Mechanisms of DNA Demethylation

Issam Mayyas

A thesis submitted for the degree of
Doctor of Philosophy

At the

University of Otago
Dunedin, New Zealand

November 2019

Acknowledgment

First and foremost, I would like to express my deep and sincere gratitude to my primary research supervisor Prof. Ian Morison for his patience, immense knowledge and continuous support during my PhD study. Ian generously provided me with his well-needed guidance and helpful advice that made my work a successful endeavour. It was great privilege to work and study under his guidance, and I greatly enjoyed our work collaboration.

I am extremely grateful for my co-supervisor Mr Robert Weeks who was with me step by step, guiding me to complete my project successfully. Rob generously shared his research expertise and taught me the methodology to carry out the experiments. Thank you for guiding me through my research journey.

Special thanks goes to my co-supervisor Prof. Mark Hampton, my advisor Dr. Tim Hore, Prof. Margreet Vissers and Dr. Rob Day for providing and sharing their knowledge and expertise with me.

I am genuinely appreciative of all my lab colleagues Dr. Erin, Dr. Abdul Alsaleh, Jackie, Dr. Karina, Claudia, Luke, Nick, Adam, and Katie for their help, suggestions and moral support during my work.

Last but not the least, my greatest appreciation goes to my lovely wife Suzan and my children Ala'a, Mohammed and Joanne, I would not have been able to complete this project without their love and continuous support.

Declaration

Except where specific reference is made, the work presented in this thesis is the work of the author. The work has not been submitted fully or partially for the award of any other degree or recognition before.

Issam Mayyas

Abstract

While every cell in an organism is genetically identical, there are marked phenotypic differences between tissues and organs that are controlled by epigenetic modifications. These epigenetic modifications provide an important role in controlling the machinery of gene expression. DNA methylation is the most stable epigenetic modification and is necessary for cellular activities. Many cancers show dysregulation of DNA methylation, with significant global loss of methylation. While the phenomenon of DNA methylation is well described in mammals, the mechanisms of reversal of methylation are incompletely understood.

In previous studies, our laboratory observed rapid DNA demethylation in Jurkat cells (T cell leukaemia) after exposure to oxidative stress. This rapid effect indicated active demethylation and suggested an activation of TET (Ten Eleven Translocation) proteins may be occurring. My research investigates the mechanism of changes in DNA methylation using a barcoded hairpin-bisulfite sequencing technique. This high throughput hairpin-bisulfite assay provides a direct assessment of the methylation status of the DNA strands by linking the complementary strands together with a barcoded DNA hairpin linker. We established and developed a bioinformatic workflow to analyse hairpin-bisulfite sequencing data, initially using the Galaxy online platform and later using UNIX command line tools.

We demonstrated hemi-methylation of densely methylated promoter regions, as early as 2 h after DNA replication in cells treated with decitabine, and this hemi-methylation increased with decitabine treatment. With the power of the hairpin-bisulfite sequencing assay, we were able to describe the kinetics and pattern of DNMT1 inhibitor treatment. Even though these results were predicted by previous studies, the kinetics of decitabine-induced hemi-methylation has not been demonstrated before in such detail and clarity. We observed an increase (40%) in unmethylated hairpin reads of the *PCDHGA12* promoter after treatment with ascorbate/decitabine compared to untreated controls. This observed rapid demethylation implicates a novel mechanism of active demethylation that has not yet been recognised by researchers in the field. Importantly, we demonstrated that decitabine is capable of erasing the epigenetic memory of somatic cells by full removal of methyl groups from both complementary DNA strands. This study is the first to report active demethylation in somatic cells treated with decitabine.

Furthermore, in both single-gene and global investigations, we found loss of DNA methylation in different leukaemia cell lines (Molt4, Nalm6 and HL60) following ascorbate and DNMT inhibitor treatment, as well as an increase in 5-hydroxymethyl cytosine in Jurkat cells treated with decitabine and combined ascorbate/decitabine. We performed a low-coverage methylation sequencing assay (PBAT) to detect changes in global DNA methylation. We observed a gradual loss in global DNA methylation in synchronised Jurkat cells treated with decitabine and combined ascorbate/decitabine, confirming the results from locus-specific demethylation. The clinical relevance of our study is that it supports the premise that ascorbate is necessary to enhance the efficacy of decitabine by promoting the function of TET. It is likely that the demethylation pathways that we are studying operate during the onset of cancer and the existence of molecules that alter TET activity may have implications for modification of this process. Cancer patients are often markedly ascorbate (vitamin C) deficient, so the addition of ascorbate to treatment protocols may increase the clinical efficacy (or toxicity) of such drugs in patients with leukaemia.

Table of Contents

List of Figures

List of Tables

List of abbreviations

5-mC	5-methylcytosine
5-hmC	5-hydroxymethylcytosine
5-fC	5-fluorocytosine
5-caC	5-carboxylcytosine
AML	Acute myeloid leukaemia
Bash	Bourne-Again SHell
BER	Base-excision repair
bp	Base pair
C	Cytosine
CpG	Cytosine-phosphate-guanine
dpc	Day post-coitum
dCTP	Deoxyribose cytosine triphosphate
Decitabine	5-aza-2'-deoxycytidine
DMEM	Dulbecco's modified Eagle's medium
DMOG	Dimethyloxalylglycine
DNA	Deoxyribose nucleic acid
DNMT	DNA methyltransferase
EDTA	Ethylenediaminetetraacetic acid
ESC	Embryonic stem cells
FS	Forward scatter
G1	Growth phase 1
G2	Growth phase 2
GC	Guanine-cytosine
gDNA	Genomic DNA
GUI	Graphical user interface
HCC	Hepatocellular carcinoma
iPSCs	induced pluripotent stem cells
MBP	Methyl-binding protein
Me	Methylation
NTD	Neural tube defect
PBAT	Post-bisulfite adapter tagging
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PI	Propidium iodide
PGC	Primordial germ cells
RNA	Ribose nucleic acid
ROS	Reactive oxygen species
S phase	Synthesis phase
SAM	S-adenosyl methionine
SS	Side scatter
T	Thymine
TDG	Thymine DNA glycosylase
TET	Ten Eleven Translocation dioxygenase

Chapter One: Literature review.

1.1 Overview of DNA methylation.

Multicellular organisms start as totipotent zygotes that undergo a series of cell divisions and differentiation events giving arise to a whole organism that contains hundreds of specialised cell types (Mitalipov & Wolf, 2009). The different cell types in a multicellular organism have differences in both structure and function while being genetically identical. This is due to the fact that cell differentiation relies on differences in gene expression rather than on changes in the nucleotide sequence of the cells' genomes. These stable changes in gene expression are called 'epigenetic' (Kohli & Zhang, 2013).

1.1.1 What is epigenetics?

The key to understanding the definition of epigenetics is the prefix "epi" which means "above", "on the top" or "over". The conventional definition of epigenetics is external modifications to DNA that change the phenotype or gene expression without changing the underlying DNA sequence. These modifications have the power to turn genes "on" or "off" and sometimes can be inherited from generation to generation (Lester et al., 2016). To date, the best understood epigenetic mechanisms are DNA methylation and histone modifications.

DNA methylation is the covalent addition of a methyl group to the 5-carbon of the cytosine ring at CpG dinucleotides, resulting in 5-methylcytosine, a stable but reversible epigenetic mark, which is sometimes referred to as the fifth base of DNA (Goll & Bestor, 2005).

Human somatic cell DNA contains approximately 1.5% 5-methylcytosine, which is mostly restricted to palindromic CpG dinucleotides. CpG dinucleotides represent 1% to 4% of genomic mammalian DNA, and are mostly located in repetitive DNA elements (Ehrlich et al., 1982). Cytosine methylation occurs in the context of paired symmetrical CpG sites in which a cytosine nucleotide is followed by guanine. Embryonic stem cells (ESCs) are an exception to this and have high levels of 5-methylcytosine (5-mC) in non CpG context (Holliday & Pugh, 1975; Scourzic et al., 2015).

Methylation of cytosine is critically important during embryogenesis and cell lineage differentiation to maintain the function and stability of the genome as well as providing a defence mechanism against transposons (Jones & Liang, 2009). The majority of gene promoters contain areas of high CpG density referred to as CpG islands. These CpG islands are mainly unmethylated permitting DNA accessibility and gene transcription. Methylation of promoter CpG islands represses gene transcription.

DNA methylation is necessary for normal development. It plays a crucial role in several diverse processes including embryonic development, genome stability and regulation of tissue-specific gene expression, genomic imprinting, X-chromosome inactivation, suppression of repetitive element transcription and transposition, and when dysregulated contributes to diseases such as cancer (Ehrlich et al., 1982; Moore, Toomire, et al., 2013; SanMiguel & Bartolomei, 2018).

1.1.2 Maintenance of DNA methylation.

In mammals, cytosine methylation seems to be the major epigenetic modification of DNA (Chiacchiera et al., 2013; Moore, Toomire, et al., 2013). The establishment of DNA methylation occurs by two catalytic activities, maintenance and *de novo* methylation, via the actions of DNA methyltransferases (DNMTs) which are responsible for catalysing the transfer of methyl group (CH_3) from the donor S-adenosyl-methionine (SAM) to the C-5 position of cytosine (Barros & Offenbacher, 2009) (Figure 1.1).

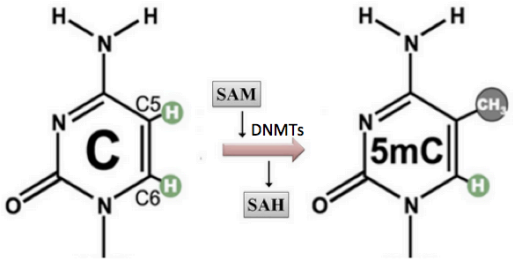


Figure 1. 1. Cytosine methylation by DNMTs.

From the six DNMTs found in mammals, only three of them participate in DNA methylation (Uysal et al., 2017). The *de novo* methylation pathway establishes DNA methylation via the activity of DNMT3A and DNMT3B independent of DNA replication. The maintenance methylation pathway mediated by DNMT1 restores the methylation of newly methylated DNA strands by converting hemi-methylated CpG sites to fully methylated ones following DNA replication (Li et al., 2016; Wu et al., 2012). Three additional methyltransferase enzyme homologues, DNMT2, DNMT3C and DNMT3L, are expressed in many cells including

embryonic stem cells and appear to act either as cofactors for DNA methylation (Goll & Bestor, 2005; Laisne et al., 2018) or as RNA cytosine methyltransferases (Goll et al., 2006).

1.1.3 The role of DNA methylation.

The first use of the term epigenetics is typically attributed to the British embryologist Conrad Waddington in an article published in 1942 (Waddington, 1942). He used the term epigenetics as the name for the study of the interaction between genes and the environment. However, in 1975 DNA methylation was proposed as a silencing epigenetic mark associated with transcriptional repression and maintenance of transcription patterns through cell division (Holliday & Pugh, 1975; Jones, 2012).

DNA modifications provide an important role in controlling the machinery of gene transcription. The role of DNA methylation in transcriptional regulation is widely demonstrated in three biological mechanisms 1) X-inactivation, 2) genomic imprinting and 3) retrotransposon inactivation. During embryonic development, X-inactivation describes the random inactivation of one of the two copies of the X chromosome in female mammals (Riggs, 1975). To maintain inactivation of the selected X chromosome, DNA methylation is required; the inactive X chromosome exhibits global hypomethylation at gene bodies, whereas hypermethylation occurs at many CpG island promoters to silence gene expression (Augui et al., 2011; Hellman & Chess, 2007; Sharp et al., 2011).

Genomic imprinting is a biological mechanism by which certain genes exhibit monoallelic expression that is inherited in a non-Mendelian manner. DNA methylation plays a crucial role in the genomic imprinting. For example, in Dnmt1-depleted mouse embryos, three imprinted genes had abnormal expression from both alleles (Li et al., 1993). Similarly, reduction in Dnmt3a in the female germ cells was found to abrogate genomic imprinting in the progeny (Kaneda et al., 2004).

Retrotransposons are genetic elements that comprise more than 40% of genome and whose expression is suppressed through DNA methylation (Karimi et al., 2011; Rebollo et al., 2011). Retrotransposons, when expressed, are capable of splicing into a new genomic site, which can alter gene expression and induce genomic instability (Groh & Schotta, 2017; Havecker et al., 2004). Alteration of DNA methylation by mutation of DNMTs exhibited upregulation of various retrotransposons, suggesting a role of DNA methylation in the repression of these repetitive elements (Bourc'his & Bestor, 2004; Walsh et al., 1998).

During carcinogenesis, the perturbation of DNA methylation, including genome-wide

hypomethylation and CpG island hypermethylation, is thought to induce activation of transposable elements that might contribute to instability of the genome (Feinberg & Tycko, 2004; Howard et al., 2008; Wild & Flanagan, 2010). Conversely, repression of tumour suppressor genes is associated with hypermethylation of promoter CpG islands that might contribute to the cancer phenotype (Feinberg & Tycko, 2004; Hansen et al., 2011; Miranda & Jones, 2007).

1.1.4 DNA-modifying enzymes: enzymes and co-factors catalysing DNA methylation

Disruption of the DNA methyltransferase gene (Dnmt3a and Dnmt3b) in murine embryonic stem cells, resulted in a decrease in the level of 5mC to about one-third of wild type cells, without any abnormalities of cell morphology or cell growth rate. Interestingly, introducing the mutation into germline cells resulted in recessive lethal phenotype: embryos with the homozygous phenotype showed delayed development and death after mid gestation (Li et al., 1992b; Okano et al., 1999). Furthermore, embryonic lethality and developmental defects were observed with other organisms including *Xenopus laevis*, *Arabidopsis thaliana* and zebrafish (Mathieu et al., 2007; Rai et al., 2010; Stancheva et al., 2002; Stancheva & Meehan, 2000).

1.1.4.1 DNMT1 and UHRF1 are key-components of maintenance methylation.

While the *de novo* methyltransferases, DNMT3a and DNMT3b are necessary for establishment of DNA methylation, the maintenance methyltransferase, DNMT1 has a major role in the preservation of this methylation during DNA replication. DNMT1 preferentially recognises hemi-methylated CpG sites, and converts hemi-methylated CpG sites to fully methylated sites (Santiago et al., 2014). Targeting of DNMT1 to the replicating DNA requires an accessory protein UHRF1 (also known nuclear protein of 95 kDa (NP95) in mouse or inverted CCAAT box-binding protein of 90 kDa (ICBP90) in human) that recognizes and binds to hemi-methylated DNA sites through its unique SRA domain (Sharif et al., 2007; Zeng & Chen, 2019). UHRF1 contained five conserved domains which are all implicated in the function of DNMT1, the RING-associated SRA domain plays an important role in loading DNMT1 onto the newly synthesised DNA strand (Bostick et al., 2007; Sharif et al., 2007).

DNMT1 was the first mammalian DNA methyltransferase gene to be discovered (Bestor et al., 1988) and is characterised by several transcription start sites and three isoforms (Mertineit et al., 1998; Rouleau et al., 1992). The first isoform is *Dnmt1s*, which is expressed in somatic cells and encodes the full length 1620 amino acid DNMT1 protein product. The second isoform, *Dnmt1o*, lacks the N-terminal 118 amino acids and is more stable and exclusively expressed in oocytes (Ding & Chaillet, 2002), whereas the third isoform *Dnmt1p,* is expressed in pachytene

spermatocytes, but has no protein product (Zeng & Chen, 2019).

1.1.4.2 DNMT3A and DNMT3B are key-components of de novo methylation machinery.

In mouse ES cells, genetic disruption of *Dnmt1* resulted in global loss of methylation (Lei et al., 1996; Li et al., 1992a). Analysis of Dnmt1 knockout mouse embryonic stem cells led to the discovery of the existence of *de novo* DNA methyltransferase activity, and identification of DNMT3A and DNMT3B by consecutive screening of expressed sequence libraries for DNMT homologues. Studies of their function have shown that both DNMT3A and DNMT3B are required for *de novo* methylation and mouse development, and they play a critical role in maintenance methylation as demonstrated by the fact that DNMT3A and DNMT3B knockout in mouse embryonic stem cells leads to gradual reduction in global DNA methylation (Chen et al., 2003; Lei et al., 1996; Okano et al., 1999; Okano et al., 1998).

New global DNA methylation is established by *de novo* DNA methyltransferases DNMT3A and DNMT3B. These enzymes interact together to form homo- and heterodimers which methylate both unmethylated and hemi-methylated CpG sites (Cheng & Blumenthal, 2008). Although, DNMT3A and DNMT3B enzymes are essential during embryonic development and their *de novo* methylation can be faithfully spread through mitotic division, their activity can be modulated by a catalytically inactive family member DNMT3L, which is involved in the regulation and targeting of DNMT3A/B for *de novo* methylation (Goll & Bestor, 2005; Hata et al., 2002).

Reduction of either Dnmt1 or both Dnmt3a and Dnmt3b in mouse ES cells resulted in hypomethylation, which led to differentiation defects and cell apoptosis, confirming that DNA methylation is necessary during differentiation. Restoration of DNA methylation levels rescued these defects (Damelin & Bestor, 2007; Jackson et al., 2004). Mutational changes in human *DNMTs* are correlated with pathological conditions (Cui & Xu, 2018; Jin & Liu, 2018). For example, mutations in DNMT1 have been identified in a neurodegenerative disorder associated with dementia and hearing loss (Klein et al., 2011). In addition, ICF (immunodeficiency centromere instability and facial abnormalities) syndrome is a rare autosomal recessive disorder that is caused by mutations in DNMT3B. ICF patients exhibit severe hypomethylation of pericentric repeats which is thought to be the cause of genomic abnormalities (Goll & Bestor, 2005; Xu et al., 1999). Mutations in DNMT3A have been identified in some leukaemias and are associated with poor prognosis (Ley et al., 2010; Yan et al., 2011).

1.1.5 Epigenetics programming and reprogramming during development.

1.1.5.1 Methylation dynamics in pre-implantation embryos.

DNA methylation is largely established during gametogenesis, and epigenetic modifications are important for programming lineage determination and cellular identity. In mammals, the genomes of mature male and female gametes are highly methylated comparable to somatic cells. After fertilisation, global methylation reprogramming in the male pronucleus occurs by active demethylation before DNA replication (Mayer et al., 2000; Wu & Zhang, 2014; Wu & Zhang, 2010; Wu & Zhang, 2017) (Figure 1.2). TET3 is responsible for mediating the active oxidation of 5-mC followed by passive dilution, resulting in restoration of unmodified cytosine (Gu et al., 2011; Wossidlo et al., 2011). A second wave of demethylation occurs at the two-cell stage, where the maternal-derived genome undergoes a gradual passive demethylation up to the blastocyst stage when the methylation level reaches the lowest point. After implantation, a wave of *de novo* methylation mediated by DNMT3A and DNMT3B, re-establishes the embryonic methylation pattern (Bochtler et al., 2017; Zeng & Chen, 2019).

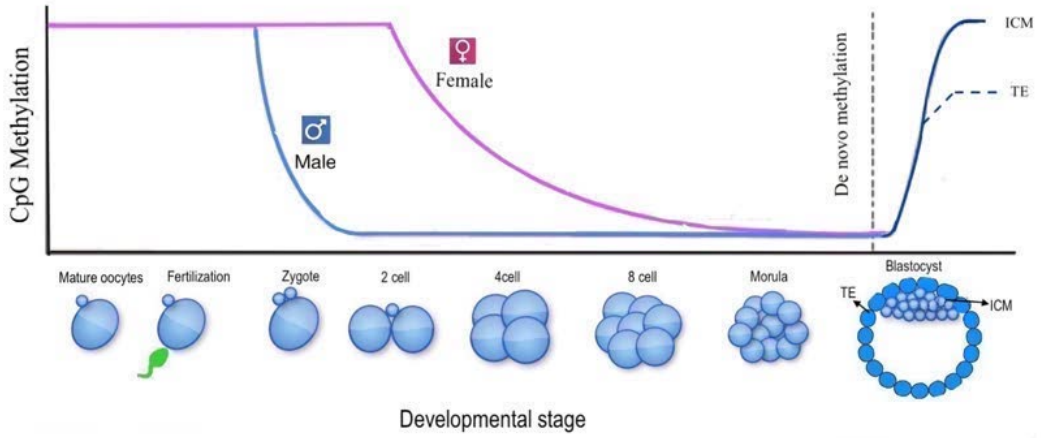


Figure 1. 2. Dynamic changes of DNA methylation during development. After fertilisation the methylation patterns for the paternal-derived genomes are erased by active demethylation and the methylation patterns of maternal-derived genome are erased by passive demethylation. TE: tropho ectoderm, ICM: inner cell mass.

1.1.5.2 Methylation dynamics in primordial germ cells (PGCs).

The primordial germ cells (PGCs) are the primary undifferentiated stem cells that will differentiate towards the gametes (spermatozoa or oocytes). The erasure of DNA methylation in primordial germ cells is the critical step in generating the totipotent epigenome in the germ line. The DNA methylation loss in primordial germ cells occurs over three stages. The first

stage shows loss of bulk DNA methylation in a TET-independent manner between mouse embryonic day 8 and 9.5. The second stage shows TET1-and also potentially TET2-mediated loss of DNA methylation via oxidation of 5-mC to 5-hmC between day 10 and 11.5, and the third stage is loss of 5-hmC through replication dependent passive dilution (Hackett et al., 2013; Kagiwada et al., 2013).

1.1.6 DNA demethylation.

DNA methylation has a profound impact on genome stability, transcription and development in somatic cells. Although DNA methylation is relatively stable and can be maintained during DNA replication (Seisenberger et al., 2013; Wu & Zhang, 2010), loss of DNA methylation has been observed in different biological contexts, either through passive or active mechanisms.

1.1.6.1 Passive DNA demethylation.

During DNA replication, the maintenance DNA methyltransferase DNMT1 and its accessory factor UHRF1 maintain DNA methylation patterns across cell division. During the S-phase, the semiconservative replication of double stranded DNA leads to formation of two daughter strands. Intriguingly, functional deficiency in maintenance methylation may lead to replication-dependent dilution of 5mC. Passive DNA demethylation occurs as a result of the loss of maintenance methyltransferase activity (Figure 1.3), through a lack of DNMT1 or loss of accessory factor UHRF1 which targets DNMT1 to the hemi-methylated DNA at the replication fork.

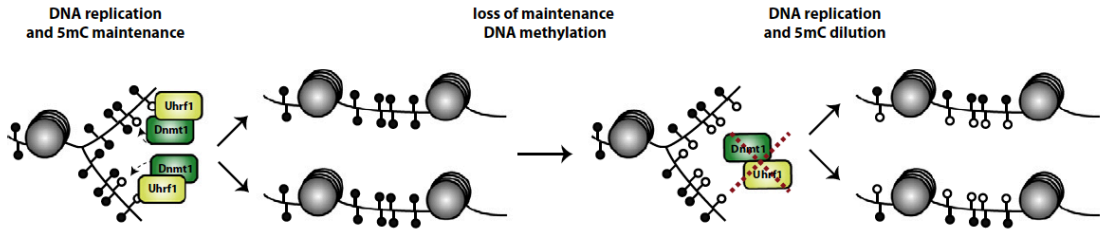


Figure 1. 3. Passive demethylation via lack of methylation maintenance during DNA replication (Hill et al. 2014 with permission).

This mechanism is fully documented in the preimplantation embryo. The erasure of DNA methylation in maternally derived genomes during cleavage division is mainly via passive demethylation due to exclusion of DNMT1 from the nuclei (Messerschmidt et al., 2014; Smallwood & Kelsey, 2012). In primordial germ cells, the first phase of global loss of methylation is passive demethylation due to silencing of the key methylation enzymes including

UHRF1 (Kagiwada et al., 2013).

Importantly, alteration of cytosine methylation patterns is usually observed in cancer cells, and may result in activation of transforming genes and silencing of tumour suppressor genes. The presence of the oxidized form 5-hydroxymethylcytosine (5-hmC) (generated from oxidation of 5-mC to 5-hmC) on the hemi-methylated parent strand reduces the DNMT1 binding affinity of the cytosine in the nascent strand *in vitro* (Valinluck & Sowers, 2007), potentially leading to passive demethylation (Figure 1.5,C).

1.1.6.2 Active DNA demethylation.

The rapid loss of DNA methylation in a replication-independent manner in slow or non-dividing cells is explained by active DNA demethylation (Figure 1.4) (Figure. 1.5.C, D). Interestingly, the discovery of rapid genome-wide loss of 5-mC of the paternal genome in the mouse zygote, which cannot be explained by replication-dependent dilution, gave support to the presence of active DNA demethylation (Gu et al., 2011; Mayer et al., 2000; Wu & Zhang, 2017). Active demethylation for a locus-specific region is well documented in post-mitotic cells in the adult brain (Guo et al., 2011b) and also during cell fate changes (Pastor et al., 2013).

1.1.7 Mechanism of TET-mediated active demethylation.

Due to the resilience of carbon-carbon bonds that link the methyl group in 5-mC, methylation was originally considered as an irreversible modification. Demethylation, by breaking this bond, was regarded as infeasible, (Kagiwada et al., 2013; Ohno et al., 2013). One step conversion of 5-methylcytosine to cytosine is unknown, and it is unlikely because of the difficultly in breaking the C-C bonds. It is now clear that DNA methylation can not only be removed by passive dilution, but also can be actively erased by enzymatic activity (Wu & Zhang, 2010).

Active DNA demethylation uses enzymatic modification of 5-mC to generate 5-hmC; this process results in indirect removal of 5-methylcytosine. Ten-eleven translocation 1,2 and 3 (TET1, TET2, TET3) proteins are the key component of the active DNA demethylation machinery (Rasmussen & Helin, 2016). The three enzymes are able to initiate the first step of active removal of 5-mC by oxidizing 5-methylcytosine into 5-hydroxymethylcytosine. This base (5-hmC) can be further processed to unmodified cytosine in two ways: firstly, by passive demethylation through DNA replication (Figure 1.5.C), or secondly, by further iterative enzymatic modifications, supporting the hypothesis that 5-hmC can be an intermediate of either passive or active DNA demethylation (Figure 1.5.D). Furthermore, iterative oxidation by TET

proteins will generate 5-formylcytosine (5-fC) or 5-carboxylcytosine (5-caC), which can be recognized and excised from DNA by thymine DNA glycosylase (TDG) and base excision repair (BER) to regenerate unmodified cytosine (He et al., 2011; Ito et al., 2011; Maiti & Drohat, 2011; Shen et al., 2013; Wu & Zhang, 2017; Zhang et al., 2012) (Figure 1.4).

All TET proteins are similar structures and are named after the ten-eleven translocation that occurs in rare cases of acute myeloid and lymphoblastic leukaemia (Seisenberger et al., 2013; Wu & Zhang, 2014). This translocation fuses the mixed-lineage leukaemia 1 (MLL1) gene located on chromosome 10 to the TET1 gene on chromosome 11 (Pastor et al., 2013).

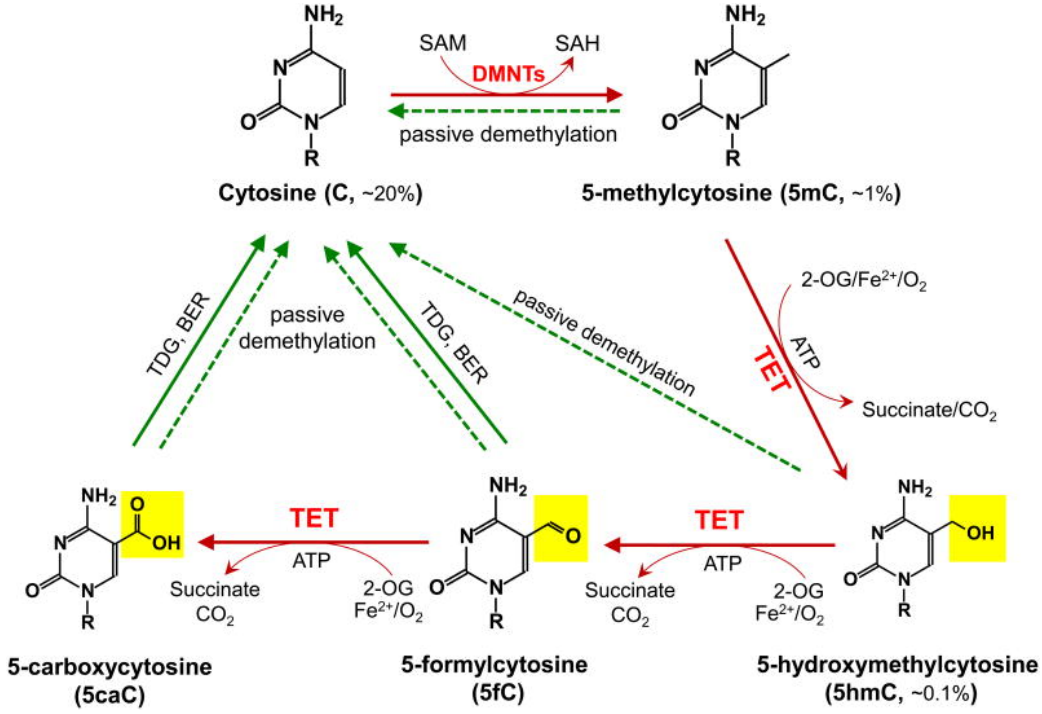


Figure 1. 4. Schematic of DNA methylation and demethylation pathways in mammals. DNA methylation occurs by transferring the methyl group from the donor SAM to cytosine by DNMTs enzymes. TET enzymes oxidise 5-mC to 5-hmC, 5-fC and 5-caC, followed by TDG mediated BER, Reproduced from (Huang & Rao, 2014) with permission.

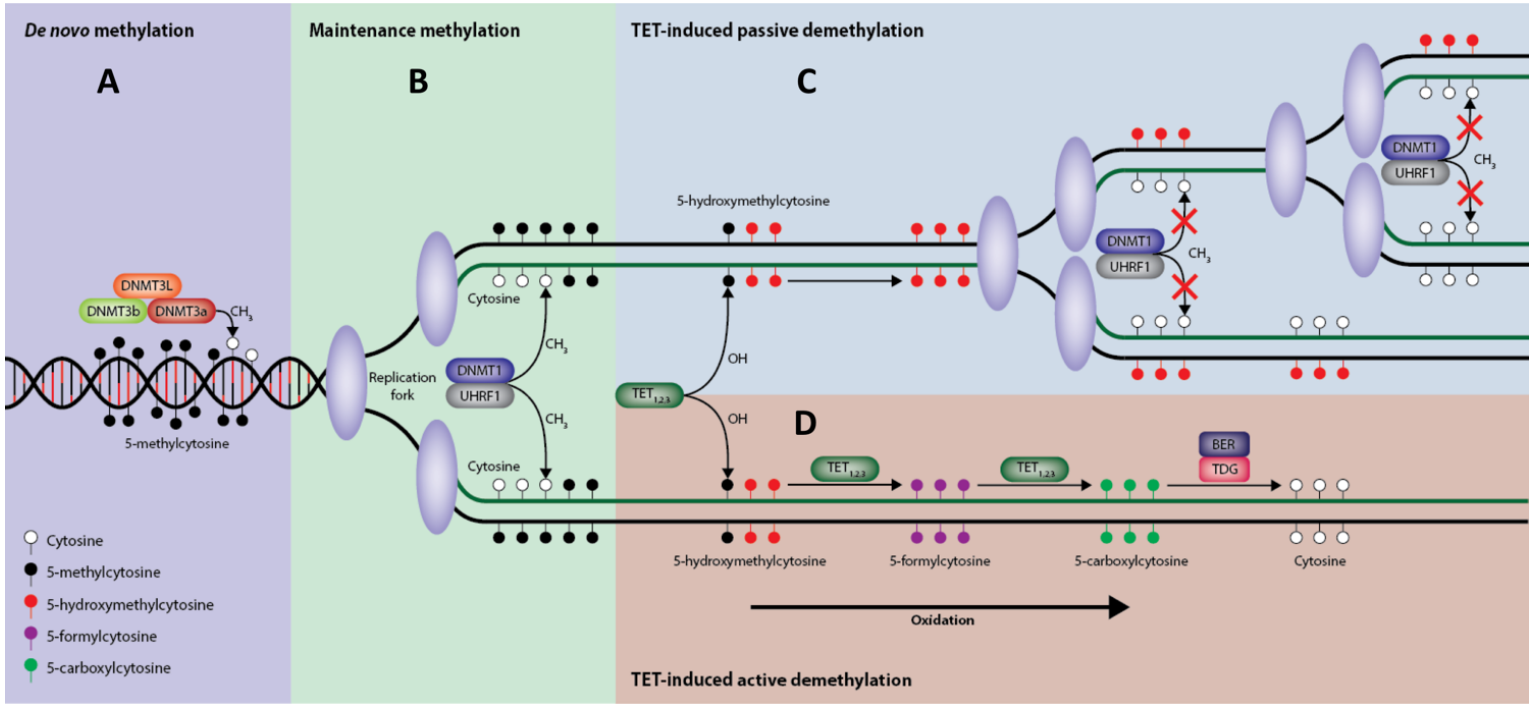


Figure 1. 5. Mechanisms of DNA methylation and DNA demethylation. A. During development, *de novo* methylation of unmethylated cytosine by DNMT3a,b proteins (left purple box). B. During cell division, maintenance DNMT1 specifically recognises hemi-methylated DNA through the interaction with UHRF1 at the replication fork and restore the methylation pattern of the newly synthesised DNA strands (middle green box). C. TET induced passive demethylation through oxidation of 5-mC to generate 5-hmC, during cell cleavage DNMT1 is unable to recognise hemi-hydroxymethylated strands, the presence of 5-hmC on the hemi-methylated DNA reduces the binding affinity of DNMT1 with methylated cytosine resulting in a passive demethylation after successive rounds of cell division. D. TET induced active demethylation, TET converts 5-mC to 5-hmC. Through further oxidations, TET proteins generates two additional modified cytosines 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC). Thymine DNA glycosylase (TDG) can efficiently recognise and excise 5-fC and 5-caC and replace them with unmodified cytosine by base excision repair (BER).

1.1.7.1 The discovery of the TET-TDG pathway.

The discovery of 5-hmC in mammalian DNA revealed a new potential mechanism for 5-mC displacement (He et al., 2011; Wu & Zhang, 2014). The three forms of TET proteins, namely TET1, TET2 and TET3 mediate multiple pathways of DNA demethylation including converting 5-methylcytosine (5-mC) to 5-hydroxymethylcytosine (5-hmC), and further, generate two additional higher oxidative products 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC), followed by thymine DNA glycosylase (TDG)- mediating base excision repair (BER). The oxidized forms 5-hmC, 5-fC and 5-caC can be diluted through successive rounds of replication dependent passive demethylation caused by the low enzymatic activity of DNMT1 on the hemi-hydroxymethylated DNA (Kohli & Zhang, 2013; Pastor et al., 2013; Wu & Zhang, 2010).

The iterative oxidative reactions by TET enzymes are analogous to a key step of thymine oxidation by thymine hydroxylase (THase) identified in yeast (Ito et al., 2011; Smiley et al., 2005). In this case, THase catalyses serial oxidation reactions to convert thymine to 5-hydroxymethyluracil (5-hmU), 5-formyluracil (5-fU) and further to iso-orotate which then can be converted by iso-orotate decarboxylase to uracil, which then can be used for nucleotide metabolism (Smiley et al., 2005). Similar to iso-orotate, 5-caC might be converted to cytosine with removal of the carboxyl group by a decarboxylation reaction, but the process has not been identified yet in mammals (Ito et al., 2011). Furthermore, Iwan and his colleages 2018 showed that 5-fC and f-caC undergo a direct C-C bond cleavage in stem cells, which leads to direct re-installation of cytosine (Iwan et al., 2018).

1.1.7.2 Nucleotide replacement is the second step of active demethylation.

Importantly, thymine DNA glycosylase (TDG) efficiently recognises and excises 5-fC and 5-caC in mouse embryonic stem cells (Rose et al.) (Gong & Zhu, 2011; He et al., 2011; Raiber et al., 2012; Zhang et al., 2012). Interestingly, TDG has high affinity for 5-fC and 5-caC, but not for 5-mC and 5-hmC (Hashimoto et al., 2012; He et al., 2011; Zhang et al., 2012). The activity of TDG in active demethylation occurs via the removal of 5-fC and 5-caC by flipping the target nucleotide from the double-strand DNA. TDG then cleaves the N-glycosidic bond, leaving an abasic site, thus activating the base excision repair pathway (BER), resulting in insertion of an unmethylated cytosine. The TDG-BER mechanism has been proposed as an active demethylation machinery to remove the oxidised forms 5-fC and 5-caC, and replacement with unmodified cytosine (Hashimoto et al., 2012; He et al., 2011; Zhang et al., 2012).

Interestingly, a recent study by Fu et al. (2019) demonstrated the mechanism of how TDG recognize and distinguish highly similar cytosine modifications for potential demethylation. Using a high-resolution structure of a dsDNA decamer containing fully symmetric 5-fC and 5-caC nucleotides, the study reported that the modified cytosines induced geometrical alteration of the DNA minor groove, which was recognized by finger residue, R275, of TDG (Fu et al., 2019).

1.1.7.3 Structure and function of TET enzymes.

TET enzymes belong to the group of α–ketoglutarate (α–KG) and Fe^{+2} dependent dioxygenases (Lorsbach et al., 2003; Tahiliani et al., 2009). These proteins provided a mechanistic basis for the active DNA demethylation pathway. TET proteins were originally identified in the context of haematologic malignancies, with TET1 discovered as a fusion partner in a rare case of acute myeloid leukaemia bearing the t(10;11)(q22;q23) (Lorsbach et al., 2003; Ono et al., 2002). This conserved gene family contains two other members (TET2 at 4q24 and TET3 at 2p12) in humans (Ko et al., 2015; Lorsbach et al., 2003).

Tahiliani et al. (2009) provided the first bona fide evidence of active DNA demethylation by demonstrating that TET1 enzymatic activity involved the oxidation of 5-mC to 5-hmC. Through bioinformatics approaches, TET enzymes were identified as human homologues of mouse Tet proteins (including Tet1, Tet2 and Tet3) and homologues of the J binding proteins 1 and 2 (JBP1 and JBP2) (in trypanosome, a unicellular protozoon), which are also able to oxidize 5-mC to 5-hmC.

Other studies confirmed that TET proteins are Fe^{+2} and α-KG-dependent dioxygenases, and the binding motif of Fe^{+2} in TET enzymes includes at least two histidine residues and one asparagine residue, which are conserved across JBP1, JBP2, human TET, and mouse Tet proteins (Shen & Zhang, 2013; Tahiliani et al., 2009). Interestingly, mutation of the Fe^{+2} binding sites in TET enzymes reduced the 5-hmC signal in cultured cells (Ito et al., 2010).

1.1.7.4 Roles of TET enzymes in regulation of DNA methylation.

TET proteins are large (~180 to 230 kDa) multidomain dioxygenase enzymes. All TET enzymes contain a conserved core catalytic domain (Cys-rich, double strand beta helix (DSBH) regions, and binding sites for the cofactors Fe^{+2} and α-ketoglutaric acid that together form the core catalytic region in the C terminal) that uses O_2 to decarboxylate α-ketoglutaric acid generating a high valent iron oxide that converts 5-mC to 5-hmC (as well as 5-hmC to 5-fC and 5-caC) (Kohli & Zhang, 2013; Tahiliani et al., 2009). However, the double strand β helix

(DSBH) fold is different for each member of α-KG – Fe^{+2} dioxygenases. In TET proteins, the DSBH domain was shown to include a conserved residue involved in the regulation of the cofactor Fe^{+2} and the co-substrate α-KG. Moreover, a cysteine-rich region identified upstream of the DSBH domain contains eight cysteine and one histidine residues proposed to be used for DNA binding activity. The core catalytical domain preferentially binds cytosine in a CpG context but does not interact with surrounding DNA bases (Tahiliani et al., 2009).

TET1 and TET3 have an N-terminal conserved CXXC zinc finger domain, though TET2 does not (Figure 1.6). While similar CXXC domains usually bind to unmethylated CpG dinucleotides, in TET enzymes they bind to 5-mC and 5-hmC (Xu, Wu, et al., 2011). Intriguingly, eleven CXXC family members have been classified and found in human: TET1 protein contains a CXXC domain type 6, whereas TET3 contains CXXC10 (Good et al., 2017; Liu, Wang, et al., 2013). The exact function of the TET CXXC domain is not fully understood, but it is most likely involved in targeting the enzyme to specific regions of the genome.

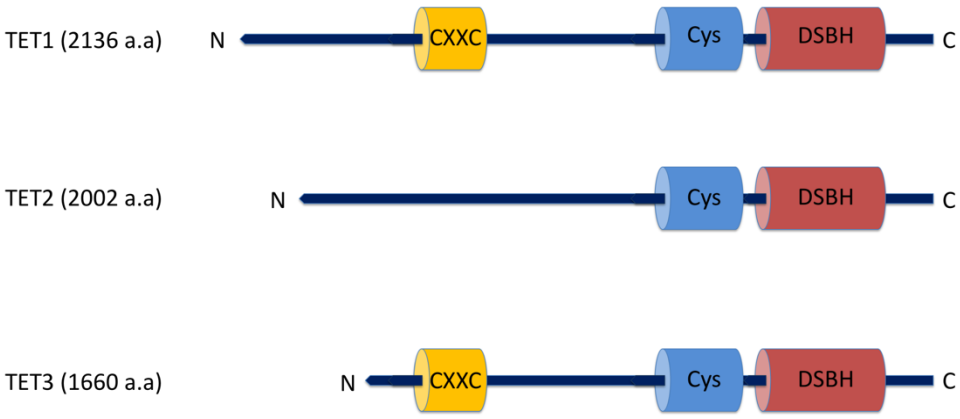


Figure 1.6. Human TET protein domain structure. TET1-3 harbour a cysteine-rich region (Cys) followed by the double-stranded β-helix (DSBH) fold characteristic of the α-KG-Fe (II) oxygenase. TET1 and TET3 also contain a CXXC domain.

1.1.7.4.1 TET1.

The discovery of 5-hmC, and the presence of high 5-hmC at the Tet1 DNA binding loci in ESC, suggesting the role of Tet1 in locus-specific 5-hmC regulation led to more studies to identify the various roles of TET1 protein (Xu, Wu, et al., 2011).

In mouse embryonic stem cells, Tet1 plays a dual function in gene regulation. Inhibition of Tet1 induced both up-regulation of some target genes and down-regulation of others. Interestingly, the percentage of Tet1 up-regulated genes was much higher than the percentage of down-regulated genes, suggesting that Tet1 might have a stronger gene repressive function (Dawlaty

et al., 2011; Wu, D'Alessio, Ito, Xia, et al., 2011). Furthermore, inhibition of Tet1 only affected Tet1 target genes without any effect on non-Tet1 binding genes, suggesting Tet1 regulation of gene-specific region (Xu, Wu, et al., 2011).

In general, Tet1 prefers to bind to exon regions rather than introns, and has increased affinity for high CpG promoters than intermediate, low cytosine promoters or gene bodies (Ficz et al., 2011; Xu, Wu, et al., 2011). In general, Tet1-binding CpG promoters are hypomethylated compared to other promoters, suggesting an inverse correlation between DNA methylation and Tet1 targets (Ficz et al., 2011; Wu, D'Alessio, Ito, Xia, et al., 2011). Inhibition of Tet1 increased the methylation of Tet1 target genes, without any effect on non-Tet1 binding genes, suggesting Tet1 has a role in regulation for these genes by maintaining the hypomethylated status of these regions (Xu, Wu, et al., 2011).

In mouse ESCs, Tet1 has a crucial role in maintenance and regulation of DNA methylation of pluripotent genes such as *Nanog*, a transcription factor involved in self renewal. Tet1 binds and regulates the expression of the *Nanog* promoter by maintaining its hypomethylated status. In contrast, inhibition of Tet1 leads to hypermethylation of the *Nanog* promoter which is associated with gene repression and a decreased proliferation rate (Ito et al., 2010). Moreover, Tet1 depletion induces up-regulation of Tet1 target genes used in development and differentiation (Wu, D'Alessio, Ito, Xia, et al., 2011), suggesting Tet1 has a dual role in maintaining the expression and repression of pluripotent and developmental genes (Ito et al., 2011; Wu, D'Alessio, Ito, Xia, et al., 2011). Furthermore, Tet1 reduction prevented embryonic specification towards the inner cell mass, suggesting an important role of Tet1 in ICM specification (Ito et al., 2010).

Notably, Tet1-knock out (KO) mice ESCs showed a partial reduction in 5-hmC levels and slight DNA methylation effects. The knock out cells maintained their pluripotent state through expression of *Oct4*, *Nanog* and *Sox2* and the cells were able to generate the three embryonic germ layers. Interestingly, Tet1 KO mice were viable and fertile but small in size (Dawlaty et al., 2011).

Inhibition of both Tet1 and Tet2 in mice was associated with a complete depletion of 5-hmC and increase in 5-mC levels in both embryonic stem cells and germ cells, where both Tet1 and Tet2 are more highly expressed than Tet3. In addition, the double mutant embryos exhibited developmental abnormalities and perinatal lethality in half the offspring, whereas the rest were viable (Dawlaty et al., 2013). Interestingly, double knockout of Tet1 and Tet2 had a stronger effect on female mice than male mice. DKO female mice exhibited smaller ovaries and reduced

fertility while males were fertile with only slight differences in testis size (Dawlaty et al., 2013).

1.1.7.4.2 TET2.

TET2 undergoes frequent somatic mutations in a wide spectrum of myeloid and lymphoid malignancies (Cimmino et al., 2011; Solary et al., 2014). In addition to oxidising 5-mC, TET2 may have another function in regulation of haematopoiesis and loss of TET2 may lead to dysregulation of the DNA repair response, increasing genome instability (Feng et al., 2019).

TET2 gene resides at chromosome 4q24 in human and contains 11 exons. Due to chromosomal inversion during evolution, TET2 lacks the CXXC domain (Iyer et al., 2009; Ko et al., 2013; Pastor et al., 2013). As with the other TET proteins, TET2 protein iteratively oxidises 5-methylcytosine to generate 5-hmC and then to 5-fC and 5-caC. Intriguingly, a study using highly sensitive isotope showed that TET2 yields 5-fC and 5-caC by an iterative process directly from 5-methylcytosine without release of the 5-hmC intermediate (Crawford et al., 2016). Normally, haematopoietic tissues such as bone marrow and spleen show high TET2 expression compared to other TET2 expressing tissues (Langemeijer et al., 2009; Li et al., 2011).

TET2 gene is located in a region that shows a high frequency of microdeletions and copy neutral loss of heterozygosity (CN-LOH) in patients with myeloid malignancies (Ko et al., 2010). Moreover, patients with myeloid cancers such as acute myeloid leukaemia (AML), secondary AML (sAML), myelodysplastic syndrome (MDS), myeloproliferative neoplasm (MPNs) and chronic myelomonocytic leukaemia (CMML) showed defects in TET2 (Ko et al., 2010; Mancini et al., 2012; Quivoron et al., 2011). Moreover, TET2 mutation was associated with the levels of 5-hmC, predicting the role of TET2 and 5-hmC in the progression of myeloid disorders (Ko et al., 2015; Quivoron et al., 2011). How TET "loss of function" promotes the progression of hematopoietic malignancies is largely unknown (Lio et al., 2019). Somatic mutations in TET2 are present in about 15% of myeloid cancer patients (Delhommeau et al., 2009).

Human B and T lymphoma cells were reported to have mutations in TET2 that affect the stem cell compartment (Quivoron et al., 2011). Intriguingly, haematopoietic stem cells were recognised with a somatic mutation in one of TET2 alleles compared to mutation in both TET2 alleles in lymphoma cells, which suggested that the mutation in the second allele was acquired during tumour progression (Quivoron et al., 2011). Furthermore, another study by Kubuki et al. 2017 showed that TET2 mutations occurred in some patients with diffuse large B-cells lymphoma (DLBCL) and most likely these mutations contributed in the pathogenesis of their malignancies (Kubuki et al., 2017).

There is evidence suggesting that TET2 mutations are drivers of tumorigenesis in blood cells, and that TET2 mutations are required at the haematopoietic stem and early progenitor cell stage. TET2 gene is frequently mutated somatically in both myeloid and lymphoid malignancies resulting in impairment of TET2 activity that interferes with the conversion of 5-mC to 5-hmC. Interestingly, TET2 is commonly mutated in clonal haematopoiesis in individual with no manifestations of blood cancer. TET2 mutation alone is insufficient to cause haematologic malignancy, suggesting that TET2 mutations represent an early event during tumorigenesis (Chiba, 2017).

Similar to other TET enzymes, the catalytical activity of TET2 enzyme depends on the co-factor α-KG that is generated from isocitrate by isocitrate dehydrogenase enzymes (IDH1 and IDH2). Interestingly, mutations in either *IDH1* or *IDH2* are observed frequently in myeloid cancers. Mutations in *IDH1* or *IDH2* impair the normal production of α-KG by reduction of α-KG to 2-hydroxyglutarate (2-HG), which is a structural analogue of α-KG (Xu, Yang, et al., 2011a). Consequently, defects in IDH1-2 enzymes induce malignant transformation through inhibition of α-KG dependent-TET2 by increasing production of 2-HG leading to a decrease in 5-hmC production (Evans & Griner, 2015; Xu, Yang, et al., 2011a).

Cancer cells are epigenetically distinct from their tissue of origin, as they frequently display global hypomethylation and gene promoter hypermethylation. These epigenetic aberrations can inactivate target genes and contribute to oncogenesis. TET enzymes that initiate DNA demethylation by oxidising 5-mC to 5-hmC are strongly involved as tumour suppressors. Interestingly, prostate cancer was characterised by TET1 mutations which also frequently occur in lymphoid cancers (Cimmino et al., 2015; Spans et al., 2016). TET2 mutations were observed frequently in myeloid cancers as well as in 12 cancer types including in solid tumours (Kandoth et al., 2013; Quivoron et al., 2011; Thienpont, Galle, et al., 2016).

Recently, a genome-wide association study revealed a common SNP near *TET2* linked to breast cancer risk. This higher risk was associated with low expression of TET2 (Guo et al., 2015). DNA methylation regulates the expression of *TET* genes: the *TET1* promoter is frequently hypermethylated in B-cell lymphomas, and the *TET2* promoter is hypermethylated in gliomas, where as *TET1* and *TET3* promoters are both hypermethylated in bladder and breast cancer (Thienpont, Galle, et al., 2016).

TET2 is expressed in various tissues, with increased expression in haematopoietic tissues such as bone marrow and spleen (Li et al., 2011). TET2 deficient mouse models show a reduction in 5-hmC especially in bone marrow and spleen (Ko et al., 2011; Quivoron et al., 2011).

Intriguingly, depletion of 5-hmC in *TET2* deficient mouse was not compensated by increase transcripts of *TET1* or *TET3*, so the relationship between TET2 mutations and the 5-hmC and 5-mC status remains unclear (Quivoron et al., 2011).

Interestingly, myelodysplastic syndrome patients with TET2 mutations treated with DNMT1 inhibitors, azacytidine and decitabine, showed more global loss of DNA methylation compared with TET2 wild-type patients (Huang & Rao, 2014).

Moreover, retinoic acid (RA) induced TET2 to positively regulate the expression of *HOXA* genes upon induction of differentiation, through the conversion of 5-mC to 5-hmC in both mouse and human embryonic carcinoma cell line NT2. Treatment of NT2 cell lines with RA increased TET2 expression and 5-hmC levels at lineage-specific loci, resulting in enhanced expression of *HOXA* gene. Alternatively, TET2 deficiency resulted in low 5-hmC and high 5-mC levels leading to reduced *HOXA* expression (Bocker et al., 2012).

1.1.7.4.3 TET3.

The human TET3 gene is located on chromosome 2 and contains 9 exons with the first exon coding for the CXXC domain. As with TET1 and TET2, TET3 dioxygenase enzyme catalyses the conversion of modified genomic base 5-mC into 5-hmC and further into 5-fC and 5-caC. However, TET3 protein is the only TET family member with reduced expression in both human and mouse embryonic stem cells and is highly expressed in the zygote whereas the other TET proteins are poorly expressed (Gu et al., 2011; Hill et al., 2014; Huang & Rao, 2014; Wu & Zhang, 2010; Wu & Zhang, 2017).

TET3 has an important role in selective active demethylation of paternal DNA in male pronucleus upon zygote formation and before cell division (Figure 1.2) (Bhutani et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011). After fertilization, TET3 rapidly mediates oxidation of 5-mC accompanied by accumulation of 5-hmC, 5-fC and 5-caC in the paternal pronuclei at the PN3 stage and these oxidised forms were diluted passively during subsequent cell divisions (Inoue et al., 2011; Inoue & Zhang, 2011; Iqbal et al., 2011; Wossidlo et al., 2011).

Interestingly, zygotic Tet3 depletion in conditional knockout mice resulted in failure of conversion of 5-mC into 5-hmC in the paternal pronuclei (Gu et al., 2011). Furthermore, Tet3 is used in demethylation of paternal *Oct4* and *Nanog* genes after fertilization; diminished Tet3 enzyme levels leads to impaired demethylation of these genes. Depletion of Tet3 in the germ cells of female mice causes a severe reduction in the fertility, and their heterozygous mutant offspring lacking maternal Tet3 showed developmental defects (Gu et al., 2011).

In contrast to the paternal genome, the maternal genome is largely protected against Tet3 demethylation by PGC7/Dppa3/stella proteins, and is passively demethylated during successive rounds of cell division (Bhutani et al., 2011; Wossidlo et al., 2011). Similarly in bovine and rabbit zygotes, Tet3 show a conserved asymmetric 5-hmC accumulation between paternal and maternal pronuclei (Wossidlo et al., 2011). In contrast to this, another study showed that Tet3 localises to the maternal pronucleus, similar to paternal pronucleus, suggesting that it can oxidise both paternal and maternal DNA in mouse zygotes, but the phenomenon was less pronounced in the female pronucleus (Tsukada et al., 2015). In addition, Tet3 knockout in oocytes had no major impact on oocyte development, maturation, fertilization or pregnancy, but resulted in neonatal sub-lethality suggesting that oxidation of zygotic 5-mC by maternal Tet3 is essential for neonatal growth rather than development (Tsukada et al., 2015).

Few data have been reported on the roles of TET3 in cancer. However, deletion of the region where TET3 is located was observed in a case of myelodysplastic syndrome with ring sideroblasts (Kinney & Pradhan, 2013). Furthermore, the methylation of TET3 in T1 and T2 oropharyngeal cancer and oral cancer patients was associated with poor survival (Misawa et al., 2018), and the methylation status of TET3 was associated with aggressive tumour behaviour in head and neck cancer (Misawa et al., 2018). Induction of Tet3 overexpression in glioblastoma by using epigenetic drugs 10 μM suberoylanilide hydroxamic acid (Sajadian et al.) and 50 μg/mL ascorbate led to markedly reduced tumour formation in an immunodeficient mice model (Carella et al., 2019). Cao and colleagues 2019 showed that a high TET3 level was associated with poor prognosis and poor survival in ovarian cancer patients (Cao et al., 2019). The few present data showed that TET3 has an important roles in cancer. These roles were associated with the progression of some types of cancer, suggesting that further investigations are needed.

1.1.7.5 Substrate preference of TET enzymes.

Three out of four modified forms of cytosine, 5-mC, 5-hmC and 5-fC can act as substrates for TET proteins (Figure 1.7), and the substrate preference for TET enzymes can be viewed at three different levels (Wu & Zhang, 2017).

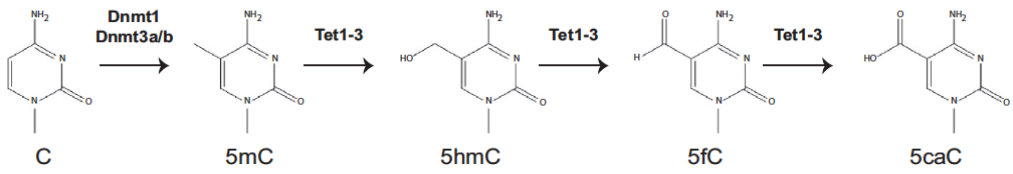


Figure 1. 7. Overview of enzymatic cytosine modifications observed in mammalian DNA.

The first level is detection of the substrate; 5-methylcytosine "the modified base formed by the addition of a methyl group" is predominantly detected in a CpG context but also found in non-CpG contexts. TET enzymes probably prefer CpG contexts over non-CpG contexts (Hu et al., 2013; Pfeifer et al., 2019).

The second level is binding of TET enzymes; TET prefers 5-mC to 5-hmC or 5-fC. In addition, enzyme kinetic analysis demonstrated that TET catalysed conversion of 5-mC to 5-hmC is 3 - 5 fold faster than 5-hmC to 5-fC and 5-fC to 5-caC (Hashimoto et al., 2014; Hu et al., 2015; Ito et al., 2011). Interestingly, mutation of the Thr1372 residue in human TET2 prevents oxidation of 5-hmC to 5-fC or 5-caC, but does not affect the 5-mC to 5-hmC conversion (Liu et al., 2017).

The third level is the symmetrical and un-symmetrical cases of CpG dyads; out of 25 different combinations, 21/25 are possible TET targets, 21/25 are acted on by TET (Figure 1.8), including the pairing of 5-mC, 5-hmC and 5-fC with unmodified cytosine as well as pairing of 5-mC with all forms of cytosine (Crawford et al., 2016; Weber et al., 2016). TET catalyses the oxidation of the substrate in the same rate regardless the status of the complementary strand (Crawford et al., 2016).

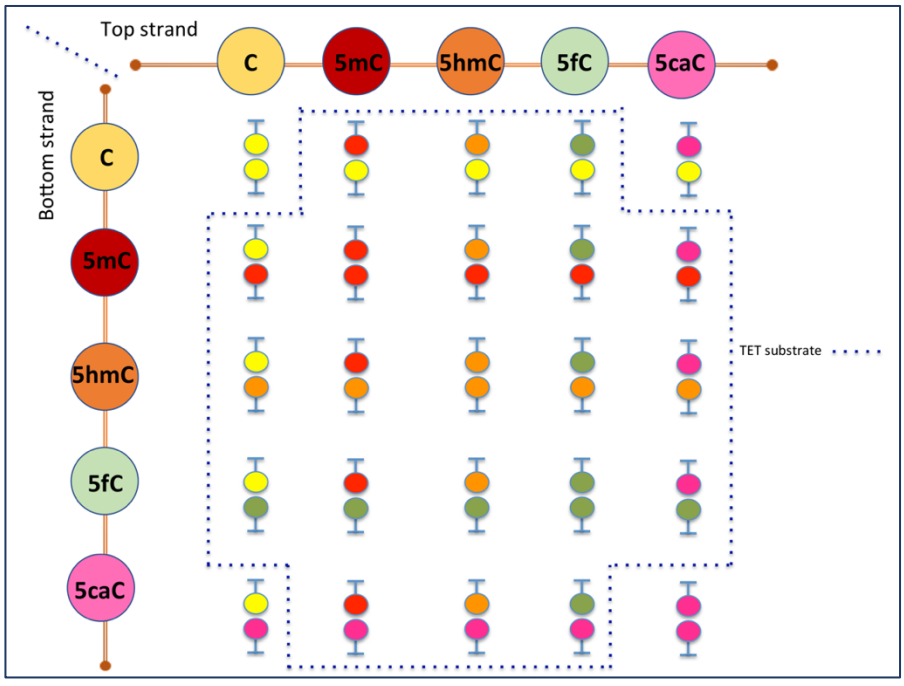


Figure 1. 8. Possible substrates for TET enzymes. CpG dyads with different combinations on the top and bottom strands. TET catalyse their substrate regardless of the modification status of the complementary strands.

1.7.5.1 5-Methylcytosine (5-mC).

Currently, 5-mC is considered a classic epigenetic mark as it fulfils the stringent criteria of an epigenetic framework as being mitotically and meiotically heritable (Riggs, 1975). Cytosine methylation typically occurs in the context of CpG dinucleotides. Typically, the transfer of methyl group is catalyzed by the DNA methyltransferases enzymes (DNMTs), which allowing 5-mC to be faithfully passed through cell division.

5-methylcytosine was first detected in the tubercle bacillus in 1925, and two decades later in calf thymus DNA (Hotchkiss, 1948; Johnson & Coghill, 1925). 5-methylcytosine was proposed as a key epigenetic modification in 1975 (Holliday & Pugh, 1975; Riggs, 1975). Intriguingly, methylation of cytosine does not affect its base pairing properties with guanine, and it is thus replicated as cytosine (Ohki et al., 2001; Pennings et al., 2005). In general, CpG dinucleotide methylation displays a bimodal distribution in mammalian DNA; with the majority of CpG sites being methylated whereas CpG sites in promoter CpG islands are usually unmethylated (Illingworth & Bird, 2009; Meissner et al., 2008; Weber et al., 2007).

Gene promoters can be classified into high-density, intermediate-density and low-density promoters based on their CpG island density. High-density promoters (HCPs) are usually unmethylated and associated with gene transcription, whereas low-density promoters (LCPs) are usually highly methylated regardless of their transcriptional activity. Intriguingly, ICPs display disparate behaviour as they are commonly unmethylated but frequently subject to *de novo* methylation which corresponds inversely with gene transcription (Weber et al., 2007).

1.7.5.2 5-Hydroxymethylcytosine (5-hmC).

5-hydroxymethylcytosine (5-hmC) is the first oxidative step in TET-mediated DNA demethylation. 5-hmC act as the "6th base" of DNA and has major roles in gene regulation and expression.

In 2009, an unidentified spot "X" was detected in a study that compared the abundance of 5-mC in Purkinje and granule cell nuclei using two-dimensional thin layer chromatography. Intriguingly, the spot was found in the xpG context and associated with decreased 5-mC in the samples. Further investigation using high pressure liquid chromatography and mass spectrophotometry demonstrated that "X" was 5-hmC. Since that time, 5-hmC has been detected in mouse brain, suggesting that 5-hmC might play a role in the epigenetic regulation of neuronal functions (Kriaucionis & Heintz, 2009).

It was not the first time that 5-hmC was identified in a genome. As early as 1952, 5-hmC was

detected in T-even bacteriophages and in 1972, 5-hmC was identified in mammalian DNA and was regarded as oxidatively damaged cytosine that might eventually be replaced by DNA repair mechanisms (Ito et al., 2011).

Typically, conversion of 5-mC to 5-hmC by TET proteins blocks the repressive methyl binding domain (MBD) containing proteins and DNMT proteins that would typically be recruited to 5-mC, and promotes gene expression through active demethylation (Branco et al., 2011). However, generation of 5-hmC by active demethylation appears to be particularly important for genes involved in brain function; for example, 5-hmC is enriched in genes involved with synaptic function in mouse and human brain.

1.1.7.5.3 5-Formylcytosine (5-fC).

5-formylcytosine (5-fC) is the second oxidized derivative of 5-methylcytosine. Genomic and proteomics studies have suggested a role for 5-fC in transcriptional regulation through chromatin remodeling. The actions of DNA repair factors and chromatin regulators in selecting and binding 5-fC in genomic sequence suggests that 5-fC may act as an epigenetic signal in its own right (Iurlaro et al., 2013).

Each oxidative derivative of TETs appears to have a different distribution. For example, genome-wide mapping of 5-fC showed that 5-fC is enriched at poised enhancers among other regulatory elements in mouse embryonic stem cells (Song et al., 2013). Raiber et al. (2015) revealed that 5-fC alters the structure of the DNA double helix and induces a unique conformation (Raiber et al., 2015). In addition, Fu et al. (2019) reported that "formyl and carboxyl" modifications of cytosine induce geometric alteration of the DNA minor groove, this alteration is recognized by (TDG) via its finger residue R275 followed by excision and DNA repair (Fu et al., 2019).

1.1.7.5.4 5-Carboxylcytosine (5-caC).

5-carboxylcytosine is the final oxidized derivative of 5-mC. Both 5-fC and 5-caC act as substrates for active demethylation and can be converted to unmodified cytosine by thymine DNA glycosylase (TDG) and base excision repair (Zhang et al., 2012).

Some evidence suggests that 5-fC and 5-caC can act as epigenetic marks, with specific reader proteins (Shi et al., 2017). A previous study demonstrated that 5-fC and 5-caC affect the rate and specificity of RNA polymerase II (RNAPII) activity. Both 5-fC and 5-caC cause increased RNAPII backtracking, increased pausing, and reduced fidelity of nucleotide incorporation (Kellinger et al., 2012).

Like many other oxidatively damaged DNA nucleotides and because both formyl and carboxyl groups are electron-withdrawing, 5-fC and 5-caC nucleotides have weak glycosidic bonds. The weak glycosidic bonds make the base pairing weaker for 5-fC:G and 5-caC:G than C:G and presumably increase the rate of spontaneous flipping (Bochtler et al., 2017; Lenz et al., 2015) creating suitable substrates for excision by TDG (Dai et al., 2016; Maiti et al., 2013).

Thymine DNA glycosylase has been named for its a monofunctional glycosylase activity against thymine bases, arising from deamination of 5-mC mispaired to Guanine. TDG was the first enzyme reported to excise 5-fC and 5-caC (Maiti & Drohat, 2011). Interestingly, the weak glycosidic bonds and high flipping rate for 5-fC and 5-caC make a favourable interaction between TDG and the flipped 5-fC and 5-caC (Hashimoto et al., 2013), and result in faster excision of 5-fC and 5-caC than from Thymine pairs with Guanine (Maiti & Drohat, 2011).

1.1.8 DNA demethylation and cancer.

Cancer growth requires the coordinated change in gene expression of different sets of genes. Alterations in DNA methylation can account for changes in gene expression (Baylin & Jones, 2011; Esteller, 2007). Hypermethylation of CpG islands in gene promoters of tumour suppressor genes is associated with abnormal gene expression resulting in gene silencing. This dysregulation of the tumor epigenome is common within different types of cancer cells (Wu & Zhang, 2011a). Typically, hypermethylation of tumour suppressor gene promoters and global hypomethylation with activation of oncogenes is involved in the progression of cancer (Cheishvili et al., 2015).

In cancer cells, the steady state of DNA methylation can be modulated by many reagents such as pharmacological reagents or other chemicals. This modulation makes DNA methylation attractive as a target for therapeutic treatment (Cheishvili et al., 2015; Szyf, 2005). It has been suggested that impairment of TET-mediated DNA demethylation may contribute to cancer (An et al., 2017; Wu & Zhang, 2011a). The initial story implicating TET proteins in cancer development came from the identification of human *TET1* as a rare fusion partner of MLL in patients with acute myeloid leukemia (AML) (Ono et al., 2002). Consequent research studies uncovered a link between *TET2* gene mutations and myelodysplastic syndromes (Delhommeau et al., 2009; Langemeijer et al., 2009). *TET2* mutations impaired enzymatic activity resulting in a consistent loss of 5-hmC in myeloid malignancies (Ko et al., 2010; Lio et al., 2019). Furthermore, bone marrow samples from patients with *TET2* mutations displayed low levels of 5-hmC in their genomic DNA compared to bone marrow samples from healthy controls (Ko et al., 2010). While this loss of 5-hmC was first associated with mutations in *TET2* genes in

various myeloid cancers, a reduction of 5-hmC was also observed in myeloid cancers with mutations in *IDH1* and *IDH2* genes (Quivoron et al., 2011).

Mutations in *IDH1* and *IDH2* genes generate high levels of 2-hydroxyglutarate (2-HG), which inhibits TET2-mediated 5-mC hydroxylation leading to reduced 5-hmC levels (Figueroa et al., 2010; Mondesir et al., 2016; Xu, Wu, et al., 2011). Interestingly, mutations in *IDH1*, *IDH2* and *TET2* genes are mutually exclusive in AML, suggesting a common role in the pathogenesis of these diseases (Figueroa et al., 2010; Pronier & Delhommeau, 2012).

Reduced expression of TET and IDH genes due to promoter methylation has been reported in many solid tumours, despite the absence of coding mutations (Hsu et al., 2012; Rasmussen & Helin, 2016; Sharma, 2018; Yang et al., 2013).

Global loss of 5-hmC can be broadly used as a biomarker for many human cancers such as melanoma, breast, liver, lung, pancreatic and prostate. Dysregulation of gene expression of both *IDH2* and *TET* genes is associated with loss of 5-hmC in melanoma cells. Interestingly, re-introduction of *IDH2* or *TET2* could restore the 5-hmC level in melanoma cells, suppressing melanoma growth and increasing tumour-free survival in animal models (Lian et al., 2012). Moreover, 5-hmC appeared to be the epigenetic marker that can distinguish melanocytes in benign naevi from those in melanoma, as 5-hmC levels were high in naevi, whereas low or absent levels were seen in all primary and metastatic melanomas (Lian et al., 2012).

Xu et al. (2011) showed a high percentage of patients with lower-grade brain tumours and secondary glioblastomas carried mutations in *IDH1* and *IDH2*. *IDH1* mutations were associated with increased histone methylation and decreased 5-hmC. Furthermore, tumours with either *IDH1* or *IDH2* mutations showed decreased α-KG levels and an increase in 2-HG (Xu, Yang, et al., 2011a). Moreover, *IDH1* and *IDH2* mutations in esophageal squamous cell carcinomas were associated with high 2HG levels, predicting that *IDH1* and *IDH2* mutations might play a limited role in the development and progression of esophageal squamous cell carcinomas (Miyake et al., 2018).

In addition to myeloid cancer, TET2 mutations are common in lymphoid malignancies, especially among patients with T cell lymphomas such as angioimmunoblastic T cell lymphoma and peripheral T cell lymphoma (Han et al., 2015; Lemonnier et al., 2012; Odejide et al., 2014; Quivoron et al., 2011). It is generally accepted that Tet2 loss of function induces myeloid transformation, but the reason is still unclear (Ko et al., 2015). Loss of Tet2 function in Tet2-deficient mice strains did not cause disease, suggesting that there is an unidentified factor that associated with Tet2 loss to drive malignancies (Quivoron et al., 2011). Concurrent depletion

of Tet2 associated with either *Asxl1* or *Ezh2* mutations accelerated the development of myeloid disease in murine models similar to human myelodysplastic syndrome (Abdel-Wahab et al., 2013; Han et al., 2015; Muto et al., 2013).

TET1 and *TET2* expression levels were found to be low in hepatocellular carcinoma (HCC) tissue and cell lines and this was associated with down regulation of global 5-hmC levels which might be a novel prognostic biomarker of hepatocellular carcinoma (Wang et al., 2019).

In summary, many mechanisms may cause depletion of 5-hmC by inactivation of TET enzymes, but the relevant molecular targets of TET proteins that link to cancer development have not yet been identified.

1.1.9 Ascorbate and epigenetics reprogramming.

Ascorbate is an essential, water-soluble micronutrient that is naturally present in some food, and available as a dietary supplement. The importance of ascorbate came from its role as an antioxidant and free radical scavenger as well as acting as a co-factor in numerous enzymatic reactions (Lee Chong et al., 2019; Young et al., 2015). Ascorbate plays an essential role in collagen crosslinking and its deficiency can cause scurvy (Camarena & Wang, 2016; Gould & Woessner, 1957; Van Robertson & Schwartz, 1953).

Most mammals, can generate ascorbate from glucose by a *de novo* biosynthetic pathway in the liver. In contrast, humans as well as primates are unable synthesise ascorbate due to the lack of the functional L-gulonolactone oxidase enzyme (Linster & Van Schaftingen, 2007). For humans, ascorbate needs to be acquired through dietary sources and supplements.

Primarily, ascorbate enters cells through sodium-dependent vitamin C transporters (SVCTs). The high-capacity, low-affinity SVCT1 transporter is primarily responsible for ascorbate absorption and reabsorption in intestinal and renal epithelial cells, whereas the high-affinity and low-capacity SVCT2 transporter is responsible for distribution of ascorbate to tissues (Camarena & Wang, 2016; Wilson, 2005; Young et al., 2015).

Ascorbate as a reducing agent is predominant in the plasma of healthy humans, and can be oxidised to dehydroascorbic acid (May, 1998), which is no longer able to enter cells through SVCTs, but through glucose transporters (GLUTs) and then reduced to ascorbate (Lykkesfeldt, 2007).

Ascorbate acts as a cofactor for many monooxygenases and dioxygenases enzymes. Ascorbate-dependent monooxygenases include dopamine β-hydroxylase and peptidyl-glycine α-

amidating monooxygenase, which require Cu^{+2} and ascorbate as cofactors (Prigge et al., 2000). In contrast, dioxygenases utilise Fe^{+2}, α-ketoglutarate and ascorbate as cofactors.

1.1.9.1 Cofactor for TET dioxygenases.

While, ascorbate is a cofactor for collagen prolyl 4-hydroxylase (P4H) which is well known for its involvement in scurvy, ascorbate also has a role in TET-mediated DNA demethylation and in other dioxygenase mediated process (Vasta & Raines, 2016).

It was shown that ascorbate has the capability to cause widespread DNA demethylation of nearly 2000 genes and modify the status of DNA methylation in mammalian embryonic stem cells (Chung et al., 2010). Standard cell culture media lacks ascorbate in its formula; therefore, the activity of ascorbate-dependent enzymes is compromised in those conditions.

A study by our collaborator Dr. Tim Hore (2016), suggested that ascorbate potentiates TET catalysis by enhancing 5-hmC production, not as a cofactor as reported previously, but rather by enhancing the recycling of Fe^{2+} (reduction of Fe^{3+} to Fe^{2+}) in the catalytic centre of TET proteins (Hore et al., 2016). Ascorbate provides indirect role to enhance TET to induce demethylation.

The Minor and Wang laboratories showed that ascorbate enhanced the generation of 5-hmC in mouse embryonic fibroblast cells (MEFs), most likely by acting as a cofactor for TET to hydroxylate 5-mC (Camarena & Wang, 2016; Dickson et al., 2013; Minor et al., 2013; Young et al., 2015). Intriguingly, MEFs expressed all TET enzymes at low but detectable levels. Thus, these cells constituted a convenient tool to analyze TET enzymatic requirement in a cell culture. The content of 5-hmC is very low in MEFs cultured in ascorbate free media, while addition of ascorbate significantly enhanced the generation of 5-hmC. The maximum level of 5-hmC was achieved when MEFs were treated with 10 μM ascorbate, possibly due to low expression of TET proteins. Surprisingly, ascorbate induced the rapid generation of 5-hmC in MEFs (< 1 hour), while TET or IDH expression was not changed. The effect of ascorbate on 5-hmC could not be attributed to its role as a reducing agent, as treatment with other agent such as glutathione did not affect the 5-hmC levels (Minor et al., 2013).

The presence of ascorbate at physiological concentrations has been shown to enhance the uptake of iron by cells, suggesting that ascorbate might have an indirect effect on 5-hmC by increasing the cellular uptake of iron (Lane et al., 2010). Additionally, using ascorbate transporter inhibitors, such as phloretin or sulfinpyrazone, decreased ascorbate's effect on 5-hmC levels, suggesting that intracellular ascorbate accumulation is necessary for the activity of

TET dioxygenases (Lane et al., 2013; Lane et al., 2010).

Yin and colleagues demonstrated that ascorbic acid could directly enhance the catalytic activity of Tet dioxygenases for the oxidation of 5-mC (Yin et al, 2013). Ascorbic acid interacted with the purified C-terminal catalytic domain of TET enzymes, thereby promoting their folding and/or the recycling of Fe^2 (Yin et al., 2013). In contrast, other reducing agent such as spermidine, vitamin B1, vitamin E, glutathione, nicotinamide adenine dinucleotide phosphate and L-cysteine did not enhance the oxidation effect on 5-mC (Yin et al., 2013). Interestingly, in cells in which Tet1 and Tet2 were deleted, ascorbic acid altered neither 5-mC oxidation nor the overall level of 5-mC, but the ascorbate effects were restored when Tet2 was re-expressed (Yin et al., 2013).

1.1.9.2 Ascorbate and DNA demethylation.

As a co-factor, ascorbate participates in the cascade oxidations of 5-mC to generate the unmodified form 'cytosine' (Young et al., 2015). In regulatory T cells, demethylation of CpG motifs of the conserved noncoding sequence 2 (CNS2) in the *Foxp3* intronic element is essential for the expression of *Foxp3* (Nair et al., 2016). An *in vitro* study found that ascorbate enhanced the demethylation of CpG motifs of CNS2 in the *Foxp3* intronic element. Additionally, the methylated CpG motifs of CNS2 became demethylated rapidly after ascorbate treatment with efficient conversion of 5-mC to 5-hmC in a TET2-dependent manner (the activity of ascorbate disappeared in TET $^{-/-}$ cells). Collectively, for Foxp3 expression, ascorbate was required for the CNS2 demethylation mediated by Tet proteins, suggesting that environmental factors, such as nutrients (ascorbate) could cause changes in immune homeostasis through epigenetic mechanisms (Nair et al., 2016).

Sajadian (2016) showed that the combination of ascorbate and 5-azacytidine (5-AZA) enhanced TET activity in hepatocellular carcinoma (HCC) cell lines leading to enhanced active demethylation (Sajadian et al., 2016b). Ascorbate enhanced the anti-proliferative and apoptotic effect of 5-AZA in these cells. Interestingly, the combination of 5-AZA and ascorbate led to downregulation of SNAIL expression and cell cycle arrest (Sajadian et al., 2016b).

Dickson (2013) showed that ascorbate can induce generation of 5-hmC independent of the levels of iron and 2-oxoglutarate in mouse embryonic fibroblasts (MEFs). However, removing iron from the culture medium did not affect the induction of 5-hmC by 10 μM ascorbate, and the effect of ascorbate did not involve increased expression of Tet1-3 or isocitrate dehydrogenases, which overall suggests that ascorbate may directly participate in the conversion of 5-mC to 5-hmC by acting as a co-factor for TET enzymes (Dickson et al., 2013).

In contrast, *in vitro* enzymatic analysis suggested that hydroxylation of 5-mC mediated by TET enzymes was not affected by ascorbate, as recombinant TET1 was able to convert 5-mC to 5-hmC at the same efficiency in the presence or absence of ascorbate (Tahiliani et al., 2009).

1.1.9.3 Epigenetic regulation by ascorbate in early embryonic development.

To date, there are few studies investigating the effect of ascorbate on DNA methylation and demethylation during embryonic development (Young et al., 2015). Women with lower leukocyte ascorbate concentration had a higher risk for neural tube defect (NTD) compared to low-risk women (Brender et al., 2011; Schorah et al., 1983).

The exact mechanism of DNA demethylation in pre-implanted embryos was unclear for a long time, but now it is known that both TET mediated oxidation and passive dilution participate in this process. A complex epigenetic reprogramming with DNA demethylation and re-methylation occurs during mammalian embryonic development. The expression level of TETs is much higher in embryonic tissues, especially at early developmental stages.

There is potential for ascorbate levels to affect these TET-mediated process, but no data or direct evidence are available. Ascorbate enhances the generation of induced pluripotent stem cells (iPSCs) from differentiated cells, which is associated with genome-wide loss of methylation (Esteban et al., 2010; Stadtfeld et al., 2012). This observation indirectly supports a similar role of ascorbate in early embryonic development.

Recent study by Gao and his colleague showed that ascorbate enhanced the expression of TET1,2,3 proteins in PA (parthenogenic) mice embryos, whereas, DMOG (Dimethyloxalylglycine) treatment inhibits TET1,2,3 expression. Furthermore, treatment with ascorbate during early embryonic development increased the blastocyst rate, in contrast, DMOG reduced the rate, (Gao et al., 2019).

The first round of epigenetic reprogramming occurs, immediately after fertilization and before the pro-nuclei merge, the paternal genome goes through a complex epigenetic remodeling process that includes rapid erasure of DNA methylation. Paternal 5-mC is rapidly replaced by 5-hmC via Tet3-mediated oxidation, leading to passive demethylation and erasing of the paternal methylation patterns (Kohli & Zhang, 2013). In contrast, maternal DNA methylation is unaffected at the one cell stage, but the maternal genome passively demethylates by dilution due to inhibition of DNMT1 (Kohli & Zhang, 2013; Wang et al., 2014).

TET-mediated active demethylation also involved in the second round of epigenetic reprogramming that happens in primordial germ cells (Kohli & Zhang, 2013). The timing of

demethylation in PGCs is started upon the entry of the embryonic day 11.5 post-coitum (dpc). PGCs exhibited dynamic changes on day 12.5 dpc, by erasure of methylation by active demethylation (Hajkova et al., 2002). Mechanistically, the phenomenon loss of DNA methylation in mouse PGCs results from the consequence suppression of the maintenance and *de novo* methylation enzymes by PRDM14 and activation of active DNA demethylation pathways (von Meyenn & Reik, 2015).

1.1.9.4 Inhibitor of TET function.

Although, TET activity can be modulated by many substances including vitamin C, Fe^{2+} and α-ketoglutarate, another citric acid cycle intermediates 2-hydroxyglutarate (2-HG) acts as a competitive inhibitor of multiple α-ketoglutarate dependent dioxygenases, including the TET enzymes (Xu, Yang, et al., 2011b). 2-HG occupies the same space as α-ketoglutarate in the active sites of TET enzymes, and reduces its activity. Similar compounds can bind the active site and competitively downregulates TET, Dimethyloxalylglycine (DMOG) is a 2-oxoglutarate (2-OG) analogue which acts as a competitive inhibitor of dioxygenase protein (Babosova et al., 2019). In addition, TET activity can be inhibited by metals such as cobalt (Rose et al., 2011). Furthermore, Yin and colleagues found that natural Ni^{2+} ion can selectively bind to the Fe^{2+}-chelating motif site and inhibit TET-mediate oxidation of 5-mC (Yin et al., 2017).

Chapter Two: Rationale for research hypothesis, design and aims.

2.1 Introduction.

It is likely that, as yet, unidentified molecular pathways of DNA demethylation occur in cancer. Global changes in DNA methylation that occur in diverse reprogramming networks are likely to need the cooperative action of several DNA demethylation mechanisms.

We know that DNA methylation is altered during development, tissue differentiation and in response to various physiological factors (Moore, Le, et al., 2013). It has been postulated that DNA methylation is dysregulated in various human diseases, involving both gain and loss of DNA methylation (Feil & Fraga, 2012). Nevertheless, the mechanisms involved in this dysregulation are unclear.

Despite the observed dysregulation of DNA methylation in human disease, methylation patterns of differentiated cells are paradoxically believed to be stable. Thus, many questions remain to be answered; for example: 1) how is DNA methylation deregulated in cancer? 2) how does passive and active demethylation interact during reprogramming that happens in the zygote and germ cells? 3) does demethylation occur in somatic cells, and if so, by what mechanism does it occur? 4) what are the mechanisms of demethylating agents in cancer patients?

Professor Mark Hampton and Dr. Karina O'Connor (University of Otago, Christchurch) observed a mark reduction in DNA methylation during the first 6 hours of DNA replication after exposing human Jurkat T cell leukaemia cells to a mild physiological oxidant glycine chloramine (O'Connor et al., 2020). They postulated that the effect of this oxidative stress might be mediated through inhibition of the maintenance DNA methyltransferase (DNMT1*)*. Using low throughput techniques, the Morison group found active loss of methylation in cultured human cells exposed to physiologic oxidative stress. They showed rapid replication-independent demethylation in these cultured cells, suggesting novel mechanisms of active demethylation that have not yet been identified.

A previous PhD student Dr. Helena Magrath, from the Morison laboratory, used hairpin-bisulphite methylation sequencing and confirmed that oxidative stress leads to passive hemi-methylation (Figure 2.1). As expected, similar hemi-methylation could also be induced by decitabine, a demethylating drug that is known to block DNMT1 activity.

In addition, Dr Magrath observed complete demethylation within two hours of the start of synchronised DNA synthesis. These demethylated hairpin-sequences were seen at genes that are consistently densely methylated in untreated Jurkat cells. This was interpreted as active enzymatic demethylation occurring in the presence of hemi-methylated strands.

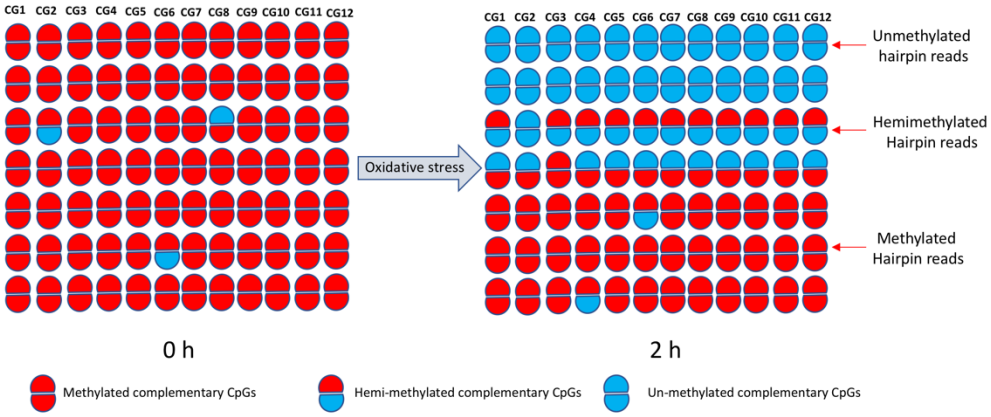


Figure 2. 1. Loss of DNA methylation in synchronized Jurkat cells, 2 hours after exposure to oxidative stress. DNA hairpin sequences showed total demethylation of some reads, which was never seen in untreated cells. Each row represents a PCDHGA12 DNA hairpin containing 12 CpG sites.

Until recently, the loss of DNA methylation was not well understood, and it was believed that somatic cells are protected from demethylation (Lee et al., 2014; Wu & Zhang, 2014). Indeed, in a comprehensive review on demethylation, Wu and Zhang made no mention of demethylation in adult tissues except to comment that: “Intriguingly, the broad expression patterns of TET enzymes and other components of active DNA demethylation pathway in somatic tissues such as brains imply that dynamic DNA demethylation may be more prevalent than previously thought” (Wu & Zhang, 2014).

2.2 Evidence of active demethylation.

Prior research has demonstrated that gain and loss of methylation is a key regulator in the process of tissue differentiation (Boland et al., 2014). For instance, global loss of DNA methylation is seen in the paternal genome after fertilisation before the first cell division (Hackett & Surani, 2013). Replication-independent demethylation in a B cell lymphomas cell lines correlated with induction of the lytic EBV viral cycle (Falk & Ernberg, 1999). The *Tat* gene is demethylated during development of the rat late fetal liver to allow expression of tyrosine aminotransferase (Thomassin et al., 2001). Many genes lose DNA methylation during differentiation of embryonic stem cells to neural progenitor cells (Meissner et al., 2008) and

during haematopoiesis (Calvanese et al., 2012; Hodges et al., 2011). During monocyte differentiation, replication-independent demethylation has been detected in some genes (Klug et al., 2010; Klug et al., 2013). Similarly, in many cancers, DNA methylation is globally decreased (Ehrlich, 2009).

2.3 Specific hypotheses.

There are known enzymes that covalently bind methyl group (CH_3) to the C5 position of the cytosine. Simultaneously, it is likely that other enzymes or mechanisms remove the methyl group from methylated cytosine.

The present work offers two hypotheses summarised in Figure 2.2:

1) **TET proteins are key components of active demethylation in human somatic cells.**
   This hypothesis will be addressed by upregulating TET enzyme activity with ascorbate (Minor et al., 2013). This will be done in synchronised Jurkat cells in G1 phase and then, at the time of release, cells will be exposed to demethylating drug.
2) **Hemi-methylated DNA markedly enhances the rate of TET-induced active demethylation.**
   This hypothesis will be addressed by inducing hemi-methylation in Jurkat cells. Hemi-methylation occurs when the activity of DNMT1 is blocked by either oxidative stress or a demethylating agent (decitabine or azacytidine). Then, TET activity can be enhanced by treating the cells with ascorbate.

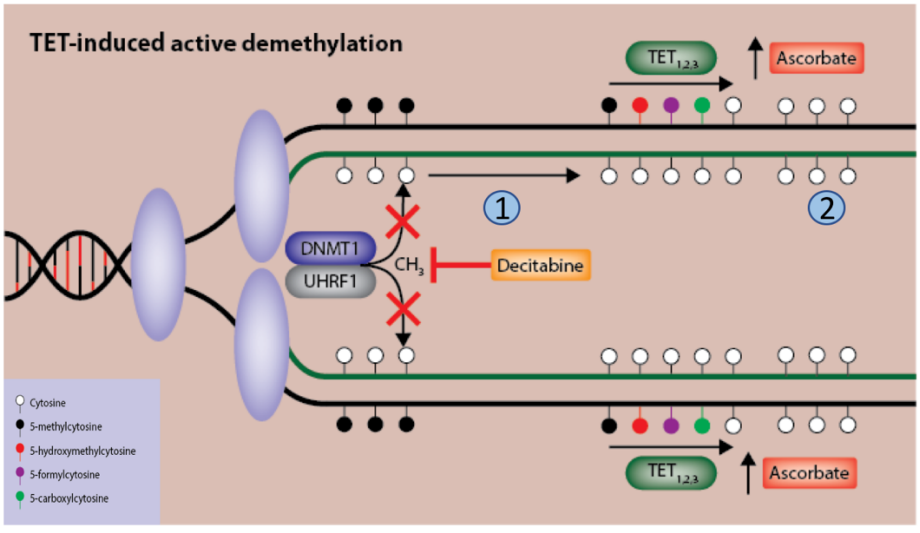


Figure 2. 2 The two hypotheses in this study. It is hypothesised that addition of ascorbate will enhance the demethylating activity of the TET enzymes.

2.4 Overall aims of the research.

The project is divided into three major goals.

The **first** goal was to optimise DNA hairpin protocols and techniques for further work.

The **second** goal was to confirm the recent observations of DNA demethylation that have been observed by the Morison laboratory, by using the high throughput sequencing (MiSeq).

The **third** goal in this study was to investigate the role of ascorbate in upregulating TET enzymes leading to methylation changes.

To achieve these goals:

1- Develop and optimise methods for assessment of the kinetics of demethylation.

First. Optimisation of the hairpin bisulfite sequencing protocols.

Second. Develop and establish bioinformatic tools to analyse DNA methylation of hairpin sequences.

2- **Investigate whether TET proteins mediate the active demethylation in somatic cells.**

This goal will be addressed by modifying TET activity and measuring whether the proportion of unmethylation is altered (see Figure 2.1).

TETs are dioxygenases that can be up-regulated by ascorbate (Blaschke et al., 2013; Minor et al., 2013) and blocked with cobalt (Rose et al., 2011) allowing manipulation of the activity of TET enzymes *in vitro*. We hypothesise that if the TET proteins are responsible for the observed active demethylation, then co-treatment with ascorbate should increase the proportion of fully demethylated (hemi- and unmethylated) sequences, whereas cobalt treatment should reduce proportion of fully demethylated sequences. Documentation that the TET proteins mediate active demethylation in somatic cells will be an important step in determining the factors that might alter the epigenomic profile of human cells.

An Illumina MiSeq sequencer will be used to sequence hairpin libraries by performing sequencing for multiple genes of multiplexed libraries within a single run. High throughput sequencing (Illumina MiSeq) will be used to generate sufficient sequence reads for accurate quantification.

3- Determine whether hemi-methylation enhances the rate of active demethylation.

Based on observations by the Morison laboratory, we hypothesise that active TET-mediated demethylation is markedly enhanced in the context of hemi-methylated DNA. To assess this

observation, Jurkat cells (synchronised at G1 phase) will be treated with decitabine (DNA methyl transferase inhibitor DNMTi) to generate hemi-methylated alleles. Based on this approach, we predict that four different methylation patterns will be detected in Jurkat cells (Figure 2.3).

First: Jurkat cells grown in standard ascorbate-deficient media, will show little or no demethylation.

Second: Jurkat cells treated with decitabine (DNMTi) will show hemi-methylation.

Third: Jurkat cells grown in media loaded with ascorbate will show enhanced demethylation as previously demonstrated in ESCs (Blaschke et al., 2013).

Fourth: Jurkat cells treated with a combination of DNMT1 inhibition (decitabine) and ascorbate will showed complete demethylation in Jurkat cells.

Confirmation of these hypotheses will help explain the kinetics of demethylation in cancer cells, and will have major implications for the current models of physiologically induced demethylation.

4- Investigate the effect of long-term exposure of decitabine, ascorbate and glycine chloramine on DNA demethylation.

To investigate DNA demethylation, we started by using glycine chloramine treated Jurkat cells from the Hampton laboratory (UOC), but this model was unreliable. Therefore to create a stable system we used a hypomethylating drug decitabine to generate hemi-methylated DNA.

This will be investigated by:

First: Examine the effects of decitabine and glycine chloramine on DNA methylation.

The experiments for aim 1 & 2 will be extended to measure the proportion of unmethylated hairpin sequences at 24 and 48 h and continued for several cell cycles (4-5 days) to compare the short-term and long-term effect after a single exposure. Cell number and viability will be measured throughout.

We predict progressive loss of methylation at genes including *RASSF1* and *PCDGHA12*. The alternative hypothesis is that, since the effect of glycine chloramine is short-lived, restoration of DNMT1 activity will occur, and the short-term changes that we have observed will not persist.

Second: Investigate the roles of TET enzymes in chronic exposure model by adding ascorbate. As above, we hypothesise that the addition of ascorbate will upregulates TET enzymes and enhance DNA demethylation.

Third: Assess the roles of TET enzymes in leukaemia cell lines (Molt4, Nalm6 and HL60).

Fourth: Perform low coverage genome-wide methylation sequencing, to assess the global impact on methylation levels if significant demethylation is observed at our test genes.

5- Determine the effect of DNMTi and ascorbate on hydroxymethylcytosine levels in Jurkat cells.

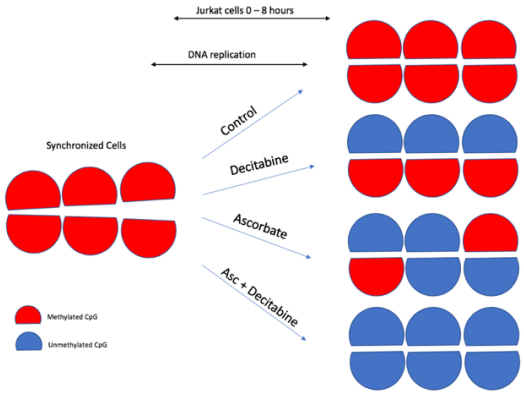


Figure 2. 3. Prediction of methylation patterns in Jurkat cells after 500 μM sodium ascorbate and DNMT1 inhibition.

2.5 Rationale behind this project.

It has been suggested that human disease may result from failure to control the methylation process, leading to gains and losses of DNA methylation (Ehrlich et al., 1982). However, in cancer DNA methylation is markedly disrupted, globally decreased and locally increased, the particular mechanisms involved in DNA demethylation are still unknown.

The potential impact of understanding the underlying mechanisms include: 1. increasing our knowledge of environmentally/physiologically-induced methylation changes, especially in the context of oxidative stress; 2. revealing factors that contribute to the observed hypomethylation in cancer; 3. understanding the interaction between passive and active demethylation reprogramming that occurs in the zygote and in germ cells; and 4. modulating the therapeutic efficacy of demethylating drugs used in cancer therapy.

Chapter Three: Materials and methods.

3.1 Materials.

3.1.1 Chemicals and kits

Chemicals and kits used in experiments are listed in Table 3.1 below.

Table 3. 1. Chemicals, kits and instruments.

Reagent	Supplier
MethylFlash™ Global DNA Hydroxymethylation ELISA Easy Kit (Colorimetric) (Colorimetric)	EPIGENETEK
Tris-hydroxymethyl aminomethane (TRIS) Ethylenediaminetetraacetic acid (EDTA) Acetic acid 2-deoxycytidine 5-aza-2'-deoxycytidine (Decitabine) Ethidium Bromide Thymidine Dimethyl Sulfoxide (DMSO) Dimethyloxalylglycine (DMOG) 10x Orange G loading Dye β-hydroxybutyrate ((±)-Sodium 3- hydroxybutyrate) D-(+)-Glucose Sodium L- Ascorbate $FeCl_3$ Sodium Pyruvate Dimethyl-Alpha-ketoglutarate	Sigma Aldrich
Agarose LE	Roche
Agarose, Certified Molecular Biology	BioRad
Ethanol	Normapur
2-mercaptoethanol Leukaemia Inhibitory Factor (LIF) Propidium Iodide + RNase Trypsin-EDTA Dulbecco's Modified Eagle Medium (DMEM) Minimum Essential Medium Non-Essential Amino Acids (MEM NEAA) RPMI 1640 RPMI + Glutamax 1 Kb+ DNA Ladder	Life Technologies
QIAamp DNA Minikit™.	Qiagen
EZ DNA Methylation-Gold™ Kit. DNA Clean & Concentrator-5™.	Zymo Research
Agencourt AMPure XP™.	Beckman Coulter
Fetal Bovine Serum	Moregate Biotech

0.4% (w/v) Trypan Blue	Gibco
Agilent 2100 Bioanalyzer	Agilent Technologies
Agilent High Sensitivity DNA Chips	Agilent Technologies
Agilent High Sensitivity DNA Reagents	Agilent Technologies
TriGene Advance	Ethical Agents International LTD
Qubit™ 4 Fluorometer	Thermofisher Scientific
UV Gel dock – Mini HD9	Unitec Cambridge

3.1.2 Enzymes

Table 3. 2. Enzymes and their suppliers.

Reagent	**Supplier**
KAPA HiFi Hotstart Uracil+ 2x Readymix	Illumina
Taq FastStart DNA Polymerase	Roche
T4 DNA ligase (400,000 units/ml; Catalogue number: M0202S)	New England BioLabs
BamHI (20,000 units/ml; Catalogue number: R0136S)	New England BioLabs
SacI (20,000 units/ml; Catalogue number: R0156S)	New England BioLabs
HpaII (20,000 units/ml; Catalogue number: R0171S)	New England BioLabs

3.1.3 Oligonucleotides.

Integrated DNA Technologies (IDT) synthesized all primers and hairpin linkers.

Table 3. 3. Primers used for amplification of regions of interest from bisulfite treated DNA. Illumina-specific tag sequences are highlighted in grey.

Gene name	Primers
RASSF1	Forward: ACGACGCTCTTCCGATCT-TATAGTTTTTGTATTTAGGTTTTTATTG
	Reverse: CGTGTGCTCTTCCGATCT-CCTACACCCAAATTTCCATTAC
PCDHGA12	Forward: ACGACGCTCTTCCGATCT-GTAAGGATTAGGTGGTGAGTAGTTT
	Reverse: CGTGTGCTCTTCCGATCT-ACCAAATAATAAACAACCTTTTCTT
NEFM	Forward: ACGACGCTCTTCCGATCT-TTATAAGTAGTTTGGGATTGAAAGG
	Reverse: CGTGTGCTCTTCCGATCT-TAACTCATCTTAAAAACCTTAAAAC
KCNA4	Forward: ACGACGCTCTTCCGATCT-TTTGGAGTTTAGTATAGGAGGGATT
	Reverse: CGTGTGCTCTTCCGATCT-CTATAAAAACCACAACCAAAAACTC
C10orf53	Forward: ACGACGCTCTTCCGATCT-GAATTTATTGTGTTTGGTTGTTGTT
	Reverse: CGTGTGCTCTTCCGATCT-AAACCCACCTTACAAACCCTAC
MARCH11_2	Forward: ACGACGCTCTTCCGATCT-GGAAGTGGGGGATTTTTTAGT
	Reverse: CGTGTGCTCTTCCGATCT-CCAAAAACTTCTATTTTAAAACTCTC
GOLSYN	Forward: ACGACGCTCTTCCGATCT-GGGTTTTGGTAGATTATGTTGTAGG
	Reverse: CGTGTGCTCTTCCGATCT-TCAATATTCAAAACTCTTCCTCTCC

3.2 Methods.

3.2.1 Leukaemia cell lines.

Jurkat T cell leukaemia and other leukaemia cell lines (MOLT4, NALM6 and HL60) were obtained from the American Type culture collection (Rockville, MD, USA) and maintained by Rob Weeks (Department of Pathology, University of Otago, Dunedin) and by Dr. Karina O'Connor (University of Otago, Christchurch). Cells were grown in RPMI 1640, DMEM and RPMI Glutamax medium supplemented with 10% heat inactivated fetal bovine serum, 4.5 g/L glucose, 1mM sodium pyruvate at 37°C in a humidified incubator with 5% CO_2.

3.2.2 Mammalian cell culture.

Cell lines were obtained from Robert Weeks (Department of Pathology, University of Otago) and Karina O'Connor (Department of Pathology, University of Otago, Christchurch). Stock vials were stored in liquid nitrogen with 3-5 x 10^6 cells in 1 mL. All the procedures were done in sterile conditions within a class II biological safety cabinet. The hood was cleaned using 70% ethanol and TriGene prior to use.

3.2.2.1 Cryopreservation and thawing.

To preserve Jurkat cells, cell suspensions (6 x 10^6 cells) from the exponential growth phase were centrifuged at 300 g for 5 minutes and resuspended in cold medium with 10% dimethyl sulfoxide (DMSO). Resuspended cells were transferred to cryo vials, placed in freezing containers at -80°C for 20 h to facilitate slow freezing, and then transferred to liquid nitrogen. For thawing, cells were thawed by placing in a 37°C water path for 2 minutes followed by dropwise addition of 1 mL of pre-warmed culture media. Cells were transferred to 15 mL plastic tube and pelleted by centrifugation (300 g for 5 min) to remove DMSO resuspended in pre-warmed media and transferred to fresh T25 flask.

3.2.2.2 Media preparation.

Cells were grown in RPMI, RPMI + glutamax and DMEM media composed of 10% v/v heat inactivated fetal calf serum, 1 µM sodium pyruvate and glucose (4.5 g/L). Cells were cultured in suspension at a starting cell density of 8 x 10^5 within a T25 or T75 flask at 37°C in a humidified atmosphere incubator with 5% CO_2. Cell density was maintained between 0.8-1.2 x 10^6 cells per mL by splitting the cells every three days.

3.2.2.3 Measurement of cell count and viability.

Cell viability was determined by utilization of the Trypan Blue method using a haemocytometer. A suspension of Jurkat cells (50 µL) was mixed with 10 µL Trypan Blue, then the mix was loaded into a haemocytometer. Viable cells in the four large corner squares were counted and the mean of the squares was calculated. Live cells appear unstained shiny and transparent. The number of viable cells calculated according to the following equation:

$$Number\ of\ viable\ cells = cell\ counts \times 60/50 \times 10^4.$$

3.2.2.4 Thymidine block procedure.

The normal cell cycle is composed of two main phases: interphase (G1, S and G2) and mitotic (M) phase (Chen & Deng, 2018b; Leger et al., 2016; Rodriguez-Ubreva et al., 2010). During interphase, the nuclear DNA is duplicated and the cell prepares itself for division, which occurs during mitosis. Cell synchronisation is particularly useful protocol to investigate process that occur during specific stages of the cell cycle. Excess thymidine is an inhibitor of DNA synthesis that can be used to arrest the cells at the G1/S checkpoint by depletion of intracellular cytosine. The use of thymidine is known as a thymidine block.

3.2.2.4.1 Double thymidine block.

Cells were grown to required concentration 1 x 10^6 /mL. Then, thymidine was added to the culture media (final concentration 2 mM) and incubated in a tissue culture incubator at 37 ℃ for 18 h. Thymidine was removed by washing the cells twice with pre-warmed media, and the cells released from thymidine block by incubation with deoxycytidine for 9 h. A second round of thymidine incubation at a final concentration 2 mM was repeated for another 14 hours followed by release from thymidine using deoxycytidine (Chen & Deng, 2018b).

Cells were collected at 0, 2, 4, 6 and 8 hours for cell cycle analysis by DNA staining using propidium iodide (PI).

3.2.2.4.2 Single thymidine block.

Cells were grown to 1 x 10^6 /mL in RPMI-Glutamax media. Cells were incubated with thymidine (2 mM) at 37℃ for 18 hours. Jurkat cells were washed two times with pre-warmed media and released from thymidine block by incubation with deoxycytidine (Chen & Deng, 2018b; Cude et al., 2007).

Cells were collected at 0, 2, 4, 6, 8, 24, 48 and 72 hours of cell cycle analysis by DNA staining using propidium iodide (PI). Cell cycle was analysed by flow cytometry with FlowJo software.

3.2.2.5 Reagent preparation.

2'-deoxythymidine (thymidine).

100 mM 2'-deoxythymidine solution was freshly prepared for each experiment, thymidine crystals were dissolved in MilliQ water on the day of use. Jurkat cells at density 1 x 10^6 cells/mL were incubated with media containing thymidine to give final concentration 2 mM. Cell suspension was gently mixed by pipetting and cells were incubated 18 hours at 37 ℃.

2'-deoxycytidine.

10 mM 2'-deoxycytidine solution was freshly prepared for each experiment, 2'-deoxycytidine crystals were dissolved in MilliQ water on the day of use. Jurkat cells were released from thymidine block by incubation with 2'-deoxycytidine at final concentration 50 µM.

5-Aza-2'-deoxycytidine (decitabine).

10 mM 5-Aza-2'-deoxycytidine (decitabine) solution was freshly prepared for each experiment, decitabine crystals were dissolved in MilliQ water on the day of use. Jurkat cells were incubated with decitabine with final concentration 5 µM.

Glycine chloramine.

Glycine chloramine was prepared by Dr Karina O'Connor on the day of use as the following: hypochlorous acid was added drop wise to five molar excess of glycine in an equivalent volume of milliQ water or PBS with vortexing. The concentration of glycine chloramine was determined spectrophotometrically using $\varepsilon_{252} = 415\ M^{-1}\ cm^{-1}$. Absorption spectra from 200 – 400 nm were also monitored to ensure the absence of dichloramine (λ_{max} 300 nm) and any unreacted hypochlorous acid (λ_{max}290 nm). Jurkat cells were incubated with glycine chloramine so that the final concentration 500 µM.

Sodium L-ascorbate.

100 mM Sodium L-ascorbate and ascorbic acid solutions were prepared on the day of use. Ascorbate and ascorbic acid were dissolved in MilliQ water and incubated with Jurkat cells at a final concentration of 500 µM.

3.2.2.6 FACS/cell cycle analysis.

Flow cytometry was used to confirm the synchronization process by comparing untreated control cells with cells treated with thymidine. Briefly, treated and control cells were centrifuged at 300x g for 5 min before washing twice with 1 mL cold PBS. Cells were resuspended in 100 µL of cold PBS and fixed with 900 µL of ice-cold 70 % ethanol added dropwise slowly and stored at 4 °C.

Prior to flow cytometry analysis, fixed cells were centrifuged at 350 g for 5 minutes and washed twice with 1 mL PBS. Then, cells were resuspended in 250 µL PBS and 250 µL of propidium iodide/RNase staining solution (catalogue number: F10797) and incubated in the dark for 30 minutes at room temperature prior to flow cytometric analysis. After incubation in darkness, DNA content was determined by using a flow cytometer (Beckman Coulter).

3.2.2.7 Cell viability assay following decitabine exposure.

A titration assay was performed to determine the effect of increasing decitabine concentration on the viability of Jurkat cells, Molt4 cells, Nalm6 cells and HL60 cells.

Cells at a density of 1 x 10^6 cells/mL were synchronized by single thymidine block described in section 3.8.5.2. Upon release from thymidine block, cells were treated with decitabine at 0, 0.5, 1, 2 or 5 µM. Samples were incubated for 24 hours followed by manual live and dead cell counts as described in section 3.2.2.3.

3.2.2.8 Experimental design and cells treatment.

The diagram below represents the tissue culture protocol. Briefly, Jurkat cells (1x 10^6 cells/mL) were divided into two main groups (untreated control cells and ascorbate treated cells), cells were synchronised by incubation with 2'-deoxythymidine at final concentration 2 mM for 18 hours, and released by 50 µM 2'-deoxycytidine (Figure 3.1).

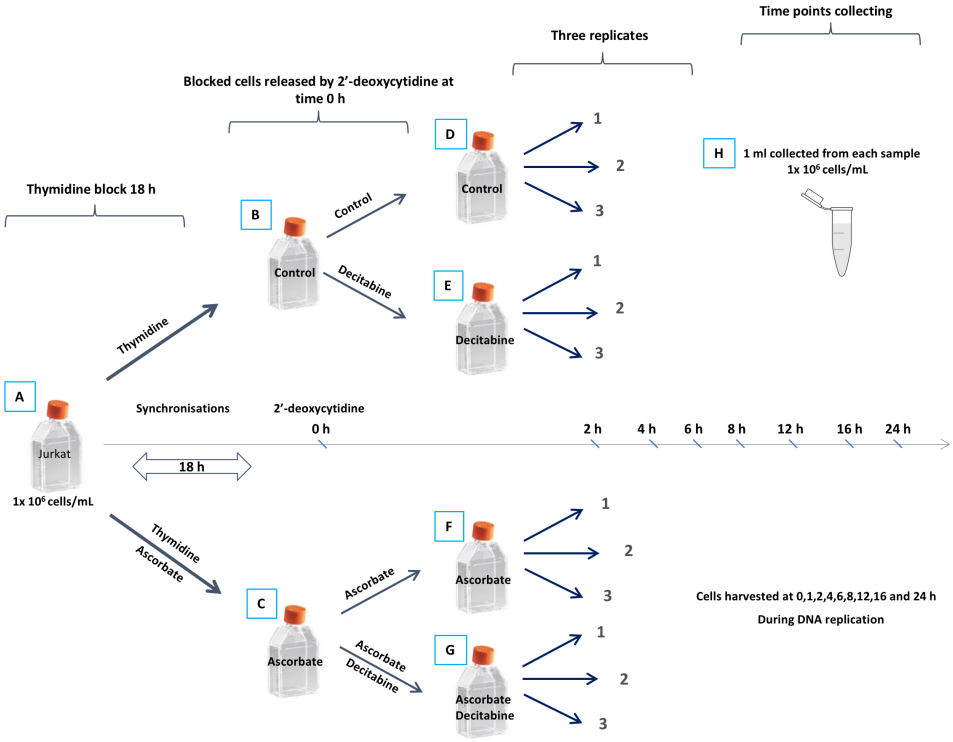


Figure 3. 1. Scheme of the experiment designed for Jurkat cell treatment. A) Flask of Jurkat cells (1x 10^6 cells/mL) were divided into two main flasks. Flask B) Jurkat cells were treated with 2'-deoxythymidine at final concentration 2 mM for 18 hours. Flask C) Jurkat cells were incubated with 2'-deoxythymidine at final concentration 2 mM and ascorbate 500 µM for 18 hours. After synchronization, Jurkat cells were washed twice and each flask divided into two flasks. Flask D) Jurkat Cells were released from thymidine block by incubation with 50 µM 2'-deoxycytidine (considered as control sample). Flask E) Jurkat Cells were released from thymidine block by incubation with 50 µM 2'-deoxycytidine and supplemented with 5 µM decitabine at the time of release. Flask F) Jurkat cells (pre-incubated with ascorbate) were released from thymidine block by incubation with 50 µM 2'-deoxycytidine and replenished with 500 µM ascorbate. Flask G) Jurkat cells (pre-incubated with ascorbate) were released from thymidine block by incubation with 50 µM 2'-deoxycytidine and replenished with 500 µM ascorbate and supplemented with 5 µM decitabine. H) Cells were harvested in triplicate at increasing time point during DNAreplication. Collected samples were divided into two groups. One group for genomic DNA extraction and sequencing, and the other group for cell cycle analysis.

3.2.3 Genomic DNA preparation.

3.2.3.1 Identification of densely methylated regions in Jurkat cells.

PCDHGA12 and *RASSF1* gene promoters were reported to be hypermethylated in haematological cancers (Gordon et al., 2012; Milani et al., 2010; Taylor et al., 2007).

The dense methylation of these gene promoters was confirmed by Rob Weeks and Helena Magrath in Jurkat cells by targeted bisulfite sequencing.

Additionally, *MARCH11*, *NEFM*, *KCNA4*, *C10orf53* and *GOLSYN* gene promoters were confirmed to be densely methylated in Jurkat cells after being identified from publicly available 450K methylation array data sets with the R package, ABC.RAP (CRAN depository; Alsaleh et al., *in preparation*).

3.2.3.2 Identification of densely methylated regions in other leukaemia cell lines.

The *PCDHGA12* gene promoter was found to be densely methylated in MOLT4, NALM6, HL60 leukaemia cell lines and Myelodysplastic syndrome MDS cell line. Whereas, *RASSF1* gene promoter was found to be densely methylated only in the MOLT4 cell line and was hemi- or un-methylated in the other cell lines.

3.2.3.3 DNA extraction.

Genomic DNA was extracted from cell lines using the Qiagen QIAamp DNA Blood Minikit™ system, according to the manufacturer's instructions. Cell pellets were resuspended in 200 µL of PBS, before the addition of 20 µL of proteinase K and 200 µL lysis buffer and incubation at 56°C for 2 hours. Then, 200 µL of 100% ethanol were added, pulse vortexed and spun. The total 600 µL was transferred to a Qiagen minispin column, and centrifuged at 6000 x g (800 rpm) for 1 minute. Columns were transferred to a new collection tube. AW1 buffer 500 µL was added, and centrifugation at 6000 x g (8000 rpm) for 1 minute. Columns were transferred to new collection tubes and 500 µL of buffer AW2 was added and centrifugation at 20,000 x g for 3 minutes. A dry spin was performed for 1 minute at 20,000 x g. To elute DNA, each column was transferred to a 1.5 mL tube and 150 µL of AE buffer was added. All columns were incubated at room temperature for 5~10 minutes prior to eluting DNA by centrifugation at 6,000 x g for 1 minute. DNA concentration was quantified using NanoDrop™.

3.2.3.4 Quantification of DNA.

DNA concentration was determined using a NanoDrop™ spectrophotometer and AE buffer as a blank. The ratio of the absorbance at 260 nm was used to assess the purity of DNA, good quality DNA has an A260/280 ratio above 1.8.

3.2.4 Generating DNA hairpin molecules.

3.2.4.1 Overview and method development.

To determine the symmetry of CpG methylation, and similar to Laird et al. (Laird et al., 2004) who first illustrated the use of hairpin linkers, the Morison laboratory developed a hairpin bisulfite technique to generate methylation data for both complementary strands

simultaneously. This technique allowed comparison of parent-daughter strand CpG methylation symmetry, enabling assessment of global methylation fidelity.

To design barcoded hairpin bisulfite linkers for genomic DNA, sequence within and adjacent to a CpG Island of interest were mapped for restriction endonuclease recognition sites. The favoured restriction enzymes are those that 1) produce long sticky ends, 2) are not affected by potentially methylated target sequence. Requirement one increases the specificity of subsequent PCRs for the region of interest, and requirement two allows the presence of hypermethylated and hypomethylated alleles in the final PCR products (Pingoud et al., 2014).

3.2.4.2 Hairpin linker design.

The hairpin DNA linker structure allows the binding of the two DNA strands together, so that both strands can be interrogated and the DNA methylation of complementary CpG sites can be determined. Hairpin needs to have a stem, sticky ends and loop structure; the presence of the loop allows us to include a random nucleotide sequence unique to each hairpin molecule, which will permit monitoring of over-amplification and contaminating molecules. The molecular barcodes are a random sequences arranged from adenine, thymine and guanine nucleotides.

Hairpin linkers were designed to be complementary to the targeted cut-site of genomic DNA (Figure 3.2). Hairpin molecules are ~ 36 base-long oligonucleotides designed to connect the parent-daughter strands together (Table 3.4). Each hairpin is composed of three regions:

1. **Hairpin loop**. The sequence in the loop had no complementary bases and incorporate a random 14-nucleotide molecular barcode consisting of three nucleotides, adenine, thymine and guanine as these nucleotides are insensitive to bisulfite treatment (Clark et al., 1994). Thus, this barcode region has ≈ 4 million (3^{14}) possible, unique sequences by which duplicate or contaminating sequences can be recognized and excluded.
2. **Complementary stem sequence.** Sequences around the hairpin barcodes that allow the molecules to loop back on itself.
3. **Sticky end.** Hairpin sticky ends are complementary to the targeted cut-site of genomic DNA. To provide successful ligation of the hairpin linker to the 3’ end of the desired DNA, a phosphate group is present at the 5’ end.

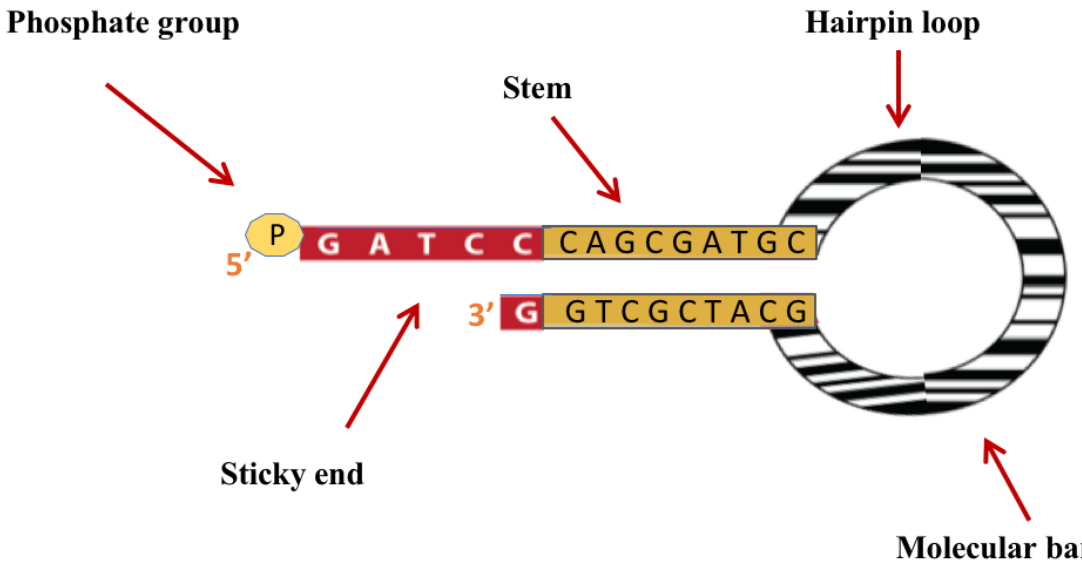


Figure 3. 2. The structure of a hairpin linker., with three components: (i) Hairpin loop with a unique 14 bp barcode embedded into the loop; (ii) Complementary stem sequence and (iii) Sticky end with a phosphate group at the 5'end.

Table 3. 4**.** Hairpin linker sequences for *RASSF1* and *PCDHGA12* promoters. The letter D represents A, G or T bases that were added randomly in the barcode sequence.

Hairpin linker	**Sequence**	**Amplicon length**
RASSF1	pCGCGATGC(D)$_{14}$GCATCGCGAGCT (SacI)	428 bp
PCDHGA12	pGATCCAGCGATGC(D)$_{14}$GCATCGCTG (BamHI)	384 bp

3.2.4.3 DNA digestion.

Hypermethylated genes were cleaved with either SacI (for the *RASSF1* promoter) or BamHI (for the *PCDHGA12* promoter) restriction enzymes in separate reactions. 1 µg DNA isolated from each sample was digested with BamHI or SacI for three hours at 37 °C, in a total volume of 50 µL, as shown in Table 3.5.

Table 3. 5. Reaction mixtures for BamHI and SacI restriction enzymes.

BamHI mixture	**Volume (µL)**	**SacI mixture**	**Volume (µL)**
BamHI	1.0 (20 units)	SacI	1.0 (20 units)
NEBuffer	5.0	NEBuffer	5.0
gDNA (100 ng/µL)	10.0	gDNA (100 ng/µL)	10.0
Milli Q water	34.0	Milli Q water	34.0
Total	**50.0**	**Total**	**50.0**
Digestion: 37 °C for 3 hours			

Successful digestion was confirmed by electrophoresing the digested DNA on 0.8% agarose gel for 60 minutes at 50 V. Figure 3.3 below shows undigested, and SacI and BamHI digested DNA.

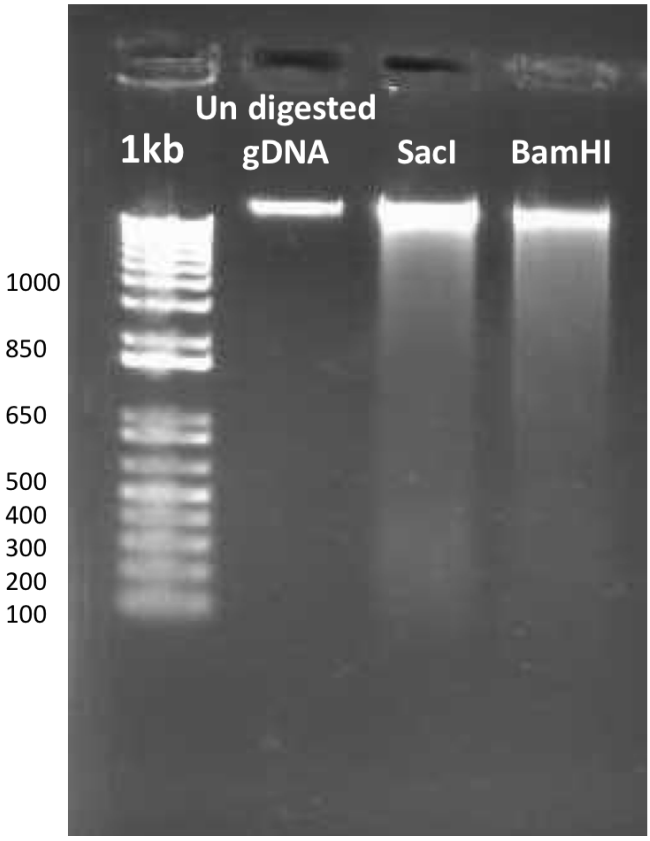


Figure 3. 3. Jurkat genomic DNA digested with SacI and BamHI restriction enzymes. The gel photo shows the restriction enzyme assay for Sac1 and BamH1 compared with undigested. The product visualized on 0.8% agarose gel for 60 minutes at 50 V.

3.2.4.4 DNA clean up.

To remove restriction enzymes and undesirable waste products of the digestion step, digested DNA samples were cleaned using a Zymo Research Clean and Concentrate kit, according to manufacturer's protocol. Briefly, digested DNA was added with binding buffer (1:5 ratio) to a Zymo-SpinTM Column in a collection tube. Tubes were centrifuged for 30 seconds at 20,000 x g. DNA samples were washed twice with washing buffer and eluted with 20 µL elution buffer.

3.2.4.5 Hairpin ligation.

Purified DNA (400 ng) with sticky complementary ends was ligated to hairpin linkers using T4 DNA ligase (Thermo Fisher Scientific). The hairpin linker concentrations were optimised using different hairpin concentration (0.1, 1, 2 and 5 µM); optimum ligation results were obtained by using 5 µM hairpin. DNA-hairpin ligation was performed in 20 µL reaction mixture at 4°C overnight (Table 3.6).

Table 3. 6**.** Ligation reaction mixture for PCDHGA12 and RASSF1 hairpins linkers.

Reagents	**PCDHGA12 Volume (µL)**	**RASSF1 Volume (µL)**
T4 ligase enzyme	0.5 (200 units)	0.5 (200 units)
10X ligase buffer	2	2
Hairpin linker (RASSF1 or PCDHGA12)	5	5
Milli Q water	4.5	4.5
DNA (400 ng digested DNA)	8	8
Total	**20 µL**	
Ligation: 4 °C overnight		

3.2.4.6 Principle of generating PCDHGA12 hairpin molecules.

The densely methylated *PCDHGA12* promoter region was cleaved with BamHI restriction enzyme and ligated with a barcoded hairpin oligonucleotide (Figure 3.4).

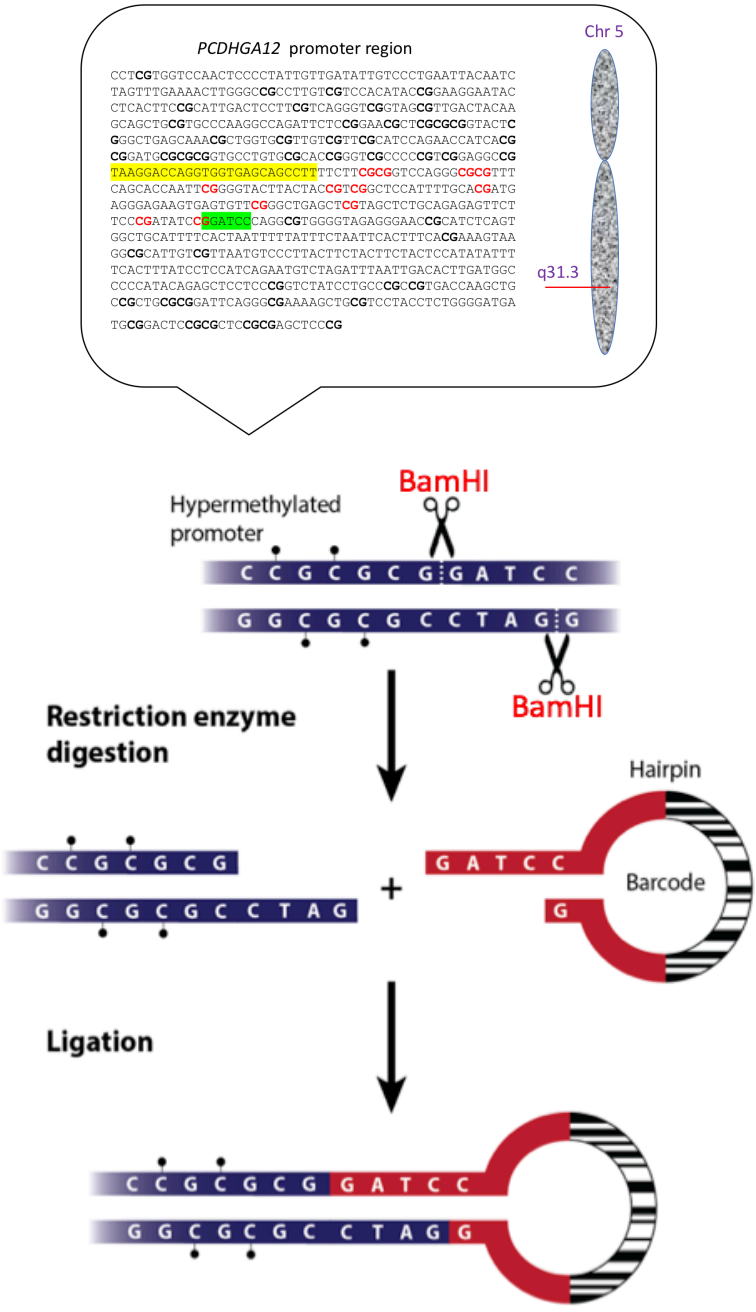


Figure 3. 4. Generation of PCDHGA12 DNA hairpins. Upper panel shows the *PCDHGA12* promoter region, green colour shows the BamHI restriction site, yellow colour shows forward primer sequence and red colour shows CpGs within *PCDHGA12* promoter. Lower panel shows the protocol of generating DNA hairpin molecule, DNA was digested with BamHI restriction enzyme. Hairpin linkers (with a unique 14 bp barcode) complementary to BamHI digested DNA were ligated to the complementary overhang of the target DNA.

3.2.4.7 Principle of generating *RASSF1* hairpin molecules.

The densely methylated *RASSF1* promoter region was digested with SacI restriction enzyme and ligated with a barcoded hairpin oligonucleotide (Figure 3.5). Every hairpin linker contained a 14 bp molecular barcode.

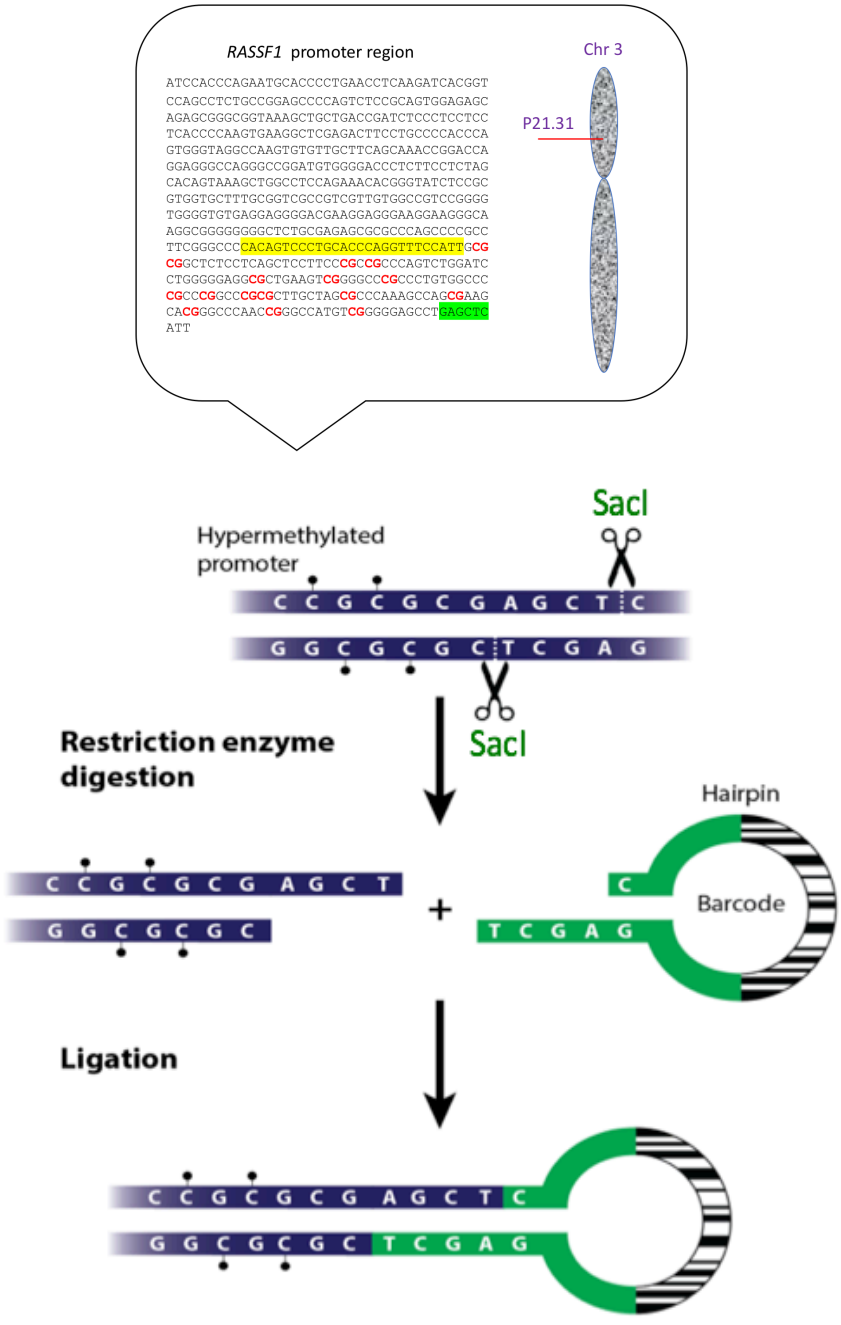


Figure 3. 5. Diagram showing generation of the RASSF1 DNA hairpins. Upper panel shows the *RASSF1* promoter region, Green colour shows SacI restriction site, yellow colour shows forward primer sequence and red colour shows CpGs within the *RASSF1* promoter. Lower panel shows the protocol of generating DNA hairpin molecule. DNA was digested with SacI restriction enzyme. The hairpin linkers (with a unique 14 bp barcode) complementary to SacI digested DNA were ligated to the complementary overhang of the target DNA.

3.2.5 Bisulfite conversion and PCR workflow.

Bisulfite conversion is the gold standard technique used to study and analyse DNA methylation. Unmethylated cytosine residues are deaminated through a series of conversion steps to generate uracil and subsequently to thymine upon PCR amplification, whilst leaving 5-methylcytosine intact (Figure 3.6) (Laird et al., 2004). Briefly, the steps are: 1) Denaturation, by incubating DNA at 98 °C; 2) Conversion, deaminates cytosine residues in DNA by incubation with sodium bisulfite at 64°C; 3) Desulphonation, by incubating the DNA at high pH at room temperature for 15-20 minutes to generate uracil; 4) PCR amplification, during which uracil is converted to thymine during the subsequent PCR.

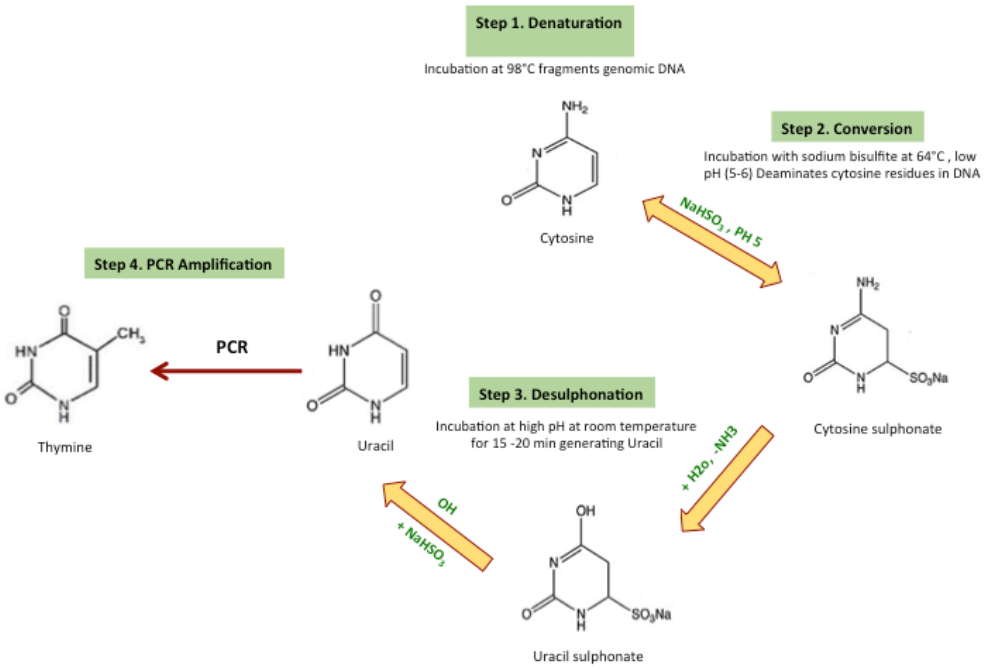


Figure 3.6. Bisulfite conversion of cytosine.

3.2.5.1 Bisulfite conversion of hairpin-ligated DNA

Bisulfite conversion for the hairpin-ligated DNA was performed using Zymo Research "EZ DNA Methylation-Gold™" kit, according to the manufacturer's instructions. Briefly, 130 µL of prepared CT sodium bisulfite conversion reagent was added to 20 µL DNA solution (500 ng). Then, DNA was denatured at 98°C for 10 minutes and incubated at 64°C for 2.5 hours, and the DNA was desulphonated for 15-20 min at room temperature. DNA was eluted in 10 µL of M-elution buffer. At the end of bisulfite conversion, unmethylated cytosine residues were

deaminated and converted to uracil, which are read as thymine in subsequent PCRs (Figure 3.7).

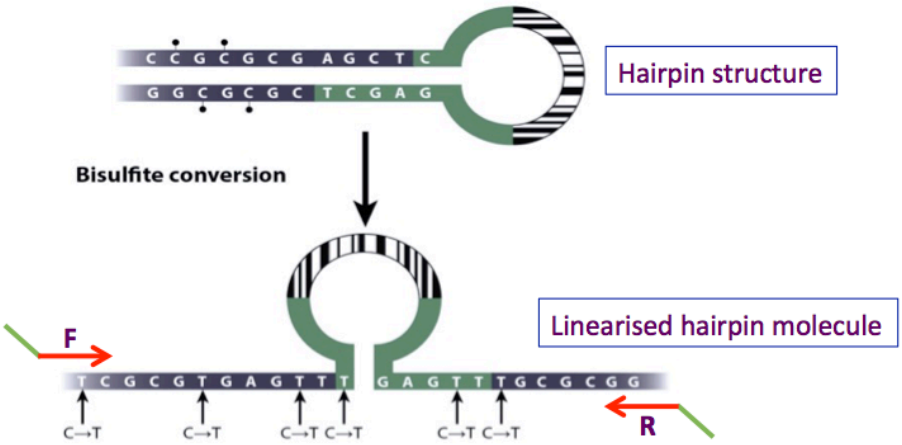


Figure 3. 7. Diagram of bisulfite conversion of hairpin molecule. The bisulfite reaction linearises the hairpin after bisulfite conversion. Primers designed for the bisulfite convert hairpin DNA within 150-200 bp of the linker.

3.2.5.2 Primers design and gene specific amplification.

Primers were designed to anneal the bisulfite genomic sequence 150-200 bp each side of the hairpin linker creating an amplimer of 350-450 bp which is ideal for Illumina MiSeq (Figure 3.7 & Figure 3.8). Primers were designed by Rob Weeks using the online software *MethPrimer* (Li & Dahiya, 2002) and the following criteria: 1) primers should not contain any CpG dinucleotides to avoid discrimination between unmethylated and methylated DNA; 2) primer length should be longer than 20 nucleotides to enhance specificity. In addition, Illumina universal adapter sequences were added to forward and reverse primers, which are required for MiSeq sequencing.

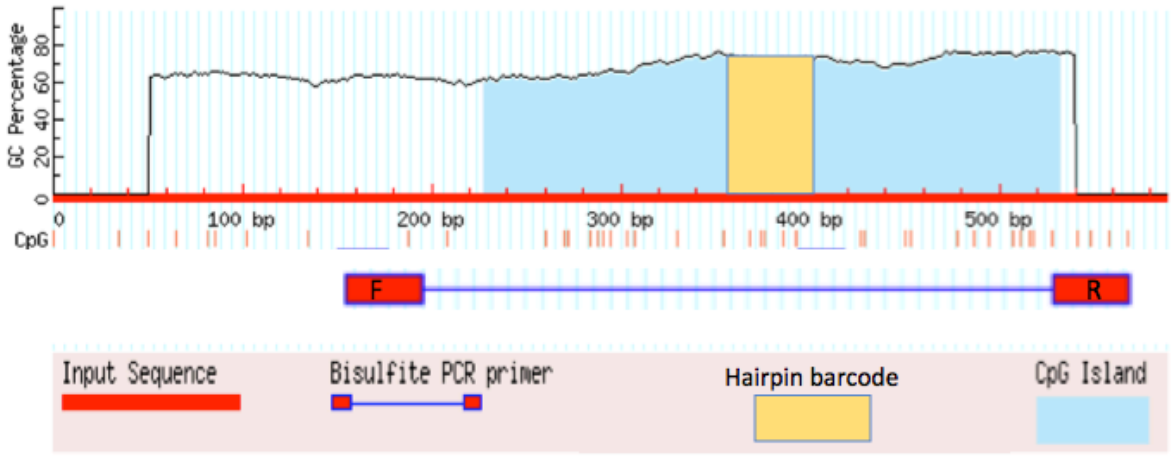


Figure 3. 8. Modified MethPrimer image of *RASSF1* hairpin sequence that was used as an input for primer design. CpG island(s) was found in the sequence (shaded in blue) and hairpin primers (non-complementary) were picked for each side of the hairpin barcode (yellow).

3.2.5.3 PCR workflow and preparation for MiSeq sequencing.

Preparation for Illumina sequencing required two separate rounds of PCR. The advantage of the two rounds PCR approach is to allow us to amplify the target amplicons using nonspecific adapters that can then be used as a template for unique adapter sequences in the second round PCR.

During the PCR reaction, index sequences on both ends of the DNA were added, that enables dual-indexed sequencing of pooled libraries on any Illumina Sequencing System (Figure 3.9).

A 2-step PCR protocol was used that enabled a small amount of input DNA (2 ng), and requires only a single set of non-indexed primers for each amplicon, whatever the number of samples to be multiplexed. This protocol provided a low-cost method as an alternative to purchasing individual sets of indexed primers for PCR amplicons of each sample. In brief,

1) DNA was amplified using gene specific primers with Kapa HiFi HotStart Uracil + ReadyMix polymerase.

2) The different amplicons for each sample were pooled, 2 µL of each amplicon from each sample were pooled. As each amplicon had the same 18 bp tag sequence at the end of each product, they could be pooled and amplified in the second PCR reaction.

3) A clean-up step. The purified products were quantified using qubit dsDNA HS kit and diluted to 1 ng/µL for the second round PCR reaction.

4) Second round PCR for adapter ligation. In the second round of PCR, unique indexes were incorporated into the PCR amplicons of sample. The primers used in the second reaction were

designed with non-specific adapter sequences. These sequences contain the necessary primer binding sites for paired end sequencing and dual index sequencing.

5) Pooling all the samples.

6) The samples were cleaned by Ampure XP clean up.

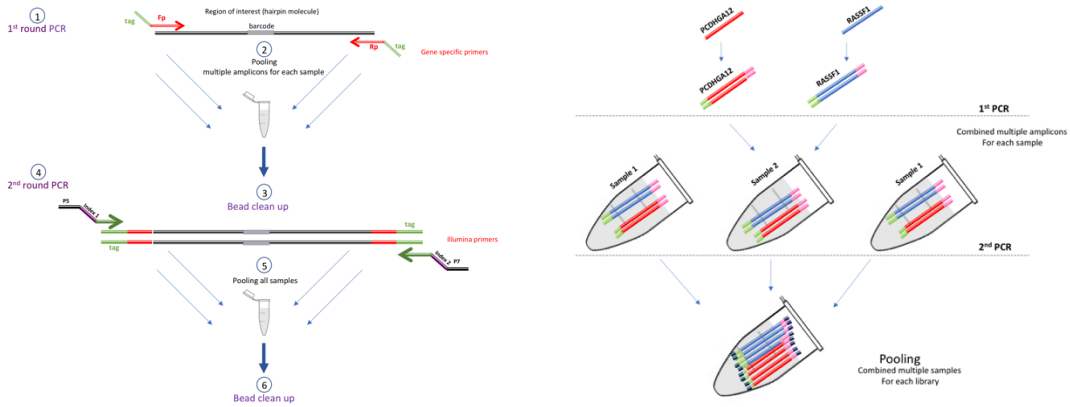


Figure 3. 9. An overall diagram of the two step PCR pooling workflow.

3.2.5.4 Thermocycler Protocols.

Table 3. 7**.** First round PCR protocol for *RASSF1* & *PCDHGA12* promoters.

Step	Temp °C	Time	Cycles
Initial denaturation	95 °C	4 min	X 1
Denaturation	98 °C	20 sec	X 34
Annealing	56.1 °C *(RASSF1)*	15 sec	
	57.2 °C *(PCDHGA12)*	30 sec	
Extension	72 °C	30 sec	
Final extension	72 °C	7 min	X 1
Hold	4 °C	Forever	

3.2.5.5 PCR verification.

Gel electrophoresis was used to verify successful amplification, absence of contamination and no secondary products. 4 µL of PCR product was stained with 1 µL blue xylene cyanol loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, 15% Ficoll in TE) and visualised on 2 % agarose gel. The DNA samples were electrophoresed in 0.5xTAE buffer with EthBr for 30 min

at 100 V. DNA samples were compared to the 1 kb Plus DNA Ladder to determine the size of DNA products. UVIDOC HD6 DNA/protein gel imaging system (Cambridge) was used to visualise products (Figure 3.10).

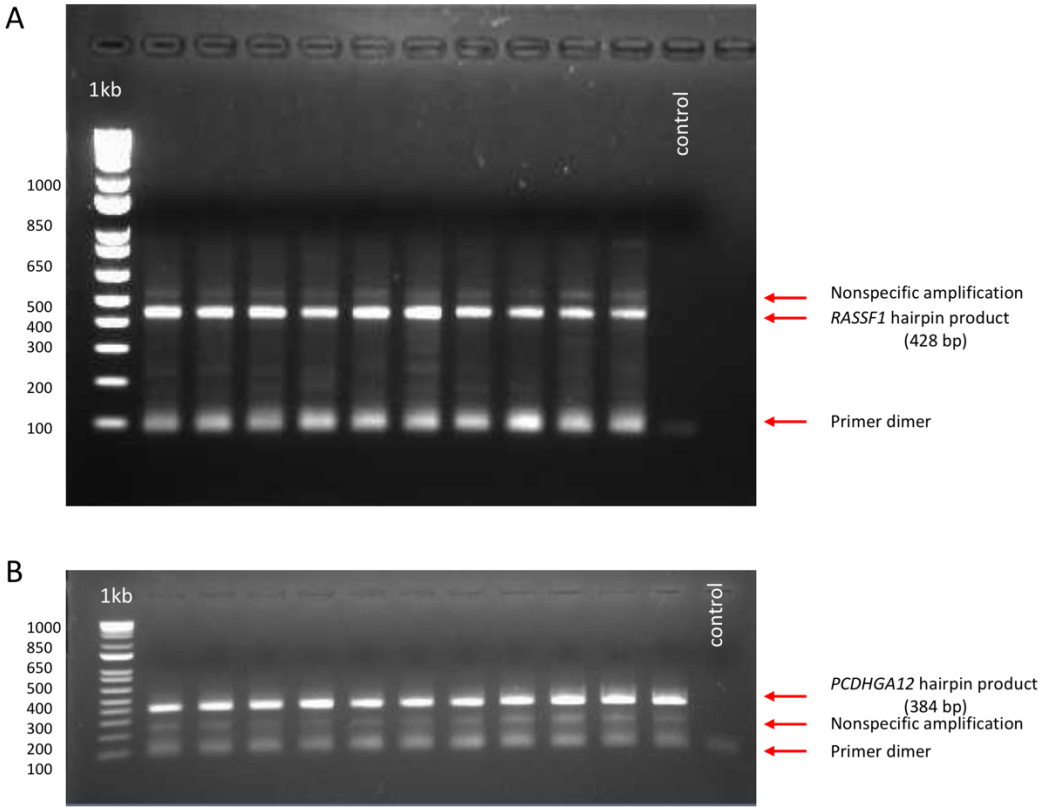


Figure 3. 10. Shows 1st round PCR product of A) *RASSF1* hairpin product (428 bp) and B) *PCDHGA12* hairpin product (384 bp) visualised on 2% agarose gel. Electrophoresed for 30 min at 100V.

3.2.5.6 Ampure XP bead clean up.

Agencourt AMPure beads are solid-phase paramagnetic beads used for high-throughput purification of PCR amplicon. The purification process requires no centrifugation or filtration. A bead to sample volume at 1:1 was used to selectively bind DNA >100 bp to paramagnetic beads and residual primers, nucleotides and primer dimers were removed by washing of the DNA.

PCR products were purified using Ampure XP beads following the manufacturer's protocols. The procedure was performed in three stages: 1) Selective binding of DNA to paramagnetic beads and beads separation using a magnet; 2) Washing with 70% ethanol to remove primer

dimers and contaminants; 3) Elution of the purified DNA from the magnetic beads (Figure 3.11).

Detailed protocol:

1. An aliquot of Ampure XP beads was taken from the fridge, mixed thoroughly and left to equilibrate to room temperature for 30 min.
2. 10 µL of specific pooled sample was added to 10 µL of Ampure beads, mixed well by pipetting and incubated for 5 min.
3. Beads were collected on the side of the tube using a neodymium magnetic plate and stand for 5 min, aspirated and the supernatant discarded.
4. Samples were washed to remove salts, dNTPs, primers and primer-dimer sequences (< 100 bp), by adding 200 µL of freshly prepared 70% ethanol, incubated for at least 30 sec, aspirated and the ethanol discarded.
5. Repeat the wash step: add 200 µL of freshly prepared 70% ethanol, incubate 30 sec, aspirated and the ethanol discarded.
6. The pellet was dried completely. Then tubes were removed from the magnetic plate.
7. Samples were resuspended in 40 µL of elution buffer. Incubate 10 min at room temperature.
8. Then the tubes were placed on the magnetic plate for 5 min.
9. The supernatant containing the DNA was removed carefully, so as not to disturb the bead pellet and stored at 4 °C until needed.

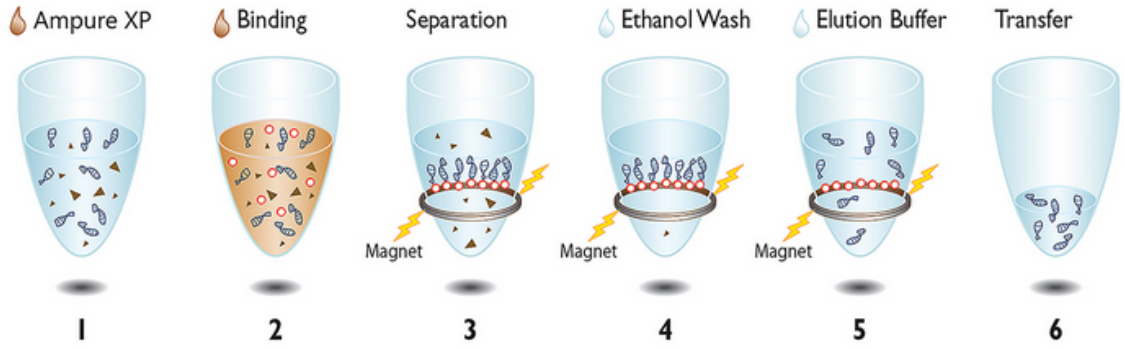


Figure 3. 11. Shows Ampure bead purification steps. 1) Mixing the Ampure beads solution with PCR products, 2) Binding the beads to DNA molecules, 3) The solution placed in the magnetic plate to draw the beads onto the side of the wells, supernatant and beads discarded to wash out primer dimers and small DNA molecules, 4) Ethanol wash followed by 5) Elution and 6) Transfer the purified DNA. Image source https://labplan.ie/product/ampure-xp/.

3.2.5.7 Quantification of PCR products.

The product of the 1st round PCR was quantified using a Qubit fluorimeter and the 1x High Sensitivity dsDNA kit (ThermoFisher), as per the manufacturer's protocol. Briefly, 2 µL of PCR product was added to 198 µL 1x dsDNA HS buffer and measured using the dsDNA program.

3.2.5.8 Second round PCR.

The aim of the second round PCR is to incorporate the Illumina indices onto the PCR product of the first round PCR, so only 10 cycles of amplification was used to label each sample with specific Illumina adapter sequence (Figure 3.12). These unique indexes identified each sample after pooling. KAPA HiFi Ready mix (Roche) was used in each reaction. Samples were diluted to 2 ng/µL. Master mix and protocol details for second round PCR are in Tables 8 and 9.

Table 3. 8. Second round PCR master reaction mix.

PCR 2nd Round reagents	**Volume (µL)**
KAPA Hi Fi Mix	5
10 mM Forward Primer	0.5
10 mM Reverse Primer	0.5
MilliQ H_2O	3
Bead-purified 1st round PCR product (2 ng total)	1
Total	**10**

Table 3. 9. Second round PCR thermocycler protocol.

Step	**Temp °C**	**Time**	**Cycles**
Initial denaturation	95	2 minutes	
Denaturation	95	20 seconds	Repeat 10 cycles
Annealing	58	20 seconds	
Extension	72	20 seconds	
Final extension	72	40 seconds	
Hold	4	∞	

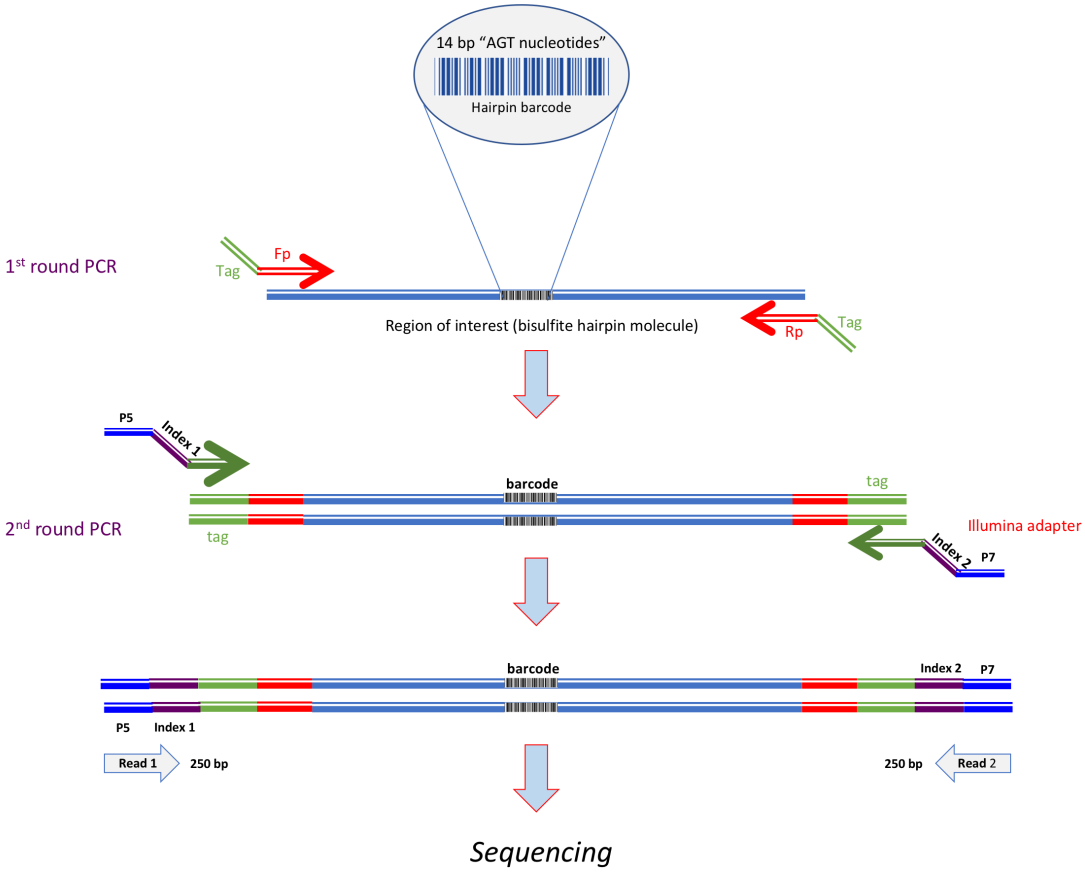


Figure 3. 12. Diagram of two PCR steps. Step one, 1st round PCR to amplify region of interest using gene specific primers tagged with non-specific universal adapter. Step two, 2nd round PCR to label the samples with a unique index using Illumina primers. There is a 96-base increase in amplicon size at this step.

3.2.5.9 Sample confirmation, pooling and purification.

PCR amplification was confirmed on 2% agarose gel (Figure 3.13). There was an increase in amplicon size during electrophoresis, the final MiSeq PCR products length for *RASSF1* and *PCDHGA12* hairpins are:

RASSF1 hairpin molecule = 428 + (68-18) + (63-17) = 524 bp

PCDHGA12 hairpin molecule = 384 + (68-18) + (63-17) = 480 bp**.**

"428 and 384 are the amplicons size" + "68 is the Illumina FP size - 18 is the overlapped adapter sequence between the forward primers of the first round and second round PCR" + "63 is the Illumina RP size" - 17 is the RP overlapped adapter sequence between the reverse primers of the first round and second round PCR".

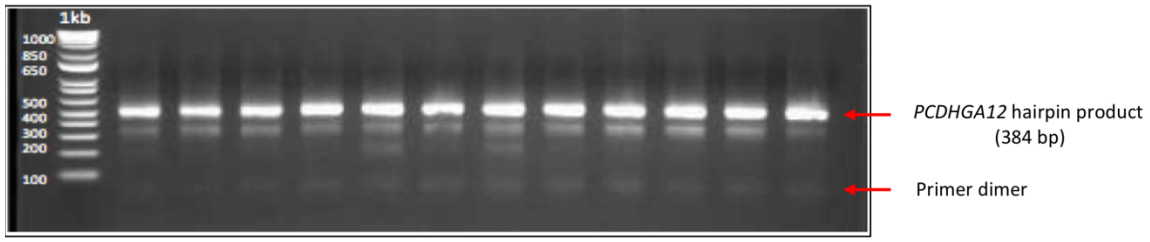


Figure 3. 13. Shows *PCDHGA12* second round PCR product visualized on 2% agarose gel. Electrophoresed for 30 min at 100V.

All samples were pooled into one single tube. 5 µL of each post PCR reaction was pooled into one tube, except where a sample had a lower band intensity in comparison to other bands so the volume of the sample increased accordingly. The intensity of the band in the agarose gel image specified the volume of the sample that was added to the pooled library.

The pooled library was cleaned up to remove primer-dimer and short non-specific amplification products using Ampure XP beads using the protocol described previously. Specific cut-off points for the size selection was obtained by altering the bead to DNA ratio during DNA purification (Figure 3.14). The ratio 0.7:1 (volume of beads : volume of DNA) was determined to be optimum for my amplicons, as this purification ratio removed non-specific amplification products < 350 bp length.

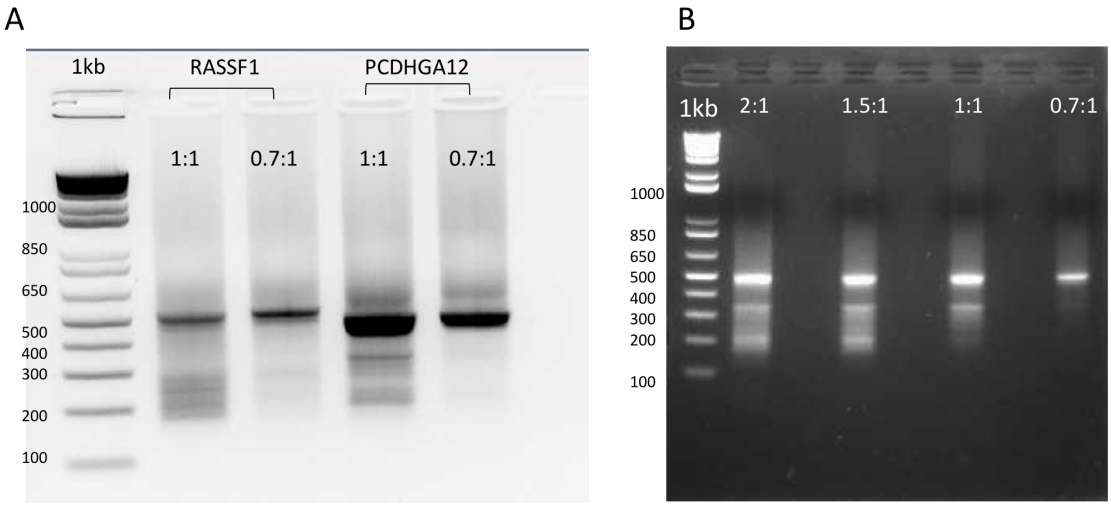


Figure 3. 14. Effect of bead : DNA ratio on DNA purification. Ampure bead to DNA ratio used for specific cut-off points for size selection (e.g. 0.7 beads : 1 DNA). A) *RASSF1* and *PCDHGA12* second round PCR products cleaned with two different beads : DNA ratios. B) optimal size exclusion of *PCDHGA12* PCR products was achieved by adjusting the beads to DNA ratio. Product visualized on 2% agarose gel. Electrophoresed for 30 min at 100V.

3. 2.5.10 Quantification of the second round PCR products.

The bead-cleaned indexed amplicons were quantified on Qubit using high sensitivity DNA Qubit kit (Invitrogen) and then diluted to 1 ng/µL.

3.2.5.11 Agilent 2100 *Bio Analyzer*.

Pooled PCR products were assessed using a High sensitivity DNA protocol (*Bio Analyser*, Agilent Technology, Germany). The Agilent *Bio Analyser* is an automated device that uses capillary electrophoresis for the analysis of DNA integrity, quality and quantity. The device generates electropherogram peaks and an electrophoretic image for the DNA samples (Figure 3.15).

All the solutions and reagents were prepared according to the manufacturer's protocol and stored in the fridge. Before use, the reagents were allowed to reach room temperature for 30 minutes. Briefly, a new high sensitivity DNA chip was placed onto the chip priming station. Then, 9 µL of the gel was loaded in a specifically marked wells of the DNA chip. The chip priming station was closed and the plunger of the syringe was set at 1.0 mL, and the plunger was pressed down until a clip held it. After 60 seconds, the clip was released and the plunger was moved back to 0.7 mL mark. After 5 seconds the plunger pulled back to 1.0 mL mark and the chip was removed. Then 9 µL of the gel was loaded in two wells. The ladder and the samples were loaded to separate wells respectively. The chip was run in the Agilent 2100 *Bio Analyzer* within 15 min.

The final molarity of each library was calculated based on the *Bio Analyser* average fragment size.

$$DNA\ concentration\ nM = \frac{Qubit\ DNA\ concentration\ (ng/\mu L)}{Bio\ Analyser\ average\ size\ x\ \ 0.00065}$$

$$= \frac{34.2\ (ng/\mu L)}{513\ x\ \ 0.00065} = 103\ (Qubit\ conc\ (nM)) \rightarrow volume\ to\ make\ 25\ \mu L\ 4\ nM = 25\ x\ (4/103)$$

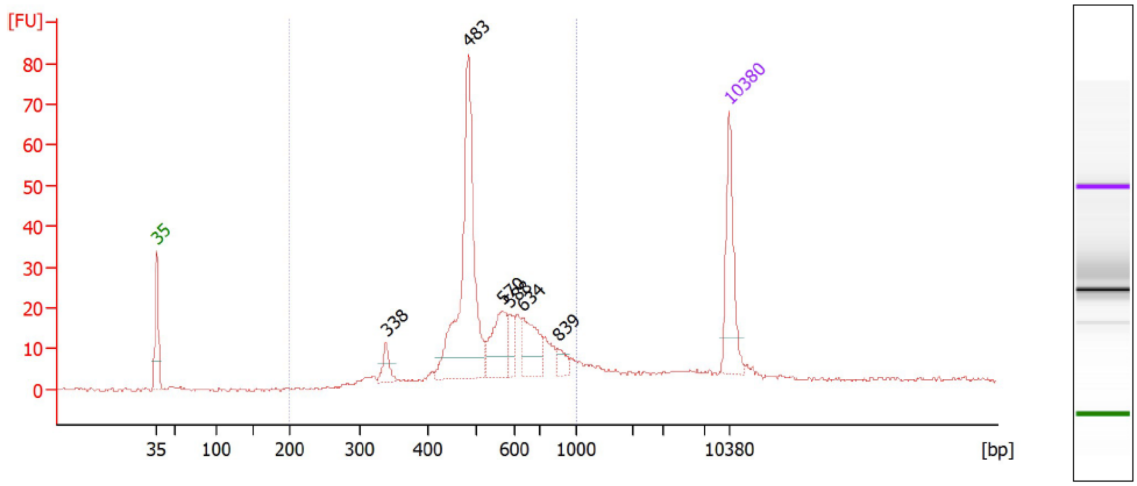
= 1 µL ➔ add water 24 =25 µL.

After that, 4 nM final concentration was prepared for each library, and then all the DNA libraries were pooled into 20 µL tube based on the number of amplicons in each library.

To calculate the ratio =Number of my samples/total number of samples=250/313=0.8.

To calculate the volume to be mixed= the ratio of my samples (0.8) x (total volume) 20 µL = 16.0 µL.

After that, each library was diluted to 4 nM, and then all the DNA libraries were pooled into a 20 µL tube based on the amount of DNA in each library. The final pool (4 nM) of DNA was taken to Dr. Rob Day (Department of Biochemistry, University of Otago) who performed the MiSeq sequencing run.



Peak		Size [bp]	Conc. [pg/µl]	Molarity [pmol/l]	Observations
1	◀	35	125.00	5,411.3	Lower Marker
2		338	23.88	107.0	
3		483	290.14	910.8	
4		570	63.39	168.4	
5		588	19.41	50.0	
6		634	57.03	136.4	
7		839	14.53	26.2	
8	▶	10,380	75.00	10.9	Upper Marker

Figure 3. 15. A representative example of Agilent 2100 Bio Analyser results of the pooled PCR products. The peak labelled "483" represents the pooled library of 480 bp *PCDHGA12* amplicons.

3.2.6 Measurement of intracellular ascorbate concentration.

Intracellular ascorbate concentration was measured by Professor Margreet Vissers' laboratory (Department of Pathology, University of Otago, Christchurch).

Jurkat cells were treated for 18 hours with 250 µM and 500 µM sodium –L-ascorbate added during the beginning of synchronization (2 mM thymidine block). After release from thymidine block, Jurkat cells were collected at 0 and 6 hr time points. Collected cells (1 million cells per sample) were washed 3 times with PBS to remove extracellular ascorbate and the pellet sent to Professor Vissers' laboratory for intracellular ascorbate measurement. Briefly, cells were lysed

in 50 µL of H_2O and 50 µL of 0.54 M perchloric acid with 50 mM diethylenetriamine-penta-acetic acid (DTPA) added to precipitate the protein and stabilize the ascorbate. Samples were centrifuged to remove the protein followed by measurement of the supernatant ascorbate content using HLPC-ECD (Aqua 5-µM C18 column and ESA Coulochem II detector; mobile phase of 80 mM sodium acetate, pH 4.8, containing 0.54 mM DTPA) (Kuiper & Vissers, 2014).

3.2.7 Global low coverage methylation sequencing (PBAT).

Whole genome bisulfite sequencing (WGBS) is a technique that can be used to assess DNA methylation across the genome. It typically requires 5 µg DNA and large sequencing runs, with large numbers of aligned reads, in order to generate sufficient read depth for analysis.

Post-bisulfite adaptor tagging (PBAT) was developed as an alternative novel and cheap method that requires sub-microgram quantities of DNA (< 100 ng). PBAT is a highly efficient procedure to create libraries for whole genome bisulfite sequencing. PBAT analysis was performed by Dr Tim Hore's laboratory (Department of Anatomy, University of Otago).

Preparation of PBAT libraries was undertaken using the method adapted from Peat et al., (2014). Briefly, 100 ng of purified DNA was subjected to bisulfite conversion, followed by heat denaturation step for additional fragmentation. Then, bisulfite-converted DNA underwent synthesis of the first strand with a biotin–labelled adapter sequence possessing seven random nucleotides at its 3' end (BioP5N7, biotin-ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNN), followed by purification using streptavidin-coated Dynabeads (Thermo, 11205D) and magnetic immobilization. Immobilized first strand was used as a template for an additional adapter that possess seven random nucleotides at its 3' (P7N7, GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNN) to create double-strand DNA. PCR was used to add a unique molecular barcode and adapter sequences necessary for binding to Illumina flow cells using KAPA HotStart Uracil+Mix. Libraries were sequenced on a MiSeq instrument (Illumina) and at least 15,000 mapped CG calls were attained per sample (Peat et al., 2014).

3.2.8 Global hydroxymethylcytosine quantification.

Quantification of global 5-hydroxymethylcytosine was done using The MethylFlash™ Global Hydroxymethylation (5-hmC) ELISA Easy Kit (Epigenetek), according to manufacturer's protocol. The protocol is summarised in 3 steps: 1) 100 ng of sample DNA was bound to a high DNA affinity treated wells; 2) the hydroxylated portion of DNA was detected using a 5hmC

mAb-based detection complex; 3) the samples were quantified calorimetrically by reading the absorbance in microplate spectrophotometer.

In brief, the working buffer and solution were prepared according to kit protocol. Standard curve was prepared by diluting 1 µL of positive control (PC) with 9 µL negative control (NC), and preparing solutions of different concentrations by combining PC, diluted PC and NC according to the chart below (Table 3.10). Standard curve was generated by plotting the optical density OD values at 450 nm versus the PC at each percentage point and the slope of the standard curve was determined using linear regression. The percentage of hydroxymethylated DNA is proportional to the OD intensity measured, and the 5-hmC standard curve can be seen in chapter seven (Section 7.4.2). All DNA samples were measured in duplicate (Table 3.11).

Table 3. 10. Positive control concentrations for preparation of standard curve.

Control		PC (1.0%) (µL)		Diluted PC (0.1%) (µL)		NC (µL)
0.02 % PC/well	=	0.0	+	1.0	+	9.0
0.04 % PC/well	=	0.0	+	1.0	+	4.0
0.10 % PC/well	=	0.0	+	3.0	+	3.0
0.20 % PC/well	=	1.0	+	0.0	+	9.0
0.40 % PC/well	=	1.0	+	0.0	+	4.0
1.00 % PC/well	=	3.0	+	0.0	+	3.0

Table 3. 11. The configuration of the strip well-plate.

Well #	Strip 1	Strip 2	Strip 3	Strip 4	Strip 5	Strip 6
A	NC	0.02%PC	Sample 2	Sample 6	Sample 10	Sample 14
B	NC	0.02%PC	Sample 2	Sample 6	Sample 10	Sample 14
C	0.02%PC	0.40%PC	Sample 3	Sample 7	Sample 11	Sample 15
D	0.02%PC	0.40%PC	Sample 3	Sample 7	Sample 11	Sample 15
E	0.04%PC	1.0%PC	Sample 4	Sample 8	Sample 12	Sample 16
F	0.04%PC	1.0%PC	Sample 4	Sample 8	Sample 12	Sample 16
G	0.10%PC	Sample 1	Sample 5	Sample 9	Sample 13	Sample 17
H	0.10%PC	Sample 1	Sample 5	Sample 9	Sample 13	Sample 17

5-hmC % Calculation

$$5\text{-}hmC\ \% = \frac{Sample\ OD - NC\ OD}{\text{Slope}\ x\ \text{S}}\ x\ 100\ \%$$

S: is the amount of input sample DNA.

Chapter Four: Establishing and optimising sequencing tools and bioinformatics protocols.

4.1 Introduction.

The hairpin bisulfite sequencing technique is a novel developed method used to determine the methylation pattern of two complementary DNA strands together. Hairpin linkers were specifically designed with a unique molecular barcodes to ligate the densely methylated *RASSF1* and *PCDHGA12* promoter regions. The methylation patterns for these genes were assessed by using bisulfite treatment followed by PCR amplified with barcoded index adaptors. Illumina MiSeq sequencing was used to sequence the hairpin molecule using MiSeq Reagent Kit v2 that provides 500 bp reads (paired end x 250 bp). The indexing allowing 200 samples to be sequenced in parallel, and the sequencer yields hundred thousand reads per sample.

To process and analyse the MiSeq data and to compare the methylation status of complementary CpG sites, several bioinformatic tools were used. Most of the tools are publicly available and extensively used like RStudio, Galaxy software, Pear merger, and FASTQC. To facilitate the data analysis we contributed to the development of an automated workflow to improve the identification of unique reads for each hairpin sample. In addition, we have developed a script run in macOS that allowed the automated piping of command-lines programs together, which replaced the user-intensive Galaxy pipeline. We also optimised a new method for comparing the methylation status of complementary CpG sites using RStudio.

4.1.1 The aim of this chapter.

To document the development of the bioinformatic workflow and to demonstrate the reproducibility of the hairpin bisulfite protocol.

Specific aims are:

A- Design automated workflow for hairpin bisulfite analysis.
B- Confirm the reproducibility of the hairpin bisulfite sequencing protocol.
C- Process hairpin methylation data to measure the percentage of hemi-methylation.
D- Optimise scripts to calculate the percentage of methylated/hemi/un-methylated hairpin reads in each sample.

4.2 High throughput sequencing (Illumina MiSeq).

4.2.1 Library multiplexing.

This part will detail the multiplexing and tools used to analyse the hairpin bisulfite sequences. Firstly, multiplexing allows pooling and sequencing of large numbers of libraries (~ 350 libraries) in a single sequencing run. Multiplexing of libraries is possible due to unique barcoded index sequences added during library preparation of each DNA sample, such that each read can be identified and assigned to its originating sample. With multiplexing, the MiSeq sequencing time for multi-sample studies can be considerably reduced. Multiplex sequencing is gaining increasing interest because it is cost effective, large numbers of amplicons can be assessed, library preparation is faster and high quality data is obtained.

For library construction, *RASSF1* and *PCDHGA12* hairpin DNA libraries were prepared for each sample using our modified protocol (details in Chapter three). Two rounds of multiplexing occurred during library preparation; the first level of multiplexing mixed all amplicons for each sample (multiplexed amplicons *RASSF1* and *PCDHGA12*), and the second level of multiplexing mixed many samples in one pooled library.

4.2.2 Sequencing of hairpin libraries.

An Illumina MiSeq System (highly automated workflow) was used to perform sequencing for multiple genes of multiplexed libraries within a single run with high yields of data. Before sequencing, libraries were quantified and validated using an Agilent 2100 *Bio Analyser* (Agilent Technologies). Quantified and pooled libraries were contained with samples from other groups, and sequencing was carried out on the MiSeq System (Illumina) using MiSeq Reagent Nano kit V2 500 bp to generate 2 x 250 bp paired end reads for each sample according to manufacturer's instructions (Quail et al., 2009).

4.2.3 Illumina MiSeq principles.

MiSeq sequencing provides information about the methylation status of many hairpin sequences by allowing parallel reading and outputting hundreds of thousands reads for each sample. Template DNA is hybridised to a small piece of single-stranded DNA (adapter) that is attached to the polymer-coated glass surface of a flowcell (Bronner et al., 2014) (Figure 4.1). Templates are amplified by flowing enzymes and reagents through the channel of flowcell creating a cluster of identical sequences for each.

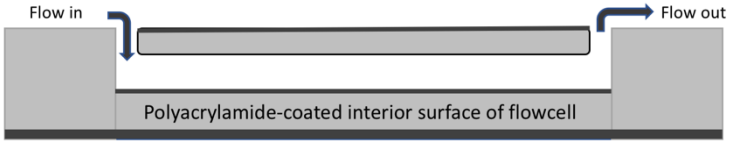


Figure 4. 1. Diagram of Illumina flowcell. Template DNA and reagents flow through a hollow glass slide. Modified from Bronner et al., 2014.

Sequencing was performed by Dr Rob Day at the Centre for Translational Cancer Research, Department of Biochemistry, University of Otago, Dunedin.

4.2.4 Libraries demultiplexing.

The sequence data was then demultiplexed based on the Illumina index sequence and saved as two sequence files (fastq.gz files) for each sample, one for each reading direction of the fragment (read 1 and read 2 files). Both files contain the same number of reads. Quality metrics of the sequencing data was then manipulated using web-based platform Galaxy (section 4.3.1).

4.3 Raw MiSeq data processing.

To process and analyse the MiSeq data, several publicly available programs were used. In addition, we contributed to the development of a framework to facilitate the methylation analysis of the complementary CpG sites of hairpin data. A brief overview of the tools and software used in this study is described in the following sections.

4.3.1 Galaxy platform and workflow.

Galaxy is an open source web-based application for interactive data analysis (https://usegalaxy.org). The system provides a base on which to build multiple analyses steps through a graphical user interface (Calvanese et al., 2012). Data can be imported to Galaxy from the user's computer or directly from many online resources such as the UCSC table browser. The application provides a collection of command line tools (FASTX-Toolkit) which are used to process Next Generation Sequencing (Pennings et al.) data.

This section summarise the stages used to process the raw hairpin MiSeq sequence data. We established a bioinformatics framework for hairpin DNA methylation analysis (Figure 4.2). The online Galaxy software was used to process the raw files, forward read R1 and reverse read R2.

1. Paired-end fastq.gz files produced by the Illumina MiSeq sequencer were uploaded onto the Galaxy server (usegalaxy.org) and converted to Illumine 1.8+ format by the 'FASTQ Groomer' function (Blankenberg et al., 2010).

2. The barcode splitter function was used to select gene specific reads. Forward and reverse primers were used to split gene specific sequences in read 1 and read 2 files respectively.
3. Barcode-split files were downloaded and checked for quality analysis per-base using FASTQC function, then uploaded files were groomed to Illumine 1.8+ format using FASTQ Groomer tool .
4. Filter by quality function was used to select the reads with PHRED score >30 with a cut-off 80% of read length. The PHRED score measures of the quality of the nucleobases generated by DNA sequencing and calculates the errors of base calling in the sequenced nucleotide.
5. If the amplicon size is more than 250 bp and less than 500 bp "Trim sequences function" were used for trimming the reads to make the size of both reads equal to the full length of sequenced amplicon.
6. Read 2 was modified by "Reverse Complement".
7. Read 1 and Read 2 were joined using "*FASTQ Joiner*" function.
8. The unique reads in FASTA file were selected by "Collapse" function, and PCR replicates were removed.
9. FASTA files were then imported to *BiQ Analyzer* for visualisation.

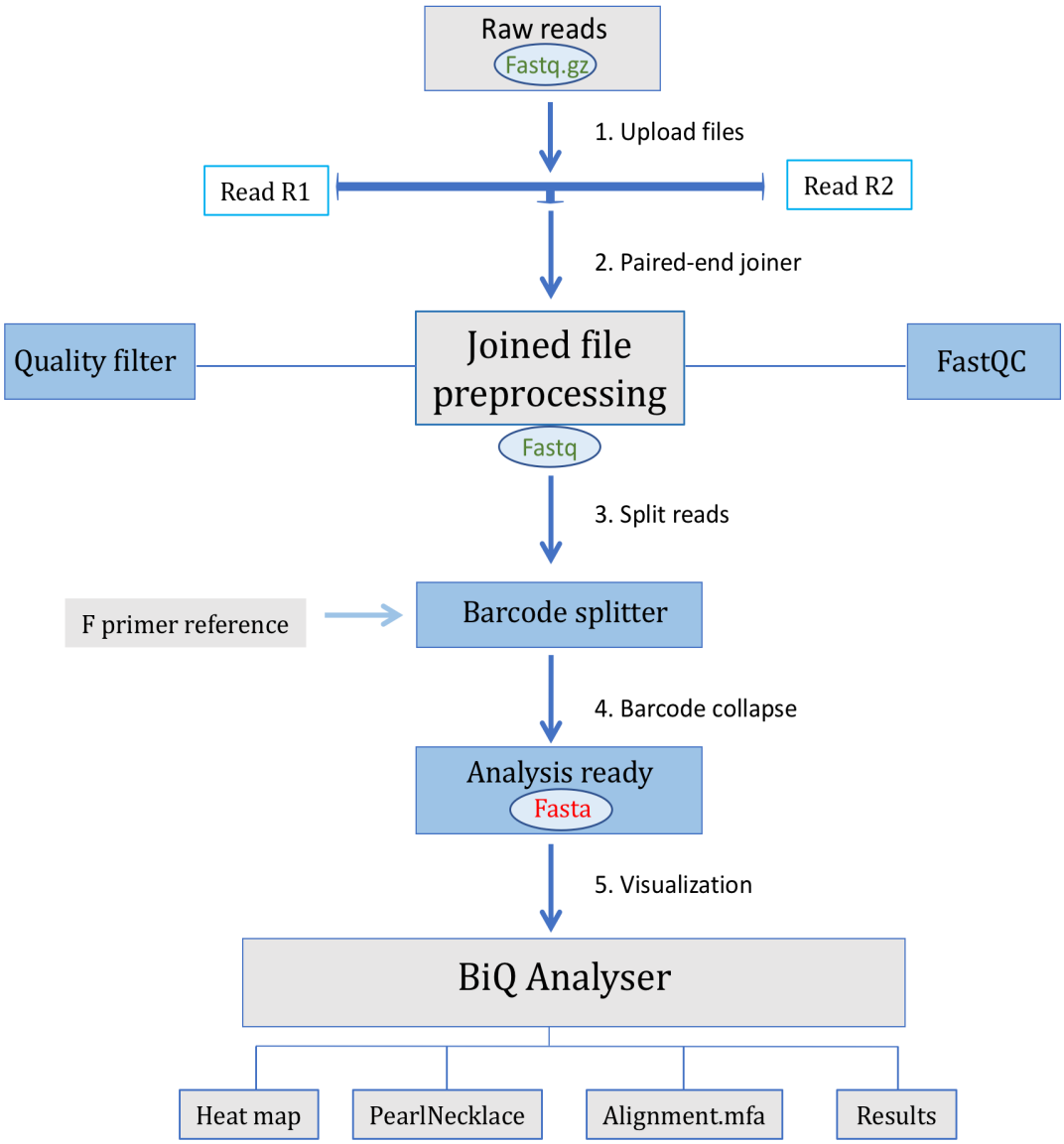


Figure 4. 2. Galaxy workflow diagram shows the main steps used in processing the MiSeq sequence data. Step 1) MiSeq sequence output R1&R2 were uploaded to Galaxy online platform. Step 2) Paired-end sequence files were joined to generate the joined file (FASTQ), then FASTQC was used to check the quality of the joined sequence, followed by quality filter to remove the poor-quality sequences. Step 3) '*Barcode splitter'* on the joined file with forward primer reference. Step 4) Barcode collapser was used to remove repeated sequence based on unique hairpin barcode and output the collapsed reads in FASTA format. Step 5) FASTA files were imported to '*BiQ Analyzer'* to facilitate locus-specific analysis and visualisation of high throughput hairpin bisulfite sequence.

4.3.2 Multiple mergers used to join paired-end reads.

When the amplicon size is more than the length of individual read and smaller than twice the length of the single-end reads, the corresponding paired-end reads can be merged into a fragment to get the full length of the sequence (Figure 4.3). The overlapping sequence between the two reads can be deployed for correcting sequencing errors and yield higher quality sequences. Several packages are available to merge paired-end sequencing reads and these include PEAR (Zhang et al., 2014), FLASH (Magoč & Salzberg, 2011), PANDAseq (Masella et al., 2012) and COPE (Liu et al., 2012). Two of the joiner packages (*PEAR* and *FLASH*) were tested in this study.

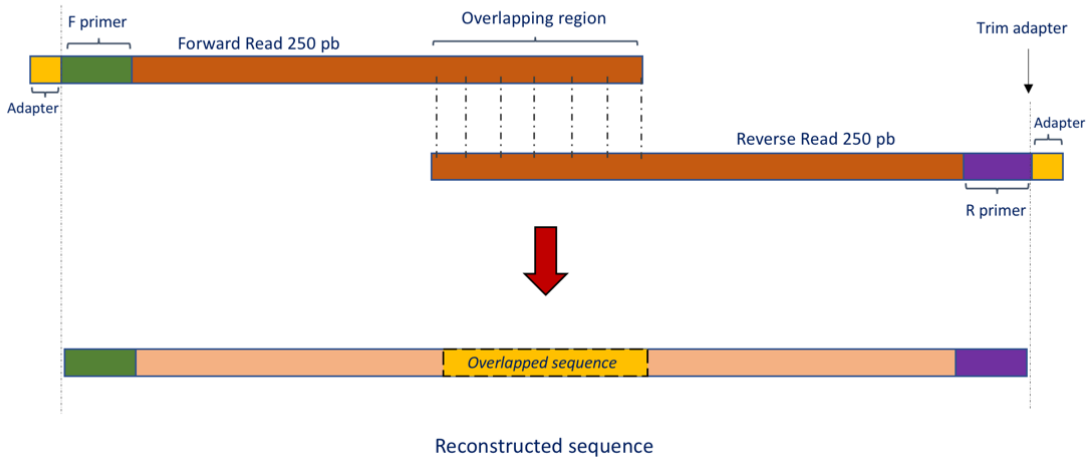


Figure 4. 3. Typical scenario of read overlapping. Forward reads (250 bp; R1) and reverse reads (250 bp; R2) overlap.

4.3.2.1 'PEAR' – Paired-End reAdmergeR (v0.9.8).

'*PEAR*' is a highly accurate paired-end merger designed to join corresponding reads based on the unique sequence identifiers (Zhang et al., 2014). *PEAR* evaluates all possible overlap sequences using a minimum overlap size of 10 bp and minimum possible length of the assembled sequences of 50 bp. *PEAR* determines highest alignment score by recording all possible overlaps for each pair of corresponding paired-end reads. Using this method, the sequence accuracy was enhanced as *PEAR* compares the base quality scores for the overlapping regions. Therefore, PEAR can merge sequences with varying fragment length. *PEAR* uses statistical testing to generate a P value for each overlap and to minimise the likelihood of false positive merging (Zhang et al., 2014). Reads that could not be joined were removed.

4.3.2.2 '*FLASH*' – Fast Length Adjustment of Short Reads (v1.2.6)

'FLASH' is another paired-end merger used to extend the short sequence length from next-generation sequencing experiments by overlapping paired-end reads of short fragments. *FLASH* requires compatible input fastq paired-end reads where the reads with a unique sequence identifier overlap the read generated from the opposite end of the same DNA sequence. *FLASH* joiner processes each pair of reads separately and searches for the correct overlap between the paired end reads (Magoč & Salzberg, 2011). There are a few limitations of the *FLASH* program: 1) *FLASH* failed to merge the majority of reads with short overlap sequence, since it requires a minimum overlap value to work correctly. 2) *FLASH* is slow when tested on a larger dataset of paired-end reads from human chromosome 14 compared to *PEAR* program. (Zhang et al., 2014).

4.3.3 Splitting reads by gene specific primers ('*FASTQ Barcode splitter*').

Barcode splitter function was used to demultiplex gene-specific reads from the pooled sequence output of high throughput sequencing. This third step of data pre-processing splits reads by the gene specific primer sequence to allow identification of the origin of the reads. The barcodes here are different than the hairpin barcodes used to identified the hairpin sequences. The barcode files are input text files that contain amplicon name and primer sequence used to split Illumina sequencing library (FASTQ file) to multiple files. This tool used forward primer sequences (with a maximum of 2 mismatches) as identifiers to match the sequences that belong to these primers. As a result, the sequences matching the forward primer sequence will be stored in FASTQ files for "matched files" created for each primer. Sequences not matching the barcode sequence will be stored in additional FASTQ files created for the un-matched sequences "un matched files" (Babraham Bioinformatics).

4.3.4 FASTQC (Babraham Bioinformatics).

FASTQC (www.bioinformatics.babraham.ac.uk/projects/fastqc) is a JAVA application used to check the quality of the raw sequence data coming from high throughput sequencing. FASTQC provides a summary of quality assessment of each file including per-base and per-sequence quality scores (better known PHRED score) (Figure 4.4). PHRED is used to assess sequencing quality by estimating the sequence call error; high quality reads are defined as having a PHRED score >30, representing a greater than '1 in 1000' probability of an incorrect base call; the higher the PHRED score the more confident the base calling. The Y-axis on the graph shows PHRED quality scores, the background of the graph divided into three parts according to the base call quality; 1) green, very good quality with PHRED score of 28 to 38; 2) orange, medium quality with PHRED score of 20 to 28; and 3) red, poor quality with PHRED score below 20 (Cock et al., 2010).

4.3.5 '*Collapsing*' (removing redundant reads).

In bisulfite sequencing, an excessive number of PCR cycles can result in over-amplification of sequences with the same hairpin barcode. To deal with the effects of PCR amplification, the tool '*collapsing*' was used to "collapse" the duplicate sequences into a single sequence. The hairpins incorporate random 14 bp barcodes with four million unique sequences by which replicated or contaminating sequences can be recognised and excluded. Since each sample will have up to only 10,000 reads, it is not expected that the same barcode will be observed, unless this was due to PCR amplification.

The Galaxy platform provides a collapsing function tool that compares the reads along their full length, and retains only one identical sequence per sample. The Galaxy collapse tool depends on the entire sequence of the reads (including any errors introduced during PCR or sequencing). The Galaxy collapse tool might recognise sequences as distinct reads in spite of having the same hairpin barcode. After collapsing of the reads, inspection of the barcodes strongly suggest that these sequences are not unique (Figure 4.5) (Figure 4.6).

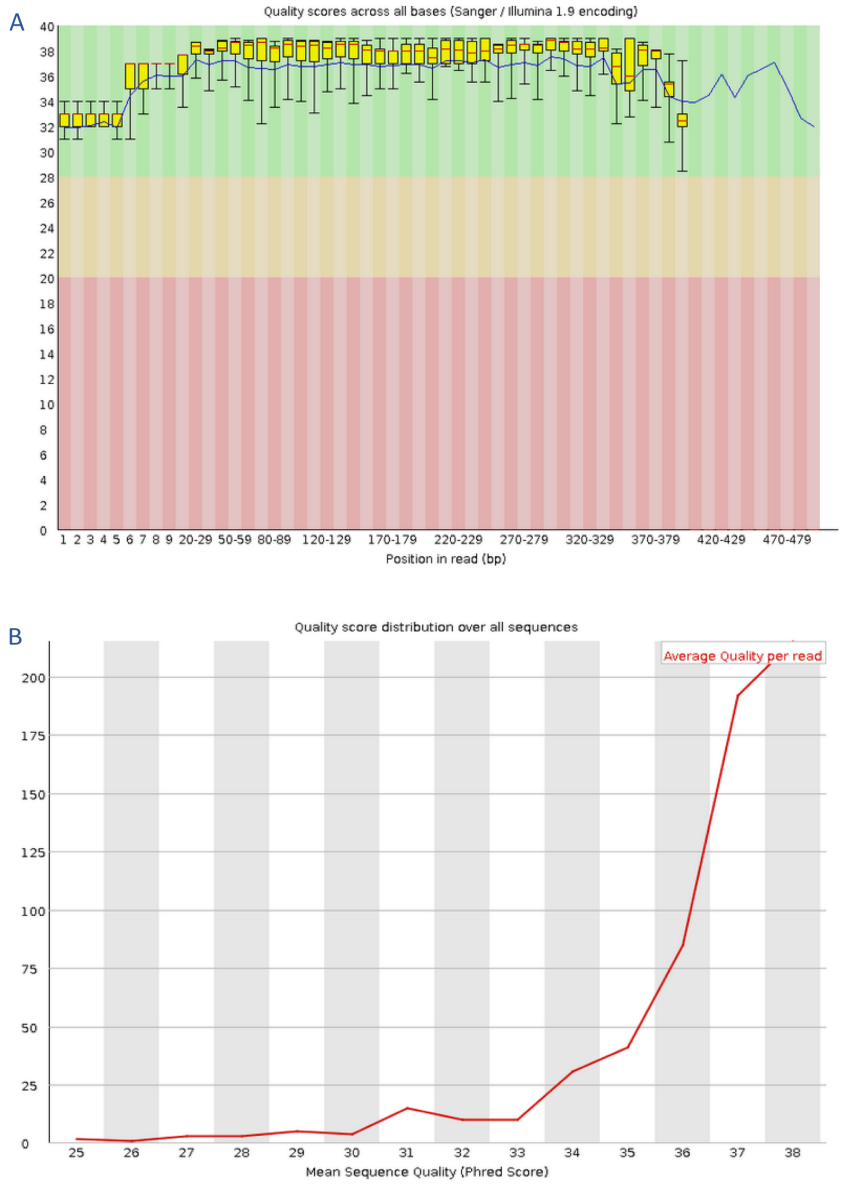


Figure 4. 4. Quality metrics interface for Illumina MiSeq data. A) FASTQC per-base sequence quality chart. The x-axis shows the base position in the reads, and the quality score is on the Y-axis. This box and whisker plot shows the quality scores for each base within the merged sequence file calculated as mean PHRED score (blue line), median PHRED score (red line), and interquartile range (yellow box). The upper and lower whiskers indicating the 10^{th} and 90^{th} percentile scores. The plot background is colour-coded to determine the good (green), acceptable (beige) and bad (red) quality scores. B) Per sequence quality assessment (PHRED score), the chart displays the number of read sequences that have the same mean sequence quality.

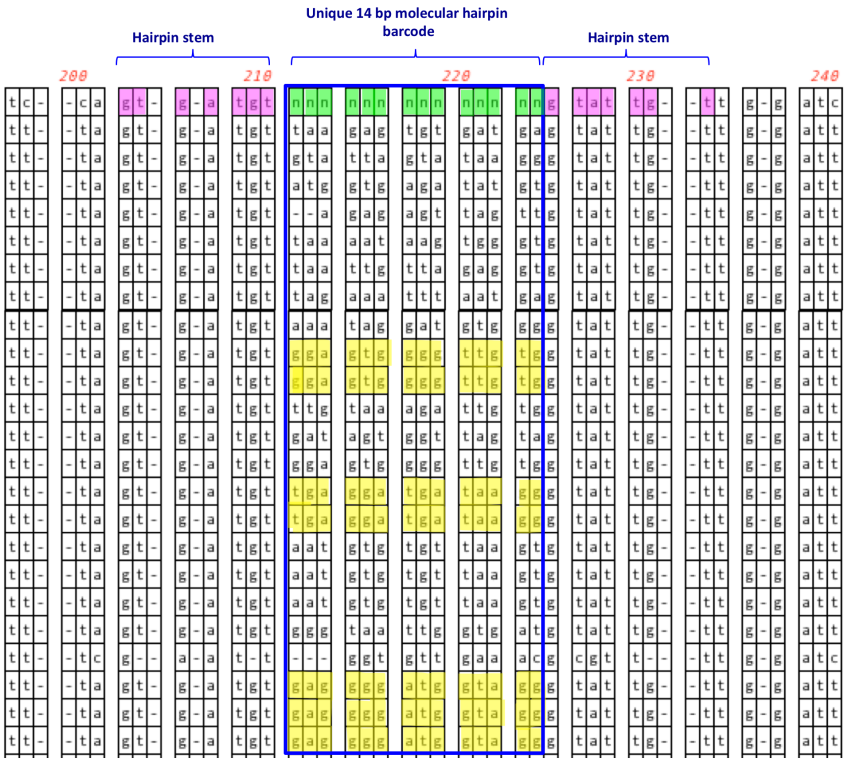


Figure 4. 5. FASTQ collapsing. Multiple *PCDHGA12* hairpin sequences from the alignment file visualised a different hairpin barcode for each sequence (labelled with blue rectangle, and the hairpin reference sequence labelled in green) with the constant region of the hairpin (GTGATGT located before the hairpin molecular barcode, labelled with purple colour). Each row represents different hairpin reads with a unique molecular barcode. A few hairpin reads share the same barcode (highlighted in yellow).

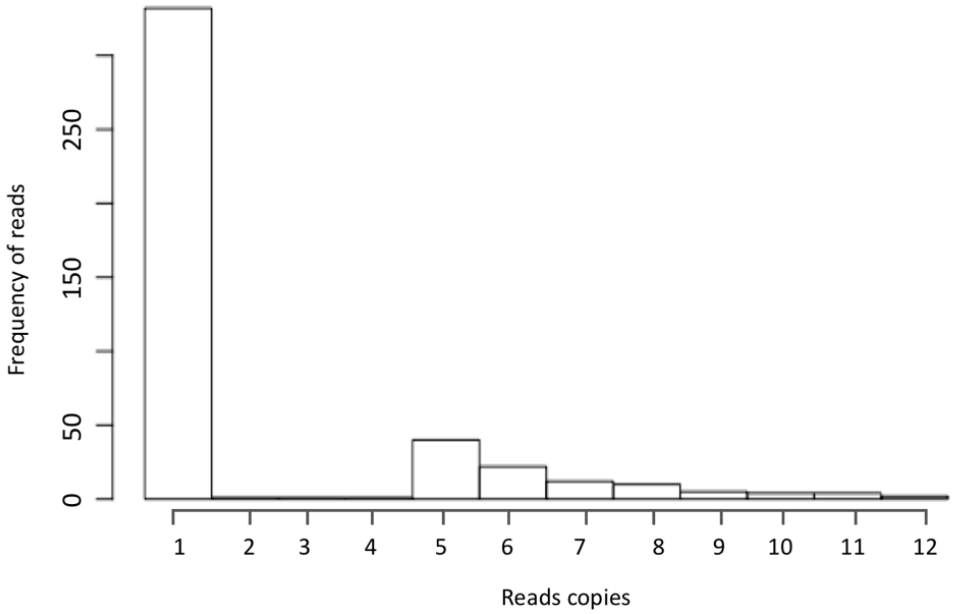


Figure 4. 6. Redundancy of hairpin sequences following FASTQ collapser. Most of the hairpin reads contain one unique molecular barcode. Few hairpin reads shared the same barcode after Galaxy *FASTQ collapser*.

In a fully collapsed file, each read should be repeated once and contain a unique hairpin molecular barcode. The nature of a unique hairpin present in our DNA sequences provides a possible pathway to search for the constant region of the hairpin (located before the 14 bp unique molecular barcode) linked to DNA fragment (Figure 4.7).

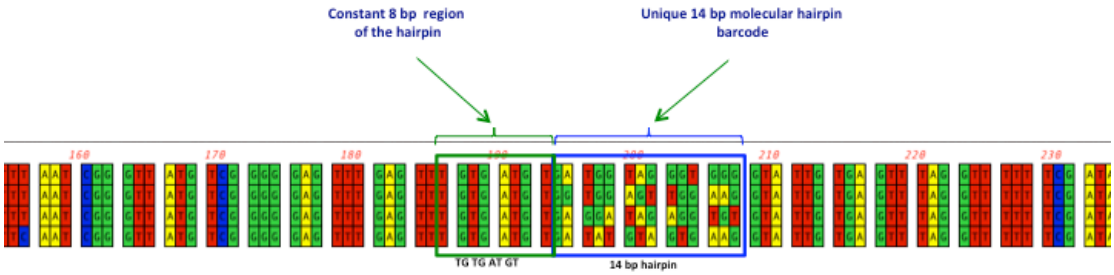


Figure 4. 7. Multiple PCDHGA12 hairpin sequences from a fasta file showing the different hairpin barcode for each sequence (blue rectangle) with the constant stem of the hairpin (TGTGATGT located before the hairpin molecular barcode, green square). Each row represents a different read with a unique molecular barcode.

4.3.6 Hairpin barcode collapsing (remove redundant barcodes) using UNIX command line.

To solve the incomplete collapsing of the hairpin sequences resulted from the online Galaxy *fastQ collapser*, a bash script command was developed to search for the hairpin stem sequence and to write a FASTA sequence file containing only the collapsed sequences with a unique barcode (one barcode for each sequence), ready for further analysis. The bash script used to collapse the *PCDHGA12* sequencing files is reproduced below.

```
for filename2 in *.fastq;
      do
            awk '{if(NR%4==2) print;}' < $filename2 |
#Prints the second of every 4 lines in the fastq file - this line contains the
sequence.

            awk'{sub(/GTGATGT|.TGATGT|G.GATGT|GT.ATGT|GTG.TGT|GTGA.GT|GT
            GAT.T|GTGATG./,".")}; 1'|
#Allowing one mismatch, searches for the constant region of the hairpin immediately
prior to the molecular barcode, and replaces it with a ".".

            sort -t . -k 2.1,2.14 -u |
#Sort each file based on the first 14 characters after the "."; and keep one unique
read per match.

            cat -n | sed 's/^/>/' |
            tr "[\t]" "\n" > ${filename2/%.fastq/.fasta};
            rm ./*unmatchedf.fasta;
#Numbers each line, then inserts a ">" before every line number, then replaces
every tab "\t" with a newline "\n" to create a .fasta format file.

      done);
done &&
```

For the *RASSF1* sequencing files, the above bash script was amended by replacing the *PCDHGA12* hairpin stem sequences with the *RASSF1* hairpin stem sequences: GTGATGT|.TGATGT|G.GATGT|GT.ATGT|GTG.TGT|GTGA.GT|GTGAT.T|GTGATG

However, using these *RASSF1* stem sequences resulted in truncated, collapsed reads, as two of these (GTGA[A]GT) and (GT[T]ATGT) hairpin stem sequences are located in other parts of *RASSF1* read sequences (Figure 4.8).

```
TAGTTTTTGTATTTAGGTTTTTATTGTGTGGTTTTTTTTAGTTTTTTTTTGTTGTTTAGTT
TGGATTTTGGGGGAGGTGTTGAAGTTGGGGTTTGTTTTGTGGTTTTGTTTGGTTTGTGTTT
GTTAGTGTTTAAAGTTAGTGA[A]GTATGGGTTTAATTGGGT[T]ATGTTGGGGGAGTTTG
AGTTTGTGATGTNNNNNNNNNNNNNNGTATTGTGAGTTTAGGTTTTTTTGATATGGTTTGG
TTGGGTTTGTGTTTTGTTGGTTTTGGGTGTTAGTAAGTGTGGGTTGGGTGGGGTTATAGGG
TGGGTTTTGATTTTAGTGTTTTTTTTAGGATTTAGATT
```

Figure 4. 8. Bisulfite converted sequence of RASSF1 hairpin molecule. Two of the sequences mismatched (GTGA[A]GT) and (GT[T]ATGT), highlighted yellow and green.

The presence of these mismatches was corrected by adding a T to the beginning of the hairpin constant sequence.

Mismatch with extra T at start:

TGTGATGT|.GTGATGT|T.TGATGT|TG.GATGT|TGT.ATGT|TGTG.TGT|TGTGA.GT|TG TGAT.T|TGTGATG.|

With the extra T at the start of the mismatched sequences, there are now no truncated sites.

```
for filename2 in *.fastq;
      do
            awk '{if(NR%4==2) print;}' < $filename2 |
#Prints the second of every 4 lines in the fastq file - this line contains the
sequence.

awk'{sub(/TGTGATGT|T.TGATGT|TG.GATGT|TGT.ATGT|TGTG.TGT|TGTGA.GT|TGTGAT.T|T
GTGATG./,".")}; 1'
#Allowing one mismatch, searches for the constant region of the hairpin immediately
prior to the molecular barcode, and replaces it with a ".".

            sort -t . -k 2.1,2.14 -u |
#Sort each file based on the first 14 characters after the "."; and keep one unique
read per match.

            cat -n | sed 's/^/>/' |
            tr "[\t]" "\n" > ${filename2/%.fastq/.fasta};
            rm ./*unmatchedf.fasta;
#Numbers each line, then inserts a ">" before every line number, then replaces
every tab "\t" with a newline "\n" to create a .fasta format file.

      done);
done &&
```

To determine the efficiency of the developed script, and its capability to resolve the incompletely collapsed sequences, hairpin barcodes were selected from a collapsed FASTA

files, and a clustering analysis was done to show that all barcode sequences were unique and there were no groups of similar barcode sequences. The analysis was done in R using Jaro-Winkler distance method (Figure 4.9). A fully collapsed file should only show one read for each unique molecular barcode. Similar sequences apparently are PCR replicates possessing the same barcode, and the generation of these similar sequences leads to a biased heatmap, which then affect the DNA methylation percentages.

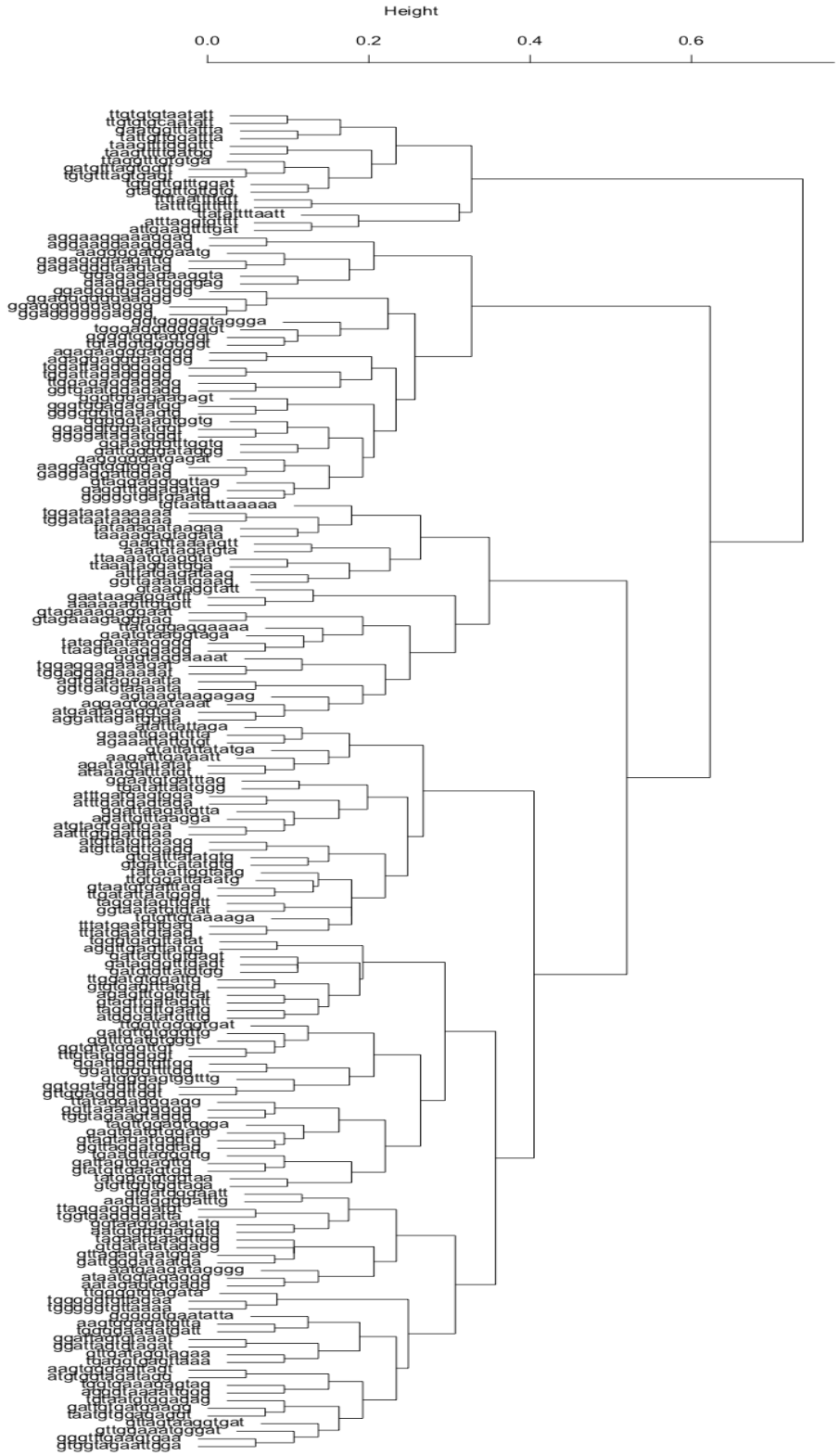


Figure 4. 9. Dendrogram showing the results of Clustering analysis for hairpin barcodes using Jaro-Winkler distance method. The clustering analysis was done to demonstrate hairpin barcode uniqueness.

4.4 High throughput data visualisation.

4.4.1 BiQ Analyzer HT.

To further analyse and visualise the FASTA files resulting from the collapse function of Galaxy, *BiQ Analyzer HT* program was used. *BiQ Analyzer* is a Java-based software tool used to process and analyse the output of high throughput bisulfite sequencing files (Lutsik et al., 2011). The application implements a data processing pipeline to generate quality-controlled output results for the inferred DNA methylation for each sample.

CG methylation is the most common epigenetic modification of eukaryotic DNA (Lee et al., 2010). By default, *BiQ Analyzer* focused on this modification by comparing the sequences at the corresponding positions of the aligned sequence read, and creates a binary methylation output directory ("1" methylated, "0" unmethylated and "X" missing value) for each sequence.

4.4.1.1 Inference of DNA methylation pattern.

BiQ Analyzer HT supports the alignment analysis of bisulfite sequencing data for many target loci in multiple samples. The program is available as a downloadable installation package and can be run on any computer with an appropriate version of Java. *BiQ Analyzer* analyses can be done using a graphical interface, or alternatively can be run from the command line with automated data processing scripts. The *BiQ Analyzer* program is designed to analyse a dataset with precise mapping to a single locus.

The galaxy and *BiQ Analyzer* workflow showing the main steps used in processing the MiSeq sequence data is shown in (Figure 4.10). To prepare for analysis,

1. High throughput bisulfite sequencing files are converted to FASTA format (Galaxy collapsed files).
2. FASTA files are imported into *BiQ Analyzer* and aligned with genomic reference sequences (Lutsik et al., 2011). Based on these alignments and by comparing the read sequence with the genomic reference sequence, the methylation status of every CpG in all sequence reads can be determined .
3. After alignment, *BiQ Analyzer* outputs DNA methylation data in several different file formats (Figure 4.10).
4. Results files can be imported into R, Excel or Prism (GraphPad) for further analysis, such as determining the frequency of hemi-methylated and un-methylated reads.

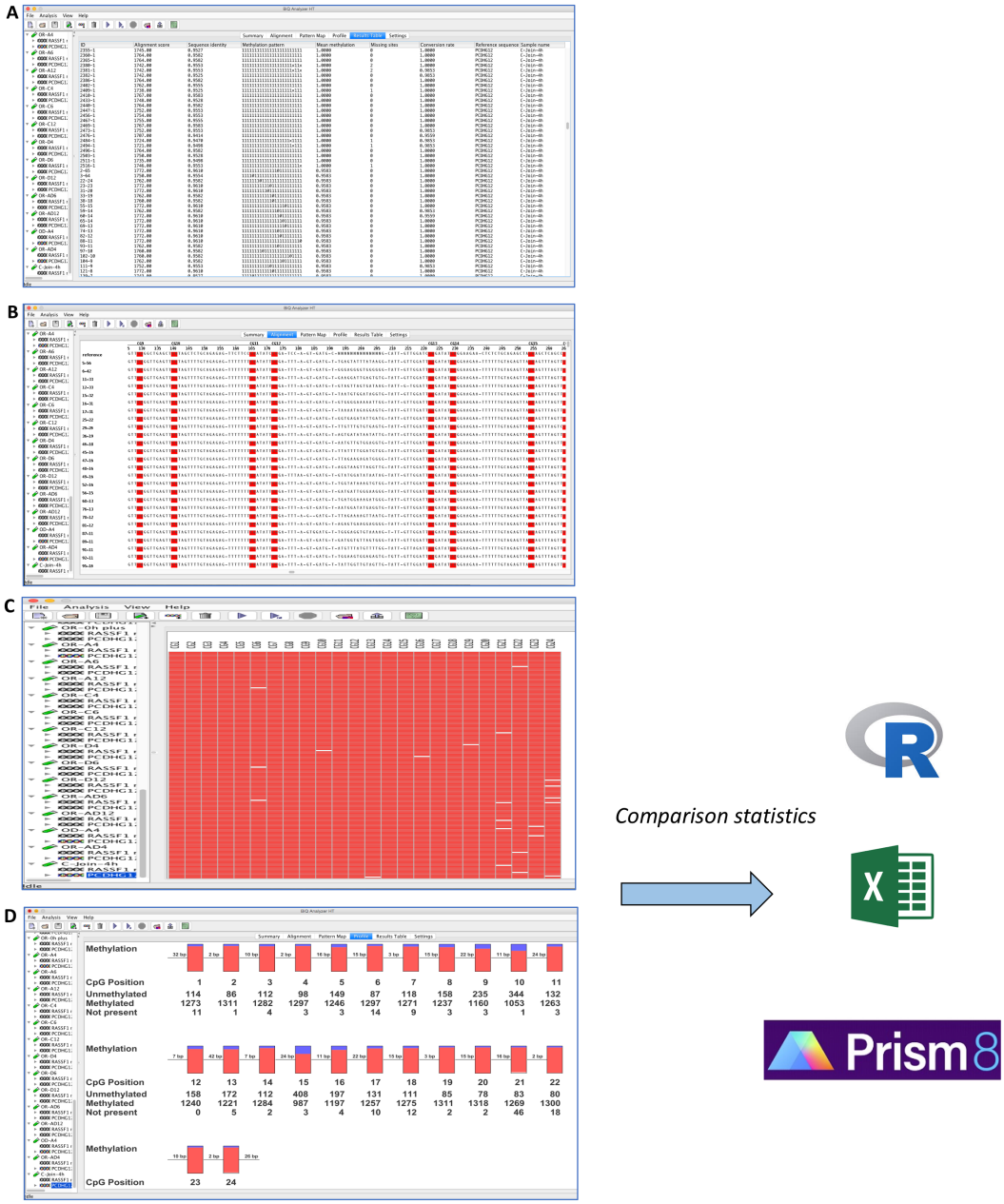


Figure 4. 10. Example of the output files produced by *BiQ Analyzer* HT. A) the results file (results.tsv) summarises all the satisfactory information for each read of demultiplexed sample after alignment, for example, alignment score, sequence identity, mean methylation score, conversion score results for each read. The binary CG- methylation output is (1 = methylated, 0 = unmethylated, X = missing value). B) Multiple alignment of sequencing reads for each amplicon in each sample. C) Example of a methylation pattern map (heatmap.png). Each row represents an individual hairpin read, while each column reflects CpG sites across different hairpin reads. Methylated CpG were labelled red, unmethylated in blue and missing value in white. D) Example of a Pearl necklace diagram (pearlNecklace.png), which summarises the methylation frequency (methylated, unmethylated and missing value) per CpG sites across all sequenced reads for each sample. Results files were uploaded to R, Excel and Prism (GraphPad) for further statistical analyses (calculate the frequency of hemi-methylated and un-methylated reads).

4.4.1.2 Alignment score inflection point.

BiQ Analyzer generates an alignment score for every read and high alignment score is essential for analysis of bisulfite sequences. To ensure only good quality reads were analysed, poorly aligned reads were removed. The modified "Needleman-Wunsch algorithm" was used to calculate the alignment score for each sequence read. The modified algorithm has highly optimised implementation and provides excellent performance for large read numbers in the order of 10^4 per CpG site (Lutsik et al., 2011).

BiQ Analyzer generated an alignment score for each hairpin read when sequences were aligned to the predicted reference sequence. The higher alignment scores reflect the sequence of interest. Cumulative frequency for reads alignment scores was calculated to determine the inflection point cut-off for each amplicon, i.e. the cut-off value obtained when the curve increases exponentially. Sequences with alignment score less than 2000 and 1800, for *RASSF1* and *PCDHGA12* respectively, were regarded as low-quality sequences and were discarded (Figure 4.11).

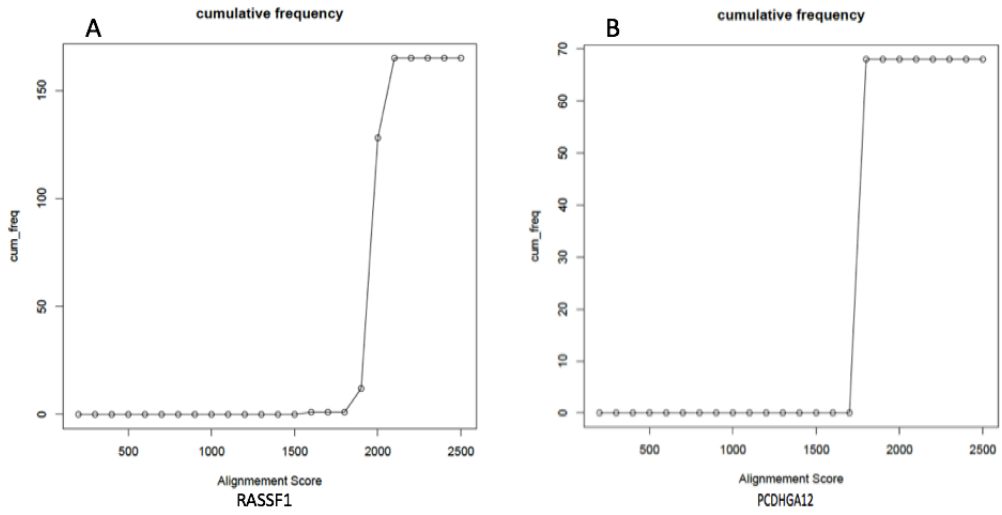


Figure 4. 11. Cumulative analysis plot for *BiQ Analyzer* alignment scores. *BiQ Analyzer* alignments score inflection point (cut-off) for A) *RASSF1* and B) *PCDHGA12* hairpin reads.

4.5 Command line sequence processing (automated analysis).

While the Galaxy workflow is reliable and successful in analysing and interpreting hairpin MiSeq data, the program depends on user-operated, online tools and manual inputs which limited the sequence processing. The data processing workflow is time consuming and repetitive with large amounts of high throughput sequence samples. In view of this limitation, the Morison laboratory group established an automated bioinformatic framework for DNA methylation analysis of the raw MiSeq sequencing data. Essentially this workflow recapitulated

the online Galaxy workflow using locally installed UNIX tools. This approach used a Mac OS terminal bash script to analyse the demultiplexed paired end fastq.gz files (read 1 and read 2). The bash script was set to perform all steps described above automatically using MacOS terminal and the output of each step was monitored to ensure accurate analysis. The full script with annotation can be found in the **Appendix 1**.

4.5.1 Quantification process and data analysis.

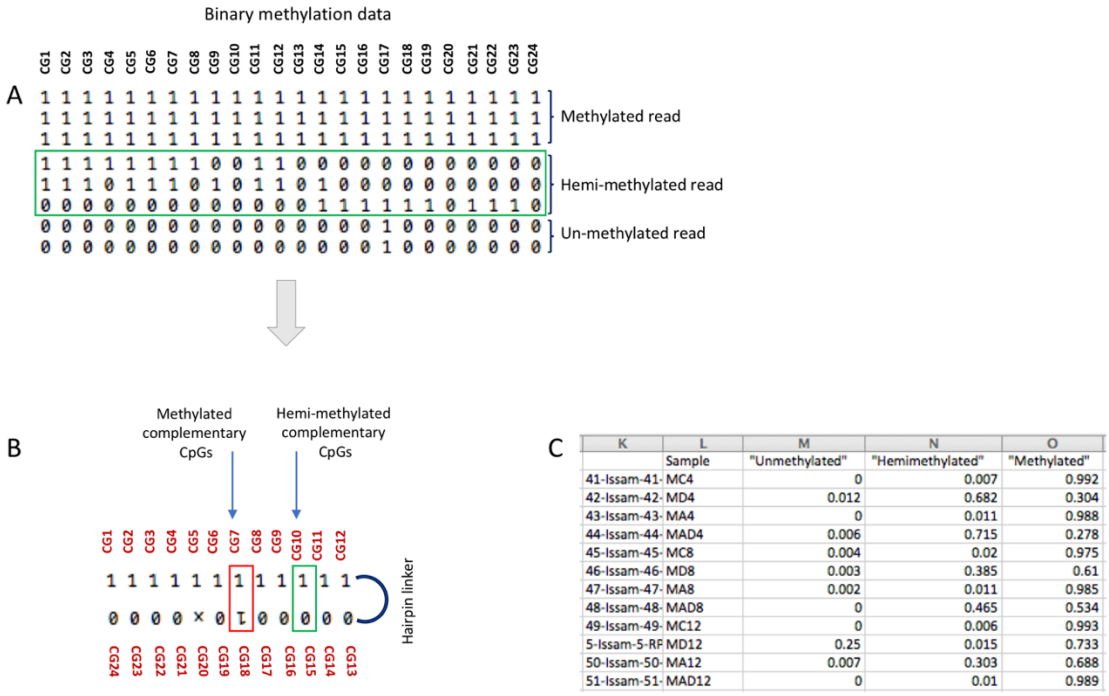
BiQ Analyzer generates a result.tsv file for each sample describing the methylation information for each read (Figure 4.10A). For calculating demethylation, it was necessary to quantify the number of meth/unmeth/hemi-methylated hairpin reads for each amplicon in each sample. The lines containing this binary methylation data ('1'- methylated, '0'-unmethylated and 'x'-undetermined) were extracted from the results.tsv file, with each character being tab-separated and output to a methylation.tsv file (Figure 4.12.A). The full terminal script with annotation can be found in the **Appendix 1**. All methylation.tsv files were uploaded to RStudio for automated analysis.

4.5.2 Processing hairpin data.

RStudio was used to process the binary methylation data (methylation.tsv files) by comparing the methylation of the complementary CpG sites across the entire hairpin strand. An R script was written to bioinformatically "fold" each read at the hairpin (Figure 4.12B) and to classify the resulting methylation pattern of complementary CpG sites as methylated, hemi-methylated or unmethylated (Figure 4.12C).

Each read was classified according to the status of complementary CpG sites: if the proportion of complementary CpGs was ≥ 0.51 the read was called 'methylated', 'hemi-methylated' if the proportion of hemi-methylated CpGs was ≥ 0.51 and 'unmethylated' if the proportion of unmethylated CpGs was ≥ 0.51. Finally, the script calculated the overall proportion of fully methylated, hemi-methylated and unmethylated reads for each sample (Figure 4.12). The full R script with annotation can be found in the **Appendix 2.**

One of the limitations in calculating the overall methylation proportion of reads was discovered for samples in which the proportion of methylated/hemi-methylated or unmethylated was (0.5:0.5) or (0.33:0.33:0.33). The proportion of these reads was negligible (< 2%) and they were discarded from analysis.



	Sample	"Unmethylated"	"Hemimethylated"	"Methylated"
41-Issam-41-	MC4	0	0.007	0.992
42-Issam-42-	MD4	0.012	0.682	0.304
43-Issam-43-	MA4	0	0.011	0.988
44-Issam-44-	MAD4	0.006	0.715	0.278
45-Issam-45-	MC8	0.004	0.02	0.975
46-Issam-46-	MD8	0.003	0.385	0.61
47-Issam-47-	MA8	0.002	0.011	0.985
48-Issam-48-	MAD8	0	0.465	0.534
49-Issam-49-	MC12	0	0.006	0.993
5-Issam-5-RF	MD12	0.25	0.015	0.733
50-Issam-50-	MA12	0.007	0.303	0.688
51-Issam-51-	MAD12	0	0.01	0.989

Figure 4. 12. Processing hairpin DNA data using terminal commands and R scripts. A) Binary methylation data ('1' and '0') extracted from 'results.txt' file and output as a 'methylation.tsv' file. B) Folding of the binary data to calculate the status of complementary CpG methylation sites. C) The 'output.tsv' file generated by the R script showing the proportions of methylation ('unmethylated', 'hemi-methylated' and 'methylated') for each sample.

4.5.3 Folding the hairpin reads using RStudio.

To visualise the methylation status of the hairpin reads, a heat map was generated using *BiQ Analyzer* to represent the methylation pattern. Each row in the heat map represents an unfolded hairpin read (Figure 4.13, left); columns represent CpG dinucleotides; blue squares show unmethylated CpGs; red squares show methylated CpGs; white squares show missing sequence information. An R script was written using ggplot in RStudio to "fold" each hairpin read, and generate a heat map to show the methylation of two complementary CpGs across hairpin reads (Figure 4.13, right). The full R script with annotation can be found in the **Appendix section.**

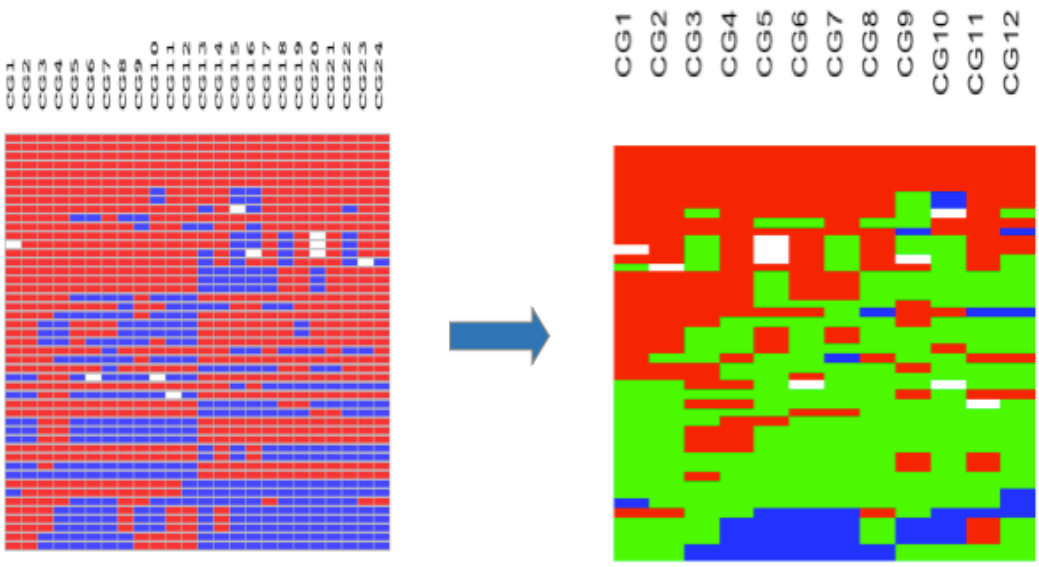


Figure 4. 13. *BiQ Analyzer* output showing the methylation pattern of the PCDHGA12 promoter. The left panel shows the results of the methylation analyses by deep bisulfite sequencing. Rows represent unfolded hairpin reads; columns represent CpG dinucleotides across hairpin reads; blue squares show unmethylated CpGs; red squares show methylated CpGs; white squares show missing sequence information. The right panel shows the methylation status of the "folded" hairpin reads, demonstrating the methylation of the two complementary CpGs at each of 12 CpG sites; red show methylated CpGs; blue show unmethylated CpGs; green show hemi-methylated CpGs; white show missing sequence information.

4.5.4 Comparison of methylation mapping between *PEAR merger* and *FASTQ merger.*

The efficiency and fidelity of different sequence joiner programs were investigated using a single sample (sample treated with decitabine for 8 h after release from thymidine block). The DNA methylation pattern by the two merger programs were compared. The figure below shows different heatmaps pattern of hairpin *PCDHGA12* promoter after using different read joiners (*PEAR* and *FASTQ*). *Fastq* merger performs a blunt join based on the unique sequence identifier (Blankenberg et al., 2010), which resulted in loss of some of the aligned CpG sites at the middle of the hairpin sequences which leads to short incomplete sequences compared to a full aligned sequences joined with *PEAR* merger.

After alignment by *BiQ Analyzer*, multiple tools were used to compare different methylation sorting across a single sample joined with two different joiners *PEAR* and FASTQ. The methylation patterns of these heatmaps depends on whether the data are sorted by alignment score, methylation level or number of missing sites tools (Figure 4.14).

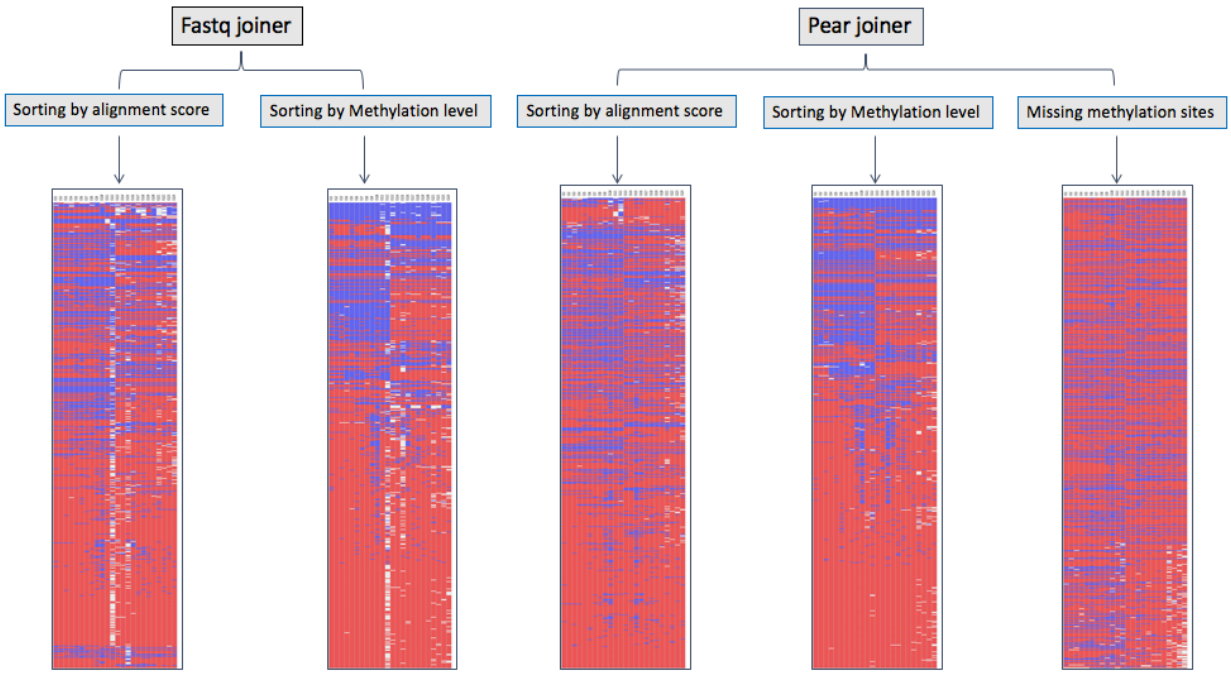


Figure 4. 14. Comparison of joiner performance. The paired reads from a single sample sequences were joined with two different joiners (*FASTQ* & *PEAR*). The heatmaps show different methylation patterns of the *PCDHGA12* promoter.

4.6 Reproducibility of MiSeq sequencing platform.

Illumina MiSeq sequencing is a popular next generation sequencer due to high accuracy, high throughput and great flexibility. The MiSeq method introduced a new type of library preparation protocol for NGS that require only small amounts of input 1 ng DNA for library construction. Illumina provided paired-end sequencing of both ends of the DNA and aligning the forward and reverse reads as paired reads.

However, numerous technical problems such as PCR bias or sequencing itself, inherent in gene sequencing errors have been reported in various sequencing platforms. Low overall reproducibility among technical replicates has been reported in samples sequenced in different runs (Wen et al., 2017).

To determine the reproducibility of the MiSeq platform, *PCDHGA12* PCR products were prepared from technical replicates of the same sample but at different stages: 1) from the same bisulfite converted DNA, 2) from the first-round PCR products; and 3) from genomic DNA with separate ligation and bisulfite conversion for each replicate (see Figure 4.15A). To confirm the amplification, 4 µL of the PCR product for each library were assessed by agarose gel electrophoresis (Figure 4.15B).

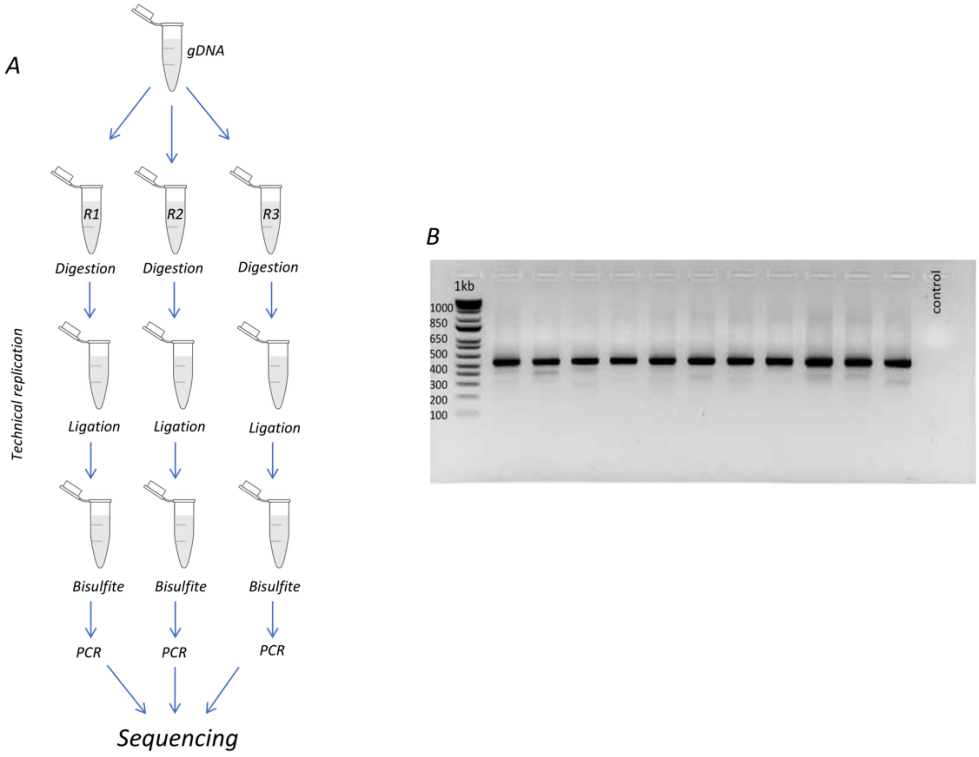


Figure 4. 15. A) Schematic representation of experimental design illustrating the technical replication. B) Gel electrophoresis image confirming the successful of *PCDHGA12* PCR amplifications.

The library was sequenced on a Illumina MiSeq sequencer using 2 x 250 bp paired end sequencing. The high throughput output data for the technical replicates were processed and the mean DNA methylation was calculated (Table 4.1)

Similar methylation patterns of the *PCDHGA12* amplicon were observed in technical replicates confirming the reproducibility of the sequencing pipeline (Figure 4.16). Interestingly, one sample showed differences in the methylation pattern in replicates repeated from genomic DNA compared to the replicates repeated from bisulfite converted DNA or replicates repeated from first PCR products (Figure 4.16 sample 5). This dissimilarity in DNA methylation indicated that the change in DNA methylation happened during sodium bisulfite conversion steps, most likely due to contamination or deamination of 5-mC which might include the conversion of 5-mC to thymine residues (Wang et al., 1980). Lapet et al. (2009) suggested that under certain condition deamination of 5-mC can occur (Labet et al., 2009). Future experiments could explore whether the well-optimised kit that we use could cause such deamination in certain conditions. Although, bisulfite conversion is the gold standard method for evaluating DNA methylation by achieving the base pair resolution of the CpG methylation, the challenge of this method is the undesired side-effects, i.e., DNA degradation and inappropriate conversion.

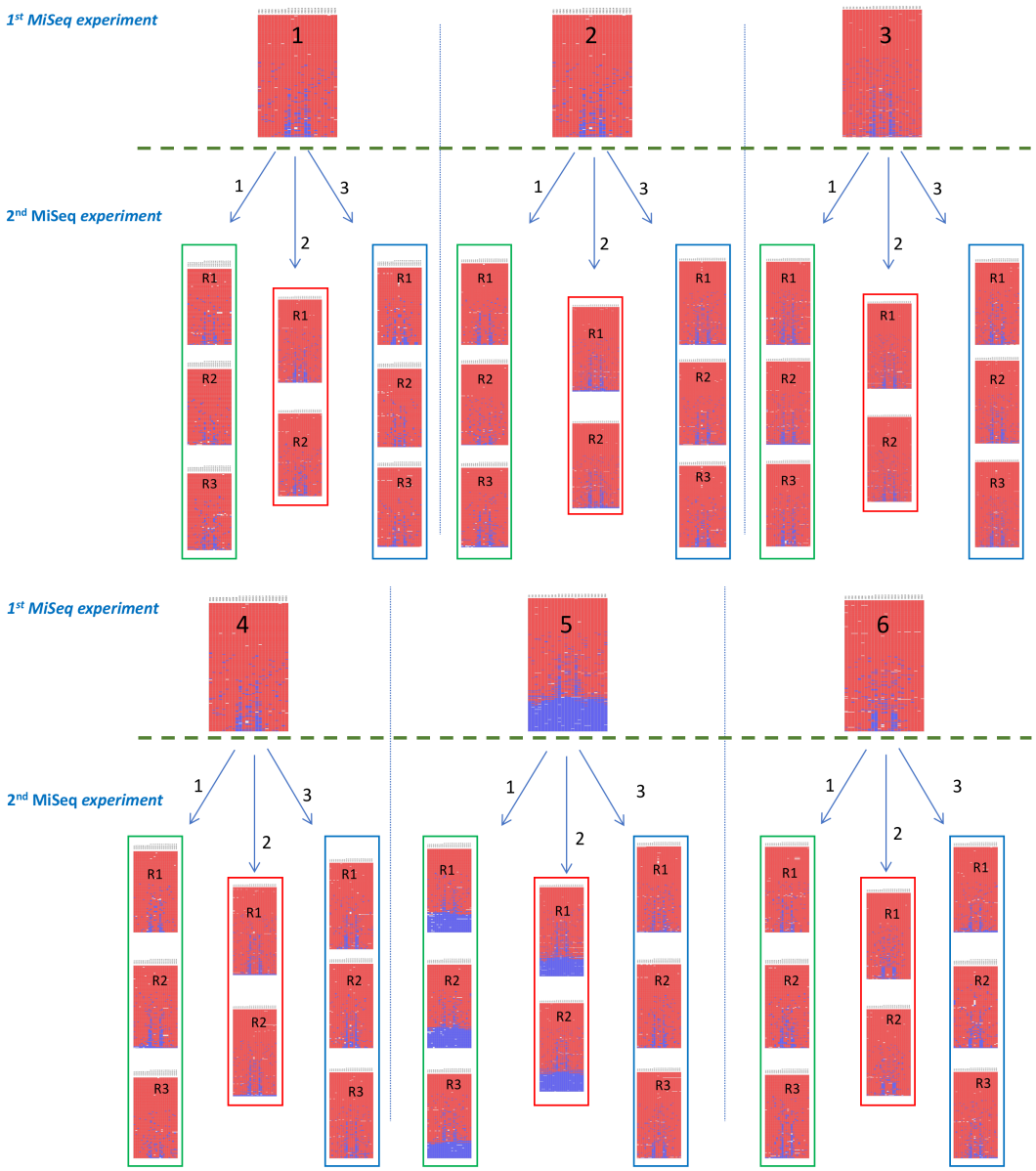


Figure 4. 16. Overall comparison of library methylation patterns between different MiSeq runs. These diagrams display the methylation pattern of hairpin-linked *PCDHGA12* sequences obtained from two independent MiSeq runs for six different samples separated by green dotted lines. Three replicates for each sample were repeated from different steps during MiSeq protocol 1) DNA sequencing was repeated from bisulfite converted DNA, 2) DNA sequencing repeated from the 1st round PCR products, and 3) DNA sequencing was repeated from genomic DNA with new ligation and bisulfite conversion. Similar mean methylation was observed in technical replicates confirming the reproducibility of the MiSeq protocol.

The precision of this approach is very high; e.g., based on 50,000 unique sequences per sample, the 95% confidence interval for the mean, in a sample containing 15% demethylated hairpin sequences, is 14.7-15.3%. In addition, the hairpins themselves incorporate random 14-bp barcodes that distinguish around four million unique sequences by which duplicate or contaminating sequences can be recognised and excluded.

Table 4. 1. Raw methylation data for the technical replicates from two independent experiments.

Sample	Replicate 1 Methylation %	Replicate 2 Methylation %							
		bisulfite			1st PCR		gDNA		
		R1	R2	R3	R1	R2	R1	R2	R3
Control at 8h	0.938	0.935	0.946	0.923	0.932	0.938	0.929	0.945	0.939
Control at 12h	0.928	0.941	0.944	0.936	0.922	0.922	0.937	0.923	0.938
Control at 24h	0.940	0.925	0.934	0.919	0.937	0.944	0.937	0.924	0.930
Dec at 8 h	0.937	0.943	0.907	0.946	0.925	0.926	0.933	0.932	0.928
Dec at 12h	0.704	0.722	0.706	0.749	0.731	0.725	0.929	0.931	0.934
Dec at 24h	0.938	0.943	0.919	0.913	0.940	0.933	0.927	0.931	0.911
Asc/Dec at 8 h	0.890	0.952	0.944	0.932	0.893	0.897	0.955	0.927	0.922
Asc/Dec at 8 h	0.935	0.941	0.898	0.945	0.935	0.943	0.927	0.944	0.943
Asc/Dec at 8 h	0.936	0.942	0.928	0.925	0.935	0.932	0.931	0.935	0.848

4.6.1 Precision of MiSeq sequencing.

Another example of baseline levels of DNA methylation from two independent MiSeq runs for the same samples, was performed in November 2018 and February 2019 (Table 4.2) The mean DNA methylation for a particular sample showed high reproducibility of MiSeq sequencing data among replicates. The Bland-Altman plot shown below (Figure 4.19) compares the mean DNA methylation value for each sample between the two MiSeq runs. Each data point shown, therefore, represents the mean DNA methylation value for each sample compared to the mean DNA methylation for that sample for the MiSeq data from 2018 and 2019 MiSeq data.

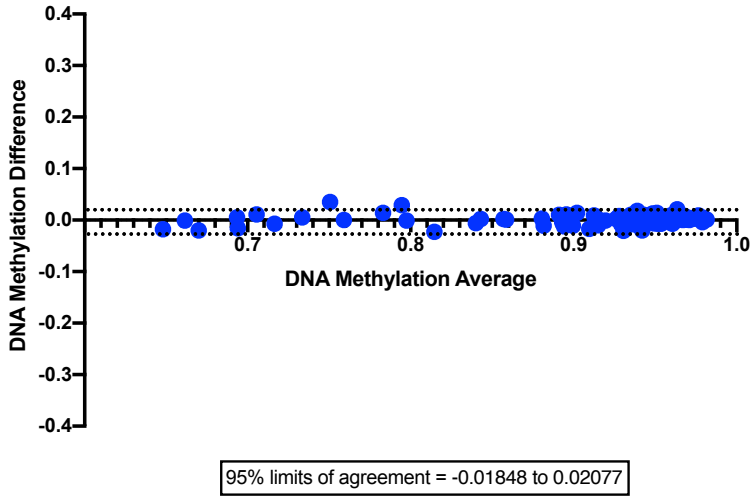


Figure 4. 17. Bland-Altman plot showing the difference in mean DNA methylation of the PCDHGA12 promoter between two independent MiSeq datasets. Y-axis represents the difference between the respective means, and x-axis represents the average of the two respective means. The two dotted lines represent the upper and lower 95% limits of agreement (-0.018 and 0.021).

Table 4. 2. Mean DNA methylation for *PCDHGA12* from two independent MiSeq runs.

No	Replicate 1 Methylation %	Replicate 2 Methylation %	No	Replicate 1 Methylation %	Replicate 2 Methylation %
1	0.937	0.938	45	0.95	0.957
2	0.925	0.933	46	0.797	0.798
3	0.888	0.9	47	0.957	0.957
4	0.91	0.92	48	0.79	0.776
5	0.948	0.951	49	0.977	0.981
6	0.982	0.978	50	0.837	0.843
7	0.972	0.964	51	0.981	0.977
8	0.976	0.973	52	0.901	0.89
9	0.955	0.942	53	0.97	0.969
10	0.9	0.897	54	0.951	0.95
11	0.938	0.932	55	0.957	0.964
12	0.917	0.908	56	0.958	0.944
13	0.97	0.969	57	0.931	0.939
14	0.964	0.964	58	0.686	0.702
15	0.973	0.971	59	0.935	0.936
16	0.956	0.957	60	0.696	0.691
17	0.971	0.971	61	0.963	0.96
18	0.946	0.952	62	0.759	0.759
19	0.974	0.97	63	0.948	0.93
20	0.948	0.955	64	0.858	0.856
21	0.948	0.938	65	0.982	0.981
22	0.92	0.941	66	0.859	0.858
23	0.958	0.959	67	0.932	0.952
24	0.916	0.912	68	0.876	0.887
25	0.902	0.895	69	0.976	0.972
26	0.66	0.68	70	0.891	0.895
27	0.912	0.914	71	0.956	0.944
28	0.713	0.72	72	0.844	0.842
29	0.974	0.953	73	0.94	0.93
30	0.882	0.879	74	0.711	0.70
31	0.957	0.961	75	0.94	0.935
32	0.894	0.905	76	0.736	0.731
33	0.971	0.969	77	0.967	0.967
34	0.932	0.924	78	0.809	0.78
35	0.981	0.972	79	0.964	0.955
36	0.901	0.918	80	0.803	0.826
37	0.972	0.97	81	0.981	0.978
38	0.95	0.941	82	0.896	0.886
39	0.972	0.969	83	0.979	0.973
40	0.946	0.952	84	0.909	0.895
41	0.927	0.925	85	0.968	0.964
42	0.639	0.657	86	0.934	0.944
43	0.919	0.92	87	0.967	0.965
44	0.661	0.662	88	0.929	0.925

Summary

In summary, we have successfully developed and optimised a high-throughput, hairpin bisulfite sequencing protocol for investigating the DNA methylation states of the *PCDHGA12* and *RASSF1* promoters. We developed an automated bioinformatics pipeline to analyse MiSeq methylation for hairpin sequencing data (using terminal and R), that can reliably collapse sequences based on unique molecular barcodes. This replaced the cumbersome, time-consuming Galaxy workflow and '*FASTQ collapser*', which collapsed sequences based on the entirety of the read sequences.

The automated pipeline using Terminal and RStudio is more user-friendly, accurate and much faster than completing the tasks by repetitively using graphical user interface (GUI) approach using Galaxy. Furthermore, we were able to confirm the reproducibility and the precision of MiSeq platform by conducting many replicates from different steps during MiSeq protocol, as well as comparing the sequencing of multiple samples in different MiSeq run.

Chapter Five: The impact of DNA methyltransferase inhibitor ‘decitabine’ on DNA methylation.

5.1. Introduction.

Decitabine or ‘5’-aza-2’deoxycytidine’ is a synthetic nucleoside analogue of cytosine used as an anti-cancer chemotherapy drug (Figure 5.1); for example, in the treatment of leukaemia (Vogler et al., 1976). Decitabine is incorporated into DNA during DNA replication or the S-phase of the cell-cycle, inhibiting DNA methylation by formation of covalent bond with DNA methyltransferase (DNMT1) and ultimately inducing DNA hypomethylation (Momparler, 2005a).

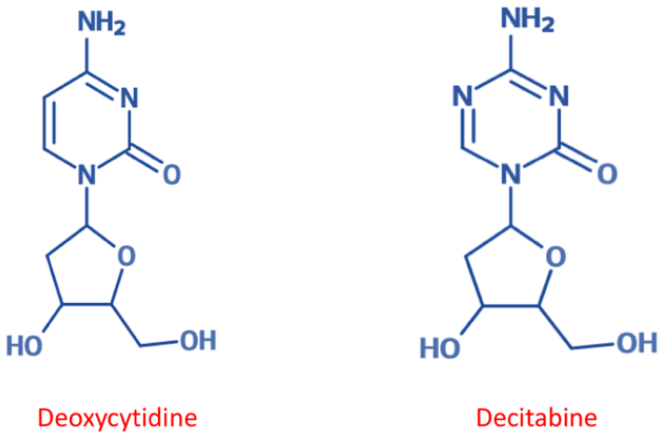


Figure 5. 1. Chemical structure of decitabine and deoxycytidine. Decitabine is an analogue of deoxycytidine in which the carbon in position 5 is replaced by nitrogen.

Once incorporated, decitabine irreversibly binds DNMT1, otherwise known as the maintenance methyltransferase, depleting DNMT1 levels and resulting in failure to reproduce methylation on newly synthesised daughter strands and loss of methylation (Ghoshal et al., 2004).

Decitabine, like natural cytosine, enters human cells through equilibrative nucleoside transporters hENT1, hENT2 and hENT3. Decitabine becomes an active nucleoside triphosphate analogue through a series of phosphorylation steps. Firstly, decitabine is phosphorylated by

deoxycytidine kinase to produce 5-aza-deoxycytidine-5'-monophosphate (5-aza-dCMP). And this monophosphorylated decitabine undergoes further phosphorylation to produce 5-aza-deoxycytidine-5′-diphosphate (5-aza-dCDP) and 5-aza-deoxycytidine-5′-triphosphate (5-aza-dCTP) via the activity of deoxycytidine-5'- monophosphate (dCMP) kinase and the nucleoside diphosphokinase enzymes, respectively (Momparler, 2005b). Finally, with the aid of DNA polymerase, the active form, 5-aza-deoxycytidine-5'-triphosphate (5-aza-dCTP), is incorporated into the replicating DNA during S phase instead of dCTP, blocking the methylation of newly synthesised DNA by trapping DNA methyltransferase (Figure 5.2).

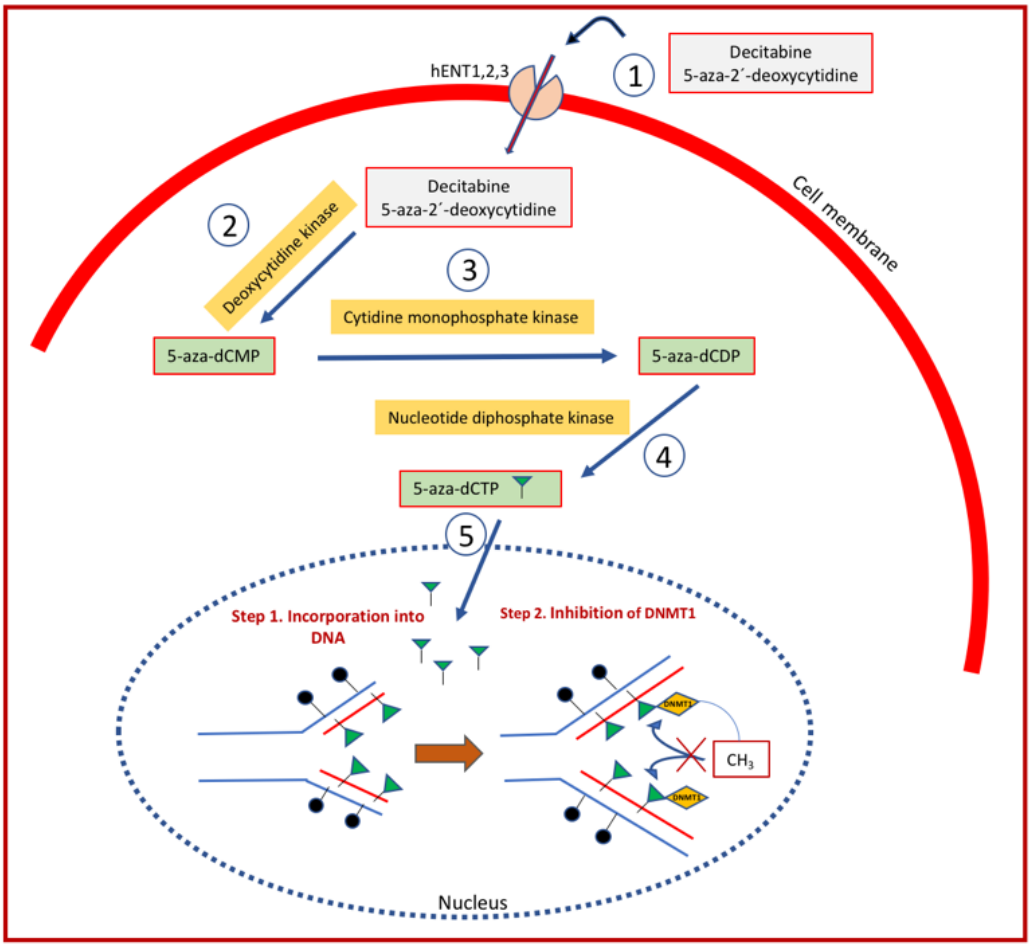


Figure 5. 2. Molecular mechanism of decitabine. Inactivation of DNMT1 by formation of covalent bond with 5-aza-dCTP present in the newly-synthesised daughter DNA strand.

This chapter is divided into two main parts. The first part discusses gene-specific methylation in response to decitabine treatment, and the second part discusses the cell cycle progression and cell viability following decitabine treatment.

5.2 Part one: Cell growth characteristics.

The aims of this part are:

1. To determine the effect of different treatments on the viability of Jurkat, Molt4, Nalm6 and HL60 cells.
2. To study the impact of decitabine on cell cycle progression.
3. To distinguish the effect of indicated decitabine concentration on cell viability.
4. To optimise the thymidine block for the four leukaemia cell lines.

The purpose of this section was to confirm the validity and to support the interpretation of results shown in part one.

5.2.1 Effect of decitabine on cell cycle progression.

Decitabine is a DNA methyltransferase inhibitor (DNMTi) which has anti-tumour activity and is approved for treatment of leukaemia and myelodysplastic syndrome (Jabbour et al., 2008). Relatively, low decitabine doses lead to passive demethylation resulting in re-expression of silent tumour suppressor genes and induction of apoptosis (Jabbour et al., 2008; Qin et al., 2007). We assessed the effect of single dose treatment of decitabine on cell cycle progression to determine whether the cells are sensitive to decitabine and determine the cell toxicity for different doses and longer incubation time.

5.2.1.1 Cell cycle analysis.

The cell cycle comprises four different phases: G1, S, G2 and M. During the G1 phase, cells grow in size and prepare for DNA replication. S phase (or synthesis phase) starts when the cells pass the G1/S checkpoint and begin DNA synthesis. Once the cells reach G2, the DNA is duplicated and the cells prepare to divide by passing the G2/M checkpoint into M-phase (or mitosis). Determining the effect of decitabine on cell cycle progression requires a reliable cell cycle assay. The DNA content of the proliferating cells can be used to determine the proportion of cells in each stage; this can be assessed by binding of nucleotide-binding dye such as propidium iodide followed by flow cytometric analysis.

To monitor the cell cycle progression and determine the success of the thymidine-block induced synchronisation, Jurkat cells were grown to the required concentration (1 million cells/mL), then fixed with 70% ethanol and stained with propidium iodide RNAse to measure the DNA content. Propidium iodide staining was measured using flow cytometry and histogram plots were used to observe the changes in the distribution of the Jurkat cell populations through the cell cycle.

The flow cytometric results for asynchronous Jurkat cells showed a large peak of cells in G1/G0 phase (diploid DNA content), a middle S-phase in which fewer cells contained increasing amounts of DNA, and G2 phase cells represented by a smaller peak (tetraploid) (Figure 5.3).

Analysis of all flow cytometry output data was performed using FlowJo® 10.4.2 software (FlowJo, LLC).

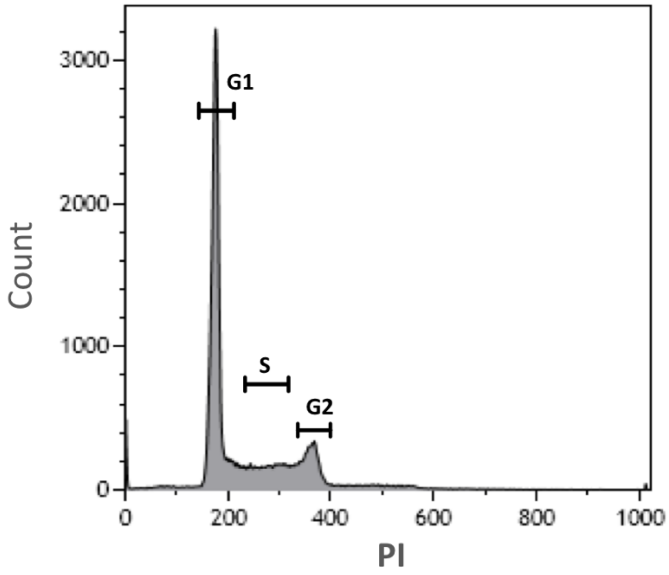


Figure 5. 3. Histogram showing DNA content distribution of live unsynchronised Jurkat cells stained with propidium iodide (PI) stain and analysed by flow cytometry.

5.2.1.2 Synchronisation of Jurkat cells.

Cell synchronisation is the process by which cells in different phases of cell cycle were brought into one phase. This step is crucial to study specific events during the cell cycle progression. To ensure the simultaneous replication of DNA, thymidine (a DNA synthesis inhibitor) was used to synchronise the cells in G1 phase boundary prior to DNA synthesis, such that all cells will go through S-phase together (Chen & Deng, 2018a).

To arrest the cells in G1 phase, a single thymidine block protocol was used. Jurkat cells were incubated with 2 mM 2′-deoxythymidine (Sigma-Aldrich) for 18 h. Thymidine depletes the intracellular cytidine levels, leading to growth arrest of most cells in the G1 phase, and therefore significantly reducing the cell numbers in the G2.

After synchronisation, Jurkat cells were washed twice with pre-warmed fresh media supplemented with 50 µM 2′- deoxycytidine (dC). Cells were harvested at increasing time point and the cells were prepared for flow cytometric analysis.

Flow cytometry histograms were used to monitor the changes in cell population during cell progression. Histograms of thymidine blocked Jurkat cells showed a large peak of resting cells in G1 phase, with a significant reduction in G2 population (Figure 5.4, 0 h). At 1-2 h synchronised cells show very slow progression, at 4 h the synchronised cells start progression into the S phase showing a large peak in the S phase and two small peaks of G1 and G2. At 6 h cells are undergoing DNA replication in S phase and some start G2 phase showing a peak combined of S phase and early G2.

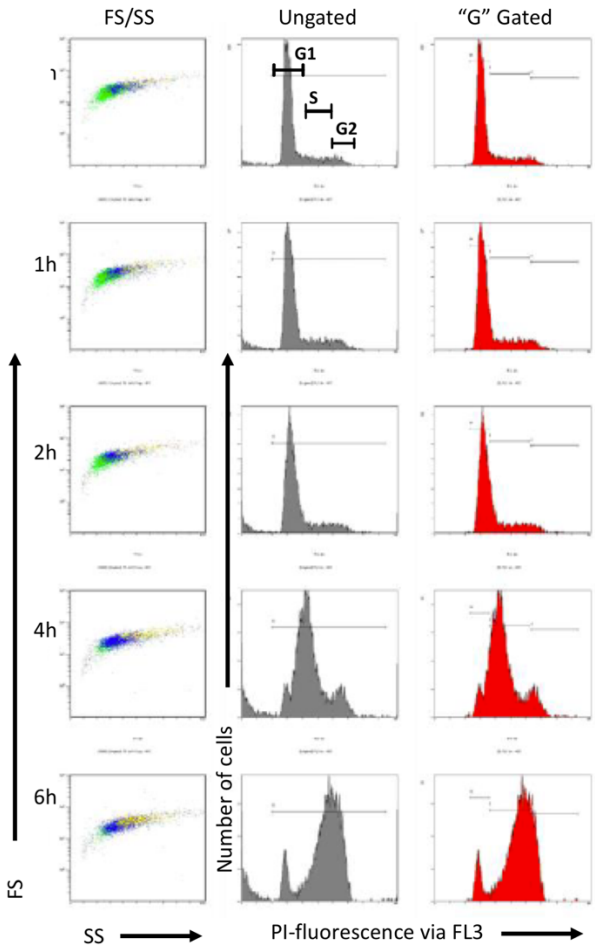


Figure 5. 4. Analysis of synchronised Jurkat cells following release of cell cycle arrest. Cells were treated with thymidine for 18 h followed by supplementation of 50 µM 2′- deoxycytidine (dC). The cells proceeded through S phase with time. The first column shows Forward scatter (FS)/Side Scatter (SS) plots of cells. The second column (Ungated) shows plots of number of cells vs PI-fluorescence across 1024 'bins', with some doublets and aggregates of dead cells. The third column (Gated) shows data from the second panel that falls within the arbitrarily labelled gate "G" from the second panel to exclude most of the doublets and aggregates.

5.2.1.3 Roles of 2′-deoxycytidine (dC) in cell cycle progression.

To show the necessity of deoxycytidine in initiating the cell cycle progression, synchronised Jurkat cells were supplemented with or without 2′-deoxycytidine, and then flow cytometry was

used to monitor cell cycle progression through S phase. Briefly, Jurkat cells were cultured in RPMI media and synchronised to G1 phase with 2 mM thymidine (as above). Cells were divided into two groups, the first group was supplemented with fresh media loaded with 50 µM 2′-deoxycytidine and the second group, supplemented with fresh media without deoxycytidine, acted as a control. Thereafter, cells were collected at 0, 2, 4, 6 h. The cell cycle transition was analysed with flow cytometry by DNA staining using propidium iodide (PI). Deoxycytidine initiates the cell cycle by providing the cytosine molecule required for DNA synthesis. Jurkat cells supplemented with 50 µM 2′-deoxycytidine start cell progression and proceeded through S phase. In contrast, Jurkat cells not supplemented with 2′-deoxycytidine remained arrested at G1 phase (Figure 5.5).

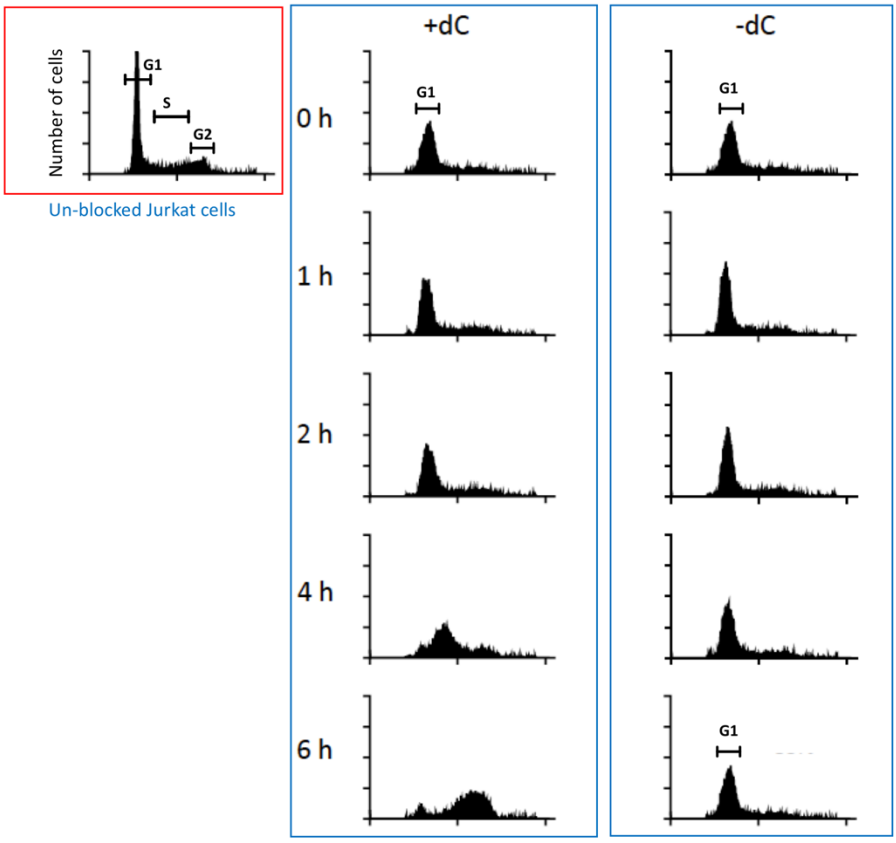


Figure 5. 5. Flow cytometry results showing the reliance on dC to allow progression through the cell cycle. The first column; NB (un-blocked) asynchronous cells with a large peak of cells in G1/G0 phase and small peak of cells in the G2/M phase. The second column shows thymidine-blocked cells supplemented with 50 µM 2′- deoxycytidine (dC). The cells proceeded through S phase. The third column shows thymidine blocked cells not supplemented with 50 µM 2′- deoxycytidine (dC). The cells remained arrested at G1/S boundary.

5.2.2 Effect of decitabine on cell cycle progression.

To examine the cytotoxic effect of decitabine on cell growth and cell cycle progression, synchronised Jurkat cells were treated with a single dose of decitabine (5 µM) and incubated for 72 h. Briefly, Jurkat cells (1x10^6/mL) were arrested in G1 using thymidine as before, washed and transferred to fresh media, supplemented with 50 µM 2´-deoxycytidine (dC). Immediately, thereafter, the cells were treated with 0 or 5 µM decitabine and collected at 0, 1, 2, 4, 6, 24, 48 and 72 h time points. Cells were pelleted by centrifugation (2000 rpm for 4 min at 4 °C), washed twice with cold PBS and suspended in 250 µL propidium iodide/RNase staining solution (Pub. No. MAN0007554) and incubated at 4 °C for 1 h in the dark. The cell cycle analysis was performed by measuring propidium iodide fluorescence using Gallios flow cytometer (Figure 5.6).

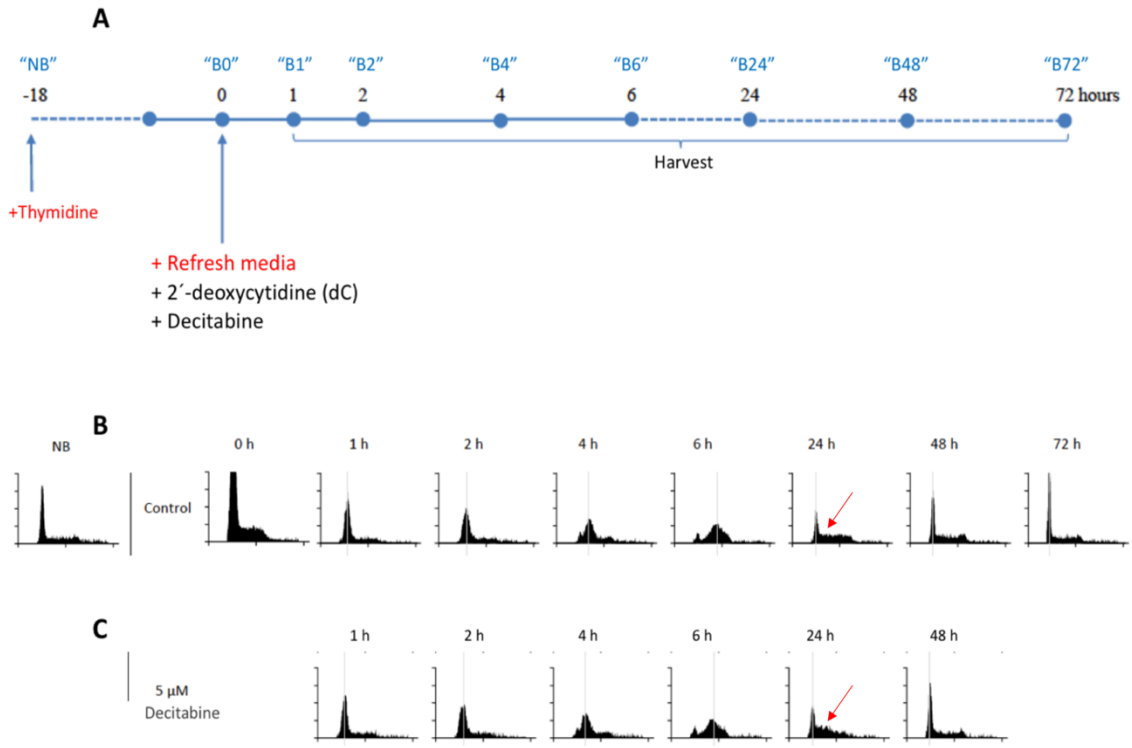


Figure 5. 6. Cell cycle analysis of synchronised Jurkat cells. A) Treatment and collecting timepoints. B) Cell cycle progression of control Jurkat cells. C) Cell cycle progression of Jurkat cells treated with 5 µM decitabine. All cells were supplemented with 50 µM dC at the time of release from the thymidine block. The arrows show the S phase (cells population) between untreated and decitabine treated cells.

Cell cycle distribution analysis revealed that decitabine had little to no effect on cell cycle kinetics compared with the control. The cells continued to progress through the S phase of the cell cycle albeit slower than control cells h. This is suggested by an increase in the percentage

of cells in S phase at 24 h timepoint in decitabine treated cells compared to the control cells (Figure 5. 6).

In conclusion, cell cycle analysis of synchronised Jurkat cells treated with 5 μM decitabine demonstrated little effect on cell cycle progression. A level of growth arrest was observed after 24 h of exposure to decitabine compared to control cells; this was accompanied by an increase in cells in the S phase of the cell cycle compared to untreated control cells.

5.2.3 Inhibition of cell viability in Jurkat cells.

To assess the efficacy of decitabine and other reagents in inhibiting cell growth and proliferation, cell viability of Jurkat cells was assessed, cells were stained by propidium iodide and analysed using flow cytometry. Cells were treated with 5 μM of decitabine, 500 μM glycine chloramine, 100 μM cobalt chloride, 500 μM ascorbate or different combinations of these reagents, and incubated for three days. Cells were harvested and washed with PBS and prepared for analysis as described in section 3.8.9.

Cell viability of Jurkat cells was examined during the treatment with different reagents at 0 h, 24 h, 48 and 72 h. Reagent concentrations were used according to two previous independent studies performed in the Morison and Hampton laboratories. Viability of Jurkat cells treated with ascorbate or decitabine was not affected and remained similar to the viability of the control cells, whereas, the combination of ascorbate and decitabine reduced the viability of Jurkat cells by 25% after 72 h. There was a 30% reduction in the viability after 24 h in cells treated with glycine chloramine. Interestingly, the combination of ascorbate and glycine chloramine increased the viability of Jurkat cells and reduced the cell toxicity induced by glycine chloramine alone. Glycine chloramine is a short-acting oxidative compound that is unstable. There was a high likelihood that most of the glycine chloramine would have been neutralised by the antioxidant ascorbate.

As expected, as it is known that cobalt is a toxic compound, cobalt chloride was detrimental to cell viability. Cell viability was significantly reduced in cells treated with cobalt chloride alone or in combination with decitabine or glycine chloramine (Figure 5.7).

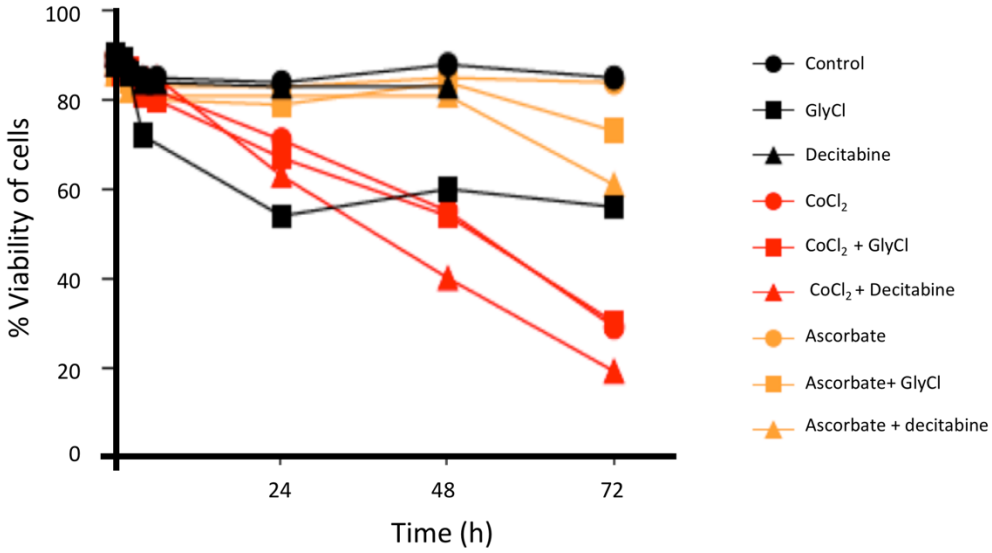


Figure 5. 7. The effect of different treatments on the viability of Jurkat cells over time. Cell viability was measured using flow cytometry. Cells were stained by propidium iodide (PI).

5.2.4 Growth inhibition of four haematopoietic cell lines by decitabine.

In addition to Jurkat cells, our study was extended to investigate the growth inhibitory effect of decitabine in another three haematopoietic cell lines: Molt4, Nalm6 and HL60. Evaluation of decitabine toxicity was performed using different concentrations of decitabine for 24 h. Cells were blocked by thymidine for 18 h, released by supplementation with fresh media loaded with dC and different concentrations of decitabine (0 to 5 µM) were added at the time of release. To measure the viability of cells, cells were stained with trypan blue and counted using a haemocytometer.

Minor changes in the proportions of live and dead cells were seen between decitabine concentrations (up to 2 µM) in the three cell lines Jurkat, Molt4 and Nalm6, while a marked change was seen in HL60 (Figure 5.8). Additionally, decreased cell number was observed in the four cell lines after exposure to ≥ 1 µM decitabine (Figure 5.9).

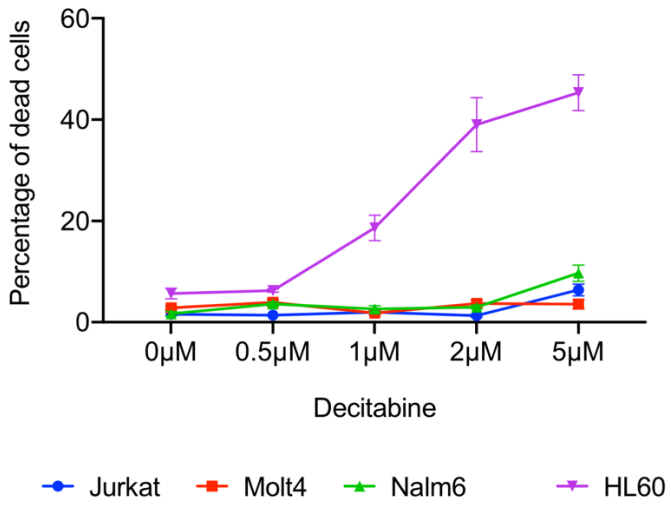


Figure 5. 8. Cell death in the four cell lines (Jurkat, Molt4, Nalm6 and HL60) following incubation with increasing concentrations of decitabine for 24 h. Viability was determined by trypan blue staining. Values plotted are means of three counts ± SEM.

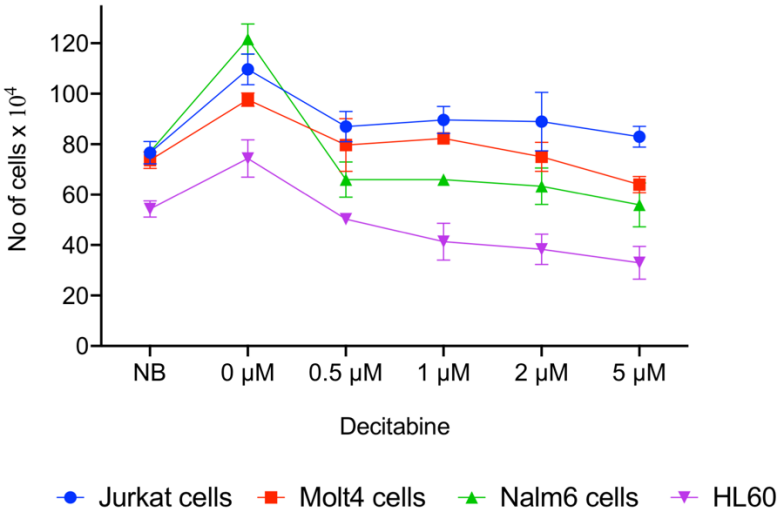


Figure 5. 9. Effect of decitabine on number of viable cells. Jurkat, Molt4, Nalm6 and HL60 cell lines were exposed to different concentrations of decitabine for 24 h. Cell viability was assessed by trypan blue. NB: represents the number of cells at 0 timepoint, Value are means of three replicates ± SEM.

The addition of 5 μM decitabine to cells resulted in a reduction of cell viability and increase in the number of dead cells. Minimal variation in the percentage of dead cells were observed in treated Jurkat, Molt4 and Nalm6 between concentrations of 0.5 and 2 μM compared to HL60 cells which found to be sensitive to decitabine concentration of 1 μM (decitabine LC50 for HL60 is 10 times less compared to Jurkat, Molt4 or Nalm6 cells). Based on the results obtained from the decitabine toxicity experiment and to ensure cell cycle progression through DNA synthesis, the decitabine concentration was decreased to 1 μM for the three cell lines and 0.5 μM for HL60.

5.3 Part two: Gene specific methylation response to decitabine.

The goal of this part is test the hypothesis that hemi-methylated DNA enhances the rate of TET enzyme induced active demethylation in somatic cells. The hypothesis was developed based on the surprising observation by Dr. Helena Magrath from the Morison laboratory, that active demethylation occurs within just two hours of exposure of cells to physiological oxidative stress.

Hypothesis:

1. The hemi-methylation induced by the DNMTi inhibitor, decitabine, enhances the rate of TET induced active demethylation.

The aims of this section:

1. To assess the impact of decitabine on DNA methylation status of *RASSF1* and *PCDHGA12* promoters.
2. To document the hemi-methylation induced by decitabine.
3. To study the role of TET enzyme in active demethylation.

5.3.1 Use of MiSeq hairpin sequencing to assess the *RASSF1* and *PCDHGA12* promoters in Jurkat cells.

Bisulfite conversion and sequencing is the gold standard method for describing DNA methylation (Feng et al., 2011). However, standard bisulfite conversion techniques result in separation of double stranded DNA and amplification of one strand only. Therefore, to investigate the methylation state of both strands together during DNA synthesis requires the development of a novel experimental technique. *RASSF1* and *PCDHGA12* gene promoters have been reported to be highly methylated in haematological malignancies (Gordon et al., 2012; Taylor et al., 2007). The hairpin bisulfite technique with a barcoded hairpin linkers was developed to determine the symmetry of CpG sites for both complementary strands in the promoter regions of *RASSF1* and *PCDHGA12* genes similar to protocol described by (Laird et al., 2004). Hairpin linkers were designed to specifically binds the two DNA strand together, so the methylation pattern of both (parental and daughter) strands can be determined.

The method is described in detail in chapter 3. In brief, the experiment was carried out on Jurkat cells. Jurkat cell DNA was digested with two different restriction enzymes SacI and BamHI and ligated with hairpin linkers. *RASSFI* and *PCDHGA12* regions were then amplified using bisulfite specific primers (section 3.4.6).

The experimental protocol requires two rounds of PCR, which allows tagging of the target amplicons with unique Illumina adapter sequences and enables multiplexing during sequencing of the pooled libraries. The libraries were sequenced on an Illumina MiSeq using Reagent Kit v3 (Illumina, San Diego, USA; 500 cycles) in paired-end mode. The output sequencing files were analysed using our established bioinformatics framework for hairpin DNA methylation analysis (details in chapter four).

5.3.2 Methylation patterns of *RASSF1* and *PCDHGA12* promoters in Jurkat cells.

The standard heatmap output of *BiQ Analyzer* represents the linear forms of the *RASSF1* and *PCDHGA12* hairpin reads, with methylated and unmethylated CpG sites represented as red and blue rectangles, respectively (Figure 5.10). Each column represents a CpG site and every row a single sequencing read, with each read being the linear form of the hairpin molecule. Each hairpin molecule most likely reflects a separate cell after PCR amplification. On the basis of the output of the *BiQ Analyzer*, the *RASSF1* and *PCDHGA12* hairpin sequences were shown to be heavily methylated in control/untreated Jurkat cells (Figure 5.10) (Figure 5.11), confirming their suitability for demethylation studies.

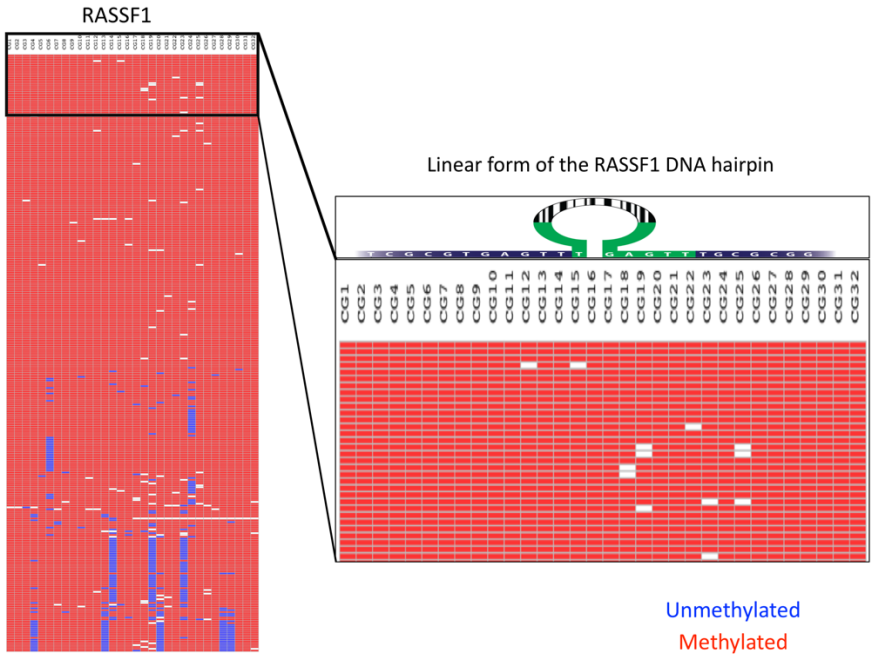


Figure 5. 10. *BiQ Analyze*r output showing the methylation pattern in untreated Jurkat cells of the linearised hairpin of the *RASSF1* promoter. Rows represent individual hairpin reads; columns represent CpG dinucleotides; blue rectangles – unmethylated CpGs; red rectangles – methylated CpGs; white rectangles – missing sequence information.

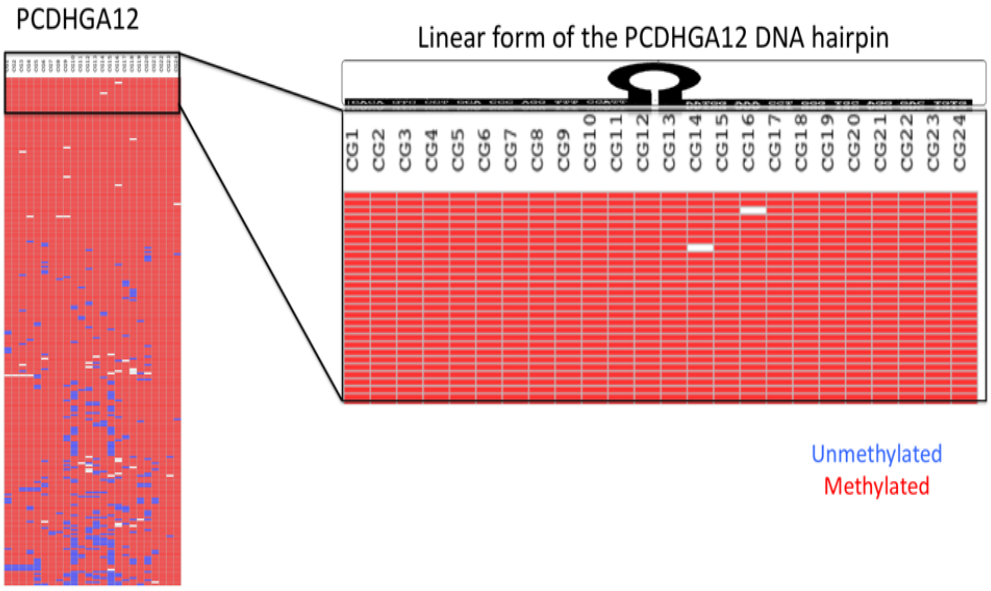


Figure 5. 11. *BiQ Analyzer* output showed the methylation pattern in untreated Jurkat cells of the linear form of the hairpin of the *PCDHGA12* promoter. Rows represent unfolded hairpin reads; columns represent CpG dinucleotides; blue rectangles – unmethylated CpGs; red rectangles – methylated CpGs; white rectangles – missing sequence information.

5.3.3 Documentation of short-term decitabine induced hemi-methylation in the *PCDHGA12* promoter.

To study the impact of short-term effect of decitabine on the DNA methylation of Jurkat cells, the methylation pattern of the *PCDHGA12* promoter was investigated in synchronised Jurkat cells treated with decitabine for 12 h compared to untreated control. Briefly, Jurkat cells were grown to a concentration of 1 x 10^6 /mL and synchronised in G1 phase by incubation with 2 mM 2′-deoxythymidine for 18 hours. Synchronised Jurkat cells were divided into two groups (control and decitabine) and released from cell cycle arrest by supplementation with 50 µM 2′-deoxycytidine. At the time of release, cells were treated either with or without 5 µM decitabine and the cells were harvested at different time points during DNA replication. In all experiments, decitabine was added at the time of cell release from synchronisation (0 h).

With the hairpin bisulfite technique we monitored the changes in DNA methylation in Jurkat cells during DNA synthesis. The percentage of hemi-methylated DNA strands gradually increased during DNA replication in a time dependent manner in synchronised Jurkat cells treated with decitabine compared to control.

The heatmaps clearly show that decitabine induced a gradual increase in hemi-methylated DNA strands in the *PCDHGA12* promoter as early as 2 h after release from synchronisation. Hemi-

methylation is shown when one side of the hairpin is methylated and the other side is unmethylated (Figure 5.12).

Based on the mechanism of decitabine, it was predicted hemi-methylation, but the hemi-methylation never been clearly shown before in the literature. We clearly described the dynamic of decitabine mechanism and demonstrated a gradual increase in hemi-methylated DNA with time in decitabine-treated, post-synchronised cells.

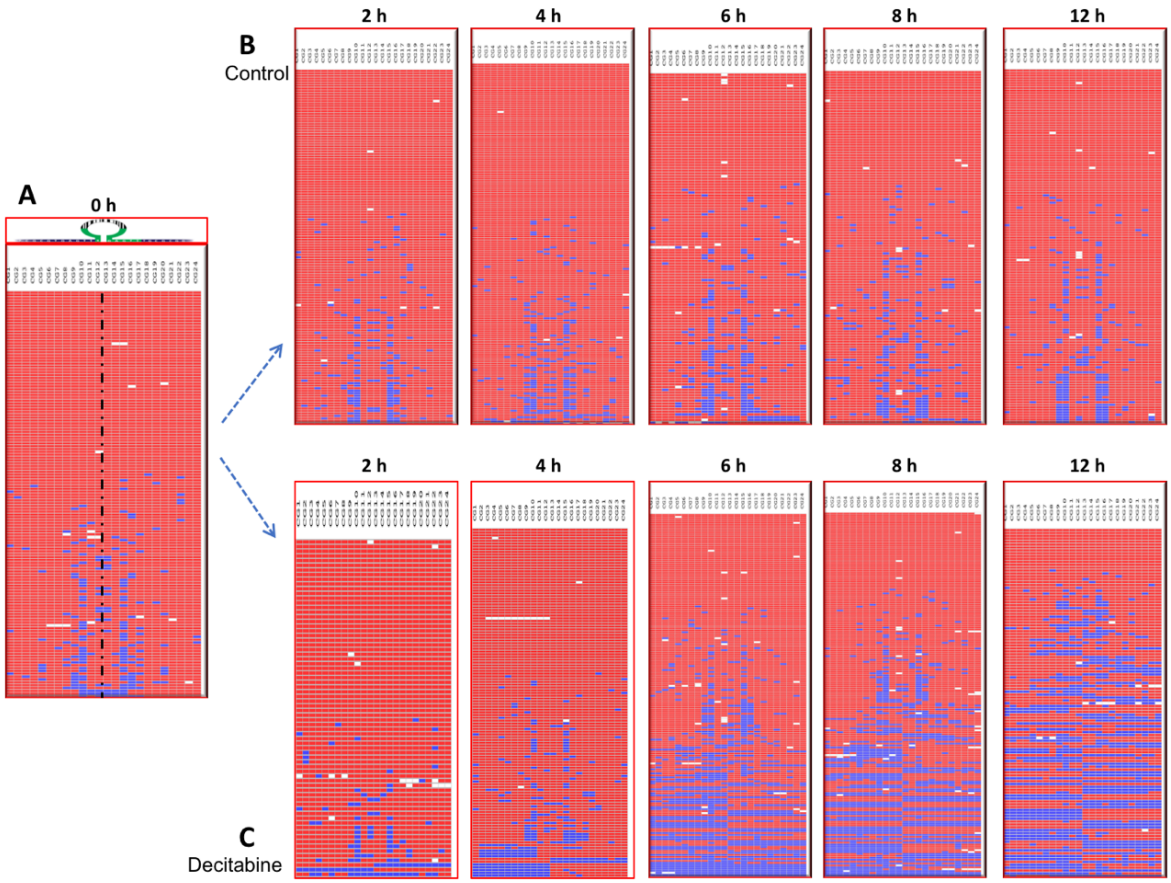


Figure 5. 12. Representative example of decitabine-induced hemi-methylation in Jurkat cells, A) Heatmap output of *PCDHGA12* promoter from synchronised Jurkat cells at time 0 h. B, C) One example showing increase in hemi-methylation with decitabine treatment during DNA replication in *PCDHGA12* promoter compared to untreated control. The dotted line represents the location of the hairpin barcode. (red = methylated CpG, blue = unmethylated CpG, white = missing sequence).

Interestingly, we were expecting fully hemi-methylated DNA reads at 12 h after inhibiting the DNMT1 by decitabine, as all the cells were initiated into the S phase. We observed 15~20% fully methylated hairpin reads at 12 h, this value represents the cells that did not respond to decitabine treatment. The incomplete hemi-methylated strands might be due to 1) dead cells

which still having the fully methylated strands, 2) death of the cell as a result of decitabine toxicity, 3) might be due to slow replication, 4) might be due to cellular repair mechanisms during decitabine treatment. No evidence of delay or cell cycle arrest was seen from cell cycle analysis at 12 h (see section 5.2.2), but we observed an increase in cell population in S phase at 24 h in decitabine treated cells compared to untreated control, which might have reflected a slower replication process (Figure 5.6 c), or 5) incomplete synchronisation, meaning that some of the cells were not at the G1 phase, and might be passed the S phase to G2 or entered mitosis.

To calculate the percentages of hemi-methylated hairpin reads, the binary methylation data (1, 0 or x) were extracted from the *BiQ Analyzer* output result files using UNIX terminal scripts. Using an R script, the data were further analysed after "folding" the hairpin reads to calculate the methylation proportion of each hairpin read. The hairpin reads were classified based on the methylation pattern of the complementary CpGs sites as methylated, hemi-methylated or unmethylated (Figure 5.13). All scripts can be found in the appendix.

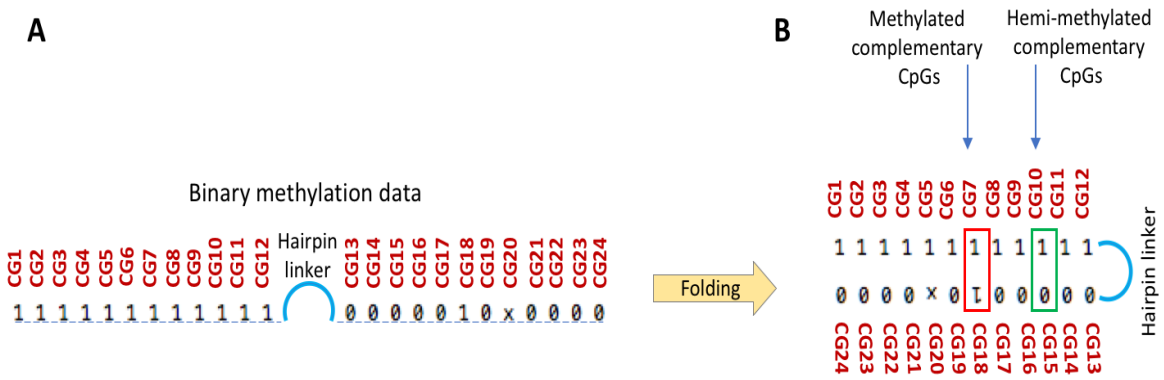


Figure 5. 13. Calculating the proportion of complementary CpG methylation by "folding" the binary methylation data. A) Linear binary methylation data. B) Folded binary methylation data.

A sequence read was classified as 'methylated' when the proportion of methylated complementary CpGs ≥ 0.51; 'hemi-methylated' if the proportion of hemi-methylated CpGs ≥ 0.51 and called 'unmethylated' if the proportion of unmethylated CpGs ≥ 0.51. The percentage of fully methylated, hemi-methylated and unmethylated CpGs sites for each sample was determined (Figure 5.14).

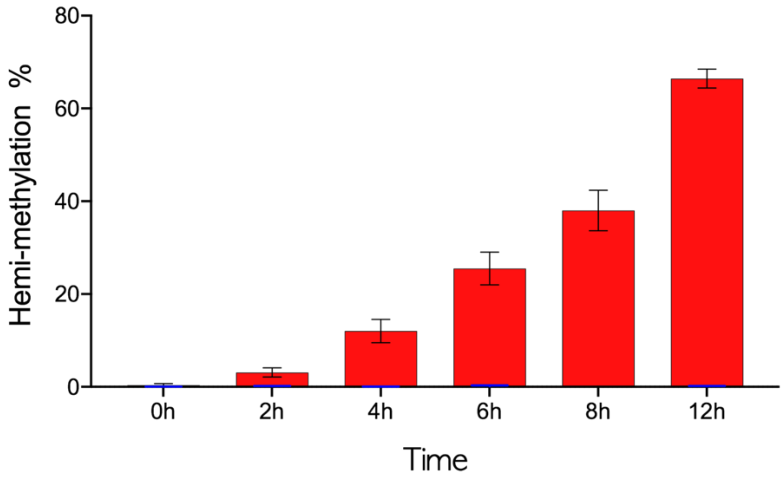


Figure 5. 14. Percentage of hemi-methylated sequence reads following treatment with decitabine. Decitabine induced hemi-methylation of *PCDHGA12* promoter in post-synchronised Jurkat cells. The data presented are means of three replicates ± SEM.

After 2 h, the hairpin reads displayed 4% fully hemi-methylated reads, which increased to 18% at 4 h, 35% at 6 h, 45% at 8 h and 65% at 12 h. Complete loss of methylation (active demethylation) was observed at 6 h time points with 7% fully unmethylated hairpin reads.

5.3.4 Long-term impact of decitabine on methylation of the *PCDHGA12* promoter.

To validate the short-term impact of decitabine, and to investigate the long-term effect of a single dose of decitabine treatment, the experiment was repeated with extended timepoints. I performed this experiment in Prof Hampton's laboratory (University of Otago, Christchurch) and the resulting cells were collected and transferred to the Morison lab for DNA extraction and sequencing. This experiment was conducted using the same protocol mentioned in section 5.3.3. Three technical replicates were prepared and sequenced for each sample (replicates were prepared from genomic DNA with separate digestions and bisulfite conversions). The figures below show the methylation patterns of the *PCDHGA12* promoter following decitabine treatment, compared to untreated control (Figure 5.15, Figure 5.16).

The percentage of fully methylated, hemi-methylated and unmethylated strands was calculated for each sample and stacked plots for the three replicates were generated (Figure 5.9). The following figures show the changes in DNA methylation levels, especially the increase in hemi-methylated reads in decitabine treated *PCDHGA12* promoter compared to untreated control. The effect of decitabine appeared as early as 4 h after release, and increased with time for all three replicates.

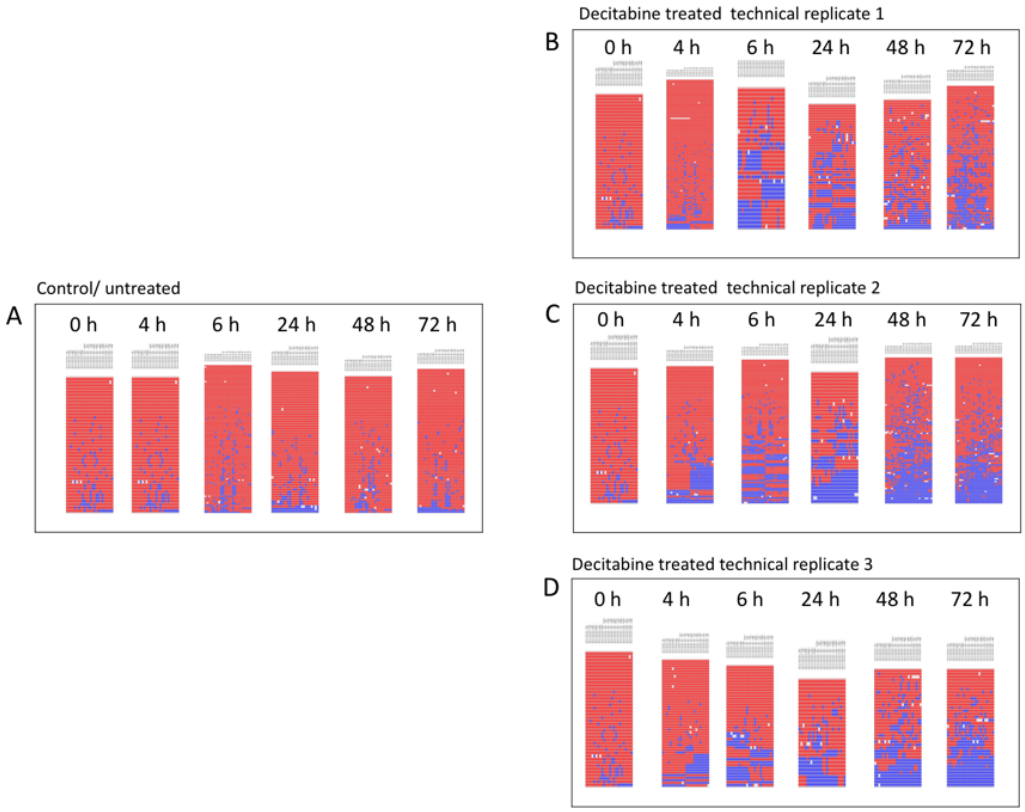


Figure 5. 15. Heatmap outputs from *BiQ Analyzer* showing the methylation pattern of the *PCDHGA12* promoter after decitabine treatment for up to 72 h. A) control Jurkat cells. B, C and D) show the results of three technical replicates.

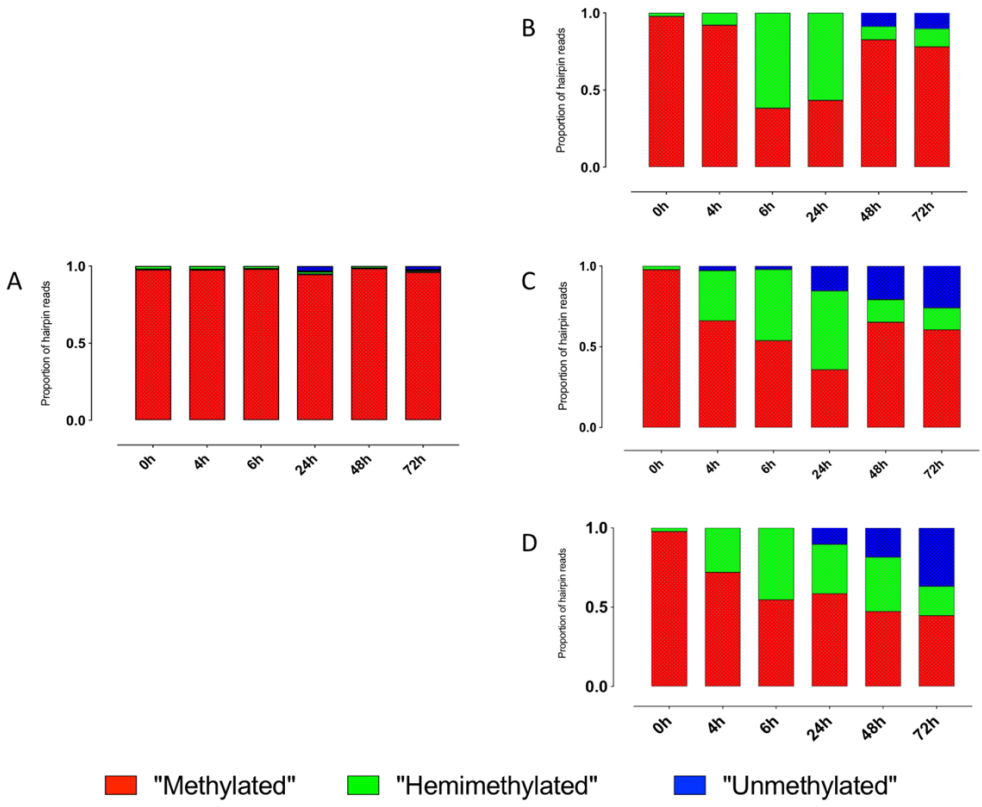


Figure 5. 16. Hairpin-based methylation of *PCDHGA12* promoter following decitabine treatment for up to 72 h. A) control Jurkat cells. B, C and D) show the results of the three technical replicates. Classification of the reads depends on the formula described above.

To show the distribution of methylation of reads within an individual sample, the methylation values for each sequence were extracted from the results file and plotted using a 'violin-points' plot. The plot presents the distribution methylation for each sample and provides a similar reflection of the methylation patterns as the heatmaps for the control and decitabine treated samples (Figure 5.17).

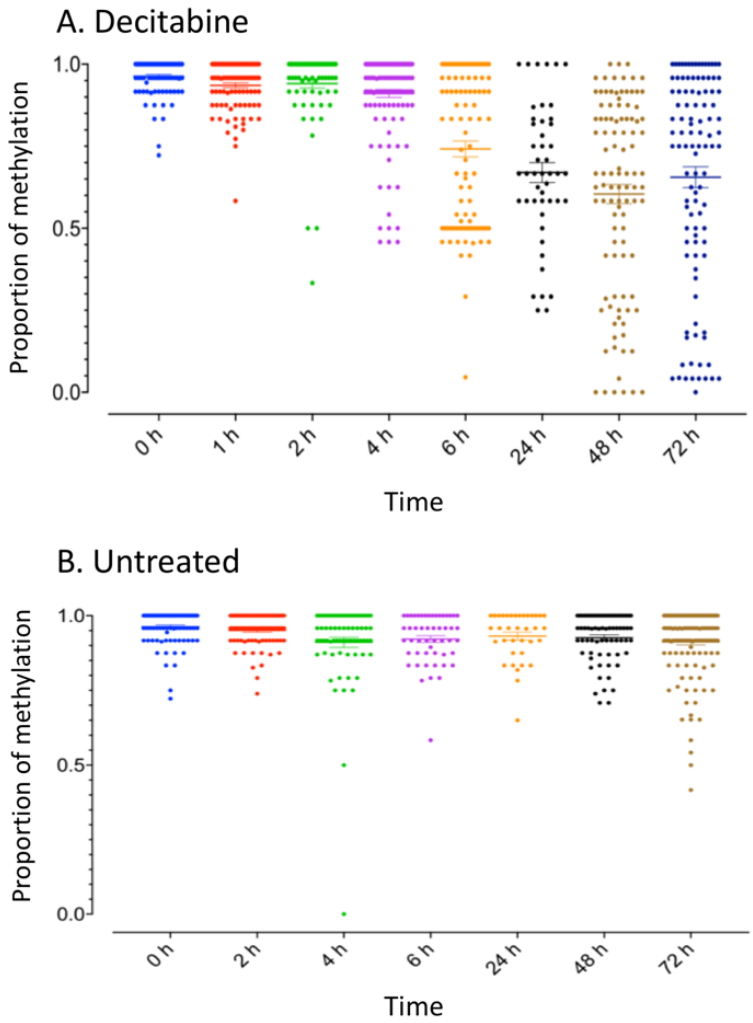


Figure 5. 17. CG methylation reads distribution of *PCDHGA12* promoter through multiple timepoints among different groups of Jurkat cells. A) decitabine treated compared to, B) untreated control. The data correspond to Figure 5.9 B. The longitudinal lines are the means ± SEM.

While short-term demethylation in response to decitabine treatment is mechanistically informative, we extended the treatment to examine the extent of long-term demethylation as a result of decitabine.

5.3.5 Long term exposure to single dose decitabine associated with methylation changes of *PCDHGA12* promoter.

To test the hypothesis that decitabine leads to long-term changes in methylation , we took three approaches, 1) we extended the decitabine treatment to 120h so that it would reflect several cell cycles after a single dose exposure, and measured the DNA methylation changes at 48 h, 72 h, 96 h and 120 h; 2) we extended the experiment to primary cell lines Molt4, Nalm6 and HL60 to confirm that the observations of decitabine-induced demethylation are widely applicable; 3) genome wide methylation (PBAT) was used to assess the global impact on methylation levels (Chapter Seven).

To investigate the chronic effect of single dose of decitabine on the methylation patterns of *PCDHGA12* promoter, synchronized Jurkat cells were incubated with 5 μM decitabine and proceeded through DNA synthesis by supplementation of 50 μM dC. Cells were collected at 0 h, 2 h, 4 h, 6 h, 8 h and 12 h and furthermore at the extended timepoints 24 h, 48 h, 72 h, 96 h and 120 h after a single dose of decitabine at 0 h. Media was changed every two days.

A substantial decrease in gene specific methylation of *PCDHGA12* promoter was observed in decitabine treated cells. The decrease in methylation was associated with an increase in the percentage of hemi-methylated reads in cells treated with decitabine compared to control untreated cells (Figure 5.18). The amount of hemi-methylation reached 65% at 12 h. Interestingly, slight re-methylation on day three of decitabine free growth was observed. This re-methylation continued through days four and five of decitabine-free growth. These results were observed and confirmed 'locally' by gene specific methylation analysis (Figure 5.19) and 'globally' by low coverage PBAT analysis (details in Chapter Seven).

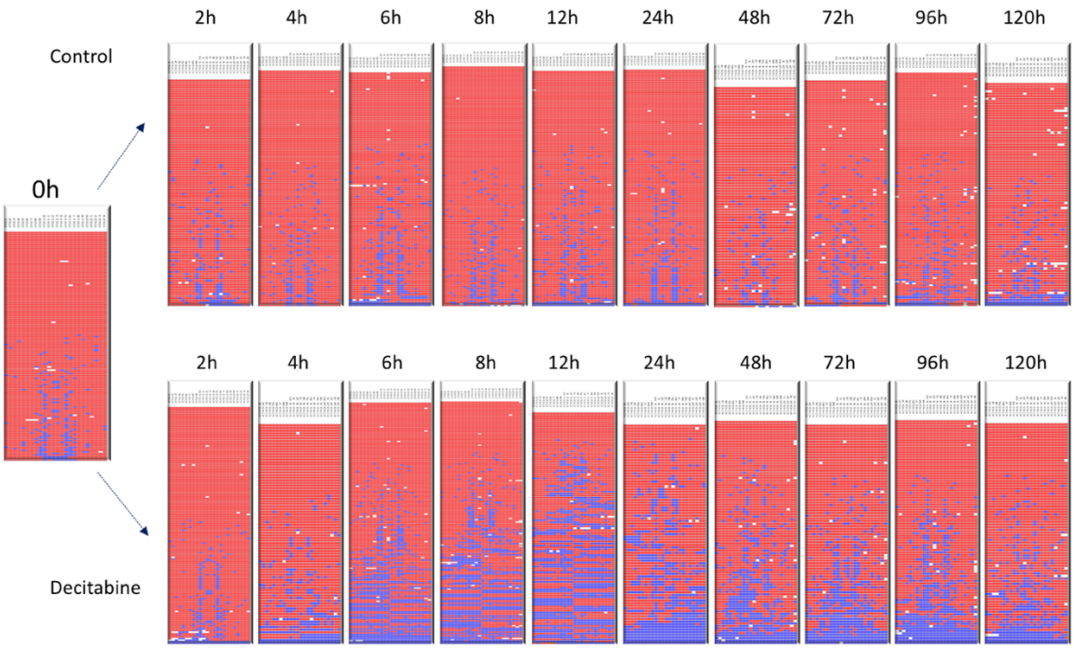


Figure 5. 18. Gene specific demethylation and re-methylation in Jurkat cells. Heatmaps of sequencing results showing the methylation pattern of *PCDHGA12* promoter at different timepoints in cells treated with decitabine compared to controls. Jurkat cells were treated with a single dose of decitabine and cultured for up to 5 days.

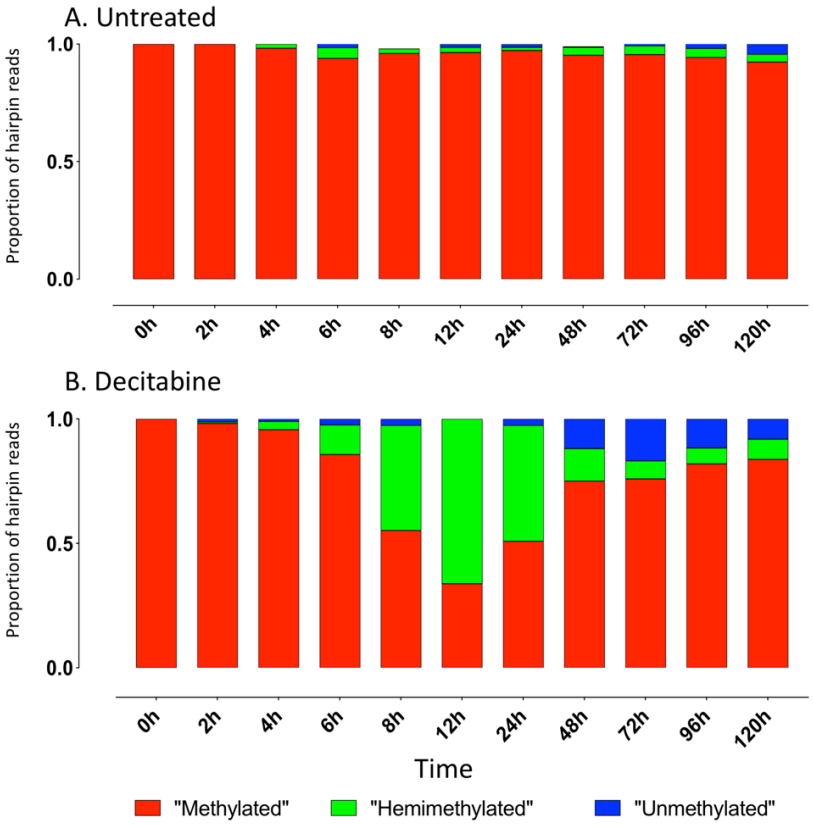


Figure 5. 19. Hairpin-based methylation of *PCDHGA12* promoter in synchronised Jurkat cells over 120 h . A) methylation pattern of *PCDHGA12* in untreated control cells, B) methylation pattern of *PCDHGA12* in decitabine treated cells. This figure shows the same data as 5.18, but shows the read classifications.

5.3.6 Impact of decitabine on *RASSF1* promoter.

To investigate the effect of decitabine on the methylation pattern of the *RASSF1* promoter, *RASSF1* hairpins were prepared according to the protocol detailed in Chapter Three. The hairpin bisulfite sequencing technique was used to demonstrate the methylation changes in the complementary CpG sites of the *RASSF1* promoter.

Similar to the methylation finding of the *PCDHGA12* promoter, the *RASSF1* promoter showed increasing decitabine-induced hemi-methylation at increasing time-points compared to control (Figure 5.20).

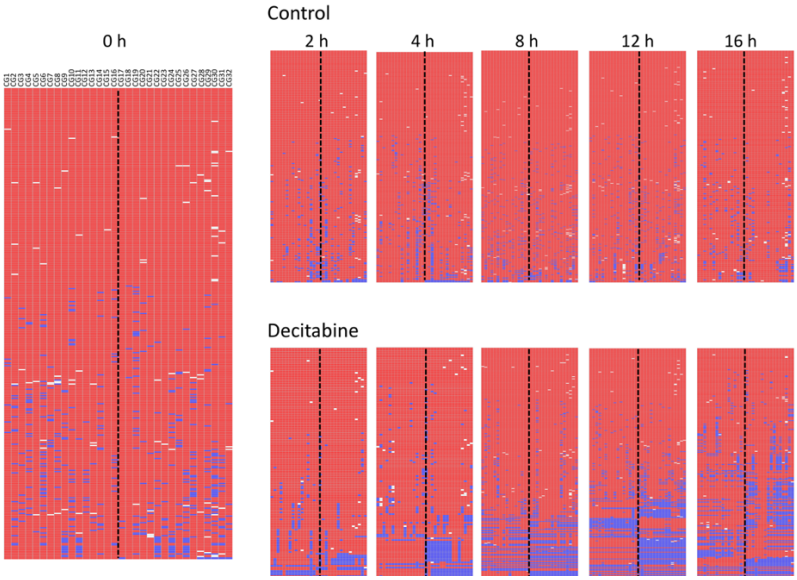


Figure 5. 20. Hairpin based methylation of the *RASSF1* promoter in synchronised Jurkat cells treated with A) decitabine (5 μM) for 24 h compared to B) untreated control cells.

In this experiment, *RASSF1* sequencing results showed that decitabine induced minor hemi-methylation percentages of only 5% and 13% at the 2 h and 4 h timepoints, respectively. Interestingly, the percentage of hemi-methylated reads increased to 29% at 8 h, 40% at 12 h and 55% at 16 h compared to 0% in the control.

5.3.7 Active DNA demethylation in synchronised Jurkat cells induced by decitabine.

The hairpin bisulfite sequencing method was used to show methylation changes in both parental and daughter DNA strands and to explore active demethylation in *RASSF1* promoter. Active demethylation can be interpreted from the hairpin results, and defined as a full removal of DNA methylation pattern of the two strands within an individual hairpin sequence. Surprisingly, decitabine displayed 50% of fully unmethylated reads in *RASSF1* hairpin sequences as early as 2 h and 6 h timepoints compared to control cells (Figure 5.21).The *RASSF1* promoter demethylated differently to the *PCDHGA12* promoter. This may be due to differences in

replication times between these genes. The results are consistent with previous results obtained by Dr. Magrath from the Morison laboratory.

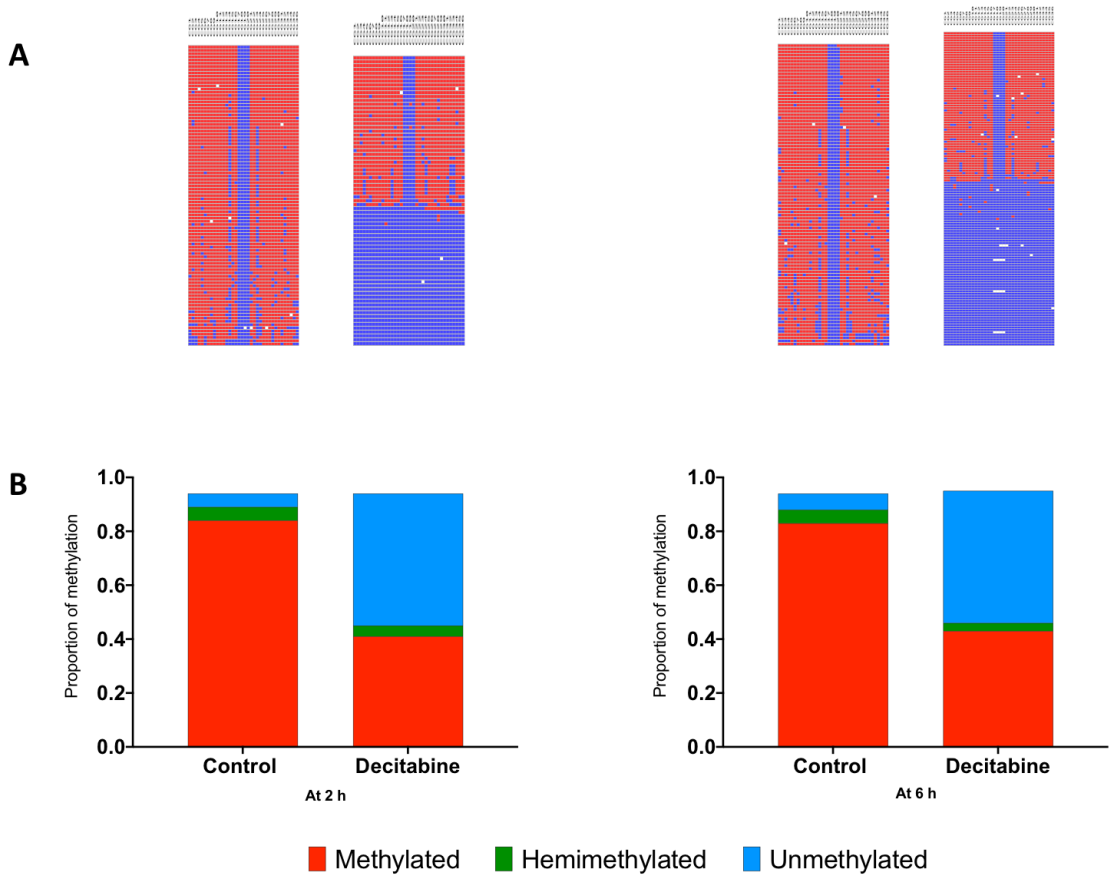


Figure 5. 21. Impact of decitabine in *RASSF1* promoter methylation. A) *BiQ Analyzer* heatmaps for control and decitabine treated DNA samples at 2 h and 6 h post release of synchronisation. B) CpG based methylation of decitabine treated *RASSF1* promoter at 2 h and 6 h post release of synchronisation compared to control.

It is highly likely this demethylation is real. Firstly, contamination was excluded as each read has a hairpin barcode, and all hairpin sequences were unique. And secondly, Jurkat cells take 20 h to complete one cycle of cell division (Fimognari et al., 2002)(Karina O'Connor 2009 unpublished data), so it is highly improbable that Jurkat cells divided twice within 2 h of release from synchronisation. We propose therefore that the parent strand must have undergone active demethylation. This active demethylation was observed in some experiments but not in others; (section 6.4.8) discusses the factors that might affect the rate of active demethylation.

Decitabine may have a direct impact on hemi-methylated strands by activating TET enzymes which enhance iterative oxidation of 5-mC to 5-fC, 5-caC or unmodified cytosine. Alternatively, decitabine could indirectly generate reactive oxygen species (ROS) that induce

epigenetic alterations and apoptosis in leukaemia cells (Shin et al., 2012). This is discussed in more detail in Chapter Eight.

5.3.8 Low decitabine concentrations induce hemi-methylation in *RASSF1* and *PCDHGA12* promoters of Jurkat cells.

Based on the results obtained from viability and toxicity experiment (Section 5.2.4), decitabine was found to be slightly toxic and associated with a reduction in cell viability at higher concentrations. To minimise the cytotoxicity, the concentration of decitabine was decreased from 5 µM to 1 µM. To test whether the lower decitabine concentration has an impact on DNA methylation, the methylation pattern of the complementary CpG sites at each of 16 CpG sites of *RASSF1* and at each of 12 CpG sites of *PCDHGA12* promoters were explored in Jurkat cells treated with 1 µM decitabine. Jurkat cells were synchronised with thymidine and released by adding fresh media loaded with 50 µM dC and low decitabine concentration compared to untreated control. Cells were harvested at 0, 4, 8, 12, 16 and 24 h after release from cell arrest.

Similar to high dose decitabine, when administered at low doses, decitabine induced substantial changes in the methylation levels of *RASSF1* and *PCDHGA12* promoters. This effect was largely attributable to an increase in hemi-methylated hairpin reads compared to control untreated cells (Figure 5.22) (Figure 5.23).

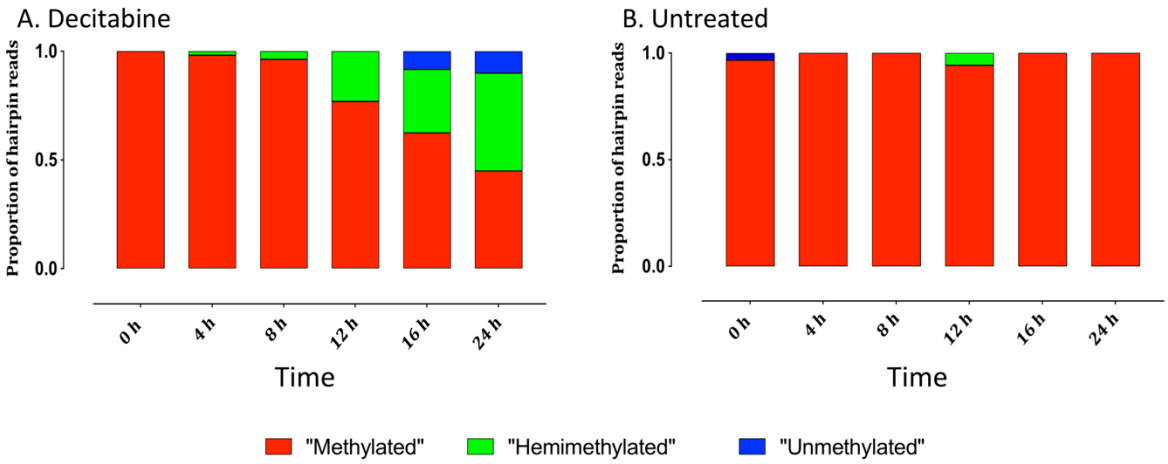


Figure 5. 22. Hairpin based methylation of *RASSF1* promoter in synchronised Jurkat cells treated with A) decitabine (1 µM) for 24 h compared to B) untreated control cells.

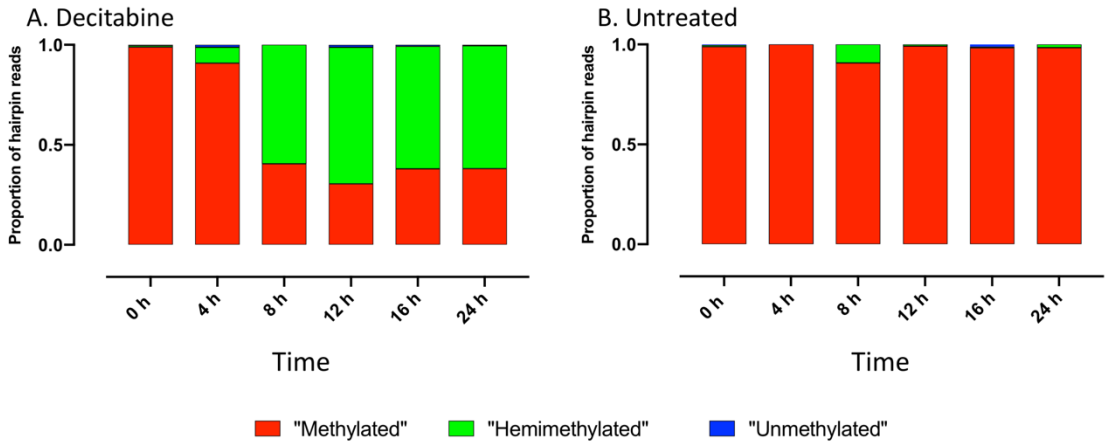


Figure 5. 23. Hairpin based methylation of *PCDHGA12* promoter in synchronised Jurkat cells treated with A) decitabine (1 μM) for 24 h compared to B) control cells.

5.3.9 Impact of decitabine on *PCDHGA12* promoter of Molt4, Nalm6 and HL60 cell lines.

Based on our observation on Jurkat cells, we expanded our scope of analysis by investigating the effect of decitabine on a further three leukaemia cell lines, Molt4, Nalm6 and HL60. The *PCDHGA12* promoter was found to be densely methylated (details in Chapter Six) in these leukaemia cell lines.

Molt4, Nalm6 and HL60 cells were arrested in G1 phase by incubation with 2 mM thymidine for 22 h. Cells were then washed twice and incubated with fresh media containing 0.5-1 μM decitabine and 50 μM dC. Cells were harvested at 0, 2, 4, 8, 12 and 24 h time-points following release from cell cycle arrest. *PCDHGA12* hairpin molecules were prepared for each cell line according to standard protocol (details in Chapter Three).

Decitabine induced demethylation was observed in the Molt4 and Nalm6 cell lines. Decitabine treatment elicited an increase in the proportion of hemi-methylated hairpin reads in both Molt4 and Nalm6 cells (Figure 5.24-25 and Figure 5.26-27) compared to control cells. However, we did not observe any noticeable decitabine-dependent increase in hemi-methylated hairpin reads in HL60 cell line compared to control untreated cells, except the anomalous observation of 10% hemi/unmethylated reads at 16 h (Figure 5.28-29).

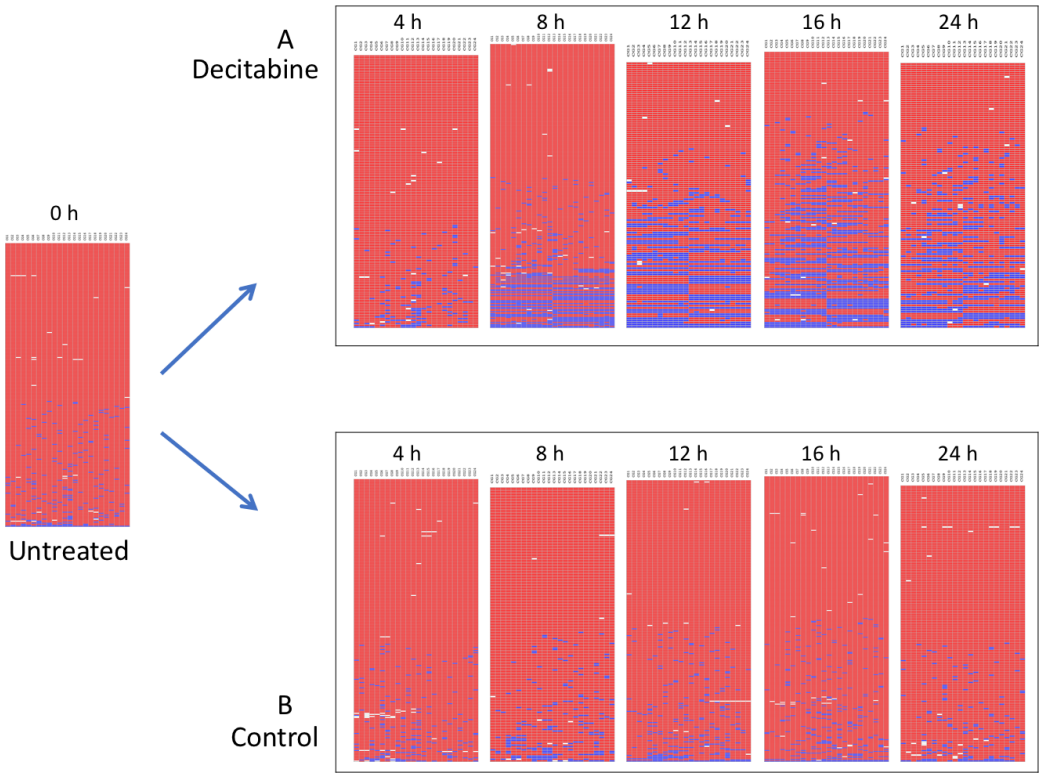


Figure 5. 24. Heatmap output of *BiQ Analyzer* showing the methylation pattern of the *PCDHGA12* promoter in Molt4 cell lines treated with A) 1 µM decitabine for 24 h compared to B) untreated control cells.

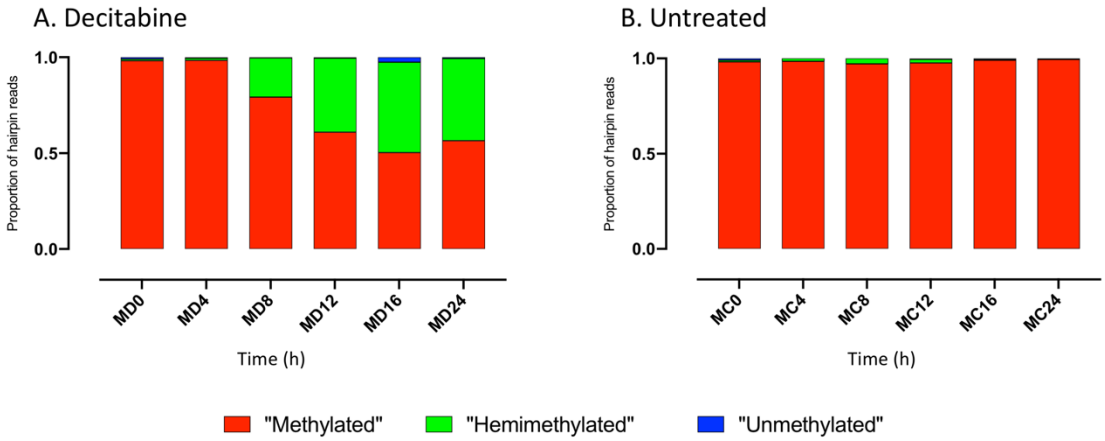


Figure 5. 25. Hairpin based methylation of *PCDHGA12* promoter in synchronised Molt4 cells treated with A) 1 µM decitabine for 24 h compared to B) untreated control cells.

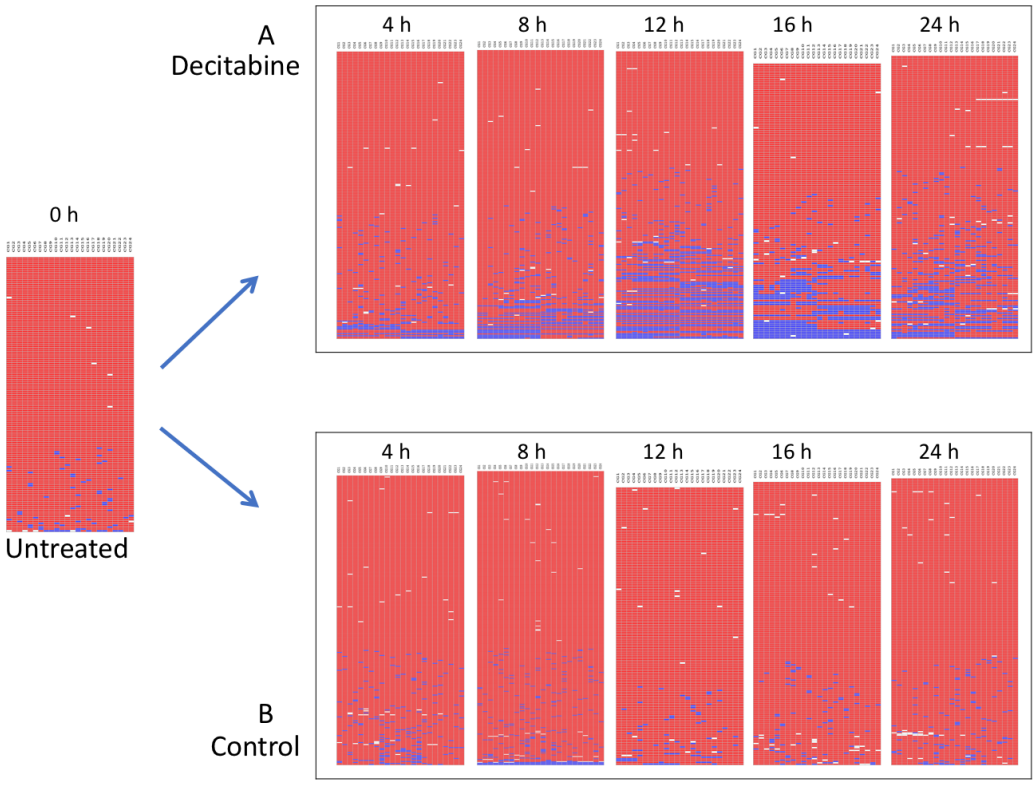


Figure 5. 26. Heatmap output of *BiQ Analyzer* showing the methylation pattern of the *PCDHGA12* promoter in Nalm6 cell lines treated with A) 1 µM decitabine for 24 h compared to B) untreated control cells.

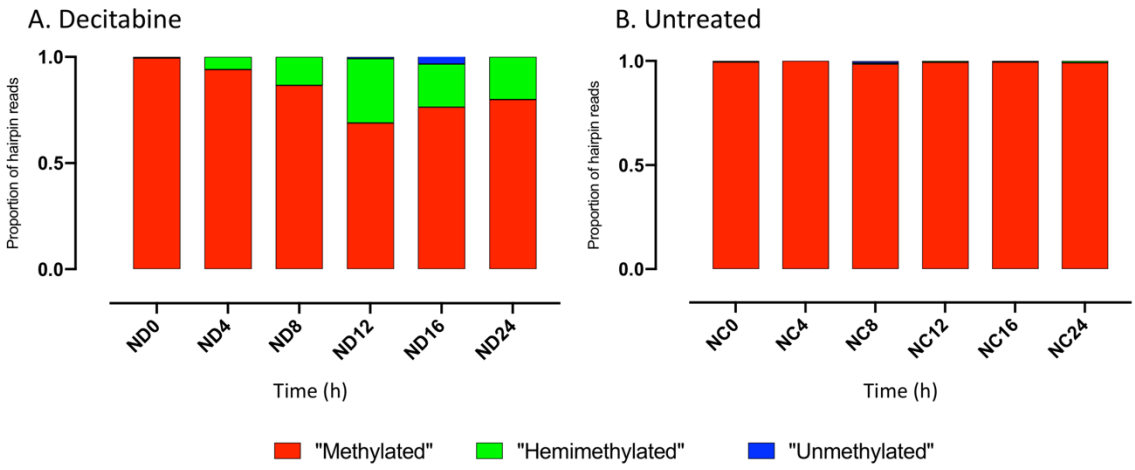


Figure 5. 27. Hairpin based methylation of *PCDHGA12* promoter in synchronised Nalm6 cells treated with A) 1 µM decitabine for 24 h compared to B) untreated control cells.

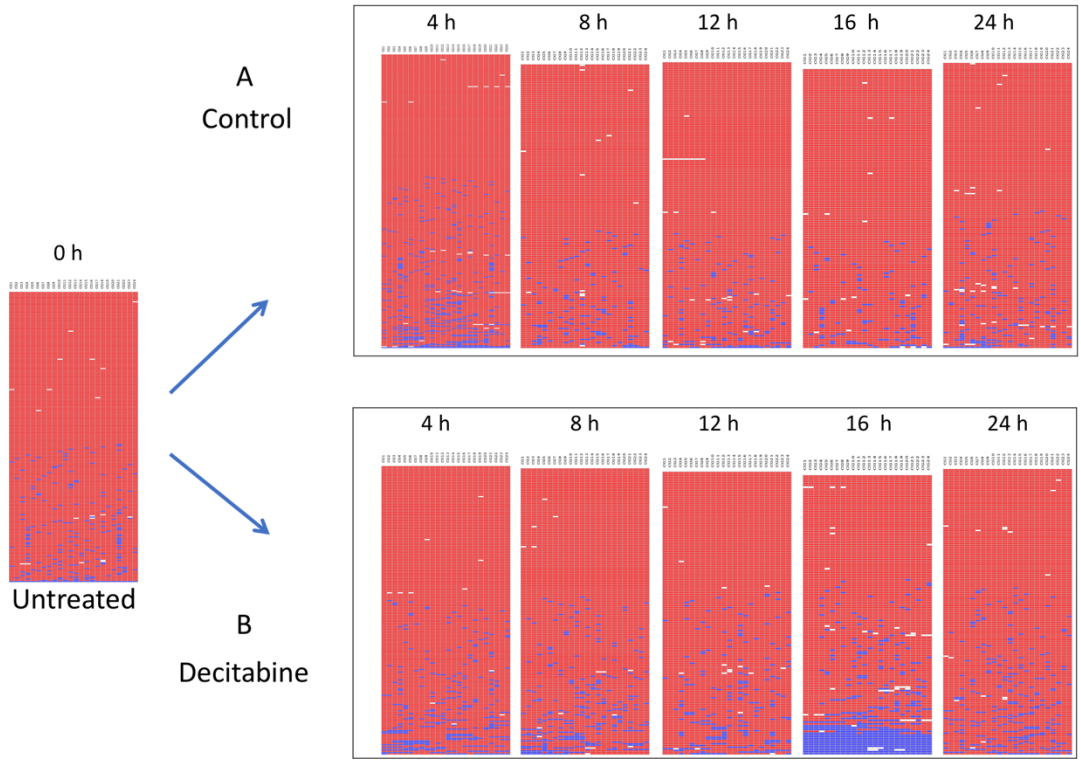


Figure 5. 28. Heatmap output of *BiQ Analyzer* showing the methylation pattern of the *PCDHGA12* promoter in HL60 cell lines treated with A) 0.5 µM decitabine for 24 h compared to B) untreated control cells.

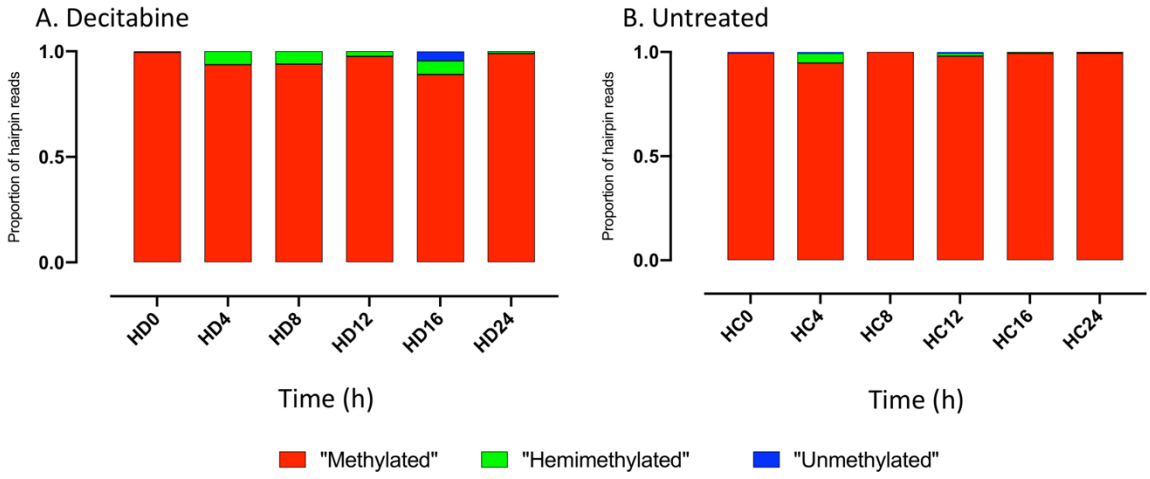


Figure 5. 29. Hairpin based methylation of *PCDHGA12* promoter in synchronised HL60 cell lines treated with A) 0.5 µM decitabine for 24 h compared to B) untreated control cells.

5.4 Validation of decitabine effect using publicly available methylation data.

This part describes the DNA demethylation induced by decitabine and combined ascorbate/decitabine from relevant data extracted from publicly available human methylation data for *RASSF1* and *PCDHGA12* promoters. The publicly available data were analysed by a pipeline (*ABC.RAP* package) developed by a previous PhD student, Abdulmonem Alsaleh from the Morison laboratory. The program is composed of a list of R scripts that utilised basic R functions to identify differentially methylated genes in publicly available data using 450K methylation array analysis. Basically, the *ABC.RAP* pipeline applies Student's T test and delta beta analysis to identify methylated genes that are both statistically significant different and have a minimum absolute difference in DNA methylation between two sets of samples such as treated and untreated control.

Our methylation sequencing results obtained from short and long-term exposure experiments are consistent with the analysis obtained from publicly available human methylation datasets for densely methylated *RASSF1* and *PCDHGA12* promoters. Datasets from Tsai et al. (2012) was examined and analysed and their result was clearly consistent with our hairpin *PCDHGA12* results. They reported a global methylation changes following treatment of three leukaemia cell lines KG1A, KG1, and Kasumi-1 and breast cancer cell lines MCF7 with 10 nM decitabine (Tsai et al., 2012).

Then, methylation data for *RASSF1* and *PCDHGA12* genes were analysed. There was a substantial decrease of DNA methylation in *PCDHGA12* gene treated with 10 nM decitabine over 72 h in the four cell lines (Figure 5.30). The *RASSF1* gene also showed a substantial decrease of DNA methylation in three cell lines treated with 10 nM decitabine over 72 h (data not shown).

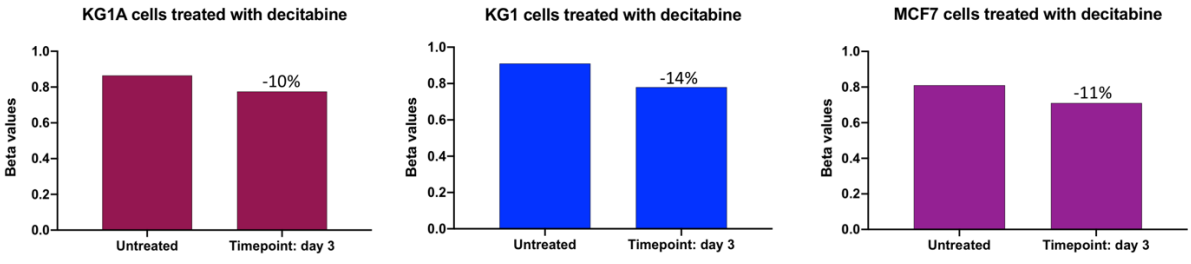


Figure 5. 30. Publicly available data showing effect of decitabine on *PCDHGA12* promoter methylation. DNA methylation changes of *PCDHGA12* in KG1A, KG1 leukaemia cell lines and MCF7 breast cancer cell lines incubated with 10 nM decitabine over 72 h. Data extracted from Tsai et al. 2012.

Kasumi-1 cell lines treated with 10 nM decitabine showed a decrease in DNA methylation of *PCDHGA12* genes during the first three days of incubation, but by the 28th day of decitabine incubation, the methylation levels of the gene approached those observed prior to decitabine treatment (Figure 5.31).

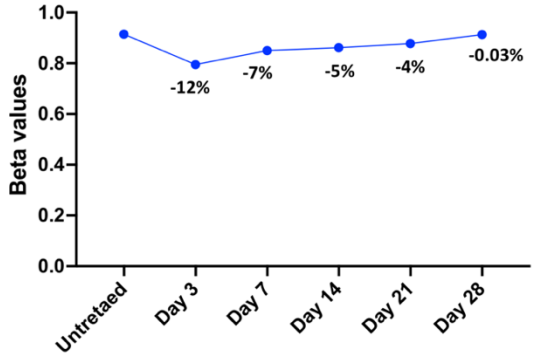


Figure 5. 31. Gene re-methylation in Kasumi-1 leukaemia cell line. Line plots show the proportion of DNA methylation for *PCDHGA12* in Kasumi-1 cells treated with 10 nM decitabine over 28 days. Data extracted from Tsai et al. 2012.

In addition, the *ABC.RAP* R package was used to analyse publicly available data from Öz et al. (2014). U937 acute myeloid leukaemia cells were treated with 100 nM decitabine for 24 h. Genomic DNA was extracted, sheared into 300 bp fragments, and adapters were added and cluster generation was performed on the Illumina cBot and sequenced (Öz et al., 2014). The *ABC.RAP* R package analysis produces two plots. The first boxplot shows the median methylation value of U937 cells treated with 100 nM decitabine compared to control (Figure 5.32), and the second plot shows the mean methylation and the difference in DNA methylation between decitabine treated and untreated control cells) for individual CpG sites within the *RASSF1* and *PCDHGA12* genes arranged from 5′ to 3′ (Figure 2.33).

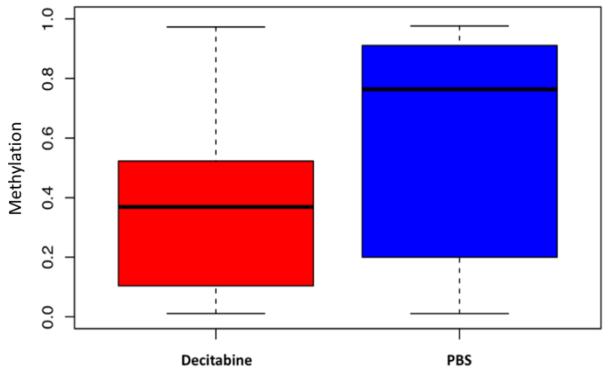


Figure 5. 32. Global methylation level of U937 AML cells treated with 100 nM decitabine for 24 h compared to control cells. The boxes show the median methylation of decitabine treated U937 cells compared to untreated control. Data extracted from Öz et al. 2014.

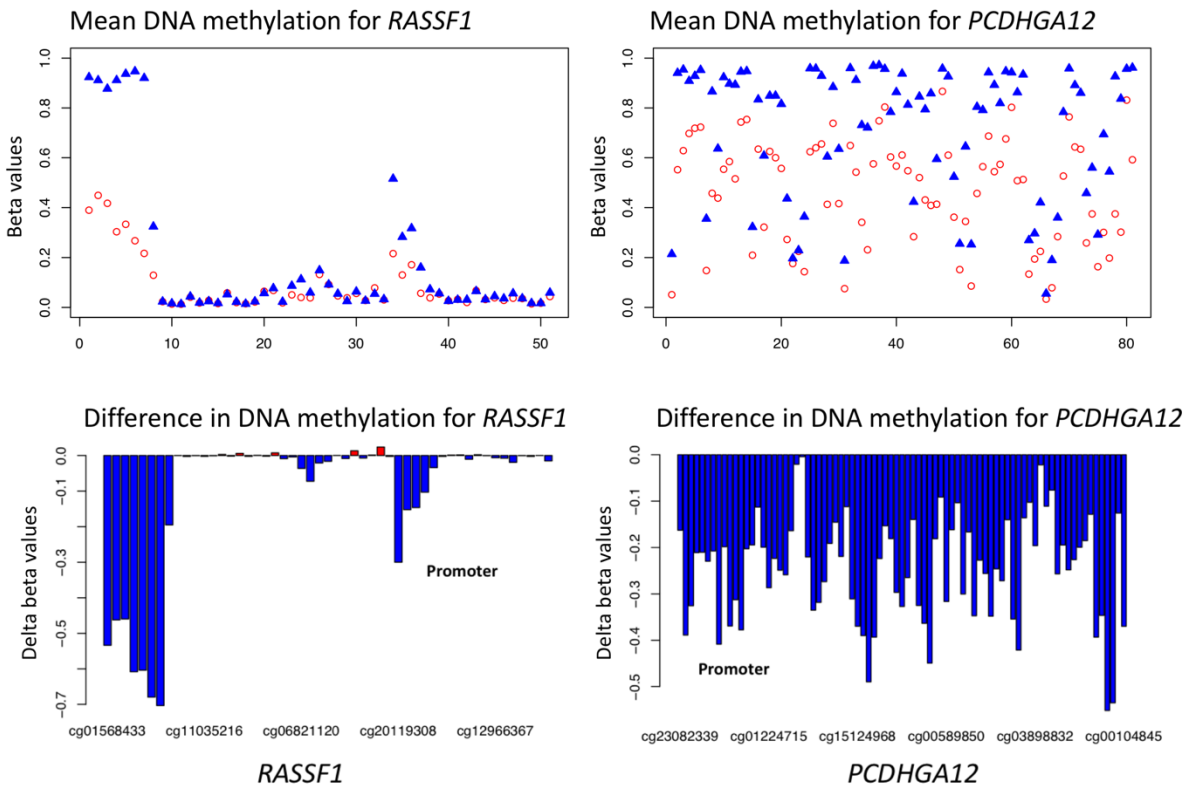


Figure 5. 33. Methylation difference between decitabine-treated U937 cells and untreated control cells. Upper panel: Mean DNA methylation for *RASSF1* and *PCDHGA12* promoters, red circles = decitabine treated and blue triangle = untreated control. Lower panel: The difference in DNA methylation for *RASSF1* and *PCDHGA12*.

Global DNA methylation levels declined from 80% and reached 40% in U937 cells line incubated with 100 nM decitabine for 24 h. The difference in DNA methylation (delta beta values) analysis showed that decitabine had a high demethylation impact on *PCDHGA12* gene, and less on *RASSF1* gene, due to the *RASSF1* gene promoter being largely unmethylated in U937 cells.

In addition, publicly available data extracted from (Imanishi et al., 2017) and analysed by the *ABC.RAP* package shows loss in DNA methylation in U937 cells incubated with 5 μM decitabine for 24 h and 72 h compared to untreated control cells (Figure 5.34) (Figure 5.35).

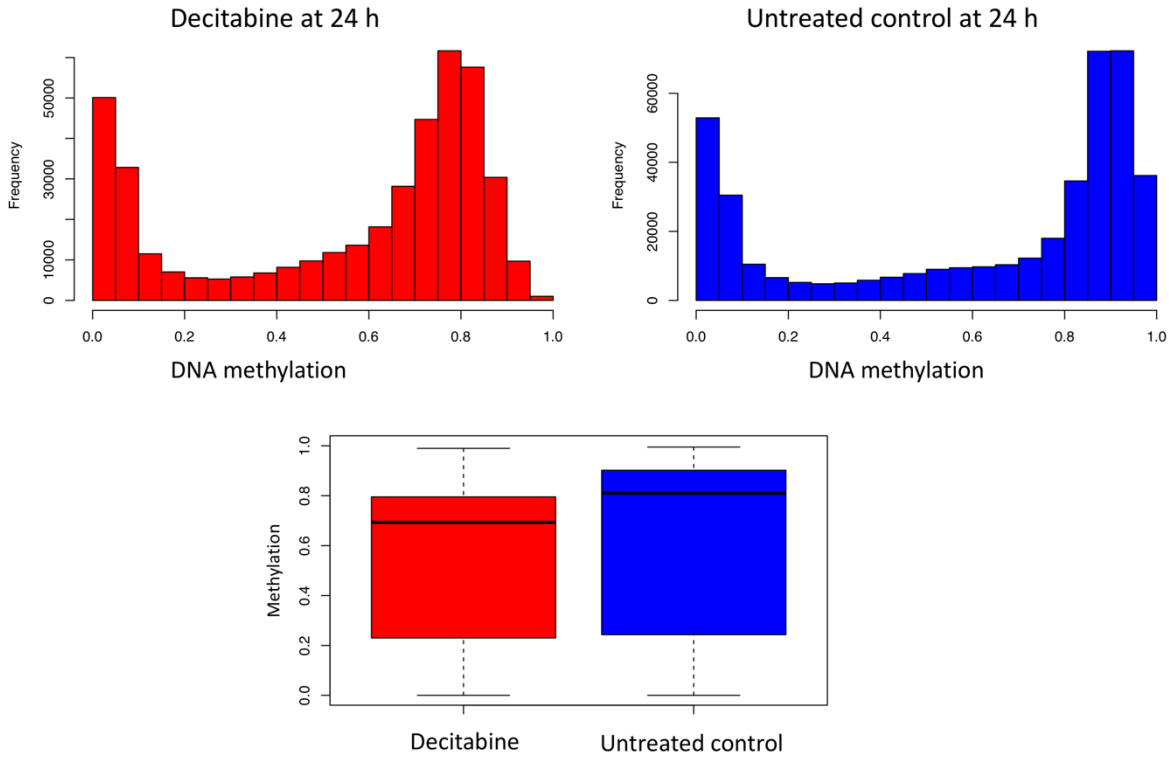


Figure 5. 34. Global methylation level of U937 cells treated with 5 μM decitabine for 24 h compared to untreated control cells. The boxes show the median methylation of decitabine treated U937 cells compared to untreated cells. Data extracted from Imanishi et al. 2017.

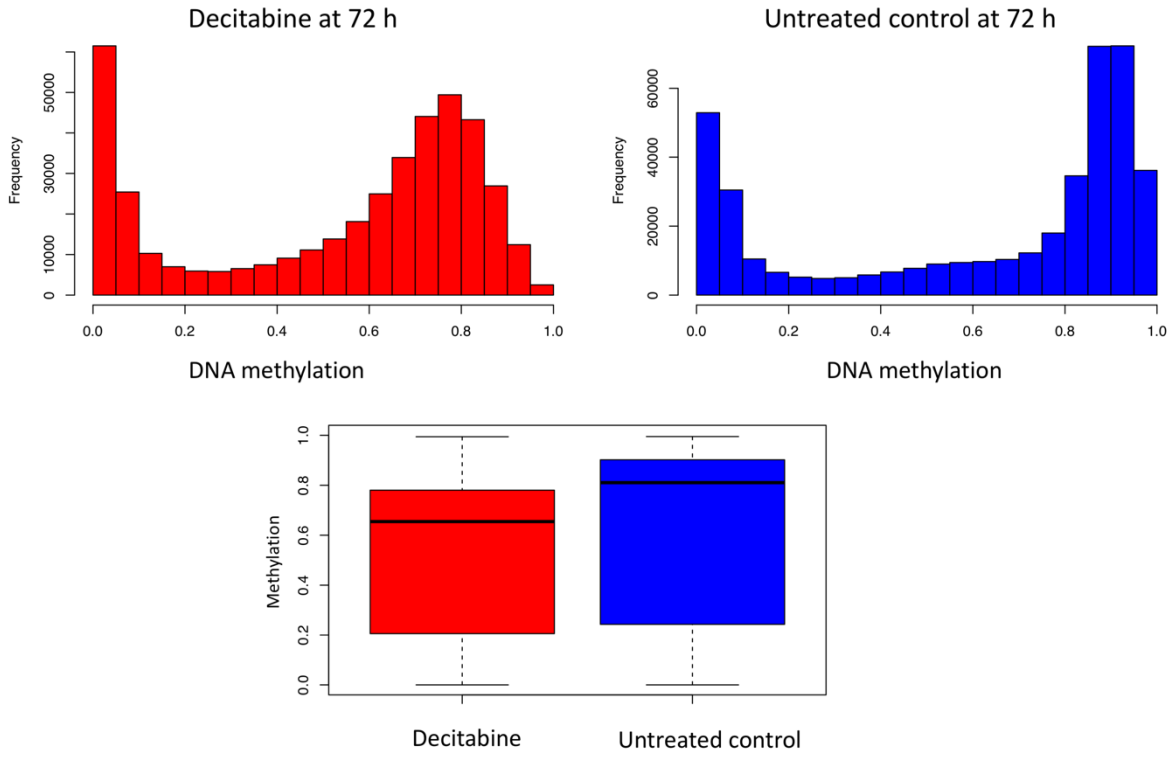


Figure 5. 35. Global methylation level of U937 cells treated with 5 μM decitabine for 72 h compared to untreated control cells. The boxes show the median methylation of decitabine treated U937 cells compared to untreated cells. Data extracted from Imanishi et al. 2017.

5.5 Discussion.

The objective of this chapter was to investigate the impact of decitabine on cell cycle progression and DNA methylation. Development of a novel method linking the two DNA strands together by a unique DNA hairpin linker permitted the study of methylation at complementary CpG sites, allowing the percentage of hemi-methylated strands to be determined. As a cytidine analogue, decitabine incorporates into DNA strands during replication and inhibits the methylation of the newly synthesised DNA strands by stopping the action of DNMT1. It has been shown that decitabine sensitivity depends on the expression of human equilibrative nucleoside transporters in the cell membrane, and that leukaemia cells lacking these nucleoside transporters are resistant to decitabine treatment (Damaraju et al., 2012). Intracellularly, decitabine is activated through a series of phosphorylation steps to form 'decitabine triphosphate' (5-aza-dCTP), which can be incorporated into DNA during DNA replication. This incorporated 5-aza-dCTP will covalently bind to and inactivate DNMT1, preventing maintenance methylation activity, resulting in DNA demethylation (Daskalakis et al., 2010; Oki et al., 2007).

In this study we investigated the effect of different decitabine concentrations on the methylation and cell cycle progression of Jurkat cells. Most of the chemical assays are only allow measuring DNA methylation on one DNA strand, making this impossible to calculate the methylation percentages of complementary strands. The hairpin bisulfite sequencing assays were developed to simultaneously assess the impact of selected decitabine concentrations on the methylation of newly synthesised DNA in multiple leukaemia cell lines. Additionally, the hairpin assay was used to address the presence of hemi-methylation-enhanced active demethylation.

Using the hypermethylated *RASSF1* and *PCDHGA12* promoters, we show clearly that decitabine induced rapid DNA demethylation in the *RASSF1* promoter as early as 2 h and 6 h post release from synchronisation. The hairpin method efficiently monitored the changes in DNA methylation by displaying fully hemi-methylated and unmethylated individual hairpin reads in cells treated with decitabine at different timepoints, and proved that the selected decitabine concentrations significantly inhibited DNA methylation during DNA synthesis. Surprisingly, we observed completely unmethylated hairpin reads at 6 h after decitabine treatment. Normally, we would expect to observe complete unmethylation only after two consecutive cell cycles via the blockage of DNMT1, which is not possible within 6 h of release from G1 arrest. A study based on gene expression, suggested the possibility of decitabine

induced active demethylation (Gius et al., 2004); however their results were not repeated and the mechanism has not been outlined.

In this study, we investigated the effect of decitabine on the cell cycle kinetics of synchronised Jurkat cells. We performed flow cytometric analysis to study the cell cycle progression of synchronised Jurkat cells following a single decitabine treatment. Flow cytometry analysis showed that synchronised cells treated with decitabine for 24 h progressed through the cell cycle slower than controls and that decitabine-treated cells remained in S-phase longer (Figure 5.32 b,c). This observation corroborates previous published work stating that the incorporation of decitabine to DNA caused stalled replication forks and changes in the cell cycle of cells treated with decitabine (Palii et al., 2008).

Chapter Six: The roles of ascorbate in modulating DNA methylation.

6.1 Introduction.

In cancer, locus-specific hypermethylation of DNA promoters leads to suppression and loss of function of tumour suppressor genes (Jones & Baylin, 2007). Methylation of DNA is controlled by the activity of DNA methyltransferase enzymes (DNMTs) which catalyse the addition of a methyl group to the C5 position of cytosine at CpG dinucleotide sites. Initially, the methylation was considered as non-reversible until the discovery of TET (Ten-Eleven Translocation) dioxygenases that act as the key mediators of active demethylation. TET enzymes are Fe^{2+} and 2-oxoglutarate-dependent dioxygenases that modify DNA methylation by initiating active demethylation through oxidation of 5-methylcytosine to 5-hydroxymethylcytosine (Tahiliani et al., 2009). Furthermore, TET enzymes catalyse the iterative oxidation of 5-hydroxymethylcytosine to 5-formylcytosine and 5-carboxylcytosine which are ultimately replaced with unmodified cytosine by the activity of thymine-DNA glycosylase (TDG) and base excision repair (Ito et al., 2011). 5-hydroxymethylcytosine deficiency is a hallmark epigenetic feature of many cancers (Delatte et al., 2014; Lian et al., 2012).

Ascorbate serves as a co-factor for TET enzymes by regenerating the essential cofactor, Fe^{2+}. Ascorbate catalyses the redox reaction of inactive Fe^{3+} to the catalytically active Fe^{2+} (Young et al., 2015). Interestingly, ascorbate dramatically modifies the status of DNA methylation of mouse embryonic fibroblasts (MEFs) by generation of 5-hydroxymethylcytosine (Minor et al., 2013).

Decitabine is an epigenetic therapeutic compound that is used in the treatment of myelodysplastic syndrome (MDS) and acute myeloid leukaemia (AML). Decitabine acts as a DNMT inhibitor by incorporating into the newly synthesised DNA strands and binding DNMT1 during cell division. Upon binding, DNMT1 is degraded resulting in loss in activity and hypomethylation.

In this study, we investigated the role of ascorbate alone or in combination with decitabine on locus-specific DNA methylation in Jurkat cells. Following previous reports of ascorbate

effects, we aimed to determine whether additive or synergistic effects on DNA methylation can be achieved when exposing cells to a DNMT1 inhibitor combined with ascorbate.

6.1.1 Hypothesis and experiments design.

The experiments were designed to address the following hypotheses:

1. In synchronised Jurkat cells, the complementary CpG sites of tumour suppressor gene promoters *RASSF1* and *PCDHGA12* are always densely methylated.
2. Due to inhibition of DNMT1, CpG demethylation of the newly synthesised daughter strands should be observed in decitabine-treated Jurkat cells.
3. In ascorbate-treated Jurkat cells, sporadic complementary CpG demethylation should be noted due to upregulation of TET enzymes.
4. In Jurkat cells treated with a combination of ascorbate and decitabine, rapid CpG demethylation of both parental and daughter strands should be observed due to enhancement of TET activity, by the presence of hemi-methylated DNA and the upregulation of TET enzymes.

6.1.2 The main aims of this chapter.

1. To investigate the short and long-term effects of combined decitabine and ascorbate on gene specific methylation.
2. To develop a novel method to measure the percentages of hemi and demethylation in newly synthesised DNA strands.
3. To compare the effects of different reagents on methylation between synchronised and unsynchronised leukaemia cell lines.
4. To determine the effect of combined decitabine and ascorbate on cell cycle progression.
5. To estimate the effects of different decitabine concentrations and ascorbate concentrations on DNA methylation.

6.2 Quantification of intracellular ascorbate concentrations.

6.2.1 Preparation for ascorbate experiments.

To examine the effect of ascorbate on upregulating TET activity, ascorbate intracellular levels were measured in Jurkat cells in culture. Jurkat cells were treated with different ascorbate concentrations (250 or 500 μM) during synchronisation (18 h). Figure 6.1 shows the intracellular ascorbate concentration in Jurkat cells loaded with 250 μM ascorbate overnight

during a thymidine block. The cells were not refreshed with ascorbate when media was replaced at time 0 h.

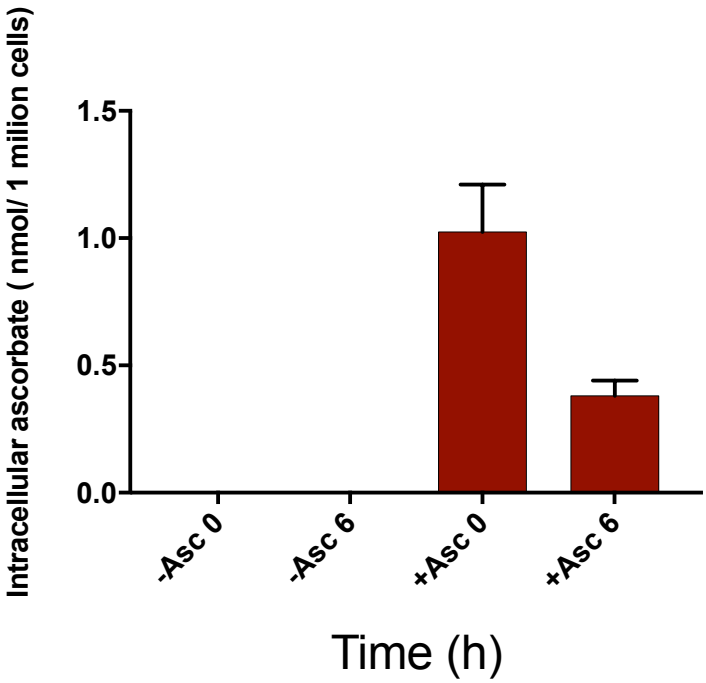


Figure 6. 1. Intracellular ascorbate concentration. Jurkat cells loaded with 250 μM ascorbate for 18 h during thymidine block, were not refreshed when media replaced. Data are means ± SEM of two independent experiments.

The Jurkat cells showed lower intracellular ascorbate concentration at the 6 h timepoint and failed to reach the intracellular ascorbate concentration levels achieved by Kuipers et al., (2014). The concentration of ascorbate was increased to 500 μM overnight during the thymidine block, and ascorbate was included in the media when the cells were released to cell cycle (Figure 6.2). Interestingly, in Jurkat cells treated with 500 μM ascorbate the intracellular ascorbate concentrations were increased significantly and cells maintained these levels at 6 h. In general, culture media is ascorbate deficient and hence cancer cells growing in culture reflect an ascorbate deficiency state. In accord with this, intracellular ascorbate concentration is undetectable in untreated control Jurkat cells (Figures 6.1 and 6.2).

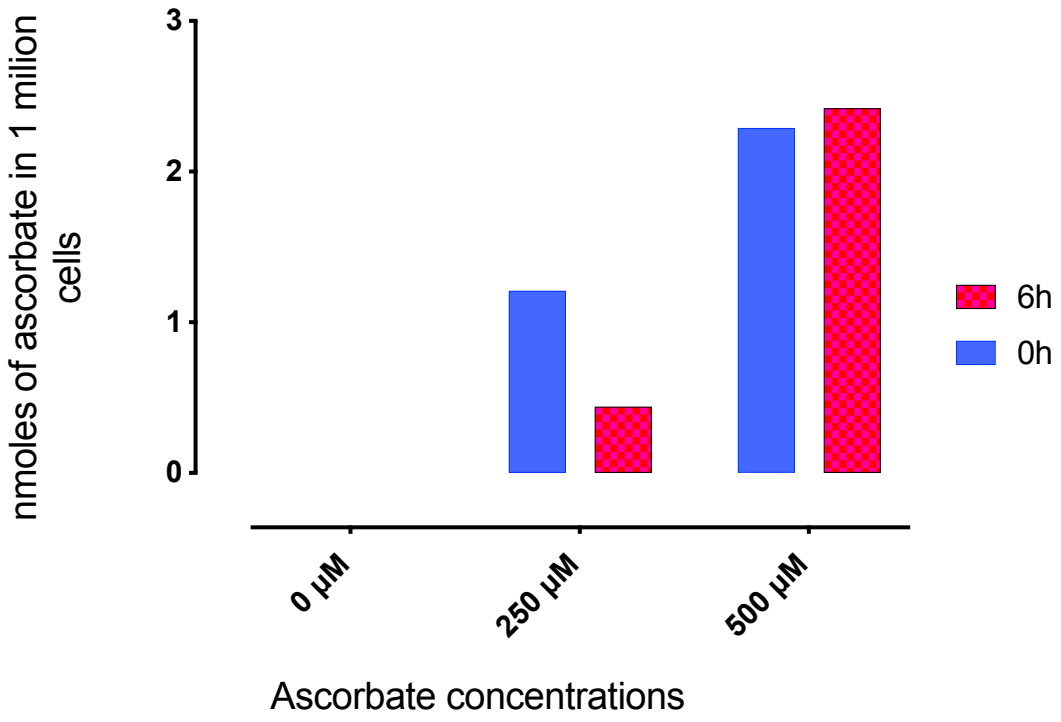


Figure 6. 2. Intracellular ascorbate concentration. Jurkat cells loaded with 0, 250, 500 µM ascorbate for 18 h during thymidine block. Cells were refreshed with ascorbate at a time 0 h.

Next, the total and reduced forms of intracellular ascorbate concentrations in Jurkat cells treated with ascorbate was measured.

Cells were preloaded with 500 µM ascorbic acid for 18 h, cells were refreshed with fresh media loaded with 500 µM ascorbic acid at time 0 h. Treated and untreated control cells were harvested at different timepoints 0 h, 4 h , 8 h and 12 h (Figure 6.3). Cells were collected and washed with PBS, and the cell pellets were transferred to laboratory of Professor Vissers (University of Otago, Christchurch) for intracellular ascorbate measurement according to standard protocol (details in Chapter Three).

Cell culture media contains no ascorbate, and supplementation optimises intracellular total and reduced ascorbate concentrations. We maintained sufficient intracellular ascorbate levels in the cells; both total and reduced ascorbate forms were increased at increasing timepoints after ascorbate treatment. The reduced form of ascorbate reduces Fe^{3+} to Fe^{2+} which can act as a co-factor to enhance TET activity.

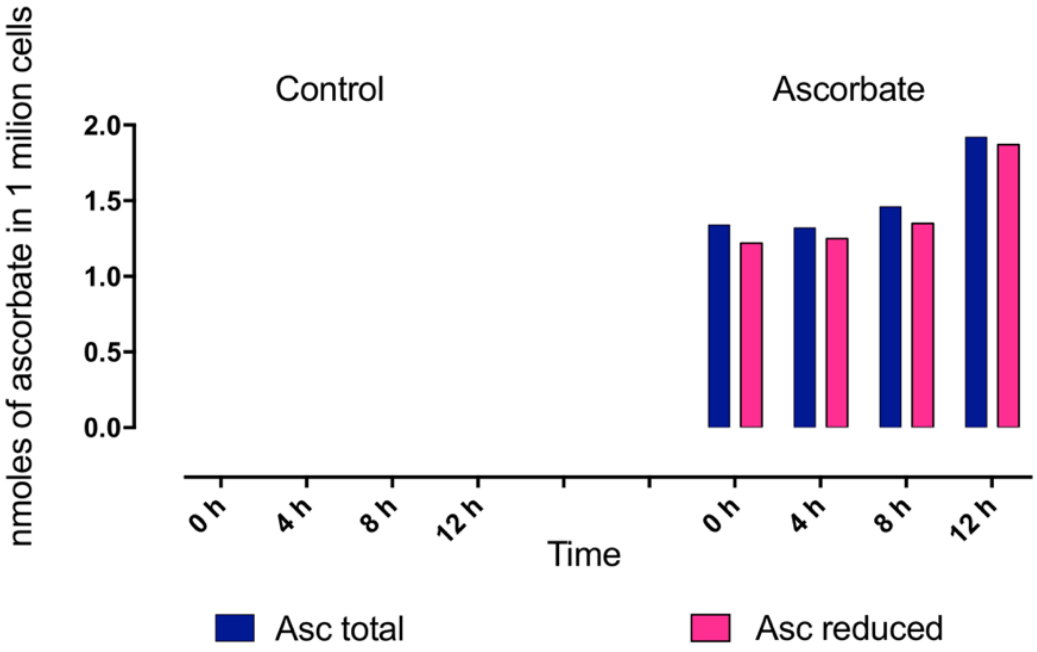


Figure 6. 3. Intracellular total and reduced ascorbate concentrations. Jurkat cells were loaded with 0 or 500 μM ascorbate for 18 h during thymidine block. Cells were refreshed with 0 or 500 μM ascorbate at the time of release from thymidine.

6.3 Cell viability following exposure to ascorbate/decitabine.

6.3.1 Cell cycle analysis.

Jurkat cells were grown to a concentration of 1 million/mL and synchronised by thymidine block for 18 h. To investigate the effect of ascorbate or combined ascorbate and other reagents on cell cycle progression, the treated Jurkat cells were collected, fixed with 70% ethanol and stained with propidium iodide (PI) and analysed using flow cytometry to measure the DNA content at different time points (details in Chapter 5). The experiment was conducted three times at different timepoints.

Short term exposure to 250 and 500 μM ascorbate or in combination with decitabine had no marked effect on cell cycle progression compared to untreated control cells (Figure 6.4 a, b, c). Interestingly, growth arrest with an increase in the number of Jurkat cells in S phase were seen following 24 h incubation with ascorbate and decitabine, compared to untreated control (Figure 6.4 d). This suggests that incorporation of decitabine into the DNA of treated cells delayed DNA replication and cells remained longer in S phase, which resulted in increased cell population in S phase compared to untreated control.

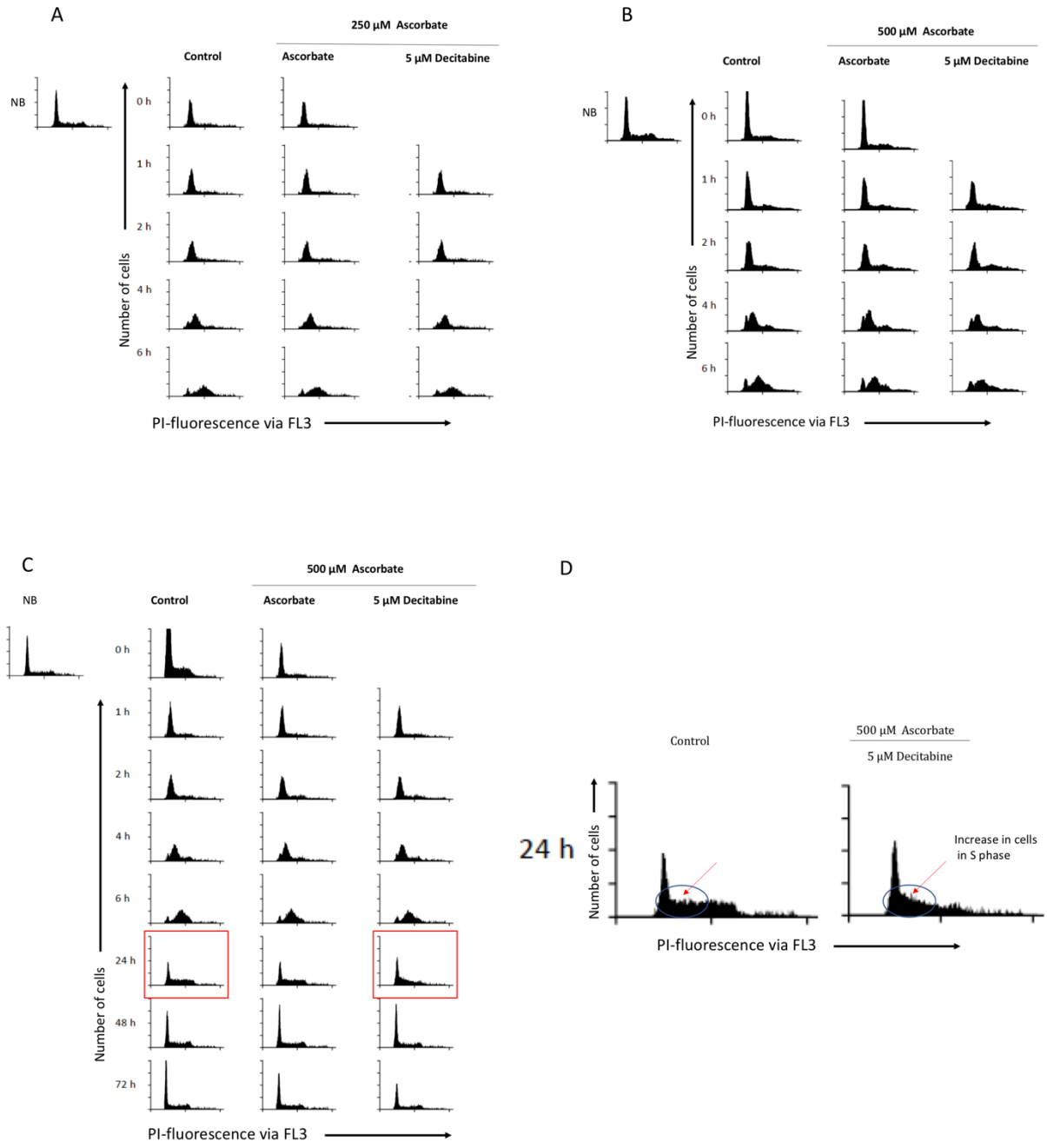


Figure 6. 4. Cell cycle analysis of synchronised Jurkat cells. A) Cell cycle progression of cells treated with 250 µM ascorbate and combined ascorbate with decitabine for up to 6 h compared to controls; B) Cell cycle progression of cells treated with 500 µM ascorbate and combined ascorbate with decitabine for up to 6 h; C) Cell cycle progression of Jurkat cells treated with 500 µM ascorbate and combined ascorbate/decitabine for up to 72 h; D) Enlargement of the untreated control and combined ascorbate/decitabine figures from the labelled figure c. All cells were supplemented with 50 µM dC at the time of release from the thymidine block (time 0 h).

6.3.2 Induction of cell cycle arrest and inhibition of cell division was enhanced by combined treatment of decitabine and ascorbate.

To determine the effect of combined decitabine and ascorbate treatment on the cell cycle progression and cell division, four leukaemia cell lines were tested. The replication time of the four cell lines was different, and to ensure that all were arrested in G1 phase, the duration of synchronisation by thymidine was increased to 22 h. Jurkat, Molt4, Nalm6 and HL60 were synchronised for 22 h by thymidine block, and cells were released with fresh media supplemented with 50 µM dC. For the 'ascorbate' group, ascorbate was added to a final concentration of 500 µM during synchronisation and to the media at thymidine block release.

Based upon the decitabine toxicity results above (section 5.4.4), the concentration of decitabine was lowered from 5 µM to 1 µM for Jurkat, Molt4 and Nalm6 cells and to 0.5 µM for the HL60 cell line. The effects of decitabine and ascorbate on cell viability were compared in these of four cell lines.

Decitabine reduced the number of viable cells at 24 h by 20% in three cell lines (Jurkat, Molt4 and Nalm6) and by 10% in HL60. However, the cytotoxic effect was enhanced with the addition of 500 µM of ascorbate, with enhancement of 45% in Jurkat cells and approximately 20% in Molt4 and Nalm6. In contrast, No additional cytotoxic effect was observed in HL60 cells with 500 µM ascorbate (Figure 6.5).

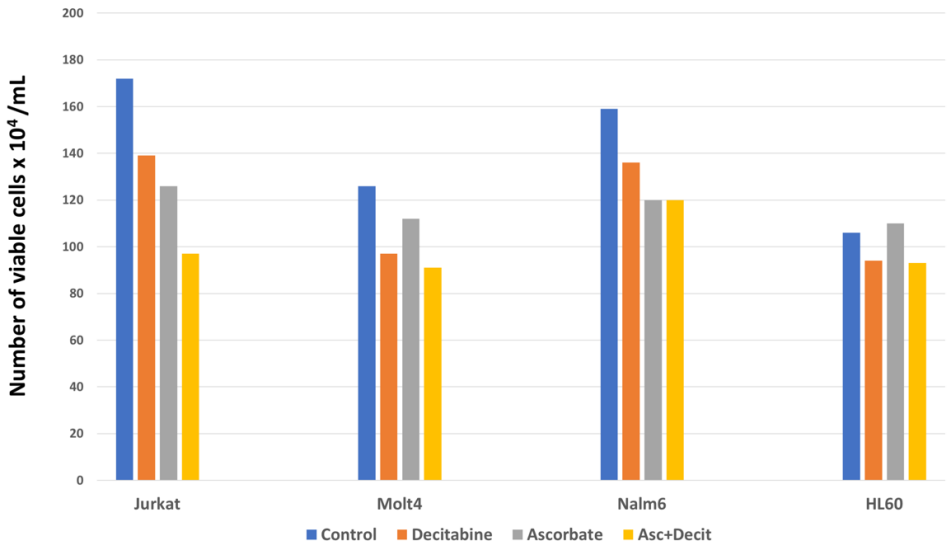


Figure 6. 5. Effect of decitabine and ascorbate on cell viability. Cell survival measurement following incubation with different reagents. Cell survival was determined by trypan blue staining. Values are means of three counts. The starting number of cells for Jurkat, Molt4 and Nalm6 is 0.8 x 10^6 cells/mL and for HL60 is 0.6 x 10^6 cells/mL.

6.4 DNA methylation profile following ascorbate/decitabine treatments.

6.4.1 Impact of different ascorbate concentrations on the DNA methylation of *PCDHGA12* promoter.

To study the impact of different ascorbate concentrations on promoter DNA methylation, synchronised Jurkat cells were incubated with different ascorbate concentrations during cell cycle progression, and collected at -18 (unblocked), 0, 4, 8, 12 and 24 h.

Treatment with 0, 100, 250, 500 µM ascorbate at different time points had little effect on DNA methylation of *PCDHGA12* promoter compared to untreated Jurkat cells, the figure below shows the mean methylation from the hairpin bisulfite MiSeq results for *PCDHGA12* promoter in synchronised Jurkat cells (Figure 6.6).

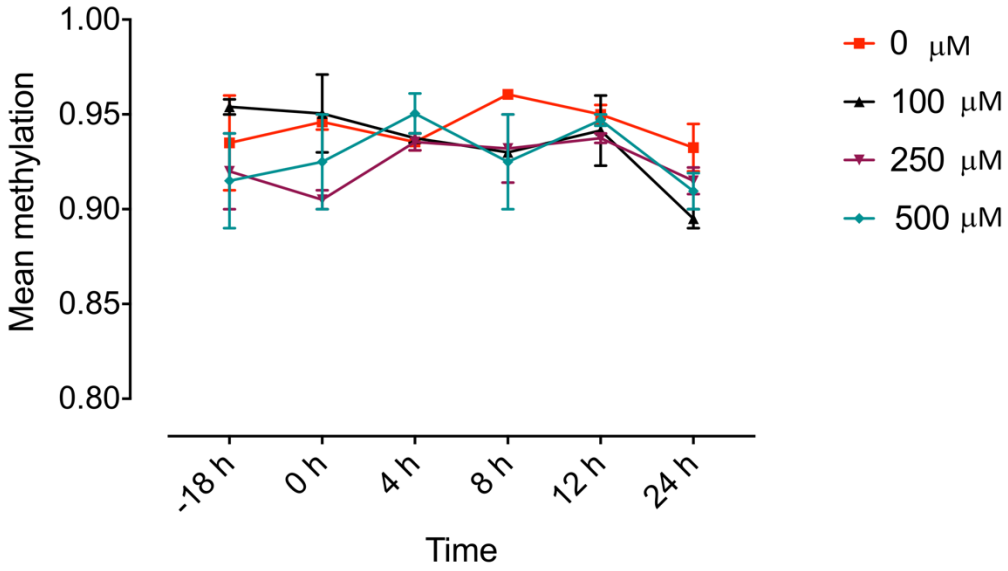


Figure 6. 6. Impact of ascorbate on 5-mC percentages of PCDHGA12 promoter. Jurkat cells were exposed to increasing concentrations of ascorbate for 24 h. Effects of different ascorbate concentrations on the mean methylation of *PCDHGA12* promoter were assessed.(-18 unblocked, 0 blocked +18).

Our results were consistent with a previous study by Gerecke et al. which demonstrated that treatment with ascorbate alone had no significant effect on global 5-methylcytosine levels in HCT116 cells (Gerecke et al., 2018).

6.4.2 Visualisation of hairpin MiSeq sequencing data.

We designed the hairpin bisulfite sequencing assay to provide information about the methylation status of the complementary DNA strands as the position of the hairpin barcode sequence generates an axis of two symmetrical hairpin arms. To visualise this methylation

status, the binary methylation data of linearised hairpin sequences were extracted from the results file of *BiQ Analyzer* HT. Each sequence contains the aligned CpG sites across the arms of the hairpin molecule. The methylation of CpG sites at each arm of the hairpin are represented by lollipops and plotted next to each other (black lollipops are methylated and white are unmethylated) (Figure 6.7). To evaluate the methylation status of each hairpin read, hairpin sequences were folded to superpose the complementary CpG sites (details in Chapter 4), and the percentages of methylated, hemi-methylated or unmethylated complementary CpG sites were calculated (Figure 6.7). The CpG sites were classified as methylated (^{m}CpG/^{m}CpG) if the two complementary CpG sites were methylated; hemi-methylated (^{m}CpG/CpG or CpG/^{m}CpG) if one of the complementary CpG site was methylated and the other is unmethylated; or unmethylated (CpG/CpG) if the two complementary CpG sites were unmethylated.

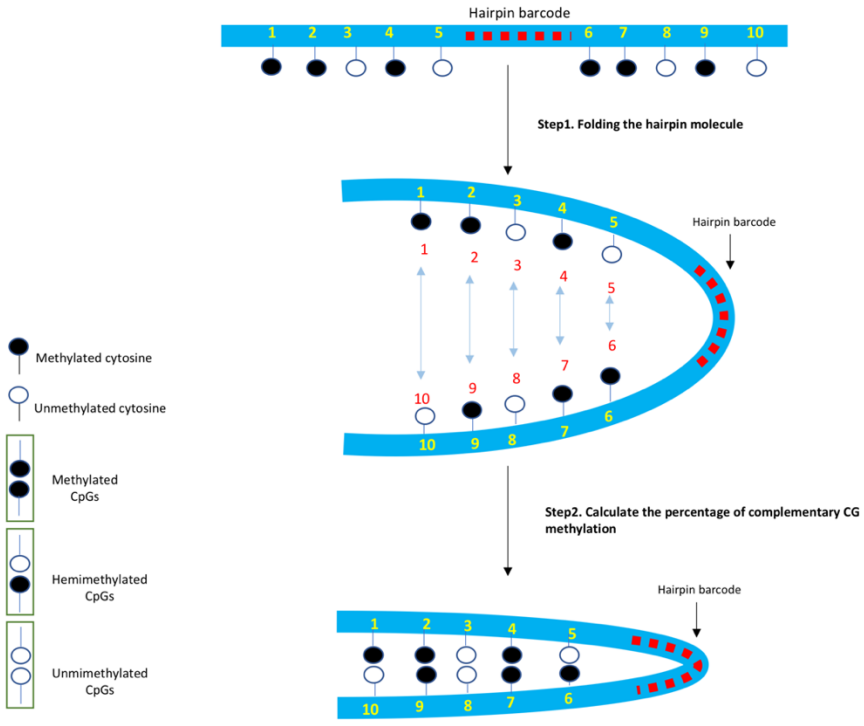


Figure 6. 7. Schematic visualisation of the hairpin molecule.

6.4.3 Surprising active DNA demethylation of *RASSF1* promoter caused by decitabine and ascorbate.

The main aim of this experiment was to confirm the surprising active demethylation resulting from decitabine treatment observed using a very low throughput technique by the Morison laboratory. The second aim was to investigate the role of TET enzymes in initiating active demethylation.

Notably, the control samples in these experiments displayed stable hypermethylation of RASSF1 and PCDHGA12 promoters with little sporadic occurrence of unmethylated CpGs. To explore the DNA methylation changes in complementary CpG sites, hairpin DNA molecules were prepared from synchronised Jurkat DNA treated with different reagents (ascorbate, decitabine and combined ascorbate/decitabine), and the percentages of methylated, hemi, and unmethylated hairpin sequences were calculated.

Briefly, Jurkat cells were arrested in G1 phase by single thymidine block, divided into two groups and incubated with or without 500 μM ascorbate. Cells were initiated into cell cycle progression by incubation with 50 μM dC and simultaneously treated with 5 μM decitabine, 500 μM ascorbate or combinations of decitabine and ascorbate. Cells were harvested for genomic DNA at 0, 2, 4 and 6 h timepoints. *RASSF1* hairpin molecules were prepared using the modified hairpin technique (detailed in Chapter 3).

The experiment was designed to explore active demethylation by monitoring the DNA methylation changes across the complementary DNA strands of the *RASSF1* promoter using the hairpin technique. Surprisingly, the *RASSF1* promoter treated with decitabine and combined ascorbate/decitabine exhibited a massive reduction in DNA methylation at 2 and 6 h, by 49% and 67% respectively (Figure 6.8). Decitabine is known as a specific DNMT1 inhibitor that generates unmethylated daughter strands during DNA replications. We were expecting to see an increase in the hemi-methylated reads in cells treated with decitabine. The substantial change in DNA methylation was due to a complete demethylation of both DNA strands as early as 2 h and 6 h after release from synchronisation. It is most likely that the rapid effects of decitabine skipped the appearance of hemi-methylation status during DNA replication, by rapidly inducing demethylation of the parent strands.

Interestingly, adding ascorbate to decitabine enhanced demethylation, with the *RASSF1* promoter showing a dramatic loss of methylation in ascorbate (62%) and combined ascorbate/decitabine treated (78%) cells at 6 h. The results displayed fully unmethylated hairpin reads, in which parental strands were demethylated. As synchronised Jurkat cells require 8-10 hours to complete interphase (S and G2 phases) and then proceed to cell division (Fimognari et al., 2002), it is improbable that Jurkat cells had replicated their DNA twice in 6 h; therefore the parent strands must have undergone active demethylation.

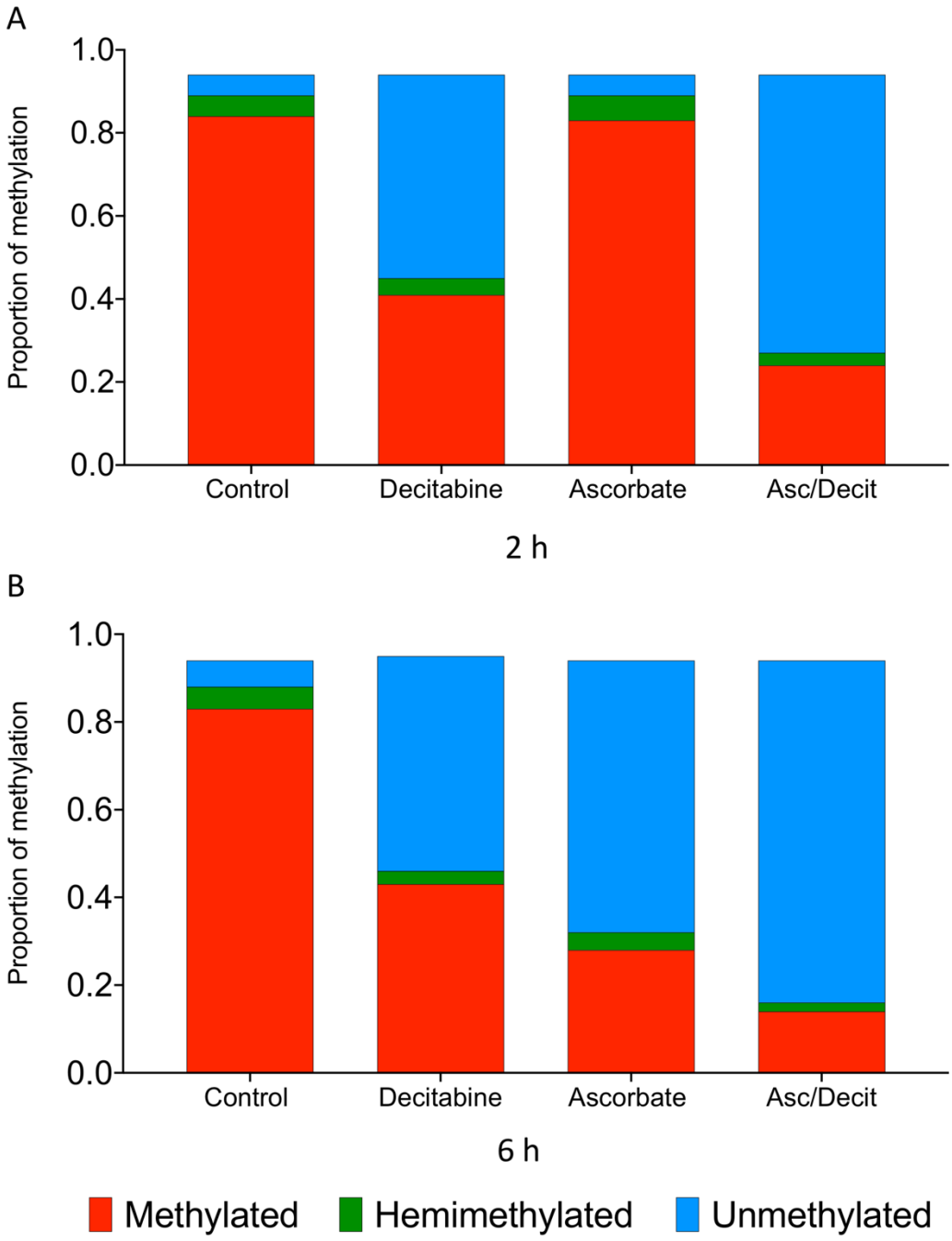


Figure 6. 8. CpG based methylation of *RASSF1* promoter at 2 and 6 h. Synchronised Jurkat cells were treated with decitabine, ascorbate and combined ascorbate/decitabine, or untreated. A) 2 h timepoint: Loss of DNA methylation in cells treated with decitabine and combined ascorbate/decitabine, B) 6 h timepoint: dramatic loss in DNA methylation in Jurkat cells treated with ascorbate and combined ascorbate/decitabine compared to untreated control.

To confirm these surprising demethylation results, the experiments were repeated with a single dose treatment of ascorbate during thymidine block, and the hypermethylated *PCDHGA12* promoter was examined.

6.4.4 Short-term treatment of single dose ascorbate/decitabine induced active demethylation in *PCDHGA12* promoter of Jurkat cells.

To test the ability of ascorbate to upregulate TET activity and induce active DNA demethylation in Jurkat cells, we assessed the methylation changes in *PCDHGA12* promoter of Jurkat cells treated with a single dose of ascorbate during cell cycle synchronisation. *PCDHGA12* hairpin molecules were prepared from DNA from synchronised Jurkat treated with a single dose of 500 µM ascorbate alone during thymidine block or combined with 5 µM decitabine at time 0 h. After synchronisation, cells were refreshed with new media without adding ascorbate. Cells were divided into four groups; control, ascorbate, decitabine and the combination of ascorbate/decitabine. Cells were initiated into cell cycle by supplementation with 50 µM dC. The treatments and the tissue culture work were done by Dr Karina O'Connor in (Prof Hampton laboratory, Christchurch), cell pellets were transported to the Morison laboratory for sequencing and further analysis. DNA extraction, digestion, ligation, bisulfite conversion and two sets of PCR for MiSeq sequencing were done in Morison laboratory.

Figure 6.9 shows the relative proportions of methylated, hemi-methylated and unmethylated hairpin reads within the *PCDHGA12* promoter region. At 6 h after treatment with decitabine, 15% of reads showed unmethylation and 15% showed hemi-methylation. Hemi-methylation was also observed at 2 and 4 h. Cells preloaded with ascorbate alone during thymidine block (for 18 h) showed 13% unmethylated reads in the *PCDHGA12* promoter. Furthermore, ascorbate treated cells showed and an increase in the percentage of unmethylated reads in the *PCDHGA12* promoter by 13%. Furthermore, 12-14% hemi-methylated and unmethylated reads, respectively, were seen at 6 h after release from cell synchronisation (Figure 6.9).

Notably, when ascorbate was added in conjunction with decitabine, evidence for upregulation of TET-mediated active demethylation was seen. Co-administration of ascorbate with decitabine increased the percentage of hemi-methylated reads to 50% at the 4 h timepoint. Interestingly, the combination of ascorbate/decitabine induced 40% unmethylation and 33% hemi-methylated at 6 h. The observed active demethylation likely occurred due to generation of Fe^{2+} by ascorbate which upregulated TET enzymatic activity on the hemi-methylated DNA strand.

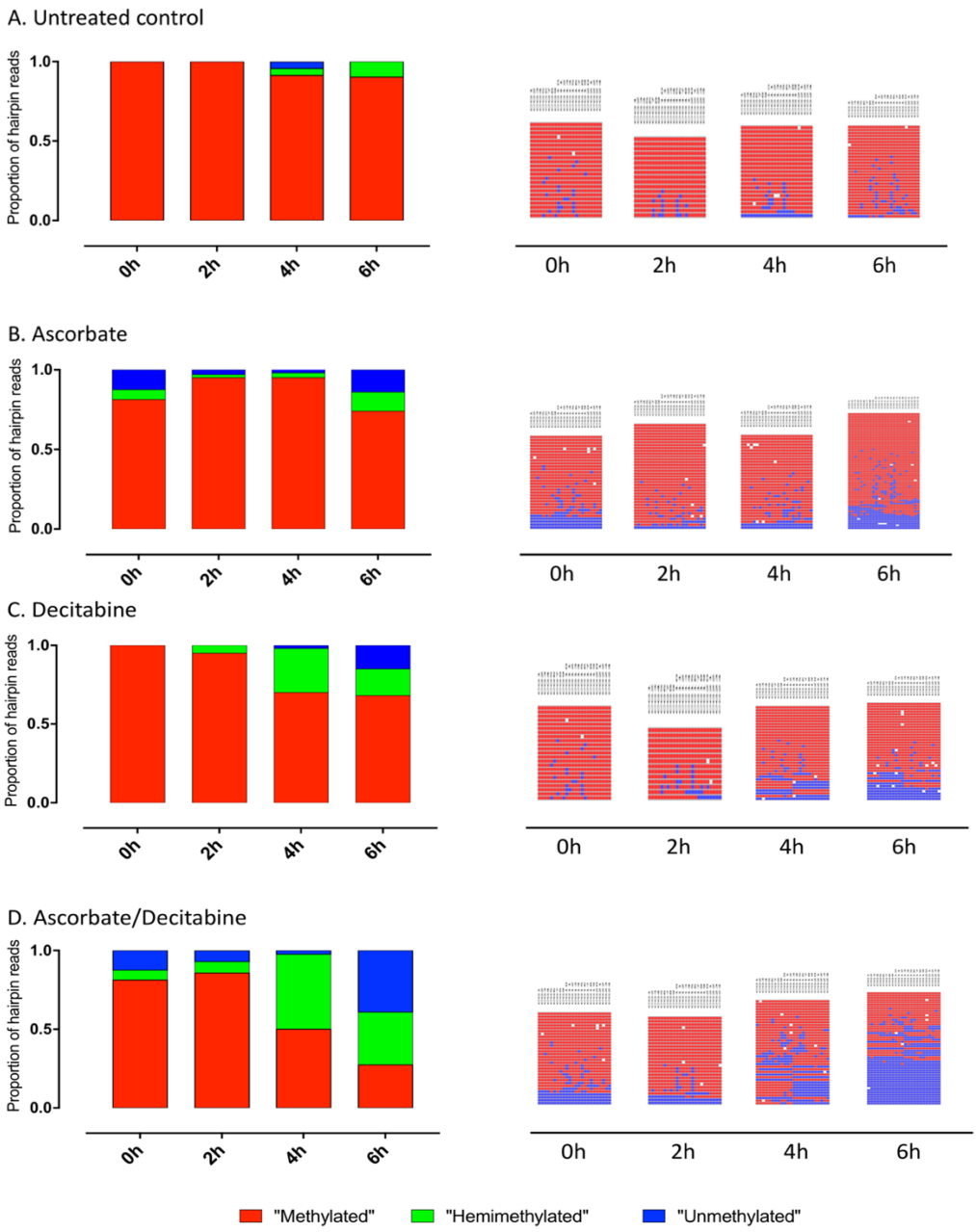


Figure 6. 9. Hairpin based methylation of *PCDHGA12* promoter in Jurkat cells treated with 500 µM ascorbate, 5 µM decitabine and combinations of ascorbate/decitabine compared to control untreated. The left panels show the proportions of methylated, hemi-methylated and unmethylated hairpin reads. The right panels show the corresponding heatmap output of *BiQ Analyzer*.

6.4.5 Impact of ascorbate/decitabine on the methylation pattern of *PCDHGA12* promoter in Jurkat cells over 72 h.

While the demethylation in response to decitabine is mechanistically informative, we wished to examine the role of TET enzymes in long-term exposure of ascorbate. We investigated the long-term effect of ascorbate/decitabine on methylation of *PCDHGA12* and *RASSF1* promoters. To assess the methylation of complementary CpG sites, hairpin DNA sequences

were prepared for *RASSF1* and *PCDHGA12* molecules from Jurkat DNA treated with ascorbate or/and decitabine at different timepoints. Briefly, Jurkat cells were synchronised for 18 h by thymidine block, preloaded with 0 or 500 µM ascorbate. At the time of release, cells were divided into groups and incubated with fresh media loaded with 1) 500 µM ascorbate; 2) 5 µM decitabine; 3) combined ascorbate/decitabine and 4) control untreated. Cells were harvested at 0, 4, 6, 24, 48 and 72 h. Cell viability was measured for all the groups during the experiment.

Upon treatment with ascorbate alone, the methylation levels of the *PCDHGA12* promoter was similar to that of the untreated control, although ascorbate was associated with demethylation of both strands of the *PCDHGA12* DNA 10% at 6 h (Figure 6.10). Application of decitabine resulted in an increase in the hemi-methylated *PCDHGA12* hairpin reads compared to control cells (Figure 6.10 c). The addition of ascorbate to decitabine did not alter the demethylating effect of decitabine; co-treatment with ascorbate slightly increased the proportion of hemi-methylated reads and did not increase the proportion of unmethylated reads (Figure 6.10 d).

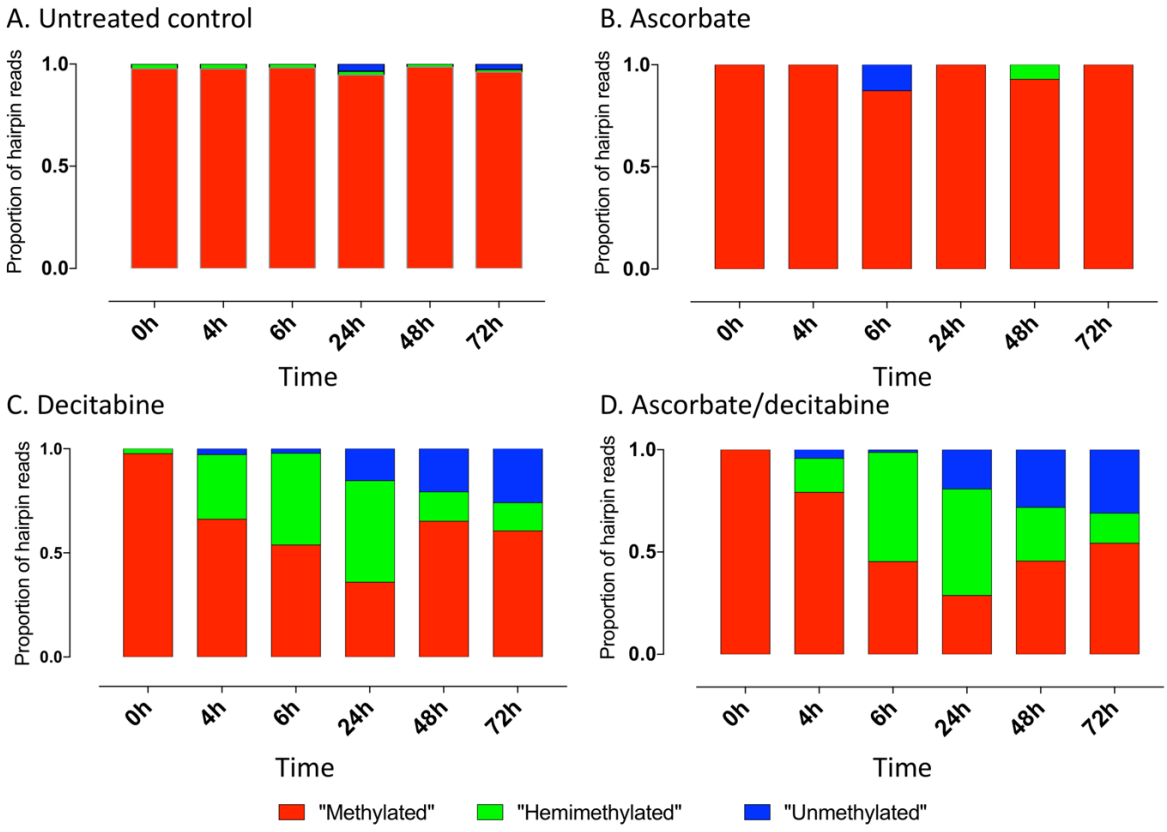


Figure 6. 10. Impact of ascorbate/decitabine on DNA methylation of Jurkat cells. Hairpin based methylation of *PCDHGA12* promoter following single or combined treatment. A) control untreated cells; B) cells treated 500 µM ascorbate; C) cells treated with 5 µM decitabine and D) cells treated with combined ascorbate/decitabine.

To further elucidate the mechanism of DNA methylation loss in cells treated with ascorbate/decitabine, three technical replicates for *PCDHGA12* were performed for treated and control samples. Using the same DNA samples as the above experiment, MiSeq libraries were prepared and sequenced from new BamHI digestions, *PCDHGA12* hairpin linker ligations, bisulfite conversions and PCR amplifications.

In the technical replicates, decitabine alone led to a similar increase in hemi-methylated DNA reads in a time dependent manner (Figure 6.11). The combined treatment of ascorbate/decitabine led to increases in the proportion of completely unmethylated DNA hairpin reads at 24, 48 and 72 h compared to untreated control.

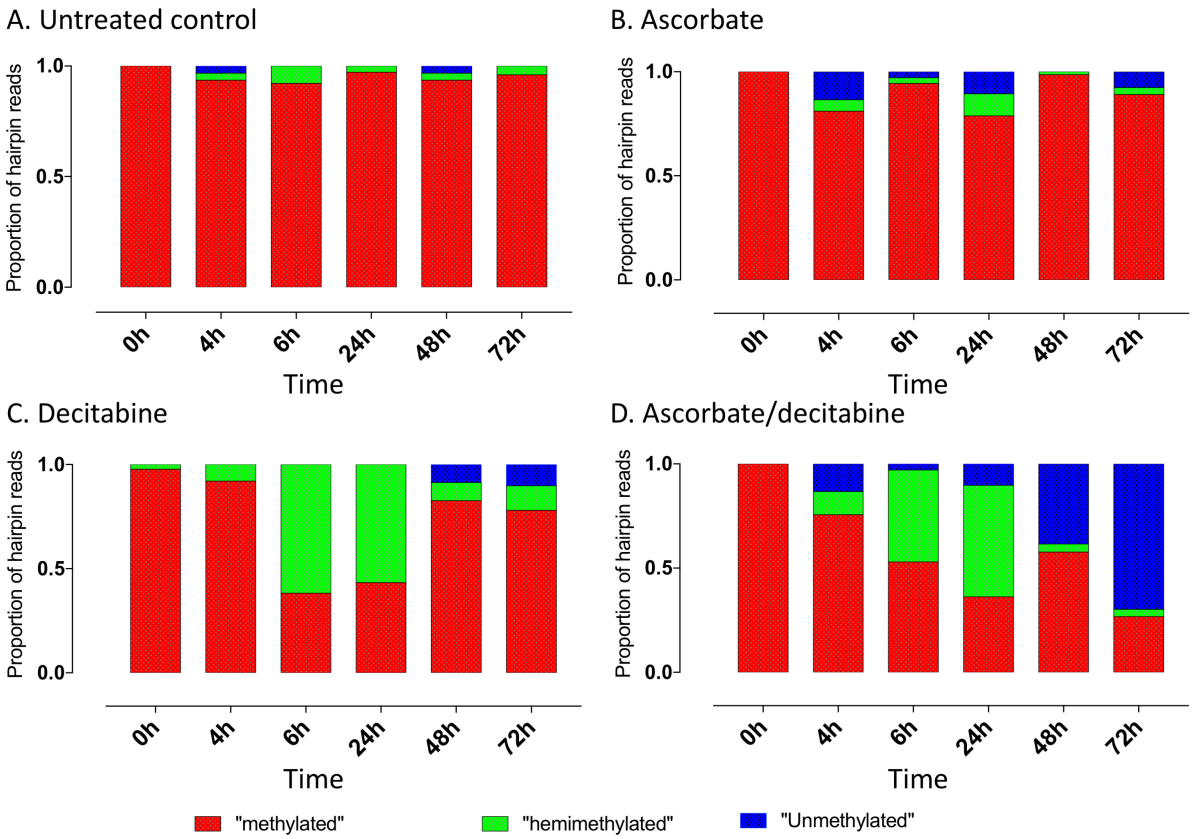


Figure 6. 11. Effect of ascorbate/decitabine on DNA methylation of *PCDHGA12* promoter. One of the technical replicates to verify the methylation changes in *PCDHGA12* promoter observed in figure 6.10.

To show the methylation value for all sequences within an individual sample for treated cells and untreated controls at different timepoints, we constructed both box and scatter plots. The box plot below shows median methylation for each sample at different timepoints (Figure 6.12). Furthermore, the scatter plot was constructed to show the methylation distribution for each individual hairpin reads within a treated sample at different timepoints (Figure 6.13). This plot

allows the visualisation of methylation of every single reads within individual sample; however with the limitation that multiple reads in the same position cannot be distinguished (Figure 6.13).

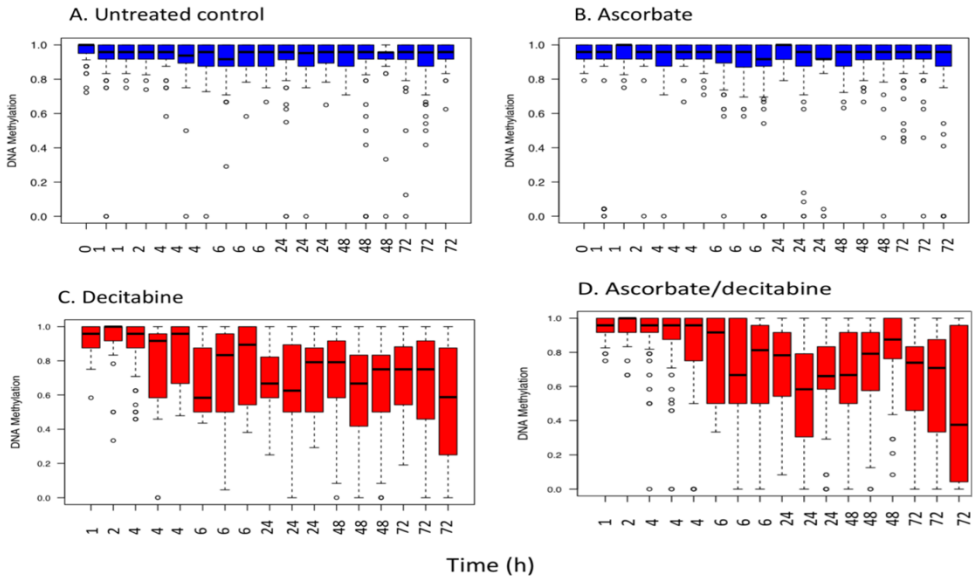


Figure 6. 12. Boxplots shows the median methylation of *PCDHGA12* in each sample of treated Jurkat cells at different timepoints compared to untreated control. Three technical replicates at 4, 6, 24, 48 and 72 h timepoints were repeated from the same genomic DNA of the first dataset.

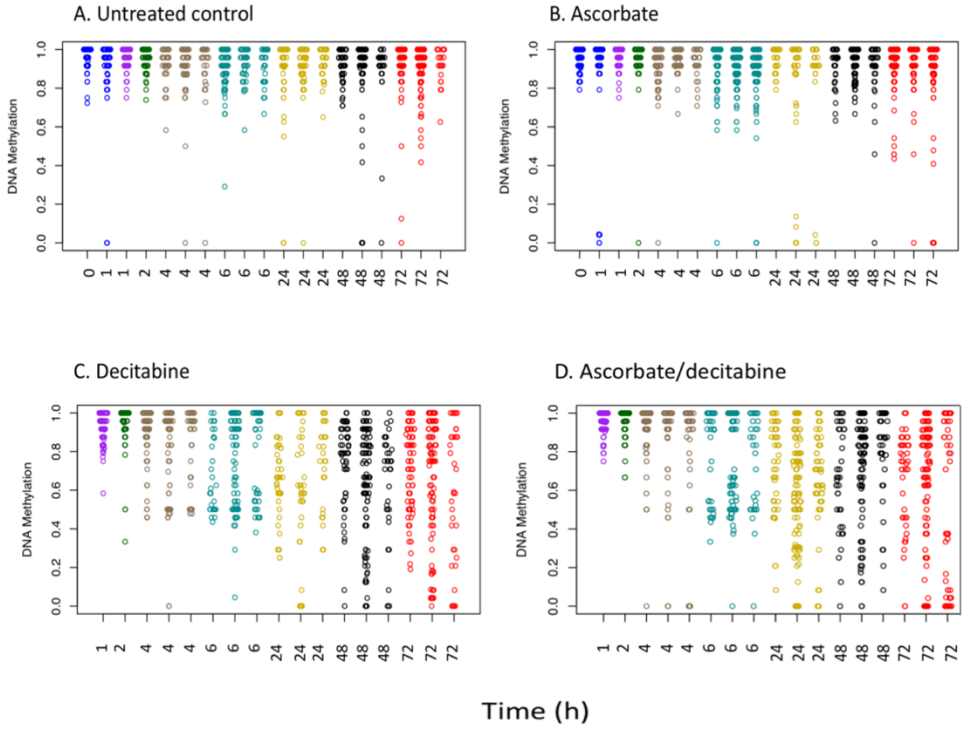


Figure 6. 13. Scatter plot shows the methylation of each read of *PCDHGA12* within individual sample at different timepoints.

6.4.6 Impact of long-term effects of ascorbate/decitabine on methylation pattern of *RASSF1* promoter.

To verify the reproducibility of the results obtained above, and to determine whether the demethylation phenomenon is observed in other genomic areas, the densely methylated *RASSF1* promoter was examined.

To investigate the impact of ascorbate and ascorbate together with decitabine on the methylation of the *RASSF1* promoter, similar methylation analysis was performed for the RASSF1 promoter.

DNA methylation was determined using the protocol described previously (Chapter 5). Briefly, Jurkat DNA was digested with SacI and ligated to *RASSF1* hairpin linkers followed by two PCR amplification steps and sequencing. Figure 2.16 shows the methylation patterns of treated *RASSF1* hairpin molecules at 0, 2, 4, 6, 24 and 48 h timepoints compared to untreated controls.

Similarly to *PCDHGA12* promoter, *RASSF1* promoter showed loss of methylation in decitabine and combined ascorbate with decitabine treated cells compared to untreated controls. Ascorbate alone induced minimal effects on methylation, while decitabine increased the proportion of hemi-methylated reads in time-dependent manner. As before, adding ascorbate to decitabine treatment resulted in an increase in the proportions of hemi-methylated and unmethylated reads at increasing timepoints compared to decitabine alone (Figure 6.14). A small increase in the number of unmethylated reads was observed in untreated and ascorbate treated samples. It was unusual to see demethylation in the untreated control and the reason for this is unknown.

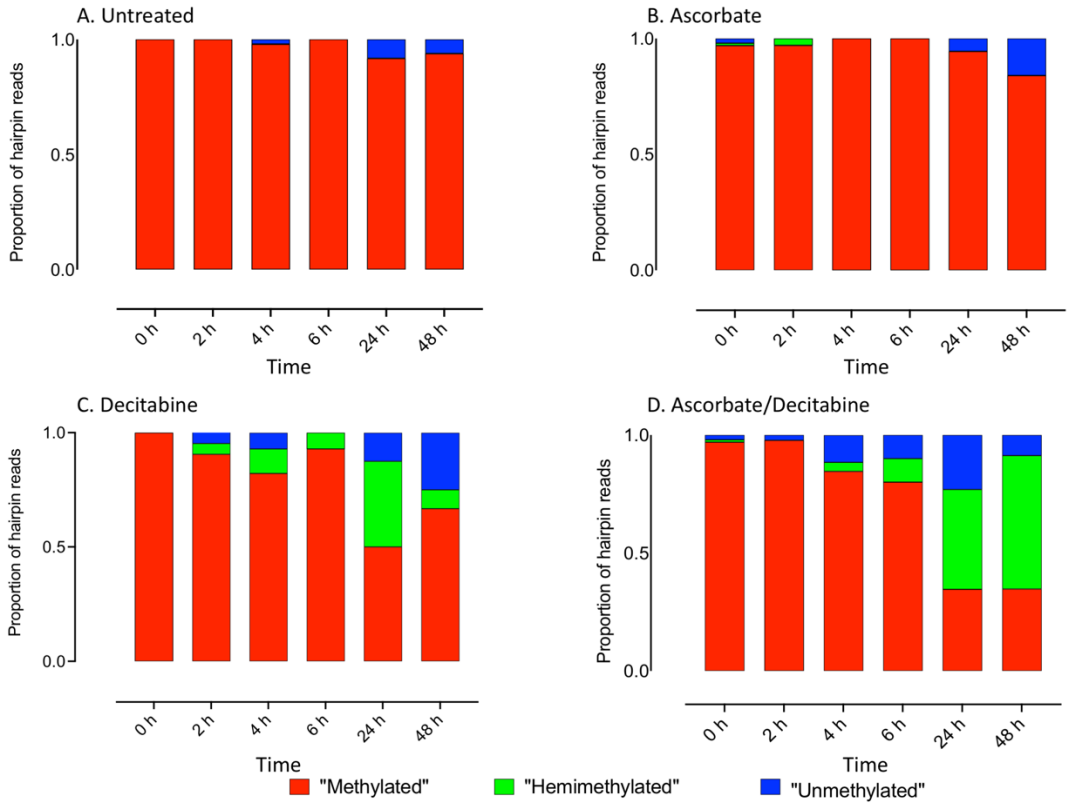


Figure 6. 14. Hairpin based analysis of *RASSF1* promoter methylation. A) control untreated cells; B) cells treated 500 µM ascorbate; C) cells treated with 5 µM decitabine and D) cells treated with combined ascorbate/decitabine.

6.4.7 Demethylation and re-methylation response to ascorbate/decitabine treatments.

To elucidate the chronic effect of ascorbate alone or combined ascorbate/decitabine on the demethylation and re-methylation dynamics of the *PCDHGA12* promoter, Jurkat cells were synchronised and incubated with a single dose of ascorbate or combined ascorbate/decitabine. Cells were released from thymidine block and collected at 0 , 2, 4 , 6 h, 12, 24, 48, 72, 96 and 120 h. The complementary CpG methylation changes were investigated and assessed (Figure 6.15). Following treatment of Jurkat cells with decitabine and ascorbate, the cells showed a substantial decrease of *PCDHGA12* methylation over 72 h compared to untreated controls. By the fifth day of additive free growth, re-methylation of the *PCDHGA12* promoter was observed. Nonetheless, cells treated with decitabine or combined ascorbate/decitabine still had lower levels of methylation compared to pre-treatment levels.

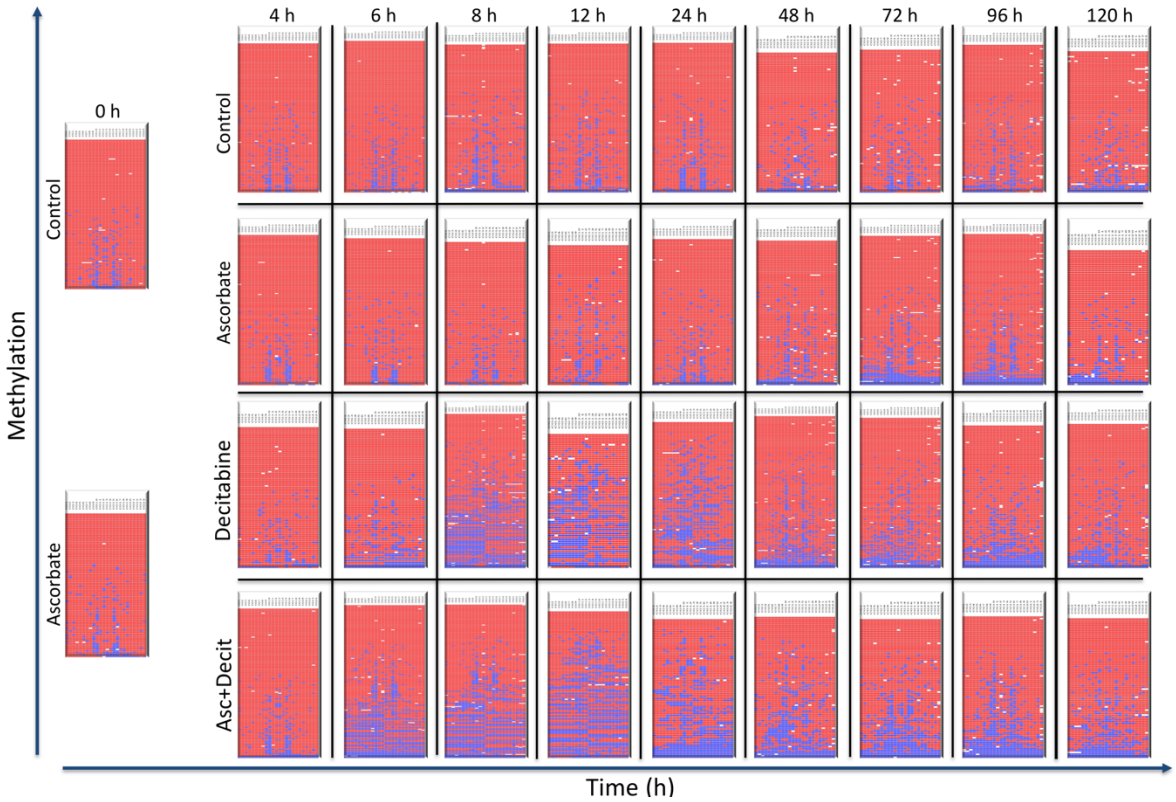


Figure 6. 15. *PCDHGA12* gene specific demethylation and re-methylation in Jurkat cells. Cells treated with a single dose decitabine or combined ascorbate/decitabine exhibited demethylation after treatment. Re-methylation occurred over the 72 h to 120 h timepoints.

The complementary CpG methylation was analysed for each sample, reads were classified as methylated, hemi-methylated or unmethylated (section 4.5.2, Chapter 4) and proportions of each were plotted (Figure 6.16).

The gene specific results were verified by global methylation analysis using low coverage methylation analysis (PBAT). The PBAT libraries were prepared from the same batch of genomic DNA, and sequenced and analysed using the standard protocol (Chapter 7). Global DNA methylation results were consistent with the locus-specific methylation results (Figure 6.16), showing gradual loss of DNA methylation with increasing timepoints in cells treated with decitabine and combined ascorbate/decitabine (Figure 6.17).

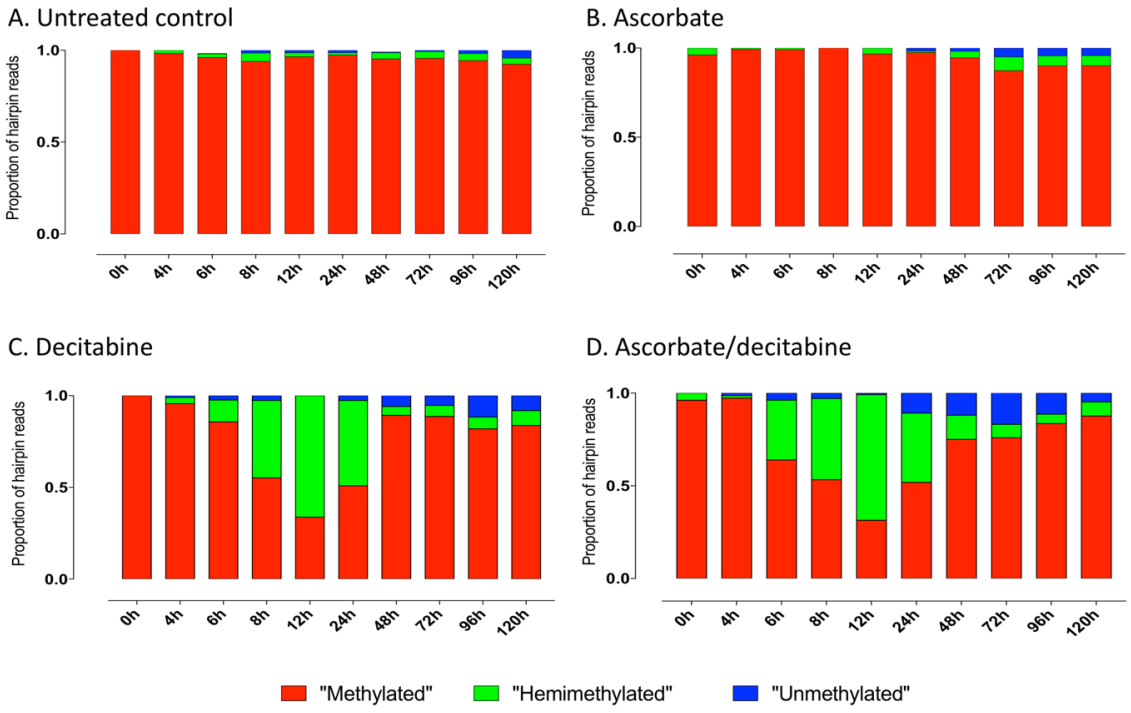


Figure 6. 16. The proportion of hemi-methylated/unmethylated or methylated *PCDHGA12* hairpin reads. These classifications reflect the raw data shown in Figure 6.15.

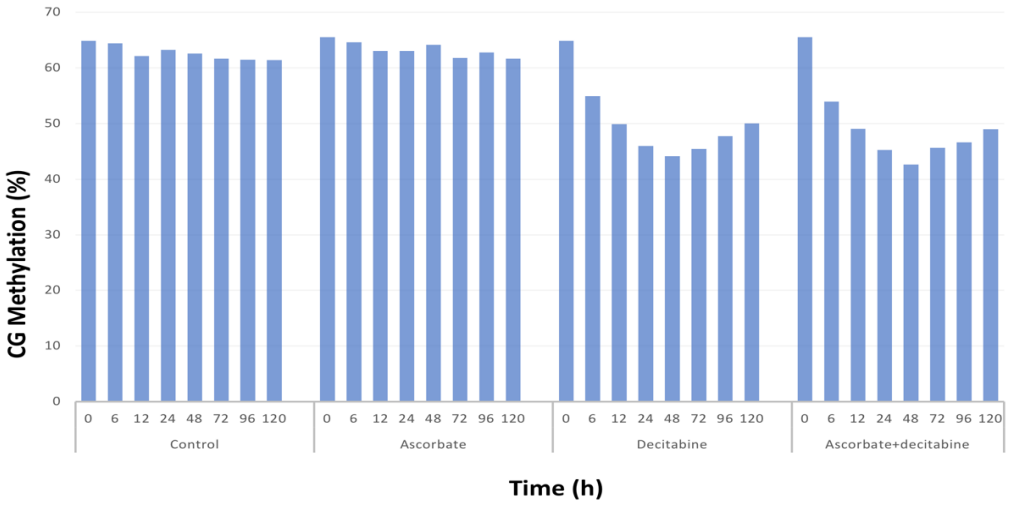


Figure 6. 17. Global CpG methylation levels (%) of Jurkat DNA treated with decitabine, ascorbic acid and combined ascorbic acid and decitabine for 120 h compared to control cells.

6.4.8 Summary and conclusion to this point.

With the power of the hairpin bisulfite technique, we were able to document the kinetics and pattern of DNMT1 inhibitor (decitabine) treatment with unparalleled clarity and detail. It was expected that decitabine would enhance hemi-methylation during DNA synthesis and we demonstrate this 2 h after release from G1 arrest. Surprisingly, we demonstrated full

demethylation of promoters or evidence of active demethylation which was never seen in untreated controls. However, these results were inconsistent and active demethylation was not always observed. For example the experiment shown in (Figure 6.8) shows substantial unmethylation, whereas minimal unmethylation is seen in Figure 6.15.

Numerous potential reasons were considered to explain the inconsistent and variable demethylation results. First we considered PCR and sequencing artefacts. PCR bias can be excluded as only reads with unique hairpin barcodes are aligned during sequencing analysis; i.e. each fragment of cellular DNA was represented only once on the final output. Additionally, sequencing artefacts can be excluded, as we confirmed the precision and the reproducibility by repeating the DNA sequencing for technical replicates and similar mean methylation was observed for the technical replicates confirming the reproducibility of the MiSeq sequencing (Figure 4.16).

While we could exclude PCR bias and sequencing artefacts, many other potential experimental variables that could affect demethylation should be examined such as the type of culture media, iron and alpha-ketoglutarate concentrations or the synergism of these variables. We decided to explore the culture conditions and different media components as a possible answer for this results variation. Different culture media contain different minerals and essential metals. Interestingly, Fe^{2+} is present at higher concentration in DMEM media, whereas alpha-ketoglutarate is a component of RPMI+glutamax media. These two compounds are co-factors for TET enzymes, and play a key role in upregulating TET activity and induction of active demethylation.

6.5 Does culture media affect DNA methylation?

We investigated the effect of different reagents on the methylation of *RASSF1* and *PCDHGA12* promoters in Jurkat cells grown in three different culture media: RPMI 1640 (basic culture media), DMEM (provides Fe^{2+}) and RPMI+glutamax (provides alpha-ketoglutarate). *RASSF1* and *PCDHGA12* hairpin molecules were prepared from treated and untreated genomic DNA at different timepoints and sequenced.

6.5.1 Comparison of RPMI 1640 vs DMEM.

To better understand the role that different media and their compositions can have in modulating DNA methylation, Jurkat cells were grown in different culture media and the effect of single or combined treatment on DNA methylation was investigated.

Towards this goal, we compared the methylation of *RASSF1* and *PCDHGA12* promoters in two different batches of Jurkat cells grown in parallel using either RPMI and DMEM. Jurkat cells that were used in this study were obtained from either Prof Morison (Dunedin) or Prof Hampton (Christchurch).

6.5.2 Jurkat cells obtained from Dunedin.

In this experiment we compared the effect of different media on the methylation of *RASSF1* and *PCDHGA12* promoters in Jurkat cells obtained from Dunedin laboratories. Briefly, Jurkat cells were grown in RPMI or DMEM media. Cells were synchronised by thymidine block for 18 h, and the cell cycle was initiated by incubation with 50 μM dC. At the time of release cells were treated with 500 μM ascorbate, 5 μM decitabine or the combination of ascorbate with decitabine (details in Chapter Three). Cells were harvested at 0, 4, 6 and 12 h. Then, *RASSF1* and *PCDHGA12* hairpin molecules were prepared from treated and untreated genomic DNA (protocol in Chapter Three) and sequenced for DNA methylation analysis. Heatmap outputs of *BiQ Analyzer* showed a high similarity of *RASSF1* methylation pattern between the two media (Figure 6.18). The proportions of methylated, hemi and unmethylated hairpin reads were calculated for each sample at different timepoints for the two experiments (Figure 6.19). We conclude that media compositions had no impact on the effect of ascorbate or decitabine reagents on DNA methylation.

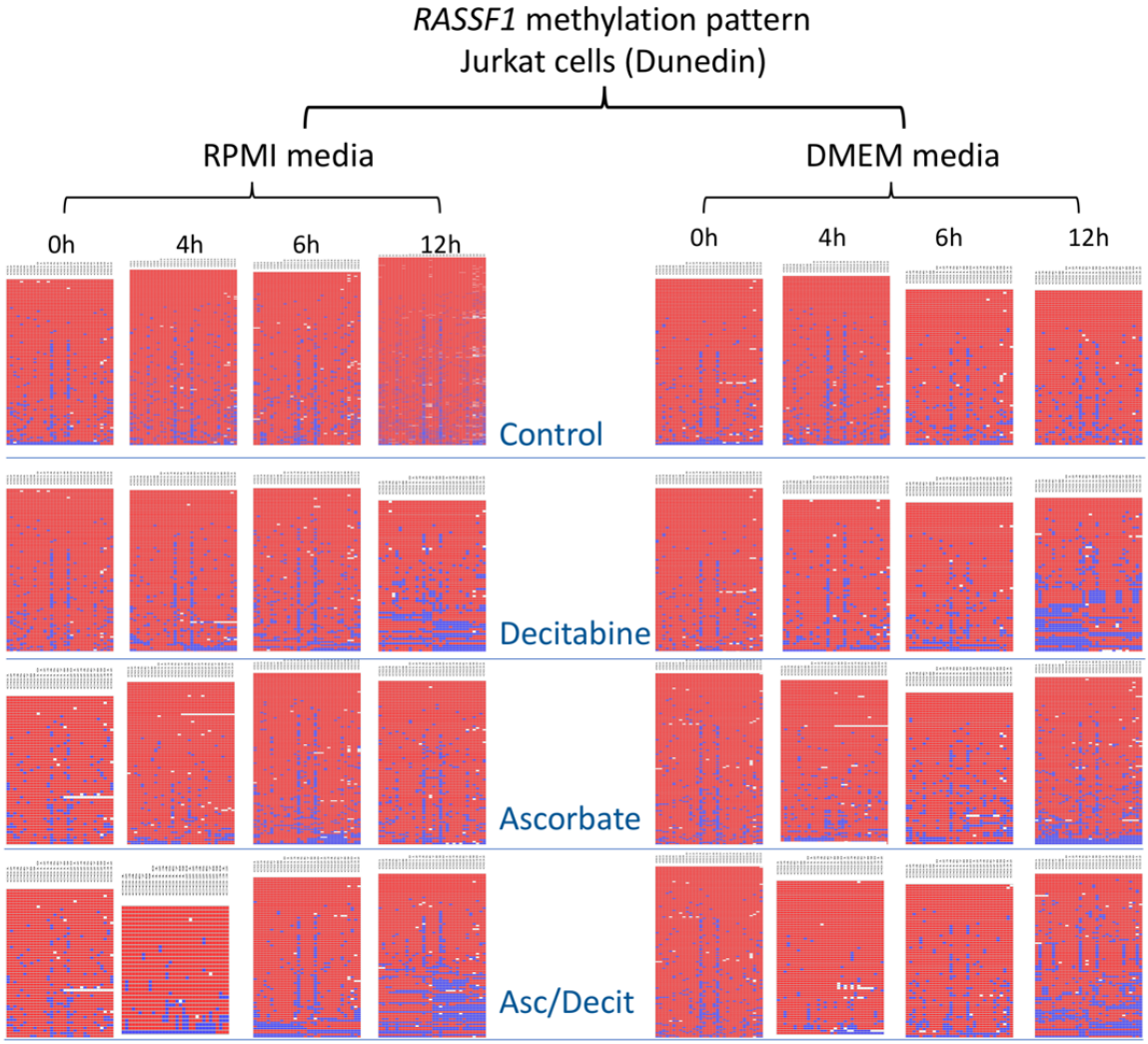


Figure 6. 18. Methylation of RASSF1 promoter in Jurkat cells grown in two different media, RPMI and DMEM, and treated with different reagents at different timepoints.

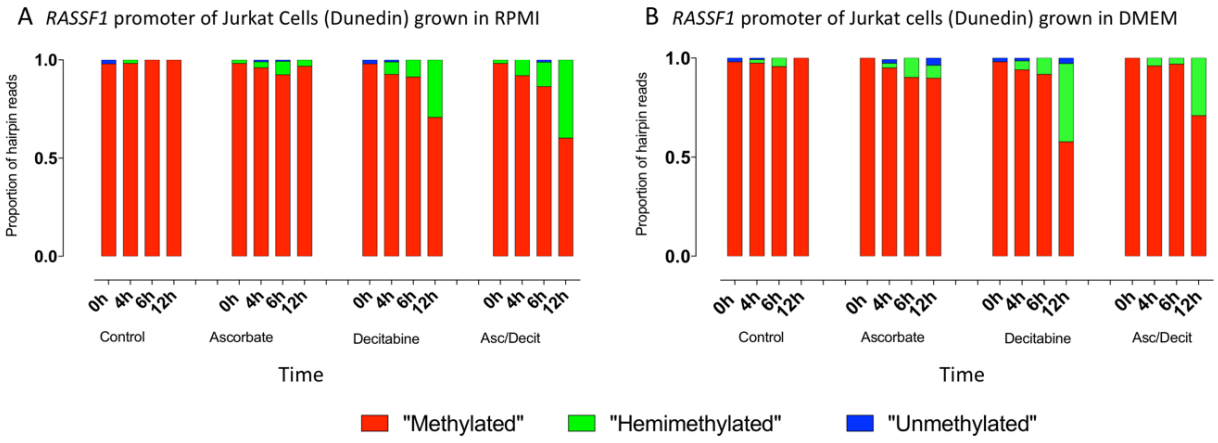


Figure 6. 19. The proportion of hemi-methylated/unmethylated or methylated *RASSF1* hairpin reads in Jurkat cells. Comparison of RPMI and DMEM media.

Similarly, the methylation of the *PCDHGA12* promoter was examined. *PCDHGA12* showed little methylation differences between the two media. Interestingly, loss of methylation in ascorbate and combined ascorbate/decitabine treated DNA samples was observed at 4 h in cells grown in DMEM (Figure 6.20). The proportions of unmethylated hairpin reads were increased in cells grown in DMEM media and treated with ascorbate and combined ascorbate/decitabine at 4 h (Figure 6.21). Active demethylation represented by complete unmethylated *PCDHGA12* hairpin reads were observed in cells treated with ascorbate and combined ascorbate /decitabine at 4, 6 and 12 h timepoints in cells grown in both media. DMEM media provides 0.25 µM (ferric nitrate). Adding ascorbate to the DMEM culture media converts the oxidised inactive Fe^{3+} to the reduced active Fe^{2+} which acts as a co-factor for TET to induce active demethylation. Because the level of iron in fetal calf serum (FCS) is not consistent and varies between manufacturers, and DMEM media contains iron, we conclude that DMEM media might facilitate DNA demethylation better than RPMI.

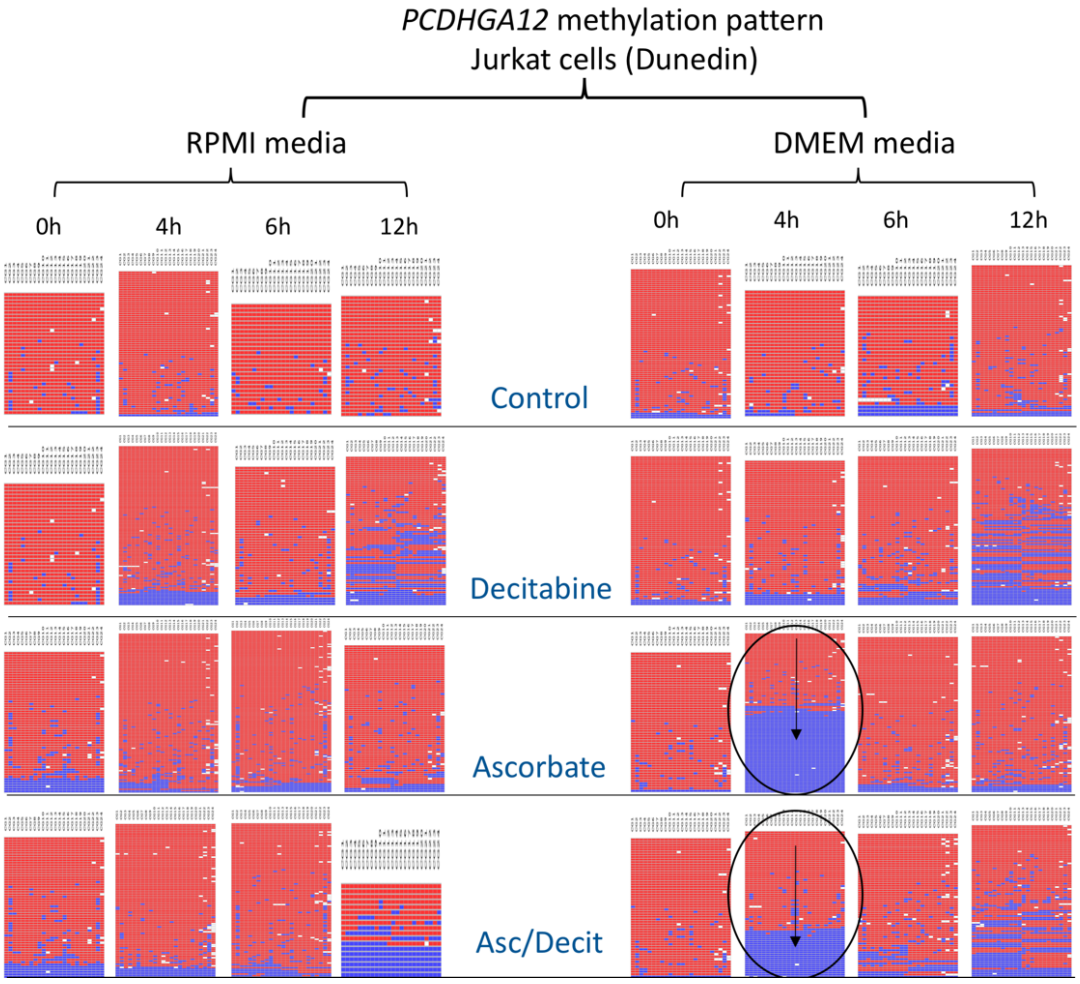


Figure 6. 20. Methylation of *PCDHGA12* promoter in Jurkat cells grown in two different media, RPMI and DMEM.

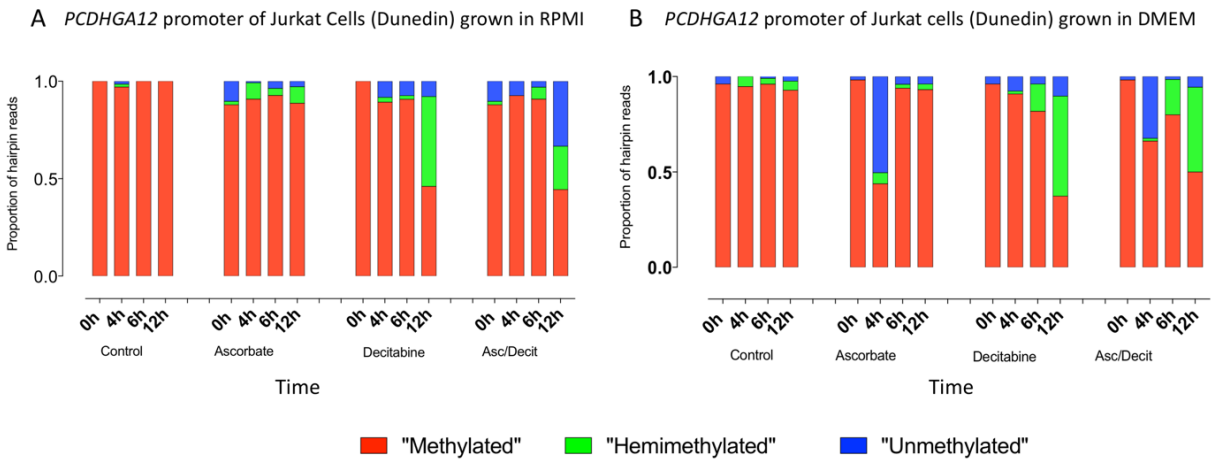


Figure 6. 21. The proportion of hemi-methylated/unmethylated or methylated *PCDHGA12* hairpin reads in Jurkat cells. Comparison of RPMI and DMEM media.

6.5.3 Jurkat cells obtained from Christchurch.

We then compared the methylation pattern of *PCDHGA12* promoter in Jurkat cells obtained from Christchurch to determine if the variability in demethylation was related to differences in the cell lines. We conducted two experiments in parallel with RPMI and DMEM media. Jurkat cells were obtained from Prof Hampton's laboratory (Christchurch). The cells were synchronised (detailed protocol in Chapter Three) and divided into groups. Cells were treated with 500 µM ascorbate and 5 µM decitabine or combined ascorbate with decitabine compared to untreated control. Cells were harvested at 0, 2, 4, 6, 12 and 24 h.

There were no differences in the methylation patterns of Christchurch Jurkat cells at the *PCDHGA12* promoter for cells grown in RPMI or DMEM (Figure 6.22). Cells grown in RPMI media displayed a slight increase in the proportion of hemi/unmethylated hairpin reads when treated with decitabine and combined ascorbate/decitabine compared to the other media. These differences, however, were comparable to those seen between technical replicates as discussed in section (section 6.4.8).

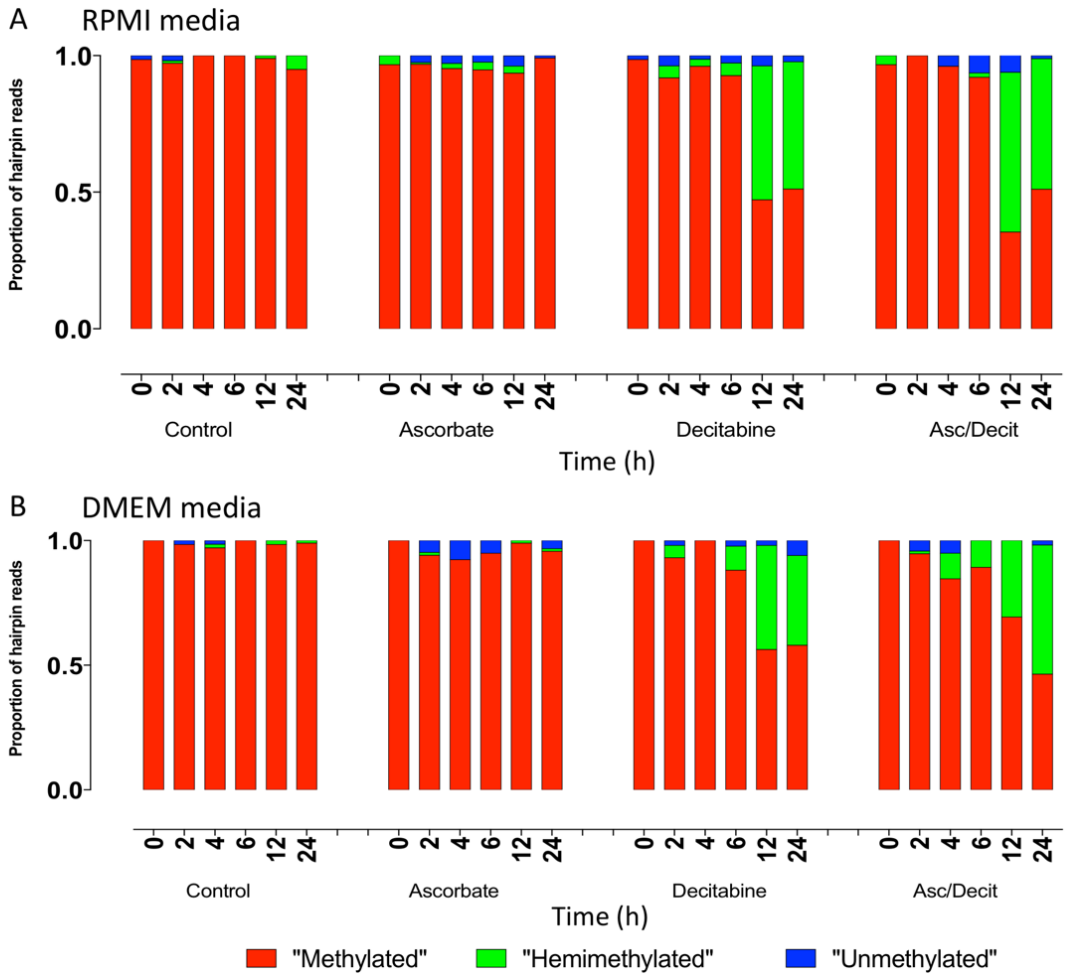


Figure 6. 22. The proportion of hemi-methylated/unmethylated or methylated *PCDHGA12* hairpin reads in Jurkat cells obtained from Christchurch. Comparison of RPMI and DMEM media.

6.5.4 Comparison of *PCDHGA12* methylation in Jurkat cells grown in RPMI Glutamax vs RPMI 1640.

To investigate whether the changes in DNA methylation are dependent on media composition, we studied the methylation pattern of *PCDHGA12* promoter in Jurkat cells grown in two different media, RPMI+glutamax and RPMI 1640. Glutamax media is characterised by supplementation with L-glutamine and L-alanine amino acid. Glutamine is an essential nutrient for energy production and source of nitrogen for nucleic acid synthesis.

The two culture media were supplemented with 10% FCS and were investigated in parallel. Jurkat cells (Christchurch) were grown, synchronised (details in Chapter Three) and treated in standard conditions.

The methylation pattern of the *PCDHGA12* promoter showed no marked differences between the two media. Decitabine and combined ascorbate with decitabine showed a decrease in

PCDHGA12 methylation levels in both media (Figure 6.23). We therefore conclude that the media composition had no impact on the effect of ascorbate or decitabine on methylation of the *PCDHGA12* promoter.

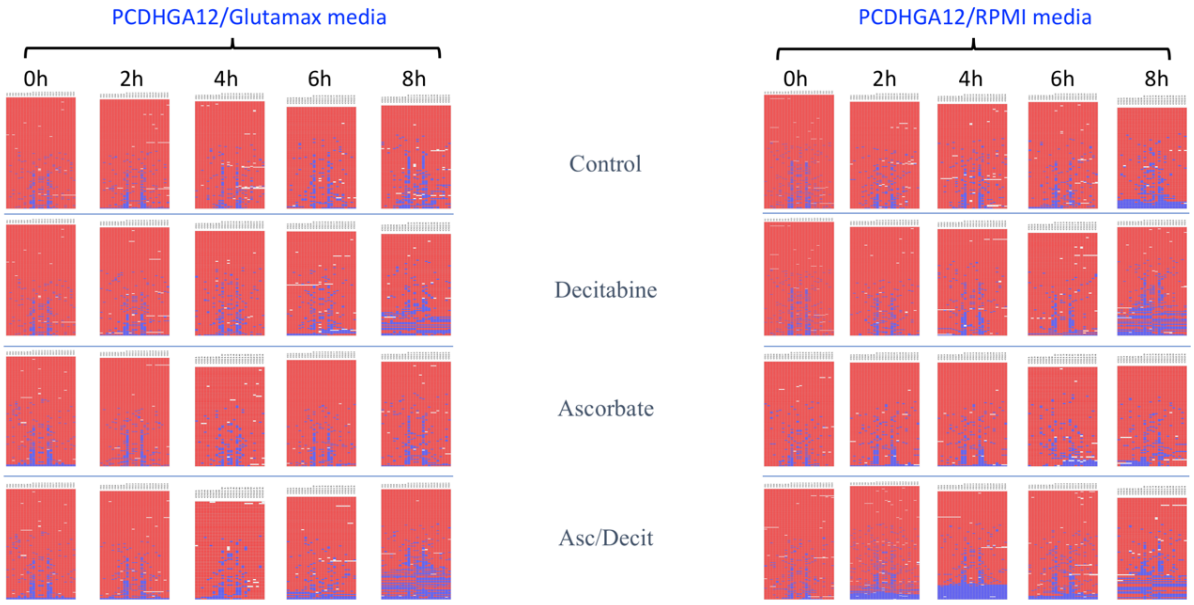


Figure 6. 23. Methylation of *PCDHGA12* promoter in Jurkat cells grown in either RPMI Glutamax or RPMI 1640 and treated with different reagents.

6.6 Thymidine synchronisation vs unsynchronisation: effect on demethylation.

6.6.1 Comparing the effect of synchronised and unsynchronised cells on DNA methylation.

Cells grown in culture exist at all stages of the cell cycle. Cell synchronisation refers to bringing cells into the same stage of the cell cycle; for example, excess thymidine will inhibit DNA synthesis and cells will be synchronised to the G1 phase.

In this experiment, two batches of Jurkat cells were prepared to examine if cell synchronisation has an impact on the effect of decitabine and ascorbate on DNA methylation. This was an attempt to understand the active demethylation using a simplified work stream and fewer variables. The experiments were conducted in duplicate.

Briefly, in the first group, Jurkat cells were synchronised for 22 h with and without ascorbate. At the time of release, cells were divided into four subgroups: untreated control; 1 µM decitabine; 500 µM ascorbate and combined ascorbate/decitabine. In the second group unsynchronised Jurkat cells were divided into four subgroups: control; 1 µM decitabine; 500

μM ascorbate and combined ascorbate/decitabine. Cells were harvested at 0, 4, 8, 12, 16 and 24 h. *PCDHGA12* hairpin molecules were prepared from DNA for all samples at all timepoints.

Interestingly, dramatic differences in the methylation pattern of the *PCDHGA12* promoter treated with ascorbate/decitabine were observed between synchronised and unsynchronised treated Jurkat cells. Synchronised cells showed a marked increase in hemi-methylation in cells treated with decitabine and combined ascorbate and decitabine during cell cycle progression (Figure 6.24), un synchronised showed no demethylation and heatmaps are not shown. The proportions of methylated, hemi-methylated and unmethylated hairpin reads were calculated for all treated and untreated samples. The *PCDHGA12* promoter of synchronised cells, treated with decitabine or combined ascorbate/decitabine, showed a decrease in methylation levels represented by an increase in hemi-methylated hairpin reads with increasing time (Figure 6.25a), compared to a complete absence of decitabine and ascorbate effects in unsynchronised cells (Figure 6.25b).

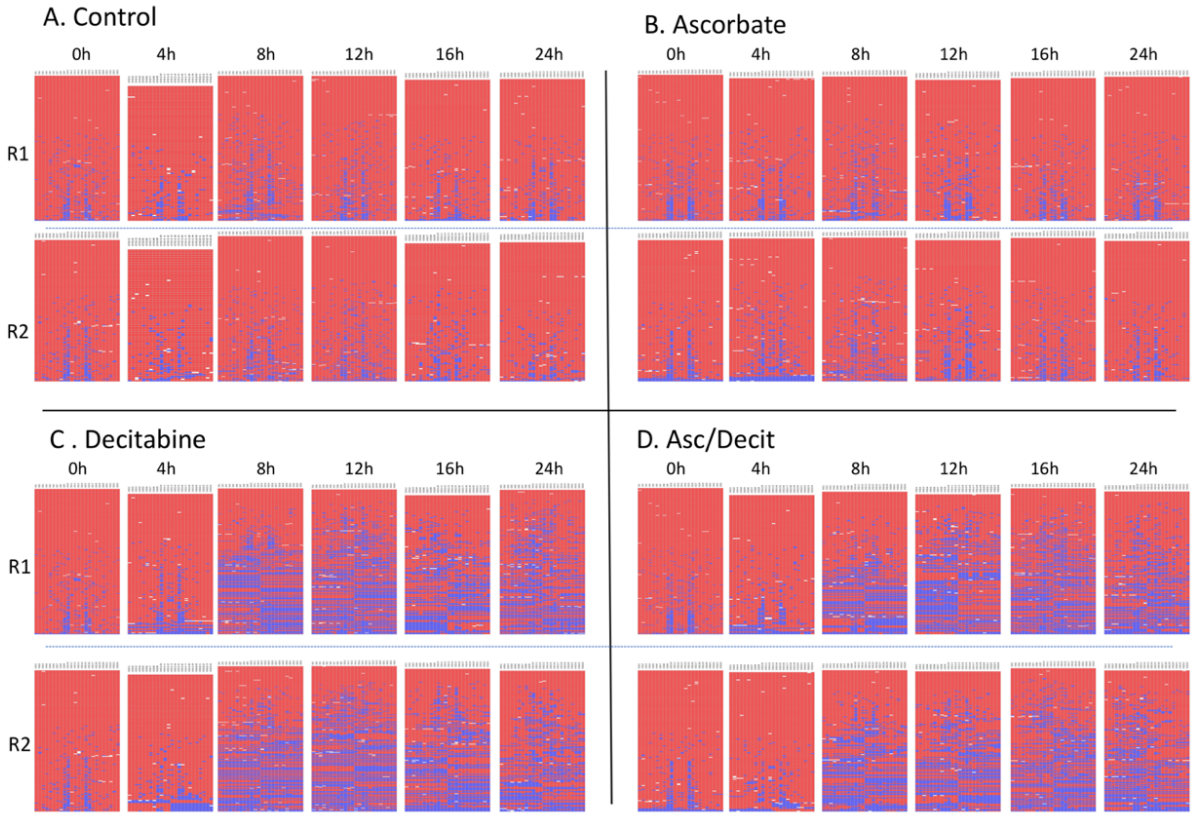


Figure 6. 24. The methylation patterns of *PCDHGA12* promoter in two biological replicates (R1 and R2) of synchronised A) control untreated cells, B) 500 μM ascorbate treated cells, C) 1 μM decitabine treated cells and D) combined ascorbate/decitabine treated cells.

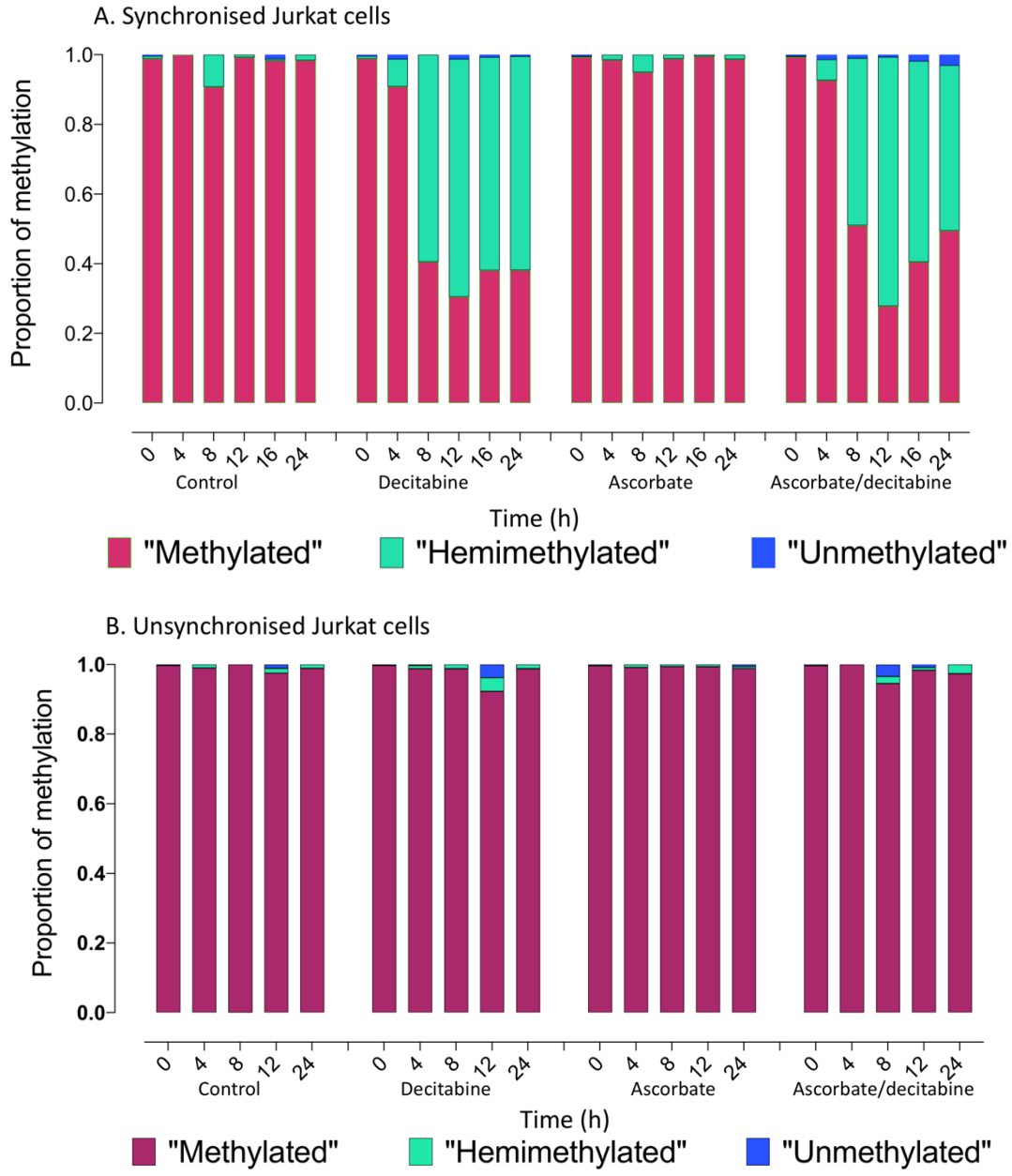


Figure 6. 25. Methylation of *PCDHGA12* promoter. Comparison of A) synchronised Jurkat cells treated with ascorbate, decitabine and combined ascorbate/decitabine at the time of release for 24, and B) unsynchronised Jurkat cells treated with ascorbate, decitabine and it is combinations for 24 h.

To see whether the phenomenon "loss of methylation" is reproducible, we investigated the effect of low dose of decitabine "1 µM decitabine" and 500 µM ascorbate on the methylation of the *RASSF1* promoter. Interestingly, *RASSF1* showed a gradual loss of methylation reflecting increase in the number of hemi-methylated reads in a time dependent manner in cells treated with decitabine and combined ascorbate/decitabine compared to untreated control cells (Figure 6.26). Cells were harvested at 0, 4, 8, 12, 16 and 24 h.

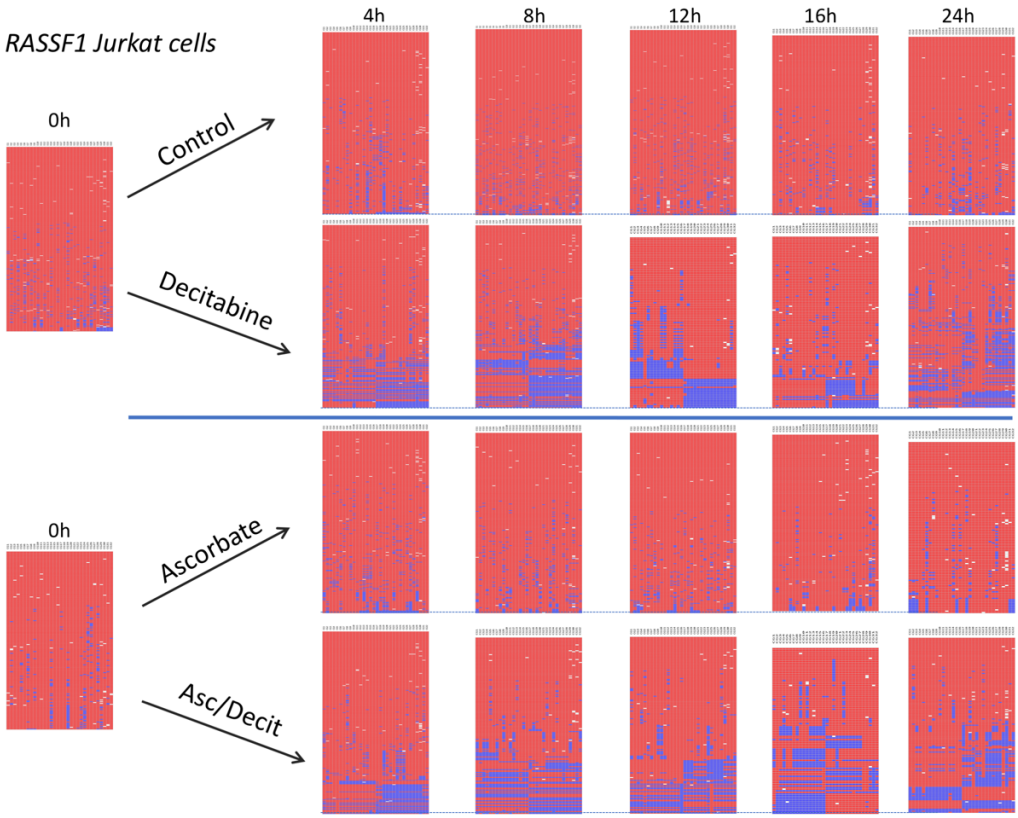


Figure 6. 26. Methylation pattern of *RASSF1* promoter. Synchronised Jurkat cells treated with 1 µM decitabine, 500 µM ascorbate and combinations of ascorbate/decitabine compared to untreated cells.

Marked hemi-methylation was seen the *RASSF1* promoter, but complete lack of active demethylation , emphasising the variable nature of the experiments.

6.6.2 Summary and conclusion to this point.

Active demethylation was observed at 6 h after treatment with ascorbate/decitabine. Decitabine induced consistent hemi-methylation percentage during DNA replication in the two tumour suppressor genes *PCDHGA12* and *RASSF1,* and the percentage of hemi-methylation increased with increasing time. Cell synchronisation is an important and essential step for the induction of demethylation. Decitabine and combination of ascorbate/decitabine had no effect on unsynchronised Jurkat cells. These were surprising results, as we were expecting a small amount of demethylation. The use of unsynchronised cells was not useful and did not make the system more simple. Changing the culture media has no significant effect on the percentages of demethylation. Similar effects were observed in RPMI1640, DMEM or RPMI Glutamax media.

6.7 Evaluating the methylation patterns of *PCDHGA12* and *RASSF1* promoters in other leukaemia cell lines.

Since, variable results were observed between experiments in Jurkat cells, we evaluated whether other cell lines would show more consistent demethylation results. The methylation signatures of *PCDHGA12* and *RASSF1* that might be serve as a methylation markers for future study were investigated in other leukaemia cell lines.

DNA samples were extracted from six leukaemia cell lines: CCRF-CEM, Molt4, Nalm6, RS4;11, HL-60 and MDS. DNA was prepared as before for *PCDHGA12* and *RASSF1* promoter sequencing (Chapter 3).

Interestingly, *PCDHGA12* promoter was densely methylated in all examined cell lines, with overall ~ 95% mean methylation (Figure 6.27 'upper panel'). In comparison, the *RASSF1* promoter was densely methylated only in Molt4 cells with ~ 95 % mean methylation, but less methylated in Nalm6 and MDS (~ 50% mean methylation), and unmethylated in RS4, CEM and HL60 cell lines (Figure 6.27 'lower panel'). The violin plot below summarises the DNA methylation patterns of *PCDHGA12* and *RASSF1* in different leukaemia cell lines (Figure 6.28). Therefore, *PCDHGA12* gene can be used as a methylated marker for all the examined cell lines, whereas *RASSF1* genes can only be used with the Molt4 cell line.

From the cell lines above, Molt4 was selected to study the methylation pattern of complementary CpG sites of *PCDHGA12* and *RASSF1* promoters and Nalm6 and HL60 cell lines were selected to study *PCDHGA12* promoter. In order to investigate the effect of decitabine and ascorbate on DNA demethylation, two experiments were performed and each included replicates.

In the first experiment, Molt4, Nalm6 and HL60 cells were synchronised by a single thymidine block for 22 h. At the time of release, cells were washed and loaded with fresh media supplemented with 50 μM dC and divided into four group: untreated control; 1 μM decitabine; 500 μM ascorbate; and combined 500 μM ascorbate and 1 μM decitabine. Cells were harvested at 0, 2, 4, 6, 12 and 24 h after release from thymidine block.

The second experiment was conducted without synchronisation. Molt4, Nalm6 and HL60 cells were treated with decitabine or combined ascorbate/decitabine for 24 h compared to untreated control cells. Cells were harvested at 0, 4, 8, 12 and 24 h.

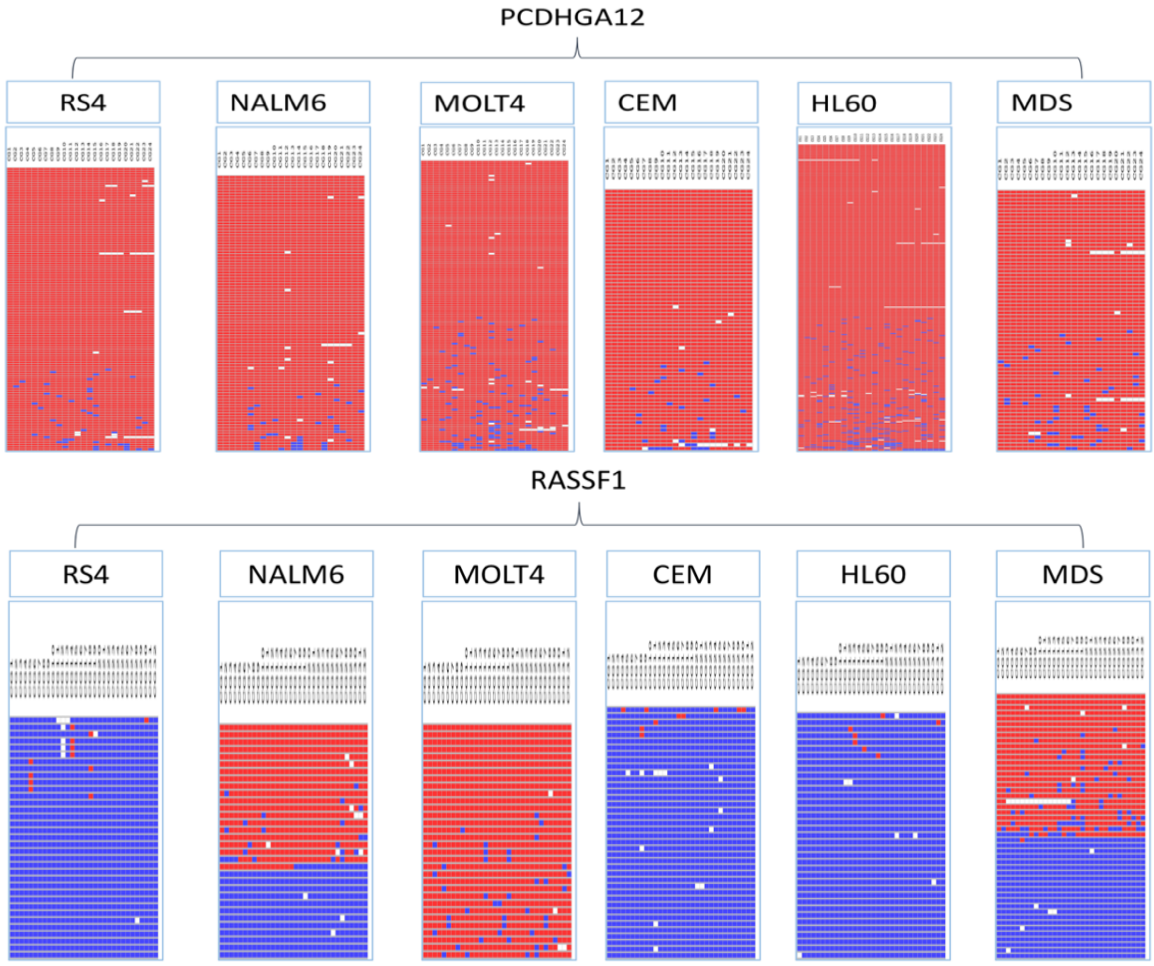


Figure 6. 27. Methylation pattern of *PCDHGA12* promoter in different leukaemia cell lines (upper panel). The lower panel shows the methylation pattern of *RASSF1* promoter across different leukaemia cell lines. (red = methylated, blue = unmethylated, white = missing sequence).

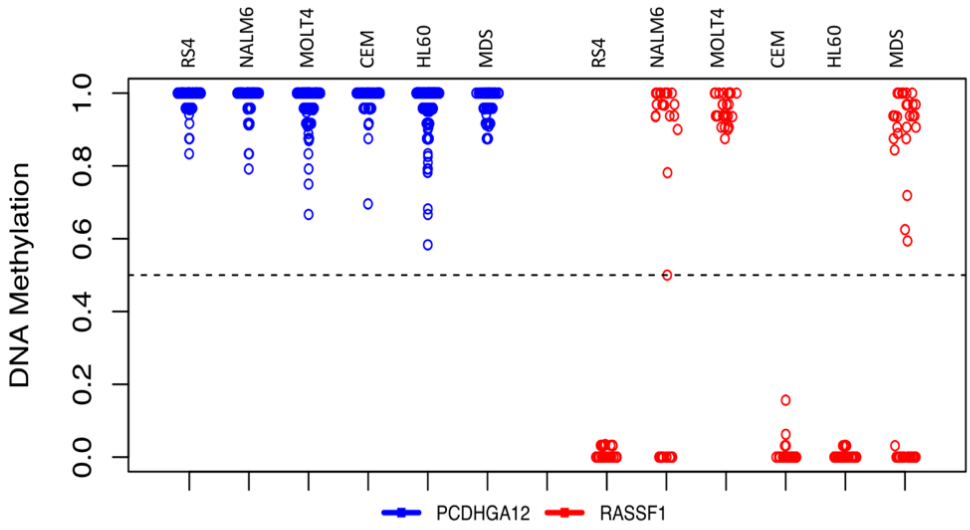


Figure 6. 28. Violin-point plot showing the methylation of all sequencing results of *PCDHGA12* (plotted in blue) and *RASSF1* promoters (plotted in red) in leukaemia cell lines. A dotted line at 0.5 was used as a cut-off to define methylation of a single sequence read.

6.7.1 Impact of ascorbate/decitabine on methylation of *PCDHGA12* promoter in Molt4 cell lines.

Synchronised Molt4 cells treated with decitabine or combined ascorbate/decitabine showed substantial changes in DNA methylation levels with an increase in the percentage of hemi-methylated hairpin reads at increasing timepoints compared to untreated control cells (Figure 6.29).

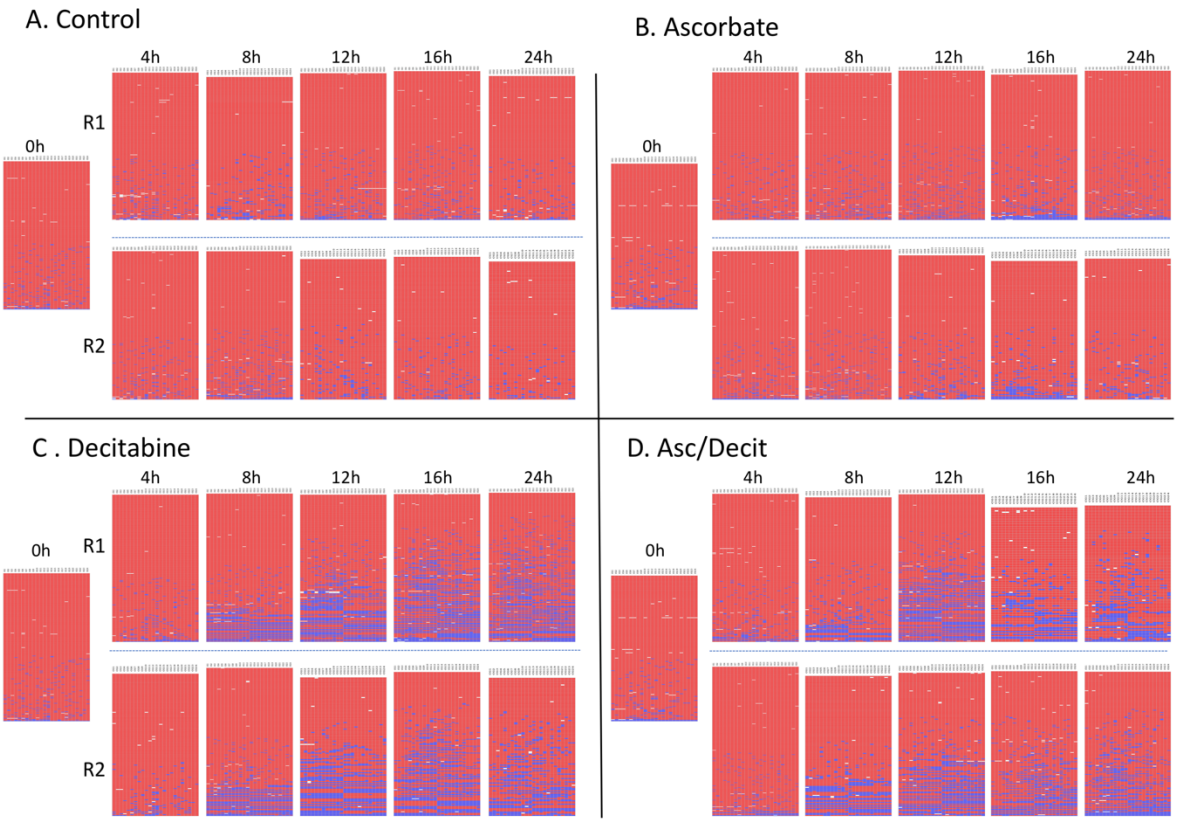


Figure 6. 29. Methylation of *PCDHGA12* promoter in synchronised Molt4 cell line. Two biological replicates (R1 and R2) showed the methylation pattern of *PCDHGA12* promoter in synchronised A) control untreated cells, B) 500 µM ascorbate treated cells, C) 1 µM decitabine treated cells and D) combined ascorbate/decitabine treated cells.

The proportions of methylated, hemi-methylated and unmethylated hairpin reads were calculated for all samples. Synchronised Molt4 cell lines treated with decitabine or combined ascorbate/decitabine showed an increase in the percentage of hemi-methylated reads at increasing timepoints. The combination of ascorbate with decitabine showed the same degree of methylation change as decitabine alone (Figure 6.30 a).

Unsynchronized Molt4 cells showed no effects of decitabine or combined ascorbate/decitabine on the methylation pattern of *PCDHGA12* except an increase in the percentage of hemi-methylated reads at 24 h in cells treated with decitabine (Figure 6.30 b).

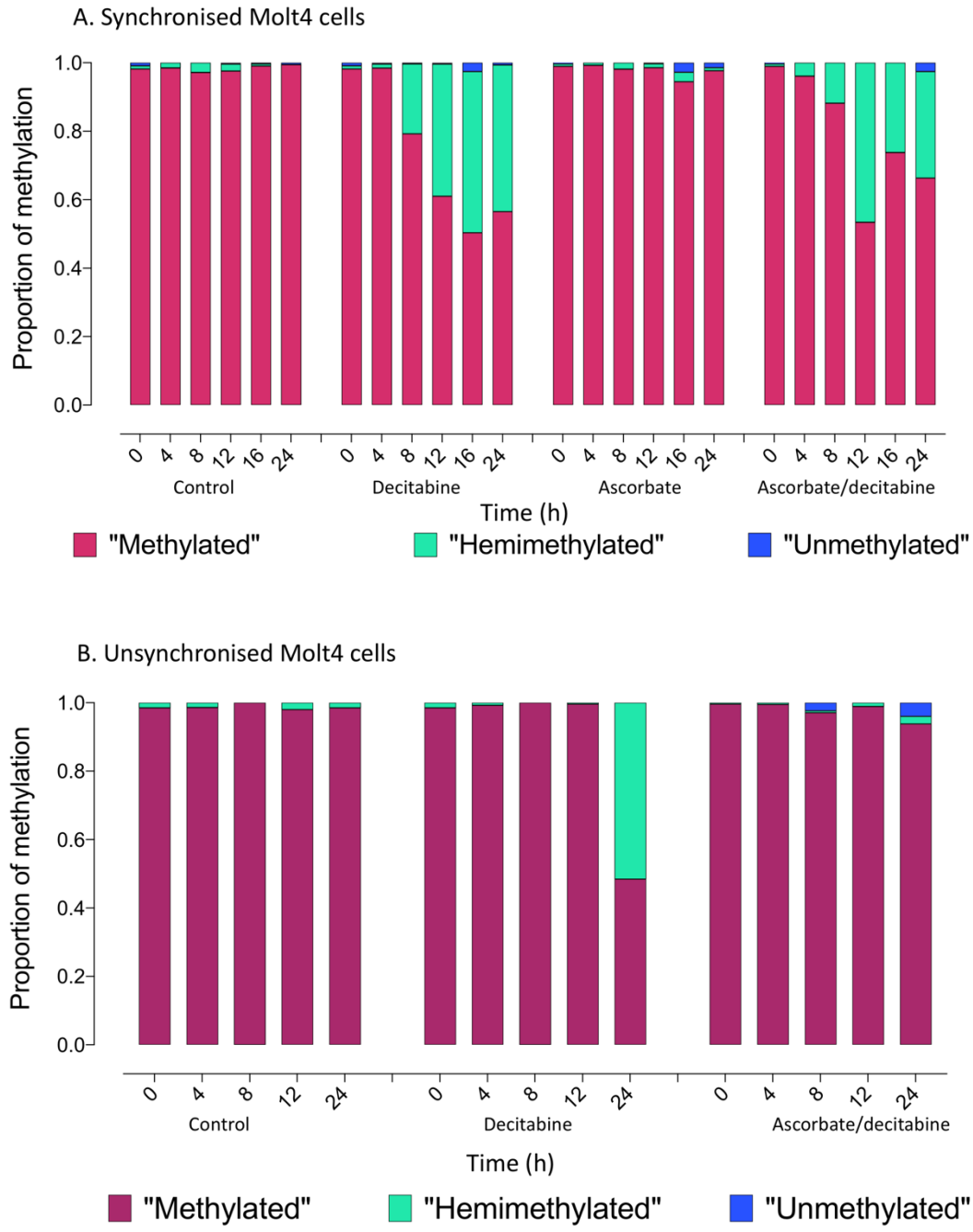


Figure 6. 30. Comparison of hairpin based methylation of *PCDHGA12* promoter in A) synchronised and B) unsynchronised Molt4 cell lines.

The experiment on Molt4 was extended to test and investigate the DNA methylation changes on the hypermethylated *RASSF1* promoter. Similarly, the *RASSF1* promoter showed a substantial increase in hemi-methylated reads in decitabine and combined ascorbate/decitabine

treated cells. The hemi-methylated effect was observed as early as 4 h after release from thymidine block (Figure 6.31).

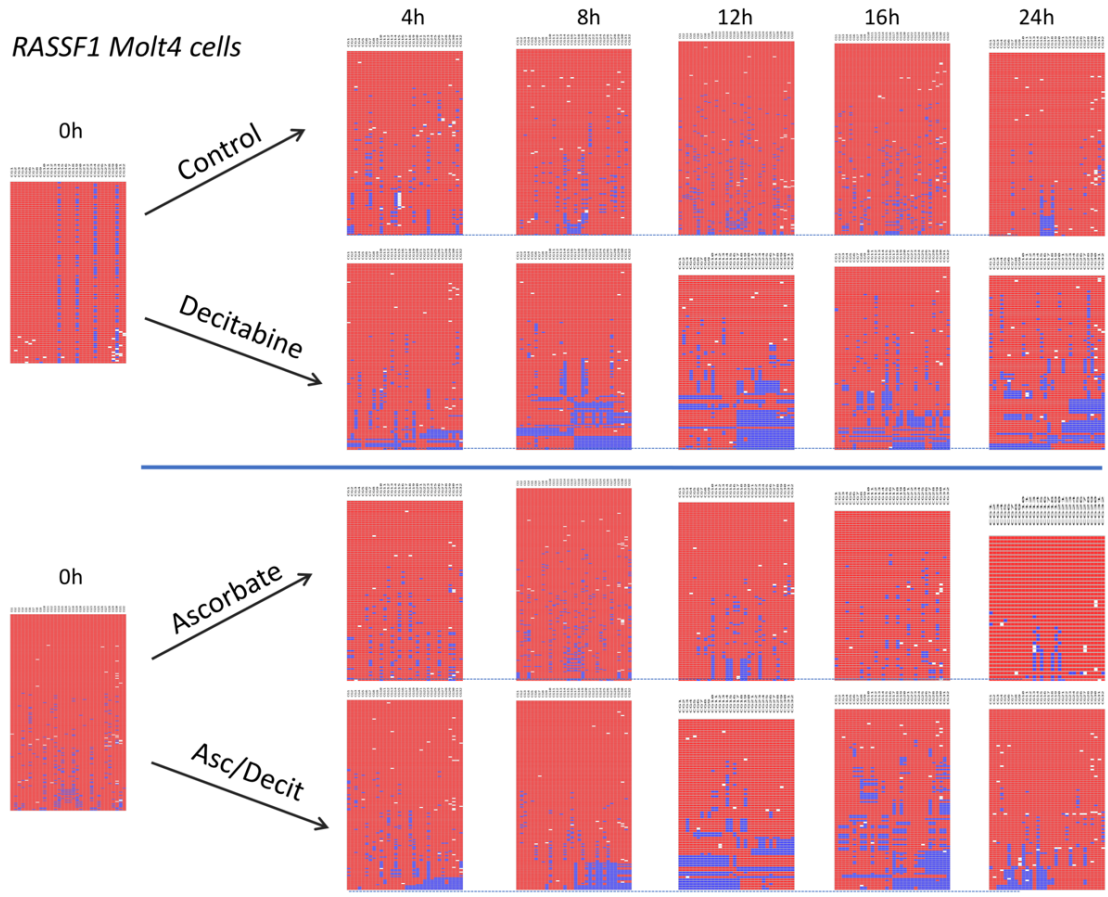


Figure 6. 31. Methylation pattern of *RASSF1* promoter. Synchronised Molt4 cells treated with 1 µM decitabine, 500 µM ascorbate and combination of ascorbate with decitabine compared to untreated cells.

6.7.2 Impact of ascorbate/decitabine on *PCDHGA12* promoter of Nalm6 cell lines.

Similar to Molt4 cell lines, synchronised Nalm6 showed marked methylation changes in the *PCDHGA12* promoter after 1 µM decitabine and combined ascorbate with decitabine treatment (Figure 6.32). The proportions of hemi-methylated hairpin reads increased with increasing time in cells treated with 1 µM decitabine compared to untreated control. Treatment with ascorbate alone had no effect on DNA methylation, and adding ascorbate to decitabine treatment did not change the decitabine effect (Figure 6.33a).

Unsynchronized Nalm6 cells showed no decitabine effect except an increase in the proportion of hemi-methylated reads at 24 h with decitabine treatment. Combined ascorbate with decitabine had no effect on the DNA methylation of unsynchronised cells (Figure 6.33b).

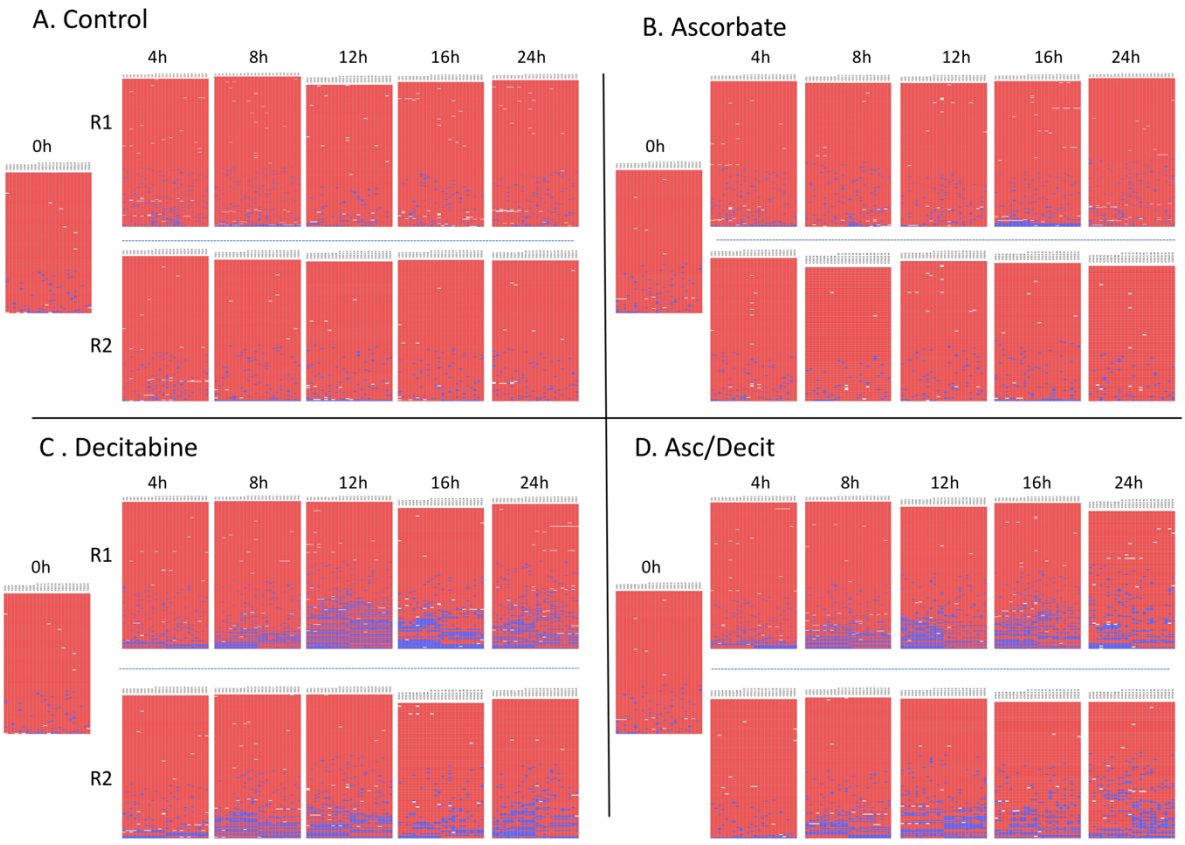


Figure 6. 32. Methylation of the *PCDHGA12* promoter in synchronised Nalm6 cells. Two biological replicates (R1 and R2) showed the methylation pattern of *PCDHGA12* promoter in synchronised A) control untreated cells, B) 500 µM ascorbate treated cells, C) 1 µM decitabine treated cells and D) combined ascorbate/decitabine treated cells.

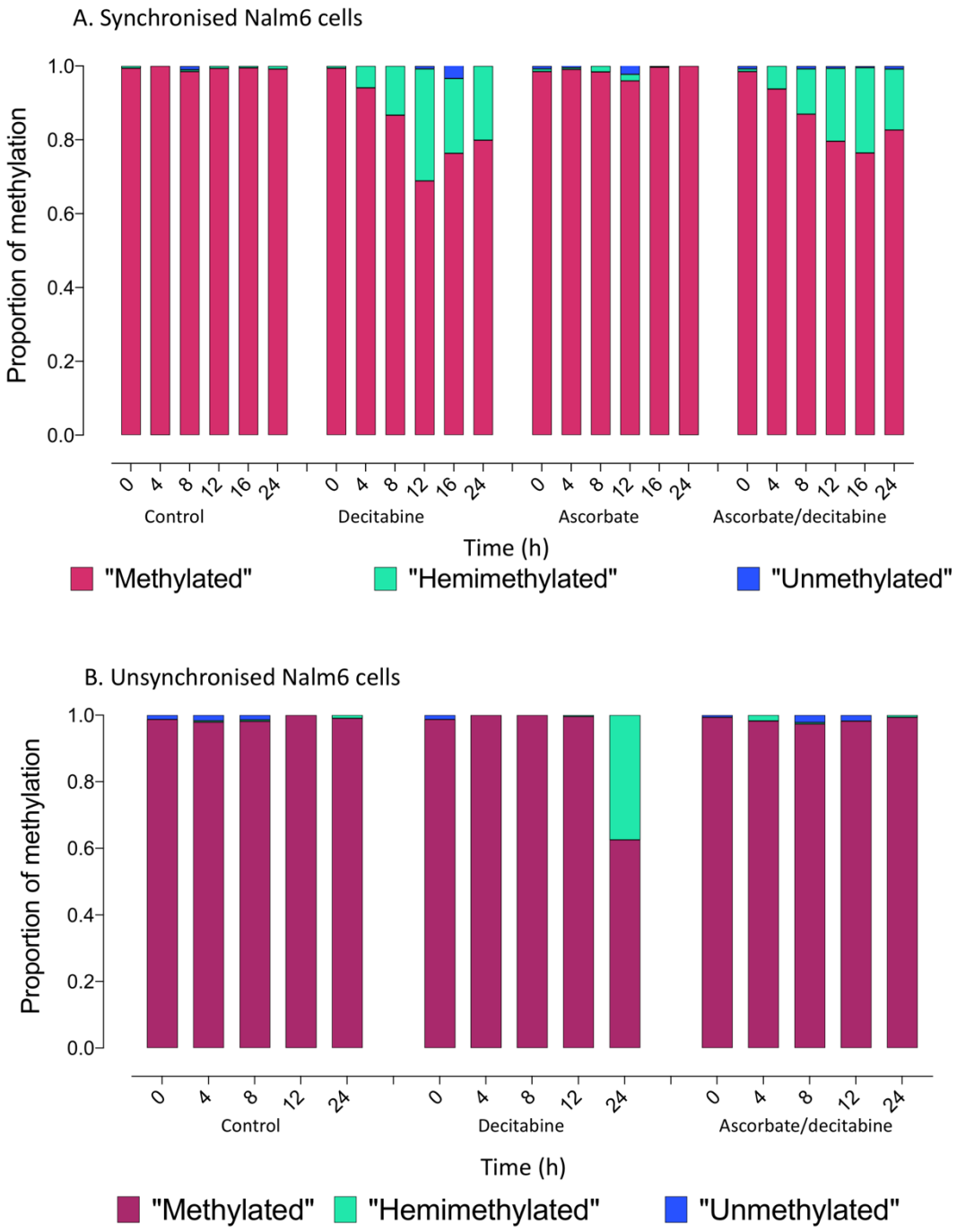


Figure 6. 33. Comparison of hairpin based methylation of *PCDHGA12* promoter in A) synchronised and B) unsynchronised Nalm6 cells.

6.7.3 Impact of ascorbate/decitabine on *PCDHGA12* promoter of HL60 cell lines.

Interestingly, synchronised HL60 cells treated with 0.5 µM decitabine or combined ascorbate/decitabine showed no marked differences on methylation pattern of *PCDHGA12* promoter compared to untreated control (Figure 6.34). A slight increase in the number of hemi-methylated and unmethylated hairpin reads was observed along decitabine and combined ascorbate/decitabine treatment at 16 h by (10% and 15% respectively).

No differences in the proportions of hemi-methylated or unmethylated *PCDHGA12* reads were seen between synchronised and unsynchronised HL60 (Figure 6.35 a,b).

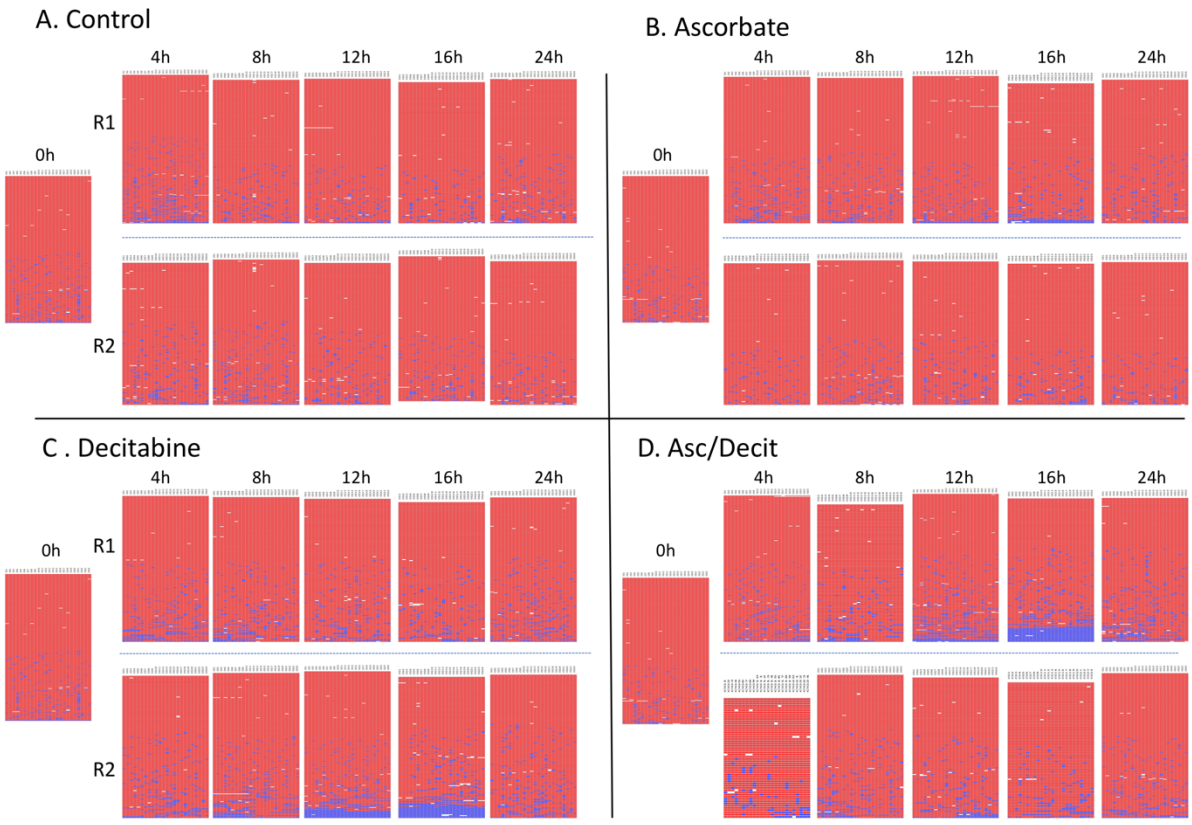


Figure 6. 34. Methylation of *PCDHGA12* promoter in synchronised *HL60* cell lines. Two biological replicates (R1 and R2) showed the methylation pattern of *PCDHGA12* promoter in synchronised A) control untreated cells, B) 500 µM ascorbate treated cells, C) 0.5 µM decitabine treated cells and D) combined ascorbate/decitabine treated cells.

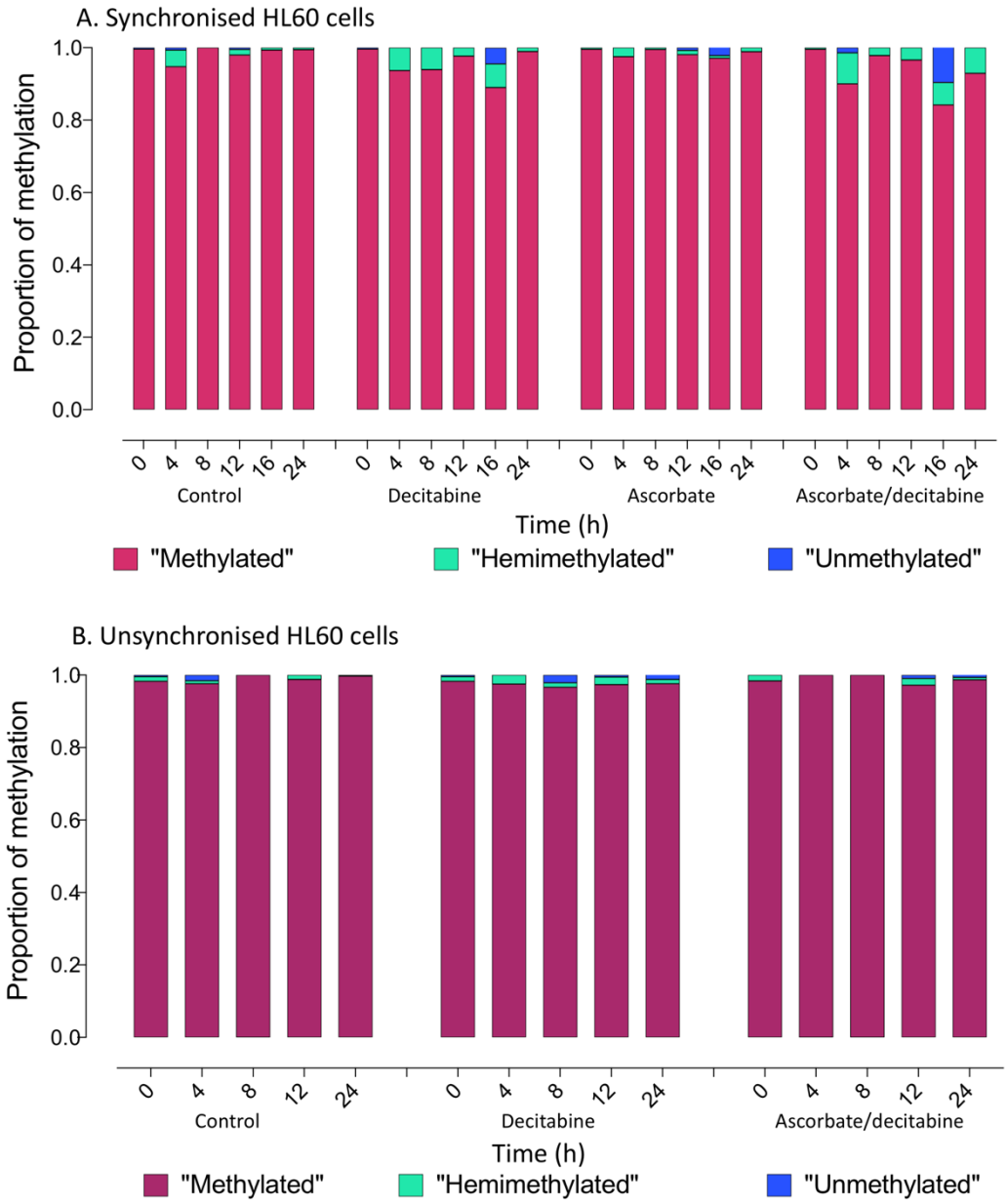


Figure 6. 35. Comparison of hairpin based methylation of *PCDHGA12* promoter among A) synchronised and B) unsynchronised HL60 cell lines.

6.7.4 Summary and conclusion to this point.

I extended the experiments to investigate the methylation signatures of *PCDHGA12* and *RASSF1* in other leukaemia cell lines. *PCDHGA12* was found to be densely methylated in all examined cell lines, whereas, *RASSF1* was densely methylated in Molt4 and less or unmethylated in the other tested cell lines. Demethylation phenomena is not unique to Jurkat cells. Similar to Jurkat cells, we observed demethylation effect of decitabine and combined ascorbate/decitabine in Molt4 and Nalm6 cell lines. Synchronisation is essential and critical step for demethylation study in Molt4 and Nalm6, as we did not see demethylation in

unsynchronised except hemi-methylation at 24 h in both cell lines. The experiments showed a concordance in the methylation results between Jurkat, Molt4 and Nalm6, which emphasised the stability of the experiments system, but active demethylation (complete unmethylation) not observed. Interestingly, the HL60 cell line seems to be resistant to demethylation following decitabine treatment, but HL60 were more sensitive to the cytotoxic effects of decitabine, as we did not observe any changes in the DNA methylation of *PCDHGA12* promoter in synchronised and unsynchronised HL60 treated with decitabine alone or combined ascorbate/decitabine. It seems that decitabine stops replication of HL60 and therefore there is no opportunity for demethylation to occur.

6.8 Hypermethylated genes identified from 450 K methylation array data set.

6.8.1 Evaluating the methylation pattern of *MARCH11*, *NEFM*, *KCNA4*, *GOLSYN* and *C10orf53* genes in Jurkat cells.

We were unable to demonstrate a consistent effect of ascorbate/decitabine in some but not all experiments on *RASSF1* and *PCDHGA12* promoters, in that active demethylation was observed. We excluded many factors that might cause the variable results such as PCR bias, DNA sequencing and co-factor components by using different culture media. We extended our investigation to other genes in Jurkat cells. The methylation patterns of *MARCH11*, *NEFM*, *KCNA*4, *GOLSYN* and *C10orf53* genes in Jurkat cells were evaluated by MiSeq sequencing to determine whether they could be utilised as new methylation markers.

The genes were identified from publicly available 450 K methylation array data set using an array-based analysis pipeline (called ABC.RAP). The pipeline is R package developed by Dr. Abdulmonem Alsaleh (previous PhD student in Morison laboratory) and available on the CRAN website (https://cran.r-project.org/web/packages/ABC.RAP/index.html). Primers were designed using MethPrimer online tool, primers were designed to amplify 357 bp region of *NEFM* gene, 345 bp region of *C10orf53* gene, 329 bp region of *GOLSYN* gene, 467 bp region of *MARCH11* gene and 432 bp region of *KCNA4* gene. Illumina universal adapter sequences were added to each primer.

To examine these five genes, DNA was extracted from ascorbate and decitabine-treated and untreated cells, then bisulfite converted using EZ DNA Methylation Gold Kit, followed by two sets of PCR amplification. Amplification of each gene was performed using KAPA HiFi DNA polymerase, and the five genes were successfully amplified. Sequencing was done using Illumina high throughput sequencing (MiSeq sequencing).

These genes were used to examine the methylation pattern and the sensitivity of these genes to decitabine and ascorbate, so the evaluation of these genes did not involve the use of the hairpins.

6.8.1.1 Evaluation of NEFM methylation.

DNA methylation analysis (MiSeq sequencing) was performed for *NEFM* in Jurkat cells. The MiSeq sequencing results showed that this gene is densely methylated with a mean methylation of 91% (Figure 6.36), suggesting that might be used as a methylated marker for DNA demethylation studies.

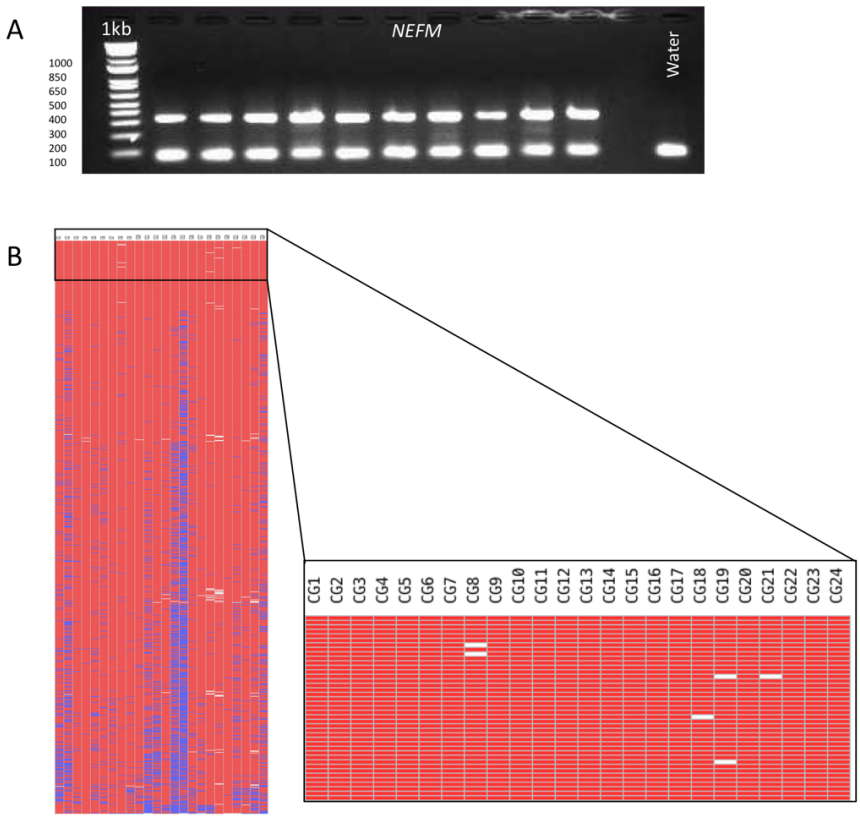


Figure 6. 36. *NEFM* methylation analysis. A) Agarose gel image confirming successful 1st round PCR amplification of *NEFM* gene (amplicon size 357 bp). B) Image of sequencing results showing densely methylated *NEFM* gene (24 CpG sites) in Jurkat cells (red = methylated CpG, blue = unmethylated CpG, white = missing sequence).

6.8.1.2 Evaluation of *C10orf53* methylation.

The sequence was selected within the promoter region of *C10orf53*. The primers were designed to amplify 345 bp of promoter region (30 CpG sites). The region was sequenced for DNA methylation analysis, and the gene was found to be completely methylated in Jurkat cells with mean methylation of 97% (Figure 6.37).

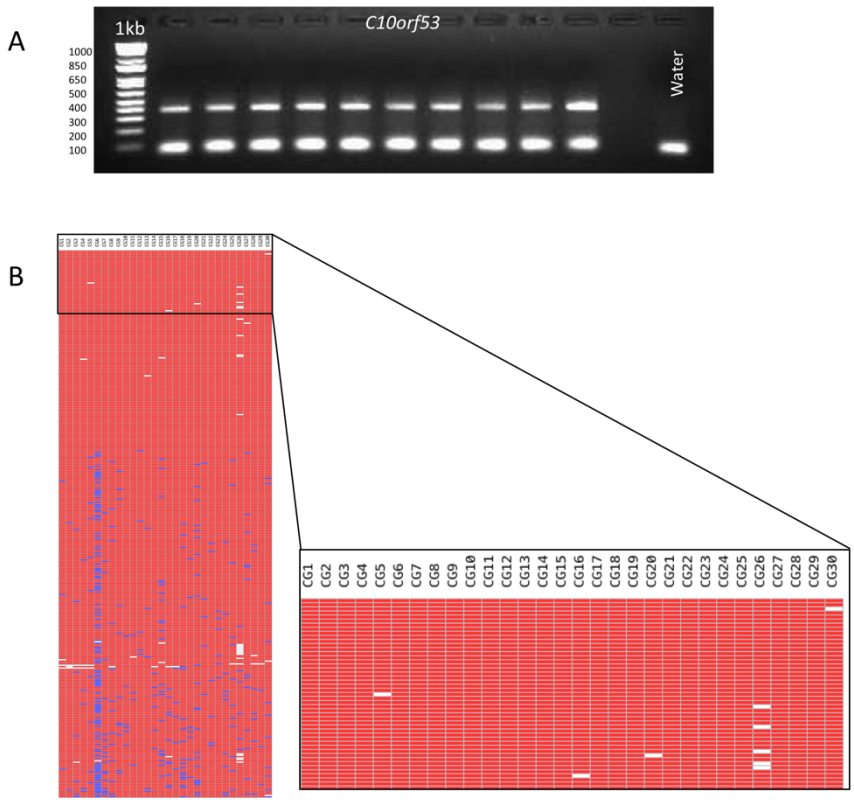


Figure 6. 37. *C10orf53* methylation. A) Agarose gel image confirming successful 1st round PCR amplification of *C10orf53* (amplicon size 345 bp). B) Image of sequencing results showing densely methylated *C10orf53* in Jurkat cells with mean methylation 97% (30 CpG sites) in Jurkat cells (red = methylated CpG, blue = unmethylated CpG, white = missing sequence).

6.8.1.3 Evaluation of *GOLSYN (SYBU)* methylation.

The target region of *GOLSYN* (also called *SYBU*) was selected within the promoter region with a size of 329 bp and 12 CpG sites. The selected region is densely methylated in B-ALL and less methylation level in normal cells (UCSC genome browser). MiSeq sequencing showed highly methylated *GOLSYN* region in untreated Jurkat cells (mean methylation 69%) (Figure 6.38).

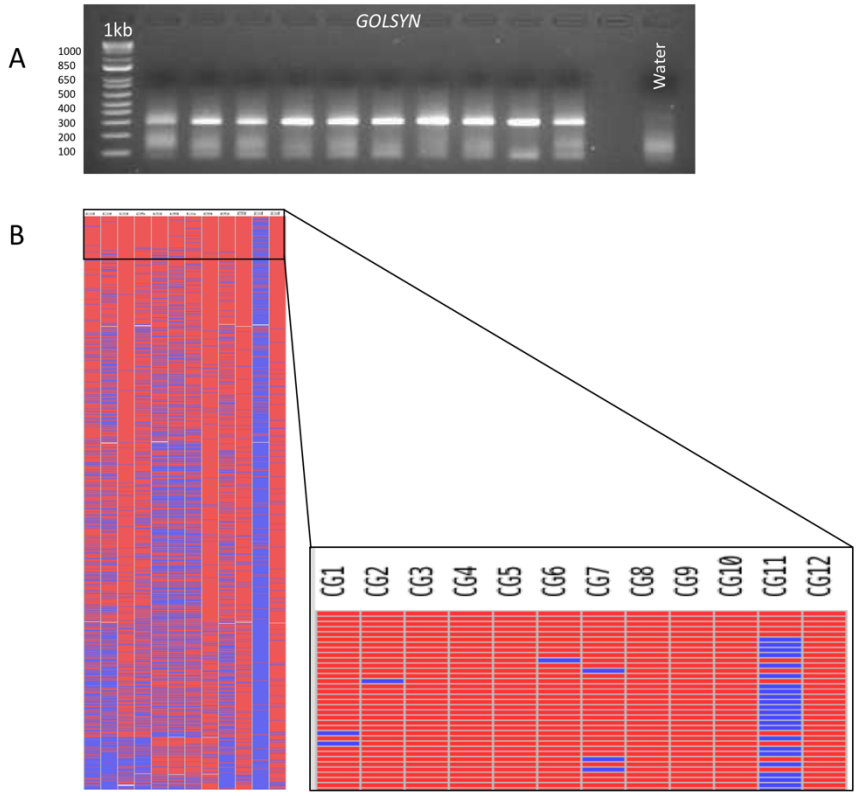


Figure 6. 38. *GOLSYN (SYBU)* methylation. A) Agarose gel image confirming successful 1st round PCR amplification of amplicon size 329 bp. B) Image of sequencing results showing densely methylated *GOLSYN* (12 CpG sites) in Jurkat cells. (red = methylated CpG, blue = unmethylated CpG, white = missing sequence).

6.8.1.4 Evaluation of *MARCH11* methylation.

A 467 bp region in *MARCH11* containing 41 CpG sites was selected. DNA methylation sequencing (MiSeq) was performed for *MARCH11* and the results showed that this gene is densely methylated in Jurkat cells (mean methylation 91%) (Figure 6.39). Therefore *MARCH11* could be an ideal gene for methylation studies.

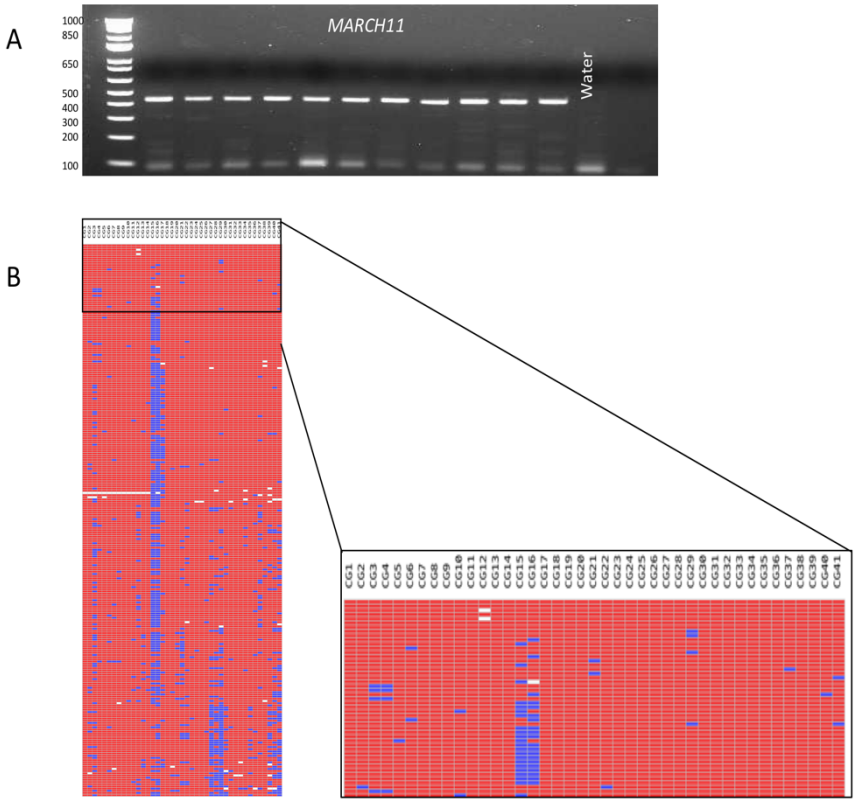


Figure 6. 39. *MARCH11* methylation. A) Agarose gel image confirming successful 1st round PCR amplification. B) Image of sequencing results showing densely methylated *MARCH11* in Jurkat cells (41 CpG sites). (red = methylated CpG, blue = unmethylated CpG, white = missing sequence)

6.8.1.5 Evaluation of *KCNA4* methylation.

A 432 bp region within the *KCNA4* promoter, contains 34 CpG sites was selected. DNA sequencing results showed highly methylated sequences (mean methylation 85%) (Figure 6.40).

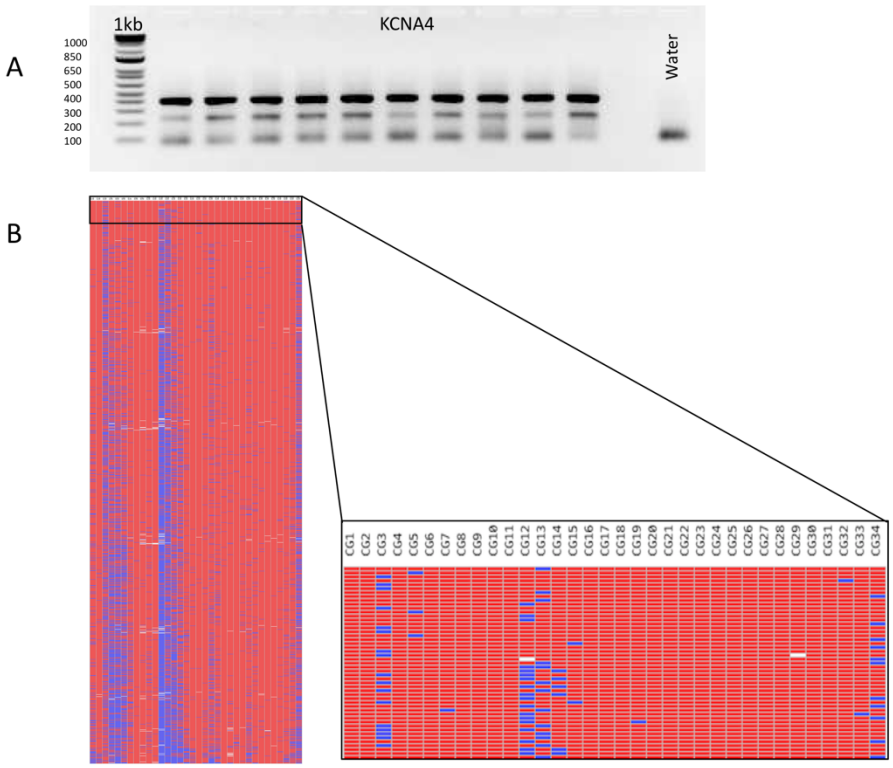


Figure 6. 40. *KCNA4* methylation. A) Agarose gel image confirming successful 1st round PCR amplification of *KCNA4*. B) Image of sequencing results showing densely methylated *NEFM* (34 CpG sites) in Jurkat cells (red = methylated CpG, blue = unmethylated CpG, white = missing sequence).

6.8.2 Effect of decitabine and ascorbate/decitabine on *NEFM*, *C10orf53, GOLSYN, MARCH11*and *KCNA4* methylation.

6.8.2.1 Impact of decitabine and combination of ascorbate/decitabine on DNA methylation of NEFM.

DNA methylation analysis was applied to evaluate the DNA methylation changes in control and treated Jurkat cells. Briefly, Jurkat cells were synchronised by thymidine block and divided into three groups (untreated control, decitabine and ascorbate/decitabine groups). After synchronisation cells were washed loaded with fresh media supplemented with 50 µM 2' deoxycytidine. At the time of release cells were treated with decitabine, ascorbate/decitabine or untreated. Cells were collected at 0, 12 and 24 h.

Incubation of Jurkat cells with 5 µM decitabine for 12 and 24 h led to a 40% decrease in methylation of *NEFM* compared to the untreated control. The combined treatment of 500 µM ascorbate and 5 µM decitabine did not alter the demethylating effects of decitabine (Figure 6.41).

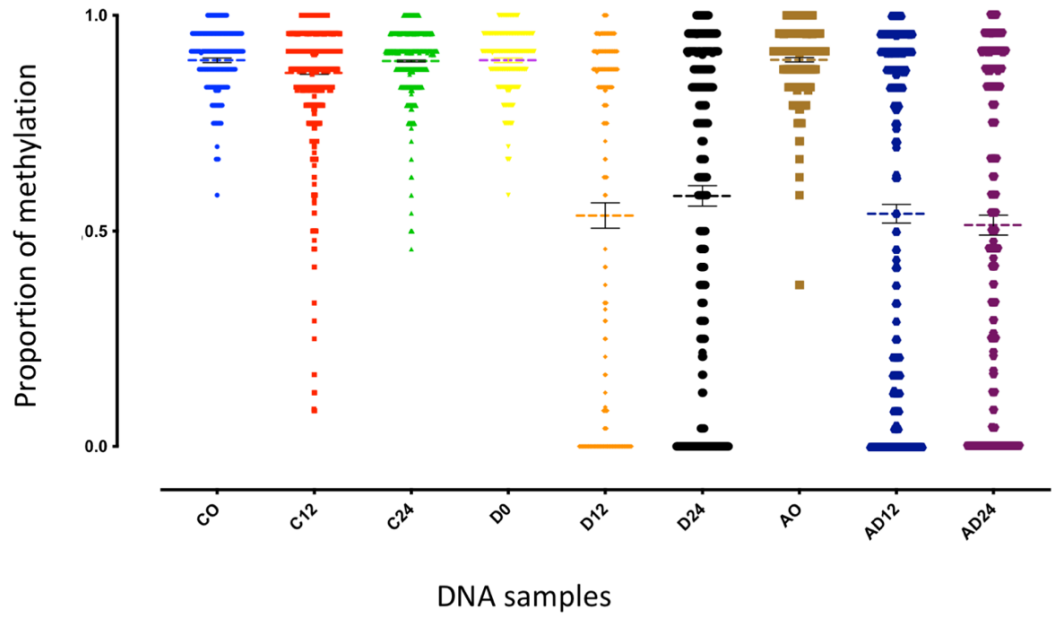


Figure 6. 41. DNA methylation of *NEFM* in Jurkat cells treated with decitabine, ascorbate, or combined ascorbate/decitabine for 24 h. (C: Control cells, D: decitabine treated cells, A: ascorbate treated cells and AD: ascorbate/decitabine treated cells.

6.8.2.2 Impact of decitabine and ascorbate on DNA methylation of C10orf53.

DNA methylation analysis was performed to evaluate the DNA methylation changes after exposure to decitabine and combination of ascorbate with decitabine. Interestingly, massive loss in DNA methylation was observed in *C10orf53* after 12 and 24 h in Jurkat cells incubated with decitabine. Co-treatment of decitabine with ascorbate had no further influence on this effect compared to the control. However, treatment with 500 µM ascorbate during synchronisation had no apparent effect (Figure 6.42).

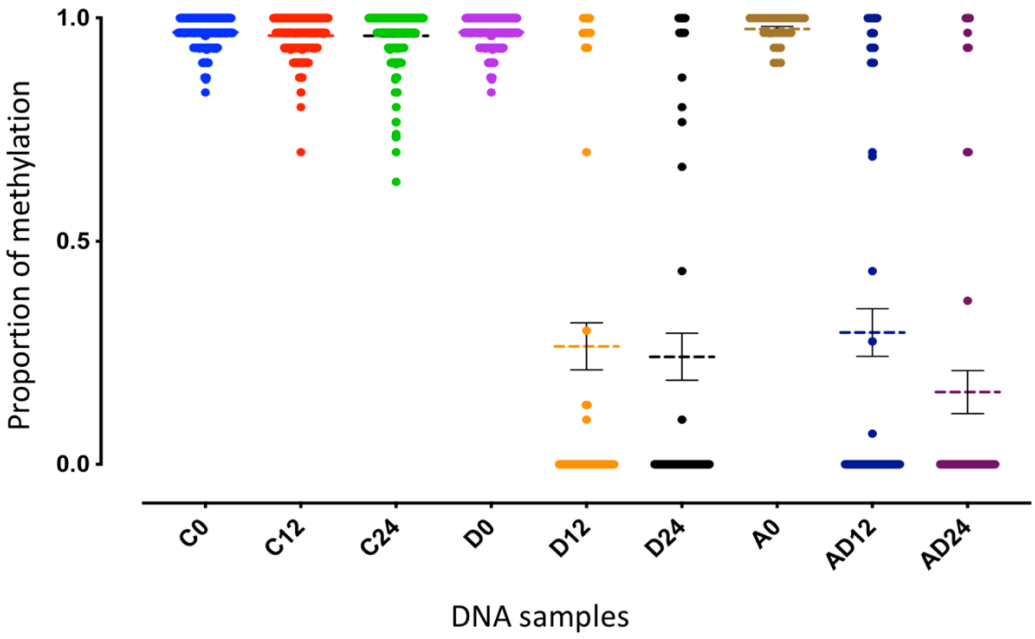


Figure 6. 42. DNA methylation of *C10orf53* in Jurkat cells treated with decitabine, ascorbate or ascorbate/decitabine for 24 h. (C: Control cells, D: decitabine treated cells, A: ascorbate treated cells and AD: ascorbate/decitabine treated cells. Error bars show SEM).

6.8.2.3 Impact of decitabine and ascorbate on GOLSYN DNA methylation.

Decitabine time dependently enhances the loss of DNA methylation in *GOLSYN* gene after 12 and 24 h incubation compared to untreated control. Co-treatment with ascorbate/decitabine had no further influence on this demethylation (Figure 6. 43).

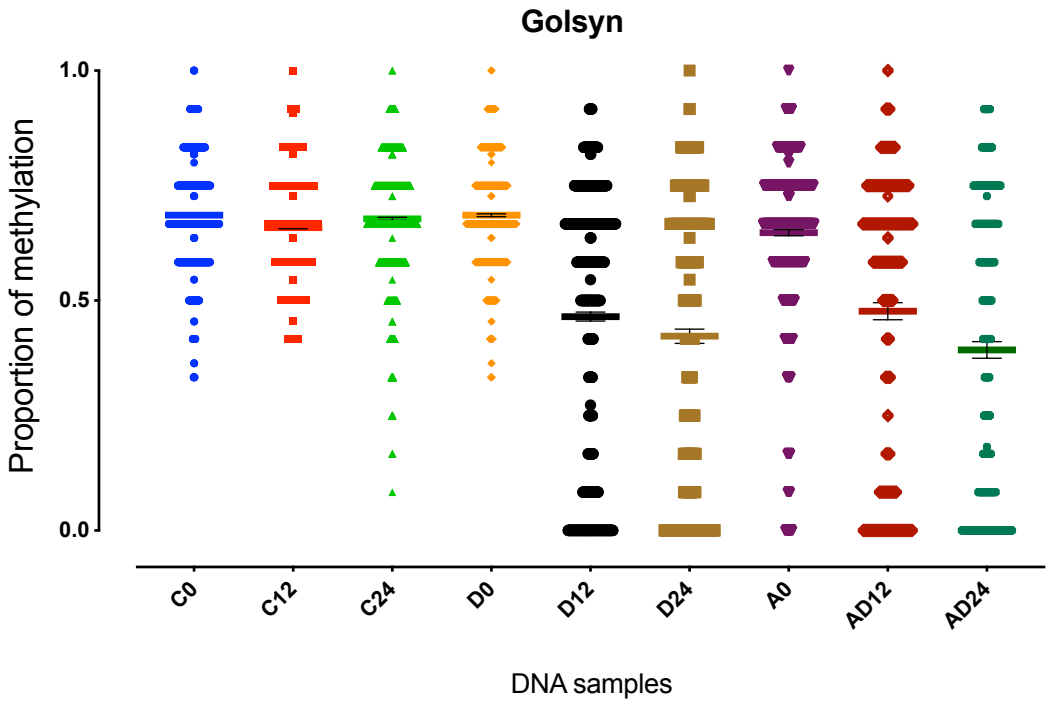


Figure 6. 43. DNA methylation of *GOLSYN* in Jurkat cells treated with decitabine, ascorbate or ascorbate/decitabine for 12 and 24 h. (C: Control cells, D: decitabine treated cells, A: ascorbate treated cells, AD: ascorbate/decitabine treated cells).

6.8.2.4 Impact of decitabine and ascorbate on MARCH11 DNA methylation.

Reduction in *MARCH11* DNA methylation levels occurred in a time dependent manner in Jurkat cells incubated with 5 µM decitabine for 12 and 24 h. The reduction in DNA methylation levels induced by decitabine was not affected by further addition of ascorbate (Figure 6.44).

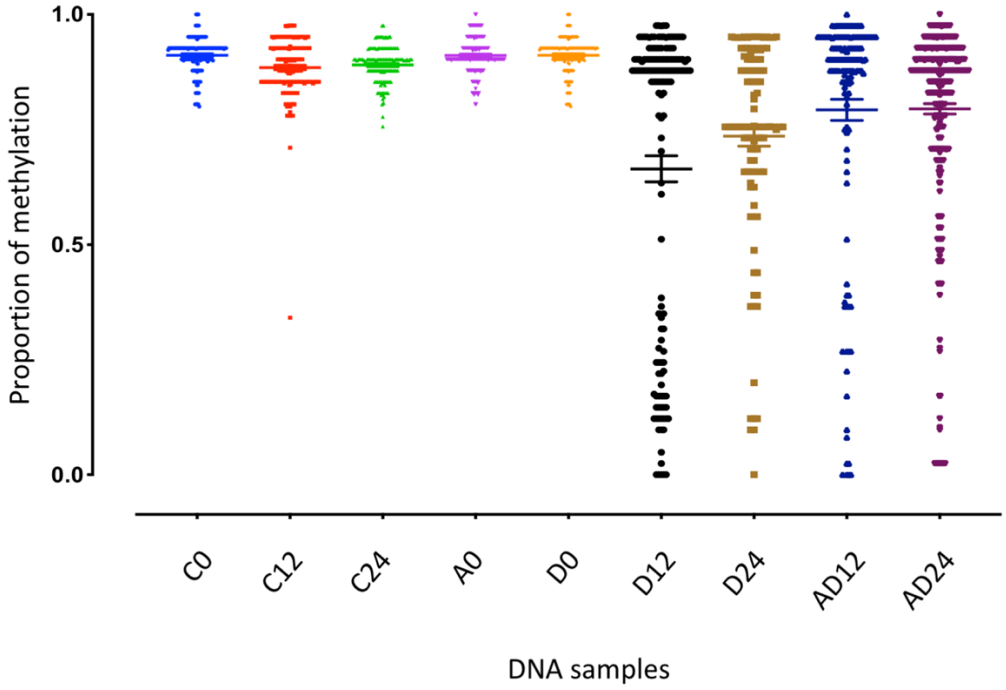


Figure 6. 44. DNA methylation of *MARCH11* in Jurkat cells treated with decitabine, ascorbate or ascorbate/decitabine for 24 h. (C: control cells, D: decitabine treated cells, A: ascorbate treated cells and AD: ascorbate/decitabine treated cells).

6.8.2.5 Impact of decitabine and ascorbate on DNA methylation of KCNA4.

The impact of decitabine and a combination of ascorbate/decitabine on *KCNA4* DNA methylation was investigated in Jurkat cells. Interestingly, rapid loss in DNA methylation was observed in *KCNA4* gene after 12 and 24 h in Jurkat cells incubated with decitabine. Co-treatment with ascorbate had no further influence on this effect compared to the control. (Figure 6.45).

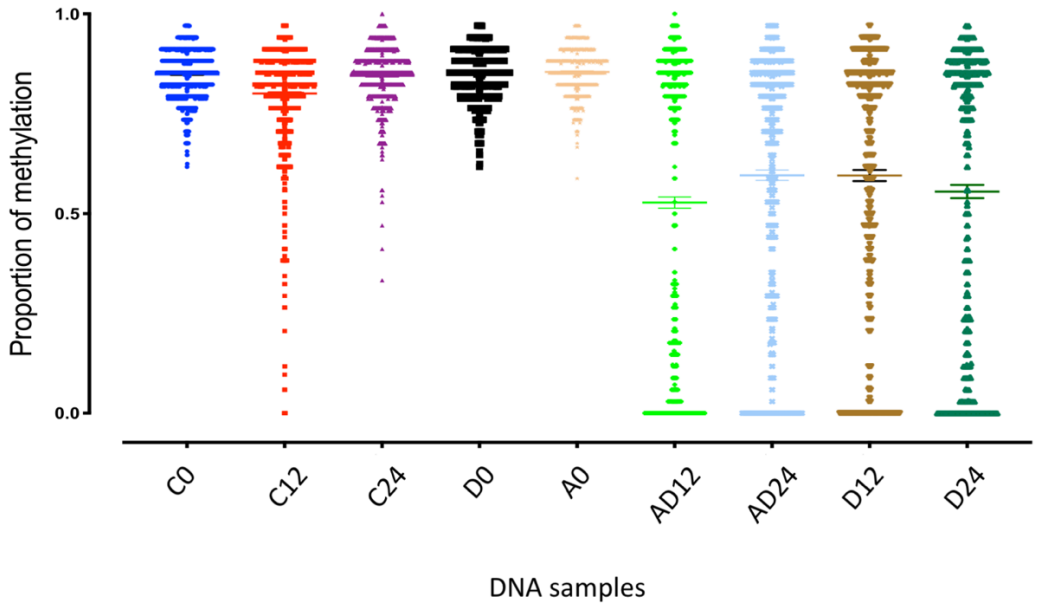


Figure 6. 45. DNA methylation of *KCNA4* gene in Jurkat cells treated with decitabine, ascorbate and ascorbate/decitabine for 12 and 24 h. (C: Control cells, D: decitabine treated cells, A: ascorbate treated cells and AD: ascorbate/decitabine treated cells).

6.9 Analysis of publicly available data using ABC.RAP R package.

To support our finding, other evidence of relevant data from publicly available data set was carried out analysed by Dr. Abdulmonem Alsaleh using an array-based analysis pipeline called ABC.RAP R package. Briefly, the package utilised probes obtained from 450K analysis, filtered non-specific probes, annotated the filtered out probes, measured significantly different CpG sites and applied delta beta analysis to measure the differences between treated and control samples. The data were extracted from different cohorts and analysed for the effect of ascorbate/decitabine. The package applied 'two sided' unequal variance Student's t-test analysis to select CpG sites with significant differences in DNA methylation 'beta value'. The delta beta value is the difference in mean DNA methylation between treated (ascorbate and

decitabine) and untreated control. These data was consistent with our results and supported our demethylation finding.

6.9.1 Global DNA methylation analysis: HCT116 CRC cell line.

Interestingly, using the ABC.RAP R package, the analysis of publicly available data extracted from Li et al.(2016) showed global loss in DNA methylation in cells treated with decitabine and combined ascorbate/decitabine compared to control cells. Ascorbate alone had no effect on global DNA methylation (Figure 6.46) (Figure 6.47) (Liu et al., 2016). The global effect of decitabine alone or in combination with ascorbate are concordant with loss of DNA methylation detected by our MiSeq sequencing (gene specific) and PBAT (global) results.

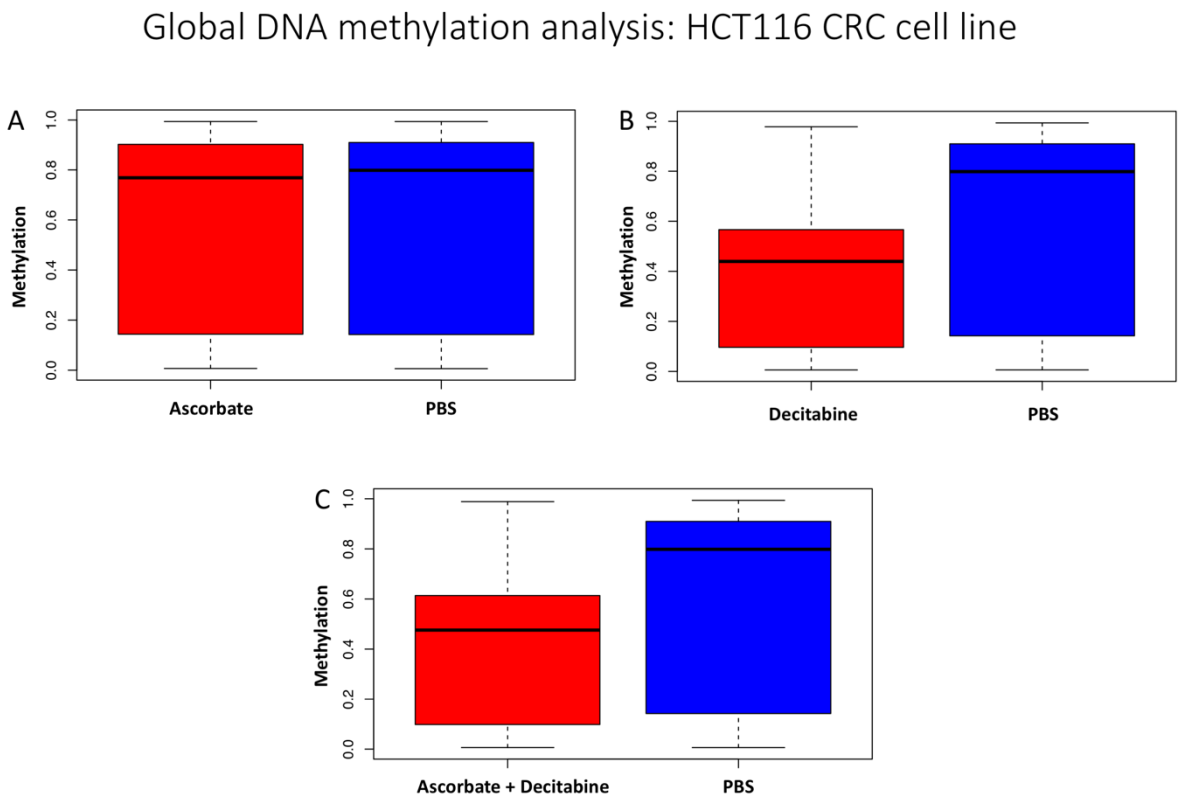


Figure 6. 46. Global DNA methylation analysis of HCT116 CRC cell lines treated with A) 57 µM ascorbate, B) 300 nM decitabine or C) combined ascorbate/decitabine for 24 h compared to untreated control cells. Data extracted from Lin et al (2016).

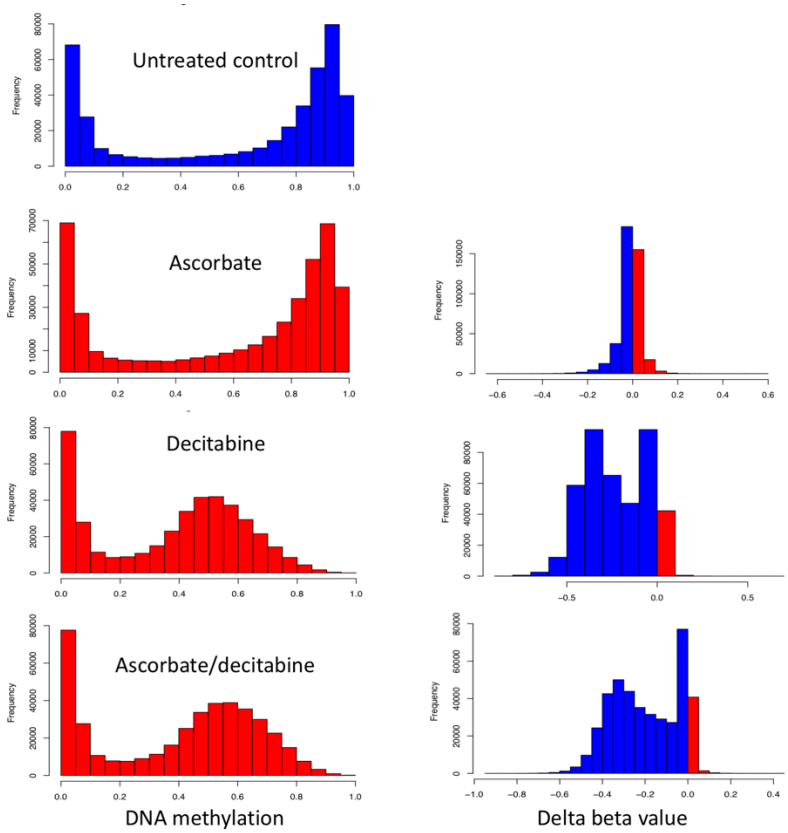


Figure 6. 47. Distribution of methylation for HCT116 CRC cell line. A) Methylation of decitabine treated cells compared to untreated control, and delta beta (the difference of DNA methylation between decitabine and untreated control. B) Methylation of ascorbate treated cells compared to untreated control, and delta beta between ascorbate treated and untreated control. C) Methylation of combined ascorbate with decitabine compared to untreated control, and delta beta value, the difference between combined ascorbate with decitabine and untreated control.

Additionally, we compared our methylation results with results from a study by Imanishi et al. (2017). This study examined global methylation of U937 haematopoietic cells incubated with decitabine for 24 and 72 h.

Using the *ABC.RAP* R package, global DNA methylation of U937 cells treated with 5 μM decitabine after 24 and 72 h decreased when compared to untreated cells (Figures 6.48 and 6.49). These results are concordant with our MiSeq sequencing results on Jurkat cells.

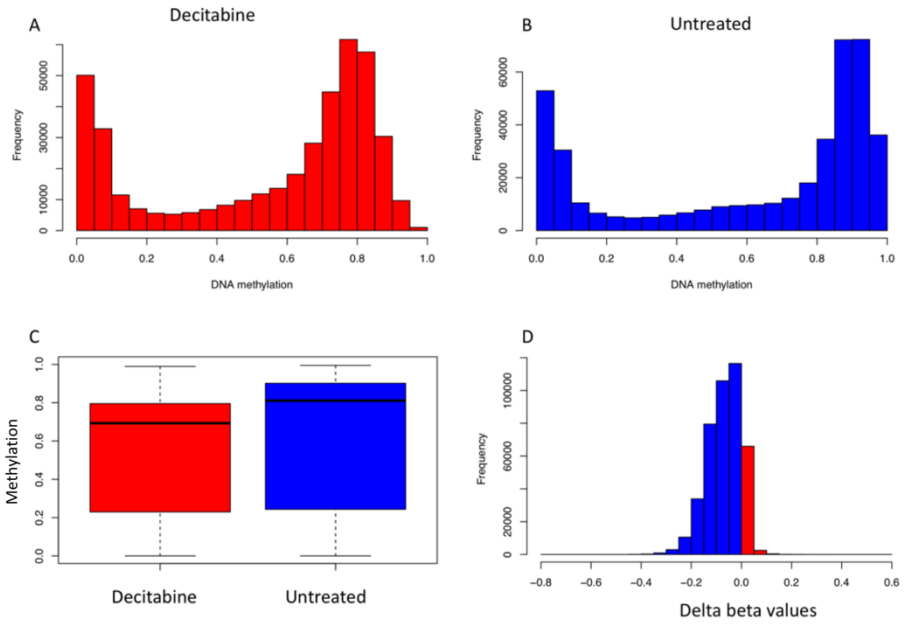


Figure 6. 48. Methylation analysis of U937 cells after 24 h decitabine treatment. A+B) Histogram showing the DNA methylation distributions of decitabine and untreated control cells, respectively. C) Boxplot showing the DNA methylation of decitabine vs untreated, D) The difference in methylation (delta beta values) between decitabine and untreated cells. Data extracted from Imanishi et al. (2017).

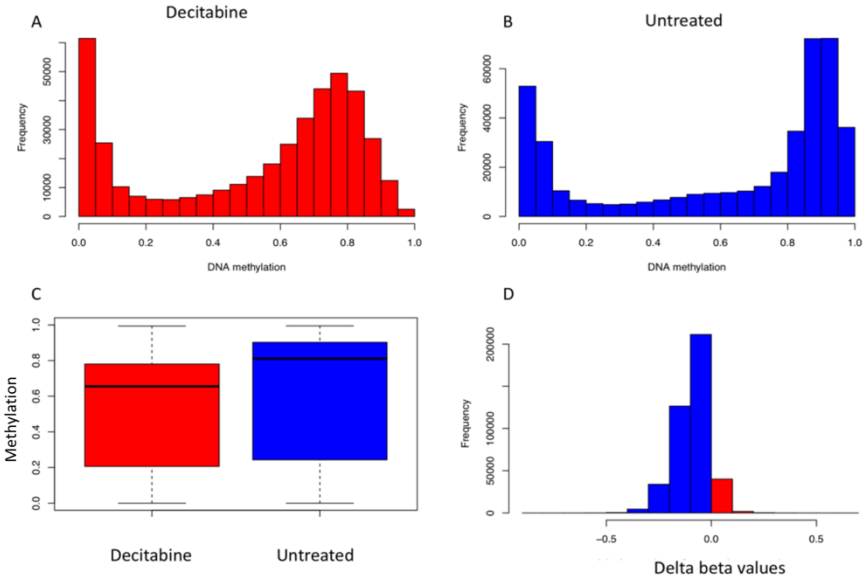


Figure 6. 49. Methylation analysis of U937 cells after 72 h decitabine treatment. A+B) Histogram showing the DNA methylation distributions of decitabine and untreated control cells, respectively. C) Boxplot showing the DNA methylation of decitabine vs untreated, D) The difference in methylation between decitabine and untreated cells. Data extracted from Imanishi et al. (2017).

6.10 Discussion.

Active demethylation is defined as the enzymatic removal of a methyl group from cytosine nucleotides in DNA. Cytosine analogues, e.g. decitabine and azacytidine, are commonly used as demethylating drugs in treating haematological cancers, such as Acute Myeloid Leukaemia (AML) and Myelodysplastic syndrome (MDS) (Pleyer & Greil, 2015).

In this study, we investigated DNA demethylation in a cancer model-Jurkat cells (T-ALL cell line). We developed a tractable hairpin bisulfite sequencing assay to investigate active demethylation. With this assay, we consistently demonstrated hemi-methylated sequences in haematopoietic cell lines treated with decitabine. In addition, we demonstrated active demethylation in cells treated with decitabine.

Through further experiments, we have shown that treatment with ascorbate occasionally enhances the demethylation efficacy of decitabine. In line with this data, we observed a significant increase in the number of unmethylated reads in Jurkat cells co-incubated with ascorbate/decitabine paralleled by an increase in the presence of 5-hydroxymethylcytosine compared to untreated cells (Chapter Seven). Combined ascorbate with decitabine induces active demethylation resulted by conversion of 5-methyl cytosine to 5-hydroxymethyl cytosine leads to upregulate tumour suppressor genes which leads to arresting the cells in G1, S or G/M phases of the cell cycle (Sajadian et al., 2016a). It is interesting to know that decitabine alone induced active demethylation by enhancing the production of 5-hydroxymethylcytosine at 12 h timepoint. Various types of cancer cells exhibited loss of 5-hydroxymethylcytosine which was suggested as a demethylation marker for cancer cells (Huang & Rao, 2014; Kroeze et al., 2015).

The hypomethylating agents decitabine and azacytidine were able to reduce 5-methylcytosine levels in colon cancer cells in a concentration-dependent manner and were not affected by the addition of ascorbate. Interestingly, the hypomethylating agents alone failed to trigger any changes to 5-hydroxymethyl cytosine levels (Gerecke et al., 2018). In contrast, another study by Chowdhury (2015) reported that decitabine increased the 5-hydroxymethylcytosine levels in HL60 (AML) and TK6 (CML) leukaemia cell lines (Chowdhury et al., 2015b).

Although, TET proteins depend on two substantial factors Fe^{2+} and alpha-ketoglutarate in order to mediate the iterative oxidation of 5-methylcytosine to 5-hydroxymethyl cytosine and mediate active demethylation (Ito et al., 2010; Tahiliani et al., 2009), in the present study, we inferred that adding ascorbate to decitabine treatment enhanced the epigenetic activity of decitabine by induction of TET proteins and rapid generation of 5-hydroxymethylcytosine.

Untreated Jurkat cells are completely ascorbate deficient, and supplementation optimises intracellular total and reduced ascorbate concentrations (Figure 6.3). The generation of 5-hydroxymethyl cytosine is completely absent in untreated Jurkat cells at different timepoints during cell cycle progression (Chapter Seven). In lines with other papers, there is a remarkable correlation between TET proteins, tumour progression and the production of 5-hydroxymethyl cytosine which suggest that TET proteins might act as a tumour suppressor in certain cancer types (Haffner et al., 2011; Xu, Wu, et al., 2011).

Sajadian et al. (2015) observed a significant induction of TET2 and TET3 and 5-hydroxymethyl cytosine in HCC cell lines at 20 μM decitabine treatment. Interestingly, combination of 10 μM decitabine with ascorbate showed the same degree of epigenetic changes similar to 20 μM decitabine, and ascorbate was able to compensate the higher dose of decitabine (Sajadian et al., 2015b).

Through further study, Sajadian et al. (2016) showed that co-treatment of hepatocellular carcinoma (HCC) cell lines with ascorbate and 5-AZA induced active demethylation and production of 5-hydroxymethylcytosine. Furthermore, they postulated that this demethylation enhanced expression of GADD45 and decreased expression of *Snail*, leading to inhibition of cyclin B and PCNA expression and induction of cell cycle arrest (Sajadian et al., 2016a).

We have shown that decitabine consistently induced cell cycle delay after 24 h and hemi-methylation started as earlier as 2 h. Cell viability decreased in the four leukaemia cell lines treated with 5 μM decitabine in a time dependent manner. Furthermore, we observed a dramatic cytotoxic increase when Jurkat cells were co-incubated with ascorbate/decitabine represented by a sharp decrease in the number of viable cells after co-treatment.

Here, we extended our study to investigate whether ascorbate can enhance the demethylation efficacy of decitabine in another three haematopoietic cell lines: Molt4, Nalm6 and HL60. Ascorbate has very little effects on DNA methylation, but with combination of ascorbate/decitabine the methylation decreased in Molt4 and Nalm6. Cell synchronisation is an essential step in our DNA demethylation experiments. HL60 is resistance to decitabine as we did not see a marked effect of decitabine or combined ascorbate/decitabine on global or gene specific methylation in both synchronised and unsynchronised.

Chapter Seven: Effect of decitabine and ascorbate on DNA methylation and hydroxymethylation in Jurkat cell lines.

7.1 Introduction.

Whole-genome bisulfite sequencing or 'WGBS' is used to measure genome-wide methylation and has led to novel discoveries (Olova et al., 2018). However this protocol is expensive and impractical for large numbers of samples and typically requires large amounts (5 μg) of DNA as starting material, which is sometimes difficult to prepare from many samples, such as early embryo and embryonic tissue.

Alternatively, low-coverage bisulfite sequencing or PBAT, is cheap, unbiased and efficient, requiring only small amounts of input DNA and is able to process large numbers of DNA samples. Using modified PBAT (Peat et al., 2014) we quantified global CpG methylation in Jurkat cells treated with different reagents.

5-hydroxymethylcytosine (5-hmC) is the first intermediate epigenetic mark in TET-mediated active demethylation (Wu & Zhang, 2011a) and is generated by TET1-3 enzymes catalysed oxidation of 5mC (Ito et al., 2010; Tahiliani et al., 2009) (Figure 7.12). As a "sixth" DNA base, 5-hmC plays an important role in phenotype and gene expression (Shi et al., 2017).

To assess whether decitabine or/and ascorbate enhance the production of 5-hmC in synchronised Jurkat cells, Global Hydroxymethylation ELISA assay was used to quantify global 5-hydroxymethylcytosine in treated Jurkat at different timepoints after release from synchronisation.

7.1.1 The aims of this chapter:

1. To determine the global 5-mC level in Jurkat cells using low coverage bisulfite sequencing.
2. To evaluate the effect of decitabine on global DNA methylation of synchronised Jurkat cells.
3. To investigate the effect of ascorbate and combined ascorbate and decitabine treatments on global DNA methylation.

4. To determine the global 5-hmC level in Jurkat cells treated with decitabine, ascorbate and combined ascorbate and decitabine.

7.2. Low coverage bisulfite sequencing (post-bisulfite adapter tagging): Experiment one.

7.2.1 Preparation of PBAT libraries.

Using PBAT, the genome-wide DNA methylation of cells treated with decitabine and other reagents was monitored over 72 h.

For this experiment, Jurkat cells were grown to 1 x 10^6 /mL and arrested in G1 phase by incubation with 2 mM 2′-deoxythymidine for 18 h. Arrested cells were washed twice with fresh media and divided into two five groups (1: control, 2: 5 μM decitabine, 3: 500 μM glycine chloramine, 4: 500 μM ascorbate and 5: 500 μM ascorbate and 5 μM decitabine). Cells were supplemented with 2′-deoxycytidine to initiate DNA synthesis (defined as 0 h). Cells were harvested at different timepoints 0, 2, 4, 6, 24, 48 and 72 h.

The treatment and tissue culture work for this experiment was performed by me in Prof. Hampton's laboratory (University of Otago, Christchurch).

After DNA purification, global 5-mC levels were determined using PBAT sequencing. Briefly, genomic DNA for each sample was bisulfite converted and the DNA was copied by primer extension with biotin-labelled random primers. After purification with streptavidin-coated DynaBeads and magnetic immobilisation, immobilized first strand was used as a template for an additional adapter to create double-strand DNA. PCR was used to add a unique molecular barcode and to add sequences necessary for binding to Illumina flow cells. Libraries were sequenced on a MiSeq instrument (Illumina) (explained in section 3.5) (Figure 7.1) and analysed by Dr. Tim Hore.

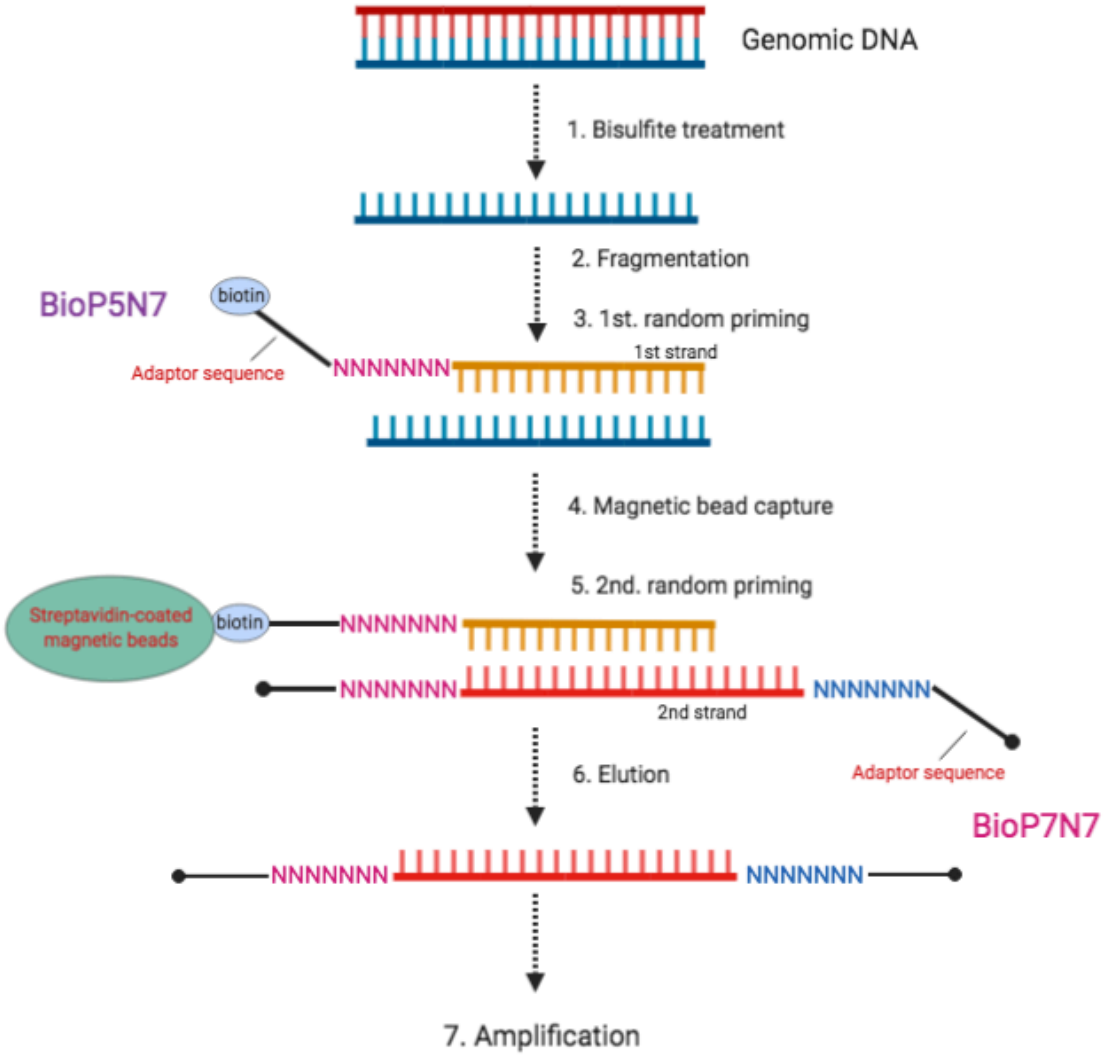


Figure 7. 1. General PBAT procedure consisting of two rounds of random priming on bisulfite-treated DNA. Modified from (Miura et al., 2012).

PBAT amplification quality and quantity was checked by running the libraries on 1% agarose gel electrophoresis at 140 V for 45 min. (Biotium Gelred stain). The resulting electrophoresed DNA shows a smear from 100 to >1000 bp; the smear consistency suggested that the DNA fragments were amplified to the same quantity (Figure 7.2 'left panel'). PCR product libraries were pooled based on the samples fluorescent intensity of the gel electrophoresis (Figure 7.2 'right panel').

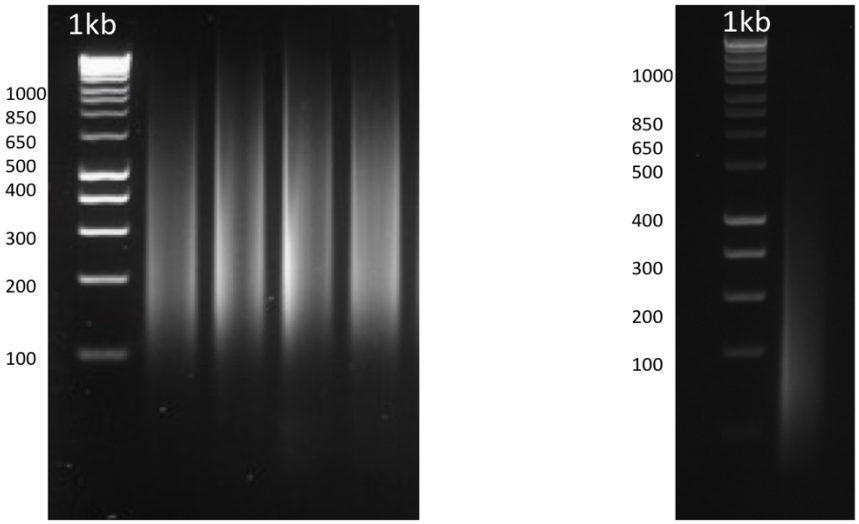


Figure 7. 2. Agarose gel showing PBAT PCR amplified products of different DNA samples. Left panel shows gel electrophoresis of PBAT libraries. Lane 1: DNA ladder (100 bp), lane 2, 3 and 4 consist of PBAT libraries from 100 bp. Right panel shows the pooled PBAT PCR libraries. (Images were taken from Dr. Tim Hore laboratory).

Low coverage 150 bp paired-end Illumina MiSeq sequencing was performed to evaluate the global CG methylation levels for all samples. Illumina MiSeq output FASTQ files were monitored for read quality by FASTQC before and after quality trimming to remove poor quality sequences, and removal of adapter sequences using TrimGalore (v0.4.0, default parameters). Trimmed reads were then mapped to the human genome (hg19) and total methylation levels in CpG and non-CpG contexts were determined using BISMARK v0.4.0, default parameters) (Krueger & Andrews, 2011) with the PBAT option specified by Hore's laboratory. The precision of the methylation measurement increases with the number of analysed CpGs. For example at 15000 CpG cells the margin of error (99% confidence interval) is ± 1.8 percentage points (Peat et al., 2017).

7.2.2 Impact of decitabine on global genomic DNA methylation in Jurkat cells.

The number of analysable CpGs ranged from 1000 to 33,500, except for two samples (ascorbate plus and ascorbate minus at 0 h), that had very low CpG calls (9 and 10 calls, respectively) and were discarded (Figure 7.3).

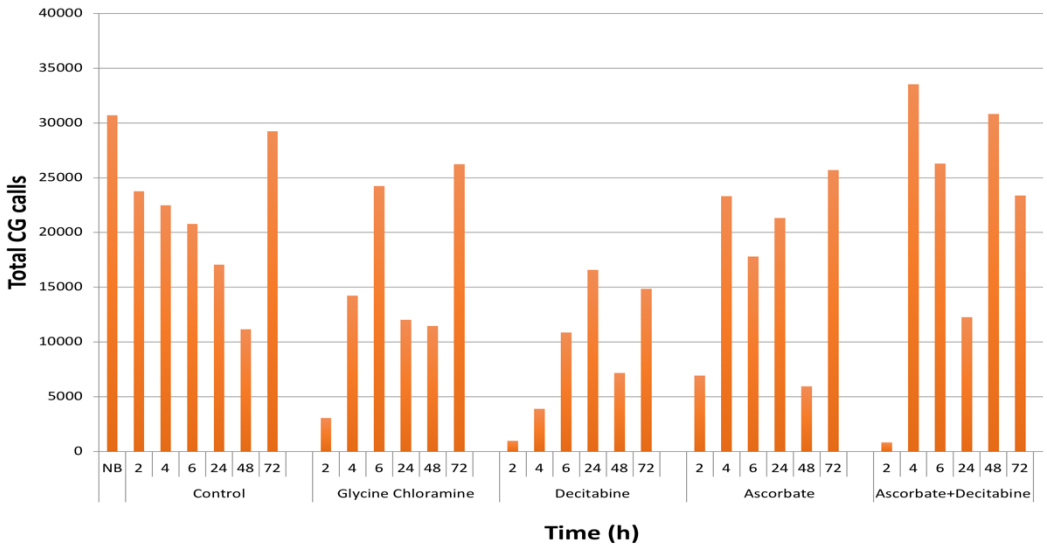


Figure 7. 3. Number of CG calls for Jurkat DNA samples examined by low-coverage sequencing.

Quantification of non-CG "methylation" is another parameter measured during PBAT analysis to examine the completeness of bisulfite conversion. Because methylation in mammals is targeted to CGs, completely converted samples are characterised by low levels of non-CG "methylation". Non-CG "methylation" < 1% in all samples (Figure 7.4).

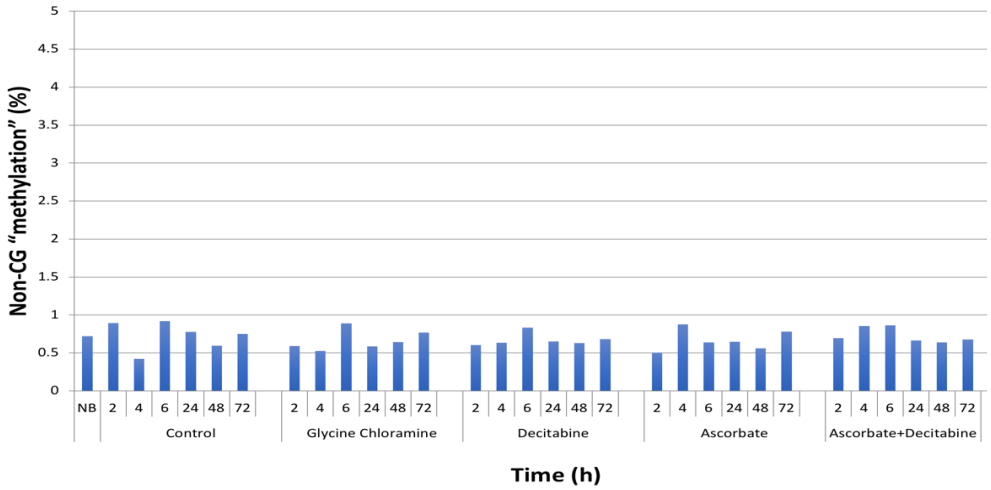


Figure 7. 4. Percent of non-CG "methylation" for Jurkat DNA samples.

All samples had mapping efficiency of 53-67%, which is consistent with previous results from Dr. Hore's laboratory (Figure 7.5).

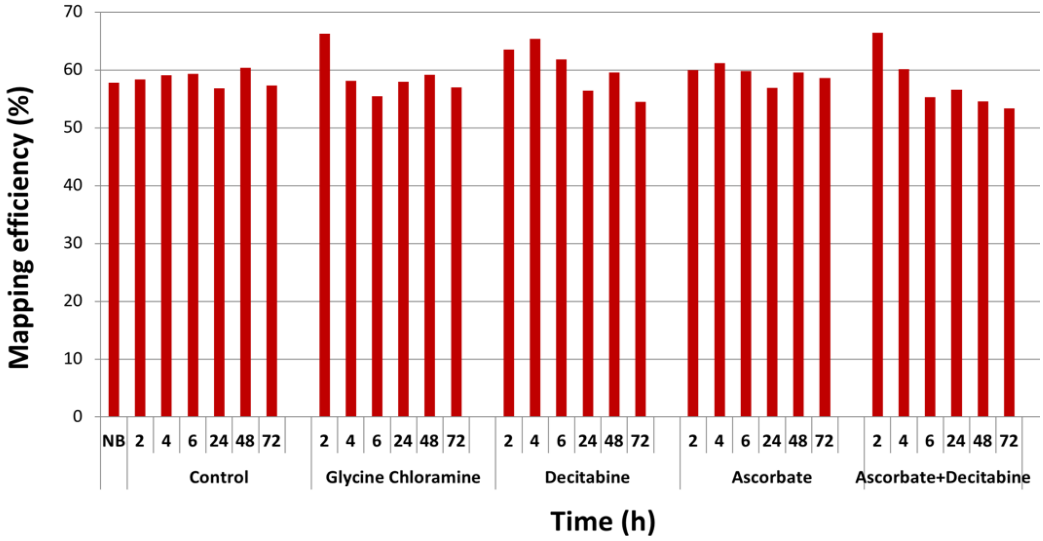


Figure 7. 5. Mapping efficiency (%) for Jurkat DNA samples.

Low-coverage genome wide sequencing results demonstrated that there is a constant CpG methylation level in non-blocked (67%) and in arrested Jurkat cells (68%). As expected, there is a gradual loss of CpG methylation in Jurkat cells treated with decitabine and combined ascorbate/decitabine in a time-dependent manner following release from cell cycle arrest (from 67% in non-blocked Jurkat cells to 30% for decitabine treated cells at 72 h) (Figure 7.6).

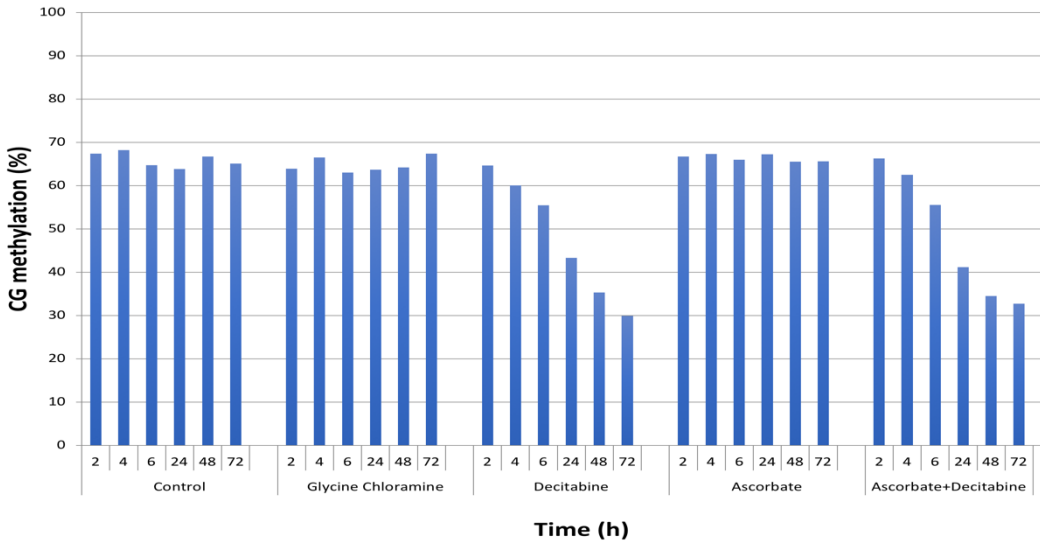


Figure 7. 6. Global CpG methylation levels (%) of Jurkat DNA treated with glycine chloramine, decitabine, ascorbate and combined ascorbate/decitabine compared to control cells.

Furthermore, incubation of Jurkat cells with 500 μM glycine chloramine or 500 μM ascorbate alone had no effect on global CG methylation levels.

In this study, we performed the low coverage bisulfite sequencing assay to confirm the gene specific hairpin bisulfite sequencing results (Figure 7.7) (previously shown in Chapter Six). The results of low coverage bisulfite sequencing showed consistent methylation results in both global methylation by PBAT and gene specific hairpin DNA methylation results. Figure 7.6 and Figure 7.7 both show reduction in methylation with decitabine and similar reduction with combined ascorbate/decitabine.

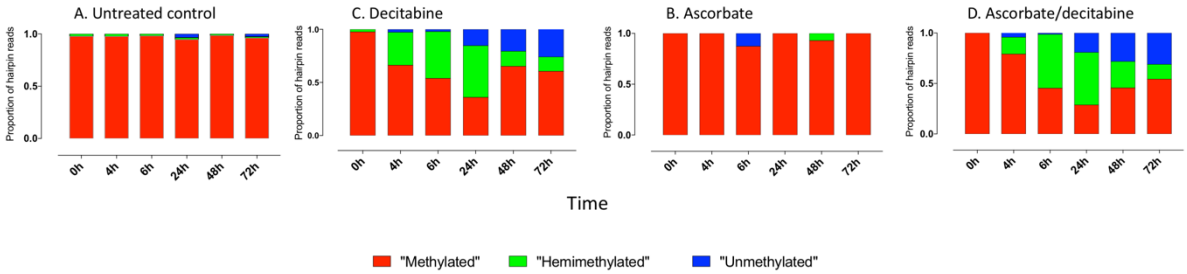


Figure 7. 7. Impact of ascorbate/decitabine on DNA methylation of Jurkat cells. Hairpin based methylation of *PCDHGA12* promoter following single or combined treatment. A) control untreated cells; B); cells treated with 5 µM decitabine C) cells treated 500 µM ascorbate and D) cells treated with combined ascorbate/decitabine.

7.3 Low-coverage genome wide sequencing PBAT: Experiment Two.

To confirm these results, PBAT was performed on samples collected from a second independent experiment.

In this experiment, performed by me in Prof Morison's laboratory (University of Otago, Dunedin), Jurkat cells were grown to 1 x 10^6 /mL and arrested at G1 phase by incubation with 2 mM 2′-deoxythymidine for 18 h. Cells were washed twice with fresh media and divided into two four groups (Group 1: control, Group 2: 5 µM decitabine, Group 3: 500 µM ascorbic acid and Group 4: 500 µM ascorbic acid and 5 µM decitabine). To initiate DNA synthesis cells were supplemented with 2′-deoxycytidine. Cells were harvested at different timepoints 0, 6, 12, 24, 48, 72, 96 and 120 h. The CG call numbers for this experiment ranging from 14,500 to 50,611 (Figure 7.8). These results showed that the non-CpG levels was less than 1%, indicating that the bisulfite conversion was at least 99% efficient (Figure 7.9).

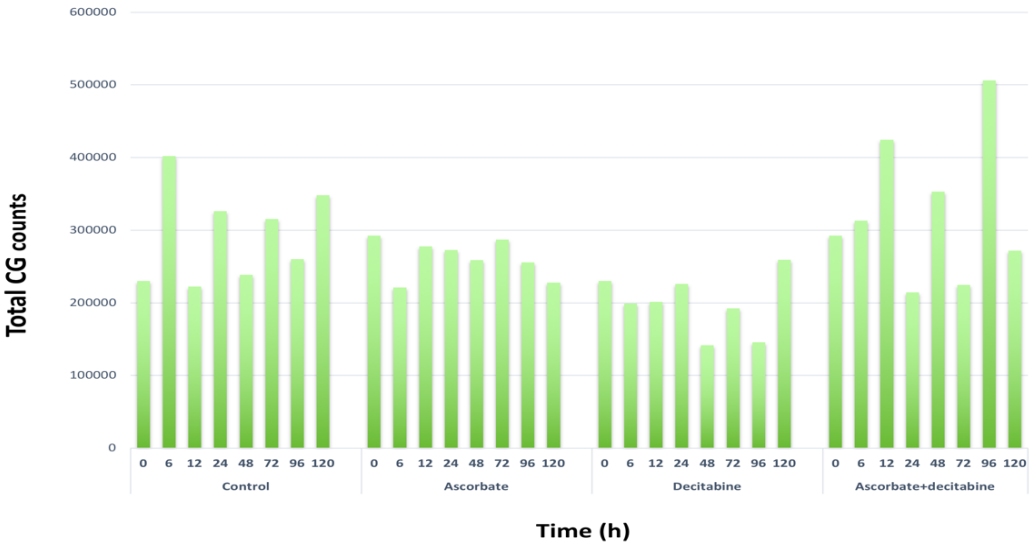


Figure 7. 8. CG count for Jurkat DNA samples.

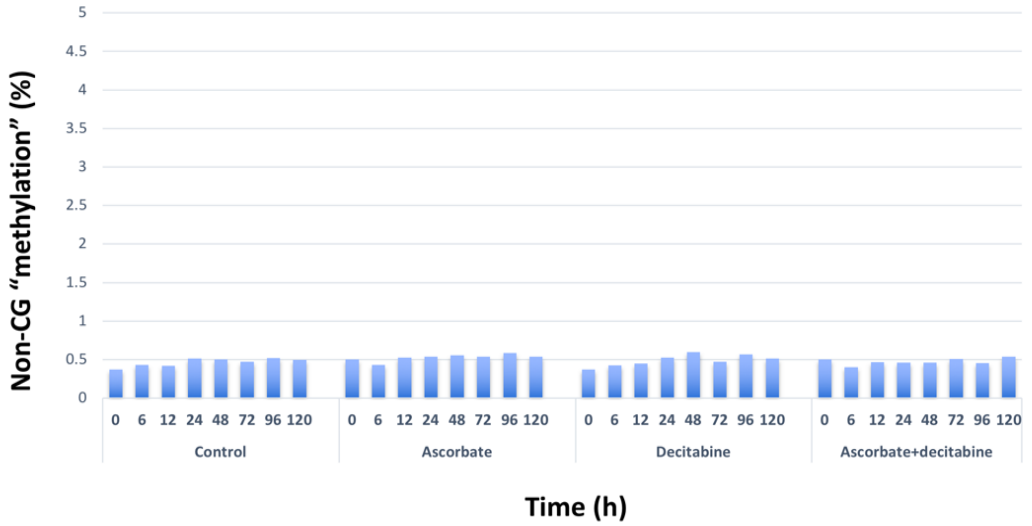


Figure 7. 9. Total non-CG "methylation" levels for Jurkat cells DNA samples.

All samples had a good mapping efficiency ranging from 59.5 – 64 %, (Figure 7.10).

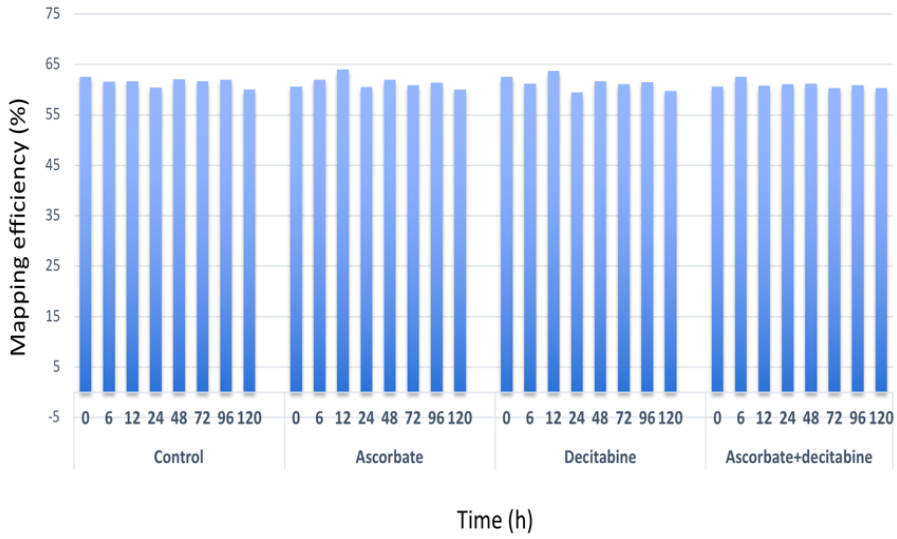


Figure 7. 10. Mapping efficiency (%) for Jurkat DNA samples.

The results of low coverage bisulfite sequencing demonstrated that there is consistent CpG methylation in untreated (control) Jurkat cells across different timepoints (~ 66%). There is a gradual loss in CpG methylation levels in Jurkat cells treated with decitabine (44% at 48 h) and combined ascorbic acid and decitabine (42% at 48 h) in a time dependent manner after a single dose treatment at time zero. There is some recovery in CpG methylation at 72, 96 and 120 h, but methylation remains lower than that in the control (Figure 7.11).

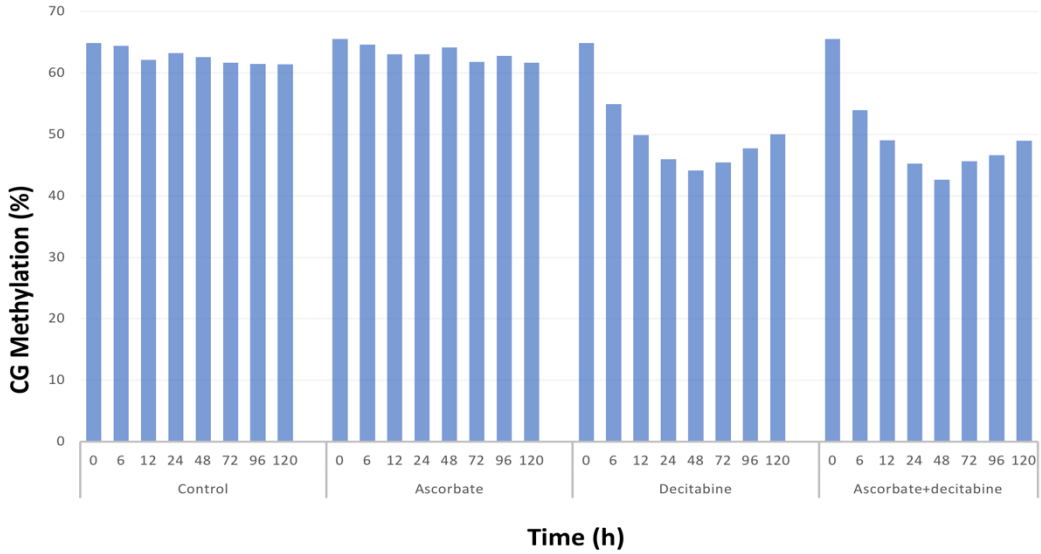


Figure 7. 11. Global CpG methylation levels (%) of Jurkat DNA treated with decitabine, ascorbic acid and combined ascorbic acid and decitabine for 120 h compared to untreated control cells.

7.4 Low-coverage genome wide sequencing PBAT: experiment three.

In this experiment, we extended our global methylation analysis to include three more haematopoietic cell lines Molt4, Nalm6 and HL60 in addition to Jurkat cells. The DNA samples were extracted and prepared for low coverage bisulfite sequencing from treated and untreated samples of the three cell lines (details of the treatment and cultural conditions in Chapter Three). The samples had a good mapping efficiency ranging between 51.7% to 59.7 %, and CG call numbers ranging from 15,000 to 25,000.

The results of low coverage bisulfite sequencing confirmed the results of gene-specific methylation. Ascorbate had no effect on global methylation of the four examined cell lines compared to untreated cells. Jurkat, Molt4 and Nalm6 showed a decrease in global methylation at 12 or/and 24 h ranging between 10% to 30% in decitabine and combined ascorbate/decitabine treatment compared to untreated. Interestingly, decitabine and combined ascorbate/decitabine had no effect on the methylation of HL60, consistent with the results obtained by gene-specific methylation sequencing (Figure 7.12).

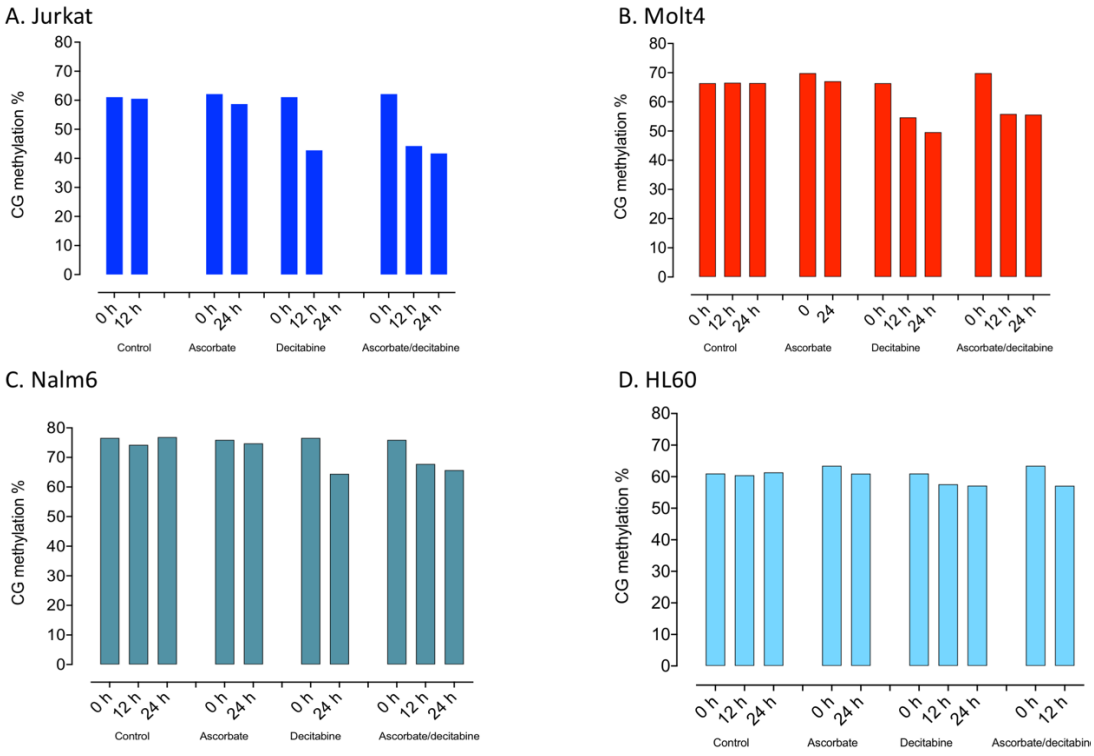


Figure 7. 12. Global CpG methylation levels of A) Jurkat cells, B) Molt4 cells, C) Nalm6 cells and D) HL60 cells treated with decitabine, ascorbic acid and combined ascorbic acid and decitabine for 24 h compared to untreated control cells.

7.5 Determination of genome-wide 5-hydroxymethylcytosine levels.

TET proteins usually catalyse the first step of active demethylation, by oxidising 5-mC into 5-hmC (Figure 7.13). The level of 5-hmC is tissue specific, ranging from undetectable levels in cultured cell lines to 0.6 % of DNA in human brain tissues (Globisch, Munzel, et al., 2010). A global decrease in 5-hmC content has been shown in nearly all cancers and has been proposed as a molecular marker and therapeutic target of cancer (Chen, Shen, et al., 2013; Haffner et al., 2011; Jin et al., 2011; Yang et al., 2013). 5-hmC was similarly depleted in bone marrow samples from patients with haematopoietic malignancies (Ko et al., 2010; Liu, Zhang, et al., 2013).

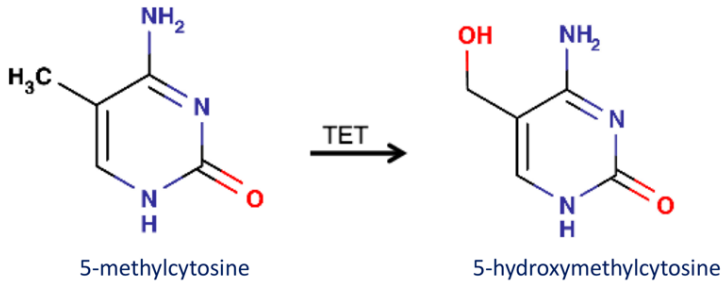


Figure 7. 13. Generation of 5-hydroxymethylcytosine by oxidation of 5-methylcytosine catalysed by the TET1, 2 or 3 enzymes.

Conventional bisulfite treatment does not differentiate between 5-mC and 5-hmC, as 5-mC is unmodified by bisulfite and 5-hmC is modified to cytosine methylene sulfonate (CMS). Due to DNA polymerase recognition, they are both 'read' as cytosine after sequencing (Figure 7.14). In contrast, cytosine, 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC) are 'read' as thymine after sequencing.

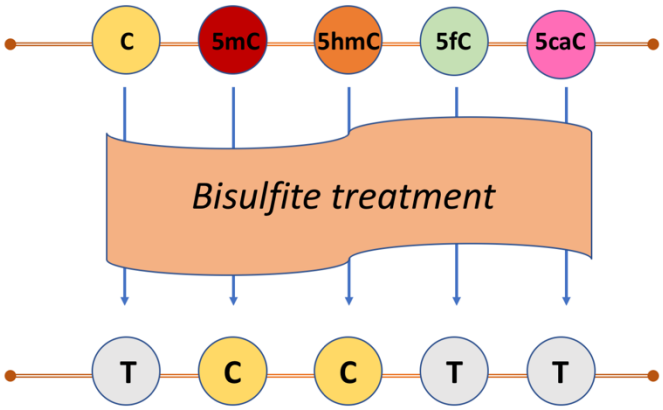


Figure 7. 14. Summary of the effect of bisulfite treatment on modified cytosines. Following bisulfite conversion and sequencing, both 5-mC (red) and 5-hmC (orange) are 'read' as cytosine (yellow), whereas cytosine, 5-fC and 5-caC are 'read' as thymine.

7.5.1 Global hydroxymethylcytosine quantification.

7.5.1.1 Method:

Quantification of global 5-hydroxymethylcytosine was done using the MethylFlash™ Global Hydroxymethylation (5-hmC) ELISA Easy Kit Colorimetric (EPIGENETIK), according to the manufacturer's protocol. This assay has high analytical sensitivity with a detection limit for hydroxymethylated DNA as low as 0.01% from 100 ng input DNA and high specificity to 5-hmC with no cross reactivity to methylated or unmethylated cytosine.

Briefly, the protocol consists of three steps:

Step 1. 100 ng of sample DNA was bound to a high DNA affinity treated wells;

Step 2. the hydroxylated portion of DNA was detected using a 5-hmC monoclonal antibody-based detection complex;

Step 3. 5-hmC was quantified by absorbance in a microplate spectrophotometer (Chapter Three, Section 3.8).

A standard curve was generated by plotting the optical density OD values at 450 nm versus the standards at each percentage point and the slope of the line was determined using linear regression (Figure 7.15).

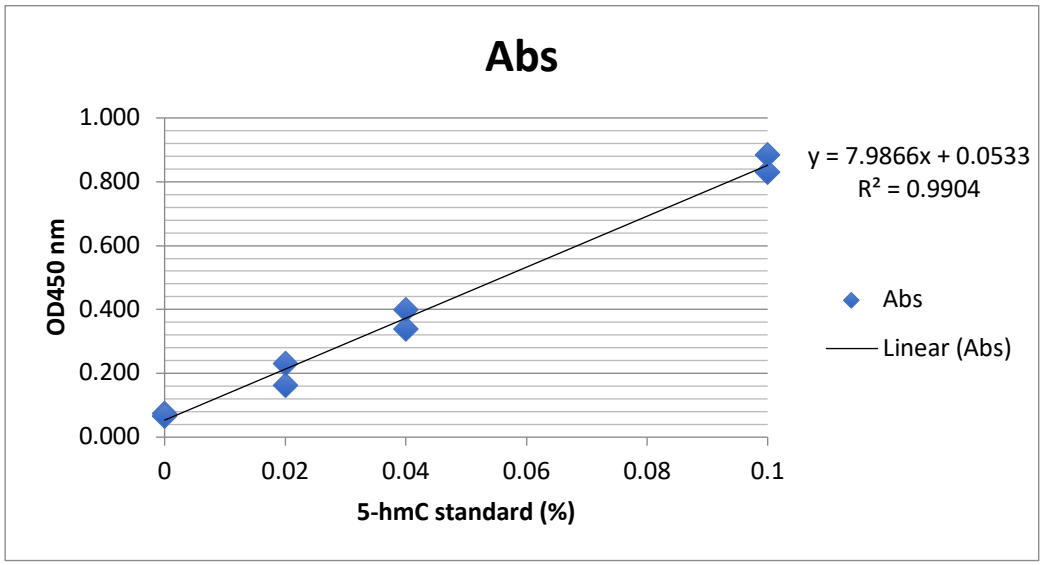


Figure 7. 15. Standard curve generated using included 5-hmC standards. The slope was determined using linear regression. The portion of the standard curve that is relevant to the measured samples is shown.

Ascorbate is a potent activator of the TET proteins, and has been proposed to lead to an increase of genome-wide hydroxymethylation levels in ascorbate treated cells (Blaschke et al., 2013). In this study, we assayed the amount of genome wide DNA hydroxymethylation in Jurkat cells after treatment with ascorbate and decitabine. Jurkat cells treated with combined ascorbate and decitabine showed a gradual increase in 5-hmC levels; 0.03% at 4 h, 0.34% at 8 h followed by reduction to 0.05% at 12 h compared to undetectable level in control cells. Decitabine alone caused negligible 5-hmC at 8h, but surprisingly, caused an increase in 5-hmC levels 0.12% (12 fold) at 12 h. Ascorbate was able to increase 5-hmC levels 0.04% at 0 h (after 18 h incubation with Jurkat cells during synchronisation) compared to control. Notably, greater 5-hmC levels were achieved with ascorbate and decitabine combinatorial treatments than decitabine alone. Interestingly, although decitabine caused only negligible 5-hmC at timepoint 8 h when applied alone, in combination with ascorbate its efficacy was greatly enhanced 37 fold (Figure 7.16).

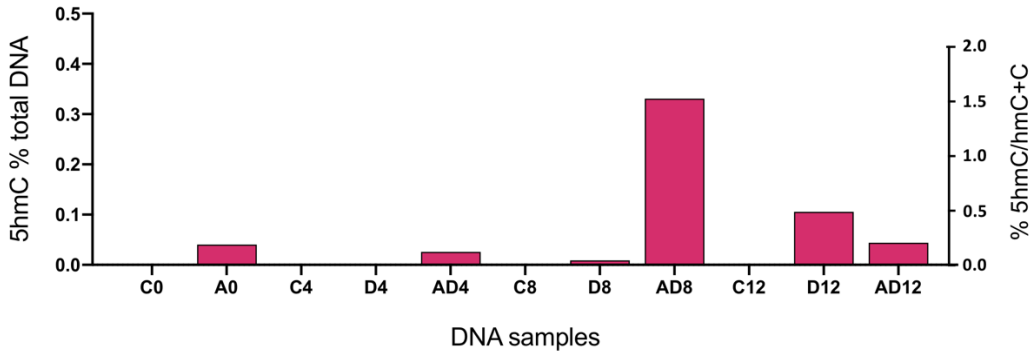


Figure 7. 16. Bar blot showing the percentage of 5-hmC in different samples at multiple timepoints (5-hmC/total DNA (A+G+C+T) (left axis). The right axis shows 5-hmC as a percentage of cytosines. Cytosine content is 21% in human. Hydroxymethylation levels (absorbance) of DNA samples that were less than the negative control were assigned to zero (C: control, A: ascorbate, D: decitabine and AD: ascorbate + decitabine treated cells. Values are means of duplicates.

The amount of 5-hmC as a percentage of cytosine was calculated by dividing the calculated 5-hmC% by the human DNA cytosine content (0.21).

7.6 Discussion.

Epigenetic modification is a well-studied hallmark of cancer development. The well-known cytosine analogue, demethylating drug decitabine, is effective against haematological malignancies such as myelodysplastic syndrome and AML (Pleyer & Greil, 2015). Previous efforts concentrated on the treatment of tumour malignancies used high doses of decitabine in cancer cells. Study by Kantarjian (2007) showed that decitabine was associated with a survival advantage compared with intensive chemotherapy on patients with high risk MDS (Kantarjian et al., 2007). In contrast, the high doses of decitabine, led to cytotoxic induction without modulating the epigenetic pattern (Fenaux et al., 2010; Pohlmann et al., 2002; Schwartsmann et al., 2000; Seymour et al., 2010).

In this study, the gene specific sequencing results obtained from hairpin-bisulfite sequencing (Chapter 5 and 6) were validated using PBAT or low coverage global bisulfite sequencing. From the PBAT results, we were able to determine global DNA methylation levels and to study the kinetics of DNA demethylation in Jurkat cells after treatment with ascorbate and/or decitabine. One limitation of PBAT is that it cannot determine the proportions of DNA hemi-methylation and therefore cannot determine the mechanism for loss of methylation, as hemi-methylated DNA strands are separated into single strands during bisulfite treatment.

Although, the WGBS method is a powerful method to analyse global DNA methylation levels, its cost is prohibitive and requires large amounts of starting DNA, whereas PBAT provides whole-genome bisulfite sequencing for low DNA input (100 ng) and at low cost. We constructed PBAT sequencing libraries for 64 samples from two experiments. The protocol yielded variable CG calls for all samples in the first experiment except for two samples where the CG call was low. In the second experiment, the protocol yielded more than 15,000 CG calls for each sample. In addition non CG "methylation" was less than 1% for both experiments indicating efficient bisulfite conversion 99% across the genome.

There were comparable levels of global DNA methylation for all control samples and timepoints within the two experiments, indicating the precision of the PBAT protocol. We observed loss of global DNA methylation in Jurkat cells, Molt4 cells and Nalm6 cells treated with decitabine or a combination of ascorbate/decitabine at increasing timepoints. The largest proportion of demethylation occurred at 72 h with around 60% loss methylation in Jurkat cells treated with decitabine or combined ascorbate/decitabine compared to untreated controls. The PBAT result corroborated the targeted, hairpin bisulfite sequencing results (Chapter 5), thus confirming the demethylating effect of decitabine.

Our results show that incubation of Jurkat cells with 500 µM glycine chloramine had little effect on global CG methylation levels. There was only a small reduction in global CpG methylation (4%) at 6 h. The reduction in methylation was then confirmed by hairpin DNA sequencing which showed an increase in the percentage of DNA hemi-methylation in Jurkat cells treated with glycine chloramine at 6 h post release from cell cycle arrest (results not shown). Interestingly, a previous study by Dr. Helena Magrath from the Morison laboratory (unpublished data) showed active demethylation in Jurkat cells as early as 2 h post release from cell cycle arrest after treatment with glycine chloramine. Another study by Dr. Karina O'Connor showed that glycine chloramine inhibited DNA methylation by 35% compared with control Jurkat cells. The difference in demethylating activity of glycine chloramine between our experiment and the previous experiments may be because glycine chloramine is a short acting oxidant and inclined to degradation which leads to failure of inhibition of DNMT1.

Therefore, in this research we studied the efficacy of hypomethylating agent decitabine and the concomitant treatment of decitabine/ascorbate to introduce epigenetic modifications in Jurkat cells. We investigated whether ascorbate can enhance the demethylating actions of decitabine. Ascorbate has been shown to act as a potent essential co-factor for the TET enzymes, which mediate the first step of active DNA demethylation by oxidation of 5-mC to generate 5-hmC (Figure 7.15). We found reduced DNA methylation levels as early as 6 h after release from cell cycle arrest in Jurkat cells treated with decitabine alone, or with ascorbate and decitabine.

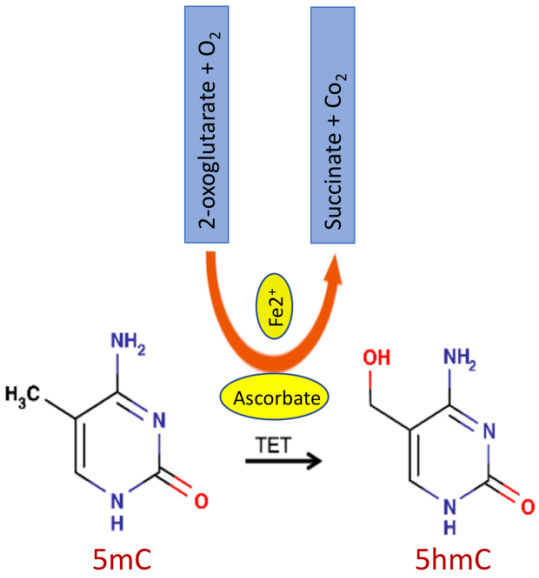


Figure 7. 17. Ascorbate act as co-factor, in addition to $Fe2^+$, for TET enzymes to oxidise 5-mC to 5-hmC.

Although, sodium bisulfite treatment is considered the gold standard for measuring DNA methylation, one primary limitation of this method is its inability to distinguish between 5-mC and 5-hmC and measuring both modified bases as methylated.
To elucidate whether the epigenetic activity of TET enzymes was increased by the combination treatment of ascorbate and decitabine, genome-wide changes in DNA hydroxymethylation states were assessed.

Remarkably, combination treatment of decitabine and ascorbate resulted in an increase in 5-hmC at 4 h and 8 h, suggesting some synergism and additive effect between decitabine and ascorbate. Decitabine alone failed to elicit any changes to 5-hmC levels at 4 h with only negligible effects at 8 h. Interestingly, decitabine alone induced an increase of 5-hmC levels in Jurkat cells after 12 h. Diminished TET activity in the absence of ascorbate suggesting further mechanisms are at play, that ascorbate is not an essential co-factor for TET enzyme activity.
The presence of hemi-methylated DNA strands after decitabine treatment, induced TET enzymes activity, which is consistent with previous study that showed decitabine induced an increase in 5-hmC levels in HL60 and TK6 leukaemia cell lines (Chowdhury et al., 2015b). Furthermore, Gerecke et al (2018) used LC-MS/MS to demonstrate that ascorbate was not involved in the passive DNA hypomethylation mediated by the action of decitabine, but was involved in the global increase of 5-hmC (Gerecke et al., 2018).

Minor et al. (2013) showed that ascorbate enhanced 5-hmC generation in a dose dependent manner in mouse embryonic fibroblasts (MEFs), and that ascorbate could rapidly (< 1 h) induce significant TET dependent generation of 5-hmC, suggesting a direct involvement of ascorbate in the catalytic activity of TET proteins (Minor et al., 2013). A similar study on MEFs, reported that ascorbate induces TET mediated production of 5-hmC, and the effect was independent of iron or 2-oxoglutarate levels (Dickson et al., 2013). Treatment with vitamin C at physiological levels increased 5-hmC in five cancer cell lines, with the biggest yield seen in MCF7 cells. Interestingly, combination treatment of ascorbate and decitabine further increased 5-hmC levels by about twofold compared with ascorbate alone in four of the cell lines, suggesting the involvement of TET enzymes in enhancing the production of 5-hmC (Liu et al., 2016). A study by Sajadian et al. showed that 5-azacytidine is able to trigger TET-dependent active demethylation in hepatocellular carcinoma HCC with concomitant significant changes in 5-hmC/mC (Sajadian et al., 2015a).

Our study clearly showed that the addition of ascorbate to decitabine treatment increased the efficacy of decitabine in production of 5-hmC at 8 h presumably by upregulating TET enzymes.

Thus, our data suggested a mechanism between ascorbate and the hemi-methylation status by decitabine: firstly, that ascorbate up-regulated TET proteins; and secondly, the hemi-methylation status enhanced TET proteins to hydroxylate 5-mC to 5-hmC at 8 h, compared to negligible 5-hmC levels in decitabine treatment alone.

The reduction in global DNA hydroxymethylation at 12 h in Jurkat cells treated with combinatorial treatments of decitabine and ascorbate suggests that ascorbate enhanced TET enzyme to induce further oxidation of 5-hmC to 5-fC and 5-caC. It is known that cancer patients are vitamin C deficient and their tumour tissues are depleted of 5-hmC compared to corresponding normal tissue, suggesting the importance of ascorbate in cancer treatment (Sant et al., 2018).

Chapter Eight: Discussion and future directions.

Previously, the Morison laboratory had observed DNA demethylation after exposure to glycine chloramine or decitabine in dividing Jurkat cells. Using a low throughput hairpin bisulfite sequencing assay, we had demonstrated hemi-methylation of densely methylated promoter regions after DNA synthesis in cells exposed to glycine chloramine or decitabine. Although this hemi-methylation was expected for decitabine, the mechanism by which glycine chloramine, an oxidative agent, is able to induce demethylation is unknown. Furthermore, using this low throughput assay, we surprisingly detected a small proportion of sequences which were fully demethylated. At the time of these results, demethylation mechanisms were largely unknown; however contemporaneous results from early embryo development studies suggested that the TET proteins were involved in active demethylation processes.

To validate these initial results, I modified the hairpin bisulfite sequencing assay to utilise high-throughput, next generation sequencing protocols, thus enabling the generation of thousands of reads per experiment and permitting documentation of methylation for both complementary DNA strands. Furthermore, we developed a bioinformatics workflow to facilitate rapid analysis of our sequencing results. An additional benefit of our assay is that examination of the random nucleotide hairpin sequence reveals any over-amplification of reads and our bioinformatic workflow will remove these non-unique (over-amplified) sequences, i.e. PCR bias.

This study is unique in both aims and methodology, with the hairpin bisulfite sequencing assay revealing complementary strand methylation states. This research is the first to have investigated the complementary effects of ascorbate and decitabine on DNA methylation in somatic cells.

Here we summarise the outcome of this study:

Firstly, we developed a unique hairpin sequencing model system to observe methylation changes in complementary DNA strands; this model can generate thousands of unique barcoded reads.

Secondly, we established and developed a bioinformatic workflow to analyse hairpin bisulfite sequencing data, initially using the online Galaxy platform, but later with more efficient UNIX command line tools in combination with the R statistical platform.

Thirdly, we documented the kinetics of DNMT inhibitor treatment, which has never before been observed in such detail and clarity.

Fourthly, we documented the presence of hemi-methylation as early as 2 h after release from cell cycle arrest.

Fifthly, we documented active DNA demethylation in somatic cells treated with ascorbate and decitabine.

Sixthly, this study is the first to document the generation of 5-hmC in Jurkat cells after combined ascorbate/decitabine treatment, suggesting that the availability of ascorbate may enhance the efficacy of decitabine treatment in cancer.

And finally, we demonstrated that a novel effect of decitabine producing hemi-methylation, is to induce TET-mediated active demethylation, as shown by increasing global 5-hmC levels.

DNA methylation is believed to be stable and necessary to maintain differentiation, regulation of gene expression and cellular identity of fully differentiated, somatic cells. Nevertheless, dysregulation of DNA methylation has been documented in many cancers and likely contributes to tumorigenesis (Lio et al., 2019). The loss or gain of methylation is a prognostic marker in various types of cancer (Bochtler et al., 2017; Rasmussen & Helin, 2016). Recently, changing the balance of DNA methylation or demethylation of aberrantly methylated tumour suppressor gene promoters has become of interest to many cancer researchers.

The phenomena of DNA methylation (de-novo and maintenance methylation) are well characterised in mammals; however, the mechanism for reversal of methylation or demethylation was less clear. Many studies have started to unravel the mechanism of DNA demethylation by focusing on the methylation changes that occur during early embryonic development. These studies have revealed two mechanisms that result in the loss of methylation. The first mechanism is termed replication-dependent or passive demethylation and relies on failure to reproduce CpG methylation and therefore, dilution of 5-mC during successive rounds of cell division, and the second mechanism is replication-independent or active demethylation, which requires enzymatic activity to modify and replace the 5-mC with unmodified cytosine (Bochtler et al., 2017). Despite many reports having documented active

DNA demethylation in early development, the extent to which it occurs in somatic cells remains unclear (Ho et al., 2017; Ooi & Bestor, 2008; Rose et al., 2014).

In our laboratory, we have focused on mechanisms of DNA demethylation. We have used a model system Jurkat cells (a T cell leukaemia cell line) that expresses all three TET enzymes (Kalender Atak et al., 2012; Ko et al., 2011).

Using this *in vitro* model, we observed a rapid DNA demethylation, as early as 2 h after release from thymidine block, in Jurkat cells exposed to oxidative stress. This observation confirmed unpublished data by Karina O'Connor (2009) indicating that oxidative stress induced by glycine chloramine treatment, reduced DNA methylation in the first 6 h of DNA replication without affecting the cell growth. This early loss of methylation is most likely due to active demethylation, and most probably occurred by the activity of the existing TET enzymes.

To investigate this process of active demethylation and to test whether TET enzymes were involved, a high throughput barcoded hairpin bisulfite sequencing technique was developed to study the methylation pattern of complementary DNA strands. This advanced version of our original hairpin bisulfite sequencing technique provides a direct assessment of the methylation status of complementary strands, by covalently linking complementary DNA strands together with a DNA hairpin linker and generating thousands of unique barcoded sequences for each amplicon. From these sequencing data, we were able to document the process of DNA demethylation throughout the cell cycle and beyond.

8.1 Decitabine enhances hemi-methylation and active demethylation.

Decitabine is a well-known hypomethylating and anti-proliferative drug frequently used in the treatment of leukaemia (Badar et al., 2015). The mechanism of decitabine action as a specific inhibitor of DNMT1 was proposed many years ago, but not fully explained in detail (Raj et al., 2007; Santi et al., 1984). Here, with the power of our hairpin bisulfite sequencing assay, we were able to describe the kinetics and pattern of DNMT1 inhibitor treatment. Our results show that decitabine, similar to glycine chloramine, was able to enhance DNA hemi-methylation in *PCDHGA12* as early as 2 h after treatment (N.B. we define CpG site hemi-methylation as the lack of a methyl group on one strand, and retention of methylation on the other strand. Additionally, hemi-methylation levels increased as DNA replication proceeded. Even though these results were predicted by previous studies, the kinetics of decitabine-induced hemi-methylation has not been demonstrated before in such detail and clarity.

In addition, a delay in cell cycle progression was observed at 24 h in decitabine-treated cells, represented by increased number of daughter cells in S phase. This delay might be due to ability of decitabine to stall DNA polymerase, and allowing time for DNA repair processes to remove the decitabine from DNA (Ubhi & Brown, 2019). Decitabine stalled DNA replications forks are transient, the cells are able to resolve them and continue replication via DNA repair response. For instance, previous reports showed that the deoxycytidine analogues cytarabine and gemcitabine can compete directly with dCTP for incorporation into newly synthesised DNA, leading to delay in replication fork progression or even termination of some replication forks (Huang et al., 1991; Richardson et al., 2004).

To confirm the loss of methylation observed in gene-specific sequencing analysis, we performed a low coverage methylation sequencing assay (PBAT) to detect changes in global DNA methylation. We observed a gradual loss in global DNA methylation in synchronised Jurkat cells treated with decitabine, confirming the results from locus-specific studies. This finding is consistent with previous reports that decitabine enhanced loss of global methylation of colorectal cancer cells and leukaemia cells (Mossman et al., 2010; Öz et al., 2014; Yang et al., 2004).

8.2 Ascorbate induces more demethylation in decitabine treated cells.

Ascorbate appears to act as a co-factor for the TET enzymes by reducing the inactive Fe^{3+} to the active Fe^{2+}, thus playing an important role in active DNA demethylation (Hore et al., 2016). Intriguingly, the normal ascorbate concentration in human plasma is ~50 µM, whereas standard cell culture media is ascorbate deficient, suggesting that the addition of ascorbate to treatment protocols may increase the clinical efficacy of such drugs in patients with leukaemia.

We have observed an increase (40%) in unmethylated hairpin reads of the *PCDHGA12* promoter after treatment with ascorbate/decitabine compared to untreated controls. To the best of our knowledge, this is the first report of ascorbate enhancing the efficacy of decitabine to induce active DNA demethylation in somatic cells. A previous study has shown that treating mouse embryonic fibroblasts (MEF) with 10 µM ascorbate enhanced a rapid and significant generation of 5-hmC after as little as 1 h. These observed results were not unique to MEFs as the same results were obtained in different cell line HEK-293T and HeLa cells treated with ascorbate (Minor et al., 2013). Additionally, our results are consistent with another previous report demonstrating that ascorbate can directly enhance the catalytic activity of TET dioxygenases for the oxidation of 5-mC (Yin et al., 2013). Ascorbate is a cofactor for a large family of iron and α-ketoglutarate dependent dioxygenases (Young et al., 2015). Probably,

ascorbate plays a role in reducing the inactive Fe^{3+} to the active Fe^{2+} and maintaining enzyme cycling (Hore et al., 2016; Ito et al., 2010; Markolovic et al., 2015; Yin et al., 2013). Interestingly, contrasting results were published by Dickson (2013), who showed that ascorbate induced demethylation by generation of 5-hmC independent of the levels of iron (Dickson et al., 2013), whereas, Tahiliani (2009) reported that hydroxylation of 5-mC mediated by TET enzymes was not affected by ascorbate, and that TET proteins can oxidise 5-mC at the same rate and efficiency in the presence or absence of ascorbate (Tahiliani et al., 2009).

It was reported that glutathione, another antioxidant reagent, was unable to affect the generation of 5-hmC in MEFs (Minor et al., 2013), thus suggesting that the effect of ascorbate on both DNA demethylation and 5-hmC levels is specifically TET-mediated, and not through its role as a general antioxidant. It is known that TET activity is intimately linked to the TCA 'tricarboxylic acid cycle' metabolite products, and dysregulation of the TCA enzymes is relevant to TET activity in cancers. For example, mutations of fumarate hydratase or succinate dehydrogenase can result in accumulation of fumarate and succinate, respectively, and impede the activity of a variety of α-ketoglutarate dependent enzymes, including TET (Xiao et al., 2012).

Treatment of cells with decitabine alone or in combination with ascorbate is only transient, as both reagents are unstable in cell culture conditions (Chepda et al., 2001; Rogstad et al., 2009). The observed re-methylation of the *PCDHGA12* hairpin molecules at 96 and 120 h after decitabine and combined ascorbate/decitabine treatments suggests re-methylation by the de novo methyltransferases DNMT3A and DNMT3B (Uysal et al., 2017), or that outgrowth of partially demethylated cells has occurred.

The clinical relevance of our study is that it supports the premise that ascorbate is necessary to enhance the efficacy of decitabine by promoting the function of TET. Some patients with haematological cancers are treated with demethylating drugs such as decitabine. Given that cancer patients are often markedly ascorbate (vitamin C) deficient, the addition of ascorbate to treatment protocols may increase the clinical efficacy (or toxicity) of such drugs in such patients.

8.3 Genome-wide 5-hmC content increased following decitabine and combined ascorbate/decitabine treatments.

While, global 5-mC constitutes 4-5% of cytosine in normal adult tissue, 5-hmC concentration usually varies between 0.03% to 0.7% and is much higher in the brain (Globisch, Münzel, et

al., 2010). However, 5-hmC levels are dysregulated in cancer, and were depleted in carcinomas from prostate, colon, liver, brain, breast, skin, lung, melanoma and in many haematopoietic malignancies (Rodger et al., 2014). In general, the distribution of 5-hmC within the genome is associated with gene transcription, and depletion is associated with poorer prognosis, suggesting that 5-hmC has a role in suppressing tumour progression (Guo et al., 2011b; Wu, D'Alessio, Ito, Wang, et al., 2011; Wu & Zhang, 2011b; Xu, Wu, et al., 2011). However, 5-hmC function was dependent on genome location; firstly, location in gene bodies was correlated positively with gene expression and secondly, location at gene promoters was associated gene repression (Wu, D'Alessio, Ito, Wang, et al., 2011).

We observed a rapid active demethylation in the densely methylated *PCDHGA12* and *RASSF1* promoters in synchronised Jurkat cells after incubation with decitabine and combined ascorbate/decitabine, and we predicted that this was mediated by TET activity. Because TET proteins mediate the first oxidation step of 5-mC to 5-hmC, we aimed to confirm the involvement of TETs in our observed locus-specific active demethylation. Conventional bisulfite conversion treatment is unable to differentiate between 5-mC and 5-hmC as both of them are read as cytosine after sequencing.

Using an enzyme immunoassay, we measured the global DNA content of 5-hmC in treated and untreated synchronised Jurkat cells at different timepoints during DNA replication. Simply, 5-hmC levels were undetectable in untreated Jurkat cells during all timepoints, consistent with previous observations of 5-hmC deficiency in cancer cells. This deficiency may be due to high intrinsic TET activity, which facilitates rapid oxidation of 5-hmC to 5-fC and 5-caC. This speculation is supported by several reports showing that the highest levels of 5-hmC were present in the least proliferative tissues, such as the brain (Globisch, Munzel, et al., 2010; Guo et al., 2011a; Kriaucionis & Heintz, 2009; Szwagierczak et al., 2010).

Decitabine-treated cells exhibited very low levels of 5-hmC at 4 and 8 h respectively, but surprisingly, a marked increase at 12 h after decitabine treatment. This increase suggested that the decitabine-induced hemi-methylation enhanced TET-induced oxidation of 5-mC to 5-hmC and facilitated active demethylation. This result is consistent with others showing that the treatment of haematopoietic cell lines with decitabine decreased 5-mC content, but increased in 5-hmC levels (Vető et al., 2018). Additionally, a study by Chowdhury and colleagues demonstrated that decitabine not only reduced the global 5-mC levels but also promoted an anomalous increase in 5-hmC, 5-fC and 5-caC in acute myeloid leukaemia (Chowdhury et al., 2015a). Furthermore, Sajadian and colleagues reported that decitabine treatment of

hepatocellular carcinoma cell line (HCC) led to a significant increase in 5-hmC positive cells after 24 and 48 h, accompanied by a significant increase in TET2 and TET3 expression (Sajadian et al., 2015b).

Intriguingly, we found an increase in global 5-hmC levels in cells treated with ascorbate at 0 h (the time of release from cell cycle arrest, but after 18 h incubation with ascorbate during synchronisation) and with combined ascorbate/decitabine treated cells at 4 h after release from cell cycle arrest. Surprisingly, the global 5-hmC level was increased markedly at 8 h in cells treated with the combined ascorbate/decitabine. This rapid increase of 5-hmC strongly supported the involvement of TET enzymes in this hydroxymethylation process, and most likely that ascorbate promotes upregulation of TET activity and accelerates the generation of 5-hmC. This finding was consistent with several reports showing an increase in the 5-hmC levels in cells incubated with ascorbate or combined ascorbate/decitabine. A study by Nair (2016) showed that ascorbate efficiently accelerated the activity of TET2 to induce iterative oxidation reactions by converting 5-mC into unmodified cytosine (Nair et al., 2016). Moreover, Chowdhury and colleagues also showed that TET enzymes have partial selectivity for hemi-methylated CpG dinucleotides which could led to such changes in 5-hmC content (Chowdhury et al., 2015a).

Importantly, we demonstrated that decitabine enhanced active demethylation, defined by the full removal of methyl groups from both complementary DNA strands. ***This study is the first to report active demethylation in somatic cells treated with decitabine.*** Our results demonstrate that decitabine is capable of erasing the epigenetic memory of somatic cells by affecting TET activity to induce hemi-methylation and active DNA demethylation.

Although we never observed demethylation in untreated, control cells, we did observe some variability between experiments which might be attributed to differences in culture conditions (media composition, O_2 levels, pH, etc.), PCR bias, sequencing artefacts or other unidentified factors combined that might have direct effect on DNA demethylation, e.g. BER, UHRF2 or / and APOBEC activities. To explore these potential variations, we conducted several replicate experiments to exclude most of the possible factors that might cause this variation. For example, PCR bias could be excluded, as we are using the hairpin barcode within each individual read, to remove all non-unique and repeated reads (reads sharing the same hairpin barcode). As we had changed the media composition from RPMI1640 to RPMI1640+glutamax and from RPMI1640 to DMEM due to changes during experimental design, we compared RPMI1640

versus DMEM media (the media provides iron), or RPMI1640 versus RPMI1640+glutamax (the media provides L-alanine and L-glutamine, and minimise toxic ammonia build up and improve cell viability and growth).

Interestingly, we observed ~50% full unmethylated reads (active demethylation) of *PCDHGA12* promoter at 4 h in Jurkat cells grown in DMEM media and treated with ascorbate and combined ascorbate/decitabine. This demethylation might be attributed to the iron content of DMEM media, and with the addition of ascorbate accelerating the recycling of Fe^{3+} into the active form Fe^{2+}, acting as a co-factor for TET activity. No differences in the DNA methylation results between Jurkat cells grown in RPMI1640 vs RPMI1640+glutamax.

To exclude 'Jurkat cell specific activity', we repeated the experiments in other haematopoietic cell lines. In three out of four studied cell lines (Jurkat, Molt4 and Nalm6), a marked gene-specific as well as genome-wide demethylation was observed. Methylation was reduced by more than 50% in Jurkat cells treated with decitabine and with combined ascorbate/decitabine at 72 h, and was in the form of hemi- and fully unmethylated hairpin reads. Methylation was reduced in Molt4 and Nalm6 cell lines treated for 24 h with decitabine or combined ascorbate/decitabine, the loss of methylation is represented by gradual increase in hemi-methylated sequences (~ 50% in Molt4 at 16 h, and 30% in Nalm6 at 12h) and few unmethylated sequences with increasing time. However, gene-specific methylation analysis for HL60 cells, showed little demethylation (~ 5-10% after 0.5 or 5 μM decitabine treatments). The reason for this resistance to demethylation of HL60 cells is unknown, but may be due to their sensitivity and observed cytotoxicity with decitabine treatment, and might be due to differences in cell surface transporters or receptors between HL60 and Jurkat, Molt4 (T-cell leukaemia) or Nalm6 (B-cells leukaemia). The combined use of decitabine and ascorbate may be an advantageous in enhancing the demethylation effect in decitabine resistant cells.

It is widely recognised that proliferative cells such as cancerous cells can metabolise glutamine as source of energy (Daye & Wellen, 2012). Glutamine metabolism goes through two deamination steps, which are sequentially catalysed by two enzymes glutaminase and glutamine dehydrogenase and result in the generation of α-ketoglutarate (Corbet & Feron, 2017; Daye & Wellen, 2012). An increased level of α-ketoglutarate provides TET enzymes with a sufficient quantity to be upregulated and initiate oxidation of 5-mC to its oxidative derivatives, suggesting the addition of α-ketoglutarate to the culture media may increase TET activity during ascorbate and decitabine treatments.

However, our observations indicate that the DNMT inhibitor decitabine might not be the only substance that affects DNA methylation and induces rapid active demethylation. Perhaps activation of existing enzymes may occur during this process which led to this demethylation phenomenon.

DNA methyltransferase enzymes DNMT1, DNMT3a and DNMT3b are the key controller of DNA methylation (Uysal et al., 2017), and these DNMTs are controlled by many factors such as folate level and the redox state of enzymes as well as the ratio of the methyl donor SAM to SAH. Changing in these regulatory factors results in up or down regulation of DNMTs or dysregulation of its function. An in vitro study by Chen (2013) reported that oxidative stress can switch the function of human and mouse DNMTs from DNA methylase activity to demethylase activity in the presence of calcium ions and in the absence of reducing agents (Chen, Wang, et al., 2013). Further investigation of factors that affect demethylation is clearly needed.

8.4. Clinical relevance.

Recent evidence suggests that there is a remarkable correlation between decreased TET expression and decreased 5-hmC content, resulted in progression of tumour, suggesting the possibility that the TETs act as tumour suppressor proteins in some cancer types (Haffner et al., 2011; Xu, Wu, et al., 2011). The TET family of proteins were identified due to TET being a frequent translocation target in leukaemia. Another study reported that TET2 gene but not TET1 or TET3 was frequently mutated in leukaemia, and acted as a relevant tumour suppressor gene (Delhommeau et al., 2009). TET1 mutations have been reported in T-cell ALL (Kalender Atak et al., 2012), and also decreased expression of TET1 has been observed in prostate and breast cancer (Hsu et al., 2012). Muller and his colleagues found that TET1 is frequently either not expressed or localised outside the nucleus, which suggested that the exclusion of TET1 from the nucleus resulted in depletion of 5-hmC in tumours, and also indicated that TET1 is required to maintain 5-hmC marks (Muller et al., 2012).

Decitabine is used in the treatment of acute myeloid leukaemia. He and colleague showed that decitabine is an effective therapeutic alternative drug with acceptable side effects in early AML patients (He et al., 2017). A recent study compared the efficacy of a 5-day and 10-day schedule treatments in old patients with newly diagnosed AML, the result suggested that 10-day schedule of decitabine treatment led to better outcomes than 5-day schedule (Short et al., 2019). As in vitro study showed that the combination of low dose ascorbate and decitabine has synergistic

effect on proliferation, apoptosis and TET expression in HL60 and NB4 human leukaemia cells compared to decitabine alone (Zhao et al., 2018).

8.5. Conclusions.

The results of this study provide evidence that supports our hypotheses. We observed active demethylation in somatic cells after treatment with decitabine. Ascorbate increased demethylation. Demethylation was accompanied by increase in 5-hmC and presumably involved TET proteins.

8.6. Future work.

We observed variability of the results which might be attributed to differences in culture conditions (media composition, O_2 levels, pH, etc.), PCR bias, sequencing artefacts or other. This is a mystery and further work needs to be undertaken to identify the missing variables. Potentially, they could be:

1. Oxygen is an essential factor for TET enzymes, and oxygen shortage (hypoxia) play an important role in tumour biology. For example, about one third of tumour area has an oxygen level below 0.5% (Thienpont, Steinbacher, et al., 2016). However, purified TET enzymes require more than 2% oxygen to function normally, and hypoxic environment (0.5% O_2) reduced the activity of TET to half (Thienpont, Steinbacher, et al., 2016). In contrast, Lin and colleagues (2017) demonstrated that hypoxia upregulates the expression of TET enzymes and catalyse the conversion of 5-mC to 5-hmC in hepatoblastoma (HepG2) cells (Lin et al., 2017). Oxygen concentration might have a direct / or indirect effect on the variation in our results, and use of hypoxia chamber is an option for future experiments to maintain the O_2 concentration.
2. Visualise changes in TET expression after different treatments using different techniques like qPCR and Western blotting. Or visualise changes in activity by using an in vitro TET activity assay. This should be investigated by measuring TET activity after decitabine and ascorbate treatments at different timepoints including early timepoints such as 15, 30, 45, 60 and 90 minutes.
3. Direct measurement of TET-mediate methylcytosine oxidation by Knocking down TET expression using siRNA.
4. Inhibition of TET activity by using iron chelators to limit supply of Fe2+ for TET activity.

5. Extend the study to design hairpin molecules to include more genes and more cell lines with extended timepoints. Some genes are more sensitive to decitabine, and cell lines may show variation in their response to decitabine.
6. Gene-specific 5-hydroxymethycytosine assay. An assay to directly and efficiently monitor and measure TET activity is lacking. Chemical modification of cytosine bases or enzymatic detection have been described and could be used with modification of the hairpin assay (Giehr et al., 2018).
7. Study the effects/kinetics of ascorbate treatment during cell growth. This could be investigated by studying the redox state of ascorbate and measuring the degradation rate of the oxidised form of ascorbate. Furthermore, more research is needed to determine the optimal timepoints at which the ascorbate should be added to the media.

Appendices

Appendix 1. Terminal analysis script.

These scripts were established and developed by Morison laboratory members Rob Weeks, Luke Bridgman, Suzan Almomani and myself. The scripts were used to process the raw files outputs of MiSeq sequencing. Each script is annotated with comments.

```
for filenameF in *_R1_001.fastq.gz;
   do
      mkdir "${filenameF%_R1_001.fastq.gz}";
      mv -f "$filenameF" "${filenameF%_R1_001.fastq.gz}";
   done;
# sets up 'for…loop' for all files ending with *_R1_001.fastq.gz, makes a directory
with the file name, and moves the files into their directory.

for filenameF in *_R2_001.fastq.gz;
   do
      mv -f "$filenameF" "${filenameF%_R2_001.fastq.gz}";
      cd ./"${filenameF%_R2_001.fastq.gz}";
# sets up 'for…loop' for files ending with *_R2_001.fastq.gz and moves them to the
previously created directory

      for filenameF in *_R1_001.fastq.gz;
         do
            for filenameR in *_R2_001.fastq.gz;
               do
                  pear -f $filenameF -r $filenameR -o ${filenameF%_R1_001.fastq.gz};
# Do pear, joins paired-end sequence files, and outputs to the current directory.

                     for filename in *.assembled.fastq;
                       do
                           mkdir ./${filename/%.assembled.fastq};
# Downloads the assembled files, makes new directory inside the present working
directory, named for the assembled files.

                           cat < $filename |
                           fastx_barcode_splitter.pl --bcfile
                           ~/Downloads/FPrimers1.txt --bol --mismatches 2 \  --
                           prefix
                           ./${filename%.assembled.fastq}/
                           ${filename%.assembled.fastq}_ --suffix .fastq;
# splits against the forward first round primer sequence (only F primer needed as
the two files have already been joined using Pear.

                           mkdir ./Filtered_by_quality;
                           mkdir ./FastQC;
                           fastqc ./${filename%.assembled.fastq}/*.fastq --
                           outdir=./FastQC;
                     done;
                  cd ../;
               done;
         done;
done &&
# Make directory with name 'filter by quality. And make another directory with name
'FASTQC' run FASTQC against the barcoded-split.fastq files and output to the
directory FASTQC.

for d in ./*/*001/ ; do (cd "$d" && for filename in *.fastq; do fastq_quality_filter
-Q33 -q 30 -p 80 -i $filename -o ${filename/%.fastq/f.fastq} -v ; done ; cd ../ ; mv
./*/*f.fastq ./Filtered_by_quality; cd ./Filtered_by_quality;

# filter the barcoded-split files by quality, for phred score >30 for 80% of read
length, for Illumina 1.3+.fastq files (-Q33).
```

```
   for filename2 in *.fastq;
        do awk '{if(NR%4==2) print;}' < $filename2 |

# for the files ending in '.fastq' print the second line of every four from each
line.


  awk

     '{sub(/GTGATGT|.TGATGT|G.GATGT|GT.ATGT|GTG.TGT|GTGA.GT|
           GTGAT.T|GTGATG./,".")}; 1'

# with 1 mismatch allow collapsing for the liker sequence immediately before the
14 bp molecular barcode.


| sort -t . -k 2.1,2.14 -u | cat -n | sed 's/^/>/' | tr "[\t]" "\n" >
${filename2/%.fastq/.fasta}; rm ./*unmatchedf.fasta; done);

     done && for d in ./*/Filtered_by_quality/ ;
         do (cd "$d" && for filename in *PCDHGf.fasta;
              do mkdir ./${filename/%f.fasta};
                    java -jar ~/Downloads/BiQ_Analyzer.jar -nogui
  -rseq ~/Downloads/PCDHG12.fasta -bseq $filename -outdir

# make new directory with file name PCDHGA12, and run BiQ Analyser java script (-
nogui).


   ./${filename/%f.fasta}
        -sortcgmeth -desc -minas 1000; done); done &&
               for d in ./*/*/*/ ; do (cd "$d" && cut -f 4 < results.tsv |
                         sed -e 's/\(.\)/\1 /g' |
                                sed 1d > methylation.tsv);
                                                done
# sort the files in the output directory, minimum alignment score (1000).

# For RASSF1 we used the same script, and changed the linker sequence
before the molecular barcode

  awk
        '{sub(/TGTGATGT|T.TGATGT|TG.GATGT|TGT.ATGT|TGTG.TGT|
                    TGTGA.GT|TGTGAT.T|TGTGATG./,".")}; 1'



# Make new directory called "filename RASSF1" and run BiQ Analyser java
in the command line (-nogui) script with reference RASSF1.fast.
```

```
# To extract the methylation data from the result files. Change directory
into  "methylation.tsv"


for d in */Filtered_by_quality/; do cd $d && for d2 in */; do cd $d2
      && mv "methylation.tsv" "${d2%/}_methylation.tsv"; cd ..; done; cd ..; cd ..; done
            && mkdir Methylation &&
                mv */Filtered_by_quality/*/*_methylation.tsv Methylation/

# Then prefix any file called "methylation.tsv" with the name of the
directory that file is in.



# To extract the heatmap from the output files. Change directory into
"heatmap.png"

for d in */Filtered_by_quality/; do cd $d && for d2 in */; do cd $d2
       && mv "heatmap.png" "${d2%/}_heatmap.png"; cd ..; done; cd ..; cd ..; done
             && mkdir Heatmap &&
                  mv */Filtered_by_quality/*/*_heatmap.png Heatmap/

# Then prefix the file called "heatmap.png" with the name of the directory
that file is in.
```

Appendix 2. R methylation script.

The script was used to calculate the proportion of strands that are methylated, hemi-methylated and unmethylated for each sample file.

```
# Creates a list of files ending with "methylation.tsv"

setwd("~/Desktop/Methylation")
filenames = list.files(pattern = '*_methylation.tsv', full.names = F)

# The first line of the R script creates a list of filenames in the directory Methylation


Meth_to_table <- function(filenames){
 # The function Meth_to_table is created, which is then applied to the list "filenames"
 results <- data.frame()
 # An empty dataframe called "results" is created
 for(i in 1:length(filenames)){
  # A for loop is invoked - this will run for every file listed in "filenames".
  dat <- read.table(filenames[i], quote="\"", comment.char="", header = FALSE)
  dat <- data.frame(lapply(dat, as.character), stringsAsFactors=FALSE)
  dat <- data.frame(lapply(dat, as.numeric), stringsAsFactors=FALSE)

  # The data from each "*_methylation.tsv" file is read into R and called "dat", and all characters changed to character form, then back to numeric - this changes all uncalled "X" values from BiQAnalyser to N/As - these can then be ignored by RStudio.

  dat_T <- data.frame(Forward = ifelse(rowSums(dat[,1:(ncol(dat)/2)], na.rm=T) > (0.25*ncol(dat)), "1", "0"))

  # Next, a new dataframe is created called dat_T - this has one column called Forward. It reads the columns from 1 to the number of columns in "dat" divided by two. For these columns it then sums the values. If the total of these columns is greater than 1/4 of the total number of columns in the dataframe "dat", then a "1" is recorded in the column "Forward". If not, then a "0" is recorded. In this way, the "parent" strand is called as being either methylated (1) or unmethylated (0).

  dat_T <- transform(dat_T, Reverse = ifelse(rowSums(dat[,(ncol(dat)/2+1):ncol(dat)], na.rm=T) > (0.25*ncol(dat)), "1", "0"))
  # Then, a new column called "Reverse" is added to the dataframe "dat_T". A function to call the "daughter" strand as being methylated (1) or unmethylated (0) is invoked, as in the previous line - except this runs for the last half of the dataframe "dat", not the first.

  dat_T <- data.frame(lapply(dat_T, as.character), stringsAsFactors=FALSE)
  dat_T <- data.frame(lapply(dat_T, as.numeric), stringsAsFactors=FALSE)
  # Then all values are converted to numeric again to force N/As.

  dat_Strand <- data.frame(Strand = rowSums(dat_T[,1:2]))
  # Next, a new dataframe called dat_Strand is created. This has only one column, called "Strand". This column contains the sum of each row from the previous table "dat_T". This gives three values: 0, 1, and 2.

  dat_Meth <- data.frame(Unmethylated = (sum(dat_Strand$Strand == "0"))/(sum(dat_Strand$Strand == "0", dat_Strand$Strand == "1", dat_Strand$Strand =="2")), Hemimethylated = (sum(dat_Strand$Strand == "1"))/(sum(dat_Strand$Strand == "0", dat_Strand$Strand == "1", dat_Strand$Strand =="2")), Methylated = (sum(dat_Strand$Strand == "2"))/(sum(dat_Strand$Strand == "0", dat_Strand$Strand == "1", dat_Strand$Strand =="2")))

  # Next, a new dataframe called dat_Meth is created. This calls the proportions of the column "Strand" that are methylated, hemimethylated, and unmethylated. It has three columns. The first is called "Unmethylated", and contains the proportion of "Strand" that is equivalent to "0" - i.e., both parent and daughter strand are unmethylated. The second is called "Hemimethylated", and contains the proportion of "Strand" that is equivalent to "1" - i.e., one of the parent or daughter strand is methylated and one unmethylated. The third is called "Methylated", and contains the proportion of "Strand" that is equivalent to "2" - i.e., both the parent and daughter strand are methylated.

  results <- rbind(results, dat_Meth)
  # Next, the proportions calculated for the dataframe "dat_Meth" are pasted onto the end of the dataframe "results".
 }
 write.table(results, file = "output3.csv", row.names = filenames, append = FALSE, col.names = NA, sep = ", ")

# Once the for loop has ended, the contents of the dataframe "results" are output to the directory "Methylation" in the file "output.csv". Each row in "output.csv" is labelled with the name of the file that its data came from, in the initial list "filenames".
filenames = list.files(pattern = '*_methylation.tsv', full.names = F)
}

Meth_to_table(filenames)
# This runs the function "Meth_to_table", on the list "filenames".

input.file <- "output2.csv"
```

Appendix 3. Supplementary table.

Supplementary Table1. Details of cell lines.

Cell line	Details:	
Jurkat DSMZ no: ACC 282	Cell type: T cell leukaemia Origin: Peripheral blood of 14-year old boy. Doubling time: 25-35 hours	Maximal density: 1.5 x 10^6 cells/ml Morphology: round Cells growing singly or clumps in suspension.
Molt4 DSMZ no: ACC 362	Cell type: T cell leukaemia Origin: Peripheral blood of 19-year old man. Doubling time: 36 hours	Maximal density: 0.5-1.3 x 10^6 cells/ml Morphology: round Cells growing in suspension.
Nalm6 DSMZ no: ACC128	Cell type: B cell precursor leukaemia Origin: Peripheral blood of 19-year old man. Doubling time: 40 hours	Maximal density: 1.0-2.0 x 10^6 cells/ml Morphology: single, small round Cells growing in suspension.
HL60 DSMZ no: ACC 282	Cell type: acute myeloid leukaemia Origin: Peripheral blood of 35-year old woman. Doubling time: about 40 hours	Maximal density: 1.5-2.0 x 10^6 cells/ml Morphology: round single Cells growing in suspension.

References:

Abdel-Wahab O, Gao J, Adli M, Dey A, Trimarchi T, Chung YR, Kuscu C, Hricik T, Ndiaye-Lobry D, Lafave LM, Koche R, Shih AH, Guryanova OA, Kim E, Li S, Pandey S, Shin JY, Telis L, Liu J, Bhatt PK, Monette S, Zhao X, Mason CE, Park CY, Bernstein BE, Aifantis I, & Levine RL. (2013). Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. *J Exp Med,* 210(12), 2641-2659. doi:10.1084/jem.20131141

An J, Rao A, & Ko M. (2017). TET family dioxygenases and DNA demethylation in stem cells and cancers. *Exp Mol Med,* 49(4), e323. doi:10.1038/emm.2017.5

Augui S, Nora EP, & Heard E. (2011). Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat Rev Genet,* 12(6), 429-442. doi:10.1038/nrg2987

Babosova O, Kapralova K, Raskova Kafkova L, Korinek V, Divoky V, Prchal JT, & Lanikova L. (2019). Iron chelation and 2-oxoglutarate-dependent dioxygenase inhibition suppress mantle cell lymphoma's cyclin D1. *J Cell Mol Med,* 23(11), 7785-7795. doi:10.1111/jcmm.14655

Badar T, Kantarjian HM, Ravandi F, Jabbour E, Borthakur G, Cortes JE, Pemmaraju N, Pierce SR, Newberry KJ, Daver N, & Verstovsek S. (2015). Therapeutic benefit of decitabine, a hypomethylating agent, in patients with high-risk primary myelofibrosis and myeloproliferative neoplasm in accelerated or blastic/acute myeloid leukemia phase. *Leuk Res,* 39(9), 950-956. doi:10.1016/j.leukres.2015.06.001

Barros SP, & Offenbacher S. (2009). Epigenetics: connecting environment and genotype to phenotype and disease. *J Dent Res,* 88(5), 400-408. doi:10.1177/0022034509335868

Baylin SB, & Jones PA. (2011). A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer,* 11(10), 726-734. doi:10.1038/nrc3130

Bestor T, Laudano A, Mattaliano R, & Ingram V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells: The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of Molecular Biology,* 203(4), 971-983. doi:https://doi.org/10.1016/0022-2836(88)90122-2

Bhutani N, Burns DM, & Blau HM. (2011). DNA demethylation dynamics. *Cell,* 146(6), 866-872. doi:10.1016/j.cell.2011.08.042

Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, & Taylor J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol,* Chapter 19, Unit 19.10.11-21. doi:10.1002/0471142727.mb1910s89

Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, Lorincz MC, & Ramalho-Santos M. (2013). Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature,* 500(7461), 222-226. doi:10.1038/nature12362

Bochtler M, Kolano A, & Xu GL. (2017). DNA demethylation pathways: Additional players and regulators. *Bioessays,* 39(1), 1-13. doi:10.1002/bies.201600178

Bocker MT, Tuorto F, Raddatz G, Musch T, Yang FC, Xu M, Lyko F, & Breiling A. (2012). Hydroxylation of 5-methylcytosine by TET2 maintains the active state of the mammalian HOXA cluster. *Nat Commun,* 3, 818. doi:10.1038/ncomms1826

Boland MJ, Nazor KL, & Loring JF. (2014). Epigenetic regulation of pluripotency and differentiation. *Circ Res,* 115(2), 311-324. doi:10.1161/circresaha.115.301517

Bostick M, Kim JK, Estève P-O, Clark A, Pradhan S, & Jacobsen SE. (2007). UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. *Science,* 317(5845), 1760-1764. doi:10.1126/science.1147939

Bourc'his D, & Bestor TH. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature,* 431(7004), 96-99. doi:10.1038/nature02886

Branco MR, Ficz G, & Reik W. (2011). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet,* 13(1), 7-13. doi:10.1038/nrg3080

Brender JD, Werler MM, Kelley KE, Vuong AM, Shinde MU, Zheng Q, Huber JC, Jr., Sharkey JR, Griesenbeck JS, Romitti PA, Langlois PH, Suarez L, & Canfield MA. (2011). Nitrosatable drug exposure during early pregnancy and neural tube defects in offspring: National Birth Defects Prevention Study. *Am J Epidemiol,* 174(11), 1286-1295. doi:10.1093/aje/kwr254

Bronner IF, Quail MA, Turner DJ, & Swerdlow H. (2014). Improved Protocols for Illumina Sequencing. *Curr Protoc Hum Genet,* 80, 18.12.11-42. doi:10.1002/0471142905.hg1802s80

Calvanese V, Fernandez AF, Urdinguio RG, Suarez-Alvarez B, Mangas C, Perez-Garcia V, Bueno C, Montes R, Ramos-Mejia V, Martinez-Camblor P, Ferrero C, Assenov Y, Bock C, Menendez P, Carrera AC, Lopez-Larrea C, & Fraga MF. (2012). A promoter DNA demethylation landscape of human hematopoietic differentiation. *Nucleic Acids Res,* 40(1), 116-131. doi:10.1093/nar/gkr685

Camarena V, & Wang G. (2016). The epigenetic role of vitamin C in health and disease. *Cell Mol Life Sci,* 73(8), 1645-1658. doi:10.1007/s00018-016-2145-x

Cao T, Pan W, Sun X, & Shen H. (2019). Increased expression of TET3 predicts unfavorable prognosis in patients with ovarian cancer-a bioinformatics integrative analysis. *Journal of Ovarian Research,* 12(1), 101. doi:10.1186/s13048-019-0575-4

Carella A, Tejedor JR, Garcia MG, Urdinguio RG, Bayon GF, Sierra M, Lopez V, Garcia-Torano E, Santamarina-Ojeda P, Perez RF, Bigot T, Mangas C, Corte-Torres MD, Saenz-de-Santa-Maria I, Mollejo M, Melendez B, Astudillo A, Chiara MD, Fernandez AF, & Fraga MF. (2019). Epigenetic downregulation of TET3 reduces genome-wide 5hmC levels and promotes glioblastoma tumorigenesis. *Int J Cancer*. doi:10.1002/ijc.32520

Cheishvili D, Boureau L, & Szyf M. (2015). DNA demethylation and invasive cancer: implications for therapeutics. *Br J Pharmacol,* 172(11), 2705-2715. doi:10.1111/bph.12885

Chen CC, Wang KY, & Shen CK. (2013). DNA 5-methylcytosine demethylation activities of the mammalian DNA methyltransferases. *J Biol Chem,* 288(13), 9084-9091. doi:10.1074/jbc.M112.445585

Chen G, & Deng X. (2018a). Cell Synchronization by Double Thymidine Block. *Bio-protocol,* 8(17), e2994. doi:10.21769/BioProtoc.2994

Chen G, & Deng X. (2018b). Cell Synchronization by Double Thymidine Block. *Bio Protoc,* 8(17). doi:10.21769/BioProtoc.2994

Chen ML, Shen F, Huang W, Qi JH, Wang Y, Feng YQ, Liu SM, & Yuan BF. (2013). Quantification of 5-methylcytosine and 5-hydroxymethylcytosine in genomic DNA from hepatocellular carcinoma tissues by capillary hydrophilic-interaction liquid chromatography/quadrupole TOF mass spectrometry. *Clin Chem,* 59(5), 824-832. doi:10.1373/clinchem.2012.193938

Chen T, Ueda Y, Dodge JE, Wang Z, & Li E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol,* 23(16), 5594-5605. doi:10.1128/mcb.23.16.5594-5605.2003

Cheng X, & Blumenthal RM. (2008). Mammalian DNA methyltransferases: a structural perspective. *Structure,* 16(3), 341-350. doi:10.1016/j.str.2008.01.004

Chepda T, Cadau M, Girin P, Frey J, & Chamson A. (2001). Monitoring of ascorbate at a constant rate in cell culture: effect on cell growth. *In Vitro Cell Dev Biol Anim,* 37(1), 26-30. doi:10.1290/1071-2690(2001)037<0026:Moaaac>2.0.Co;2

Chiacchiera F, Piunti A, & Pasini D. (2013). Epigenetic methylations and their connections with metabolism. *Cell Mol Life Sci,* 70(9), 1495-1508. doi:10.1007/s00018-013-1293-5

Chiba S. (2017). Dysregulation of TET2 in hematologic malignancies. *Int J Hematol,* 105(1), 17-22. doi:10.1007/s12185-016-2122-z

Chowdhury B, McGovern A, Cui Y, Choudhury SR, Cho I-H, Cooper B, Chevassut T, Lossie AC, & Irudayaraj J. (2015a). The hypomethylating agent Decitabine causes a paradoxical increase in 5-hydroxymethylcytosine in human leukemia cells. *Scientific Reports,* 5, 9281. doi:10.1038/srep09281

https://www.nature.com/articles/srep09281#supplementary-information

Chowdhury B, McGovern A, Cui Y, Choudhury SR, Cho IH, Cooper B, Chevassut T, Lossie AC, & Irudayaraj J. (2015b). The hypomethylating agent Decitabine causes a paradoxical increase in 5-hydroxymethylcytosine in human leukemia cells. *Sci Rep,* 5, 9281. doi:10.1038/srep09281

Chung TL, Brena RM, Kolle G, Grimmond SM, Berman BP, Laird PW, Pera MF, & Wolvetang EJ. (2010). Vitamin C promotes widespread yet specific DNA demethylation of the epigenome in human embryonic stem cells. *Stem Cells,* 28(10), 1848-1855. doi:10.1002/stem.493

Cimmino L, Abdel-Wahab O, Levine RL, & Aifantis I. (2011). TET family proteins and their role in stem cell differentiation and transformation. *Cell Stem Cell,* 9(3), 193-204. doi:10.1016/j.stem.2011.08.007

Cimmino L, Dawlaty MM, Ndiaye-Lobry D, Yap YS, Bakogianni S, Yu Y, Bhattacharyya S, Shaknovich R, Geng H, Lobry C, Mullenders J, King B, Trimarchi T, Aranda-Orgilles B, Liu C, Shen S, Verma AK, Jaenisch R, & Aifantis I. (2015). Erratum: TET1 is a tumor suppressor of hematopoietic malignancy. *Nat Immunol,* 16(8), 889. doi:10.1038/ni0815-889a

Clark SJ, Harrison J, Paul CL, & Frommer M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res,* 22(15), 2990-2997. doi:10.1093/nar/22.15.2990

Cock PJ, Fields CJ, Goto N, Heuer ML, & Rice PM. (2010). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res,* 38(6), 1767-1771. doi:10.1093/nar/gkp1137

Corbet C, & Feron O. (2017). Cancer cell metabolism and mitochondria: Nutrient plasticity for TCA cycle fueling. *Biochim Biophys Acta Rev Cancer,* 1868(1), 7-15. doi:10.1016/j.bbcan.2017.01.002

Crawford DJ, Liu MY, Nabel CS, Cao XJ, Garcia BA, & Kohli RM. (2016). Tet2 Catalyzes Stepwise 5-Methylcytosine Oxidation by an Iterative and de novo Mechanism. *J Am Chem Soc,* 138(3), 730-733. doi:10.1021/jacs.5b10554

Cude K, Wang Y, Choi HJ, Hsuan SL, Zhang H, Wang CY, & Xia Z. (2007). Regulation of the G2-M cell cycle progression by the ERK5-NFkappaB signaling pathway. *J Cell Biol,* 177(2), 253-264. doi:10.1083/jcb.200609166

Cui D, & Xu X. (2018). DNA Methyltransferases, DNA Methylation, and Age-Associated Cognitive Function. *Int J Mol Sci,* 19(5). doi:10.3390/ijms19051315

Dai Q, Sanstead PJ, Peng CS, Han D, He C, & Tokmakoff A. (2016). Weakened N3 Hydrogen Bonding by 5-Formylcytosine and 5-Carboxylcytosine Reduces Their Base-Pairing Stability. *ACS Chem Biol,* 11(2), 470-477. doi:10.1021/acschembio.5b00762

Damaraju VL, Mowles D, Yao S, Ng A, Young JD, Cass CE, & Tong Z. (2012). Role of human nucleoside transporters in the uptake and cytotoxicity of azacitidine and decitabine. *Nucleosides Nucleotides Nucleic Acids,* 31(3), 236-255. doi:10.1080/15257770.2011.652330

Damelin M, & Bestor TH. (2007). Biological functions of DNA methyltransferase 1 require its methyltransferase activity. *Mol Cell Biol,* 27(11), 3891-3899. doi:10.1128/mcb.00036-07

Daskalakis M, Blagitko-Dorfs N, & Hackanson B. (2010). Decitabine. *Recent Results Cancer Res,* 184, 131-157. doi:10.1007/978-3-642-01222-8_10

Dawlaty MM, Breiling A, Le T, Raddatz G, Barrasa MI, Cheng AW, Gao Q, Powell BE, Li Z, Xu M, Faull KF, Lyko F, & Jaenisch R. (2013). Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell,* 24(3), 310-323. doi:10.1016/j.devcel.2012.12.015

Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, Cheng AW, Gao Q, Kim J, Choi SW, Page DC, & Jaenisch R. (2011). Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell,* 9(2), 166-175. doi:10.1016/j.stem.2011.07.010

Daye D, & Wellen KE. (2012). Metabolic reprogramming in cancer: unraveling the role of glutamine in tumorigenesis. *Semin Cell Dev Biol,* 23(4), 362-369. doi:10.1016/j.semcdb.2012.02.002

Delatte B, Deplus R, & Fuks F. (2014). Playing TETris with DNA modifications. *Embo j,* 33(11), 1198-1211. doi:10.15252/embj.201488290

Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, Kosmider O, Le Couedic JP, Robert F, Alberdi A, Lecluse Y, Plo I, Dreyfus FJ, Marzac C, Casadevall N, Lacombe C, Romana SP, Dessen P, Soulier J, Viguie F, Fontenay M, Vainchenker W, & Bernard OA. (2009). Mutation in TET2 in myeloid cancers. *N Engl J Med,* 360(22), 2289-2301. doi:10.1056/NEJMoa0810069

Dickson KM, Gustafson CB, Young JI, Zuchner S, & Wang G. (2013). Ascorbate-induced generation of 5-hydroxymethylcytosine is unaffected by varying levels of iron and 2-oxoglutarate. *Biochem Biophys Res Commun,* 439(4), 522-527. doi:10.1016/j.bbrc.2013.09.010

Ding F, & Chaillet JR. (2002). In vivo stabilization of the Dnmt1 (cytosine-5)-methyltransferase protein. *Proc Natl Acad Sci U S A,* 99(23), 14861-14866. doi:10.1073/pnas.232565599

Ehrlich M. (2009). DNA hypomethylation in cancer cells. *Epigenomics,* 1(2), 239-259. doi:10.2217/epi.09.33

Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, & Gehrke C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res,* 10(8), 2709-2721. doi:10.1093/nar/10.8.2709

Esteban MA, Wang T, Qin B, Yang J, Qin D, Cai J, Li W, Weng Z, Chen J, Ni S, Chen K, Li Y, Liu X, Xu J, Zhang S, Li F, He W, Labuda K, Song Y, Peterbauer A, Wolbank S, Redl H, Zhong M, Cai D, Zeng L, & Pei D. (2010). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell,* 6(1), 71-79. doi:10.1016/j.stem.2009.12.001

Esteller M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet,* 8(4), 286-298. doi:10.1038/nrg2005

Evans B, & Griner E. (2015). Registered report: Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Elife,* 4, e07420. doi:10.7554/eLife.07420

Falk KI, & Ernberg I. (1999). Demethylation of the Epstein-barr virus origin of lytic replication and of the immediate early gene BZLF1 is DNA replication independent. Brief report. *Arch Virol,* 144(11), 2219-2227.

Feil R, & Fraga MF. (2012). Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet,* 13(2), 97-109. doi:10.1038/nrg3142

Feinberg AP, & Tycko B. (2004). The history of cancer epigenetics. *Nature Reviews Cancer,* 4(2), 143-153. doi:10.1038/nrc1279

Fenaux P, Mufti GJ, Hellstrom-Lindberg E, Santini V, Gattermann N, Germing U, Sanz G, List AF, Gore S, Seymour JF, Dombret H, Backstrom J, Zimmerman L, McKenzie D, Beach CL, & Silverman LR. (2010). Azacitidine prolongs overall survival compared with conventional care regimens in elderly patients with low bone marrow blast count acute myeloid leukemia. *J Clin Oncol,* 28(4), 562-569. doi:10.1200/jco.2009.23.8329

Feng S, Rubbi L, Jacobsen SE, & Pellegrini M. (2011). Determining DNA methylation profiles using sequencing. *Methods Mol Biol,* 733, 223-238. doi:10.1007/978-1-61779-089-8_16

Feng Y, Li X, Cassady K, Zou Z, & Zhang X. (2019). TET2 Function in Hematopoietic Malignancies, Immune Regulation, and DNA Repair. *Front Oncol,* 9, 210. doi:10.3389/fonc.2019.00210

Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, Marques CJ, Andrews S, & Reik W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature,* 473(7347), 398-402. doi:10.1038/nature10008

Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, Tallman MS, Sun Z, Wolniak K, Peeters JK, Liu W, Choe SE, Fantin VR, Paietta E, Löwenberg B, Licht JD, Godley LA, Delwel R, Valk PJ, Thompson CB, Levine RL, & Melnick A. (2010). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell,* 18(6), 553-567. doi:10.1016/j.ccr.2010.11.015

Fimognari C, Nusse M, Cesari R, Iori R, Cantelli-Forti G, & Hrelia P. (2002). Growth inhibition, cell-cycle arrest and apoptosis in human T-cell leukemia by the isothiocyanate sulforaphane. *Carcinogenesis,* 23(4), 581-586. doi:10.1093/carcin/23.4.581

Fu T, Liu L, Yang QL, Wang Y, Xu P, Zhang L, Liu S, Dai Q, Ji Q, Xu GL, He C, Luo C, & Zhang L. (2019). Thymine DNA glycosylase recognizes the geometry alteration of minor grooves induced by 5-formylcytosine and 5-carboxylcytosine. *Chem Sci,* 10(31), 7407-7417. doi:10.1039/c9sc02807b

Gao W, Yu X, Hao J, Wang L, Qi M, Han L, Lin C, & Wang D. (2019). Ascorbic acid improves parthenogenetic embryo development through TET proteins in mice. *Bioscience Reports,* 39(1). doi:10.1042/bsr20181730

Gerecke C, Schumacher F, Edlich A, Wetzel A, Yealland G, Neubert LK, Scholtka B, Homann T, & Kleuser B. (2018). Vitamin C promotes decitabine or azacytidine induced DNA hydroxymethylation and subsequent reactivation of the epigenetically silenced tumour suppressor CDKN1A in colon cancer cells. *Oncotarget,* 9(67), 32822-32840. doi:10.18632/oncotarget.25999

Ghoshal K, Datta J, Majumder S, Kutay H, & Jacob ST. (2004). Molecular mechanism of decitabine -induced degradation of DNA methyltransferase 1 (DNMT1) in cancer cells. *Cancer Research,* 64(7 Supplement), 367-367.

Giehr P, Kyriakopoulos C, Lepikhov K, Wallner S, Wolf V, & Walter J. (2018). Two are better than one: HPoxBS - hairpin oxidative bisulfite sequencing. *Nucleic Acids Res,* 46(15), e88. doi:10.1093/nar/gky422

Gius D, Cui H, Bradbury CM, Cook J, Smart DK, Zhao S, Young L, Brandenburg SA, Hu Y, Bisht KS, Ho AS, Mattson D, Sun L, Munson PJ, Chuang EY, Mitchell JB, & Feinberg AP. (2004). Distinct effects on gene expression of chemical and genetic manipulation of the cancer epigenome revealed by a multimodality approach. *Cancer Cell,* 6(4), 361-371. doi:10.1016/j.ccr.2004.08.029

Globisch D, Munzel M, Muller M, Michalakis S, Wagner M, Koch S, Bruckl T, Biel M, & Carell T. (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One,* 5(12), e15367. doi:10.1371/journal.pone.0015367

Globisch D, Münzel M, Müller M, Michalakis S, Wagner M, Koch S, Brückl T, Biel M, & Carell T. (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One,* 5(12), e15367. doi:10.1371/journal.pone.0015367

Goll MG, & Bestor TH. (2005). Eukaryotic cytosine methyltransferases. *Annu Rev Biochem,* 74, 481-514. doi:10.1146/annurev.biochem.74.010904.153721

Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, Golic KG, Jacobsen SE, & Bestor TH. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science,* 311(5759), 395-398. doi:10.1126/science.1120976

Gong Z, & Zhu JK. (2011). Active DNA demethylation by oxidation and repair. *Cell Res,* 21(12), 1649-1651. doi:10.1038/cr.2011.140

Good CR, Madzo J, Patel B, Maegawa S, Engel N, Jelinek J, & Issa JJ. (2017). A novel isoform of TET1 that lacks a CXXC domain is overexpressed in cancer. *Nucleic Acids Res,* 45(14), 8269-8281. doi:10.1093/nar/gkx435

Gordon M, El-Kalla M, & Baksh S. (2012). RASSF1 Polymorphisms in Cancer. *Mol Biol Int,* 2012, 365213. doi:10.1155/2012/365213

Gould BS, & Woessner JF. (1957). Biosynthesis of collagen; the influence of ascorbic acid on the proline, hydroxyproline, glycine, and collagen content of regenerating guinea pig skin. *J Biol Chem,* 226(1), 289-300.

Groh S, & Schotta G. (2017). Silencing of endogenous retroviruses by heterochromatin. *Cell Mol Life Sci,* 74(11), 2055-2065. doi:10.1007/s00018-017-2454-8

Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, Xie ZG, Shi L, He X, Jin SG, Iqbal K, Shi YG, Deng Z, Szabo PE, Pfeifer GP, Li J, & Xu GL. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature,* 477(7366), 606-610. doi:10.1038/nature10443

Guo JU, Su Y, Zhong C, Ming GL, & Song H. (2011a). Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle,* 10(16), 2662-2668. doi:10.4161/cc.10.16.17093

Guo JU, Su Y, Zhong C, Ming GL, & Song H. (2011b). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell,* 145(3), 423-434. doi:10.1016/j.cell.2011.03.022

Guo X, Long J, Zeng C, Michailidou K, Ghoussaini M, Bolla MK, Wang Q, Milne RL, Shu XO, Cai Q, Beesley J, Kar SP, Andrulis IL, Anton-Culver H, Arndt V, Beckmann MW, Beeghly-Fadiel A, Benitez J, Blot W, Bogdanova N, Bojesen SE, Brauch H, Brenner H, Brinton L, Broeks A, Bruning T, Burwinkel B, Cai H, Canisius S, Chang-Claude J, Choi JY, Couch FJ, Cox A, Cross SS, Czene K, Darabi H, Devilee P, Droit A, Dork T, Fasching PA, Fletcher O, Flyger H, Fostira F, Gaborieau V, Garcia-Closas M, Giles GG, Grip M, Guenel P, Haiman CA, Hamann U, Hartman M, Hollestelle A, Hopper JL, Hsiung CN, Ito H, Jakubowska A, Johnson N, Kabisch M, Kang D, Khan S, Knight JA, Kosma VM, Lambrechts D, Le Marchand L, Li J, Lindblom A, Lophatananon A, Lubinski J, Mannermaa A, Manoukian S, Margolin S, Marme F, Matsuo K, McLean CA, Meindl A, Muir K, Neuhausen SL, Nevanlinna H, Nord S, Olson JE, Orr N, Peterlongo P, Putti TC, Rudolph A, Sangrajrang S, Sawyer EJ, Schmidt MK, Schmutzler RK, Shen CY, Shi J, Shrubsole MJ, Southey MC, Swerdlow A, Teo SH, Thienpont B, Toland AE, Tollenaar RA, Tomlinson IP, Truong T, Tseng CC, van den Ouweland A, Wen W, Winqvist R, Wu A, Yip CH, Zamora MP, Zheng Y, Hall P, Pharoah PD, Simard J, Chenevix-Trench G, Dunning AM, Easton DF, & Zheng W.

(2015). Fine-scale mapping of the 4q24 locus identifies two independent loci associated with breast cancer risk. *Cancer Epidemiol Biomarkers Prev,* 24(11), 1680-1691. doi:10.1158/1055-9965.Epi-15-0363

Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, & Surani MA. (2013). Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science,* 339(6118), 448-452. doi:10.1126/science.1229277

Hackett JA, & Surani MA. (2013). DNA methylation dynamics during the mammalian life cycle. *Philos Trans R Soc Lond B Biol Sci,* 368(1609), 20110328. doi:10.1098/rstb.2011.0328

Haffner MC, Chaux A, Meeker AK, Esopi DM, Gerber J, Pellakuru LG, Toubaji A, Argani P, Iacobuzio-Donahue C, Nelson WG, Netto GJ, De Marzo AM, & Yegnasubramanian S. (2011). Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. *Oncotarget,* 2(8), 627-637. doi:10.18632/oncotarget.316

Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J, & Surani MA. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of Development,* 117(1), 15-23. doi:https://doi.org/10.1016/S0925-4773(02)00181-8

Han JA, An J, & Ko M. (2015). Functions of TET Proteins in Hematopoietic Transformation. *Mol Cells,* 38(11), 925-935. doi:10.14348/molcells.2015.0294

Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, Wen B, Wu H, Liu Y, Diep D, Briem E, Zhang K, Irizarry RA, & Feinberg AP. (2011). Increased methylation variation in epigenetic domains across cancer types. *Nat Genet,* 43(8), 768-775. doi:10.1038/ng.865

Hashimoto H, Hong S, Bhagwat AS, Zhang X, & Cheng X. (2012). Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. *Nucleic Acids Res,* 40(20), 10203-10214. doi:10.1093/nar/gks845

Hashimoto H, Pais JE, Zhang X, Saleh L, Fu ZQ, Dai N, Correa IR, Jr., Zheng Y, & Cheng X. (2014). Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. *Nature,* 506(7488), 391-395. doi:10.1038/nature12905

Hashimoto H, Zhang X, & Cheng X. (2013). Activity and crystal structure of human thymine DNA glycosylase mutant N140A with 5-carboxylcytosine DNA at low pH. *DNA Repair (Amst),* 12(7), 535-540. doi:10.1016/j.dnarep.2013.04.003

Hata K, Okano M, Lei H, & Li E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development,* 129(8), 1983-1993.

Havecker ER, Gao X, & Voytas DF. (2004). The diversity of LTR retrotransposons. *Genome Biol,* 5(6), 225. doi:10.1186/gb-2004-5-6-225

He PF, Zhou JD, Yao DM, Ma JC, Wen XM, Zhang ZH, Lian XY, Xu ZJ, Qian J, & Lin J. (2017). Efficacy and safety of decitabine in treatment of elderly patients with acute myeloid leukemia: A systematic review and meta-analysis. *Oncotarget,* 8(25), 41498-41507. doi:10.18632/oncotarget.17241

He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, & Xu GL. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science,* 333(6047), 1303-1307. doi:10.1126/science.1210944

Hellman A, & Chess A. (2007). Gene body-specific methylation on the active X chromosome. *Science,* 315(5815), 1141-1143. doi:10.1126/science.1136352

Hill PW, Amouroux R, & Hajkova P. (2014). DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: an emerging complex story. *Genomics,* 104(5), 324-333. doi:10.1016/j.ygeno.2014.08.012

Ho JJ, Cattoglio C, McSwiggen DT, Tjian R, & Fong YW. (2017). Regulation of DNA demethylation by the XPC DNA repair complex in somatic and pluripotent stem cells. *Genes Dev,* 31(8), 830-844. doi:10.1101/gad.295741.116

Hodges E, Molaro A, Dos Santos CO, Thekkat P, Song Q, Uren PJ, Park J, Butler J, Rafii S, McCombie WR, Smith AD, & Hannon GJ. (2011). Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. *Mol Cell,* 44(1), 17-28. doi:10.1016/j.molcel.2011.08.026

Holliday R, & Pugh JE. (1975). DNA modification mechanisms and gene activity during development. *Science,* 187(4173), 226-232.

Hore TA, von Meyenn F, Ravichandran M, Bachman M, Ficz G, Oxley D, Santos F, Balasubramanian S, Jurkowski TP, & Reik W. (2016). Retinol and ascorbate drive erasure of epigenetic memory and enhance reprogramming to naive pluripotency by complementary mechanisms. *Proc Natl Acad Sci U S A,* 113(43), 12202-12207. doi:10.1073/pnas.1608679113

Hotchkiss RD. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J Biol Chem,* 175(1), 315-332.

Howard G, Eiges R, Gaudet F, Jaenisch R, & Eden A. (2008). Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene,* 27(3), 404-408. doi:10.1038/sj.onc.1210631

Hsu CH, Peng KL, Kang ML, Chen YR, Yang YC, Tsai CH, Chu CS, Jeng YM, Chen YT, Lin FM, Huang HD, Lu YY, Teng YC, Lin ST, Lin RK, Tang FM, Lee SB, Hsu HM, Yu JC, Hsiao PW, & Juan LJ. (2012). TET1 suppresses cancer invasion by activating the tissue inhibitors of metalloproteinases. *Cell Rep,* 2(3), 568-579. doi:10.1016/j.celrep.2012.08.030

Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, Shi YG, Zhu J, Wang P, & Xu Y. (2013). Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell,* 155(7), 1545-1555. doi:10.1016/j.cell.2013.11.020

Hu L, Lu J, Cheng J, Rao Q, Li Z, Hou H, Lou Z, Zhang L, Li W, Gong W, Liu M, Sun C, Yin X, Li J, Tan X, Wang P, Wang Y, Fang D, Cui Q, Yang P, He C, Jiang H, Luo C, & Xu Y. (2015). Structural insight into substrate preference for TET-mediated oxidation. *Nature,* 527(7576), 118-122. doi:10.1038/nature15713

Huang P, Chubb S, Hertel LW, Grindey GB, & Plunkett W. (1991). Action of 2',2'-difluorodeoxycytidine on DNA synthesis. *Cancer Res,* 51(22), 6110-6117.

Huang Y, & Rao A. (2014). Connections between TET proteins and aberrant DNA modification in cancer. *Trends Genet,* 30(10), 464-474. doi:10.1016/j.tig.2014.07.005

Illingworth RS, & Bird AP. (2009). CpG islands--'a rough guide'. *FEBS Lett,* 583(11), 1713-1720. doi:10.1016/j.febslet.2009.04.012

Imanishi S, Takahashi R, Katagiri S, Kobayashi C, Umezu T, Ohyashiki K, & Ohyashiki JH. (2017). Teriflunomide restores 5-azacytidine sensitivity via activation of pyrimidine salvage in 5-azacytidine-resistant leukemia cells. *Oncotarget,* 8(41), 69906-69915. doi:10.18632/oncotarget.19436

Inoue A, Shen L, Dai Q, He C, & Zhang Y. (2011). Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. *Cell Res,* 21(12), 1670-1676. doi:10.1038/cr.2011.189

Inoue A, & Zhang Y. (2011). Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science,* 334(6053), 194. doi:10.1126/science.1212483

Iqbal K, Jin SG, Pfeifer GP, & Szabo PE. (2011). Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci U S A,* 108(9), 3642-3647. doi:10.1073/pnas.1014033108

Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, & Zhang Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature,* 466(7310), 1129-1133. doi:10.1038/nature09303

Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, & Zhang Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science,* 333(6047), 1300-1303. doi:10.1126/science.1210597

Iurlaro M, Ficz G, Oxley D, Raiber EA, Bachman M, Booth MJ, Andrews S, Balasubramanian S, & Reik W. (2013). A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. *Genome Biol,* 14(10), R119. doi:10.1186/gb-2013-14-10-r119

Iwan K, Rahimoff R, Kirchner A, Spada F, Schroder AS, Kosmatchev O, Ferizaj S, Steinbacher J, Parsa E, Muller M, & Carell T. (2018). 5-Formylcytosine to cytosine conversion by C-C bond cleavage in vivo. *Nat Chem Biol,* 14(1), 72-78. doi:10.1038/nchembio.2531

Iyer LM, Tahiliani M, Rao A, & Aravind L. (2009). Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle,* 8(11), 1698-1710. doi:10.4161/cc.8.11.8580

Jabbour E, Issa JP, Garcia-Manero G, & Kantarjian H. (2008). Evolution of decitabine development: accomplishments, ongoing investigations, and future strategies. *Cancer,* 112(11), 2341-2351. doi:10.1002/cncr.23463

Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J, & Ramsahoye B. (2004). Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol Cell Biol,* 24(20), 8862-8871. doi:10.1128/mcb.24.20.8862-8871.2004

Jin SG, Jiang Y, Qiu R, Rauch TA, Wang Y, Schackert G, Krex D, Lu Q, & Pfeifer GP. (2011). 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Res,* 71(24), 7360-7365. doi:10.1158/0008-5472.Can-11-2023

Jin Z, & Liu Y. (2018). DNA methylation in human diseases. *Genes Dis,* 5(1), 1-8. doi:10.1016/j.gendis.2018.01.002

Johnson TB, & Coghill RD. (1925). RESEARCHES ON PYRIMIDINES. C111. THE DISCOVERY OF 5-METHYL-CYTOSINE IN TUBERCULINIC ACID, THE NUCLEIC ACID OF THE TUBERCLE BACILLUS1. *Journal of the American Chemical Society,* 47(11), 2838-2844. doi:10.1021/ja01688a030

Jones PA. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet,* 13(7), 484-492. doi:10.1038/nrg3230

Jones PA, & Baylin SB. (2007). The epigenomics of cancer. *Cell,* 128(4), 683-692. doi:10.1016/j.cell.2007.01.029

Jones PA, & Liang G. (2009). Rethinking how DNA methylation patterns are maintained. *Nature Reviews Genetics,* 10(11), 805-811. doi:10.1038/nrg2651

Kagiwada S, Kurimoto K, Hirota T, Yamaji M, & Saitou M. (2013). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *Embo j,* 32(3), 340-353. doi:10.1038/emboj.2012.331

Kalender Atak Z, De Keersmaecker K, Gianfelici V, Geerdens E, Vandepoel R, Pauwels D, Porcu M, Lahortiga I, Brys V, Dirks WG, Quentmeier H, Cloos J, Cuppens H, Uyttebroeck A, Vandenberghe P, Cools J, & Aerts S. (2012). High accuracy mutation detection in leukemia on a selected panel of cancer genes. *PLoS One,* 7(6), e38463. doi:10.1371/journal.pone.0038463

Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, Leiserson MDM, Miller CA, Welch JS, Walter MJ, Wendl MC, Ley TJ, Wilson RK, Raphael BJ, & Ding L. (2013). Mutational landscape and

significance across 12 major cancer types. *Nature,* 502(7471), 333-339. doi:10.1038/nature12634

Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, & Sasaki H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature,* 429(6994), 900-903. doi:10.1038/nature02633

Kantarjian HM, O'Brien S, Huang X, Garcia-Manero G, Ravandi F, Cortes J, Shan J, Davisson J, Bueso-Ramos CE, & Issa JP. (2007). Survival advantage with decitabine versus intensive chemotherapy in patients with higher risk myelodysplastic syndrome: comparison with historical experience. *Cancer,* 109(6), 1133-1137. doi:10.1002/cncr.22508

Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, Shinkai Y, Mager DL, Jones S, Hirst M, & Lorincz MC. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell,* 8(6), 676-687. doi:10.1016/j.stem.2011.04.004

Kellinger MW, Song CX, Chong J, Lu XY, He C, & Wang D. (2012). 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat Struct Mol Biol,* 19(8), 831-833. doi:10.1038/nsmb.2346

Kinney SR, & Pradhan S. (2013). Ten eleven translocation enzymes and 5-hydroxymethylation in mammalian development and cancer. *Adv Exp Med Biol,* 754, 57-79. doi:10.1007/978-1-4419-9967-2_3

Klein CJ, Botuyan MV, Wu Y, Ward CJ, Nicholson GA, Hammans S, Hojo K, Yamanishi H, Karpf AR, Wallace DC, Simon M, Lander C, Boardman LA, Cunningham JM, Smith GE, Litchy WJ, Boes B, Atkinson EJ, Middha S, PJ BD, Parisi JE, Mer G, Smith DI, & Dyck PJ. (2011). Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss. *Nat Genet,* 43(6), 595-600. doi:10.1038/ng.830

Klug M, Heinz S, Gebhard C, Schwarzfischer L, Krause SW, Andreesen R, & Rehli M. (2010). Active DNA demethylation in human postmitotic cells correlates with activating histone modifications, but not transcription levels. *Genome Biol,* 11(6), R63. doi:10.1186/gb-2010-11-6-r63

Klug M, Schmidhofer S, Gebhard C, Andreesen R, & Rehli M. (2013). 5-Hydroxymethylcytosine is an essential intermediate of active DNA demethylation processes in primary human monocytes. *Genome Biol,* 14(5), R46. doi:10.1186/gb-2013-14-5-r46

Ko M, An J, Bandukwala HS, Chavez L, Aijo T, Pastor WA, Segal MF, Li H, Koh KP, Lahdesmaki H, Hogan PG, Aravind L, & Rao A. (2013). Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature,* 497(7447), 122-126. doi:10.1038/nature12052

Ko M, An J, Pastor WA, Koralov SB, Rajewsky K, & Rao A. (2015). TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunol Rev,* 263(1), 6-21. doi:10.1111/imr.12239

Ko M, Bandukwala HS, An J, Lamperti ED, Thompson EC, Hastie R, Tsangaratou A, Rajewsky K, Koralov SB, & Rao A. (2011). Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc Natl Acad Sci U S A,* 108(35), 14566-14571. doi:10.1073/pnas.1112317108

Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, An J, Lamperti ED, Koh KP, Ganetzky R, Liu XS, Aravind L, Agarwal S, Maciejewski JP, & Rao A. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature,* 468(7325), 839-843. doi:10.1038/nature09586

Kohli RM, & Zhang Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature,* 502(7472), 472-479. doi:10.1038/nature12750

Kriaucionis S, & Heintz N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science,* 324(5929), 929-930. doi:10.1126/science.1169786

Kroeze LI, van der Reijden BA, & Jansen JH. (2015). 5-Hydroxymethylcytosine: An epigenetic mark frequently deregulated in cancer. *Biochim Biophys Acta,* 1855(2), 144-154. doi:10.1016/j.bbcan.2015.01.001

Krueger F, & Andrews SR. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics,* 27(11), 1571-1572. doi:10.1093/bioinformatics/btr167

Kubuki Y, Yamaji T, Hidaka T, Kameda T, Shide K, Sekine M, Kamiunten A, Akizuki K, Shimoda H, Tahira Y, Nakamura K, Abe H, Miike T, Iwakiri H, Tahara Y, Sueta M, Yamamoto S, Hasuike S, Nagata K, Kitanaka A, & Shimoda K. (2017). TET2 mutation in diffuse large B-cell lymphoma. *J Clin Exp Hematop,* 56(3), 145-149. doi:10.3960/jslrt.56.145

Kuiper C, & Vissers MC. (2014). Ascorbate as a co-factor for fe- and 2-oxoglutarate dependent dioxygenases: physiological activity in tumor growth and progression. *Front Oncol,* 4, 359. doi:10.3389/fonc.2014.00359

Labet V, Morell C, Cadet J, Eriksson LA, & Grand A. (2009). Hydrolytic deamination of 5-methylcytosine in protic medium--a theoretical study. *J Phys Chem A,* 113(11), 2524-2533. doi:10.1021/jp808902j

Laird CD, Pleasant ND, Clark AD, Sneeden JL, Hassan KMA, Manley NC, Vary JC, Morgan T, Hansen RS, & Stöger R. (2004). Hairpin-bisulfite PCR: Assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proceedings of the National Academy of Sciences,* 101(1), 204-209. doi:10.1073/pnas.2536758100

Laisne M, Gupta N, Kirsh O, Pradhan S, & Defossez PA. (2018). Mechanisms of DNA Methyltransferase Recruitment in Mammals. *Genes (Basel),* 9(12). doi:10.3390/genes9120617

Lane DJ, Chikhani S, Richardson V, & Richardson DR. (2013). Transferrin iron uptake is stimulated by ascorbate via an intracellular reductive mechanism. *Biochim Biophys Acta,* 1833(6), 1527-1541. doi:10.1016/j.bbamcr.2013.02.010

Lane DJ, Robinson SR, Czerwinska H, Bishop GM, & Lawen A. (2010). Two routes of iron accumulation in astrocytes: ascorbate-dependent ferrous iron uptake via the divalent metal transporter (DMT1) plus an independent route for ferric iron. *Biochem J,* 432(1), 123-132. doi:10.1042/bj20101317

Langemeijer SM, Kuiper RP, Berends M, Knops R, Aslanyan MG, Massop M, Stevens-Linders E, van Hoogen P, van Kessel AG, Raymakers RA, Kamping EJ, Verhoef GE, Verburgh E, Hagemeijer A, Vandenberghe P, de Witte T, van der Reijden BA, & Jansen JH. (2009). Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat Genet,* 41(7), 838-842. doi:10.1038/ng.391

Lee Chong T, Ahearn EL, & Cimmino L. (2019). Reprogramming the Epigenome With Vitamin C. *Front Cell Dev Biol,* 7, 128. doi:10.3389/fcell.2019.00128

Lee HJ, Hore TA, & Reik W. (2014). Reprogramming the methylome: erasing memory and creating diversity. *Cell Stem Cell,* 14(6), 710-719. doi:10.1016/j.stem.2014.05.008

Lee TF, Zhai J, & Meyers BC. (2010). Conservation and divergence in eukaryotic DNA methylation. *Proc Natl Acad Sci U S A,* 107(20), 9027-9028. doi:10.1073/pnas.1005440107

Leger K, Hopp AK, Fey M, & Hottiger MO. (2016). ARTD1 regulates cyclin E expression and consequently cell-cycle re-entry and G1/S progression in T24 bladder carcinoma cells. *Cell Cycle,* 15(15), 2042-2052. doi:10.1080/15384101.2016.1195530

Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R, & Li E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development,* 122(10), 3195-3205.

Lemonnier F, Couronne L, Parrens M, Jais JP, Travert M, Lamant L, Tournillac O, Rousset T, Fabiani B, Cairns RA, Mak T, Bastard C, Bernard OA, de Leval L, & Gaulard P. (2012). Recurrent TET2 mutations in peripheral T-cell lymphomas correlate with TFH-like features and adverse clinical parameters. *Blood,* 120(7), 1466-1469. doi:10.1182/blood-2012-02-408542

Lenz SA, Kellie JL, & Wetmore SD. (2015). Glycosidic Bond Cleavage in DNA Nucleosides: Effect of Nucleobase Damage and Activation on the Mechanism and Barrier. *J Phys Chem B,* 119(51), 15601-15612. doi:10.1021/acs.jpcb.5b10337

Lester BM, Conradt E, & Marsit C. (2016). Introduction to the Special Section on Epigenetics. *Child Dev,* 87(1), 29-37. doi:10.1111/cdev.12489

Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J, Harris CC, Lichti CF, Townsend RR, Fulton RS, Dooling DJ, Koboldt DC, Schmidt H, Zhang Q, Osborne JR, Lin L, O'Laughlin M, McMichael JF, Delehaunty KD, McGrath SD, Fulton LA, Magrini VJ, Vickery TL, Hundal J, Cook LL, Conyers JJ, Swift GW, Reed JP, Alldredge PA, Wylie T, Walker J, Kalicki J, Watson MA, Heath S, Shannon WD, Varghese N, Nagarajan R, Westervelt P, Tomasson MH, Link DC, Graubert TA, DiPersio JF, Mardis ER, & Wilson RK. (2010). DNMT3A mutations in acute myeloid leukemia. *N Engl J Med,* 363(25), 2424-2433. doi:10.1056/NEJMoa1005143

Li E, Beard C, & Jaenisch R. (1993). Role for DNA methylation in genomic imprinting. *Nature,* 366(6453), 362-365. doi:10.1038/366362a0

Li E, Bestor TH, & Jaenisch R. (1992a). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell,* 69(6), 915-926. doi:https://doi.org/10.1016/0092-8674(92)90611-F

Li E, Bestor TH, & Jaenisch R. (1992b). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell,* 69(6), 915-926. doi:10.1016/0092-8674(92)90611-f

Li H, Li W, Liu S, Zong S, Wang W, Ren J, Li Q, Hou F, & Shi Q. (2016). DNMT1, DNMT3A and DNMT3B Polymorphisms Associated With Gastric Cancer Risk: A Systematic Review and Meta-analysis. *EBioMedicine,* 13, 125-131. doi:10.1016/j.ebiom.2016.10.028

Li LC, & Dahiya R. (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics,* 18(11), 1427-1431. doi:10.1093/bioinformatics/18.11.1427

Li Z, Cai X, Cai CL, Wang J, Zhang W, Petersen BE, Yang FC, & Xu M. (2011). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood,* 118(17), 4509-4518. doi:10.1182/blood-2010-12-325241

Lian CG, Xu Y, Ceol C, Wu F, Larson A, Dresser K, Xu W, Tan L, Hu Y, Zhan Q, Lee CW, Hu D, Lian BQ, Kleffel S, Yang Y, Neiswender J, Khorasani AJ, Fang R, Lezcano C, Duncan LM, Scolyer RA, Thompson JF, Kakavand H, Houvras Y, Zon LI, Mihm MC, Jr., Kaiser UB, Schatton T, Woda BA, Murphy GF, & Shi YG. (2012). Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell,* 150(6), 1135-1146. doi:10.1016/j.cell.2012.07.033

Lin G, Sun W, Yang Z, Guo J, Liu H, & Liang J. (2017). Hypoxia induces the expression of TET enzymes in HepG2 cells. *Oncol Lett,* 14(6), 6457-6462. doi:10.3892/ol.2017.7063

Linster CL, & Van Schaftingen E. (2007). Vitamin C. Biosynthesis, recycling and degradation in mammals. *Febs j,* 274(1), 1-22. doi:10.1111/j.1742-4658.2006.05607.x

Lio CJ, Yuita H, & Rao A. (2019). Dysregulation of the TET family of epigenetic regulators in hematopoietic malignancies. *Blood*. doi:10.1182/blood.2019791475

Liu B, Yuan J, Yiu SM, Li Z, Xie Y, Chen Y, Shi Y, Zhang H, Li Y, Lam TW, & Luo R. (2012). COPE: an accurate k-mer-based pair-end reads connection tool to facilitate genome assembly. *Bioinformatics,* 28(22), 2870-2874. doi:10.1093/bioinformatics/bts563

Liu M, Ohtani H, Zhou W, Orskov AD, Charlet J, Zhang YW, Shen H, Baylin SB, Liang G, Gronbaek K, & Jones PA. (2016). Vitamin C increases viral mimicry induced by 5-aza-2'-deoxycytidine. *Proc Natl Acad Sci U S A,* 113(37), 10238-10244. doi:10.1073/pnas.1612262113

Liu MY, Torabifard H, Crawford DJ, DeNizio JE, Cao XJ, Garcia BA, Cisneros GA, & Kohli RM. (2017). Mutations along a TET2 active site scaffold stall oxidation at 5-hydroxymethylcytosine. *Nat Chem Biol,* 13(2), 181-187. doi:10.1038/nchembio.2250

Liu N, Wang M, Deng W, Schmidt CS, Qin W, Leonhardt H, & Spada F. (2013). Intrinsic and extrinsic connections of Tet3 dioxygenase with CXXC zinc finger modules. *PLoS One,* 8(5), e62755. doi:10.1371/journal.pone.0062755

Liu X, Zhang G, Yi Y, Xiao L, Pei M, Liu S, Luo Y, Zhong H, Xu Y, Zheng W, & Shen J. (2013). Decreased 5-hydroxymethylcytosine levels are associated with TET2 mutation and unfavorable overall survival in myelodysplastic syndromes. *Leuk Lymphoma,* 54(11), 2466-2473. doi:10.3109/10428194.2013.778408

Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, & Downing JR. (2003). TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia,* 17(3), 637-641. doi:10.1038/sj.leu.2402834

Lutsik P, Feuerbach L, Arand J, Lengauer T, Walter J, & Bock C. (2011). BiQ Analyzer HT: locus-specific analysis of DNA methylation by high-throughput bisulfite sequencing. *Nucleic Acids Res,* 39(Web Server issue), W551-556. doi:10.1093/nar/gkr312

Lykkesfeldt J. (2007). Ascorbate and dehydroascorbic acid as reliable biomarkers of oxidative stress: analytical reproducibility and long-term stability of plasma samples subjected to acidic deproteinization. *Cancer Epidemiol Biomarkers Prev,* 16(11), 2513-2516. doi:10.1158/1055-9965.Epi-07-0639

Magoč T, & Salzberg SL. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics,* 27(21), 2957-2963. doi:10.1093/bioinformatics/btr507

Maiti A, & Drohat AC. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem,* 286(41), 35334-35338. doi:10.1074/jbc.C111.284620

Maiti A, Michelson AZ, Armwood CJ, Lee JK, & Drohat AC. (2013). Divergent mechanisms for enzymatic excision of 5-formylcytosine and 5-carboxylcytosine from DNA. *J Am Chem Soc,* 135(42), 15813-15822. doi:10.1021/ja406444x

Mancini M, Veljkovic N, Leo E, Aluigi M, Borsi E, Galloni C, Iacobucci I, Barbieri E, & Santucci MA. (2012). Cytoplasmatic compartmentalization by Bcr-Abl promotes TET2 loss-of-function in chronic myeloid leukemia. *J Cell Biochem,* 113(8), 2765-2774. doi:10.1002/jcb.24154

Markolovic S, Wilkins SE, & Schofield CJ. (2015). Protein Hydroxylation Catalyzed by 2-Oxoglutarate-dependent Oxygenases. *J Biol Chem,* 290(34), 20712-20722. doi:10.1074/jbc.R115.662627

Masella AP, Bartram AK, Truszkowski JM, Brown DG, & Neufeld JD. (2012). PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics,* 13, 31. doi:10.1186/1471-2105-13-31

Mathieu O, Reinders J, Caikovski M, Smathajitt C, & Paszkowski J. (2007). Transgenerational stability of the Arabidopsis epigenome is coordinated by CG methylation. *Cell,* 130(5), 851-862. doi:10.1016/j.cell.2007.07.007

May JM. (1998). Ascorbate function and metabolism in the human erythrocyte. *Front Biosci,* 3, d1-10. doi:10.2741/a262

Mayer W, Niveleau A, Walter J, Fundele R, & Haaf T. (2000). Demethylation of the zygotic paternal genome. *Nature,* 403(6769), 501-502. doi:10.1038/35000656

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, & Lander ES. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature,* 454(7205), 766-770. doi:10.1038/nature07107

Mertineit C, Yoder JA, Taketo T, Laird DW, Trasler JM, & Bestor TH. (1998). Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development,* 125(5), 889-897.

Messerschmidt DM, Knowles BB, & Solter D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev,* 28(8), 812-828. doi:10.1101/gad.234294.113

Milani L, Lundmark A, Kiialainen A, Nordlund J, Flaegstad T, Forestier E, Heyman M, Jonmundsson G, Kanerva J, Schmiegelow K, Soderhall S, Gustafsson MG, Lonnerholm G, & Syvanen AC. (2010). DNA methylation for subtype classification and prediction of treatment outcome in patients with childhood acute lymphoblastic leukemia. *Blood,* 115(6), 1214-1225. doi:10.1182/blood-2009-04-214668

Minor EA, Court BL, Young JI, & Wang G. (2013). Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *J Biol Chem,* 288(19), 13669-13674. doi:10.1074/jbc.C113.464800

Miranda TB, & Jones PA. (2007). DNA methylation: the nuts and bolts of repression. *J Cell Physiol,* 213(2), 384-390. doi:10.1002/jcp.21224

Misawa K, Imai A, Mochizuki D, Mima M, Endo S, Misawa Y, Kanazawa T, & Mineta H. (2018). Association of TET3 epigenetic inactivation with head and neck cancer. *Oncotarget,* 9(36), 24480-24493. doi:10.18632/oncotarget.25333

Mitalipov S, & Wolf D. (2009). Totipotency, pluripotency and nuclear reprogramming. *Adv Biochem Eng Biotechnol,* 114, 185-199. doi:10.1007/10_2008_45

Miura F, Enomoto Y, Dairiki R, & Ito T. (2012). Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res,* 40(17), e136. doi:10.1093/nar/gks454

Miyake K, Baba Y, Ishimoto T, Hiyoshi Y, Iwatsuki M, Miyamoto Y, Yoshida N, Watanabe M, Ogata Y, Nagayama M, Silsirivanit A, Kobayashi D, Araki N, & Baba H. (2018). Isocitrate dehydrogenase gene mutations and 2-hydroxyglutarate accumulation in esophageal squamous cell carcinoma. *Med Oncol,* 36(1), 11. doi:10.1007/s12032-018-1229-x

Momparler RL. (2005a). Epigenetic therapy of cancer with 5-aza-2'-deoxycytidine (decitabine). *Semin Oncol,* 32(5), 443-451. doi:10.1053/j.seminoncol.2005.07.008

Momparler RL. (2005b). Pharmacology of 5-Aza-2'-deoxycytidine (decitabine). *Semin Hematol,* 42(3 Suppl 2), S9-16.

Mondesir J, Willekens C, Touat M, & de Botton S. (2016). IDH1 and IDH2 mutations as novel therapeutic targets: current perspectives. *J Blood Med,* 7, 171-180. doi:10.2147/jbm.S70716

Moore LD, Le T, & Fan G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology,* 38(1), 23-38. doi:10.1038/npp.2012.112

Moore SP, Toomire KJ, & Strauss PR. (2013). DNA modifications repaired by base excision repair are epigenetic. *DNA Repair (Amst),* 12(12), 1152-1158. doi:10.1016/j.dnarep.2013.10.002

Mossman D, Kim KT, & Scott RJ. (2010). Demethylation by 5-aza-2'-deoxycytidine in colorectal cancer cells targets genomic DNA whilst promoter CpG island methylation persists. *BMC Cancer,* 10, 366. doi:10.1186/1471-2407-10-366

Muller T, Gessi M, Waha A, Isselstein LJ, Luxen D, Freihoff D, Freihoff J, Becker A, Simon M, Hammes J, Denkhaus D, zur Muhlen A, Pietsch T, & Waha A. (2012). Nuclear exclusion of TET1 is associated with loss of 5-hydroxymethylcytosine in IDH1 wild-type gliomas. *Am J Pathol,* 181(2), 675-683. doi:10.1016/j.ajpath.2012.04.017

Muto T, Sashida G, Oshima M, Wendt GR, Mochizuki-Kashio M, Nagata Y, Sanada M, Miyagi S, Saraya A, Kamio A, Nagae G, Nakaseko C, Yokote K, Shimoda K, Koseki H, Suzuki Y, Sugano S, Aburatani H, Ogawa S, & Iwama A. (2013). Concurrent loss of Ezh2 and Tet2 cooperates in the pathogenesis of myelodysplastic disorders. *J Exp Med,* 210(12), 2627-2639. doi:10.1084/jem.20131144

Nair VS, Song MH, Ko M, & Oh KI. (2016). DNA Demethylation of the Foxp3 Enhancer Is Maintained through Modulation of Ten-Eleven-Translocation and DNA Methyltransferases. *Mol Cells,* 39(12), 888-897. doi:10.14348/molcells.2016.0276

O'Connor KM, Das AB, Winterbourn CC, & Hampton MB. (2020). Inhibition of DNA methylation in proliferating human lymphoma cells by immune cell oxidants. *J Biol Chem*. doi:10.1074/jbc.RA120.013092

Odejide O, Weigert O, Lane AA, Toscano D, Lunning MA, Kopp N, Kim S, van Bodegom D, Bolla S, Schatz JH, Teruya-Feldstein J, Hochberg E, Louissaint A, Dorfman D, Stevenson K, Rodig SJ, Piccaluga PP, Jacobsen E, Pileri SA, Harris NL, Ferrero S, Inghirami G, Horwitz SM, & Weinstock DM. (2014). A targeted mutational landscape of angioimmunoblastic T-cell lymphoma. *Blood,* 123(9), 1293-1296. doi:10.1182/blood-2013-10-531509

Ohki I, Shimotake N, Fujita N, Jee J, Ikegami T, Nakao M, & Shirakawa M. (2001). Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell,* 105(4), 487-497. doi:10.1016/s0092-8674(01)00324-5

Ohno R, Nakayama M, Naruse C, Okashita N, Takano O, Tachibana M, Asano M, Saitou M, & Seki Y. (2013). A replication-dependent passive mechanism modulates DNA demethylation in mouse primordial germ cells. *Development,* 140(14), 2892-2903. doi:10.1242/dev.093229

Okano M, Bell DW, Haber DA, & Li E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell,* 99(3), 247-257. doi:10.1016/s0092-8674(00)81656-6

Okano M, Xie S, & Li E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet,* 19(3), 219-220. doi:10.1038/890

Oki Y, Aoki E, & Issa JP. (2007). Decitabine--bedside to bench. *Crit Rev Oncol Hematol,* 61(2), 140-152. doi:10.1016/j.critrevonc.2006.07.010

Olova N, Krueger F, Andrews S, Oxley D, Berrens RV, Branco MR, & Reik W. (2018). Comparison of whole-genome bisulfite sequencing library preparation strategies identifies sources of biases affecting DNA methylation data. *Genome Biol,* 19(1), 33. doi:10.1186/s13059-018-1408-2

Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, & Hayashi Y. (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). *Cancer Res,* 62(14), 4075-4080.

Ooi SK, & Bestor TH. (2008). The colorful history of active DNA demethylation. *Cell,* 133(7), 1145-1148. doi:10.1016/j.cell.2008.06.009

Öz S, Raddatz G, Rius M, Blagitko-Dorfs N, Lübbert M, Maercker C, & Lyko F. (2014). Quantitative determination of decitabine incorporation into DNA and its effect on mutation rates in human cancer cells. *Nucleic Acids Res,* 42(19), e152. doi:10.1093/nar/gku775

Palii SS, Van Emburgh BO, Sankpal UT, Brown KD, & Robertson KD. (2008). DNA methylation inhibitor 5-Aza-2'-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B. *Mol Cell Biol,* 28(2), 752-771. doi:10.1128/mcb.01799-07

Pastor WA, Aravind L, & Rao A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol,* 14(6), 341-356. doi:10.1038/nrm3589

Peat JR, Dean W, Clark SJ, Krueger F, Smallwood SA, Ficz G, Kim JK, Marioni JC, Hore TA, & Reik W. (2014). Genome-wide bisulfite sequencing in zygotes identifies demethylation targets and maps the contribution of TET3 oxidation. *Cell Rep,* 9(6), 1990-2000. doi:10.1016/j.celrep.2014.11.034

Peat JR, Ortega-Recalde O, Kardailsky O, & Hore TA. (2017). The elephant shark methylome reveals conservation of epigenetic regulation across jawed vertebrates. *F1000Res,* 6, 526. doi:10.12688/f1000research.11281.1

Pennings S, Allan J, & Davey CS. (2005). DNA methylation, nucleosome formation and positioning. *Brief Funct Genomic Proteomic,* 3(4), 351-361. doi:10.1093/bfgp/3.4.351

Pfeifer GP, Szabo PE, & Song J. (2019). Protein Interactions at Oxidized 5-Methylcytosine Bases. *J Mol Biol*. doi:10.1016/j.jmb.2019.07.039

Pingoud A, Wilson GG, & Wende W. (2014). Type II restriction endonucleases--a historical perspective and more. *Nucleic Acids Res,* 42(12), 7489-7527. doi:10.1093/nar/gku447

Pleyer L, & Greil R. (2015). Digging deep into "dirty" drugs - modulation of the methylation machinery. *Drug Metab Rev,* 47(2), 252-279. doi:10.3109/03602532.2014.995379

Pohlmann P, DiLeone LP, Cancella AI, Caldas AP, Dal Lago L, Campos O, Jr., Monego E, Rivoire W, & Schwartsmann G. (2002). Phase II trial of cisplatin plus decitabine, a new DNA hypomethylating agent, in patients with advanced squamous cell carcinoma of the cervix. *Am J Clin Oncol,* 25(5), 496-501.

Prigge ST, Mains RE, Eipper BA, & Amzel LM. (2000). New insights into copper monooxygenases and peptide amidation: structure, mechanism and function. *Cell Mol Life Sci,* 57(8-9), 1236-1259.

Pronier E, & Delhommeau F. (2012). Role of TET2 mutations in myeloproliferative neoplasms. *Curr Hematol Malig Rep,* 7(1), 57-64. doi:10.1007/s11899-011-0108-8

Qin T, Youssef EM, Jelinek J, Chen R, Yang AS, Garcia-Manero G, & Issa JP. (2007). Effect of cytarabine and decitabine in combination in human leukemic cell lines. *Clin Cancer Res,* 13(14), 4225-4232. doi:10.1158/1078-0432.Ccr-06-2762

Quail MA, Swerdlow H, & Turner DJ. (2009). Improved protocols for the illumina genome analyzer sequencing system. *Curr Protoc Hum Genet,* Chapter 18, Unit 18.12. doi:10.1002/0471142905.hg1802s62

Quivoron C, Couronne L, Della Valle V, Lopez CK, Plo I, Wagner-Ballon O, Do Cruzeiro M, Delhommeau F, Arnulf B, Stern MH, Godley L, Opolon P, Tilly H, Solary E, Duffourd Y, Dessen P, Merle-Beral H, Nguyen-Khac F, Fontenay M, Vainchenker W, Bastard C, Mercher T, & Bernard OA. (2011). TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell,* 20(1), 25-38. doi:10.1016/j.ccr.2011.06.003

Rai K, Jafri IF, Chidester S, James SR, Karpf AR, Cairns BR, & Jones DA. (2010). Dnmt3 and G9a cooperate for tissue-specific development in zebrafish. *J Biol Chem,* 285(6), 4110-4121. doi:10.1074/jbc.M109.073676

Raiber EA, Beraldi D, Ficz G, Burgess HE, Branco MR, Murat P, Oxley D, Booth MJ, Reik W, & Balasubramanian S. (2012). Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. *Genome Biol,* 13(8), R69. doi:10.1186/gb-2012-13-8-r69

Raiber EA, Murat P, Chirgadze DY, Beraldi D, Luisi BF, & Balasubramanian S. (2015). 5-Formylcytosine alters the structure of the DNA double helix. *Nat Struct Mol Biol,* 22(1), 44-49. doi:10.1038/nsmb.2936

Raj K, John A, Ho A, Chronis C, Khan S, Samuel J, Pomplun S, Thomas NS, & Mufti GJ. (2007). CDKN2B methylation status and isolated chromosome 7 abnormalities predict responses to treatment with 5-azacytidine. *Leukemia,* 21(9), 1937-1944. doi:10.1038/sj.leu.2404796

Rasmussen KD, & Helin K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev,* 30(7), 733-750. doi:10.1101/gad.276568.115

Rebollo R, Karimi MM, Bilenky M, Gagnier L, Miceli-Royer K, Zhang Y, Goyal P, Keane TM, Jones S, Hirst M, Lorincz MC, & Mager DL. (2011). Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms. *PLoS Genet,* 7(9), e1002301. doi:10.1371/journal.pgen.1002301

Richardson KA, Vega TP, Richardson FC, Moore CL, Rohloff JC, Tomkinson B, Bendele RA, & Kuchta RD. (2004). Polymerization of the triphosphates of AraC, 2',2'-difluorodeoxycytidine (dFdC) and OSI-7836 (T-araC) by human DNA polymerase alpha and DNA primase. *Biochem Pharmacol,* 68(12), 2337-2346. doi:10.1016/j.bcp.2004.07.042

Riggs AD. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet,* 14(1), 9-25. doi:10.1159/000130315

Rodger EJ, Chatterjee A, & Morison IM. (2014). 5-hydroxymethylcytosine: a potential therapeutic target in cancer. *Epigenomics,* 6(5), 503-514. doi:10.2217/epi.14.39

Rodriguez-Ubreva FJ, Cariaga-Martinez AE, Cortes MA, Romero-De Pablos M, Ropero S, Lopez-Ruiz P, & Colas B. (2010). Knockdown of protein tyrosine phosphatase SHP-1 inhibits G1/S progression in prostate cancer cells through the regulation of components of the cell-cycle machinery. *Oncogene,* 29(3), 345-355. doi:10.1038/onc.2009.329

Rogstad DK, Herring JL, Theruvathu JA, Burdzy A, Perry CC, Neidigh JW, & Sowers LC. (2009). Chemical decomposition of 5-aza-2'-deoxycytidine (Decitabine): kinetic analyses and identification of products by NMR, HPLC, and mass spectrometry. *Chem Res Toxicol,* 22(6), 1194-1204. doi:10.1021/tx900131u

Rose CM, van den Driesche S, Sharpe RM, Meehan RR, & Drake AJ. (2014). Dynamic changes in DNA modification states during late gestation male germ line development in the rat. *Epigenetics & Chromatin,* 7(1), 19. doi:10.1186/1756-8935-7-19

Rose NR, McDonough MA, King ON, Kawamura A, & Schofield CJ. (2011). Inhibition of 2-oxoglutarate dependent oxygenases. *Chem Soc Rev,* 40(8), 4364-4397. doi:10.1039/c0cs00203h

Rouleau J, Tanigawa G, & Szyf M. (1992). The mouse DNA methyltransferase 5'-region. A unique housekeeping gene promoter. *J Biol Chem,* 267(11), 7368-7377.

Sajadian SO, Ehnert S, Vakilian H, Koutsouraki E, Damm G, Seehofer D, Thasler W, Dooley S, Baharvand H, Sipos B, & Nussler AK. (2015a). Induction of active demethylation and 5hmC formation by 5-azacytidine is TET2 dependent and suggests new treatment strategies against hepatocellular carcinoma. *Clin Epigenetics,* 7, 98. doi:10.1186/s13148-015-0133-x

Sajadian SO, Ehnert S, Vakilian H, Koutsouraki E, Damm G, Seehofer D, Thasler W, Dooley S, Baharvand H, Sipos B, & Nussler AK. (2015b). Induction of active demethylation and 5hmC formation by 5-azacytidine is TET2 dependent and suggests new treatment

strategies against hepatocellular carcinoma. *Clin Epigenetics,* 7(1), 98. doi:10.1186/s13148-015-0133-x

Sajadian SO, Tripura C, Samani FS, Ruoss M, Dooley S, Baharvand H, & Nussler AK. (2016a). Vitamin C enhances epigenetic modifications induced by 5-azacytidine and cell cycle arrest in the hepatocellular carcinoma cell lines HLE and Huh7. *Clinical Epigenetics,* 8(1), 46. doi:10.1186/s13148-016-0213-6

Sajadian SO, Tripura C, Samani FS, Ruoss M, Dooley S, Baharvand H, & Nussler AK. (2016b). Vitamin C enhances epigenetic modifications induced by 5-azacytidine and cell cycle arrest in the hepatocellular carcinoma cell lines HLE and Huh7. *Clin Epigenetics,* 8, 46. doi:10.1186/s13148-016-0213-6

SanMiguel JM, & Bartolomei MS. (2018). DNA methylation dynamics of genomic imprinting in mouse development. *Biol Reprod,* 99(1), 252-262. doi:10.1093/biolre/ioy036

Sant DW, Mustafi S, Gustafson CB, Chen J, Slingerland JM, & Wang G. (2018). Vitamin C promotes apoptosis in breast cancer cells by increasing TRAIL expression. *Sci Rep,* 8(1), 5306. doi:10.1038/s41598-018-23714-7

Santi DV, Norment A, & Garrett CE. (1984). Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. *Proc Natl Acad Sci U S A,* 81(22), 6993-6997. doi:10.1073/pnas.81.22.6993

Santiago M, Antunes C, Guedes M, Sousa N, & Marques CJ. (2014). TET enzymes and DNA hydroxymethylation in neural development and function — How critical are they? *Genomics,* 104(5), 334-340. doi:https://doi.org/10.1016/j.ygeno.2014.08.018

Schorah CJ, Wild J, Hartley R, Sheppard S, & Smithells RW. (1983). The effect of periconceptional supplementation on blood vitamin concentrations in women at recurrence risk for neural tube defect. *Br J Nutr,* 49(2), 203-211. doi:10.1079/bjn19830026

Schwartsmann G, Schunemann H, Gorini CN, Filho AF, Garbino C, Sabini G, Muse I, DiLeone L, & Mans DR. (2000). A phase I trial of cisplatin plus decitabine, a new DNA-hypomethylating agent, in patients with advanced solid tumors and a follow-up early phase II evaluation in patients with inoperable non-small cell lung cancer. *Invest New Drugs,* 18(1), 83-91.

Scourzic L, Mouly E, & Bernard OA. (2015). TET proteins and the control of cytosine demethylation in cancer. *Genome Med,* 7(1), 9. doi:10.1186/s13073-015-0134-6

Seisenberger S, Peat JR, & Reik W. (2013). Conceptual links between DNA methylation reprogramming in the early embryo and primordial germ cells. *Curr Opin Cell Biol,* 25(3), 281-288. doi:10.1016/j.ceb.2013.02.013

Seymour JF, Fenaux P, Silverman LR, Mufti GJ, Hellstrom-Lindberg E, Santini V, List AF, Gore SD, Backstrom J, McKenzie D, & Beach CL. (2010). Effects of azacitidine compared with conventional care regimens in elderly (>/= 75 years) patients with higher-risk myelodysplastic syndromes. *Crit Rev Oncol Hematol,* 76(3), 218-227. doi:10.1016/j.critrevonc.2010.04.005

Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M, & Koseki H. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature,* 450(7171), 908-912. doi:10.1038/nature06397

Sharma H. (2018). Development of Novel Therapeutics Targeting Isocitrate Dehydrogenase Mutations in Cancer. *Curr Top Med Chem,* 18(6), 505-524. doi:10.2174/1568026618666180518091144

Sharp AJ, Stathaki E, Migliavacca E, Brahmachary M, Montgomery SB, Dupre Y, & Antonarakis SE. (2011). DNA methylation profiles of human active and inactive X chromosomes. *Genome Res,* 21(10), 1592-1600. doi:10.1101/gr.112680.110

Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, Zhang K, & Zhang Y. (2013). Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell,* 153(3), 692-706. doi:10.1016/j.cell.2013.04.002

Shen L, & Zhang Y. (2013). 5-Hydroxymethylcytosine: generation, fate, and genomic distribution. *Curr Opin Cell Biol,* 25(3), 289-296. doi:10.1016/j.ceb.2013.02.017

Shi DQ, Ali I, Tang J, & Yang WC. (2017). New Insights into 5hmC DNA Modification: Generation, Distribution and Function. *Front Genet,* 8, 100. doi:10.3389/fgene.2017.00100

Shin DY, Park YS, Yang K, Kim GY, Kim WJ, Han MH, Kang HS, & Choi YH. (2012). Decitabine, a DNA methyltransferase inhibitor, induces apoptosis in human leukemia cells through intracellular reactive oxygen species generation. *Int J Oncol,* 41(3), 910-918. doi:10.3892/ijo.2012.1546

Short NJ, Kantarjian HM, Loghavi S, Huang X, Qiao W, Borthakur G, Kadia TM, Daver N, Ohanian M, Dinardo CD, Estrov Z, Kanagal-Shamanna R, Maiti A, Benton CB, Bose P, Alvarado Y, Jabbour E, Kornblau SM, Pemmaraju N, Jain N, Gasior Y, Richie MA, Pierce S, Cortes J, Konopleva M, Garcia-Manero G, & Ravandi F. (2019). Treatment with a 5-day versus a 10-day schedule of decitabine in older patients with newly diagnosed acute myeloid leukaemia: a randomised phase 2 trial. *Lancet Haematol,* 6(1), e29-e37. doi:10.1016/s2352-3026(18)30182-0

Smallwood SA, & Kelsey G. (2012). De novo DNA methylation: a germ cell perspective. *Trends Genet,* 28(1), 33-42. doi:10.1016/j.tig.2011.09.004

Smiley JA, Kundracik M, Landfried DA, Barnes VR, Sr., & Axhemi AA. (2005). Genes of the thymidine salvage pathway: thymine-7-hydroxylase from a Rhodotorula glutinis cDNA library and iso-orotate decarboxylase from Neurospora crassa. *Biochim Biophys Acta,* 1723(1-3), 256-264. doi:10.1016/j.bbagen.2005.02.001

Solary E, Bernard OA, Tefferi A, Fuks F, & Vainchenker W. (2014). The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. *Leukemia,* 28(3), 485-496. doi:10.1038/leu.2013.337

Song CX, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, Street C, Li Y, Poidevin M, Wu H, Gao J, Liu P, Li L, Xu GL, Jin P, & He C. (2013). Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell,* 153(3), 678-691. doi:10.1016/j.cell.2013.04.001

Spans L, Van den Broeck T, Smeets E, Prekovic S, Thienpont B, Lambrechts D, Karnes RJ, Erho N, Alshalalfa M, Davicioni E, Helsen C, Gevaert T, Tosco L, Haustermans K, Lerut E, Joniau S, & Claessens F. (2016). Genomic and epigenomic analysis of high-risk prostate cancer reveals changes in hydroxymethylation and TET1. *Oncotarget,* 7(17), 24326-24338. doi:10.18632/oncotarget.8220

Stadtfeld M, Apostolou E, Ferrari F, Choi J, Walsh RM, Chen T, Ooi SS, Kim SY, Bestor TH, Shioda T, Park PJ, & Hochedlinger K. (2012). Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. *Nat Genet,* 44(4), 398-405, s391-392. doi:10.1038/ng.1110

Stancheva I, El-Maarri O, Walter J, Niveleau A, & Meehan RR. (2002). DNA methylation at promoter regions regulates the timing of gene activation in Xenopus laevis embryos. *Dev Biol,* 243(1), 155-165. doi:10.1006/dbio.2001.0560

Stancheva I, & Meehan RR. (2000). Transient depletion of xDnmt1 leads to premature gene activation in Xenopus embryos. *Genes Dev,* 14(3), 313-327.

Szwagierczak A, Bultmann S, Schmidt CS, Spada F, & Leonhardt H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res,* 38(19), e181. doi:10.1093/nar/gkq684

Szyf M. (2005). DNA methylation and demethylation as targets for anticancer therapy. *Biochemistry (Mosc),* 70(5), 533-549. doi:10.1007/s10541-005-0147-7

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, & Rao A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* 324(5929), 930-935. doi:10.1126/science.1170116

Taylor KH, Pena-Hernandez KE, Davis JW, Arthur GL, Duff DJ, Shi H, Rahmatpanah FB, Sjahputera O, & Caldwell CW. (2007). Large-scale CpG methylation analysis identifies novel candidate genes and reveals methylation hotspots in acute lymphoblastic leukemia. *Cancer Res,* 67(6), 2617-2625. doi:10.1158/0008-5472.Can-06-3993

Thienpont B, Galle E, & Lambrechts D. (2016). TET enzymes as oxygen-dependent tumor suppressors: exciting new avenues for cancer management. *Epigenomics,* 8(11), 1445-1448. doi:10.2217/epi-2016-0126

Thienpont B, Steinbacher J, Zhao H, D'Anna F, Kuchnio A, Ploumakis A, Ghesquière B, Van Dyck L, Boeckx B, Schoonjans L, Hermans E, Amant F, Kristensen VN, Peng Koh K, Mazzone M, Coleman M, Carell T, Carmeliet P, & Lambrechts D. (2016). Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature,* 537(7618), 63-68. doi:10.1038/nature19081

Thomassin H, Flavin M, Espinas ML, & Grange T. (2001). Glucocorticoid-induced DNA demethylation and gene memory during development. *Embo j,* 20(8), 1974-1983. doi:10.1093/emboj/20.8.1974

Tsai HC, Li H, Van Neste L, Cai Y, Robert C, Rassool FV, Shin JJ, Harbom KM, Beaty R, Pappou E, Harris J, Yen RW, Ahuja N, Brock MV, Stearns V, Feller-Kopman D, Yarmus LB, Lin YC, Welm AL, Issa JP, Minn I, Matsui W, Jang YY, Sharkis SJ, Baylin SB, & Zahnow CA. (2012). Transient low doses of DNA-demethylating agents exert durable antitumor effects on hematological and epithelial tumor cells. *Cancer Cell,* 21(3), 430-446. doi:10.1016/j.ccr.2011.12.029

Tsukada Y, Akiyama T, & Nakayama KI. (2015). Maternal TET3 is dispensable for embryonic development but is required for neonatal growth. *Sci Rep,* 5, 15876. doi:10.1038/srep15876

Ubhi T, & Brown GW. (2019). Exploiting DNA Replication Stress for Cancer Treatment. *Cancer Res,* 79(8), 1730-1739. doi:10.1158/0008-5472.Can-18-3631

Uysal F, Ozturk S, & Akkoyunlu G. (2017). DNMT1, DNMT3A and DNMT3B proteins are differently expressed in mouse oocytes and early embryos. *J Mol Histol,* 48(5-6), 417-426. doi:10.1007/s10735-017-9739-y

Valinluck V, & Sowers LC. (2007). Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res,* 67(3), 946-950. doi:10.1158/0008-5472.Can-06-3123

Van Robertson WB, & Schwartz B. (1953). Ascorbic acid and the formation of collagen. *J Biol Chem,* 201(2), 689-696.

Vasta JD, & Raines RT. (2016). Human Collagen Prolyl 4-Hydroxylase Is Activated by Ligands for Its Iron Center. *Biochemistry,* 55(23), 3224-3233. doi:10.1021/acs.biochem.6b00251

Vető B, Szabó P, Bacquet C, Apró A, Hathy E, Kiss J, Réthelyi JM, Szeri F, Szüts D, & Arányi T. (2018). Inhibition of DNA methyltransferase leads to increased genomic 5-hydroxymethylcytosine levels in hematopoietic cells. *FEBS Open Bio,* 8(4), 584-592. doi:10.1002/2211-5463.12392

Vogler WR, Miller DS, & Keller JW. (1976). 5-Azacytidine (NSC 102816): a new drug for the treatment of myeloblastic leukemia. *Blood,* 48(3), 331-337.

von Meyenn F, & Reik W. (2015). Forget the Parents: Epigenetic Reprogramming in Human Germ Cells. *Cell,* 161(6), 1248-1251. doi:https://doi.org/10.1016/j.cell.2015.05.039

Waddington CH. (1942). The epigenotype. *Endeavour,* 1, 18-20.

Walsh CP, Chaillet JR, & Bestor TH. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet,* 20(2), 116-117. doi:10.1038/2413

Wang L, Zhang J, Duan J, Gao X, Zhu W, Lu X, Yang L, Zhang J, Li G, Ci W, Li W, Zhou Q, Aluru N, Tang F, He C, Huang X, & Liu J. (2014). Programming and inheritance of parental DNA methylomes in mammals. *Cell,* 157(4), 979-991. doi:10.1016/j.cell.2014.04.017

Wang P, Yan Y, Yu W, & Zhang H. (2019). Role of ten-eleven translocation proteins and 5-hydroxymethylcytosine in hepatocellular carcinoma. *Cell Proliferation,* 52(4), e12626. doi:10.1111/cpr.12626

Wang RY, Gehrke CW, & Ehrlich M. (1980). Comparison of bisulfite modification of 5-methyldeoxycytidine and deoxycytidine residues. *Nucleic Acids Res,* 8(20), 4777-4790. doi:10.1093/nar/8.20.4777

Weber AR, Krawczyk C, Robertson AB, Kusnierczyk A, Vagbo CB, Schuermann D, Klungland A, & Schar P. (2016). Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nat Commun,* 7, 10806. doi:10.1038/ncomms10806

Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, & Schubeler D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet,* 39(4), 457-466. doi:10.1038/ng1990

Wen X, Pique-Regi R, & Luca F. (2017). Integrating molecular QTL data into genome-wide genetic association analysis: Probabilistic assessment of enrichment and colocalization. *PLoS Genet,* 13(3), e1006646. doi:10.1371/journal.pgen.1006646

Wild L, & Flanagan JM. (2010). Genome-wide hypomethylation in cancer may be a passive consequence of transformation. *Biochim Biophys Acta,* 1806(1), 50-57. doi:10.1016/j.bbcan.2010.03.003

Wilson JX. (2005). Regulation of vitamin C transport. *Annu Rev Nutr,* 25, 105-125. doi:10.1146/annurev.nutr.25.050304.092647

Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W, & Walter J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun,* 2, 241. doi:10.1038/ncomms1240

Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, Sun YE, & Zhang Y. (2011). Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev,* 25(7), 679-684. doi:10.1101/gad.2036011

Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Sun YE, & Zhang Y. (2011). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature,* 473(7347), 389-393. doi:10.1038/nature09934

Wu H, & Zhang Y. (2011a). Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev,* 25(23), 2436-2452. doi:10.1101/gad.179184.111

Wu H, & Zhang Y. (2011b). Tet1 and 5-hydroxymethylation: a genome-wide view in mouse embryonic stem cells. *Cell Cycle,* 10(15), 2428-2436. doi:10.4161/cc.10.15.16930

Wu H, & Zhang Y. (2014). Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell,* 156(1-2), 45-68. doi:10.1016/j.cell.2013.12.019

Wu Q, Lu S, Wang L, Hu J, Qiao F, Qiu X, Zhao C, Lao Y, Song Y, & Fan H. (2012). DNMT3A rs36012910 A>G polymorphism and gastric cancer susceptibility in a Chinese population. *Mol Biol Rep,* 39(12), 10949-10955. doi:10.1007/s11033-012-1996-y

Wu SC, & Zhang Y. (2010). Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol,* 11(9), 607-620. doi:10.1038/nrm2950

Wu X, & Zhang Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet,* 18(9), 517-534. doi:10.1038/nrg.2017.33

Xiao M, Yang H, Xu W, Ma S, Lin H, Zhu H, Liu L, Liu Y, Yang C, Xu Y, Zhao S, Ye D, Xiong Y, & Guan KL. (2012). Inhibition of α-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev,* 26(12), 1326-1338. doi:10.1101/gad.191056.112

Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ, & Viegas-Pequignot E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature,* 402(6758), 187-191. doi:10.1038/46052

Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim SH, Ito S, Yang C, Wang P, Xiao MT, Liu LX, Jiang WQ, Liu J, Zhang JY, Wang B, Frye S, Zhang Y, Xu YH, Lei QY, Guan KL, Zhao SM, & Xiong Y. (2011a). Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell,* 19(1), 17-30. doi:10.1016/j.ccr.2010.12.014

Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim SH, Ito S, Yang C, Wang P, Xiao MT, Liu LX, Jiang WQ, Liu J, Zhang JY, Wang B, Frye S, Zhang Y, Xu YH, Lei QY, Guan KL, Zhao SM, & Xiong Y. (2011b). Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer Cell,* 19(1), 17-30. doi:10.1016/j.ccr.2010.12.014

Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, Barbera AJ, Zheng L, Zhang H, Huang S, Min J, Nicholson T, Chen T, Xu G, Shi Y, Zhang K, & Shi YG. (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell,* 42(4), 451-464. doi:10.1016/j.molcel.2011.04.005

Yan XJ, Xu J, Gu ZH, Pan CM, Lu G, Shen Y, Shi JY, Zhu YM, Tang L, Zhang XW, Liang WX, Mi JQ, Song HD, Li KQ, Chen Z, & Chen SJ. (2011). Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet,* 43(4), 309-315. doi:10.1038/ng.788

Yang AS, Estécio MR, Doshi K, Kondo Y, Tajara EH, & Issa JP. (2004). A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res,* 32(3), e38. doi:10.1093/nar/gnh032

Yang H, Liu Y, Bai F, Zhang JY, Ma SH, Liu J, Xu ZD, Zhu HG, Ling ZQ, Ye D, Guan KL, & Xiong Y. (2013). Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. *Oncogene,* 32(5), 663-669. doi:10.1038/onc.2012.67

Yin R, Mao SQ, Zhao B, Chong Z, Yang Y, Zhao C, Zhang D, Huang H, Gao J, Li Z, Jiao Y, Li C, Liu S, Wu D, Gu W, Yang YG, Xu GL, & Wang H. (2013). Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc,* 135(28), 10396-10403. doi:10.1021/ja4028346

Yin R, Mo J, Dai J, & Wang H. (2017). Nickel(II) Inhibits Tet-Mediated 5-Methylcytosine Oxidation by High Affinity Displacement of the Cofactor Iron(II). *ACS Chemical Biology,* 12(6), 1494-1498. doi:10.1021/acschembio.7b00261

Young JI, Zuchner S, & Wang G. (2015). Regulation of the Epigenome by Vitamin C. *Annu Rev Nutr,* 35, 545-564. doi:10.1146/annurev-nutr-071714-034228

Zeng Y, & Chen T. (2019). DNA Methylation Reprogramming during Mammalian Development. *Genes (Basel),* 10(4). doi:10.3390/genes10040257

Zhang J, Kobert K, Flouri T, & Stamatakis A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics,* 30(5), 614-620. doi:10.1093/bioinformatics/btt593

Zhang L, Lu X, Lu J, Liang H, Dai Q, Xu GL, Luo C, Jiang H, & He C. (2012). Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat Chem Biol,* 8(4), 328-330. doi:10.1038/nchembio.914

Zhao H, Zhu H, Huang J, Zhu Y, Hong M, Zhu H, Zhang J, Li S, Yang L, Lian Y, Wang S, Mao J, Chen Y, Li J, & Qian S. (2018). The synergy of Vitamin C with decitabine activates TET2 in leukemic cells and significantly improves overall survival in elderly patients with acute myeloid leukemia. *Leuk Res,* 66, 1-7. doi:10.1016/j.leukres.2017.12.009