crucial during mammalian development. Before the discovery of the Tet proteins and their catalytic activity, loss of DNA methylation was considered a purely passive process, resulting from prevention of DNA methylation maintenance by Dnmt1. Consistently, decreasing Dnmt1 levels or disruption of its interaction with PCNA and Uhrf1 were shown to result in hypomethylation and even death of corresponding mouse models (Gaudet *et al.*, 2003; Hervouet *et al.*, 2010). Loss of methylation in the maternal pronucleus was initially attributed to replication depended dilution effects, since Dnmt1 is retained in the cytoplasm during early embryogenesis (Cardoso and Leonhardt, 1999). Meanwhile it could be shown that retention of Dnmt1 and accompanied replication dependent dilution effects, but also Tet3 mediated 5mC oxidation concomitantly contribute to the loss of DNA methylation in the maternal, but also the paternal pronucleus (Shen *et al.*, 2014).

After the removal of 5mC during fertilization, DNA methylation marks have to be set again by Dnmt3A and Dnmt3B to ensure faithful embryonic development. While Dnmt3A was shown to play an important role in setting parental imprints (Okano *et al.*, 1999), Dnmt3B was linked to the correct methylation pattern of pericentric DNA regions and hence chromosome structure (Kaneda *et al.*, 2004). Accordingly, mutations of Dnmt3B, whose loss results in embryonic death, were linked to the ICF (Immunodeficiency, Centromere instability, Facial abnormalities) syndrome. This disease is, on a molecular level, characterized by hypomethylation of the normally highly methylated pericentric repeats in human chromosomes 1, 9 and 16 (Xu *et al.*, 1999).

Similar to the lethality observed in embryos lacking Dnmt3B, failure to maintain the DNA methylome by Dnmt1 also results in non-viable embryos (Li, Bestor, and Jaenisch, 1992). Mutations in the TS domain of Dnmt1, repeatedly found in the autosomal neurodegenera-

tive disease HSN1E (hereditary sensory and autonomic neuropathy with dementia and hearing loss type 1E) abrogate the interaction with its facilitating factor Uhrf1 and thereby DNA methylation maintenance (Smets *et al.*, 2017). Furthermore, mutations or aberrant expression patterns of Uhrf1 are often associated with cancer phenotypes and a constant overexpression of Uhrf1 is described in many malignant neoplasms. Even though, Uhrf1 plays an important role in DNA methylation maintenance and the accompanied silencing of repetitive and retrotransposable elements (Bostick *et al.*, 2007), a recurring loss of DNA methylation in LINE1 is observed in cancer cells that overexpress Uhrf1 (Nakamura *et al.*, 2016; Hong *et al.*, 2018). The overexpression of Uhrf1 and also its homolog Uhrf2 in cancer cells and the accompanied decrease of DNA methylation levels was attributed to the E3-ligase targeting and subsequent proteasomal degradation of Dnmt3A (Jia *et al.*, 2016). Similar to cancer cells, a loss of DNA methylation and accompanied reactivation of retroviral elements is observed in neural stem cells upon loss of Uhrf1 which results in impaired neonatal neurodegeneration (Ramesh *et al.*, 2016).

Despite their different roles during earliest embryonic development and gametogenesis, where mostly Tet3 plays a role, all three Tet family members are expressed throughout neurogenesis and in the adult brain, where the highest 5hmC levels are detected in comparison to other tissues. 5hmC is in general the most abundant 5mC oxidation product with an abundance of around 0.5% of all cytosines in cells of the central nervous system and around 0.05% of all cytosines in different internal organs (Globisch *et al.*, 2010). 5fC and 5caC, on the other hand, are found with up to 100-fold lower levels than 5hmC and often fall below the detection limit (Bachman *et al.*, 2015).

Tet3 is highly expressed in preimplantation embryos but rapidly downregulated with progressing embryonic development (Gu *et al.*, 2011). In contrast to Tet1 and Tet2, whose individual or combined loss has no direct effects on embryonic development and manifests only during advanced postnatal development, loss of Tet3 is lethal in neonatal mice (Gu *et al.*, 2011; Dawlaty *et al.*, 2011). Despite the dispensable biological role for Tet1 and Tet2 in early embryonic development, their loss can have severe effects in postnatal organisms. Hematopoietic stem cells, for instance, were found to frequently give rise to myeloid and lymphoid abnormalities upon loss of Tet1 but mostly Tet2 (Ko *et al.*, 2011; Li *et al.*, 2011). Tet1, on the other hand, was also assigned an important role in the maintenance of neural

progenitor cells (Zhang *et al.*, 2013) and its loss was found to impair the expression of activity-regulated genes and synaptic plasticity (Rudenko *et al.*, 2013).

Table 1: Phenotypes of cytosine modifier knockout mice

Phenotypes of cytosine modifier knockout mice		
Dnmt3A null mice	Postnatal growth retardation and dysplasia neonatal mortality	(Okano *et al.*, 1999)
Dnmt3B null mice	Not viable	(Okano *et al.*, 1999; Ueda, 2006)
Dnmt3C null mice	Hypogonadism and aberrant retrotransposon expression	(Barau *et al.*, 2016)
Dnmt3L null mice	Early embryonic lethality in females and sterile males	(Hata *et al.*, 2002)
Dnmt1 null mice	Not viable	(Li, Bestor, and Jaenisch, 1992)
Tet1 null mice	Mostly normal embryonic and postnatal development abnormal long-term depression	(Dawlaty *et al.*, 2014) (Rudenko *et al.*, 2013)
Tet2 null mice	Increased development of myeloid malignancies and early life mortality	(Ko *et al.*, 2011; Li *et al.*, 2011)
Tet3 null mice	Normal embryonic development but neonatal mortality	(Gu *et al.*, 2011; Tsukada, Akiyama, and Nakayama, 2015)
Uhrf1 null mice	Embryonic growth retardation and ultimately lethality Global genomic hypomethylation Development to expanded blastocyst stage severely impaired upon maternal KO	(Sharif *et al.*, 2007) (Bostick *et al.*, 2007) (Maenohara *et al.*, 2017)
Uhrf2 null mice	Viable but increased risk of spontaneous seizures	(Liu *et al.*, 2017)

2.6 DNA replication and epigenome maintenance

As just discussed, epigenetic DNA modifications can severely impact mammalian embryonic and also postnatal development, which makes the faithful propagation of epigenetic modifications during mitotic cell division a crucial step in the cell cycle. To ensure the correct inheritance of epigenetic modifications to a daughter cell, epigenome replication is tightly coupled to the replication of DNA itself.

Genome duplication in eukaryotic and especially mammalian cells before mitotic cell division is a temporally and spatially tightly orchestrated process. The foundation for DNA replication during S-phase is laid already during late mitosis and in G1 phase, when the MCM2-7 helicase complex is loaded onto DNA at origins of replication. These origins are distributed throughout the whole nucleus, which ensures replication of the entire genome. Upon being licensed and starting replication, an origin becomes inactive to avoid re-replication of DNA (Blow and Dutta, 2005). The transition from MCM2-7 chromatin loading to origin licensing and eventually firing as well as the prevention of re-licensing in S-phase is tightly regulated by cyclin-dependent kinases (CDK) (Satyanarayana and Kaldis, 2009). Upon initiation of DNA replication from an origin, replication forks move away bidirectionally from the origin. A plethora of proteins concomitantly ensures the progression of the replication fork and faithful DNA duplication. Amongst them are helicases, different DNA polymerases, DNA single strand binding proteins and the processivity factor PCNA (Proliferating cell nuclear antigen) that serves as a loading platform for proteins associated with DNA replication, but also epigenetic inheritance (Chagin, Stear, and Cardoso, 2010; Moldovan, Pfander, and Jentsch, 2007).
In somatic cells, S-phase progression follows the gene-density and GC-content of the underlying DNA (Costantini and Bernardi, 2008). In early S-phase, gene-rich and transcriptionally active euchromatin is replicated, while in mid S-phase, replication of facultative heterochromatin takes place. Facultative heterochromatin comprises silenced genic regions and can be highly variable between different cells. Constitutive heterochromatin is highly methylated and replicated in late S-phase in somatic cells. It comprises many repetitive sequences like tandem satellite repeats and furthermore fulfils structural tasks (Saksouk, Simboeck, and Déjardin, 2015). Large heterochromatic clusters, so-called chromocenters, are observed in DNA-stainings of murine cells, which are characterized by large focal patterns during DNA replication (**Figure 6**).

However, not only the DNA but also associated proteins and chromatin modifications have to be duplicated. The most prominent DNA-associated proteins are the core histones, which make up the nucleosome, hence the fundamental structure of chromatin organization. It is by now well established that the replicated nascent DNA strands receive newly

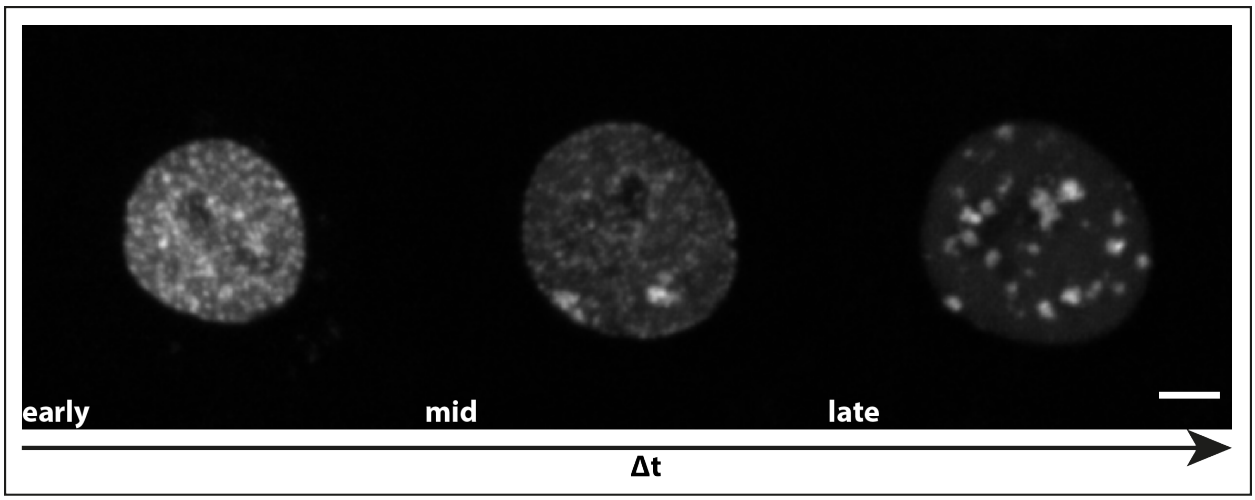


Figure 6: Progression of DNA replication in somatic cells. Shown are maximum Z-projections of a C2C12 mouse myoblasts that was transiently transfected with mRFP-PCNA and subjected to live-cell time lapse microscopy. In early S-phase, euchromatic DNA is replicated, which is characterized by a homogeneous distribution of PCNA foci. Facultative heterochromatin is replicated in mid S-phase, which shows PCNA patterns in the nuclear and nucleolar periphery. Furthermore, in female cells, like the shown tetraploid mouse myoblasts, the two inactive X-chromsomes are replicated in mid S-phase and are visible as large, synchronously replicating structures. During late S-phase, constitutive heterochromatin which forms large clusters in murine cells, so-called chromocenters, is replicated. Scale bar = 5 µm.

synthesized and recycled parental histones in equal amounts (Campos, Stafford, and Reinberg, 2014) ensuring the correct inheritance of post-translational histone modifications. For this purpose, histones H3 and H4, which contain a majority of the regulatory PTMs (Rando, 2007), from the parental DNA strand are divided symmetrically by the replicative helicase MCM2 with a minor bias for the leading strand over the lagging strand (Petryk *et al.*, 2018).

Besides histones and DNA itself, also DNA base modifications have to be duplicated, as all known mammalian DNA modifications are added post-synthetically.

2.6.1 Maintaining the DNA methylome

Considerable efforts were undertaken to unravel the cell cycle dependent inheritance of the repressive DNA modification 5mC. The most important players in maintaining 5mC are the just introduced maintenance DNA methyltransferase Dnmt1 and its facilitating factor Uhrf1. Although, Dnmt1 harbours an N-terminal PCNA binding domain, association with the replication machinery via PCNA, was shown to be dispensable for DNA methylation maintenance (Spada *et al.*, 2007). Due to an autoinhibitory conformation, Dnmt1 needs a

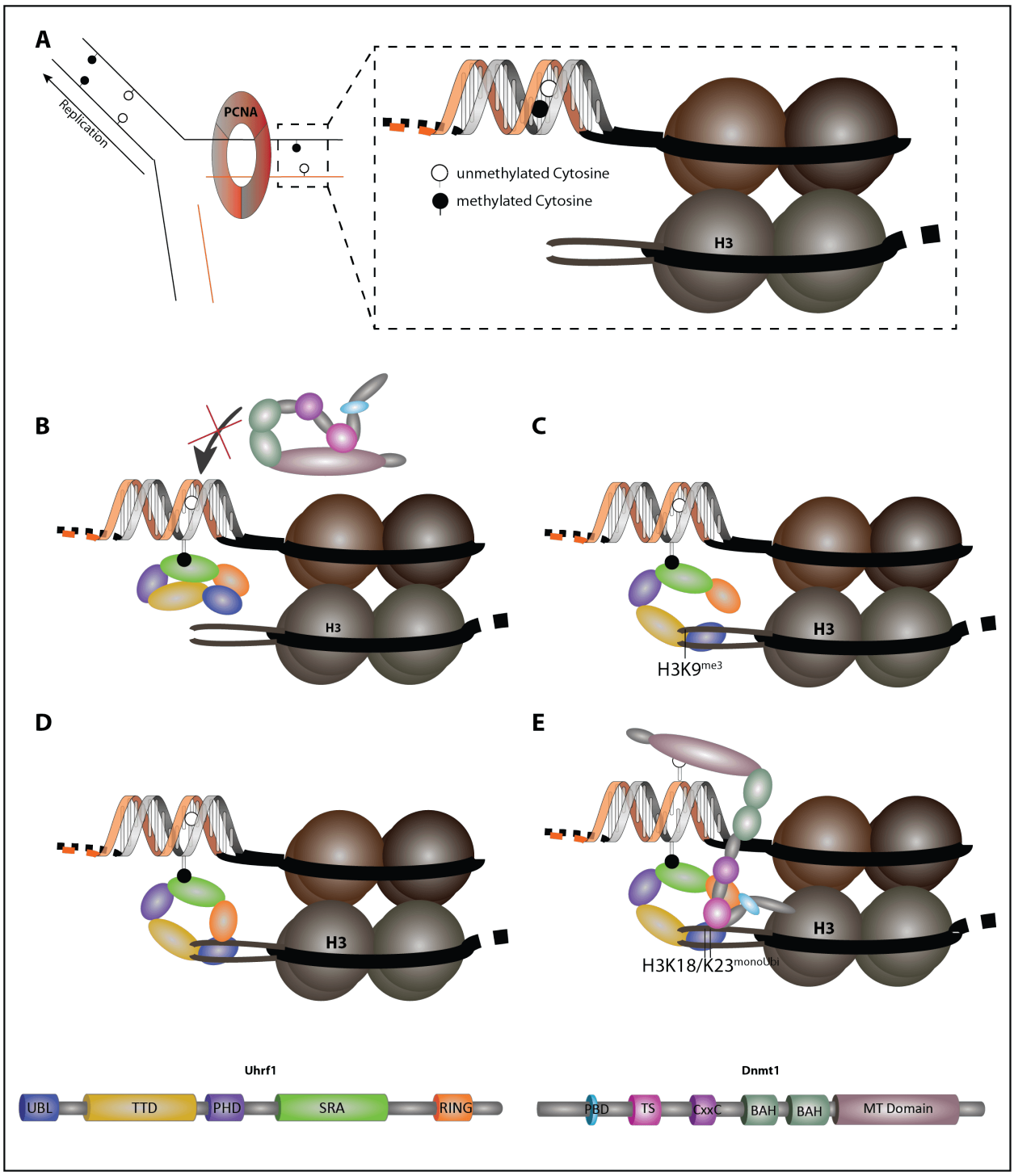


Figure 7: Molecular mechanism of DNA methylation maintenance. (**A**) Schematic, simplified overview of a DNA replication fork and a detailed view of newly synthesized piece of DNA, harbouring a hemimethylated CpG adjacent to a nucleosome. (**B**) Dnmt1 access to the unmethylated cytosine is prevented by its autoinhibitory conformation. Uhrf1 recognizes and binds the opposing methylated cytosine via its SRA domain. (**C**) Binding to methylated cytosine in a hemimethylated CpG-context triggers a conformational change of Uhrf1, which in turn binds H3K9me3. (**D**) Interaction with the histone H3 tail activates the E3 ubiquitin-ligase function of Uhrf1 and it deposits monoubiquitination marks on lysines 18 and 23 in histone H3, via the interplay of its UBL and RING domain. (**E**) Histone H3K18/K23 monoubiquitination and physical interaction with Uhrf1 triggers a conformational change of Dnmt1 and recruits it for DNA methylation maintenance.

regulatory cofactor to stimulate its enzymatic activity, which is facilitated by Uhrf1 (Bashtrykov *et al.*, 2013).

Uhrf1 is initially recruited to chromatin by recognizing hemimethylated DNA via its SRA domain. This binding triggers a conformational change, enabling it to bind H3K9me3 in the vicinity of hemimethylated DNA through its tandem tudor domain (Liu *et al.*, 2013b; Fang *et al.*, 2016). The thereby mediated binding to the histone H3 tail activates the E3-ligase activity of Uhrf1 and results in monoubiquitination of lysines 18 and 23 in the histone H3 tail (Qin *et al.*, 2015; Harrison *et al.*, 2016; Foster *et al.*, 2018). These monoubiquitinations are in turn bound by the TS domain of Dnmt1 and crucial for its maintenance methytransferase activity (Qin *et al.*, 2015). The protein-protein interaction of chromatin bound Uhrf1 and Dnmt1, via the Uhrf1 SRA- and UBL-domains, results in a conformational change in Dnmt1 and its recruitment. The TS domain of Dnmt1 allosterically blocks the catalytic center, rendering Dnmt1 inactive. This blocking is reversed by Uhrf1 binding to Dnmt1, which promotes its maintenance methyltransferase activity (Bashtrykov *et al.*, 2013; Li *et al.*, 2018) (**Figure 7**).

Although, it has been found that Uhrf1-recruitment for replication coupled methylation maintenance depends also on its binding to a trimethylated H3K9me3-mimicking peptide within ligase1, complete loss of ligase1 did not affect DNA methylation maintenance, suggesting a backup mechanism for this interaction (Ferry *et al.*, 2017).

As briefly outlined here, extensive efforts have been undertaken to elucidate the exact mechanism, how 5mC is inherited during mitotic cell division and the picture of this crucial biological mechanism becomes ever more complex. In contrast to this, the maintenance of the oxidative derivatives of 5mC remains elusive, although it was shown that 5hmC and 5fC are largely stable throughout the cell cycle and follow similar kinetics as 5mC (Bachman *et al.*, 2014; Bachman *et al.*, 2015).

3 Aim of the study

Although the proteins of the Tet dioxygenase family have been the subject of extensive research over the past years, the mechanisms by which they are recruited in the course of 5mC oxidation is still poorly understood. Furthermore, it remains elusive how Tet proteins maintain the 5mC oxidation products that were shown to be more than mere intermediates in the active removal of DNA methylation and to furthermore follow the cell cycle dynamics of 5mC.

With regard to these findings, I addressed the question of a potential role for the Tet proteins in DNA replication-coupled 5mC oxidation product maintenance and therefore addressed the following points:

i) The localization of the different Tet proteins during ongoing DNA replication

ii) Potential interactions with proteins that are implicated in (epi)genome replication

iii) Potential physiological consequences of a Tet recruitment in DNA replication

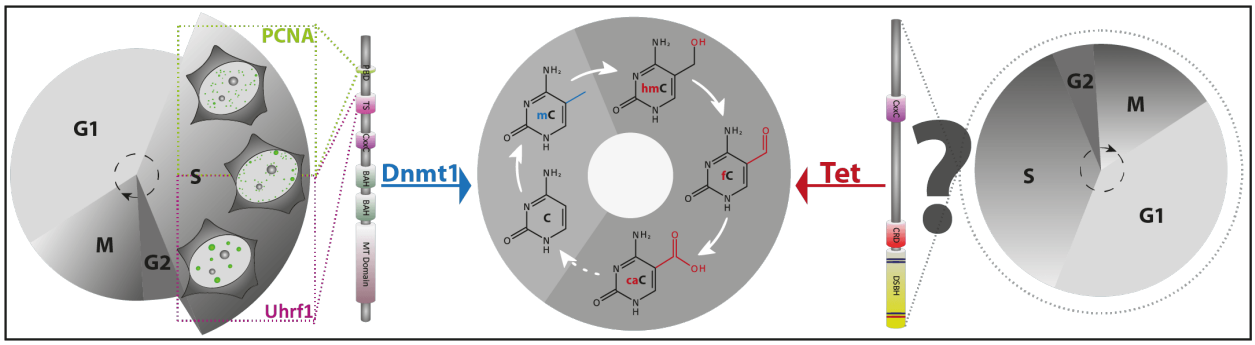


Figure 8: Aim of the study. The mechanism how DNA methylation is maintained throughout the cell cycle in proliferating cells is well established. The maintenance DNA methyltransferase Dnmt1 is recruited to sites of-going DNA replication predominantly by interactions with the multidomain protein Uhrf1 but also, with the clamp loader protein PCNA. Dnmt1 interaction with these two proteins ensures faithful inheritance of the DNA methylome in proliferating cells (**Section 2.6**). The members of the Tet dioxygenase family on the other hand, are known to oxidise 5mC to 5hmC, 5fC and 5caC in an iterative manner. The exact mechanism, how Tet proteins are recruited for catalytic activity or how the levels of these predominantly stable DNA modifications are maintained remains, however, elusive. The aim of this study was therefore, to characterize the cell cycle localization of Tet proteins, interactions with other proteins that are involved in (epi)genome maintenance and physiological consequences of a putative cell cycle dependent Tet recruitment.

4 Material and Methods

4.1 Expression plasmids

All mammalian expression vectors were obtained by using standard restriction enzyme- and PCR-based molecular cloning techniques, if not stated otherwise. All oligonucleotides that were used in the cloning procedures are given in table 1.

Plasmids encoding human EGFP- and mRFP-tagged PCNA have been described in previous publications (Leonhardt *et al.*, 2000; Sporbert *et al.*, 2005). To create miRFP-PCNA, the sequence of miRFP670 was amplified from pmiRFP670-N1 (Addgene plasmid # 79987) (Shcherbakova *et al.*, 2016)and used to replace the mRFP sequence in pENmRFPCNAL2 (Sporbert *et al.*, 2005) using BsrGI and AgeI restriction endonucleases sites.

EGFP- or mcherry-tagged catalytically active and inactive mouse Tet1CD constructs as well as mcherry-tagged catalytic domains of Tet2 and Tet3 and the three full-length proteins have been described before (Frauer *et al.*, 2011a; Liu *et al.*, 2013a; Müller *et al.*, 2014; Ludwig *et al.*, 2017; Zhang *et al.*, 2017c).

The EGFP-tagged short isoform of Tet1 was generated by amplifying the respective coding sequence from EGFP-Tet1 (Frauer *et al.*, 2011a) and replacing the full-length Tet1 sequence with the Tet1s sequence by restriction with AsiSI and NotI. Mcherry-tagged Tet1s was generated by replacing the coding-sequence of Uhrf1 from mcherry-Uhrf1 (Qin, Leonhardt, and Spada, 2011)with the Tet1s sequence by AsiSI and NotI restriction.

EGFP-tagged Tet1-CRD and Tet1-DSBH were generated by amplifying the respective domains from full-length Tet1 and replacing the EGFP-Tet1 (Frauer *et al.*, 2011a) sequence by AsiSI and NotI restriction.

EGFP-Tet1Δ1-389 and EGFP-Tet1Δ566-833 were obtained by overlap-extension PCR and the thereby generated products were used to replace the Tdg-coding sequence in EGFP-TDG (Müller *et al.*, 2014) with AsiSI and NotI restriction.

To obtain a EGFP-tagged peptide encoding the putative Tet1-PBD, oligo-cloning was performed and MCMT-PBD in an EGFP-N2 backbone (Mortusewicz *et al.*, 2005)was replaced by restriction with XmaI and EcoRI.

The CRD of EGFP-Tet1s or EGFP-Tet1 was deleted by overlap-extension PCR(Ho *et al.*, 1989) and the obtained amplicon was used to replace the Tet1 coding sequence in EGFP-

Tet1 (Frauer *et al.*, 2011a).

Lysine 852 in Tet1s was mutated to either arginine or glutamate using a slightly modified protocol of a previously described sequence and ligation independent cloning(SLIC) approach (Jeong *et al.*, 2012). In brief, two overlapping DNA fragments that contained the desired mutations in their respective complementary 5' and 3' regions were generated by PCR, using the Tet1s coding sequence as template. The Tet1 coding sequence was excised from EGFP-Tet1 (Frauer *et al.*, 2011a) with AsiSI and NotI and the remainder used as backbone. 100 ng of the backbone were mixed with the two amplicons in a ratio of 1:2:2 and a T4 polymerase reaction with 0.6 units of T4 polymerase was set up in a volume of 10 µL. The reaction was incubated at room temperature for 2.5 minutes and further incubated on ice for 10 minutes before being transformed into chemically competent Top10 cells.

EGFP-tagged Tet1CD harbouring the just described K to R or K to E mutations within the CRD was generated by PCR amplifying the respective Tet1CD coding sequence from the corresponding Tet1s vectors. The amplicons were than used to replace the Tet1 coding sequence from EGFP-Tet1 (Frauer *et al.*, 2011a).

Full-length mouse Uhrf1 fused to EGFP or mcherry and mouse EGFP-Uhrf2 constructs have been described before (Qin, Leonhardt, and Spada, 2011; Pichler *et al.*, 2011). Single domain deletion constructs or single domains of Uhrf1 fused to EGFP were described in previous studies (Rottach *et al.*, 2010; Pichler *et al.*, 2011; Qin *et al.*, 2015).

Mcherry-tagged VprBP was cloned from murine ESC cDNA by overlap extension PCR (Ho *et al.*, 1989) and the obtained PCR fragment used to replace the Uhrf1 sequence from mcherry-Uhrf1 (Qin, Leonhardt, and Spada, 2011) with AsiSI and NotI restriction.

A major satellite repeat recognizing polydactyl zinc-finger fused to EGFP or a EGFP-recognizing nanobody was described in previous publications (Lindhout *et al.*, 2007; Casas-Delucchi *et al.*, 2012b).

4.1.1 List of oligonucleotides

Table 2: List of oligonucleotides used for plasmid generation

Name	Oligonucleotide sequence (5' –> 3')
Tet1s_fwd	ATAAGCAGCGATCGCATGGACTGCAGTAGAC
Tet1_rev	ATAAGCAGCGGCCGCTTAGACCCAACGATTG
Tet1CD_fwd	AAAAGCGATCGCATGGAAGCTGCACCCTGTGACTG
Tet1CD_rev	CGTTAACTGCGGCCGCTTAGACCCAACG
Tet1-CRD_fwd	AAAAGCGATCGCATGGAAGCTGCACCCTGTGACTG
Tet1-CRD_rev	GAATGCGGCCGCTTACCCAAACTTACAGCC
Tet1-DSBH_fwd	AAGCGATCGCCCAAGGAAATTTAGGCTAC
Tet1-DSBH_rev	AAAAGCGGCCGCTTAGAGGGAACGATTG
EGFP-Tet1Δ1-389_fwd1	AAAAGCGATCGCATGGCTATTACTATGCTAAACCAA
EGFP-Tet1Δ1-389_rev1	TGCAGCTTCCTGTTCCTTCAATCCAGCTATCAGGTT
EGFP-Tet1Δ1-389_fwd2	ATAGCTGGATTGAAGGAACAGGAAGCTGCACCCTGT
EGFP-Tet1Δ1-389_rev2	GAATTCGTTAACTGCGGCCGC
EGFP-Tet1Δ1-389_fwdfinal	AAAAGCGATCGCATGGCTATTACTATGCTAAACCAA
EGFP-Tet1Δ1-389_revfinal	GAATTCGTTAACTGCGGCCGC
EGFP-Tet1Δ566-833_fwd1	ATGTCTCGGTCCCGCCCCGCAATGTCTCGGTCCCGCCCCGCA
EGFP-Tet1Δ566-833_rev1	AGGTCGACTCCCAGCCACCATTGGCGGCGTAGAATT
EGFP-Tet1Δ566-833_fwd2	ACGCCGCCAATGGTGGCTGGGAGTCGACCTCACCCT
EGFP-Tet1Δ566-833_rev2	TGCAGCTTCCTGTTCCTTCAATCCAGCTATCAGGTT
EGFP-Tet1Δ566-833_fwd3	ATAGCTGGATTGAAGGAACAGGAAGCTGCACCCTGT
EGFP-Tet1Δ566-833_rev3	GAATTCGTTAACTGCGGCCGC
EGFP-Tet1Δ566-833_fwdfinal	ATGTCTCGGTCCCGCCCCGCAATGTCTCGGTCCCGCCCCGCA
EGFP-Tet1Δ566-833_revfinal	GAATTCGTTAACTGCGGCCGC
Tet1-PBD_fwd	AATTCCATGAAACGGCGGCAAAAAGAAAAAGGCCCATATTATCACAAACGCCGGCTGC
Tet1-PBD_rev	CCGGGCAGCCGGCGTTTGTGATAATATGGGCCTTTTTCTTTTTGCCGCCGTTTCATGG
Tet1Fl-ΔCRD-outer_fwd	CTCGAGGCGATCGCATGTCTCGGTCC
Tet1s-ΔCRD-outer_fwd	ATAAGCAGCGATCGCATGGACTGCAGTAGAC
Tet1-ΔCRD-inner_fwd	TAGCTGGATTGAAGGAACAGAGGAGTGAAAACCCCAGAAA
Tet1-ΔCRD-inner_rev	TTTCTGGGGTTTTCACTCCTCTGTTCCTTCAATCCAGCTA
Tet1-ΔCRD-outer_rev	CGTTAACTGCGGCCGCTTAGACCCAAC
Tet1s-K852X/Q693X-outer_fwd	GATCTCGAGGCGATCGCATGGAC
Tet1s-K852X/Q693X-outer_rev	CGTTAACTGCGGCCGCTTAGAC
Tet1s-K852E-inner_fwd	CAACGGCTGTGAGTTTGGGAGGAG
Tet1s-K852E-inner_rev	CTCCTCCCAAACTCACAGCCGTTG
Tet1s-K852R-inner_fwd	CAACGGCTGTCGGTTTGGGAGGAG
Tet1s-K852R-inner_rev	CTCCTCCCAAACCGACAGCCGTTG
Tet1s-Q693E-inner_fwd	GATGGAGGTACAGAGAAAGAAAAAGGCCC
Tet1s-Q693E-inner_rev	GGGCCTTTTTCTTTCTCTGTACCTCCATC
miRFP670_fwd	ATCCACCGGTCGCCACCATGGTA
miRFP670_rev	AAATGTACAGGCTCTCAAGCGCGGTGAT
VprBP_fwd	AAAAGCGATCGCATGACTACAGTAGTGGTAC
VprBP-inner_fwd	GTCTCTCTTCCCCGGACCCCTCG
VprBP-inner_rev	CGAGGGGTCCGGGGAAGAGAGAC
VprBP_rev	AAAAGCGGCCGCATCACTCATTCAGAGATAAG

4.2 Cell culture and transfection

All mammalian cells were cultured in a humidified atmosphere under 5% CO2 at 37 °C. Cells were split every two to three days in the case of somatic cells and embryonic stem cells were split every two days with daily medium changes. All media were supplemented with 1x Penicillin-Streptomycin (Sigma-Aldrich Chemie GmbH, Steinheim, Germany, Cat.No.: P4333) and 1x L-glutamine (Sigma-Aldrich Chemie GmbH, Steinheim, Germany, Cat.No.: G7513).

4.2.1 Cell lines

C2C12 mouse myoblasts (Blau *et al.*, 1985)were cultured as described previously (Ludwig *et al.*, 2017) in DMEM (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: D6429) containing 20% FCS.

Mouse embryonic fibroblasts (MEF) deficient for Dnmt1 and p53 (MEF-PM, Dnmt1$^{-/-}$, p53$^{-/-}$), deficient for p53 (MEF-P, p53$^{-/-}$), deficient for SUV39 (MEF-D15, SUV39H1$^{-/-}$, SUV39H2$^{-/-}$) and corresponding wildtype cells were cultured under previously described conditions in DMEM containing 15% or 10% FCS, respectively (Lande-Diner *et al.*, 2007; Peters *et al.*, 2001).

MCF10a non-tumorigenic mammary gland cells were cultured in DMEM/F12 (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: D8900) supplemented with final concentrations of 5% horse serum (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: H0146), 20 ng/mL EGF (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: E9644), 0.5 mg/mL hydrocortisone (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: H0888), 100 ng/mL cholera toxin (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: C8052) and 10 µg/mL insulin (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: I2643) and MCF7 breast cancer cells were cultured in DMEM (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: D6429) containing 10% FCS.

Mouse embryonic stem cells E14 and stem cells deficient for Uhrf1 (Uhrf1$^{-/-}$) (Sharif *et al.*, 2007) and ligase1 (Lig1$^{-/-}$) (Ferry *et al.*, 2017) were cultured under feeder-free, 2i/LIF conditionsin culture dishes that were coated with gelatin (0.2%; Sigma-Aldrich Chemie GmbH, Steinheim, Germany, Cat.No.: G2500). DMEM (Sigma-Aldrich Chemie GmbH,

Steinheim, Germany; Cat.No.: D6429) for embryonic stem cell culture contained 16% FCS and was, in addition to L-glutamine and Pen/Strep, supplemented with 1x non-essential amino acids (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: M7145), 0.1 mM β-mercaptoethanol (Carl Roth, Karlsruhe, Germany, Cat.No.: 4227), 0.1 μM PD 0325901 (Axon Medchem BV, Groningen, The Netherlands, Cat.No.: Axon 1408), 0.3 μM CHiR 99021 (Axon Medchem BV, Groningen, The Netherlands, Cat.No.: Axon 1386), 1,000 U/ml LIF (Enzo Life Sciences GmbH, Lörrach, Germany, Cat.No.: ALX-201-242).
HEK-293 cells (Invitrogen; catalog # 620-07, Paisley PA49RF, UK) were cultured in DMEM (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: D6429) containing 10% FCS.

4.2.2 Transfection

C2C12, MEF, MCF7 and MCF10a cells were transfected by electroporation with the AMAXA Nucleofector II system (Lonza, Cologne, Germany), using a self-made buffer (5 mM KCl, 15 mM $MgCl_2$, 120 mM Na_2HPO_4/NaH_2PO_4 pH 7.2, 50 mM Mannitol) with default programs B032, A024, P020 or T024, respectively.
Mouse embryonic stem cells were transfected with the Neon nucleofection system (Thermo Fisher Scientific, Waltham, MA, USA; Cat.No.: MPK5000), according to the manufacturer's protocol.
HEK-293 cells were transfected with polyethylenimine (PEI) pH 7.0 (Sigma-Aldrich Chemie GmbH, Steinheim, Germany; Cat.No.: 408727) as previously described (Agarwal *et al.*, 2007).

4.3 Immunofluorescence staining of cells

For immunofluorescent staining of modified nucleotides, Tet1 or replication associated proteins, all cells were seeded on gelatine-coated glass coverslips and fixed in 3.7% formaldehyde (Sigma-Aldrich Chemie GmbH, Steinheim, Germany, Cat.No.: F8775) in 1x PBS for 10 minutes. After three washing steps with PBS-T (1x PBS, 0.01% Tween-20), cells were permeabilized with 0.5% Triton X-100 in 1x PBS for 20 minutes, incubated in

ice-cold 88% methanol for 5 minutes and washed again.

In the case of immunostainings of modified nucleotides, cells were incubated with 10 µg /mL RNaseA in 1x PBS for 30 minutes at 37 °C. After three more washing steps, cells were blocked with 1% BSA in 1x PBS at 37 °C for 30 minutes. The primary antibody solution contained final concentration of 0.5% BSA, 1x DNaseI reaction buffer (60 mM Tris/HCl pH 8.1, 0.66 mM $MgCl_2$, 1 mM β-mercapthoethanol) and 0.1 U/µL DNaseI (Sigma-Aldrich Chemie GmbH, Steinheim, Germany, Cat.No.: D5025). Antibodies against 5hmC (Active Motif, La Hulpe, Belgium, Cat.No.: 39769) and 5mC (mouse monoclonal clone 32E2, Manuscript in preparation: Weichmann F. *et al.*, "Generation and validation of monoclonal antibodies specific to modified nucleotides"), were each used in a dilution of 1:200. The primary antibody mix was incubated at 37 °C for 70 minutes and afterwards washed three times with PBS-TE (PBS-T + 100 mM EDTA).

Cells that were immunostained against Tet1, PCNA and ligase1 were blocked in 1% BSA for 30 minutes after fixation and incubated with primary antibodies diluted in 1% BSA for 60 minutes at room temperature. As antibodies, α-Tet1 clone 5D8 (Bauer *et al.*, 2015), αPCNA clone PC10 (Agilent Technologies, Inc. / DAKO, Santa Clara, USA, Cat.No.: M0879) and αLigaseI (rabbit, polyclonal against N-terminal amino acids 1-23, Ab-S977) were used in dilutions of 1:2, 1:100 and 1:250, respectively. After incubation with the primary antibodies, cells were washed three times with PBS-T.

For the detection of the primary antibodies, cells were incubated with following secondary antibodies diluted in 1% BSA, Alexa Fluor 488-conjugated goat α-rabbit IgG (H+L) (1:500; ThermoFisher Scientific, Invitrogen, Carlsbad CA, USA, Cat.No.: A-11008), Alexa Fluor 488-conjugated goat α-mouse IgG(H+L) (1:250; ThermoFisher Scientific, Invitrogen, Carlsbad CA, USA, Cat.No.: A-11001), Alexa Fluor 594-conjugated goat α-rabbit IgG (H+L) (1:250; ThermoFisher Scientific, Invitrogen, Carlsbad CA, USA, Cat.No.: R37117), Cy5-conjugated donkey α-mouse IgG(H+L) (1:250; The Jackson Laboratory, Bar Harbor, ME, USA, Cat.No.: 715-715-150).

After an incubation of 45 minutes at room temperature, cells were washed three times with PBS-T, counterstained with 10 µg/ml DAPI (4',6- diamidin-2-phenylindol, Sigma-Aldrich Chemie GmbH, Steinheim, Germany, Cat.No.: D9542) and mounted in Mowiol (Sigma-Aldrich Chemie GmbH, Steinheim, Germany, Cat.No.: 81381).

4.4 Microscopy and image analysis

4.4.1 Live cell time-lapse microscopy

Live cell time-lapse experiments were performed with a Nikon Ti-E microscope equipped with a UltraVIEW VoX spinning disk confocal unit (PerkinElmer, UK), controlled by Volocity 6.3 (PerkinElmer, UK), and a live cell chamber (ACU control, Olympus), at 37 °C, with 5% CO2 and 60% air humidity. Z-stacks were acquired with a 60x/1.49 NA CFI Apochromat TIRF oil immersion objective (voxel size: 0.12 x 0.12 x 0.5-1 µm; Nikon, Tokyo, Japan) or a 100x/1.49 NA CFI Apochromat TIRF oil immersion objective (voxel size, 0.071 x 0.071 x 0.5-1 µm; Nikon, Tokyo, Japan) and a cooled 14-bit CCD camera (Hamamatsu Photonics K.K., Hamamatsu City, Japan, Cat.No.: C9100-50). Z-stack images were analysed Using Volocity 6.3 (PerkinElmer, UK) and [ImageJ].

4.4.2 Protein accumulation analysis

Heterochromatin accumulation ability of ectopically overexpressed fluorescently-tagged proteins, during ongoing DNA replication, was assesed by transfecting cells with the respective plasmids. 8-12 hours post transfection, confocal Z-stacks (voxel size, 0.12 x 0.12 x 0.5 µm) were acquired using the aforementioned Nikon-Ti-E setup. Z-stacks were analysed using ImageJ. For this purpose, three circular regions with a radius of 4 pixels, in the nucleoplasm and in chromocenters, marked by PCNA, were measured in maximum intensity Z-projections images. The ratio of the averaged mean signal intensities of the protein of interest in PCNA marked replication sites and in the nucleoplasm was plotted using RStudio (Version 1.1.447).

4.4.3 High content screening microscopy

Endogenous 5mC and 5hmC levels and levels of ectopically expressed EGFP-tagged proteins were measured with the Operetta high-content screening system (Perkin Elmer, UK) in wide-field mode, equipped with a Xenon fibre-optic light source and a 20x/0.45 NA long working distance or a 40x/0.95 NA objective. For excitation and emission, following filtercombinations were used, 360-400 nm and 410-480 nm for DAPI or AMCA, 460-490 nm and 500-550 nm for EGFP as well as 560-580 nm and 590-640 nm for TexasRed.

Fluorescence intensity levels were quantified with the Harmony software (Version 3.5.1, PerkinElmer, UK) (**Figure 9**).

For the analysis of cells that were transfected with EGFP or EGFP fusion encoding plasmids and stained against 5hmC and counterstained with DAPI, cell nuclei were first identified according to their EGFP fluorescence and evaluated for morphological properties like roundness and size. These parameters as well as the mean and sum nuclear fluorescence intensities of DAPI, EGFP and Alexa 594-labelled 5hmC were calculated in cells that fit these morphological criteria. Measurement results were further processed with RStudio (Version 1.1.447). For further analysis, sum nuclear 5hmC levels were divided by the sum nuclear DAPI intensity, to compensate for potential cell cycle dependent 5hmC fluctuations. The normalized 5hmC values were grouped according to the mean EGFP intensities of the respective cells and cells with mean EGFP levels below 50 were considered as background. The before normalized 5hmC levels of the background-expressing cells were averaged and the 5hmC values of cells above background level, divided by this average. Cells above EGFP background expression levels were then grouped according to their mean EGFP fluorescence levels in low (50-100 AU), mid (100-500 AU), and high (500-1000 AU) expressing groups and plotted with RStudio (Version 1.1.447).

The nuclei of MCF10a and MCF7 cells that were simultaneously stained for 5mC and 5hmC and counterstained with DAPI, were identified by their DAPI-staining and further grouped according to morphological criteria. The mean and sum nuclear fluorescence intensities of DAPI, Alexa-488 labelled 5mC and Alexa-594 labelled 5hmC were calculated in cells that matched this criteria. Exported measurement results were further analysed and plotted with RStudio (Version 1.1.447). For this purpose, the sum nuclear Alexa-488 and Alexa-594 levels of cells that were only stained with the secondary antibodies, were averaged to obtain a background fluorescence value. The sum nuclear fluorescence values of 5mC and 5hmC in cells that were incubated with primary and secondary antibody were each normalized to the respective average background fluorescence and further normalized to the DAPI fluorescence intensity and the normalized values plotted.

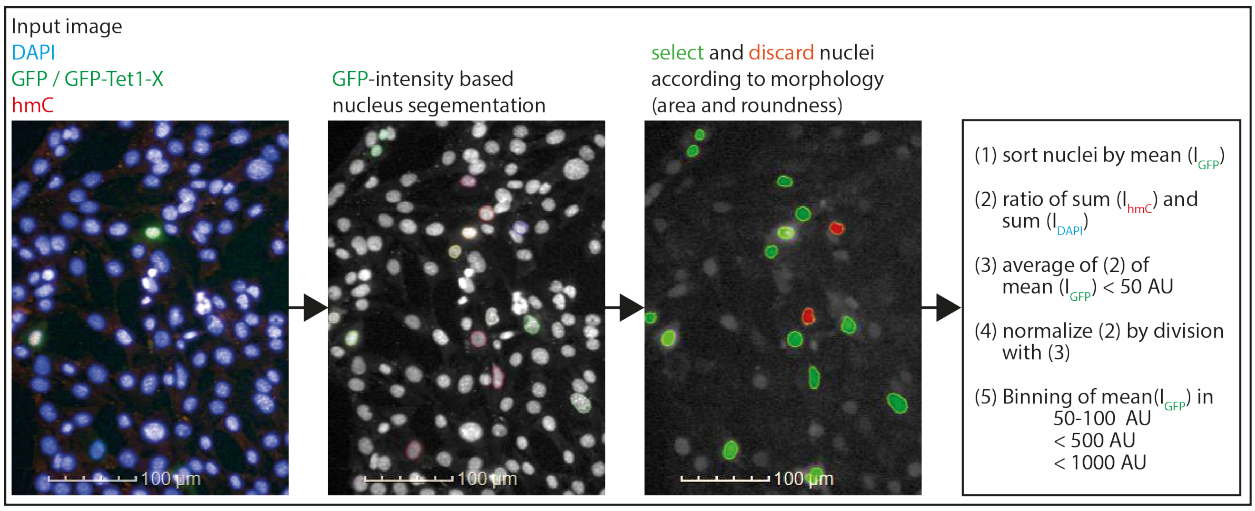


Figure 9: Workflow of the Harmony-based nuclear 5hmC quantification. To quantify sum nuclear 5hmC levels in C2C12 mouse myoblasts that ectopically overexpressed different fluorescently tagged Tet variants, cells were imaged in wide-field mode and analysed with the Harmony software. Cell nuclei in the input image were segmented according to their EGFP fluorescence and further categorized according to morphological properties. Sum and mean nuclear fluorescence intensities were calculated in all three channels and further processed with RStudio (Version 1.1.447). Sum nuclear 5hmC levels were normalized to sum nuclear DAPI levels. The normalized 5hmC levels of cells wih a mean nuclear EGFP intensity below 50 were averaged and considered as background values. Cells that were normalized to the background 5hmC intensity were grouped according to their mean nuclear EGFP fluorescence intensity into low (50-100 AU), mid (100-500 AU) and high (500-1000 AU) and plotted.

4.4.4 *In situ* DNA modification quantification

To quantify 5hmC levels in chromocenters and the nucleoplasm, C2C12 cells were transfected with EGFP-tagged Tet1-variants or EGFP alone and immunostained against 5hmC and counterstained with DAPI. Confocal Z-stacks with a Z-step length of 0.5 µm were acquired with the before described Nikon Ti-E spinningdisk setup. Images were analysed with a self-written semiautomated macro routine for ImageJ (**Figure 19A**). In brief, maximum Z-projections were generated and binary nuclear masks were created of the respective 16-bit images, based on the DAPI signal. For this, images were thresholded with the triangle method and grey values above 3000 were considered for the mask. A second binary mask for the chromocenters was created with the triangle threshold method for all pixels with intensities above 7000. Based on these two masks a third mask for the nucleoplasm was calculated and the respective mean 5hmC levels were measured and plotted with RStudio (Version 1.1.447).

To address the endogenous levels of 5mC and 5hmC in heterochromatic domains of MCF10a or MCF7 cells, cells were immunostained against 5mC and 5hmC and counter-

stained with DAPI. Cells were imaged with the before described confocal spinningdisk setup and analysed with ImageJ. For this, maximum Z-projections were generated and circular ROIs with a radius of 5 pixel were created around DAPI-dense regions and the respective mean fluorescence intensities measured. Mean fluorescence intensities were normalised to the DAPI intensity and plotted with RStudio (Version 1.1.447).

4.5 Statistical analysis

Statistical significance was tested with a paired two-samples Wilcoxon test with RStudio (Version 1.1.447). *P*-values below 0.005 were assumed to be statistically significant. All *p*-values are given in the figure legends or respective plots directly. Boxplots represent the median with the box depicting the 25–75 percentile and the lines and circles denote the 95% confidence interval and outliers, respectively.

4.6 Western blotting and co-immunoprecipitation

Cell pellets for western blotting and co-immunoprecipitation were obtained by harvesting cells by centrifugation for 5 minutes at 500 xg and 4 °C. For general western blotting, cell lysates were prepared by resuspending the cell pellet in lysis buffer and (20 mM Tris-HCl pH 8, 150 mM NaCl, 0.5 mM EDTA, 0.5 % NP-40) supplemented with Pepstatin A (1 µM final concentration; Sigma-Aldrich Chemie GmbH, Steinheim, Germany, Cat.No.: P5318), PMSF (10 µM final concentration, Sigma-Aldrich Chemie GmbH, Steinheim, Germany, Cat.No.: 78830), AEBSF (1 mM final concentration, AppliChem, Darmstadt, Germany, Cat.No.: A1421,0100). Cells were disrupted by syringe treatment with 21 G needles and incubated on ice for 30 minutes with repeated vortexing in-between. Cell lysates were cleared by centrifugation for 15 minutes at 13000 xg and 4 °C and the supernatant resolved by SDS-PAGE and subjected to western blotting analysis.

Co-immunoprecipitation experiments were essentially performed as described before (Becker *et al.*, 2016). In brief, HEK-293 cells were PEI-transfected with EGFP and mRFP- or mcherry-fusion protein encoding plasmids and harvested approximately 36 hours later by centrifugation for 5 minutes at 500 xg and 4 °C. The cells were washed with ice-cold washing buffer (20 mM Tris-HCl pH 8, 150mM NaCl, 0.5 mM EDTA) and centrifuged again for 5 minutes at 500 xg and 4 °C. The resulting cell pellet was resuspended in 200

µL lysis buffer and cells were disrupted and the cell lysate cleared as described earlier. 15% percent of the cell lysate was used as input control and the rest was incubated with equilibrated GFP- or RFP-trap beads (Chromotek GmbH, Planegg-Martinsried, Germany) on a rotator for 90 minutes at 4 °C. Afterwards, beads were washed 3 times with 300 µL washing buffer. Input and bound fraction were separated on 6% SDS-PA gels and transferred to a nitrocellulose membrane (Bio-Rad Laboratories, Hercules, CA, USA, Cat.No.: 1620112).

Membranes were blocked with 5% milk in 1x PBS-T (0.1% Tween). For the detection of EGFP- or mRFP- and mcherry-tagged proteins, rat monoclonal anti-GFP clone 3H9 (ChromoTek GmbH, Planegg-Martinsried, Germany, Cat.No.: 3h9) and rat monoclonal anti-RFP clone 5F8 (Rottach *et al.*, 2008); (ChromoTek GmbH, Planegg-Martinsried, Germany, Cat.No.: 5f8) were used as primary antibodies. Tet1 was detected with monoclonal rat 5D8 (1:2, (Bauer *et al.*, 2015)). Primary antibodies were diluted in blocking buffer and incubated for 2 hours at room temperature. After washing three times with 1x PBS supplemented with 0.1 Tween-20, a horseradish peroxidase conjugated goat anti-rat IgG secondary antibody (1:5000; The Jackson Laboratory, Bar Harbor, ME, USA, Cat.No.: 112-035-068) was incubated in blocking buffer for 1 hour at room temperature. The secondary antibody was visualized with ECL Plus Western Blotting Substrate (ThermoFisher Scientific, Waltham, MA, USA, Cat.No.:32132) and signals were detected with a chemiluminescence imager.

5 Results

5.1 The short isoform of Tet1 localizes to pericentric heterochromatin during ongoing DNA replication

5.1.1 Tet1 and its short isoform show different S-phase localizations

Methylated cytosine serves as a substrate for the dioxygenases of the Tet protein family, which can oxidise this repressive DNA modification and thereby change and diversify the information content of the epigenome. The Tet catalytic activity thus modulates crucial cellular processes like transcription, DNA replication or repair and in turn affects embryonic as well as postnatal organismic development but also plays a role in tumorigenesis. Despite extensive efforts to unravel the regulation of Tet proteins, the mechanism how cellular levels of 5hmC and its oxidative derivatives are maintained in proliferating cells, remains elusive. Tet1, the founding member of the Tet dioxygenase family, has previously been implicated in maintaining euchromatic 5hmC levels in differentiated cells, throughout the cell cycle and to thereby prevent DNA methylation spreading to euchromatin (Jin *et al.*, 2014). However, also this study failed to present a mechanism on how this maintenance is achieved. Recently, a short isoform of Tet1 was described to be expressed in somatic cells upon differentiation, while expression of full-length Tet1 is mainly restricted to embryonic stem cells and early embryogenesis (Zhang *et al.*, 2016b). Said short isoform was found to arise from an alternate promoter and to lack a major part of its N-terminus, including the CxxC-domain, which targets the full-length Tet1 protein to non-modified DNA (Frauer *et al.*, 2011a). The short isoform of Tet1 was furthermore shown to be overexpressed in various cancers, which correlates with poor patient survival (Good *et al.*, 2017). In line with the finding that Tet1 can act as a maintenance dioxygenase, the Tet-mediated 5mC oxidation products 5hmC and 5fC were found to be stable throughout the cell cycle in proliferating cells with almost similar cell cycle-dependent dynamics as described for 5mC (Bachman *et al.*, 2014; Bachman *et al.*, 2015).

Based on the question, how the oxidative derivatives of 5mC are maintained throughout the cell cycle, I first addressed the cellular localization of full-length Tet1 and its cancer-associated, N-terminally truncated, short isoform during ongoing DNA replication in somatic

cells. For this purpose I transfected C2C12 mouse myoblasts with plasmids encoding EGFP-tagged fusion proteins of full-length Tet1 and its short isoform, hereafter called Tet1s, (**Figure 10A**) together with mRFP-tagged PCNA, as a marker for cell cycle progression and, in particular, S-phase substages and active replication sites. C2C12 cells were chosen, as their S-phase and its spatio-temporal progression have been studied extensively and are well characterized (Casas-Delucchi *et al.*, 2011). In their undifferentiated state, C2C12 cells moreover express comparatively low to undetectable levels of Mbd proteins like MeCP2 (Brero *et al.*, 2005), which we could previously show to counteract Tet catalytic activity by restricting its access to DNA (Ludwig *et al.*, 2017). Besides the low levels of Mbd proteins, undifferentiated C2C12 cells also show low expression of all three Tet proteins, which makes them a suitable model for ectopic Tet overexpression experiments (Zhong *et al.*, 2017). To avoid secondary effects from prolonged Tet1 overexpression and 5mC oxidation, cells were subjected to live cell time-lapse microscopy eight to ten hours after transfection. For time-lapse imaging, cells in early S-phase, defined by a homogeneously distributed focal PCNA pattern, were chosen and confocal Z-stacks acquired every 20 minutes over the course of several hours.

While full-length Tet1 showed no noteworthy accumulation during either of the S-phase substages or in G2, a clear localization between Tet1s and PCNA could be observed at sites of ongoing DNA replication in pericentric heterochromatin during late S-phase (**Figure 10B**). This observation led me to quantify the relative accumulation of Tet1 and Tet1s during late S-phase. For this purpose, C2C12 cells were again co-transfected with either of the two EGFP-tagged Tet1 variants, together with mRFP-tagged PCNA and imaged live approximately eight hours post transfection. Live-cell imaging was chosen, because the observed localization of Tet1s during late S-phase was lost upon fixation, although different fixation protocols were tested. Cells that showed a late S-phase PCNA pattern, characterized by strong focal accumulations, were imaged. For the analysis of Tet1 or Tet1s accumulation, maximum intensity Z-projections were generated and three nuclear areas in PCNA marked, replicating pericentric heterochromatin (PHC) domains and in the nucleoplasm outside of DNA replication foci, were chosen. The fluorescence intensity signals of the three respective spots were averaged and the ratio of the mean Tet1 fluorescence signals inside and outside of PCNA foci was calculated. In this way, a

significant accumulation of Tet1s at sites of late S-phase, in comparison to full-length Tet1 was found (**Figure 10**C).

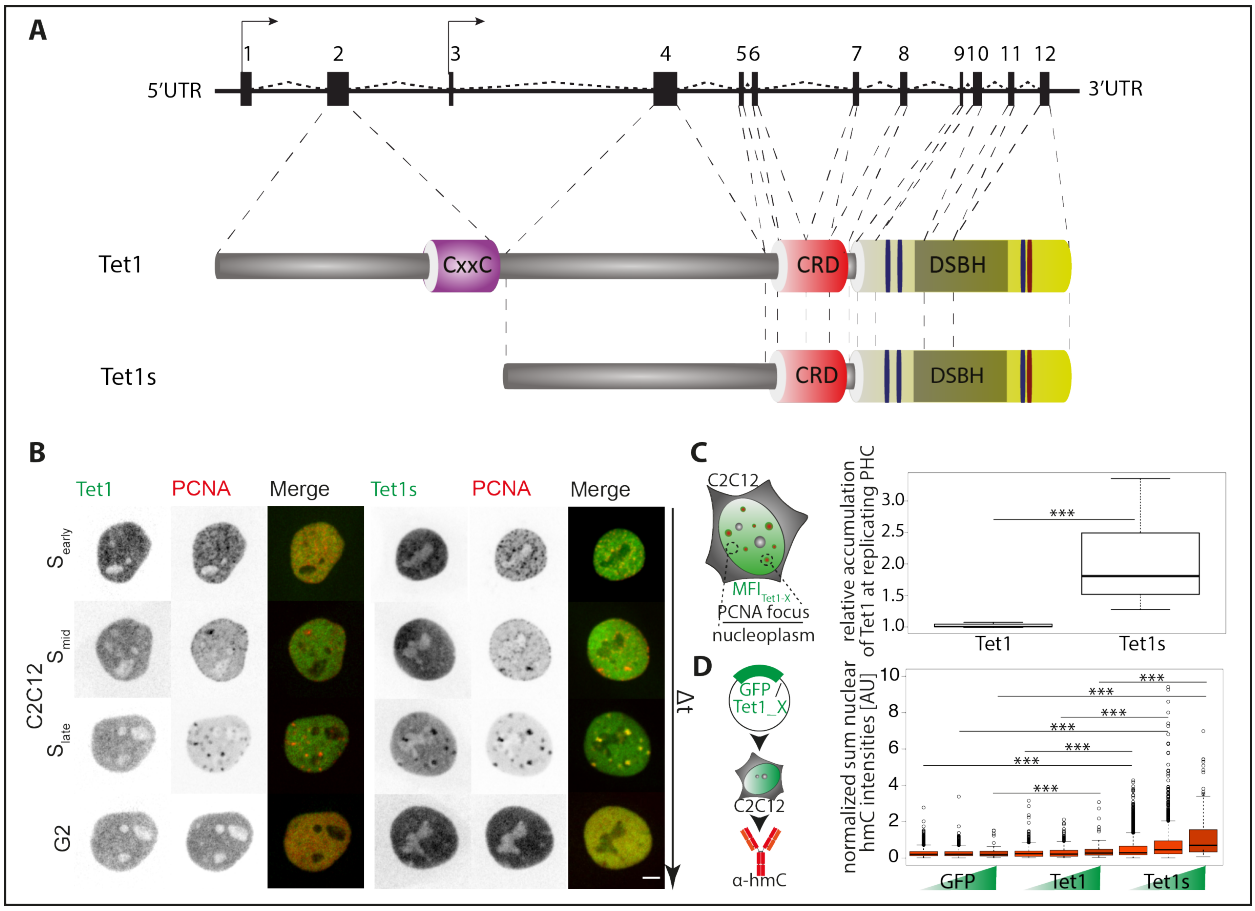


Figure 10: The short isoform of Tet associates with pericentric heterochromatin during DNA replication and increases 5hmC levels. (**A**) Gene structure of the human Tet1 locus. Exons 1 and 3 are non-coding and each harbour a distinct transcription start site, of which each is targeted by a different promoter. Translation of full-length Tet1 and its short isoform starts in exon 2 and 4, respectively (Good *et al.*, 2017). (**B**) Confocal mid Z-planes of C2C12 mouse myoblasts that were transfected with mRFP-PCNA and EGFP-Tet1 or EGFP-Tet1s and subjected to time-lapse live cell microscopy eight hours post transfection. Cells that showed an early S-phase pattern were chosen and followed over time with time intervals of 20 minutes. While no noteworthy accumulation or PCNA colocalization was observed for Tet1, Tet1s showed a clear colocalization with PCNA during ongoing replication in pericentric heterochromatin in late S-phase. (**C**) To analyse the accumulation of Tet1 and Tet1s at PCNA marked replication sites, C2C12 cells were transfected with EGFP-tagged Tet1 or Tet1s and mRFP-PCNA and imaged live eight hours post transfection. For the quantification of the Tet1 or Tet1s accumulation at late replication PHC, maximum intensity Z-projections were generated and the mean fluorescence intensities of three nuclear areas inside and outside of sites of ongoing DNA replication were measured. The respective values were averaged and the mean fluorescence intensities of Tet1 in PCNA foci divided by the mean fluorescence in nucleoplasmic regions. (n-numbers: Tet1: 10, Tet1s: 13). (**D**) Ectopic overexpression of Tet1s significantly increases 5hmC levels in C2C12 cells. C2C12 mouse myoblasts were transfected with EGFP-Tet1, EGFP-Tet1s or EGFP alone and 24 hours later fixed and immunostained against 5hmC. Fluorescence intensity levels of overexpressed proteins and 5hmC were measured with a wide-field high content screening microscopy system. Sum nuclear 5hmC levels were normalized to the sum nuclear DAPI intensity and grouped according to their respective mean EGFP fluorescence (low: 50-100 AU, mid:100-500 AU, high: 500-1000 AU; n-numbers: GFP=6255, Tet1=3169, Tet1s=5960). Experiments were performed at least in triplicate and statistical significance was tested with a paired two-samples Wilcoxon test, using R-studio version 1.1.447. Scale bars = 5 µm.

Prompted by the finding that Tet1s localized to pericentric heterochromatin during ongoing DNA replication, I was furthermore interested, to see how global 5hmC levels changed upon overepression of either Tet1 or Tet1s. To answer this question, I transfected C2C12 mouse myoblasts with either EGFP-tagged Tet1 or Tet1s or EGFP as a control and fixed the cells 24 hours after transfection and performed immunofluorescence staining against 5hmC. Stained cells were imaged with a wide-field high content screening microscope system and sum and mean nuclear fluorescence intensities quantified with the respective analysis software (**Figure 9**). For further analysis, sum 5hmC levels were normalized to the sum nuclear DAPI intensity as a proxz for DNA content. EGFP levels, were normalized to the average of the lowest expressing cells and than grouped in low, mid and high expressing cells, according to their mean nuclear EGFP fluorescence levels. Interestingly, Tet1s overexpression led to significant 5hmC increases in all groups compared to the EGFP-control, but also Tet1-transfected cells. In contrast, full-length Tet1 failed to markedly increase 5hmC levels in low- and mid-expressing cells and, with exception for the high-expressing group, showed no significantly higher 5hmC levels as the respective EGFP-control groups (**Figure 10**D).

5.1.2 The CxxC domain of Tet1 prevents its S-phase localization

A domain, distinguishing Tet1 from its short isoform is the N-terminal zinc-finger that is lacking in Tet1s (**Figure 10**A). Besides the CxxC-domain one more regulatory domain in the very N-terminus was recently identified and termed BC (before CxxC) domain (Zhang *et al.*, 2016b). The CxxC- and the BC-domain were implicated in concomitantly facilitating chromatin binding of full-length Tet1 and their combined deletion was shown to result in markedly decreased chromatin loading (Zhang *et al.*, 2016b). These findings prompted me to investigate the accumulation of different N-terminal deletion mutants of Tet1 and also its CxxC-domain alone, during late S-phase. Hence, I transfected C2C12 cells with mRFP-PCNA and EGFP-Tet1Δ1-389 that lacks the BC-domain, EGFP-Tet1Δ566-833 that lacks the CxxC-domain or EGFP-Tet1-CxxC (**Figure 11A**). Cells in late S-phase were imaged live eight to ten hours post transfection and the relative accumulation in PCNA foci of late S-phase cells quantified as described before. While the CxxC domain of Tet1

showed a very homogeneous nuclear pattern and no accumulation at all, Tet1Δ1-389 showed a slightly (p=0.1823) and Tet1Δ566-833 a significantly (p=0.002756) increased accumulation compared to full-length Tet1 (**Figure 11B**). In regard to this finding, I set out to investigate the spatio-temporal dynamics of the CxxC-lacking Tet1Δ566-833 fusion-protein during S-phase. For this, I transfected the respective EGFP-tagged construct together with mRFP-PCNA and performed live-cell time lapse microscopy as described earlier. Interestingly, in stark contrast to full-length Tet1 or Tet1s, strong focal accumulations were observed already during early S-phase and persisted throughout the whole S-phase and G2. The characteristic colocalization with PCNA at PHC in late S-phase was also observed (**Figure 11C**).

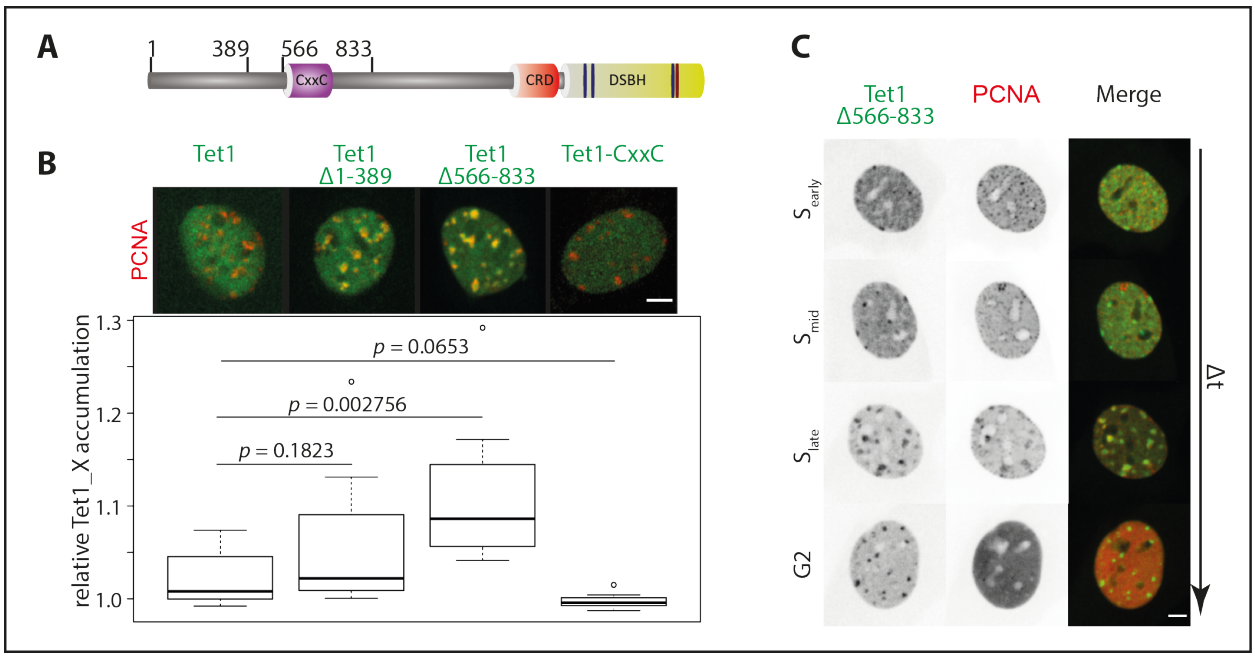


Figure 11: The CxxC domain of Tet1 prevents PHC association during DNA replication. (**A**) Domain organization of full-length Tet1. Indicated are the the locations of amino acids from N-terminal deletion mutants, used in this study. (**B**) Representative confocal mid-Z-plane images of C2C12 mouse myoblasts in late S-phase that were transfected with EGFP-tagged full-length Tet1, Tet1Δ1-389, Tet1Δ566-833 or the Tet1-CxxC domain and mRFP-PCNA. The corresponding boxplot shows the relative accumulation of the EGFP signal at PCNA marked DNA replication sites in pericentric heterochromatin. Tet1Δ566-833 shows a significant accumulation compared to full-length Tet1, while Tet1Δ1-389 accumulation is only slightly increased and the Tet1-CxxC does not show any accumulation (n-numbers: Tet1=10, Tet1Δ1-389=11, Tet1Δ566-833=10, Tet1-CxxC=11). Statistical significance was tested with a paired two-samples Wilcoxon test, using R-studio version 1.1.447. (**C**) Representative confocal mid-Z-sections of a C2C12 mouse myoblast cell that was co-transfected with EGFP-Tet1Δ566-833 and mRFP-PCNA and subjected to live-cell time lapse microscopy 8 hours post transfection. Tet1Δ566-833 shows a strong focal pattern in the nucleus throughout the entire S-phase and in G2, but also colocalizes with PCNA during ongoing DNA replication in PHC in late S-phase. Scale bars = 5 µm.

5.1.3 DNA replication association of Tet1 is independent of its catalytic activity

Many studies, including our own that investigated the effects of Tet catalytic activity, used not only the full-length Tet proteins, but also the respective C-terminal catalytic domains. The catalytic domain of all Tet proteins comprises the cysteine-rich domain (CRD) and the double-stranded beta helix (DSBH) that together are sufficient for Tet-mediated 5hmC formation *in vivo* and also *in vitro* (Ludwig *et al.*, 2017; Zhang *et al.*, 2017b; Zhang *et al.*, 2017c; Zhang *et al.*, 2017a). To investigate the cell cycle distribution of the catalytic domain of Tet1 (Tet1CD), I transfected C2C12 mouse myoblasts with mcherry-Tet1CD, miRFP-PCNA and EGFP-MaSat, imaged cells in S-phase and scored the accumulation in late S-phase as described earlier (**Figure 10C**). Similar to the observation made for Tet1s, no notewhorty nuclear Tet1CD-patterns were observed throughout early or mid S-phase (not shown). However, in late S-phase Tet1CD showed a clear colocalization with PCNA as also observed for Tet1s (**Figure 12B, upper panel**). This prompted me to further test, whether single amino acid substitutions in the cofactor binding sites (H1652Y, D1654A, (**Figure 12A**)) of Tet1CD can change the observed localization. These mutations render Tet1 catalytically inactive but do not affect its DNA-binding affinity (Zhang *et al.*, 2017b).

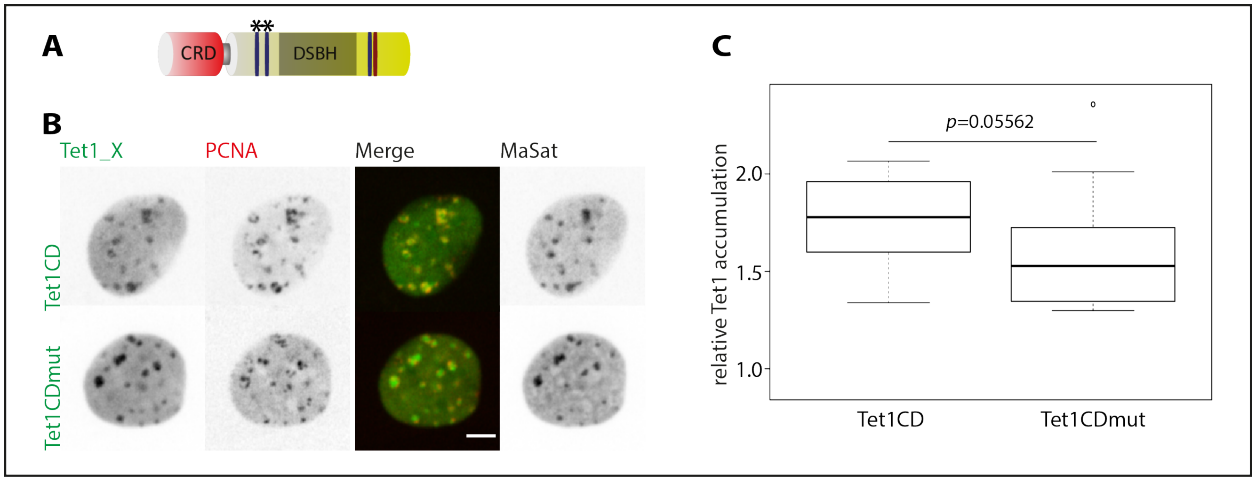


Figure 12: DNA replication association of Tet1 is independent of its catalytic activity. (**A**) Schematic of the catalytic domain of Tet1. Stars indicate the location of cofactor binding site mutations that render it catalytically inactive. (**B**) Representative pseudocolored mid-Z-sections of C2C12 mouse myoblasts in late S-phase, transiently expressing fluorescently tagged Tet1CD or Tet1CDmut as well as PCNA and MaSat. (**B**) Quantification result of the accumulation of Tet1CD and Tet1CDmut at late replicating PHC, marked by PCNA (n-number for both = 15). For accumulation quantification, mean fluorescence intensities in three circular regions, inside and outside of sites of ongoing DNA replication were measured and the average of the ratio plotted. Statistical significance was tested with a paired two-samples Wilcoxon test, using R-studio version 1.1.447. Scale bar = 5 µm.

For this, I again transfected C2C12 myoblasts with miRFP-PCNA, EGFP-MaSat and the mcherry-tagged catalytically inactive mutant of Tet1CD (Tet1CDmut). This mutant showed the same nuclear localization pattern as previously observed for the short isoform of Tet1 or Tet1CD and the accumulation at replication sites in PHC was not significantly changed, compared to Tet1CD. (**Figure 12B, lower panel, C**).

With respect to the finding that the catalytic domain of Tet1 is sufficient for its association with DNA replication at PHC and the high sequence homology to the catalytic domains of Tet2 and Tet3 (Bauer *et al.*, 2015), I investigated also the S-phase distribution of full-length Tet2 and Tet3 and their respective catalytic domains, employing the protein accumulation assay (**Figure 10C**). With this approach, no accumulation throughout the whole S-phase, including during the replication of pericentric heterochromatin in late S-phase was observed for either the full-length proteins or the catalytic domains of Tet2 or Tet3 (**Figure 13A, B**).

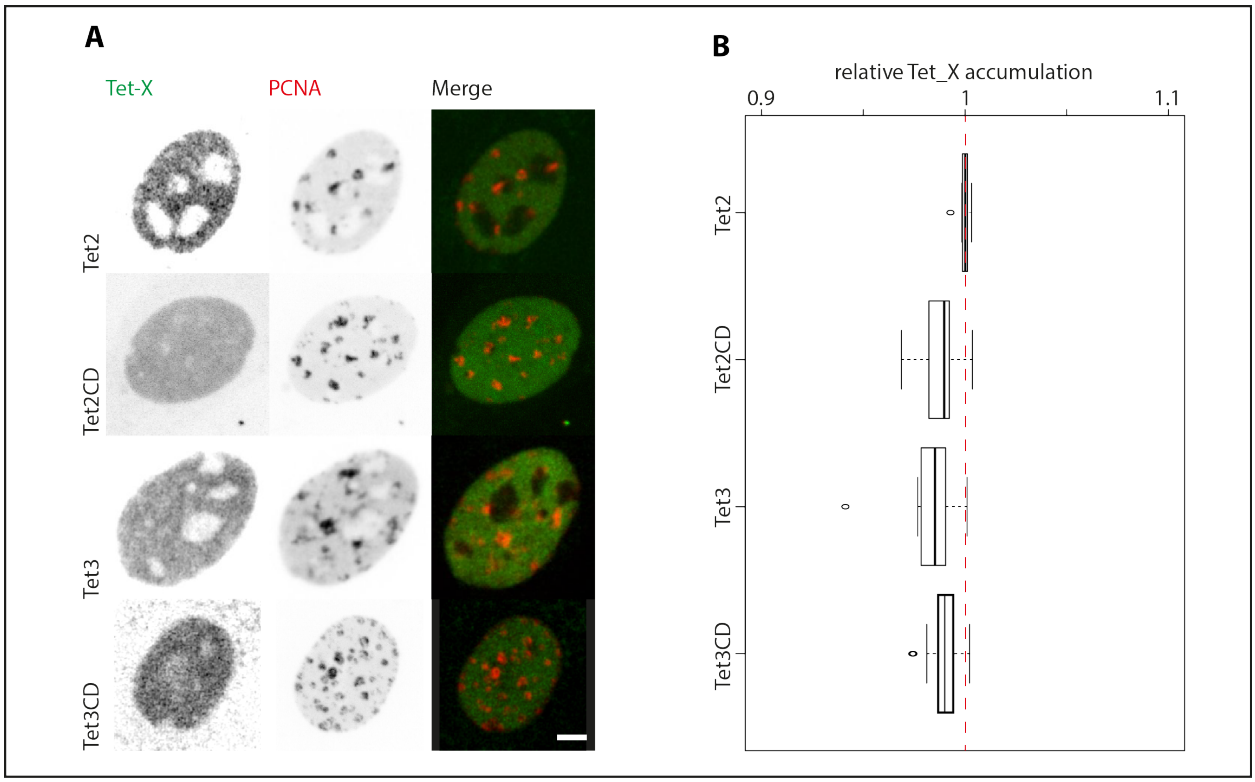


Figure 13: Tet2 and Tet3 do not localize to pericentric heterochromatin during late DNA replication. (**A**) Representative pseudocolored confocal mid-Z-sections of C2C12 mouse myoblasts in late S-phase transiently transfected with fluorescently tagged PCNA and Tet2 and Tet3 full-length proteins and respective catalytic domains. (**B**) Boxplots show the relative accumulation of Tet2 and Tet3 or their respective catalytic domains in PCNA marked chromocenters. Statistical significance was tested with a paired two-samples Wilcoxon test, using R-studio version 1.1.447. Scale bar = 5 µm.

5.1.4 The cysteine-rich domain of Tet1 is essential for its S-phase association and catalytic activity

The minimal domain that is needed for the dioxygenase activity of the Tet proteins, is their C-terminal catalytic domain, which comprises the CRD and the DSBH. While these domains together are sufficient for catalytic activity and localization to PHC during ongoing DNA replication, I was interested to test, whether either of the two domains alone can localize to late-replicating DNA or at least still be catalytically active. I therefore first tested the localization of these two domains during ongoing DNA replication in PHC. For this purpose, I co-transfected C2C12 cells with mRFP-tagged PCNA and EGFP-tagged CRD or DSBH and acquired confocal Z-stacks of cells with a late S-phase PCNA pattern approximately eight hours post transfection and quantified their accumulation.

In this way, I found no cells, which overexpressed the DSBH, to show focal patterns that resembled those of PCNA. Cells that were co-transfected with EGFP-CRD, however, occasionally showed weak accumulations of CRD at PCNA marked PHC. When I quantified the accumulation of the CRD and the DSBH at late-replicating PHC, values close to 1 were measured, indicating almost no accumulation. However, the CRD, for which occasional accumulations were observed, showed a significantly increased accumulation compared to the DSBH (**Figure 14A, B**). To test, whether either of the domains can increase global 5hmC levels, C2C12 cells were transfected with EGFP-Tet1CD, EGFP-CRD or EGFP-DSBH and 24 hours later fixed and immunostained against 5hmC and counterstained with DAPI. Stained cells were imaged with a wide-field high content screening microscope and the sum and mean nuclear values of DAPI, EGFP and Alexa-594-labelled 5hmC measured. Sum nuclear 5hmC levels were normalized to the sum DAPI intensity and grouped according to their mean EGFP fluorescence as an indicator for Tet1-X expression levels. While both, CRD and DSBH, failed to increase global 5hmC levels compared to control EGFP-transfected cells, Tet1CD significantly (p for all groups = 2.2×10^{-16}) increased 5hmC levels independent of its respective expression levels, compared to EGFP, EGFP-CRD or EGFP-DSBH transfected cells (**Figure 14C**).

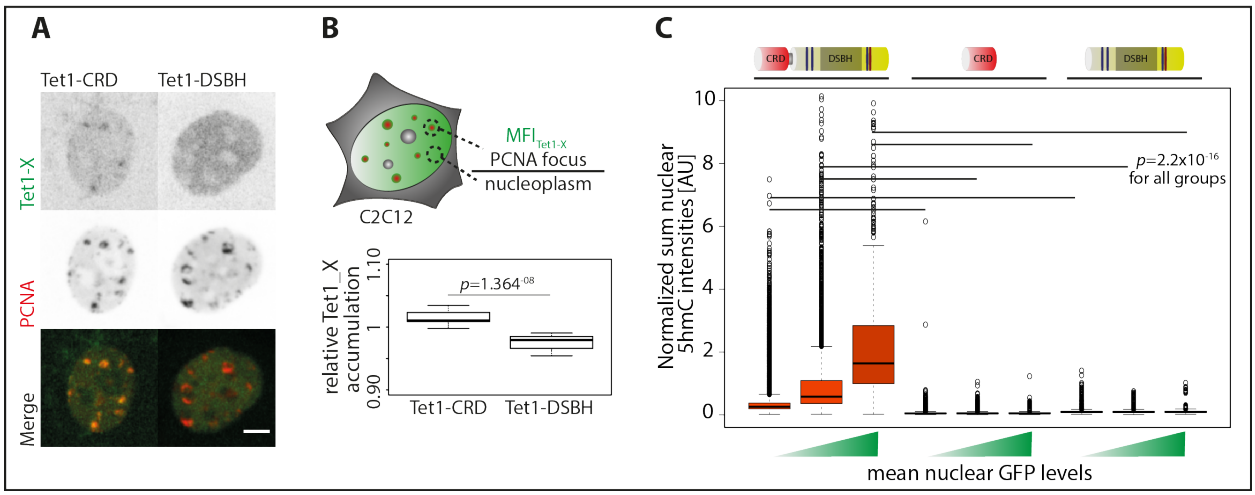


Figure 14: The cysteine-rich domain of Tet1 colocalizes with PCNA during PHC replication but is catalytically inert. (**A**) Representative confocal mid-Z-sections of C2C12 mouse myoblasts, co-transfected with plasmids, encoding fluorescently tagged PCNA and Tet1-CRD or Tet1-DSBH. (**B**) Schematic of the accumulation quantification of Tet1-CRD or Tet1-DSBH in PCNA marked PHC replication sites and the corresponding boxplot that summarizes the quantification results (n-numbers: CRD=14, DSBH=13) (**C**) Boxplot showing the normalized sum nuclear 5hmC levels in C2C12 mouse myoblasts that were transfected with Tet1CD, Tet1-CRD or Tet1-DSBH and 24 hours later immunostained against 5hmC. 5hmC levels were normalized to background fluorescence levels and the sum nuclear DAPI levels. Cells were grouped in low (<50-100 AU), mid (<100-500 AU) and high (<500-1000 AU) expressing cells according to their mean nuclear EGFP fluorescence (n-numbers: CD=15893, CRD=15835, DSBH=8121). Three stars indicate a *p*-value > $2.2x10^{-16}$. Statistical significance was tested with a paired two-samples Wilcoxon test, using R-studio version 1.1.447 Scale bar = 5 µm.

The very weak, but focal accumulation that was occasionally observed for Tet1-CRD, prompted me to delete the CRD from Tet1s and investigate the effects that the deletion of this domain has on S-phase Tet1 localization, as well as global 5hmC production. I generated an EGFP-tagged Tet1sΔCRD-construct and transfected it into C2C12 mouse myoblasts together with mRFP-PCNA, to perform live-cell time lapse microscopy. In stark contrast to the localization of Tet1s that showed a clear accumulation in late S-phase at PCNA marked chromocenters, no accumulation at PCNA marked replication foci was observed throughout S-phase. Investigating the global 5hmC levels of C2C12 cells that overexpressed EGFP-Tet1sΔCRD gave a similar picture as observed before for the CRD or DSBH alone. No 5hmC increase was found for either low, mid or high Tet1sΔCRD overexpressing cells, while Tet1s overexpression significantly increased global 5hmC levels compared to EGFP-control transfected cells.

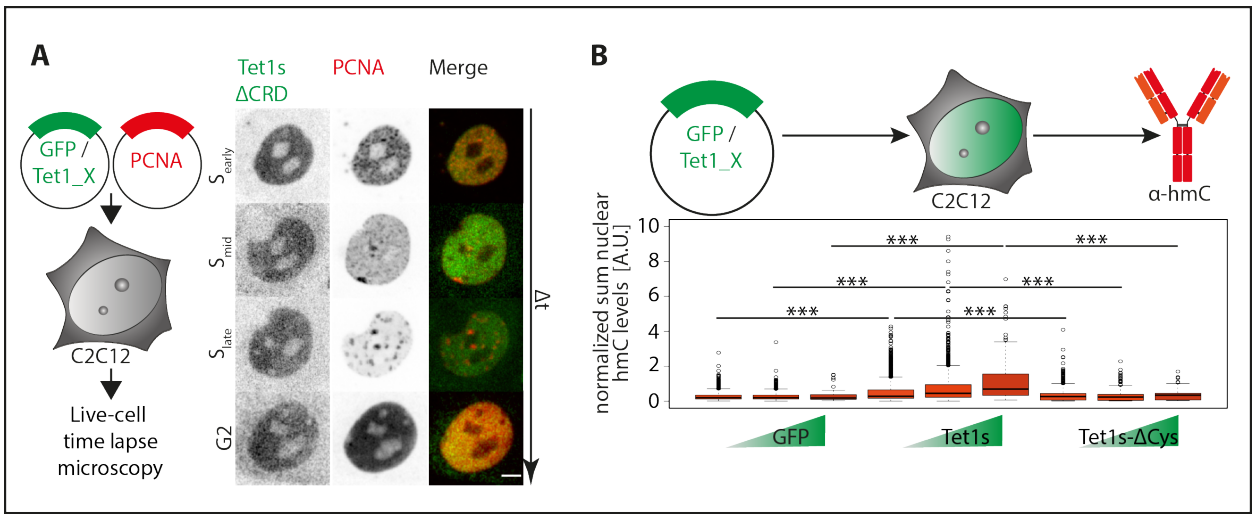


Figure 15: The CRD of Tet1s is required for its catalytic activity and S-phase localization. (**A**) Representative confocal mid-Z-section images of a C2C12 mouse myoblast cell that was co-transfected with EGFP-Tet1sΔCRD and mRFP-PCNA and subjected to live-cell time lapse microscopy. (**B**) Schematic of the experimental procedure and respective boxplots. Mouse myoblasts were transfected with EGFP, EGFP-Tet1s or EGFP-Tet1sΔCRD and 24 hours later stained against 5hmC. Sum nuclear 5hmC were normalized to the 5hmC levels in the lowest expressing cells and the sum nuclear DAPI intensity and grouped, according to their mean nuclear EGFP fluorescence intensity in low (<50-100 AU), mid (<100-500 AU) and high (<500-1000 AU) expressing (n-numbers: GFP=6255, Tet1s=5960, Tet1sΔCRD=2070). Experiments were performed at least in triplicate and statistical significance was tested with a paired two-samples Wilcoxon test, using R-studio version 1.1.447. Three stars indicate a *p*-value > $2.2x10^{-16}$. Scale bar = 5 µm.

Many proteins that associate with sites of ongoing DNA replication, do so via the interaction with the clamp loader protein PCNA. Dnmt1, for example, harbours a so-called PCNA binding domain (PBD), a short peptide with a conserved sequence motif that facilitates this interaction. The consensus sequence of classical PBDs is characterized by an initial glutamine (Q), followed by two variable amino acids, a hydrophobic amino acid, like leucine (L) or isoleucine (I), two variable residues and finally two aromatic amino acids, like phenylalanine (F), tryptophan (W), tyrosine (Y) or histidine (H). This consensus sequence is often followed by basic residues like arginine (R) or lysine (K) (Easwaran, Leonhardt, and Cardoso, 2007) (**Figure 16A**). Interestingly, PCNA was found to be a strong interactor of Tet1 and to regulate its dioxygenase activity throughout the cell cycle and thereby protect cells from aberrant DNA methylation (Cartron *et al.*, 2013). To identify a putative PBD in Tet1, but also Tet2 and Tet3, the respective amino acids sequences were searched for a glutamine followed by five variable amino acids and finally two aromatic amino acids with the pattern finder tool from www.bioinformatics.org. While I did not get any hit in the respective protein sequences of Tet2 or Tet3, a short peptide that harbours the desired sequence was identified in the CRD of Tet1. The major difference of the putative PBD of

Tet1 and the consensus sequence, is the hydrophobic middle amino acid. The putative PBD of Tet1 harbours a hydrophilic lysine residue, instead of a hydrophobic leucine or isoleucine (**Figure 16A**). To test, whether the PBD-like sequence of Tet1 can nevertheless associate with PCNA and consequently the DNA replication machinery, I performed oligo-cloning to obtain a EGFP-tagged Tet1-"PBD". This construct was transfected into C2C12 mouse myoblasts together with mRFP-PCNA and confocal Z-stacks were acquired approximately 8 hours post transfection and cells in different S-phase substages, defined by their respective PCNA pattern. Strikingly, Tet1-"PBD" showed a distribution throughout the cell with a strong nuclear signal. This distribution was attributed to the fact that the flanking sequences of the Tet1-"PBD" contain many arginines that are also found in nuclear localization sequences (Kalderon *et al.*, 1984). Despite the strong nuclear signal, no accumulation of Tet1-"PBD" at PCNA labelled sites of ongoing DNA replication was observed in either of the S-phase substages, which was also confirmed by line-profile analysis (**Figure 16B**).

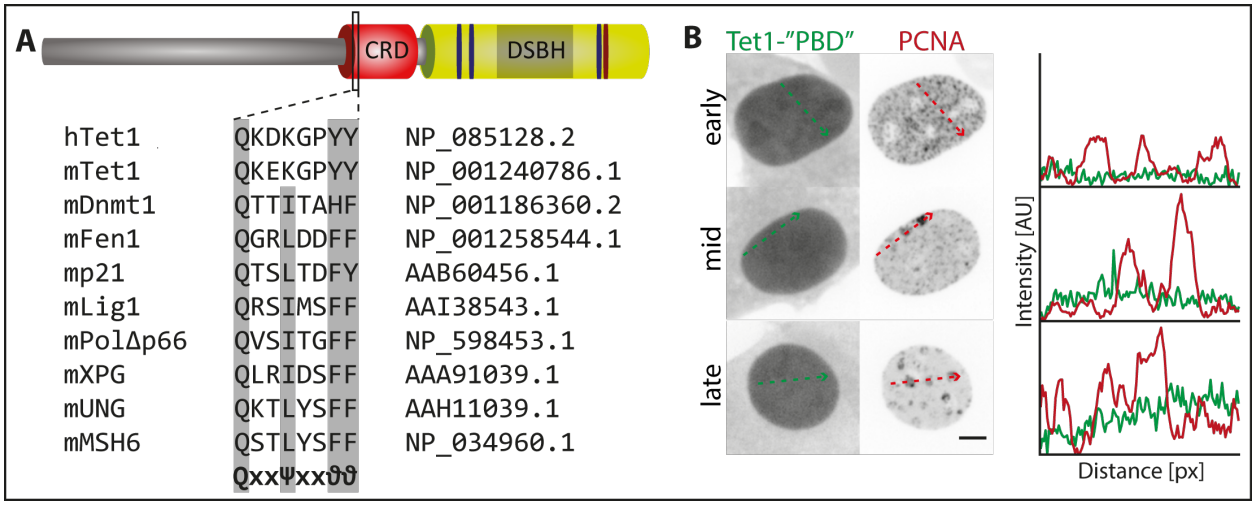


Figure 16: The CRD of Tet1 harbours an inert putative PBD. (**A**) Domain organization of the short isoform of Tet1 with the location of the putative PCNA binding domain (PBD) identified in the N-terminal region of its cysteine-rich domain (CRD). The putative PBD sequence is aligned to known PBDs from different proteins, associated with DNA replication and repair, with their respective accession numbers. The consensus PBD sequence consisting of a glutamine followed by two variable amino acids, a hydrophobic amino acid, two variable amino acids and two aromatic amino acid residues, is given at the bottom. (**B**) Exemplary confocal mid-Z-sections of C2C12 mouse myoblasts in different S-phase substages co-transfected with the EGFP-tagged putative PBD of Tet1 and mRFP-tagged PCNA. Line profiles of the indicated arrows are shown next to the respective images. Scale bar = 5 µm.

5.1.5 A conserved lysine in Tet1s CRD is crucial for its S-phase localization and targeted catalytic activity

In addition to the non-functional putative PBD of Tet1 just described, two recent studies independently identified a conserved lysine residue within the CRD of all Tet proteins. Monoubiquitination of this lysine was found to modulate Tet catalytic activity and to be important during oocyte development and for Tet2 chromatin binding (Yu *et al.*, 2013; Nakagawa *et al.*, 2015). The lysine itself resides within a peptide that is conserved between all three Tet protein family members and also between human and mouse Tet (**Figure 17A**). This short amino acid stretch was shown to stabilize the DNA around the modified cytosine target, by interacting with the phosphate backbone (**Figure 17B**). Furthermore, mutations of the conserved lysine residue to a glutamate resulted in loss of Tet2 catalytic activity, indicating its importance for the correct function of Tet proteins. (Hu *et al.*, 2013).

I therefore mutated lysine (K) 852 in Tet1s to a glutamate (E) or an arginine (R) to invert or keep the respective charge and abrogate a putative ubiquitination. I continued to investigate the DNA replication association capability of the respective constructs and the effect of their overexpression on global 5hmC generation.

To test, whether the lysine mutations affect the accumulation during DNA replication in PHC, EGFP-tagged Tet1s-constructs that harbour the K852R or K852E mutation were co-transfected with mRFP-PCNA into C2C12 mouse myoblasts. Confocal Z-stacks were acquired approximately eight hours later and the accumulation of the respective constructs in late replicating DNA, marked by PCNA, was assessed as described earlier (**Figure 10C**). With this approach, no accumulation was found for either of the two Tet1s-mutants during late S-phase, indicating that the respective lysine residue is indeed crucial for Tet1s recruitment during DNA replication (**Figure 17C, D**).

To further test, whether the lost DNA replication association of Tet1s-K852R and Tet1s-K852E (**Figure 17A**) affects global 5hmC levels, C2C12 myoblasts were transfected with the respective EGFP-tagged constructs and immunostained against 5hmC 24 hours later (**Figure 18A**). Tet1s, Tet1s-K852R, Tet1s-K582E all increased global 5hmC levels significantly, compared to EGFP-control transfected cells (**Figure 18B**). However, while a clear and significant increase of global 5hmC levels in Tet1s and Tet1s-K852R transfected

cells was observed with increasing EGFP levels and hence Tet1s levels, increasing Tet1s-K852E levels failed to further increase global 5hmC levels (**Figure 18B**).

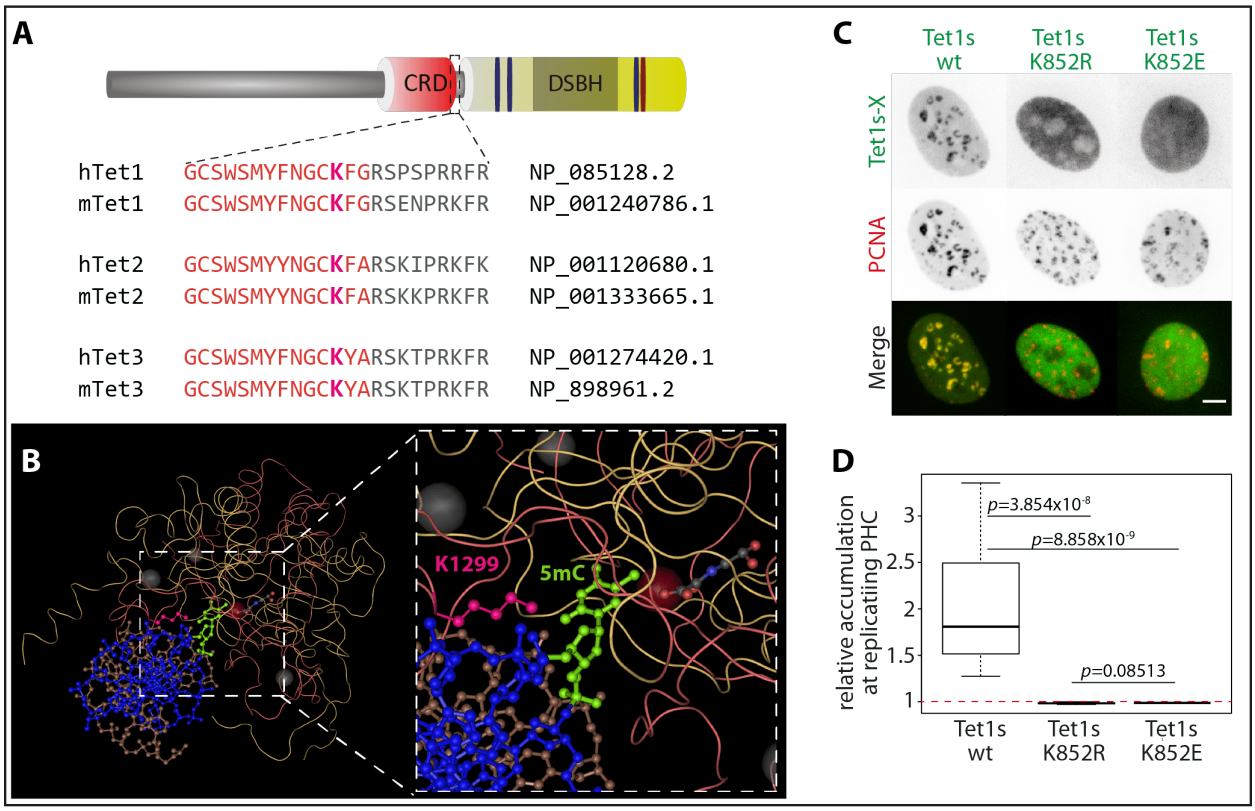


Figure 17: A conserved lysine in Tet1s CRD is crucial for its S-phase localization. (**A**) Localization within Tet1s and sequence alignment of a peptide from human and mouse Tet1, Tet2 and Tet3 surrounding a conserved lysine residue that was shown to be monoubiquitination and thereby modulate Tet chromatin binding and catalytic activity. The accession numbers are given next to the respective sequences. (**B**) Crystal structure of human Tet2 bound to a 12 base pair long stretch of methylated DNA at a resolution of 2.03 Å (PDB 4NM6) (Hu *et al.*, 2013). The CRD is shown as faint red tube worms, with the conserved lysine 1492 shown in pink, and the DSBH as faint yellow tube worms. The bound DNA is depicted as blue and brown stick representation helix with the methylated cytosine flipped out of the double helix shown in green. Zinc and iron atoms are shown as grey and red spheres, respectively. The 2-oxoglutarate analog, N-oxalylglycine is shown in stick representation. The model was visualized with Cn3D (Wang *et al.*, 2000). (**C**) Representative confocal mid-Z-sections of mouse myoblasts transiently expressing mRFP-PCNA together with EGFP-tagged Tet1s, Tet1s-K852R or Tet1s-K852E. (**D**) Boxplots showing the relative accumulation of Tet1s and its lysine mutants at PCNA marked DNA replication foci in C2C12 mouse myoblasts. The accumulation of the respective Tet1s-constructs was quantified by measuring the mean fluorescence intensity levels of Tet1s or its lysine mutants in PCNA marked PHC and dividing it by the mean levels in nucleoplasmic regions (n-numbers: Tet1s=13 Tet1s-K852R=18, Tet1s-K852E=21). Scale bar = 5 µm.

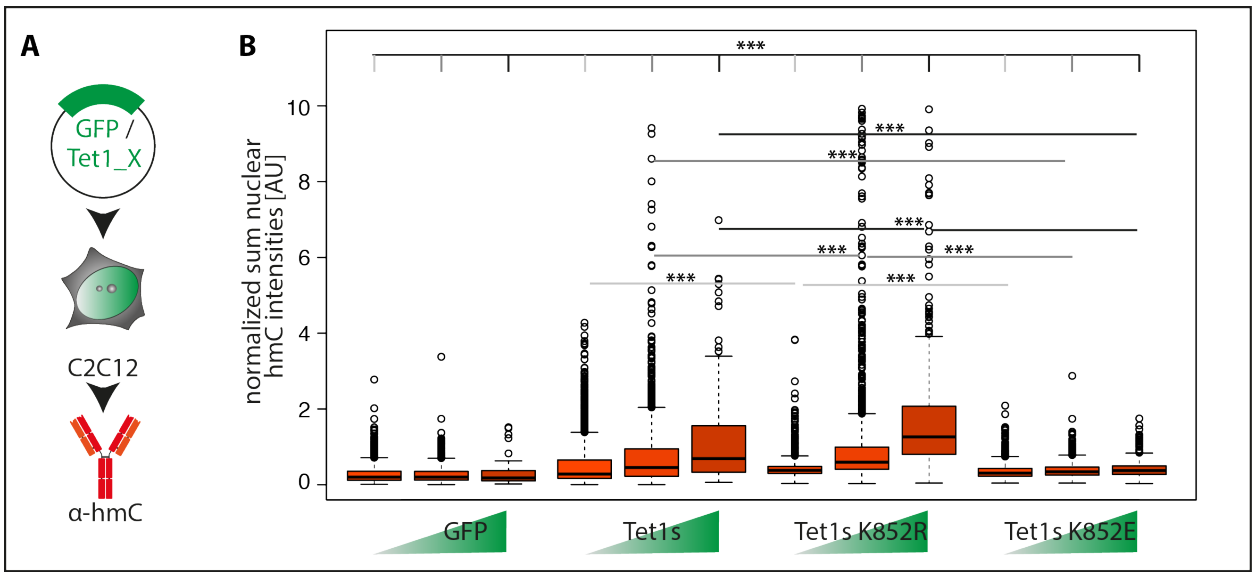


Figure 18: Tet1s lysine 852 mutations affect global 5hmC levels. (**A**) Schematic rationale of the experiment. C2C12 mouse myoblasts were transfected with EGFP, EGFP-Tet1s, EGFP-Tet1s-K852R or EGFP-Tet1s-K852E and 24 hours later immunostained against 5hmC and the DNA counterstained with DAPI. (**B**) Boxplots show the normalized sum nuclear 5hmC levels of C2C12 mouse myoblasts that transiently overexpressed the indicated EGFP-tagged proteins. Sum nuclear 5hmC levels were normalized to the averaged 5hmC averaged levels of lowest expressing cells and further against the sum nuclear DAPI intensity. Cells were grouped acording to their respective mean EGFP fluorescence intensities in low (50-100 AU), mid (100-500) and high (500-1000) expressing. Circles above boxes denote outliers. All experiments were done at least in triplicate and respective *p* values are denoted in the plot and outliers are depicted as circles above and below boxes (n-numbers: GFP=6255, Tet1s=5960, Tet1s-K852R=7482, Tet1s-K852E=16901). Statistical significance was tested with a paired two-samples Wilcoxon test, using R-studio version 1.1.447.

The finding that Tet1s-K852R overexpression significantly increased global 5hmC levels, compared to wildtype Tet1s overexpressing cells (**Figure 18B**) prompted me to examine the nuclear 5hmC distribution in closer detail. For this purpose, I acquired high-magnification confocal Z-stacks of randomly chosen cells that were previously imaged with the wide-field high content screening system. To analyse the sub-nuclear 5hmC deposition upon Tet1s-overexpression, a self-written, semiautomated ImageJ macro was used (**Figure 19A**). With this analysis method, mean fluorescence intensity levels of 5hmC in chromocenters and the surrounding nucleoplasm were measured and plotted. Similar to the observations, made for global 5hmC levels (**Figure 18B**), I found that Tet1s, Tet1s-K852R and also Tet1s-K852E showed significantly increased 5hmC levels in the nucleoplasm and chromocenters, compared to EGFP control-transfected cells (**Figure 19B, C**). While Tet1s transfected cells showed a significant increase of 5hmC in chromocenters, compared to Tet1s-K852E transfected cells, no major difference was found in cells that overexpressed Tet1s-K852R. In contrast to this, 5hmC levels in the nucleoplasm of cells that overexpressed Tet1s-K852R,

were significantly higher than in cells that overexpressed wildtype Tet1s (**Figure 19C**).

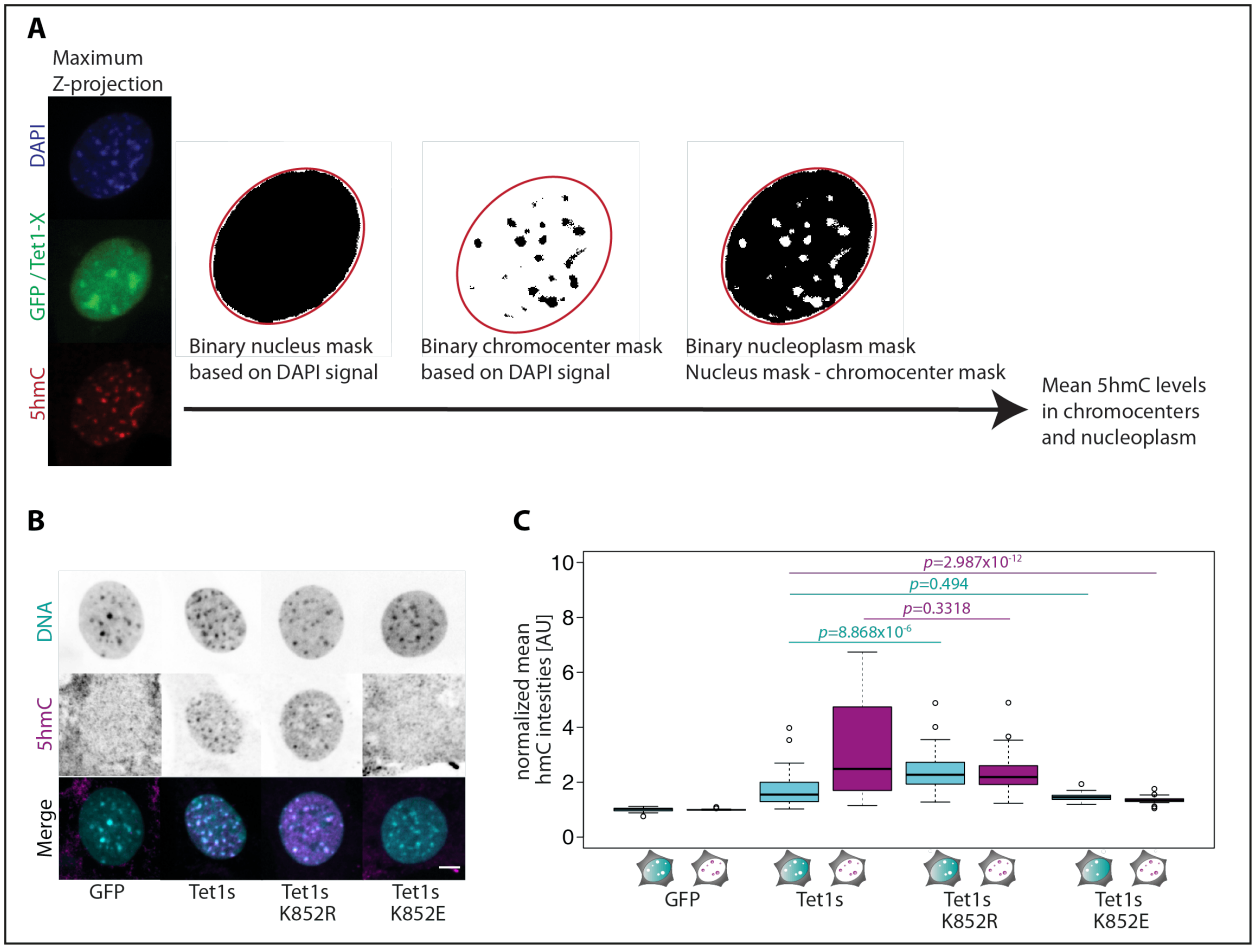


Figure 19: Tet1s lysine 852 mutations affect local 5hmC levels. (**A**) Schematic of the ImageJ-based *in situ* 5hmC analysis procedure. Maximum intensity projections of confocal Z-stacks were generated. Binary nuclear and chromocenter masks were generated based on the thresholded DAPI signals and a binary nucleoplasm mask was calculated from the DAPI and chromocenter masks. Mean fluorescence intensities in the respective areas, were measured. (**B**) Representative confocal, pseudocoloured mid-Z-sections of C2C12 cells that transiently overexpressed the indicated Tet1s variants or EGFP. (**C**) Boxplots show the quantification results of (A), hence 5hmC levels in C2C12 mouse myoblasts in chromocenters (cyan) and the surrounding nucleoplasm (purple), (n-numbers: GFP=19, S=50, K852R=39, K852E=33). Outliers are depicted as circles above and below boxes, *p*-values are given in the plot directly and were calculated with a paired two-samples Wilcoxon test, using R-studio version 1.1.447. Scale bar = 5 µm.

5.1.6 The CRL4(VprBP) complex might facilitate a crucial role in the S-phase association of Tet1

VprBP/DCAF1 (Viral protein r Binding Protein/DDB1 And CUL4 Associated Factor 1) is a threonine kinase that was shown to phosphorylate histone H3T120 (Kim *et al.*, 2013), but is most prominent for its role as an adapter protein for the DDB1-Cul4 E3 ubiquitin-ligase complex (Angers *et al.*, 2006). This complex and especially the E3 ubiquitin-ligases of the CRL (CUL4-RING E3 ubiquitn-ligase) family have been implicated in the regulation of

various cellular processes like DNA repair, but also replication. They were, for example, shown to prevent DNA re-replication in *C. elegans* by restraining origin licensing (Jackson and Xiong, 2009). Besides this role, recent studies found the proteins of the Tet dioxygenase family to be a common substrate for the CRL4(VprBP)-complex. Monoubiquitination of a conserved lysine residue in the CRD of all three Tet proteins, via the CRL4(VprBP)-complex, was linked to increased catalytic activity and in the case of Tet2, also chromatin loading (Yu *et al.*, 2013; Nakagawa *et al.*, 2015). Despite this, it was also found that the CRL4(VprBP)-complex can polyubiquitinate Tet2 upon HIV-1 infection, which results in proteasomal degradation of Tet2 and facilitates HIV-1 replication (Lv *et al.*, 2018). Little is known about the cell cycle- and in particular S-phase dynamics of the CRL4(VprBP)-complex or VprBP in general.

I, therefore, addressed the S-phase localization of VprpBP, with respect to the previous finding that mutations of the conserved lysine residue, known to be monoubiquitinated, abrogate the S-phase localization of Tet1s (**Figure 17C, D**). To this end, C2C12 mouse myoblasts were co-transfected with EGFP-Tet1, mcherry-VprBP and miRFP-PCNA encoding plasmids and subjected to live cell time-lapse microscopy. Cells in early S-phase were chosen, according to their respective PCNA pattern and followed over time while acquiring confocal Z-stacks every 20 minutes. Although VprBP showed a pancellular distribution, a slight cytoplasmic-to-nuclear translocation was observed from early to late S-phase (**Figure 20A, B**). The translocation was quantified by calculating the ratio of the mean nuclear to the mean cytoplasmic VprBP fluorescence intensity. In addition to this, line profile analysis of a late S-phase replication focus, showed that the VprBP signal follows the distribution of PCNA and Tet1s, despite its apparently homogeneous nuclear distribution (**Figure 20C**).

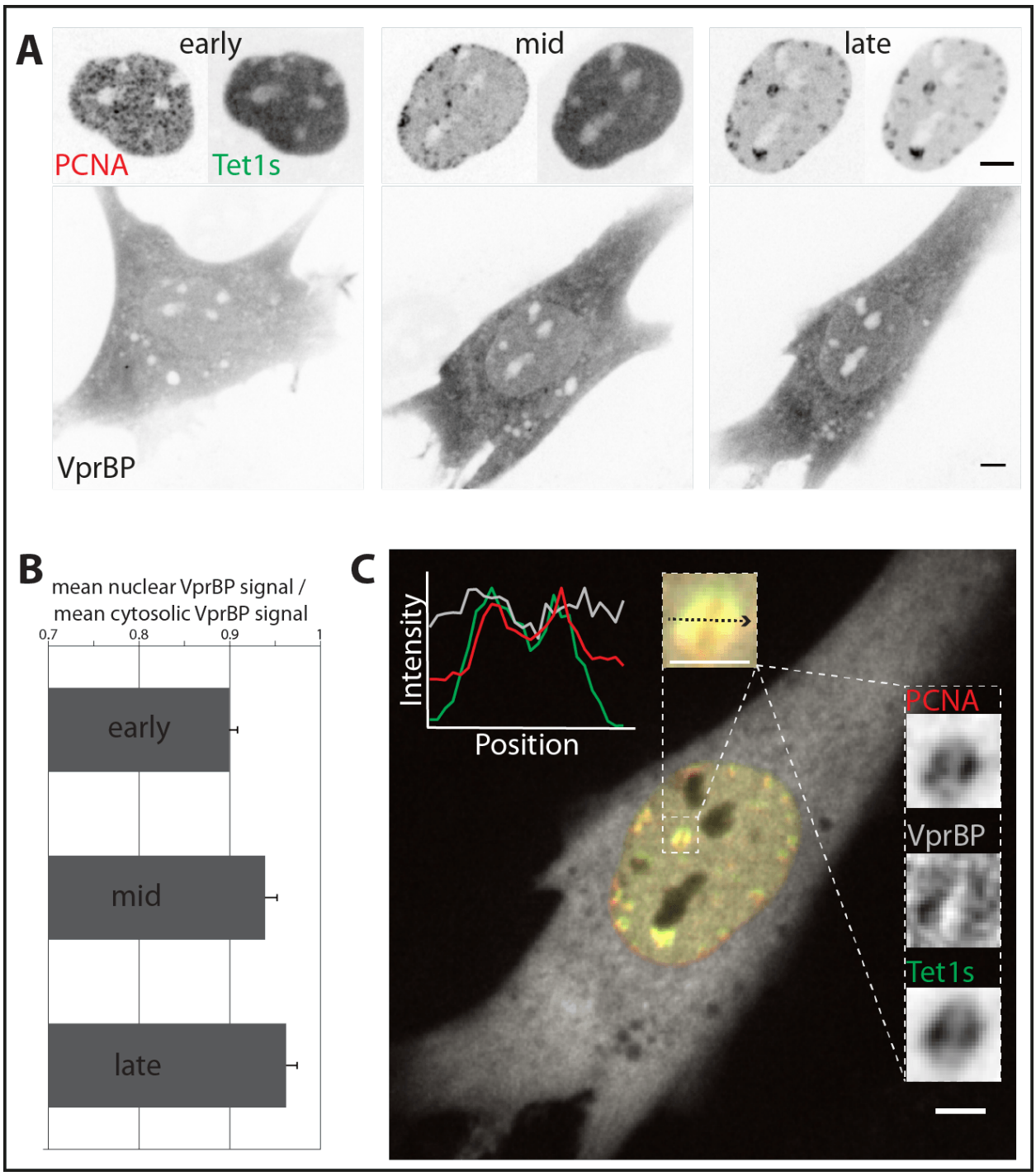


Figure 20: The CRL4-adapter protein VprBP colocalizes with Tet1s during late S-phase. (**A**) Representative mid-Z-plane images of C2C12 mouse myoblasts that were transiently transfected with EGFP-Tet1s, miRFP-PCNA and mRFP-VprBP and subjected to live-cell time lapse microscopy. Snapshots of the respective channels during early, mid and late S-phase are shown. (**B**) Bar graphs represent the ratio of the mean nuclear to mean cytoplasmic fluorescence intensity of mRFP-VprBP. (**C**) Pseudocoloured, detailed examination of the late S-phase cell shown in (A). A chromocenter that was marked by PCNA and Tet1s was examined in detail by line-profile analysis. Scale bars = 5 µm, except scale bar in blow-up in (C) = 2 µm.

5.2 Uhrf1 is crucial for Tet1s replication association

5.2.1 Loss of Uhrf1 but not Dnmt1 affects Tet1s localization

As the catalytic inactive mutant of Tet1 still localizes to DNA replication sites and therefore seems to be dispensable for its association with sites of ongoing DNA replication during late S-phase, I set out to test whether loss of the Tet substrate 5mC can affect the observed localization. For this purpose I made use of cells that are deficient for either the maintenance methyltransferase Dnmt1 or its facilitating factor Uhrf1, which both colocalize or even associate with PCNA, during DNA replication in pericentric heterochromatin.
Mouse embryonic fibroblasts deficient for Dnmt1 were also deficient for p53 (MEF-PM), as primary fibroblasts from Dnmt1 negative and p53 proficient embryos proved non-viable (Lande-Diner *et al.*, 2007). As loss of Uhrf1 is eventually lethal during embryonic development and differentiation, the effects of Uhrf1 deficiency were tested in mouse embryonic stem cells (Bostick *et al.*, 2007; Sharif *et al.*, 2007). Cells deficient for Dnmt1 and cells deficient for Uhrf1 are characterized by global hypomethylation and exhibit only residual genomic 5mC levels and are therefore suitable models to study the effects of 5mC depletion on Tet1s localization during the cell cycle. Dnmt1-deficient cells are characterized by decondensed pericentric heterochromatin and therefore often morphologically distorted chromocenters (Casas-Delucchi *et al.*, 2012a). For this reason, I co-transfected MEF-PM and MEF control cells that are only deficient for p53, with mcherry-Tet1s, miRFP-PCNA and EGFP-MaSat, to mark pericentric heterochromatin and imaged the cells live approximately eight hours later, to qualitatively assess the localization of Tet1s during late S-phase. The p53 deficient cells, and also the cells deficient for p53 and Dnmt1, showed a clear accumulation of Tet1s at PCNA labelled chromocenters (**Figure 21A**), indicating that neither the loss of the maintenance DNA methyltransferase Dnmt1 nor the accompanied global loss of DNA methylation does affect the association of Tet1s with sites of ongoing DNA replication in chromocenters.
The multi-domain protein Uhrf1 is mostly implicated in serving as a facilitating factor for Dnmt1 mediated DNA methylation maintenance. This is achieved by interpreting the combined information of the DNA methylation status and different histone modifications in the vicinity of hemimethylated CpGs during ongoing DNA replication. This triggers the E3

ubiquitin-ligase activity of Uhrf1 towards lysines in the histone H3 tail, which recruits the TS of Dnmt1 (Hashimoto *et al.*, 2009; Qin *et al.*, 2015). The loss of Uhrf is accompanied by severe global hypomethylation, similar to the loss of Dnmt1. To qualitatively address the effect of Uhrf1 loss, I co-transfected wildtype and Uhrf1-deficient embryonic stem cells with mcherry-Tet1s, miRFP-PCNA and EGFP-MaSat and imaged them live approximately 10 hours later. While E14 wildtype cells showed a clear colocalization of Tet1s, PCNA and MaSat, no accumulation of Tet1s was observed in Uhrf1-deficient E14 cells, at PCNA marked chromocenters (**Figure 21B**).

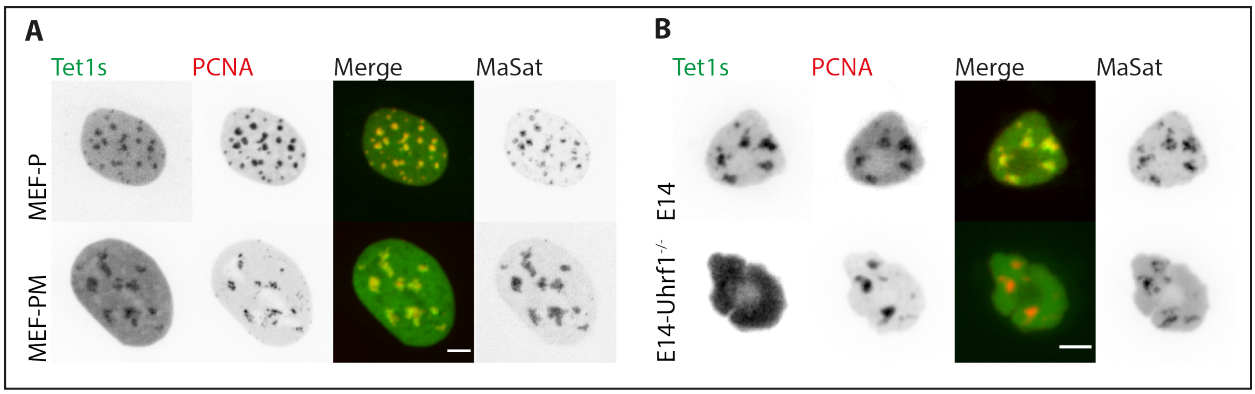


Figure 21: Loss of Uhrf1 prevents Tet1s accumulation at replicating PHC. (**A**) Representative mid-Z-plane sections of mouse embryonic fibroblasts deficient for p53 (MEF-P; upper panel) or p53 and Dnmt1 (MEF-PM; lower panel) transiently expressing mcherry-Tet1s, miRFP-PCNA and EGFP-MaSat. (**B**) Representative mid-Z-plane sections of wild type E14 mouse embryonic stem cells (upper panel) or E14 mouse embryonic stem cells deficient for Uhrf1 (E14-Uhrf1$^{-/-}$; lower panel) transiently expressing mcherry-Tet1s, miRFP-PCNA and EGFP-MaSat. Scale bars = 5 µm.

5.2.2 Uhrf1 physically interacts with Tet1 and can rescue the S-phase localization of its short isoform

Based on the finding that Tet1s accumulation is lost in Uhrf1-deficient cells, I was interested to see, whether Tet1 and Uhrf1 can physically interact. The putative protein-protein interaction was tested by co-immunoprecipitation experiments with single-domain deletions of Uhrf1 and the minimal catalytically active domain of Tet1, Tet1CD that still localizes to sites of ongoing DNA replication in PHC. For this purpose, HEK-293 cells were co-transfected with mcherry-tagged Tet1CD and EGFP-tagged single-domain deletion mutants of Uhrf1 and harvested approximately 36 hours after transfection. Cell extracts were prepared and incubated with GFP-binding nanobody immobilized on sepharose beads. The input fraction and the bound fraction were separated by SDS-PAGE and transferred to

nitrocellulose membranes. The fusion proteins were detected with monoclonal antibodies against GFP and red fluorescent proteins, respectively. While full-length Uhrf1 as well as all of its single-domain deletion mutants could pull down mcherry-Tet1CD, EGFP alone failed to so (**Figure 22**).

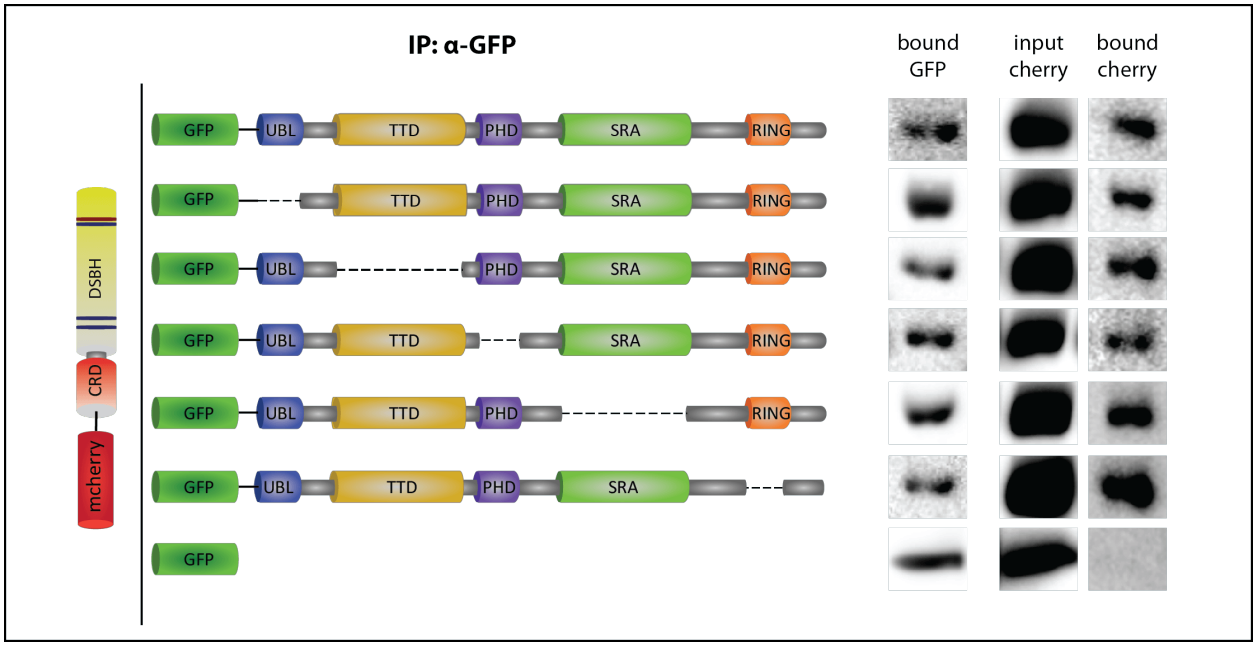


Figure 22: Uhrf1 physically interacts with Tet1. EGFP or EGFP-tagged Uhrf1 constructs were ectopically overexpressed in HEK-293 cells together with mcherry-Tet1CD and cell extracts were analysed by immunoprecipitation using an immobilized GFP-binding nanobody followed by western blotting with antibodies against GFP and red fluorescent proteins. Schematic protein structures of the used fusion-proteins are depicted on the left. The cut-outs on the right show the bound GFP-fractions and the input and bound mcherry fractions.

As no Tet1s-recruitment to replication foci in PHC was observed in Uhrf1-deficient E14 ESCs, I was prompted to test if Tet1s localization can be rescued by ectopically reintroducing Uhrf1. Besides this, I also tested, whether its highly similar homologue Uhrf2 can substitute Uhrf1 in Tet1s recruitment. Uhrf2 was shown to specifically bind 5hmC but was not implicated in DNA methylation maintenance due to a different chromatin modification readout (Vaughan *et al.*, 2018). Besides this, Uhrf2 is mostly absent from embryonic stem cells and found to be upregulated in somatic tissues (Pichler *et al.*, 2011). The effects of ectopic re-expression of Uhrf1 or Uhrf2 were tested by co-transfecting E14-Uhrf1$^{-/-}$ ESCs with miRFP-PCNA, mcherry-Tet1s and either EGFP, EGFP-Uhrf1 or EGFP-Uhrf2. After approximately 12 hours confocal Z-stacks were acquired and the accumulation of Tet1s in large PCNA foci that were assumed to represent ongoing PHC replication, assessed with

ImageJ as described earlier (**Figure 23B**). Cells that co-expressed EGFP-Uhrf1 showed a clear colocalization between Tet1s and PCNA and the relative accumulation was similar to the one that was measured in E14 wildtype cells, miRFP-PCNA, mcherry-Tet1s and Uhrf1. In contrast to this finding, EGFP and Uhrf2 failed to restore the accumulation of Tet1s in a Uhrf1-deficient background (**Figure 23A, C**).

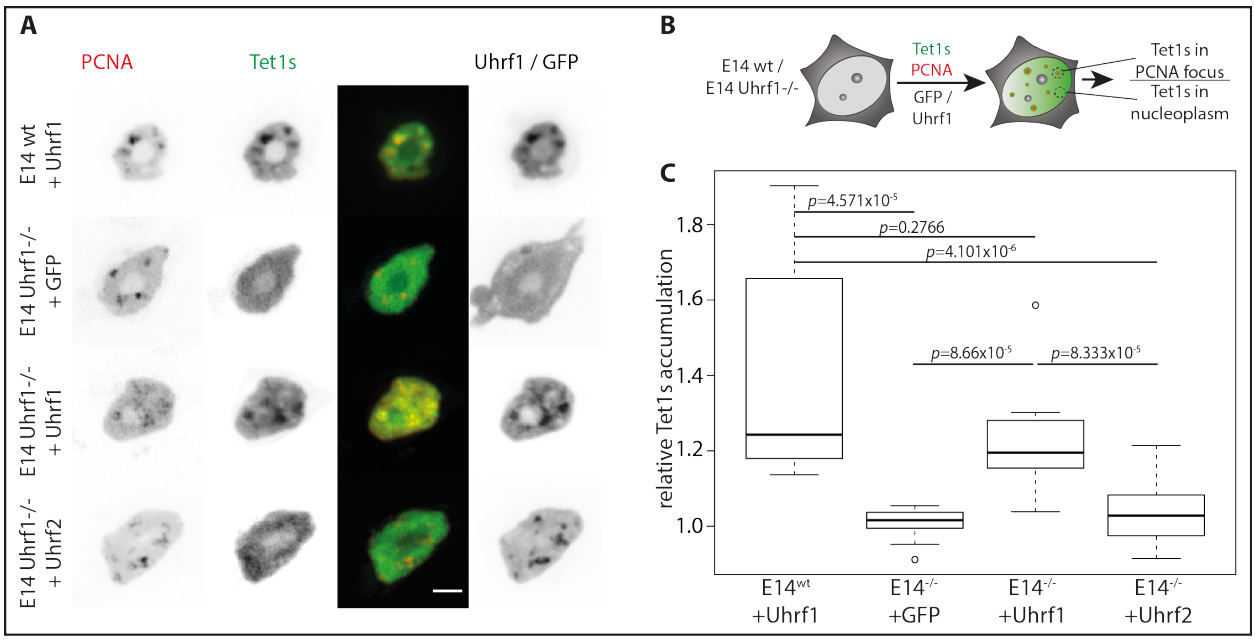


Figure 23: Uhrf1 but not Uhrf2 is required for Tet1s S-phase localization. (**A**) Representative pseudo-coloured confocal mid-Z-sections of E14 and E14$^{-/-}$ co-transfected with fluorescently tagged PCNA, Tet1s and either Uhrf1, Uhrf2 or EGFP. (**B**) Rational of the Tet1s accumulation quantification. Relative mcherry-Tet1s fluorescence intensity in PCNA marked chromocenters and surrounding nucleoplasm was quantified by measuring three areas each and putting them into relation. (**C**) Boxplots depict the quantification results of the accumulation measurements of E14 wildtype and E14-Uhrf1$^{-/-}$ co-transfected with fluorescently tagged PCNA, Tet1s and either Uhrf1, Uhrf2 or EGFP (n-numbers: E14+EGFP-Uhrf1=9, E14-Uhrf1$^{-/-}$+EGFP=12, E14-Uhrf1$^{-/-}$+EGFP-Uhrf1=10, E14-Uhrf1$^{-/-}$+EGFP-Uhrf2=23). Outliers are depicted as circles above boxes. *P*-values are given in the plot directly and were calculated with a paired two-samples Wilcoxon test in RStudio version 1.1.447. Scale bar = 5 µm.

From the co-immunoprecipitation experiments (**Figure 22**), I deduced that at least two domains of Uhrf1 can facilitate the interaction with Tet1, as the protein-protein interaction was observed with full-length Uhrf1, but also all of its single domain deletion mutants. This finding and my previous finding that lysine 852 of Tet1s, which was shown to be monoubiquitinated (Nakagawa *et al.*, 2015), is crucial for its S-phase localization (**Figure 17C, D**), led me to investigate, which effect the deletion of the UBL or the RING domains of Uhrf1 have on Tet1s S-phase localization. These two domains have been shown to concomitantly act on different protein targets, like the histone H3-tail, by forming a stable

E2-E3 complex during ubiquitin-ligation (Foster *et al.*, 2018). To test whether UBL- or RING-deletion mutants can rescue Tet1s recruitment, Uhrf1-deficient E14 cells were co-transfected with plasmids, encoding fluorescently tagged PCNA and Tet1s in combination with Uhrf1-ΔUBL or Uhrf1-ΔRING. Confocal Z-stacks were acquired approximately 12 hours later and analysed for Tet1s acumulation in PCNA foci in PHC, as described earlier (**Figure 23B**). Uhrf1-ΔUBL failed to significantly increase the accumulation of Tet1s in replicating PHC over the levels observed in EGFP control transfected Uhrf1 knock-out cells. In contrast, a significant increase in Tet1s accumulation was observed upon co-overexpression of Uhrf1-ΔRING in comparison to E14-Uhrf1$^{-/-}$ co-transfected with EGFP or EGFP-Uhrf1-ΔUBL encoding plasmids (**Figure 24B**). In regard to the finding that deletion of the E3 ubiqutin-ligase mediating RING domain failed to rescue the S-phase localization of Tet1s and the fact that lysine 852 in Tet1s was shown to be monoubiquitinated, I performed a CoIP of Uhrf1 with Tet1s or its respective lysine mutants. For this, HEK-293 cells were co-transfected with mcherry-tagged Uhrf1 and EGFP or EGFP-tagged Tet1CD, Tet1CD-K171R or Tet1CD-K171E mutants and protein lysates were prepared approximately 36 hours later. Cell extracts were incubated with an antibody against multiple red fluorescent proteins coupled to sepharose beads. The input and bound fractions were seperated by SDS-PAGE and transferred to nitrocellulose membranes. The fusion proteins were detected with monoclonal antibodies against GFP and red fluorescent proteins. While mcherry-Uhrf1 could not pull down GFP alone, EGFP-tagged wildtype Tet1CD as well as Tet1CD-K171R and Tet1CD-K171E could all be pulled down by Uhrf1 (**Figure 24C**).

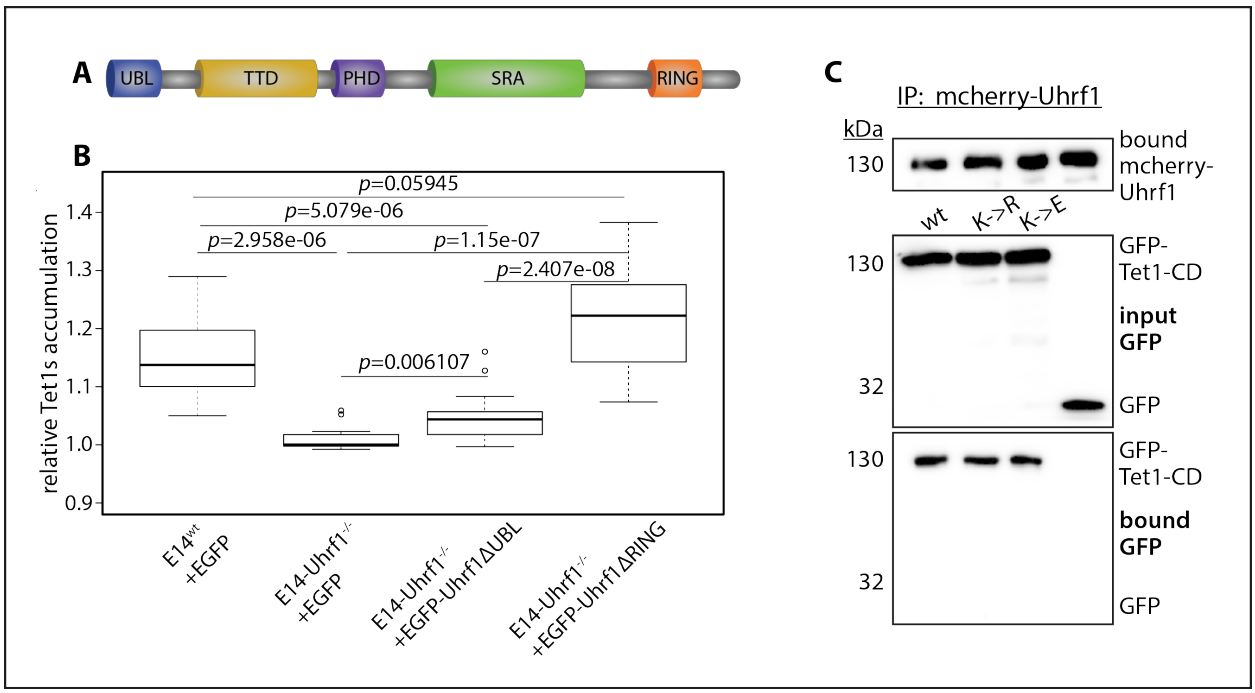


Figure 24: The UBL-domain of Uhrf1 is requiered for Tet1s S-phase association. (**A**) Schematic of full-length Uhrf1 indicating domain architecture. (**B**) Relative mcherry-Tet1s fluorescence intensity in PCNA marked chromocenters and surrounding nucleoplasm was quantified by measuring three areas each and putting them into relation. Boxplots depict the quantification results of the accumulation measurements of E14 wildtype and E14$^{-/-}$ co-transfected with fluorescently tagged PCNA, Tet1s and either Uhrf1 or EGFP (n-numbers: E14+EGFP=12, E14$^{-/-}$+EGFP=12, E14$^{-/-}$+EGFP-Uhrf1ΔUBL=21, E14$^{-/-}$+EGFP-Uhrf1ΔRING=21). Outliers are depicted as circles above boxes. *P*-values are given in the plot directly and were calculated with a paired two-samples Wilcoxon test in RStudio version 1.1.447. (**C**) EGFP or EGFP-tagged Tet1CD, Tet1CD-KtoR or Tet1CD-KtoE were ectopically overexpressed in HEK-293 cells together with mcherry-Uhrf1. Cell extracts were analysed by immunoprecipitation using an immobilized antibody against multiple red fluorescent proteins followed by western blotting with antibodies against GFP and red fluorescent proteins.

5.2.3 Tet1s localization is S-phase substage-dependent and independent of ligase1

Besides its binding to hemimethylated DNA, Uhrf1 was shown to recognize and bind trimethylated lysine 9 in the histone H3 tail (H3K9me3) (Liu *et al.*, 2013b). This modification was, however, shown to be dispensable for Uhrf1 recruitment to pericentric heterochromatin during ongoing DNA replication and subsequent DNA methylation maintenance by Dnmt1 (Lu *et al.*, 2015). I tested whether Tet1s recruitment is affected by the deficiency of the histone methyltransferase Suv39, which is responsible for di- and trimethylation of H3K9 (Peters *et al.*, 2002; Lehnertz *et al.*, 2003).

To address, whether Tet1s is still recruited during late S-phase in Suv39-deficient mouse embryonic fibroblasts (MEF), I first tested whether Uhrf1 can still be recruited. For this, I transfected wildtype MEF-W8 cells and Suv39-deficient MEF-D15 cells with fluores-

cently tagged Uhrf1, PCNA and MaSat encoding plasmids and imaged the cells live, approximately 8 hours later. In the wildtype as well as the Suv39-deficient cells, a clear colocalization of Uhrf1 and PCNA at chromocenters was observed (**Figure 25A**), as it was also published (Lu *et al.*, 2015). To further address the binding of Tet1s in these cells, they were again co-transfected with plasmids encoding fluorescently tagged PCNA, Uhrf1 and Tet1s and imaged live after about 8 hours. Again, a clear colocalization between PCNA, Uhrf1 and also Tet1s was observed (**Figure 25B**).

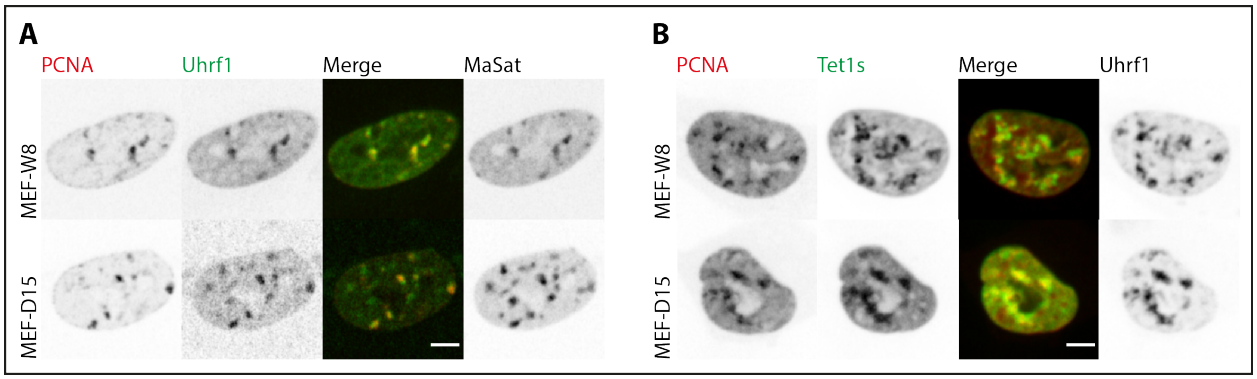


Figure 25: Tet1s localization in late S-phase is independent of the H3K9 methylation status. (**A**) Representative, pseudocoloured confocal mid-Z-sections of wildtype MEF-W8 and Suv39-deficient MEF-D15, cotransfected with fluorescently tagged PCNA, MaSat and Uhrf1. (**B**) Representative, pseudocoloured confocal mid-Z-sections of wildtype MEF-W8 cells and Suv39-deficient MEF-D15 cells, cotransfected with fluorescently tagged PCNA, Uhrf1 and Tet1s. Scale bars = 5 µm.

As I found Tet1s recruitment to sites of late DNA replication to be unaffected by DNA hypomethylation and abscence of the H3K9 methylation marks, I deemed replication association to be more important than general heterochromatin binding. To further address this, I made use of a fluorescent three hybrid (F3H) assay, using the MEF-PM cells that are deficient for Dnmt1 and, hence, retain only residual levels of 5mC within their major satellite repeats (Casas-Delucchi *et al.*, 2012a). I first assessed the Uhrf1 S-phase distribution in MEF-PM cells, by transfecting the cells with EGFP-Uhrf1 and mRFP-PCNA. While in early or mid S-phase only very few focal, if at all, Uhrf1 accumulations were observed, a clear, chromocenter-like, pattern was observed in late S-phase (**Figure 26A**). For the MEF-PM based F3H assay, the major satellite recognizing zinc-finger protein MaSat, fused to a GFP-binding nanobody (GBP-MaSat) (Casas-Delucchi *et al.*, 2012b)was co-transfected with a EGFP fusion protein as bait and a different fluorescently tagged protein as prey. EGFP or EGFP-fusion proteins are tethered to chromocenters via the GBP-MaSat and, thus, independent of the DNA methylation status and a colocalization with the prey protein

is observed in case of protein-protein interactions. If cell-cycle independent interactions are observed, this colocalization persists throughout the different S-phase and non-S-phase substages as we found for example, for Tet1 and the methyl-CpG-binding domain protein Mbd1 (Zhang *et al.*, 2017c).

To monitor the cell cycle stage, I co-transfected miRFP-PCNA together with the minimal Uhrf1-interacting Tet1 domain and EGFP or EGFP-Uhrf1 with or without GBP-MaSat in MEF-PM cells. Cells were imaged live 8 hours post transfection and all cells that showed EGFP-labelled chromocenter-like patterns were counted. Interestingly, around 50% of all cells with an EGFP-labelled chromocenter pattern showed a colocalization of mcherry-Tet1CD with the EGFP signal, independent of whether EGFP or EGFP-Uhrf1 were co-transfected. In contrast, almost 90% of cells without GBP-MaSat co-transfection but EGFP-Uhrf1 foci showed a colocalization between the mcherry-Tet1CD and EGFP-Uhrf1 signals (**Figure 26B**). When I analysed the respective PCNA patterns in these cells, it was observed that non-targeted cells with a clear colocalization between Tet1 and Uhrf1 showed a PCNA late S-phase pattern in most of the cases. However, cells co-expressing GBP-MaSat that showed a clear EGFP-Uhrf1 chromocenter localization, but no mcherry-Tet1CD accumulation were found to also be in early or mid S-phase or non-S-phase (not shown). Taken together, this *in situ* interaction assay emphasizes an S-phase substage dependent interaction between Tet1 and Uhrf1.

Tethering Uhrf1 to chromocenters could not provoke a premature Tet1CD recruitment, before late S-phase (**Figure 26B**). This furthermore hinted to a replication-dependent recruitment mechanism for Tet1s. A recent study proposed that Uhrf1 recruitment during DNA replication-coupled methylation maintenance, is achieved by binding to a H3K9me3-mimicking peptide in DNA ligase1. It was found that Uhrf1 recruitment, as well as global 5mC levels are severly decreased, upon mutating the respective peptide (Ferry *et al.*, 2017). Based on this finding, I was interested to test, whether abscence of ligase1 would affect the recruitment of Tet1s. For this, I made use of recently published ligase1-deficient E14 mouse ESCs and corresponding wildtype cells (Ferry *et al.*, 2017) and validated the ligase1 knockout by immunostaining against ligase1 (**Figure 27A**). To adress, whether Tet1s can still be recruited to DNA replication sites in PHC, I cotransfected the cells with plasmids that encode fluorescently tagged PCNA, Tet1s and MaSat and acquired confocal

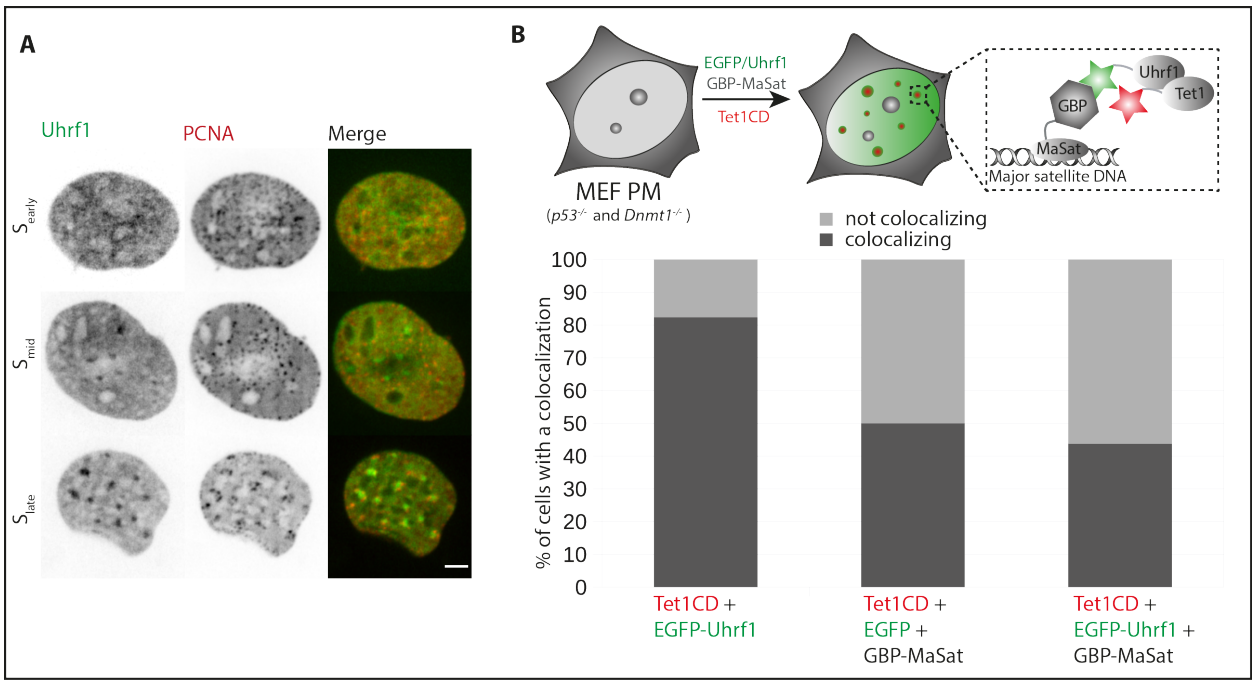


Figure 26: Tet1 localization to pericentric heterochromatin is restricted to late S-phase. (**A**) Confocal mid-Z-sections of MEF-PM cells, transiently expressing EGFP-Uhrf1 and mRFP-PCNA. Cells were imaged live approximately eight hours post transfection. Shown are cells in representative S-phase substages. (**B**) Schematic of the MEF-PM based F3H-assay. MEF-PM cells were transfected with mcherry-Tet1CD and EGFP or EGFP-Uhrf1 with or without GBP-MaSat. Approximately eight hours post transfection, cells with EGFP-labelled chromocenters were imaged live. The bargraph depicts the percentage of cells that showed a colocalization between EGFP-labelled chromocenters and mcherry-Tet1CD (n-numbers: EGFP-Uhrf1=17, GBP-MaSat+EGFP=20, GBP-MaSat+EGFP-Uhrf1=32). Scale bar = 5 µm.

Z-stacks live, approximately ten hours later. The accumulation of Tet1s in PCNA marked pericentric heterochromatin was assessed as previously described (**Figure 23B**). I found that this accumulation was largely unaffected between wildtype and ligase1-deficient ESCs (**Figure 27B, C**).

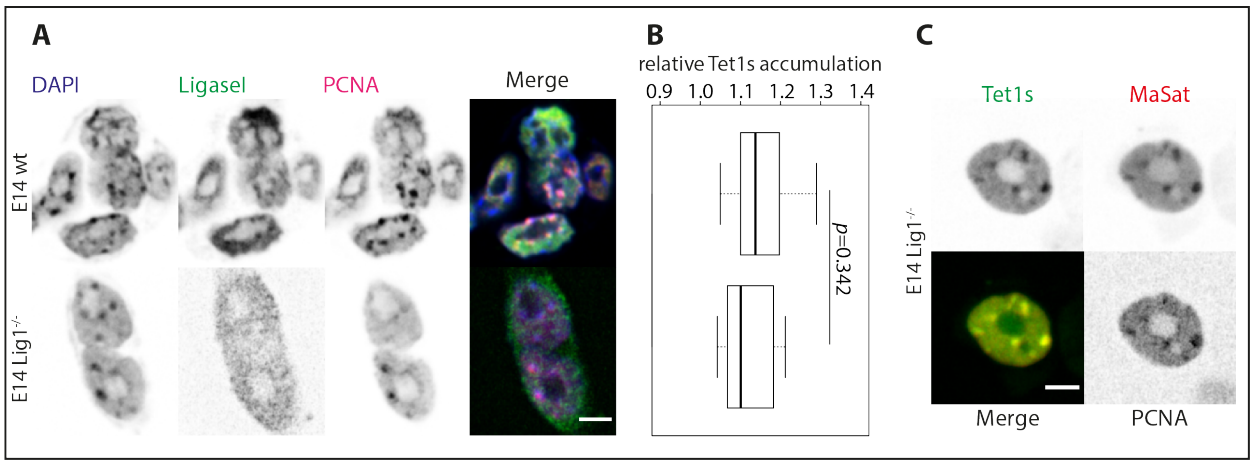


Figure 27: Tet1 localization to pericentric heterochromatin is not affected by the loss of DNA ligase1. (**A**) Confocal mid-Z-sections of E14 wildtype and E14-ligase1$^{-/-}$ mouse embryonic stem cells, immunostained against PCNA, ligase1 and counterstained with DAPI. (**B**) Boxplots depict the quantification results of the accumulation measurements in E14 wildtype and E14-ligase1$^{-/-}$ co-transfected with fluorescently tagged PCNA, Tet1s and MaSat (n-numbers: E14=12, E14-ligase1$^{-/-}$=8). (**C**) Pseudocolored, representative mid Z-section of a E14-ligase1$^{-/-}$ ESC, co-transfected with fluorescently tagged PCNA, Tet1s and MaSat. Scale bars = 5 µm.

5.3 Tet1s overexpressing cells show aberrant cytosine modification levels

5.3.1 Tet1s localization is conserved between mouse and human

The short isoform of Tet1 was identified to be expressed predominantly in somatic cells (Zhang *et al.*, 2016b) and was furthermore found to be overexpressed in different cancer cells (Good *et al.*, 2017). A hallmark of most cancer types are aberrant cytosine modification levels. On one hand, a global hypomethylation is observed in most cancers, while local hypermethylation is often found in the context of promoter and coding regions of tumour suppressor genes that are usually hypomethylated in normal tissues (Vidal *et al.*, 2017). In addition, many cancers also show a global loss of genomic 5hmC (Yang *et al.*, 2013). Interestingly, the canonical Tet1 promoter is found to become hypermethylated in cancer cells (Li *et al.*, 2016), which leads to the use of an alternate promoter and TSS and consequently the expression of a N-terminally truncated, but catalytically active isoform of Tet1, Tet1s (Good *et al.*, 2017). I therefore wanted to address the localization of Tet1s in human cancer cells.

MCF7 breast cancer cells were shown to overexpress Tet1s (Good *et al.*, 2017). As I performed all my previous experiments in murine cells, I first tested whether the described

S-phase localization pattern of Tet1s (**Figure 10B**) is conserved between mouse and human. For this, MCF7 breast cancer cells and MCF10a mammary gland cells that serve as non-tumorigenic control cells, were co-transfected with EGFP-Tet1s and mRFP-PCNA encoding plasmids and the localization of Tet1s analyzed. Cells in early, mid and late S-phase, identified by the respective PCNA pattern, were imaged live approximately 12 hours after transfection. In the MCF10a control cells, as well as the MCF7 breast cancer cells, a clear Tet1s accumulation at PCNA foci during late S-phase was observed, while no accumulation was found in early and mid S-phase cells (**Figure 28A**). As it was published that the MCF7 cells overexpress Tet1s (Good *et al.*, 2017) (**Figure 28B**), I immunostained them with a Tet1 specific antibody, which is directed against the unstructured C-terminal region that splits the DSBH and therefore can detect full-length Tet1 and also Tet1s. Surprisingly, a nuclear, focal Tet1 pattern was observed which, however, did not colocalize with PCNA and was, moreover, found in nuclear regions with a weak DAPI-staining (**Figure 28C**).

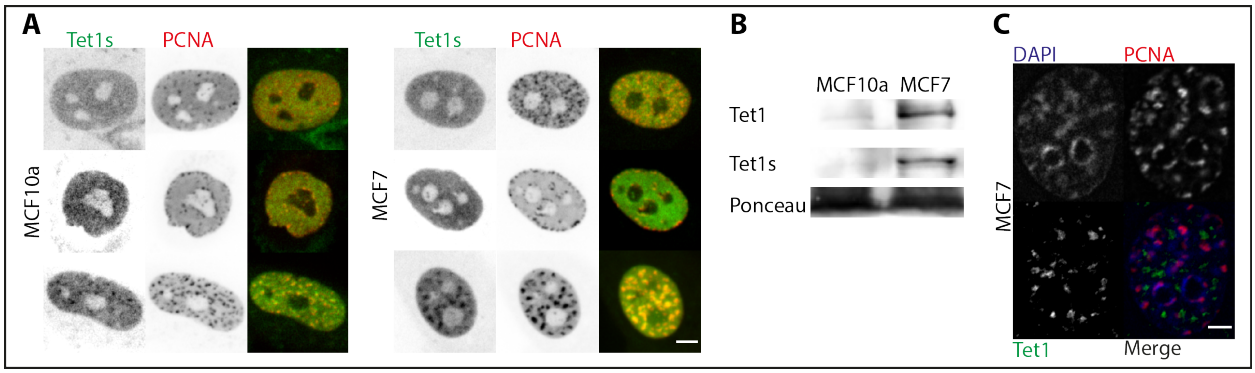


Figure 28: Tet1s is expressed in MCF7 breast cancer cells. (**A**) Shown are mid-Z sections of non-tumorigenic MCF10a breast epithelial cells and MCF7 breast cancer cells that were transfected with EGFP-Tet1s and mRFP-PCNA. Cells were imaged live approximately 12 hours post transfection and confocal Z-stacks were acquired. (**B**) Representative western blot of MCF10a and MCF7 cells probed with an antibody against Tet1. Cutouts show the respective Tet1 bands at approximately 235 kDa (Tet1) and 165 kDa (Tet1s). A cutout from the Ponceau stained membrane is shown as loading control. (**C**) Confocal mid-Z-section image of the nucleus of a MCF7 breast cancer cell that was immunostained against PCNA and Tet1 and the DNA counterstained with DAPI. Scale bars = 5 µm.

5.3.2 Cancer cells that overexpress Tet1s show aberrant cytosine modification levels

Many cancer cells are characterized by aberrant cytosine modification levels. In regard to the finding that the short isoform of Tet1 can also localize to late replication chromatin in cancer cells (**Figure 28A**), I was interested to see, whether global 5mC and 5hmC levels

are altered between MCF7 breast cancer cells and the non-tumorigenic MCF10a mammary gland cells. To test this, cells were seeded on glass coverslips and simultaneously immunostained against 5mC and 5hmC and counterstained with DAPI. Immunostained cells were imaged with a wide-field high content screening microscope and the respective sum nuclear fluorescence levels were quantified. To control, for the potential of the two antibodies blocking each others epitope access, I also performed single stainings for either 5mC or 5hmC. I found no intensity differences compared to double stained cells and therefore continued to analyse the fluorescence intensities of DAPI, Alexa-488 labelled 5mC and Alexa-594 labelled 5hmC in double-stained cells. For the analysis the sum nuclear 5mC and 5hmC levels were normalized to the averaged sum

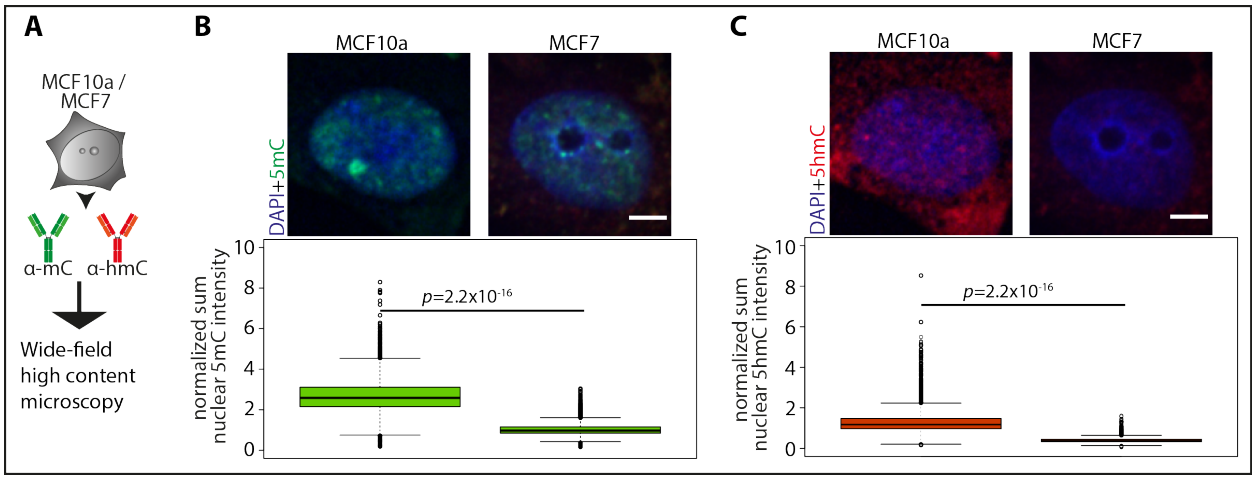


Figure 29: MCF7 breast cancer cells show aberrant cytosine modification levels. (**A**) Scheme of the experimental procedure. MCF10a and MCF7 cells were seeded on glass coverslips and the next day fixed and immunostained against 5mC and 5hmC and counterstained with DAPI. Cells were imaged with a wide-field high content screening microscope and sum nuclear fluorescence levels were quantified. (**B**) Representative confocal mid-Z-sections of 5mC and DAPI stained nuclei of MCF10a and MCF7 cells and corresponding boxplots showing the normalized sum nuclear 5mC intensity. Nuclear 5mC levels were normalized to the sum DAPI signal and the averaged fluorescence intensity signal of cells that were only incubated with the respective secondary antibody (n-numbers: MCF10a=9635, MCF7=32715). (**C**) Representative confocal mid-Z-sections of 5hmC and DAPI stained nuclei of MCF10a and MCF7 cells and corresponding boxplots showing the normalized sum nuclear 5hmC intensity. Nuclear 5hmC levels were normalized as just described (n-numbers: MCF10a=9635, MCF7=32715). Statistical significance was tested with paired two-samples Wilcoxon test in RStudio version 1.1.447. Scale bars = 5 µm.

nuclear Alexa-488 and Alexa594 fluorescence of cells that were only incubated with the secondary antibodies, but not the primary ones, to exclude background fluorescence from the analysis. To furthermore compensate for cell cycle effects and differences in DNA content, the signal was normalized to the sum nuclear DAPI intensity. Analysing the respective sum nuclear fluorescence levels in this way, I found that the MCF7 breast

cancer cells show significantly reduced global levels of 5mC and also 5hmC in contrast to the non-tumorigenic MCF10a cells. (**Figure 29B, C**).

5.3.3 Pericentric heterochromatin is depleted of 5mC and 5hmC in cancer cells

I have shown in this work that the S-phase recruitment of Tet1s to chromocenters in mouse myoblasts is accompanied by a massive 5hmC increase in pericentric heterochromatin (**Figure 19C**). Based on this finding and the reduction of global 5mC and 5hmC levels in MCF7 breast cancer cells (**Figure 29A, B**), I set out to quantify the respective cytosine modification levels in heterochromatic regions of MCF7 and MCF10a cells. In contrast to the acrocentric chromosomes of mice that all harbour a large pericentric heterochromatin domain (Guenatri *et al.*, 2004), large clusters of highly methylated pericentric heterochromatin in humans are predominantly found in chromosomes 1, 9, 16 and in male cells also the Y chromosome (Miller *et al.*, 1974). Hence, human cells show fewer and smaller heterochromatic subnuclear domains.

To address the levels of 5mC and 5hmC *in situ*, confocal Z-stack images were acquired of MCF10a and MCF7 cells that were simultaneously immunostained against 5mC and 5hmC and counterstained with DAPI. For the analysis of the 5mC and 5hmC levels in DAPI dense regions, maximum intensity Z-projections were generated and a circular region of interest with a radius of 5 pixels was chosen and the respective mean fluorescence intensities measured (**Figure 30A**). The mean levels of 5mC and 5hmC were normalized to the respective DAPI levels and plotted. Similar to the observations made for global 5mC and 5hmC levels, MCF7 cells showed a significant reduction of both modifications compared to MCF10a cells (**Figure 30C, D**), with a major decrease in 5mC levels. This is also reflected in the respective ratios of 5hmC to 5mC levels in the measured regions that almost doubled in MCF7 cells, compared to MCF10a cells (**Figure 30B**).

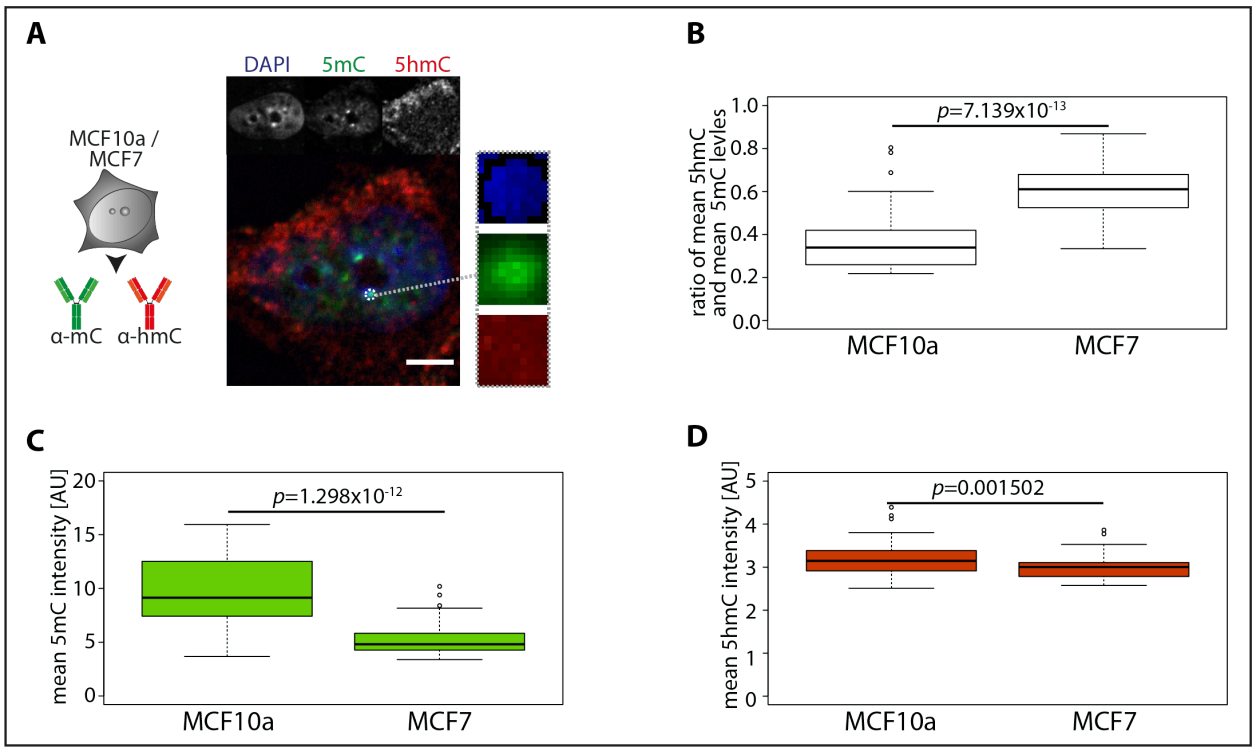


Figure 30: Levels of 5mC and 5hmC in heterochromatin-rich regions in breast cancer cells. (**A**) Rational of the experimental procedure. MCF10a and MCF7 cells were seeded on glass cover slips and the next day fixed and immunostained against 5mC and 5hmC and counterstained with DAPI. Confocal Z-stacks were acquired and analysed with ImageJ. Maximum intensity Z-projections were generated and the mean nuclear levels of 5mC and 5hmC in DAPI dense focal regions that represent heterochromatin-rich regions, measured in a ROI with a radius of 5 pixels and normalized to the respective DAPI signal. (**B**) The ratio of mean 5mC and 5hmC in DAPI-dense regions of MCF10a and MCF7 cells was plotted. (**C**) Mean fluorescence intensity of 5mC in DAPI-dense focal regions in MCF10a and MCF7 cell nuclei (n-numbers: MCF10a=32, MCF7=38). (**D**) Mean fluorescence intensity of 5hmC in DAPI-dense focal regions in MCF10a and MCF7 cell nuclei (n-numbers: MCF10a=32, MCF7=38). Statistical significance was tested with paired two-samples Wilcoxon test in RStudio version 1.1.447. Scale bars = 5 µm.

6 Discussion & Perspective

The main objective of the present study was to elucidate, whether Tet proteins can be recruited during ongoing DNA replication, which would emphasize a role in the maintenance of 5mC oxidation products during the cell cycle and in particular during S-phase. I therefore addressed the nuclear localization of the different Tet proteins and found a late S-phase recruitment of the short isoform of Tet1, but not for full-length Tet1, Tet2 or Tet3, during ongoing DNA replication in pericentric heterochromatin. I continued to investigate the domains of Tet1s that mediate S-phase recruitment and found that the cysteine-rich domain within the catalytic domain of Tet1 is needed not only for catalytic activity but also for the association with ongoing DNA replication in late S-phase. Based on this finding, I further dissected the CRD and found a conserved lysine to be crucial for the observed S-phase localization of Tet1s. S-phase localization was abrogated upon mutating the said lysine to arginine (Tet1s-K852R). However, global 5hmC levels were significantly increased in cells that transiently overexpressed Tet1s-K852R, compared to wildtype Tet1s overexpressing cells. This observation was explained by analysing 5hmC levels *in situ* and I found that Tet1s-mediated 5hmC formation was mostly restricted to pericentric heterochromatin, whereas Tet1s-K852R overexpression increased 5hmC levels on a more global scale. (**Section 5.1, 6.1**)

I could moreover show that Tet1s recruitment to PHC during DNA replication is independent of its catalytic activity and was also not affected by a strong decrease of global DNA methylation levels or the absence of the maintenance DNA methyltransferase Dnmt1. In contrast, I found that Uhrf1, an important facilitating factor of Dnmt1, physically interacts with Tet1. Besides this, I could show that Tet1s recruitment to DNA replication in pericentric heterochromatin is lost in Uhrf1-deficient cells, but can be rescued by ectopically reintroducing Uhrf1. (**Section 5.2, 6.2**)

Finally, I found the described S-phase localization of Tet1s to be conserved also in cancer cells that are known to overexpress Tet1s. I could show that these cancer cells, in comparison to non-tumorigenic cells, are characterized by a significant decrease of global 5mC and 5hmC levels globally and also in pericentric heterochromatin, whereby especially 5mC is depleted in heterochromatic domains of breast cancer cells. (**Section 5.3, 6.3**)

6.1 The short isoform of Tet1 localizes to pericentric heterochromatin during ongoing DNA replication

The proteins of the Tet dioxygenase family, Tet1, Tet2 and Tet3, have all been shown to oxidise the repressive DNA modification 5mC to create 5hmC, 5fC and 5caC and thereby diversify the epigenetic information in eukaryotic cells (Tahiliani *et al.*, 2009; Ito *et al.*, 2011; Pfaffeneder *et al.*, 2011). Similar to 5mC, the Tet-mediated oxidation products were shown to be stable DNA modifications throughout the cell cycle, which is in line with their specific signalling functions (Bachman *et al.*, 2014; Bachman *et al.*, 2015). Recently, a short isoform of Tet1, the founding member of the Tet protein family, was identified and shown to be overexpressed in cancer cells, which show loss of non-modified euchromatic regions, due to aberrant DNA methylation spreading (Good *et al.*, 2017). This observation was in part attributed to the N-terminal, chromatin-binding BC- and CxxC-domains of Tet1 that are not present in its short isoform, which therefore shows a significantly reduced chromatin-association (Zhang *et al.*, 2016b). The chromatin-binding ability of full-length Tet1 has before been implicated in the maintenance of euchromatin integrity by preventing DNA methylation spreading into non-methylated CpG-islands in somatic cells (Jin *et al.*, 2014).

Despite these findings, little is known on how the 5mC oxidation products are maintained throughout the cell cycle. This prompted me to investigate the cell cycle localization of the putative maintenance dioxygenase Tet1 (Jin *et al.*, 2014) and its short isoform, but also of Tet2 and Tet3. By addressing the spatio-temporal nuclear localization of Tet1 and Tet1s with live-cell time lapse microscopy, I observed a significant accumulation of Tet1s at sites of ongoing DNA replication in pericentric heterochromatin during late S-phase (**Figure 10B, C**). Full-length Tet1 on the other hand, did not show any focal accumulations during ongoing DNA replication in either of the S-phase substages. These difference in localization can be explained by the retention of full-length Tet1 in euchromatic regions via its CxxC- and BC-domains, which have both been shown to concomitantly facilitate the strong Tet1 chromatin association. Tet1s, on the other hand, which lacks both of these domains (**Figure 10A**), was shown to be found predominantly in the soluble fraction in

cellular fractionation experiments and furthermore to be highly susceptible to detergent extraction (Zhang *et al.*, 2016b). These findings in part also explain my observation that the late S-phase association of Tet1s can not be fixed, necessitating to address this localization in live-cell microscopy experiments. To further dissect the role of Tet1 and Tet1s in the nucleus during different S-phase stages and in different nuclear regions, FRAP (fluorescence recovery after photobleaching) experiments could provide insight into binding kinetics. Fast binding kinetics of Tet1s in chromocenters during DNA replication would further explain, why it could not be fixed. For full-length Tet1 on the other hand, a strong euchromatic binding is to be expected, while its kinetics in PHC are likely faster, due to its different mode of chromatin binding via the N-terminus.

In this line, FRAP experiments could also be used to further elucidate the role of the N-terminal BC- and CxxC-domains. The observation that the deletion of the CxxC-domain was sufficient to target the respective Tet1 deletion mutant to replicating chromocenters, hints to the CxxC-domain as major facilitator of the full-length Tet1 euchromatin localization (**Figure 11**), as it was also speculated before (Jin *et al.*, 2014). The observed premature heterochromatin association of the Tet1 CxxC deletion mutant, in comparison to Tet1s (**Figure 11B**), hints to a general chromatin association via the BC-domain, with a potential bias for methylated DNA. This could be addressed by testing the binding affinities of respective N-terminal deletion-mutants to differentially modified DNA substrates *in vitro*.

Interestingly, Tet1s overexpression markedly increased global 5hmC levels in comparison to cells that overexpressed full-length Tet1 (**Figure 10D**. In line with this finding, we could previously show that ectopic overexpression of the catalytic domain of Tet1, which shows the same S-phase localization as Tet1s (**Figure 12 A**), significantly increases global 5hmC levels. The significantly increased 5hmC levels, upon Tet1CD overexpression, were accompanied by increased reexpression and reactivation of major satellite repeats and LINE1 retrotransposable elements in mouse and human cells (Ludwig *et al.*, 2017; Zhang *et al.*, 2017b). Fittingly, a previous study found that overexpression of full-length Tet1 resulted only in minor transcriptional changes, whereas Tet1s overexpression resulted in fatal transcription defects during development (Zhang *et al.*, 2016b).

The observation that Tet1CD is sufficient for late S-phase localization, led me to investigate also the S-phase localization of the structurally, highly similar (**Figure 5A**) (Bauer *et al.*,

2015) catalytic domains of Tet2 and Te3, but also the respective full-length proteins. For neither of the variants of Tet2 or Tet3, S-phase accumulations, similar to those observed for Tet1s or Tet1CD, were found (**Figure 13**). This observation hints to a distinct biological role for the short isoform of Tet1. Accordingly, although all three Tet proteins share the same catalytic activity, distinct roles in development have been implicated (Santiago *et al.*, 2014).

Based on the finding that its catalytic domain is sufficient for the observed Tet1s-localization in late S-phase, I further dissected Tet1CD and tested whether its main domains, the CRD and the DSBH, can individually localize to late replicating DNA or be catalytically active. Both domains alone failed to increase global 5hmC levels, in stark contrast to Tet1CD (**Figure 14C**), which fits to structural studies on the catalytic domain of Tet2. This study found that a peptide within the CRD, which is conserved between all three Tet proteins, is needed to stabilize the DNA around the substrate base, in the catalytic center. The actual substrate and cofactor binding and the accompanied flipping of the substrate base are facilitated by the DSBH (Hu *et al.*, 2013). These findings underline that neither the CRD, nor the DSBH can act on 5mC or its oxidative derivatives by themselves.

When I addressed the respective late S-phase localization of the CRD and the DSBH, I found only a minor accumulation for the CRD, while the DSBH did not accumulate at all (**Figure 14A**). Fittingly, a Tet1s-mutant that lacks the CRD also failed to increase 5hmC levels and to localize to sites of ongoing DNA replication in PHC (**Figure 15**).

With regard to the observation that the deletion of the CRD renders Tet1s catalytically inactive and abrogates its S-phase localization, it would be interesting to see whether the fusion protein that results from the eponymous MLL-Tet1 genomic translocation (Ono, Taki, and Taketani, 2002; Lorsbach *et al.*, 2003) is catalytically active or shows any noteworthy S-phase localization patterns. The genomic breakpoint in intron 8 of Tet1 results in a fusion protein where the N-terminus of MLL/KMT2 that harbours a DNA-binding AT-hook and a CxxC-domain, is fused to the C-terminal part of the catalytic domain of Tet1, hence its DSBH without the CRD (Allen *et al.*, 2006; Hamidi, Singh, and Chen, 2015). As demonstrated in this work and known from structural work on Tet2 (Hu *et al.*, 2013), the CRD of Tet proteins is essential for their catalytic activity and in the case of Tet1s also for the S-phase recruitment. Hence, the CxxC-domain and/or the AT-hook of MLL could either

replace the CRD and its chromatin-binding ability in the MLL-Tet1 fusion protein or it would be rendered catalytically inert by the underlying genomic translocation.

As the CRD was identified as the domain that mediates late S-phase association of Tet1s, I examined it for motifs that could facilitate this accumulation. A previous study found Tet1 to interact with the replication-associated clamp loader protein PCNA, which was shown to mediate Tet1 activity (Cartron *et al.*, 2013). In regard to this finding, I screened the proteinsequences of full-length Tet1, Tet2 and Tet3 for a PCNA binding domain (PBD) that is characterized by a conserved consensus sequence. I found a highly similar sequence in the CRD of Tet1 (**Figure 16A**) but not Tet2 or Tet3 and proceeded to test its PCNA association capability. The major difference between the putative PBD of Tet1 and the PBD-consensus sequence, is a hydrophilic lysine that replaces the central hydrophobic amino acid residue in canonical PBDs and the putative PBD of Tet1 did to not colocalize with PCNA during ongoing DNA replication (**Figure 16B**). This finding can most likely be explained with the just described central amino acid substitution and it would be interesting to see, whether the putative PBD of Tet1 could be rendered active by exchanging the lysine to a leucine or isoleucine. A successful reactivation of the putative PBD, hence a restored interaction with PCNA, would hint to an initial function that was lost during evolution or replaced by a different mode of recruitment, respectively.

The before described conserved peptide in the CRD of all Tet proteins that stabilizes the DNA around the substrate was found to harbour a central lysine (K852), whose CRL4(VprBP)-mediated monoubiquitination was linked to the modulation of the catalytic activity in all three Tet proteins (**Figure 17A, B**) (Yu *et al.*, 2013; Nakagawa *et al.*, 2015). Mutating this conserved lysine in Tet1s to arginine or glutamate abrogated replication association during late S-phase completely (**Figure 17C, D**), emphasizing that it indeed plays a crucial role during Tet1s replication association. Surprisingly, global 5hmC levels in cells that overexpressed Tet1s-K852R were significantly increased compared to cells that overexpressed wildtype Tet1s or Tet1s-K852E (**Figure 18**). Overexpression of Tet1s-K852E resulted only in a very minor 5hmC increase, which can be explained by the DNA-binding role of the surrounding peptide (Hu *et al.*, 2013). Whereas the positively charged lysine or arginine can stabilize the negatively charged phosphate backbone of the DNA, substitution to a negatively charged glutamate could significantly distort this interaction and only a

residual catalytic activity is retained. To further examine the markedly increased 5hmC levels in Tet1s-K852R overexpressing cells, I addressed the nuclear 5hmC distribution *in situ* (**Figure 19A**). In the course of this experiment, I found that Tet1s overexpressing cells showed a significant 5hmC increase in chromocenters, hence, pericentric heterochromatin (**Figure 19B, C**). In contrast to this, 5hmC levels in cells that overexpressed Tet1s-K852R were also significantly increased in the nucleoplasm, outside of chromocenters, compared to Tet1s overexpressing cells. This hints to a targeted mode of Tet1s catalytic activity, which could be further addressed by investigating cells that overexpress lower levels of Tet1s or Tet1s-K852R. On the other hand, the respective mutants could be introduced into the endogenous Tet1 locus to observe the effects of physiological expression levels on 5hmC production. This would of course also necessitate to ensure the expression of only the short isoform of Tet1s. The expression of only the short isoform could be achieved by either deleting exon 1 from the endogenous locus or render the canonical promoter inactive, to trigger expression of only the short isoform.

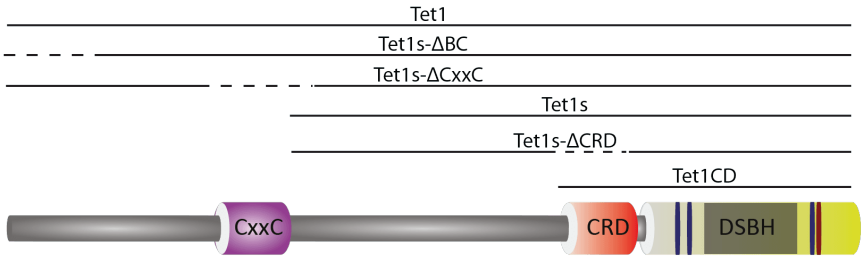
On the basis of these findings, I was interested in the cell cycle localization of the proteins of the CRL4(VprBP)-complex that were found to be responsible for the monoubiquitination of the conserved lysine (**Figure 17A**) in all Tet proteins (Yu *et al.*, 2013; Nakagawa *et al.*, 2015). Although, the proteins of the CRL4(VprBP)-complex have been implicated in various cellular processes, like DNA replication or repair (Angers *et al.*, 2006; Jackson and Xiong, 2009), little is known about their subcellular dynamics. The histone H3 directed kinase VprBP serves as the central adapter molecule in the CRL4(VprBP)-complex, bringing together the E2 ubiquitin-conjugating enzymes, the E3 ubiquitin-ligases and their respective targets (Angers *et al.*, 2006). Examining the localization of VprBP throughout S-phase by live-cell time lapse microscopy, showed a cytoplasmatic to nuclear translocation with S-phase progression, besides its pancellular localization (**Figure 20A, B**). Interestingly, when late S-phase DNA replication foci in cells that co-overexpressed fluorescently tagged PCNA, Tet1s and VprBP were examined in close detail, a colocalization of the three proteins was discovered. The observed cytoplasmatic to nuclear translocation of VprBP, as well as the replication foci association in late S-phase is in line with previous findings. It was demonstrated that the chromatin association of VprBP, but also its associated E3 ubiquitn ligase Cul4A, which targets Tet proteins (Yu *et al.*, 2013; Nakagawa *et al.*, 2015),

increases throughout S-phase and is strongest during late S-phase and G2 (McCall *et al.*, 2008). Considering that VprBP bridges the interaction between Tet proteins and the CUL4 E3 ubiquitn-ligases, it is very likely that Tet1s ubiquitination by the CRL4(VprBP) complex takes place during late S-phase. This would eventually recruit Tet1s to sites of ongoing DNA replication for targeted catalytic activity in pericentric heterochromatin. Mutations of the conserved lysine residue (**Figure 17C, D**) resulted in the loss of S-phase recruitment of Tet1s and furthermore support this idea. Consistently, knockdown of Cul4A, Cul4B or VprBP was shown to negatively affect the catalytic activity of Tet proteins (Yu *et al.*, 2013; Nakagawa *et al.*, 2015).

However, it still needs to be tested, whether depleting cells of Cul4A, Cul4B or VprBP also abrogates the observed S-phase localization of Tet1s. In order to address this question, lentiviral knockdowns of different components of the CRL4(VprBP) complex are being established, to test the effect of depletion of the respective proteins. If the monoubiquitantion of lysine 852 in Tet1s plays a role in its late S-phase recruitment, this pattern might be lost upon knockdown of the respective E3-ligases. If the ubiquitination of lysine 852 is also required for targeted catalytic activity, depletion of the Tet1s-K852 monoubiquitination facilitators might result in a similar nuclear 5hmC pattern, as it was observed for the Tet1s-K852R mutant that also lost its S-phase localization (**Figure 17C, D, 19**).

Taken together, I found that the short isoform of Tet1 is recruited to sites of ongoing DNA replication in pericentric heterochromatin, which is accompanied by a massive increase of 5hmC levels in chromocenters. Full-length Tet1 recruitment to these sites, is likely prevented by its binding to non-modified euchromatin via its N-terminal CxxC-domain in interplay with the adjacent BC-domain (**Figure 31**).

A



B

	Tet1	Tet1s	Tet1s K852R	Tet1s K852E	Tet1 ΔBC	Tet1 ΔCxxC	Tet1 CD	Tet1 CDmut	Tet1s ΔCRD	Tet1 CRD	Tet1 DSBH	Tet2	Tet2 CD	Tet3	Tet3 CD
accumulation at PHC during DNA replication	—	+++	—	—	+	++	+++	+++	—	+	—	—	—	—	—
global 5hmC increase	+/-	++	+++	+/-	n.d.	n.d.	+++	— (n.d.)	—	—	—	n.d.	+++ (n.d.)	n.d.	+++ (n.d.)

C

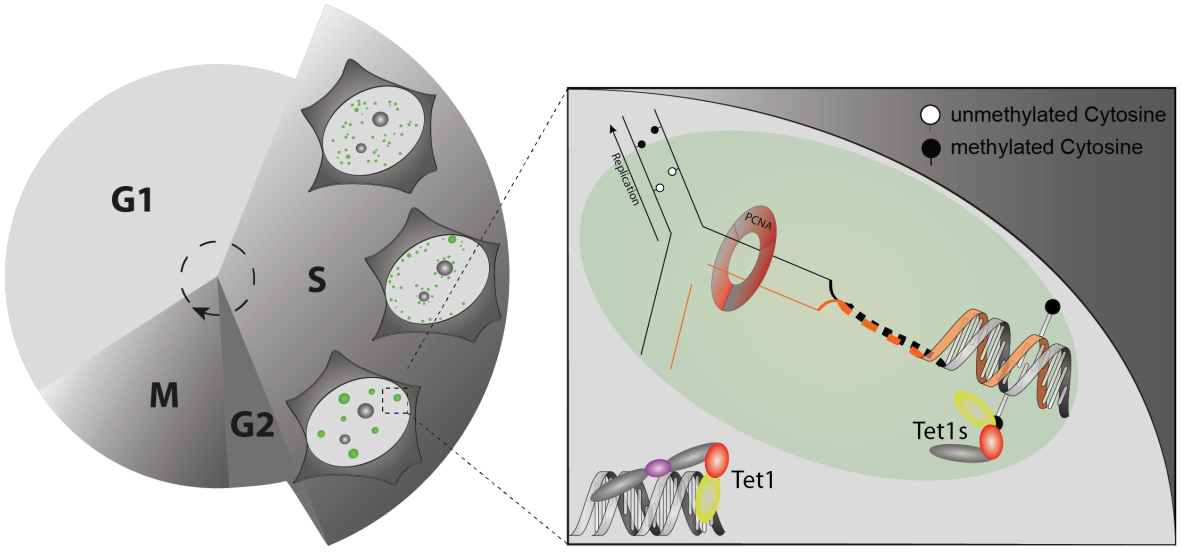


Figure 31: Summary of S-phase behaviour and catalytic activity of different Tet proteins and subdomains. (**A**) Schematic protein structure of Tet1. Lines above the protein structure denote different constructs that were used in this study. (**B**) Table summarizing the association capability of different ectopically expressed Tet1, Tet2 and Tet3 constructs at PCNA marked late replicating pericentric heterochromatin. Furthermore the effect on global 5hmC levels, upon overexpression is summarized (n.d.=not determined in this study. Catalytic activity of Tet2CD and Tet3CD was addressed previously (Zhang *et al.*, 2017c).) (**C**) Schematic summary of the key findings of (**Section 5.1**). Tet1s is recruited to sites of ongoing DNA replication in late S-phase, which is accompanied by an increased 5mC oxidation. This recruitment depends on a conserved lysine within the CRD of Tet1. Full-length Tet1 is not recruited during DNA replication and is known to bind to non-modified DNA in euchromatic regions via its N-terminal BC- and CxxC-domain (Zhang *et al.*, 2016b).

6.2 Uhrf1 is crucial for Tet1s replication association

Uhrf1 has been implicated as a crucial factor in DNA methylation maintenance, as it recognizes and binds hemimethylated DNA and interacts with the histone H3-tail in the vicinity of hemimethylated CpGs (Hashimoto *et al.*, 2008; Qin *et al.*, 2015). This multimodal

chromatin interaction recruits Dnmt1 for methylation maintenance during DNA replication. Loss of Dnmt1 or Uhrf1 results in severe phenotypes (**Section 2.5.5**), underlining their crucial role in cellular homeostasis. As cells that lack Dnmt1 or Uhrf1 are characterized by global DNA hypomethylation (Lande-Diner *et al.*, 2007; Sharif *et al.*, 2007), I was interested to see, whether loss of either of these two proteins or the consequent loss of DNA methylation could affect the observed Tet1s localization during late S-phase.

While qualitatively assessing the localization of Tet1s in Dnmt1 and Uhrf1 deficient cells, I found that Tet1s recruitment to ongoing DNA replication in pericentric heterochromatin is unaffected by the loss of Dnmt1 and the drastically reduced 5mC levels (**Figure 21A**). The qualitative assessment of Tet1s localization during ongoing DNA replication in PHC in Uhrf1-deficient mouse embryonic stem cells, on the other hand, showed no accumulation, while a clear Tet1s recruitment in the corresponding wildtype cells was observed (**Figure 21B**, **Figure 23**).

Tet1s was previously shown not to be expressed in ESCs, but only in somatic cells upon differentiation (Zhang *et al.*, 2016b). Hence, this indicates that the general recruitment mechanism that underlies the S-phase localization of Tet1s is not restricted to somatic cells, but also active in ESCs and therefore independent of the differentiation status of the cell.

In many of the live-cell experiments, I made use of a triple transfection with fluorescently tagged Tet1s, PCNA and also a major-satellite repeat recognizing polydactyl zinc-finger protein (Lindhout *et al.*, 2007), to mark pericentric heterochromatin in addition to DNA replication sites (**Figure 21, 12**). Interestingly, we could show in a previous publication that Tet1CD mediated 5hmC formation in pericentric heterochromatin is counteracted by blocking DNA access through sequence-specific DNA-binding proteins, like said major-satellite recognizing zinc-finger (Ludwig *et al.*, 2017). Despite this finding, co-overexpression with this zinc-finger protein did not prevent Tet1s or Tet1CD localization to sites of ongoing DNA replication in PHC (**Figure 21, 12**). This hints to a protein-protein interaction mediated replication association, via players, whose chromatin-association is unaffected by the occupancy of binding sites in major satellite repeat-containing heterochromatic DNA by another protein. With respect to the multimodal heterochromatin association of Uhrf1 during ongoing DNA replication and the observed absence of Tet1s from DNA replication

sites in PHC in Uhrf1-deficient cells, this underlines the importance for Uhrf1 in Tet1s replication association.

Prompted by the finding that Uhrf1-deficiency abrogated Tet1s replication association, I tested, whether Uhrf1 and Tet1 can physically interact. For this purpose, I performed co-immunoprecipitation experiments with Tet1CD, hence the minimal domain of Tet1 that still colocalizes with ongoing DNA replication and is also still catalytically active (**Figure 12A, 14C**) and full-length Uhrf1 as well as different single-domain deletion mutants. With this approach I found that immobilized full-length Uhrf1, as well as, all of its single domain deletion mutants were able to pull down Tet1CD (**Figure 22**).

The findings that Uhrf1-deficiency abrogates Tet1s recruitment to DNA replication sites in PHC and that Uhrf1 and Tet1 physically interact, led me to test, whether the absent Tet1s recruitment in Uhrf1-deficient cells can be rescued by reintroducing Uhrf1. With this approach, I found that ectopic overexpression of Uhrf1 in Uhrf1-deficient cells could restore the accumulation of Tet1s to similar levels as observed in wildtype ESCs (**Figure 23A, C**). I have shown initially that Tet1s accumulation was not affected by Dnmt1-deficiency (**Figure 21A**) or the accompanied, severely decreased DNA methylation levels (Lande-Diner *et al.*, 2007). It can, therefore, be assumed that potentially changing DNA methylation levels, upon Uhrf1 reintroduction, did not play a role in the restored Tet1s accumulation. Interestingly, the structurally highly similar Uhrf1 homologue, Uhrf2 failed to rescue the Tet1s accumulation phenotype. Uhrf2 was identified to specifically recognize and bind 5hmC in neuronal progenitor cells (NPCs) via its SRA domain (Spruijt *et al.*, 2013b; Zhou *et al.*, 2014). The failure to restore Tet1s localization in an Uhrf1-deficient background, despite its affinity for 5hmC, can likely be explained by the different chromatin readout in comparison to Uhrf1 (Vaughan *et al.*, 2018). The findings that the E3 ubiquitin-ligase activity of Uhrf2 towards histone H3 is promoted by hydroxymethylated DNA (Vaughan *et al.*, 2018) however pose an interesting starting point, to further investigate, how 5hmC is maintained in general during the cell cycle in somatic cells, where Uhrf2 is predominantely expressed (Pichler *et al.*, 2011).

From the findings in the co-immunoprecipitation experiments that all single-domain deletion mutants of Uhrf1 can physically interact with Tet1, I deduced that at least two domains of Uhrf1 are responsible for this interaction. In the present work, I could show that a

conserved lysine within the CRD of Tet1s that is known to be monoubiquitinated, is needed for the association of Tet1s with ongoing DNA replication in PHC (**Figure 17C, D**). In regard to my finding and the previously described monoubiquitination of this lysine (Yu *et al.*, 2013; Nakagawa *et al.*, 2015), I reasoned that the UBL and the RING domain of Uhrf1 are most likely the domains responsible for the protein-protein interaction with Tet1. The interplay of the UBL and the RING domain of Uhrf1 was recently shown to be crucial during the maintenance of DNA methylation in DNA replication (Foster *et al.*, 2018). This study found that the UBL of Uhrf1 connects the E2 ubiquitin-conjugating enzyme UbcH5a with the RING-domain of Uhrf1 to facilitate histone H3-tail monoubiquitination, which in turn results in faithful DNA methylation maintenance via Dnmt1 (Qin *et al.*, 2015).

When I tried to ectopically rescue the localization of Tet1s in Uhrf1-deficient cells with either a UBL- or a RING-domain deficient Uhrf1 construct, I found that deletion of the UBL failed to restore Tet1s localization. Deletion of the RING domain, on the other hand, resulted in a significantly increased Tet1s accumulation compared to control-transfected Uhrf1-deficient ESCs (**Figure 24B**). This observation hints to the UBL of Uhrf1 as the major Tet1s interacting domain. I furthermore tested whether the interaction of Uhrf1 and Tet1 is mediated via the conserved lysine residue of the CRD of Tet1 by co-immunoprecipitation experiments. However, immobilized Uhrf1 could successfully bind to wildtype Tet1CD and also its respective KtoR or KtoE mutants (**Figure 24C**). Taken together these results indicate that the UBL of Uhrf1 plays a crucial role in Tet1s recruitment during ongoing DNA replication, but does not bind to Tet1 via the conserved lysine in its CRD.

Prompted by the physical interaction between Tet1 and Uhrf1, I tested whether Tet1s can be recruited to chromocenters outside of DNA replication, by tethering Uhrf1 to major-satellite repeats (**Figure 26B**). With this approach I could show that the *in situ* interaction of Tet1 and Uhrf1 is S-phase substage dependent, as Uhrf1 tethering could not increase the number of cells with Tet1-Uhrf1 co-localization compared to EGFP-control transfected cells. A recent study could show that pericentric heterochromatin can be re-localized and tethered to the nuclear periphery. This was shown to affect the histone-modification pattern, but also to change replication timing of pericentric heterochromatin to a more mid S-phase like state (Heinz *et al.*, 2018). On the basis of the finding that tethering Uhrf1 to chromocenters could not localize Tet1s to pericentric heterochromatin outside

of late S-phase, it would be interesting to see, whether Tet1s is recruited to re-localized chromocenters at the nuclear lamina already during mid S-phase. This would indicate that not only pericentric DNA and associated histones are being localized to the nuclear lamina, but also that this re-localization dramatically changes the signalling cascade that is responsible for Tet1s recruitment during late S-phase.

The S-phase localization and activity of Uhrf1 was recently shown to depend not only on different chromatin modifications, but also on the binding to a H3K9me3-mimicking peptide in ligase1 (Ferry *et al.*, 2017). However, testing Tet1s accumulation capabilities in ligase1 deficient mouse ESCs showed no difference compared to wildtype cells. This could be explained by a putative backup pathway via ligase3 (Le Chalony *et al.*, 2012). Accordingly, cells that are deficient for ligase1 show no DNA methylation defects, while cells that express ligase1 mutants that lack the H3K9me3-mimic, show a massive loss of global DNA methylation levels (Ferry *et al.*, 2017). In the future it will therefore be tested, if cells that only express H3K9-mimic deficient mutants of ligase1 show the same Tet1s accumulation pattern as ligase1-deficient or wildtype ESCs.

Taken together, I found Uhrf1 to be crucial for Tet1s recruitment during ongoing DNA replication in pericentric heterochromatin. I furthermore found that the UBL-domain is crucial for Tet1s localization and that the interaction between Tet1 and Uhrf1 is independent of a conserved lysine that is known to be monoubiquitinated and to modulate Tet dioxygenase activity. It was recently found that the UBL of Uhrf1 serves as a stabilizing factor that brings together E2 ubiquitin-conjugating enzymes and the E3 ubiquitin-ligase RING domain of Uhrf1, as well as the respective substrate (Foster *et al.*, 2018). Moreover, I found that Uhrf1 tethering to chromocenters can not induce a premature Tet1s recruitment outside of late S-phase. These recent findings on the role of the UBL of Uhrf1 and my own observations prompt me to speculate that the Uhrf1 UBL domain fulfils a stabilizing task also in the monoubiquitination of Tet1s by bringing it together with the CRL4(VprBP)-complex. The main components of this complex, VprBP, undergoes a cytoplasmatic to nuclear translocation with progressing S-phase, which indicates a nuclear activity mostly during late S-phase and G2. This would also explain, why Uhrf1 tethering failed to recruit Tet1s outside of late S-phase, as the physical interaction might be disrupted upon failure to monoubiquitinate Tet1s. The final recruitment of monoubiquitinated Tet1s dur-

ing DNA replication is likely facilitated by a factor that still needs to be identified (**Figure 32**).

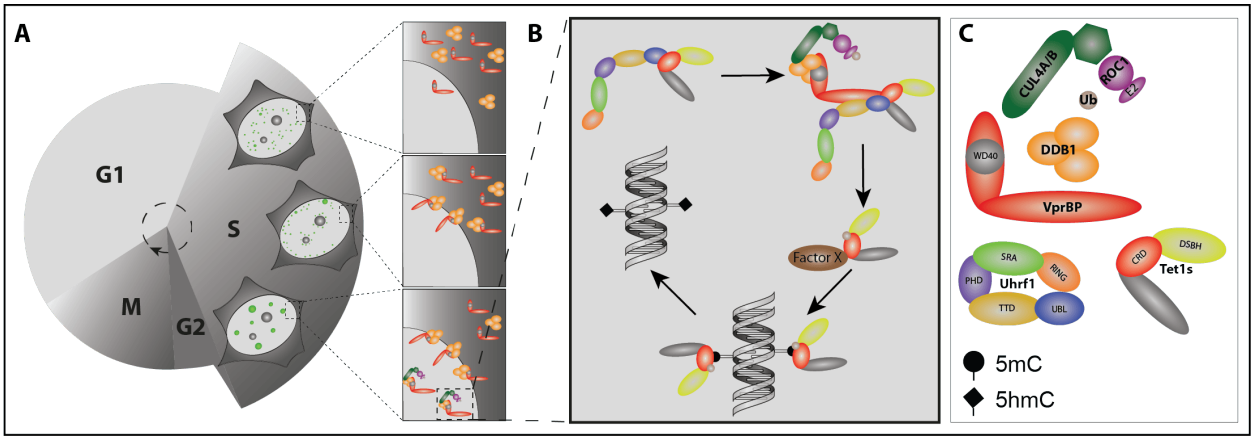


Figure 32: Potential mode of Tet1s recruitment during DNA replication. (**A**) Different members of the CRL4(VprBP)-complex, like DDB1 or VprBP itself, undergo a cytoplasmatic to nuclear translocation during progressing S-phase and the CRL4(VprBP)-complex is eventually assembled in the nucleus. (**B**) Schematic model of Tet1s recruitment incorporating the findings from the present study and literature data that were discussed here. Uhrf1 binds to Tet1s via its UBL-domain and recruits it to the assembled CRL4(VprBP)-complex. The UBL of Uhrf1 stabilizes Tet1s, while a monoubiquitnation of lysine 852 takes place. Monoubiquitinated Tet1s is recruited to sites of ongoing DNA replication via a yet unknown factor. Tet1s binds methylated cytosine in the context of pericentric heterochromatin and oxidizes it to 5hmC. (**C**) Schematic overview of the different proteins that are involved in the hypothesized Tet1s recruitment during DNA replication.

6.3 Tet1s-overexpressing cells show aberrant cytosine modification levels

Global DNA hypomethylation is a characteristic epigenetic alteration in cancer cells (Gaudet *et al.*, 2003). However, despite the strong reduction of global 5mC levels, local hypermethylation events are frequently observed in the context of promoters or coding sequences of tumour suppressor genes, which results in their transcriptional silencing (Baylin and Jones, 2011). It has been speculated, whether this feature of malignant neoplasms arises from a stochastic *de novo* deposition of aberrant DNA methylation marks or in fact represents a tumour-specific epigenetic code, which would hint to an instructive model for *de novo* DNA methylation in cancer (Struhl, 2014). It could, for example, be shown that genomically adjacent regions and functionally related genes are often found to be novo methylated in a mode that opposes a stochastic model (Keshet *et al.*, 2006; Serra *et al.*, 2014).
Interestingly, also the canonical Tet1 promoter is found to become hypermethylated in cancer cells (Li *et al.*, 2016), which leads to the use of an alternative promoter and a

thereby targeted alternative transcription start site (**Figure 10A**). This eventually results in a N-terminally truncated, but catalytically active short isoform of Tet1 (Good *et al.*, 2017), whose biological relevance I characterized in the present work.

The S-phase localization of Tet1s that I characterized in murine cells (**Figure 10B**), is conserved also in human cells and even in cancer cells (**Figure 28A**). This finding hints to a conserved molecular mechanism for the Tet1s recruitment that is not affected by the malignant transformation in breast cancer cells.

Interestingly, no colocalization of the Tet1 signal and the PCNA signal was found in late S-phase, when the localization of endogenous Tet1 was addressed by immunostaining Tet1 in MCF7 breast cancer cells that overexpress Tet1s (Good *et al.*, 2017). This observation might be attributed to the fact that the short isoform of Tet1 lacks the N-terminal BC- and CxxC-domains, which together contribute to chromatin binding in hypomethylated, euchromatic regions (Zhang *et al.*, 2016b). The loss of these domains and the thereby increased kinetics and decreased fixability of Tet1s could again be addressed by FRAP experiments, as discussed before (**Section 6.1**). The overexpression of Tet1s in cancer cells was found to be accompanied by a failure to prevent DNA methylation- and hence heterochromatin spreading into euchromatic domains (Good *et al.*, 2017). However, cancer cells are in general found to be hypomethylated (Gaudet *et al.*, 2003). Besides my finding that MCF7 breast cancer cells show a strong global reduction of 5mC and also 5hmC levels, compared to non-tumorigenic mammary gland cells (**Figure 29B, C**), it has previously been published that these cells overexpress the short isoform of Tet1s, which I could confirm (**Figure 28B**). Non-tumorigenic MCF10a mammary gland cells, on the other hand, expressed only minor levels of Tet1s. Considering my initial findings on the localization and the targeted catalytic activity of Tet1s in pericentric heterochromatin (**Figure 10B, 19B, C**), the global 5mC and 5hmC reduction in MCF7 cells could be accounted for by the massive Tet1s overexpression. By further addressing, differences in 5mC and 5hmC levels *in situ* in heterochromatic regions of MCF10a and MCF7 cells, I again found significant differences between the non-tumorigenic and the breast cancer cell lines (**Figure 30C, D**). However, while normalized 5mC levels in pericentric heterochromatin regions of MCF7 cells were reduced by almost 50% in comparison to non-tumorigenic MCF10a cells, 5hmC levels showed only a small, but still significant, reduction in MCF7 cells. This minor decrease in 5hmC, compared to

the strong loss of 5mC is also underlined by the ratio of 5hmC to 5mC levels in PHC. The 5hmC/5mC ratio is significantly increased in MCF7 cells compared to MCF10a cells, which is mostly due to the massive 5mC decrease (**Figure 30B**). Furthermore, the comparatively minor changes of the 5hmC levels could, besides the cancer phenotype, also be attributed to normal DNA modification fluctuations between different individuals, as the MCF7 and MCF10a cells were obtained from two different donors (Palumbo *et al.*, 2018). In regard to the estimate that about 5% of all cytosines in the human genome are methylated of which a large number is found in a heterochromatic CpG-context (Smith and Meissner, 2013), a decrease of about 50% in methylation levels accounts for a considerable number of cytosines that can not be considered as normal variation between individuals. Overall, the observed 5mC reduction, especially in heterochromatin, supports my initial findings from murine cells that Tet1s overexpression and its targeted catalytic activity in pericentric heterochromatin, results in a significant 5mC oxidation, hence 5hmC production (**Figure 10D, Figure 19**).

Interestingly, Uhrf1 has been found to contribute to global hypomethylation in the cancer context, despite its crucial role in DNA methylation maintenance. Accordingly, it was shown that the DNA *de novo* methyltransferase Dnmt3A is targeted for ubiquitin-dependent proteasomal degradation by Uhrf1, which was found to be frequently overexpressed in various cancers (Jia *et al.*, 2016). In line with these findings, Uhrf1 overexpression in different cancer types like squamous cell carcinoma or gastric cancer was found to be correlated with poor prognosis for affected patients. Moreover and again in stark contrast to the role of Uhrf1 in DNA methylation maintenance, decreased DNA methylation levels in LINE1 and its accompanied reactivation, are found to be inversely correlated to Uhrf1 expression in these cancers (Nakamura *et al.*, 2016; Hong *et al.*, 2018).

We have shown before that Tet1CD mediated 5mC oxidation can reactivate LINE1 retrotransposition in human cells (Zhang *et al.*, 2017b). Besides this, we found that loss of the crucial 5mC binding protein MeCP2, also results in the lost protection of PHC against 5mC oxidation by Tet1. This results in significantly increased 5hmC levels and accompanied increased reexpression of major satellite repeats in mouse brain (Ludwig *et al.*, 2017). The reexpression of LINE1 and satellite repeats is a commonly observed feature of many epithelial cancer types (Ting *et al.*, 2011). Moreover, genomic insertions upon LINE1

reactivation are found to be disruptive to genes, where they are newly inserted and in this way can render tumour suppressor genes inactive. Besides this, the LINE1 encoded proteins, ORF1p and ORF2p, have also been implicated in the activation of oncogenic transcription factors (Xiao-Jie *et al.*, 2016).

Taken together, my findings allow me to speculate on the involvement of Tet1s and also Uhrf1 in an expanded instructive model for a cancer-specific epigenetic code. Aberrant *de novo* DNA methylation in cancer cells is thought to be deposited, in part, in a directed mode to eventually silence tumour suppressor genes (Serra *et al.*, 2014). The role for Uhrf1 in Tet1s recruitment for DNA replication association in highly methylated PHC and the accompanied increase of 5mC oxidation in those regions hints to a secondary mechanism, besides aberrant local *de novo* hypermethylation. As discussed, we have shown that the catalytic activity of Tet1CD in highly methylated heterochromatic regions is accompanied by the reactivation of retrotransposable elements that are also reactivated in Uhrf1-overexpressing cells. The finding that Tet1s is overexpressed in many cancer types (Good *et al.*, 2017) makes a co-overexpression of Uhrf1 and Tet1s a likely scenario, which would, based on my findings, result in a targeted loss of DNA methylation in pericentric heterochromatin.

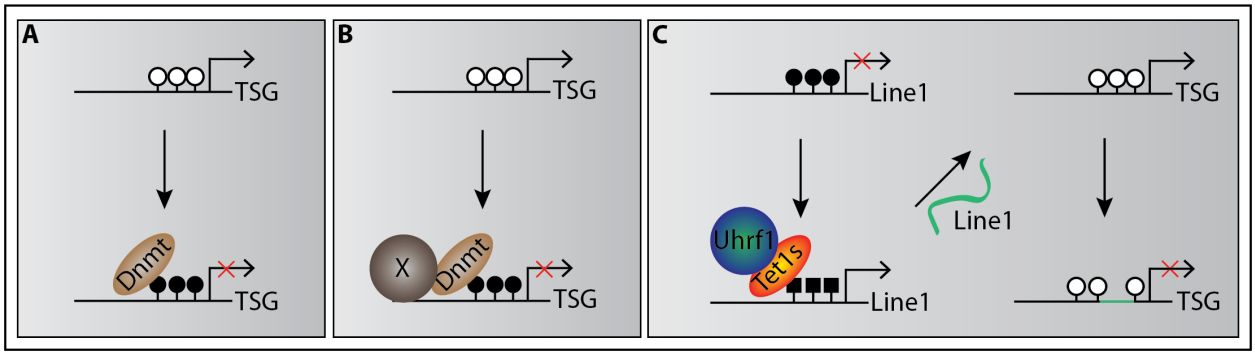


Figure 33: Potential role of Tet1s in an instructive cancer epigenetics model. (**A**) Stochastic model for DNA methylation in cancer. *De novo* DNA methylation is deposited by a DNA methyltransferase randomly across the genome, eventually silencing a tumour suppressor gene (TSG). (**B**) Instructive model for DNA methylation in cancer. A DNA methyltransferase is recruited to the genomic locus of a tumour suppressor gene (TSG) for *de novo* DNA methylation, by a sequence specific protein (X). This eventually silences the tumour suppressor gene. (**C**) Putative instructive model for the role of Tet1s in cancer epigenetics. Highly methylated heterochromatin that contains LINE1 is targeted by the Uhrf1-Tet1s axis. This results in the oxidation of the methylated DNA and eventually the reexpression of LINE1. Random or directed retrotransposition in loci of tumour suppressor genes (TSG), disrupts their expression. White circles: non-modified cytosine, black circles: methylated cytosine, black squares: hydroxymethylated cytosine.

6.4 Conclusion & Outlook

The proteins of the Tet dioxygenase family have been implicated in crucial developmental processes and in this line also in various diseases and recurringly in different cancer types **Section 2.5**. In my study, I could show that the cancer-associated short isoform of Tet1s can be recruited to ongoing DNA replication in an Uhrf1-dependent manner (**Figure 23**), which is very likely also dependent on the proteins of the CRL4(VrpBP) E3 ubiquitin-ligase complex, given my finding of the the subcellular dynamics of the central protein VprBP (**Figure 20**). To further address this, knockdowns of the different components of the CRL4(VprBP)-complex will be used to dissect the interplay of this complex, Uhrf1 and also the conserved lysine residue 852 within the CRD of Tet1s, during Tet1s recruitment. The interplay of Uhrf1 and the catalytic domain of Tet1 poses another interesting question, as mutations of the conserved lysine residue did not affect the interaction of the two proteins. Structural data hint to the low-complexity insert within the DSBH of Tet1 as a protein-protein interaction scaffold, which will be further addressed (Hu *et al.*, 2013) (**Section 6.1, 6.2**).
Similar to the previously discussed FRAP experiments that can be used to examine the nuclear kinetics of Tet1 and Tet1s and address the role of different domains of Tet1, live cell experiments with different inhibitors will be used to dissect the role of ubiquitination in replication association of Tet1. First attempts to alter Tet1s recruitment during S-phase, by treatment with the E1 ubiquitin-activating enzyme inhibitor PYR-41 (Yang *et al.*, 2007) proved also toxic to the progression of DNA replication in general. Therefore, a more directed mode for abrogation of Tet1s ubiquitination would be the use of a specific inhibitor for Cul4 E3 ubiquitin-ligases, assesing if they are responsible for lysine 852 monoubiquitination (Lan *et al.*, 2016).
The overexpression of Tet1s in cancer cells was found to be accompanied by a failure to prevent DNA methylation and hence heterochromatin spreading into euchromatic domains (Good *et al.*, 2017). Moreover, I could show that Tet1s is recruited to, normally highly methylated heterochromatic domains in murine cells during DNA replication, which resulted in a major increase of 5hmC levels in these domains. This observation led me to speculate on the involvement of Tet1s in an instructive cancer epigenetics model. (**Section 6.3**) This hypothesis could be addressed by disrupting the canonical Tet1 promoter in non-transformed cells, which would result in the use of the alternative promoter and sole

expression of Tet1s. By investigating the resulting transcriptional changes and epigenetic alterations, a first idea of the Tet1s potential in malignant transformation could be gained. Finally, as described earlier, it was striking that the S-phase localization of Tet1s was lost upon fixation. Consequently, no colocalization of Tet1 and PCNA was found, when the localization of endogenous Tet1 was addressed by immunostaining Tet1 in MCF7 breast cancer cells that are known to overexpress the short isoform of Tet1 (**Figure 28C**) (Good *et al.*, 2017). This observation might be attributed to the fact that the short isoform lacks the N-terminal BC- and CxxC-domains, which together contribute to chromatin binding in hypomethylated, euchromatic regions (Zhang and Xu, 2017), thereby increasing kinetics and decreasing fixability. However, a portion of endogenous Tet1 could be fixed and showed focal patterns in immunostainings. Interestingly, these patterns were found to be anti-correlated with DAPI-dense regions and in follow-up experiments, I found a colocalization of the fixable fraction of Tet1 and the splicing factor SC-35/SFRS2 (Splicing factor, arginine/serine-rich 2). In this line, I could moreover show that this colocalization persists throughout different developmental stages, from mouse embryonic stem cells to adult mouse fibroblasts and is susceptible to RNase treatment. These observations are in line with a previous finding in *Drosophila*, where a Tet homologue was shown to oxidise 5mC in RNA and, thereby, positively affected translation efficiency (Delatte *et al.*, 2016). Ectopic overexpression of Tet proteins was furthermore found to increase the levels of 5hmC in RNA from HEK cells (Fu *et al.*, 2014). Taken together, published data and my observations on the nuclear whereabouts of Tet ptoteins, hint to a dual mode of gene regulation, on the DNA but likely also on the RNA level, which both will be addressed further.

7 Bibliography

Agarwal, Noopur, Tanja Hardt, Alessandro Brero, Danny Nowak, Ulrich Rothbauer, Annette Becker, Heinrich Leonhardt, and M. Cristina Cardoso (2007). “MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation”. In: *Nucleic Acids Research* 35.16, pp. 5402–5408.

Allen, Mark D., Charles G. Grummitt, Christine Hilcenko, Sandra Young Min, Louise M. Tonkin, Christopher M. Johnson, Stefan M. Freund, Mark Bycroft, and Alan J. Warren (2006). “Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase”. In: *EMBO Journal* 25.19, pp. 4503–4512.

Angers, Stephane, Ti Li, Xianhua Yi, Michael J MacCoss, Randall T Moon, and Ning Zheng (2006). “Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery.” In: *Nature* 443.7111, pp. 590–3.

Arita, Kyohei, Mariko Ariyoshi, Hidehito Tochio, Yusuke Nakamura, and Masahiro Shirakawa (2008). “Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism”. In: *Nature* 455.7214, pp. 818–821.

Arita, Kyohei, Shin Isogai, Takashi Oda, Motoko Unoki, Kazuya Sugita, Naotaka Sekiyama, Keiko Kuwata, R. Hamamoto, H. Tochio, M. Sato, M. Ariyoshi, and M. Shirakawa (2012). “Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1”. In: *Proceedings of the National Academy of Sciences of the United States of America* 109.32, pp. 12950–12955.

Avvakumov, George V., John R. Walker, Sheng Xue, Yanjun Li, Shili Duan, Christian Bronner, Cheryl H. Arrowsmith, and Sirano Dhe-Paganon (2008). “Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1”. In: *Nature* 455.7214, pp. 822–825.

Bachman, Martin, Santiago Uribe-Lewis, Xiaoping Yang, Michael Williams, Adele Murrell, and Shankar Balasubramanian (2014). “5-Hydroxymethylcytosine is a predominantly stable DNA modification”. en. In: *Nature Chemistry* 6.12, pp. 1049–1055.

Bachman, Martin, Santiago Uribe-Lewis, Xiaoping Yang, Heather E Burgess, Mario Iurlaro, Wolf Reik, Adele Murrell, and Shankar Balasubramanian (2015). “5-Formylcytosine can be a stable DNA modification in mammals”. In: *Nature Chemical Biology* 11.8, pp. 555–557.

Bannister, Andrew J. and Tony Kouzarides (2011). “Regulation of chromatin by histone modifications”. In: *Cell Research* 21.3, pp. 381–395.

Barau, Joan, Aurélie Teissandier, Natasha Zamudio, Stéphanie Roy, Valérie Nalesso, Yann Hérault, Florian Guillou, and Déborah Bourc’his (2016). “The DNA methyltransferase DNMT3C protects male germ cells from transposon activity”. In: *Science* 354.6314, pp. 909–912.

Bashtrykov, Pavel, Gytis Jankevicius, Renata Z Jurkowska, Sergey Ragozin, and Albert Jeltsch (2013). “The Uhrf1 protein stimulates the activity and specificity of the maintenance DNA methyltransferase Dnmt1 by an allosteric mechanism.” In: *The Journal of biological chemistry*, pp. 0–18.

Bauer, Christina, Klaus Göbel, Nagarjuna Nagaraj, Christian Colantuoni, Mengxi Wang, Udo Müller, Elisabeth Kremmer, Andrea Rottach, and Heinrich Leonhardt (2015). “Phosphorylation of TET Proteins Is Regulated via O -GlcNAcylation by the O -Linked N -Acetylglucosamine Transferase (OGT)”. In: *Journal of Biological Chemistry* 290.8, pp. 4801–4812.

Baumbusch, L O, T Thorstensen, V Krauss, A Fischer, K Naumann, R Assalkhou, I Schulz, G Reuter, and R B Aalen (2001). “The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes.” In: *Nucleic acids research* 29.21, pp. 4319–33.

Baylin, Stephen B. and Peter A. Jones (2011). “A decade of exploring the cancer epigenome — biological and translational implications”. In: *Nature Reviews Cancer* 11.10, pp. 726–734.

Becker, Annette, Peng Zhang, Lena Allmann, Daniela Meilinger, Bianca Bertulat, Daniel Eck, Maria Hofstaetter, Giody Bartolomei, Michael O. Hottiger, Valérie Schreiber, Heinrich Leonhardt, and M. Cristina Cardoso (2016). “Poly(ADP-ribosyl)ation of methyl CpG binding domain protein 2 regulates chromatin structure”. In: *Journal of Biological Chemistry* 291.10, pp. 4873–4881.

Berkyurek, Ahmet Can, Isao Suetake, Kyohei Arita, Kohei Takeshita, Atsushi Nakagawa, Masahiro Shirakawa, and Shoji Tajima (2014). “The DNA Methyltransferase Dnmt1 Directly Interacts with the SET and RING Finger-associated (SRA) Domain of the Multifunctional Protein Uhrf1 to Facilitate Accession of the Catalytic Center to Hemi-methylated DNA.” In: *The Journal of biological chemistry* 289.1, pp. 379–86.

Bird, A. (2002). “DNA methylation patterns and epigenetic memory”. In: *Genes & Development* 16.1, pp. 6–21.

Bird, Adrian P. (1986). “CpG-rich islands and the function of DNA methylation”. In: *Nature* 321, pp. 209–213.

Blau, Helen M, Grace K Pavlath, Edna C Hardeman, Choy-Pik Chiu, Laura Silberstein, Steven G Webster, Steven C Miller, and Cecelia Webster (1985). “Plasticity of the differentiated state.” In: *Science (New York, N.Y.)* 230.4727, pp. 758–66.

Blow, J Julian and Anindya Dutta (2005). “Preventing re-replication of chromosomal DNA”. In: *Nature Reviews Molecular Cell Biology* 6.6, pp. 476–486.

Bostick, Magnolia, Jong Kyong Kim, Pierre-olivier Estève, Amander Clark, Sriharsa Pradhan, Steven E. Jacobsen, P.-O. Esteve, Amander Clark, Sriharsa Pradhan, and Steven E. Jacobsen (2007). “UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells”. In: *Science* 317.August, pp. 1760–1764.

Boyarchuk, Ekaterina, Rocío Montes de Oca, and Geneviève Almouzni (2011). “Cell cycle dynamics of histone variants at the centromere, a model for chromosomal landmarks”. In: *Current Opinion in Cell Biology* 23.3, pp. 266–276.

Boyarchuk, Ekaterina, Dan Filipescu, Isabelle Vassias, Sylvain Cantaloube, and Geneviève Almouzni (2014). “The histone variant composition of centromeres is controlled by the pericentric heterochromatin state during the cell cycle.” In: *Journal of cell science* 127.Pt 15, pp. 3347–59.

Brero, Alessandro, Hariharan P. Easwaran, Danny Nowak, Ingrid Grunewald, Thomas Cremer, Heinrich Leonhardt, and M. Cristina Cardoso (2005). “Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation.” In: *The Journal of cell biology* 169.5, pp. 733–43.

Bronner, Christian, Mayada Achour, Yoshimi Arima, Thierry Chataigneau, Hideyuki Saya, and Valérie B Schini-Kerth (2007). “The UHRF family: oncogenes that are drugable targets for cancer therapy in the near future?” In: *Pharmacology & therapeutics* 115.3, pp. 419–34.

Campos, Eric I., James M. Stafford, and Danny Reinberg (2014). “Epigenetic inheritance: Histone bookmarks across generations”. In: *Trends in Cell Biology* 24.11, pp. 664–674.

Capuano, Floriana, Michael Mülleder, Robert Kok, Henk J Blom, and Markus Ralser (2014). “Cytosine DNA Methylation Is Found in Drosophila melanogaster but Absent in Saccharomyces cerevisiae, Schizosaccharomyces pombe , and Other Yeast Species”. In: *Analytical Chemistry* 86.8, pp. 3697–3702.

Cardoso, M C and H Leonhardt (1999). “DNA methyltransferase is actively retained in the cytoplasm during early development.” In: *The Journal of cell biology* 147.1, pp. 25–32.

Cartron, Pierre-François, Arulraj Nadaradjane, Fiona Lepape, Lisenn Lalier, Betty Gardie, and François M Vallette (2013). “Identification of TET1 Partners That Control Its DNA-Demethylating Function.” In: *Genes & cancer* 4.5-6, pp. 235–41.

Casas-Delucchi, Corella S., Alessandro Brero, Hans-Peter Rahn, Irina Solovei, Anton Wutz, Thomas Cremer, Heinrich Leonhardt, and M. Cristina Cardoso (2011). “Histone acetylation controls the inactive X chromosome replication dynamics.” In: *Nature communications* 2, p. 222.

Casas-Delucchi, Corella S., Joke G. van Bemmel, Sebastian Haase, Henry D. Herce, Danny Nowak, Daniela Meilinger, Jeffrey H. Stear, Heinrich Leonhardt, and M. Cristina Cardoso (2012a). “Histone hypoacetylation is required to maintain late replication timing of constitutive heterochromatin.” In: *Nucleic acids research* 40.1, pp. 159–69.

Casas-Delucchi, Corella S., Annette Becker, Janine J. Bolius, and M. Cristina Cardoso (2012b). “Targeted manipulation of heterochromatin rescues MeCP2 Rett mutants and re-establishes higher order chromatin organization”. In: *Nucleic Acids Research* 40.22.

Chagin, Vadim O., Jeffrey H. Stear, and M. Cristina Cardoso (2010). “Organization of DNA replication.” In: *Cold Spring Harbor perspectives in biology* 2.4, a000737.

Chedin, Frederic, Michael R Lieber, and Chih-Lin Hsieh (2002). “The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a.” In: *Proceedings of the National Academy of Sciences of the United States of America* 99.26, pp. 16916–21.

Cheng, Jingdong, Yi Yang, Jian Fang, Jianxiong Xiao, Tingting Zhu, Fei Chen, Ping Wang, Ze Li, Huirong Yang, and Yanhui Xu (2013). “Structural insight into coordinated recognition of trimethylated histone H3 lysine 9 (H3K9me3) by the plant homeodomain (PHD) and tandem tudor domain (TTD) of UHRF1 (ubiquitin-like, containing PHD and RING finger domains, 1) protein.” In: *The Journal of biological chemistry* 288.2, pp. 1329–39.

Citterio, Elisabetta, Roberto Papait, and Francesco Nicassio (2004). “Np95 is a histone-binding protein endowed with ubiquitin ligase activity”. In: *. . . and cellular biology* 24.6, pp. 2526–2535.

Clapier, Cedric R., Janet Iwasa, Bradley R. Cairns, and Craig L. Peterson (2017). “Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes”. In: *Nature Reviews Molecular Cell Biology* 18.7, pp. 407–422.

Costantini, M. and G. Bernardi (2008). “Replication timing, chromosomal bands, and isochores”. In: *Proceedings of the National Academy of Sciences* 105.9, pp. 3433–3437.

Dawlaty, Meelad M., Kibibi Ganz, Benjamin E. Powell, Yueh-Chiang Hu, Styliani Markoulaki, Albert W. Cheng, Qing Gao, Jongpil Kim, Sang-Woon Choi, David C. Page, and Rudolf Jaenisch (2011). “Tet1 Is Dispensable for Maintaining Pluripotency and Its Loss Is Compatible with Embryonic and Postnatal Development”. In: *Cell Stem Cell* 9.2, pp. 166–175.

Dawlaty, Meelad M, Achim Breiling, Thuc Le, M Inmaculada Barrasa, Günter Raddatz, Qing Gao, Benjamin E Powell, Albert W Cheng, Kym F Faull, Frank Lyko, and Rudolf Jaenisch (2014). “Loss of Tet enzymes compromises proper differentiation of embryonic stem cells.” In: *Developmental cell* 29.1, pp. 102–11.

Delatte, Benjamin, Fei Wang, Long Vo Ngoc, Evelyne Collignon, Elise Bonvin, Rachel Deplus, Emilie Calonne, Bouchra Hassabi, Pascale Putmans, Stephan Awe, Collin Wetzel, Judith Kreher, Romuald Soin, Catherine Creppe, Patrick A. Limbach, Cyril Gueydan, Véronique Kruys, Alexander Brehm, Svetlana Minakhina, Matthieu Defrance, Ruth Steward, and François Fuks (2016). “Transcriptome-wide distribution and function of RNA hydroxymethylcytosine”. In: *Science* 351.6270, pp. 282–285.

Easwaran, Hariharan P, Heinrich Leonhardt, and M Cristina Cardoso (2007). “Distribution of DNA replication proteins in Drosophila cells.” In: *BMC cell biology* 8, p. 42.

Easwaran, Hariharan P, Lothar Schermelleh, Heinrich Leonhardt, and M Cristina Cardoso (2004). “Replication-independent chromatin loading of Dnmt1 during G2 and M phases.” In: *EMBO reports* 5.12, pp. 1181–6.

Fang, Jian, Jingdong Cheng, Jiaolong Wang, Qiao Zhang, Mengjie Liu, Rui Gong, Ping Wang, Xiaodan Zhang, Yangyang Feng, Wenxian Lan, Zhou Gong, Chun Tang, Jiemin Wong, Huirong Yang, Chunyang Cao, and Yanhui Xu (2016). “Hemi-methylated DNA opens a closed conformation of UHRF1 to facilitate its histone recognition”. en. In: *Nature Communications* 7, p. 11197.

Ferry, Laure, Alexandra Fournier, Takeshi Tsusaka, Guillaume Adelmant, Tadahiro Shimazu, Shohei Matano, Olivier Kirsh, Rachel Amouroux, Naoshi Dohmae, Takehiro Suzuki, Guillaume J. Filion, Wen Deng, Maud de Dieuleveult, Lauriane Fritsch, Srikanth Kudithipudi, Albert Jeltsch, Heinrich Leonhardt, Petra Hajkova, Jarrod A. Marto, Kyohei Arita, Yoichi Shinkai, and Pierre-Antoine Antoine Defossez (2017). “Methylation of DNA Ligase 1 by G9a/GLP Recruits UHRF1 to Replicating DNA and Regulates DNA Methylation”. In: *Molecular Cell* 67.4, 550–565.e5.

Foster, Benjamin M, Paul Stolz, Christopher B Mulholland, Alex Montoya, Holger Kramer, Sebastian Bultmann, Till Bartke Correspondence, and Till Bartke (2018). “Critical Role of the UBL Domain in Stimulating the E3 Ubiquitin Ligase Activity of UHRF1 toward Chromatin”. In: *Molecular Cell* 72, pp. 1–14.

Frauer, Carina, Andrea Rottach, Daniela Meilinger, Sebastian Bultmann, Karin Fellinger, Stefan Hasenöder, Mengxi Wang, Weihua Qin, Johannes Söding, Fabio Spada, and Heinrich Leonhardt (2011a). “Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1.” In: *PloS one* 6.2, e16627.

Frauer, Carina, Thomas Hoffmann, Sebastian Bultmann, Valentina Casa, M. Cristina Cardoso, Iris Antes, and Heinrich Leonhardt (2011b). “Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain.” In: *PloS one* 6.6, e21306.

Fu, Lijuan, Candace R Guerrero, Na Zhong, Nicholas J Amato, Yunhua Liu, Shuo Liu, Qian Cai, Seung-Gi Jin, Debin Ji, Laura J Niedernhofer, Gerd P Pfeifer, Guo-Liang Xu, Yinsheng Wang, Seung-Gi Jin, Laura J Niedernhofer, Gerd P Pfeifer, Guo-Liang Xu, and Yinsheng Wang (2014). “Tet-mediated Formation of 5-Hydroxymethylcytosine in RNA.” In: *Journal of the American Chemical Society* 136.33, pp. 11582–11585.

Fujimori, Akira, Yoichi Matsuda, Yoshihiro Takemoto, Yasuhiro Hashimoto, Eiko Kubo, Ryoko Araki, Ryutaro Fukumura, Kazuei Mita, Kouichi Tatsumi, and Masahiro Muto (1998). “Cloning and mapping of Np95 gene which encodes a novel nuclear protein associated with cell proliferation”. In: *Mammalian Genome* 9.12, pp. 1032–1035.

Gaudet, François, J Graeme Hodgson, Amir Eden, Laurie Jackson-Grusby, Jessica Dausman, Joe W Gray, Heinrich Leonhardt, and Rudolf Jaenisch (2003). “Induction of tumors in mice by genomic hypomethylation.” In: *Science (New York, N.Y.)* 300.5618, pp. 489–92.

Globisch, Daniel, Martin Münzel, Markus Müller, Stylianos Michalakis, Mirko Wagner, Susanne Koch, Tobias Brückl, Martin Biel, and Thomas Carell (2010). “Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates.” In: *PloS one* 5.12, e15367.

Goll, Mary Grace, Finn Kirpekar, Keith A. Maggert, Jeffrey A. Yoder, Chih Lin Hsieh, Xiaoyu Zhang, Kent G. Golic, Steven E. Jacobsen, and Timothy H. Bestor (2006). “Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2”. In: *Science* 311.5759, pp. 395–398.

Good, Charly R., Jozef Madzo, Bela Patel, Shinji Maegawa, Nora Engel, Jaroslav Jelinek, and Jean-Pierre J. Issa (2017). “A novel isoform of TET1 that lacks a CXXC domain is overexpressed in cancer”. In: *Nucleic Acids Research* 45.14, pp. 8269–8281.

Gu, Tian-Peng, Fan Guo, Hui Yang, Hai-Ping Wu, Gui-Fang Xu, Wei Liu, Zhi-Guo Xie, Linyu Shi, Xinyi He, Seung-gi Jin, Khursheed Iqbal, Yujiang Geno Shi, Zixin Deng, Piroska E Szabó, Gerd P Pfeifer, Jinsong Li, and Guo-Liang Xu (2011). “The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes”. In: *Nature* 477.7366, pp. 606–610.

Guenatri, Mounia, Delphine Bailly, Christèle Maison, and Geneviève Almouzni (2004). “Mouse centric and pericentric satellite repeats form distinct functional heterochromatin”. In: *Journal of Cell Biology* 166.4, pp. 493–505.

Hamidi, Tewfik, Anup Kumar Singh, and Taiping Chen (2015). “Genetic alterations of DNA methylation”. In: 7, pp. 247–265.

Han, Pei, Wei Li, Chiou-Hong Lin, Jin Yang, Ching Shang, Sylvia T. Nurnberg, Kevin Kai Jin, Weihong Xu, Chieh-Yu Lin, Chien-jung Lin, Yiqin Xiong, Huan-Chieh Chien, Bin Zhou, Euan Ashley, Daniel Bernstein, Peng-Sheng Chen, Huei-Sheng Vincent Chen, Thomas Quertermous, and Ching-pin Chang (2014). “A long noncoding RNA protects the heart from pathological hypertrophy”. In: *Nature* 514.7520, pp. 102–106.

Harrison, Joseph S, Evan M Cornett, Dennis Goldfarb, Paul A DaRosa, Zimeng M Li, Feng Yan, Bradley M Dickson, Angela H Guo, Daniel V Cantu, Lilia Kaustov, Peter J Brown, Cheryl H Arrowsmith, Dorothy A Erie, Michael B Major, Rachel E Klevit, Krzysztof Krajewski, Brian Kuhlman, Brian D Strahl, and Scott B Rothbart (2016). “Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1”. In: *eLife* 5, pp. 818–821.

Hashimoto, Hideharu, John R. Horton, Xing Zhang, Magnolia Bostick, Steven E. Jacobsen, and Xiaodong Cheng (2008). “The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix.” In: *Nature* 455.7214, pp. 826–9.

Hashimoto, Hideharu, John R. Horton, Xing Zhang, and Xiaodong Cheng (2009). “UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications”. In: *Epigenetics* 4.1, pp. 8–14.

Hata, Kenichiro, Masaki Okano, Hong Lei, and En Li (2002). “Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice.” In: *Journal of embryology and experimental morphology* 129.8, pp. 1983–1993.

Heinz, Kathrin S., Corella S. Casas-Delucchi, Timea Török, Dusan Cmarko, Alexander Rapp, Ivan Raska, and M. Cristina Cardoso (2018). “Peripheral re-localization of constitutive heterochromatin advances its replication timing and impairs maintenance of silencing marks”. In: *Nucleic Acids Research* 46.12, pp. 6112–6128.

Hergeth, S. P. and R. Schneider (2015). “The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle”. In: *EMBO reports* 16.11, pp. 1439–1453.

Hervouet, Eric, Lisenn Lalier, Emilie Debien, Mathilde Cheray, Audrey Geairon, Hélène Rogniaux, Delphine Loussouarn, and Stéphane Cartron (2010). “Disruption of Dnmt1 / PCNA / UHRF1 Interactions Promotes Tumorigenesis from Human and Mice Glial Cells”. In: *PloS one* 5.6, e11333.

Ho, Steffan N., Henry D. Hunt, Robert M. Horton, Jeffrey K. Pullen, and Larry R. Pease (1989). “Site-directed mutagenesis by overlap extension using the polymerase chain reaction”. In: *Gene* 77.1, pp. 51–59.

Hong, Jang Hee, Eun Heui Jin, Soyeon Kim, Kyu Sang Song, and Jae Kyu Sung (2018). “LINE-1 hypomethylation is inversely correlated with UHRF1 overexpression in gastric cancer”. In: *Oncology Letters* 15.5, pp. 6666–6670.

Hopfner, Raphaël, Marc Mousli, JM Jeltsch, and A Voulgaris (2000). “ICBP90, a novel human CCAAT binding protein, involved in the regulation of topoisomerase IIα expression”. In: *Cancer research*, pp. 121–128.

Hotchkiss, Rollin D (1948). “The Quantitative Separation of Purines, Pyrimidines, and Nucleosides by Paper Chromatography”. In: *J. Biol. Chem.* 175, pp. 315–332.

Hu, Lulu, Ze Li, Jingdong Cheng, Qinhui Rao, Wei Gong, Mengjie Liu, Yujiang Geno Shi, Jiayu Zhu, Ping Wang, and Yanhui Xu (2013). “Crystal Structure of TET2-DNA Complex: Insight into TET-Mediated 5mC Oxidation.” In: *Cell*, pp. 1–11.

Hübner, Michael R. and David L. Spector (2010). “Chromatin Dynamics”. In: *Annual Review of Biophysics* 39.1, pp. 471–489.

Ito, Shinsuke, Li Shen, Qing Dai, Susan C Wu, Leonard B Collins, James a Swenberg, Chuan He, and Yi Zhang (2011). “Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine.” In: *Science (New York, N.Y.)* 333.6047, pp. 1300–3.

Iwata, Atsushi, Yu Nagashima, Lumine Matsumoto, Takahiro Suzuki, Tomoyuki Yamanaka, Hidetoshi Date, Ken Deoka, Nobuyuki Nukina, and Shoji Tsuji (2009). “Intranuclear Degradation of Polyglutamine Aggregates by the Ubiquitin-Proteasome System”. In: *Journal of Biological Chemistry* 284.15, pp. 9796–9803.

Jackson, Sarah and Yue Xiong (2009). “CRL4s: the CUL4-RING E3 ubiquitin ligases”. In: *Trends in Biochemical Sciences* 34.11, pp. 562–570.

Jenuwein, Thomas and C David Allis (2001). “Translating the Histone Code \r10.1126/science.1063127 ”. In: *Science* 293.August, pp. 1074–1081.

Jeong, Jae Yeon, Hyung Soon Yim, Ji Young Ryu, Hyun Sook Lee, Jung Hyun Lee, Dong Seung Seen, and Sung Gyun Kang (2012). “One-step sequence-and ligation-independent cloning as a rapid and versatile cloning method for functional genomics Studies”. In: *Applied and Environmental Microbiology* 78.15, pp. 5440–5443.

Jia, Yuanhui, Pishun Li, Lan Fang, Haijun Zhu, Liangliang Xu, Hao Cheng, Junying Zhang, Fei Li, Yan Feng, Yan Li, Jialun Jiwen Jialun Li, Ruiping Wang, James X Du, Jialun Jiwen Jialun Li, Taiping Chen, Hongbin Ji, Jackie Han, Wenqiang Yu, Qihan Wu, and Jiemin Wong (2016). “Negative regulation of DNMT3A de novo DNA methylation by frequently overexpressed UHRF family proteins as a mechanism for widespread DNA hypomethylation in cancer”. en. In: *Cell Discovery* 2.1, p. 16007.

Jin, Chunlei, Yue Lu, Jaroslav Jelinek, Shoudan Liang, Marcos R.H. H. Estecio, Michelle Craig Barton, and Jean-Pierre J. Issa (2014). “TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells”. In: *Nucleic Acids Research* 42.11, pp. 6956–6971.

Jin, Seung-Gi, Zhi-Min Zhang, Thomas L. L Dunwell, Matthew R. R Harter, Xiwei Wu, Jennifer Johnson, Zheng Li, Jiancheng Liu, Piroska E. E Szabó, Qiang Lu, Guo-Liang Xu, Jikui Song, and Gerd P. P Pfeifer (2016). “Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration”. In: *Cell Reports* 14.3, pp. 493–505.

Joffe, Boris, Heinrich Leonhardt, and Irina Solovei (2010). “Differentiation and large scale spatial organization of the genome”. In: *Current Opinion in Genetics & Development* 20.5, pp. 562–569.

Jones, Peter A. (2012). “Functions of DNA methylation: Islands, start sites, gene bodies and beyond”. In: *Nature Reviews Genetics* 13.7, pp. 484–492.

Jurkowska, Renata Zofia, Tomasz Piotr Jurkowski, and Albert Jeltsch (2011). “Structure and Function of Mammalian DNA Methyltransferases”. In: *ChemBioChem* 12.2, pp. 206–222.

Kalderon, Daniel, Bruce L. Roberts, William D. Richardson, and Alan E. Smith (1984). “A short amino acid sequence able to specify nuclear location”. In: *Cell* 39.3 PART 2, pp. 499–509.

Kaneda, Masahiro, Masaki Okano, Kenichiro Hata, Takashi Sado, Naomi Tsujimoto, En Li, and Hiroyuki Sasaki (2004). “Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting”. In: *Nature* 429.6994, pp. 900–903.

Karagianni, Panagiota, Larbi Amazit, Jun Qin, and Jiemin Wong (2008). “ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation.” In: *Molecular and cellular biology* 28.2, pp. 705–17.

Karlic, R., H.-R. Chung, J. Lasserre, K. Vlahovicek, and M. Vingron (2010). “Histone modification levels are predictive for gene expression”. In: *Proceedings of the National Academy of Sciences* 107.7, pp. 2926–2931.

Keshet, Ilana, Yeshayahu Schlesinger, Shlomit Farkash, Eyal Rand, Merav Hecht, Eran Segal, Eli Pikarski, Richard A. Young, Alain Niveleau, Howard Cedar, and Itamar Simon (2006). “Evidence for an instructive mechanism of de novo methylation in cancer cells”. In: *Nature Genetics* 38.2, pp. 149–153.

Kim, Kyunghwan, Jin-Man Kim, Joong-Sun Kim, Jongkyu Choi, Yong Suk Lee, Nouri Neamati, Jin Sook Song, Kyu Heo, and Woojin An (2013). “VprBP Has Intrinsic Kinase Activity Targeting Histone H2A and Represses Gene Transcription”. In: *Molecular Cell* 52.3, pp. 459–467.

Klutstein, Michael, Deborah Nejman, Razi Greenfield, and Howard Cedar (2016). “DNA Methylation in Cancer and Aging”. In: *Cancer Research* 76.12, pp. 3446–3450.

Ko, Myunggon, Hozefa S. Bandukwala, Jungeun An, Edward D. Lamperti, Elizabeth C. Thompson, Ryan Hastie, Angeliki Tsangaratou, Klaus Rajewsky, Sergei B. Koralov, and Anjana Rao (2011). “Ten-Eleven-Translocation 2 (TET2) negatively regulates homeosta-

sis and differentiation of hematopoietic stem cells in mice." In: *Proceedings of the National Academy of Sciences of the United States of America* 108.35, pp. 14566–71.

Ko, Myunggon, Jungeun An, Hozefa S Bandukwala, Lukas Chavez, Tarmo Äijö, William A Pastor, Matthew F Segal, Huiming Li, Kian Peng Koh, Harri Lähdesmäki, Patrick G Hogan, L Aravind, and Anjana Rao (2013). "Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX". In: *Nature* 497.7447, pp. 122–126.

Kouzarides, Tony (2007). "Chromatin modifications and their function." In: *Cell* 128.4, pp. 693–705.

Kraushaar, Daniel C., Zuozhou Chen, Qingsong Tang, Kairong Cui, Junfang Zhang, and Keji Zhao (2018). "The gene repressor complex NuRD interacts with the histone variant H3.3 at promoters of active genes". In: *Genome Research*.

Kriaucionis, Skirmantas and Nathaniel Heintz (2009). "The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain". In: *Science* 324.5929, pp. 929–930.

Lan, Huiyin, Zaiming Tang, Hongchuan Jin, and Yi Sun (2016). "Neddylation inhibitor MLN4924 suppresses growth and migration of human gastric cancer cells". In: *Scientific Reports* 6, pp. 1–12.

Lande-Diner, Laura, Jianmin Zhang, Ittai Ben-Porath, Ninette Amariglio, Ilana Keshet, Merav Hecht, Veronique Azuara, Amanda G. Fisher, Gideon Rechavi, and Howard Cedar (2007). "Role of DNA methylation in stable gene repression". In: *Journal of Biological Chemistry* 282.16, pp. 12194–12200.

Langemeijer, Saskia M C, Roland P Kuiper, Marieke Berends, Ruth Knops, Mariam G Aslanyan, Marion Massop, Ellen Stevens-Linders, Patricia van Hoogen, Ad Geurts van Kessel, Reinier A P Raymakers, Eveline J Kamping, Gregor E Verhoef, Estelle Verburgh, Anne Hagemeijer, Peter Vandenberghe, Theo de Witte, Bert A van der Reijden, and Joop H Jansen (2009). "Acquired mutations in TET2 are common in myelodysplastic syndromes". In: *Nature Genetics* 41, p. 838.

Le Chalony, Catherine, Françoise Hoffschir, Laurent R. Gauthier, Julia Gross, Denis S. Biard, François D. Boussin, and Vincent Pennaneach (2012). "Partial complementation of a DNA ligase I deficiency by DNA ligase III and its impact on cell survival and telomere stability in mammalian cells". In: *Cellular and Molecular Life Sciences* 69.17, pp. 2933–2949.

Lehnertz, Bernhard, Yoshihide Ueda, Alwin A H A Derijck, Ulrich Braunschweig, Laura Perez-burgos, Stefan Kubicek, Taiping Chen, En Li, Thomas Jenuwein, and Antoine H F M Peters (2003). "Suv39h -Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin". In: 13, pp. 1192–1200.

Leonhardt, Heinrich, Andrea W Page, Heinz-Ulrich Weier, and Timothy H Bestor (1992). "A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei". In: *Cell* 71.5, pp. 865–873.

Leonhardt, Heinrich, Hans-Peter Rahn, Peter Weinzierl, Anje Sporbert, Thomas Cremer, Daniele Zink, and M Cristina Cardoso (2000). "Dynamics of DNA Replication Factories in Living Cells". In: *The Journal of Cell Biology* 149.2, pp. 271–280.

Li, En, Timothy H Bestor, and Rudolf Jaenisch (1992). "Targeted Mutation of the DNA Methyltransferase Gene Results in Embryonic Lethality". In: *Cell* 69, pp. 915–926.

Li, Haitao, Serge Ilin, Wooikoon Wang, Elizabeth M Duncan, Joanna Wysocka, C David Allis, and Dinshaw J Patel (2006). "Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF." In: *Nature* 442.7098, pp. 91–5.

Li, Lili, Chen Li, Haitao Mao, Zhenfang Du, Wai Yee Chan, Paul Murray, Bing Luo, Anthony T.C. Chan, Tony S.K. Mok, Francis K.L. Chan, Richard F. Ambinder, and Qian Tao (2016). "Epigenetic inactivation of the CpG demethylase TET1 as a DNA methylation feedback loop in human cancers". In: *Scientific Reports* 6.January, pp. 1–12.

Li, Tao, Linsheng Wang, Yongming Du, Si Xie, Xi Yang, Fuming Lian, Zhongjun Zhou, and Chengmin Qian (2018). "Structural and mechanistic insights into UHRF1-mediated DNMT1 activation in the maintenance DNA methylation". In: *Nucleic Acids Research* 46.6, pp. 3218–3231.

Li, Zhe, X. Cai, C.-L. Cai, Jiapeng Wang, Wenyong Zhang, Bruce E Petersen, F.-C. Yang, and Mingjiang Xu (2011). "Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies". In: *Blood* 118.17, pp. 4509–4518.

Lindhout, Beatrice I., Paul Fransz, Federico Tessadori, Tobias Meckel, Paul J.J. Hooykaas, and Bert J. van der Zaal (2007). "Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins". In: *Nucleic Acids Research* 35.16.

Liu, Nan, Mengxi Wang, Wen Deng, Christine S Schmidt, Weihua Qin, Heinrich Leonhardt, and Fabio Spada (2013a). "Intrinsic and extrinsic connections of Tet3 dioxygenase with CXXC zinc finger modules." In: *PloS one* 8.5, e62755.

Liu, Xiaoli, Qinqin Gao, Pishun Li, Qian Zhao, Jiqin Zhang, Jiwen Li, Haruhiko Koseki, and Jiemin Wong (2013b). “UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9.” In: *Nature communications* 4.1, p. 1563.

Liu, Yidan, Bin Zhang, Xiaoyu Meng, Matthew J. Korn, Jack M. Parent, Lin-Yu Lu, and Xiaochun Yu (2017). “UHRF2 regulates local 5-methylcytosine and suppresses spontaneous seizures”. In: *Epigenetics* 12.7, pp. 551–560.

Lorsbach, R B, J Moore, S Mathew, S C Raimondi, S T Mukatira, and J R Downing (2003). “TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23)”. In: *Leukemia* 17.3, pp. 637–641.

Lu, Lin-Yu, Henry Kuang, Gautam Korakavi, and Xiaochun Yu (2015). “Topoisomerase II Regulates the Maintenance of DNA Methylation”. In: *Journal of Biological Chemistry* 290.2, pp. 851–860.

Ludwig, Anne K., Peng Zhang, and M. C. Cardoso (2016). “Modifiers and Readers of DNA Modifications and Their Impact on Genome Structure, Expression, and Stability in Disease”. In: *Frontiers in Genetics* 7.June, pp. 1–24.

Ludwig, Anne K., Peng Zhang, Florian D. Hastert, Stephanie Meyer, Cathia Rausch, Henry D. Herce, Udo Müller, Anne Lehmkuhl, Ines Hellmann, Carina Trummer, Christian Storm, Heinrich Leonhardt, and M. Cristina Cardoso (2017). “Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise”. In: *Nucleic Acids Research* 45.5, pp. 2438–2457.

Luger, K and T J Richmond (1998). “DNA binding within the nucleosome core”. In: *Curr Opin Struct Biol* 8.1, pp. 33–40.

Luger, Karolin, Armin W Mäder, Robin K Richmond, David F Sargent, and Timothy J Richmond (1997). “Crystal structure of the nucleosome resolution core particle at 2 . 8 A”. In: *Nature* 389, pp. 251–260.

Lv, Lei, Qi Wang, Yanping Xu, Li-Chung Chung Tsao, Tadashi Nakagawa, Haitao Guo, Lishan Su, and Yue Xiong (2018). “Vpr Targets TET2 for Degradation by CRL4VprBPE3 Ligase to Sustain IL-6 Expression and Enhance HIV-1 Replication”. In: *Molecular Cell* 70.5, 961–970.e5.

Maenohara, Shoji, Motoko Unoki, Hidehiro Toh, Hiroaki Ohishi, Jafar Sharif, Haruhiko Koseki, and Hiroyuki Sasaki (2017). “Role of UHRF1 in de novo DNA methylation in oocytes and maintenance methylation in preimplantation embryos”. In: *PLOS Genetics* 13.10. Ed. by Paula E. Cohen, e1007042.

Maiti, Atanu and Alexander C Drohat (2011). “Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine”. In: *Journal of Biological Chemistry* 286.41, pp. 35334–35338.

Margot, J B, a M Aguirre-Arteta, B V Di Giacco, S Pradhan, R J Roberts, M C Cardoso, and H Leonhardt (2000). “Structure and function of the mouse DNA methyltransferase gene: Dnmt1 shows a tripartite structure.” In: *Journal of molecular biology* 297.2, pp. 293–300.

Marina, Ryan J, David Sturgill, Marc A Bailly, Morgan Thenoz, Garima Varma, Maria F Prigge, Kyster K Nanan, Sanjeev Shukla, Nazmul Haque, and Shalini Oberdoerffer (2015). “TET-catalyzed oxidation of intragenic 5-methylcytosine regulates CTCF-dependent alternative splicing.” In: *The EMBO journal*.

McCall, Chad M, Paula L Miliani de Marval, Paul D Chastain, Sarah C Jackson, Yizhou J He, Yojiro Kotake, Jeanette Gowen Cook, and Yue Xiong (2008). “Human immunodeficiency virus type 1 Vpr-binding protein VprBP, a WD40 protein associated with the DDB1-CUL4 E3 ubiquitin ligase, is essential for DNA replication and embryonic development.” In: *Molecular and cellular biology* 28.18, pp. 5621–33.

Miller, Orlando J, Wolfgang Schnedl, Julian Allen, and Bernard F Erlanger (1974). “5-Methylcytosine localised in mammalian constitutive heterochromatin.” In: *Nature* 251.5476, pp. 636–7.

Moldovan, George Lucian, Boris Pfander, and Stefan Jentsch (2007). “PCNA, the Maestro of the Replication Fork”. In: *Cell* 129.4, pp. 665–679.

Mori, Tsutomu, Yuanyuan Li, Hiroaki Hata, and Hideo Kochi (2004). “NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a PEST-containing nuclear protein”. In: *FEBS Letters* 557.1-3, pp. 209–214.

Mortusewicz, Oliver, Lothar Schermelleh, Joachim Walter, M Cristina Cardoso, and Heinrich Leonhardt (2005). “Recruitment of DNA methyltransferase I to DNA repair sites.” In: *Proceedings of the National Academy of Sciences of the United States of America* 102.25, pp. 8905–9.

Müller, Udo, Christina Bauer, Michael Siegl, Andrea Rottach, and Heinrich Leonhardt (2014). “TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation.” In: *Nucleic acids research* 42.13, pp. 8592–604.

Nakagawa, Tadashi, Lei Lv, Makiko Nakagawa, Yanbao Yu, Chao Yu, Ana C D'Alessio, Keiko Nakayama, Heng-Yu Fan, Xian Chen, Yue Xiong, Ana C. D'Alessio, Keiko Nakayama, Heng-Yu Fan, Xian Chen, and Yue Xiong (2015). "CRL4VprBP E3 Ligase Promotes Monoubiquitylation and Chromatin Binding of TET Dioxygenases". In: *Molecular Cell* 57.2, pp. 247–260.

Nakamura, Kenichi, Yoshifumi Baba, Keisuke Kosumi, and Kazuto Harada (2016). "UHRF1 regulates global DNA hypomethylation and is associated with poor prognosis in esophageal squamous cell carcinoma". In: *Oncotarget* 7.36, pp. 57821–57831.

Neri, Francesco, Danny Incarnato, Anna Krepelova, Stefania Rapelli, Andrea Pagnani, Riccardo Zecchina, Caterina Parlato, and Salvatore Oliviero (2013). "Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells". In: *Genome Biology* 14.8, R91.

Ngo, Thuy T.M., Jejoong Yoo, Qing Dai, Qiucen Zhang, Chuan He, Aleksei Aksimentiev, and Taekjip Ha (2016). "Effects of cytosine modifications on DNA flexibility and nucleosome mechanical stability". In: *Nature Communications* 7, pp. 1–9.

Okano, M, D W Bell, D a Haber, and E Li (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." In: *Cell* 99.3, pp. 247–57.

O'Leary, Valerie Bríd, Sarah Hain, Doris Maugg, Jan Smida, Omid Azimzadeh, Soile Tapio, Saak Victor Ovsepian, and Michael John Atkinson (2017). "Long non-coding RNA PARTICLE bridges histone and DNA methylation". In: *Scientific Reports* 7.1, p. 1790.

Ono, Ryoichi, Tomohiko Taki, and Takeshi Taketani (2002). "LCX , Leukemia-associated Protein with a CXXC Domain , Is Fused to MLL in Acute Myeloid Leukemia with Trilineage Dysplasia Having t (10 ; 11)(q22 ; q23)". In: *Cancer research* 62, pp. 4075–4080.

Ooi, Steen K. T., Chen Qiu, Emily Bernstein, Keqin Li, Da Jia, Zhe Yang, Hediye Erdjument-Bromage, Paul Tempst, Shau-Ping Lin, C. David Allis, Xiaodong Cheng, and Timothy H. Bestor (2007). "DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA". In: *Nature* 448.7154, pp. 714–717.

Palazzo, Alexander F. and Eliza S. Lee (2015). "Non-coding RNA: what is functional and what is junk?" In: *Frontiers in Genetics* 6.JAN, pp. 1–11.

Palumbo, Domenico, Ornella Affinito, Antonella Monticelli, and Sergio Cocozza (2018). "DNA Methylation variability among individuals is related to CpGs cluster density and evolutionary signatures". In: *BMC Genomics* 19.1, pp. 1–9.

Pena, PV, RA Hom, T Hung, H Lin, AJ Kuo, RP Wong, OM Subach, KS Champagne, R Zhao, VV Verkhuscha, G Li, O Gozani, and TG Kutateladze (2009). "Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppresor". In: *J Mol Biol* 380.2, pp. 303–312.

Penn, N W, R Suwalski, C O'Riley, K Bojanowski, and R Yura (1972). "The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid." In: *The Biochemical journal* 126.4, pp. 781–90.

Peters, Antoine H F M, Dó Nal O 'carroll, Harry Scherthan, Karl Mechtler, Stephan Sauer, Christian Schö, Klara Weipoltshammer, Michaela Pagani, Monika Lachner, Alexander Kohlmaier, Susanne Opravil, Michael Doyle, Maria Sibilia, O Carroll, Harry Scherthan, Karl Mechtler, Stephan Sauer, Christian Scho, Klara Weipoltshammer, Michaela Pagani, Monika Lachner, Alexander Kohlmaier, Susanne Opravil, Michael Doyle, Maria Sibilia, and Thomas Jenuwein (2001). "Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability". In: *Cell* 107, pp. 323–337.

Peters, Antoine H F M, Jacqueline E Mermoud, Dónal O'Carroll, Michaela Pagani, Dieter Schweizer, Neil Brockdorff, and Thomas Jenuwein (2002). "Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin." In: *Nature genetics* 30.1, pp. 77–80.

Petryk, Nataliya, Maria Dalby, Alice Wenger, Caroline B. Stromme, Anne Strandsby, Robin Andersson, and Anja Groth (2018). "MCM2 promotes symmetric inheritance of modified histones during DNA replication". In: *Science* 361.6409, pp. 1389–1392.

Pfaffeneder, Toni, Benjamin Hackner, Matthias Truss, Martin Münzel, Markus Müller, Christian a Deiml, Christian Hagemeier, and Thomas Carell (2011). "The discovery of 5-formylcytosine in embryonic stem cell DNA." In: *Angewandte Chemie (International ed. in English)* 50.31, pp. 7008–12.

Pfaffeneder, Toni, Fabio Spada, Mirko Wagner, Caterina Brandmayr, Silvia K Laube, David Eisen, Matthias Truss, Jessica Steinbacher, Benjamin Hackner, Olga Kotljarova, David Schuermann, Stylianos Michalakis, Olesea Kosmatchev, Stefan Schiesser, Barbara Steigenberger, Nada Raddaoui, Gengo Kashiwazaki, Udo Müller, Cornelia G Spruijt, Michiel Vermeulen, Heinrich Leonhardt, Primo Schär, Markus Müller, and Thomas Carell (2014). "Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA." In: *Nature chemical biology* 10.7, pp. 574–81.

Pichler, Garwin, Patricia Wolf, Christine S. Schmidt, Daniela Meilinger, Katrin Schneider, Carina Frauer, Karin Fellinger, Andrea Rottach, and Heinrich Leonhardt (2011). “Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways”. In: *Journal of Cellular Biochemistry* 112.9, pp. 2585–2593.

Qin, Weihua, Heinrich Leonhardt, and Fabio Spada (2011). “Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1”. In: *Journal of Cellular Biochemistry* 112.2, pp. 439–444.

Qin, Weihua, Patricia Wolf, Nan Liu, Stephanie Link, Martha Smets, Federica La Mastra, Ignasi Forné, Garwin Pichler, David Hörl, Karin Fellinger, Fabio Spada, Ian Marc Bonapace, Axel Imhof, Hartmann Harz, and Heinrich Leonhardt (2015). “DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination”. In: *Cell Research* 25.8, pp. 911–929.

Rajakumara, Eerappa, Zhentian Wang, Honghui Ma, Lulu Hu, Hao Chen, Yan Lin, Rui Guo, Feizhen Wu, Haitao Li, Fei Lan, Yujiang Geno Geno Yang Shi, Yanhui Xu, Dinshaw J. J Patel, and Yujiang Geno Geno Yang Shi (2011). “PHD finger recognition of unmodified histone H3R2 links UHRF1 to regulation of euchromatic gene expression.” In: *Molecular cell* 43.2, pp. 275–84.

Rajavelu, Arumugam, Cristiana Lungu, Max Emperle, Michael Dukatz, Alexander Bröhm, Julian Broche, Ines Hanelt, Edris Parsa, Sarah Schiffers, Rahul Karnik, Alexander Meissner, Thomas Carell, Philipp Rathert, Renata Z Jurkowska, and Albert Jeltsch (2018). “Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2”. In: *Nucleic Acids Research* 46.17, pp. 9044–9056.

Ramesh, Vidya, Efil Bayam, Filippo M. Cernilogar, Ian M. Bonapace, Markus Schulze, Markus J. Riemenschneider, Gunnar Schotta, and Magdalena Götz (2016). “Loss of Uhrf1 in neural stem cells leads to activation of retroviral elements and delayed neurodegeneration”. In: *Genes & Development* 30.19, pp. 2199–2212.

Rando, Oliver J. (2007). “Global patterns of histone modifications”. In: *Current Opinion in Genetics and Development* 17.2, pp. 94–99.

Rottach, Andrea, Elisabeth Kremmer, Danny Nowak, Heinrich Leonhardt, M Cristina Cardoso, Cardoso Mc, and M Cristina Cardoso (2008). “Generation and characterization of a rat monoclonal antibody specific for multiple red fluorescent proteins.” In: *Hybridoma (2005)* 27.5, pp. 337–43.

Rottach, Andrea, Carina Frauer, Garwin Pichler, Ian Marc Bonapace, Fabio Spada, and Heinrich Leonhardt (2010). “The multi-domain protein Np95 connects DNA methylation and histone modification.” In: *Nucleic acids research* 38.6, pp. 1796–804.

Rudenko, Andrii, MeeladM Dawlaty, Jinsoo Seo, AlbertW Cheng, Jia Meng, Thuc Le, KymF Faull, Rudolf Jaenisch, and Li Huei Tsai (2013). “Tet1 is critical for neuronal activity-regulated gene expression and memory extinction”. In: *Neuron* 79.6, pp. 1109–1122.

Saksouk, Nehmé, Elisabeth Simboeck, and Jérôme Déjardin (2015). “Constitutive heterochromatin formation and transcription in mammals”. In: *Epigenetics and Chromatin* 8.1, pp. 1–17.

Sanchez-Mut, J. V., H. Heyn, E. Vidal, S. Moran, S. Sayols, R. Delgado-Morales, M. D. Schultz, B. Ansoleaga, P. Garcia-Esparcia, M. Pons-Espinal, M. M. De Lagran, J. Dopazo, A. Rabano, J. Avila, M. Dierssen, I. Lott, I. Ferrer, J. R. Ecker, and M. Esteller (2016). “Human DNA methylomes of neurodegenerative diseases show common epigenomic patterns”. In: *Translational Psychiatry* 6.1, e718–8.

Santiago, Mafalda, Claudia Antunes, Marta Guedes, Nuno Sousa, and C. Joana Marques (2014). “TET enzymes and DNA hydroxymethylation in neural development and function - How critical are they?” In: *Genomics* 104.5, pp. 334–340.

Satyanarayana, A. and P. Kaldis (2009). “Mammalian cell-cycle regulation: several Cdks, numerous cyclins and diverse compensatory mechanisms.” In: *Oncogene* 28.33, pp. 2925–39.

Schaefer, Matthias, Tim Pollex, Katharina Hanna, Francesca Tuorto, Madeleine Meusburger, Mark Helm, and Frank Lyko (2010). “RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage”. In: *Genes & Development* 24.15, pp. 1590–1595.

Serra, Ryan W., Minggang Fang, Sung Mi Park, Lloyd Hutchinson, and Michael R. Green (2014). “A KRAS-directed transcriptional silencing pathway that mediates the CpG island methylator phenotype”. In: *eLife* 2014.3, pp. 1–22.

Sharif, Jafar, Masahiro Muto, Shin-ichiro Takebayashi, Isao Suetake, Akihiro Iwamatsu, Takaho A. Endo, Jun Shinga, Yoko Mizutani-Koseki, Tetsuro Toyoda, Kunihiro Okamura, Shoji Tajima, Kohzoh Mitsuya, Masaki Okano, and Haruhiko Koseki (2007). “The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA.” In: *Nature* 450.7171, pp. 908–12.

Shcherbakova, Daria M., Mikhail Baloban, Alexander V. Emelyanov, Michael Brenowitz, Peng Guo, and Vladislav V. Verkhusha (2016). “Bright monomeric near-infrared fluorescent proteins as tags and biosensors for multiscale imaging”. In: *Nature Communications* 7, p. 12405.

Shen, Li, Azusa Inoue, Jin He, Yuting Liu, Falong Lu, and Yi Zhang (2014). “Tet3 and DNA Replication Mediate Demethylation of Both the Maternal and Paternal Genomes in Mouse Zygotes”. In: *Cell Stem Cell* 15.4, pp. 459–471.

Shi, Feng Tao, Hyeung Kim, Weisi Lu, Quanyuan He, Dan Liu, Margaret A. Goodell, Ma Wan, and Zhou Songyang (2013). “Ten-eleven translocation 1 (Tet1) is regulated by o-linked n-acetylglucosamine transferase (ogt) for target gene repression in mouse embryonic stem cells”. In: *Journal of Biological Chemistry* 288.29, pp. 20776–20784.

Shukla, Sanjeev, Ersen Kavak, Melissa Gregory, Masahiko Imashimizu, Bojan Shutinoski, Mikhail Kashlev, Philipp Oberdoerffer, Rickard Sandberg, and Shalini Oberdoerffer (2011). “CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing.” In: *Nature* 479.7371, pp. 74–9.

Smets, Martha, Stephanie Link, Patricia Wolf, Katrin Schneider, Veronica Solis, Joel Ryan, Daniela Meilinger, Weihua Qin, and Heinrich Leonhardt (2017). “DNMT1 mutations found in HSANIE patients affect interaction with UHRF1 and neuronal differentiation”. In: *Human Molecular Genetics* 26.8, pp. 1522–1534.

Smith, Zachary D. and Alexander Meissner (2013). “DNA methylation: roles in mammalian development.” In: *Nature reviews. Genetics* 14.3, pp. 204–20.

Song, Jikui, Olga Rechkoblit, Timothy H Bestor, and Dinshaw J Patel (2011). “Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation.” In: *Science (New York, N.Y.)* 331.6020, pp. 1036–40.

Song, Jikui, Marianna Teplova, Satoko Ishibe-Murakami, and Dinshaw J. Patel (2012). “Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation”. In: *Science* 335.6069, pp. 709–712.

Spada, Fabio, Andrea Haemmer, David Kuch, Ulrich Rothbauer, Lothar Schermelleh, Elisabeth Kremmer, Thomas Carell, Gernot Längst, and Heinrich Leonhardt (2007). “DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells”. In: *The Journal of Cell Biology* 176.5, pp. 565–571.

Sporbert, Anje, Petra Domaing, Heinrich Leonhardt, and M. Cristina Cardoso (2005). “PCNA acts as a stationary loading platform for transiently interacting Okazaki fragment maturation proteins.” In: *Nucleic acids research* 33.11, pp. 3521–8.

Spruijt, Cornelia G, Felix Gnerlich, Arne H Smits, Toni Pfaffeneder, Pascal W T C Jansen, Christina Bauer, Mirko Wagner, Markus Mu, Fariha Khan, H Christian Eberl, Anneloes Mensinga, Rolf Boelens, Hugo Van Ingen, Heinrich Leonhardt, Arie B Brinkman, Konstantin Lephikov, Udo Mu, Thomas Carell, and Michiel Vermeulen (2013a). “Dynamic Readers for 5- (Hydroxy) Methylcytosine and Its Oxidized Derivatives”. In: *Cell*, pp. 1–14.

Spruijt, Cornelia G., Felix Gnerlich, Arne H. Smits, Toni Pfaffeneder, Pascal W.T.C. Jansen, Christina Bauer, Martin Münzel, Mirko Wagner, Markus Müller, Fariha Khan, H. Christian Eberl, Anneloes Mensinga, Arie B. Brinkman, Konstantin Lephikov, Udo Müller, Jörn Walter, Rolf Boelens, Hugo van Ingen, Heinrich Leonhardt, Thomas Carell, and Michiel Vermeulen (2013b). “Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives”. In: *Cell* 152.5, pp. 1146–1159.

Strahl, Brian D and C David Allis (2000). “The language of covalent histone modifications”. In: *Nature* 403.6765, pp. 41–45.

Struhl, Kevin (2014). “Is DNA methylation of tumour suppressor genes epigenetic?” In: *eLife* 3.3, pp. 1–3.

Tahiliani, Mamta, Kian Peng KP Koh, Yinghua Shen, William a Pastor, Hozefa Bandukwala, Yevgeny Brudno, Suneet Agarwal, Lakshminarayan M Iyer, David R Liu, L Aravind, and Anjana Rao (2009). “Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.” In: *Science* 324.5929, pp. 930–935.

Ting, David T, Doron Lipson, Suchismita Paul, Brian W Brannigan, S. Akhavanfard, Erik J Coffman, Gianmarco Contino, Vikram Deshpande, A. J. Iafrate, Stan Letovsky, Miguel N Rivera, Nabeel Bardeesy, Shyamala Maheswaran, and Daniel A Haber (2011). “Aberrant Overexpression of Satellite Repeats in Pancreatic and Other Epithelial Cancers”. In: *Science* 331.6017, pp. 593–596.

Tsai, Miao-Chih, Ohad Manor, Yue Wan, Nima Mosammaparast, Jordon K Wang, Fei Lan, Yang Shi, Eran Segal, and Howard Y Chang (2010). “Long noncoding RNA as modular scaffold of histone modification complexes.” In: *Science (New York, N.Y.)* 329.5992, pp. 689–93.

Tsukada, Yu-ichi, Tomohiko Akiyama, and Keiichi I. Nakayama (2015). “Maternal TET3 is dispensable for embryonic development but is required for neonatal growth”. In: *Scientific Reports* 5.1, p. 15876.

Ueda, Yoshihide (2006). “Roles for Dnmt3b in mammalian development: a mouse model for the ICF syndrome”. In: *Development* 133.6, pp. 1183–1192.

Unoki, Motoko, Toshihiko Nishidate, and Yusuke Nakamura (2004). “ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain.” In: *Oncogene* 23.46, pp. 7601–10.

Vaughan, Robert M, Bradley M Dickson, Evan M Cornett, Joseph S Harrison, Brian Kuhlman, and Scott B Rothbart (2018). “Comparative biochemical analysis of UHRF proteins reveals molecular mechanisms that uncouple UHRF2 from DNA methylation maintenance”. In: *Nucleic Acids Research* 46.9, pp. 4405–4416.

Vidal, E., S. Sayols, S. Moran, A. Guillaumet-Adkins, M. P. Schroeder, R. Royo, M. Orozco, M Gut, I. Gut, N. Lopez-Bigas, H. Heyn, and M. Esteller (2017). “A DNA methylation map of human cancer at single base-pair resolution”. In: *Oncogene* 36.40, pp. 5648–5657.

Waddington, Conrad H. (1942). “The Epigenotype”. In: *Endeavor* 1.1, pp. 18–20.

Wade, Paul A, Anne Gegonne, Peter L Jones, Esteban Ballestar, Florence Aubry, and Alan P Wolffe (1999). “Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation”. In: 23.september, pp. 1–5.

Wang, Fei, Svetlana Minakhina, Hiep Tran, Neha Changela, Joseph Kramer, and Ruth Steward (2018). “Tet protein function during Drosophila development”. In: *PLOS ONE* 13.1. Ed. by Sriharsa Pradhan, e0190367.

Wang, Yanli, Lewis Y. Geer, Colombe Chappey, Jonathan A. Kans, and Stephen H. Bryant (2000). “Cn3D: Sequence and structure views for Entrez”. In: *Trends in Biochemical Sciences* 25.6, pp. 300–302.

Wang, Yu and Yi Zhang (2014). “Regulation of TET protein stability by calpains.” In: *Cell reports* 6.2, pp. 278–84.

Wang, Zhibin, Chongzhi Zang, Jeffrey a Rosenfeld, Dustin E Schones, Artem Barski, Suresh Cuddapah, Kairong Cui, Tae-young Roh, Weiqun Peng, Michael Q Zhang, and Keji Zhao (2008). “Combinatorial patterns of histone acetylations and methylations in the human genome”. In: *Nature Genetics* 40.7, pp. 897–903.

Watson, James D and Francis HC Crick (1953). “Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid”. In: *Nature* 171.4356, pp. 737–738.

Williams, Kristine, Jesper Christensen, MT Marianne Terndrup Pedersen, Jens V. Johansen, Paul A. C. Cloos, Juri Rappsilber, and Kristian Helin (2011). “TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity”. In: *Nature* 473.7347, pp. 343–348.

Wysocka, Joanna, Tomek Swigut, Hua Xiao, Thomas a Milne, So Yeon Kwon, Joe Landry, Monika Kauer, Alan J Tackett, Brian T Chait, Paul Badenhorst, Carl Wu, and C David Allis (2006). “A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling.” In: *Nature* 442.7098, pp. 86–90.

Xiao-Jie, Lu, Xue Hui-Ying, Xiaolong Qi, Xu Jiang, and Ma Shi-Jie (2016). “LINE-1 in cancer: Multifaceted functions and potential clinical implications”. In: *Genetics in Medicine* 18.5, pp. 431–439.

Xu, Chao, Ke Liu, Ming Lei, Ally Yang, Yanjun Li, Timothy R. Hughes, and Jinrong Min (2018). “DNA Sequence Recognition of Human CXXC Domains and Their Structural Determinants”. In: *Structure* 26.1, 85–95.e3.

Xu, Guo-Liang, Timothy H. Bestor, Déborah Bourc’his, Chih-Lin Hsieh, Niels Tommerup, Merete Bugge, Maj Hulten, Xiaoyan Qu, James J. Russo, and Evani Viegas-Péquignot (1999). “Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene”. In: *Nature* 402.6758, pp. 187–191.

Yang, H, Y Liu, F Bai, J-Y Zhang, S-H Ma, J Liu, Z-D Xu, H-G Zhu, Z-Q Ling, D Ye, K-L Guan, and Y Xiong (2013). “Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation”. In: *Oncogene* 32.5, pp. 663–669.

Yang, Yili, Jirouta Kitagaki, Ren Ming Dai, Che Tsai Yien, Kevin L. Lorick, Robert L. Ludwig, Shervon A. Pierre, Jane P. Jensen, Ilia V. Davydov, Pankaj Oberoi, Chou Chi H Li, John H. Kenten, John A. Beutler, Karen H. Vousden, and Allan M. Weissman (2007). “Inhibitors of ubiquitin-activating enzyme (E1), a new class of potential cancer therapeutics”. In: *Cancer Research* 67.19, pp. 9472–9481.

Yu, Chao, Yin-Li Li Zhang, Wei-Wei Wei Pan, Xiao-Meng Meng Li, Zhong-Wei Wei Wang, Zhao-Jia Jia Ge, Jian-Jie Jie Zhou, Yong Cang, Chao Tong, Qing-Yuan Yuan Sun, and Heng-Yu Yu Fan (2013). “CRL4 complex regulates mammalian oocyte survival and reprogramming by activation of TET proteins.” In: *Science (New York, N.Y.)* 342.6165, pp. 1518–21.

Zhang, Hao-Ying, Jun Xiong, Bao-Ling Qi, Y.-Q. Yu-Qi Feng, and Bi-Feng Yuan (2016a). “The existence of 5-hydroxymethylcytosine and 5-formylcytosine in both DNA and RNA in mammals”. en. In: *Chem. Commun.* 52.4, pp. 737–740.

Zhang, Jiqin, Qinqin Gao, Pishun Li, Xiaoli Liu, Yuanhui Jia, Weicheng Wu, Jiwen Li, Shuo Dong, Haruhiko Koseki, and Jiemin Wong (2011). “S phase-dependent interaction with DNMT1 dictates the role of UHRF1 but not UHRF2 in DNA methylation maintenance.” In: *Cell research* 21.12, pp. 1723–39.

Zhang, Peng, Florian D. Hastert, Anne K. Ludwig, Kai Breitwieser, Maria Hofstätter, and M. Cristina Cardoso (2017a). “DNA base flipping analytical pipeline”. In: *Biology Methods and Protocols* 2.1, pp. 1–12.

Zhang, Peng, Anne K. Ludwig, Florian D. Hastert, Cathia Rausch, Anne Lehmkuhl, Ines Hellmann, Martha Smets, Heinrich Leonhardt, and M. Cristina Cardoso (2017b). “L1 retrotransposition is activated by Ten-eleven-translocation protein 1 and repressed by methyl-CpG binding proteins”. In: *Nucleus* 1034.May, pp. 1–15.

Zhang, Peng, Cathia Rausch, Florian D. Hastert, Boyana Boneva, Alina Filatova, Sujit J. Patil, Ulrike A. Nuber, Yu Gao, Xinyu Zhao, and M. Cristina Cardoso (2017c). “Methyl-CpG binding domain protein 1 regulates localization and activity of Tet1 in a CXXC3 domain-dependent manner”. In: *Nucleic Acids Research* 45.12, pp. 7118–7136.

Zhang, Run Rui, Qing Yan Cui, Kiyohito Murai, Yen Ching Lim, Zachary D. Smith, Shengnan Jin, Peng Ye, Luis Rosa, Yew Kok Lee, Hai Ping Wu, Wei Liu, Zhi Mei Xu, Lu Yang, Yu Qiang Ding, Fuchou Tang, Alexander Meissner, Chunming Ding, Yanhong Shi, and Guo Liang Xu (2013). “Tet1 regulates adult hippocampal neurogenesis and cognition”. In: *Cell Stem Cell* 13.2, pp. 237–245.

Zhang, Wenhao, Weikun Xia, Qiujun Wang, Aaron J. Towers, Jiayu Chen, Rui Gao, Yu Zhang, Chia-an Yen, Ah Young Lee, Yuanyuan Li, Chen Zhou, Kaili Liu, Jing Zhang, Tian-Peng Gu, Xiuqi Chen, Zai Chang, Danny Leung, Shaorong Gao, Yong-hui Jiang, and Wei Xie (2016b). “Isoform Switch of TET1 Regulates DNA Demethylation and Mouse Development”. In: *Molecular Cell* 64.6, pp. 1062–1073.

Zhang, Wu and Jie Xu (2017). “DNA methyltransferases and their roles in tumorigenesis”. In: *Biomarker Research* 5.1, p. 1.

Zhong, Xia, Qian-Qian Wang, Jian-Wei Li, Yu-Mei Zhang, Xiao-Rong An, and Jian Hou (2017). “Ten-Eleven Translocation-2 (Tet2) Is Involved in Myogenic Differentiation of Skeletal Myoblast Cells in Vitro”. In: *Scientific Reports* 7, p. 43539.

Zhou, Ting, Jun Xiong, Mingzhu Wang, Na Yang, Jiemin Wong, Bing Zhu, and Rui-Ming Xu (2014). “Structural basis for hydroxymethylcytosine recognition by the SRA domain of UHRF2.” In: *Molecular cell* 54.5, pp. 879–86.

8 Annex

8.1 Honorary Declaration - Ehrenwörtliche Erklärung

Ehrenwörtliche Erklärung:

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe.

Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht.

Darmstadt, den 28. November 2018

Florian Dieter Hastert

8.2 Own and external contributions

All experiments were conceived and performed by me, except for the following:
Susanne Zimbelmann cloned the Tet1s-K852R plasmid and performed the CoIP in Figure 22 during an internship under my direct supervision. Plasmids encoding EGFP-tagged Tet1-CRD, Tet1-DSBH, Tet1-Δ1-389, Tet1-Δ566-833 and mRFP-tagged VprBP were kind gifts from the AG Leonhardt at the LMU Munich.

8.3 *Curriculum vitae*

Personal information

Name:	Florian Dieter Hastert
Date of birth:	10.06.1986
Place of birth:	Eberbach
Nationality:	German
Current work address:	TU Darmstadt Schnittspahnstr. 10 64287 Darmstadt, Germany

Education

PhD student, TU Darmstadt, Cardoso Group	since 11/2013
MSc Technical Biology, TU Darmstadt	10/2011-09/2013
BSc Biology, TU Darmstadt	10/2008-09/2011
& GSI Helmholtz Center for Heavy Ion Research	
Secondary school, Gymnasium Michelstadt	07/2007
Abitur / high school diploma	

Teaching positions

Instructor in basic and advanced practical courses:

Stem cell biology & epigenetics (MTB101)	2013, 2016, 2017
Methods in molecular cell biology (BB33)	2012, 2014
Methods in cell biology (BB02)	2012, 2013
Animal biodiversity (BB03)	2010

Advisor of undergraduate and graduate theses:

Bachelor's thesis, Susanne Zimbelmann	08/2016-10/2016
Master's thesis, Jasmin Weber	11/2015-05/2016
Bachelor's thesis, Jasmin Weber	06/2014-08/2014

8.4 Publications and Conference contributions

8.4.1 Journal articles

Zhang P, **Hastert FD**, Ludwig AK, Breitwieser K, Hofstätter M, Cardoso MC (2017). "DNA base flipping analytical pipeline". *Biology Methods & Protocols*. Volume 2, Issue 1, 1 January 2017, bpx010. doi: 10.1093/biomethods/bpx010.

Zhang P, Ludwig AK, **Hastert FD**, Rausch C, Lehmkuhl A, Hellmann I, Smets M, Leonhardt H, Cardoso MC (2017). "L1 retrotransposition is activated by Ten-eleven-translocation protein 1 and repressed by methyl-CpG binding proteins". *Nucleus*. 2017 May 19:0. doi: 10.1080/19491034.2017.1330238.

Zhang P, Rausch C, **Hastert FD**, Boneva B, Filatova A, Patil SJ, Nuber UA, Gao Y, Zhao X, Cardoso MC (2017). "Methyl-CpG binding domain protein 1 regulates localization and activity of Tet1 in a CXXC3 domain-dependent manner". *Nucleic Acids Res*. 2017 Apr 25. doi: 10.1093/nar/gkx281.

Ludwig, AK, Zhang P, **Hastert FD**, Meyer S, Rausch C, Herce HD, Müller U, Lehmkuhl A, Hellmann I, Trummer C, Storm C, Leonhardt H, Cardoso MC (2017). "Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise". *Nucleic Acids Res*. 2017 Mar 17;45(5):2438-2457. doi: 10.1093/nar/gkw1197.

Rajan M, Mortusewicz O, Rothbauer U, **Hastert FD**, Schmidthals K, Rapp A, Leonhardt H, Cardoso MC (2015). "Generation of an alpaca-derived nanobody recognizing y-H2AX". *FEBS Open Bio*, 5(1), 779-788. doi: 10.1016/j.fob.2015.09.005.

8.4.2 Manuscripts in preparation

Hastert FD, Ludwig AK, Zimbelmann S, Cardoso MC. "Role and Regulation of Tet1 during DNA replication".

Rausch C, Zhang P, **Hastert FD**, Cardoso MC. "Cytosine base interactions regulate DNA metabolism".

Weichmann F, Hett R, Flatley A, Slama K, Schepers A, **Hastert FD**, Cardoso MC, Dieterich C, Helm M, Feederle R, Meister G. "Generation and validation of monoclonal antibodies specific to modified nucleotides".

8.4.3 Talks and Posters

1st Symposium on Nucleic Acid Modifications

Poster: Methyl Magic - Magicians of Life.

Mainz, Germany, September 4-6, 2017

International Meeting of the German Society for Cell Biology: DGZ 2016

Flash presentation and Poster: A putative role for TET dioxygenases in RNA splicing regulation.

Munich, Germany, March 14-16, 2016

4th Munich Symposium of Chromatin Dynamics

Poster: A putative role for TET dioxygenases in RNA splicing regulation.

Munich, Germany, March 10-12, 2016

Graduate College 1657 - Radiation Biology Scientific Retreat

Short talk and Poster: Dynamic interplay of TET & UHRF proteins and their putative role in cytosine modification maintenance.

Hirschegg, Austria, June 13-18, 2014 - **awarded Poster prize**

International Meeting of the German Society for Cell Biology: DGZ 2014

Poster: Dynamic interplay of TET & UHRF proteins and their putative role in cytosine modification maintenance.

Regensburg, Germany; March 18-21, 2014

8.5 Abbreviations

2i	2 inhibitors (Gsk3β and MEK inhibitors)
2OG	2-oxoglutrate
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5hmU	5-hydroxymethyluracil
5mC	5-methylcytosine
AA	Amino acid
Aid	Activation-induced cytidine deaminase
Alu	Arthrobacter luteus
AML	Acute myeloid leukemia
Apobec	Apolipoprotein B mRNA Editing Enzyme Catalytic Polypeptide-Like
AU	Arbitrary Units
BER	Base excision repair
BP	base pair
C	Cytosine
CC	Chromocenter
CGI	CpG island
CpG	5'—cytosine—phosphate—guanine—3' dinucleotide
CTD	C-terminal domain
Cul4A/B	Cullin-4A/B
CRD	Cystein-rich domain
CTCF	CCCTC-Binding factor
CxxC	CxxC-type zinc finger domain
DAPI	4',6-diamidino-2-phenylindole
DCAF1	DDB1- and CUL4-associated factor 1
DDB1	DNA damage-binding protein 1
Dmap	Dnmt1-associated protein
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
Dnmt	DNA methyltransferase
dNTP	deoxyribonucleoside 5'-triphosphate
DSBH	Double-stranded beta helix
E	Glutamate
EDTA	Ethylenediaminetetraacetic acid

ESC	Embryonic stem cell
FRAP	Fluorescence recovery after photobleaching
Fe	Iron
GBP	GFP binding protein
GFP	Green fluorescent protein
G-phase	Gap-phase
H3K9me3	Tri-methylation of histone 3 at lysine 9
HDAC	Histone deacetylase
HEK	Human embryonic kidney
HRP	Horseradish peroxidase
HSAN1E	Hereditary Sensory and Autonomic Neuropathy Type 1E
ICBP90	Inverted CCAAT box binding protein of 90 kDa
ICF	Immunodeficiency, chromosomal instability, and facial anomalies
Idax	Inhibition of the Dvl and Axin complex
JBP	J-binding protein
K	Lysine
kDa	Kilo Dalton
KO	Knock out
LIF	Leukemia inhibitory factor
LINE1	Long intersperesed nuclear element
MaSat	Major Satellite
MCF7	Michigan Cancer Foundation 7
MCF10a	Michigan Cancer Foundation 10a
Mbd1-6	Methyl-CpG binding domain proteins 1-6
Mecp2	Methyl-CpG binding protein 2
MLL	Mixed lineage leukemia
Neil	Nei-like
NIRF	Np95/ICBP90-like RING finger protein
NuRD	Nucleosome Remodeling Deacetylase
Oct4	Octamer-binding transcription factor 4
Paf15	PCNA associated factor 15
PBD	PCNA binding domain
PBS	Phosphate buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PHC	Pericentric heterochromatin
PHD	Plant homeodomain

PTM	Post-translational modification
mRFP	Monomeric red fluorescent protein
RNA	Ribonucleic acid
RING	Really interesting new gene
RT-qPCR	reverse transcription quantitative polymerase chain reaction
RTT	Rett syndrome
Sin3a	SIN3 transcription regulator family member A
SLIC	Sequence and ligase independent cloning
Sox2	SRY (sex determining region Y)-box 2
SRA	SET and RING finger associated
R	Arginine
Tdg	Thymine DNA glycosylase
Tet	Ten eleven translocation
Tet1CD	Catalytic domain of Tet1
Tet1CDmut	Catalytic domain of Tet1
Tet1s	short isoform of Tet1
TSS	Transcription start site
TTD	Tandem tudor domain
Ubi	Ubiquitin
Ubl	Ubiquitin-like domain
Uhrf1	Ubiquitin-like with plant homeodomain and ring finger domains 1
Uhrf2	Ubiquitin-like with plant homeodomain and ring finger domains 2
VPR	Viral protein R
VPRBP	Viral protein R binding protein
Zbtb	Zinc finger and BTB domain containing

8.6 Acknowledgements

First, I have to thank Prof. Dr. M. Cristina Cardoso. From our very first encounter in the first semester in Darmstadt, Cristina impacted my (scientific) life and should continue to do so in multiple ways. You may not want to admit it, but the way you bring harmony to the lab is unparalleled and refreshing. Besides being the best Doktormutter imaginable, you are definitely the best co-bandmanager one can ask for.

I am very thankful to Prof. Dr. Bodo Laube for being my second corrector and even more, for never questioning my visits to his lab to clear my head by distracting his great people!

I would like to thank Prof. Dr. Beatrix Süß for taking over the role of being an examiner and always being super friendly!

I am deeply indebted to Prof. Dr. Gisela Taucher-Scholz, who accompanied almost my whole academic life and always had a couple of nice words and good advice to share!

Anne & Peng, to put it simply: This was very special!

Anne again, Cathia, Kathrin, Paddy, thanks for the laughs, the walks, the friendship!

My Bachelor-/Masterrettes Jasmin and Susi - I hope, I was half as good a supervisor, as you were students. It was a pleasure to work, think & simply spend time with you!

Anne Lehmkuhl, there is definitely not enough chocolate in the world, to ever thank you for your incredible support in so many ways!

Alex, what is there that you do not know? And thanks for all the Lilien games!

Bianca & Manu, thanks a lot for being something like big (lab)siblings.

Katrin, thank you so much for always dealing with my organizational chaos!

All the other current and past lab members, you made this odyssey something really special!

I want to say a big thank you to everybody in the Leonhardt lab in Munich for materials and expertise!

Thanks to everybody in the BiophoniX for the distraction. I still can not comprehend how big a stupid idea in the Schlossgraben during the Heinerfest can become.

All the people outside the lab, who shared cold beers and open ears, big hearts and big smiles:
Pinky Starfish - Thanks Guys!, Daniel & Bayram - UupsiUupsi!, CS-VVK - you crazy, beautiful people!, Papachristos-Family - Always there!, Tag&Nacht - Crew Love is True Love!

Kiki, thanks for calling me up for a beer one night, in the long term this definitely changed my life and this thesis!

Daniel, I don't know where to start or where to end, but I am forever grateful that you preferred beer over breakfast!

Isi, thanks for always being there when I was in need of a sane person!

I am deeply indebted to the Kielenz family, who welcomed me like one of their own!

Anne, thanks for being my best friend, my partner in crime and a great tortoise mum! I am really looking forward to all the leisure stress and baloney that is still to come!

Tobi & Moritz - I am ever more proud to call you family!

Mama & Papa, ich kann euch nicht oft genug dafür danken, dass ihr seit Tag eins allen Quatsch unterstützt, den ich und auch Tobi und Moritz in unserem Leben machen!

Die Schönheit der Chance, dass wir unser Leben lieben - So spät es auch ist.
Das ist nicht die Sonne die untergeht, sondern die Erde die sich dreht.
-Thees Uhlmann-