Investigating a personalised nutrition approach for modulating epigenetic profiles using over-expressed DNMTs in cell lines

Chanachai Sae-Lee, BSc. (1st Hons), MSc.

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Human Nutrition Research Centre

Population Health Sciences

June 2020

Abstract

DNA methylation is an epigenetic mechanism that enables heritable changes in gene expression without changes in DNA sequence. Methyl groups are transferred from the methyl donor *S*-adenosyl methionine (SAM) to the 5-carbon of cytosine by DNA methyltransferases (DNMTs).

The DNMT family comprises a set of DNA-modifying enzymes and uses a similar catalytic mechanism to form a covalent reaction intermediate between the substrate base and the enzyme. Food-derived bioactive compounds are among the exogenous factors that can modulate the DNA methylation patterns, either via generating SAM through one-carbon metabolism or by inhibiting the activity of DNMTs.

In this study, cell lines with stable over-expression of each of 13 DNMTs isoform (DNMT3A1, DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT3L, and DNMT1) were generated by lentiviral transduction of human embryonic kidney cells (HEK293T). DNA methylation patterns in these 13 cell lines were analysed by Illumina Infinium Methylation EPIC BeadChip, which interrogates more than 850,000 CpG sites across the genome. The sensitivity and specificity of each DNMT isoform to selected food constituents (caffeic acid (CA), (-)-Epigallocatechin-3-gallate (EGCG), curcumin, vitamin C, and theaflavin) were investigated by quantification of DNA methylation at specific CpG sites targeted by the DNMTs, using pyrosequencing. DNA methylation patterns for each DNMT isoform were obtained and the potential underlying biological mechanisms for DNMT-target CpGs were explored.

At the selected CpG sites, DNA methylation was decreased with CA and vitamin C in DNMTΔ3B4 and DNMT3A2 cells, respectively. In addition, the enzymatic activity of DNMTΔ3B4 decreased after CA treatment. In summary, despite similarity of their protein structures, DNMT isoforms show regional specificity in the maintenance of DNA methylation patterns. This study also revealed that the activity of DNMTΔ3B4 and DNMT3A2 can be specifically modulated by CA and vitamin C, respectively, in a dose-response manner. These observations further understanding of nutrition-epigenetic mechanisms, especially interactions with enzymatic activity, could be applied to modulate DNA methylation profiles using food-derived bioactive compounds in personalised nutrition.

Acknowledgements

First of all, I would like to thank my two supervisors Dr Hyang-Min Byun and Professor John C. Mathers, who have been an enormous help and guided me throughout my long PhD journey at Newcastle University. Both have been a great supporter and source of encouragement. I would also like to thank all the people involved in this project: Dr Choi, Dr Viktor, Dr Akin, Dr Elena, and all members in Dr Viktor's laboratory. I would like to thank all my colleagues, and now good friends: Julien, Stella, Fiona, Khalil, Dan, and AJ.

I would also like to thank my family and Siriraj Hospital, who supported my education. Massive thank to my mom who make me strong and curious about life and this world.

Table of Contents

List of Figures

List of Tables

Abbreviation

450K	Illumina HumanMethylation450 BeadChip
5azaC	5-azacytidine
5azadC	5-aza-2'-deoxycytidine
5-caC	5-carboxylcytosine
5-fC	5-formylcytosine
5-hmC	Hydroxymethylcytosine
5-mC	5-methylcytosine
ACTB	Actin Beta
ADD	ATRX-DNMT3-DNMT3L
ADIPOQ	Adiponectin
AdoHcy	Adenosylhomocysteine
AdoMet	Adenosylmethionine
ALL	Acute lymphoblastic leukaemia
AM-AR	Active modification-active removal
AML	Acute myeloid leukaemia
AM-PD	Active modification-passive dilution
B2M	Beta-2-Microglobulin
BAH1/2	Bromo-adjacent homology
BF	Bright field
CA	Caffeic acid
CDH11	Cadherin 11
CGIs	CpG islands
CML	Chronic myeloid leukaemia
CMML	Chronic myelomonocytic leukaemia
CVD	Cardiovascular disease
DEPC	Diethyl pyrocarbonate
DMAP	DNMT1-associated protein
DMAP1	DNA methyltransferase 1-associated protein 1

DMPs	Differentially Methylated Position
DNMT	DNA methyltransferase
DNMTi	DNMT inhibitors
DW	Deionised water
EGCG	(-)-Epigallocatechin-3-gallate
EPIC	Illumina HumanMethylationEPIC BeadChip
EYA4	EYA Transcriptional Coactivator And Phosphatase 4
FACS	Fluorescence-activated cell sorting
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GFP	Green fluorescent protein
GUSB	Glucuronidase Beta
HATs	Histone acetyltransferases
HCT116	Human colorectal carcinoma cell line
HDACs	Histone deacetylases
HEK293FT	Human embryonic kidney 293FT cell line
HEK293T	Human embryonic kidney 293T cell line
HK	Housekeeping gene
HOXA11	Homeobox A11
HPRT1	Hypoxanthine Phosphoribosyltransferase 1
IC_{50}	The half-maximal inhibitory concentration
ICF	Immunodeficiency, chromosome abnormalities and facial anomalies
IGF2AS	IGF2 Antisense RNA
IL6	Interleukin 6
IncRNAs	Long non-coding RNAs
IPA	Ingenuity Pathway Analysis
KLF13	Kruppel-like Factor-13
LB	Lysogeny broth
LINE-1	Long interspersed nuclear element
MDS	Myelodysplastic syndrome
MEG-01	Megakaryoblast cell line

miRNAs	MicroRNAs
MTFHR	Methylene tetrahydrofolate reductase
ncRNAs	Non-coding RNAs
PBMCs	Peripheral blood mononuclear cells
PHD	Plant homology domain
piRNAs	Piwi-interacting RNAs
PN	Personalised nutrition
PPIA	Peptidylprolyl Isomerase A
PTCH1	Patched 1
PWWP	Proline-tryptophan-tryptophan-proline
qPCR	Quantitative polymerase chain reaction
RA	Rheumatoid arthritis
RFLP	Restriction fragment length polymorphism
RFTS	Replication foci targeting sequence
RPL13A	Ribosomal Protein L13A
RPLP0	Ribosomal Protein Lateral Stalk Subunit P0
RPS13	Ribosomal Protein S13
RT	Reverse transcription
SAH	S-adenosylhomocysteine
SAM	S-adenosyl methionine
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
siRNAs	Small interfering RNAs
SLE	Systemic lupus erythematosus
TDG	Thymine DNA glycosylase
TET	Ten-eleven translocation
TFRC	Transferrin Receptor
TRD	Target recognition domain

Chapter 1: General Introduction

1.1 Overview of epigenetics

Epigenetics is a mechanism of modifications in gene expression without changing DNA sequences (1). Examples of epigenetic mechanisms are DNA methylations, RNA modifications, and histone modifications (Figure 1.1) (2). DNA methylations, particularly on cytosine residue that lie next to a guanine base, are commonly related to repressing gene expression. Histone modifications, however, can lead to either the activation or deactivation of gene expression. For example, di-methylation of histone3 (H3) at lysine 4 (H3K4) and acetylation of H3K9 play an active role in transcription, however tri-methylation of H3K9 (H3K9me3) and H3K27me3 are inactive markers. Each of these epigenetic markers may induce independent modifications, but commonly interplay reciprocally in controlling gene expression during stem cell differentiation, embryonic development (3), and carcinogenesis (4). Those epigenetic markers are inherited between generations, but also the level of epigenetic markers at each successive generation is modifiable by internal and external stimuli, such as pro-inflammatory foods or antioxidant-rich diets. Later in this chapter, DNA methylations, which are amongst the well-studied epigenetic mechanisms, will be discussed with differential expression of DNA methyltransferases (DNMTs) and modulating its activity with dietary factors.

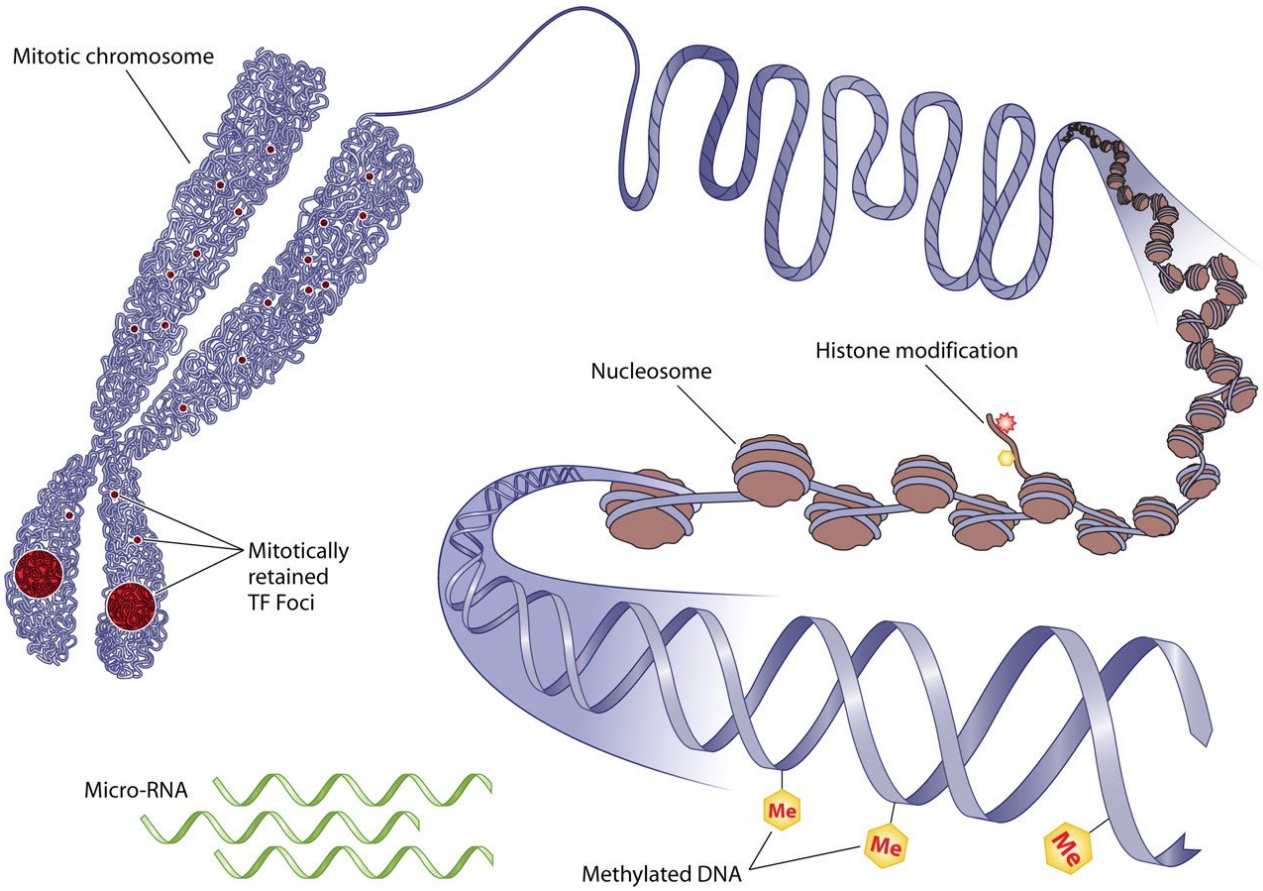


Figure 1.1 Epigenetic machinery. DNA methylations, histone tail modifications, and non-coding RNAs (ncRNAs). Transcriptional factor, TF; Methylation, Me. (taken from Sayyed K.Z. *et al.*) (2).

Histone modification

Lysine (K) and arginine (R) residues are the main amino acids in histone tails that are subject to post-translational modifications such as acetylation and methylation. Histone octamers are wrapped by DNA and each contains two copies of the histone variants: H2A, H2B, H3, and H4. There are numerous histone methyltransferases and demethylases, including histone acetyltransferases (HAT) and histone deacetylases (HDACs) (5). Which particular lysine residues on each of the histones are methylated under particular circumstances is determined by the actions of epigenetic writers (methyltransferases) and erasers (demethylases) (5). Histone H3 methylation at lysine 9 and lysine 27 (H3K9 and H3K27), as well as Histone H4 methylation at lysine 20 (H4K20) are well known markers of inactivation, since, these are associated with heterochromatin formation; in contrast, H3K36 and H3K4 methylation are known activation markers (6). Concomitant tri-methylations of H3K4 and H3K27 (bivalent domains) modulate lineage-specific gene expression from a poised state in embryonic stem cells (7). Also, a genome-wide study of histone methylation patterns showed that gene

expression was positively correlated with levels of H3K4 methylation and negatively correlated with H3K27 methylation (8).

RNA modifications

Non-coding RNAs are not translated into protein and they can be separated into regulatory non-coding RNAs and housekeeping non-coding RNAs. The regulatory role of RNA is mainly based on size; short chain non-coding RNAs (miRNAs, siRNAs, and piRNAs) and long non-coding RNAs (IncRNAs) (9-11). Non-coding RNAs play an important role in regulating gene expression (12, 13). Among non-coding RNAs, miRNAs act as post-transcriptional regulators via base-pairing with complementary sequences within messenger RNA (mRNA) targets resulting in mRNA degradation or translational repression (14) resulting in regulated gene expression. These noncoding RNAs generally work with components of DNA methylations and chromatin (DNMT, H3K27 methylation, histone deacetylase 4) (6), to sustain or establish silencing that contribute to transformation, tumour development, and tumour progression (15).

DNA methylation

DNA methylation may also inhibit the binding of transcription factors and, therefore, repress gene expression to modify cellular phenotype (16, 17). All the known DNMTs use *S*-adenosyl methionine (SAM), derived from methionine, as the methyl donor. Although epigenetic maintenance requires the interplay of many epigenetic components (*e.g.* chromatin and histone modifications) as well as proteins to control gene expression and cellular function, the availability of methyl donors for DNMTs and the activity of DNMT enzymes are critical to maintain normal DNA methylation patterns (18). Further detail of the key players in DNA methylations is discussed in the DNA methylations and DNMT sections (see section 1.2 and 1.4).

Patterns of histone modifications, DNA methylations, and expression of noncoding RNA act together to establish the epigenetic signatures that convey regulatory information and so control the phenotype by activating or suppressing of gene expression. Information given by epigenetic signatures may be considered to be as important as sequence information because these signatures are proven useful biomarkers to predict the pathological conditions and biological outcome (19). In contrast with genetic information, epigenetic information shows a certain degree of plasticity and it is inherently reversible (20).

1.2 Mechanisms of DNA methylation and demethylation

1.2.1 DNA methylation

DNA methylation is an epigenetic modification in which a DNMT catalyses the transfer of a methyl group from SAM to the 5-carbon of cytosine to form 5-methylcytosine (5-mC), and in which SAM is converted to *S*-adenosylhomocysteine (SAH) (Figure 1.2) (21).

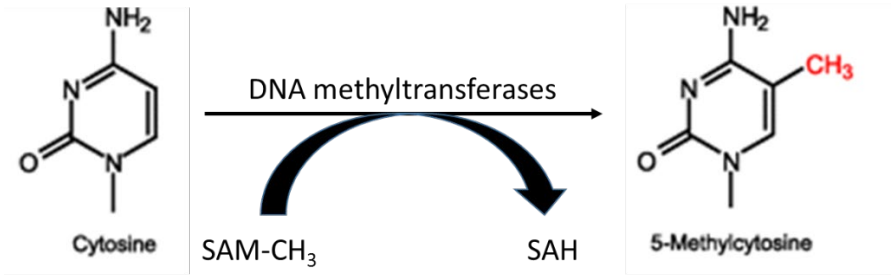


Figure 1.2 DNMTs catalyse the covalent addition of -CH_3 to 5-carbon of cytosine (22).

In mammals, the majority of DNA methylation occurs on cytosines preceding guanines (CpG dinucleotides) (23). Approximately $3x10^7$ residues of 5-methylcytosine (5-mC) are located within CpG dinucleotides 5'-m^5CG-3' in mammalian genomes (24, 25). Over 60% of human genes at promoters have CpG islands (CGIs) (26) which are short interspersed sequences. CGI is defined as a location with at least 200 bp, which have CG content greater than 50% and CpG ratio greater than 0.6 (27). Most CpGs in CGI regions are in a non-methylated state and permit transcription initiation, leading to stable activation of the associated promoter (28). The cytosines in CpG dinucleotides within CpG islands tend to be hypomethylated compared with those in non-island CpGs. The regions surrounding CGIs are known as "Shores" (North and South), followed by "Shelves" (North and South) and "Open sea" (Figure 1.3) (29). Transcription factors containing a CXXC zinc finger domain bind to unmethylated CpGs via a CXXC binding domain motif, contributing to generation of a transcriptionally competent chromatin configuration that prevents DNA methylation (30). Closer investigation of DNA methylation has revealed that exons show markedly higher methylation than introns (31).

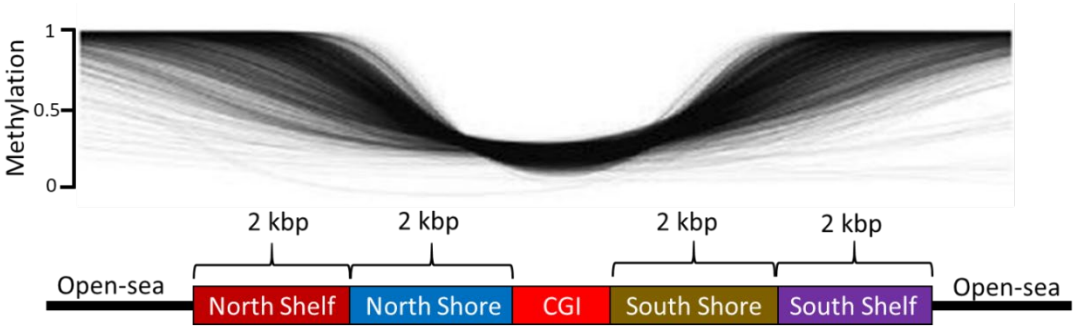


Figure 1.3 The distribution of DNA methylation levels by CpG density (adapted from Edgar R *et al.* (29)).

DNA methylation also functions to repress retrotransposons, long interspersed nuclear element (LINE-1) and *Alu* elements, which are highly methylated to keep them quiet in the human genome (32, 33). DNA methylation is also involved in genomic imprinting, silencing of repetitive elements and X chromosome inactivation (34).

DNA methylation has dual roles, both permissive and inhibitory, depending on the genomic location. DNA methylation of intragenic or distal regulatory elements with different degrees of CpG density are involved in the gene regulation (35). Gene body methylation have been found in *Arabidopsis* and this methylation was associated with active genes. In mammal, gene body methylation has been found in the active human X chromosome compared to the inactive X chromosome (36). The hypomethylated sites in the gene body have been correlated with low expression genes in cancer cell lines, also methylated of CpG-rich regions in gene bodies have been correlated with higher gene expression in human B cells (37). DNA methylation in a promoter is also an important epigenetic mechanism for the suppression of gene expression especially of tumour suppressor genes at early stage of carcinogenesis and occurs with various frequencies (38).

Enhancers are important roles in gene regulation as the enhancer-promoter interaction can increase the gene expression of downstream (39). The alteration of enhancer methylation plays an important role in cancer progression (40). Additionally, enhancer methylation accounts for a significant portion of intertumour expression heterogeneity (41). Moreover, enhancer methylation can be used to distinct breast cancer lineages (42).

Gene regulation by methylation

All housekeeping genes contain CpG islands in the proximal transcription start sites. These genes are frequently correlated with CpG islands in intronic regions downstream of the first exon, which tends to be smaller than the promoter-related CpG islands (43) and are often methylated (44). In mammals, the methylation of region 2 at CpG island in the second intron

of insulin like growth factor 2 receptor gene is the imprinting signal, which maintains expression of the maternal allele (45). In cancer, aberrant methylation at the transcription start sites leads to the silencing of tumour suppressor genes, namely *p16, VHL*, and *Rb* (46-48). It has been postulated that tumour suppressor genes might be overexpressed during tumour development and triggered hypermethylation of downstream CpG islands. Next, the methylation would spread from these sites to the CpG islands at the transcription start sites, therefore silencing the tumour suppressor genes (49). Moreover, methylation is used to silence one X chromosome in female by showing low DNA methylation on inactive X-chromosome than active X at intragenic and intergenic regions for genes subject to X-chromosome inactivation, but not at genes, which escape from inactivation (50).

1.2.2 DNA demethylation

DNA demethylation can be accomplished passively (in the new copy of DNA after replication) or actively (in a replication-independent enzymatic process) (22). Active DNA demethylation refers to an enzymatic mechanism for removal of the methyl group from 5mC. However, passive DNA demethylation refers to the lack of DNA methylation maintenance, which could be due to the absence of DNMT1 or the presence of DNMT inhibitors during DNA replication (23). Ten-eleven translocation (TET) methylcytosine dioxygenases are a family of enzymes that catalyses the conversion of 5-mC to 5-hydroxymethylcytosine (5-hmC) through oxidation (Figure 1.4) (51, 52). 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC) are produced subsequently from 5-mC during further consecutive oxidation steps (53). 5-caC is present at extremely low levels in DNA and can be excised by either a base excision repair enzyme or by thymine DNA glycosylase (TDG) to regenerate unmodified cytosine (54). 5-fC and 5-caC are much less abundant in the mammalian genome than 5-hmC is because TDG is more efficient at selection and excision of 5-fc and 5-caC, while the conversion of 5-hmC to 5-fC and 5-caC is less efficient (55).

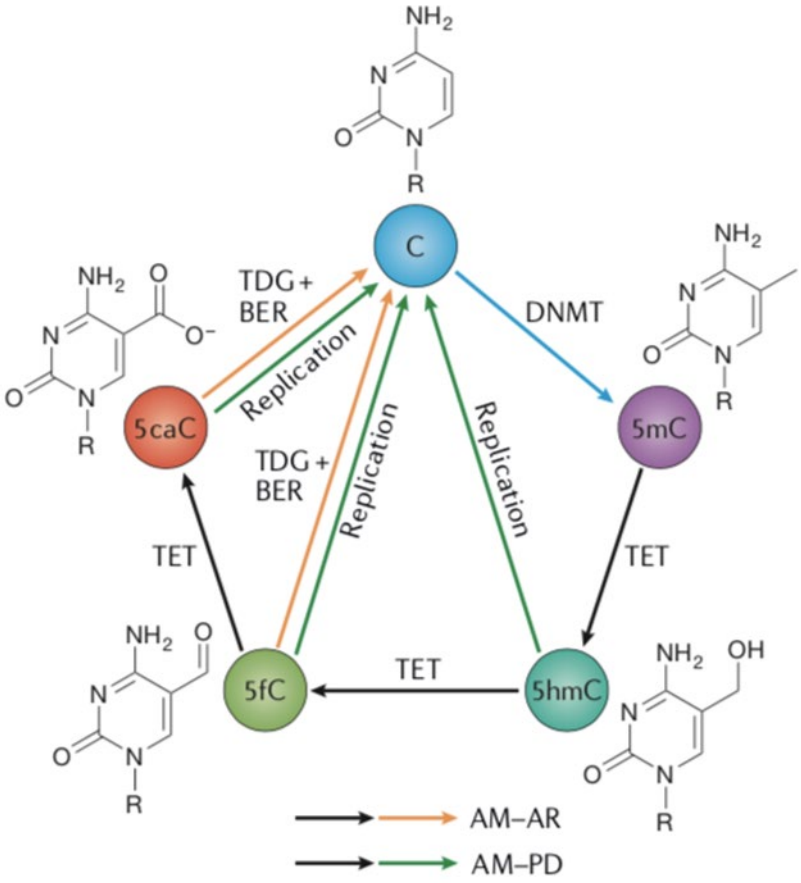


Figure 1.4 The cycle of active DNA demethylation and chemical structure of 5-methylcytosine (5-mC) and its oxidation products 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC). DNA methyltransferases (DNMTs) covert unmodified cytosine to 5-mC. 5-mC can be converted back to cytosine by ten-eleven translocations (TETs) to 5-hmC, 5-fC, and 5-caC followed by excision of 5-fC or 5-caC mediated by thymine DNA glycosylase (TDG) coupled with base excision repair (BER) (the process of active modification-active removal (AM-AR) or replication-dependent dilution of 5-hmC, 5-fC or 5-caC (the process of active modification-passive dilution (AM-PD) (taken from Xioaoji W. and Yi Z.) (55).

Demethylation occurs during early gametogenesis especially during development of primordial germ cells and also occurs postnatal (56). Aberrant demethylation has been observed in cancer. The active DNA demethylation is required for differentiation through modulation the enhancer's activities and promoters (55). Also, active DNA demethylation regulates cell reprogramming (57, 58) and cell fate transitions (59, 60).

1.3 DNA methylation in human diseases

When epigenetic information is not properly maintained or established, epigenetic abnormalities may arise. This section provides an insight into the pathological consequences of abnormalities due to DNA methylation.

Cancer

There are two types of epigenetic abnormalities: hypomethylation or hypermethylation. Hypermethylation in cancer has been found most often at CGIs (61), while very often hypomethylation has been observed in both highly and moderately repeated DNA sequences in cancers (62). Aberrant patterns of DNA methylation have been discovered in various cancer types, *e.g.* colon, liver, breast, prostate, ovarian, bone, bladder, and oesophageal cancers. Typically, cancer cells show hypomethylation of intergenic regions where there is normally high methyl-cytosine content (63). This may activate transposable elements contributing to the genome instability observed in cancer cells (64). Hypermethylation was observed within the promoter regions of CGI-associated tumour suppresser genes in cancer cells (65).

Autoimmune diseases

Autoimmune diseases are caused by over activity of the immune system leading to damage of own tissues. Genome-wide DNA methylation analysis of peripheral blood mononuclear cells (PBMCs) revealed that aberrant patterns of DNA methylation of human leukocyte antigen class II lead to high risk of developing rheumatoid arthritis (66). Hypermethylation of patched 1 gene (*PTCH1*), accompanied by low expression of patched 1 protein, activated the Hedgehog signalling pathway, resulting in increased secretion of tumour necrosis factor alpha and interleukin 6 (IL6) in the arthritic rat (67). Expression of *IL6* mRNA was significantly higher in systemic lupus erythematosus (SLE) patients than in healthy controls and this was associated with decreased DNA methylation of the *IL6* promoter in SLE patients (68). Moreover, DNA hypomethylation of LINE-1 was found in neutrophils from SLE patients (69).

Obesity

Epigenome-wide association studies showed the disturbances in DNA methylation of genes involved in lipid and lipoprotein metabolism, and inflammatory pathways are predictors of future development of type 2 diabetes, cardiovascular disease (CVD), and other adverse clinical consequences of obesity (70, 71). Genome-wide analysis has identified loci at which DNA methylation is altered in obesity (72). For example, DNA hypermethylation on adiponectin (*ADIPOQ*)-associated promoter suppressed adiponectin expression in adipocytes of obese patients (73). Moreover, CpG sites within ATF-motifs on hepatic glycolysis and insulin resistance genes were hypomethylated in both, non-diabetic and type 2 diabetic obese patients (74). DNA methylation profiling, which is a high throughput technology, has been

performed to discover novel genes and markers for obesity. The DNA methylation profiles from the blood of obese and normal individuals analysed using the Illumina HumanMethylation27 BeadChip kit showed the different pattern of DNA methylation on the promoters of *Tripartite Motif Containing3* (lower methylation levels in the obese cases) and *Ubiquitin Associated and SH3 Domain Containing A* (higher methylation levels in obese cases) (75). Additionally, the obesity-related cg07814318 methylation encoding Kruppel-like Factor-13 (*KLF13*) gene was found in the blood of childhood obesity (76). Aberrant DNA methylation patterns have been identified in the young obese generation compared to control individuals (77) and specific epigenetic marks have been identified in severely cases of child obesity (78). Recently, the study of DNA methylation between subcutaneous and omental adipose from obese individuals before and after gastric bypass showed 3,239 loci in subcutaneous and 7,722 in omental adipose that were significantly differentially methylated (79). DNA methylation on *SLC19A1* in obesity could be an epigenetic biomarker for obesity-related insulin resistance (80).

1.4 DNA methyltransferases

DNMT is responsible for the transfer of the methyl group to 5C-position of cytosine residues in DNA. DNMTs can either introduce new methylation marks or maintain them during genome replication (81). There are five mammalian DNMT families, which are DNMT1, DNMT2, DNMT3A, DNMT3B, and DNMT3L. DNMT3A, DNMT3B and DNMT1 are canonical cytosine-5 DNMTs, while DNMT2 and DNMT3L are non-canonical members as they are inactive forms. DNMT1 is responsible for the maintenance of DNA methylation patterns during DNA replication (Figure 1.5A), whereas DNMT3A and DNMT3B are known to be *de novo* methyltransferase which transfer the methyl groups onto unmethylated DNA elements (Figure 1.5B). Nevertheless, most DNMTs share similar domains with two functional parts: the *N*-terminal regulatory domain and the *C*-terminal catalytic domain (82). The *N*-terminal regulatory domain guides the nuclear translocation of the enzymes and mediates their interaction with the DNA and chromosome. The *C*-terminal catalytic part is conserved between eukaryotes and prokaryotes (83). Due to splicing and/or promotor usage, there are different isoforms of DNMT3A and DNMT3B (84, 85) (Figure 1.6).

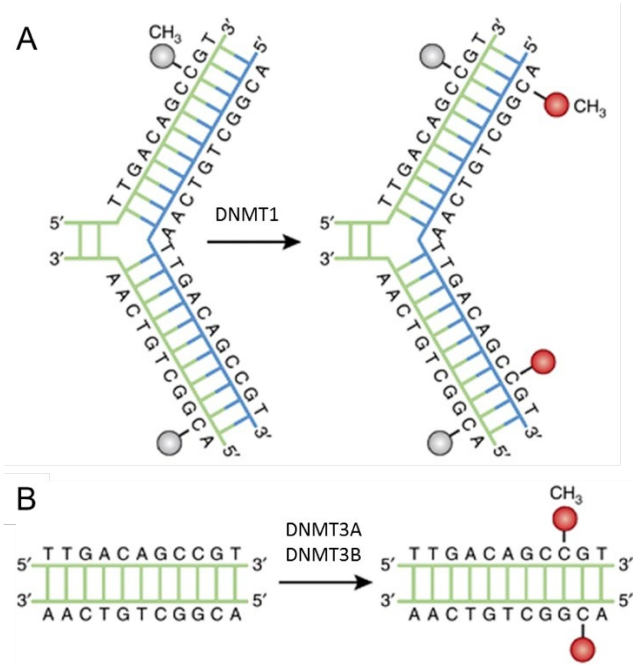


Figure 1.5 DNA methylation pathways. A) DNMT3A and DNMT3B are the *de novo* DNMTs which transfers -CH_3 onto naked DNA B) DNMT1 is the maintenance DNMT which maintains DNA methylation levels during DNA replication (Modified from Moore L.D. *et al.*) (86).

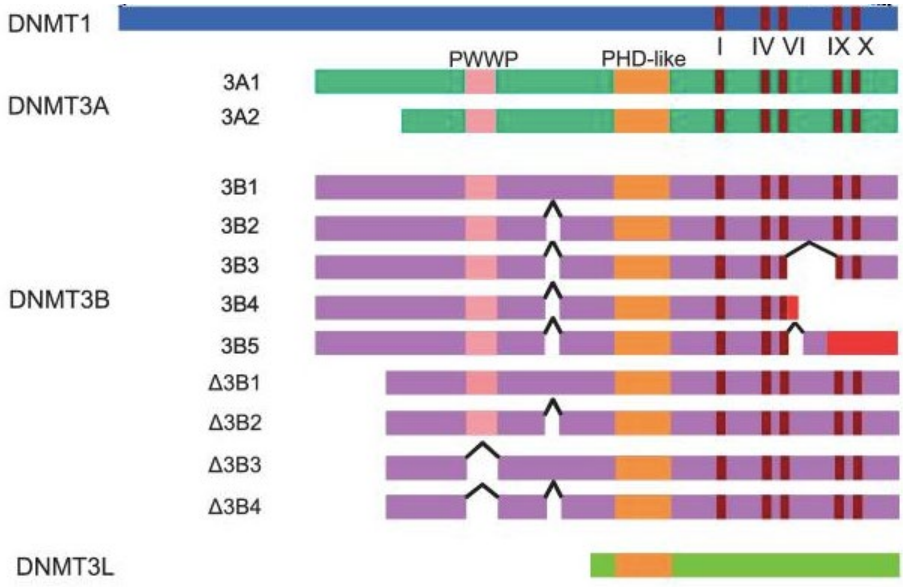


Figure 1.6 Structure of DNMT isoforms. Their proline-tryptophan-tryptophan-proline (PWWP), plant homology domain (PHD)-like and catalytic motifs (I, IV, VI, IX and X) and alternative splicing sites were shown in all DNMT isoforms (Taken from Choi SH. *et al.*) (87).

1.4.1 DNMT family members

DNMT1 and DNMT2

DNMT1 is a multimodular protein comprising 1,620 amino acids and it contains both the *N*-terminal regulatory domain and *C*-terminal catalytic domain. At *C*-terminal domain, there are two sub-domains: the target recognition domain (TRD) and the methyltransferase domain (88). The *N*-terminal regulatory domain contains a DNA methyltransferase 1-associated protein 1 (DMAP1) binding domain, a CXXC zing finger domain, a replication foci targeting sequence (RFTS) domain, and a tandem bromo-adjacent homology (BAH1/2) domain (89). Human and mouse DNMT1 share approximately 80% of their sequence identity (89). Loss of DNMT1 function leads to neurological abnormality and genome-wide demethylation (23, 90). Disruption of DNMT1 in the human colorectal carcinoma cell line (HCT116) causes loss of cell proliferation and led to cell death (91). Consistent with this, knockout of *DNMT1* leads to cell death and global hypomethylation (92). DNMT1 depletion in mice contributed to chromosome instability and tumour development (90, 93). Paradoxically, inhibition of DNMT1 by genetic knockdown or treatment with 5-aza 2' deoxycytidine has reduced tumour formation of colon cancer in a mouse model (94, 95). Loss of DNMT1 disrupts imprinted X chromosome inactivation and associated phenotypic abnormalities (96). Also, loss of methylation by depletion of *DNMT1* leads to demethylation of probes located in gene involved in neuroepithelial differentiation, fat homoeostasis/body mass, and olfactory receptor genes (97).

The human *DNMT2* located on chromosome 10p13 encodes a protein of 384 amino acids and it is homologous to yeast *pmt1* (98). DNMT2 consists only in a catalytic domain and it lacks the DNA methyltransferase activity. However, DNMT2 was found to methylate cytosine 38 of transfer RNA at anticodon loop (99). DNMT2-family enzymes affect the metazoan development and are involved in RNA methylation in the cytoplasm. Knockdown of *Dnmt2* resulted in defecting in differentiation of the retina, brain, and liver (100).

DNMT 3A and 3B

DNMT3A and *DNMT3B* are highly expressed in embryonic stem cells, early embryos, and developing germ cells, whereas low expression is observed in differentiated somatic cells (84, 101). DNMT3A encodes at least two known isoforms, which are DNMT3A1 (about 130 kDa) and DNMT3A2 (100 kDa) (Figure 1.6). DNMT3A2 lacks the *N*-terminal region and is the major isoform in embryonic stem cells (84). In humans, *DNMT3A2* is transcribed from an alternative promoter in intron 6 of *DNMT3A*. Unlike DNMT3A that was concentrated on

heterochromatin, DNMT3A2 showed a pattern suggestive of euchromatin association (84). Moreover, the specific *de novo* DNA methylation target sites of *DNMT3A1* are correlated with H3K4me3 modification (transcriptionally active) (87). *DNMT3A2* was detectable in breast/ovarian cancer cell lines and it also was expressed in testis, ovary, thymus and spleen (84). There are two subfamilies for DNMT3B; DNMT3B and DNMTΔ3B. DNMT3B has more than 30 isoforms as a result of alternative pre-mRNA splicing leading to various transcriptional variants (102-105). Only DNMT3B1 and DNMT3B2 are catalytically active, while DNMT3B3, DNMT3B4, and DNMT3B5 are not (Figure 1.6) (106). ΔDNMT3B lacks 200 amino acids from the *N*-terminal area of DNMT3B and there are seven isoforms; ΔDNMT3B1-ΔDNMT4 are catalytically active, while ΔDNMT3B5-ΔDNMT7 lack the catalytic domain (107).

Additionally, DNMT3B shows high homology with DNMT3A by conservation proline-tryptophan-tryptophan-proline (PWWP) domain, *C*-terminal catalytic domain, and cysteine-rich PHD zinc finger domain. DNMT3B is required by DNMT3A for restoration of methylation in somatic cells and DNMT3A and DNMT3B are essential for establishment of global DNA methylation patterns (108). *Dnmt3a* knockout significantly delayed the loss of gene expression driven by the CMV promoter which was highly methylated (109).

DNMT3L

The human *DNMT3L* located on chromosome 21q22.3 encodes a 387 amino acid containing a cystine-rich region with a novel-type zinc finger domain (110). DNMT3L is expressed only in embryonic stem cells and germ cells, and it regulates the expression of repetitive elements and imprinting genes in germ cells (111, 112). DNMT3L interacts with histone H3 tails and recruits or activates DNMT3A2 leading to *de novo* DNA methylation (113). Dnmt3L is a regulator of methylation in the gene body of housekeeping genes (114). Dnmt3L was also reported to cooperate with Dnmt3a and Dnmt3b to methylate the DNA (115, 116). Knockout DNMT3L in human embryonal carcinoma cells caused cell apoptosis and inhibited cell growth (117). Co-expression of DNMT3L and DNMT3A led to a stimulation of DNA methylation of imprinted genes by DNMT3A (118).

The identity matrix of the previously mentioned 13 DNMT isoforms plus DNMT2 was analysed by Fernanda I. *et al.* (Figure 1.7) (119). The amino acid sequences were identical at the catalytic site of DNMT3A1 and DNMT3A2 isoforms, and at that of DNMT3B1, DNMT3B2, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4.

Although the majority of DNMTs was identical within subfamilies, the amino acid sequences at the catalytic site of DNMT1, DNMT2, and DNMT3L showed a significant difference with the rest of the isoforms. The DNMTs also differ by size, which is governed by alternative splicing events and chromosome location. These factors contribute to changes in preferential DNA binding sites of DNMTs leading to different biological function expressed from target genes.

The roles of DNMTs in DNA methylation has been increased and partners of each DNMT have been reported. Numerous studies showed that DNMT1 and proliferative cell nuclear antigen (PCNA) interaction was important for DNMT1 activity (120). Also, the DNMT1/PCNA interaction could modify the structure of the replication focus targeting sequence domain and by a ricochet increase in both its activity and the DNMT1 affinity for DNA (121). Interaction of DNMT1 and CysxxCys finger protein1 presents a high affinity for unmethylated DNA. The inhibition of this interaction strongly decreased tumour growth of glioma cells in nude mice (122). Methyl-CpG-binding domain protein (MBD) interacts with DNMT1 and this interaction presents at methylated DNA and also involves in the recruitment of DNMT1 (123). Interaction of DNMT1 and DNMT-associated protein1 was involved in both early and late S phase of DNA replication and in the recruitment of PCNA (124). Recruitment of DNA methylation machineries by polycomb proteins has been found in highly methylated area to silence specific loci such as *HOX* genes (125). The recruitment of MBD and DNMT3A from the orphan nuclear receptor germ cell nuclear factor leads to *Oct4* repression (126). Moreover, double knockout of *DNMT1* and *DNMT3B* resulted in hypomethylation of *photocadherin10* (127).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1:DNMT1		20.5	12.8	18.3	18.3	17.6	18.1	16.1	16.1	16.1	16.1	16.1	16.1	17.1
2:DNMT2:	17.4		14.3	19.0	19.0	20.9	22.2	19.4	19.4	19.4	19.4	19.4	19.4	21.8
3:DNMT3L	5.4	7.2		23.3	23.3	27.2	29.6	23.3	23.3	23.3	23.3	23.3	23.3	33.5
4:DNMT3A1	11.1	13.6	33.2		100.0	81.6	83.3	80.6	80.6	80.6	80.6	80.6	80.6	81.8
5:DNMT3A2	11.1	13.6	33.2	100.0		81.6	83.3	80.6	80.6	80.6	80.6	80.6	80.6	81.8
6:DNMT3B5	9.1	12.8	33.2	69.9	69.9		100.0	85.7	85.7	85.7	85.7	85.7	85.7	99.4
7:DNMT3B3	8.5	12.3	32.7	64.5	64.5	90.4		77.4	77.4	77.4	77.4	77.4	77.4	99.4
8:DNMT3B1	9.8	13.8	33.2	80.6	80.6	100.0	100.0		100.0	100.0	100.0	100.0	100.0	99.4
9:DNMT3B2	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0		100.0	100.0	100.0	100.0	99.4
10:DNMTΔ3B1	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0	100.0		100.0	100.0	100.0	99.4
11:DNMTΔ3B2	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0	100.0	100.0		100.0	100.0	99.4
12:DNMTΔ3B3	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0	100.0	100.0	100.0		100.0	99.4
13:DNMTΔ3B4	9.8	13.8	33.2	80.6	80.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0		99.4
14:DNMT3B4	6.3	9.5	29.1	49.8	49.8	70.7	78.2	60.6	60.6	60.6	60.6	60.6	60.6	

Figure 1.7 Identity matrix of the catalytic site of 14 DNMT isoforms (taken from Fernanda I. *et al.*) (119). The degree of identity, from 0% to 100%, is represented by a colour gradient with dark red corresponding to 0% and dark blue to 100%.

1.4.2 The role of DNMTs in cancers and other diseases

Links between alterations in DNMT gene expression and protein activities are widely reported in multiple cancers (Table 1.1) (128). An increased DNMT activity is one mechanism that can lead to disease development through DNA hypermethylation. Hypermethylation can be found not only in the promoter region but also at the actively transcribed gene bodies and enhancer regions (129-131). Various studies show that aberrant DNA methylation contributes to cancer (62). Hypomethylation is a common characteristic of the cancer epigenome. Generally, hypomethylation of retrotransposon elements is found in tumour cells contributing to aberrant integrity of the genome (132). In a rat model, methyl donor deficiency promoted liver tumours and was associated with hypomethylation of oncogenes (133). Mice carrying a hypomorphic Dnmt1 allele developed aggressive T cell lymphomas at 4-8 months of age (90).

Table 1.1 Studies linked between DNMT alterations and various cancer types (adapted from Zhang and Xu) (128).

Tumour type	DNMT subtype	Model studied	Alteration
AML	DNMT3A	Human	Mutation
	DNMT3A	Mouse tumour	Mutation
	DNMT3A	Mouse tumour	Deletion
	DNMT3B	Mouse tumour	Deletion
	DNMTs	Human	Overexpression
MDS	DNMT3A	Human	Mutation
CMML	DNMT3A	Mouse	Mutation
CML	DNMTs	Human	Overexpression
ALL	DNMT3A	Human	Mutation
Lymphoma	DNMT1	Mouse tumour	Deletion
	DNMT3A	Mouse	Deletion
	DNMT3B	Cell line	Overexpression
Breast	DNMT1	Mouse tumour	Deletion
	DNMT1	Human	Overexpression
	DNMT3B	Human	Overexpression
Lung	DNMT3A	Mouse tumour	Deletion
	DNMT1	Cell line	Overexpression
Colon	DNMT1	Human	Mutation
	DNMT3B	Human	Overexpression
	DNMT3B	Mouse tumour	Overexpression
Liver	DNMT1	Human	Overexpression
	DNMT3A	Human	Overexpression
Melanoma	DNMT3A	Mouse tumour	Overexpression
Pancreas	DNMT1	Human	Overexpression
Prostate	DNMT3B	Human	Overexpression
Oesophagus	DNMT1	Human	Overexpression

Acute myeloid leukaemia (AML), Myelodysplastic syndrome (MDS), Chronic myelomonocytic leukaemia (CMML), Chronic myeloid leukaemia (CML), Acute lymphoblastic leukaemia (ALL).

Hypermethylation of the oncogenes caused by the overexpression of DNMTs (DNMT3A, DNMT3B, and DNMT1) is displayed in a variety of tumours (134). The over-expression of DNMT1 has been linked to the development of lung cancer (135). High activity of DNMT1 promotes tumour cell proliferation (136). DNMT1 deletion showed DNA demethylation and delayed lymphomagenesis and impairing tumour cell proliferation (137). DNMT1 and DNMT3B are over-expressed in T-cell acute lymphoblastic leukaemia and Burkitt's lymphoma (138). Around 15% of acute myeloid leukaemia (AML) cases display a heterozygous somatic missense mutation in *DNMT3A* (139). This missense mutation affects

amino acid R882 and is highly recurrent in these AML patients (139). DNMT3A is also involved in hepatocellular carcinogenesis and the depletion of *DNMT3A* in hepatocellular carcinoma inhibits cell proliferation and colony formation (140). Transient transfection of *DNMT3B1* and *DNMT3B2* in primary prostate cells increased methylation of tumour specific CpG sites such as CpG island hypermethylation of *Zinc Finger Protein 296* (141). Moreover, somatic mutations in DNMTs are one of the factors contributing to malignant transformation (142). The variation of DNMT3A is frequently mutated in myelodysplastic syndrome, AML and this has also been correlated with therapeutic resistance and increased disease aggressiveness (139, 143, 144). Dnmt3a-knockout mice led to increased proliferation of hematopoietic stem cells (145).

Aberrant DNA methylation or aberrant DNMT expression have also been associated with autoimmune disease (146), type 1 diabetes (147), SLE (148), vitiligo (149), and rheumatoid arthritis (RA) (150). Global methylation was significantly decreased in RA patients compared with control subjects, despite higher *DNMT1* expression in RA (150). Loss of Dnmt3a in the nervous system led to degeneration in adulthood and lethality (151). In another study, a tissue-specific Dnmt triple mutant (Dnmt1, Dnmt3a, aDnmt3b) mouse model demonstrated global genomic hypomethylation with reorganisation of the photoreceptor and synaptic layers within the retina (152). Mutations in the human *DNMT3B* gene have been shown to demethylate of classic satellite sequences and responsible for subsequent abnormalities such as multi-radiate chromosomes observed in immunodeficiency, chromosome abnormalities and facial anomalies (ICF) syndrome (153). Mutations in *DNMT3A* gene cause microcephalic dwarfism, a hypocellular disorder of extreme global growth failure (154). Additionally, multiple dominant germ line mutations clustered in a single small domain of DNMT1 result in a heterogeneous group of adults-onset neurological disorders including sensorineural deafness, ataxia, narcolepsy, dementia, and psychosis (155).

1.5 DNMT inhibitors (DNMTi) and their function

Several approaches have been investigated for inhibiting DNMT enzyme activity, with the goal of controlling aberrant DNA methylation. These can predominantly be divided into two categories: nucleoside and non-nucleoside compounds.

1.5.1 Nucleoside hypomethylating agents

Azacytidine and decitabine are cytidine analogues modified in position 5 of the pyrimidine ring in which a nitrogen replaces the carbon (Figure 1.8) (156). Azacytidine, a ribonucleoside,

is incorporated into both RNA and DNA, while decitabine, a deoxyribose analogue, is incorporated into only into DNA. DNMTs are incapable of performing methyl group transfer on incorporated cytidine analogues, so covalent enzyme-DNA adducts (DNMT-analogue complexes) remain blocked and are later destroyed by the proteasome complex, which recognises these complexes as sites of DNA damage (157, 158). Azanucleosides (azacytidine and decitabine) are the first hypomethylating agents approved by the US Food and Drug Administration. These medicines have functioned as the archetypal DNMT inhibitors and have been used for the treatment of patients with AML and myelodysplastic syndrome. Also, azacytidine has been approved by the European Medicines Agency for the treatment of patients with chronic myelomonocytic leukaemia (159). The therapeutic potential of DNMT inhibitors is not only limited to haematological malignancies, but have shown effectiveness in other cancers (160). Zebularine is the third novel nucleoside DNMT inhibitor family and it has been investigated in *in vitro* studies (161). Zebularine is metabolised in the same way as azacytidine and is activated after incorporation into a DNA stand (156). The demethylating effect of azacytidine and decitabine is stronger than zebularine but their mechanisms of action partially overlap. Collectively, these current nucleoside DNMT inhibitors are cytotoxic, mutagenic and show lack of specificity, which might limit the clinical application.

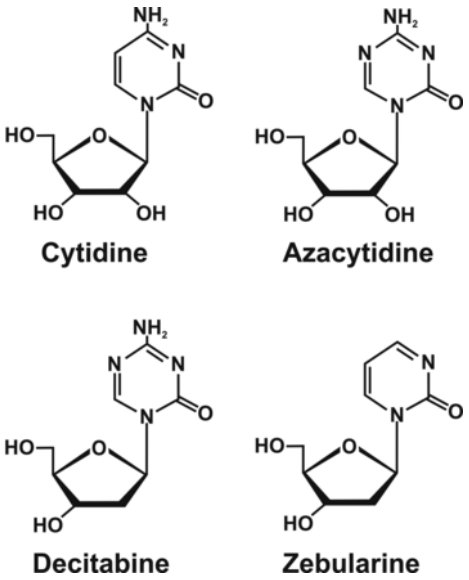


Figure 1.8 Nucleoside analogues inhibitors of DNMT (taken from Gnyszka A. *et al.*) (156).

1.5.2 Non-nucleoside inhibitor of DNMTs from nutritional compounds

Due to the toxicity of nucleoside DNMTi, researchers have discovered non-nucleoside DNMTi with improved specificity and toxicity. Another benefit of non-nucleoside DNMTi is that they are not incorporated into the DNA. This characteristic welcomes the possibility of

selective inhibition of different DNMTs with minimal consequences of side effects. Naturally occurring substances and dietary compounds have been investigated for their ability to inhibit DNMT enzymes. Several studies show the effect of polyphenols as epigenetic modulators through binding to the catalytic site of DNMTs in cancer chemoprevention (162). Polyphenolic compounds gained intensive interest due to their anti-carcinogenic, anti-oxidative, and anti-inflammatory activities (163, 164).

In this section, the inhibitory effects of five dietary constituents: (-)-epigallocatechin-3-gallate (EGCG), caffeic acid (CA), curcumin, theaflavin, and vitamin C, will be introduced. Their respective DNMT inhibitory effects are summarised in Table 1.2.

Table 1.2 Studies outline the interaction between bioactive dietary components and DNMT activity (adapted from Fatma Z.K. *et al.* (165)).

Studies	Dietary components	Concentration of dietary components	Enzymes inhibited	Epigenetic outcomes
Lee, W.J., *et al.* (166)	EGCG	7 μM (*in vitro*), 20 μM (*in vivo*)	DNMT1	EGCG inhibited human DNMT1 activity by binding to the catalytic core region
Fang *et al.* (167)	EGC–EGCG	50 μM (*in vivo*)	DNMT1	EGC and EGCG showed competitive inhibition of DNMT1 in treatment of the KYSE 510 cell line. EGCG showed a dose and time-dependent reversal of hypermethylation and re-expression of mRNA of *p16INK4a*, *RARβ*, *MGMT*, and *hMLH1* genes
Nandakumar, V., *et al.* (168)	EGC–EGCG	25 μM (*in vivo*)	DNMTs	EGCG reduced the activity of DNMTs by decreasing the mRNA levels and protein expression of DNMTs
Zhang, B. K., *et al.* (169)	EGCG	20 μM (*in vivo*)	DNMT1	EGCG inhibited the mRNA and protein expression activity of DNMT1 and downregulated binding to the promoter of DDAH2.
Shukla, S., *et al.* (170)	EGCG	20 μM (*in vivo*)	DNMT1	EGCG decreased the mRNA and protein expression activity of DNMT1 and increased the expression of unmethylation-specific GSTP1 promoter.
Morris J. *et al.* (171)	EGCG	100 μM (*in vitro*), 150 μM (*in vivo*)	DNMTs (DNMT1, DNMT3a, and DNMT3b)	EGCG treatment decreased promoter methylation of RXRα and decreased DNMT activity of DNMTs.

Pandey, M., *et al.* (172)	Green tea polyphenols, EGCG	20 µM (*in vivo*)	DNMT1	A dose and time-dependent inhibition of DNMT activity and protein expression was observed.
Lee and Zhu (173)	Caffeic acid, Chlorogenic acid	20 µM (*in vivo*)	DNMT1, M.Sssl DNMT	The caffeic acid and chlorogenic acid inhibited the DNA methylation that was catalysed by prokaryotic M.Sssl DNMT and human DNMT1, and increased levels of SAH.
Liu, Z., *et al.* (174)	Curcumin	100 µM (*in vivo*)	DNMT1	Curcumin blocked covalently the catalytic thiolate of DNMT1 to exert its inhibitory effect on DNA methylation.
Arumugam R., *et al.* (175)	Theaflavin	100 µM (*in vivo*)	DNMT3a *C*-terminal domain	Theaflavin showed inhibitory effect on DNMT3a-C with a physiologically and nutritionally relevant IC_{50} value.
Venturelli S., *et al.* (176)	Ascorbate (Vitamin C)	8 mM (*in vivo*)	DNMTs,	Ascorbate inhibited DNMTs activity.

The most abundant catechin in green tea is the EGCG, which is a competitive inhibitor of DNMTs through binding within the DNMT active site, leading to decreased global DNA methylation (167). EGCG has been found to promote vascular health through epigenetic reprogramming of endothelial-immune cell signalling and conversing systemic low-grade inflammation (177).

CA (3, 4-dihydroxy-cinnamic acid) is an organic compound and a type of polyphenol. The consumption of CA-rich foods is protective against carcinogenesis through preventing the formation of nitrosamines and nitrosamides, which are the main inducers (178, 179).

Curcumin is a bioactive compound and it is present in turmeric. Curcumin is also widely used in Asian countries as a yellow colour food additive. Curcumin has shown broad-spectrum epigenetic modulation through manipulating the activity of DNMTs, HATs, and HDACs (180). Curcumin has antioxidant activity through its influence on acetylation and deacetylation (181).

Theaflavin is another group of polyphenols and it is enriched in both black and oolong tea (182). Theaflavin have recently gained significant attention due to its biological and health promoting benefits. Research has shown that this compound might contribute to the positive benefits on diseases (183).

Apart from polyphenols, vitamin C (L-ascorbate) is a water-soluble vitamin and antioxidant. Vitamin C has recently been implicated in epigenetic regulation through its contribution in the demethylation cycle as a cofactor of TET (184). Oral vitamin C

supplementation to patients with myeloid cancer resulted in an increase in the 5-hmC/5-mC ratio compared to placebos (185).

The chemical structures of EGCG, curcumin, CA, theaflavin, and vitamin C are shown in the Figure 1.9. Several molecular modelling studies have demonstrated the interaction between some of these dietary compounds with DNMTs. For example, EGCG exerted its inhibitory effect on DNMT1 activity via blocking the entry of the key nucleotide cytosine into its active site (167). Curcumin has the potential to inhibit the DNMT1 activity by either blocking the catalytic domain, C1226 (174).

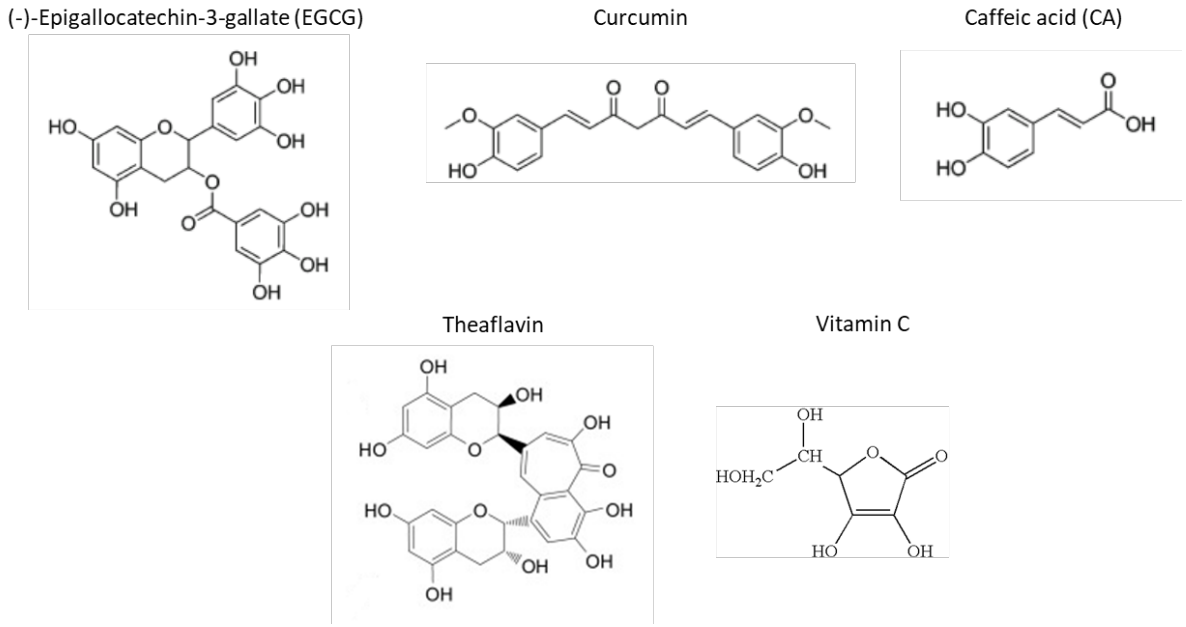


Figure 1.9 Chemical structures of EGCG, curcumin, CA, theaflavin, and vitamin C.

1.6 Nutrition and epigenetics

It is now widely known that bioactive food compounds alter molecular and cellular processes with individual genetic variations, lifestyle, physiological conditions, or environment. The recognition of individual variations in response to nutrients is the result of advances in genome-wide analysis, which allows the whole genome to be queried. Also, epigenetics can be modulated by bioactive dietary compounds and metabolites without changes in DNA sequences (186). Additionally, most epigenetic modifications are dynamic and plastic, indicating that epigenetic modifications are sensitive to environmental factors. Here, the effect of nutrition on DNA methylation and human diseases will be discussed.

1.6.1 Epigenetics and lifestyle

Lifestyle is used to explain the way of life or the manner of living, which comprises behavioural habits (*e.g.* diet, physical activity, smoking, and alcohol consumption) and other factors affecting well-being such as working environment and stress. Numerous studies have provided supporting evidence of the influence of lifestyle in epigenetic regulation. Table 1.3 presents examples of lifestyle factors that are recognised influencers of epigenetics.

Table 1.3 Lifestyle factors implicated in epigenetic modulation (taken from Jorge A. *et al*.) (187).

Factors	Example
Nutrition	Folate, phytoestrogen, polyphenols, selenium
Physical activity	Exercise
Tobacco smoke	Cigarette smoke, cigarette smoke condensate
Alcohol	High alcohol intake
Pollutants	Arsenic, PM10, black carbon, benzene, PAHs, POPs
Emotional	Stressful experiences
Shift work	Working at night

This section will focus on the impact of nutrition factors on epigenetics. Several dietary components are known to modify epigenetic marks. Folic acid and vitamin B_{12} play a vital role in DNA metabolism, and are found in a wide range of foods such as dairy products, meat, liver, and fruits. These vitamins are required for SAM production, which is a methyl donor for DNA methylation in the one-carbon metabolism (Figure 1.10) (188). Also, the SAM and *S*-adenosylhomocysteine (SAH) is generated from SAM by methionine adonosyltransferase, and this reaction is influenced by methylation events, making the Sam and SAH ratio a determinant of methylation efficiency. Folate and vitamin B_{12} are important cofactors in the regeneration of methionine by methylation of homocysteine (189). 5-Methyl tetrahydrofolate, the circulating form of folate, is depleted under the condition of folate deficiency, in which DNA methylation activity is reduced through lower availability of SAM. Folate depletion significantly increased plasma homocysteine and lymphocyte DNA hypomethylation in healthy postmenopausal women (190). High dose of oral choline (500 mg/day) for 12 weeks have compensated for the folate depletion (191).

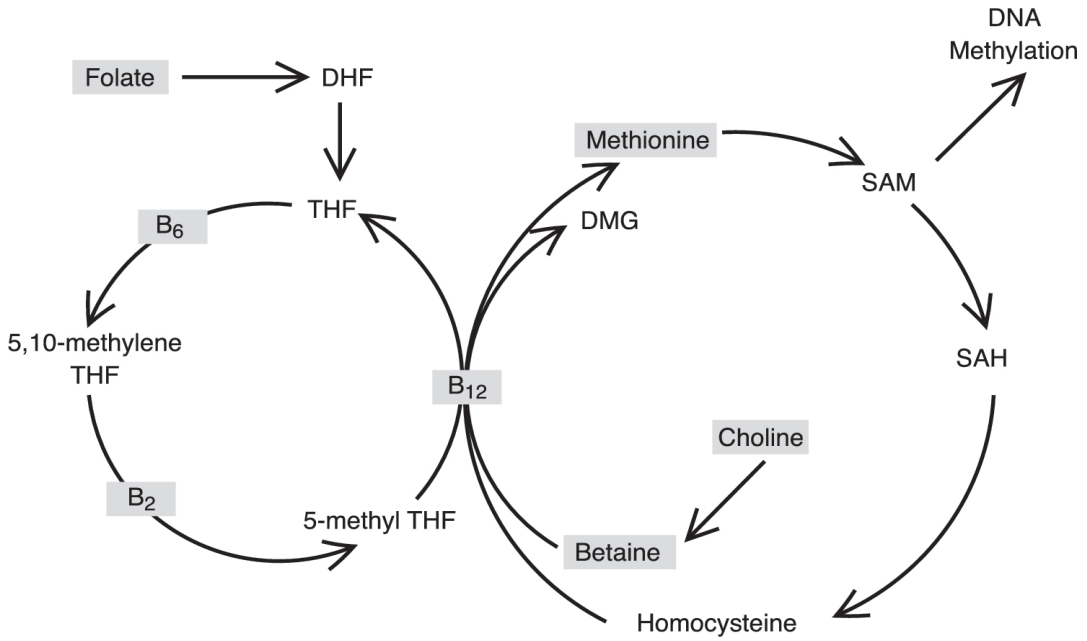


Figure 1.10 One-carbon metabolism. DHF, dihydrofolate; DMG, dimethyl glycine; SAH, *S*-adenosylhomocysteine; SAM, *S*-adenosylmethionine; THF, tetrahydrofolate (taken from James A.C. *et al.*) (192).

Selenium reactivates the pi-class glutathione-S-transferase through partial promoter DNA demethylation in prostate cancer cells (193). Selenium also decreases the activity of histone deacetylase and levels of DNMT1, while the levels of acetylated lysine 9 on histone H3 increase (193). However, methylation of the promoter of p53, a tumour oncogene, was decreased in a human colon cancer cell line treated with selenite (194). In addition to demethylating activity, polyphenols have also been shown to modify the activity of histone acetylases and histone deacetylases (195-197).

1.6.2. Methyl donors and DNMTi diet

Dietary factors can influence DNA methylation patterns through either, 1) donating a methyl group, known as a methyl donor or 2) inhibiting the activity of DNMT, known as a DNMTi (198). For example, folate, choline, betaine, vitamin B_{12}, and methionine are the methyl donors involved in the one-carbon metabolism and DNA methylation (Figure 1. 10) (199). In contrast, polyphenols can interfere with DNA methylation levels through inhibition DNMT activity.

Dietary factors as methyl donors - effects on phenotype

As for most methylation reactions within mammalian cells, the key methyl donor for the epigenome is SAM, which is generated during the methionine cycle via the one-carbon

metabolism (Figure 1.10) (200). Folate, vitamin B_{12}, choline, betaine and methionine act as methyl donors in this pathway while vitamin B_2 and vitamin B_6 are cofactors for one- carbon transfers (199). Folate is an essential vitamin for SAM generation. Folate can intervenes with global DNA methylation by impacting *DNMT* expression, or by providing more SAM (201). The ability of methyl donors to modify both the phenotype and/or the epigenetic status of the successive generations in animal models has led to a surge in increasing the knowledge surrounding methyl donors in human diets (202-205). A compelling example of folate on phenotype is its role in the study of the A^{vy} mouse model, whereby maternal folate dietary deficiency correlated with reduced DNA methylation and yellow coat colour in offspring (206). The *Agouti* gene in the yellow A^{vy} offspring is constitutively expressed due to hypomethylation of its upstream repetitive elements, giving rise to elevated yellow colour production, as well as increased obesity when compared with genetically identical brown mice (206). In another study, higher rates of brown wild type mice were observed in the offspring of pregnant mice fed with methyl donor rich diets compared with a control group fed with a standard diet (207). The impact of maternal dietary methyl donors on foetal growth was highlighted in a separate study whereby a methylating micronutrients-rich maternal diet containing methionine, folate choline, vitamin B_6, and vitamin B_{12} increased foetal weight (208).

Not only site-specific CpG but also global DNA methylation is modulated by methyl donor diets. The global DNA methylation level was increased by two year folic acid and vitamin B_{12} supplementation in elderly (209). Similarly, in randomised clinical trial, global DNA methylation was increased in post-polypectomy patients supplemented with 600 µg/day (210). Further studies regarding folate status have reinforced the necessity of folate inclusive diets. For instance, insufficient folate intake has been correlated with a reduction in the global DNA methylation in postmenopausal women (190). Similar findings of lower global DNA methylation were also observed in the offspring of female C57B16/J mice fed with low folate diets (211).

Folate supplementation was associated with increased DNA methylation of *Zinc finger protein 57* in maternal blood samples (212). Moreover, folate interferes with DNMTs in both RNA and protein expressions. For example, in the HCT116 cell line, *DNMT3A* expression was significantly reduced in high folate states (4 mg/L or 16 mg/L) compared with 0 mg/L folate, while *DNMT1* expression significantly increased in 16 mg/L of folate treated cells (213). Pseudo-pregnant mice fed a folate depleted diet (0 mg/L) showed a significant upregulation of RNA expression and protein expression of both Dnmt3a and Dnmt1, but no effects were observed for Dnmt3b (214). In addition to folate, choline and betaine are

precursors for SAM and supplementation of these also modulate DNA methylation (215). In animal studies, global DNA hypomethylation or hypomethylation of site-specific CpGs of *cyclin-dependent kinase 3*, *vascular endothelial growth factor-C, angiopoietin 2*, and *calbindin 1* genes were observed in the offspring to maternal choline deficient rodents (216-218). Rats fed by a low choline and methionine diet were deficient in 5-methyldeoxycytidine content of hepatic DNA (219). Methionine serves as a precursor for SAM in the one-carbon metabolism (220). The effect of the short-term feeding of methionine-supplementation in mice demonstrated increased methylation patterning of LINE-1 transposable elements (221).

Dietary factors that act as DNMT inhibitors (DNMTi)

This section was discussed in 1.5.2 about the inhibitors of DNMTs from nutritional compounds. Briefly, several phenolic acids, stilbenes, and flavonoids can inhibit the activity of DNMT, either directly by interaction with active site of DNMTs or through indirect mechanisms such as inhibition of DNA methylation by increasing SAH (196, 222-224). EGCG is the major metabolite (59%) of catechin (225) and has been shown to inhibit DNMT activity and to downregulate hTERT expression (226). *In vitro* studies have shown that catechin inhibits the enzyme activity of DNMT3A and DNMT3B (167, 226); however the biological consequences of this effect have not yet been examined. Genistein showed a dose-dependent inhibition of DNMT activity (227). Other phenolic compounds such as luteolin, apigenin, naringin, and hesperetin indirectly regulate DNMT activity via modulating the ratio of SAM and SAH (166, 228, 229). Although some nutrients are known to target a specific DNMT isoform (Table 1.2), there are no studies on the effects of potential DNMT inhibitors or methyl donors on specific DNMT sub-isoforms.

1.6.3 Effect of nutrition on human diseases

Nutrition is an important factor in health and diseases. Insufficient or excessive dietary intake results in a reduced state of health conditions. In a systemic analysis from the Global Burden of Disease Study in 2017, the proportion of disease-specific burden induced by dietary risk factors was estimated, detailing 11 million deaths and 255 million disability-adjusted life-years attributed to dietary risk factors such as high sodium intake and low intake of whole grains and fruits (230). Undernutrition, or nutrition deficiency, and overnutrition were the prevailing diet induced disease states. It is important to consume a balanced diet to prevent disease or reduce disease risk. For example, concentrated sugars and refined flour products impair glucose metabolism leading to obesity or diabetes (231). Low fibre intake, imbalance

of omega 3 and omega 6 fatty acids, and consumption of red meat contribute to increased cancer risk (232-235). Vitamin deficiency is well known to increase disease likelihood: vitamin D deficiency contributes to osteopenia and osteoporosis (236); xerophthalmia is an ocular manifestation of vitamin A deficiency (237); low vitamin B_{12} has been known as a potential risk factor of neural tube defects (238), scurvy is a condition caused by vitamin C deficiency. Some evidence shows the protective effect of vitamins on cancer risk. Vitamin B_{12}, for instance, is essential for DNA synthesis and cellular energy production (239). Methylcobalamin, a form of vitamin $B_{12,}$ inhibited tumour growth of oestrogen-sensitive malignant cells in mice (240).

Nutrition influences health through multiple molecular mechanisms including changes in DNA methylation patterns (241), which modulate biological functions that impact health and ageing. As mentioned in the section 1.6.2, higher intake or status of several micronutrients participate in the one-carbon metabolism, including folate, choline, betaine, methionine, vitamin B_6, and vitamin B_{12}; all of which contribute to SAM production and are associated with increased DNA methylation (242, 243). For example, dietary methyl donor restriction and polymorphisms in genes encoding key components of the one-carbon metabolism pathway have been associated with aberrant DNMT expression, decreased global DNA methylation, and increased cancer risk (244). In addition, mice fed a folate deficient diet showed chromosomal damage in nucleated erythrocyte precursors (245). Hypomethylated DNA and double DNA strand breaks were observed in the liver of rats that were fed a methionine, choline, and folic acid deficient diet (246). Protein restriction in pregnant C57BL/6J mice caused hypermethylation of the liver X-receptor alpha promoter in liver tissue of their offspring, which is a nuclear receptor controlling cholesterol and fatty acid metabolism (247). In a human study, the incorporation of [^{3}H] methyl groups was found to be increased in postmenopausal women who consumed a moderately folate-depleted diet (118 µg folate per day) for seven weeks, suggesting low methylation of DNA (248). Reduced DNA methylation contributes to carcinogenesis by altering the gene expression of tumour suppressor genes or proto-oncogenes (249). Folate depletion also leads to an increase in homocysteine, which is a risk factor of CVD (250). Concomitantly, increasing plasma homocysteine correlated with elevated plasma SAH as well as DNA hypomethylation in lymphocytes (251). Maternal folate deficiency caused global and LINE-1 hypomethylation resulting in increased retrotransposition in foetuses, which in turn increased the risk of intrauterine growth retardation (252). Mice fed a high methionine and low B vitamin diet showed hypermethylation of netrin-1 gene promoter leading to low gene expression of netrin-1 and association with memory loss (253).

Methionine adenosyltransferase converts methionine to adenosylmethionine (AdoMet) and adenosylhomocysteine (AdoHcy) is generated following methyl donation before forming homocysteine. The elevated AdoHcy concentration results in a lower AdoMet/AdoHcy ratio, thus inhibiting DNA methyltransferase in liver (254) and human fibroblasts cell lines (255). In a mice study, a folate/choline-deficient diet significantly decreased AdoMet concentration but there was a significant increase in AdoHcy, and this consequently decreased the AdoMet/AdoHcy ratio (256). The effect of homocysteine on AdoMet and AdoHcy concentrations has been found to be tissue type dependent leading to within-tissue gene-specific hypermethylation, but low global DNA methylation, all of which might promote CVD risk through an altered gene expression profile (257, 258).

1.7 Personalised nutrition

Data on relationships between dietary intake and health outcomes are used to build public health dietary recommendations. However, individuals respond differently to the same dietary intake and/ or nutrients because of inter-individual variations in their genetic makeup and other phenotypic factors. This has led to the idea that it may be possible to use individual characteristics to personalise nutritional advice.

The concept of personalised nutrition (PN) (259, 260) is based on the idea that the generic "one size fits all" notion may not be appropriate for everyone in the population and that tailoring advice for each individual may be more effective on health and disease outcomes. PN can be based on biological characteristics, on current eating behaviour and/or on psychosocial parameters of the individual (261). The potential benefits of PN in promoting health include improved efficacy and reduced costs of healthcare. Also, PN can be applied either in management of patients with specific diseases or in improving public health. The differential response of individuals to the same nutritional components provides the motivation and strategy development for PN.

A better understanding of how nutrients or diet plays an important role in health by using technology, *e.g.* Next-generation sequencing methods and mass-spectrophotometry enables biological differences between individuals to be observed for tailoring PN regimens, through genetic and epigenetic assessment. The concept of PN was proved by the Food4Me study to providing personalised dietary advice based on individual dietary intake, phenotypic and genotypic data (262, 263). In the Food4Me study after six months, those randomised to PN had bigger improvements in their diet than those randomised to the control (generic dietary advice) (262, 263). However, there was no evidence of added benefit of using

genotypic and/or phenotypic information in generating the PN advice (262). However, in a smaller Canadian intervention study, genotype-based dietary advice was more effective than general dietary advice in reducing salt intake in healthy young adults (264). Advances in data science, analytical technologies, molecular science and nutritional knowledge will allow researchers to refine PN. However, to be useful in improving public health, PN will also need to address the psychological, social, economic and cultural factors that influence eating patterns in order to ensure that advice is converted into action and that improved dietary habits are sustained in perpetuity (261).

1.8 Research gap

Since the discovery of epigenetic factors as drivers of diseases, tools and strategies have been developed to investigate the epigenome. This implementation of epigenetic tools has proven useful in designating treatment. Numerous DNMT isoforms were previously discovered (see section 1.4). High similarity of the catalytic domain among DNMT isoforms was found through the assembly of structures (119). To date, it remains a controversy over which DNMT isoform would be the best maker for diseases. Although preferential target sites for DNMTs have been identified, these target sites were limited only the cancer-related genes, due to the availability of an outdate version of the Illumina DNA methylation assay (87). Moreover, a cell model to assist with the unveiling of novel disease associated CpGs in DNMT genes is unavailable. Therefore, cell model systems that each overexpress one of 13 different DNMTs can be investigated to reveal preferential DNMT target CpG sites that are isoform-specific using Illumina Infinium Methylation EPIC BeadChip (EPIC arrays).

The role of DNMTs in human disease is important for therapeutic purposes. It is necessary to develop selective DNMTi toward a specific DNMT isoform, which could identify an appropriate DNMT isoform to target in specific diseases. However, this remains a grey area due to lack of materials and techniques. It will be valuable data for all potential DNMTi to be screened for their efficacy across a selection of conditions. Also, there remains controversy over which DNMTi can be inhibited through the action of a DNMTi. Although the catalytic pockets of DNMTs are conserved, amino acid residues are dissimilar. For example, Cys662 in DNMT3a is replaced by Trp1173 in DNMT1, Arg887 by Asn1580, and Trp889 by Val582. Additionally, the SAM cofactor can receive a different conformation in its binding pocket based on the type of methyltransferase (265).

Based on the studies discussed in this chapter, there is growing evidence that diet and bioactive food compounds play an important role in the epigenome through modulating DNA

methylation, DNMT activity, and DNMT expression. Assessment of these epigenetic determinants provides valuable information on how individuals respond to dietary factors. Therefore, nutrition and epigenetics collectively bring forward the prospect of dietary intervention for the health-promotion, disease prevention, and as a combatant therapy in diet-related disorders. Epigenetic biomarkers to help govern PN remain largely unknown or very limited. Thus, the concept of a balanced diet between methyl donor and DNMTi diet can bring us closer to the goal of effective PN.

1.9 Hypotheses, aims and objectives

1.9.1 Hypotheses

1. Overexpression of DNMTs leads to aberrant DNA methylation patterns.
2. Each DNMT isoform characteristically targets different CpG sites.
3. Certain micronutrients and other food constituents can alter global DNA methylation or methylation at specific CpG sites.
4. Polyphenols and vitamin C inhibit specific DNMTs and, as a consequence, modulate patterns of DNA methylation.

1.9.2 Aim

- To test these hypotheses using short-term cultures of cells models.

1.9.3 Objectives

1. To generate 13 cell lines that each over-express a different DNMT isoforms.
2. To identify the CpG sites that are methylated differentially by each DNMT isoform, I will undertake genome-wide DNA methylation analysis using the Illumina Infinium Methylation EPIC BeadChip.
3. To assess CpG specific DNMTs in relation to biological mechanism by pathway analysis according to enrichment statistics of the difference of gene.
4. To investigate interactions between selected food constituents, *i.e.* theaflavin, EGCG, CA, curcumin, and vitamin C, and specific DNMT isoforms in modulating DNA methylation patterns.

Chapter 2: Materials and methods

2.1 Generation of cell lines with stable over-expression of specific DNMT isoforms

2.1.1 Cell culture conditions

MEG-01 cell line

The megakaryoblast cell line (MEG-01) was purchased from Merck, UK. This cell was cultured in Roswell Park Memorial Institute (RPMI) medium (Sigma, UK) supplemented with 10% foetal bovine serum (FBS) (Sigma, UK), and 2 µM glutamine, at 37 °C in a 5% CO_2 humidified atmosphere.

HEK293T cell line

The human embryonic kidney 293T cell line (HEK293T) (gifted from Dr Viktor Korolchuk, Newcastle University) was cultured in Dulbecco's Modified Eagle's Medium (DMEM) - high glucose (Sigma, UK) supplemented with 10% FBS (Sigma, UK), and 2 mM glutamine, at 37 °C in a 5% CO_2 humidified atmosphere.

HEK293FT cell line

The human embryonic kidney 293FT cell line (HEK293FT), a fast growing HEK293T strain, (gifted from Dr Viktor Korolchuk, Newcastle University) was cultured in DMEM (Sigma, UK) supplemented with 10% FBS (Sigma, UK), 1% sodium pyruvate, 1% non-essential amino acid, and 1% G418, at 37 °C in a 5% CO_2 humidified atmosphere.

2.1.2 Retransformation of 13 DNMT-pIRES puro3 plasmids to *E.coli* DH5

E.coli DH5α cell (Bioline, UK) (25 µL) was thawed on ice, then 2.5 µL of pIRES puro3 plasmids (gifted from Dr Si Ho Choi) were added, followed by gentle swirling of the tube for a few seconds to mix and then kept on ice for 30 minutes. The tube containing the mixture was then placed in a 42 °C water bath for 45 seconds without shaking, then replaced on ice for 2 minutes. The mixture was diluted by addition 1 mL of pre-warmed LB medium and incubated at 37 °C for 1 hour with shaking. 100 µL of the transformed cells were plated out on lysogeny broth (LB) agar containing 100 µg/mL ampicillin and incubated at 37 °C overnight.

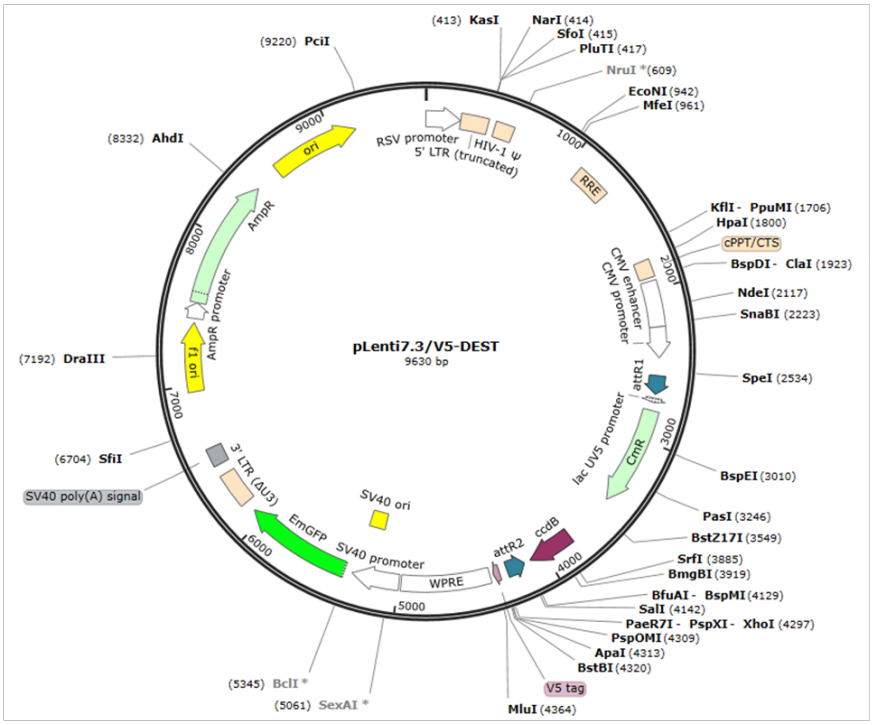
2.1.3 Plasmid construction for viral system

Plasmid preparation

The pLenti7.3/V5 DEST gateway vector (gifted from Dr Viktor Korolchuk, Newcastle University) (Figure 2.1A) was modified to add more restriction cloning sites using oligo nucleotide sequences (containing XbaI, NheI, ClaI, EcoRI, SwaI, PspOMI, and MluI sites) (IDT DNA, UK) (Figure 2.1B).

The pLenti7.3/V5 DEST gateway vector was cut by XbaI and MluI restriction enzymes (Thermo Fisher Scientific, UK) to open the plasmid using the mixture shown in Table 2.1.

A



B

```
XbaI  NheI  ClaI  EcoRI  SwaI  PspOMI MluI
CTAGAGCTAGCATCGATGAATTCATTTAAATGGGCCCA
    TCGATCGTAGCTACTTAAGTAAATTTACCCGGGTGCGC
```

Figure 2.1 Cloning sites within pLenti7.3/V5-DEST. A) Plasmid map of pLenti7.3/V5-DEST. B) Diagrammatic representation of new multiple cloning sites: NheI, ClaI, EcoRI, SwaI, and PspOMI with cutting by XbaI and MluI.

Table 2.1 Volume of reagents for restriction enzyme reaction.

Reagent	Volume per reaction
10x Fast digest green buffer (Thermo Fisher Scientific, UK)	2 µL
XbaI (Thermo Fisher Scientific, UK)	1 µL
MluI (Thermo Fisher Scientific, UK)	1 µL
Nuclease-free water (QIAGEN, USA)	Up to the total volume
pLenti7.3/V5 DEST gateway vector	1 µg
Total	20 µL

Mixture of vectors and restriction enzymes was incubated at 37 °C for 3 hours. These mixture solutions underwent gel electrophoresis on a 1% agarose gel at 90 volts for 1 hour. Then, all target vectors' bands were cut out of the gel by hand under UV light and then purified using QIAquick gel extraction kit (QIAGEN, USA). Firstly, gel slices were weighed in a 1.5 mL microcentrifuge tube and 3 volumes of buffer QG were added per volume of gel. Then, this tube was incubated at 50 °C for 10 minutes until the gel slice had dissolved completely. After that, 1 gel volume of isopropanol was added and mixed by vortexing. The mixture was applied to the QIAquick column and this column was centrifuged at 12,000g for 1 minute. The filtrate was discarded from the collection tube. 0.5 mL of buffer QG was added into the column and this column was centrifuged at 12,000g for 1 minute. For the washing step, 0.75 mL of PE buffer was added into the column and this column was centrifuged at 12,000g for 1 minute. The filtrate was discarded from the collection tube and then the empty QIAquick column was centrifuged at 12,000g for 2 minutes. For the elution step, 50 µL of buffer EB (10 mM Tris-Cl, pH 8.5) were added into the column which was centrifuged at 12,000g for 1 minute.

Annealing oligonucleotide

Two oligonucleotides for new multiple cloning sites: NheI, ClaI, EcoRI, SwaI, and PspOMI were annealed together using protocol and condition outlined in Table 2.2 and 2.3.

Table 2.2 Volume of each oligonucleotide.

Reagent	Volume per reaction
Oligonucleotide forward sequence (10 µM) 5'CTAGAGCTAGCATCGATGAATTCATTTAAATGGGCCCA3'	10 µL
Oligonucleotide Reverse sequence (10 µM) 5'TCGATCGTAGCTACTTAAGTAAATTTACCCGGGTGCGC3'	10 µL
Total	20 µL

Table 2.3 Temperatures and times for annealing condition.

Temperature	Time
95 °C	30 seconds
72 °C	2 minutes
37 °C	2 minutes
25 °C	2 minutes
4 °C	Hold

After incubation, annealing products were diluted by 1 µL of annealing product and 99 µL of Nuclease-free water.

Ligation between target DNA or oligomers and target plasmid

The annealing product or target DNA was ligated into the pLenti7.3/V5 DEST gateway vector using the protocol outlined below (Table 2.4).

Table 2.4 Volume of reagents for ligation reaction.

Reagent	Volume per reaction
5x ligase reaction buffer (Thermofisher Scientific, UK)	2 µL
Insert (diluted annealing product)	Up to the total volume
T4 ligase (5U/µL)	0.3 µL
pLenti7.3/V5 DEST gateway vector	1 µg
Total	20 µL

The components detailed in Table 2.4 were added to the mixture which was incubated at 22 °C for 3 hours and them transformed into *E.coli* DH5α using the heat shock protocol described in 2.1.2.

Plasmid DNA extraction

A single colony was cultured in 7 mL and 100 mL of LB containing 100 µg/mL of ampicillin at 37 °C overnight with vigorous shaking (220 rpm). After incubation, positive bacteria were extracted using E.Z.N.A.® Endo-free Plasmid DNA Mini KitII (VWR, USA) and PureYield™ Plasmid Midiprep System (Promega, UK).

Use of the E.Z.N.A.® Endo-free Plasmid DNA Mini KitII

Plasmid was extracted from bacteria using E.Z.N.A.® Endo-free Plasmid DNA Mini KitII, (VWR, USA) according to manufacturers' instruction. Briefly, 7 mL of suspension containing bacteria were centrifuged at 5,000g for 10 minutes at room temperature and the medium was discarded. 500 µL of solution I/RNase A were added to the cell pellet which was resuspended by vortexing and then this mixture was transferred into a new 2 mL microcentrifuge tube. 500 µL of solution II were added to the mixture in this microcentrifuge tube and mixed by inverting and gentle rotation of the tube 6 times. Next, for protein precipitation, 250 µL of N3 buffer were added and the mixture was mixed gently by inverting until a flocculent white precipitate formed. This mixture was centrifuged at 12,000g for 10 minutes. A compact white pellet was formed and the clear lysate (supernatant) was transferred to a new 1.5 mL microcentrifuge tube. Then, one volume ETR binding buffer was added and this tube was inverted 10 times to mix thoroughly. A HiBind DNA mini column II was prepared by adding 100 µL of 3M NaOH into this column and the column was centrifuged at 12,000g for 60 seconds. The filtrate was discarded from the collection tube. The clear supernatant (described above) was transferred into the HiBind DNA mini column II carefully (not to disturb the pellet). This column was centrifuged at 12,000g for 1 minute and the filtrate was discarded. This step was repeated until all of the clear lysate (supernatant) had been transferred. Next, 500 µL of ETR wash buffer were added into the column and the column was centrifuged at 12,000g for 1 minute. Then, the filtrate was discarded, 500 µL of HBC buffer were added, the column was centrifuged at 12,000g for 1 minute and the filtrate was discarded. In the washing step, 700 µL DNA wash buffer were added into the column, which was centrifuged at 12,000g for 1 minute. Then, the filtrate was discarded, and this step was repeated. After washing, the column was dried by centrifugation of the empty HiBind DNA Mini Column II

matrix at 12,000g for 2 minutes. For the elution step, 50 µL of elution buffer were added into the column and incubated at room temperature for 10 minutes. Then, this column was centrifuged at 12,000g for 1 minute. The plasmid DNA was stored at -20 °C.

Use of the PureYield™ Plasmid Midiprep System

Plasmid was extracted from bacteria using PureYield™ Plasmid Midiprep System (Promega, UK) according to manufacturers' instruction. Briefly, 100 mL of suspension containing bacteria were transferred to 50 mL centrifuged tube and centrifuged at 5,000g for 10 minutes at room temperature and the medium was discarded. 3 mL of resuspension solution were added to the cell pellet which was resuspended by vortexing and then added 3 mL of lysis solution and mixed by inverting and gentle rotation of the tube five times. Next, for protein precipitation, 5 mL of neutralization solution were added, and the mixture was mixed gently by inverting and gentle rotation of the tube ten times. This mixture tube was centrifuged at 10,000g for 30 minutes. Next, Column stack was prepared by placing a blue PureYield™ clearing column on top of a white PureYield™ binding column and placed onto a vacuum manifold. The clear lysate (supernatant) was transferred carefully into column stack and then all liquid was passed through both the clearing and binding columns after applying vacuum. The blue clearing column was removed while the binding column was left on the manifold. Next, 5 mL of endotoxin removal wash buffer were added into the binding column and allowed to pass through the binding column after applying vacuum. Then, 20 mL of column wash solution were added into the binding column and allowed to pass through the binding column after applying vacuum. After this, binding column was dried by applying a vacuum for 1 minute and removed from manifold. For the elution step, 450 µL nuclease-free water were added into the binding column which was placed into a new 50 mL centrifuged tube. This tube was centrifuged at 2,000g for 5 minutes. The plasmid DNA was stored at -20 °C.

Colony screening by PCR and cutting by restriction fragment length polymorphism (RFLP)

E. coli colony containing a positive vector was screened by PCR and RFLP using the mixture of reagents and condition described in Tables 2.5-2.7.

The mixture tube was incubated at 37 °C for 1 hour and run gel electrophoresis on a 1% agarose gel at 90 volts for 1 hour. To confirm the positive colony by PCR, each colony was picked up separately and diluted in 20 µL of water. PCR master mix was prepared as shown below (Table 2.6).

Table 2.5 Volume of the reagents for RFLP.

Reagent	Volume per reaction
10x Fast digest green buffer (Thermofisher Scientific, UK)	2 µL
EcoRI (Thermofisher Scientific, UK)	1 µL
KpnI (Thermofisher Scientific, UK)	1 µL
Nuclease-free water (QIAGEN, USA)	Up to the total volume
pLenti7.3/V5 DEST gateway vector	1 µg
Total	20 µL

Table 2.6 Volumes of the reagents for colony PCR.

Reagent	Volume per reaction
2x GoTaq® Green Master Mix (Promega, UK)	12 µL
CMV forward primer (10 µM) : 5’ CGCAAATGGGCGGTAGGCGTG 3’	1 µL
Myc reverse primer (10 µM) : 5’ CTGAGATCAGCTTCTGCTC 3’	1 µL
Nuclease-free water (QIAGEN, USA)	9 µL
Diluted colony	1 µL
Total	24 µL

Table 2.7 Condition of colony PCR using thermocycler PCR.

Procedure	Temperature	Time	Number of cycles
Pre-denaturation	95 °C	10 minutes	1
Denaturation	95 °C	40 seconds	35
Annealing	58 °C	40 seconds	

Extension	72 °C	30 seconds	
Final extension	72 °C	5 minutes	1
	4 °C	Hold	

Confirmation the insert's sequences by Sanger sequencing

All positive colonies were cultured and extracted using PureYield™ Plasmid Midiprep System (see page 33). These plasmids were cut by PvuI to confirm the target vectors. These plasmids were sent to MRC PPU DNA Sequencing and Services at Dundee University for sequence analysis.

2.1.4 Virus production

HEK293FT cells were seeded into 6-well plate by one million cells per well and incubated at 37 °C in a 5% CO_2 humidified atmosphere for 24 hours. After 24 hours, the confluence of HEK293FT cells were around 60-70%. All plasmids were prepared as the following protocol. REV PLP2, GAG PLP1, and VSV PLP1 plasmids were prepared with 600 ng of final concentration and mixed with 600 ng of plenti7.3/V5-DEST gateway-DNMTs or plenti7.3/V5-DEST gateway-Myc. The mixed solution volume was then adjusted to be 250 µL using Opti-MEM® Medium (Sigma, UK). Diluted Lipofectamine® 2000 (Thermo Fisher Scientific, UK) was prepared using 7.2 µL Lipofectamine® 2000, and 242.8 µL Opti-MEM® Medium and mixed by vortexing and incubated at room temperature for 5 minutes. Diluted plasmids and diluted Lipofectamine® 2000 were combined and incubated at room temperature for 20 minutes. 500 µL DNA-lipid complex were added into MEG-01 or HEK293T cells without changing media and incubated at 37 °C in a 5% CO_2 humidified atmosphere for 16-18 hours. After incubation, all media were removed and replaced with fresh media.

MEG-01 or HEK293T cells were seeded by 300,000 cells per well in 12-well plate for MEG-01 cells and 400,000 cells per well in 6-well plate for HEK293T cells, then all cells were incubated at 37 °C in a 5% CO_2 humidified atmosphere for 16-18 hours. After 48 hours of virus production in HEK293FT cells, 1.5 mL of cultured media were harvested and filtered through a 0.45 µm filter. 500 µL of filtered media containing virus were mixed with 500 µL of fresh media and polybrene (final concentration: 5 µg/mL) and transferred into the target cells. After that, the target cells were incubated at 37 °C in a 5% CO_2 humidified atmosphere for 16-18 hours. Target cells were changed media, replaced with fresh media, and kept growing.

2.1.5 Confirmation of positive cells using fluorescent microscopy

Positive cells were observed under fluorescent microscopy, Leica DMi8 (Leica microsystems, UK), as the positive cells can express green fluorescent protein (GFP) due to a GFP sequence in the pLenti7.3/V5 DEST gateway vector.

2.1.6 Single cell selection by Fluorescence-activated cell sorting (FACS)

All transfected cells were harvested and centrifuged at 800g for 5 minutes and filtered through a 30 μm filter into FACS tubes. Each cell was sorted and dispensed in a single drop into each well of a 96-well plate using a FACS Fusion Sorter.

2.1.7 Confirmation the positive cells by qPCR and western blot assays

Identification of housekeeping (HK) genes

Eleven HK genes (Table 2.8) (266, 267) were used to identify the most stable HK genes for from cell lines overexpressing individual DNMT isoforms in MEG-01 or HEK293T cells. A set of tested candidate HK genes was performed, and a gene expression normalization factor was calculated using geNorm (268). Two most stable HK genes of each experimental were set for normalisation of gene expression based on geNorm analysis. Briefly, the cycle threshold (Ct) values were transformed to quantities by the comparative Ct method. The highest relative quantities for each gene were set to 1. These raw HK gene quantities were the required data input for geNorm to generate normalised factor (NF). In example (Figure 2.2), geNorm analysis would indicate that HK1 and HK2 were the most stable gene. Thus, after calculation of NF, the normalised gene of interest (GOI) expression levels can be calculated by dividing the raw GOI quantities for each sample by the appropriate NF.

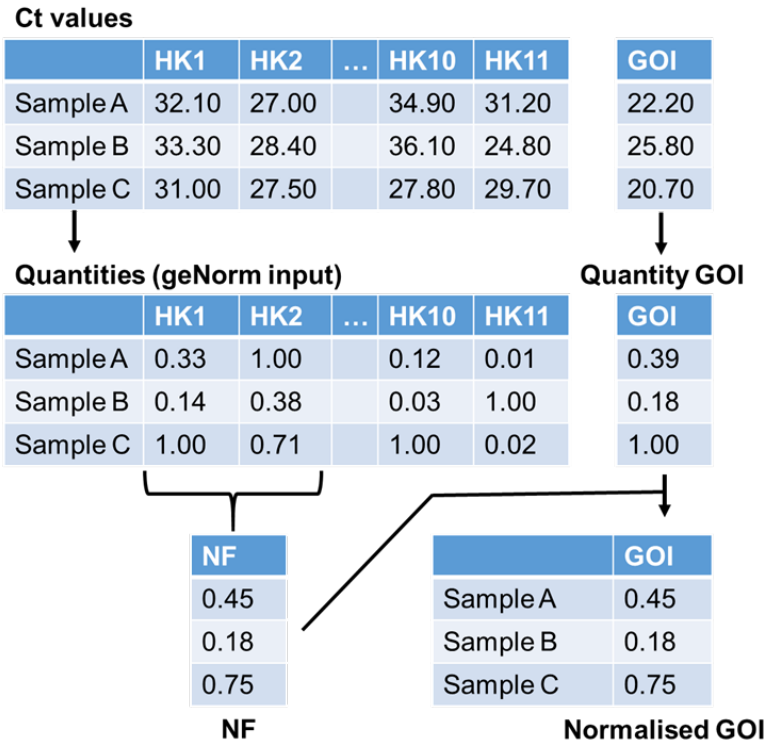


Figure 2.2 Normalisation flow chart. Cycle threshold: Ct, Gene of interest: GOI, Housekeeping gene: HK, Normalisation factor: NF.

Table 2.8 Primer lists of housekeeping genes.

Genes	Primer sequence
Ribosomal Protein S13 (RPS13)	Forward: 5' GCTTTACCCTATCGACGCA 3'
	Reverse: 5' TTGTGCAACACCATGTGAATC 3'
Transferrin Receptor (TFRC)	Forward: 5' AACTGGACAGCACAGACTTC 3'
	Reverse: 5' ACGCCAGACTTTGCTGAGTT 3'
Hypoxanthine Phosphoribosyltransferase 1 (HPRT1)	Forward: 5' GGACTAATTATGGACAGGA 3'
	Reverse: 5' TTTGATGTAATCCAGCAGG 3'
Ribosomal Protein Lateral Stalk Subunit P0 (RPLP0).	Forward: 5' CATGTTGCTGGCCAATAAGG 3'
	Reverse: 5' TAGTGGTGATACCTAAAGCCT 3'
Ribosomal Protein L13A (RPL13A),	Forward: 5' GTACGCTGTGAAGGCATCAA 3'
	Reverse: 5' ACGGTCCGCCAGAAGATG 3'
Peptidylprolyl Isomerase A (PPIA)	Forward: 5' ATGGACAAGATGCCAGGAC 3'
	Reverse: 5' TCCAGGGTTTATGTGTCAGG 3'
Beta-2-Microglobulin (B2M)	Forward: 5' TACTCTCTCTTTCTGGCCTG 3'
	Reverse: 5' GGATGGATGAAACCCAGACA 3'

Actin Beta (ACTB)	Forward: 5' GAGGCACTCTTCCAGCCTT 3'
	Reverse: 5' CGCCAGACAGCACTGTGTT 3'
Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH)	Forward: 5' CATGTTCGTCATGGGTGTGA 3'
	Reverse: 5' GAGTCCTTCCACGATACCAA 3'
Glucuronidase Beta (GUSB)	Forward: 5' CCATTCCTATGCCATCGTGT 3'
	Reverse: 5' GATGGCGATAGTGATTCGGA 3'
18sRNA	Forward: 5' CCATTCCTATGCCATCGTGT 3'
	Reverse: 5' TAGTAGCGACGGGCGGTGTG 3'

RNA extraction

All DNMT cell types were cultured by 1,000,000 cells per well in a 6-well plate for MEG-01 and 400,000 cells per well in a 6-well plate for HEK293T cells, at 37 °C in a 5% CO_2 humidified atmosphere for 16-18 hours. Total RNA was extracted by E.Z.N.A.® Total RNA kit I.

Use of the E.Z.N.A.® Total RNA kit I:

Cell culture media were aspirated and discarded. Cells were then washed with 1% PBS and 0.1-0.25% trypsin were added to allow cells to detach. After detachment, DMEM was added as the same volume as trypsin to inactivate the trypsin. All mixture solution was transferred to 15 mL tubes and centrifuged at 800g for 5 minutes. The number of cells per sample was estimated ($5 \times 10^6 - 1 \times 10^7$ cells) and the appropriate volume of TRK lysis buffer (700 µL) was added to each pellet followed by vortexing for 20 seconds. 700 µL of 70% ethanol were added and followed by vortexing. The sample was loaded onto a Hibind RNA mini column inserted into a collection tube and centrifuged at 10,000g for 1 minute. The eluate was discarded and this step was repeated until all samples had been transferred to the column. 500 µL of RNA wash buffer I were added to the Hibind RNA mini column and centrifuged at 10,000g for 30 seconds. The eluate was discarded and 500 µL of RNA wash buffer II were added to the Hibind RNA mini column and centrifuged at 10,000g for 1 minute. The washing step was repeated and the empty column centrifuged at 13,000g for 2 minutes to completely dry the Hibind RNA mini column. The columns were transferred to new microcentrifuge tubes and the total RNA was eluted by centrifuging at 13,000g for 2 minutes with 50 µL of diethyl pyrocarbonate (DEPC) water. Total RNA quantity and purity were assessed by Nanodrop1000 (Thermo Fisher Scientific, UK). Total RNA was kept at -80 °C.

cDNA synthesis

Reverse transcription (RT) was carried out using the Omniscript RT kit (QIAGEN, USA) according to manufacturers' instruction. The reaction volume was 20 µL, adjusted with molecular grade water. The remaining components were 2 µL 10x buffer RT, 2 µL dNTP mix (5 mM each dNTP), 2 µL random hexamer primer, 1 µL Omniscript reverse transcriptase, and 2 µL (1.5 µg) sample total RNA. The reaction was incubated at 37 °C for 1 hour.

Quantitative polymerase chain reaction (qPCR)

PCR was used to amplify the target genes (*DNMT* isoforms), *GFP* gene, and HK genes using primer oligonucleotide (0.01 µmol) (Tables 2.8 and 2.9). Only primer sets of hDNMT1, hDNMT3A1, hDNMT3A2, hDNMT1, and hDNMT3L were specific to their respective genes, while hDNMT3B1-3B5, hDNMTΔ3B1- Δ3B2, and hDNMTΔ3B3- Δ3B4 were able to amplify the multiple DNMT isoforms of their respective genes. For primer sets of exogenous DNMTs, they were able to bind to Myc sequence introduced as marker of exogenous sequence. 10 µL 2x QuantiTect SYBR green PCR master mix (QIAGEN, USA), 0.4 µL forward and reverse primers (final concentration 10 µM), 8.2 µL nuclease-free water (QIAGEN, USA), and 1 µL cDNA was prepared and PCR condition was carried out as following (Table 2.10).

Table 2.9 Primer lists of the total and exogenous *DNMT* isoforms.

Genes	Primer sequence
Total *DNMTs*	
hDNMT1	Forward: 5' AGGCGGCTCAAAGATTTGGA 3'
	Reverse: 5' GGGATTTGACTTTAGCCAGG 3'
hDNMT3A1	Forward: 5' AGCGGTGACACGCCAAAGGA 3'
	Reverse: 5' CTTCAGGCAGGGTCTCAGCTG 3'
hDNMT3A2	Forward: 5' AATGCTGTGGAAGAAAACCAG 3'
	Reverse: 5' ATCGCCTGCTTTGGTGGCAT 3'
hDNMT3B1-3B5	Forward: 5' AAGACTCGATCCTCGTCAAC 3'
	Reverse: 5' ATGACTGGGGTGTCAGAGCC 3'
hDNMTΔ3B1- Δ3B2	Forward: 5' TATCAGGATGGGAAGGAGTTT 3'
	Reverse: 5' CACCAGTTTGTCTGCAGAGA 3'
hDNMTΔ3B3- Δ3B4	Forward: 5' TCTCTGCAGACAAACTGGTG 3'
	Reverse: 5' GCTGGTCCTCCAATGAGTC 3'

hDNMT3L	Forward: 5’ TGAGCTCTCAAGCTCCGTTT 3’
	Reverse: 5’ GTAGGATTGGTACCCGTCAT 3’
Target genes (Exogenous *DNMTs*)	
Myc	Forward: 5’ AGAAGCTGATCTCAGAGGAG 3’
Myc-DNMT1	Reverse: 5’ ATCGTCGGGCAGCGAGAT 3’
Myc-DNMT3A1	Reverse: 5’ CGCTCCGCAGCAGAGCT 3’
Myc-DNMT3A2	Reverse: 5’ ATCGTCGGGCAGCGAGAT 3’
Myc-DNMT3B1-3B5	Reverse: 5’ GTTGACGAGGATCGAGTCTT 3’
Myc-DNMTΔ3B1- Δ3B2	Reverse: 5’ CTTCCCATCCTGATACTCTG 3’
Myc-DNMTΔ3B3- Δ3B4	Reverse: 5’ TGCAGAGACCTGATACTCTG 3’
Myc-DNMT3L	Reverse: 5’ CACTGGATCCCACCAAAATC 3’

Table 2.10 qPCR condition.

Procedure	Temperature	Time	Number of cycles
PCR initial activation step	95 °C	15 minutes	1
Denaturation	95 °C	15 seconds	40
Annealing (acquisition of fluorescence)	60 °C	30 seconds	
Extension	72 °C	30 seconds	
Melting curve step	72 - 95 °C	15 seconds	1

Protein extraction

The cell pellets were re-suspended in ice-cold Radioimmunoprecipitation assay (RIPA) buffer (1 mL per 100 mm dish) (Thermo Fisher Scientific, UK) and centrifuged at 13,000g for 20 minutes in a 4 °C precooled centrifuge. A clear supernatant was transferred into a fresh centrifuge tube and kept on ice.

Western blot using c-Myc antibody

Protein concentration was established using a Biorad protein assay kit (BIORAD, USA). Standards were prepared from bovine serum albumin (BSA) prepared as 1 mg/mL stock solution in deionised water (DW). 5 µL of standards or samples were pipetted into microplate. Reagent A was mixed with reagent S and pipetted 25 µL per well into microplate. Reagent B,

then, was added into microplate by 195 µL per well and incubated at room temperature for 15 minutes. Microplate was read for absorbance at 750 nm. Standard protein was plotted, and the unknown protein concentration of the samples was determined from the standard curve.

To cast 8% SDS-polyacrylamide gel electrophoresis (SDS-PAGE), 2.4 mL 30% acrylamide, 2.4 mL 1.5M Tris, pH 8.8, 90 µL 10% SDS, 90 µL 10% ammonium persulfate, 3.6 µL TEMED, and 4.6 mL DW were prepared, poured and left for 45 minutes at room temperature under DW until polymerisation was completed. The water was discarded, and the top of the gel washed with DW. A 5% PAGE (0.43 mL 30% acrylamide, 0.33 mL 1m Tris, pH 6.8, 26.5 µL 10% SDS, 26.5 µL 10% ammonium persulfate, 2.65 µL TEMED, and 1.4 mL DW) was prepared and poured into the resolving gel and a comb inserted and left for 30 minutes at room temperature until polymerisation was completed. 25-70 µg of protein was mixed with 4x LaemmLi sample buffer to make 48 µL final volume. Each sample was boiled at 95 °C for 5 minutes and loaded into the gel. The gel was run at 150 volts for 1 hour.

Gel was placed in 1x transfer buffer for 10 minutes and blotting pads were soaked in 1x transfer buffer. The membrane was soaked in methanol until no more bubbles were visible and washed in 1x transfer buffer. The transfer sandwich was assembled by gel and membrane in the middle of the sandwich without any bubbles and placed on the Trans-Blot SD semi-dry transfer cell and run at 17 volts for 1 hour.

Membrane was washed by DW and stained with Ponceau S solution for 5 minutes and washed by DW to check the transfer quality. Proteins were observed on the membrane. Membrane was washed by 1x PBS for 10 minutes, PBS-Tween (1x PBS mixed with 0.05% Tween 20) for 10 minutes, and the last with 1x PBS for 10 minutes. After that, the membrane blot was placed in 100 mL of blocking solution (5% non-fat dry milk) and incubated at room temperature for 1 hour with gentle shaking. Blocking solution was removed and added primary antibody, which was c-Myc antibody diluted (1:10000) in 5% milk and incubated at 4 °C overnight with gentle shaking. The blot was washed in 1x PBS for 10 minutes, PBS-Tween (1x PBS mixed with 0.05% Tween 20) for 10 minutes, and the last with 1x PBS for 10 minutes. The second antibody, then, was added and incubated at room temperature for 1 hour with gentle shaking followed by washing. The chemiluminescent substrate (Bio-rad, USA) was applied to the blot and captured the chemiluminescent signals using a Fujifilm LAS4000 luminescence imager (Fujifilm Life Science, USA).

2.2 DNA methylation microarray with DNMT overexpressing cells

2.2.1 DNA extraction and the Illumina Methylation EPIC array

Total and exogenous expression levels of DNMT isoforms were determined for all single cells from step 2.1.6. Each duplicated DNMT overexpressing cell was selected from the exogenous *DNMT* expression, which showed the similar levels of exogenous *DNMT* and DNA methylation levels were measured using Infinium Methylation EPIC arrays (Illumina, USA).

Firstly, DNA was isolated from DNMT overexpressing cells using QIAamp® DNA mini and blood mini kit (QIAGEN, USA). Cells were harvested and the number of cells per sample was estimated ($5 \times 10^6 – 1 \times 10^7$ cells). 20 µL QIAGEN protease were added into a tube containing the cells. 200 µL AL were added to sample tube and mixed, then it was incubated at 56 °C for 10 minutes. 200 µL ethanol (100%) were added and mixed, then transferred to a QiAamp Mini Spin column with collection tube. The column was centrifuged at 6,000g for 1 minute and the eluate was discarded. 500 µL buffer AW1 were added into the column and centrifuged at 6,000g for 1 minute. Then, 500 µL buffer AW2 were added into the column and centrifuged at 13,000g for 3 minutes. The columns were transferred to new centrifuge tubes and the DNA was eluted by centrifuging at 13,000g for 1 minute with 50 µL of AE. DNA quantity and purity were assessed by Nanodrop1000. DNA was kept at -20 °C.

Bisulphite converted DNA was hybridised to Infinium Methylation EPIC arrays (Illumina, USA) to measure DNA methylation in more than 850,000 CpG sites across the genome. All DNA methylation array processing was conducted at the Eurofins Genomics.

2.2.2 Data analysis for Infinium Methylation EPIC data

Unlike the Illumina Infinium Human Methylation27 BeadChip, in which only one probe type is measured, the Illumina Infinium Methylation EPIC BeadChip includes two probe types, Infinium I (n = 142,262) and Infinium II (n = 721,642) (269). EPIC microarray still contains 93.3% of loci contained on the Illumina Infinium Human Methylation450 BeadChip (269). The direct output from Illumina iScan system is an IDAT file which contained BeadArray data. The diagram of the framework is shown in Figure 2.3 and the detail of each step will be introduced below.

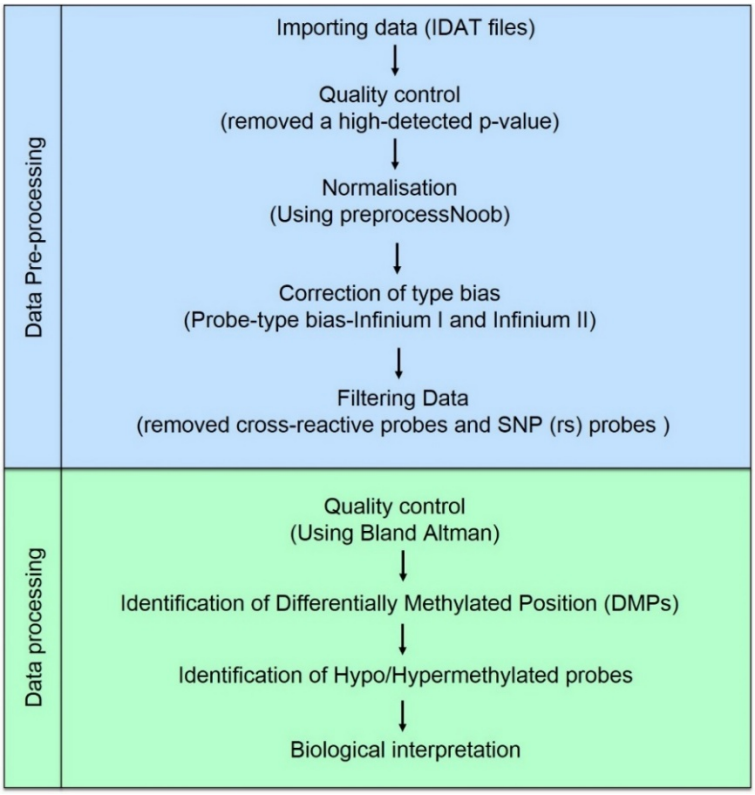


Figure 2.3 A framework of Illumina Methylation EPIC array data analysis.

All IDAT files were imported and analysed in R studio using the Bioconductor package. The quality of data was evaluated by filtering out a high-detected *p*-value (*p*-value > 0.01). Normalisation concern the removal of experimental artefact, technical and systematic variation, and random noise caused by microarray technology. PreprocessNoob was utilised for normalisation and 866,091 probes was left after this process. This step assesses and accounts for variation that is not caused by biological differences but by external variation. EPIC BeadChips use two fluorescent dyes that are linked to the nucleotides used in the single-base extension step. A and T nucleotides are linked with a red dye (the red color channel), G and C nucleotides are linked with a green dye (green color channel). Uncorrected data usually feature higher intensities in the red color channel was called dye bias. For probes of Infinium type II design, which use separate color channels to measure the methylated and unmethylated signal, this results in a shifted distribution of β-values but probes of Infinium design type I are not affected, as they measure both signals in the same color channel. Dye-bias correction normalizes the red and green color channel. The 866,091 probes were adjusted for probe-type bias Infinium I (type I) and Infinium II (type II) (270). The cross-reactive probes (43,254 loci) were also removed from this data (271). Moreover, EPIC microarray contained 59 explicit SNP probes ("rs" probes) (272) and these probes were removed from the dataset. Finally, 814,341 probes remained after quality control and filtering.

Bland Altman plot (install.Packages ("BlandAltmanLeh")) was used to identify the mean difference between biologically duplicate samples. The 95% limits of agreement were calculated (mean ± 1.96 standard deviation) to select only those CpG sites showing high correlation of DNA methylation between biological duplicate samples. The Δβ was calculated by subtracting the DNA methylation levels at each CpG site in the cell over-expressing each DNMT isoform from the corresponding DNA methylation of the control cell (Myc cell). Hypo/hypermethylated probes were identified in each DNMT overexpressing cell. Differentially Methylated Positions (DMPs) were explored across DNMT isoform dataset. The cut-offs of Δβ were set at $\Delta\beta \geq 0.4$, or ≥ 0.3 or ≥ 0.2 and $\Delta\beta \leq -0.4$, or ≤ -0.3, or ≤ -0.2 with FDR-adjusted *p*-value ≤ 0.05. FDR can be calculated by generating the empirical null distribution of test statistics, it is most commonly applied using the approach introduced by Benjamini and Hochberg. However, $\Delta\beta \geq 0.4$ and ≤ -0.4, and FDR-adjusted *p*-value ≤ 0.05 were used to select the target CpG sites of each DNMT as these cut-offs were higher than usual cut-offs in coronary heart disease (273) and cancer study (274). DNA methylation patterns by genomic location and CpG density were examined in DNMT overexpressing dataset. All hypo/hypermethylation of DMPs were analysed using pathway analysis known as functional enrichment analysis by Ingenuity Pathway Analysis (IPA) (QIAGEN, USA) and the Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.8 (275). Canonical pathway analysis identified from IPA Knowledge Base that were most significant to the data set gene that met the FDR ≤ 0.001 and post hoc *p*-value < 0.001 cut-off and were correlated to a canonical pathway in IPA, was considered for the analysis. KEGG pathway was analysed using DAVID.

2.3 Assessing specificity and sensitivity between dietary factors and DNMT isoform using *in vitro* model with measuring DNA methylation changes by pyrosequencing

2.3.1 Cell viability and dietary compound dose selection

Cells overexpressing each DNMT separately were seeded on opaque-walled multi-well plates, 6 well plates, in culture medium (DMEM- high glucose (Sigma, UK) supplemented with 10% FBS (Sigma, UK), and 2 mM glutamine, at 37 °C in a 5% CO_2 humidified atmosphere). Aqueous solution of vitamin C and dietary polyphenols (curcumin, theaflavin, CA, and EGCG) were prepared in several concentrations, *i.e.* 10, 25, 50, 100, and 200 µM. Cytotoxicity was measured using CellTiter- Glo® 2.0 assay (Promega, USA), according to manufacturers' instruction. Briefly, 70,000 cells were subjected to selected concentration of

vitamin C, curcumin, theaflavin, CA, or EGCG and cells were incubated at 37 °C in a 5% CO_2 humidified atmosphere for several incubation times, *i.e.* 12, 24, 48, and 72 hours. This reaction between cells and each concentration of compounds was equilibrated at room temperature for 30 minutes, then; a volume of CellTiter-Glo® Reagent was added equal to the volume of cell culture medium in each well. The plate was shaken to induce cell lysis and incubated at room temperature for 10 minutes to stabilise the luminescent signal. The plate was read, and luminescence was recorded. The half-maximal inhibitory concentration (IC_{50}) was calculated for each compound.

2.3.2 Treatment of cells, overexpressing specific DNMTs, with vitamin C and dietary polyphenols

The optimum concentrations of each dietary compound were selected based on percentage of cell viability, which was higher than IC_{50}: vitamin C (100 and 200 µM), curcumin (10 and 25 µM), CA (100 and 200 µM), EGCG (50 and 100 µM), and theaflavin (80.5 and 161 µM). Cells overexpressing DNMTs were cultured with the specific concentration of each dietary compound for 48 hours at 37 °C in a 5% CO_2 humidified atmosphere.

2.3.3 Quantification of DNA methylation specific CpG sites by pyrosequencing

2.3.3.1 DNA extraction and bisulphite conversion

DNA was isolated using QIAamp® DNA mini and blood mini kit (QIAGEN, USA) described in 2.1.2. DNA samples were treated with bisulphite under alkaline conditions using the EZ DNA methylation-Gold™ kit (Zymo research, USA) which coverts unmethylated cytosine residues in DNA into uracil. 500 ng of DNA were prepared and the volume was adjusted to 20 µL with nuclease-free water. 130 µL of CT conversion reagent solution were added into the DNA samples in PCR tube. The mixture was mixed and centrifuged briefly to ensure no droplets were in the cap or on the sides of the tube. The PCR tube was incubated in a thermocycler at 98 °C for 10 minutes, 64 °C for 2.5 hours, and 4 °C for storage up to 20 hours. 600 µL of M-Binding buffer were added into a Zymo-spin IC column fitted with a collection tube. The bisulphite-treated samples were loaded into the Zymo-spin IC column containing M-Binding buffer, which was mixed by inverting the column several times following by centrifugation at 13,500g for 30 seconds. The flow-through was discarded. 100 µL of M-Wash buffer were added into the column and the column was centrifuged at 13,500g for 30 seconds. 200 µL of M-Desulphonation buffer were added into the column and let stand

at room temperature for 15-20 minutes. After incubation, the column was centrifuged at 13,500g for 30 seconds. 200 µL of M-Wash buffer were added into the column and the column was centrifuged at full speed for 30 seconds and this step was repeated once more. The column was placed into a 1.5 mL microcentrifuge tube and 20 µL of M-Elution buffer were added directly to the column matrix. The column was centrifuged at full speed for 30 seconds to elute the DNA.

2.3.3.2 Pyrosequencing

PCR master mix was prepared as the following protocol (Table 2.11 and 2.12). All primers used in this experiment, were designed using MethPrimer and the original design was showed in Appendix C.

Table 2.11 Volumes of the reagents for PCR.

Reagent	Volume per reaction
2x GoTaq® Green Master Mix (Promega, UK)	12 µL
Forward primer (10 µM) (Table 2.13)	1 µL
Biotinylated reverse primer (10 µM) (Table 2.13) (This primer will be captured by streptavidin-sepharose beads)	1 µL
Nuclease-free water (QIAGEN, USA)	9 µL
Bisulphite-treated DNA	1 µL
Total	24 µL

Table 2.12 PCR condition.

Procedure	Temperature	Time	Number of cycles
Pre-denaturation	95 °C	10 minutes	1
Denaturation	95 °C	40 seconds	40
Annealing	56 °C	40 seconds	
Extension	72 °C	30 seconds	
Final extension	72 °C	5 minutes	1
	4 °C	Hold	

DNA methylation levels were quantified for all PCR products by pyrosequencing, which detects and calculates the percentage of methylation at individual CpG as a ratio of cytosine:thymine. 10 µL of PCR were mixed with 2 µL of streptavidin-sepharose beads, 40 µL of 1x binding buffer, and 28 µL of water, in a PCR plate. The PCR plate was sealed with polyester PCR sealing (Starlab, UK) and shaken at room temperature at 1,400 rpm for 10 minutes. 11.64 µL of annealing buffer was mixed with 10 µM sequencing primer into a white PSQ-HS plate (QIAGEN, USA). This plate was transferred to the vacuum prep station and the beads with PCR products were captured by slowly lowering the vacuum prep tool into the wells allowing the solution to flush through the probe tips for 10 seconds. The vacuum prep tool was transferred into 70% ethanol, denaturing solution, and wash buffer, then, the vacuum prep tool was disconnected from the station. The vacuum prep tool was placed into the PSQ-HS plate to release the PCR products from the probe tips by gentle shaking the vacuum prep tool for 30 seconds. The PSQ-HS plate was sealed and placed on the 80 °C heat block for 2 minutes and allowed it to cool at room temperature for 10 minutes. Pyrosequencing was performed by PyroMark Q96 MD (QIAGEN, USA).

Table 2.13 Primer lists of potential CpG targets of DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMT1, and DNMT3L.

Illumina probe ID	Target of DNMT	Primer sequence
cg02732111	DNMT3A2	Forward: 5’ TTGTTTAGGTTTATTATAGTTTG 3’
		Reverse: biotin-5’ TCAATAACACATTTCAACAAATAC 3’
		Sequencing: 5’ TTGTTTAGGTTTATTATAGTTTG 3’
cg16204524	DNMT3A2	Forward: 5’ ATTTTGTTATTAAGTGATGTATGATTGTAT 3’
		Reverse: biotin-5’ ACCTCCTAAAATAAAATTTAAAAAC 3’
		Sequencing: 5’ ATTTTGTGAATTTTAAATT 3’
cg02788195	DNMT3B4	Forward: 5’ GGTTATTGTAAAAATAGATTTAGTTAGATT 3’
		Reverse: biotin-5’ AATCTCCTTCCCATTACCTTTTATTAA 3’
		Sequencing: 5’ AAGTTTTATTTAGTAGATA 3’
cg26286826	DNMT3B4	Forward: 5’ TTTTTTAAAGTGTTGGGATTATAGG 3’
		Reverse: biotin-5’ CTAAACCAACTAAAAAAATCCTCTC 3’

		Sequencing: 5’ GGTTAGAGTATGAATTA 3’
cg25533247	DNMTΔ3B2	Forward: 5’ AGAGATTTTGTAATAGTGTAGT 3’
		Reverse: biotin-5’ ATACAACTCTATCATCTCTAAA 3’
		Sequencing: 5’ AGAGATTTTGTAATAGTGTAGT 3’
cg21808287	DNMTΔ3B2	Forward: 5’ TTTTATGTTATGATTTTTTAATTTG 3’
		Reverse: biotin-5’ CTAAAAACAACCCTTAAACTACA 3’
		Sequencing: 5’ TATATTTGTGAAATAAGGTGG 3’
cg20364776	DNMTΔ3B3	Forward: 5’ TGAAAATTATTTTATTTATAAGTTAGAA 3’
		Reverse: biotin-5’ ACTTAAAAAAACACTTTCCCATCTC 3’
		Sequencing: 5’ TGAAAATTATTTTATTTATAAGTTAGAA 3’
cg08927738	DNMTΔ3B3	Forward: 5’ AGGTGGTGTTTTGAAGTTAGTAGATAGA 3’
		Reverse: biotin-5’ CTTCCTAATATAAAACTACCCTCCCA 3’
		Sequencing: 5’ ATTATAAAATTTTATAGAA 3’
cg01065960	DNMT1	Forward: 5’ AGGTTAGGTTTTTGGAAGGAG 3’
		Reverse: biotin-5’ CCTCCTTTACAAACCCTCTAA 3’
		Sequencing: 5’ AGGTTAGGTTTTTGGAAGGAG 3’
cg20540357	DNMT3L	Forward: 5’ ATTGATTATTAGGATTATGTTTGG 3’
		Reverse: biotin-5’ AAACCACCACCCACACTCAT 3’
		Sequencing: 5’ ATTGATTATTAGGATTATGTTTGG 3’
cg12150401	DNMT3L	Forward: 5’ TGGGTAGAGAATGGTTGTAAG 3’
		Reverse: biotin-5’ CCCAAATAATTATTAAATTACAAAAT 3’
		Sequencing: 5’ TTATTAGTTTGGGTATTT 3’
cg25843713	DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, and DNMT1	Forward: 5’ TTTGTTTAGTGTTTTTAAGGGTTTT 3’
		Reverse: biotin-5’ ACTATCTTATATCACCATTTCCCTC 3’
		Sequencing: 5’ GGTTTTGTTGTTATTTTTAT 3’
cg04458645		Forward:

	DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, and DNMT1	5’ TTTTTGAGTGATAATTTAAGAGAAGTAAGA 3’
		Reverse: biotin-5’ ATAAACCAAAATCACACCACTACAC 3’
		Sequencing: 5’ TTTTTGAGTGATAATTTAAGAGAAGTAAGA 3’

2.4 DNMT activity/inhibitor assay

2.4.1 Protein extraction

Cells overexpressing DNMTΔ3B4 were cultured in 100 mm dish for 20 dishes until reaching 80% confluence. Cells were lysed using Mammalian protein extraction reagent (M-PER) (Thermo Fisher Scientific, UK). Lysate was collected and transferred into a microcentrifuge tube. The lysate was then centrifuged at 14,000g for 10 minutes. The clear supernatant was transferred to a new tube and kept at -80 °C.

2.4.2 Immunoprecipitation/Co-immunoprecipitation (IP/Co-IP)

The exogenous DNMT protein tagged with c-Myc was isolated using Pierce c-Myc-Tag IP/Co-IP kit (Thermo Fisher Scientific, UK) according to manufacturers’ instruction. The bottom plug was placed on the Pierce Spin Column and 200 µL of protein lysate were added to the spin column. The anti-c-Myc agarose was thoroughly resuspended by inverting the vial several times immediately before dispensing. 10 µL of anti-c-Myc agarose slurry (5 µg anti-c-Myc antibody) were dispensed into each labelled spin column using a wide-bore pipette tip. The vial mixture was incubated with gentle end-over-end mixing for overnight at 4 ^{o}C. A wash solution of Tris Buffered Saline plus 0.05% Tween 20 (TBS-T) was prepared. The top cap on the column was loosened and then the bottom plug was removed. A collection tube was put under the column and centrifuged for 10 seconds. The flow-through was discarded. 0.5 mL of TBS-T was added to column and the cap was screwed following by gently inverting the column with collection tube 2-3 times. The column was centrifuged for 10 seconds. The flow-through was discarded and this step was repeated 2 times. 500 µL of 1x conditioning buffer were added into column to wash the resin. The spin column was placed in the new collection tube and 10 µL of Elution buffer were added into the anti-c-Myc agarose with loose screwing on the cap and mixing. The column was centrifuged for 10 seconds and this step was repeated 2 times.

2.4.3 DNMT inhibition assay

Inhibition of DNMTΔ3B4 was quantified using DNMT activity assay (Abcam, UK) according to manufacturers' instruction. Briefly, the blank wells were prepared by adding 50 µL of AdoMet working buffer per well. The positive control wells were prepared by adding 50 µL of AdoMet working buffer and 1 µL of DNMT Enzyme Control per well. The sample wells without inhibitor were prepared by adding 45 µL of AdoMet working buffer and 5 µL of purified DNMT enzymes per well. The sample wells with CA were prepared by adding 40 µL of AdoMet working buffer, 5 µL of purified DNMT enzyme, and 5 µL of inhibitor solution per well. The strip-well microplate was tightly covered with the adhesive covering film to avoid evaporation and incubated at 37 ºC for 90 minutes. After incubation, the reaction solution was removed from each well and washed with 150 µL of 1x wash buffer each time; this process was repeated 3 times. 50 µL of the diluted capture antibody were added into each well, then the plate was covered with Parafilm M and incubated at room temperature for 1 hour. The diluted capture antibody solution was removed from each well and 150 µL of the 1x wash buffer solution were added to wash each well for 3 times. 50 µL of the diluted detection antibody were added to each well and the plate was covered with Parafilm M and incubated at room temperature for 30 minutes. The diluted detection antibody solution was removed from each well and 150 µL of the 1x wash buffer solution were added to wash each well 4 times. 50 µL of the diluted enhancer solution were added to each well and the plate was covered with Parafilm M and incubated for 30 minutes. The diluted enhancer solution was removed from each well. 150 µL of 1x wash buffer solution were added to wash each well for 5 times. 100 µL of developer solution were added to each well and incubated at room temperature for 10 minutes away from direct light. The developer solution then turned to blue colour in the presence of enough methylated DNA. 100 µL of stop solution were added to each well to stop enzyme reaction. The colour changed to yellow after adding the stop solution and the absorbance was read on a Thermo Scientific Muliskan GO microplate spectrophotometer (Thermo Fisher Scientific, UK) within 2 to 10 minutes at 450 nm with an optimal reference wavelength of 655 nm.

2.5 Statistical analysis

All statistical analyses were performed in IBM® SPSS statistical software programme (version 24) and R studio (version 1.1.442). Data were presented as mean ± standard deviation (SD) from three independent experiments for RT-PCR, cell viability, pyrosequencing, and DNMT inhibitory assay, and two biological duplicates for DMPs. A *p*-

value ≤ 0.05 was considered statistically significant. Also, the *p*-values were adjusted for multiple hypothesis testing by the Benjamini-Hochberg method, with $P_{FDR} \leq 0.05$ considered significant. Paired-sample *t*-tests were used to identify DMPs. Two-sample Kolmogorov-Smirnov test was used to identify significant changes of DNA methylation, global methylation and DNMT inhibition of selected cells overexpressing DNMTs after treatment with dietary constituents. Additionally, ANOVA test was applied to test the difference β values before and after treatment with different concentration of dietary constituents.

Chapter 3: Generation of stable overexpressing DNMT cell lines

3.1 Introduction

DNMT1, DNMT3A, and DNMT3B are enzymatically active DNA methyltransferases, while DNMT3L is an enzymatically inactive regulatory factor (87). Additionally, DNMT1, DNMT3A, and DNMT3B are essential for survival, for example, knockout alleles of *Dnmt1* and *Dnmt3b* in mice leads to embryonic lethality (101, 276). Alternative splicing and/or promoter usage are the main factors for different isoforms of DNMT3A (84) and DNMT3B (85). More than 30 isoforms of DNMT3B were discovered in human and mouse cells (105-107, 277) and some isoforms (DNMT3B3, DNMT3B4, and DNMT3B5) were found in cancer cells rather than normal cells (85, 106, 278). The structures of the most common DNMT isoforms are shown in Figure 1.6 (see Introduction section). DNMT1 is a multi-domain enzyme composed of a *C*-terminal catalytic domain and a *N*-terminal regulatory domain that consists of DNMT1-associated protein (DMAP) binding, replication foci, zing binding, and bromo-adjacent homology (BAH) (279, 280). DNMT3A consists of PWWP domain, ATRX-DNMT3-DNMT3L (ADD) domain, and *C*-terminal catalytic domain but the DNMT3A2 lacks of 200 amino acids at the *N*-terminal regulatory domain of DNMT3A1 (279, 280). DNMT3B also contains *C*-terminal catalytic domain and the *N*-terminal regulatory domain with alternative splicing in both domains (279).

In mouse cells, overexpression of Dnmt1 resulted in increased methylation of the imprinted regions of *Insulin-like growth factor 2* and *H19* and, more generally, genomic hypermethylation (281). DNMT1 overexpression occurs frequently in older cancer (breast and lung) patients with advanced clinical stages as well as unfavourable prognosis (282). DNMT3B overexpression also contributes to hypermethylation in breast cancer (283). Overexpression of *DNMT3A* and *DNMT3B* increased DNA methylation levels at CpG island and non-CpG islands, while overexpression of *DNMT3B3*, *DNMT3B4*, and *DNMT3B5* did not change DNA methylation in those regions (87). In $Apc^{Min/+}$ mice, overexpression of Dnmt3b1 enhanced the number of colon tumours and increased the size of colonic microadenomas with loss of imprinting (284). Overexpression of the *DNMT3B4* isoform is correlated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis (285). DNA methylation of Long Interspersed Nuclear Element 1 (LINE-1) is increased by the overexpression of *DNMT3B1, DNMT3B2*, and *DNMTΔ3B* isoforms, but not by overexpression of *DNMT3A* isoforms (87).

Choi SH. *et al.* unveiled that the 13 most common DNMT isoforms have both specificity and overlap in their DNA methylation target profiles (87). They found that each

DNMT isoform changed methylation of a different number of CpG sites, and that DNMT3A1 and DNMT3B1 showed the most distinct DNA methylation patterns. In contrast, DNMT3B3, DNMT3B4, and DNMT3B5 induced no methylation changes and these isoforms seemed to be inactive DNMTs. Furthermore, Choi SH. *et al.* (2011) found the specific target genes of DNMT3A1 were correlated with H3K4me3 modification, while the specific target genes of DNMT3B1 were correlated with H3K27me3 modification. However, their study investigated DNA methylation patterns in cells that over-expressed each DNMT isoform using the Illumina GoldenGate Methylation Cancer Panel I, which detects the DNA methylation of 1,505 CpG sites only. Therefore, given that there are 28 million CpG sites in the human genome (286), the information available on target CpG sites of each DNMT isoform by using this technique is limited. High-throughput sequencing is a cutting-edge method that offers more extensive interrogation of methylation across the human genome. The Illumina Infinium Methylation EPIC BeadChip is the latest version of this array-based approach and measures methylation at more than 850,000 CpG sites. Use of such EPIC arrays allows investigation of more potential CpG sites that are targets of specific DNMT isoforms.

The locus-specific methylation by individual DNMT isoforms is important information that might be applied to manipulate patterns of DNA methylation. This data can be applied to reveal the interaction of nutrients and individual DNMTs by measuring the DNA methylation changes on the locus-specific CpGs of each individual DNMT. Moreover, the locus-specific loci of DNMTs suggests a possibility of a mechanism of cancer or other diseases by which DNA methylation patterns are measured. From Choi's study (87), *EYA Transcriptional Coactivator And Phosphatase 4 (EYA4)* and *Homeobox A11 (HOXA11)* were reported as the target genes of the DNMT3A1, and *IGF2 Antisense RNA (IGF2AS)* and Cadherin 11 (*CDH11)* were reported as the target genes of the DNMT3B1. Although, the aberration of DNA methylation patterns in each DNMT was revealed, the target sites of individual DNMT sub-isoforms are poorly understood.

In the present study, I investigated the locus-specific target CpGs of each DNMT isoform by using molecular cloning to generate cell lines that over-expressed each of individual DNMT isoform. The efficacy of the overexpression was investigated by quantification of the expressed genes at the mRNA and protein levels by qPCR and western blot, respectively.

3.2 Hypotheses

The hypotheses for this study were:

1. The mRNA of DNMT isoform can be integrated into MEG-01 and HEK293T genomes using transient transfection or viral system.
2. DNMT-transduced cells express the specific exogenous DNMTs in both RNA and protein levels.

3.3 Aim

- To generate the stable transduced single-cell derived clonal lines in order to identify *de novo* DNA methylation target sites of specific DNMT isoforms.

3.4 Objectives

The objectives of this study were:

1. To transport the mRNA of each of the *DNMT* isoforms into MEG-01 and HEK293T cells using transfection reagents and viral system.
2. To select the cells that over-expressed each DNMT using a GFP marker.
3. To generate single cell clones that over-expressed each DNMT isoform and to quantify expression of each specific DNMT in the selected clone at the mRNA and protein levels.

3.5 Overview of the methods

A detail description of the experimental procedures and methods for molecular cloning (2.1.3 Plasmid construction for viral system, page 30), generating stable cell lines overexpressing individual DNMT isoforms (2.1.4 Viral production, page 36), quantifying exogenous mRNA and protein of each DNMT (2.1.7 Confirmation the positive cells by qPCR and western blot assays, page 37) can be found in the Method chapters.

In brief, DNMT-pIRES puro3 plasmids were retransformed and extracted from *E.coli* DH5α cells (see section 2.1.2 in Methods section, page 29). All plasmids were sub-cloned into the pLenti7.3/V5 DEST gateway vector with modifying at multiple cloning sites (see section 2.1.3 in Methods section, page 30). DNMT-pIRES puro3 plasmids were transported into MEG-01 cells using transfection reagents. Lentivirus was produced using the modified pLenti7.3/V5 DEST gateway vector (see section 2.1.4 in Methods section, page 36). This virus incorporated the specific DNMT into MEG-01 and HEK293T genomes. GFP and Myc-tag were identified as a marker for the successful transfection or transduction. Positive cell

lines, which overexpress each of the DNMTs individually, were detected using fluorescent microscopy and sorted using FACS (see section 2.1.5 (page 37) and 2.1.6 (page 37) in Methods section). Single clone of each DNMT was generated by a single cell after cell sorting. Expression of the DNMT isoform was quantified at the mRNA level using qPCR and at the protein level using western blotting in all GFP positive cells (see section 2.1.7 in Methods section, page 37).

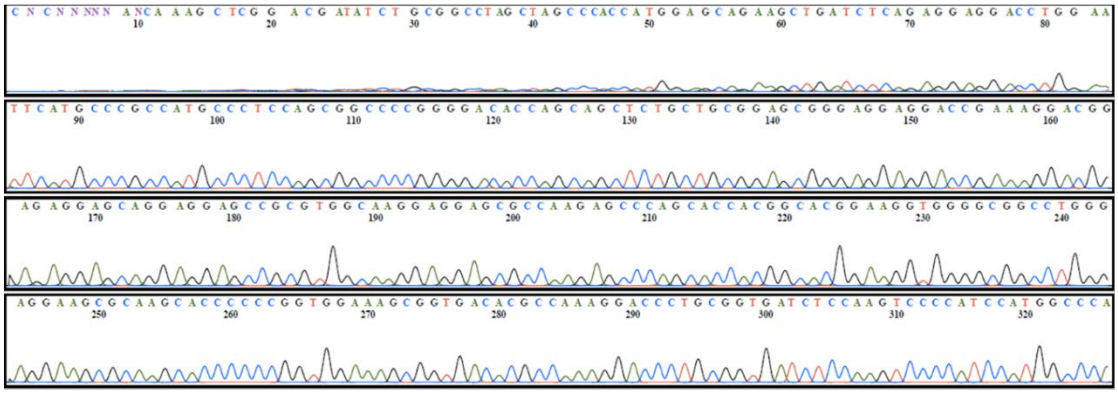
3.6 Results

3.6.1 Molecular cloning of DNMT isoforms

3.6.1.1 Retransformation and confirmation of DNMT sequences

All 14 DNMT-pIRES puro3 plasmids containing *DNMT3A1, DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT3L, DNMT1*, or Myc sequence only (control), were re-transformed into *E.coli* DH5α. These plasmids were extracted from the bacteria clones grown on an ampicillin selective agar. To confirm the sequence of each isoform, these plasmids were sent to MRC PPU DNA Sequencing and Services at Dundee University. All sequences were aligned with the relevant DNA sequence templates from the commercial vectors for each DNMT isoform using the DNASTAR Lasergene bioinformatics software. Example results are shown in Figure 3.1 and in Appendix A.

A



B

```
Majority -CCCXXXTXXXXACAGAGCTCGGGX-TCGA-TATCTGCGGCCTAGCTAGCCCACCATGGAGCAGAAGCTGATCTCAGAGG
                    10        20        30        40        50        60        70        80
3B1-T7   --CCNNNTTTNNACAGAGCTCGGGA-TCGA-TATCTGCGGCCTAGCTAGCCCACCATGGAGCAGAAGCTGATCTCAGAGG 76
3B2-T7   -CNNCNNTNNAANCAGAGCTCGGA---CGA-TATCTGCGGCCTAGCTAGCCCACCATGGAGCAGAAGCTGATCTCAGAGG 75
Myc-T7   CCCCTGTAACCCAAAAACCTCGGGGATCGAATAACTGCGGCCTAGCTAGCCCACCATGGAGCAGAAGCTGATCTCAGAGG 80

Majority AGGACCTGGAATTCATGAAGGGAGACACCAGGCATCTCAATGGAGAGGAGGACGCCGGCGGGAGGGAAGACTCGATCCTC
                   90       100       110       120       130       140       150       160
3B1-T7   AGGACCTGGAATTCATGAAGGGAGACACCAGGCATCTCAATGGAGAGGAGGACGCCGGCGGGAGGGAAGACTCGATCCTC 156
3B2-T7   AGGACCTGGAATTCATGAAGGGAGACACCAGGCATCTCAATGGAGAGGAGGACGCCGGCGGGAGGGAAGACTCGATCCTC 155
Myc-T7   AGGACCTGGAATTC--G------GATCC---GCGGCCGCATAGATAACTGATCCAGTGTGCTGGAATTAATTCGCTGTCT 149

Majority GTCAACGGGGCCTGCAGCGACCAGTCCTCCGACTCGCCCCCAATCCTGGAGGCTATCCGCACCCCGGAGATCAGAGGCCG
                   170       180       190       200       210       220       230       240
3B1-T7   GTCAACGGGGCCTGCAGCGACCAGTCCTCCGACTCGCCCCCAATCCTGGAGGCTATCCGCACCCCGGAGATCAGAGGCCG 236
3B2-T7   GTCAACGGGGCCTGCAGCGACCAGTCCTCCGACTCGCCCCCAATCCTGGAGGCTATCCGCACCCCGGAGATCAGAGGCCG 235
Myc-T7   GCGAGGGCCAGCTGTTGGGGTGAGTACTCC----CTCTCAAAAGCGGGCATGACTTCTGCGCTAAGATTGTCAGTTTCCA 225

Majority CCCCCACGGGGGGCTTTCAAGCCCCACTGGGGAATCGAAGGGCCTTCAAACCCCAACAACACGCCACCCAGTGGTTAATA
                  1130      1140      1150      1160      1170      1180      1190      1200
3B1-T7   CCC--ACGGGGGGCTT-CAAGCCC-ACTGGGGA-TCGAAGGGCCT-CAAACCCCAACAACACGCCACCCAGTGGTTAATA 1174
3B2-T7   NCCCCACGGGGGGCTTTCAAGCCCCACTGGGGAATCGAAGGGCCTTCAAACCCCAACAAC------------------- 1174

Majority AAGTCGAAAGGGGCGTCCGTGCATGCCGTAAGGAAATTAAAAATCCAGGGAAATACCGAGAAACAAGAACTCCGAAGAAC
                  1210      1220      1230      1240      1250      1260      1270      1280
3B1-T7   AAGTCGAAAGGGGCGTCCGTGCATGCCGTAAGGAAATTAAAAATCCAGGGAAATACCGAGAAACAAGAACTCCAAAGAAC 1254
3B2-T7   --------------------CACGCC-------------AACCCAG----------AGAAACAAGAACTCCGAAGAAC 1209

Majority GCCCCGGTTGGACGAATTCACGCCACCTTCTGGATTACTGGCCCCGGACCCCAACGCGCCTTCAAGAACAAATTTGCTAT
                  1290      1300      1310      1320      1330      1340      1350      1360
3B1-T7   GCCCAGGTTGAA-GAACTCACCCCACCTT-TGGACTAATGGCCCCGGACCCCAACG-GCCTTCAAGA-CAAATTTGCTAT 1330
3B2-T7   GCCCCGGCTGGACGAATTCAAGCCACCTTCTGANTTACTGGCCCCGGAACCCAAAGCGCCTTCAAGAACAAATTGGCTAT 1289

Majority AACCAACGGGCCAAAGAACCGGAGGGGGATGGAAGAATCCAAATCCTGGAAAAACAAATGGGCTTTCCGAATGGTTGGCC
                  1370      1380      1390      1400      1410      1420      1430      1440
3B1-T7   AAACAACGGGA-AAAGAACCG-AGGGGGATGAAA-AATCCAAATCCTAGAAAAA--AATGGGCTT--CAAAAGGTTGGCC 1403
3B2-T7   AACCAACCGGCCAAAGAACCGGAGGGGGATGGAAGAATCCAAAACCCGAAAAAACAAATGGGCTTTCCGAATGGTTGGCC 1369
```

Figure 3.1 Sequence analysis and alignment analysis. A) The sequence of DNMT3A1 from Sanger sequencing and B) the alignment among DNMT3B1, DNMT3B2, and Myc; majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

3.6.1.2 Sub-clone DNMT isoform into virus system

pLenti7.3V5-DEST was modified to add the restriction sites for NheI, ClaI, EcoRI, SwaI, and PspOMI. pLenti7.3V5-DEST was cut by XbaI and MluI, and the digest was run on gel electrophoresis. Three bands from original plasmid (sizes: 61bp, 1,781bp, and 7,788 bp) were separated on the gel and the largest band was extracted from the gel to be used as new backbone (Figure 3.2).

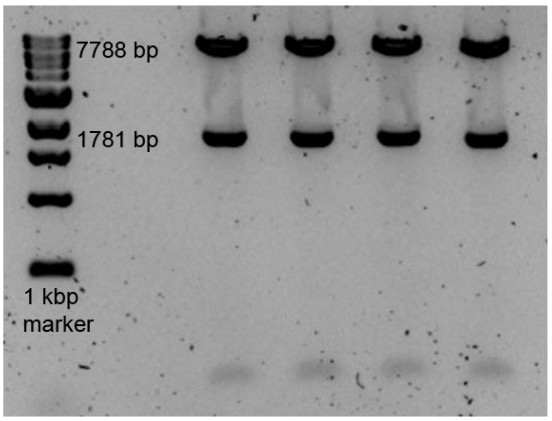


Figure 3.2 pLenti7.3/V5-DEST backbone after being cut by XbaI and MluI. Three bands (61, 1781, and 7,788 bp) were separated on the gel electrophoresis (four replicate samples).

The pLenti7.3V5-DEST backbone was ligated with oligonucleotides containing the new restriction sites (NheI, ClaI, EcoRI, SwaI, and PspOMI). Plasmids were extracted from bacteria grown on an ampicillin selective agar. For screening for positive clones, these plasmids were cut by EcoRI and KpnI, which resulted in three bands on the gel electrophoresis (sizes: 647, 1,093, and 6,086 bp), whereas the negative plasmid was cut into four bands (sizes: 481, 1,093, 1,977, and 6,097 bp) (Figure 3.3).

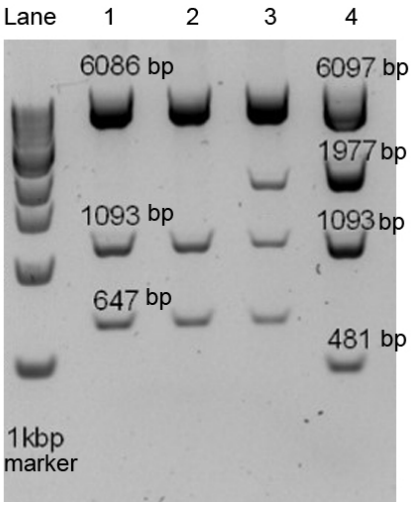


Figure 3.3 Screening new backbone of pLenti7.3/V5-DEST with new restriction sites by restriction enzymes; EcoRI and KpnI. Lane 1 and 2 showed positive clones (there were three bands: 674, 1,093, and 6,086 bp). Lane 3 and 4 showed negative clones.

The modified pLenti7.3/V5-DEST (positive clone) bacteria were grown and plasmids were extracted using the E.Z.N.A.® Endo-free Plasmid DNA Mini KitII. These plasmids were cut by NheI and PspOMI, which produce compatible cohesive ends with NotI, to make a new backbone. After running on the gel electrophoresis, there was only one band on the gel (Figure 3.4A). All pIRESpuro3-DNMTs were also cut by NheI and NotI to generate the sticky ends of DNMT sequences. Each of the different inserts gave a different size on the gel electrophoresis and these were separated completely from the starting plasmid (pIRESpuro3) (Figure 3.4B) apart from DNMT1 which could not be separated from the backbone. DNMT1 was repeated and run on a low percentage (0.7%) gel electrophoresis. Although, there were non-specific bands, this showed that the DNMT1 product was separated into three bands (Figure 3.4C). Both backbone and target inserts were cut from the gel and DNA were purified using QIAquick gel extraction kit.

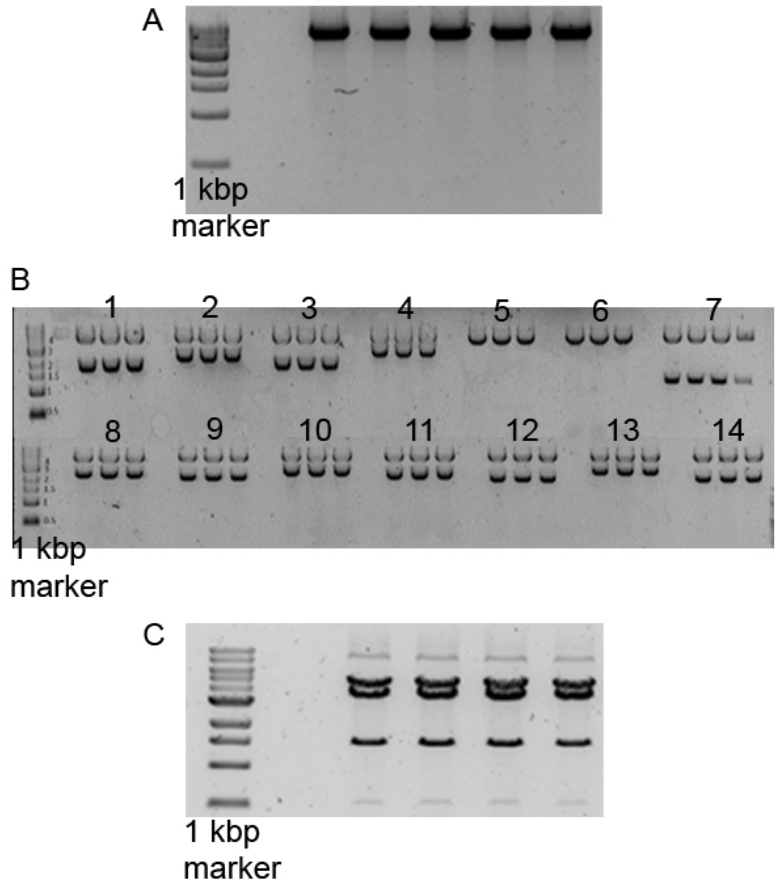


Figure 3.4 Gel electrophoresis for backbone plasmid and target inserts. A) the single band of the modified pLenti7.3/V5-DEST was linearised by NheI and PspOMI (5 replicate samples), B) the insert targets of each DNMT isoform were cut by NheI and NotI (3 replicate samples); 1-14 represent DNMTΔ3B4, DNMT3B3, DNMTΔ3B3, DNMT3A1, DNMT1, Myc, DNMT3L, DNMT3B5, DNMT3A2, DNMT3B2, DNMT3B4, DNMTΔ3B2, DNMT3B1, and DNMTΔ3B1, respectively and, C) DNMT1 band was run on a low percentage (0.7%) gel electrophoresis (four replicate samples).

After ligation and incubation, the bacteria were screened by colony PCR to select the positive clones containing each target DNMT and Myc sequence. The PCR product was amplified using CMV forward primer and Myc reverse primer; the sizes of PCR product were 191 bp for the negative clone and 159 bp for the positive clone (Figure 3.5).

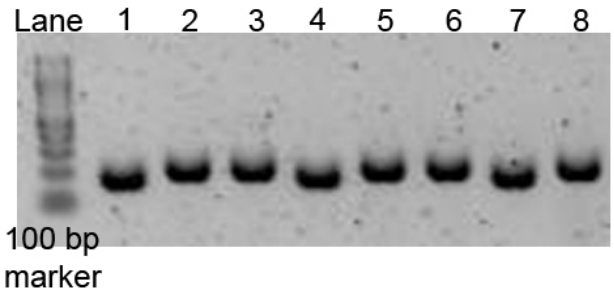


Figure 3.5 Gel electrophoresis of PCR products from colony PCR. Lanes 1, 4, 7 showed the negative clone, while lanes 2, 3, 5, 6, 8 showed the positive clone.

All positive plasmids were then sent to MRC PPU DNA Sequencing and Services at Dundee University to confirm the sequences. All sequences had been analysed and aligned with DNA sequence templates for each DNMT isoform using the DNASTAR Lasergene bioinformatics software (Figure 3.6 and in Appendix B).

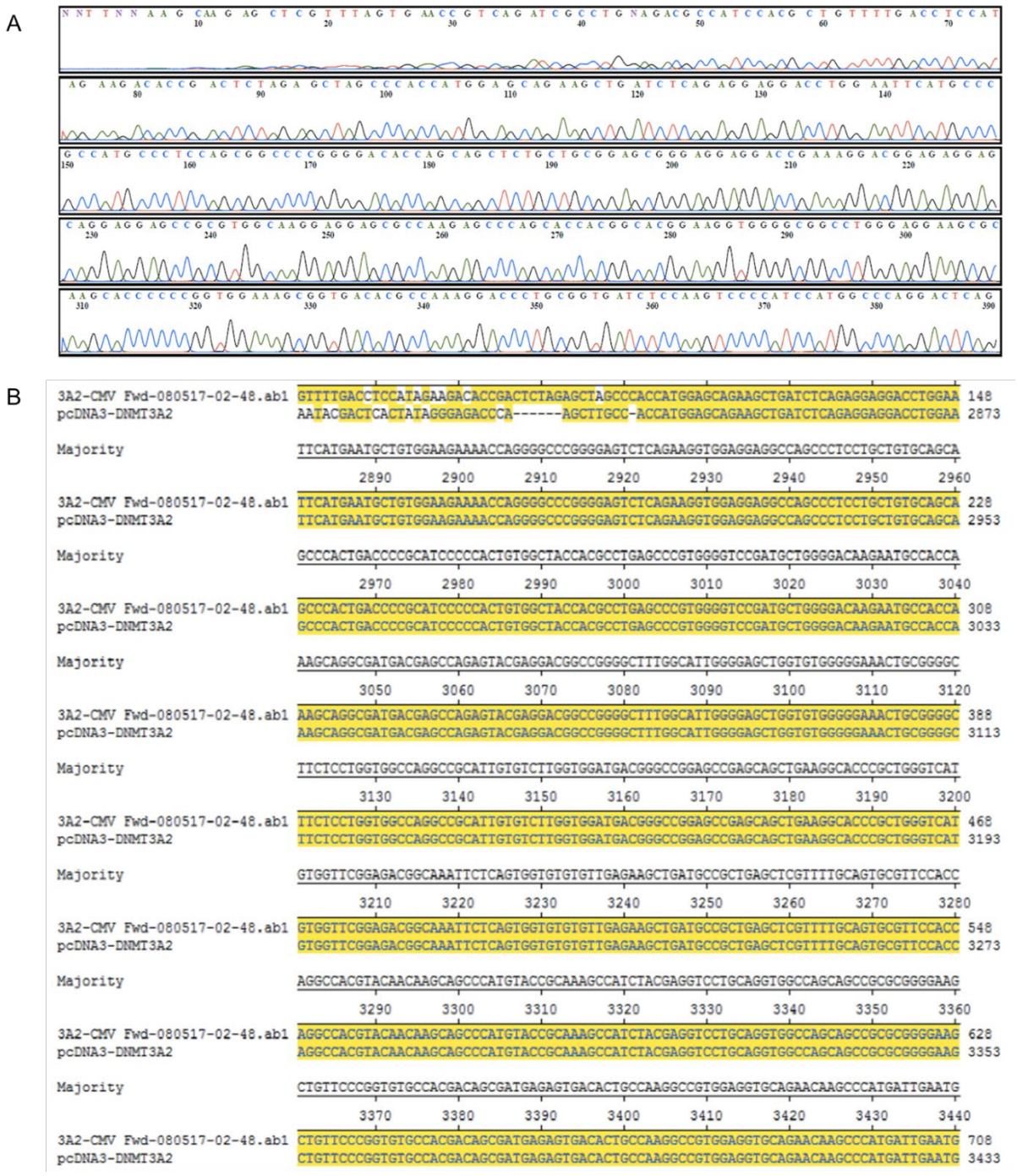


Figure 3.6 Sequence alignment. A) The sequence of DNMT3A1 from Sanger sequencing and B) the alignment between DNMT3A2 and commercial pcDNA-DNMT3A2; majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

3.6.1.3. Optimising transfection conditions between transient transfection and a virus system

Firstly, MEG-01 cells were transfected with GFP plasmids using a range of transfection reagents; Hiperfect, GeneCellin, Lipofectamine2000, and Lipofectamine3000. After 48 hours, the GFP fluorescence was observed and recorded by photography (Figure 3.7). This showed

that both doses Lipofectamin2000 were effective in transferring the plasmids into the MEG-01 cells, while Dharmafect, Hiperfect, GeneCellin, and Lipofectamine3000 were not effective when used for this transfection due to toxicity.

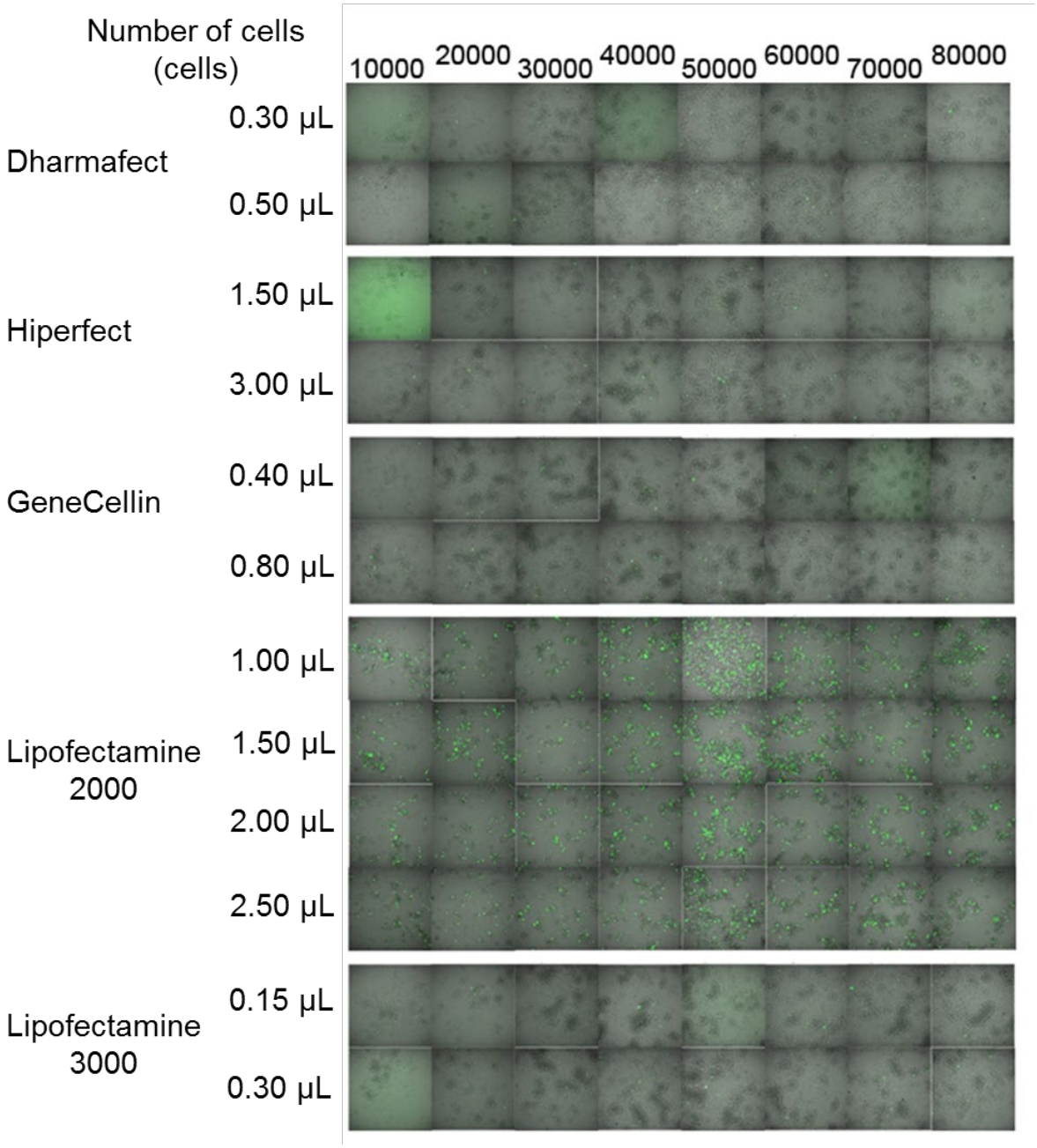


Figure 3.7 Green fluorescent protein signal of transfected MEG-01 cells. MEG-01 cells were transfected with different transfection reagents: Dharmafect, Hiperfect, GeneCellin, Lipofectamine2000, and Lipofectamine3000 with different volumes.

For the viral system, lentivirus with target DNMTs and GFP were applied to MEG-01 cells for 24 hours with or without 5 µg/mL Hexadimethrine bromide (polybrene), which can enhance the efficiency of infection. The GFP signal was determined using inverted Leica Dmi8 wide field fluorescent microscopy. The MEG-01 cells' morphology was normal, showing a round shape with a clearly defined (healthy) cell membrane (Figure 3.8A). This cell was incubated for 10 and 17 days with/without polybrene, and which the GFP signal was captured and compared with the non-transduced MEG-01 cells (Figure 3.8B).

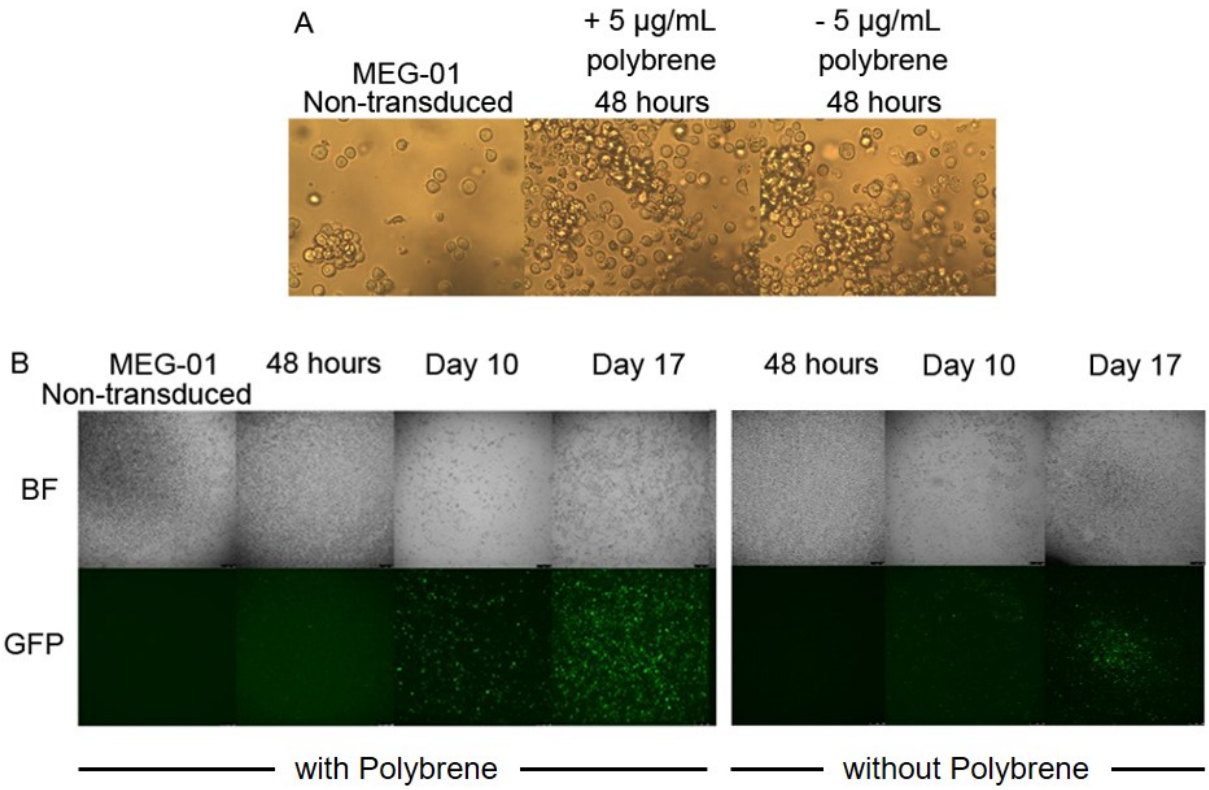


Figure 3.8 Cell morphology and green fluorescent protein (GFP) signal after MEG-01 cell transduction with a lentivirus system for 48 hours, 7 days, and 10 days. A) MEG-01 cell morphology after transduction with/without 5 µg/mL polybrene for 48 hours and B) the bright field (BF) and GFP channel images of transduced MEG-01 cells after 48 hours, 7 days, and 10 days with/without 5 µg/mL polybrene compared with non-transduced MEG-01 cell.

3.6.2 Using transient transfection to generate cell lines that over-expressed each DNMT isoform separately

MEG-01 cells were transiently transfected with pIRESpuro3 contained individually DNMT sequences and c-Myc: DNMT3A1, DNMT3A2, DNMTΔ3B1, and DNMTΔ3B2. Lipofectamine2000 was used to transfer these plasmids into the MEG-01 cells. After transfection, cells were selected using puromycin treatment. However, there were a large

proportion of dead cells after puromycin treatment so the plan for generating cell lines that overexpress each of the DNMTs individually was changed to that using the virus system.

3.6.3 Using lentiviral system to generate cell lines that over-expressed each DNMT isoform separately

MEG-01 and HEK239T cells were transduced using lentivirus for 48 hours and images of GFP expression in the cells were captured using fluorescent microscopy. There was a few positive GFP cells in transduced MEG-01 cells (Figure 3.9) compared with HEK293T cells (Figure 3.10).

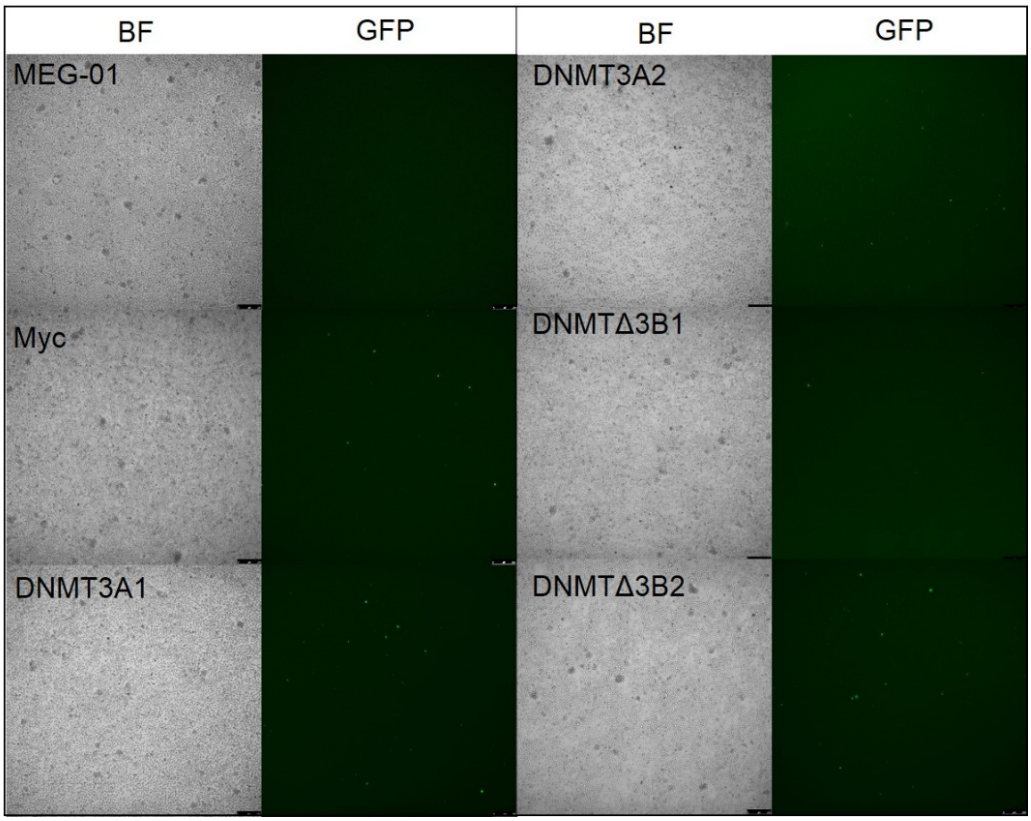


Figure 3.9 Bright field (BF) and green fluorescent protein (GFP) signals from cells overexpressing DNMTs in MEG-01 cells after 48 hours transduction.

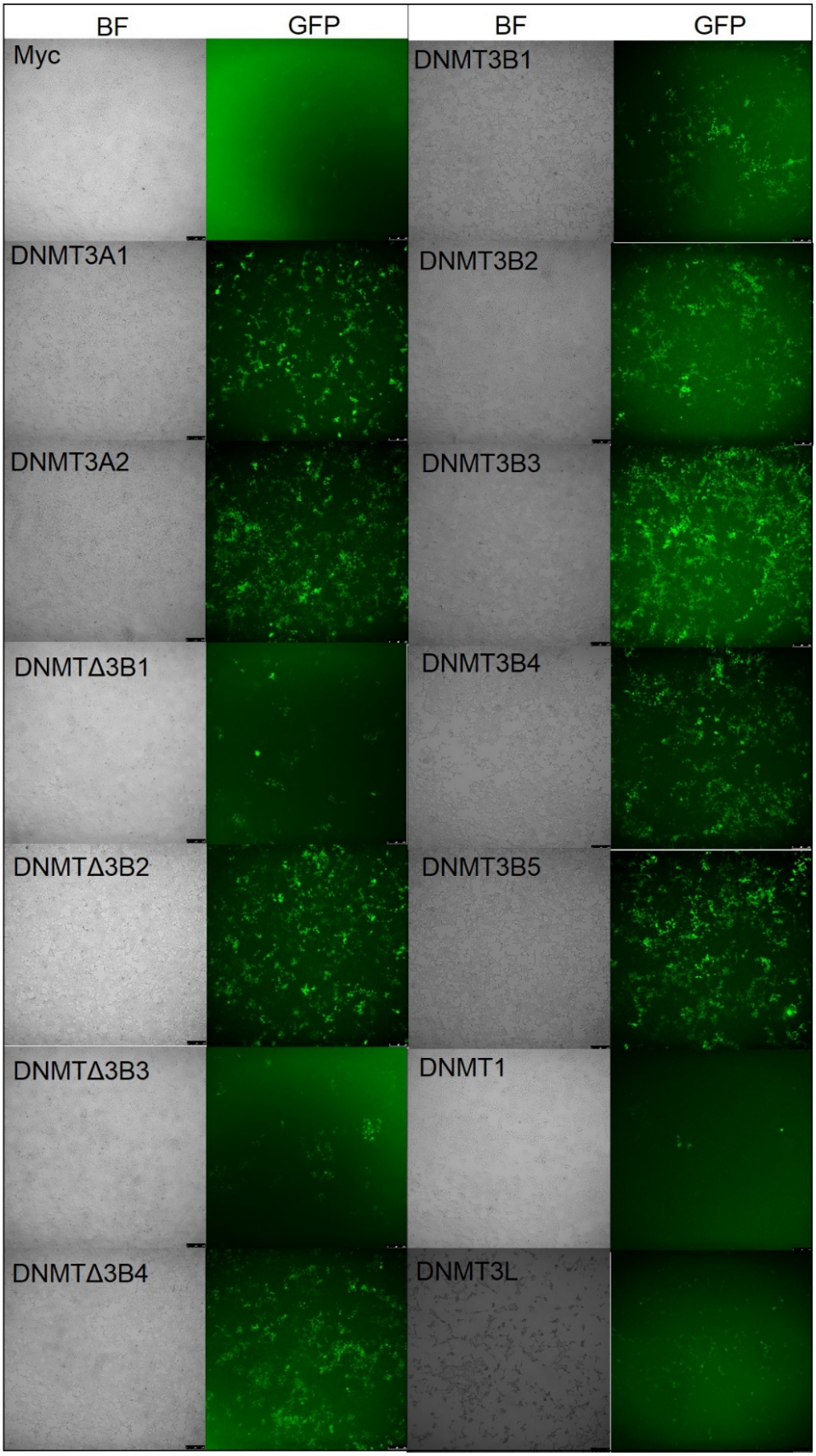


Figure 3.10 Bright field (BF) and green fluorescent protein (GFP) signals from cells overexpressing DNMTs in HEK293T cells after 48 hours transduction.

Transduced MEG-01 and HEK293T cells were grown for 17 days and then sorted using FACS, which selected for the positive GFP cells (Figure 3.11). Cells with a high GFP signal were sorted into a 96-well plate adding a single cell per well for HEK293T cells and into a flask adding a mix population for MEG-01 cells.

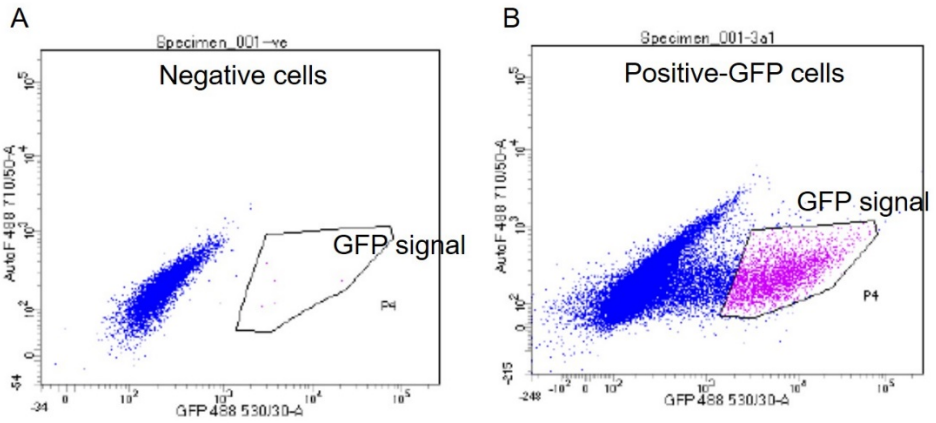


Figure 3.11 Scatter plots of green fluorescent protein (GFP) signal from transduced cells. A) non-transduced cells (negative cells) and B) positive-GFP cells; blue colour represents non-GFP cells or low GFP signal, pink colour represents high GFP signal.

Each single cell of DNMT-overexpressing HEK293T cells was cultured for 14 days, then the morphology of the cells and the number of live cells was observed. In 96-well plate of each DNMT cell, there were 7 clones for DNMT3A1, 8 clones for DNMT3A2, 14 clones for DNMT3B1, 14 clones for DNMT3B2, 29 clones for DNMT3B3, 31 clones for DNMT3B4, 22 clones for DNMT3B5, 10 clones for DNMTΔ3B1, 11 clones for DNMTΔ3B2, 15 clones for DNMTΔ3B3, 40 clones for DNMTΔ3B4, 31 clones for DNMT3L, 21 clones for DNMT1, and 30 clones for Myc. The example of morphology of stably transduced single cell-derived clonal lines was showed in Figure 3.12.

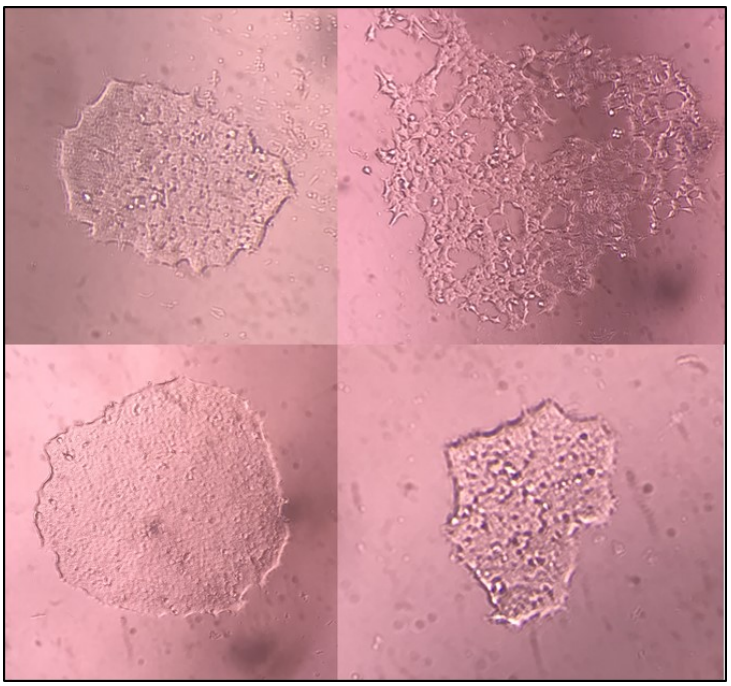


Figure 3.12 Morphology of single cell colonies from overexpression of DNMTΔ3B4 in HEK293T cells in four different wells of 96-well plate.

After expansion of single cell clones, overexpression of the DNMTs in MEG-01 and HEK293T cells was confirmed by GFP signal. Mixed cell population of the DNMT-overexpressing MEG-01 cells showed GFP signal (Figure 3.13). In addition, the stably transduced single cell-derived clonal lines of DNMT-overexpressing HEK293 cells showed GFP signal (Figure 3.14).

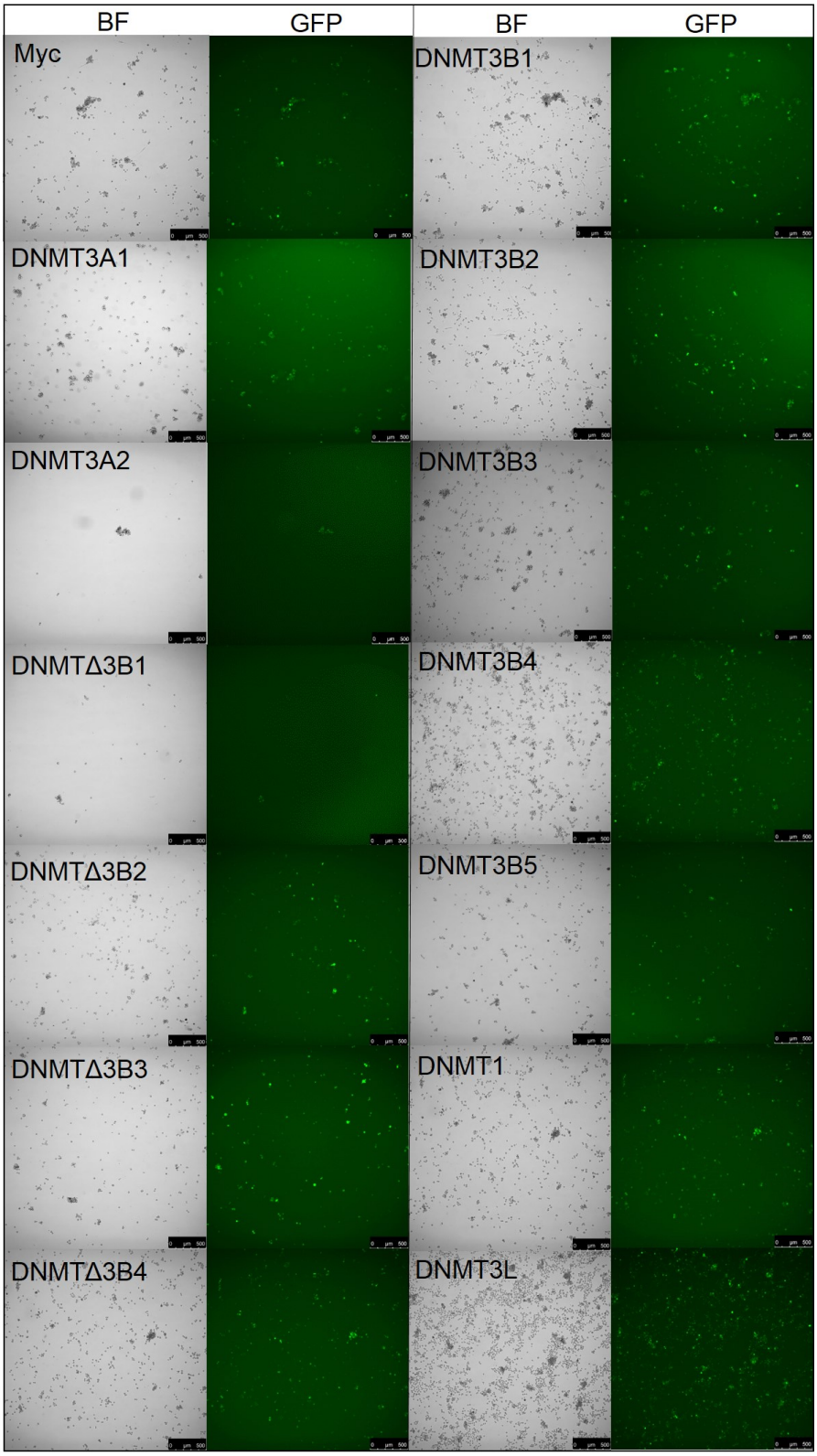


Figure 3.13 Bright field (BF) and green fluorescent protein (GFP) signals from DNMT-overexpressing MEG-01 cells derived from sorted cells.

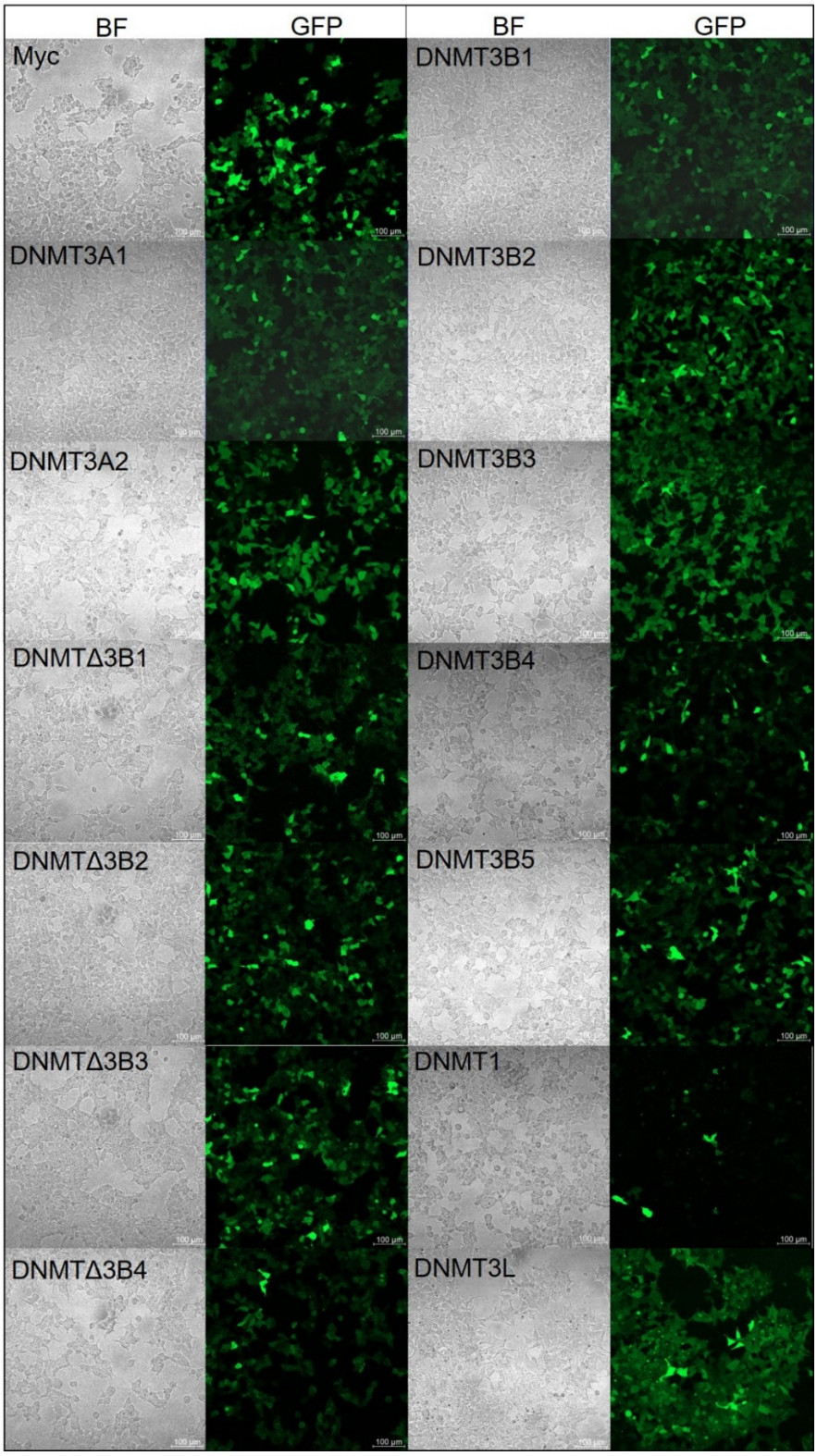


Figure 3.14 Bright field (BF) and green fluorescent protein (GFP) signals from DNMT-overexpressing HEK293 cells derived from the sorted cells.

Use of the Invitrogen™ Tali™ Image-based Cytometer showed that all DNMT-overexpressing HEK293T cells were 100% GFP positive (Table 3.1).

Table 3.1 Numbers of green fluorescent protein (GFP) and non-GFP cells in the clones expressing each of the DNMT isoforms assessed using the Invitrogen™ Tali™ Image-based Cytometer.

DNMT-overexpressing Cells	**GFP cells**	**Non GFP cells**	**Live cells**	**Dead cells**
DNMT3A1	$1.15x10^7 \pm 6.12x10^3$	0	$1.15x10^7 \pm 6.12x10^3$	$0.62x10^4 \pm 209$
DNMT3A2	$7.75x10^6 \pm 7.23x10^2$	0	$7.75x10^6 \pm 7.23x10^2$	0
DNMT3B1	$1.16x10^7 \pm 5.34x10^3$	0	$1.16x10^7 \pm 5.34x10^3$	0
DNMT3B2	$9.34x10^6 \pm 7.84x10^2$	0	$9.34x10^6 \pm 7.84x10^2$	$0.05x10^4 \pm 603$
DNMT3B3	$7.53x10^6 \pm 9.23x10^2$	0	$7.53x10^6 \pm 9.23x10^2$	$0.71x10^4 \pm 487$
DNMT3B4	$1.13x10^7 \pm 4.76x10^3$	0	$1.13x10^7 \pm 4.76x10^3$	0
DNMT3B5	$8.49x10^6 \pm 5.67x10^2$	0	$8.49x10^6 \pm 5.67x10^2$	0
DNMTΔ3B1	$8.65x10^6 \pm 5.34x10^2$	0	$8.65x10^6 \pm 5.34x10^2$	$0.81x10^4 \pm 564$
DNMTΔ3B2	$1.5x10^7 \pm 7.23x10^3$	0	$1.5x10^7 \pm 7.23x10^3$	0
DNMTΔ3B3	$1.03x10^7 \pm 8.23x10^3$	0	$1.03x10^7 \pm 8.23x10^3$	$0.64x10^4 \pm 785$
DNMTΔ3B4	$1.35x10^7 \pm 7.21x10^3$	0	$1.35x10^7 \pm 7.21x10^3$	0
DNMT1	$1.01x10^7 \pm 7.42x10^3$	0	$1.01x10^7 \pm 7.42x10^3$	0
DNMT3L	$1.32x10^7 \pm 2.34x10^3$	0	$1.32x10^7 \pm 2.34x10^3$	0
Myc	$8.11x10^6 \pm 8.21x10^3$	0	$8.11x10^6 \pm 8.21x10^3$	0
HEK293T	0	$7.52x10^6 \pm 0.23x10^2$	$7.52x10^6 \pm 0.23x10^2$	0

3.6.4 Confirmation of total and exogenous *DNMTs* expression at the RNA levels in the overexpressing cells

To confirm the expression of exogenous *DNMTs* and understand the endogenous levels of *DNMTs* in overexpressing cells, RNA was extracted from all cell lines that overexpress each of the DNMTs individually for measurement of total and exogenous *DNMTs*. Firstly, expression of eleven housekeeping (HK) genes (*ACTB, GUBS, RPL13A, RPS13, TFRC, HPRT1, 18sRNA, B2M, GAPDH, PPIA,* and *RPLP10*) was determined to find the most stable HK genes for each experimental setting to be used for normalisation of gene expression based on geNorm analysis (see Methods section 2.1.7). The PCR products of HK genes for MEG-01 cells were run on the 1% gel electrophoresis (Figure 3.15). The HK genes *18sRNA* and *RPL13A* for MEG-01 cells, and *PPIA* and *GAPDH* for HEK293T cells were used in further quantification of *DNMT* expression at the RNA level.

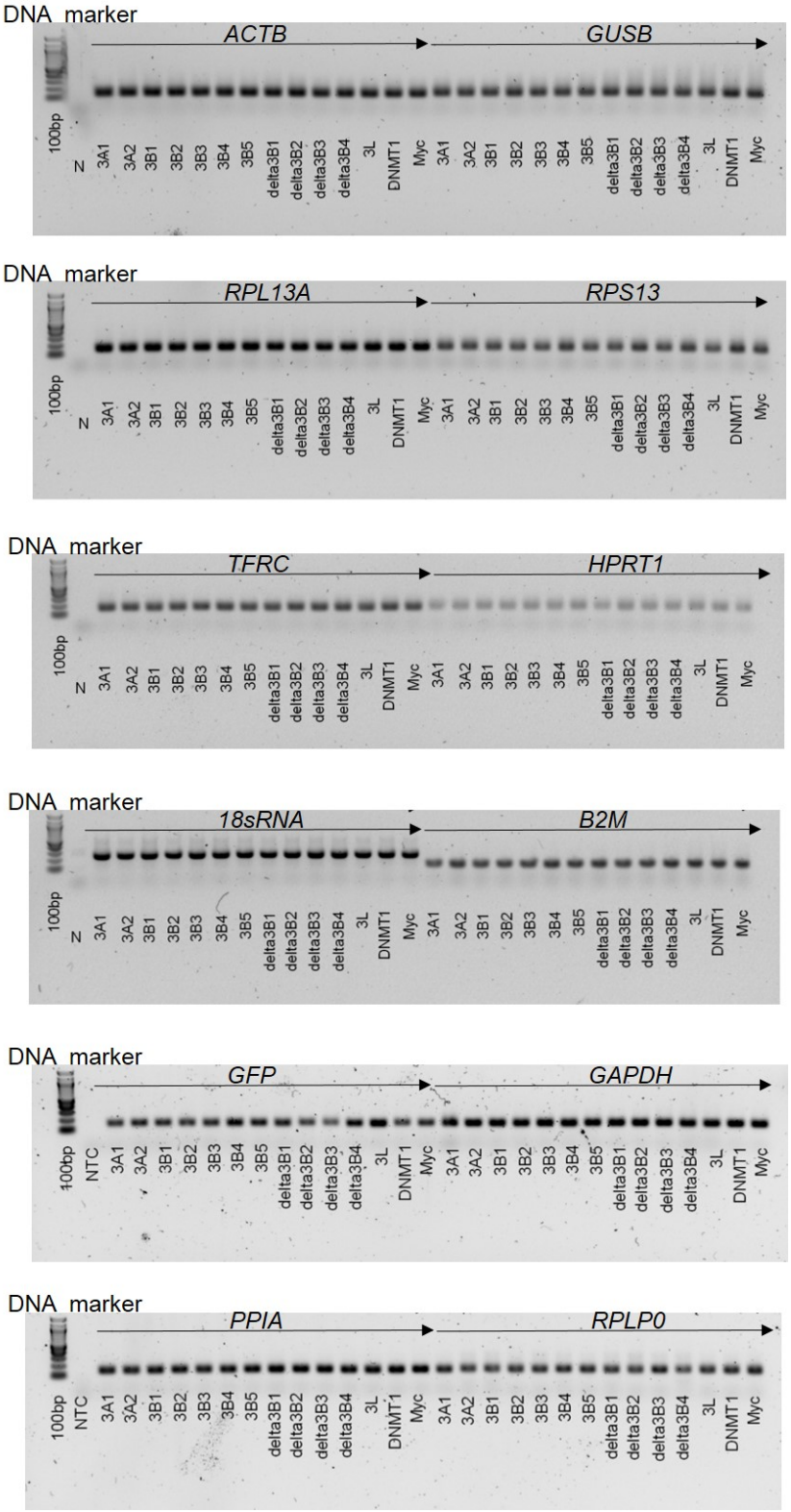


Figure 3.15 PCR products of housekeeping genes and green fluorescent protein (GFP). Actin Beta (ACTB), Glucuronidase Beta (GUSB), Ribosomal Protein L13A (RPL13A), Ribosomal Protein S13 (RPS13), Transferrin Receptor (TFRC), Hypoxanthine Phosphoribosyltransferase 1 (HPRT1), 18sRNA, Beta-2-Microglobulin (B2M), Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH), Peptidylprolyl Isomerase A (PPIA), Ribosomal Protein Lateral Stalk Subunit P0 (RPLP0).

Total *DNMT* (endo + exogenous) and exogenous *DNMT* expression were measured in all cell lines that over-expressed each of the DNMTs individual ly including the Myc-expressing control cell. The total expression of *DNMT* was analysed using the specific primer pairs to DNMT3A1, DNMT3A2, DNMT3B (captured *DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4,* and *DNMT3B5*), DNMTΔ3B1-2 (captured *DNMTΔ3B1* and *DNMTΔ3B2*), DNMTΔ3B3-4 (captured *DNMTΔ3B3* and *DNMTΔ3B4*), DNMT1, and DNMT3L. Each set of primers was designed to amplify unique mRNA transcripts avoiding non-specific target from other DNMT isoforms. Exogenous expression of *DNMTs* was analysed using the Myc primer as the forward primer and the DNMT isoform specific primers as the reverse primer. There was a signal of exogenous *DNMTs* in each DNMT cell but not in the Myc control cell. The total expression (endo + exogenous) of *DNMTΔ3B1, DNMTΔ3B2*, and *DNMT1* was lower than the endogenous expression in the MEG-01 cells (Figure 3.16), whereas, total expression of *DNMT3A2* in HEK293T cells showed low amount compared with a basal expression in Myc control cell (Figure 3.17). However, other total expressions of *DNMTs* expressed higher than endogenous *DNMTs* of Myc control cell (Figure 3.17). As expected, *DNMT3L* was expressed in both transgenic MEG-01 and HEK293T cells as this isoform did not express in MEG-01 cells as same as HEK293T cells at the basal expression.

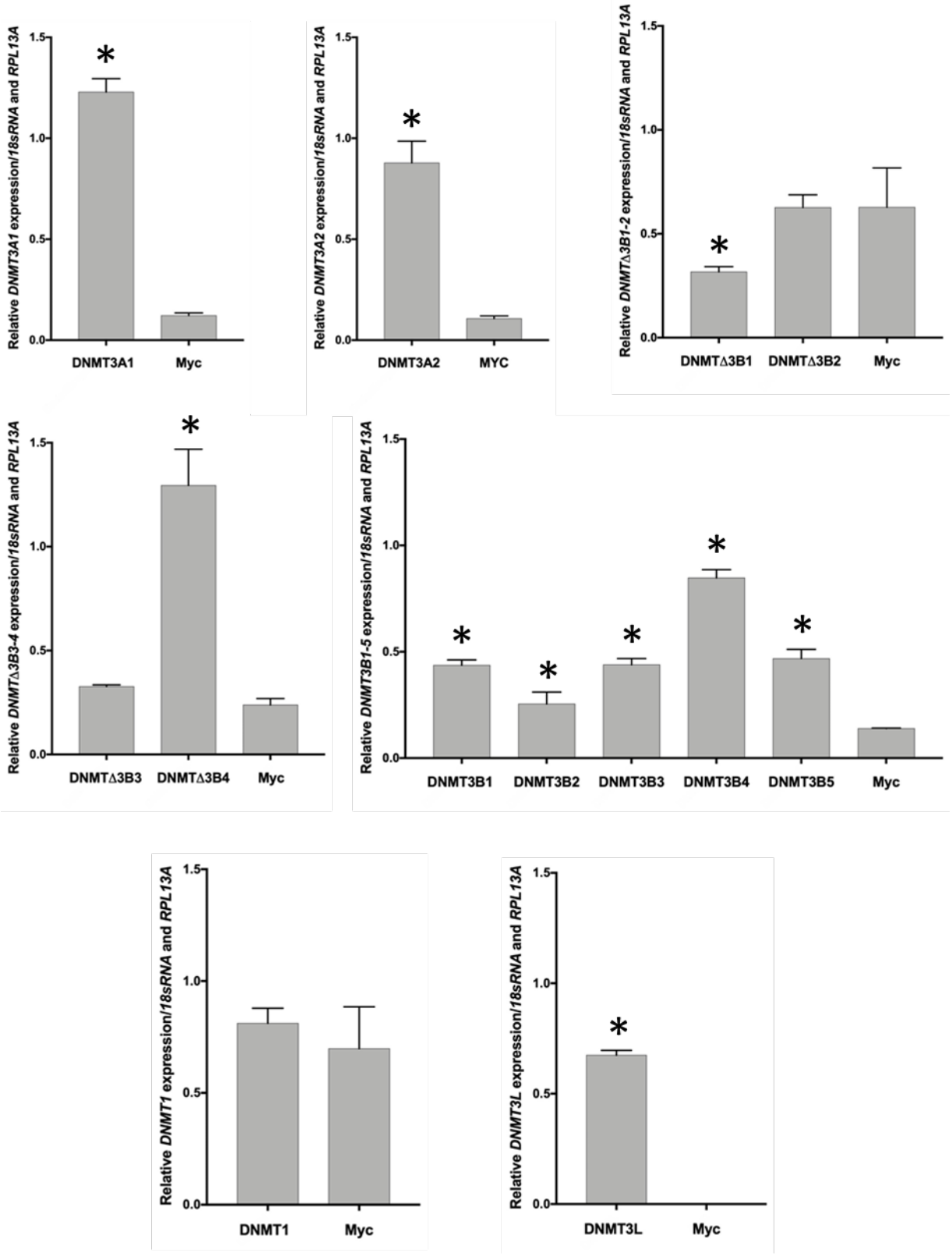


Figure 3.16 Total *DNMT* expression in the mixed cell population of overexpressing-DNMT MEG-01 cells. The relative expression of each *DNMT* isoform was compared between overexpressing-DNMT MEG-01 cells and Myc control cell. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with Myc cell.

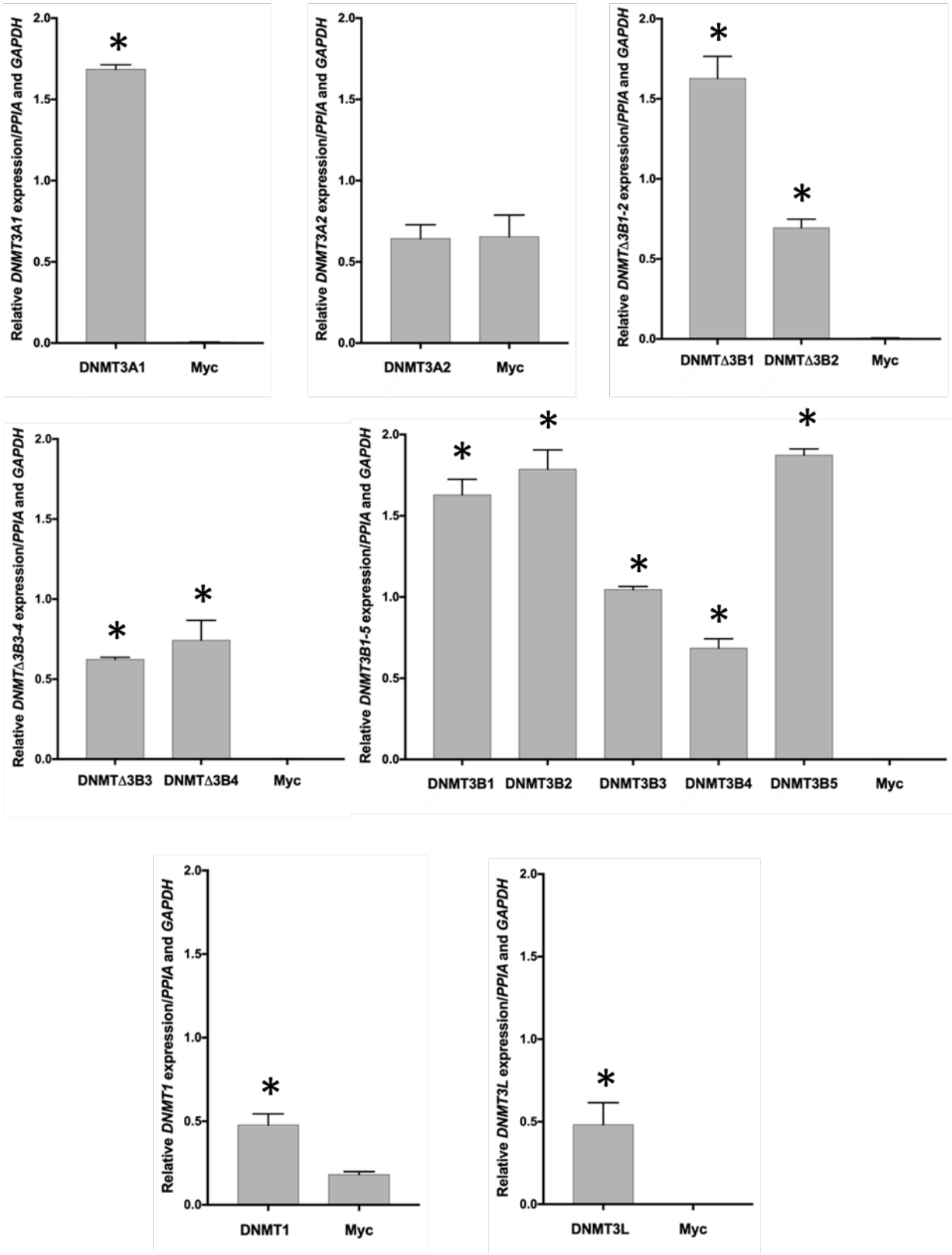


Figure 3.17 Total *DNMT* expression in the mixed cell population of overexpressing-DNMT HEK293T cells. The relative expression of each *DNMT* isoform was compared between overexpressing-DNMT HEK293T cells and Myc control cell. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with Myc cell.

3.6.5 Confirmation of the expression of DNMT isoforms at the protein levels in the DNMT-overexpressing MEG-01 and HEK293T cells

In order to confirm the protein expression of DNMT isoform in overexpressing cells, the exogenous protein expression of each DNMT were measured in both MEG-01 and HEK293T cells using the c-Myc antibody. It was difficult to detect the expression of any DNMT proteins in MEG-01 cells even if the input of protein concentration was increased (data not availablen). However, except for DNMTΔ3B3 (where the protein was not detected), expression of the other DNMT proteins was detected in DNMT-overexpressing HEK293T cells. HEK293T cells transfected stably with DNMT3L expressed relatively large amounts of DNMT3L (Figure 3.18A). DNMT3B2 and DNMT1, DNMT3A1, DNMT3B4, DNMT3B5, DNMTΔ3B2, DNMTΔ3B4 proteins were expressed at relatively low levels, while DNMTΔ3B3 protein expressed very low levels (Figure 3.18B-D).

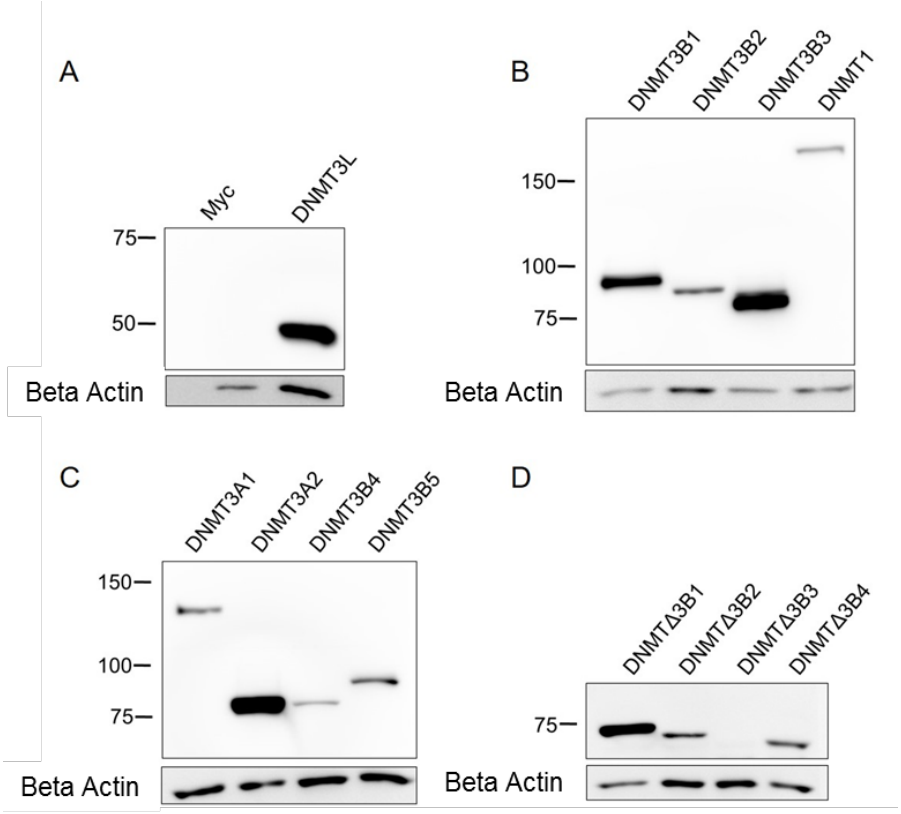


Figure 3.18 Expression of DNMT isoforms at the protein levels in DNMT-overexpressing HEK293T cells. A) DNMT3L protein (25 µg of total protein) in DNMT3L cells compared with Myc control cell; B) DNMT3B1 (50 µg of total protein), DNMT3B2 (70 µg of total protein), DNMT3B3 (50 µg of total protein), and DNMT1 proteins (50 µg of total protein); C) DNMT3A1 (50 µg of total protein), DNMT3A2 (50 µg of total protein), DNMT3B4 (50 µg of total protein), and DNMT3B5 proteins (50 µg of total protein); and D) DNMTΔ3B1 (70 µg of total protein), DNMTΔ3B2 (70 µg of total protein), DNMTΔ3B3 (70 µg of total protein), and DNMTΔ3B4 proteins (70 µg of total protein).

3.7 Discussion

3.7.1 Main findings

The chosen viral system was an effective approach for creating cell lines that over-expressed each of the DNMT isoforms in MEG-01 and HEK293T cells. However, single clones of cells that over-expressed each of the DNMT isoforms were generated successfully only for the HEK293T cells. Each of these clones expressed the specific exogenous DNMT isoform at both RNA and protein levels.

3.7.2 Overexpression of DNMT isoforms in MEG-01 cells

MEG-01 cells, which can be differentiated to platelets that play a critical role in diseases such as cardiovascular disease (287), were over-expressed individually DNMT isoforms using both transient transfection and viral system. Using transient transfection of pIRESpuro3 contained individually DNMT sequences was unsuccessful to generate the stable transgenic cells overexpressing DNMTs in MEG-01 cells because transfected MEG-01 cells no longer expressed the puromycin resistance gene leading to the large number of dead cells mixing with a few live cells. Also, it was difficult to harvest the positive cells containing target DNMTs from this mixed population, as MEG-01 cells being suspension cells, dead and live cells were not separated. However, the transduction of DNMTs into MEG-01 cells was successful using lentivirus. Lentivirus depends on an active transport of the viral pre-integration complex into the cell nucleus through the nucleopore (288). This ability allows lentiviruses to infect both non-dividing and diving cells.

The exogenous *DNMTs* expression was lower than endogenous in the transduced MEG-01 cells, especially *DNMTΔ3B1, DNMTΔ3B2*, and *DNMT1* possibly due to polyclones. Polyclones would include the major cell population of transgenic cells expressed those *DNMTs* at low levels. Furthermore, the number of GFP-positive cells was lower than that of HEK293T cells. This shows the low efficiency of transduction system in MEG-01 cells. Burstein's study (289) revealed that approximately 62% of megakaryocytes were positive after retroviral transduction. The reason of the low percentage of cell transduction may be caused from the doubling time which is 36-48 hours in MEG-01 cells. This long doubling time might affect the virus transduction leading to low expression of *DNMTs*. Although the exogenous expression of each *DNMT* in MEG-01 cells was detected, the protein expressions of each DNMT were very low or could not be detected even if the input of amount of total protein was increased. Therefore, for further work, I focussed on the DNMT-overexpressing

HEK293T cells and used these cell lines for investigation of 1) DNA methylation patterns and 2) interactions between selected food constituents and DNMTs, on DNA methylation at target CpG sites and on activity of the DNMT enzyme.

3.7.3 Overexpression of DNMT isoforms in HEK293T cells

This study is the first to generate the stable single clones of 13 DNMT isoforms in HEK293T cells using an easily handled system with high transduction efficiency. This approach diminishes cell line heterogeneity and improves product yield of gene expression (290). The expression of *GFP* was detected in all single clone cells and this confirmed that the transferring of exogenous genes was successful in HEK293T cells using lentiviral system. All target cells grown and GFP signal was re-measured after cell expanding by fluorescent microscopy and image-based cytometer. Also, the expression of the exogenous *DNMTs* in HEK293T cells was detected using Myc primer to capture the exogenous sequences. However, it is important to note that the basal endogenous expression of each *DNMT* isoform had different levels in Myc control cell (Figure 3.19). For DNMT3A2, exogenous *DNMT3A2* was detected in cells overexpressing DNMT3A2 but the basal expression of endogenous *DNMT3A2* was higher. As shown in the Figure 3.19, the expression of *DNMT3A2*, also highly expressed in Myc control cell compared to other DNMT isoforms.

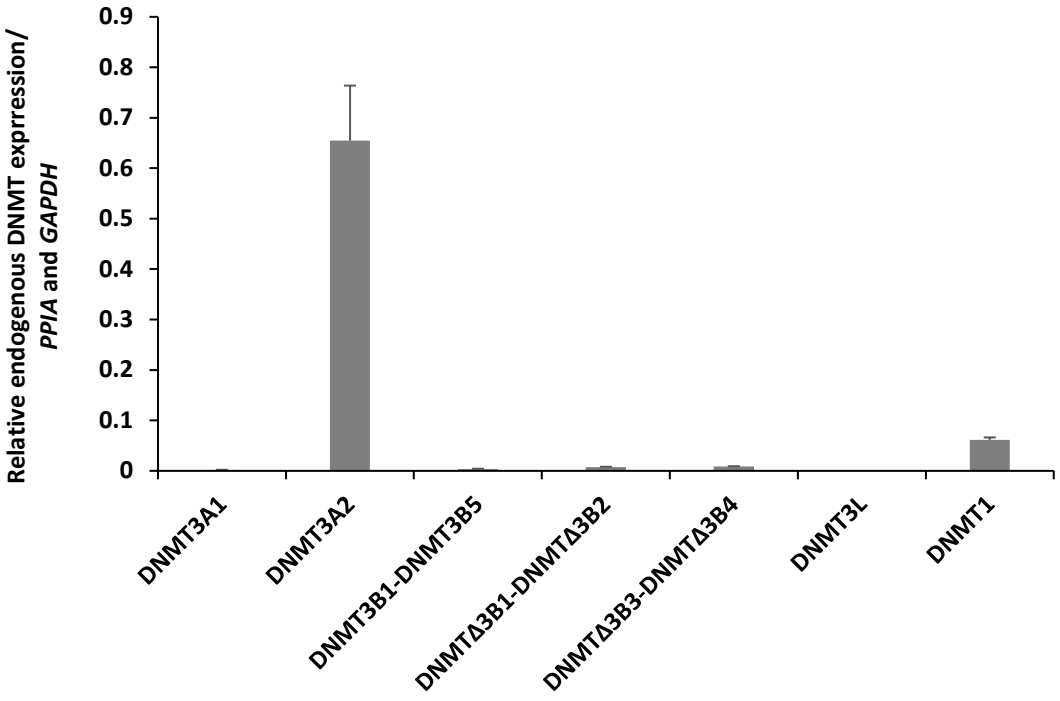


Figure 3.19 The basal endogenous expression of DNMTs. The gene expression of *DNMT3A1, DNMT3A2, DNMT3B1-DNMT3B5, DNMTΔ3B1-DNMTΔ3B2, DNMTΔ3B3-DNMTΔ3B4, DNMT3L,* and *DNMT1* was quantified in Myc control cell.

Due to random integration within the HEK293T genome following the use of the lentivirus system, the protein products of exogenous *DNMTs* were expressed at different levels. This finding was in concordance with results from Torres R. *et al.* (291) indicating that random and site-specific integration showed a difference GFP copy number. It was noted that the mRNA expression of *DNMT3A2* was lower than endogenous level in HEK293T cells, but the protein of this isoform was highly expressed. Also, the mRNA expression of *DNMTΔ3B3* was higher than endogenous level but the protein expression could not be detected. It is likely that the correlation of these proteins and mRNA was poor, possibly reflecting differential stabilities of protein. Also, posttranslational modification delayed protein synthesis leading to reduction of the correlation between RNA and protein expressions (292). The concordance between RNA and protein levels was found to have a correlation coefficient of approximately 0.3 (293). This correlation was in the line with Edfors F. *et al*. (294) showing correlation coefficient of 0.39 in HEK293 cells.

Furthermore, a false positive GFP clone cell was found after sorting by FACS as *GFP* expression could not be detected in this cell by qPCR. Accurate sorting of mixed cell populations is challenging due to noise from the overlapping target and background signals. This problem can be overcome by re-sorting. However, this study sorted single clones of positive GFP cells in 96 well plates and selected five clones to check the GFP in both mRNA and protein expression. By doing this, only positive GFP cells were selected and further expanded for DNA methylation measurement.

3.8 Conclusion

Cell lines that over-expressed each of 13 DNMT isoforms were successfully generated from both MEG-01 (generated mix cell population) and HEK293T cells (generated single cell clone) using lentiviral system. These cell models can be expanded and grown normally with stable DNMT expression. However, DNMT-overexpressing HEK293T cells were further investigated instead of MEG-01 cells due to low exogenous expression of both mRNA and protein levels in MEG-01 cells.

Chapter 4: Identification of specific target CpG sites of each DNMT isoform

4.1 Introduction

Epigenome-wide association studies (EWAS) assess DNA methylation and such studies have become a powerful approach to elucidate association between epigenetic variations and biological traits. With the rapid development of high-throughput microarray, the Illumina HumanMethylation450 BeadChip (450K) is the high-throughput method for characterisation of DNA methylation and this chip measures DNA methylation of more than 450,000 CpG sites throughout the human genome. However, it was replaced with the Illumina HumanMethylationEPIC BeadChip (EPIC), which measures DNA methylation at more than 850,000 CpG sites and covers 90% of the CpG sites of the 450K chip (295). This new chip is non-bias whole epigenome-wide approach and this chip contains new probes targeting gene, intergenetic and non-CpG island regions, covering distal regulatory elements (271). Therefore, this EPIC array can be used to screen the DNA methylation on CpG target sites of DNMTs. Additionally, this new chip requires a low quantity of DNA and it is cost effective.

Mammals express five different families of DNMTs with a number of subfamilies with each protein sharing similar structural features (see introductory section 1.4). Although the catalytic sites of each DNMT are similar, the amino acid residues in these areas are different and this may determine the preferential target CpG sites of each DNMT. Also, the way of entering the DNA major groove for CpG recognition of DNMT1 and DNMT3A was different, despite the conformation similarity in their catalytic loop (296). Moreover, the catalytic activity of DNMT3B was not required for the induced methylation in DNMT3B-deficient cell lines (277).

Choi SH. *et.al* (87) showed the preferential CpG target of each DNMT subfamily. The *de novo* DNA methylation target sites of DNMTs were identified using the Illumina GoldenGate Methylation Cancer Panel I, which contains 1,505 CpG sites from 808 cancer-related genes (87). The clustering analysis of 514 CpGs induced by any DNMT isoforms showed that the DNA methylation patterns induced by DNMTs were clustered according to the structural similarity of the DNMT variants, for example, DNMT3A1 and DNMT3A2, DNMT3B1 and DNMT3B2, DNMTΔ3B1 and DNMTΔ3B2, and DNMTΔ3B3 and DNMTΔ3B4 (87). In another study from Duymich C.E. *et al.*, the target sites of DNMT3B isoforms on a genome-wide level and their function in DNMT3B-deficient cells (3BKO and DKO8 derivatives of the HCT116 colon cancer cell line), were identified using the 450K chip (277). They found DNMT3L restored DNA methylation patterns in DNMT3B-deficient cells

and DNMT3L showed the strongest overall induction of DNA methylation compared with the DNMT3Bs (277). However, none of these studies mentioned the specific DMPs of each DNMT isoform. Moreover, the partners of DNMTs were studied to reveal the factor of DNMT recruitment on DNA, but no study describing the underlying mechanism for any specific isoforms has yet been published. For example, SP1 complex recruits DNMT1 to target on the promoter of *Slit Guidance Ligand 2* to maintain DNA methylation inheritance (297). DNMT3A and DNMT3B cooperate with oncoprotein EV1 to methylate the *miRNA-124-3 promoter* leading to the repression of this gene (298). Preferable sequences for DNMT1, DNMT3A, and DNMT3B were predicted respectively (A/G/T)(T/G/A)(T/A/C)**CG**(T/G/A)(C/A/T)(A/T/C), (T/A/C)(A/T)(T/G/A)**CG**(T/G/C)G(G/C/A), and (A/C)(C/G/A)(A/G)**CG**T(C/G)(A/G) (123).

Therefore, this study focused on the DNA methylation patterns and *de novo* target DMPs of the DNMTs individually by EPIC array to allow the interrogation of methylation patterns at genome-wide and site-specific methylation from the stable transduced single-cell derived clonal lines.

4.2 Hypotheses

The hypotheses for this study were;

1. The pattern of DNA methylation in cells that overexpress individual DNMT isoforms is altered by the structure of each DNMT isoform.
2. Despite the conformational similarity in the catalytic site of DNMTs, each DNMT isoform targets different CpG sites across the human genome.
3. The DNA methylation changes in cells that overexpress each DNMT are involved in biological pathway associated with diseases.

4.3 Aims

The aims of this study were:

- To test the above hypotheses by quantifying 1) the DNA methylation patterns of each DNMT, 2) the DMPs of the preferential target sites of each DNMT using EPIC array and 3) the implication of target DMPs of each DNMT isoform to determine a possible mechanism pathway involved in diseases by pathway analysis.

4.4 Objectives

The objectives of this study were:

1. Using EPIC arrays, to determine the DNA methylation patterns of cell lines that overexpress each of the DNMTs individually.
2. To quantify the methylation levels of each DMP and to select the DMPs that are specific for each DNMT isoform.
3. To analyse the implication of differential methylation of the target DMPs of each individual DNMT isoform using pathway analysis to investigate potential associations with disease.

4.5 Overview of the methods

A detailed description of the experimental procedures and methods for quantifying DNA methylation levels of cell lines that overexpress each of the DNMTs individually can be found in the Methods chapters (2.2 DNA methylation microarray with DNMT overexpressing cells, page 43).

In brief, DNA was extracted from cell lines that overexpress each of the DNMTs individually (see section 2.2.1 in Methods section, page 43). After DNA was extracted, all DNA samples were measured for the concentration and purity using Nanodrop and all samples were sent to Eurofins Genomics, Germany, to perform EPIC array. Statistical analysis and data analysis for EPIC data were performed using R studio (R version 3.6.0) with the Bioconductor package (see section 2.2.2 in Methods section, page 43). The $\Delta\beta$ values were set the cut-offs at $\Delta\beta \leq -0.2$ and $\Delta\beta \geq 0.2$, $\Delta\beta \leq -0.3$ and $\Delta\beta \geq 0.3$, and $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$ with FDR adjusted p-value ≤ 0.05. DNMT over expressing dataset was explored in order to determine the DNA methylation patterns and CpG target sites of DNMTs. Moreover, IPA was conducted to investigate the enrichment of pathways, biological functions, and potential diseases.

4.6 Results

4.6.1 Selection of duplicate cell clones for each overexpressing DNMT cell

To select biological duplicates of each overexpressing DNMT cell for EPIC array analysis, five cell clones of each overexpressing DNMT cell were measured the levels of exogenous DNMTs by qPCR. Two clones were selected with the similar levels of exogenous DNMT, individually (Figure 4.1): clone 2 and 5 for DNMT3A1, clone 2 and 5 for DNMT3A2, clone number 1 and 5 for DNMT3B1, clone number 2 and 3 for DNMT3B2, clone number 1 and 5 for DNMT3B3, clone number 3 and 4 for DNMT3B4, clone number 4 and 5 for DNMT3B5,

clone number 4 and 5 for DNMTΔ3B1, clone number 4 and 5 for DNMTΔ3B2, clone number 1 and 3 for DNMTΔ3B3, clone number 2 and 5 for DNMTΔ3B4, clone number 1 and 2 for DNMT3L, and clone number 2 and 5 for DNMT1.

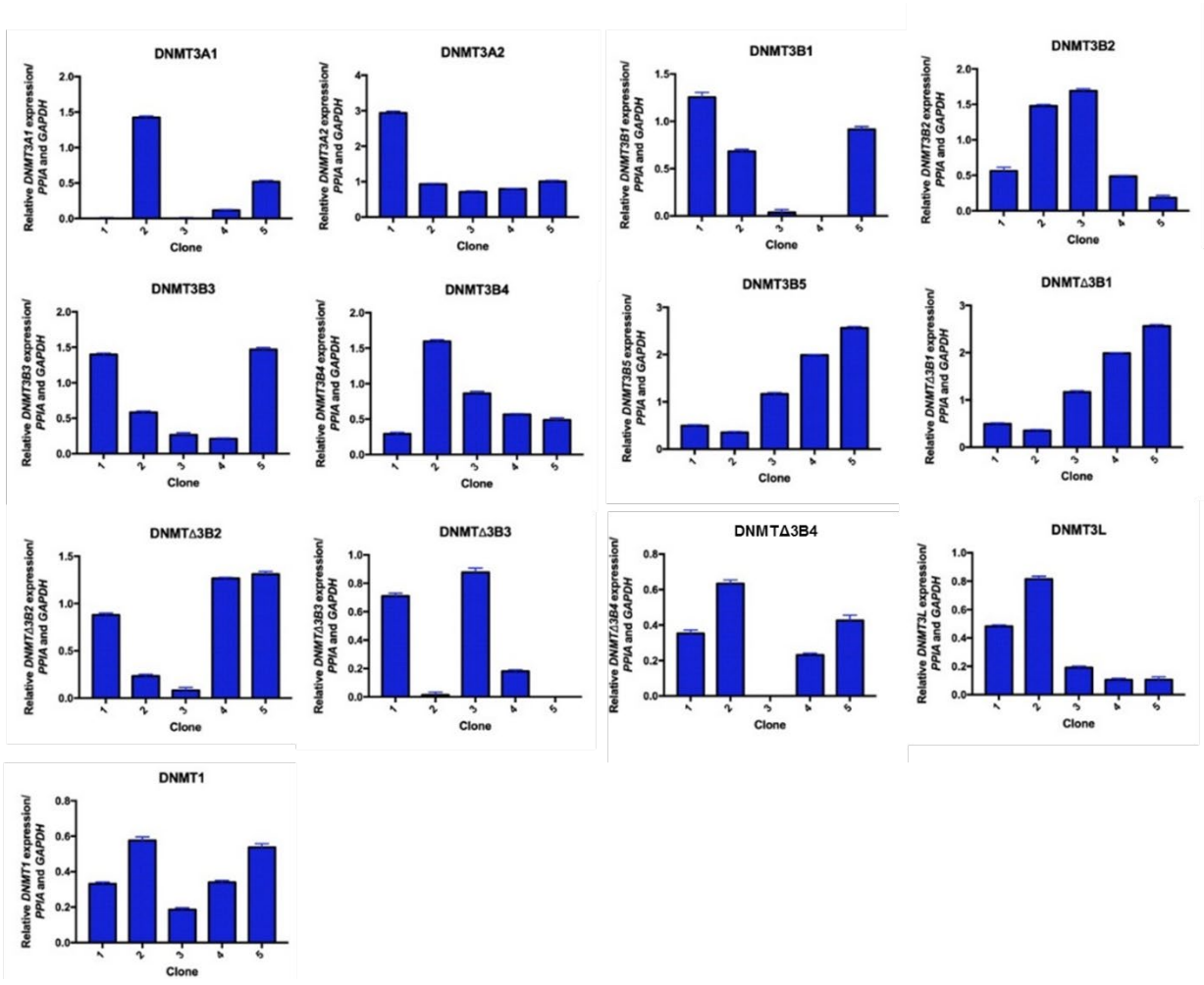


Figure 4.1 Exogenous expression of individual DNMTs in overexpressing DNMT cells.

4.6.2 Data quality checking and preprocessing

To normalise data and check the quality of EPIC datasets, all datasets from cell lines overexpressing individual DNMT isoforms were analysed in R studio using the Bioconductor package. After normalisation using the ssNoob method, β values of Infinitum I and II probes were adjusted and the probe biases were removed (Figure 4.2). The β values expressed the level of DNA methylation, ranging from 0 to 1.

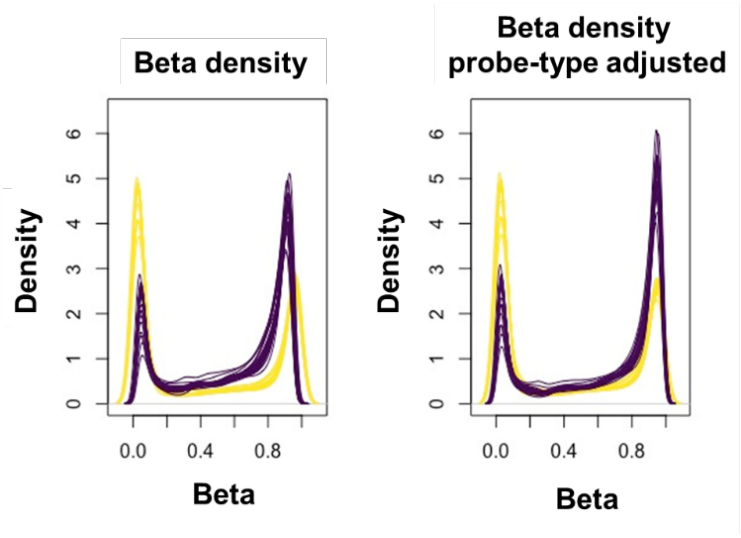


Figure 4.2 β density of all 14 datasets. The distributions of β value from the Infinium I (yellow lines) and II (purple lines) probes were showed before (left graph) and after (right graph) applying the ssNoob normalisation method.

After filtering a high detected *p*-value (*p*-value ≤ 0.05), normalising, adjusting probe-type bias, and removing rs probes, 814,341 probes remained. Bland Altman analysis was performed to illustrate the similarity, and differences, between duplicate biological samples and the 95% limits of agreement (mean ± 1.96SD) were calculated (Figure 4.3). Probes that had a difference value greater than 95% limits of agreement, were filtered out (approximately 6.8% of 814,341 probes) (Table 4.1). Therefore, only high correlation probes were selected from duplicates of each DNMT cell and Myc cell to get reliable results. Among these probes, those of each over-expressed DNMT cell that were found in common with Myc cell were selected. The final number of remaining probes is shown in Table 4.1.

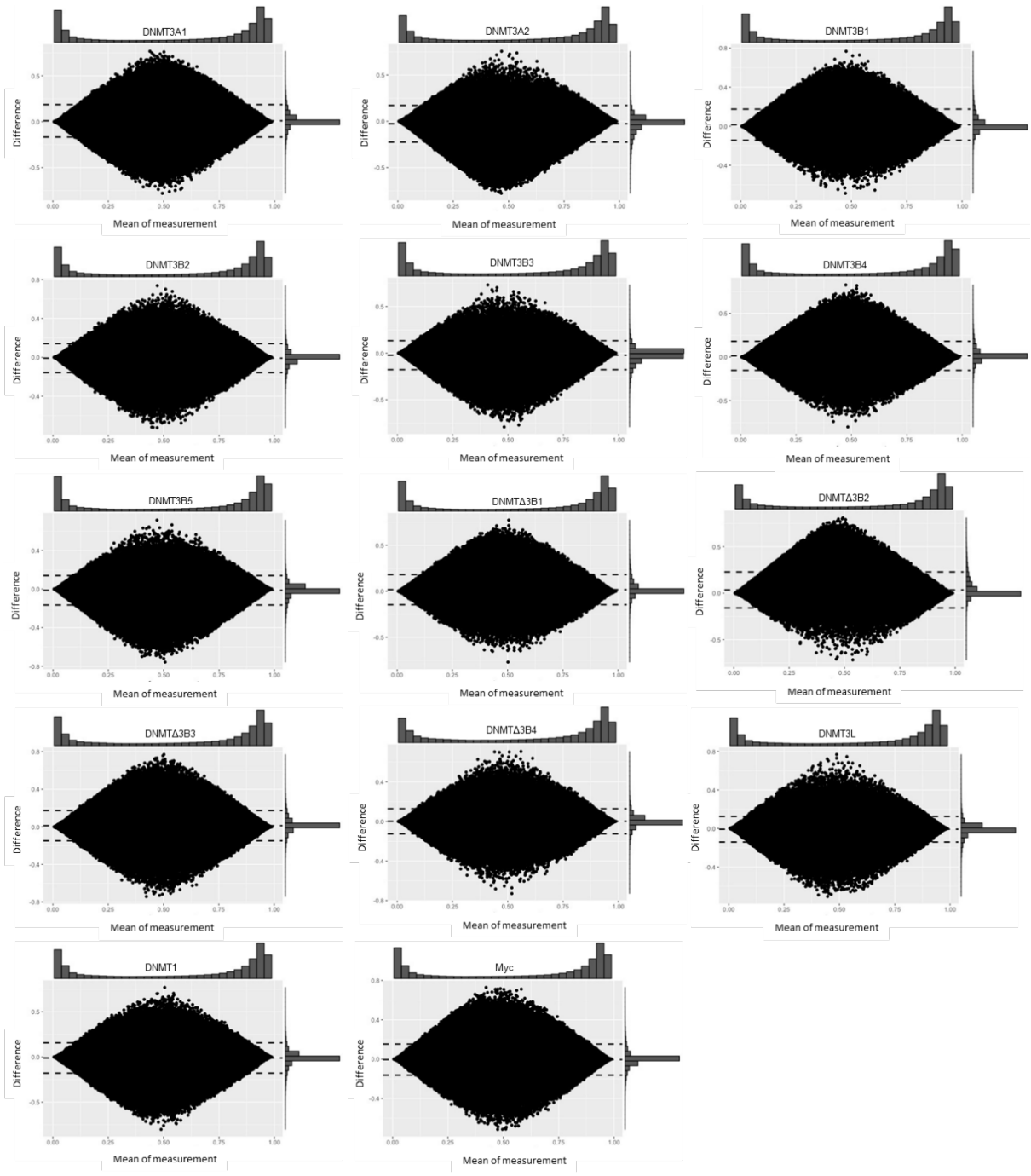


Figure 4.3 Bland Altman plots of each DNMT cell. The scatter plots showed the distribution of mean β values and the difference of DNA methylation levels from each biological duplicate (single cell clones that over-expressed each DNMT isoform and Myc).

Table 4.1 The number of remaining probes after filtering. The probes were selected by mean ± 1.96SD and then those of each over-expressed DNMT cell that were found in common with Myc cell were selected.

DNMT isoforms	The number of the remaining probes after selecting by mean ± 1.96SD	The number of the remaining probes that were found in common with Myc cell
DNMT3A1	755,841	719,204
DNMT3A2	759,760	721,122
DNMT3B1	759,105	722,459
DNMT3B2	758,044	721,013
DNMT3B3	759,484	721,685
DNMT3B4	758,196	721,175
DNMT3B5	758,100	720,891
DNMTΔ3B1	756,788	719,958
DNMTΔ3B2	761,685	723,441
DNMTΔ3B3	758,191	720,896
DNMTΔ3B4	760,585	722,149
DNMT3L	763,014	725,148
DNMT1	761,073	723,090

DNA methylation changes were analysed by subtracting the β value for each CpG in the Myc control from the β value for the corresponding CpG site in each individual DNMT dataset. The methylation changes of each DNMT isoform were indicated in the volcano plots with *p*-value ≤ 0.05 (Figure 4.4).

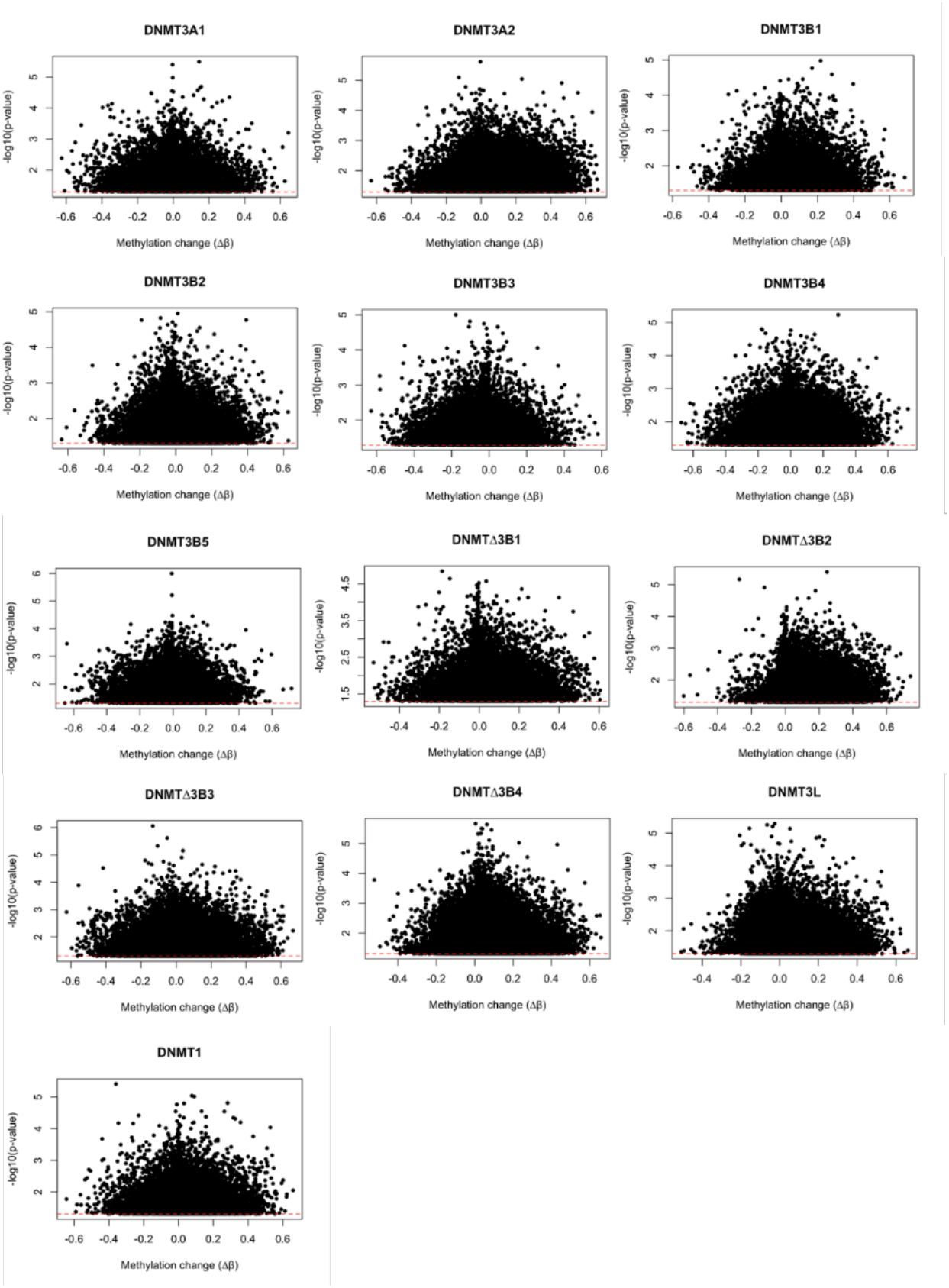


Figure 4.4 Volcano plots of each DNMT isoform. Volcano plots demonstrated the distribution of the DNA methylation changes (Δβ; DNMTs - Myc control) with -log10 (*p*-value) of significant loci (*p*-value ≤ 0.05) in cell lines over-expressing individual DNMT isoforms.

Three different cut-offs: $\Delta\beta \leq -0.2$ and $\Delta\beta \geq 0.2$, $\Delta\beta \leq -0.3$ and $\Delta\beta \geq 0.3$, and $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$ with FDR adjusted *p*-value ≤ 0.05, were set to observe the significant loci of each DNMT (Table 4.2). Narrowing the cut-offs did not significantly change the ranking of DNMTs by number of loci; DNMT3A2 and DNMTΔ3B4. Also, the distribution of significant loci by genomic region was not significantly different among the cut-offs (Figure 4.5-4.7). The significant probes were significantly annotated at the gene body and TSS1500 (*p*-value $\leq$ 0.05) in DNMT cells; the significant probes in DNMT3L and DNMTΔ3B2 were annotated at the gene body (*p*-value ≤ 0.05), and the significant probes in DNMT3A1, DNMT3A2, DNMT3B1, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B3, DNMTΔ3B4, and DNMT1 were annotated at TSS1500 (*p*-value ≤ 0.05).

Table 4.2 The number of significant loci of each DNMT after using three different cut-offs: $\Delta\beta \leq -0.2$ and $\Delta\beta \geq 0.2$, $\Delta\beta \leq -0.3$ and $\Delta\beta \geq 0.3$, and $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$ with FDR adjusted p-value ≤ 0.05.

DNMT isoforms	$\Delta\beta \leq -0.2$ and $\Delta\beta \geq 0.2$	$\Delta\beta \leq -0.3$ and $\Delta\beta \geq 0.3$	$\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$
DNMT1	2,576	896	272
DNMT3L	4,433	1,517	404
DNMT3A1	2,149	702	199
DNMT3A2	6,612	2,874	983
DNMT3B1	2,564	825	226
DNMT3B2	2,380	722	189
DNMT3B3	2,211	700	178
DNMT3B4	4,430	1,625	495
DNMT3B5	2,251	783	202
DNMTΔ3B1	2,682	909	247
DNMTΔ3B2	4,459	1,913	719
DNMTΔ3B3	4,775	1,749	536
DNMTΔ3B4	6,121	2,128	607

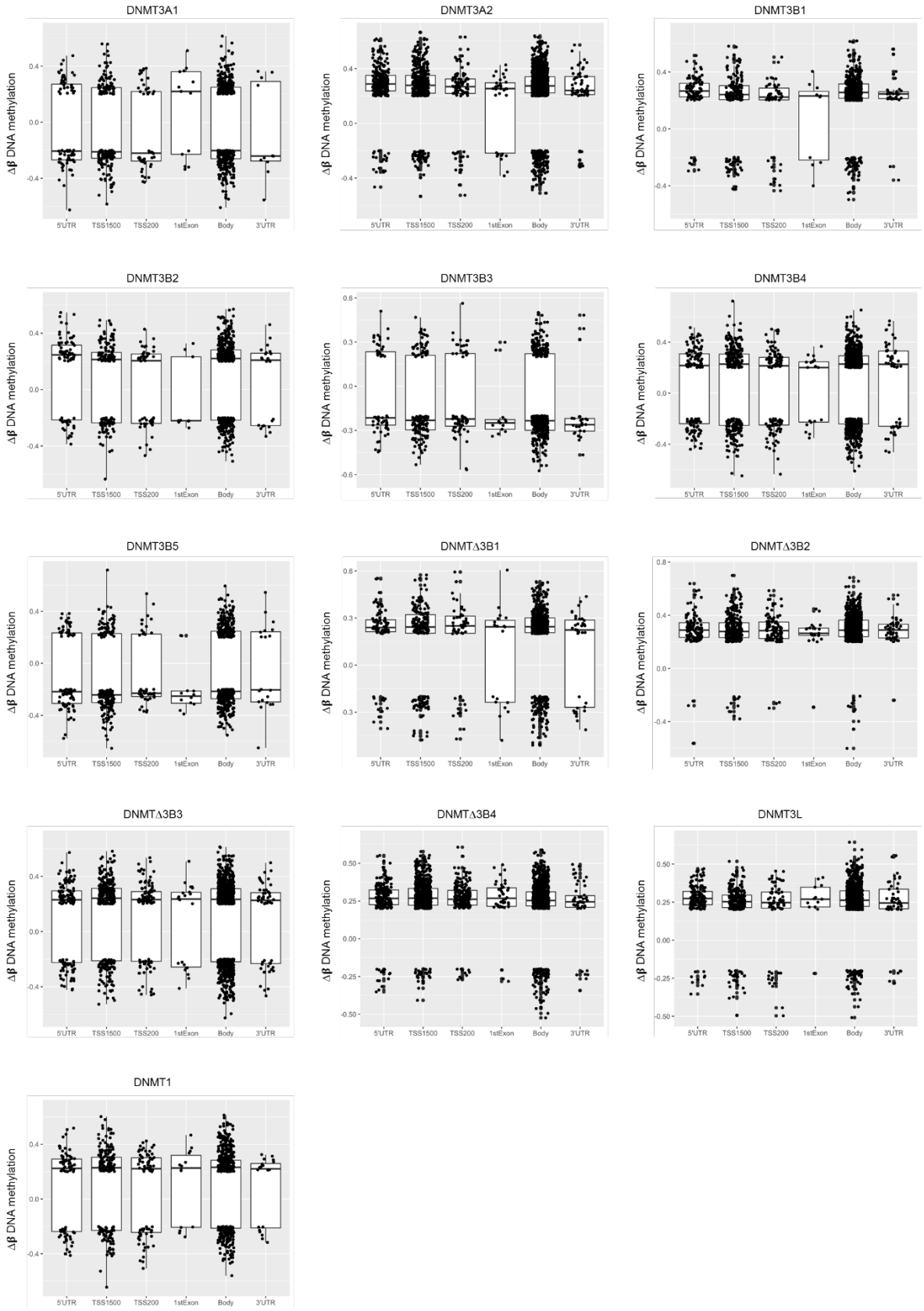


Figure 4.5 Mean Δβ DNA methylation pattern (setting cut-of at $\Delta\beta \leq -0.2$ and $\Delta\beta \geq 0.2$) by genomic location of cells overexpressing DNMTs: 3'UTR: 3' untranslated region, TSS1500: transcription start site 1500, TSS200: transcription start site 200, 1stExon: the first exon, body: gene body, 5'UTR: 5' untranslated region.

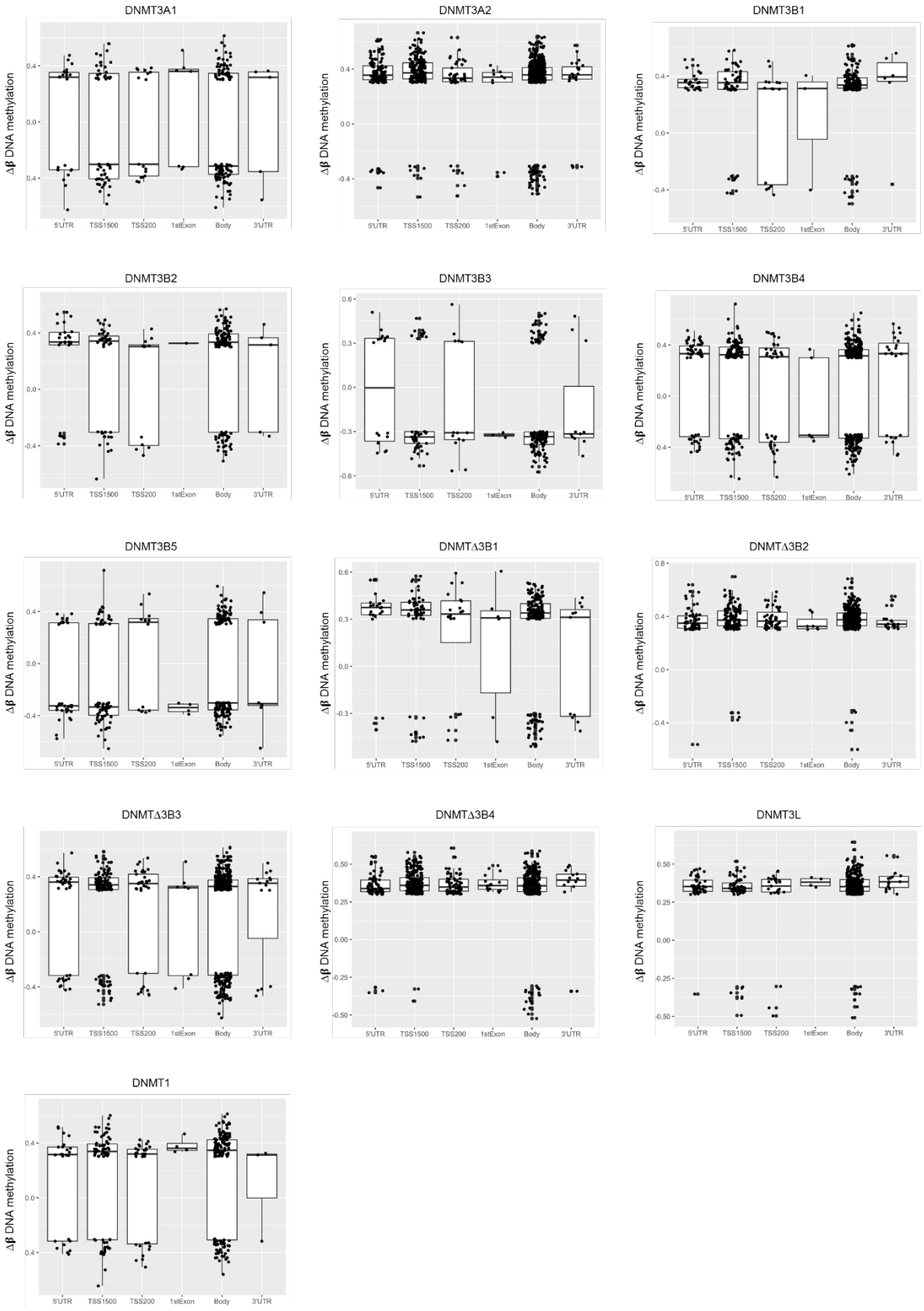


Figure 4.6 Mean Δβ DNA methylation pattern (setting cut-of at $\Delta\beta \leq -0.3$ and $\Delta\beta \geq 0.3$) by genomic location of cells overexpressing DNMTs: 3'UTR: 3' untranslated region, TSS1500: transcription start site 1500, TSS200: transcription start site 200, 1stExon: the first exon, body: gene body, 5'UTR: 5' untranslated region.

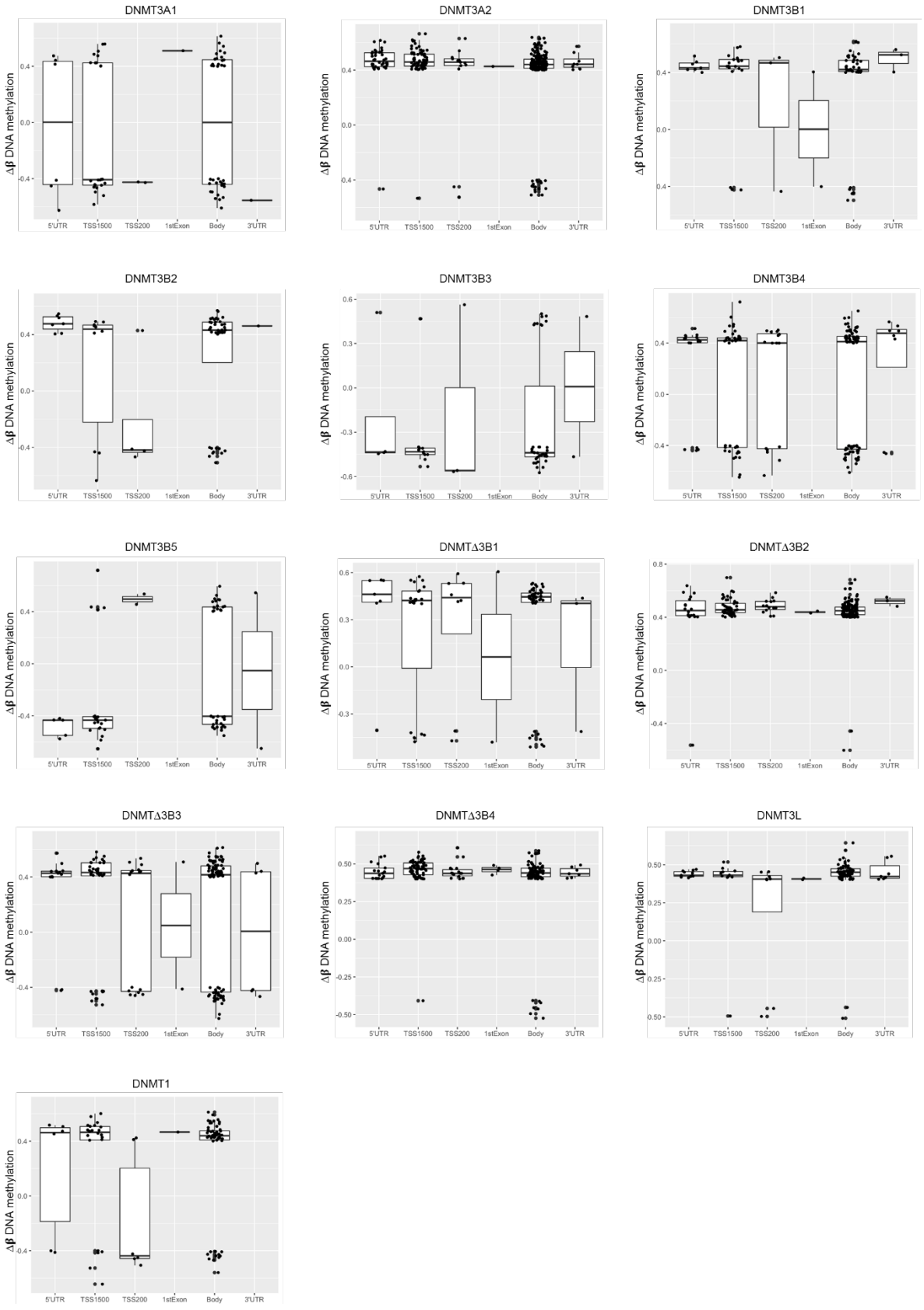


Figure 4.7 Mean Δβ DNA methylation pattern (setting cut-of at $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$) by genomic location of cells overexpressing DNMTs: 3'UTR: 3' untranslated region, TSS1500: transcription start site 1500, TSS200: transcription start site 200, 1stExon: the first exon, body: gene body, 5'UTR: 5' untranslated region.

CpG sites that showed $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$ were retained for further analysis and the distribution of those CpGs is shown as volcano plots with the number of hyper- and hypo-methylated loci (Figure 4.5). The expected methylation increase was observed for the majority of DNMT isoforms, where hypermethylated probes were in higher number than hypomethylated ones. However, DNMT3A1, DNMT3B3, and DNMT3B5 showed lower DNA methylation levels (more hypomethylated probes, in proportion).

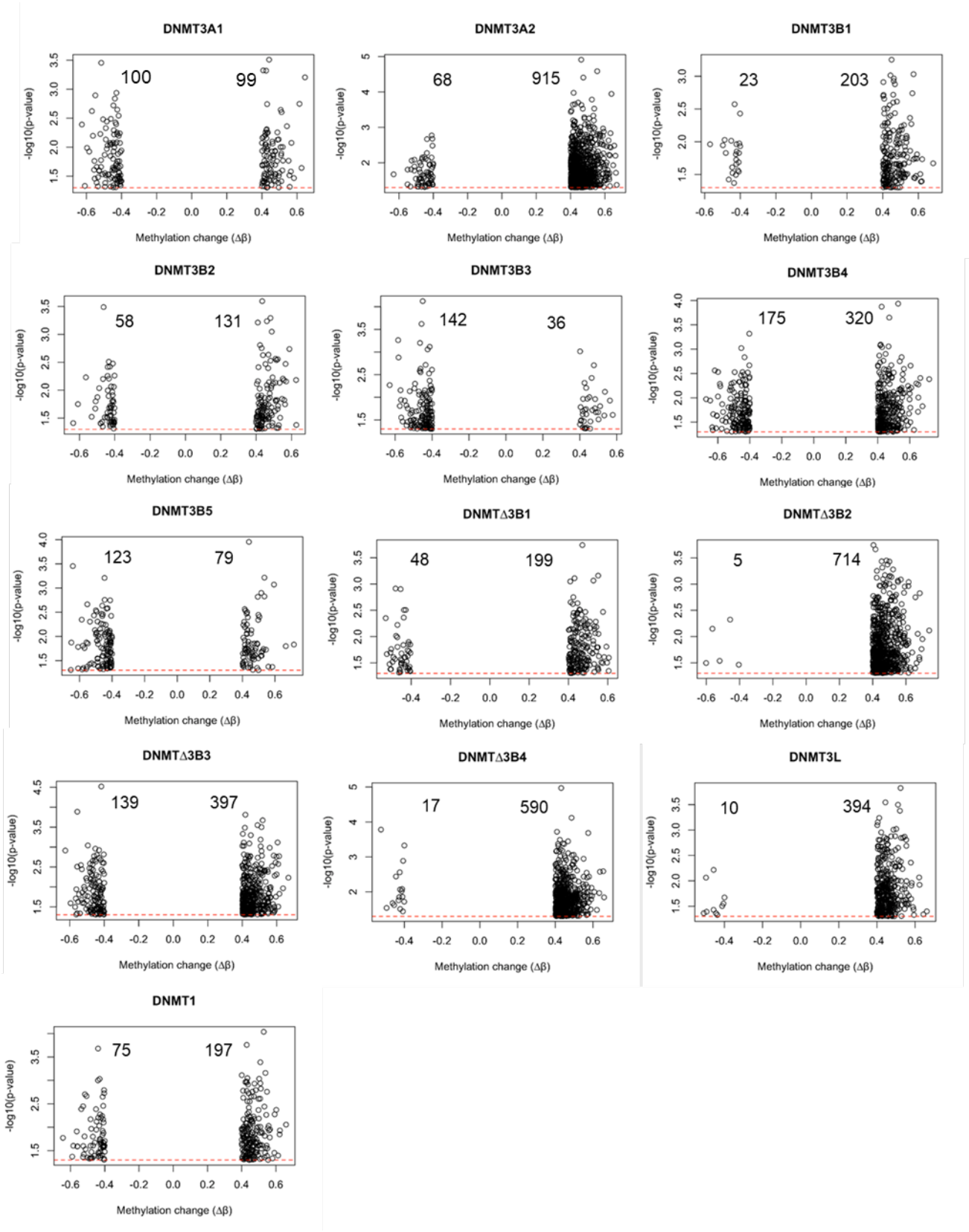


Figure 4.8 Volcano plots of each DNMT isoform. Volcano plots demonstrated the distribution of the methylation change ($\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$) with -log10 (p-value) of significant loci (p-value ≤ 0.05) in cell lines over-expressing individual DNMT isoforms. The number of hypomethylated and hypermethylated loci was labelled in each plot.

4.6.3 The effect of over-expressed DNMTs on the DNA methylation pattern

PCA was performed using β values for all probes to visualise methylation pattern differences by DNMT isoforms (Figure 4.9). PC1 explained 16.04% of the variance between samples while 9.8% of the variance was explained by PC2. There was no noticeable clustering pattern by DNMT isoforms (Figure 4.9A) but the β values of duplicate samples were not significantly different (Figure 4.9B).

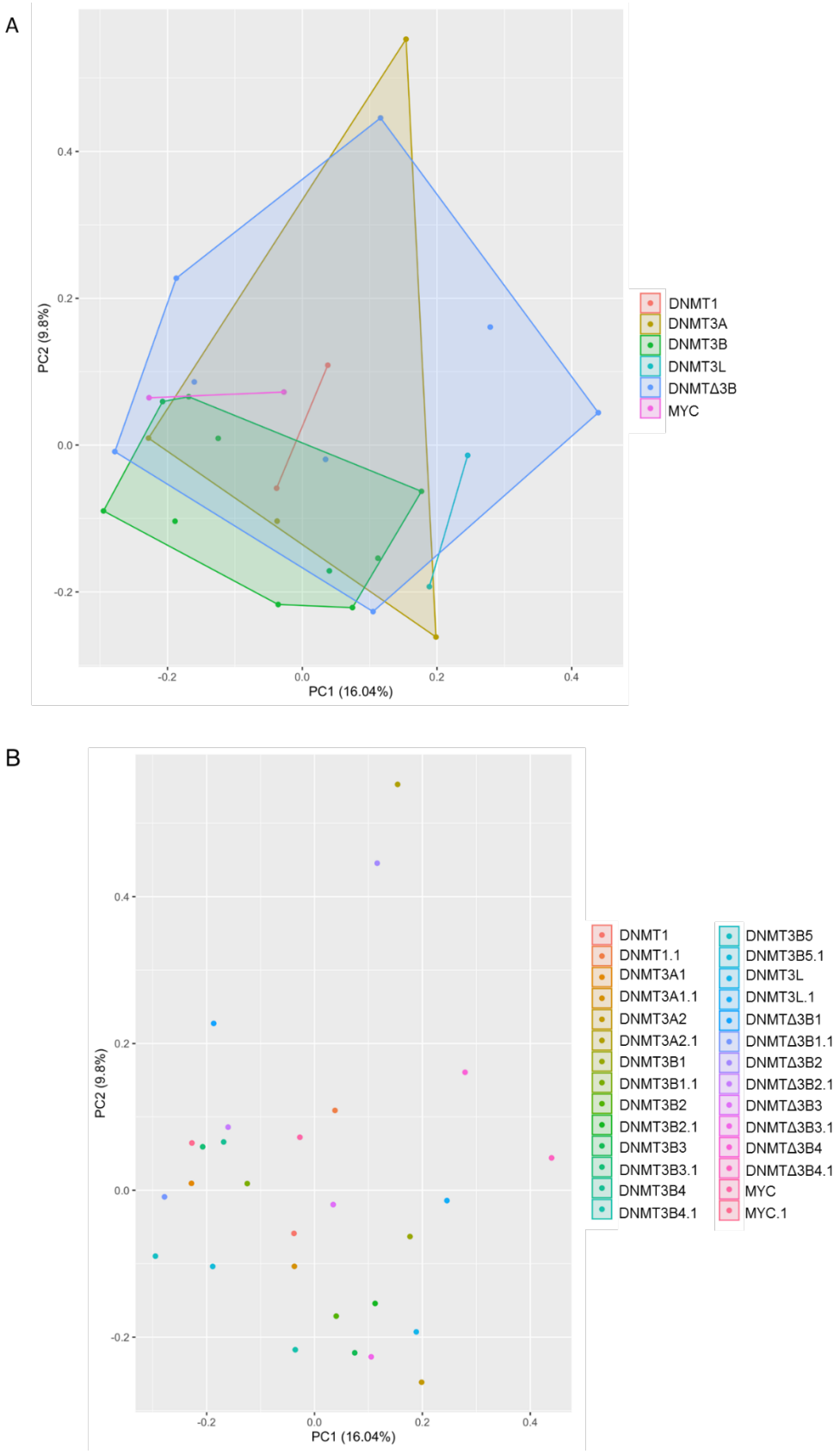


Figure 4.9 Principal component analysis (PCA). PCA was performed using β values from all probes passed the quality control for all overexpressing DNMT cells. A) colour dots indicate individual DNMT isoforms B) multicolour dots indicate individual DNMT isoforms with duplicates (indicated by .1 labelling).

To explore the DNA methylation patterns in cell lines that overexpress each of the DNMTs individually, a two-dimensional unsupervised hierarchical cluster analysis of all hyper/hypo-methylated loci (after selecting from cut-offs, $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$), compared with Myc control, was conducted. The majority of DNA methylation patterns of DNMTs did not depended on the structurally similar DNMTs, but there were two clustering induced by structurally similar DNMTs: 1) DNMTΔ3B4 and DNMTΔ3B2, 2) DNMT3B2 and DNMT3B3 (Figure 4.5). The DNA methylation patterns of DNMT3A2, DNMTΔ3B4 and DNMTΔ3B2, presented more hypermethylated patterns than other isoforms. The DNA methylation patterns of DNMT3B3 appeared to be hypomethylated pattern. The smallest isoform missing catalytic domains on C terminus, DNMT3L, shown a unique pattern and it was clustered in the same group with DNMT1 and DNMTΔ3B3.

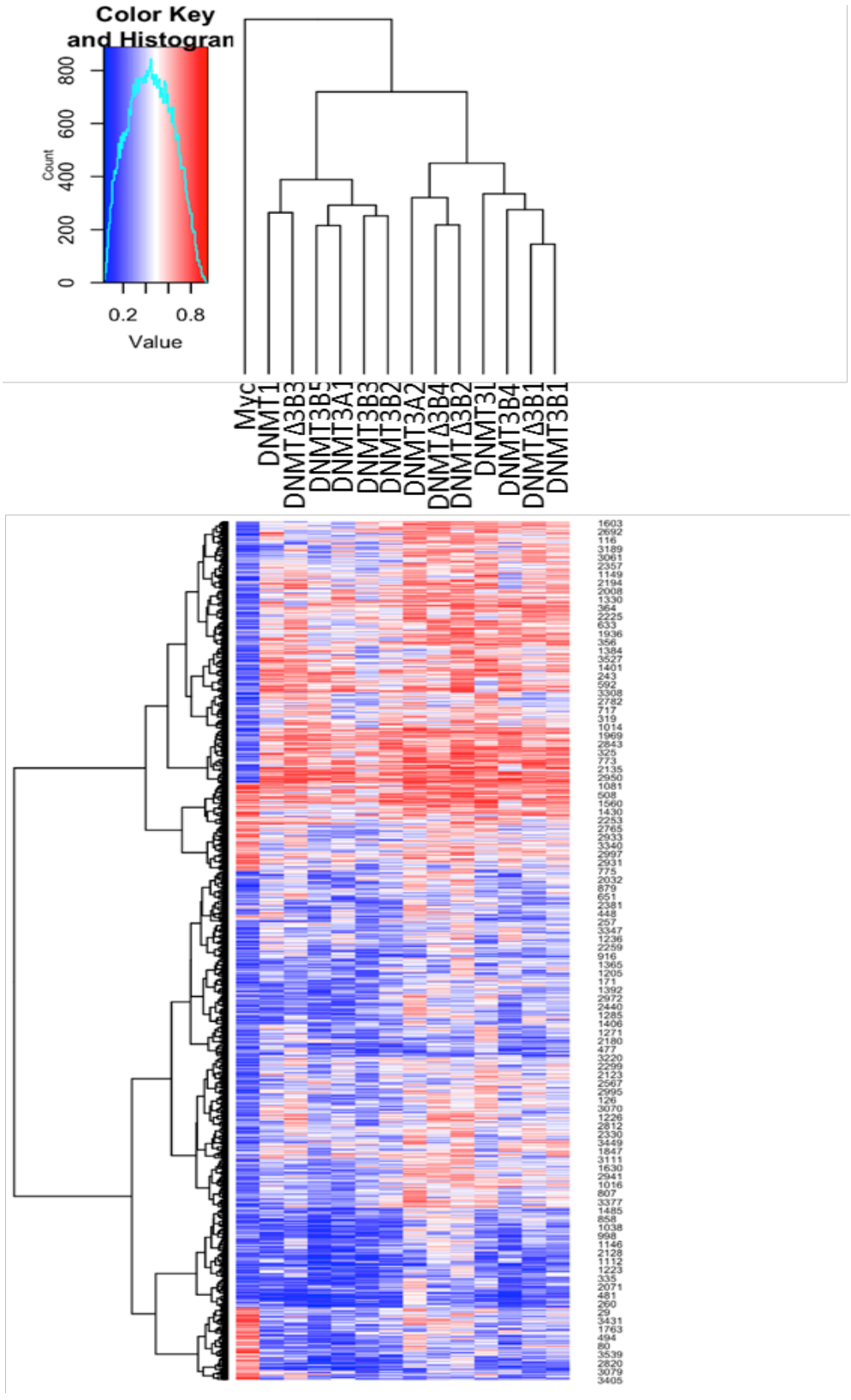


Figure 4.10 Heatmap of DNA methylation data with 3,544 CpG sites (after selecting from cut-offs, $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$ in cell lines over-expressing individual DNMT isoforms). CpGs were selected as they were targeted by at least one DNMT isoform. Myc was a control. Red colour represents high methylation β values and blue indicates low methylation β values.

To observe a clear hypermethylation patterns by DNMTs, a two-dimensional unsupervised hierarchical cluster analysis using only hypermethylated probes (2,833 loci) of each isoform was computed. The DNA methylation patterns of DNMT3B5, DNMT3A1, DNMT3B2, and DNMT3B3 were correlated with Myc control (Figure 4.6). DNMT3A2, DNMTΔ3B4 and DNMTΔ3B2 showed the hypermethylation patterns as the same as in Figure 4.5. DNMT1 and DNMTΔ3B3 were grouped together and were clustered with DNMT3L.

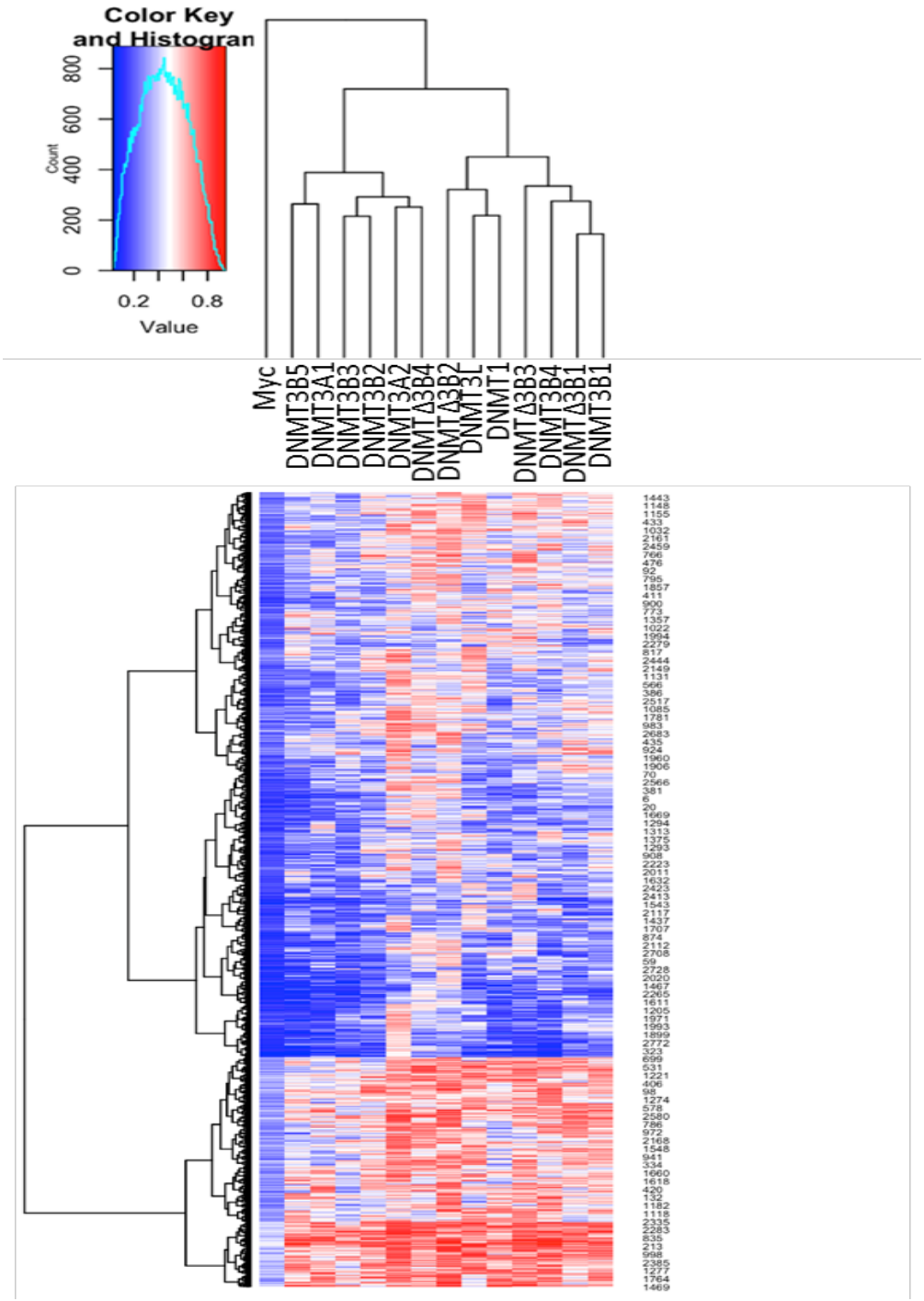


Figure 4.11 Heatmap of DNA methylation data with 2,833 CpG sites ($\Delta\beta \geq 0.4$). CpGs were selected as they were targeted by at least one DNMT isoform. Myc was a control. Red colour represents high methylation β values and blue indicates low methylation β values.

4.6.4 *De novo* target CpG sites of each DNMT isoform

Significant (FDR-adjusted *p*-value ≤ 0.05) CpG sites with $\Delta\beta \geq 0.4$ were identified in each DNMT isoform. Although there is a similarity of catalytic sites among DNMT isoforms, unique target CpGs of each isoform were identified (Figure 4.7). 23.2% of the hypermethylated probes of DNMT3A1 overlapped with the hypermethylated probes of DNMT3A2: this was 2.5% when it was calculated from the total of hypermethylated probes of DNMT3A2. 27% of the hypermethylated probes of DNM3B1, 30.5% of the hypermethylated probes of DNM3B2, 41.7% of the hypermethylated probes of DNM3B3, 25% of the hypermethylated probes of DNM3B4, and 27.8% of the hypermethylated probes of DNM3B5 were overlapped with other DNMT3Bs. The percentage of the hypermethylated probes of DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4 overlapping among DNMTΔ3Bs was 42%, 25.5%, 37.5%, and 31.7%, respectively. Top five of *de novo* hypermethylated probes with gene name of each DNMT isoform were showed in the Table 4.3. Additionally, other target genes of each DNMT isoform were showed in Appendix C.

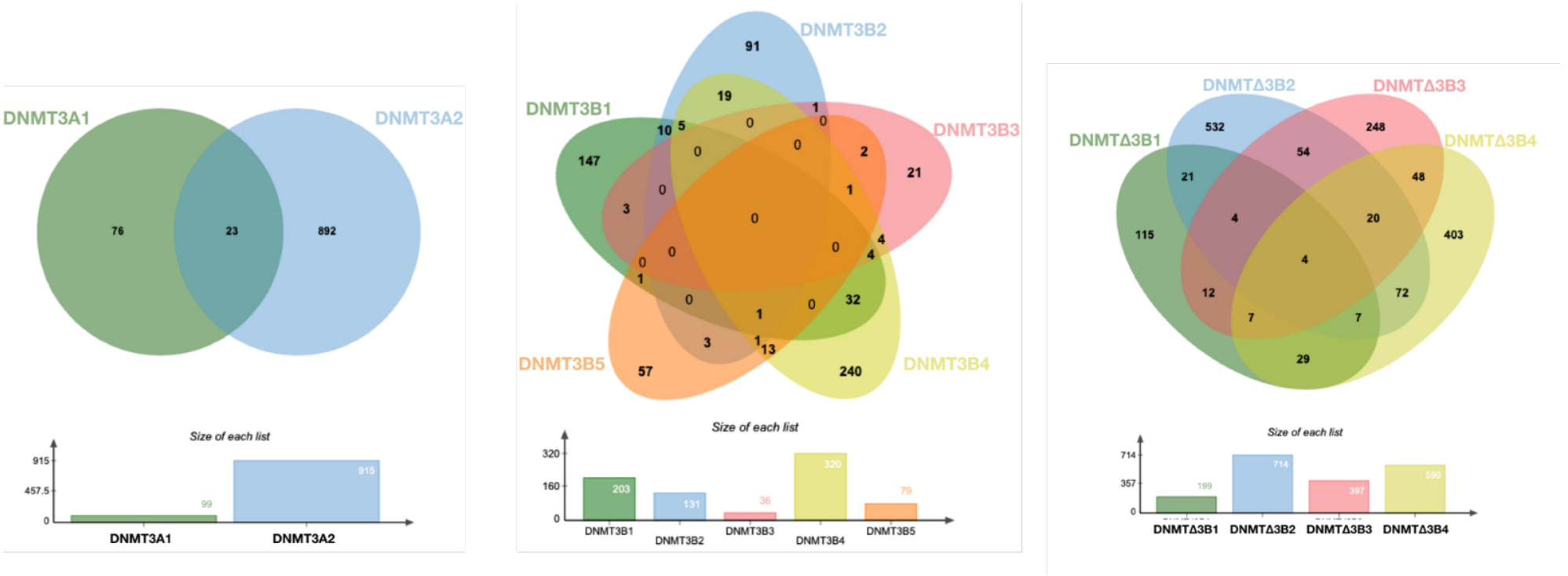


Figure 4.12 Venn diagrams of specific and overlapping target CpG sites of cell lines over-expressing individual DNMT isoforms. The number of target hypermethylated probes for each DNMT family; DNMT3A, DNMT3B, and DNMTΔ3B, overlapping with target CpG sites within those DNMT families is depicted.

Table 4.3 Top five hypermethylated loci of each DNMT isoform.

DNMT isoforms	CpG	Chromosome	Relation to CpG island	β DNA methylation		Δβ	Gene name
				DNMT	Myc		
DNMT3A1	cg21534766	chr21	OpenSea	0.64	0.24	0.41	NCAM2
	cg06041732	chr18	OpenSea	0.78	0.13	0.65	LINC00907
	cg17001806	chr8	OpenSea	0.82	0.40	0.42	-
	cg08969304	chr10	CpG Island	0.67	0.24	0.43	-
	cg01460940	chr14	OpenSea	0.58	0.16	0.42	RP11-123M6.2
DNMT3A2	cg02732111	chr4	OpenSea	0.77	0.21	0.56	-
	cg15930380	chr1	South Shelf	0.63	0.16	0.47	MIER1
	cg16204524	chr2	North Shelf	0.61	0.20	0.42	GPR113
	cg07018561	chr3	South Shore	0.65	0.12	0.52	ENTPD3-AS1
	cg03647393	chr13	OpenSea	0.63	0.19	0.44	LINC01069
DNMT3B1	cg09102332	chr7	OpenSea	0.62	0.22	0.40	-
	cg01529359	chr8	OpenSea	0.52	0.08	0.44	-
	cg23967739	chr12	OpenSea	0.61	0.19	0.43	A2ML1
	cg23491841	chr2	North Shelf	0.55	0.10	0.44	C2orf85
	cg06329392	chr13	OpenSea	0.55	0.12	0.43	-
DNMT3B2	cg18425254	chr13	CpG Island	0.67	0.15	0.52	-
	cg01986486	chr3	OpenSea	0.55	0.07	0.47	C3orf67
	cg02286819	chr1	OpenSea	0.52	0.09	0.43	-
	cg17764572	chr10	CpG Island	0.68	0.27	0.41	-
	cg11457308	chr14	OpenSea	0.53	0.10	0.43	GALNT16
DNMT3B3	cg06580033	chr17	South Shore	0.64	0.10	0.54	ZNF830; CCT6B
	cg12173487	chr16	OpenSea	0.63	0.20	0.43	CPNE2
	cg22354874	chr5	OpenSea	0.87	0.45	0.42	-
	cg12583095	chr10	South Shore	0.57	0.16	0.41	HTR7
DNMT3B4	cg26286826	chr18	OpenSea	0.54	0.12	0.42	-
	cg08954783	chr2	OpenSea	0.60	0.13	0.47	-
	cg24424219	chr15	OpenSea	0.57	0.14	0.43	-
	cg10483660	chr13	South Shelf	0.66	0.11	0.55	-
	cg26432128	chr8	South Shelf	0.80	0.38	0.41	-

Table 4.3 (continue) Top five hypermethylated loci of each DNMT isoform.

DNMT isoforms	CpG	Chromosome	Relation_to_Island	β DNA methylation		Δβ	Gene name
				DNMT	Myc		
DNMT3B5	cg15923359	chr8	OpenSea	0.79	0.27	0.52	CSGALCT1
	cg11439662	chr8	OpenSea	0.74	0.32	0.42	-
	cg13081981	chr4	OpenSea	0.62	0.12	0.50	-
	cg20268758	chr15	North Shelf	0.68	0.25	0.43	FAM189A1
	cg13659542	chr2	OpenSea	0.77	0.36	0.41	-
DNMTΔ3B1	cg26763380	chr8	OpenSea	0.73	0.30	0.43	EIF3E
	cg16318688	chr1	OpenSea	0.64	0.19	0.45	EPHX4
	cg13054800	chr4	OpenSea	0.50	0.07	0.44	-
	cg03894789	chr5	OpenSea	0.71	0.27	0.44	MIR874; KLHL3
	cg19945464	chr11	South Shore	0.47	0.07	0.40	GAS2
DNMTΔ3B2	cg01910804	chr14	North Shore	0.48	0.08	0.40	TTC6
	cg10617796	chr5	OpenSea	0.54	0.12	0.42	EDIL3
	cg21808287	chr2	OpenSea	0.58	0.12	0.45	TNP1
	cg04502126	chr4	OpenSea	0.67	0.21	0.46	LINC01258
	cg25533247	chr19	South Shore	0.63	0.12	0.50	AKAP8L
DNMTΔ3B3	cg08927738	chr20	OpenSea	0.60	0.18	0.42	BCAS1
	cg04751120	chr8	OpenSea	0.68	0.27	0.40	RP1L1
	cg03883300	chr8	OpenSea	0.56	0.15	0.41	-
	cg20364776	chr4	OpenSea	0.80	0.36	0.44	EEF1AL7
	cg13491907	chr9	OpenSea	0.87	0.46	0.41	-
DNMTΔ3B4	cg26295057	chr5	CpG Island	0.55	0.12	0.43	GDNF
	cg22976313	chr14	North Shore	0.77	0.36	0.41	TMEM179
	cg07504154	chr4	OpenSea	0.53	0.10	0.42	ASB5
	cg04220501	chr14	OpenSea	0.53	0.09	0.44	-
	cg06440519	chr19	CpG Island	0.65	0.21	0.45	SYT3
DNMT3L	cg07383415	chr2	OpenSea	0.79	0.34	0.44	FAM49A
	cg12150401	chr7	OpenSea	0.74	0.33	0.41	TAS2R16
	cg20540357	chr5	OpenSea	0.69	0.19	0.50	PPP2R2B
	cg16796997	chr7	OpenSea	0.71	0.17	0.54	-
	cg12958892	chr2	OpenSea	0.62	0.16	0.45	LOC101927285
DNMT1	cg06877649	chr8	OpenSea	0.71	0.20	0.51	-
	cg20943641	chr12	South Shore	0.47	0.07	0.40	CCDC53
	cg09559735	chr22	OpenSea	0.49	0.05	0.43	-
	cg21298915	chr4	OpenSea	0.64	0.21	0.43	-
	cg26096646	chr12	OpenSea	0.64	0.22	0.42	-

All hypermethylated loci of each DNMT isoform were explored the DNA methylation levels by CpG density (Figure 4.8). The majority of the hypermethylated probes was annotated at CpG Island and Opensea (p-value ≤ 0.05) in DNMT cells. Also, the hypermethylated probes after overexpression of *DNMT3A2, DNMT3B4, DNMTΔ3B4*, and *DNMT3L* were significantly annotated at North shore and South Shore (p-value ≤ 0.05). Only the hypermethylated probes after overexpression of DNMTΔ3B2 was significantly annotated at CpG island, North Shelf, and Opensea.

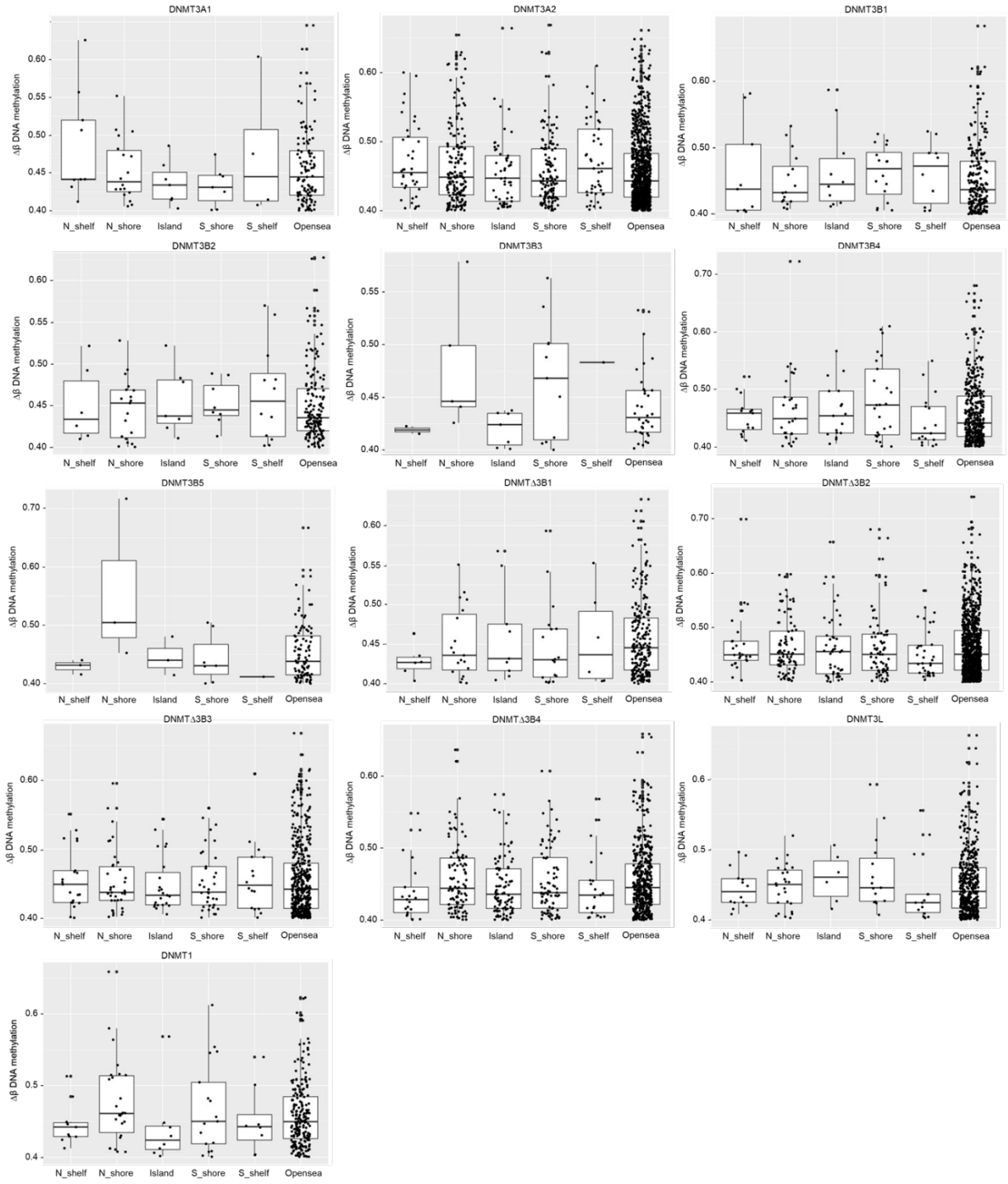


Figure 4.13 Mean Δβ DNA methylation pattern by CpG density of cells overexpressing DNMTs. N_shelf: North Shelf, N_shore: North Shore, Island: CpG island, S_shore: South Shore, S_shelf: South Shelf.

4.6.5 The implication of target DMPs of each DNMT isoform using pathway analysis

To investigate possible links with mechanisms underlying human diseases, the target DMPs (significantly hyper- and hypomethylated probes) were performed in DAVID (for data after setting cut-off at $\Delta\beta \leq -0.2$ and $\Delta\beta \geq 0.2$) and IPA (for data after setting cut-off at $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$. DAVID was used to highlighted pathway members within the biochemical pathways provided by KEGG. The *p*-value and the Benjamini-Hochberg FDR were used to determine significant of enrichment or overrepresentation of terms for each annotation in DAVID. Canonical pathway analysis identified the top pathways from IPA knowledge Base and Fischer's exact test was used to calculate a *p*-value determining the probability of the association between the significant loci and canonical pathway. The top enriched categories of KEGG pathway and canonical pathways with a *p*-value ≤ 0.05 were listed in Table 4.4 and 4.5, respectively.

For KEGG pathway in Table 4.3, the significant probes after overexpression of *DNMT3A1, DNMT3B3, DNMT3B4, DNMTΔ3B1*, and *DNMTΔ3B3*, were associated with neuroactive ligand-receptor interaction. The significant loci in DNMT3A2, DNMTΔ3B4, and DNMT3L cells were correlated with focal adhesion. Calcium signalling was associated with the target DMPs of DNMT3A1, DNMT3B1, and DNMT3B3. PI3K-Akt signalling was enriched in DNMT3B3, DNMTΔ3B4, and DNMT3L. Ascorbate and aldarate metabolism were correlated with the target DMPs of DNMT3B4 and DNMTΔ3B3. The target DMPs of DNMTΔ3B1 and DNMTΔ3B4 were associated with axon guidance. The significant loci in DNMT3B3, DNMT3B5, DNMTΔ3B3, and DNMT3L were correlated with glutamatergic synapse. Chondroitin sulfate biosynthesis was enriched in the target DMPs of DNMT3A1 and DNMT1.

For canonical pathway in Table 4.5, the significant probes after overexpression of *DNMT3B2, DNMT3B3*, and *DNMTΔ3B1*, were associated with Synaptogenesis signalling pathway. The significant probes in DNMTΔ3B3 and DNMT3B3 cells were correlated with CREB signalling in neurons and glutamate receptor signalling. IL-15 production was enriched in the target DMPs of DNMT3A2 and DNMT3L. Citrulline-nitric oxide cycle was associated with the target DMPs of DNMT3B1 and DNMT3L. GP6 signalling pathway was enriched in DNMTΔ3B2 and DNMTΔ3B4. GPCR-mediated integration of enteroendocrine signalling exemplified by an L cell was correlated with the target DMPs of DNMT3B2 and DNMT1. The target DMPs of DNMTΔ3B1 and DNMT1 involved in Renin-angiotensin signalling. The

significantly target probes of DNMT3A2, DNMTΔ3B2, and DNMT3L were associated with sperm motility.

Table 4.4 List of top KEGG pathways of each DNMT isoform.

DNMT isoforms	Top Canonical pathways	*p*-value	Benjamini-Hochberg FDR
DNMT3A1	Cell adhesion molecules	0.016	0.230
	Neuroactive ligand-receptor interaction	0.019	0.140
	Calcium signalling	0.068	0.310
	Chondroitin sulfate biosynthesis	0.077	0.270
	NOD-like receptor signalling	0.690	0.900
DNMT3A2	Focal adhesion	0.00000052	0.000096
	Regulation of actin cytoskeleton	0.00000120	0.000110
	Adherens junction	0.00002300	0.001300
	Neurotrophin signalling	0.00012000	0.005300
	Tight junction	0.00013000	0.004800
DNMT3B1	Calcium signalling	0.00000700	0.001800
	Insulin secretion	0.00003800	0.005
	Circadian entrainment	0.00004900	0.004300
	Rap1 signalling	0.00015000	0.009500
	Pentose and glucoronate interconversions	0.00028000	0.012
DNMT3B2	Protein digestion and absorption	0.00035000	0.088
	Mucin type O-glycan biosynthesis	0.00140000	0.160
	Rap1 signalling	0.00160000	0.130
	Regulation of actin cytoskeleton	0.00340000	0.200
	Ras signalling	0.004	0.190
DNMT3B3	Calcium signalling	0.00004600	0.012
	ECM-receptor interaction	0.00014000	0.017
	PI3K-Akt signalling	0.00020000	0.017
	Neuroactive ligand-receptor interaction	0.00050000	0.032
	Glutamatergic synapse	0.00075000	0.038
DNMT3B4	Pentose and glucoronate interconversions	0.00000360	0.000930
	Ascorbate and aldarate metabolism	0.00001300	0.003600
	Steroid hormone biosynthesis	0.00014000	0.018
	Drug metabolism-cytochrome P450	0.00015000	0.013
	Neuroactive ligand-receptor interaction	0.00015000	0.013

Table 4.5 (continue) List of top KEGG pathways of each DNMT isoform.

DNMT isoforms	Top Canonical pathways	*p*-value	Benjamini-Hochberg FDR
DNMT3B5	Focal adhesion	0.0001100	0.029
	Oestrogen signalling	0.0025000	0.280
	Amyotrophic lateral sclerosis	0.0028000	0.220
	Glutamatergic synapse	0.0036000	0.210
	Regulation of actin cytoskeleton	0.0041000	0.200
DNMTΔ3B1	Neuroactive ligand-receptor interaction	0.0000370	0.00620
	Axon guidance	0.0002800	0.073
	Morphine addiction	0.0003100	0.026
	Hippo signalling	0.0003800	0.050
	Wnt signalling	0.0008200	0.070
DNMTΔ3B2	Phosphatidylinositol signalling	0.0000270	0.00640
	Cholinergic synapse	0.0004100	0.048
	Rap1 signalling	0.0004100	0.035
	Retrograde endocannabinoid signalling	0.0004500	0.032
	Salivary secretion	0.0005400	0.130
DNMTΔ3B3	ECM-receptor interaction	0.0000020	0.00049
	Neuroactive ligand-receptor interaction	0.0001700	0.044
	Ascorbate and aldarate metabolism	0.0002300	0.029
	Glutamatergic synapse	0.0003900	0.032
	Dopaminergic synapse	0.0004500	0.028
DNMTΔ3B4	Focal adhesion	0.0000050	0.00130
	Ras signalling	0.0000110	0.00150
	Proteoglycans in cancer	0.0000380	0.00960
	Axon guidance	0.0000440	0.00560
	PI3K-Akt signalling	0.0004100	0.051
DNMT3L	Focal adhesion	0.0000400	0.010
	ECM-receptor interaction	0.0000540	0.007
	Amoebiasis	0.0007300	0.018
	PI3K-Akt signalling	0.0009500	0.120
	Glutamatergic synapse	0.0015000	0.120
DNMT1	MAPK signalling	0.009	0.770
	Dilated cardiomyopathy	0.026	0.880
	Chondroitin sulfate biosynthesis	0.028	0.790
	ECM-receptor interaction	0.036	0.780
	Pancreatic cancer	0.039	0.730

Genes were categorised to KEGG pathways from gene enrichment analysis was used to calculate *p*-value and Benjamini-Hochberg FDR determining the probability of the association between the significant genes and KEGG pathways.

Table 4.5 List of top canonical pathways of each DNMT isoform.

DNMT isoforms	Top Canonical pathways	*p*-value
DNMT3A1	LXR/RXR activation	0.003
	IL-10 signalling	0.005
	RhoGDI signalling	0.011
	Circadian rhythm signalling	0.011
	GABA receptor signalling	0.011
DNMT3A2	Sperm motility	0.003
	PTEN signalling	0.006
	Myo-inositol biosynthesis	0.009
	IL-15 production	0.010
	SAPK/JNK signalling	0.012
DNMT3B1	Superpathway of citrulline metabolism	0.004
	Lipoate biosynthesis and incorporation II	0.013
	Serotonin and melatonin biosynthesis	0.025
	Citrulline-nitric oxide cycle	0.032
	Urea cycle	0.038
DNMT3B2	PFKFB4 signalling pathway	0.001
	GPCR-mediated integration of enteroendocrine signalling exemplified by an L cell	0.005
	Trehalose degradation II (Trehalase)	0.015
	Synaptogenesis signalling pathway	0.018
	Adenine and adenosine salvage III	0.025
DNMT3B3	Chondroitin sulfate biosynthesis	0.002
	Dermatan sulfate biosynthesis	0.002
	Glutamate receptor signalling	0.002
	CREB signalling in neurons	0.002
	Synaptogenesis signalling pathway	0.003
DNMT3B4	GDP-glucose biosynthesis	0.002
	Glucose and glucose-1-phosphate degradation	0.003
	UDP-N-acetyl-D-galactosamine biosynthesis II	0.005
	PFKFB4 signalling pathway	0.020
	cAMP-mediated signalling	0.020

Table 4.5 (continue) List of top canonical pathways of each DNMT isoform.

DNMT isoforms	Top Canonical pathways	*p*-value
DNMT3B5	HER-2 signalling in breast cancer	0.001
	ErbB signalling	0.014
	iCOS-iCOSL signalling in T helper cells	0.017
	Triacylglycerol degradation	0.018
	Chronic myeloid leukaemia signalling	0.019
DNMTΔ3B1	FGF signalling	0.003
	RANK signalling in osteoclasts	0.004
	Synaptogenesis signalling pathway	0.007
	Endocannabinoid developing neuron pathway	0.010
	Renin-angiotensin signalling	0.010
DNMTΔ3B2	T helper cell differentiation	<0.001
	Calcium signalling	<0.001
	Cellular effects of sildenafil (Viagra)	0.001
	GP6 signalling pathway	0.002
	Sperm motility	0.004
DNMTΔ3B3	Netrin signalling	<0.001
	Amyotrophic lateral sclerosis signalling	<0.001
	Glutamate receptor signalling	0.009
	CREB signalling in neurons	0.001
	Opioid signalling pathway	0.005
DNMTΔ3B4	p53 signalling	<0.001
	Amyotrophic lateral sclerosis signalling	<0.001
	Human embryonic stem cell pluripotency	0.001
	GP6 signalling pathway	0.001
	Leukocyte extravasation signalling	0.001
DNMT3L	Citrulline-nitric oxide cycle	0.001
	Sperm motility	0.004
	Neuroinflammation signalling pathway	0.005
	IL-15 production	0.005
	Dopamine receptor signalling	0.005
DNMT1	EGF signalling	0.008
	Rac signalling	0.008
	Renin-angiotensin signalling	0.010
	ERK/MAPK signalling	0.012
	GPCR-mediated integration of enteroendocrine Signalling exemplified by an L cell	0.015

Genes were categorised to canonical pathways from IPA knowledge Base and Fischer's exact test was used to calculate a *p*-value determining the probability of the association between the significant genes and canonical pathways.

DNA methylation of the protocadherin genes was explored as loss of methylation at the protocadherin gamma gene cluster particularly affects the A and B class variable genes (97). Overexpression of DNMTs increased DNA methylation levels of probes located on PCDHG (Figure 4.14). However, there are some probes showed low methylation after overexpression of DNMTs, *i.e.* cg16541259, cg 16626420, cg23347399, cg00118365, cg00888801, cg07802710, cg21185686, cg27639030, and cg12145907.

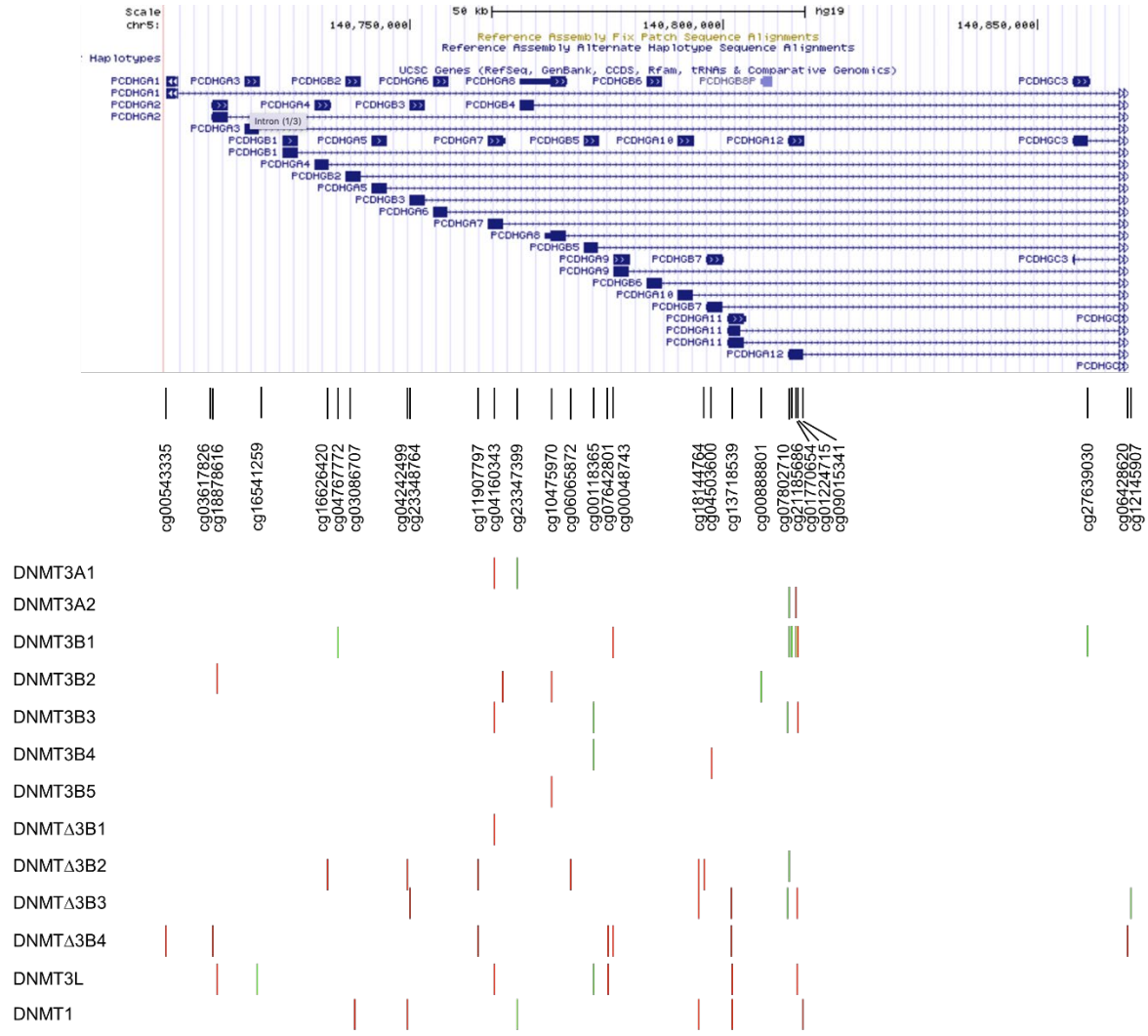


Figure 4.14 DNA methylation levels of PCDHG by overexpressing DNMT cells. The significant probes of PCDHG were represented in green colour (low methylation compared with Myc control) and red colour (hypermethylation compared with Myc control).

In addition to pathway analysis, differentially methylated loci were also categorised to related diseases. Similar as the results from canonical pathway analysis, disease with a *p*-value ≤ 0.05 were listed in Table 4.6. Most DNMTs were associated with cancer and organismal injury and abnormalities. Hereditary disorder and metabolic disorder were associated only with DNMT3B1. Also, only DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, and DNMTΔ3B4 were associated with CVD, psychological disorders, inflammatory response, infectious diseases, and hepatic system disease, respectively. However, these markers could not be specifically linked.

Table 4.6 List of top diseases of each DNMT isoform.

Diseases	3A1	3A2	3B1	3B2	3B3	3B4	3B5	Δ3B1	Δ3B2	Δ3B3	Δ3B4	DNMT1	3L
Cancer													
Organismal injury and abnormalities													
Reproductive system disease													
Gastrointestinal disease													
Auditory disease													
Endocrine system disorders													
Hereditary disorder													
Metabolic disease													
Dermatological disease and conditions													
Respiratory disease													
Cardiovascular disease													
Neurological disease													
Psychological disorders													
Inflammatory response													
Immunological disease													
Connective injury and abnormalities													
Infectious Diseases													
Hepatic System Disease													

Red boxes represent the significant diseases at *p*-value ≤ 0.05 and white boxes are not analysed. Genes were categorised to related disease from IPA knowledge Base and Fischer's exact test was used to calculate a *p*-value determining the probability of the association between the significant genes and diseases.

4.7 Discussion

4.7.1 Main findings

Investigation of the *de novo* targets of DNMT isoforms revealed that DNMT isoforms that are over-expressed in HEK293T cells can influence DNA methylation with increased methylation and decreased methylation at different CpG sites. Although there was some overlap in the target DMPs within DNMT families (e.g. DNMT3A, DNMT3B, and DNMTΔ3B), the majority of the hypermethylated probes was unique to individual DNMTs. This suggests that a splicing variance of DNMTs results in different target sites of each DNMT. Hypermethylation level changes were lower than 0.7, indicating this was a limitation of increase methylation by DNMTs. Moreover, the profiling of *de novo* DNA methylation target CpG sites for each DNMT isoform was revealed and annotated to potential biological pathways and diseases.

4.7.2 *De novo* DNA methylation target DMPs of DNMT isoforms

Nowadays, the activity of DNMT1, DNMT3A, and DNMT3B can be measured using a commercial kit but it cannot measure specific sub-isoforms of DNMTs due to the close similarity of the structure of each DNMT sub-family. In this study, 13 different DNMT cell lines were created and these cells were used to explore the DNA methylation patterns including identifying the *de novo* methylation sites. Also, the use of EPIC array can enable the identification of *de novo* DNA methylation target sites of DNMT isoforms.

This study is the first one to identify the *de novo* methylation sites using biological duplicates from single clonal cells. In contrast, a previous publication investigated the over-expressed DNMTs from a mixed cell population (87). Over 850,000 CpGs were measured for DNA methylation levels in this study and the cut-offs of $\Delta\beta$ value were set at three different values: $\Delta\beta \leq -0.2$ and $\Delta\beta \geq 0.2$, $\Delta\beta \leq -0.3$ and $\Delta\beta \geq 0.3$, and $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$. However, there was no significant different of the distribution of significant loci along with genomic regions among three cut-offs. Additionally, 1,505 loci from the Illumina GoldenGate Methylation Cancer Panel I were explored in this study in order to compare the DNA methylation levels with the previous study. However, only 436 probes out of 1,505 probes passed the quality check and more than 95% of these probes showed $-0.4 < \Delta\beta < 0.4$. Also, the top loci, which were found in the previous study, did not show significant changes in this study. This is due to the fact that the data analysis including normalisation and filtering in this study, was more robust with statistic analysis.

Generally, DNA was hypermethylated in cells that over-expressed each individual DNMT. In this study, hypermethylated probes of each DNMT isoform were significantly located on CpG island and Opensea areas where low methylation levels are, suggesting that these areas were sensitive to induce DNA methylation. Additionally, the significant loci of both hyper and hypomethylated probes were enriched in gene body and TSS1500 areas where may block the initiation of transcription (299, 300) and also may be more important for controlling gene splicing (300) and control of developmental timing of expression (301).

Obviously, there is no difference in the DNMT isoforms after PCA analysis from the whole loci, which passed the quality control and filtering. However, after selecting only the significant loci using cut-offs at $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$, the results of clustering analysis revealed that DNMT3A2, DNMTΔ3B2, and DNMTΔ3B4 contributed to DNA hypermethylation and were clustered in the same group. This suggests that there is a specificity of modulating DNA methylation levels by DNMT isoforms. Not only hypermethylation but hypomethylation was observed after DNMT overexpression. After setting the cut-offs at $\Delta\beta \leq -0.4$ and $\Delta\beta \geq 0.4$, the number of hypomethylated probes exceeded those of hypermethylated probes in cells that over-expressed DNMT3A1, DNMT3B3, and DNMT3B5. The hypotheses of this observation are that 1) trans-regulation of DNMTs can control the promoter activity of certain DNMTs leading to inducing hypermethylation or hypomethylation at its target CpGs 2) the different structure of each DNMT isoform resulted in different function of DNMTs. It suggests that lack of catalytic *C*-terminal of DNMT3B3 and DNMT3B5 is associated with DNA hypomethylation. Similarly, the DNA methylation activities of purified Dnmt3b1-Dnmt3b3 were determined, and results showed that Dnmt3b3 has no activity toward the methyl acceptors (302). However, DNMT3B4 also lacks the conserved methyltransferase motifs IX and X as well as DNMT3B5 does, but DNMT3B4 showed more hypermethylated probes than DNMT3B5. Furthermore, lacking 200 amino acids from the *N*-terminal area of DNMT3B, DNMTΔ3B1-DNMTΔ3B4 showed higher number of hypermethylated probes than DNMT3B families, suggesting that lacking those areas enhanced DNA methylation levels.

DNMT3A2 acts as an active DNMT and displays a localisation pattern suggestive of euchromatin association, while DNMT3A1 is a full-length DNMT3A and is concentrated on heterochromatin (84). DNMT3A2 is highly expressed in embryonic stem cells and embryonal carcinoma cells, while DNMT3A1 is expressed at low levels ubiquitously (84). In this study, the basal DNMT3A2 was highly expressed in HEK293T cell. This might lead to increase a number of DNMT3A2 protein after overexpression and DNMT3A2 increased DNA

methylation of several CpG targets. Another reason is that DNMT3A2 shares binding regions with DNMT3A1 and DNMT3B1, while DNMT3A1 has distinct binding targets (303).

DNMTΔ3B2 showed the highest number of hypermethylated probes and 25% of these probes overlapped with the target probes of other DNMTΔ3Bs. Although, the structure of DNMTΔ3B2 is similar to DNMTΔ3B1, the number of overlapped hypermethylated probes was 5%. Only DNMTΔ3B2 and DNMTΔ3B4 showed low number of hypomethylated probes compared to other DNMT3Bs. It suggests that lacking 200 amino acid *N*-terminal and exon 6 of DNMT3B in both isoforms induces hypermethylation within 103 shared target CpG sites. DNMT1 and DNMT3L induced hypermethylation and this result was concomitant with previous publication (87). Although DNMT3L is an inactive form, DNMT3L plays an important role in modulating DNA methylation through activating or recruiting *de novo* DNMTs (111-113). In this study, DNMT3L showed an activity to increase DNA methylation. This is concordant with Duymich C.E.'s study, which showed that DNMT3L demonstrated the ability to restore DNA methylation in DNMT3B deficient cells (277). Additionally, DNMT3L enhanced DNMT activity of DNMT3A2 and increased SAM binding by binding to DNMT3A2 promoters resulting in reorganisation of DNMT3A2 subunits (304).

4.7.3 Implication of *de novo* DNA methylation targets of DNMTs to biological pathways and diseases

In the introduction chapter (see section 1.4.2, page 14), I reviewed evidence that DNMTs play an important role in human diseases and especially in cancers. Increased hypermethylation of tumour suppressor genes in cancer cells is attributable to the aberrant expression or activity of DNMT1, DNMT3A, or DNMT3B (283, 305-307). The *de novo* target DMPs of each DNMT were located in the genes enriched in cancer and in cancer-related pathways such as IL-10 signalling (308), RhoGDI signalling (309), PTEN signalling (310), HER-2 signalling (311, 312), focal adhesion (313), PI3K-Akt (314), calcium signalling (315), and ERK/MAPK signalling (316). To my knowledge no previous publication has investigated the prediction or association of specific DNMTs to other diseases apart from cancers. These pathways identified in this study, that were predicted from the *de novo* DNA methylation target of individual DNMT isoforms, showed the potential for associations between individual DNMTs and a number of other diseases. The target CpGs of DNMT3B1 was associated with hereditary disorder and metabolic disease, while CVD, psychological disorders, and inflammatory were associated with DNMT3B3, DNMT3B4, and DNMT3B5, respectively.

Infectious diseases, hepatic system disease, and connective injury and abnormalities were specific associated with DNMTΔ3B1, DNMTΔ3B4, and DNMT3L, respectively.

The identified target CpGs would be a good marker to measure DNA methylation changes by specific DNMT isoforms in order to identify specific health conditions. For example, the dysregulation of DNMT3B contributes to ICF syndrome and the depletion of DNMT3B was associated with hypomethylation of *PCDHG* (97, 317). In this study, the overexpression of DNMT isoforms especially DNMTΔ3B can increase DNA methylation levels of *PCDHG*. However, further studies need to investigate and research the function of those targets on specific health conditions using the different cohort studies.

4.8 Conclusion

Using an unbiased genome-wide technique, this study examined the *de novo* target DMPs of 13 DNMT isoforms from stably transduced single-cell-derived clonal lines. The DNA methylation patterns of each DNMT were not dependant on the structure of DNMT variants. Although some DNMT isoforms showed similar catalytic sites, they modulated DNA methylation levels in different CpG sites. Also, some DNMT isoforms might show a trans-regulated activity on other DNMTs leading to hypomethylation. The majority of the preferential target sites of DNMTs were located on CpG island and Opensea where there is a space for increased DNA methylation. Not only the similarity of cancer-associated genes was identified from target DNMTs, but there were some unique genes associated with specific health condition. Therefore, there is a possible way for future mechanistic research to predict specific health condition from specific DNMT activity.

Chapter 5: The effect of dietary constituents (polyphenols and vitamin C) on DNA methylation levels in over-expressed DNMT cell lines

5.1 Introduction

As reviewed in the introduction section (see section 1.6.2, page 22), nutrition modulates the DNA methylation patterns either at the global scale or at locus specific CpG sites. Specific nutrients can alter DNA methylation in two possible ways: 1) through the provision of substrates or cofactors required for one-carbon metabolism (Figure 1.10), which generates SAM; and 2) through inhibition the activity of DNMT. The methyl donor (SAM) for DNMT is synthesised via one-carbon metabolism using nutrients such as folate, choline, betaine, vitamin B_{12} and B_6 (198). Diets that are deficient in methyl donors such as folate cause reduced concentration of SAM and increased concentration of SAH (318, 319), which have been correlated with global DNA hypomethylation (251). In older people (age 65-75 years), supplementation with folic acid (400 μg) and vitamin B_{12} (500 μg) for two years increased DNA methylation at hypomethylated areas in the genome including North Shore, South Shore, and CpG islands (320). In addition, these supplementations increased methylation of locus-specific loci related pre-symptomatic dementia (one of the papers I published as the first-author during my PhD training). Altering the enzyme activity of DNMTs is the second possible mechanism by which nutrients may modulate patterns of DNA methylation.

The potential of specific nutrients and other food constituents to inhibit DNMT activity has been evaluated (see introduction section 1.5.2, page 17). For example, tea polyphenols and bioflavonoids (quercetin, fisetin, and myricetin) have been found to be DNMT1 inhibitors and the structural model of DNMT1 catalytic domain revealed a substantial interaction with polyphenols (166, 167). In studies with cell lines, treatment with EGCG reduced mRNA levels and protein expression of DNMTs resulting in decreased DNA methylation and increased histone acetylation (168, 169). Genistein inhibits the effects of DNMT1 activity in a dose-dependent manner, but has no effect on expression of DNMTs at the mRNA level (227). CA and chlorogenic acid, common coffee polyphenols, increased SAH and inhibited of CpG methyltransferase (M.SssI)-mediated DNA methylation (173). Analysis of molecular docking showed that curcumin blocked the catalytic domain of DNMT1 and it also inhibited the activity of M.SssI (174). However, there have been few investigations of the effects of nutrients and other food constituents on the activity of specific DNMT isoforms. In this study, the sensitivity and specificity of DNMT inhibitors were investigated with polyphenols and vitamin C in various concentrations in order to understand

the effect of these compounds on DNMT sub-isoforms. EGCG, curcumin, theaflavin, CA, and vitamin C were investigated in this study as these compounds were found the inhibitory effect on DNMTs (see the Table 1.2 in Introduction section 1.5.2, page 17). DNA methylation of the target DNMTs (hypermethylated probes from EPIC array analysis described in Chapter 4) was quantified with/without exposure to each of polyphenols and vitamin C separately in cell lines that over-expressed each of the individual DNMT isoforms compared with control cells. From EPIC data analysis, DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B4, and DNMT1 showed more than 300 hypermethylated probes, and they were selected for dietary treatment. Although DNMT1 showed less than 300 hypermethylated probes, DNMT1 was selected as it is the full length isoform. Furthermore, for CA only, DNMT activity was measured to confirm the inhibitory effect of the food derived-polyphenol on target DNMTs. This research aims to find the specific effects of food constituents on activity of individual DNMTs as a first step towards being able to target, and inhibit, the activities of these DNMT isoforms.

5.2 Hypotheses

The hypotheses for this study were:

1. The global DNA methylation is reduced in treated cells with food-derived polyphenols (theaflavin, EGCG, CA, curcumin) and vitamin C, compared with untreated cells.
2. DNMT inhibitors attenuate DNA methylation with specificity to target CpG of DNMTs.
3. DNMT activity is decreased by DNMT inhibitor including food-derived polyphenols (theaflavin, EGCG, CA, curcumin) and vitamin C.

5.3 Aim

- To test the above hypotheses, the sensitivity and specificity of the effects of food-derived polyphenols (theaflavin, EGCG, CA, and curcumin) and vitamin C on methylation will be measured using target CpG sites of each DNMT isoform as a marker.

5.4 Objectives

The objectives of this study were:

1. To perform pyrosequencing to confirm the DNA methylation level results from EPIC array (in chapter 4) in the overexpressing DNMT isoform.
2. To optimise the concentration of dietary constituents and time for treatment into the cell lines that overexpress each of the DNMTs individually.
3. To use pyrosequencing to provide DNA methylation levels of global and specific CpGs in treated cells with DNMTi (food-derived compounds) compared with untreated cells.

5.5 Overview of methods

A detailed description of the experimental procedure and methods for measuring cell viability, DNA methylation, and DNMT activity can be found in the Method chapter sections 2.3 and 2.4, pages 45 and 50, respectively.

In brief, seven cell lines that overexpress each of the DNMTs individually (DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1, and DNMT3L) were treated with several concentrations of food-derived polyphenols (theaflavin, EGCG, CA, curcumin) and vitamin C (see section 2.3.1 in Methods section, page 45) to find the optimum concentrations of each of these compounds for use in the cell culture experiments. After that, cells overexpressing each individual DNMT isoform were cultured with the specific concentration of each compound for 48 hours (see section 2.3.2 in Methods section, page 46). Global DNA methylation and methylation at specific CpG sites was quantified using pyrosequencing (see section 2.3.3 in Methods section, page 46). Finally, DNMTΔ3B4 was extracted from the cell line that over-expressed this DNMT isoform (see section 2.4.1 in Methods section, page 50), precipitated using IP/Co-IP (see section 2.4.2 in Methods section, page 50) and its activity was determined using a DNMT inhibition assay (see section 2.4.3 in Methods section, page 51).

5.6 Results

5.6.1 Viability of cells overexpressing DNMTΔ3B2 and Myc after treatment with theaflavin, EGCG, CA, curcumin, and vitamin C

Cells overexpressing DNMTΔ3B2 and Myc control cell were selected as the experimental sample in order to determine the cell viability after treatment with a range of concentrations of theaflavin, EGCG, CA, curcumin, and vitamin C. This choice was made because cells

overexpressing DNMTΔ3B2 showed the highest number of hypermethylated CpG sites (using data from the EPIC array – see Chapter 4, section 4.6.1, page 84) and the cell-line overexpressing Myc was used as the control. Cell viability was assessed with theaflavin concentrations of 0, 8.1, 20.1, 40.3, 80.5, 161.0 μM and with concentrations of 0, 10, 25, 50, 100, 200 μM of vitamin C, EGCG, CA, and curcumin (Figure 5.1).

Since the average doubling time of HEK293T cells is 24 hours, cell viability was assessed after treatment with food constituents for up to 72 hours to select the optimum of incubation time. The viability of cells overexpressing DNMTΔ3B2 and of Myc cells after treatment with various concentrations of food constituents showed similar trends in both cell lines (Figure 5.1). The proportion of viable cells declined linearly with increasing dose of theaflavin, EGCG, CA and, especially, curcumin, but there was no evidence of any effect of vitamin C on viability of either cell line at any of the doses tested (Figure 5.1). In addition, for curcumin, but not for any of the other food constituents, there was a strong effect of time with greater loss of cell viability with increasing duration of exposure (Figure 5.1). The concentrations of food constituents to be used in further experiments were selected so that the percentage of cell viability was higher than 50%. These were: EGCG at 0, 50, 100 μM, curcumin at 0, 10, 25 μM, CA at 0, 100, 200 μM, theaflavin at 0, 80.5, 161 μM and vitamin C at 0, 100, 200 μM. Also, the optimum of incubation time was 48 hours.

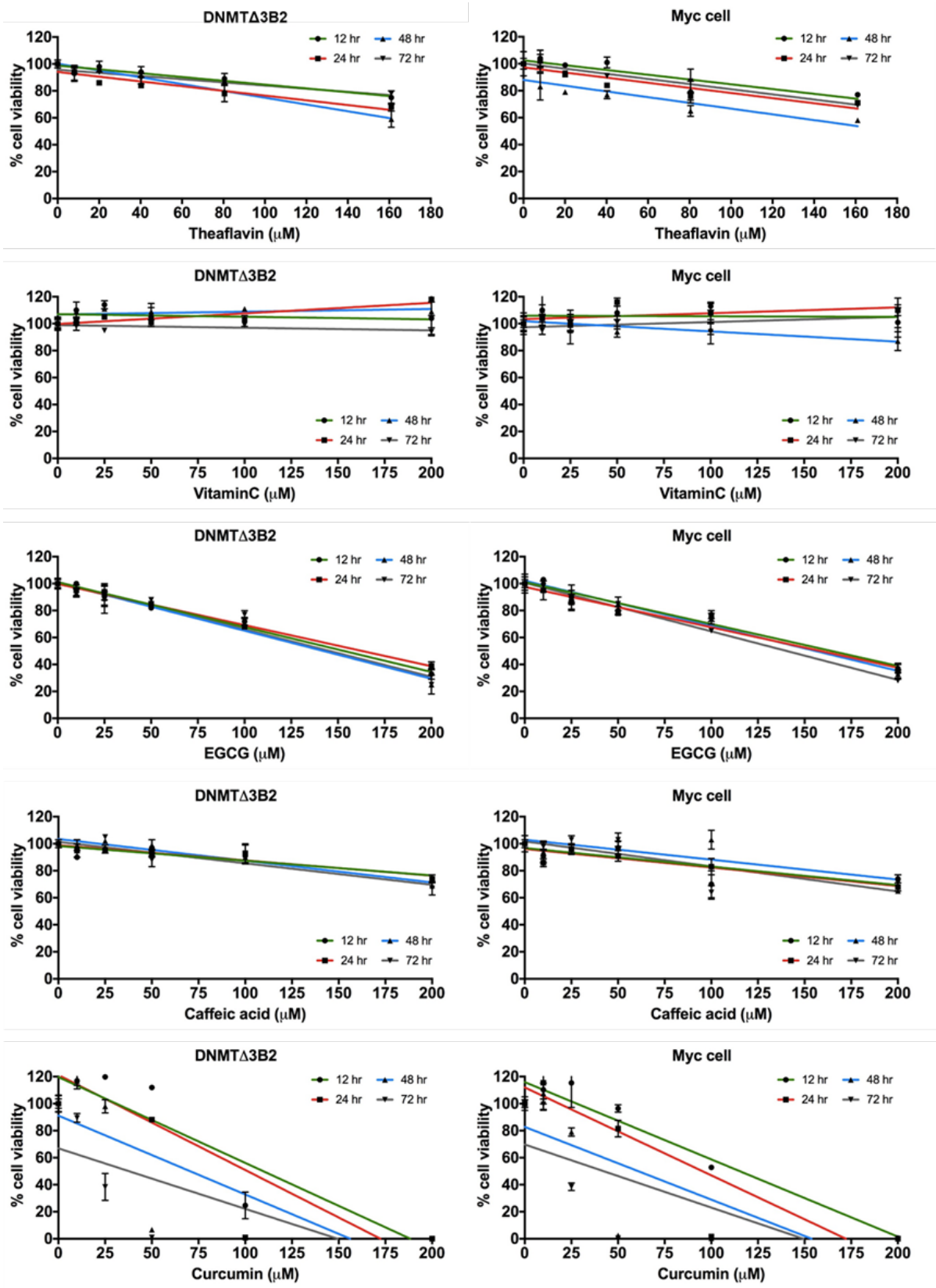


Figure 5.1 Viability of the cell lines overexpressing DNMTΔ3B2 and Myc after treatment with selected dietary constituents. The concentrations of selected dietary constituents were theaflavin at 0, 8.1, 20.1, 40.3, 80.5, 161.0 μM and vitamin C, EGCG, CA, curcumin, at 0, 10, 25, 50, 100, 200 μM. Cells were treated for 12, 24, 48, and 72 hours. Error bars represent standard deviation from triplicates.

5.6.2 Effects of treatment with food constituents on global DNA methylation

To understand global DNA methylation level, LINE-1 assay was measured in Myc (control) and cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1 and DNMT3L. It was found for these cells that global DNA methylation was significantly (p-value ≤ 0.05) increased compared with the control Myc cell (Figure 5.2).

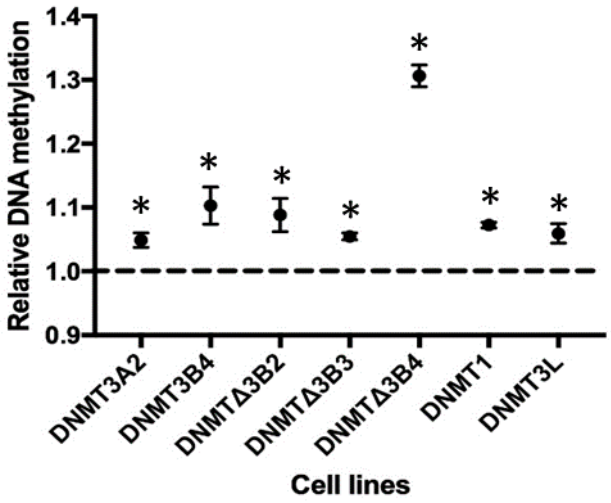


Figure 5.2 The level of LINE-1 methylation in DNMT cells. Global DNA methylation level (using LINE-1 assay) was measured in the selected cell lines (DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1, and DNMT3L) compared with that in the control Myc cell line. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05.

Global DNA methylation (assessed used the LINE-1 assay) was quantified in cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1, and DNMT3L after treatment with theaflavin, vitamin C, EGCG, CA, and curcumin. After the 48-hour treatment with CA, global DNA methylation of cells overexpressing DNMTΔ3B4 decreased significantly from 66.2% to 63.2% (p-value ≤ 0.05) at 100 μM CA and to 60.6% (p-value ≤ 0.05) at 200 μM CA (Figure 5.3) but there were no significant changes global DNA methylation for the other cell lines at either concentration of CA.

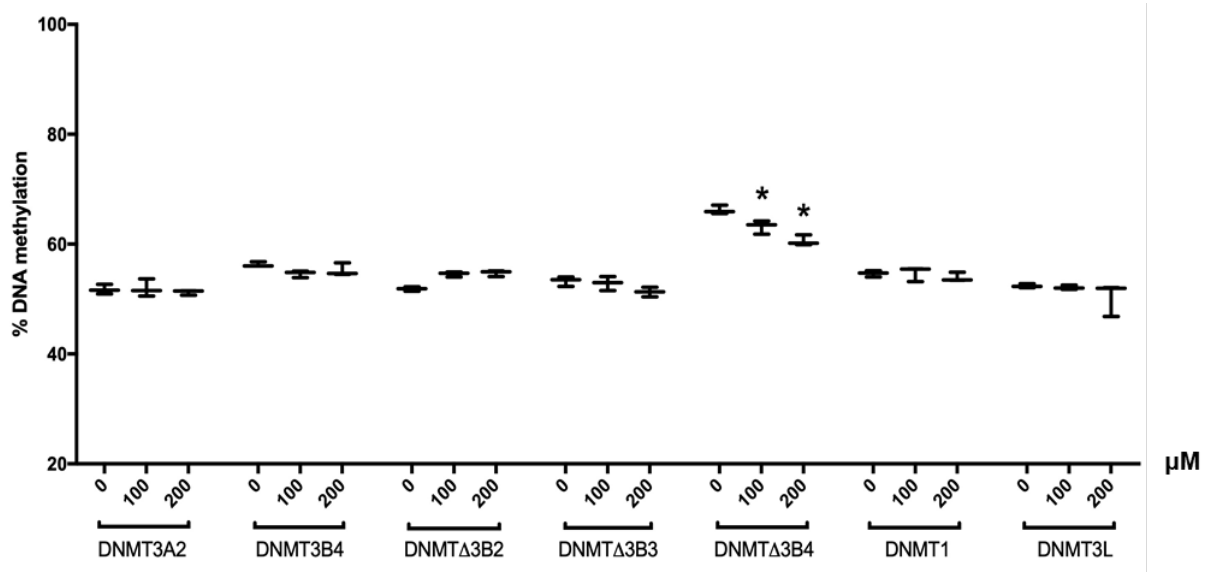


Figure 5.3 Global DNA methylation changes after treatment with CA. The global DNA methylation levels (using LINE-1 assay) were measured after treatment with CA at 100 and 200 μM for 48 hours in cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1, and DNMT3L. Error bars represent standard deviation from triplicates and * represents *p*-value ≤ 0.05 compared with untreated condition.

As for EGCG, the concentration at 100 μM reduced significantly of the global DNA methylation from 53.7% to 48.3% (*p*-value ≤ 0.05) in cells overexpressing DNMT1 (Figure 5.4).

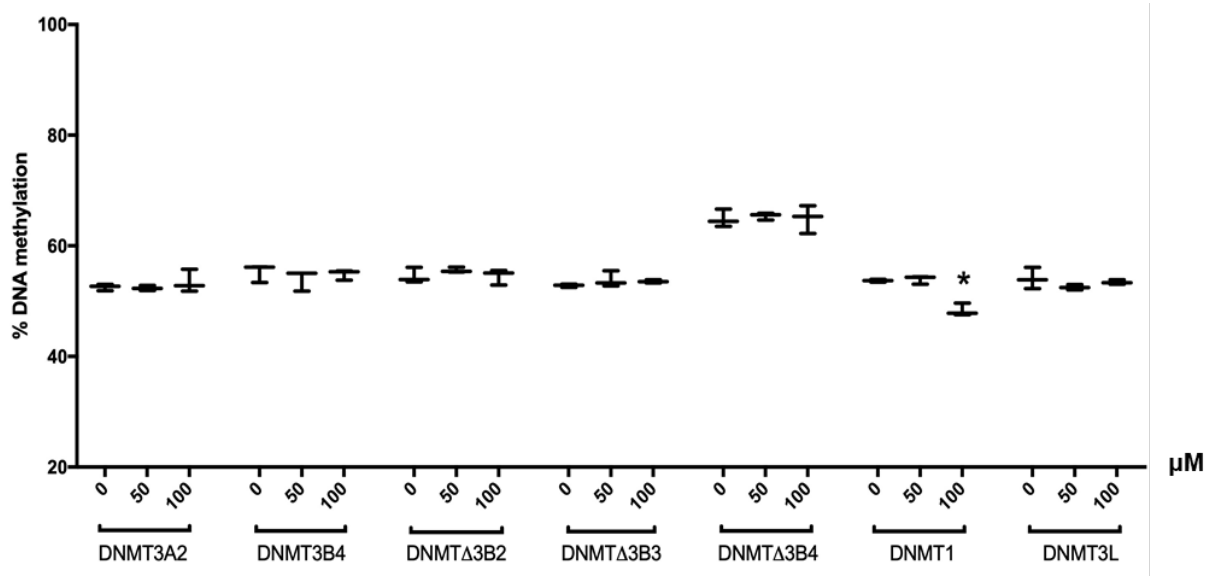


Figure 5.4 Global DNA methylation changes after treatment with EGCG. The global DNA methylation levels (using LINE-1 assay) were measured after treatment with EGCG at 50 and 100 μM for 48 hours in cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1, and DNMT3L. Error bars represent standard deviation from triplicates and * represents *p*-value ≤ 0.05 compared with untreated condition.

For curcumin and vitamin C, there were no significant changes after treatment (Figure 5.5 and 5.6). Global DNA methylation level in cells overexpressing DNMT1 was decreased significantly from 53.7% to 48.0% (p-value ≤ 0.05) after treatment with theaflavin at 80.5 μM for 48 hours (Figure 5.7). I cannot detect the global methylation from DNA of cells overexpressing DNMT3L treated with curcumin at 25 μM and of cells overexpressing DNMTΔ3B3, DNMTΔ3B4, DNMT1, and DNMT3L treated with theaflavin at 161 μM for 48 hours (data not available).

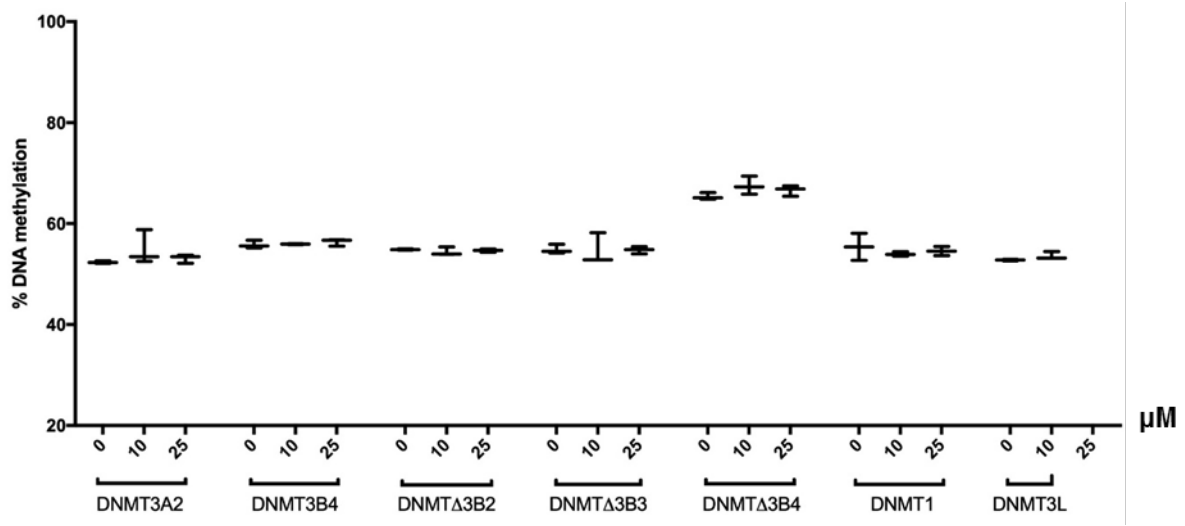


Figure 5.5 Global DNA methylation changes after treatment with curcumin. The global DNA methylation levels (using LINE-1 assay) were measured after treatment with curcumin at 10 and 25 μM for 48 hours in cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1, and DNMT3L. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with untreated condition.

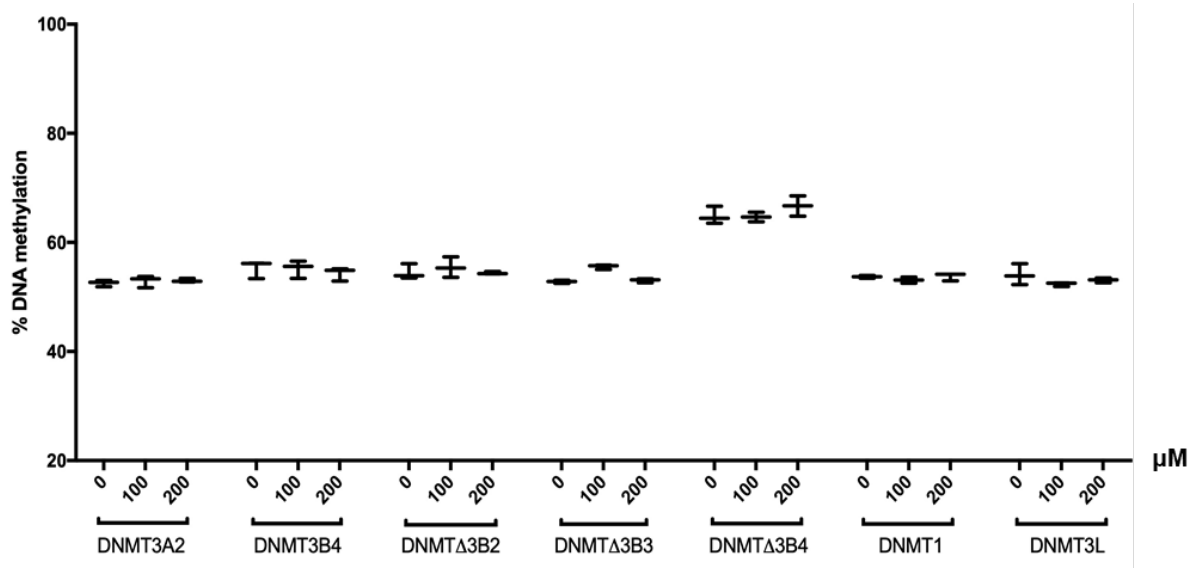


Figure 5.6 Global DNA methylation changes after treatment with vitamin C. The global DNA methylation levels (using LINE-1 assay) were measured after treatment with vitamin C at 100 and 200 μM for 48 hours in cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1, and DNMT3L. Error bars represent standard deviation from triplicates and * represents *p*-value ≤ 0.05 compared with untreated condition.

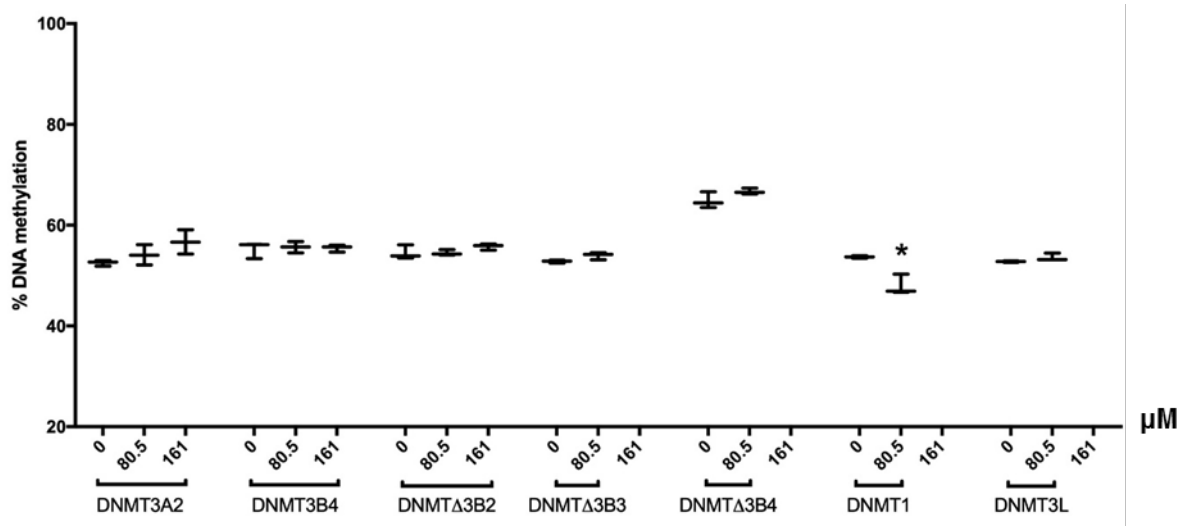


Figure 5.7 Global DNA methylation changes after treatment with theaflavin. The global DNA methylation levels (using LINE-1 assay) were measured after treatment with theaflavin at 80.5 and 161 μM for 48 hours in cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1, and DNMT3L. Error bars represent standard deviation from triplicates and * represents *p*-value ≤ 0.05 compared with untreated condition.

5.6.3 Screening of the effect of selected food constituents on DNA methylation changes of target CpGs for across DNMT isoforms

To identify the specificity of inhibitory effects from theaflavin, EGCG, CA, curcumin, and vitamin C on any of the five DNMT isoforms (DNMT3A2, DNMT3B4, DNMTΔ3B2,

DNMTΔ3B3, and DNMTΔ3B4), DNMT-specific hypermethylated CpG sites were selected from EPIC array analysis. Methylation at cg25843713 (Figure 5.8A) and cg04458645 (Figure 5.8B) in five selected cells was higher than in Myc control cell and cells overexpressing DNMT3L (non-target cell). Therefore, cg25843713 and cg04458645 were selected to quantify the methylation in across cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4, including cells overexpressing DNMT3L as non-target cell, after treatment with selected food constituents.

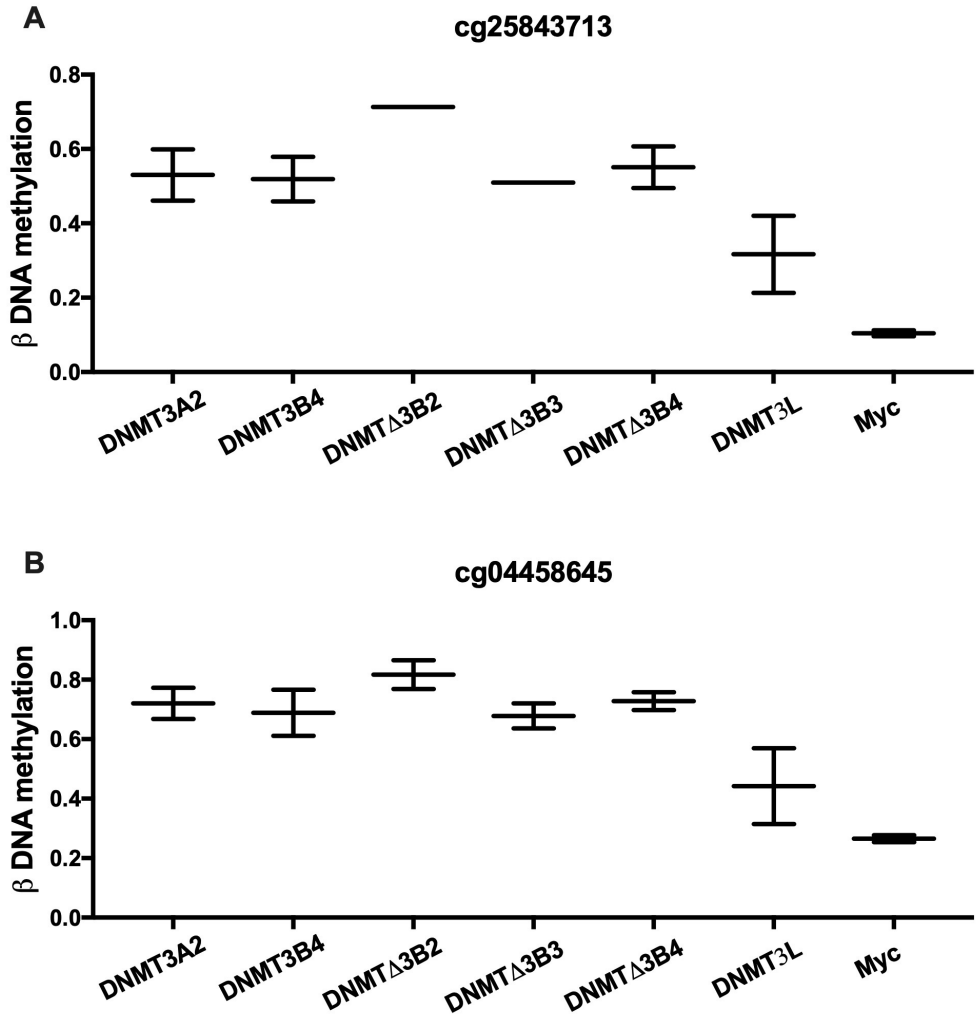


Figure 5.8 DNA methylation levels (data from EPIC array) of target CpGs for across DNMT isoforms. A) cg25843713 and B) cg04458645 were specific across 5 DNMTs (DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4) compared with cells overexpressing DNMT3L and Myc control cell. Error bars represent standard deviation from duplicates.

After 48 hours of CA treatment, CA at 100 and 200 μM exhibited a significant inhibitory effect on cg25843713 (p-value ≤ 0.05) in cells overexpressing DNMTΔ3B2 (from 59.1% to 56.3% at 100 μM and 56.1% at 200 μM) and cells overexpressing DNMTΔ3B4 (from 72.5% to 66.6% at 100 μM and 64.5% at 200 μM) (Figure 5.9A) and cg04458645 in

cells overexpressing DNMTΔ3B4 (from 77.2% to 71.2% at 100 µM and 68.5% at 200 µM) (Figure 5.9B).

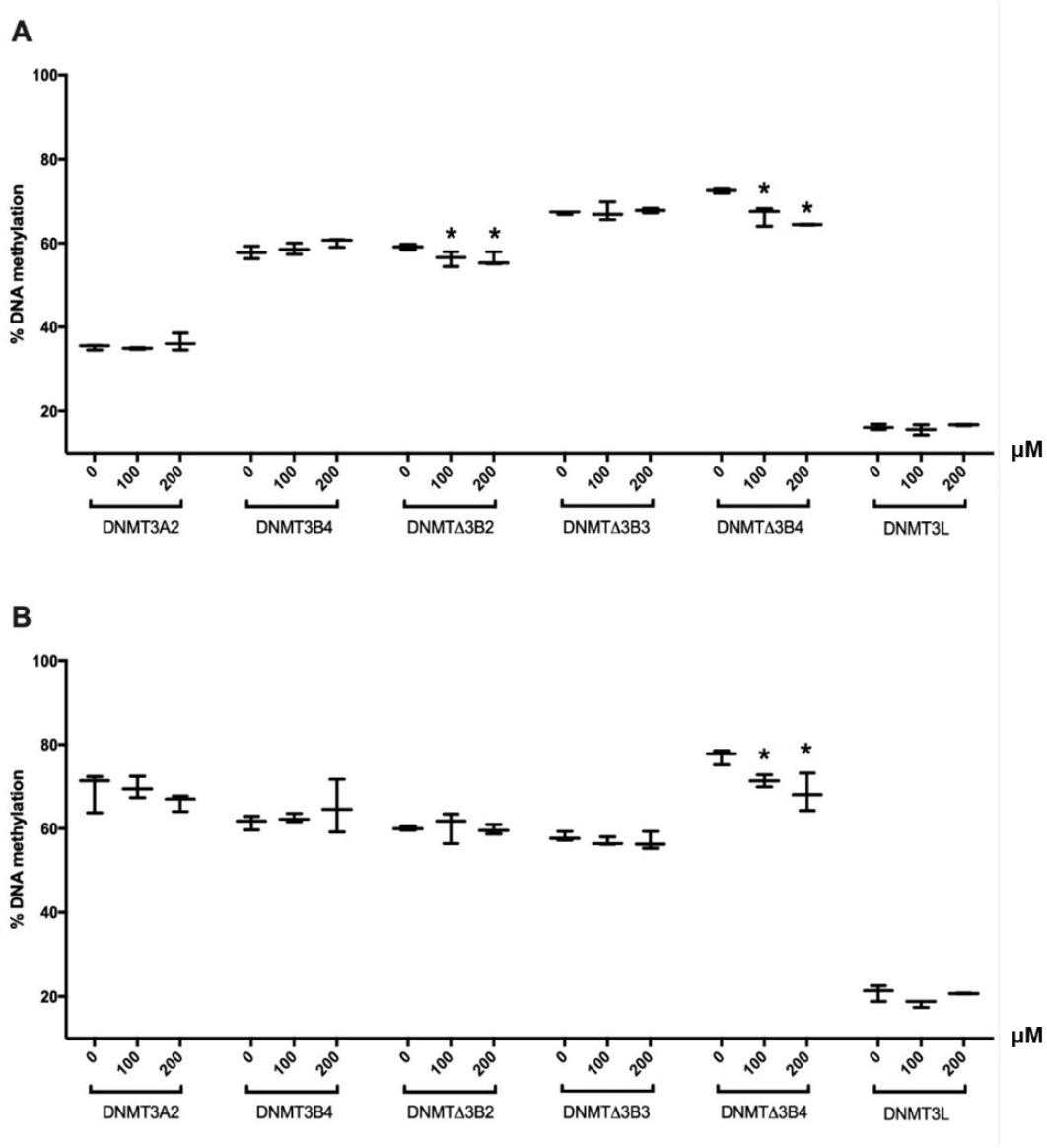


Figure 5.9 Effect of CA at 100 and 200 µM on specific CpG sites for 48 hours. A) cg25843713 and B) cg04458645 were specific to across 5 DNMTs (DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4). Cells overexpressing DNMT3L was a non-target cell of these CpGs. Error bars represent standard deviation from triplicates and * represents *p*-value ≤ 0.05 compared with untreated condition.

EGCG at 50 µM showed a significant increase of DNA methylation (*p*-value ≤ 0.05) at cg25843713 in cells overexpressing DNMTΔ3B4 (from 69.0% to 73.1%) and DNMT3L (from 12.5% to 17.2%) (Figure 5.10A) and also at cg04458645 in cells overexpressing DNMTΔ3B3 (from 58.5% to 63.2%), while this concentration exhibited a significant decrease (*p*-value ≤ 0.05) of DNA methylation at cg04458645 in cells overexpressing DNMTΔ3B2 (from 64.4% to 59%) (Figure 5.10B). Moreover, EGCG at 100 µM increased significantly

DNMT methylation (p-value ≤ 0.05) at cg25843713 in cells overexpressing DNMT3L (from 12.5% to 16.5%) (Figure 5.10A) and cg04458645 in cells overexpressing DNMTΔ3B4 (from 72.5% to 77.8%), but this concentration decreased DNA methylation at this loci in cells overexpressing DNMT3A2 (from 70.7% to 64.8%) and DNMT3B4 (from 60.8% to 53.8%) (Figure 5.10B).

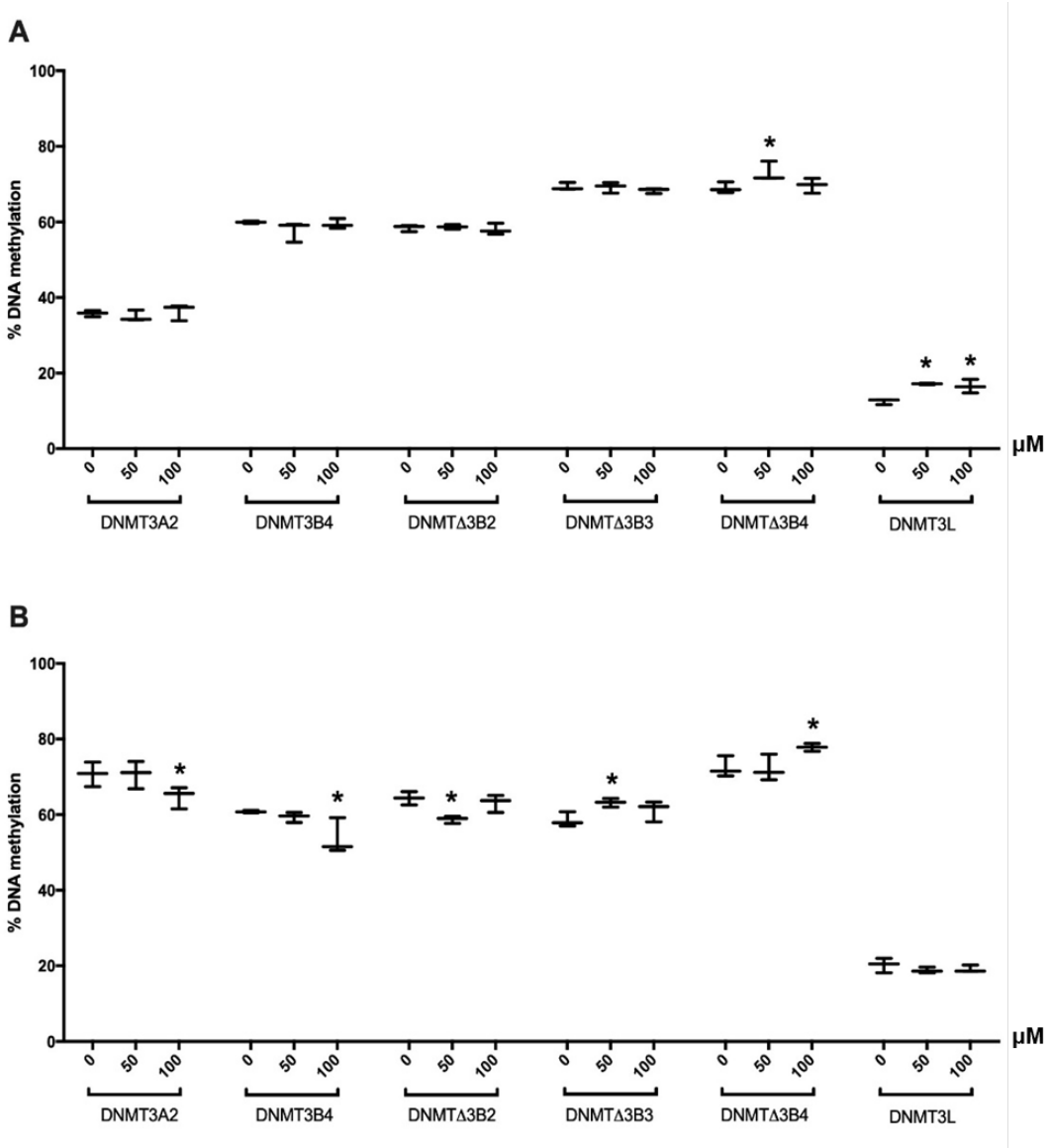


Figure 5.10 Effect of EGCG at 50 and 100 μM on specific CpG sites for 48 hours. A) cg25843713 and B) cg04458645 were specific to across 5 DNMTs (DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4). Cells overexpressing DNMT3L was a non-target cell of these CpGs. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with untreated condition.

Curcumin at 10 μM showed a significant increase of DNA methylation (p-value ≤ 0.05) at cg25843713 in cells overexpressing DNMT3B4 (from 45.6% to 60.0%) (Figure 5.11A). Additionally, at 25 μM of curcumin, it significantly enhanced DNA methylation levels (p-value ≤ 0.05) at cg25843713 in cells overexpressing DNMT3A2 (from 34.44% to

36.35%) and DNMT3B4 (from 45.6% to 64.4%) (Figure 5.11A). This concentration also significantly increased of DNA methylation levels at cg04458645 in cells overexpressing DNMTΔ3B2 (from 60.3% to 64.0%) (Figure 5.11B). However, curcumin at 25 µM exhibited a significant inhibitory effect on cg04458645 (p-value ≤ 0.05) in cells overexpressing DNMTΔ3B3 (from 60.4% to 55.1%) (Figure 5.11B). Data is not available from cg25843713 and cg04458645 in cells overexpressing DNMT3L treated with curcumin at 25 µM for 48 hours due to technical limitations.

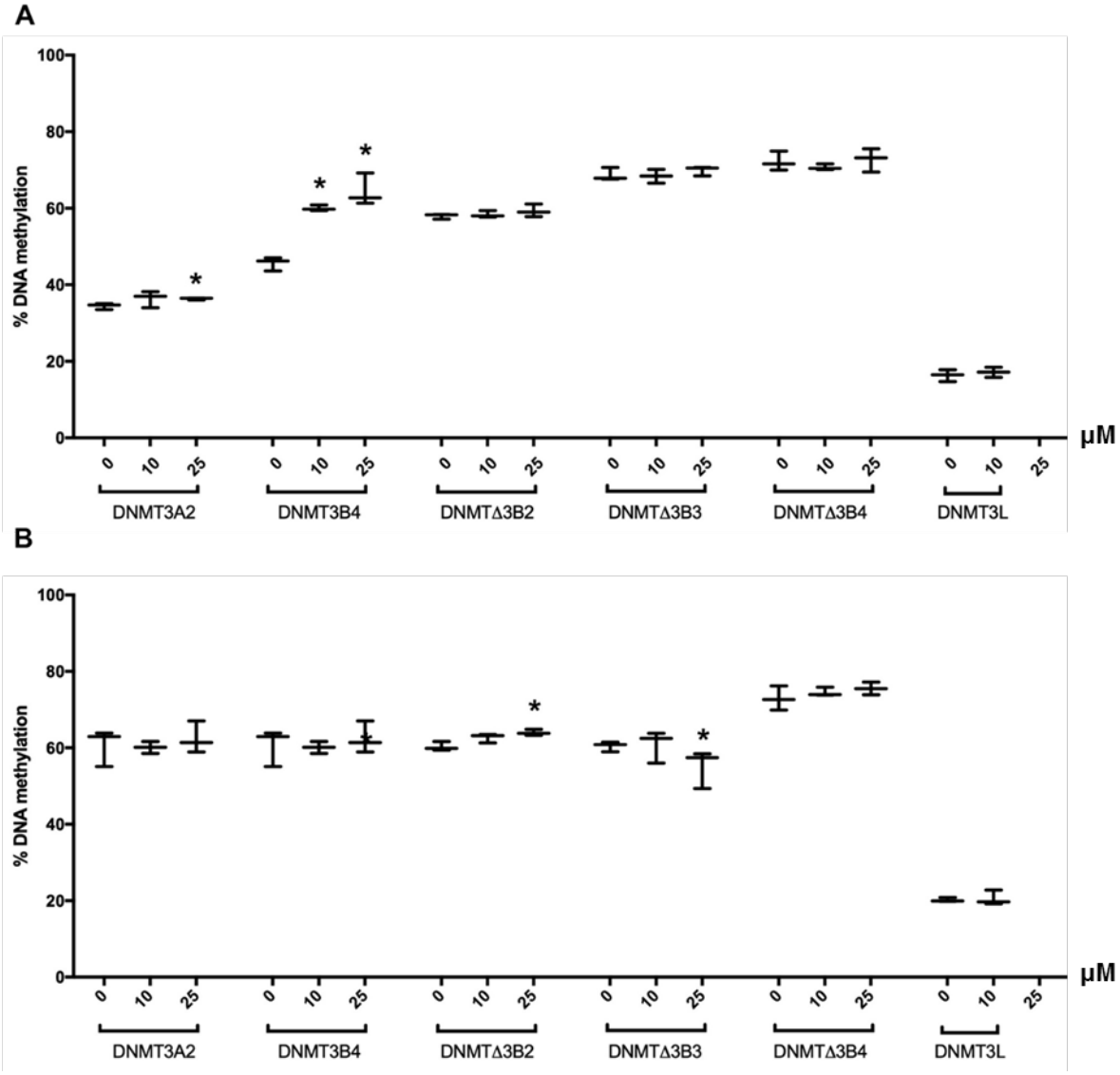


Figure 5.11 Effect of curcumin at 10 and 25 µM on specific CpG sites for 48 hours. A) cg25843713 and B) cg04458645 were specific to across 5 DNMTs (DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4). Cells overexpressing DNMT3L was a non-target cell of these CpGs. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with untreated condition.

At 100 and 200 µM of vitamin C, it decreased significantly DNA methylation levels at cg25843713 in cells overexpressing DNMT3A2 (from 35.8% to 33.4% and 32.8%), DNMT3B4 (from 63.7% to 51.2% and 53.3%), DNMTΔ3B2 (from 58.4% to 52.7% and 53.5%), DNMTΔ3B3 (from 69.3% to 63.8% and 64.9%), and DNMTΔ3B4 (from 69.0% to 48.5% to 61.5%) but these concentrations also decreased significantly DNA methylation levels at this loci in cells overexpressing DNMT3L (from 12.5% to 9.7% to 9.2%, respectively) (Figure 5.12A). Also, at a concentration of 100 µM, it significantly reduced DNA methylation levels at cg04458645 (p-value ≤ 0.05) in cells overexpressing DNMTΔ3B2 (from 64.4% to 59.2%) (Figure 5.12B).

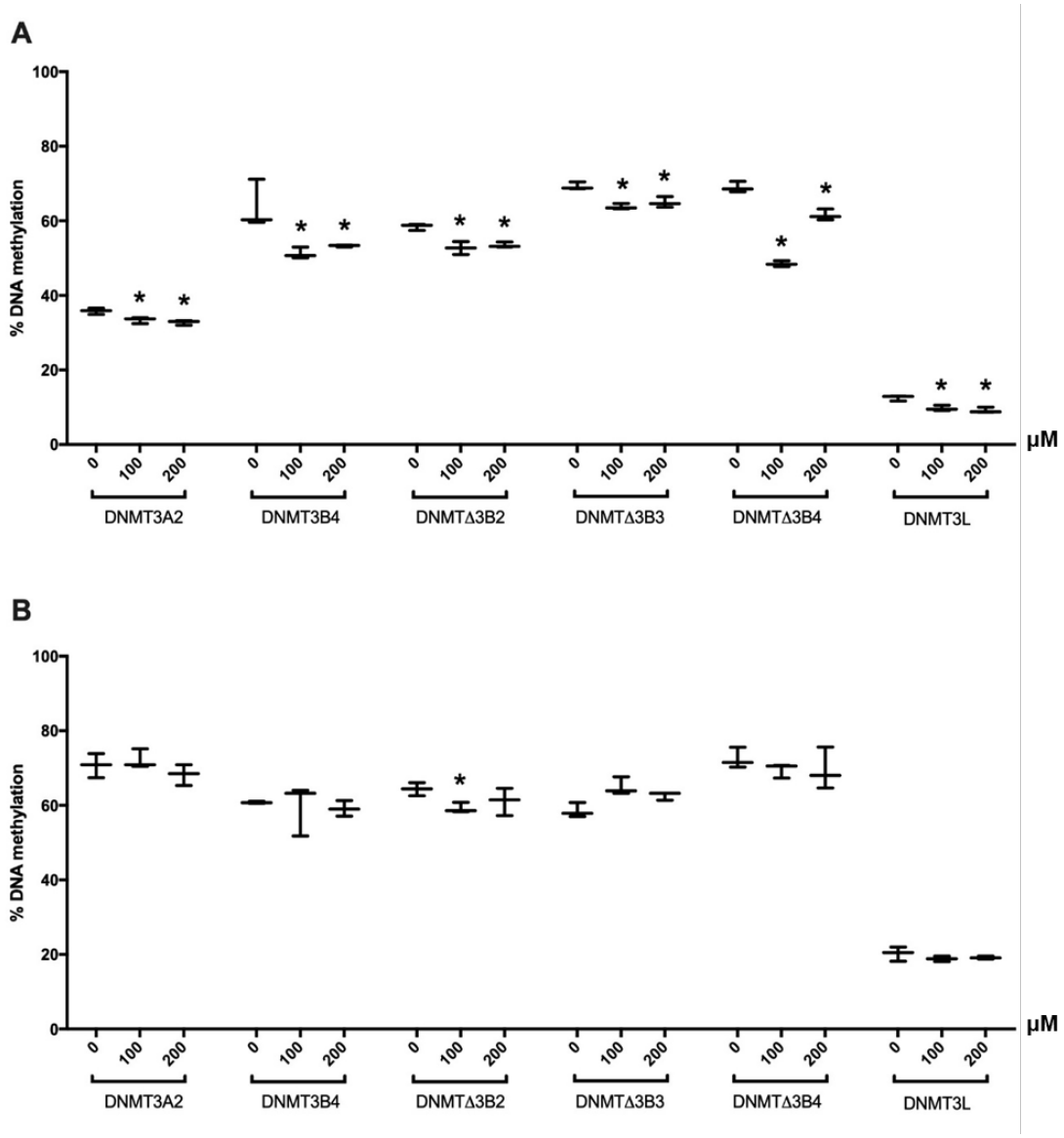


Figure 5.12 Effect of vitamin C at 100 and 200 µM on specific CpG sites for 48 hours. A) cg25843713 and B) cg04458645 were specific to across 5 DNMTs (DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4). Cells overexpressing DNMT3L was a non-target cell of these CpGs. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with untreated condition.

As for theaflavin, its concentration at 161 μM increased significantly DNA methylation levels at cg25843713 in cells overexpressing DNMT3A2 (from 35.8% to 38.1%) and DNMTΔ3B2 (from 58.4% to 61.9%) (p-value ≤ 0.05) (Figure 5.13A). At 80.5 μM of theaflavin, the DNA methylation of cg25843713 also increased significantly in cells overexpressing DNMT3L (from 12.5% to 16.0%) (p-value ≤ 0.05), but it was significantly decreased in cells overexpressing DNMT3B4 (from 60.0% to 54.1%) (p-value ≤ 0.05) (Figure 5.13A). This nutrient at 161 μM enhanced DNA methylation level of cg04458645 in cells overexpressing DNMT3A2 (from 70.7% to 74.8%) but this concentration decreased DNA methylation at cg04458645 in cells overexpressing DNMT3B4 (from 60.79 to 60.08) (p-value ≤ 0.05) (Figure 5.13B). However, DNA methylation level of this loci was significantly decreased (p-value ≤ 0.05) in cells overexpressing DNMT3L (from 20.2% to 17.8%) after treatment with theaflavin at 80.5 μM (Figure 5.13B). Moreover, I cannot detect the percentage of DNA methylation on both cg25843713 and cg04458645 in cells overexpressing DNMTΔ3B3, DNMTΔ3B4, and DNMT3L treated with theaflavin at 161 μM for 48 hours (data not available).

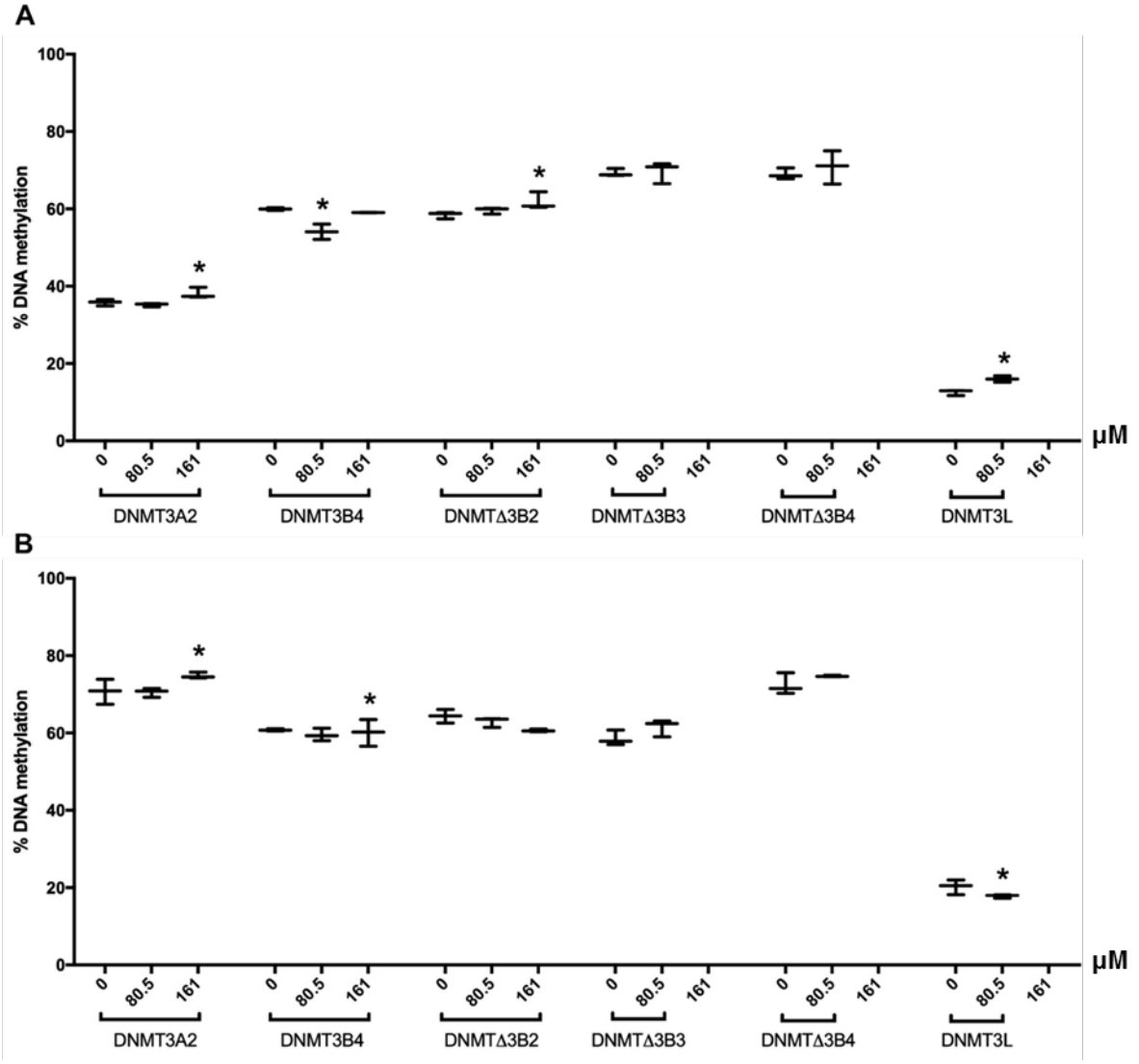


Figure 5.13 Effect of theaflavin at 80.5 and 161 μM on specific CpG sites for 48 hours. A) cg25843713 and B) cg04458645 were specific to across 5 DNMTs (DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4). Cells overexpressing DNMT3L was a non-target cell of these CpGs. Error bars represent standard deviation from triplicates and * represents *p*-value ≤ 0.05 compared with untreated condition.

5.6.4 The specificity and sensitivity of DNMT isoforms on selected food constituents

To identify the specificity and sensitivity of each DNMT isoform to each of the food constituents (CA, EGCG, curcumin, vitamin C, and theaflavin), methylation at one or two target CpG sites from top 10 DMPs for each DNMT isoform was quantified by pyrosequencing after treatment with selected food constituents. cg02372111 and cg16204524 were specific targets of DNMT3A2, cg02788195 for DNMT3B4, cg21808287 and cg

25533247 for DNMTΔ3B2, cg08927738 and cg20364776 for DNMTΔ3B3, cg22976313 and cg07504154 for DNMTΔ3B4, cg01065960 for DNMT1, cg12150401 and cg20540357 for DNMT3L (see Chapter 4, 4.6.3, Table 4.2). Most of these specific loci for each individual DNMT showed no differences of the DNA methylation levels measured by EPIC array and pyrosequencing (Figure 5.14) but the DNA methylation levels, measured by pyrosequencing, of cg02788195, cg16204524, cg21808287, cg25533247, cg08927738, cg07504154, cg12150401, and cg20540357 were lower than measuring by EPIC array. In each case, the cell lines were treated, separately, with 100 and 200 μM of CA and vitamin C, 50 and 100 of EGCG, 10 and 25 μM of curcumin, 80.5 and 161 μM of theaflavin for 48 hours.

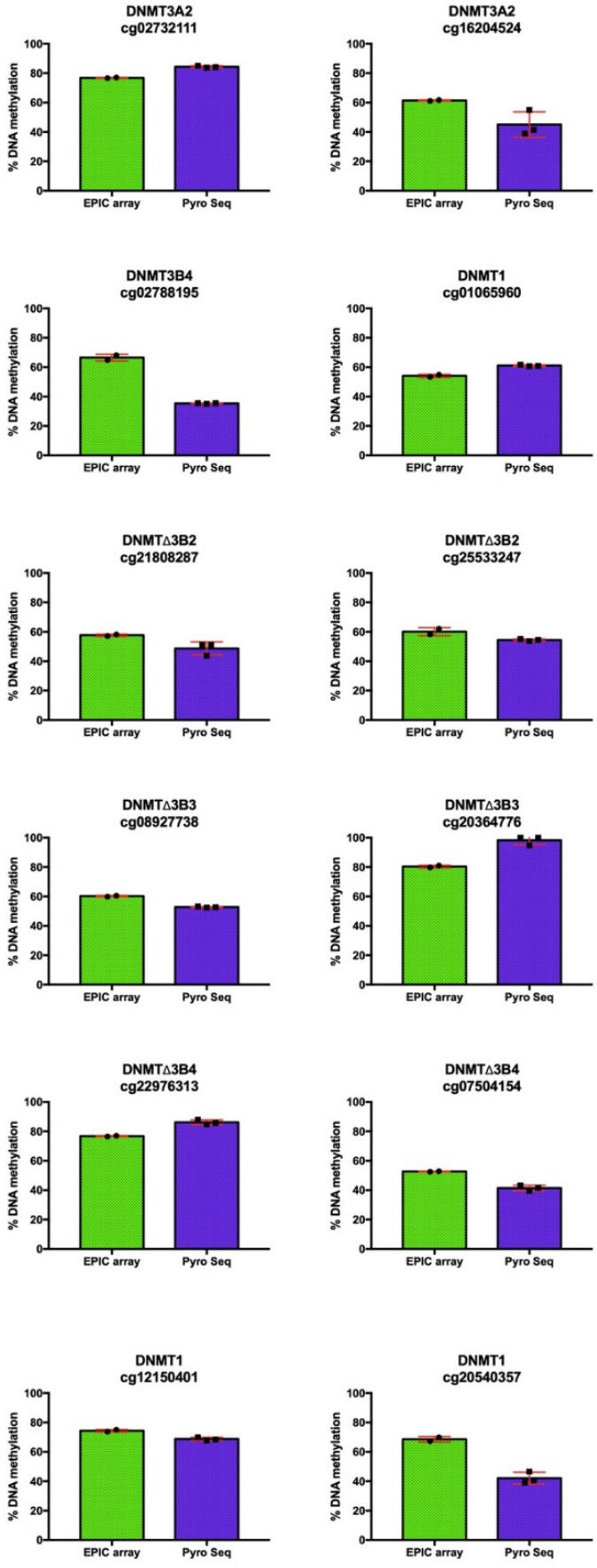


Figure 5.14 The DNA methylation levels of each target loci for each DNMT between EPIC array and pyrosequencing methods.

DNA methylation at cg02372111 in cells overexpressing DNMT3A2 was increased significantly (p-value ≤ 0.05) after treatment with EGCG at 200 μM (from 83.8% to 86.1%), curcumin at 10 and 25 μM (from 84.2% to 85.2% and 87.5%, respectively), and theaflavin at 80.5 and 161 μM (from 83.8% to 86.6% and 87.1%, respectively) (Figure 5.15A). However, methylation at this locus was increased significantly (from 84.2% to 85.2%) after treatment with 10 μM curcumin (p-value ≤ 0.05). Methylation at cg16204524 increased significantly after treatment with EGCG at 200 μM (from 46.6% to 51.3%) (Figure 5.15B). In contrast, methylation at this locus in the DNMT3A2 overexpressing cell line decreased substantially, and significantly, after treatment with 25 μM curcumin (from 56.4% to 32.8%), 100 and 200 μM vitamin C (from 46.6% to 38.5% and 39.4%, respectively), and 161 μM theaflavin (from 46.6% to 41.7%) (p-value ≤ 0.05) (Figure 5.15B).

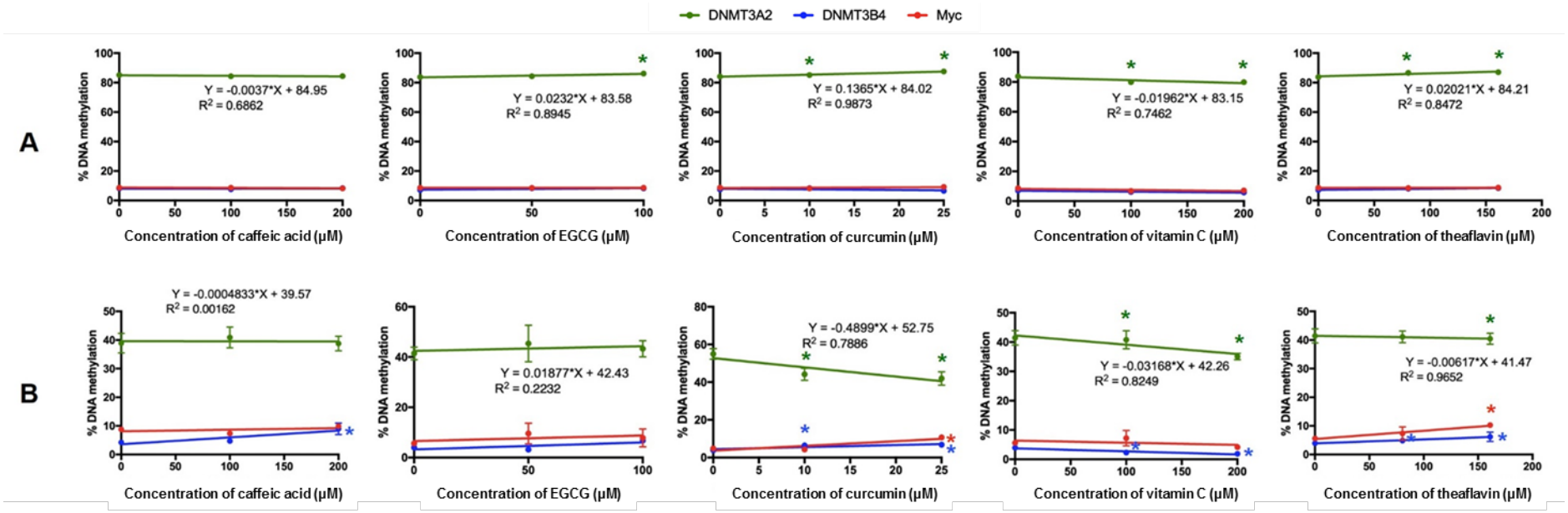


Figure 5.15 DNA methylation changes in specific CpG sites of cells overexpressing DNMT3A2 after treatment with selected food constituents for 48 hours. The effects of 100 and 200 μM CA, 50 and 100 μM EGCG, 10 and 25 μM curcumin, 100 and 200 μM vitamin C, 80.5 and 161 μM theaflavin, on A) cg02732111, B) cg16204524 were measured by pyrosequencing. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with untreated condition.

Methylation at cg02788195 in cells overexpressing DNMT3B4 was decreased significantly from 35.5% to 33.3% after treatment with 100 μM CA (p-value ≤ 0.05) (Figure 5.16) with a similar fall in methylation at this CpG (from 35.2% to 32.8%) after treatment with 100 μM EGCG (p-value ≤ 0.05) (Figure 5.16). However, methylation at cg02788195 in cells overexpressing DNMT3A2 was increased from 12.4% to 15.7% (p-value ≤ 0.05) after treatment with 25 μM curcumin (Figure 5.16).

In cells overexpressing DNMTΔ3B2, treatment with 100 μM EGCG and 80.5 μM theaflavin, increased methylation at cg21808287 from 47.67% to 52.51% and 47.67% to 51.44%, respectively (p-value ≤ 0.05) (Figure 5.17A). After treatment with 200 μM CA, 10 μM curcumin, and 200 μM vitamin C, methylation at this locus decreased significantly (p-value ≤ 0.05) from 51.2% to 46.6%, 51.3% to 49.3%, and 55.1% to 50.2% respectively (Figure 5.17A). Also, cells overexpressing DNMTΔ3B2 showed decrease of methylation at cg25533247 after treatment with 100 μM and 200 μM (from 54.5% to 52.6% and 48.5%, respectively) (p-value ≤ 0.05) (Figure 5.17B), while methylation at this locus was increased from 47.7% to 52.5% after treatment with 200 μM EGCG (Figure 5.17B). Moreover, I cannot detect the DNA methylation of both target sites of DNMTΔ3B2 from DNA of cells overexpressing DNMT3L treated with curcumin at 25 μM and with theaflavin at 161 μM for 48 hours (data not available).

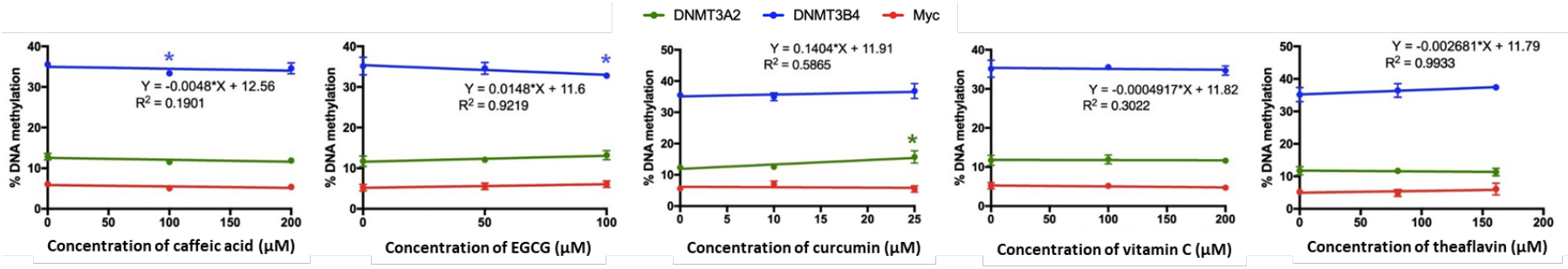


Figure 5.16 DNA methylation changes in specific CpG site of cells overexpressing DNMT3B4 after treatment with selected food constituents for 48 hours. The effects of 100 and 200 µM CA, 50 and 100 µM EGCG, 10 and 25 µM curcumin, 100 and 200 µM vitamin C, 80.5 and 161 µM, on cg02788195 were measured by pyrosequencing. Error bars represent standard deviation from triplicates and * represents *p*-value ≤ 0.05 compared with untreated condition.

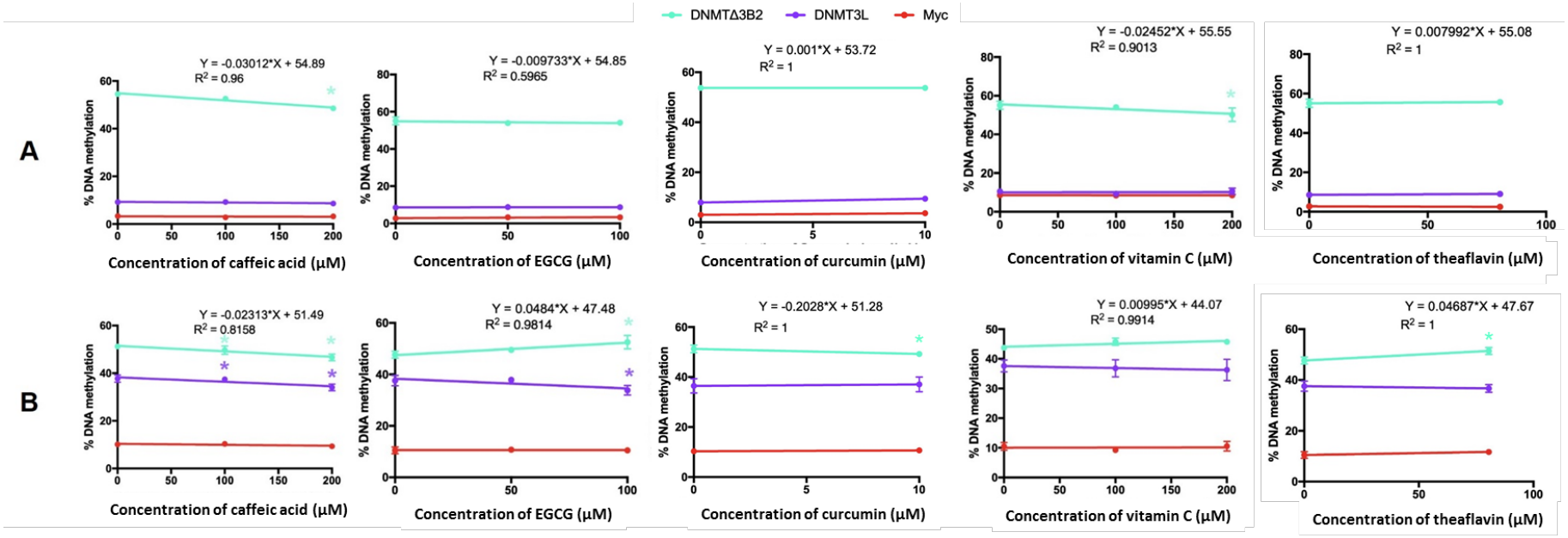


Figure 5.17 DNA methylation changes in specific CpG sites of cells overexpressing DNMTΔ3B2 after treatment with selected food constituents for 48 hours. The effects of 100 and 200 μM CA, 50 and 100 μM EGCG, 10 and 25 μM curcumin, 100 and 200 μM vitamin C, 80.5 and 161 μM theaflavin, on A) cg21808287 B) cg25533247 were measured by pyrosequencing. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with untreated condition.

Methylation at cg08927738 in cells overexpressing DNMTΔ3B3 were increased significantly after treatment with 25 μM curcumin (from 52.7% to 63.2%) and 80.5 μM theaflavin (from 53.2% to 57.1%) (p-value ≤ 0.05) (Figure 5.18A). In contrast, vitamin C exhibited an inhibitory effect on cg20364776 in cells overexpressing DNMTΔ3B3 (from 100% to 85.85% and 85.91% at 100 and 200 μM, respectively) (p-value ≤ 0.05) (Figure 5.18B). However, I cannot detect the DNA methylation of both target sites of DNMTΔ3B3 from DNA of cells overexpressing DNMT3L treated with curcumin at 25 μM and with theaflavin at 161 μM for 48 hours, and of cells overexpressing DNMTΔ3B3 treated with theaflavin at 161 μM for 48 hours (data not available).

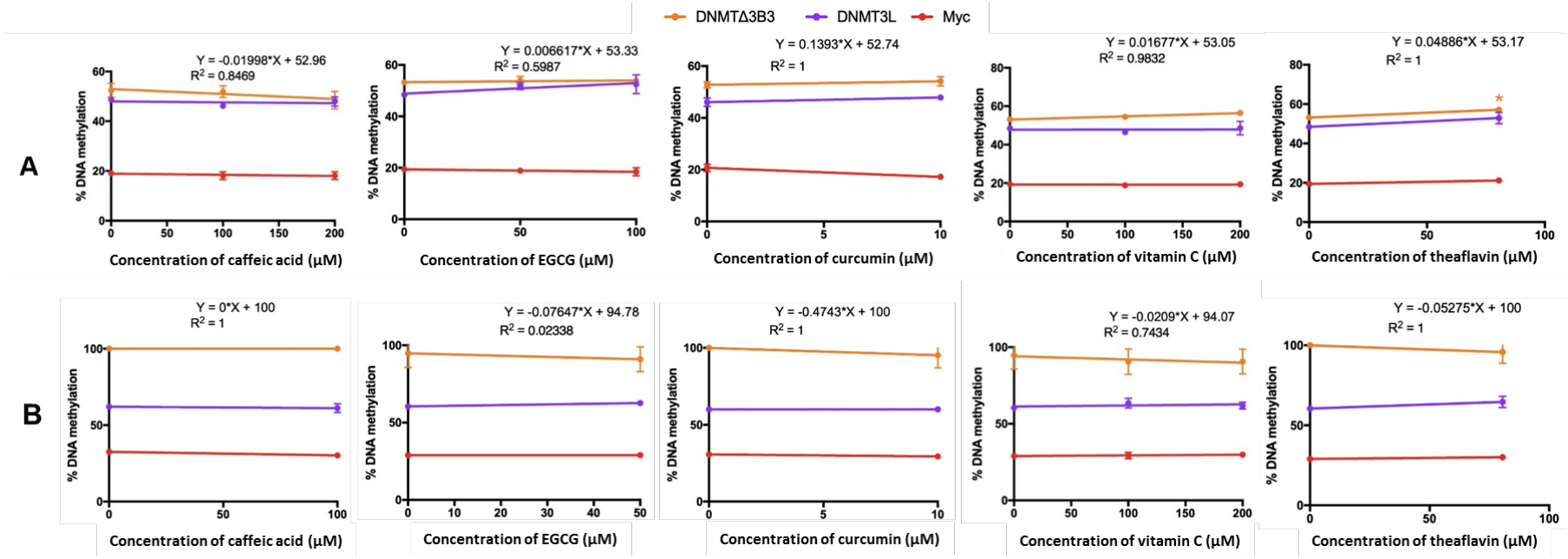


Figure 5.18 DNA methylation changes in specific CpG sites of cells overexpressing DNMTΔ3B3 after treatment with selected food constituents for 48 hours. The effects of 100 and 200 µM CA, 50 and 100 µM EGCG, 10 and 25 µM curcumin, 100 and 200 µM vitamin C, 80.5 and 161 µM theaflavin, on A) cg08927738 B) cg20364776 were measured by pyrosequencing. Error bars represent standard deviation from triplicates and * represents *p*-value ≤ 0.05 compared with untreated condition.

Methylation at cg22976313 in cells overexpressing DNMTΔ3B4 was decreased significantly after treatment with 200 μM CA (from 84.9% to 74.9%) (p-value ≤ 0.05) (Figure 5.19A). Also, CA showed the inhibitory effect on another loci (cg07504154) in cells overexpressing DNMTΔ3B4 (from 43.29% to 26.62% and 36.70% at 100 and 200 μM, respectively) (p-value ≤ 0.05) (Figure 5.19B). Additionally, methylation at cg07504154 was decreased significantly after treatment with 100 and 200 μM vitamin C (from 39.5% to 23.0% and 18.9%, respectively) (p-value ≤ 0.05) (Figure 5.19B). In contrast, methylation at cg07504154 in the DNMTΔ3B4 overexpressing cell line increased substantially, and significantly, after treatment with 100 and 200 μM EGCG (from 39.5% to 42.2% and 42.8%, respectively), 10 and 25 μM curcumin (from 41.6% to 46.3% and 45.0%, respectively), and 80.5 μM theaflavin (from 39.5% to 43.7%) (p-value ≤ 0.05) (Figure 5.19B). Furthermore, 200 μM curcumin increased methylation at cg22976313 from 85.8% to 89.3% (p-value ≤ 0.05) (Figure 5.19A). However, I cannot detect the DNA methylation of both target sites of DNMTΔ3B4 from DNA of cells overexpressing DNMT3L treated with curcumin at 25 μM and with theaflavin at 161 μM for 48 hours, and of cells overexpressing DNMTΔ3B4 treated with theaflavin at 161 μM for 48 hours (data not available).

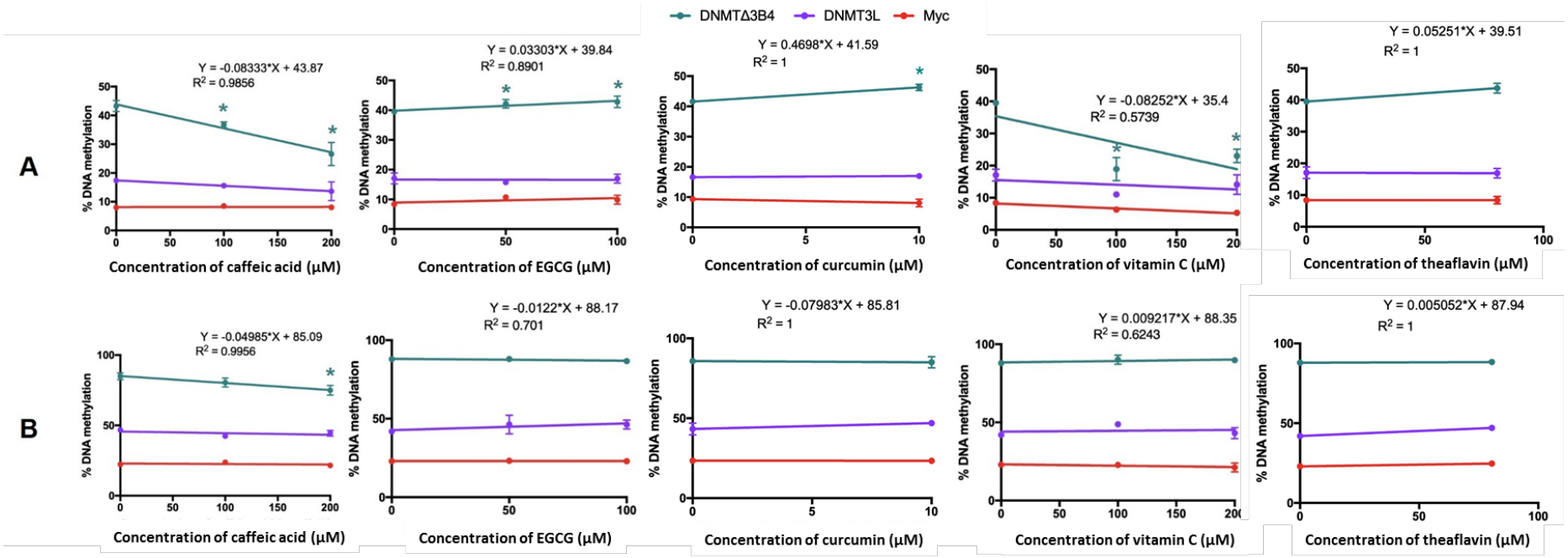


Figure 5.19 DNA methylation changes in specific CpG sites of cells overexpressing DNMTΔ3B4 after treatment with selected food constituents for 48 hours. The effects of 100 and 200 μM CA, 50 and 100 μM EGCG, 10 and 25 μM curcumin, 100 and 200 μM vitamin C, 80.5 and 161 μM theaflavin, on A) cg22976313 B) cg07504154 were measured by pyrosequencing. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with untreated condition.

Methylation at cg01065960 in cells overexpressing DNMT1 was decreased significantly after treatment with 100 and 200 μM vitamin C (from 60.9% to 52.9% and 55.2%, respectively) (p-value ≤ 0.05) (Figure 5.20). However, I cannot detect the DNA methylation of target site of DNMT1 from DNA of cells overexpressing treated with theaflavin at 161 μM for 48 hours (data not available).

Methylation at cg12150401 in cells overexpressing DNMT3L was increased significantly after treatment with 200 μM EGCG (from 67.8% to 74.2%) and 80.5 μM theaflavin (from 67.8% to 71.7%) (p-value ≤ 0.05) (Figure 5.21A). Moreover, methylation at cg20540357 was increased in cells overexpressing DNMT3L after treatment with 100 μM CA (from 40.7% to 46.5%) and 80.5% μM theaflavin (from 40.6% to 46.5%) (p-value ≤ 0.05) (Figure 5.21B). However, I cannot detect the DNA methylation of both target sites of DNMT3L from DNA of cells overexpressing DNMT3L treated with curcumin at 25 μM and with theaflavin at 161 μM for 48 hours, and of cells overexpressing DNMTΔ3B3 treated with theaflavin at 161 μM for 48 hours (data not available).

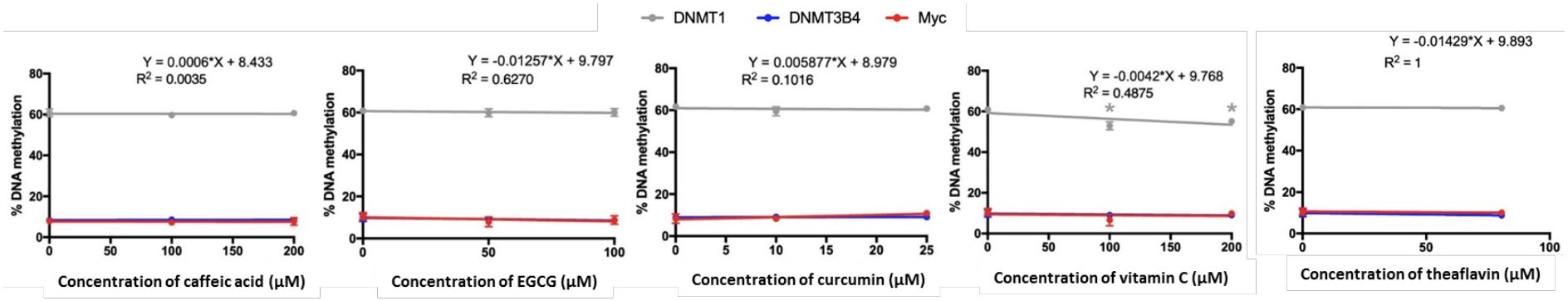


Figure 5.20 DNA methylation changes in specific CpG site (cg01065960) of cells overexpressing DNMT1 after treatment with selected food constituents for 48 hours. The effects of 100 and 200 µM CA, 50 and 100 µM EGCG, 10 and 25 µM curcumin, 100 and 200 µM vitamin C, 80.5 and 161 µM theaflavin, on cg01065960 were measured by pyrosequencing. Error bars represent standard deviation from triplicates and * represents *p*-value ≤ 0.05 compared with untreated condition.

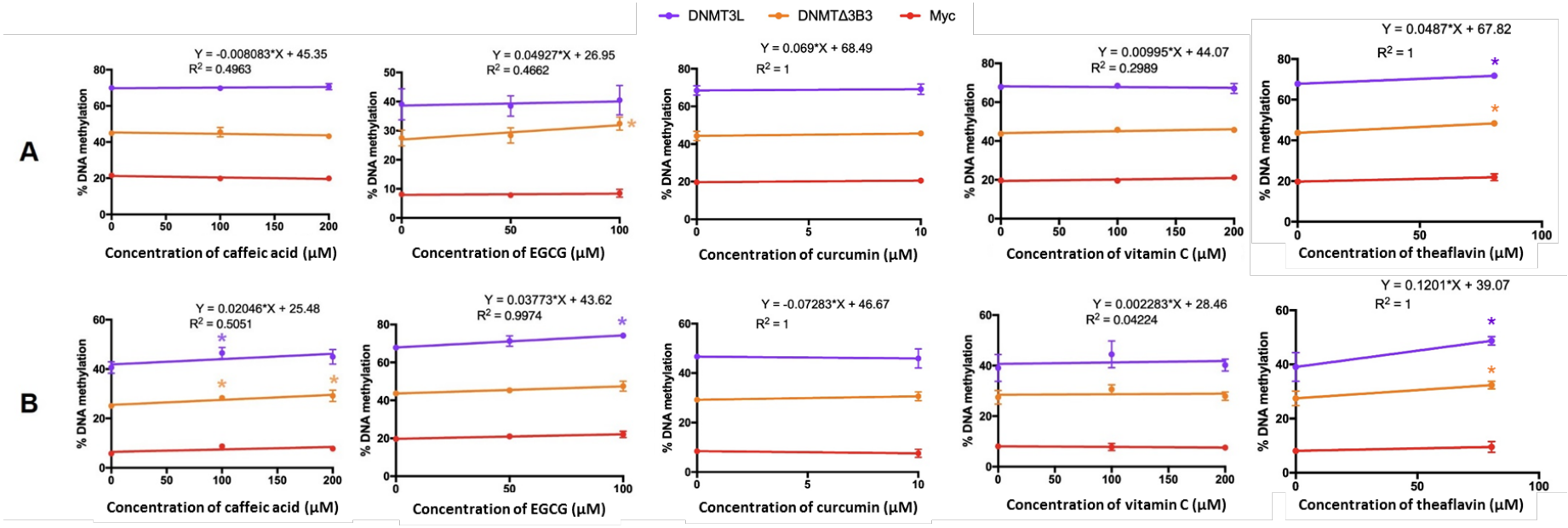


Figure 5.21 DNA methylation changes in specific CpG sites of cells overexpressing DNMT3L after treatment with selected food constituents for 48 hours. The effects of 100 and 200 µM CA, 50 and 100 µM EGCG, 10 and 25 µM curcumin, 100 and 200 µM vitamin C, 80.5 and 161 µM theaflavin, on A) cg12150401 B) cg20540357 were measured by pyrosequencing. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with untreated condition.

5.6.5 Effect of CA on DNMT enzymatic activity

In the previous section, I reported that DNMTΔ3B4 was targeted specifically by CA and that this led to reduced methylation levels of its target CpG sites (see section 5.6.4, Figure 5.19). To investigate the possible mechanism for this effect, the enzymatic activity of the DNMT was quantified after treatment with a range of concentrations of CA. Cells overexpressing DNMTΔ3B4 were selected to quantify the inhibitory effect of CA, as these cells showed low methylation in both methylation of global and site-specific loci after treatment with CA. DNMTΔ3B4 proteins were extracted and purified from cells overexpressing DNMTΔ3B4. The method used (described in the Methods chapter, section 2.4) ensured that only DNMT proteins tagged with Myc *i.e.* those derived from the over-expressed *DNMTΔ3B4* isoform were captured by Myc-antibody. These proteins were eluted and used to perform the DNMT activity assay. CA exhibited the inhibitory effect on DNMTΔ3B4 activity at all concentrations tested from 25 to 300 µM (p-value ≤ 0.05) (Figure 5.22A). The degree of inhibition decreased linearly with increasing CA concentrations (Figure 5.22B).

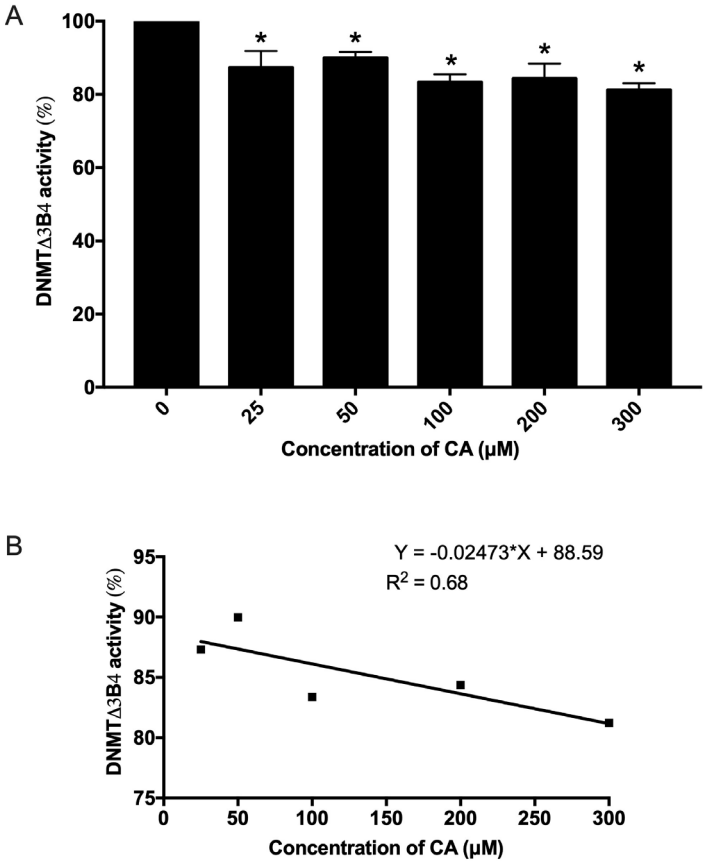


Figure 5.22 DNMT activity (%) of DNMTΔ3B4 after treatment with CA. DNMTΔ3B4 was treated with CA at 25, 50, 100, 200, and 300 µM; A) the percentage of DNMTΔ3B4 activity after treatment with CA B) linear regression of concentration of CA and DNMTΔ3B4 activity. Error bars represent standard deviation from triplicates and * represents p-value ≤ 0.05 compared with untreated condition.

5.7 Discussion

5.7.1 Main findings

The primary objectives of this study were to quantify the effects of food-derived polyphenols (theaflavin, EGCG, CA, curcumin) and of vitamin C on cell lines over-expressing individual DNMT isoforms and to determine whether these effects of these food constituents were specific to particular DNMT isoforms. Target CpG sites of each DNMT isoform were selected from the EPIC array data (described in Chapter 4) to test the effect of each polyphenol and of vitamin C, separately. Table 5.1 shows the summary of finding all significant methylation changes after treatment with selected dietary constituents. The results show that curcumin at 100 and 200 μM and theaflavin at 80.5 decreased global DNA methylation in cells that over-expressed DNMTΔ3B4 and DNMT1, respectively. For screening the effect of food constituents on cells overexpressing DNMTs, CA showed the inhibitory effect on cells overexpressing DNMTΔ3B2 and DNMTΔ3B4 by decreasing the DNA methylation of their target CpGs. Vitamin C also decreased the DNA methylation of target CpGs in cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, and DNMT1. EGCG, curcumin, and theaflavin showed both enhancement and reduction of the DNA methylation levels of target CpGs across the DNMT isoforms. After investigation of each DNMT isoform to dietary constituents, DNA methylation changes of their target CpG sites in DNMT3A2 and DNMT1 were unique to vitamin C treatment, and DNMTΔ3B2 and DNMTΔ3B4 were to CA. The confirmation of inhibitory effect on DNMT activity revealed that CA decreased DNMTΔ3B4 activity.

Table 5.1 Summary the significant decrease or increase of methylation after treatment with dietary constituents.

Over-expressed DNMT cells	CA (μM)		EGCG (μM)		Curcumin (μM)		Vitamin C (μM)		Theaflavin (μM)	
	100	200	50	100	10	25	100	200	80.5	161
Global DNA methylation										
DNMTΔ3B4	↓	↓								
DNMT1				↓					↓	

Table 5.1 (Continue) Summary the significant decrease or increase of methylation after treatment with dietary constituents.

Over-expressed	CA (μM)		EGCG (μM)		Curcumin (μM)		Vitamin C (μM)		Theaflavin (μM)	
DNMT cells	100	200	50	100	10	25	100	200	80.5	161
Screening of the effect of selected food constituents on DNA methylation changes of target CpGs for across DNMT isoforms										
cg25843713										
DNMT3A2						↑	↓	↓		↑
DNMT3B4					↑	↑	↓	↓	↓	
DNMTΔ3B2	↓	↓					↓	↓		↑
DNMTΔ3B3							↓	↓		
DNMTΔ3B4	↓	↓	↑				↓	↓		
cg04458645										
DNMT3A2				↓						
DNMT3B4				↓						
DNMTΔ3B2			↓			↑	↓			
DNMTΔ3B3			↑			↓				
DNMTΔ3B4	↓	↓		↑						
The effect of selected food constituents on DNA methylation changes of target CpGs for each individual DNMT isoform										
DNMT3A2										
cg02732111				↑	↑	↑	↓	↓	↑	↑
cg16204524					↓	↓	↓	↓		↓
DNMT3B4										
cg02788195	↓			↓						
DNMTΔ3B2										
cg21808287		↓						↓		
cg25533247	↓	↓		↑	↓				↑	
DNMTΔ3B3										
cg08927738						↑			↑	
DNMTΔ3B4										
cg22976313	↓	↓	↑	↑	↑	↑	↓	↓	↑	
cg07504154		↓				↑				
DNMT1										
cg01065960							↓	↓		
DNMT3L										
cg12150401									↑	
cg20540357	↑			↑					↑	

↑: hypermethylation; ↓: hypomethylation.

5.7.2 The effect of selected food constituents on DNMT isoforms through modulating DNA methylation in global DNA methylation and site-specific target CpGs

Epigenetic studies have tried to reveal the link between diet and DNA methylation. The results from computer models (119) and *in vitro* studies (165) show that nutrition interferes with DNA methylation by changing the cofactors and substrates in one-carbon metabolism and food constituents such as polyphenols inhibits DNMT activity (see Introductory section 1.5.2 and 1.6.2). Hypermethylation of DNA was one of the epigenetic mechanisms for silencing the gene expression, including tumour suppressor genes (321, 322). The inhibition of DNMTs is a possible therapeutic mechanism that contributes to protection against cancer. Although DNMT subfamilies have been identified, most studies focus on the main DNMT isoforms (DNMT3L, DNMT1, DNMT3A, and DNMT3B). There is a limitation of the study about the sensitivity and specificity of DNMT inhibitors to DNMT subfamilies as there is no specific antibody to each DNMT subfamily due to the similarity of the structure.

In this study, five dietary constituents: CA, curcumin, EGCG, vitamin C, and theaflavin, were selected as these nutrients showed a potential inhibitory effect on DNMT activity (see Introduction section 1.5.2, page 17). To my knowledge, apart from the major DNMT isoforms (DNMT1, DNMT3A, and DNMT3B), there is no published study that has investigated the effects of these particular food constituents on methylation of target CpGs by individual DNMT isoforms. This is the first one to investigate the effect of polyphenols and vitamin C on DNMT sub-isoforms. Cells overexpressing DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, and DNMT3L were selected in this study based on the number of significant hypermethylation probes, which were more than 300 probes. Also DNMT1 was selected as it is a full-length isoform and several studies showed that DNMT1 was inhibited by polyphenols (166, 223).

To investigate the effect of selected compounds on DNMT sub-isoforms, the DNA methylation levels were measured at the target CpGs of DNMTs before and after treatment with polyphenols and vitamin C. Firstly, the specificity of all the target CpG sites (across DNMT isoforms across DNMTs and specific DNMT targets) selected from EPIC data analysis, was confirmed by pyrosequencing as the baseline of DNA methylation level of target CpG sites was higher than the control cell. The global methylation was also investigated using LINE-1 assay to observe levels as the baseline and after treatment with dietary constituents.

CA

In this study, DNA methylation of LINE-1 was reduced in the over-expressed DNMTΔ3B4 cell after treatment with CA at 100 and 200 μM. CA also showed specific inhibitory effect on DNMTΔ3B2 and DNMTΔ3B4 at specific CpG target sites and DNMTΔ3B4 activity by dose-dependence. CA was found as DNMTi through inhibiting the activity of M.SssI and DNMT1 (173). Dibutyltin (IV) complex of CA decrease global DNA methylation using imprint DNA methylation kit (323). However, no previous study has explored inhibition of DNMTΔ3B4 by CA. It was interesting that the catalytic sites of DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4 are identical (see introduction section, Figure 1.7), but only DNMTΔ3B2 and DNMTΔ3B4 can be inhibited by CA. This finding indicates that lacking 200 amino acids at the *N*-terminal and exon 6 of DNMT3B in DNMTΔ3B2 and DNMTΔ3B4 was important for interaction with CA.

EGCG

Molecular modelling studies demonstrate that EGCG is well accommodated in a hydrophilic pocket of DNMT1 by at least four hydrogen bonds (167). In this study, EGCG at 100 μM showed the inhibitory effect on global DNA methylation in the over-expressed DNMT1 cell and this concentration decreased the methylation of the target CpG of DNMT3B4. Concomitantly, this dose of EGCG showed the potent anti-proliferative effects on HCT-116 cells (324). However, Fang MZ. *et al.* found that 20 and 50 μM EGCG are the effective concentration to re-activate of methylation-silenced genes (Retinoic acid receptor beta and p16INK4a genes) (167).

EGCG has demonstrated beneficial effects in cancer prevention through the inhibition of DNMT (167). In addition, expression of DNMT1, DNMT3a, and DNMT3b at both RNA and protein levels in human skin cancer cells was inhibited after EGCG treatment (168). Further, EGCG inhibit DNMT activity in human cervical HeLa cells (325). Also, EGCG treatment did not affect the mRNA expression of *DNMT3A, DNMT3B*, and *DNMT1* (167, 326). There is no study showing an enhancing effect of EGCG on locus-specific CpGs. However, the results from this study showed the enhancing effect of EGCG on the target DMPs of DNMTΔ3B4 and DNMT3L. This effect is hypothesised that the endogenous expression of DNMT3B4, which was inhibited specifically by EGCG, might be inhibited in cells overexpressing DNMTΔ3B4 and DNMT3L after treatment. This effect leads to hypermethylation at the target DMPs of DNMTΔ3B4 and DNMT3L by activating these isoforms.

Vitamin C

Vitamin C is a co-factor or the catalytic activity of Tet enzymes that are involved in demethylation of DNA (327). Vitamin C does not change *Tet* expression (328) but it enhances the activity of Fe (II)-2-oxoglutarate dioxygenases, including TETs, as a critical step in DNA demethylation (329). Vitamin C also was demonstrated to prevent hypermethylation and prevent loss of *DlK1-Dio3* imprinting gene in iPS cell (330). Vitamin C accumulated in haematopoietic stem cells promotes Tet function and suppressing leukaemogenesis (331).

According to the formulations of DMEM and serum used in this study, the culture media does not contain a detectable amount of vitamin C (332). This study exhibits that vitamin C decreased DNA methylation of some CpG target sites of DNMT3A2, DNMTΔ3B2, DNMTΔ3B4, and DNMT1. However, other over-expressed DNMT cells (DNMT3B4, DNMTΔ3B3, and DNMT3L) did not change the levels of DNA methylation at the specific CpG sites after treatment with vitamin C. For the concentration of vitamin C in this study, the highest concentration was 200 μM and this concentration does not harm cell viability. Also, this concentration can inhibit the methylation of both target sites of DNMT3A2. Concomitantly, the high concentration (8 mM) of vitamin C inhibited the DNMT activity in nuclear extracts of melanoma cells (176). However, it is important to note that the concentration of vitamin C in plasma for supplementation, following oral administration of high doses of vitamin C (>500 mg/day), does not exceed 150 μM, because of the intestinal absorption (333).

Curcumin

Curcumin has beneficial activities such as antioxidant and anti-cancer properties (334). Liu Z. *et al.* (174) showed the molecular docking analysis between catalytic pocket of DNMT1 and curcumin. In this study, curcumin increased DNA methylation of the target CpGs of DNMT3A2 and also decreased DNA methylation at cg16204524, which was another target site of DNMT3A2, suggesting that modulating activity on DNA methylation by curcumin is independent of its direct effect on DNMTs. Concomitantly with previous publication, DNMT1 expression and activity analyses following treatment with curcumin failed to observe any significantly changes (335). Also, curcumin can restore the DNMT activity and *DNMT1* expression in the retinal pigment epithelial cell of diabetes mice (336). Its effects also downregulated the oxidative stress-induced expression of miR-302, which is an inhibitor of DNMT1 (337). The antioxidative effect of curcumin might therefore be a potential factor in modulating DNMT functions, a hypothesis that needs to be explored in future studies.

Theaflavin

Theaflavin and its galloyl esters are the main red pigment in black tea. From *in vitro* study, the theaflavin has been found to reduce the viability of ovarian cancer cells and it appears to mediate apoptosis pathway via both intrinsic and extrinsic pathways. In this study, theaflavin increased DNA methylation of target CpG sites of DNMTΔ3B2 and DNMT3L. It is possible that theaflavin might modulate the SAM and SAH intracellular ratio, thus this ratio should be further measured after treatment with theaflavin in cells overexpressing DNMTΔ3B2 and DNMT3L. From *in vitro* study, theaflavin found to be an inhibitor of Dnmt3a (175). In this study, theaflavin showed both hypermethylation and hypomethylation on the CpG targets of DNMT3A2, suggesting that theaflavin might inhibit other DNMTs in cells overexpressing DNMT3A2 leading to hypermethylation at DNMT3A2's target CpG.

5.8 Conclusion

Here, for the first time, the target CpG sites of each DNMT were used to explore the sensitivity and specificity of polyphenols and vitamin C to DNMTs. This study also shows that CA inhibits DNMTΔ3B2 and DNMTΔ3B4 through decreased DNA methylation of the target CpGs of DNMTΔ3B2 and DNMTΔ3B4, and decreased DNMTΔ3B4 enzyme activity. Vitamin C also decreased DNA methylation levels of the target CpGs of DNMT3A2, DNMTΔ3B2, DNMTΔ3B4, and DNMT1. However, EGCG, curcumin, and theaflavin showed both increased and decreased DNA methylation, suggesting that these compounds act independently or inhibit endogenous DNMT targets resulting in hypermethylation at some target CpGs of DNMTΔ3B4, DNMT3L, or DNMT3A2.

Chapter 6: General discussion

The objectives of my PhD project were

1) to generate the cell lines that over-expressed 13 individual DNMT isoforms (DNMT3A1, DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT3L, and DNMT1) and to identify the *de novo* DNA methylation target sites of each specific DNMT isoform,

2) to undertake genome-wide DNA methylation analysis using the Illumina Infinium Methylation EPIC BeadChip,

3) to assess CpG specific DNMTs in relation to biological mechanism by pathway analysis according to enrichment statistics of the difference of gene and

4) to investigate interactions between selected food constituents, *i.e.* theaflavin, EGCG, CA, curcumin, and vitamin C, and specific DNMT isoforms in modulating DNA methylation patterns.

6.1 Generating the over-expressed DNMT cells

There are three enzymatically active DNMTs *i.e.* DNMT3A, DNMT3B, and DNMT1, and an enzymatically inactive regulatory factor, DNMT3L (338). In addition, multiple subfamilies of DNMT3A and DNMT3B have also been identified (106). In this study, 13 isoforms were investigated. Of these isoforms, nine contained the catalytically active region (see Introduction section 1.4, page 9), whilst the other four *i.e.* DNMT3B3, DNMT3B4, DNMT3B5, and DNMT3L did not. However, DNMT3B3 and DNMT3B4 are over-expressed in many tumour cell lines where they bind to, and regulate the activity of DNMT3A or DNMT3B (85). In addition, DNMT3B4 and DNMT3B5 are over-expressed in gastric carcinoma (278). DNMT1 acts primarily as a maintenance methyltransferase and it can maintain DNA methylation even when uncoupled from the DNA replication machinery (339). All DNMTs facilitate a similar catalytic mechanism in which a covalent reaction intermediate is formed between the substrate and its catalytic pocket. The identity matrix of the catalytic sites of the previously mentioned 13 DNMT isoforms plus DNMT2 was analysed by Fernanda I. *et al.* (119) (see Introduction section 1.4.1, page 11). Nonetheless, DNMT isoforms have specific and overlapping target CpG sites (shown in Choi's study (87)). This

study used an earlier version of the Illumina DNA methylation assay which contained probes for only 1,505 CpG sites from 808 cancer-related genes to identify the CpG sites that were the targets of each of the DNMTs (87). To enable my investigation of *de novo* target CpG sites of each DNMT isoform, I generated cell lines that over-expressed each of 13 individual DNMT isoforms (*DNMT3A1, DNMT3A2, DNMT3B1, DNMT3B2, DNMT3B3, DNMT3B4, DNMT3B5, DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT3L,* and *DNMT1*) in a separate cell line. These cell lines were generated successfully by integrating the mRNA of each *DNMT* isoform into HEK293T genome using a lentiviral system.

Use of the lentiviral system was an effective approach for incorporating the exogenous *DNMT*s into both the HEK293T genome and the MEG-01 genome. However, the amount of positive GFP cells was higher in HEK293T cells compared with MEG-01 cells because of their reliable growth and propensity for incorporating exogenous DNA. Another factor is that the doubling time of HEK293T (24 hours) (340) is shorter than MEG-01 cells (36-48 hours) (341). The reduced cell population after a governed time point in MEG-01 populations is likely to impact the number of transgene-expressing cells. mRNA expression of exogenous *DNMTs* in HEK239T cells was detected by qPCR using a pair of primers that bound to the Myc-tag sequence and to the relevant *DNMT* isoform.

Exogenous *DNMT* expression was clearly observed in GFP positive cells compared with Myc control cell. However, the mRNA expression of *DNMT3A2* in Myc control cell was higher than in cells overexpressing DNMT3A2, possibly due to high expression of endogenous *DNMT3A2* in HEK293T cells. The expression of exogenous *DNMTs* was not correlated with protein expression levels, especially *DNMT3ΔB3* (see Discussion section 3.7.3 in the Chapter 3, page 79). This observation was consistent with a previous publication (87). Additionally, Duymich, C.E. *et al.* (277) reported that the different mRNA expression of *DNMT3Bs* was seen for each of *DNMT3B*, but the protein levels showed a large variation, indicating the different stabilities of the DNMT variants. Also, Fredrik E. *et al.* found the correlation between total RNA expression (assessed using transcriptomics) and protein expression (assessed using proteomics) of the selected 55 genes (DNMT was not included) in HEK293 cells to be 0.39 (294). However, cell lines that over-expressed DNMT showed a significant increase in global methylation (Figure 5.2, page 122), suggesting that the extent of over-expression of each DNMT isoform is sufficient to modulate DNA methylation patterns.

6.2 Identifying *de novo* DNA methylation target sites of each specific DNMT isoform

After generating single cell clone of individually overexpressing DNMTs, the target CpG sites of each DNMT isoform were explored using the Illumina EPIC array, which has probes for more than 850,000 CpG sites. As mentioned earlier, due to the structural similarities between each DNMT, a limitation is to targeting individual isoforms is the lack of specific antibodies to the sub-families of DNMT isoforms. There are a small selection of specific antibodies on the market available for DNMT1, DNMT3A, DNMT3B and DNMT3A2. Also, the specificity of each splice variant of DNMT isoforms using specific primers to discriminate their mRNA levels has not been fully investigated. Therefore, it remains unclear over which sub-family of DNMTs would be a specific marker in disease conditions as well as which DNMTs are inhibited by a given DNMTi. Over-expressed DNMT cells can fill in this gap and provide epigenetic information about their target CpGs. The results from the current cannot be compared with the previous work from Choi SH. *et.al* (87) because I used a different method for generating the cells overexpressing DNMTs and a different parameter for statistical analysis. In this study, the cut-offs for $\Delta\beta$ values were set at different values: ≥ 0.2 and ≤ -0.2, ≥ 0.3 and ≤ -0.3, ≥ 0.4 and ≤ -0.4 relative to the β values of the Myc control cell. However, the results from different cut-offs showed the same proportion of significant loci along the genomic location and these results strengthen the analysis of the sites showing ≥ 0.4 difference in β values. These stringent criteria were selected due to the limitation of only running biological duplicates and the variation this may introduce, but also to considerably reduce the incidence of false results, *i.e.* readings that cannot be reaffirmed by pyrosequencing. The number of significantly hypermethylated probes was distributed from 36-915 probes across the cells overexpressing DNMTs. The lowest numbers of hypermethylated probes (36 and 79 probes respectively) were observed in cells lines that over-expressed the catalytically inactive DNMT3B3 and DNMT3B5. Also, the numbers of hypomethylated probes in both DNMT3B3 and DNMT3B5 (142 and 123, respectively) were higher than the numbers of hypermethylated probes, suggesting that these isoforms might inhibit, or compete for binding to, the same CpG positions as other DNMT3 isoforms. Another explanation might be the possibility of DNMT3B3 and DNMT3B5 as trans-regulatory elements, leading to hypermethylation of target promoters to other DNMTs. This could result in hypomethylation of the target CpG sites of those DNMTs (illustrated conceptually in Figure 6.1).

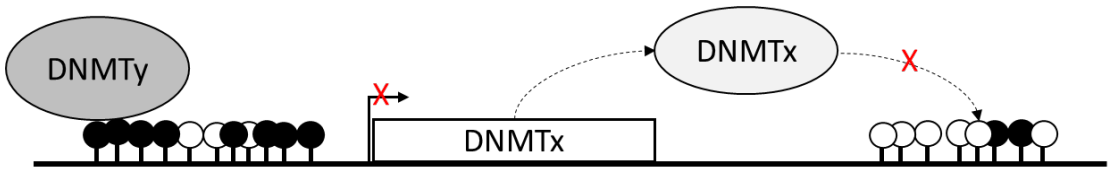


Figure 6.1 The diagrammatic overview of trans-regulation by DNMTy on the promoter of DNMTx leading to hypomethylation on the target CpG of DNMTx.

DNMT3B4, shortest isoform compared with DNMT3B1-DNMT3B5, showed 320 hypermethylated probes. This corresponded to the mRNA levels of *DNMT3B4*, which showed the highest expression among DNMT3B cells. However, in a previous study, the over-expression of *DNMT3B4* using transient transfection in HEK293 cells led to DNA hypomethylation in pericentromeric satellite regions analysed by Southern blotting, depending on the expression level of *DNMT3B4* (285). From pathway analysis (see Table 4.3, page 109), the target CpGs of DNMT3B4 involved in cell growth and cell proliferation through DP-glucose Biosynthesis, Glucose and Glucose-1-phosphate Degradation, and PFKFB4 Signalling Pathway. This observation is supported by finding Shao G. *et al.* who observed that DNMT3B4 inhibited cell proliferation in HEK293A cell line through inducing p21, a cyclin-dependent kinase inhibitor (342).

Moreover, in my study, in cell lines that over-expressed DNMT3L and DNMT1, they showed identified 394 and 197 hypermethylated probes, respectively (see Figure 4.4, page 94). Since there were approximately twice as many hypermethylated probes with DNMT3L than with DNMT1, this suggests that DNMT3L might play a regulatory role in DNA methylation at certain CpG sites via interaction with endogenous DNMT3A and DNMT3B (116). This finding supports the activity of DNMT3L to restore DNA methylation in DNMT3B-deficient cells (3BKO and DKO8 derivatives of the HCT116 colon cancer cell line) (277).

Furthermore, in this study, the cell line that over-expressed DNMT3A2 cells showed the highest number of hypermethylated probes (915 probes). With this finding, it is plausible to postulate that high endogenous DNMT3A2 might cause aberrant DNA methylation in over-expressed DNMT3A2 cells. Also, DNMT3A2 shares binding targets with DNMT3A1 and DNMT3B1 (303) and this sharing caused an increase of hypermethylated loci. Furthermore, the majority of $\Delta\beta$ levels in the over-expressed DNMT cell lines showed hypermethylation below 0.7. These results are consistent with those observed in a previous publication (87), suggesting a limitation to increasing DNA methylation of the human genome by DNMT.

6.3 Determining the sensitivity and specificity of theaflavin, EGCG, CA, curcumin, and vitamin C in interactions with specific DNMT isoforms

Natural products known as DNMTi or demethylating agents have been extensively investigated (119, 343) because there is a great need for the development of nontoxic and effective inhibitors of DNMTs. Natural DNMTi products have distinct chemical scaffolds, providing a starting point in determining their mechanism to inhibit DNMT activity. Although the specificity of these DNMTi products to DNMT sub-isoforms is poor understood, the structural variation amongst DNMTs supports the concept of each natural DNMTi specifically targeting a DNMT isoform(s) as opposed to all isoforms. The chemical structures of these natural products and food chemicals have been computerised in search of their inhibitory mechanism on DNMTs. However, the lack of appropriate cell models and antibodies specific to each DNMT sub-family isoform have hindered the progression in efforts towards identifying the mechanistic effect of DNMTi. Therefore, over-expressed DNMT cells, as created for this study, are important tools to boost the research in this field.

In this study, theaflavin, EGCG, CA, curcumin, and vitamin C were selected to test the sensitivity and specificity to each DNMT isoform. These dietary constituents have shown the modulating effects on DNMT activity or expression (see Introduction section, 1.5.2, page 17). Based on previous publications, these compounds affect DNMTs in a concentration-dependent manner, and with cell availability greater than 50%. However, there is no study to specify the activity of those compounds on DNMT subfamilies. There is evidence of differential expression of some DNMT variants in diseases and in distinctive cell types. For example, DNMTΔ3B4 was highly expressed in non-small cell lung cancer (104). DNMT3L is expressed in only embryonic stem cells and germ cells (111, 112). DNMT1 and DNMT3A1 are expressed more abundantly than DNMT3B and DNMT3A2 (84, 344, 345). Therefore, the study of DNMTi on each DNMT isoform will bring closer the understanding of the specific targets of each DNMTi and to which DNMT isoform they act upon. CA was found to be a specific DNMTi to DNMTΔ3B2 and DNMTΔ3B4 but not specific to DNMTΔ3B3. This implies that, despite these isoforms exhibiting a similar catalytic structure, the interaction between CA with affected DNMTs may be different. It suggests that lacking 200 amino acid *N*-terminal and exon 6 of DNMT3B in DNMTΔ3B2 and DNMTΔ3B4 isoforms are sensitive and specific to CA. Vitamin C showed inhibitory effects to DNMT3A2, DNMTΔ3B2, DNMTΔ3B4, and DNMT1. This finding supports the DNMT inhibitory activity of vitamin C, but the doses of this vitamin performed in this study were lower than in the Venturelli's study (176). Theaflavin enhanced the DNA methylation of the target CpGs of DNMT3L in over-

expressed DNMT3L, DNMTΔ3B3, and DNMTΔ3B4 cells, suggesting that this compound increases DNA methylation levels in the over-expressed inactive-DNMT cells. EGCG and curcumin showed both types of methylation marks in over-expressed DNMT cells; hypermethylation and hypomethylation. This can be explained by trans-regulation effects of DNMTs. In the illustration (Figure 6.2), EGCG or curcumin might inhibit DNMTy, leading to hypomethylation in the promoter area of DNMTx, and in turn, increased DNMTx expression and methylation of its target CpG sties.

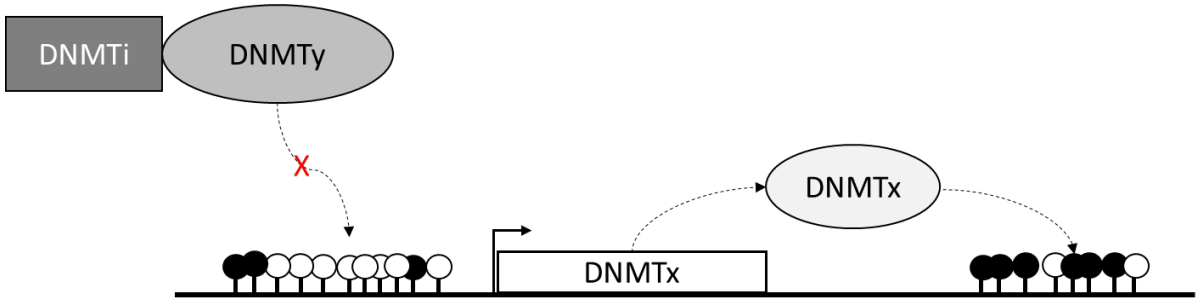


Figure 6.2 The diagrammatic overview of the effect of DNMTi (EGCG and curcumin) on trans-regulation of DNMTy on the promoter of DNMTx. This effect led to increase DNMTx and hypermethylation on the target CpG of DNMTx.

6.4 Strengths and limitations of the overall project

The main strength of this study was the ability to generate novel stable cell lines that over-expressed each of the DNMT isoforms, individually. In addition, the use of single cell clones provided a homogeneous population of cells and avoid the likely variation that would have occurred with use of heterogeneous cell lines. Also, this study covered all the major DNMT isoforms which allowed the understanding of the different structure on DNA methylation patterns. Biological duplicates were used in this study, each from a different single cell clone. This approach was taken to reduce sample bias. Additionally, cell lines that overexpress each of the DNMTs individually, were generated using a lentivirus system, which is a highly efficient method for introducing foreign mRNA into the host genome. This exogenous mRNA is inherited to subsequent cell generations. In this study, DNMTs were over-expressed in two different cell types to begin with MEG-01 and HEK293T cells. HEK293T was the selected cell type for the main study due to its more reliable expression of exogenous DNMTs, compared with MEG-01 cells (see section 3.6.4, page 72). The reference genes for making a comparative quantification were calculated and selected using eleven house-keeping genes calculated by geNorm programme. By this calculation, the stable house-keeping genes across

over-expressed DNMT cells and the Myc control were selected to obtain a more accurate expression. Furthermore, this study measured DNA methylation changes using the EPIC array, which interrogates >850,000 CpGs. Pyrosequencing was conducted to confirm the DNA methylation levels from EPIC array findings and to measure the DNA methylation of CpG sites targeted by DNMTs. Discordant duplicates from the EPIC array data were filtered out of the analysis. Additionally, the cut-offs for $\Delta\beta$ were set at $\Delta\beta \leq -0.4$ and ≥ 0.4 to capture true changes in DNA methylation, which were confirmed by pyrosequencing. All measurements of DNA methylation changes and DNMT activity were performed with technical triplicates.

The limitations of this study are that 1) the single cell clone cannot be generated for MEG-01 cells because of technical problems including separating live and dead cells, and 2) the number of samples replicates carried forward for EPIC array analysis was small (two samples represented each cell line that overexpresses each of the DNMTs individually) due to the budget of the research grant 3) the bioinformatic tool for pathway analysis was changed due to the current inability to access IPA from home during the Covid-19 pandemic. However, the vulnerabilities in the study design from point 2 above were somewhat negated by the statistical approach to select only for significant or true positive results, through employing the cut-offs ($\Delta\beta \leq -0.4$ and ≥ 0.4). Difficulties were experienced in achieving distinct bands in western blots, due to each DNMT isoform differing in expression level and in protein stability. Thus, the western blot was performed across a concentration gradient. Although DNMTΔ3B3 was not clearly detected on the blot, the function of this isoform can be measured using LINE-1 assay. Global methylation was increased by this isoform (see section 5.6.2, page 122).

6.5 Future studies

The panel of cell lines created in this project is a valuable tool to facilitate in studying of the function of individual DNMT isoforms because these cell lines make it possible to observe the changes in DNA methylation patterns for each DNMT individually. Over-expression of DNMTs is associated with a number of diseases, especially cancer (see Introduction section, 1.4.2 and Table 1.1). Some leukaemia cases, as well as, pancreatic cancer, colon and lung cancers show elevated expression of DNMT1 and DNMT3B (306, 346-348). Furthermore, high expression of DNMT1 might correlate with the fast growth rate of some tumours,

including cancer cell lines, *e.g*. H69 and U937 (306). Concurrently, Dnmt1 overexpression in ES cells showed a significant change of the genomic DNA methylation level (281).

Studies investigating all potential DNMTi and methyl donors are needed to understand the biological link and broader picture of the function of DNMTs under micronutrient exposure. Epigenetic traits are cell and tissue specific, thus the study of epigenetics in model organisms may provide extrapolative findings. This study provides some clues about the concentration and timing of treatment with nutritional compounds. These can be applied in the setup of an epigenetics intervention study in disease prevention, since the timing of epigenetic modification through exposure to a medicinal compound might be of significance. Conducting further investigation of downstream analyses will give a deeper understanding of the underlying epigenetic modifying mechanism of each nutrient. It would be beneficial to carry out gene expression profiles and protein quantitation of DNMT target genes and then, predict potential pathways which may form as potential targets for therapy. In future work, methylation profiles could be recorded prior to, and after, treatments with the nutritional compounds, where correlation may be drawn with the data obtained from this study. This type of dataset would provide detail on the epigenetic profiles in other CpG sites rather than target CpG sites of DNMTs. Furthermore, the relationship between DNA methylation and histone modification can be explored as both mechanisms influence transcription. Yang C. *et al*. found that epigenetically methylated CpG sites of Cellular Retinoic Acid Binding Protein 2 might affect histone binding and associated this with gene silencing (349).

Given the opposite DNA methylation patterns measured at the CpGs targeted by DNMTs in dietary constituents treated HEK293T cells (*e.g.* from methylated to unmethylated), it is worth examining the effects of these compounds in a broader range of cell types to enhance our understanding on the effects of nutrition by cell type. The interaction of DNMTs and CA or vitamin C should be investigated by ICM assay described by Kui L. *et al.* (350) to track DNMT's binding to genomic DNA. Moreover, the computerisation analysis should be performed to see how CA and vitamin C binds to DNMTs to confirm the inhibitory effects observed in this study. Another proposal for future examination is the research of the reversible epigenetic modifications through diet or specific nutrients, which could also help health maintenance and disease prevention. Moreover, the outcomes of this study can form the basis of finding a suitable balance of DNA methylation levels from dietary DNMTi and methyl donors. Tailoring nutriepigenomics will bring us steps closer to the ultimate goal of PN, where intake is adjusted to each individual to achieve the optimal health status (Figure 6.3).

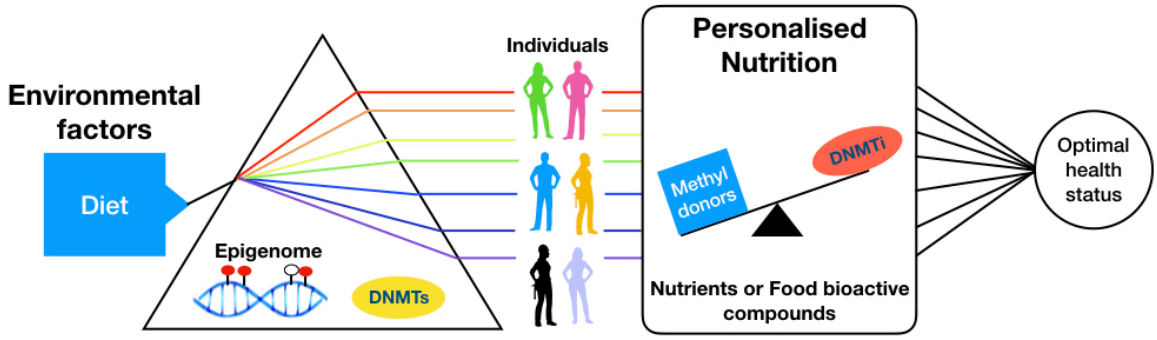


Figure 6.3 Personalised nutriepigenomics. In a typical interaction between environment and genome, diet is fed into the body (referred to as a light-dispersing from prism). The individual genome and epigenome, including the difference of either the expression or the activity of DNMT isoforms, response to diet in a multiway (referred to as individual dispersed light). The nutriepigenomic response is a function of the characteristics of the individual reaction to nutrient or food bioactive compounds. When taken through a personal nutrition regime, the balance diet between methyl donor and DNMT inhibitor diets will be ultimately be controlled by the personalised diet intervention (referred to as the last converging lens), leading to an optimal health status taking individual characteristics into consideration.

6.6 Conclusion

This study was successful to generate the novel single stable cell lines with 13 individually over-expressed DNMTs from HEK293T cells. The data derived from the EPIC data analysis supports my hypothesis that the pattern of DNA methylation in cells that overexpress individual DNMT isoforms is altered by the structure of each DNMT isoform. Also, the outcome from EPIC analysis allows to identify the *de novo* target CpG sites of each individual DNMT isoform. Results obtained from the outcome measure on DNA methylation in over-expressed DNMT cells after treatment with polyphenols and vitamin C support my hypothesis that global DNA methylation and locus-specific DMPs were modulated by these dietary constituents. Additionally, the activity of DNMTΔ3B4 and DNMT3A2 can be specifically modulated by CA and AA, respectively, in a dose-response manner. These observations further understanding of nutrition-epigenetic mechanisms, especially interactions with enzymatic activity and these understanding provides novel evidence on the effect of DNMTi diet on DNMT sub-isoforms. These findings highlight the further need to investigate the intervention study using CA and vitamin C supplementation to manipulate epigenome, especially the DNMT activity. This understanding could be applied to modulate DNA methylation profiles using food-derived bioactive compounds in PA. This application of this concept in humans remains to be demonstrated.

References

1. Handy DE, Castro R, Loscalzo J. Epigenetic modifications: basic mechanisms and role in cardiovascular disease. Circulation. 2011;123(19):2145-56.
2. Zaidi SK, Young DW, Montecino M, Lian JB, Stein JL, van Wijnen AJ, et al. Architectural epigenetics: mitotic retention of mammalian transcriptional regulatory information. Mol Cell Biol. 2010;30(20):4758-66.
3. Atlasi Y, Stunnenberg HG. The interplay of epigenetic marks during stem cell differentiation and development. Nat Rev Genet. 2017;18(11):643-58.
4. Choi JD, Lee JS. Interplay between Epigenetics and Genetics in Cancer. Genomics Inform. 2013;11(4):164-73.
5. Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. Exp Mol Med. 2017;49(4):e324.
6. Wei JW, Huang K, Yang C, Kang CS. Non-coding RNAs as regulators in epigenetics (Review). Oncol Rep. 2017;37(1):3-9.
7. Vastenhouw NL, Schier AF. Bivalent histone modifications in early embryogenesis. Curr Opin Cell Biol. 2012;24(3):374-86.
8. Araki Y, Wang Z, Zang C, Wood WH, 3rd, Schones D, Cui K, et al. Genome-wide analysis of histone methylation reveals chromatin state-based regulation of gene transcription and function of memory CD8+ T cells. Immunity. 2009;30(6):912-25.
9. Zaratiegui M, Irvine DV, Martienssen RA. Noncoding RNAs and gene silencing. Cell. 2007;128(4):763-76.
10. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell. 2009;136(4):629-41.
11. Morris KV, Mattick JS. The rise of regulatory RNA. Nat Rev Genet. 2014;15(6):423-37.
12. Costa FF. Non-coding RNAs, epigenetics and complexity. Gene. 2008;410(1):9-17.
13. Yu H. [Epigenetics: advances of non-coding RNAs regulation in mammalian cells]. Yi Chuan. 2009;31(11):1077-86.
14. Catalanotto C, Cogoni C, Zardo G. MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions. Int J Mol Sci. 2016;17(10).
15. Guil S, Esteller M. DNA methylomes, histone codes and miRNAs: tying it all together. Int J Biochem Cell Biol. 2009;41(1):87-95.
16. Bell O, Tiwari VK, Thomä NH, Schübeler D. Determinants and dynamics of genome accessibility. Nature Reviews Genetics. 2011;12(8):554-64.
17. Tate PH, Bird AP. Effects of DNA methylation on DNA-binding proteins and gene expression. Current opinion in genetics & development. 1993;3(2):226-31.
18. Mahmoud AM, Ali MM. Methyl Donor Micronutrients that Modify DNA Methylation and Cancer Outcome. Nutrients. 2019;11(3).
19. Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer. 2006;6(11):857-66.
20. Probst AV, Dunleavy E, Almouzni G. Epigenetic inheritance during the cell cycle. Nat Rev Mol Cell Biol. 2009;10(3):192-206.
21. Robertson KD. DNA methylation and human disease. Nat Rev Genet. 2005;6(8):597-610.
22. Chen ZX, Riggs AD. DNA methylation and demethylation in mammals. J Biol Chem. 2011;286(21):18347-53.
23. Li E, Zhang Y. DNA methylation in mammals. Cold Spring Harb Perspect Biol. 2014;6(5):a019133.
24. Fouse SD, Nagarajan RO, Costello JF. Genome-scale DNA methylation analysis. Epigenomics. 2010;2(1):105-17.
25. Rollins RA, Haghighi F, Edwards JR, Das R, Zhang MQ, Ju J, et al. Large-scale structure of genomic methylation patterns. Genome Res. 2006;16(2):157-63.
26. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484-92.
27. Wu H, Caffo B, Jaffee HA, Irizarry RA, Feinberg AP. Redefining CpG islands using hidden Markov models. Biostatistics. 2010;11(3):499-514.

28. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011;25(10):1010-22.
29. Edgar R, Tan PP, Portales-Casamar E, Pavlidis P. Meta-analysis of human methylomes reveals stably methylated sequences surrounding CpG islands associated with high gene expression. Epigenetics Chromatin. 2014;7(1):28.
30. Blattler A, Farnham PJ. Cross-talk between site-specific transcription factors and DNA methylation states. J Biol Chem. 2013;288(48):34287-94.
31. Gelfman S, Cohen N, Yearim A, Ast G. DNA-methylation effect on cotranscriptional splicing is dependent on GC architecture of the exon-intron structure. Genome Res. 2013;23(5):789-99.
32. Kochanek S, Renz D, Doerfler W. DNA methylation in the Alu sequences of diploid and haploid primary human cells. EMBO J. 1993;12(3):1141-51.
33. Alves G, Tatro A, Fanning T. Differential methylation of human LINE-1 retrotransposons in malignant cells. Gene. 1996;176(1-2):39-44.
34. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002;16(1):6-21.
35. Anastasiadi D, Esteve-Codina A, Piferrer F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. Epigenetics Chromatin. 2018;11(1):37.
36. Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. Science. 2007;315(5815):1141-3.
37. Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP. A human B cell methylome at 100-base pair resolution. Proc Natl Acad Sci U S A. 2009;106(3):671-8.
38. Yang M, Park JY. DNA methylation in promoter region as biomarkers in prostate cancer. Methods Mol Biol. 2012;863:67-109.
39. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. Nature. 2012;489(7414):109-13.
40. Bell RE, Golan T, Sheinboim D, Malcov H, Amar D, Salamon A, et al. Enhancer methylation dynamics contribute to cancer plasticity and patient mortality. Genome Res. 2016;26(5):601-11.
41. Aran D, Hellman A. DNA methylation of transcriptional enhancers and cancer predisposition. Cell. 2013;154(1):11-3.
42. Fleischer T, Tekpli X, Mathelier A, Wang S, Nebdal D, Dhakal HP, et al. DNA methylation at enhancers identifies distinct breast cancer lineages. Nat Commun. 2017;8(1):1379.
43. Paulsen M, Ferguson-Smith AC. DNA methylation in genomic imprinting, development, and disease. J Pathol. 2001;195(1):97-110.
44. Larsen F, Gundersen G, Lopez R, Prydz H. CpG islands as gene markers in the human genome. Genomics. 1992;13(4):1095-107.
45. Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP. Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature. 1997;389(6652):745-9.
46. Gonzalez-Zulueta M, Bender CM, Yang AS, Nguyen T, Beart RW, Van Tornout JM, et al. Methylation of the 5' CpG island of the p16/CDKN2 tumor suppressor gene in normal and transformed human tissues correlates with gene silencing. Cancer Res. 1995;55(20):4531-5.
47. Herman JG, Latif F, Weng Y, Lerman MI, Zbar B, Liu S, et al. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. Proc Natl Acad Sci U S A. 1994;91(21):9700-4.
48. Sakai T, Toguchida J, Ohtani N, Yandell DW, Rapaport JM, Dryja TP. Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. Am J Hum Genet. 1991;48(5):880-8.
49. Jones PA. The DNA methylation paradox. Trends Genet. 1999;15(1):34-7.
50. Cotton AM, Price EM, Jones MJ, Balaton BP, Kobor MS, Brown CJ. Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. Hum Mol Genet. 2015;24(6):1528-39.
51. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324(5929):930-5.
52. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. Nature. 2013;502(7472):472-9.
53. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science. 2011;333(6047):1300-3.

54. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science. 2011;333(6047):1303-7.
55. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. Nat Rev Genet. 2017;18(9):517-34.
56. Tricarico R, Bellacosa A. Active DNA Demethylation in Development, Human Disease, and Cancer. DNA Replication, Recombination, and Repair: Springer; 2016. p. 517-48.
57. Okashita N, Kumaki Y, Ebi K, Nishi M, Okamoto Y, Nakayama M, et al. PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. Development. 2014;141(2):269-80.
58. Hu X, Zhang L, Mao SQ, Li Z, Chen J, Zhang RR, et al. Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. Cell Stem Cell. 2014;14(4):512-22.
59. Wheldon LM, Abakir A, Ferjentsik Z, Dudnakova T, Strohbuecker S, Christie D, et al. Transient accumulation of 5-carboxylcytosine indicates involvement of active demethylation in lineage specification of neural stem cells. Cell Rep. 2014;7(5):1353-61.
60. Lio CW, Zhang J, Gonzalez-Avalos E, Hogan PG, Chang X, Rao A. Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. Elife. 2016;5.
61. Sproul D, Meehan RR. Genomic insights into cancer-associated aberrant CpG island hypermethylation. Brief Funct Genomics. 2013;12(3):174-90.
62. Ehrlich M. DNA methylation in cancer: too much, but also too little. Oncogene. 2002;21(35):5400-13.
63. Feinberg AP, Gehrke CW, Kuo KC, Ehrlich M. Reduced genomic 5-methylcytosine content in human colonic neoplasia. Cancer Res. 1988;48(5):1159-61.
64. Howard G, Eiges R, Gaudet F, Jaenisch R, Eden A. Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. Oncogene. 2008;27(3):404-8.
65. Schuebel KE, Chen W, Cope L, Glockner SC, Suzuki H, Yi JM, et al. Comparing the DNA hypermethylome with gene mutations in human colorectal cancer. PLoS Genet. 2007;3(9):1709-23.
66. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nat Biotechnol. 2013;31(2):142-7.
67. Sun ZH, Liu YH, Liu JD, Xu DD, Li XF, Meng XM, et al. MeCP2 Regulates PTCH1 Expression Through DNA Methylation in Rheumatoid Arthritis. Inflammation. 2017;40(5):1497-508.
68. Cai X, Lu Y, Tang C, Lin X, Ye J, Li W, et al. Effect of interleukin-6 promoter DNA methylation on the pathogenesis of systemic lupus erythematosus. Zhonghua yi xue za zhi. 2017;97(19):1491-5.
69. Sukapan P, Promnarate P, Avihingsanon Y, Mutirangura A, Hirankarn N. Types of DNA methylation status of the interspersed repetitive sequences for LINE-1, Alu, HERV-E and HERV-K in the neutrophils from systemic lupus erythematosus patients and healthy controls. J Hum Genet. 2014;59(4):178-88.
70. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017;541(7635):81-6.
71. Mendelson MM, Marioni RE, Joehanes R, Liu C, Hedman AK, Aslibekyan S, et al. Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. PLoS Med. 2017;14(1):e1002215.
72. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aissi D, Wahl S, et al. DNA methylation and body-mass index: a genome-wide analysis. Lancet. 2014;383(9933):1990-8.
73. Kim AY, Park YJ, Pan X, Shin KC, Kwak SH, Bassas AF, et al. Obesity-induced DNA hypermethylation of the adiponectin gene mediates insulin resistance. Nat Commun. 2015;6:7585.
74. Kirchner H, Sinha I, Gao H, Ruby MA, Schonke M, Lindvall JM, et al. Altered DNA methylation of glycolytic and lipogenic genes in liver from obese and type 2 diabetic patients. Mol Metab. 2016;5(3):171-83.
75. Wang X, Zhu H, Snieder H, Su S, Munn D, Harshfield G, et al. Obesity related methylation changes in DNA of peripheral blood leukocytes. BMC Med. 2010;8:87.

76. Koh IU, Lee HJ, Hwang JY, Choi NH, Lee S. Obesity-related CpG Methylation (cg07814318) of Kruppel-like Factor-13 (KLF13) Gene with Childhood Obesity and its cis-Methylation Quantitative Loci. Sci Rep. 2017;7:45368.
77. Rhee JK, Lee JH, Yang HK, Kim TM, Yoon KH. DNA Methylation Profiles of Blood Cells Are Distinct between Early-Onset Obese and Control Individuals. Genomics Inform. 2017;15(1):28-37.
78. Fradin D, Boelle PY, Belot MP, Lachaux F, Tost J, Besse C, et al. Genome-Wide Methylation Analysis Identifies Specific Epigenetic Marks In Severely Obese Children. Sci Rep. 2017;7:46311.
79. Macartney-Coxson D, Benton MC, Blick R, Stubbs RS, Hagan RD, Langston MA. Genome-wide DNA methylation analysis reveals loci that distinguish different types of adipose tissue in obese individuals. Clin Epigenetics. 2017;9:48.
80. Day SE, Coletta RL, Kim JY, Garcia LA, Campbell LE, Benjamin TR, et al. Potential epigenetic biomarkers of obesity-related insulin resistance in human whole-blood. Epigenetics. 2017;12(4):254-63.
81. Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. Annu Rev Biochem. 2005;74:481-514.
82. Jurkowska RZ, Jurkowski TP, Jeltsch A. Structure and function of mammalian DNA methyltransferases. Chembiochem. 2011;12(2):206-22.
83. Bachman KE, Rountree MR, Baylin SB. Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. J Biol Chem. 2001;276(34):32282-7.
84. Chen T, Ueda Y, Xie S, Li E. A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. J Biol Chem. 2002;277(41):38746-54.
85. Gordon CA, Hartono SR, Chedin F. Inactive DNMT3B splice variants modulate de novo DNA methylation. PLoS One. 2013;8(7):e69486.
86. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38(1):23-38.
87. Choi SH, Heo K, Byun HM, An W, Lu W, Yang AS. Identification of preferential target sites for human DNA methyltransferases. Nucleic Acids Res. 2011;39(1):104-18.
88. Song J, Teplova M, Ishibe-Murakami S, Patel DJ. Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. Science. 2012;335(6069):709-12.
89. Zhang ZM, Liu S, Lin K, Luo Y, Perry JJ, Wang Y, et al. Crystal Structure of Human DNA Methyltransferase 1. J Mol Biol. 2015;427(15):2520-31.
90. Gaudet F, Hodgson JG, Eden A, Jackson-Grusby L, Dausman J, Gray JW, et al. Induction of tumors in mice by genomic hypomethylation. Science. 2003;300(5618):489-92.
91. Chen T, Hevi S, Gay F, Tsujimoto N, He T, Zhang B, et al. Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. Nat Genet. 2007;39(3):391-6.
92. Liao J, Karnik R, Gu H, Ziller MJ, Clement K, Tsankov AM, et al. Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. Nat Genet. 2015;47(5):469-78.
93. Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. Science. 2003;300(5618):455.
94. Laird PW, Jackson-Grusby L, Fazeli A, Dickinson SL, Jung WE, Li E, et al. Suppression of intestinal neoplasia by DNA hypomethylation. Cell. 1995;81(2):197-205.
95. Cormier RT, Dove WF. Dnmt1N/+ reduces the net growth rate and multiplicity of intestinal adenomas in C57BL/6-multiple intestinal neoplasia (Min)/+ mice independently of p53 but demonstrates strong synergy with the modifier of Min 1(AKR) resistance allele. Cancer Res. 2000;60(14):3965-70.
96. McGraw S, Oakes CC, Martel J, Cirio MC, de Zeeuw P, Mak W, et al. Loss of DNMT1o disrupts imprinted X chromosome inactivation and accentuates placental defects in females. PLoS Genet. 2013;9(11):e1003873.
97. O'Neill KM, Irwin RE, Mackin SJ, Thursby SJ, Thakur A, Bertens C, et al. Depletion of DNMT1 in differentiated human cells highlights key classes of sensitive genes and an interplay with polycomb repression. Epigenetics Chromatin. 2018;11(1):12.
98. Okano M, Xie S, Li E. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Res. 1998;26(11):2536-40.
99. Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, et al. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science. 2006;311(5759):395-8.

100. Rai K, Chidester S, Zavala CV, Manos EJ, James SR, Karpf AR, et al. Dnmt2 functions in the cytoplasm to promote liver, brain, and retina development in zebrafish. Genes Dev. 2007;21(3):261-6.
101. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99(3):247-57.
102. Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet. 1998;19(3):219-20.
103. Xie S, Wang Z, Okano M, Nogami M, Li Y, He WW, et al. Cloning, expression and chromosome locations of the human DNMT3 gene family. Gene. 1999;236(1):87-95.
104. Wang J, Walsh G, Liu DD, Lee JJ, Mao L. Expression of Delta DNMT3B variants and its association with promoter methylation of p16 and RASSF1A in primary non-small cell lung cancer. Cancer Res. 2006;66(17):8361-6.
105. Gopalakrishnan S, Van Emburgh BO, Shan J, Su Z, Fields CR, Vieweg J, et al. A novel DNMT3B splice variant expressed in tumor and pluripotent cells modulates genomic DNA methylation patterns and displays altered DNA binding. Mol Cancer Res. 2009;7(10):1622-34.
106. Ostler KR, Davis EM, Payne SL, Gosalia BB, Exposito-Cespedes J, Le Beau MM, et al. Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins. Oncogene. 2007;26(38):5553-63.
107. Wang L, Wang J, Sun S, Rodriguez M, Yue P, Jang SJ, et al. A novel DNMT3B subfamily, DeltaDNMT3B, is the predominant form of DNMT3B in non-small cell lung cancer. Int J Oncol. 2006;29(1):201-7.
108. Sharma S, De Carvalho DD, Jeong S, Jones PA, Liang G. Nucleosomes containing methylated DNA stabilize DNA methyltransferases 3A/3B and ensure faithful epigenetic inheritance. PLoS Genet. 2011;7(2):e1001286.
109. Jia YL, Guo X, Lu JT, Wang XY, Qiu LL, Wang TY. CRISPR/Cas9-mediated gene knockout for DNA methyltransferase Dnmt3a in CHO cells displays enhanced transgenic expression and long-term stability. J Cell Mol Med. 2018;22(9):4106-16.
110. Aapola U, Kawasaki K, Scott HS, Ollila J, Vihinen M, Heino M, et al. Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. Genomics. 2000;65(3):293-8.
111. Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature. 2004;431(7004):96-9.
112. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. Science. 2001;294(5551):2536-9.
113. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature. 2007;448(7154):714-7.
114. Neri F, Krepelova A, Incarnato D, Maldotti M, Parlato C, Galvagni F, et al. Dnmt3L antagonizes DNA methylation at bivalent promoters and favors DNA methylation at gene bodies in ESCs. Cell. 2013;155(1):121-34.
115. Suetake I, Shinozaki F, Miyagawa J, Takeshima H, Tajima S. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J Biol Chem. 2004;279(26):27816-23.
116. Chen ZX, Mann JR, Hsieh CL, Riggs AD, Chedin F. Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. J Cell Biochem. 2005;95(5):902-17.
117. Minami K, Chano T, Kawakami T, Ushida H, Kushima R, Okabe H, et al. DNMT3L is a novel marker and is essential for the growth of human embryonal carcinoma. Clin Cancer Res. 2010;16(10):2751-9.
118. Chedin F, Lieber MR, Hsieh CL. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. Proc Natl Acad Sci U S A. 2002;99(26):16916-21.
119. Saldivar-Gonzalez FI, Gomez-Garcia A, Chavez-Ponce de Leon DE, Sanchez-Cruz N, Ruiz-Rios J, Pilon-Jimenez BA, et al. Inhibitors of DNA Methyltransferases From Natural Sources: A Computational Perspective. Front Pharmacol. 2018;9:1144.
120. Vilkaitis G, Suetake I, Klimasauskas S, Tajima S. Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. J Biol Chem. 2005;280(1):64-72.
121. Schermelleh L, Haemmer A, Spada F, Rosing N, Meilinger D, Rothbauer U, et al. Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. Nucleic Acids Res. 2007;35(13):4301-12.

122. Cheray M, Nadaradjane A, Bonnet P, Routier S, Vallette FM, Cartron PF. Specific inhibition of DNMT1/CFP1 reduces cancer phenotypes and enhances chemotherapy effectiveness. Epigenomics. 2014;6(3):267-75.
123. Hervouet E, Peixoto P, Delage-Mourroux R, Boyer-Guittaut M, Cartron PF. Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma. Clin Epigenetics. 2018;10:17.
124. Negishi M, Chiba T, Saraya A, Miyagi S, Iwama A. Dmap1 plays an essential role in the maintenance of genome integrity through the DNA repair process. Genes Cells. 2009;14(11):1347-57.
125. Wu X, Gong Y, Yue J, Qiang B, Yuan J, Peng X. Cooperation between EZH2, NSPc1-mediated histone H2A ubiquitination and Dnmt1 in HOX gene silencing. Nucleic Acids Res. 2008;36(11):3590-9.
126. Gu P, Xu X, Le Menuet D, Chung AC, Cooney AJ. Differential recruitment of methyl CpG-binding domain factors and DNA methyltransferases by the orphan receptor germ cell nuclear factor initiates the repression and silencing of Oct4. Stem Cells. 2011;29(7):1041-51.
127. Ying J, Li H, Seng TJ, Langford C, Srivastava G, Tsao SW, et al. Functional epigenetics identifies a protocadherin PCDH10 as a candidate tumor suppressor for nasopharyngeal, esophageal and multiple other carcinomas with frequent methylation. Oncogene. 2006;25(7):1070-80.
128. Zhang W, Xu J. DNA methyltransferases and their roles in tumorigenesis. Biomark Res. 2017;5:1.
129. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317-30.
130. Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, et al. Intragenic DNA methylation prevents spurious transcription initiation. Nature. 2017;543(7643):72-7.
131. Maunakea AK, Chepelev I, Cui K, Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. Cell Res. 2013;23(11):1256-69.
132. Ehrlich M. DNA hypomethylation in cancer cells. Epigenomics. 2009;1(2):239-59.
133. Bhave MR, Wilson MJ, Poirier LA. c-H-ras and c-K-ras gene hypomethylation in the livers and hepatomas of rats fed methyl-deficient, amino acid-defined diets. Carcinogenesis. 1988;9(3):343-8.
134. Esteller M. Epigenetics in cancer. N Engl J Med. 2008;358(11):1148-59.
135. Wu J, Issa JP, Herman J, Bassett DE, Jr., Nelkin BD, Baylin SB. Expression of an exogenous eukaryotic DNA methyltransferase gene induces transformation of NIH 3T3 cells. Proc Natl Acad Sci U S A. 1993;90(19):8891-5.
136. Oridate N, Lotan R. Suppression of DNA methyltransferase 1 levels in head and neck squamous carcinoma cells using small interfering RNA results in growth inhibition and increase in Cdk inhibitor p21. Int J Oncol. 2005;26(3):757-61.
137. Peters SL, Hlady RA, Opavska J, Klinkebiel D, Novakova S, Smith LM, et al. Essential role for Dnmt1 in the prevention and maintenance of MYC-induced T-cell lymphomas. Mol Cell Biol. 2013;33(21):4321-33.
138. Poole CJ, Zheng W, Lodh A, Yevtodiyenko A, Liefwalker D, Li H, et al. DNMT3B overexpression contributes to aberrant DNA methylation and MYC-driven tumor maintenance in T-ALL and Burkitt's lymphoma. Oncotarget. 2017;8(44):76898-920.
139. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010;363(25):2424-33.
140. Zhao Z, Wu Q, Cheng J, Qiu X, Zhang J, Fan H. Depletion of DNMT3A suppressed cell proliferation and restored PTEN in hepatocellular carcinoma cell. J Biomed Biotechnol. 2010;2010:737535.
141. Kobayashi Y, Absher DM, Gulzar ZG, Young SR, McKenney JK, Peehl DM, et al. DNA methylation profiling reveals novel biomarkers and important roles for DNA methyltransferases in prostate cancer. Genome Res. 2011;21(7):1017-27.
142. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer. 2011;11(10):726-34.
143. Yan XJ, Xu J, Gu ZH, Pan CM, Lu G, Shen Y, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. Nat Genet. 2011;43(4):309-15.
144. Walter MJ, Ding L, Shen D, Shao J, Grillot M, McLellan M, et al. Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. Leukemia. 2011;25(7):1153-8.

145. Challen GA, Sun D, Jeong M, Luo M, Jelinek J, Berg JS, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. Nat Genet. 2011;44(1):23-31.
146. Richardson B. DNA methylation and autoimmune disease. Clin Immunol. 2003;109(1):72-9.
147. Li Y, Zhao M, Hou C, Liang G, Yang L, Tan Y, et al. Abnormal DNA methylation in CD4+ T cells from people with latent autoimmune diabetes in adults. Diabetes Res Clin Pract. 2011;94(2):242-8.
148. Luo Y, Li Y, Su Y, Yin H, Hu N, Wang S, et al. Abnormal DNA methylation in T cells from patients with subacute cutaneous lupus erythematosus. Br J Dermatol. 2008;159(4):827-33.
149. Zhao M, Gao F, Wu X, Tang J, Lu Q. Abnormal DNA methylation in peripheral blood mononuclear cells from patients with vitiligo. Br J Dermatol. 2010;163(4):736-42.
150. Liu CC, Fang TJ, Ou TT, Wu CC, Li RN, Lin YC, et al. Global DNA methylation, DNMT1, and MBD2 in patients with rheumatoid arthritis. Immunol Lett. 2011;135(1-2):96-9.
151. Nguyen S, Meletis K, Fu D, Jhaveri S, Jaenisch R. Ablation of de novo DNA methyltransferase Dnmt3a in the nervous system leads to neuromuscular defects and shortened lifespan. Dev Dyn. 2007;236(6):1663-76.
152. Singh RK, Mallela RK, Hayes A, Dunham NR, Hedden ME, Enke RA, et al. Dnmt1, Dnmt3a and Dnmt3b cooperate in photoreceptor and outer plexiform layer development in the mammalian retina. Exp Eye Res. 2017;159:132-46.
153. Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, et al. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc Natl Acad Sci U S A. 1999;96(25):14412-7.
154. Heyn P, Logan CV, Fluteau A, Challis RC, Auchynnikava T, Martin CA, et al. Gain-of-function DNMT3A mutations cause microcephalic dwarfism and hypermethylation of Polycomb-regulated regions. Nat Genet. 2019;51(1):96-105.
155. Baets J, Duan X, Wu Y, Smith G, Seeley WW, Mademan I, et al. Defects of mutant DNMT1 are linked to a spectrum of neurological disorders. Brain. 2015;138(Pt 4):845-61.
156. Gnyszka A, Jastrzebski Z, Flis S. DNA methyltransferase inhibitors and their emerging role in epigenetic therapy of cancer. Anticancer Res. 2013;33(8):2989-96.
157. Santi DV, Norment A, Garrett CE. Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. Proc Natl Acad Sci U S A. 1984;81(22):6993-7.
158. Stresemann C, Lyko F. Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. Int J Cancer. 2008;123(1):8-13.
159. Gros C, Fahy J, Halby L, Dufau I, Erdmann A, Gregoire JM, et al. DNA methylation inhibitors in cancer: recent and future approaches. Biochimie. 2012;94(11):2280-96.
160. Cowan LA, Talwar S, Yang AS. Will DNA methylation inhibitors work in solid tumors? A review of the clinical experience with azacitidine and decitabine in solid tumors. Epigenomics. 2010;2(1):71-86.
161. Taweechaipaisankul A, Kim GA, Jin JX, Lee S, Qasim M, Kim EH, et al. Enhancement of epigenetic reprogramming status of porcine cloned embryos with zebularine, a DNA methyltransferase inhibitor. Mol Reprod Dev. 2019;86(8):1013-22.
162. Vanden Berghe W. Epigenetic impact of dietary polyphenols in cancer chemoprevention: lifelong remodeling of our epigenomes. Pharmacol Res. 2012;65(6):565-76.
163. Yang CS, Landau JM, Huang MT, Newmark HL. Inhibition of carcinogenesis by dietary polyphenolic compounds. Annu Rev Nutr. 2001;21:381-406.
164. Li Y, Tollefsbol TO. Impact on DNA methylation in cancer prevention and therapy by bioactive dietary components. Curr Med Chem. 2010;17(20):2141-51.
165. Kadayifci FZ, Zheng S, Pan YX. Molecular Mechanisms Underlying the Link between Diet and DNA Methylation. Int J Mol Sci. 2018;19(12).
166. Lee WJ, Shim JY, Zhu BT. Mechanisms for the inhibition of DNA methyltransferases by tea catechins and bioflavonoids. Mol Pharmacol. 2005;68(4):1018-30.
167. Fang MZ, Wang Y, Ai N, Hou Z, Sun Y, Lu H, et al. Tea polyphenol (-)-epigallocatechin-3-gallate inhibits DNA methyltransferase and reactivates methylation-silenced genes in cancer cell lines. Cancer Res. 2003;63(22):7563-70.
168. Nandakumar V, Vaid M, Katiyar SK. (-)-Epigallocatechin-3-gallate reactivates silenced tumor suppressor genes, Cip1/p21 and p16INK4a, by reducing DNA methylation and increasing histones acetylation in human skin cancer cells. Carcinogenesis. 2011;32(4):537-44.

169. Zhang BK, Lai YQ, Niu PP, Zhao M, Jia SJ. Epigallocatechin-3-gallate inhibits homocysteine-induced apoptosis of endothelial cells by demethylation of the DDAH2 gene. Planta Med. 2013;79(18):1715-9.
170. Shukla S, Trokhan S, Resnick MI, Gupta S. Epigallocatechin-3-gallate causes demethylation and activation of GSTP1 gene expression in human prostate cancer LNCaP cells. AACR; 2005.
171. Morris J, Moseley VR, Cabang AB, Coleman K, Wei W, Garrett-Mayer E, et al. Reduction in promotor methylation utilizing EGCG (epigallocatechin-3-gallate) restores RXRalpha expression in human colon cancer cells. Oncotarget. 2016;7(23):35313-26.
172. Pandey M, Shukla S, Gupta S. Promoter demethylation and chromatin remodeling by green tea polyphenols leads to re-expression of GSTP1 in human prostate cancer cells. Int J Cancer. 2010;126(11):2520-33.
173. Lee WJ, Zhu BT. Inhibition of DNA methylation by caffeic acid and chlorogenic acid, two common catechol-containing coffee polyphenols. Carcinogenesis. 2006;27(2):269-77.
174. Liu Z, Xie Z, Jones W, Pavlovicz RE, Liu S, Yu J, et al. Curcumin is a potent DNA hypomethylation agent. Bioorg Med Chem Lett. 2009;19(3):706-9.
175. Rajavelu A, Tulyasheva Z, Jaiswal R, Jeltsch A, Kuhnert N. The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols. BMC Biochem. 2011;12:16.
176. Venturelli S, Sinnberg TW, Berger A, Noor S, Levesque MP, Bocker A, et al. Epigenetic impacts of ascorbate on human metastatic melanoma cells. Front Oncol. 2014;4:227.
177. Milenkovic D, Declerck K, Guttman Y, Kerem Z, Claude S, Weseler AR, et al. (-)-Epicatechin metabolites promote vascular health through epigenetic reprogramming of endothelial-immune cell signaling and reversing systemic low-grade inflammation. Biochem Pharmacol. 2019:113699.
178. Sidoryk K, Jaromin A, Filipczak N, Cmoch P, Cybulski M. Synthesis and Antioxidant Activity of Caffeic Acid Derivatives. Molecules. 2018;23(9).
179. Damasceno SS, Dantas BB, Ribeiro-Filho J, Antonio MAD, Galberto MdCJ. Chemical Properties of Caffeic and Ferulic Acids in Biological System: Implications in Cancer Therapy. A Review. Curr Pharm Des. 2017;23(20):3015-23.
180. Gupta SC, Prasad S, Kim JH, Patchva S, Webb LJ, Priyadarsini IK, et al. Multitargeting by curcumin as revealed by molecular interaction studies. Nat Prod Rep. 2011;28(12):1937-55.
181. Reuter S, Gupta SC, Park B, Goel A, Aggarwal BB. Epigenetic changes induced by curcumin and other natural compounds. Genes Nutr. 2011;6(2):93-108.
182. Leung LK, Su Y, Chen R, Zhang Z, Huang Y, Chen ZY. Theaflavins in black tea and catechins in green tea are equally effective antioxidants. J Nutr. 2001;131(9):2248-51.
183. Zhang S, Yang C, Idehen E, Shi L, Lv L, Sang S. Novel Theaflavin-Type Chlorogenic Acid Derivatives Identified in Black Tea. J Agric Food Chem. 2018;66(13):3402-7.
184. Camarena V, Wang G. The epigenetic role of vitamin C in health and disease. Cell Mol Life Sci. 2016;73(8):1645-58.
185. Gillberg L, Orskov AD, Nasif A, Ohtani H, Madaj Z, Hansen JW, et al. Oral vitamin C supplementation to patients with myeloid cancer on azacitidine treatment: Normalization of plasma vitamin C induces epigenetic changes. Clin Epigenetics. 2019;11(1):143.
186. Tammen SA, Friso S, Choi SW. Epigenetics: the link between nature and nurture. Mol Aspects Med. 2013;34(4):753-64.
187. Alegria-Torres JA, Baccarelli A, Bollati V. Epigenetics and lifestyle. Epigenomics. 2011;3(3):267-77.
188. Crider KS, Bailey LB, Berry RJ. Folic acid food fortification-its history, effect, concerns, and future directions. Nutrients. 2011;3(3):370-84.
189. Miller AL. The methionine-homocysteine cycle and its effects on cognitive diseases. Altern Med Rev. 2003;8(1):7-19.
190. Jacob RA, Gretz DM, Taylor PC, James SJ, Pogribny IP, Miller BJ, et al. Moderate folate depletion increases plasma homocysteine and decreases lymphocyte DNA methylation in postmenopausal women. J Nutr. 1998;128(7):1204-12.
191. Shin W, Yan J, Abratte CM, Vermeylen F, Caudill MA. Choline intake exceeding current dietary recommendations preserves markers of cellular methylation in a genetic subgroup of folate-compromised men. J Nutr. 2010;140(5):975-80.

192. Chamberlain JA, Dugue PA, Bassett JK, Hodge AM, Brinkman MT, Joo JE, et al. Dietary intake of one-carbon metabolism nutrients and DNA methylation in peripheral blood. Am J Clin Nutr. 2018;108(3):611-21.
193. Xiang N, Zhao R, Song G, Zhong W. Selenite reactivates silenced genes by modifying DNA methylation and histones in prostate cancer cells. Carcinogenesis. 2008;29(11):2175-81.
194. Davis CD, Uthus EO, Finley JW. Dietary selenium and arsenic affect DNA methylation in vitro in Caco-2 cells and in vivo in rat liver and colon. J Nutr. 2000;130(12):2903-9.
195. Fini L, Selgrad M, Fogliano V, Graziani G, Romano M, Hotchkiss E, et al. Annurca apple polyphenols have potent demethylating activity and can reactivate silenced tumor suppressor genes in colorectal cancer cells. J Nutr. 2007;137(12):2622-8.
196. Link A, Balaguer F, Goel A. Cancer chemoprevention by dietary polyphenols: promising role for epigenetics. Biochemical pharmacology. 2010;80(12):1771-92.
197. Qin W, Zhu W, Shi H, Hewett JE, Ruhlen RL, MacDonald RS, et al. Soy isoflavones have an antiestrogenic effect and alter mammary promoter hypermethylation in healthy premenopausal women. Nutr Cancer. 2009;61(2):238-44.
198. McKay JA, Mathers JC. Diet induced epigenetic changes and their implications for health. Acta Physiol (Oxf). 2011;202(2):103-18.
199. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. The Journal of nutritional biochemistry. 2012;23(8):853-9.
200. Friso S, Udali S, De Santis D, Choi SW. One-carbon metabolism and epigenetics. Mol Aspects Med. 2017;54:28-36.
201. Wasson GR, McGlynn AP, McNulty H, O'Reilly SL, McKelvey-Martin VJ, McKerr G, et al. Global DNA and p53 region-specific hypomethylation in human colonic cells is induced by folate depletion and reversed by folate supplementation. J Nutr. 2006;136(11):2748-53.
202. Dolinoy DC, Huang D, Jirtle RL. Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. Proc Natl Acad Sci U S A. 2007;104(32):13056-61.
203. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol. 2003;23(15):5293-300.
204. Cropley JE, Dang TH, Martin DI, Suter CM. The penetrance of an epigenetic trait in mice is progressively yet reversibly increased by selection and environment. Proc Biol Sci. 2012;279(1737):2347-53.
205. Barua S, Chadman KK, Kuizon S, Buenaventura D, Stapley NW, Ruocco F, et al. Increasing maternal or post-weaning folic acid alters gene expression and moderately changes behavior in the offspring. PLoS One. 2014;9(7):e101674.
206. Dolinoy DC. The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. Nutr Rev. 2008;66 Suppl 1:S7-11.
207. Wolff GL, Kodell RL, Moore SR, Cooney CA. Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. FASEB J. 1998;12(11):949-57.
208. Oster M, Nuchchanart W, Trakooljul N, Murani E, Zeyner A, Wirthgen E, et al. Methylating micronutrient supplementation during pregnancy influences foetal hepatic gene expression and IGF signalling and increases foetal weight. Eur J Nutr. 2016;55(4):1717-27.
209. Kok DE, Dhonukshe-Rutten RA, Lute C, Heil SG, Uitterlinden AG, van der Velde N, et al. The effects of long-term daily folic acid and vitamin B12 supplementation on genome-wide DNA methylation in elderly subjects. Clin Epigenetics. 2015;7:121.
210. O'Reilly SL, McGlynn AP, McNulty H, Reynolds J, Wasson GR, Molloy AM, et al. Folic Acid Supplementation in Postpolypectomy Patients in a Randomized Controlled Trial Increases Tissue Folate Concentrations and Reduces Aberrant DNA Biomarkers in Colonic Tissues Adjacent to the Former Polyp Site. J Nutr. 2016;146(5):933-9.
211. McKay JA, Waltham KJ, Williams EA, Mathers JC. Folate depletion during pregnancy and lactation reduces genomic DNA methylation in murine adult offspring. Genes Nutr. 2011;6(2):189-96.
212. Irwin RE, Thursby SJ, Ondicova M, Pentieva K, McNulty H, Richmond RC, et al. A randomized controlled trial of folic acid intervention in pregnancy highlights a putative methylation-regulated control element at ZFP57. Clin Epigenetics. 2019;11(1):31.

213. Farias N, Ho N, Butler S, Delaney L, Morrison J, Shahrzad S, et al. The effects of folic acid on global DNA methylation and colonosphere formation in colon cancer cell lines. J Nutr Biochem. 2015;26(8):818-26.
214. Ding YB, He JL, Liu XQ, Chen XM, Long CL, Wang YX. Expression of DNA methyltransferases in the mouse uterus during early pregnancy and susceptibility to dietary folate deficiency. Reproduction. 2012;144(1):91-100.
215. Zeisel S. Choline, Other Methyl-Donors and Epigenetics. Nutrients. 2017;9(5).
216. Niculescu MD, Craciunescu CN, Zeisel SH. Dietary choline deficiency alters global and gene-specific DNA methylation in the developing hippocampus of mouse fetal brains. FASEB J. 2006;20(1):43-9.
217. Mehedint MG, Niculescu MD, Craciunescu CN, Zeisel SH. Choline deficiency alters global histone methylation and epigenetic marking at the Re1 site of the calbindin 1 gene. FASEB J. 2010;24(1):184-95.
218. Mehedint MG, Craciunescu CN, Zeisel SH. Maternal dietary choline deficiency alters angiogenesis in fetal mouse hippocampus. Proc Natl Acad Sci U S A. 2010;107(29):12834-9.
219. Wilson MJ, Shivapurkar N, Poirier LA. Hypomethylation of hepatic nuclear DNA in rats fed with a carcinogenic methyl-deficient diet. Biochem J. 1984;218(3):987-90.
220. Zhang N. Role of methionine on epigenetic modification of DNA methylation and gene expression in animals. Anim Nutr. 2018;4(1):11-6.
221. Miousse IR, Pathak R, Garg S, Skinner CM, Melnyk S, Pavliv O, et al. Short-term dietary methionine supplementation affects one-carbon metabolism and DNA methylation in the mouse gut and leads to altered microbiome profiles, barrier function, gene expression and histomorphology. Genes Nutr. 2017;12:22.
222. Lee WJ, Zhu BT. Inhibition of DNA methylation by caffeic acid and chlorogenic acid, two common catechol-containing coffee polyphenols. Carcinogenesis. 2006;27(2):269-77.
223. Fang M, Chen D, Yang CS. Dietary polyphenols may affect DNA methylation. The Journal of nutrition. 2007;137(1):223S-8S.
224. Rangel-Huerta OD, Pastor-Villaescusa B, Aguilera CM, Gil A. A systematic review of the efficacy of bioactive compounds in cardiovascular disease: phenolic compounds. Nutrients. 2015;7(7):5177-216.
225. Steinmann J, Buer J, Pietschmann T, Steinmann E. Anti-infective properties of epigallocatechin-3-gallate (EGCG), a component of green tea. Br J Pharmacol. 2013;168(5):1059-73.
226. Meeran SM, Patel SN, Chan TH, Tollefsbol TO. A novel prodrug of epigallocatechin-3-gallate: differential epigenetic hTERT repression in human breast cancer cells. Cancer Prev Res (Phila). 2011;4(8):1243-54.
227. Fang MZ, Chen D, Sun Y, Jin Z, Christman JK, Yang CS. Reversal of hypermethylation and reactivation of p16INK4a, RARbeta, and MGMT genes by genistein and other isoflavones from soy. Clin Cancer Res. 2005;11(19 Pt 1):7033-41.
228. Fang M, Chen D, Yang CS. Dietary polyphenols may affect DNA methylation. J Nutr. 2007;137(1 Suppl):223S-8S.
229. Mukherjee N, Kumar AP, Ghosh R. DNA Methylation and Flavonoids in Genitourinary Cancers. Curr Pharmacol Rep. 2015;1(2):112-20.
230. Collaborators GBDD. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958-72.
231. Aller EE, Abete I, Astrup A, Martinez JA, van Baak MA. Starches, sugars and obesity. Nutrients. 2011;3(3):341-69.
232. Holmes MD, Liu S, Hankinson SE, Colditz GA, Hunter DJ, Willett WC. Dietary carbohydrates, fiber, and breast cancer risk. Am J Epidemiol. 2004;159(8):732-9.
233. Goodstine SL, Zheng T, Holford TR, Ward BA, Carter D, Owens PH, et al. Dietary (n-3)/(n-6) fatty acid ratio: possible relationship to premenopausal but not postmenopausal breast cancer risk in U.S. women. J Nutr. 2003;133(5):1409-14.
234. Maillard V, Bougnoux P, Ferrari P, Jourdan ML, Pinault M, Lavillonniere F, et al. N-3 and N-6 fatty acids in breast adipose tissue and relative risk of breast cancer in a case-control study in Tours, France. Int J Cancer. 2002;98(1):78-83.
235. Bingham SA, Hughes R, Cross AJ. Effect of white versus red meat on endogenous N-nitrosation in the human colon and further evidence of a dose response. J Nutr. 2002;132(11 Suppl):3522S-5S.

236. Gani LU, How CH. PILL Series. Vitamin D deficiency. Singapore Med J. 2015;56(8):433-6; quiz 7.
237. Sommer A. Xerophthalmia and vitamin A status. Prog Retin Eye Res. 1998;17(1):9-31.
238. Kirke PN, Molloy AM, Daly LE, Burke H, Weir DG, Scott JM. Maternal plasma folate and vitamin B12 are independent risk factors for neural tube defects. Q J Med. 1993;86(11):703-8.
239. O'Leary F, Samman S. Vitamin B12 in health and disease. Nutrients. 2010;2(3):299-316.
240. Nishizawa Y, Yamamoto T, Terada N, Fushiki S, Matsumoto K, Nishizawa Y. Effects of methylcobalamin on the proliferation of androgen-sensitive or estrogen-sensitive malignant cells in culture and in vivo. Int J Vitam Nutr Res. 1997;67(3):164-70.
241. Mathers JC, Strathdee G, Relton CL. Induction of epigenetic alterations by dietary and other environmental factors. Adv Genet. 2010;71:3-39.
242. Pirouzpanah S, Taleban FA, Mehdipour P, Atri M. Association of folate and other one-carbon related nutrients with hypermethylation status and expression of RARB, BRCA1, and RASSF1A genes in breast cancer patients. J Mol Med (Berl). 2015;93(8):917-34.
243. Piyathilake CJ, Johanning GL, Macaluso M, Whiteside M, Oelschlager DK, Heimburger DC, et al. Localized folate and vitamin B-12 deficiency in squamous cell lung cancer is associated with global DNA hypomethylation. Nutr Cancer. 2000;37(1):99-107.
244. Fang JY, Xiao SD. Folic acid, polymorphism of methyl-group metabolism genes, and DNA methylation in relation to GI carcinogenesis. J Gastroenterol. 2003;38(9):821-9.
245. MacGregor JT, Schlegel R, Wehr CM, Alperin P, Ames BN. Cytogenetic damage induced by folate deficiency in mice is enhanced by caffeine. Proc Natl Acad Sci U S A. 1990;87(24):9962-5.
246. Pogribny IP, Basnakian AG, Miller BJ, Lopatina NG, Poirier LA, James SJ. Breaks in genomic DNA and within the p53 gene are associated with hypomethylation in livers of folate/methyl-deficient rats. Cancer Res. 1995;55(9):1894-901.
247. van Straten EM, Bloks VW, Huijkman NC, Baller JF, van Meer H, Lutjohann D, et al. The liver X-receptor gene promoter is hypermethylated in a mouse model of prenatal protein restriction. Am J Physiol Regul Integr Comp Physiol. 2010;298(2):R275-82.
248. Rampersaud GC, Kauwell GP, Hutson AD, Cerda JJ, Bailey LB. Genomic DNA methylation decreases in response to moderate folate depletion in elderly women. Am J Clin Nutr. 2000;72(4):998-1003.
249. Choi SW, Mason JB. Folate and carcinogenesis: an integrated scheme. J Nutr. 2000;130(2):129-32.
250. Homocysteine Studies C. Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis. JAMA. 2002;288(16):2015-22.
251. Yi P, Melnyk S, Pogribna M, Pogribny IP, Hine RJ, James SJ. Increase in plasma homocysteine associated with parallel increases in plasma S-adenosylhomocysteine and lymphocyte DNA hypomethylation. J Biol Chem. 2000;275(38):29318-23.
252. Li B, Chang S, Liu C, Zhang M, Zhang L, Liang L, et al. Low Maternal Dietary Folate Alters Retrotranspose by Methylation Regulation in Intrauterine Growth Retardation (IUGR) Fetuses in a Mouse Model. Med Sci Monit. 2019;25:3354-65.
253. Kalani A, Chaturvedi P, Kalani K, Kamat PK, Chaturvedi P. A high methionine, low folate and vitamin B6/B12 containing diet can be associated with memory loss by epigenetic silencing of netrin-1. Neural Regen Res. 2019;14(7):1247-54.
254. Hoffman DR, Marion DW, Cornatzer WE, Duerre JA. S-Adenosylmethionine and S-adenosylhomocystein metabolism in isolated rat liver. Effects of L-methionine, L-homocystein, and adenosine. J Biol Chem. 1980;255(22):10822-7.
255. De Cabo SF, Hazen MJ, Molero ML, Fernandez-Piqueras J. S-adenosyl-L-homocysteine: a non-cytotoxic hypomethylating agent. Experientia. 1994;50(7):658-9.
256. MacFarlane AJ, Liu X, Perry CA, Flodby P, Allen RH, Stabler SP, et al. Cytoplasmic serine hydroxymethyltransferase regulates the metabolic partitioning of methylenetetrahydrofolate but is not essential in mice. J Biol Chem. 2008;283(38):25846-53.
257. Devlin AM, Arning E, Bottiglieri T, Faraci FM, Rozen R, Lentz SR. Effect of Mthfr genotype on diet-induced hyperhomocysteinemia and vascular function in mice. Blood. 2004;103(7):2624-9.
258. Choumenkovitch SF, Selhub J, Bagley PJ, Maeda N, Nadeau MR, Smith DE, et al. In the cystathionine beta-synthase knockout mouse, elevations in total plasma homocysteine increase tissue S-adenosylhomocysteine, but responses of S-adenosylmethionine and DNA methylation are tissue specific. J Nutr. 2002;132(8):2157-60.

259. Nizel AE. Personalized nutrition counseling. ASDC J Dent Child. 1972;39(5):353-60.
260. Ordovas JM, Ferguson LR, Tai ES, Mathers JC. Personalised nutrition and health. BMJ. 2018;361:bmj k2173.
261. Mathers JC. Paving the way to better population health through personalised nutrition. EFSA Journal. 2019;17:e170713.
262. Celis-Morales C, Livingstone KM, Marsaux CF, Forster H, O'Donovan CB, Woolhead C, et al. Design and baseline characteristics of the Food4Me study: a web-based randomised controlled trial of personalised nutrition in seven European countries. Genes Nutr. 2015;10(1):450.
263. Celis-Morales C, Livingstone KM, Marsaux CF, Macready AL, Fallaize R, O'Donovan CB, et al. Effect of personalized nutrition on health-related behaviour change: evidence from the Food4Me European randomized controlled trial. Int J Epidemiol. 2017;46(2):578-88.
264. Nielsen DE, El-Sohemy A. A randomized trial of genetic information for personalized nutrition. Genes Nutr. 2012;7(4):559-66.
265. Yost JM, Korboukh I, Liu F, Gao C, Jin J. Targets in epigenetics: inhibiting the methyl writers of the histone code. Curr Chem Genomics. 2011;5(Suppl 1):72-84.
266. de Jonge HJ, Fehrmann RS, de Bont ES, Hofstra RM, Gerbens F, Kamps WA, et al. Evidence based selection of housekeeping genes. PLoS One. 2007;2(9):e898.
267. Provenzano M, Mocellin S. Complementary techniques: validation of gene expression data by quantitative real time PCR. Adv Exp Med Biol. 2007;593:66-73.
268. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 2002;3(7):RESEARCH0034.
269. Fortin JP, Triche TJ, Jr., Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2017;33(4):558-60.
270. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013;14:293.
271. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol. 2016;17(1):208.
272. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Res. 2017;45(4):e22.
273. Nazarenko MS, Markov AV, Lebedev IN, Freidin MB, Sleptcov AA, Koroleva IA, et al. A comparison of genome-wide DNA methylation patterns between different vascular tissues from patients with coronary heart disease. PLoS One. 2015;10(4):e0122601.
274. Naumov VA, Generozov EV, Zaharjevskaya NB, Matushkina DS, Larin AK, Chernyshov SV, et al. Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips. Epigenetics. 2013;8(9):921-34.
275. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57.
276. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 1992;69(6):915-26.
277. Duymich CE, Charlet J, Yang X, Jones PA, Liang G. DNMT3B isoforms without catalytic activity stimulate gene body methylation as accessory proteins in somatic cells. Nat Commun. 2016;7:11453.
278. Li M, Wang Y, Song Y, Bu R, Yin B, Fei X, et al. Expression profiling and clinicopathological significance of DNA methyltransferase 1, 3A and 3B in sporadic human renal cell carcinoma. Int J Clin Exp Pathol. 2014;7(11):7597-609.
279. Gujar H, Weisenberger DJ, Liang G. The Roles of Human DNA Methyltransferases and Their Isoforms in Shaping the Epigenome. Genes (Basel). 2019;10(2).
280. Tajima S, Suetake I, Takeshita K, Nakagawa A, Kimura H. Domain Structure of the Dnmt1, Dnmt3a, and Dnmt3b DNA Methyltransferases. Adv Exp Med Biol. 2016;945:63-86.
281. Biniszkiewicz D, Gribnau J, Ramsahoye B, Gaudet F, Eggan K, Humpherys D, et al. Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. Mol Cell Biol. 2002;22(7):2124-35.
282. Amara K, Ziadi S, Hachana M, Soltani N, Korbi S, Trimeche M. DNA methyltransferase DNMT3b protein overexpression as a prognostic factor in patients with diffuse large B-cell lymphomas. Cancer Sci. 2010;101(7):1722-30.

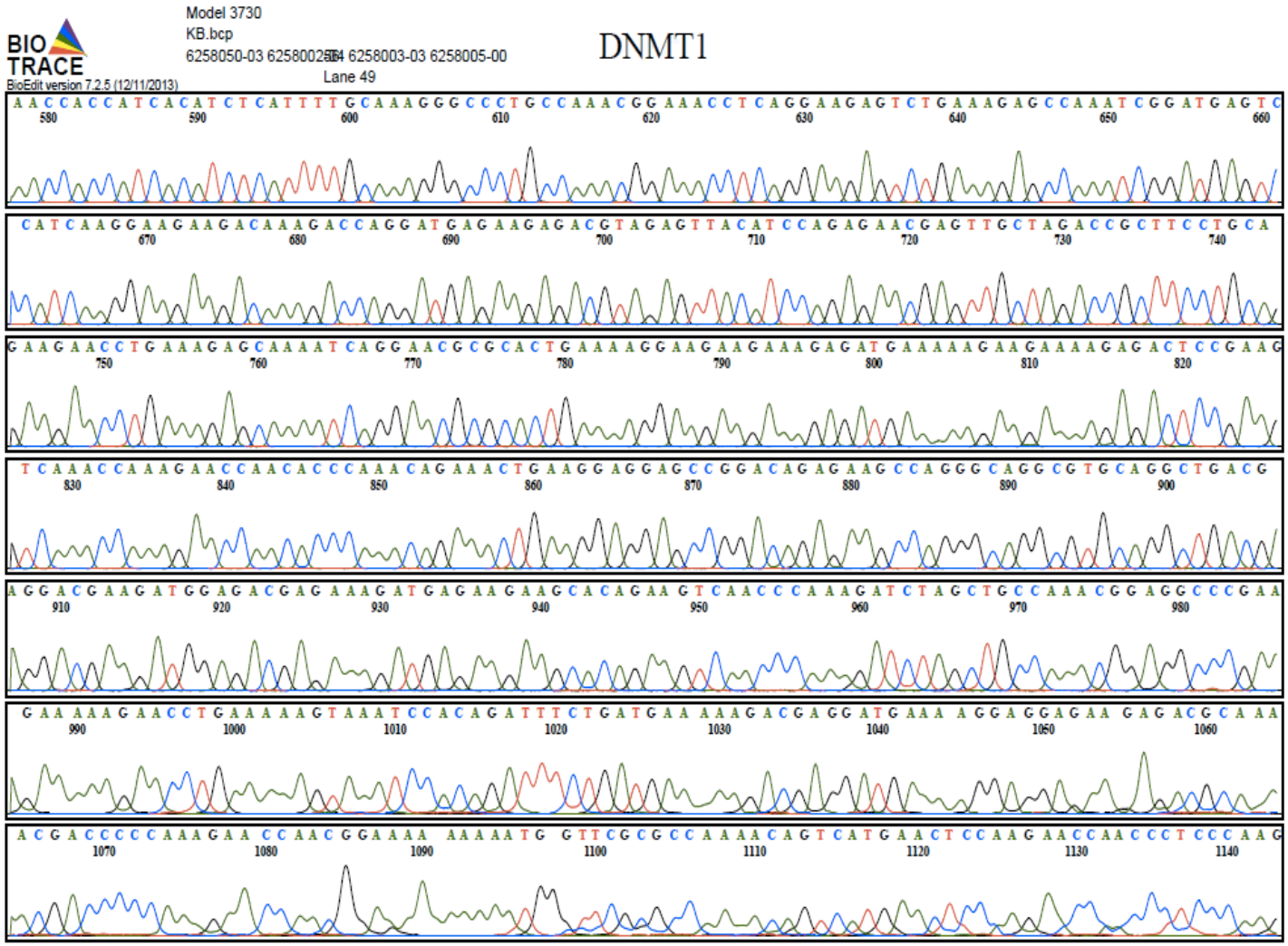
283. Roll JD, Rivenbark AG, Jones WD, Coleman WB. DNMT3b overexpression contributes to a hypermethylator phenotype in human breast cancer cell lines. Mol Cancer. 2008;7:15.
284. Linhart HG, Lin H, Yamada Y, Moran E, Steine EJ, Gokhale S, et al. Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. Genes Dev. 2007;21(23):3110-22.
285. Saito Y, Kanai Y, Sakamoto M, Saito H, Ishii H, Hirohashi S. Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis. Proc Natl Acad Sci U S A. 2002;99(15):10060-5.
286. Lovkvist C, Dodd IB, Sneppen K, Haerter JO. DNA methylation in human epigenomes depends on local topology of CpG sites. Nucleic Acids Res. 2016;44(11):5123-32.
287. Baccarelli AA, Byun HM. Platelet mitochondrial DNA methylation: a potential new marker of cardiovascular disease. Clin Epigenetics. 2015;7:44.
288. Bukrinsky MI, Haffar OK. HIV-1 nuclear import: in search of a leader. Front Biosci. 1999;4:D772-81.
289. Burstein SA, Dubart A, Norol F, Debili N, Friese P, Downs T, et al. Expression of a foreign protein in human megakaryocytes and platelets by retrovirally mediated gene transfer. Exp Hematol. 1999;27(1):110-6.
290. Zitzmann J, Schreiber C, Eichmann J, Bilz RO, Salzig D, Weidner T, et al. Single-cell cloning enables the selection of more productive Drosophila melanogaster S2 cells for recombinant protein expression. Biotechnol Rep (Amst). 2018;19:e00272.
291. Torres R, Garcia A, Jimenez M, Rodriguez S, Ramirez JC. An integration-defective lentivirus-based resource for site-specific targeting of an edited safe-harbour locus in the human genome. Gene Ther. 2014;21(4):343-52.
292. Gedeon T, Bokes P. Delayed protein synthesis reduces the correlation between mRNA and protein fluctuations. Biophys J. 2012;103(3):377-85.
293. Gry M, Rimini R, Stromberg S, Asplund A, Ponten F, Uhlen M, et al. Correlations between RNA and protein expression profiles in 23 human cell lines. BMC Genomics. 2009;10:365.
294. Edfors F, Danielsson F, Hallstrom BM, Kall L, Lundberg E, Ponten F, et al. Gene-specific correlation of RNA and protein levels in human cells and tissues. Mol Syst Biol. 2016;12(10):883.
295. Solomon O, MacIsaac J, Quach H, Tindula G, Kobor MS, Huen K, et al. Comparison of DNA methylation measured by Illumina 450K and EPIC BeadChips in blood of newborns and 14-year-old children. Epigenetics. 2018;13(6):655-64.
296. Ren W, Gao L, Song J. Structural Basis of DNMT1 and DNMT3A-Mediated DNA Methylation. Genes (Basel). 2018;9(12).
297. Hervouet E, Vallette FM, Cartron PF. Dnmt1/Transcription factor interactions: an alternative mechanism of DNA methylation inheritance. Genes Cancer. 2010;1(5):434-43.
298. Senyuk V, Premanand K, Xu P, Qian Z, Nucifora G. The oncoprotein EVI1 and the DNA methyltransferase Dnmt3 co-operate in binding and de novo methylation of target DNA. PLoS One. 2011;6(6):e20793.
299. Farthing CR, Ficz G, Ng RK, Chan CF, Andrews S, Dean W, et al. Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. PLoS Genet. 2008;4(6):e1000116.
300. Han H, Cortez CC, Yang X, Nichols PW, Jones PA, Liang G. DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter. Hum Mol Genet. 2011;20(22):4299-310.
301. Aran D, Toperoff G, Rosenberg M, Hellman A. Replication timing-related and gene body-specific methylation of active human genes. Hum Mol Genet. 2011;20(4):670-80.
302. Aoki A, Suetake I, Miyagawa J, Fujio T, Chijiwa T, Sasaki H, et al. Enzymatic properties of de novo-type mouse DNA (cytosine-5) methyltransferases. Nucleic Acids Res. 2001;29(17):3506-12.
303. Gu T, Lin X, Cullen SM, Luo M, Jeong M, Estecio M, et al. DNMT3A and TET1 cooperate to regulate promoter epigenetic landscapes in mouse embryonic stem cells. Genome Biol. 2018;19(1):88.
304. Kareta MS, Botello ZM, Ennis JJ, Chou C, Chedin F. Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. J Biol Chem. 2006;281(36):25893-902.
305. Wong KK, Lawrie CH, Green TM. Oncogenic Roles and Inhibitors of DNMT1, DNMT3A, and DNMT3B in Acute Myeloid Leukaemia. Biomark Insights. 2019;14:1177271919846454.

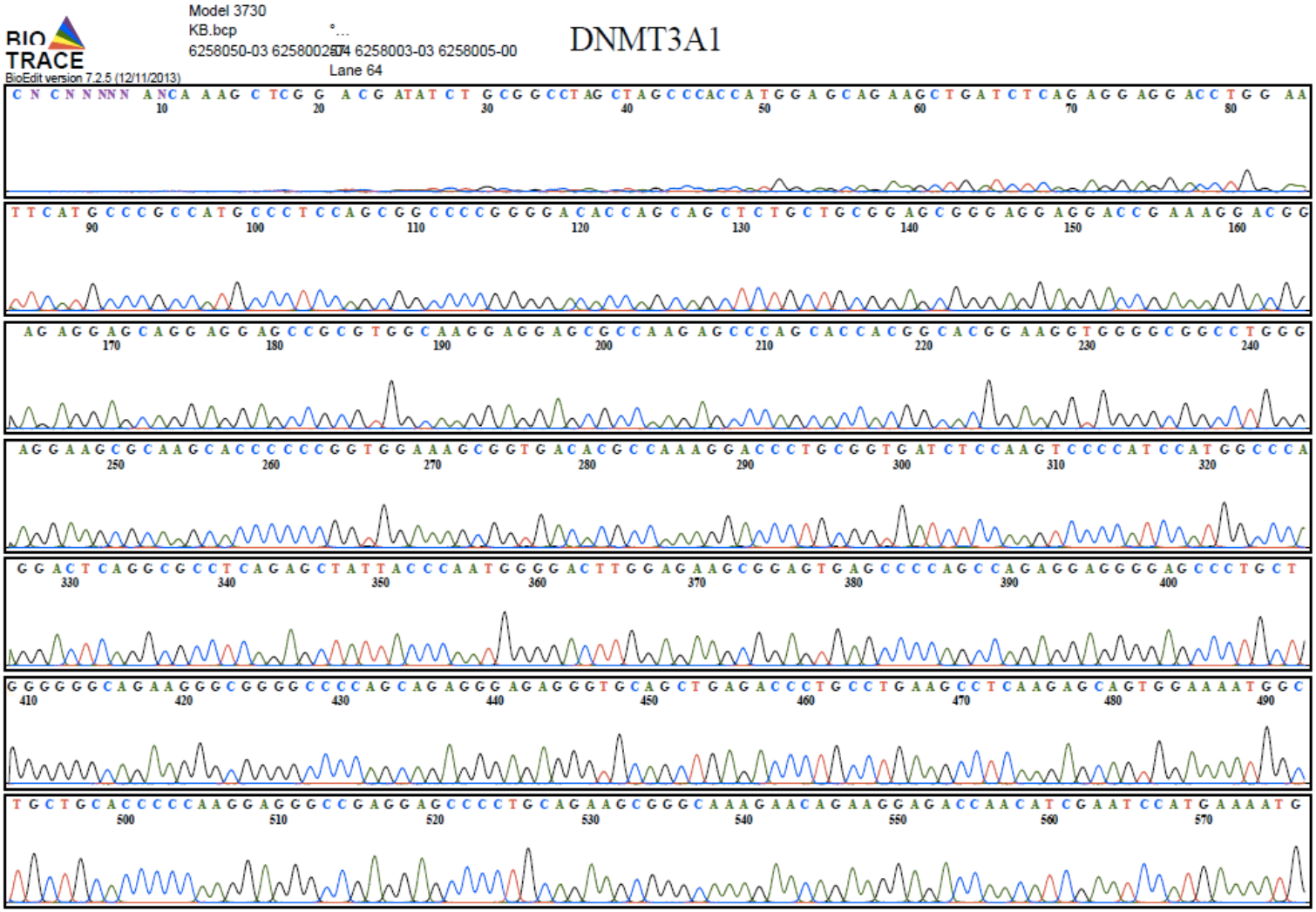
306. Honeywell RJ, Sarkisjan D, Kristensen MH, de Klerk DJ, Peters GJ. DNA methyltransferases expression in normal tissues and various human cancer cell lines, xenografts and tumors. Nucleosides Nucleotides Nucleic Acids. 2018;37(12):696-708.
307. Girault I, Tozlu S, Lidereau R, Bieche I. Expression analysis of DNA methyltransferases 1, 3A, and 3B in sporadic breast carcinomas. Clin Cancer Res. 2003;9(12):4415-22.
308. Dennis KL, Blatner NR, Gounari F, Khazaie K. Current status of interleukin-10 and regulatory T-cells in cancer. Curr Opin Oncol. 2013;25(6):637-45.
309. Cho HJ, Kim JT, Lee SJ, Hwang YS, Park SY, Kim BY, et al. Protein phosphatase 1B dephosphorylates Rho guanine nucleotide dissociation inhibitor 1 and suppresses cancer cell migration and invasion. Cancer Lett. 2018;417:141-51.
310. Gkountakos A, Sartori G, Falcone I, Piro G, Ciuffreda L, Carbone C, et al. PTEN in Lung Cancer: Dealing with the Problem, Building on New Knowledge and Turning the Game Around. Cancers (Basel). 2019;11(8).
311. Greally M, Kelly CM, Cercek A. HER2: An emerging target in colorectal cancer. Curr Probl Cancer. 2018;42(6):560-71.
312. Elster N, Collins DM, Toomey S, Crown J, Eustace AJ, Hennessy BT. HER2-family signalling mechanisms, clinical implications and targeting in breast cancer. Breast Cancer Res Treat. 2015;149(1):5-15.
313. Tai YL, Chen LC, Shen TL. Emerging roles of focal adhesion kinase in cancer. Biomed Res Int. 2015;2015:690690.
314. Jiang N, Dai Q, Su X, Fu J, Feng X, Peng J. Role of PI3K/AKT pathway in cancer: the framework of malignant behavior. Mol Biol Rep. 2020;47(6):4587-629.
315. Azimi I, Roberts-Thomson SJ, Monteith GR. Calcium influx pathways in breast cancer: opportunities for pharmacological intervention. Br J Pharmacol. 2014;171(4):945-60.
316. Zhou G, Yang J, Song P. Correlation of ERK/MAPK signaling pathway with proliferation and apoptosis of colon cancer cells. Oncol Lett. 2019;17(2):2266-70.
317. Toyoda S, Kawaguchi M, Kobayashi T, Tarusawa E, Toyama T, Okano M, et al. Developmental epigenetic modification regulates stochastic expression of clustered protocadherin genes, generating single neuron diversity. Neuron. 2014;82(1):94-108.
318. Shivapurkar N, Poirier LA. Tissue levels of S-adenosylmethionine and S-adenosylhomocysteine in rats fed methyl-deficient, amino acid-defined diets for one to five weeks. Carcinogenesis. 1983;4(8):1051-7.
319. Pogribny IP, Ross SA, Wise C, Pogribna M, Jones EA, Tryndyak VP, et al. Irreversible global DNA hypomethylation as a key step in hepatocarcinogenesis induced by dietary methyl deficiency. Mutat Res. 2006;593(1-2):80-7.
320. Sae-Lee C, Corsi S, Barrow TM, Kuhnle GGC, Bollati V, Mathers JC, et al. Dietary Intervention Modifies DNA Methylation Age Assessed by the Epigenetic Clock. Mol Nutr Food Res. 2018;62(23):e1800092.
321. Esteller M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. Oncogene. 2002;21(35):5427-40.
322. Nandakumar V, Vaid M, Tollefsbol TO, Katiyar SK. Aberrant DNA hypermethylation patterns lead to transcriptional silencing of tumor suppressor genes in UVB-exposed skin and UVB-induced skin tumors of mice. Carcinogenesis. 2011;32(4):597-604.
323. Pellerito C, Morana O, Ferrante F, Calvaruso G, Notaro A, Sabella S, et al. Synthesis, chemical characterization, computational studies and biological activity of new DNA methyltransferases (DNMTs) specific inhibitor. Epigenetic regulation as a new and potential approach to cancer therapy. J Inorg Biochem. 2015;150:18-27.
324. Du GJ, Zhang Z, Wen XD, Yu C, Calway T, Yuan CS, et al. Epigallocatechin Gallate (EGCG) is the most effective cancer chemopreventive polyphenol in green tea. Nutrients. 2012;4(11):1679-91.
325. Khan MA, Hussain A, Sundaram MK, Alalami U, Gunasekera D, Ramesh L, et al. (-)-Epigallocatechin-3-gallate reverses the expression of various tumor-suppressor genes by inhibiting DNA methyltransferases and histone deacetylases in human cervical cancer cells. Oncol Rep. 2015;33(4):1976-84.
326. Detich N, Theberge J, Szyf M. Promoter-specific activation and demethylation by MBD2/demethylase. J Biol Chem. 2002;277(39):35791-4.

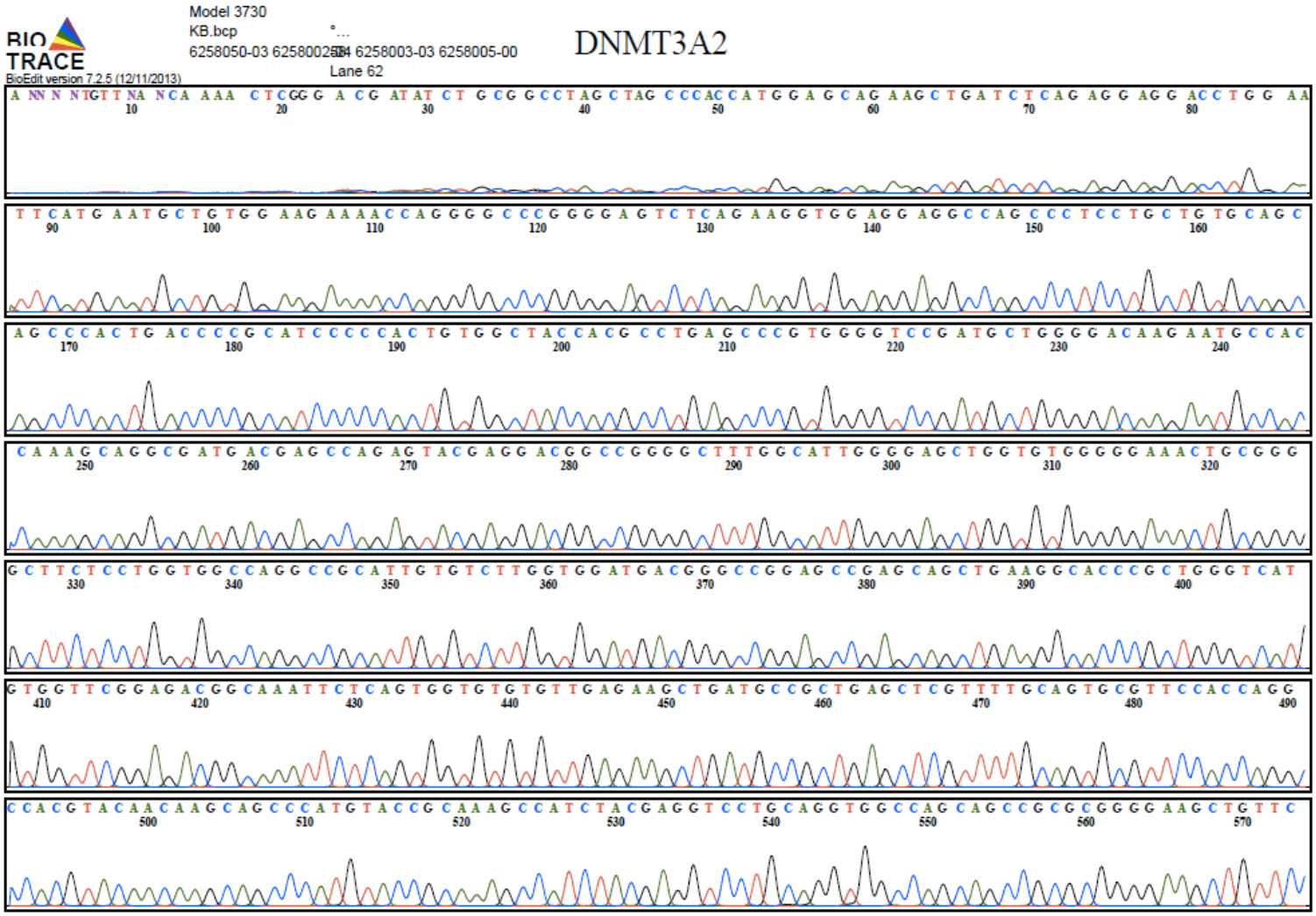
327. Yin R, Mao SQ, Zhao B, Chong Z, Yang Y, Zhao C, et al. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. J Am Chem Soc. 2013;135(28):10396-403.
328. Minor EA, Court BL, Young JI, Wang G. Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. J Biol Chem. 2013;288(19):13669-74.
329. Loenarz C, Schofield CJ. Expanding chemical biology of 2-oxoglutarate oxygenases. Nat Chem Biol. 2008;4(3):152-6.
330. Stadtfeld M, Apostolou E, Ferrari F, Choi J, Walsh RM, Chen T, et al. Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. Nat Genet. 2012;44(4):398-405, S1-2.
331. Agathocleous M, Meacham CE, Burgess RJ, Piskounova E, Zhao Z, Crane GM, et al. Ascorbate regulates haematopoietic stem cell function and leukaemogenesis. Nature. 2017;549(7673):476-81.
332. Azqueta A, Costa S, Lorenzo Y, Bastani NE, Collins AR. Vitamin C in cultured human (HeLa) cells: lack of effect on DNA protection and repair. Nutrients. 2013;5(4):1200-17.
333. Young JI, Zuchner S, Wang G. Regulation of the Epigenome by Vitamin C. Annu Rev Nutr. 2015;35:545-64.
334. Meeran SM, Ahmed A, Tollefsbol TO. Epigenetic targets of bioactive dietary components for cancer prevention and therapy. Clin Epigenetics. 2010;1(3-4):101-16.
335. Link A, Balaguer F, Shen Y, Lozano JJ, Leung HC, Boland CR, et al. Curcumin modulates DNA methylation in colorectal cancer cells. PLoS One. 2013;8(2):e57709.
336. Maugeri A, Mazzone MG, Giuliano F, Vinciguerra M, Basile G, Barchitta M, et al. Curcumin Modulates DNA Methyltransferase Functions in a Cellular Model of Diabetic Retinopathy. Oxid Med Cell Longev. 2018;2018:5407482.
337. Howell JC, Chun E, Farrell AN, Hur EY, Caroti CM, Iuvone PM, et al. Global microRNA expression profiling: curcumin (diferuloylmethane) alters oxidative stress-responsive microRNAs in human ARPE-19 cells. Mol Vis. 2013;19:544-60.
338. Lyko F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. Nat Rev Genet. 2018;19(2):81-92.
339. Spada F, Haemmer A, Kuch D, Rothbauer U, Schermelleh L, Kremmer E, et al. DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. J Cell Biol. 2007;176(5):565-71.
340. Cervera L, Gutierrez S, Godia F, Segura MM. Optimization of HEK 293 cell growth by addition of non-animal derived components using design of experiments. BMC Proc. 2011;5 Suppl 8:P126.
341. Ogura M, Morishima Y, Ohno R, Kato Y, Hirabayashi N, Nagura H, et al. Establishment of a novel human megakaryoblastic leukemia cell line, MEG-01, with positive Philadelphia chromosome. Blood. 1985;66(6):1384-92.
342. Shao G, Zhang R, Zhang S, Jiang S, Liu Y, Zhang W, et al. Splice variants DNMT3B4 and DNMT3B7 overexpression inhibit cell proliferation in 293A cell line. In Vitro Cell Dev Biol Anim. 2013;49(5):386-94.
343. Zwergel C, Valente S, Mai A. DNA Methyltransferases Inhibitors from Natural Sources. Curr Top Med Chem. 2016;16(7):680-96.
344. Chedin F. The DNMT3 family of mammalian de novo DNA methyltransferases. Prog Mol Biol Transl Sci. 2011;101:255-85.
345. Simmons RK, Stringfellow SA, Glover ME, Wagle AA, Clinton SM. DNA methylation markers in the postnatal developing rat brain. Brain Res. 2013;1533:26-36.
346. Park HJ, Yu E, Shim YH. DNA methyltransferase expression and DNA hypermethylation in human hepatocellular carcinoma. Cancer Lett. 2006;233(2):271-8.
347. Mizuno S, Chijiwa T, Okamura T, Akashi K, Fukumaki Y, Niho Y, et al. Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. Blood. 2001;97(5):1172-9.
348. Robertson KD, Uzvolgyi E, Liang G, Talmadge C, Sumegi J, Gonzales FA, et al. The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. Nucleic Acids Res. 1999;27(11):2291-8.

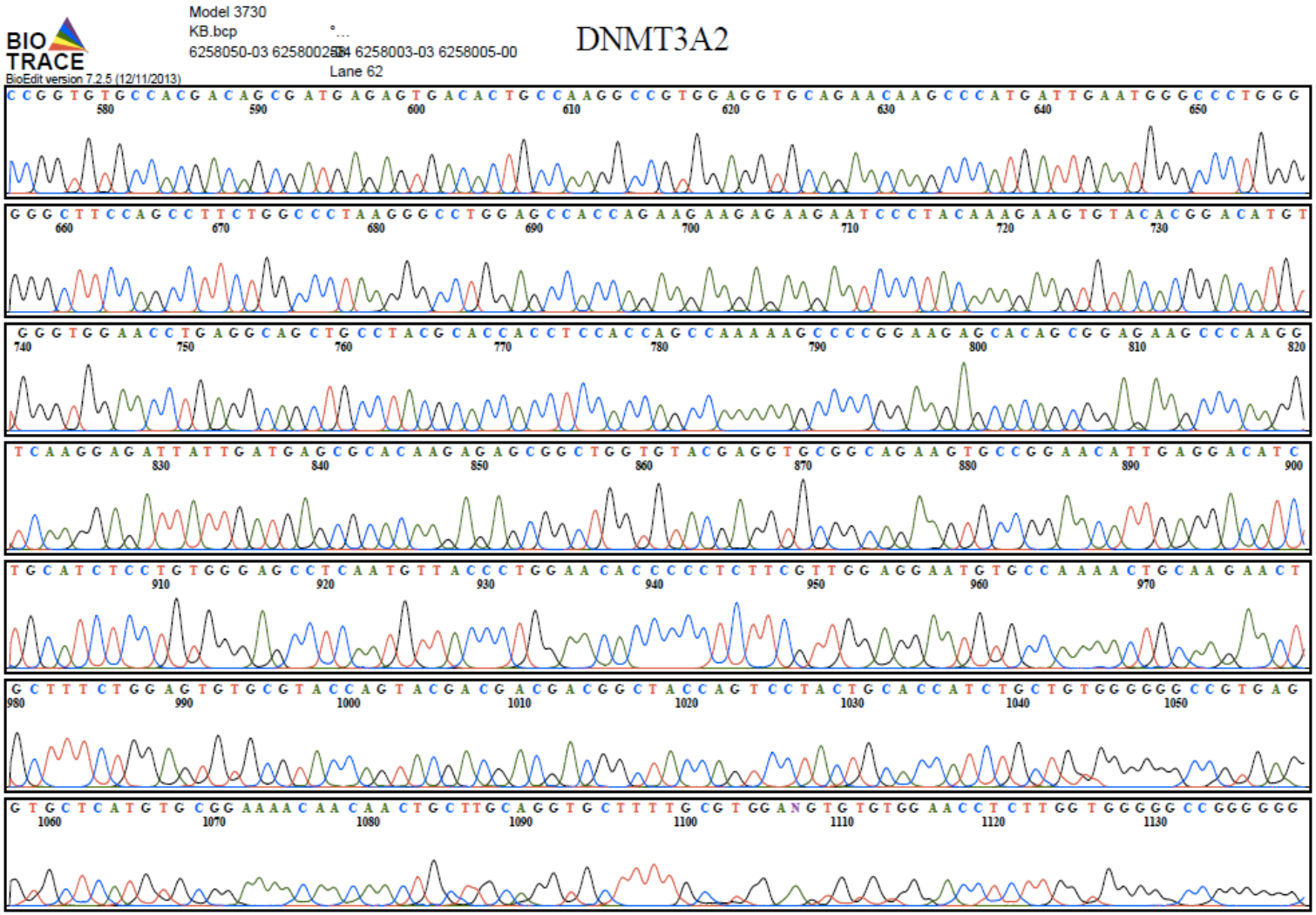
349. Chen Y, Liao LD, Wu ZY, Yang Q, Guo JC, He JZ, et al. Identification of key genes by integrating DNA methylation and next-generation transcriptome sequencing for esophageal squamous cell carcinoma. Aging (Albany NY). 2020;12.
350. Liu K, Wang YF, Cantemir C, Muller MT. Endogenous assays of DNA methyltransferases: Evidence for differential activities of DNMT1, DNMT2, and DNMT3 in mammalian cells in vivo. Mol Cell Biol. 2003;23(8):2709-19.

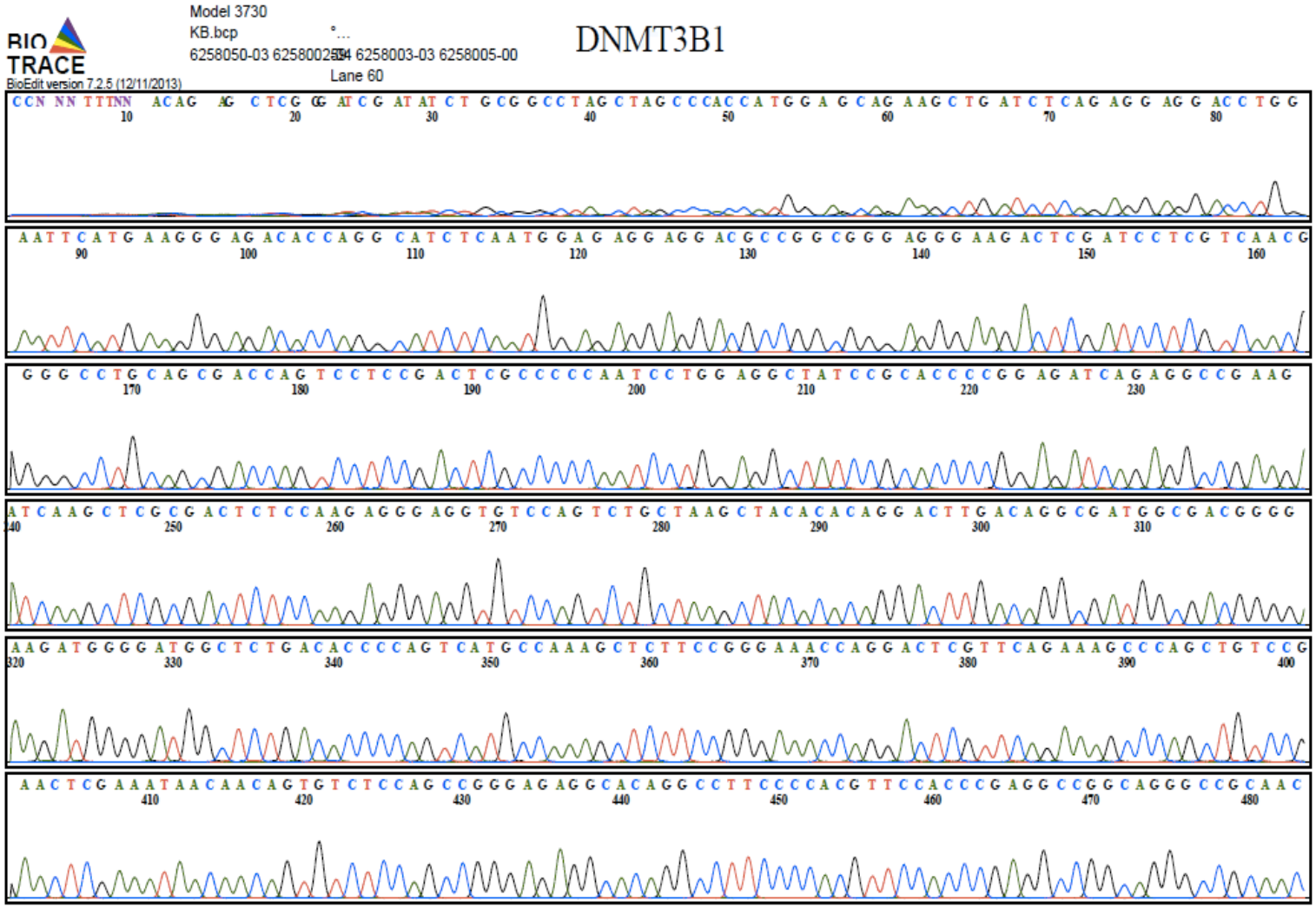
Appendix A: Sequencing by Sanger analysis of each DNMT isoform and alignment analysis

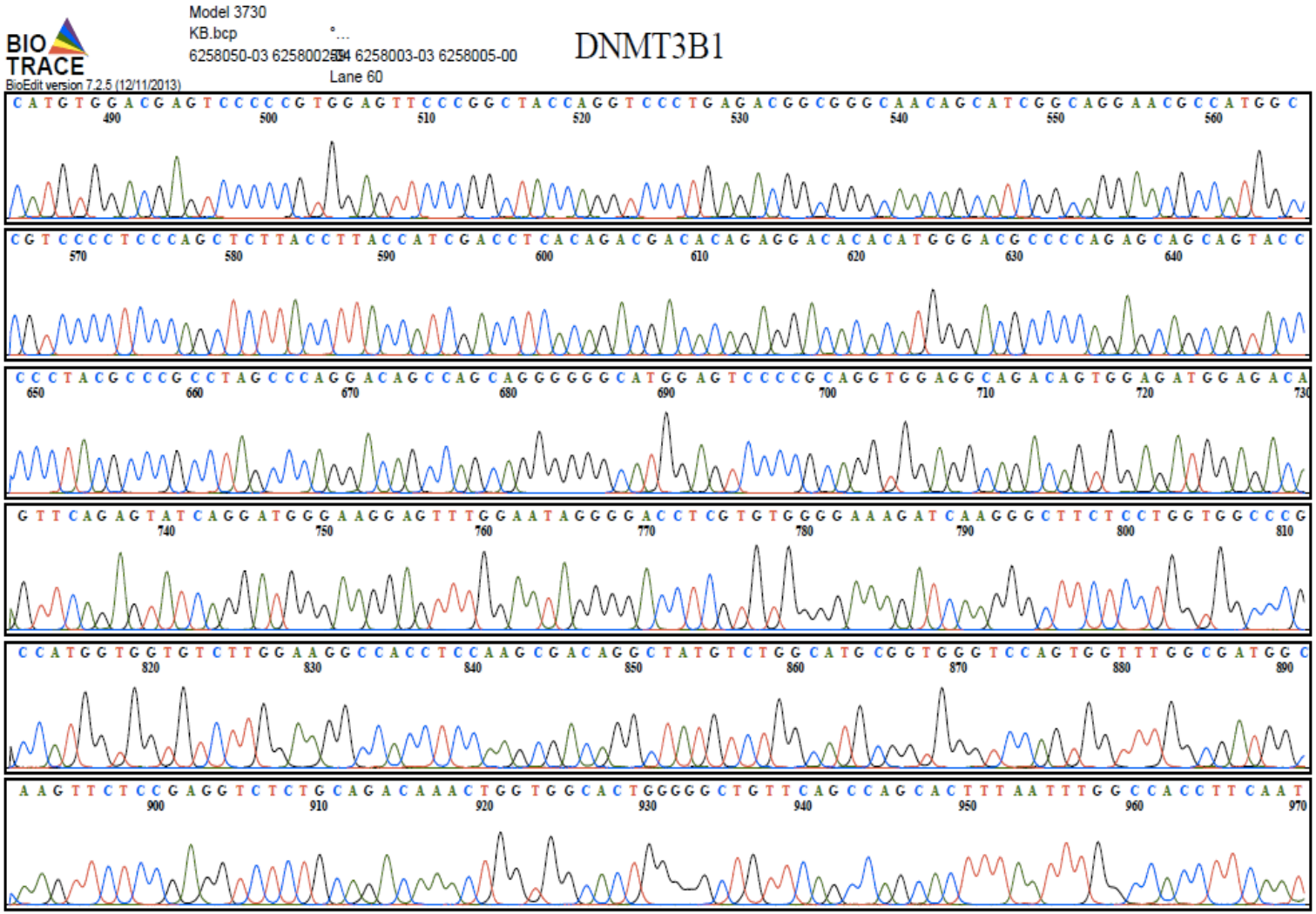


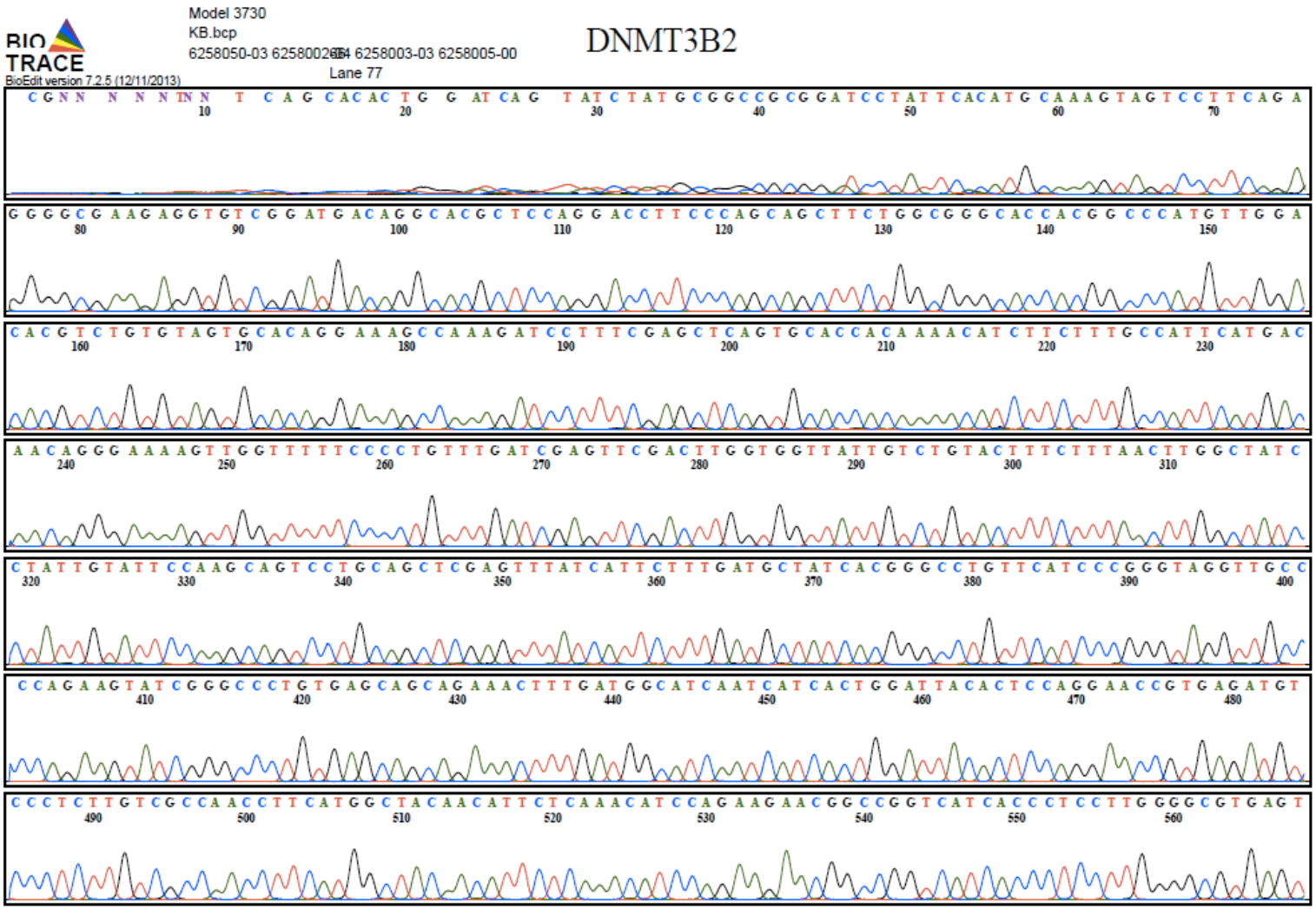

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258003-03 6258005-00
Lane 64
DNMT3A1

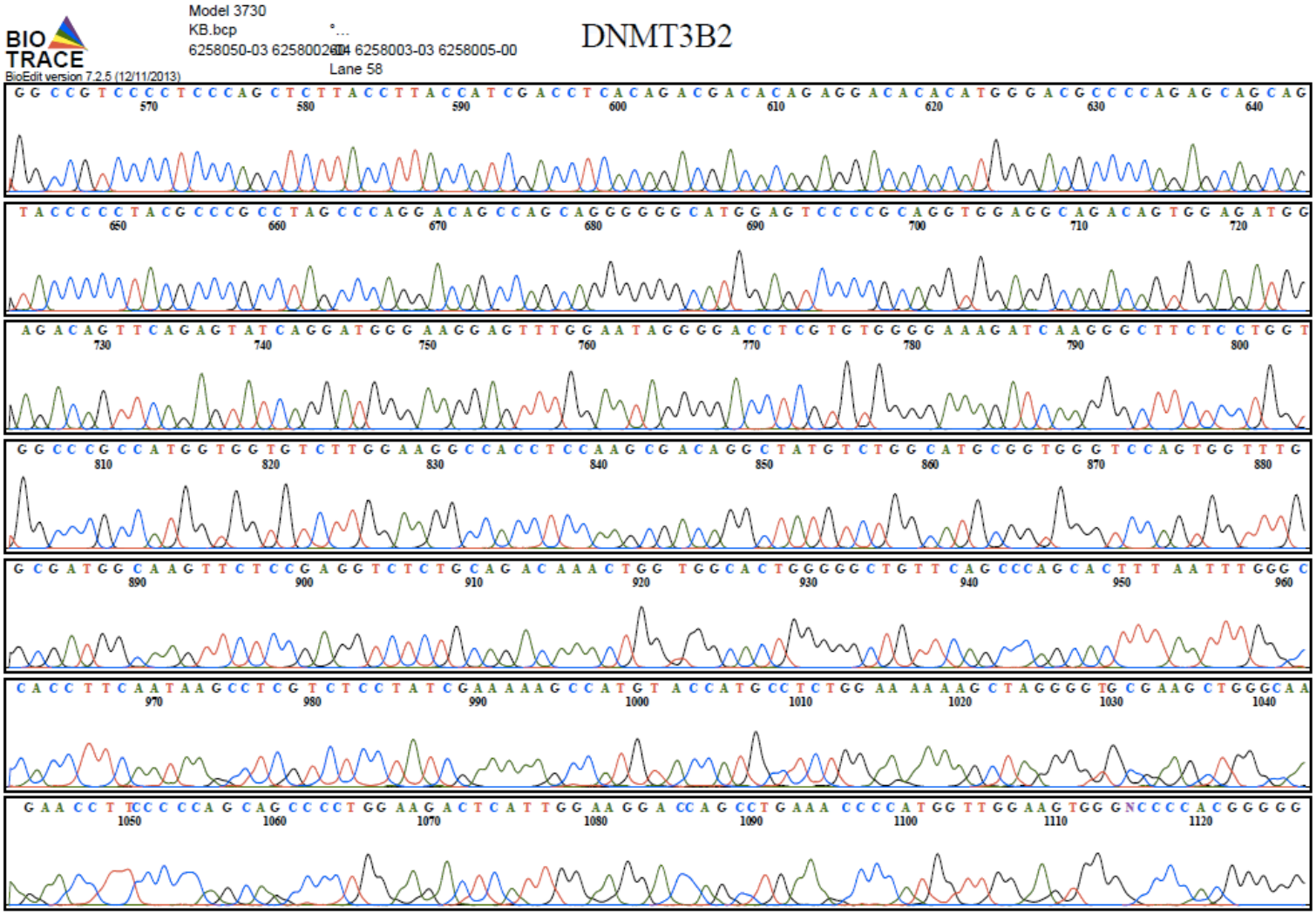

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-03 6258003-03 6258005-00
Lane 62
DNMT3A2

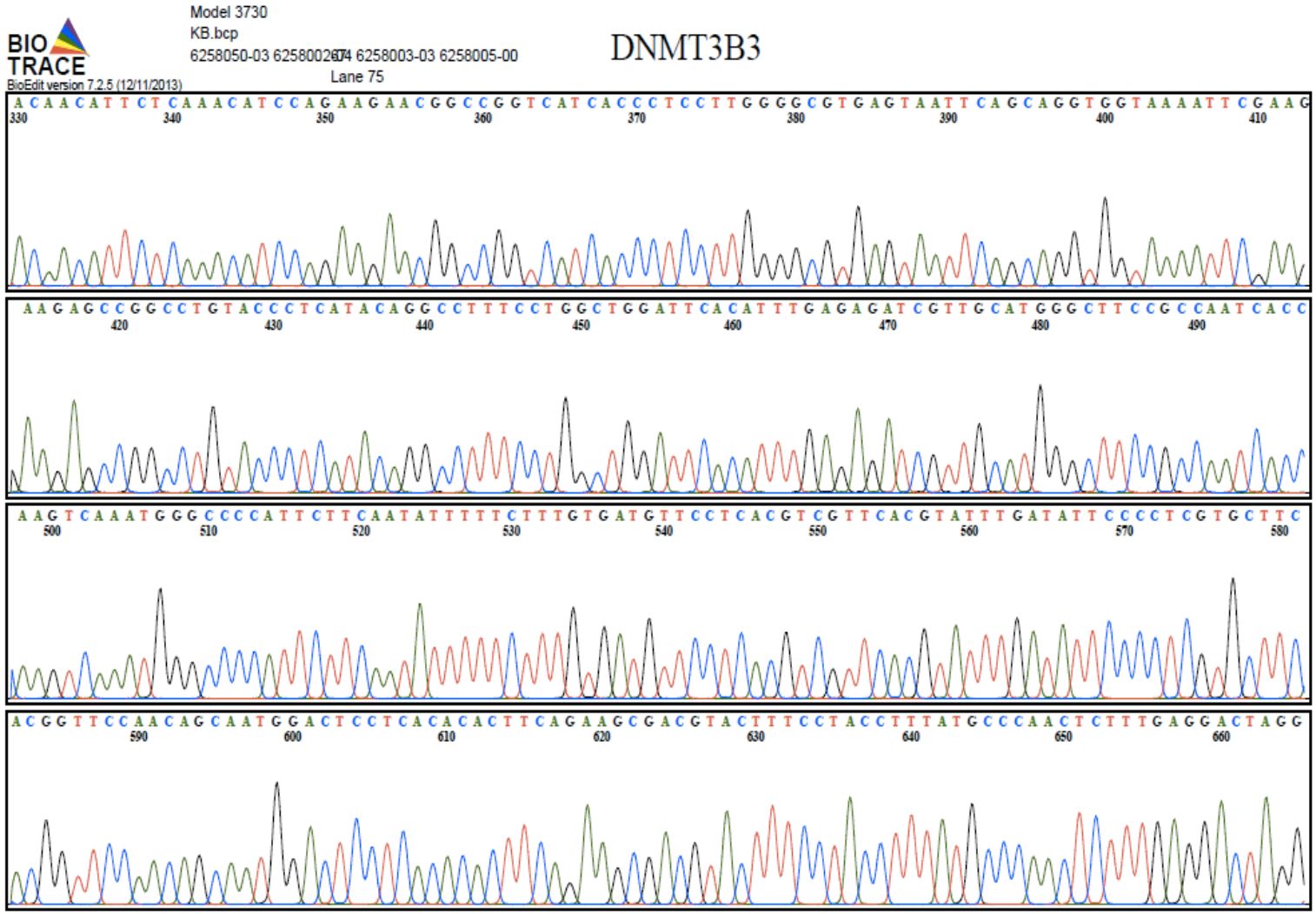

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258003-03 6258005-00
Lane 62
DNMT3A2

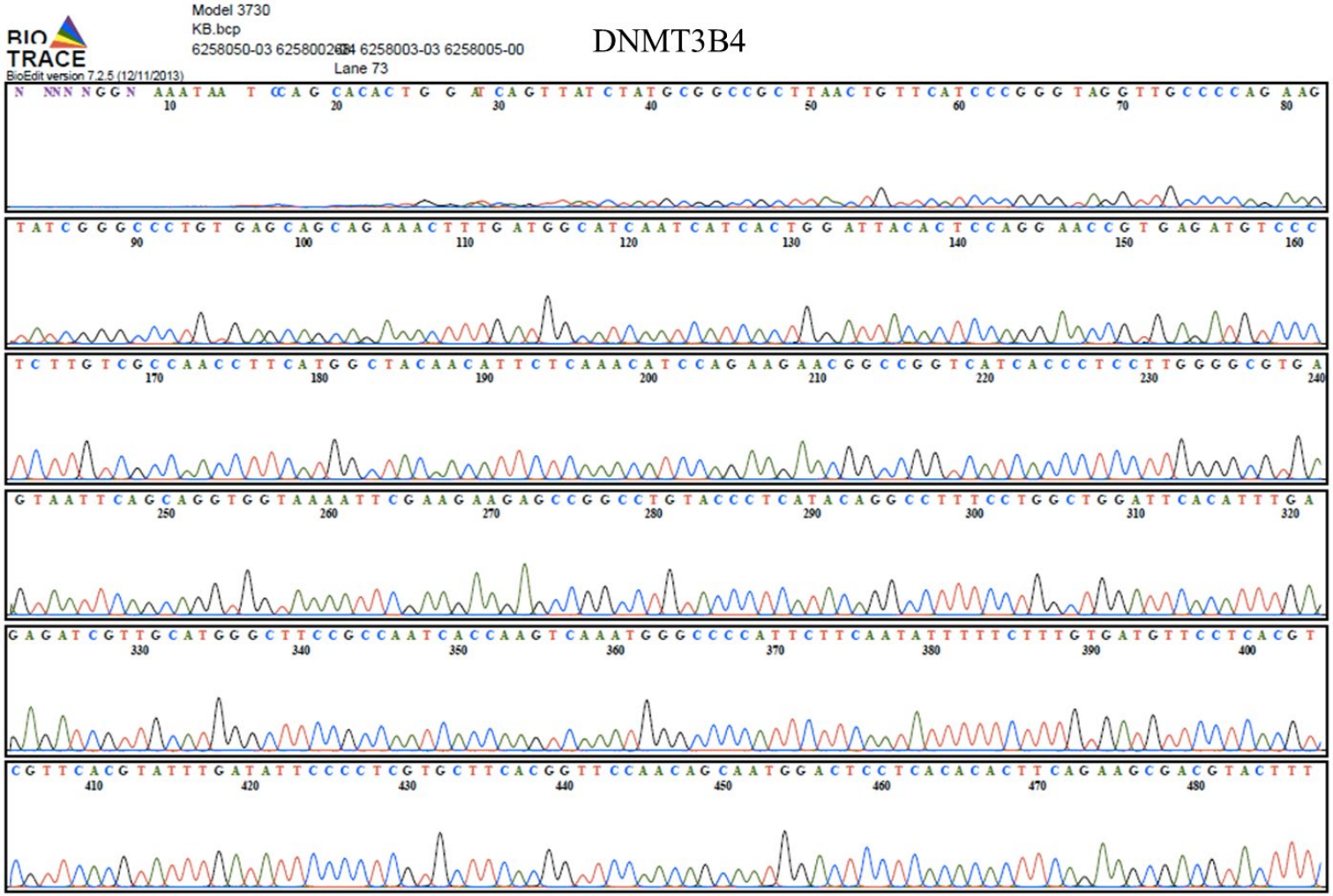

BIO
TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258003-03 6258005-00
Lane 60
DNMT3B1

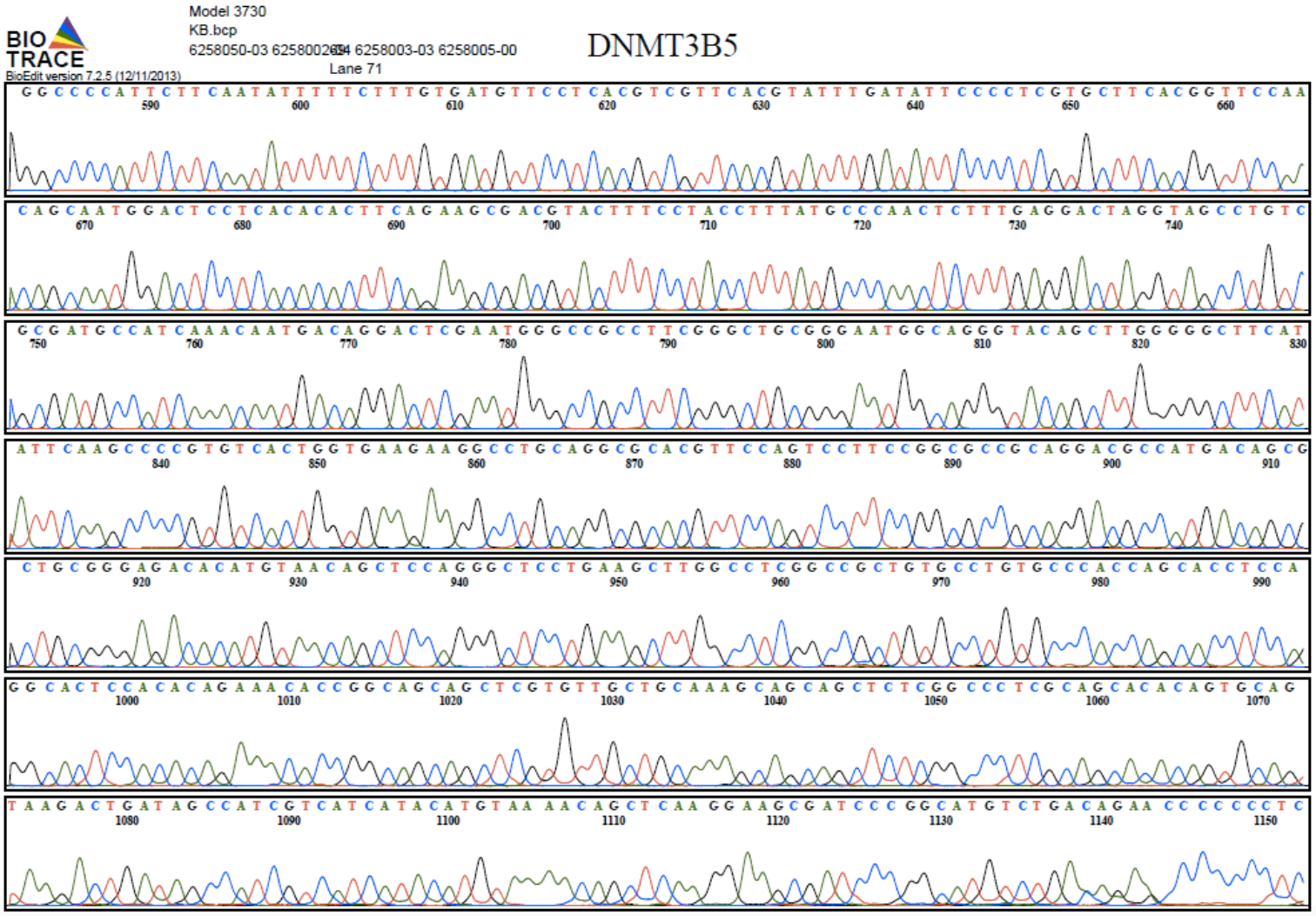

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258003-03 6258005-00
Lane 60
DNMT3B1

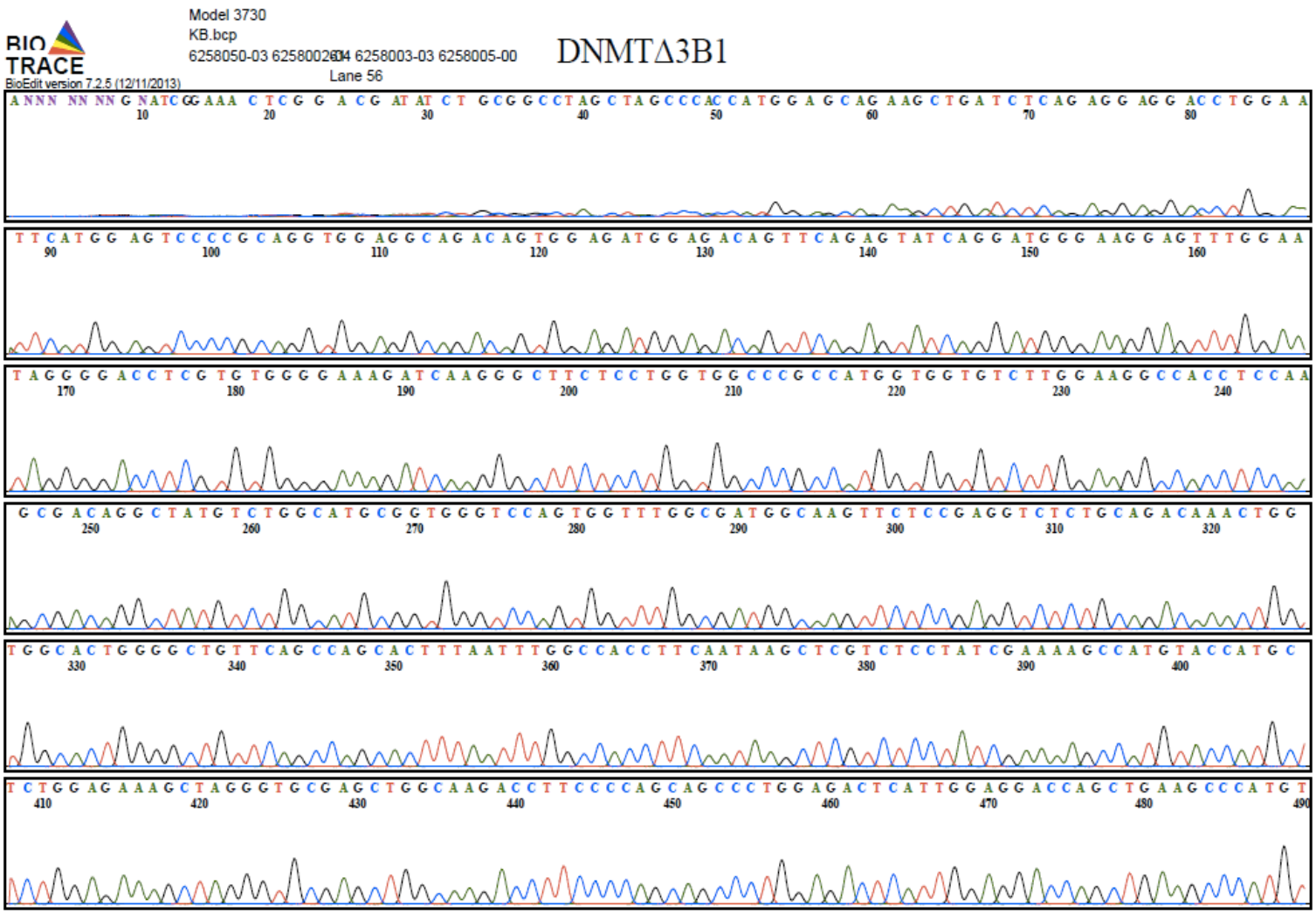

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 77
DNMT3B2

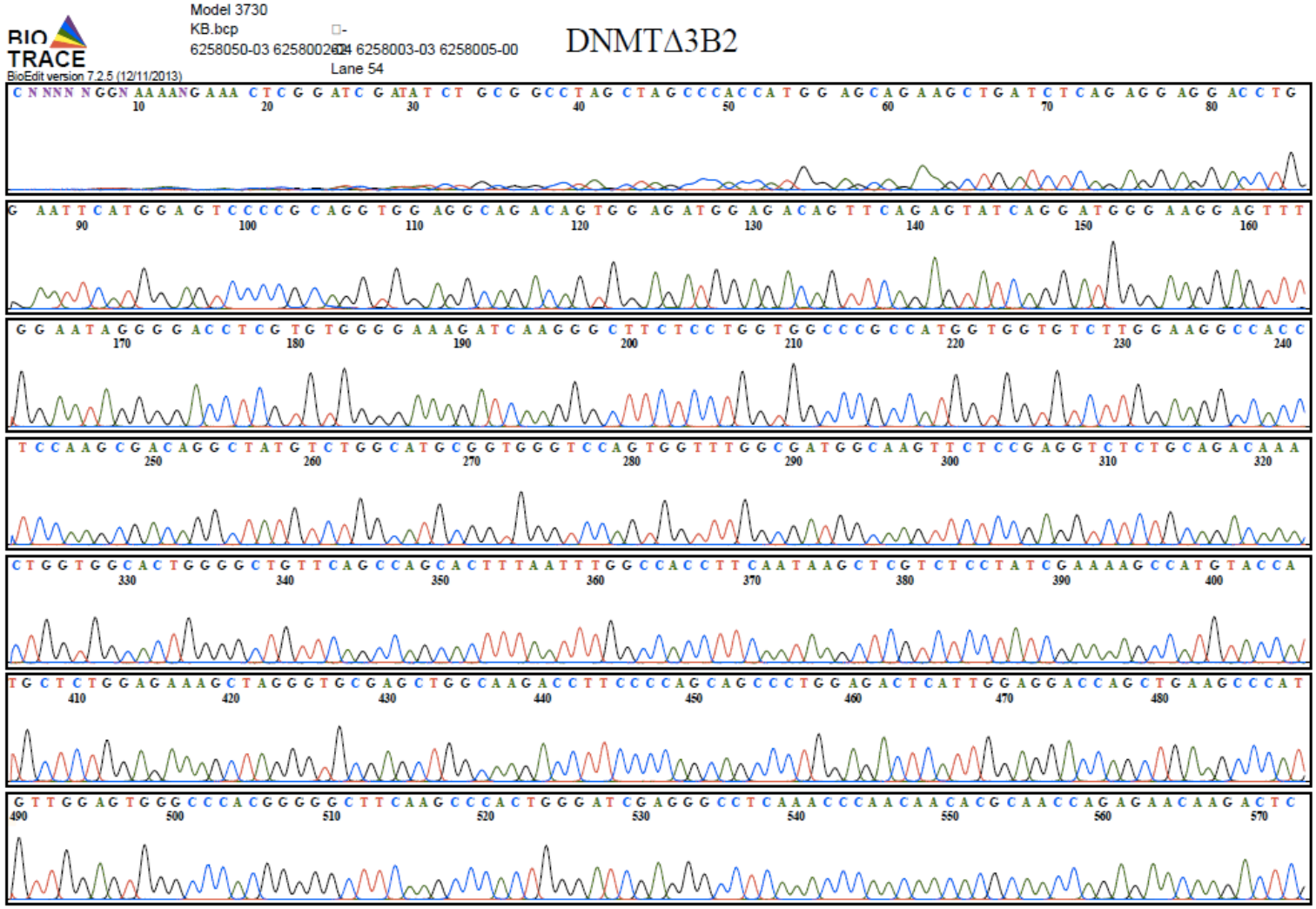

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002604 6258003-03 6258005-00
Lane 58
DNMT3B2

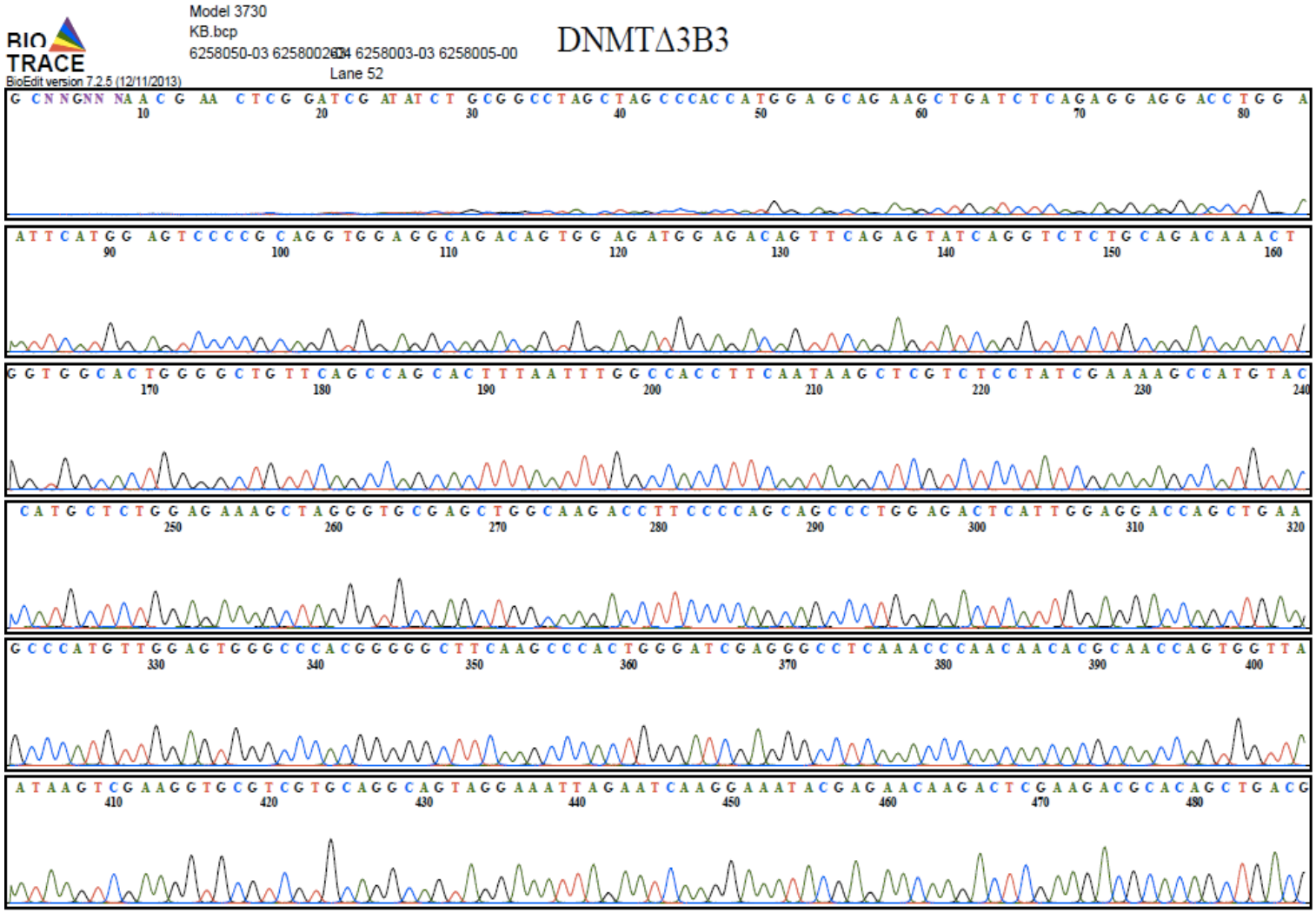

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002674 6258003-03 6258005-00
Lane 75
DNMT3B3

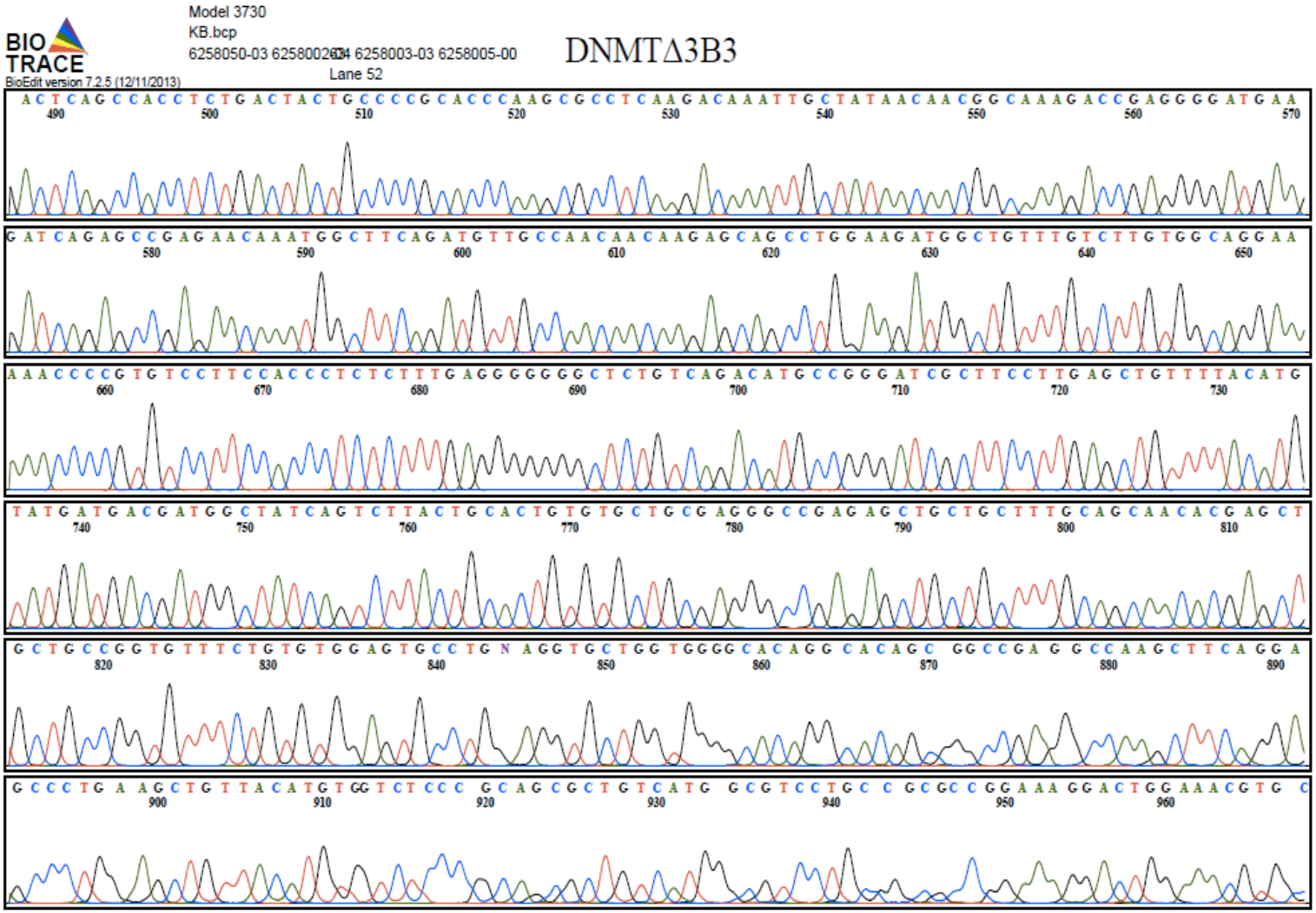

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 73
DNMT3B4

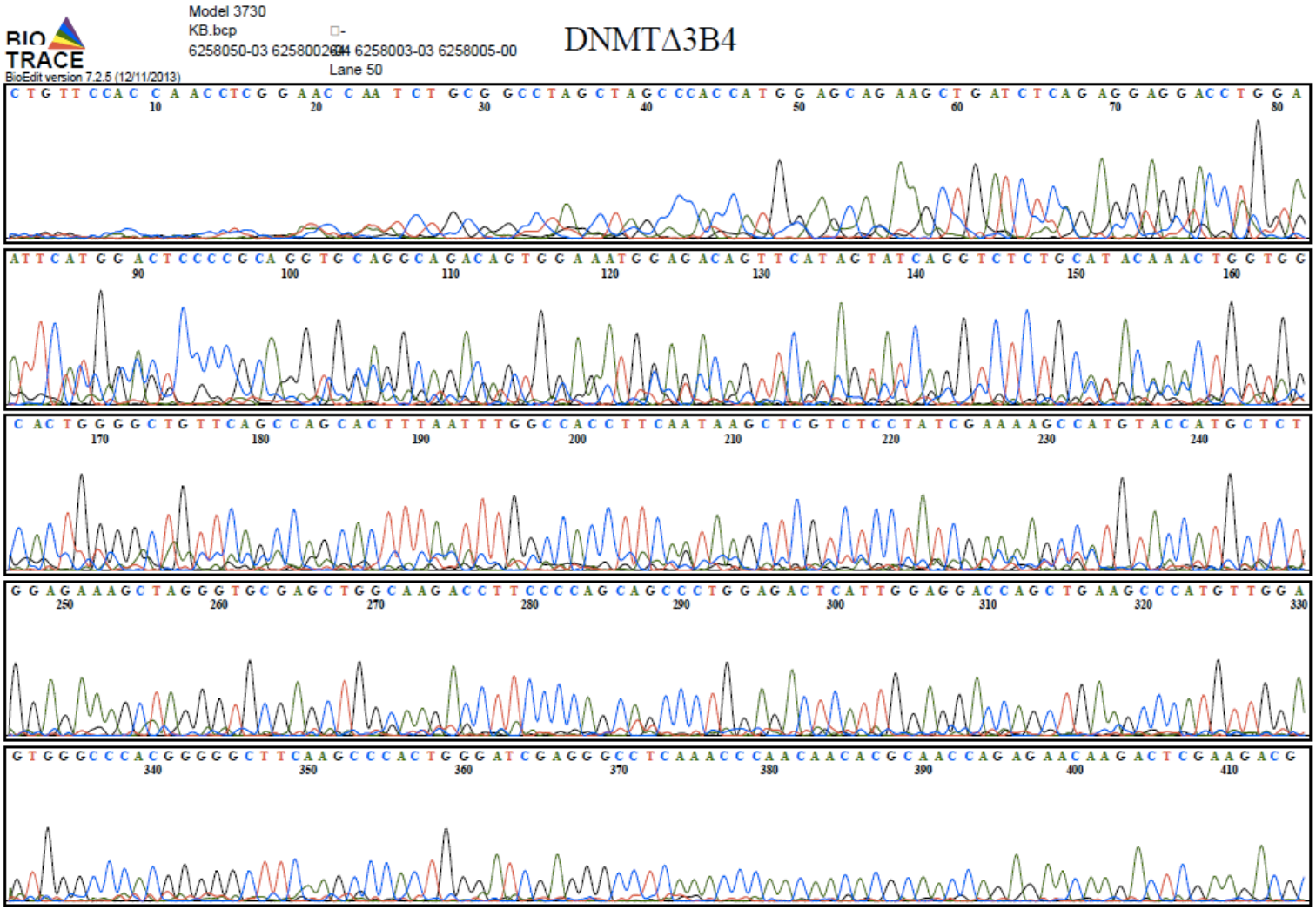

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002654 6258003-03 6258005-00
Lane 71
DNMT3B5

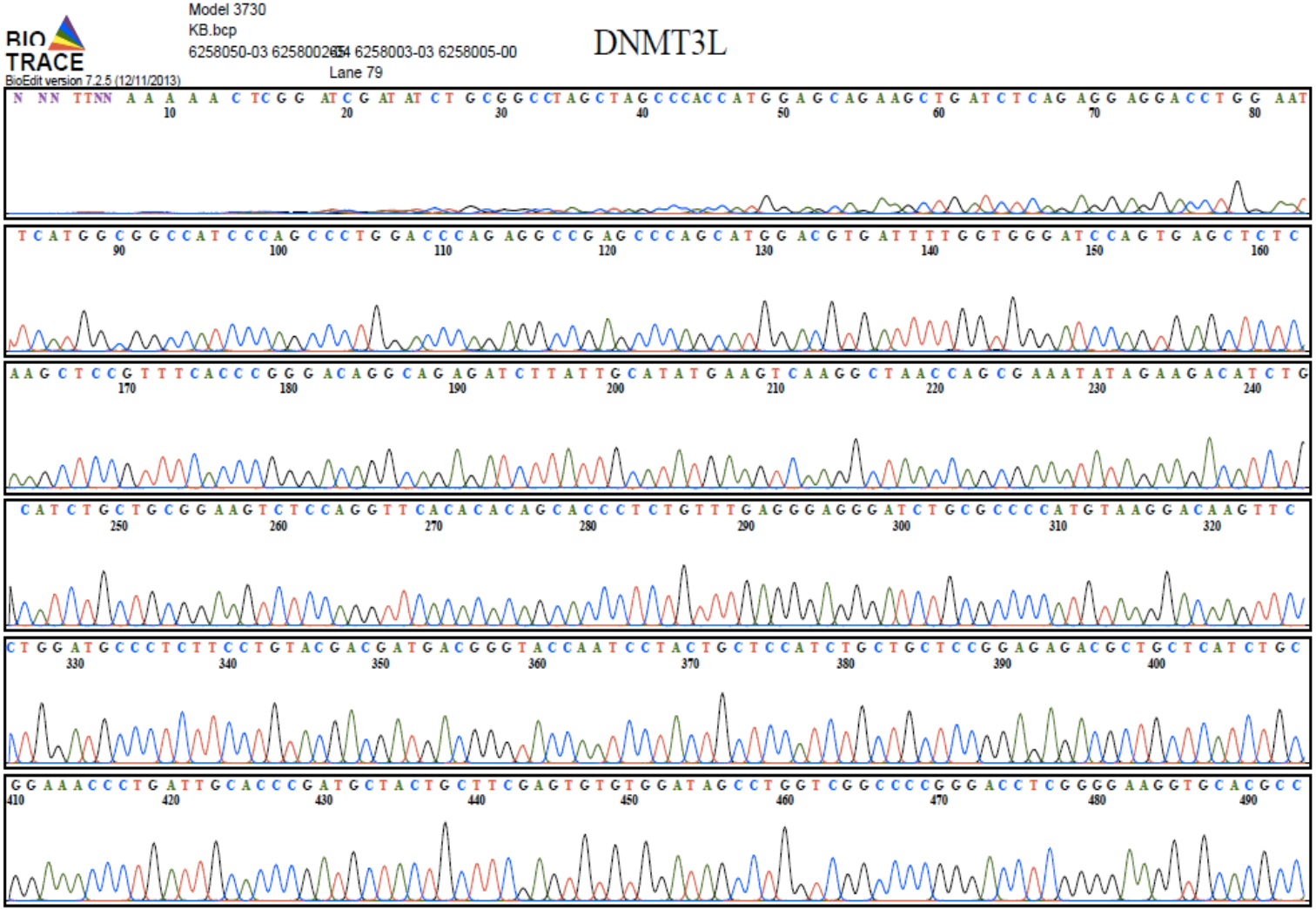

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002604 6258003-03 6258005-00
Lane 56
DNMTΔ3B1


Model 3730
KB.bcp
6258050-03 6258002624 6258003-03 6258005-00
Lane 54
DNMTΔ3B2
BioEdit version 7.2.5 (12/11/2013)


BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002624 6258003-03 6258005-00
Lane 52
DNMTΔ3B3


BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002634 6258003-03 6258005-00
Lane 52
DNMTΔ3B3


BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 50
DNMTΔ3B4


BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002654 6258003-03 6258005-00
Lane 79
DNMT3L

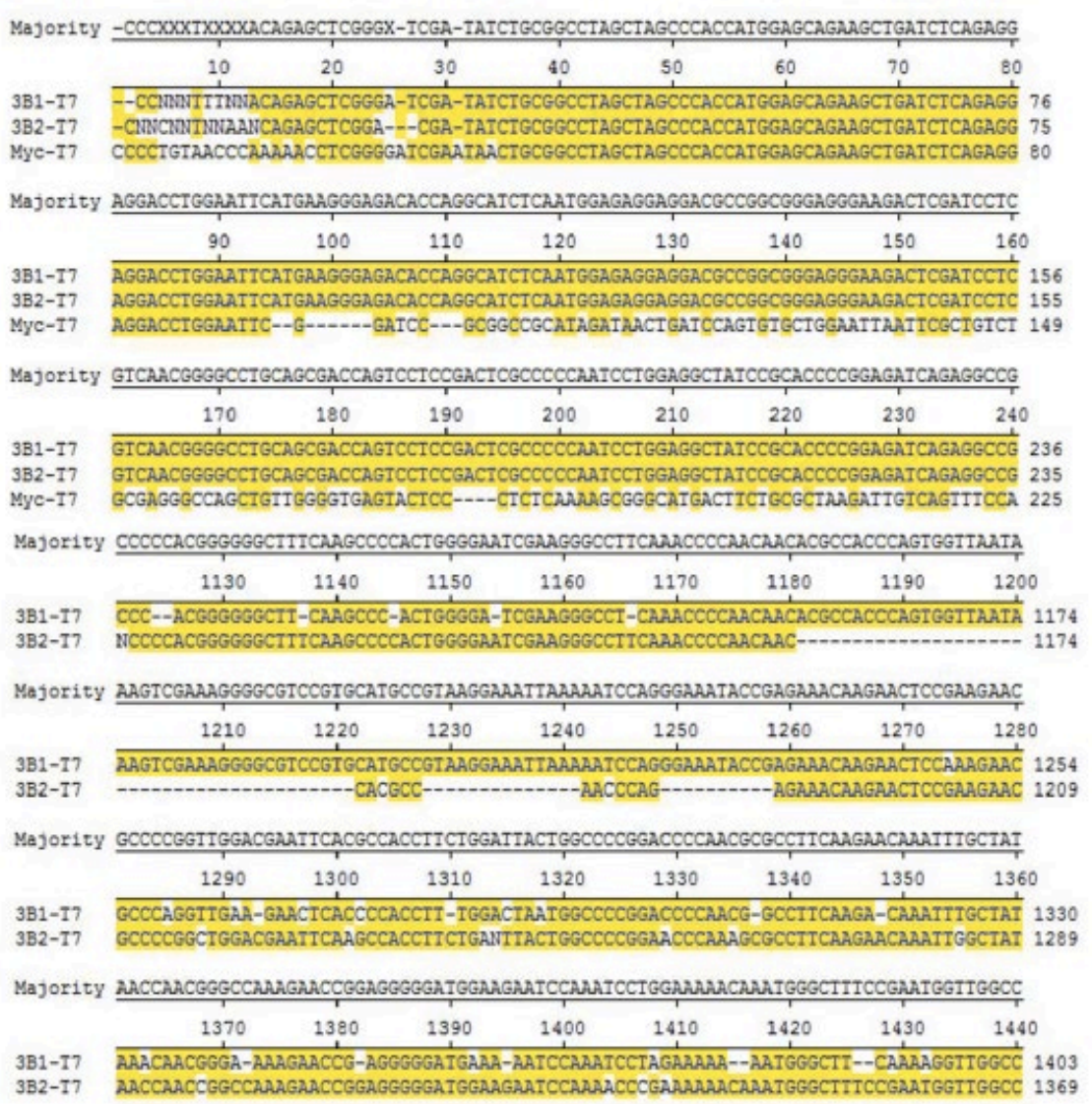


Figure A1 Alignment result between DNMT3B1 and DNMT3B2. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

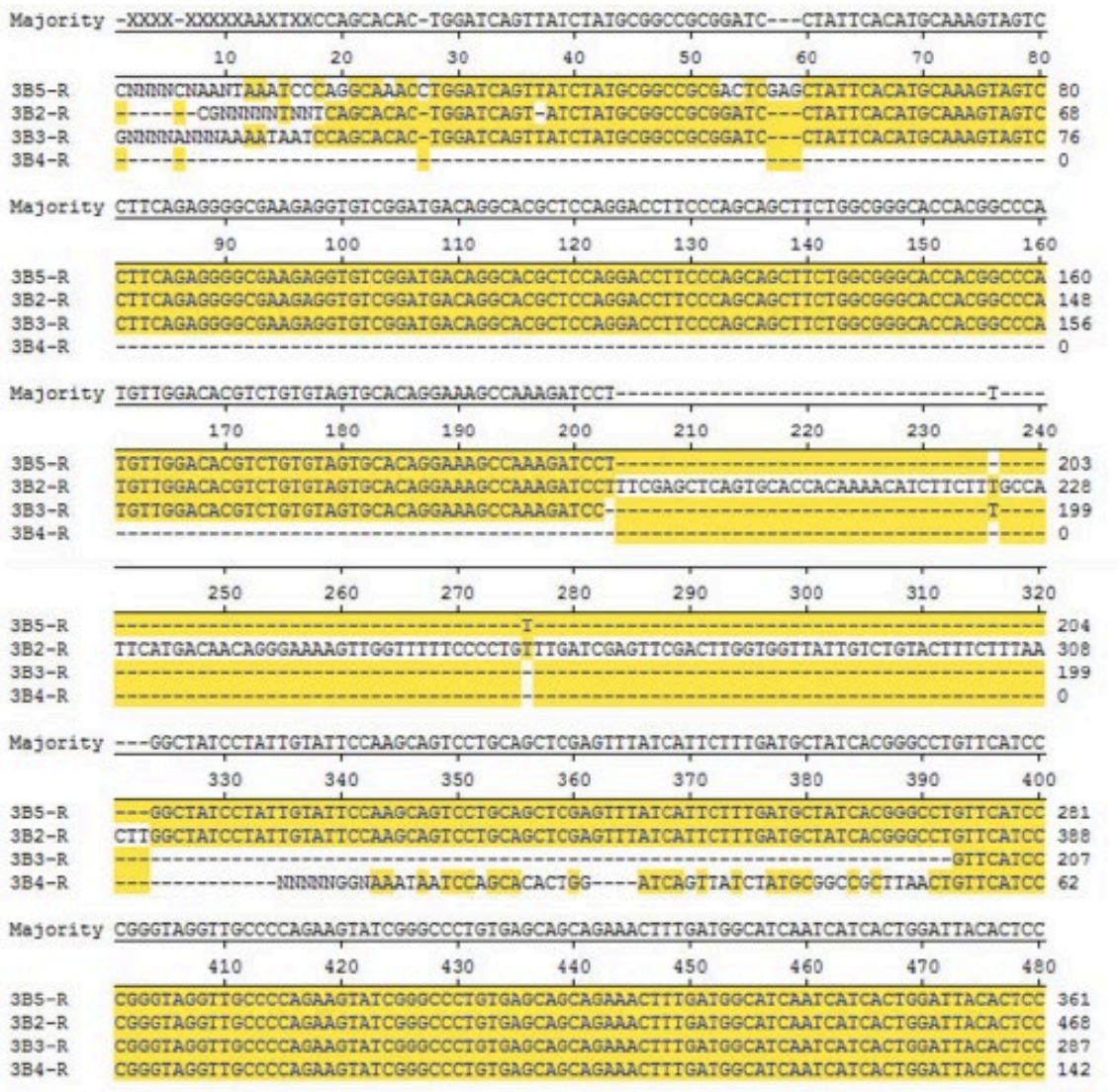


Figure A2 Alignment result among DNMT3B2, DNMT3B3, DNMT3B4 and DNMT3B5. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

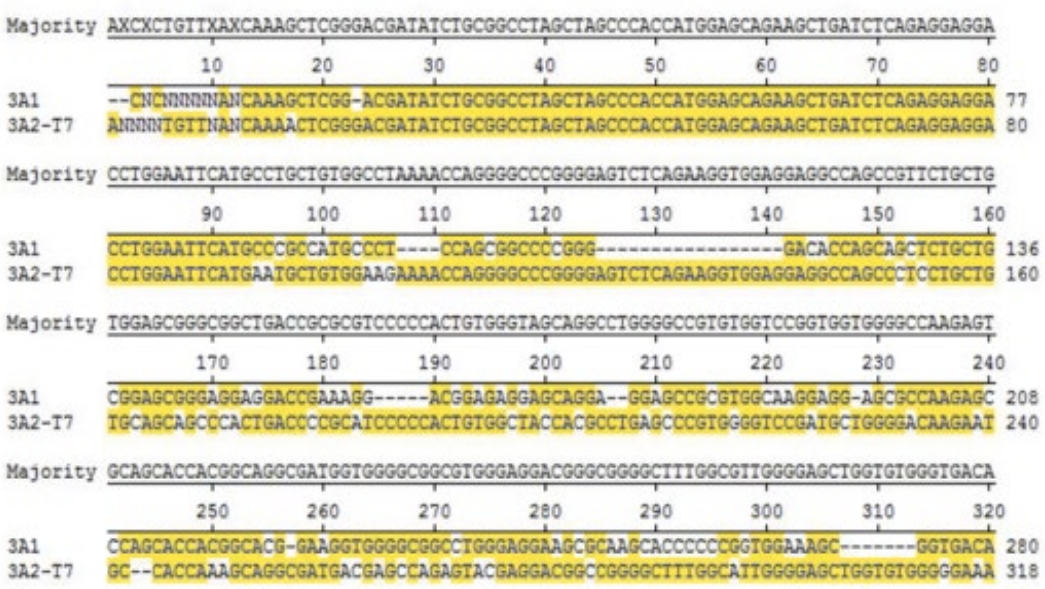


Figure A3 Alignment result between DNMT3A1 and DNMT3A2. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

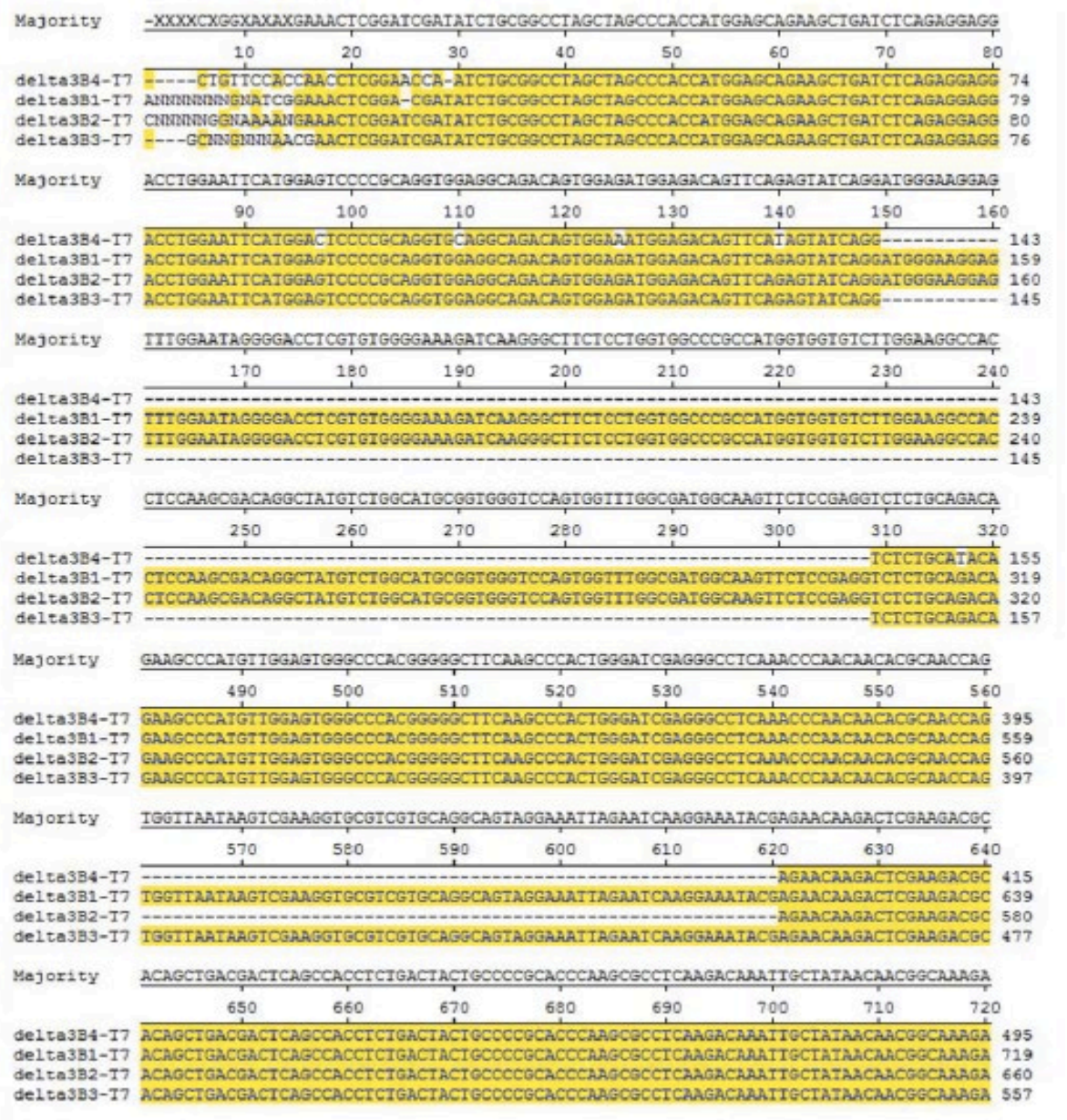


Figure A4 Alignment result among DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

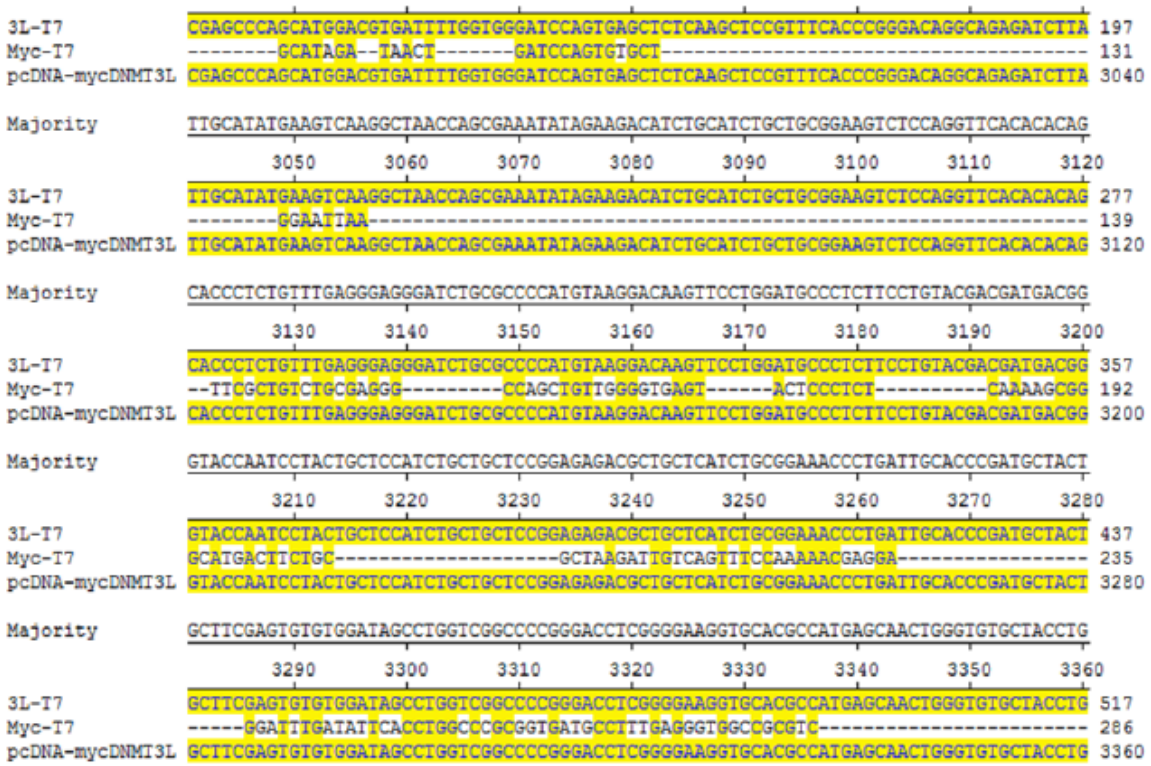


Figure A5 Alignment result between DNMT3L and pcDNA-MycDNMT3L (commercial plasmid). Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

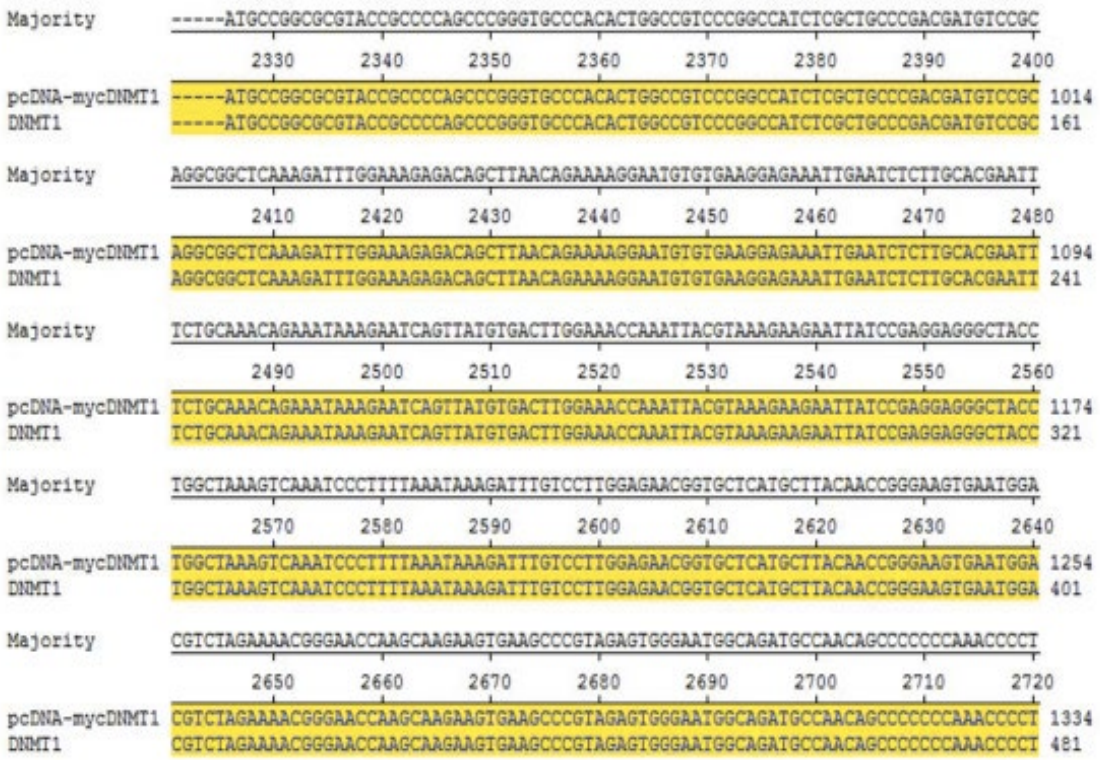


Figure A6 Alignment result between DNMT1 and pcDNA-MycDNMT1 (commercial plasmid). Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

DNMT3A1:

Homo sapiens DNA methyltransferase 3 alpha (DNMT3A), transcript variant 3, mRNA
Sequence ID: NM_022552.5 Length: 9421Number of Matches: 1

Score	Expect	Identities
1317 bits(713)	0.0	713/713(100%)

```
Query  1    ATGCCCGCCATGCCCTCCAGCGGCCCCGGGGACACCAGCAGCTCTGCTGCggagcgggag  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  278  ATGCCCGCCATGCCCTCCAGCGGCCCCGGGGACACCAGCAGCTCTGCTGCGGAGCGGGAG  337

Query  61   gaggaccgaaaggacggagaggagcaggaggagccgcgtggcaaggaggagcgccaagag  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  338  GAGGACCGAAAGGACGGAGAGGAGCAGGAGGAGCCGCGTGGCAAGGAGGAGCGCCAAGAG  397

Query  121  cCCAGCACCACGGCACGGAAGGTGGGGCGGCCTGGGAGGAAGCGCAAGCACCCCCCGGTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  398  CCCAGCACCACGGCACGGAAGGTGGGGCGGCCTGGGAGGAAGCGCAAGCACCCCCCGGTG  457

Query  181  GAAAGCGGTGACACGCCAAAGGACCCTGCGGTGATCTCCAAGTCCCCATCCATGGCCCAG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  458  GAAAGCGGTGACACGCCAAAGGACCCTGCGGTGATCTCCAAGTCCCCATCCATGGCCCAG  517

Query  241  GACTCAGGCGCCTCAGAGCTATTACCCAATGGGGACTTGGAGAAGCGGAGTGAGCCCCAG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  518  GACTCAGGCGCCTCAGAGCTATTACCCAATGGGGACTTGGAGAAGCGGAGTGAGCCCCAG  577

Query  301  CCAGAGGAGGGGAGCCCTGCTGGGGGGCAGAAGGGCGGGGCCCCAGCAGAGGGAGAGGGT  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  578  CCAGAGGAGGGGAGCCCTGCTGGGGGGCAGAAGGGCGGGGCCCCAGCAGAGGGAGAGGGT  637

Query  361  GCAGCTGAGACCCTGCCTGAAGCCTCAAGAGCAGTGGAAAATGGCTGCTGCACCCCCAAG  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  638  GCAGCTGAGACCCTGCCTGAAGCCTCAAGAGCAGTGGAAAATGGCTGCTGCACCCCCAAG  697

Query  421  GAGGGCCGAGGAGCCCCTGCAGAAGCGGGCAAAGAACAGAAGGAGACCAACATCGAATCC  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  698  GAGGGCCGAGGAGCCCCTGCAGAAGCGGGCAAAGAACAGAAGGAGACCAACATCGAATCC  757

Query  481  ATGAAAATGGAGGGCTCCCGGGGCCGGCTGCGGGGTGGCTTGGGCTGGGAGTCCAGCCTC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  758  ATGAAAATGGAGGGCTCCCGGGGCCGGCTGCGGGGTGGCTTGGGCTGGGAGTCCAGCCTC  817
```

```
Query  541  CGTCAGCGGCCCATGCCGAGGCTCACCTTCCAGGCGGGGGACCCCTACTACATCAGCAAG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  818  CGTCAGCGGCCCATGCCGAGGCTCACCTTCCAGGCGGGGGACCCCTACTACATCAGCAAG  877

Query  601  CGCAAGCGGGACGAGTGGCTGGCACGCTGGAAAAGGGAGGCTGAGAAGAAAGCCAAGGTC  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  878  CGCAAGCGGGACGAGTGGCTGGCACGCTGGAAAAGGGAGGCTGAGAAGAAAGCCAAGGTC  937

Query  661  ATTGCAGGAATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAA  713
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  938  ATTGCAGGAATGAATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAA  990
```

DNMT3A2:

Homo sapiens DNA methyltransferase 3 alpha (DNMT3A), transcript variant 3, mRNA
Sequence ID: NM_022552.5Length: 9421Number of Matches: 1

Score	Expect	Identities
1491 bits(807)	0.0	807/807(100%)

```
Query  1     AATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGCCCT  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  950   AATGCTGTGGAAGAAAACCAGGGGCCCGGGGAGTCTCAGAAGGTGGAGGAGGCCAGCCCT  1009

Query  61    CCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAGCCC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1010  CCTGCTGTGCAGCAGCCCACTGACCCCGCATCCCCCACTGTGGCTACCACGCCTGAGCCC  1069

Query  121   GTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAGTAC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1070  GTGGGGTCCGATGCTGGGGACAAGAATGCCACCAAAGCAGGCGATGACGAGCCAGAGTAC  1129

Query  181   GAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTCTCC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1130  GAGGACGGCCGGGGCTTTGGCATTGGGGAGCTGGTGTGGGGGAAACTGCGGGGCTTCTCC  1189

Query  241   TGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAAGGC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1190  TGGTGGCCAGGCCGCATTGTGTCTTGGTGGATGACGGGCCGGAGCCGAGCAGCTGAAGGC  1249

Query  301   ACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAGCTG  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1250  ACCCGCTGGGTCATGTGGTTCGGAGACGGCAAATTCTCAGTGGTGTGTGTTGAGAAGCTG  1309

Query  361   ATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCCATG  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1310  ATGCCGCTGAGCTCGTTTTGCAGTGCGTTCCACCAGGCCACGTACAACAAGCAGCCCATG  1369

Query  421   TACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTGTTC  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1370  TACCGCAAAGCCATCTACGAGGTCCTGCAGGTGGCCAGCAGCCGCGCGGGGAAGCTGTTC  1429

Query  481   CCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAACAAG  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1430  CCGGTGTGCCACGACAGCGATGAGAGTGACACTGCCAAGGCCGTGGAGGTGCAGAACAAG  1489
```

```
Query  541   CCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAGCCA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1490  CCCATGATTGAATGGGCCCTGGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCCTGGAGCCA  1549

Query  601   CCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCTGAG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1550  CCAGAAGAAGAGAAGAATCCCTACAAAGAAGTGTACACGGACATGTGGGTGGAACCTGAG  1609

Query  661   GCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAGAAG  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1610  GCAGCTGCCTACGCACCACCTCCACCAGCCAAAAAGCCCCGGAAGAGCACAGCGGAGAAG  1669

Query  721   CCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTGCGG  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1670  CCCAAGGTCAAGGAGATTATTGATGAGCGCACAAGAGAGCGGCTGGTGTACGAGGTGCGG  1729

Query  781   CAGAAGTGCCGGAACATTGAGGACATC  807
             |||||||||||||||||||||||||||
Sbjct  1730  CAGAAGTGCCGGAACATTGAGGACATC  1756
```

DNMT3B1:

Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 1, mRNA
Sequence ID: NM_006892.4Length: 4336Number of Matches: 1

Score	Expect	Identities
1803 bits(976)	0.0	984/987(99%)

```
Query  1     AAGGGAGACACCAGGCATCTCAATGGAGAGGAGGACGCCGGCGGGAGGGAAGACTCGATC  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  325   AAGGGAGACACCAGGCATCTCAATGGAGAGGAGGACGCCGGCGGGAGGGAAGACTCGATC  384

Query  61    CTCGTCAACGGGGCCTGCAGCGACCAGTCCTCCGACTCGCCCCCAATCCTGGAGGCTATC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  385   CTCGTCAACGGGGCCTGCAGCGACCAGTCCTCCGACTCGCCCCCAATCCTGGAGGCTATC  444

Query  121   CGCACCCCGGAGATCAGAGGCCGAAGATCAAGCTCGCGACTCTCCAAGAGGGAGGTGTCC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  445   CGCACCCCGGAGATCAGAGGCCGAAGATCAAGCTCGCGACTCTCCAAGAGGGAGGTGTCC  504

Query  181   AGTCTGCTAAGCTACACACAGGACTTGACAGGCGATGGCGACGGGGAAGATGGGGATGGC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  505   AGTCTGCTAAGCTACACACAGGACTTGACAGGCGATGGCGACGGGGAAGATGGGGATGGC  564

Query  241   TCTGACACCCCAGTCATGCCAAAGCTCTTCCGGGAAACCAGGACTCGTTCAGAAAGCCCA  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  565   TCTGACACCCCAGTCATGCCAAAGCTCTTCCGGGAAACCAGGACTCGTTCAGAAAGCCCA  624

Query  301   GCTGTCCGAACTCGAAATAACAACAGTGTCTCCAGCCGGGAGAGGCACAGGCCTTCCCCA  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  625   GCTGTCCGAACTCGAAATAACAACAGTGTCTCCAGCCGGGAGAGGCACAGGCCTTCCCCA  684

Query  361   CGTTCCACCCGAGGCCGGCAGGGCCGCAACCATGTGGACGAGTCCCCCGTGGAGTTCCCG  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  685   CGTTCCACCCGAGGCCGGCAGGGCCGCAACCATGTGGACGAGTCCCCCGTGGAGTTCCCG  744

Query  421   GCTACCAGGTCCCTGAGACGGCGGGCAACAGCATCGGCAGGAACGCCATGGCCGTCCCCT  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  745   GCTACCAGGTCCCTGAGACGGCGGGCAACAGCATCGGCAGGAACGCCATGGCCGTCCCCT  804

Query  481   CCCAGCTCTTACCTTACCATCGACCTCACAGACGACACAGAGGACACACATGGGACGCCC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  805   CCCAGCTCTTACCTTACCATCGACCTCACAGACGACACAGAGGACACACATGGGACGCCC  864
```

```
Query  541   CAGAGCAGCAGTACCCCCTACGCCCGCCTAGCCCAGGACAGCCAGCAGGGGGGCATGGAG  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  865   CAGAGCAGCAGTACCCCCTACGCCCGCCTAGCCCAGGACAGCCAGCAGGGGGGCATGGAG  924

Query  601   TCCCCGCAGGTGGAGGCAGACAGTGGAGATGGAGACAGTTCAGAGTATCAGGATGGGAAG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  925   TCCCCGCAGGTGGAGGCAGACAGTGGAGATGGAGACAGTTCAGAGTATCAGGATGGGAAG  984

Query  661   GAGTTTGGAATAGGGGACCTCGTGTGGGGAAAGATCAAGGGCTTCTCCTGGTGGCCCGCC  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  985   GAGTTTGGAATAGGGGACCTCGTGTGGGGAAAGATCAAGGGCTTCTCCTGGTGGCCCGCC  1044

Query  721   ATGGTGGTGTCTTGGAAGGCCACCTCCAAGCGACAGGCTATGTCTGGCATGCGGTGGGTC  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1045  ATGGTGGTGTCTTGGAAGGCCACCTCCAAGCGACAGGCTATGTCTGGCATGCGGTGGGTC  1104

Query  781   CAGTGGTTTGGCGATGGCAAGTTCTCCGAGGTCTCTGCAGACAAACTGGTGGCACTGGGG  840
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  1105  CAGTGGTTTGGCGATGGCAAGTTCTCCGAGGTCTCTGCAGACAAACTGGTGGCACT-GGG  1163

Query  841   GCTGTTCAGCCAGCACTTTAATTTGGCCACCTTCAATAAGCTCGTCTCCTATCGAAAAGC  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1164  GCTGTTCAGCCAGCACTTTAATTTGGCCACCTTCAATAAGCTCGTCTCCTATCGAAAAGC  1223

Query  901   CATGTACCATGCTCTGGAGAAAGCTAGGGTGCGAGCTGGCAAGACCTTCCCCAGCAGCCC  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1224  CATGTACCATGCTCTGGAGAAAGCTAGGGTGCGAGCTGGCAAGACCTTCCCCAGCAGCCC  1283

Query  961   TGGAAGACTCATTGGAAGGACCAGCTG  987
             |||| |||||||||| ||||||||||
Sbjct  1284  TGGA-GACTCATTGGA-GGACCAGCTG  1308
```

DNMT3B2:

Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 1, mRNA
Sequence ID: NM_006892.4Length: 4336Number of Matches: 1

Score	Expect	Identities
1829 bits(990)	0.0	990/990(100%)

```
Query  1     AGGCACAGCGGCCGAGGCCAAGCTTCAGGAGCCCTGGAGCTGTTACATGTGTCTCCCGCA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1851  AGGCACAGCGGCCGAGGCCAAGCTTCAGGAGCCCTGGAGCTGTTACATGTGTCTCCCGCA  1910

Query  61    GCGCTGTCATGGCGTCCTGCGGCGCCGGAAGGACTGGAACGTGCGCCTGCAGGCCTTCTT  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1911  GCGCTGTCATGGCGTCCTGCGGCGCCGGAAGGACTGGAACGTGCGCCTGCAGGCCTTCTT  1970

Query  121   CACCAGTGACACGGGGCTTGAATATGAAGCCCCCAAGCTGTACCCTGCCATTCCCGCAGC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1971  CACCAGTGACACGGGGCTTGAATATGAAGCCCCCAAGCTGTACCCTGCCATTCCCGCAGC  2030

Query  181   CCGAAGGCGGCCCATTCGAGTCCTGTCATTGTTTGATGGCATCGCGACAGGCTACCTAGT  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2031  CCGAAGGCGGCCCATTCGAGTCCTGTCATTGTTTGATGGCATCGCGACAGGCTACCTAGT  2090

Query  241   CCTCAAAGAGTTGGGCATAAAGGTAGGAAAGTACGTCGCTTCTGAAGTGTGTGAGGAGTC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2091  CCTCAAAGAGTTGGGCATAAAGGTAGGAAAGTACGTCGCTTCTGAAGTGTGTGAGGAGTC  2150

Query  301   CATTGCTGTTGGAACCGTGAAGCACGAGGGGAATATCAAATACGTGAACGACGTGAGGAA  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2151  CATTGCTGTTGGAACCGTGAAGCACGAGGGGAATATCAAATACGTGAACGACGTGAGGAA  2210

Query  361   CATCACAAAGAAAAATATTGAAGAATGGGGCCCATTTGACTTGGTGATTGGCGGAAGCCC  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2211  CATCACAAAGAAAAATATTGAAGAATGGGGCCCATTTGACTTGGTGATTGGCGGAAGCCC  2270

Query  421   ATGCAACGATCTCTCAAATGTGAATCCAGCCAGGAAAGGCCTGTATGAGGGTACAGGCCG  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2271  ATGCAACGATCTCTCAAATGTGAATCCAGCCAGGAAAGGCCTGTATGAGGGTACAGGCCG  2330

Query  481   GCTCTTCTTCGAATTTTACCACCTGCTGAATTACTCACGCCCCAAGGAGGGTGATGACCG  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2331  GCTCTTCTTCGAATTTTACCACCTGCTGAATTACTCACGCCCCAAGGAGGGTGATGACCG  2390
```

```
Query  541   GCCGTTCTTCTGGATGTTTGAGAATGTTGTAGCCATGAAGGTTGGCGACAAGAGGGACAT  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2391  GCCGTTCTTCTGGATGTTTGAGAATGTTGTAGCCATGAAGGTTGGCGACAAGAGGGACAT  2450

Query  601   CTCACGGTTCCTGGAGTGTAATCCAGTGATGATTGATGCCATCAAAGTTTCTGCTGCTCA  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2451  CTCACGGTTCCTGGAGTGTAATCCAGTGATGATTGATGCCATCAAAGTTTCTGCTGCTCA  2510

Query  661   CAGGGCCCGATACTTCTGGGGCAACCTACCCGGGATGAACAGGCCCGTGATAGCATCAAA  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2511  CAGGGCCCGATACTTCTGGGGCAACCTACCCGGGATGAACAGGCCCGTGATAGCATCAAA  2570

Query  721   GAATGATAAACTCGAGCTGCAGGACTGCTTGGAATACAATAGGATAGCCAAGTTAAAGAA  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2571  GAATGATAAACTCGAGCTGCAGGACTGCTTGGAATACAATAGGATAGCCAAGTTAAAGAA  2630

Query  781   AGTACAGACAATAACCACCAAGTCGAACTCGATCAAACAGGGGAAAAACCAACTTTTCCC  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2631  AGTACAGACAATAACCACCAAGTCGAACTCGATCAAACAGGGGAAAAACCAACTTTTCCC  2690

Query  841   TGTTGTCATGAATGGCAAAGAAGATGTTTTGTGGTGCACTGAGCTCGAAAGGATCTTTGG  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2691  TGTTGTCATGAATGGCAAAGAAGATGTTTTGTGGTGCACTGAGCTCGAAAGGATCTTTGG  2750

Query  901   CTTTCCTGTGCACTACACAGACGTGTCCAACATGGGCCGTGGTGCCCGCCAGAAGCTGCT  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2751  CTTTCCTGTGCACTACACAGACGTGTCCAACATGGGCCGTGGTGCCCGCCAGAAGCTGCT  2810

Query  961   GGGAAGGTCCTGGAGCGTGCCTGTCATCCG  990
             ||||||||||||||||||||||||||||||
Sbjct  2811  GGGAAGGTCCTGGAGCGTGCCTGTCATCCG  2840
```

DNMT3B3:

Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 1, mRNA
Sequence ID: NM_006892.4Length: 4336Number of Matches: 1

Score	Expect	Identities
1576 bits(853)	0.0	853/853(100%)

```
Query  1     GACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGATGACGATGGCTATCAGTC  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1689  GACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGATGACGATGGCTATCAGTC  1748

Query  61    TTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCAGCAACACGAGCTGCTG  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1749  TTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCAGCAACACGAGCTGCTG  1808

Query  121   CCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGTGGGCACAGGCACAGCGGCCGAGGC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1809  CCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGTGGGCACAGGCACAGCGGCCGAGGC  1868

Query  181   CAAGCTTCAGGAGCCCTGGAGCTGTTACATGTGTCTCCCGCAGCGCTGTCATGGCGTCCT  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1869  CAAGCTTCAGGAGCCCTGGAGCTGTTACATGTGTCTCCCGCAGCGCTGTCATGGCGTCCT  1928

Query  241   GCGGCGCCGGAAGGACTGGAACGTGCGCCTGCAGGCCTTCTTCACCAGTGACACGGGGCT  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1929  GCGGCGCCGGAAGGACTGGAACGTGCGCCTGCAGGCCTTCTTCACCAGTGACACGGGGCT  1988

Query  301   TGAATATGAAGCCCCCAAGCTGTACCCTGCCATTCCCGCAGCCCGAAGGCGGCCCATTCG  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1989  TGAATATGAAGCCCCCAAGCTGTACCCTGCCATTCCCGCAGCCCGAAGGCGGCCCATTCG  2048

Query  361   AGTCCTGTCATTGTTTGATGGCATCGCGACAGGCTACCTAGTCCTCAAAGAGTTGGGCAT  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2049  AGTCCTGTCATTGTTTGATGGCATCGCGACAGGCTACCTAGTCCTCAAAGAGTTGGGCAT  2108

Query  421   AAAGGTAGGAAAGTACGTCGCTTCTGAAGTGTGTGAGGAGTCCATTGCTGTTGGAACCGT  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2109  AAAGGTAGGAAAGTACGTCGCTTCTGAAGTGTGTGAGGAGTCCATTGCTGTTGGAACCGT  2168

Query  481   GAAGCACGAGGGGAATATCAAATACGTGAACGACGTGAGGAACATCACAAAGAAAAATAT  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2169  GAAGCACGAGGGGAATATCAAATACGTGAACGACGTGAGGAACATCACAAAGAAAAATAT  2228
```

```
Query  541   TGAAGAATGGGGCCCATTTGACTTGGTGATTGGCGGAAGCCCATGCAACGATCTCTCAAA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2229  TGAAGAATGGGGCCCATTTGACTTGGTGATTGGCGGAAGCCCATGCAACGATCTCTCAAA  2288

Query  601   TGTGAATCCAGCCAGGAAAGGCCTGTATGAGGGTACAGGCCGGCTCTTCTTCGAATTTTA  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2289  TGTGAATCCAGCCAGGAAAGGCCTGTATGAGGGTACAGGCCGGCTCTTCTTCGAATTTTA  2348

Query  661   CCACCTGCTGAATTACTCACGCCCCAAGGAGGGTGATGACCGGCCGTTCTTCTGGATGTT  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2349  CCACCTGCTGAATTACTCACGCCCCAAGGAGGGTGATGACCGGCCGTTCTTCTGGATGTT  2408

Query  721   TGAGAATGTTGTAGCCATGAAGGTTGGCGACAAGAGGGACATCTCACGGTTCCTGGAGTG  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2409  TGAGAATGTTGTAGCCATGAAGGTTGGCGACAAGAGGGACATCTCACGGTTCCTGGAGTG  2468

Query  781   TAATCCAGTGATGATTGATGCCATCAAAGTTTCTGCTGCTCACAGGGCCCGATACTTCTG  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2469  TAATCCAGTGATGATTGATGCCATCAAAGTTTCTGCTGCTCACAGGGCCCGATACTTCTG  2528

Query  841   GGGCAACCTACCC  853
             |||||||||||||
Sbjct  2529  GGGCAACCTACCC  2541
```

DNMT3B4:

Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 1, mRNA
Sequence ID: NM_006892.4Length: 4336Number of Matches: 1

Score	Expect	Identities
1838 bits(995)	0.0	995/995(100%)

```
Query  1     GACCGAGGGGATGAAGATCAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACAACAAG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1546  GACCGAGGGGATGAAGATCAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACAACAAG  1605

Query  61    AGCAGCCTGGAAGATGGCTGTTTGTCTTGTGGCAGGAAAAACCCCGTGTCCTTCCACCCT  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1606  AGCAGCCTGGAAGATGGCTGTTTGTCTTGTGGCAGGAAAAACCCCGTGTCCTTCCACCCT  1665

Query  121   CTCTTTGAgggggggCTCTGTCAGACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1666  CTCTTTGAGGGGGGGCTCTGTCAGACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATG  1725

Query  181   TATGATGACGATGGCTATCAGTCTTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1726  TATGATGACGATGGCTATCAGTCTTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTG  1785

Query  241   CTTTGCAGCAACACGAGCTGCTGCCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGTG  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1786  CTTTGCAGCAACACGAGCTGCTGCCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGTG  1845

Query  301   GGCACAGGCACAGCGGCCGAGGCCAAGCTTCAGGAGCCCTGGAGCTGTTACATGTGTCTC  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1846  GGCACAGGCACAGCGGCCGAGGCCAAGCTTCAGGAGCCCTGGAGCTGTTACATGTGTCTC  1905

Query  361   CCGCAGCGCTGTCATGGCGTCCTGCGGCGCCGGAAGGACTGGAACGTGCGCCTGCAGGCC  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1906  CCGCAGCGCTGTCATGGCGTCCTGCGGCGCCGGAAGGACTGGAACGTGCGCCTGCAGGCC  1965

Query  421   TTCTTCACCAGTGACACGGGGCTTGAATATGAAGCCCCCAAGCTGTACCCTGCCATTCCC  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1966  TTCTTCACCAGTGACACGGGGCTTGAATATGAAGCCCCCAAGCTGTACCCTGCCATTCCC  2025

Query  481   GCAGCCCGAAGGCGGCCCATTCGAGTCCTGTCATTGTTTGATGGCATCGCGACAGGCTAC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2026  GCAGCCCGAAGGCGGCCCATTCGAGTCCTGTCATTGTTTGATGGCATCGCGACAGGCTAC  2085
```

```
Query  541   CTAGTCCTCAAAGAGTTGGGCATAAAGGTAGGAAAGTACGTCGCTTCTGAAGTGTGTGAG  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2086  CTAGTCCTCAAAGAGTTGGGCATAAAGGTAGGAAAGTACGTCGCTTCTGAAGTGTGTGAG  2145

Query  601   GAGTCCATTGCTGTTGGAACCGTGAAGCACGAGGGGAATATCAAATACGTGAACGACGTG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2146  GAGTCCATTGCTGTTGGAACCGTGAAGCACGAGGGGAATATCAAATACGTGAACGACGTG  2205

Query  661   AGGAACATCACAAAGAAAAATATTGAAGAATGGGGCCCATTTGACTTGGTGATTGGCGGA  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2206  AGGAACATCACAAAGAAAAATATTGAAGAATGGGGCCCATTTGACTTGGTGATTGGCGGA  2265

Query  721   AGCCCATGCAACGATCTCTCAAATGTGAATCCAGCCAGGAAAGGCCTGTATGAGGGTACA  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2266  AGCCCATGCAACGATCTCTCAAATGTGAATCCAGCCAGGAAAGGCCTGTATGAGGGTACA  2325

Query  781   GGCCGGCTCTTCTTCGAATTTTACCACCTGCTGAATTACTCACGCCCCAAGGAGGGTGAT  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2326  GGCCGGCTCTTCTTCGAATTTTACCACCTGCTGAATTACTCACGCCCCAAGGAGGGTGAT  2385

Query  841   GACCGGCCGTTCTTCTGGATGTTTGAGAATGTTGTAGCCATGAAGGTTGGCGACAAGAGG  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2386  GACCGGCCGTTCTTCTGGATGTTTGAGAATGTTGTAGCCATGAAGGTTGGCGACAAGAGG  2445

Query  901   GACATCTCACGGTTCCTGGAGTGTAATCCAGTGATGATTGATGCCATCAAAGTTTCTGCT  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2446  GACATCTCACGGTTCCTGGAGTGTAATCCAGTGATGATTGATGCCATCAAAGTTTCTGCT  2505

Query  961   GCTCACAGGGCCCGATACTTCTGGGGCAACCTACC  995
             ||||||||||||||||||||||||||||||||||
Sbjct  2506  GCTCACAGGGCCCGATACTTCTGGGGCAACCTACC  2540
```

DNMT3B5:

Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 1, mRNA
Sequence ID: NM_006892.4Length: 4336Number of Matches: 2

Score	Expect	Identities
1733 bits(938)	0.0	938/938(100%)

```
Query  1     GTCAGACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGATGACGATGGCTATC  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1685  GTCAGACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGATGACGATGGCTATC  1744

Query  61    AGTCTTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCAGCAACACGAGCT  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1745  AGTCTTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCAGCAACACGAGCT  1804

Query  121   GCTGCCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGTGGGCACAGGCACAGCGGCCG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1805  GCTGCCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGTGGGCACAGGCACAGCGGCCG  1864

Query  181   AGGCCAAGCTTCAGGAGCCCTGGAGCTGTTACATGTGTCTCCCGCAGCGCTGTCATGGCG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1865  AGGCCAAGCTTCAGGAGCCCTGGAGCTGTTACATGTGTCTCCCGCAGCGCTGTCATGGCG  1924

Query  241   TCCTGCGGCGCCGGAAGGACTGGAACGTGCGCCTGCAGGCCTTCTTCACCAGTGACACGG  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1925  TCCTGCGGCGCCGGAAGGACTGGAACGTGCGCCTGCAGGCCTTCTTCACCAGTGACACGG  1984

Query  301   GGCTTGAATATGAAGCCCCCAAGCTGTACCCTGCCATTCCCGCAGCCCGAAGGCGGCCCA  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1985  GGCTTGAATATGAAGCCCCCAAGCTGTACCCTGCCATTCCCGCAGCCCGAAGGCGGCCCA  2044

Query  361   TTCGAGTCCTGTCATTGTTTGATGGCATCGCGACAGGCTACCTAGTCCTCAAAGAGTTGG  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2045  TTCGAGTCCTGTCATTGTTTGATGGCATCGCGACAGGCTACCTAGTCCTCAAAGAGTTGG  2104

Query  421   GCATAAAGGTAGGAAAGTACGTCGCTTCTGAAGTGTGTGAGGAGTCCATTGCTGTTGGAA  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2105  GCATAAAGGTAGGAAAGTACGTCGCTTCTGAAGTGTGTGAGGAGTCCATTGCTGTTGGAA  2164

Query  481   CCGTGAAGCACGAGGGGAATATCAAATACGTGAACGACGTGAGGAACATCACAAAGAAAA  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2165  CCGTGAAGCACGAGGGGAATATCAAATACGTGAACGACGTGAGGAACATCACAAAGAAAA  2224
```

```
Query  541   ATATTGAAGAATGGGGCCCATTTGACTTGGTGATTGGCGGAAGCCCATGCAACGATCTCT  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2225  ATATTGAAGAATGGGGCCCATTTGACTTGGTGATTGGCGGAAGCCCATGCAACGATCTCT  2284

Query  601   CAAATGTGAATCCAGCCAGGAAAGGCCTGTATGAGGGTACAGGCCGGCTCTTCTTCGAAT  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2285  CAAATGTGAATCCAGCCAGGAAAGGCCTGTATGAGGGTACAGGCCGGCTCTTCTTCGAAT  2344

Query  661   TTTACCACCTGCTGAATTACTCACGCCCCAAGGAGGGTGATGACCGGCCGTTCTTCTGGA  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2345  TTTACCACCTGCTGAATTACTCACGCCCCAAGGAGGGTGATGACCGGCCGTTCTTCTGGA  2404

Query  721   TGTTTGAGAATGTTGTAGCCATGAAGGTTGGCGACAAGAGGGACATCTCACGGTTCCTGG  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2405  TGTTTGAGAATGTTGTAGCCATGAAGGTTGGCGACAAGAGGGACATCTCACGGTTCCTGG  2464

Query  781   AGTGTAATCCAGTGATGATTGATGCCATCAAAGTTTCTGCTGCTCACAGGGCCCGATACT  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2465  AGTGTAATCCAGTGATGATTGATGCCATCAAAGTTTCTGCTGCTCACAGGGCCCGATACT  2524

Query  841   TCTGGGGCAACCTACCCGGGATGAACAGGCCCGTGATAGCATCAAAGAATGATAAACTCG  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2525  TCTGGGGCAACCTACCCGGGATGAACAGGCCCGTGATAGCATCAAAGAATGATAAACTCG  2584

Query  901   AGCTGCAGGACTGCTTGGAATACAATAGGATAGCCAAG  938
             ||||||||||||||||||||||||||||||||||||||
Sbjct  2585  AGCTGCAGGACTGCTTGGAATACAATAGGATAGCCAAG  2622
```

DNMTΔ3B1:

Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 1, mRNA
Sequence ID: NM_006892.4Length: 4336Number of Matches: 1

Score	Expect	Identities
1600 bits(866)	0.0	869/870(99%)

```
Query  1     GAGTCCCCGCAGGTGGAGGCAGACAGTGGAGATGGAGACAGTTCAGAGTATCAGGATGGG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  922   GAGTCCCCGCAGGTGGAGGCAGACAGTGGAGATGGAGACAGTTCAGAGTATCAGGATGGG  981

Query  61    AAGGAGTTTGGAATAGGGGACCTCGTGTGGGGAAAGATCAAGGGCTTCTCCTGGTGGCCC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  982   AAGGAGTTTGGAATAGGGGACCTCGTGTGGGGAAAGATCAAGGGCTTCTCCTGGTGGCCC  1041

Query  121   GCCATGGTGGTGTCTTGGAAGGCCACCTCCAAGCGACAGGCTATGTCTGGCATGCGGTGG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1042  GCCATGGTGGTGTCTTGGAAGGCCACCTCCAAGCGACAGGCTATGTCTGGCATGCGGTGG  1101

Query  181   GTCCAGTGGTTTGGCGATGGCAAGTTCTCCGAGGTCTCTGCAGACAAACTGGTGGCACTG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1102  GTCCAGTGGTTTGGCGATGGCAAGTTCTCCGAGGTCTCTGCAGACAAACTGGTGGCACTG  1161

Query  241   GGGCTGTTCAGCCAGCACTTTAATTTGGCCACCTTCAATAAGCTCGTCTCCTATCGAAAA  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1162  GGGCTGTTCAGCCAGCACTTTAATTTGGCCACCTTCAATAAGCTCGTCTCCTATCGAAAA  1221

Query  301   GCCATGTACCATGCTCTGGAGAAAGCTAGGGTGCGAGCTGGCAAGACCTTCCCCAGCAGC  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1222  GCCATGTACCATGCTCTGGAGAAAGCTAGGGTGCGAGCTGGCAAGACCTTCCCCAGCAGC  1281

Query  361   CCTGGAGACTCATTGGAGGACCAGCTGAAGCCCATGTTGGAGTGGGCCCACGGGGGCTTC  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1282  CCTGGAGACTCATTGGAGGACCAGCTGAAGCCCATGTTGGAGTGGGCCCACGGGGGCTTC  1341

Query  421   AAGCCCACTGGGATCGAGGGCCTCAAACCCAACAACACGCAACCAGTGGTTAATAAGTCG  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1342  AAGCCCACTGGGATCGAGGGCCTCAAACCCAACAACACGCAACCAGTGGTTAATAAGTCG  1401

Query  481   AAGGTGCGTCGTGCAGGCAGTAGGAAATTAGAATCAAGGAAATACGAGAACAAGACTCGA  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1402  AAGGTGCGTCGTGCAGGCAGTAGGAAATTAGAATCAAGGAAATACGAGAACAAGACTCGA  1461
```

```
Query  541   AGACGCACAGCTGACGACTCAGCCACCTCTGACTACTGCCCCGCACCCAAGCGCCTCAAG  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1462  AGACGCACAGCTGACGACTCAGCCACCTCTGACTACTGCCCCGCACCCAAGCGCCTCAAG  1521

Query  601   ACAAATTGCTATAACAACGGCAAAGACCGAGGGGATGAAGATCAGAGCCGAGAACAAATG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1522  ACAAATTGCTATAACAACGGCAAAGACCGAGGGGATGAAGATCAGAGCCGAGAACAAATG  1581

Query  661   GCTTCAGATGTTGCCAACAACAAGAGCAGCCTGGAAGATGGCTGTTTGTCTTGTGGCAGG  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1582  GCTTCAGATGTTGCCAACAACAAGAGCAGCCTGGAAGATGGCTGTTTGTCTTGTGGCAGG  1641

Query  721   AAAAACCCCGTGTCCTTCCACCCTCTCTTTGAgggggggCTCTGTCAGACATGCCGGGGA  780
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  1642  AAAAACCCCGTGTCCTTCCACCCTCTCTTTGAGGGGGGGCTCTGTCAGACATGCC-GGGA  1700

Query  781   TCGCTTCCTTGAGCTGTTTTACATGTATGATGACGATGGCTATCAGTCTTACTGCACTGT  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1701  TCGCTTCCTTGAGCTGTTTTACATGTATGATGACGATGGCTATCAGTCTTACTGCACTGT  1760

Query  841   GTGCTGCGAGGGCCGAGAGCTGCTGCTTTG  870
             ||||||||||||||||||||||||||||||
Sbjct  1761  GTGCTGCGAGGGCCGAGAGCTGCTGCTTTG  1790
```

DNMTΔ3B2:

Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 2, mRNA
Sequence ID: NM_175848.2Length: 4276Number of Matches: 1

Score	Expect	Identities
1415 bits(766)	0.0	769/770(99%)

```
Query  1     GAGTCCCCGCAGGTGGAGGCAGACAGTGGAGATGGAGACAGTTCAGAGTATCAGGATGGG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  922   GAGTCCCCGCAGGTGGAGGCAGACAGTGGAGATGGAGACAGTTCAGAGTATCAGGATGGG  981

Query  61    AAGGAGTTTGGAATAGGGGACCTCGTGTGGGGAAAGATCAAGGGCTTCTCCTGGTGGCCC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  982   AAGGAGTTTGGAATAGGGGACCTCGTGTGGGGAAAGATCAAGGGCTTCTCCTGGTGGCCC  1041

Query  121   GCCATGGTGGTGTCTTGGAAGGCCACCTCCAAGCGACAGGCTATGTCTGGCATGCGGTGG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1042  GCCATGGTGGTGTCTTGGAAGGCCACCTCCAAGCGACAGGCTATGTCTGGCATGCGGTGG  1101

Query  181   GTCCAGTGGTTTGGCGATGGCAAGTTCTCCGAGGTCTCTGCAGACAAACTGGTGGCACTG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1102  GTCCAGTGGTTTGGCGATGGCAAGTTCTCCGAGGTCTCTGCAGACAAACTGGTGGCACTG  1161

Query  241   GGGCTGTTCAGCCAGCACTTTAATTTGGCCACCTTCAATAAGCTCGTCTCCTATCGAAAA  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1162  GGGCTGTTCAGCCAGCACTTTAATTTGGCCACCTTCAATAAGCTCGTCTCCTATCGAAAA  1221

Query  301   GCCATGTACCATGCTCTGGAGAAAGCTAGGGTGCGAGCTGGCAAGACCTTCCCCAGCAGC  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1222  GCCATGTACCATGCTCTGGAGAAAGCTAGGGTGCGAGCTGGCAAGACCTTCCCCAGCAGC  1281

Query  361   CCTGGAGACTCATTGGAGGACCAGCTGAAGCCCATGTTGGAGTGGGCCCACGGGGGCTTC  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1282  CCTGGAGACTCATTGGAGGACCAGCTGAAGCCCATGTTGGAGTGGGCCCACGGGGGCTTC  1341

Query  421   AAGCCCACTGGGATCGAGGGCCTCAAACCCAACAACACGCAACCAGAGAACAAGACTCGA  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1342  AAGCCCACTGGGATCGAGGGCCTCAAACCCAACAACACGCAACCAGAGAACAAGACTCGA  1401

Query  481   AGACGCACAGCTGACGACTCAGCCACCTCTGACTACTGCCCCGCACCCAAGCGCCTCAAG  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1402  AGACGCACAGCTGACGACTCAGCCACCTCTGACTACTGCCCCGCACCCAAGCGCCTCAAG  1461
```

```
Query  541   ACAAATTGCTATAACAACGGCAAAGACCGAGGGGATGAAGATCAGAGCCGAGAACAAATG  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1462  ACAAATTGCTATAACAACGGCAAAGACCGAGGGGATGAAGATCAGAGCCGAGAACAAATG  1521

Query  601   GCTTCAGATGTTGCCAACAACAAGAGCAGCCTGGAAGATGGCTGTTTGTCTTGTGGCAGG  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1522  GCTTCAGATGTTGCCAACAACAAGAGCAGCCTGGAAGATGGCTGTTTGTCTTGTGGCAGG  1581

Query  661   AAAAACCCCGTGTCCTTCCACCCTCTCTTTGAgggggggCTCTGTCAGACATGCCGGGGA  720
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  1582  AAAAACCCCGTGTCCTTCCACCCTCTCTTTGAGGGGGGGCTCTGTCAGACATGCC-GGGA  1640

Query  721   TCGCTTCCTTGAGCTGTTTTACATGTATGATGACGATGGCTATCAGTCTT  770
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1641  TCGCTTCCTTGAGCTGTTTTACATGTATGATGACGATGGCTATCAGTCTT  1690
```

DNMTΔ3B3:

Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 1, mRNA
Sequence ID: NM_006892.4Length: 4336Number of Matches: 2

Score	Expect	Identities
1380 bits(747)	0.0	751/753(99%)

```
Query  53    AGGTCTCTGCAGACAAACTGGTGGCACTGGGGCTGTTCAGCCAGCACTTTAATTTGGCCA  112
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1133  AGGTCTCTGCAGACAAACTGGTGGCACTGGGGCTGTTCAGCCAGCACTTTAATTTGGCCA  1192

Query  113   CCTTCAATAAGCTCGTCTCCTATCGAAAAGCCATGTACCATGCTCTGGAGAAAGCTAGGG  172
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1193  CCTTCAATAAGCTCGTCTCCTATCGAAAAGCCATGTACCATGCTCTGGAGAAAGCTAGGG  1252

Query  173   TGCGAGCTGGCAAGACCTTCCCCAGCAGCCCTGGAGACTCATTGGAGGACCAGCTGAAGC  232
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1253  TGCGAGCTGGCAAGACCTTCCCCAGCAGCCCTGGAGACTCATTGGAGGACCAGCTGAAGC  1312

Query  233   CCATGTTGGAGTGGGCCCACGGGGGCTTCAAGCCCACTGGGATCGAGGGCCTCAAACCCA  292
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1313  CCATGTTGGAGTGGGCCCACGGGGGCTTCAAGCCCACTGGGATCGAGGGCCTCAAACCCA  1372

Query  293   ACAACACGCAACCAGTGGTTAATAAGTCGAAGGTGCGTCGTGCAGGCAGTAGGAAATTAG  352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1373  ACAACACGCAACCAGTGGTTAATAAGTCGAAGGTGCGTCGTGCAGGCAGTAGGAAATTAG  1432

Query  353   AATCAAGGAAATACGAGAACAAGACTCGAAGACGCACAGCTGACGACTCAGCCACCTCTG  412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1433  AATCAAGGAAATACGAGAACAAGACTCGAAGACGCACAGCTGACGACTCAGCCACCTCTG  1492

Query  413   ACTACTGCCCCGCACCCAAGCGCCTCAAGACAAATTGCTATAACAACGGCAAAGACCGAG  472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1493  ACTACTGCCCCGCACCCAAGCGCCTCAAGACAAATTGCTATAACAACGGCAAAGACCGAG  1552

Query  473   GGGATGAAGATCAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACAACAAGAGCAGCC  532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1553  GGGATGAAGATCAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACAACAAGAGCAGCC  1612

Query  533   TGGAAGATGGCTGTTTGTCTTGTGGCAGGAAAAACCCCGTGTCCTTCCACCCTCTCTTTG  592
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1613  TGGAAGATGGCTGTTTGTCTTGTGGCAGGAAAAACCCCGTGTCCTTCCACCCTCTCTTTG  1672
```

```
Query  593   AgggggggCTCTGTCAGACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGATG  652
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1673  AGGGGGGGCTCTGTCAGACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGATG  1732

Query  653   ACGATGGCTATCAGTCTTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCA  712
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1733  ACGATGGCTATCAGTCTTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCA  1792

Query  713   GCAACACGAGCTGCTGCCGGTGTTTCTGTGTGGAGTGCCTGNAGGTGCTGGTGGGGCACA  772
             |||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Sbjct  1793  GCAACACGAGCTGCTGCCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGT-GGGCACA  1851

Query  773   GGCACAGCGGCCGAGGCCAAGCTTCAGGAGCCC  805
             |||||||||||||||||||||||||||||||||
Sbjct  1852  GGCACAGCGGCCGAGGCCAAGCTTCAGGAGCCC  1884
```

DNMTΔ3B4:

Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 2, mRNA
Sequence ID: NM_175848.2Length: 4276Number of Matches: 1

Score	Expect	Identities
1199 bits(649)	0.0	651/652(99%)

```
Query  53    AGGTCTCTGCATACAAACTGGTGGCACTGGGGCTGTTCAGCCAGCACTTTAATTTGGCCA  112
             |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1133  AGGTCTCTGCAGACAAACTGGTGGCACTGGGGCTGTTCAGCCAGCACTTTAATTTGGCCA  1192

Query  113   CCTTCAATAAGCTCGTCTCCTATCGAAAAGCCATGTACCATGCTCTGGAGAAAGCTAGGG  172
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1193  CCTTCAATAAGCTCGTCTCCTATCGAAAAGCCATGTACCATGCTCTGGAGAAAGCTAGGG  1252

Query  173   TGCGAGCTGGCAAGACCTTCCCCAGCAGCCCTGGAGACTCATTGGAGGACCAGCTGAAGC  232
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1253  TGCGAGCTGGCAAGACCTTCCCCAGCAGCCCTGGAGACTCATTGGAGGACCAGCTGAAGC  1312

Query  233   CCATGTTGGAGTGGGCCCACGGGGGCTTCAAGCCCACTGGGATCGAGGGCCTCAAACCCA  292
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1313  CCATGTTGGAGTGGGCCCACGGGGGCTTCAAGCCCACTGGGATCGAGGGCCTCAAACCCA  1372

Query  293   ACAACACGCAACCAGAGAACAAGACTCGAAGACGCACAGCTGACGACTCAGCCACCTCTG  352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1373  ACAACACGCAACCAGAGAACAAGACTCGAAGACGCACAGCTGACGACTCAGCCACCTCTG  1432

Query  353   ACTACTGCCCCGCACCCAAGCGCCTCAAGACAAATTGCTATAACAACGGCAAAGACCGAG  412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1433  ACTACTGCCCCGCACCCAAGCGCCTCAAGACAAATTGCTATAACAACGGCAAAGACCGAG  1492

Query  413   GGGATGAAGATCAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACAACAAGAGCAGCC  472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1493  GGGATGAAGATCAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACAACAAGAGCAGCC  1552

Query  473   TGGAAGATGGCTGTTTGTCTTGTGGCAGGAAAAACCCCGTGTCCTTCCACCCTCTCTTTG  532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1553  TGGAAGATGGCTGTTTGTCTTGTGGCAGGAAAAACCCCGTGTCCTTCCACCCTCTCTTTG  1612

Query  533   AgggggggCTCTGTCAGACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGATG  592
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1613  AGGGGGGGCTCTGTCAGACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGATG  1672
```

```
Query  593   ACGATGGCTATCAGTCTTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCA  652
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1673  ACGATGGCTATCAGTCTTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCA  1732

Query  653   GCAACACGAGCTGCTGCCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGT  704
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1733  GCAACACGAGCTGCTGCCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGT  1784
```

DNMT1:

Homo sapiens DNA methyltransferase 1 (DNMT1), transcript variant 2, mRNA
Sequence ID: NM_001379.4Length: 5226Number of Matches: 1

Score	Expect	Identities
1664 bits(901)	0.0	901/901(100%)

```
Query  1    CCGGCGCGTACCGCCCCAGCCCGGGTGCCCACACTGGCCGTCCCGGCCATCTCGCTGCCC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  58   CCGGCGCGTACCGCCCCAGCCCGGGTGCCCACACTGGCCGTCCCGGCCATCTCGCTGCCC  117

Query  61   GACGATGTCCGCAGGCGGCTCAAAGATTTGGAAAGAGACAGCTTAACAGAAAAGGAATGT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  118  GACGATGTCCGCAGGCGGCTCAAAGATTTGGAAAGAGACAGCTTAACAGAAAAGGAATGT  177

Query  121  GTGAAGGAGAAATTGAATCTCTTGCACGAATTTCTGCAAACAGAAATAAAGAATCAGTTA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  178  GTGAAGGAGAAATTGAATCTCTTGCACGAATTTCTGCAAACAGAAATAAAGAATCAGTTA  237

Query  181  TGTGACTTGGAAACCAAATTACGTAAAGAAGAATTATCCGAGGAGGGCTACCTGGCTAAA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  238  TGTGACTTGGAAACCAAATTACGTAAAGAAGAATTATCCGAGGAGGGCTACCTGGCTAAA  297

Query  241  GTCAAATCCCTTTTAAATAAAGATTTGTCCTTGGAGAACGGTGCTCATGCTTACAACCGG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  298  GTCAAATCCCTTTTAAATAAAGATTTGTCCTTGGAGAACGGTGCTCATGCTTACAACCGG  357

Query  301  GAAGTGAATGGACGTCTAGAAAACGGGAACCAAGCAAGAAGTGAAGCCCGTAGAGTGGGA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  358  GAAGTGAATGGACGTCTAGAAAACGGGAACCAAGCAAGAAGTGAAGCCCGTAGAGTGGGA  417

Query  361  ATGGCAGATGCCAACAGccccccAAACCCCTTTCCAAACCTCGCACGCCCAGGAGGAGC  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  418  ATGGCAGATGCCAACAGCCCCCCCAAACCCCTTTCCAAACCTCGCACGCCCAGGAGGAGC  477

Query  421  AAGTCCGATGGAGAGGCTAAGCCTGAACCTTCACCTAGCCCCAGGATTACAAGGAAAAGC  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  478  AAGTCCGATGGAGAGGCTAAGCCTGAACCTTCACCTAGCCCCAGGATTACAAGGAAAAGC  537

Query  481  ACCAGGCAAACCACCATCACATCTCATTTTGCAAAGGGCCCTGCCAAACGGAAACCTCAG  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  538  ACCAGGCAAACCACCATCACATCTCATTTTGCAAAGGGCCCTGCCAAACGGAAACCTCAG  597
```

```
Query  541  GAAGAGTCTGAAAGAGCCAAATCGGATGAGTCCATCAAGGAAGAAGACAAAGACCAGGAT  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  598  GAAGAGTCTGAAAGAGCCAAATCGGATGAGTCCATCAAGGAAGAAGACAAAGACCAGGAT  657

Query  601  GAGAAGAGACGTAGAGTTACATCCAGAGAACGAGTTGCTAGACCGCTTCCTGCAGAAGAA  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  658  GAGAAGAGACGTAGAGTTACATCCAGAGAACGAGTTGCTAGACCGCTTCCTGCAGAAGAA  717

Query  661  CCTGAAAGAGCAAAATCAGGAACGCGCACTgaaaaggaagaagaaagagatgaaaaagaa  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  718  CCTGAAAGAGCAAAATCAGGAACGCGCACTGAAAAGGAAGAAGAAAGAGATGAAAAAGAA  777

Query  721  gaaaagagaCTCCGAAGTCAAACCAAAGAACCAACACCCAAACAGAAACTGAAGGAGGAG  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  778  GAAAAGAGACTCCGAAGTCAAACCAAAGAACCAACACCCAAACAGAAACTGAAGGAGGAG  837

Query  781  CCGGACAGAGAAGCCAGGGCAGGCGTGCAGGCTGACGAGGACGAAGATGGAGACGAGAAA  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  838  CCGGACAGAGAAGCCAGGGCAGGCGTGCAGGCTGACGAGGACGAAGATGGAGACGAGAAA  897

Query  841  GATGAGAAGAAGCACAGAAGTCAACCCAAAGATCTAGCTGCCAAACGGAGGCCCGAAGAA  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  898  GATGAGAAGAAGCACAGAAGTCAACCCAAAGATCTAGCTGCCAAACGGAGGCCCGAAGAA  957

Query  901  A  901
            |
Sbjct  958  A  958
```

DNMT3L:

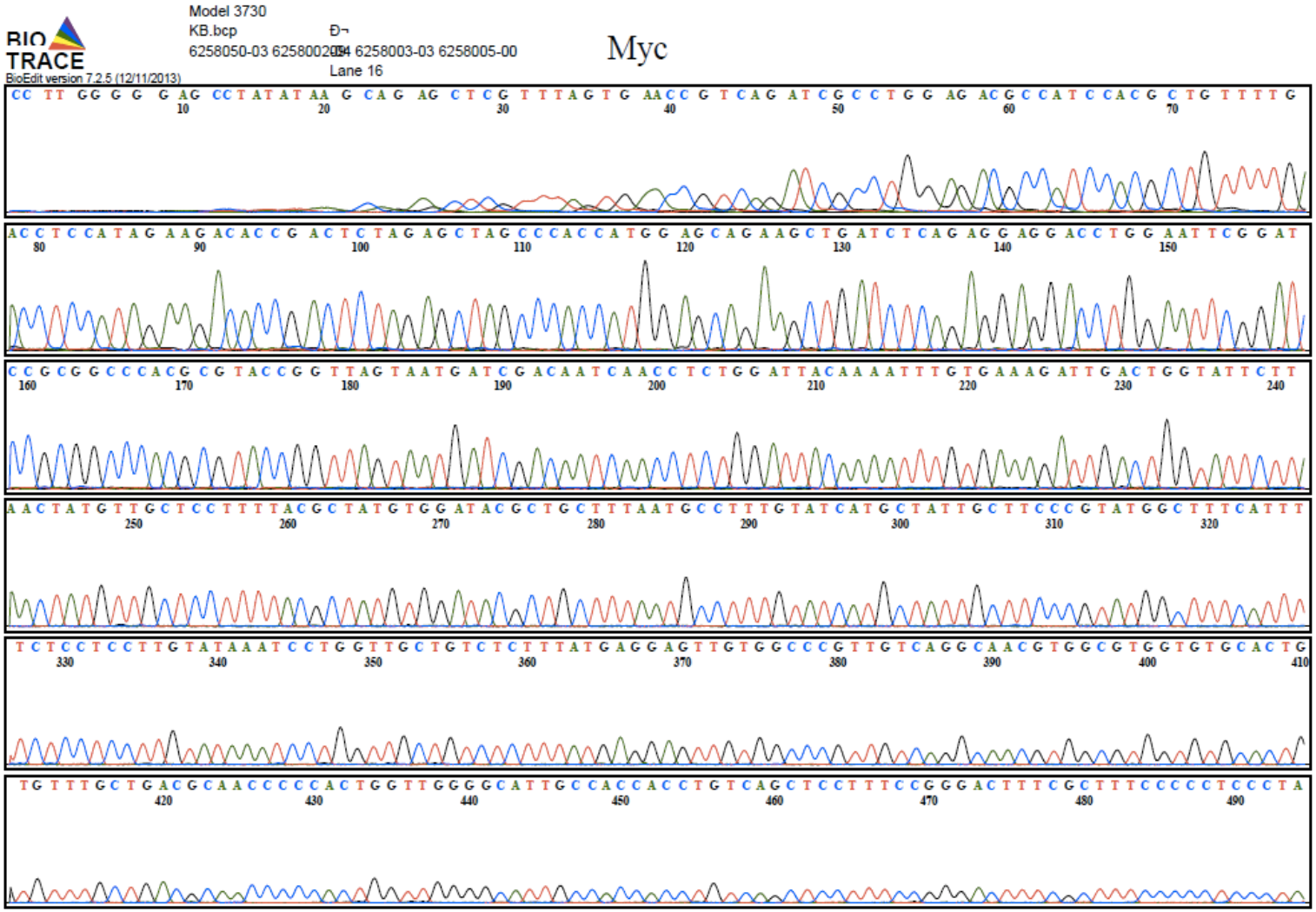
Homo sapiens DNA methyltransferase 3 like (DNMT3L), transcript variant 1, mRNA
Sequence ID: NM_013369.4Length: 1387Number of Matches: 1

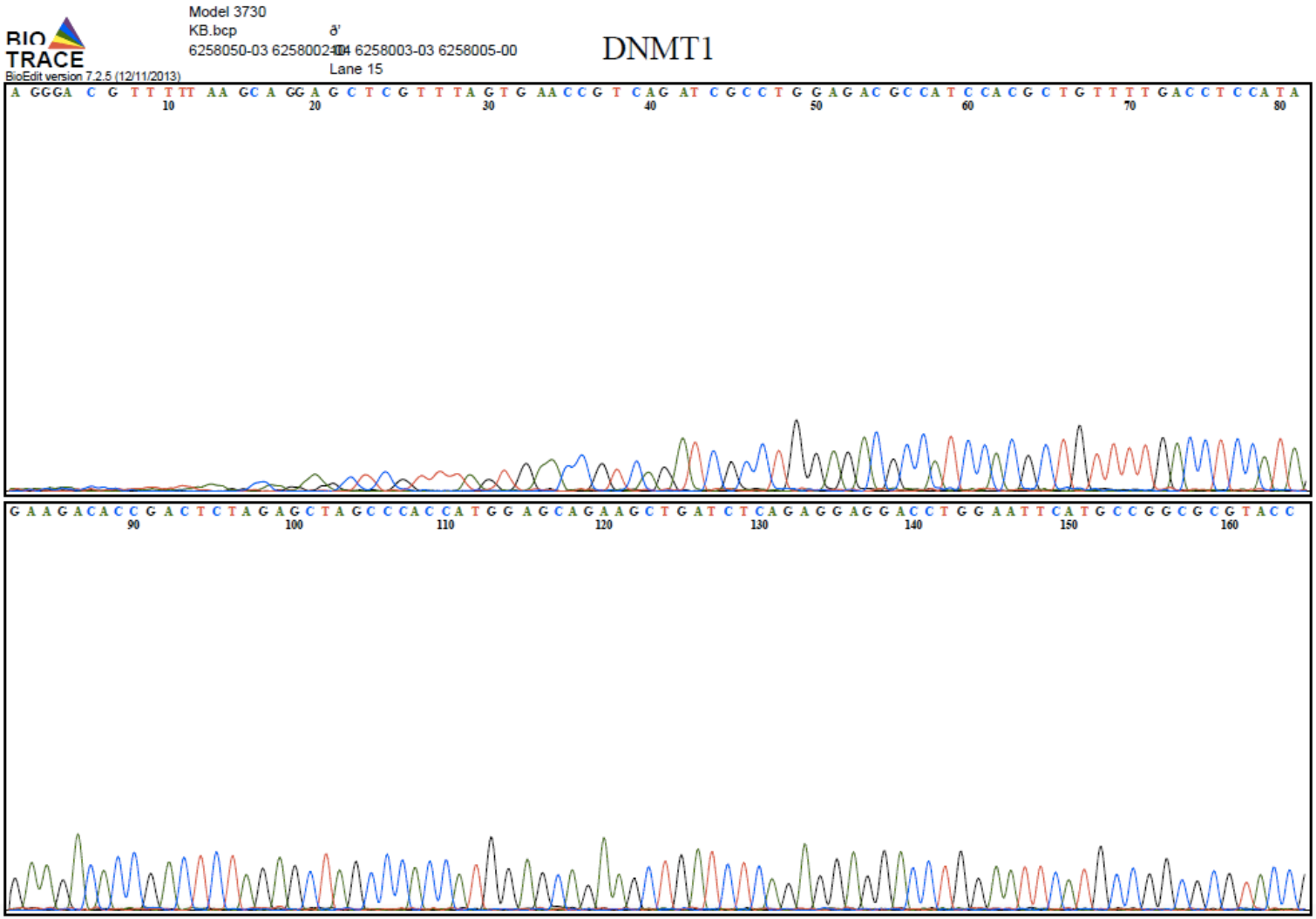
Score	Expect	Identities
1650 bits(893)	0.0	898/900(99%)

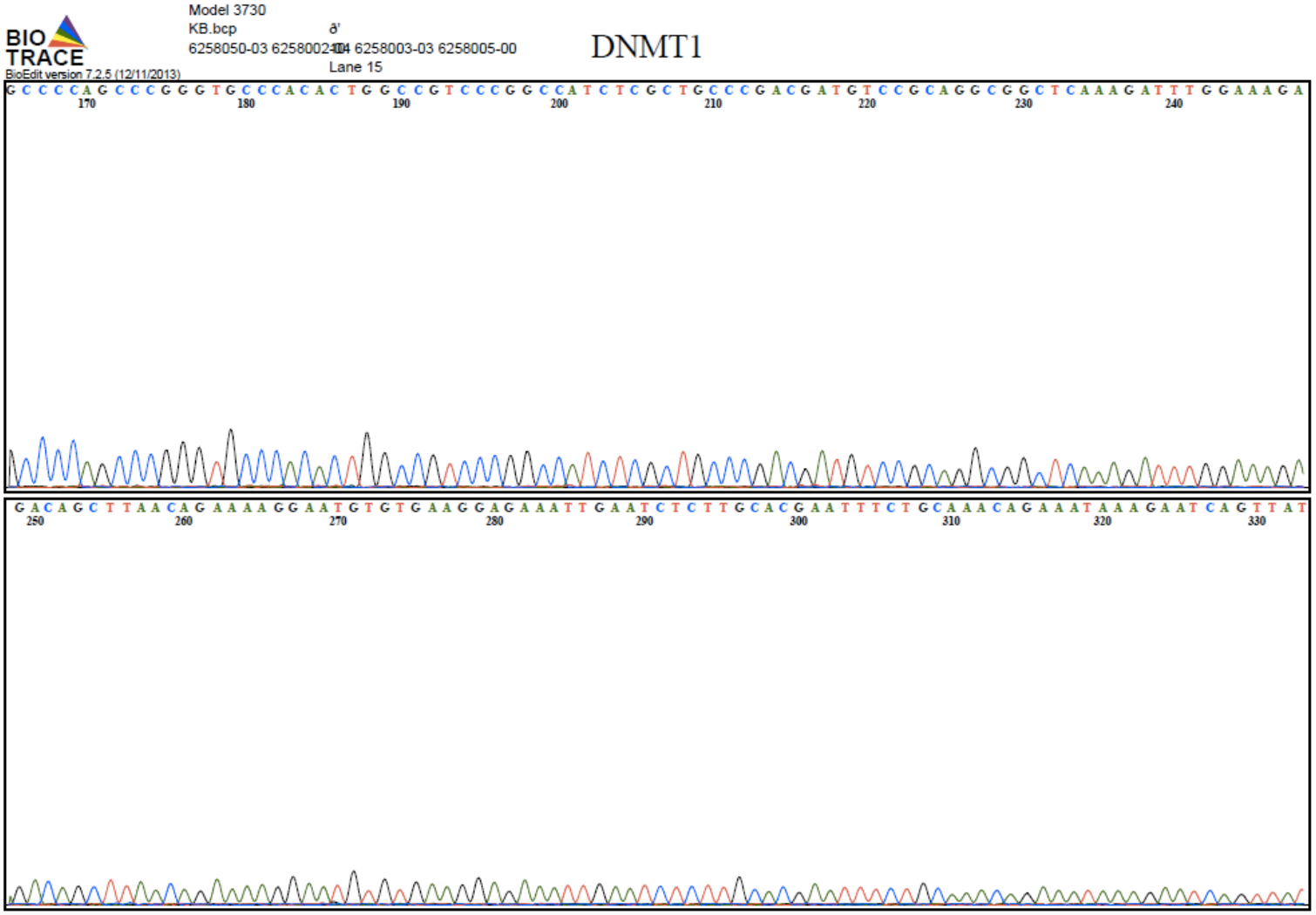
```
Query  1     GCGGCCATCCCAGCCCTGGACCCAGAGGCCGAGCCCAGCATGGACGTGATTTTGGTGGGA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  169   GCGGCCATCCCAGCCCTGGACCCAGAGGCCGAGCCCAGCATGGACGTGATTTTGGTGGGA  228

Query  61    TCCAGTGAGCTCTCAAGCTCCGTTTCACCCGGGACAGGCAGAGATCTTATTGCATATGAA  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  229   TCCAGTGAGCTCTCAAGCTCCGTTTCACCCGGGACAGGCAGAGATCTTATTGCATATGAA  288

Query  121   GTCAAGGCTAACCAGCGAAATATAGAAGACATCTGCATCTGCTGCGGAAGTCTCCAGGTT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  289   GTCAAGGCTAACCAGCGAAATATAGAAGACATCTGCATCTGCTGCGGAAGTCTCCAGGTT  348

Query  181   CACACACAGCACCCTCTGTTTGAGGGAGGGATCTGCGCCCCATGTAAGGACAAGTTCCTG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  349   CACACACAGCACCCTCTGTTTGAGGGAGGGATCTGCGCCCCATGTAAGGACAAGTTCCTG  408

Query  241   GATGCCCTCTTCCTGTACGACGATGACGGGTACCAATCCTACTGCTCCATCTGCTGCTCC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  409   GATGCCCTCTTCCTGTACGACGATGACGGGTACCAATCCTACTGCTCCATCTGCTGCTCC  468

Query  301   GGAGAGACGCTGCTCATCTGCGGAAACCCTGATTGCACCCGATGCTACTGCTTCGAGTGT  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  469   GGAGAGACGCTGCTCATCTGCGGAAACCCTGATTGCACCCGATGCTACTGCTTCGAGTGT  528

Query  361   GTGGATAGCCTGGTCGGCCCCGGGACCTCGGGGAAGGTGCACGCCATGAGCAACTGGGTG  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  529   GTGGATAGCCTGGTCGGCCCCGGGACCTCGGGGAAGGTGCACGCCATGAGCAACTGGGTG  588

Query  421   TGCTACCTGTGCCTGCCGTCCTCCCGAAGCGGGCTGCTGCAGCGTCGGAGGAAGTGGCGC  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  589   TGCTACCTGTGCCTGCCGTCCTCCCGAAGCGGGCTGCTGCAGCGTCGGAGGAAGTGGCGC  648

Query  481   AGCCAGCTCAAGGCCTTCTACGACCGAGAGTCGGAGAATCCCCTTGAGATGTTCGAAACC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  649   AGCCAGCTCAAGGCCTTCTACGACCGAGAGTCGGAGAATCCCCTTGAGATGTTCGAAACC  708
```

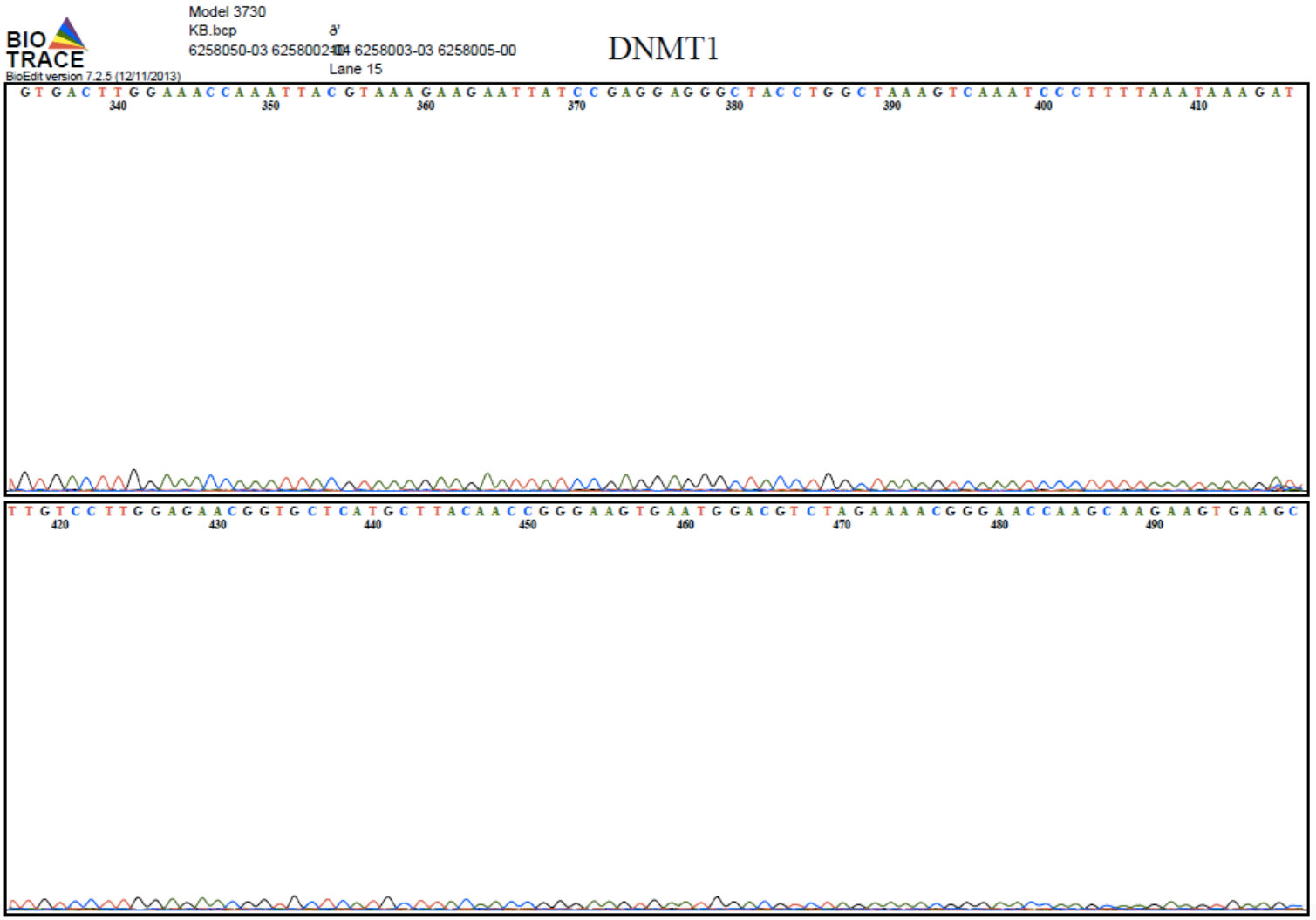
```
Query  541   GTGCCTGTGTGGAGGAGACAGCCAGTCCGGGTGCTGTCCCTTTTTGAAGACATCAAGAAA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  709   GTGCCTGTGTGGAGGAGACAGCCAGTCCGGGTGCTGTCCCTTTTTGAAGACATCAAGAAA  768

Query  601   GAGCTGACGAGTTTGGGCTTTTTGGAAAGTGGTTCTGACCCGGGACAACTGAAGCATGTG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  769   GAGCTGACGAGTTTGGGCTTTTTGGAAAGTGGTTCTGACCCGGGACAACTGAAGCATGTG  828

Query  661   GTTGATGTCACAGACACAGTGAGGAAGGATGTGGAGGAGTGGGGACCCTTCGATCTTGTG  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  829   GTTGATGTCACAGACACAGTGAGGAAGGATGTGGAGGAGTGGGGACCCTTCGATCTTGTG  888

Query  721   TACGGCGCCACACCTCCCCTGGGCCACACCTGTGACCGTCCTCCCAGCTGGTACCTGTTC  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  889   TACGGCGCCACACCTCCCCTGGGCCACACCTGTGACCGTCCTCCCAGCTGGTACCTGTTC  948

Query  781   CAGTTCCACCGGCTCCTGCAGTACGCACGGCCCAAGCCAGGCAGCCCCAGGCCCTTCTTC  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  949   CAGTTCCACCGGCTCCTGCAGTACGCACGGCCCAAGCCAGGCAGCCCCAGGCCCTTCTTC  1008

Query  841   TGGATGTTCGTGGACAATCTGGTGCTGAAACAAGGAAGACCTGGAACGTCGCATCTCGCT  900
             |||||||||||||||||||||||||| ||||||||||||||| ||||||||||||||||
Sbjct  1009  TGGATGTTCGTGGACAATCTGGTGCTG-AACAAGGAAGACCTGG-ACGTCGCATCTCGCT  1066
```

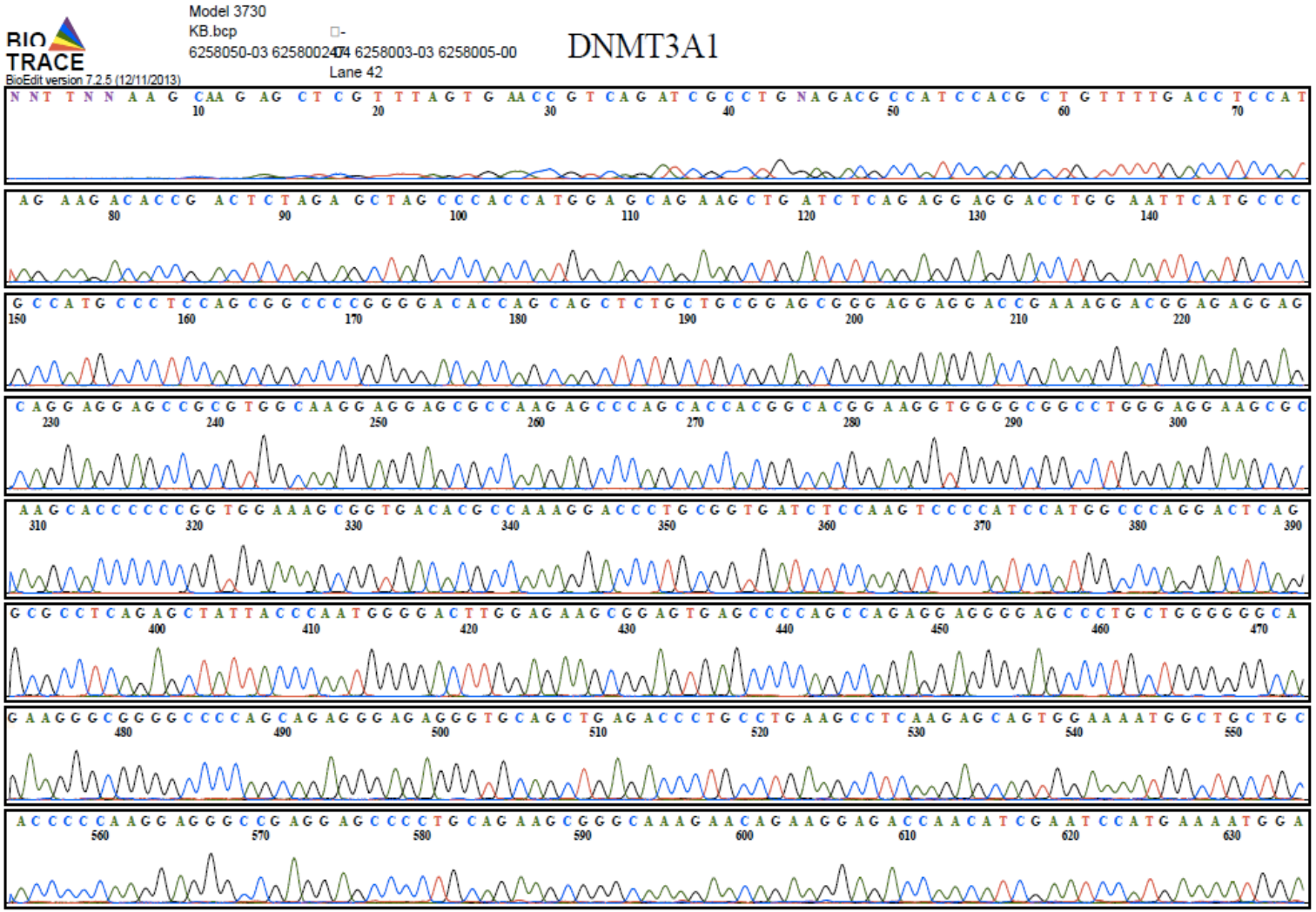
Appendix B: Sequencing by Sanger analysis of each DNMT isoform and alignment analysis

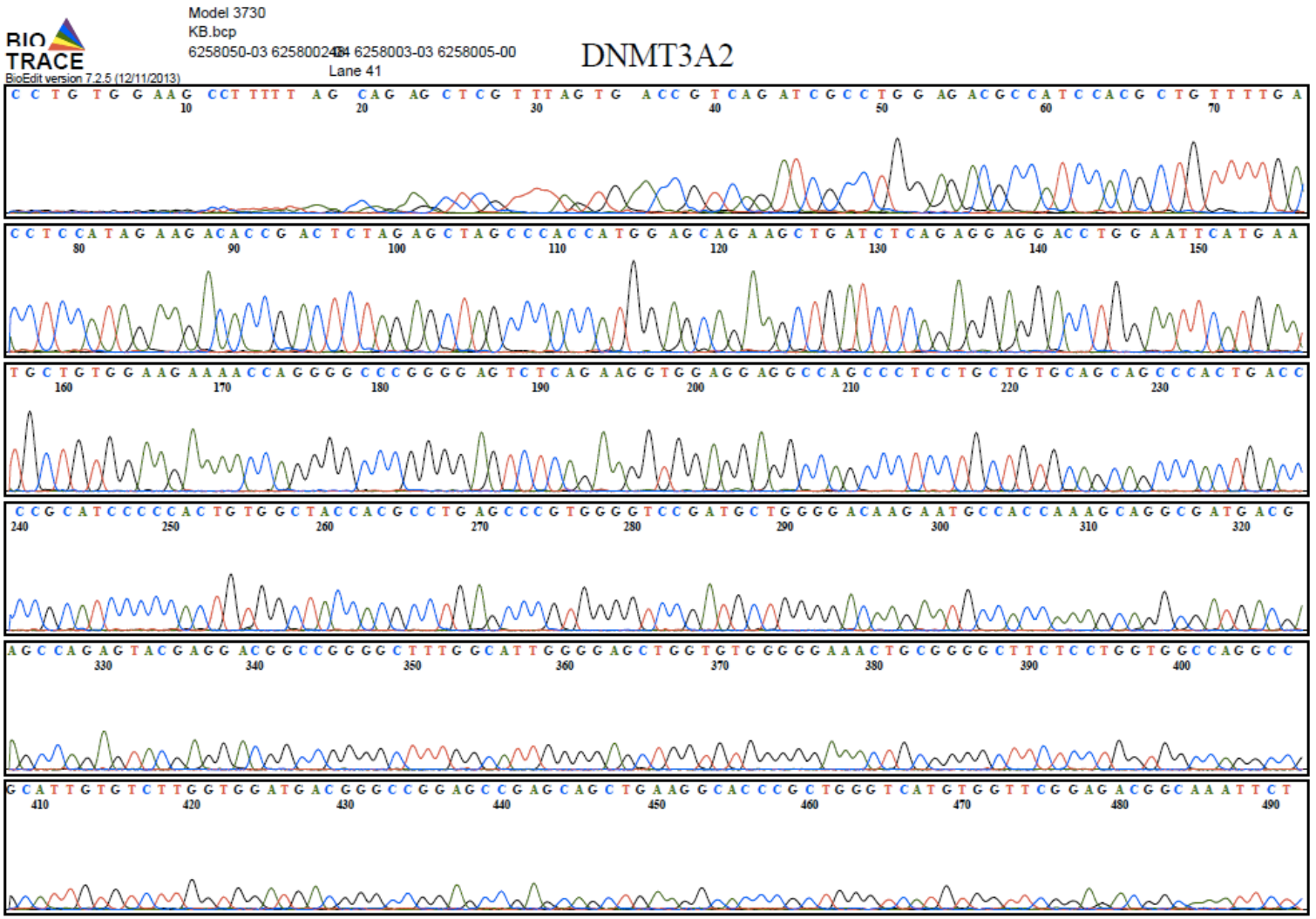


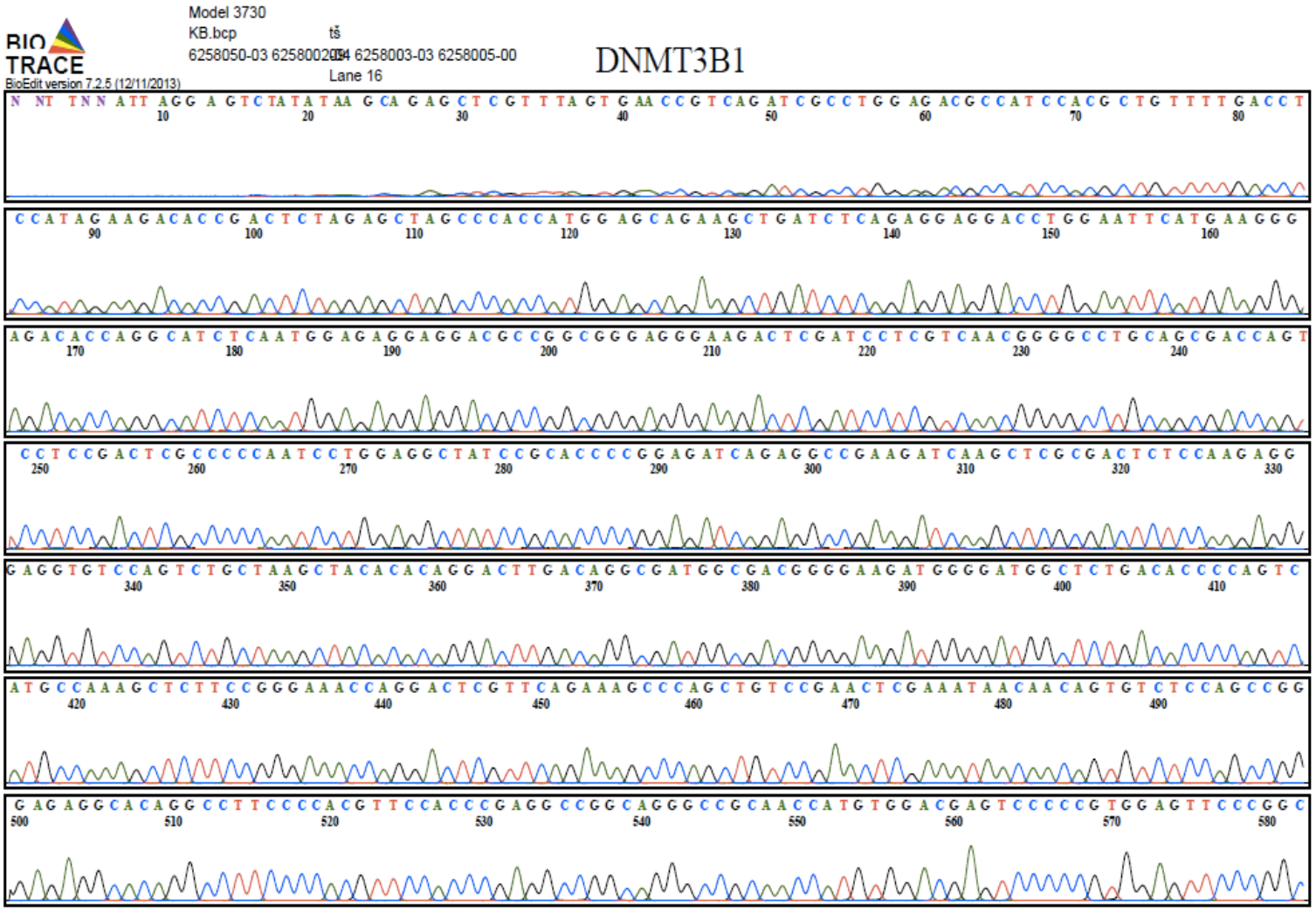

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 15
DNMT1

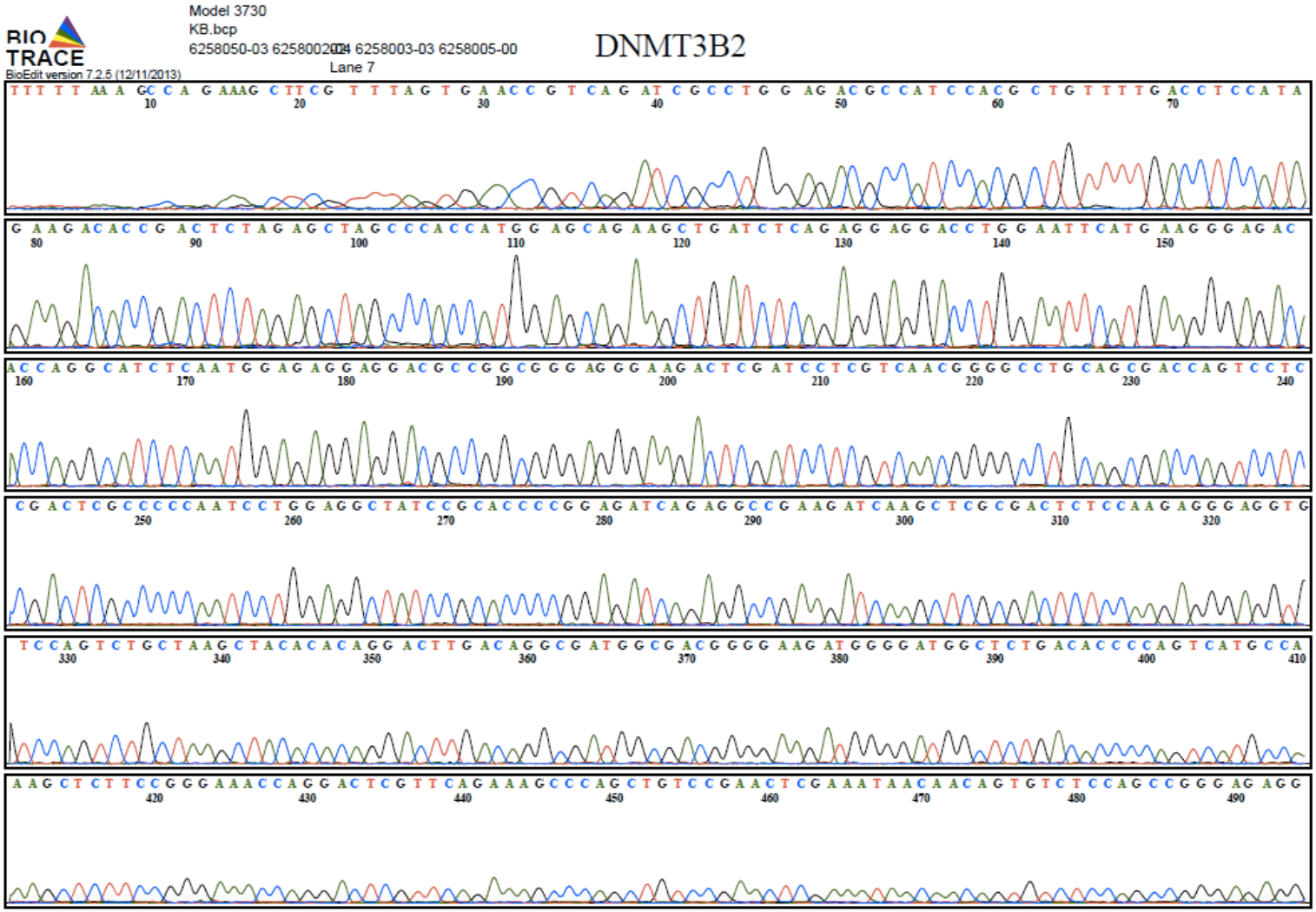

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 15
DNMT1
GCCCCAGCCCGGGTGCCCACACTGGCCGTCCCGGCCATCTCGCTGCCCGACGATGTCCGCAGGCGGCTCAAAGATTTGGAAAGA
170 180 190 200 210 220 230 240
GACAGCTTAACAGAAAAGGAATGTGTGAAGGAGAAATTGAATCTCTTGCACGAATTTCTGCAAACAGAAATAAAGAATCAGTTAT
250 260 270 280 290 300 310 320 330

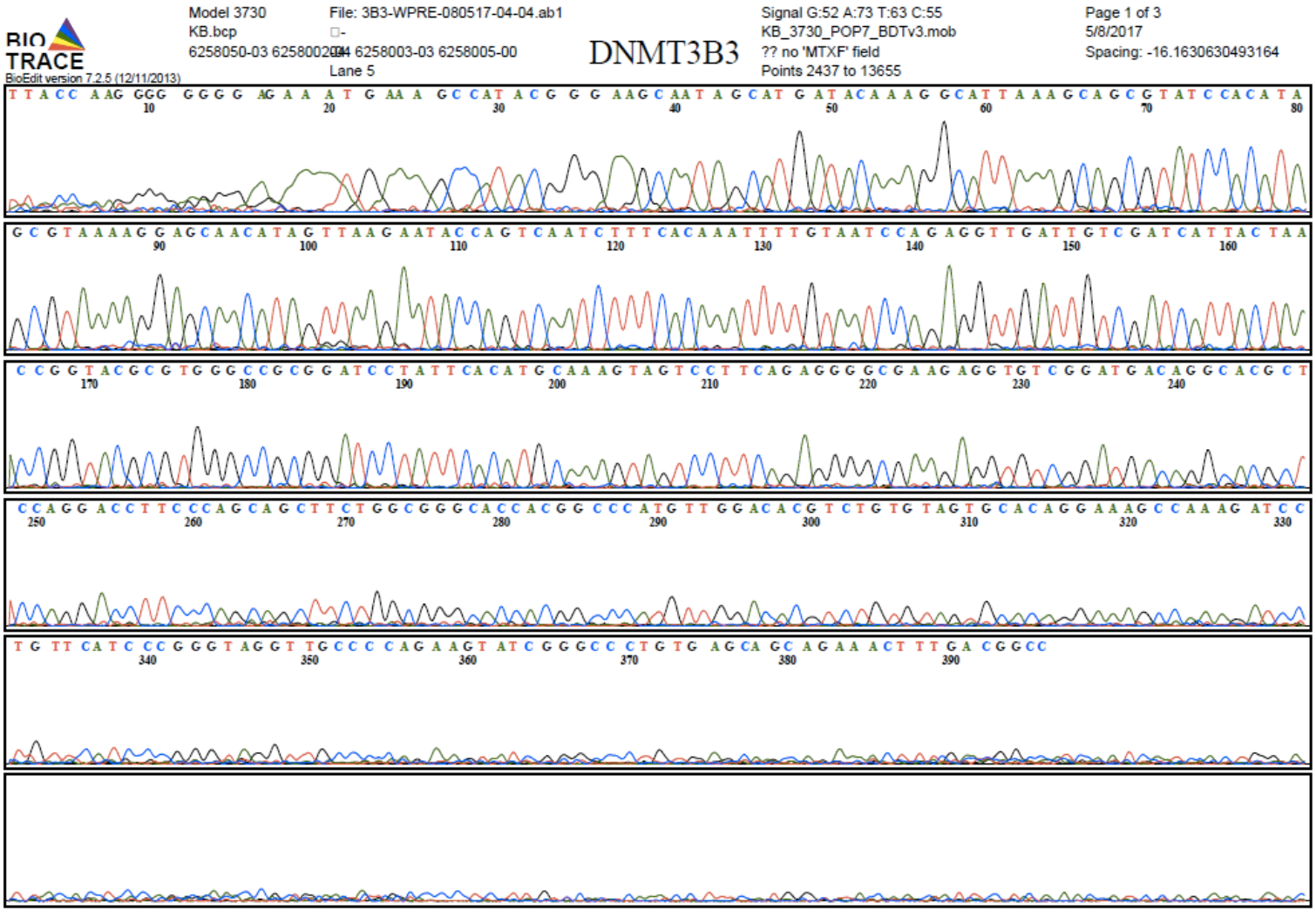

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 15
DNMT1
GTGACTTGGAAACCAAATTACGTAAAGAAGAATTATCCGAGGAGGGCTACCTGGCTAAAGTCAAATCCCTTTTAAATAAAGAT
340 350 360 370 380 390 400 410
TTGTCCTTGGAGAACGGTGCTCATGCTTACAACCGGGAAGTGAATGGACGTCTAGAAAACGGGAACCAAGCAAGAAGTGAAGC
420 430 440 450 460 470 480 490

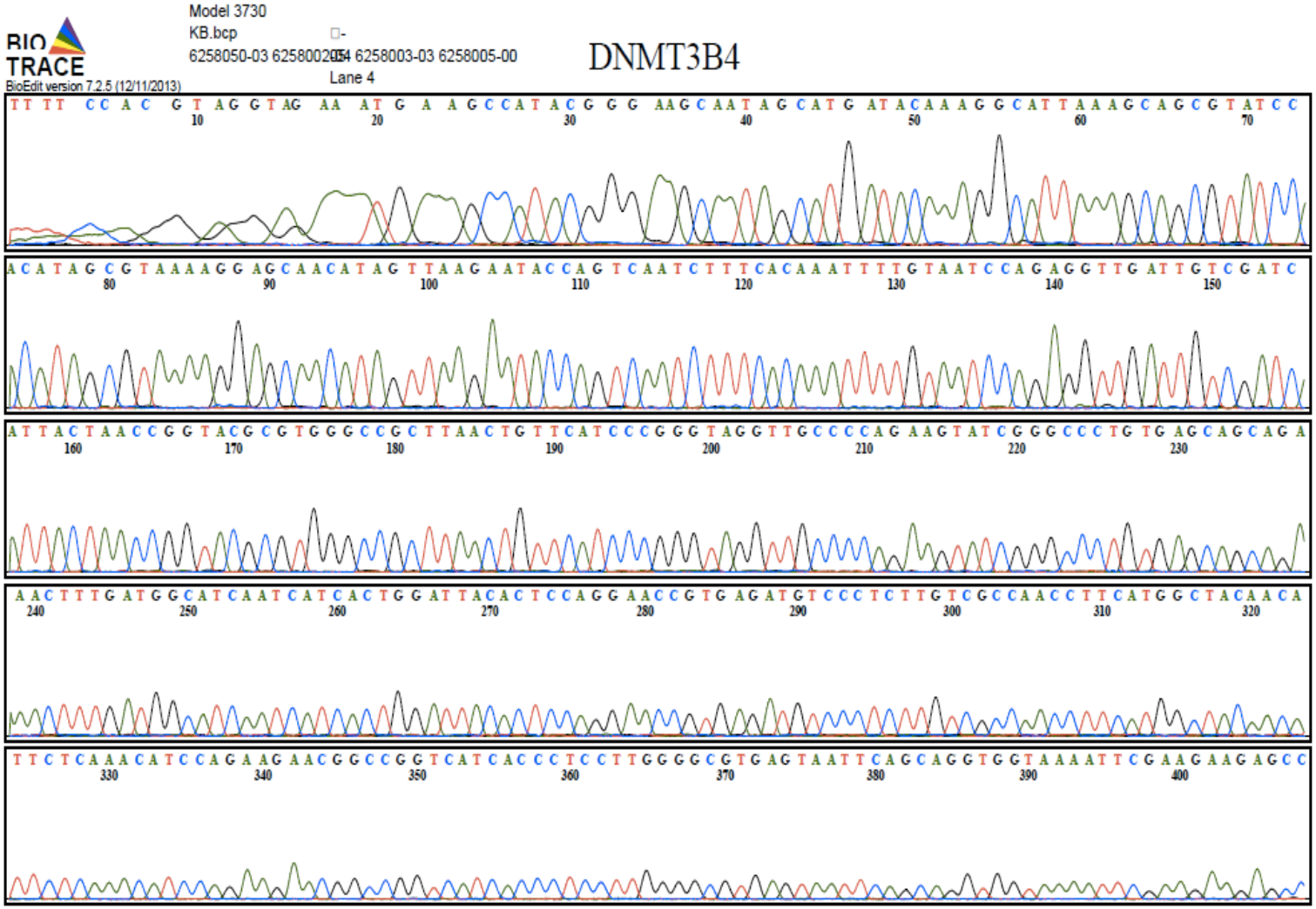

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 42
DNMT3A1

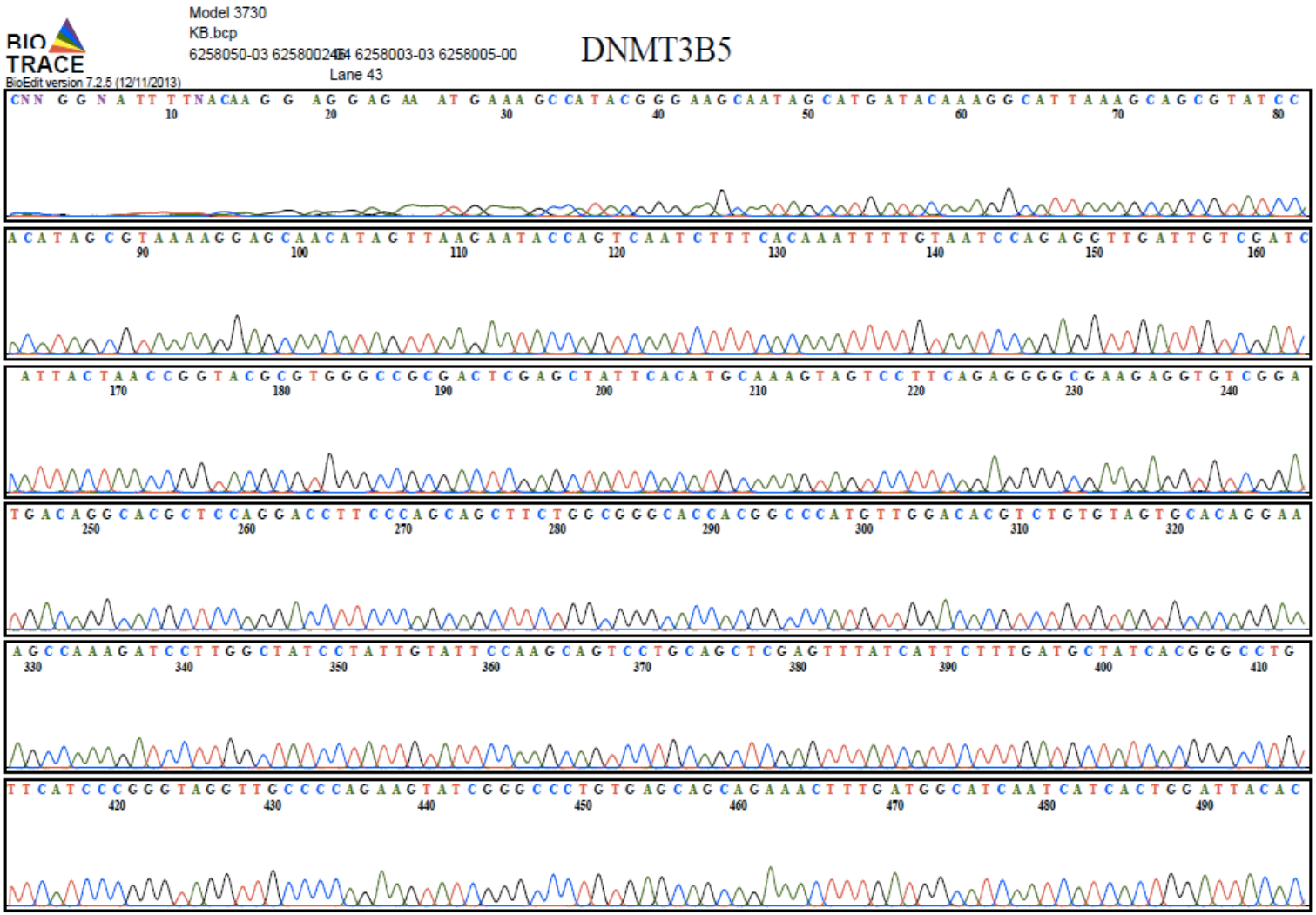

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 41
DNMT3A2

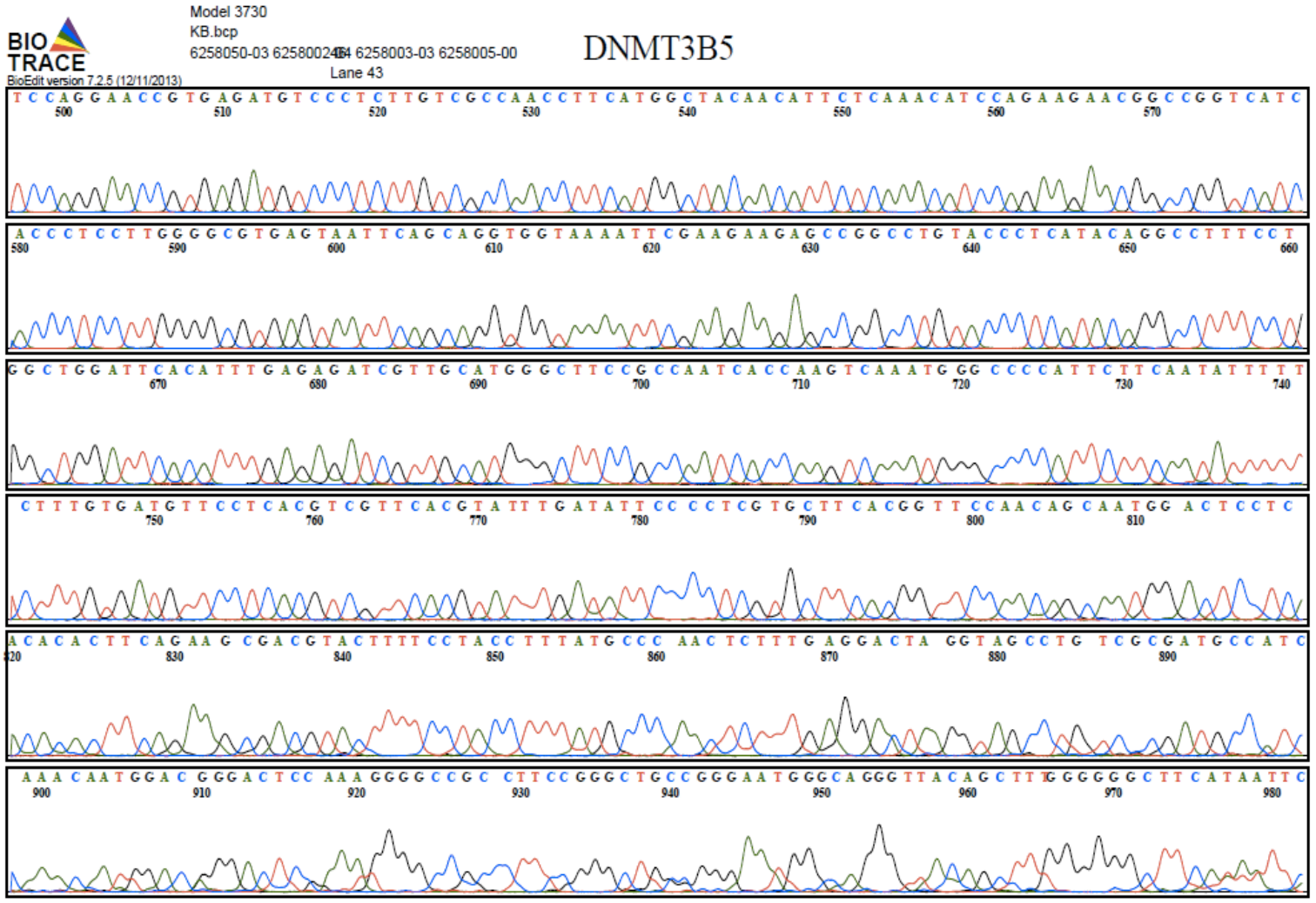

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258003-03 6258005-00
Lane 16
DNMT3B1

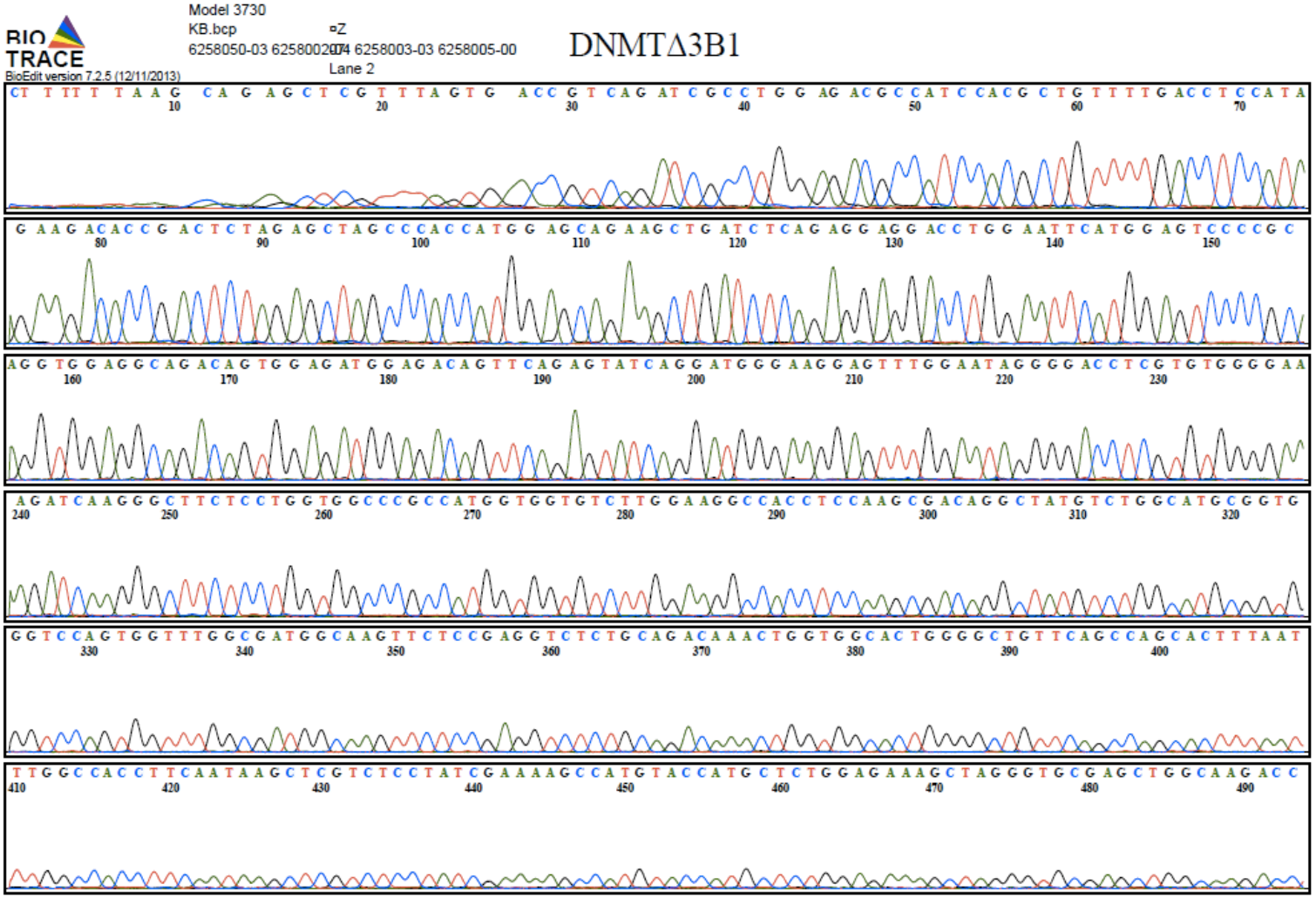

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 7
DNMT3B2

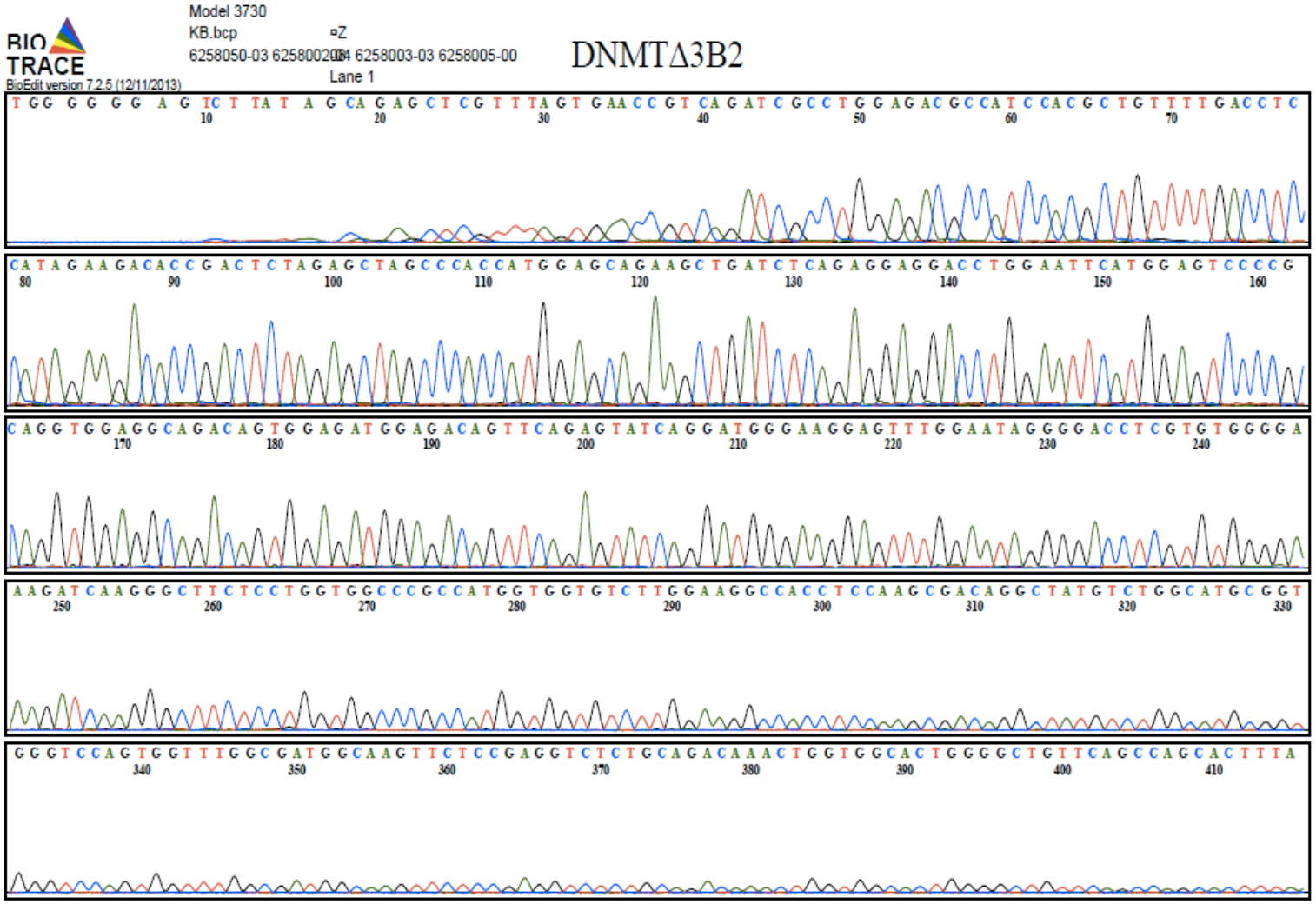

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04
File: 3B3-WPRE-080517-04-04.ab1
6258003-03 6258005-00
Lane 5
DNMT3B3
Signal G:52 A:73 T:63 C:55
KB_3730_POP7_BDTv3.mob
?? no 'MTXF' field
Points 2437 to 13655
Page 1 of 3
5/8/2017
Spacing: -16.1630630493164

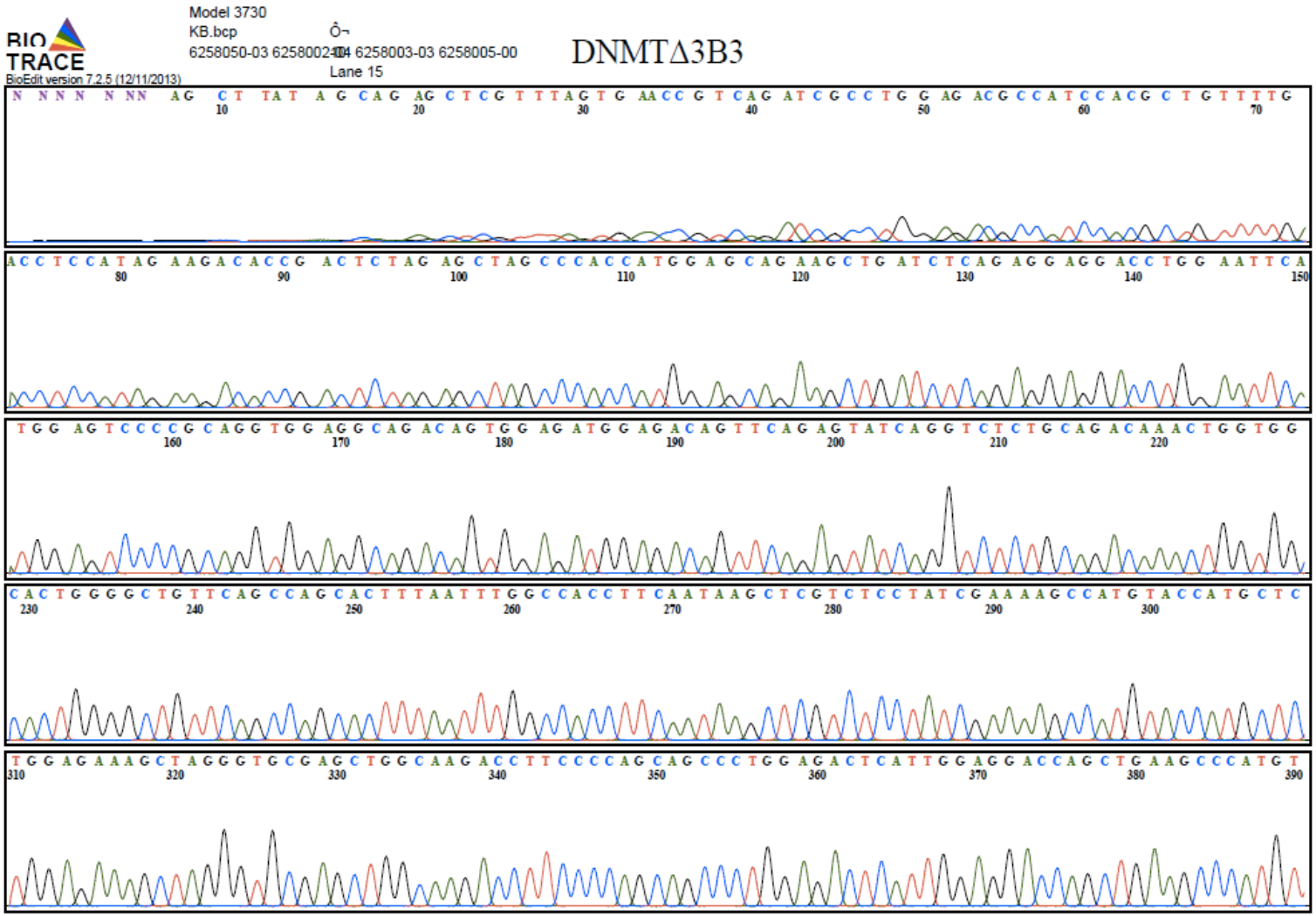

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258003-03 6258005-00
Lane 4
DNMT3B4

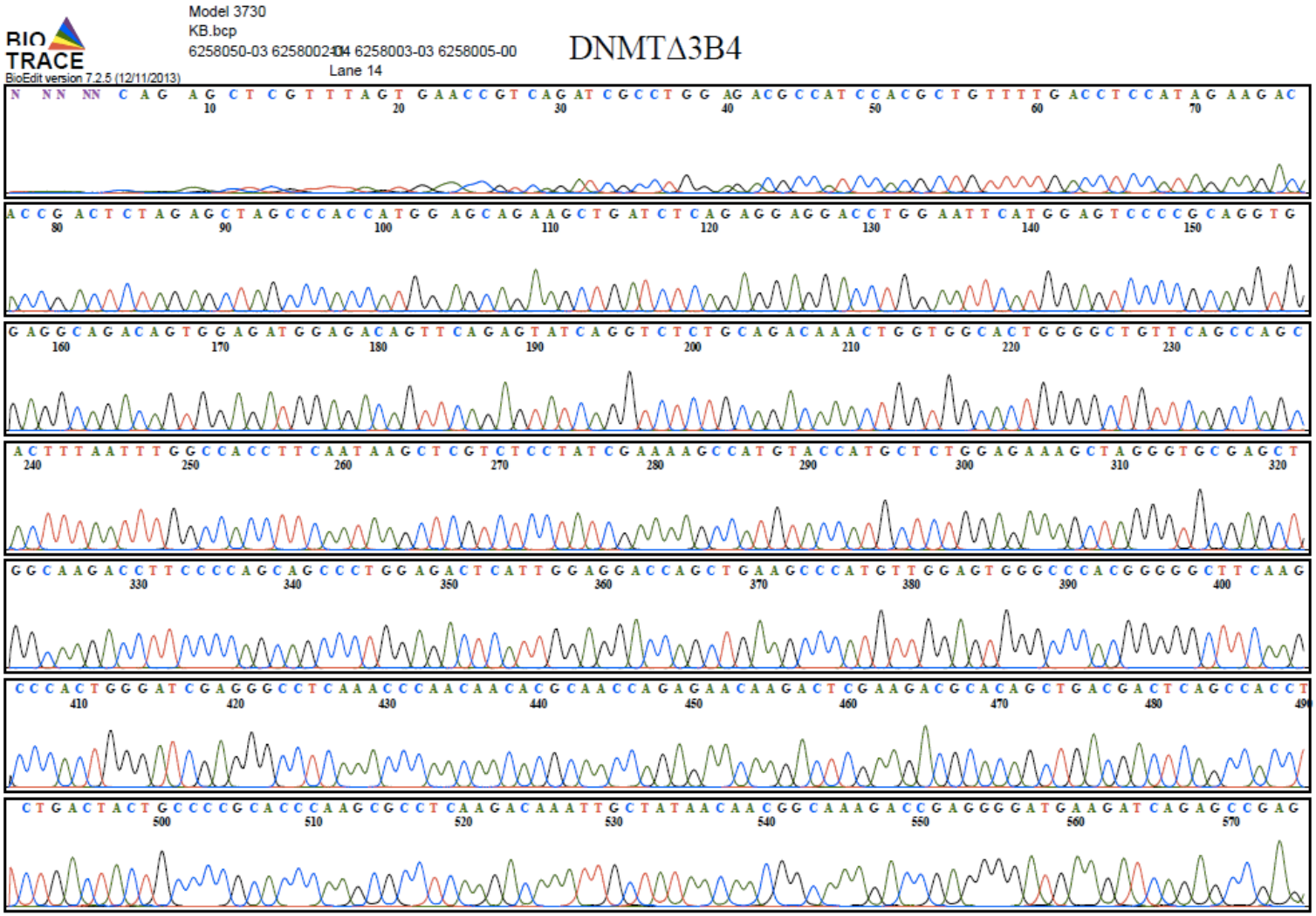

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 43
DNMT3B5

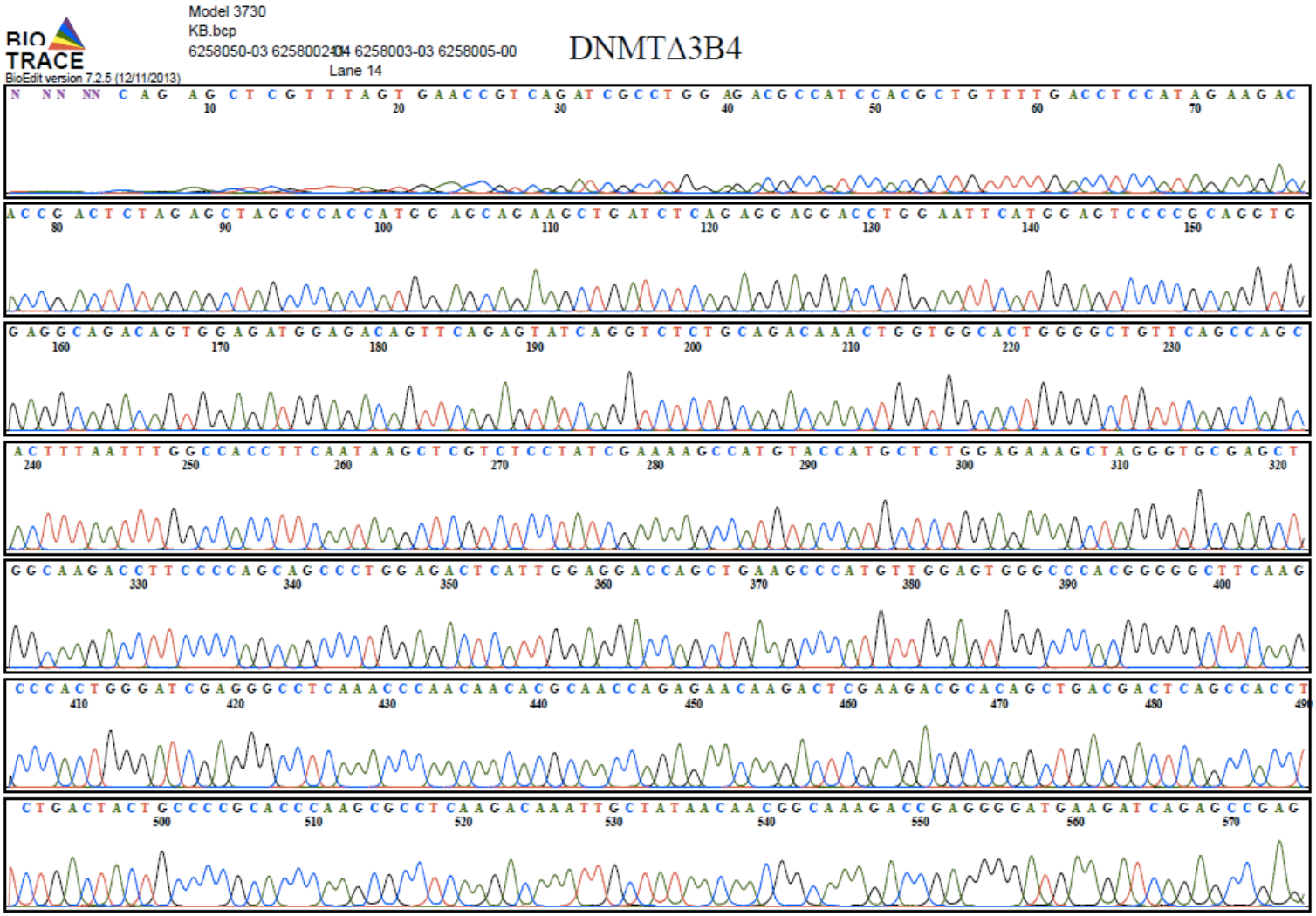

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 43
DNMT3B5

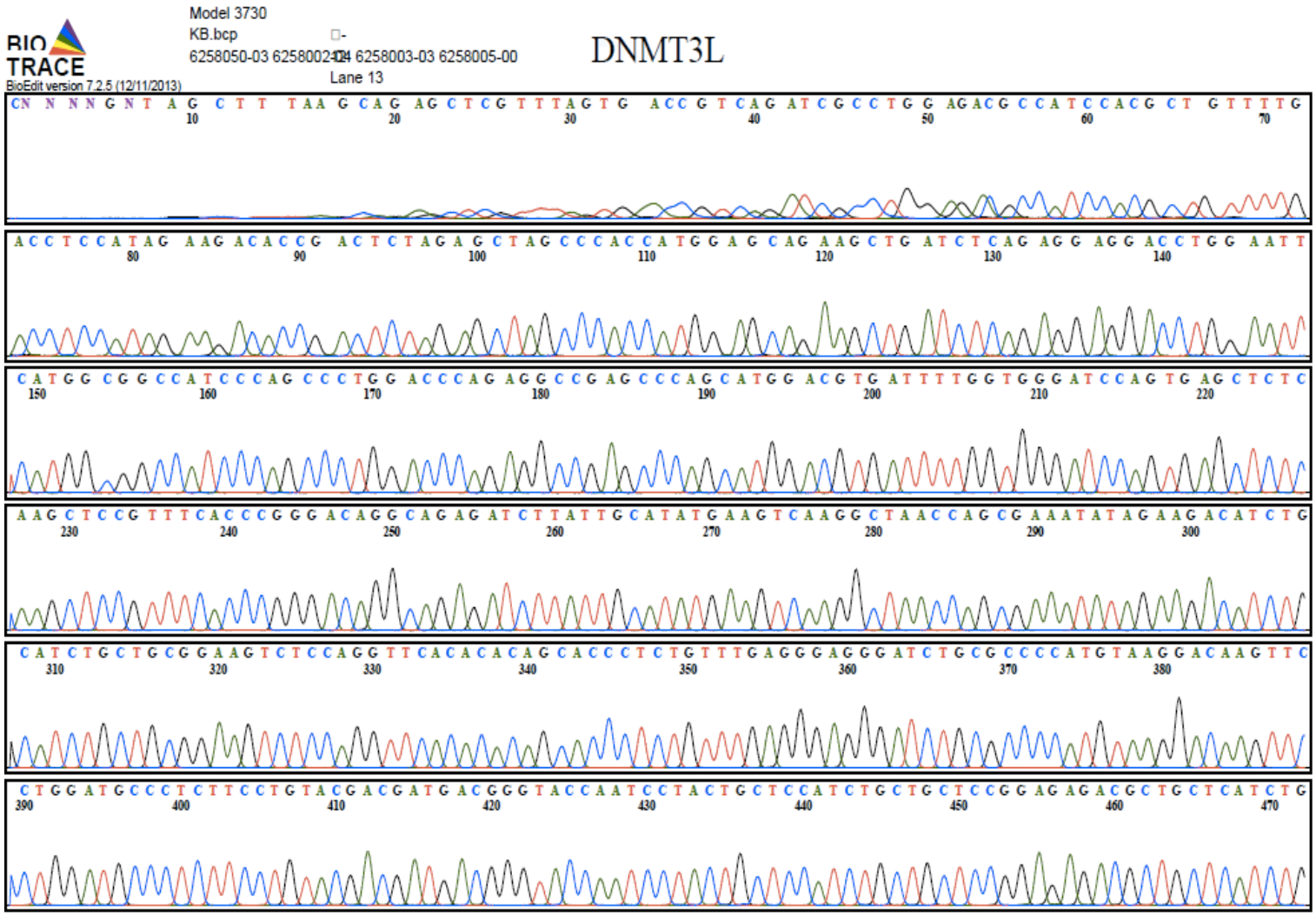

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 2
DNMTΔ3B1

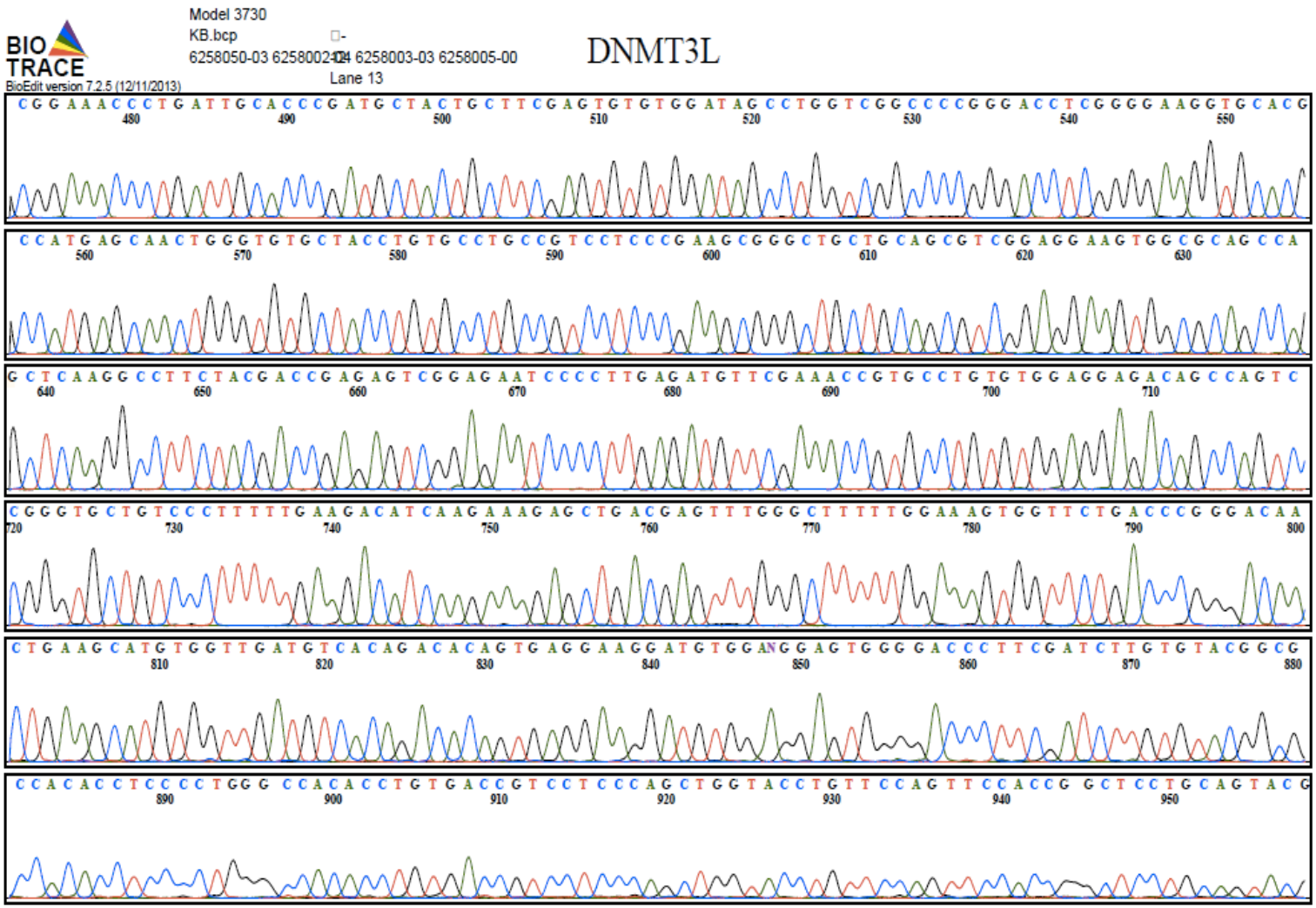

BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258003-03 6258005-00
Lane 1
DNMTΔ3B2


BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 15
DNMTΔ3B3


BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 14
DNMTΔ3B4


BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 14
DNMTΔ3B4


BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 13
DNMT3L


BIO TRACE
BioEdit version 7.2.5 (12/11/2013)
Model 3730
KB.bcp
6258050-03 6258002-04 6258003-03 6258005-00
Lane 13
DNMT3L

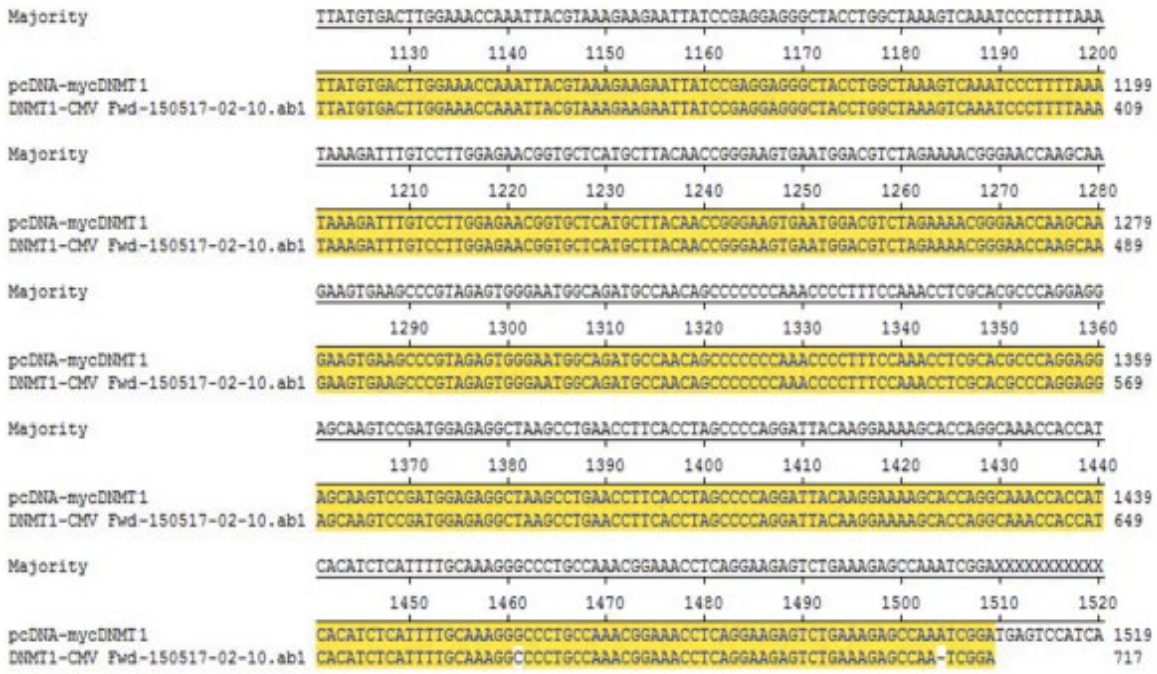


Figure B1 Alignment result between DNMT1 and pcDNA-MycDNMT1 (commercial plasmid). Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

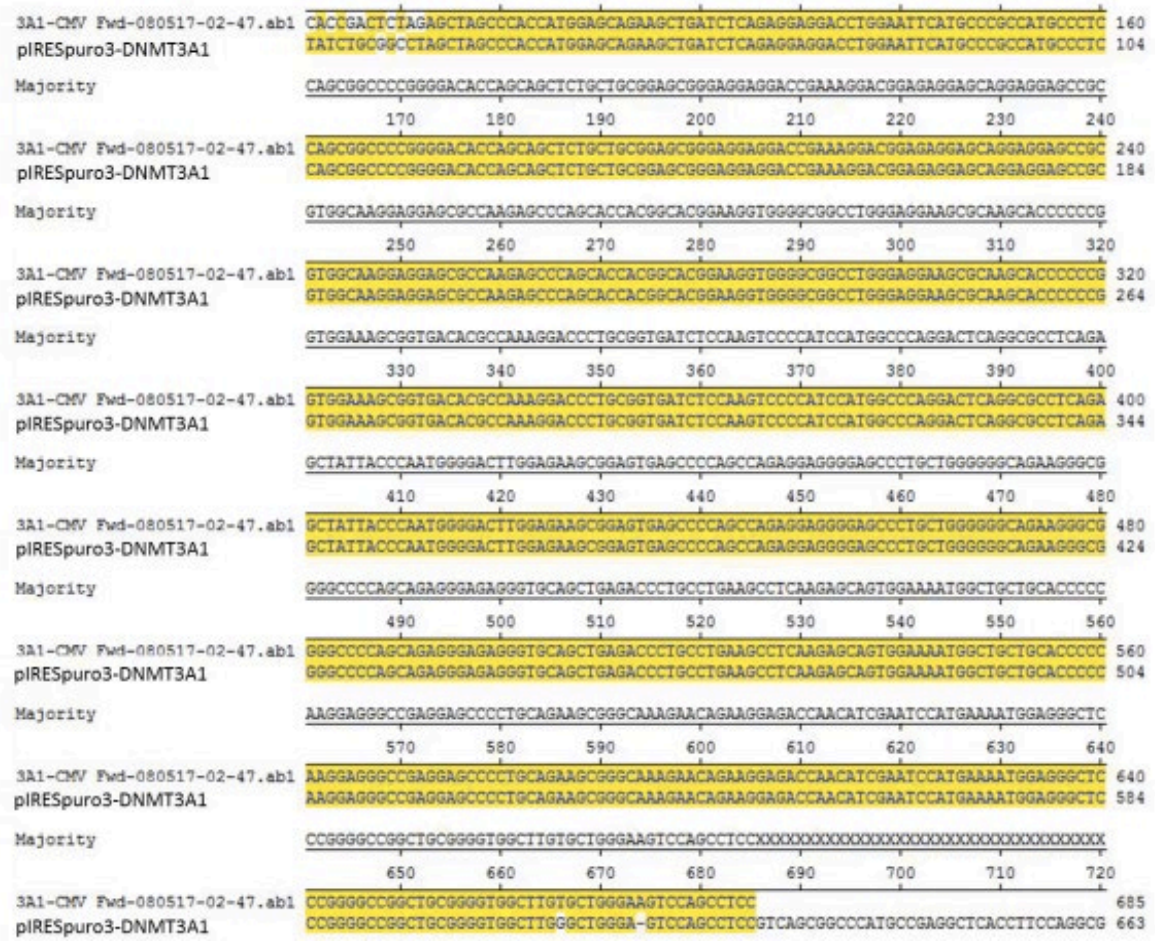


Figure B2 Alignment result between DNMT3A1 and pIRESpuro3-DNMT3A1. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

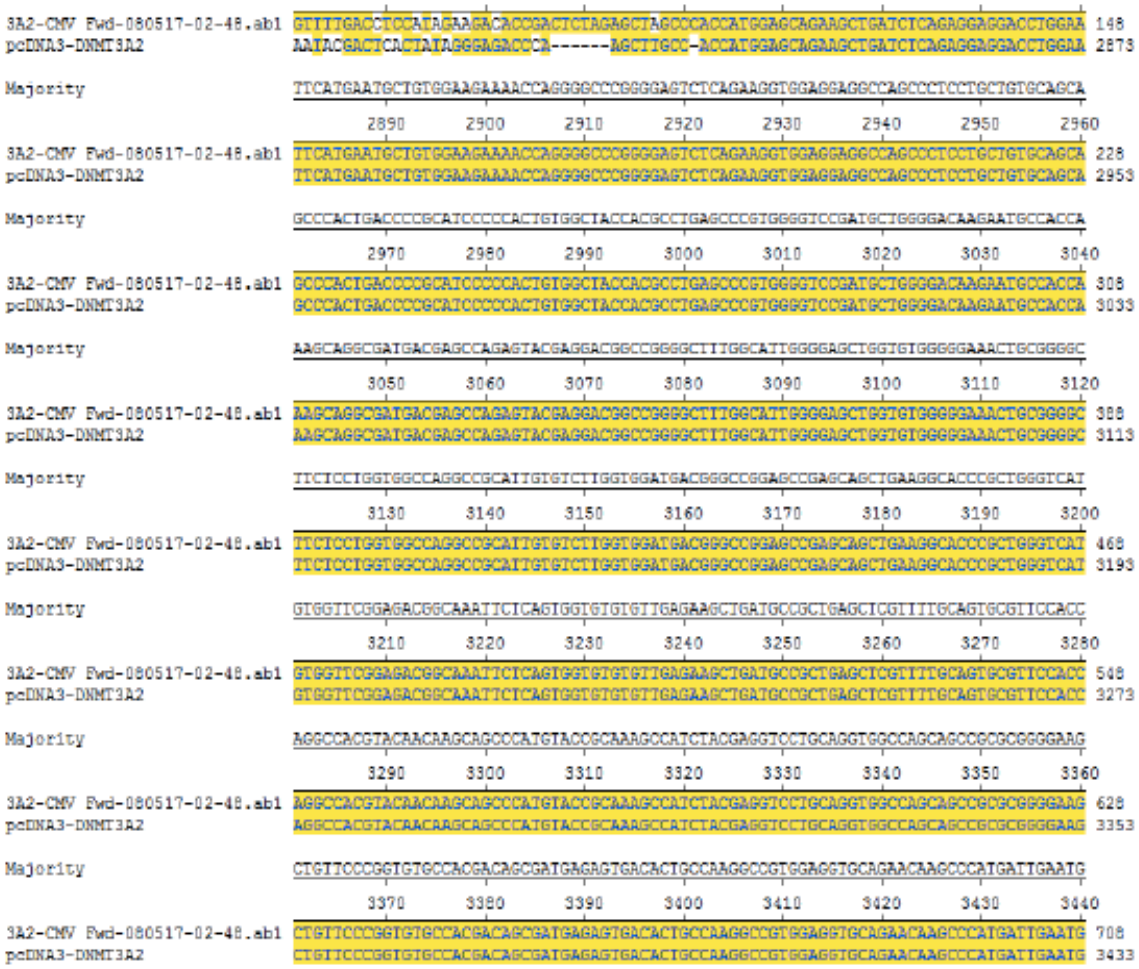


Figure B3 Alignment result between DNMT3A2 and pcDNA3-DNMT3A2. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

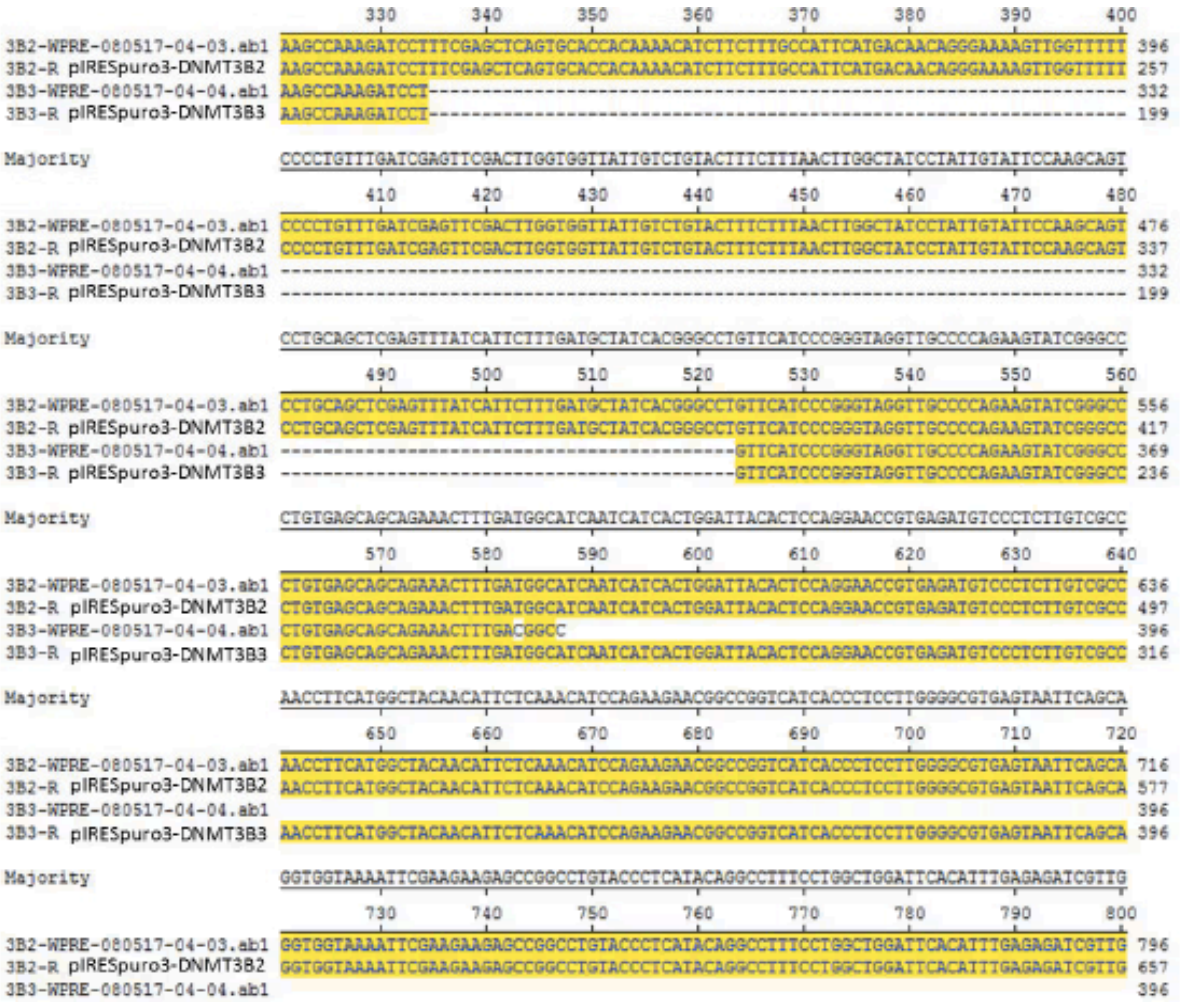


Figure B4 Alignment result among DNMT3B2, pIRESpuro3-DNMT3B2, DNMT3B3, and pIRESpuro3-DNMT3B3. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

Figure B5 Alignment result among DNMT3B2, DNMT3B3, DNMT3B4, and DNMT3B5. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

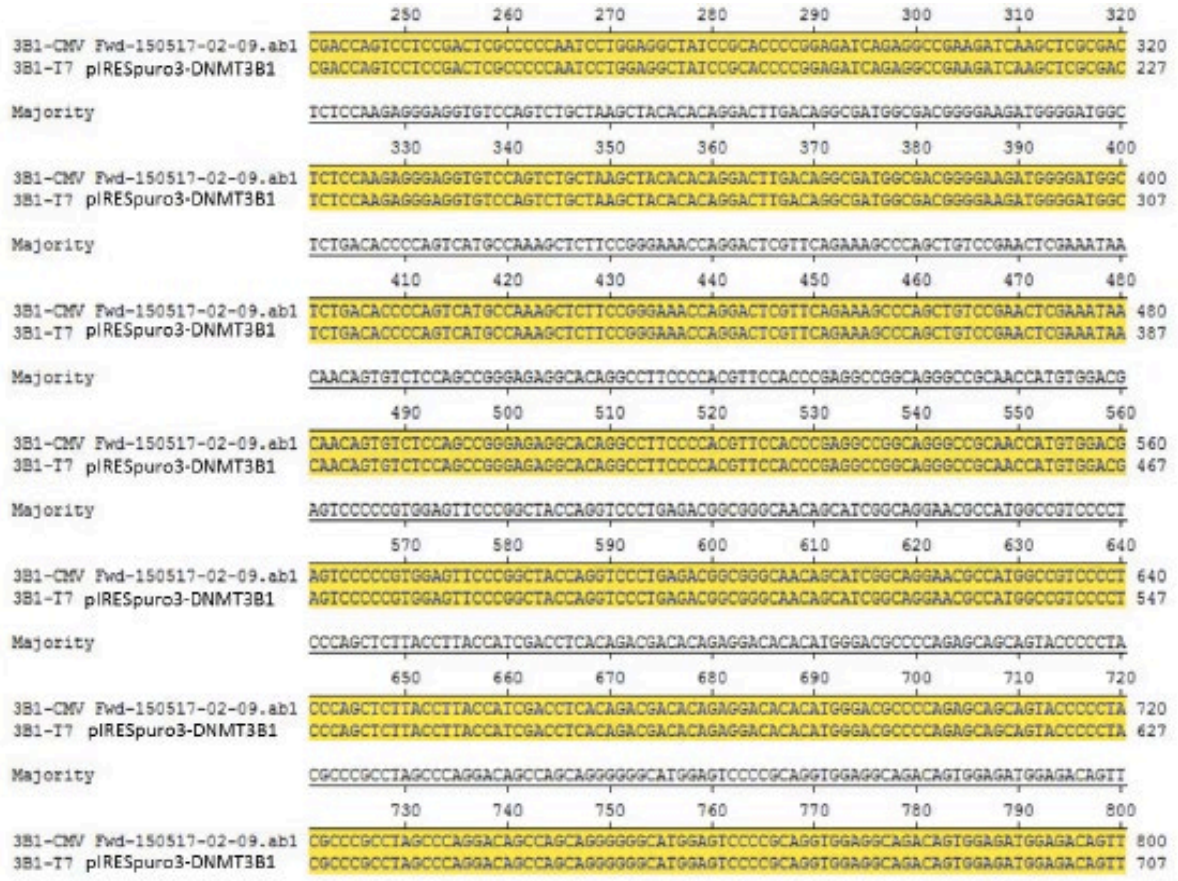


Figure B6 Alignment result between DNMT3B1 and pIRESpuro3-DNMT3B1. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

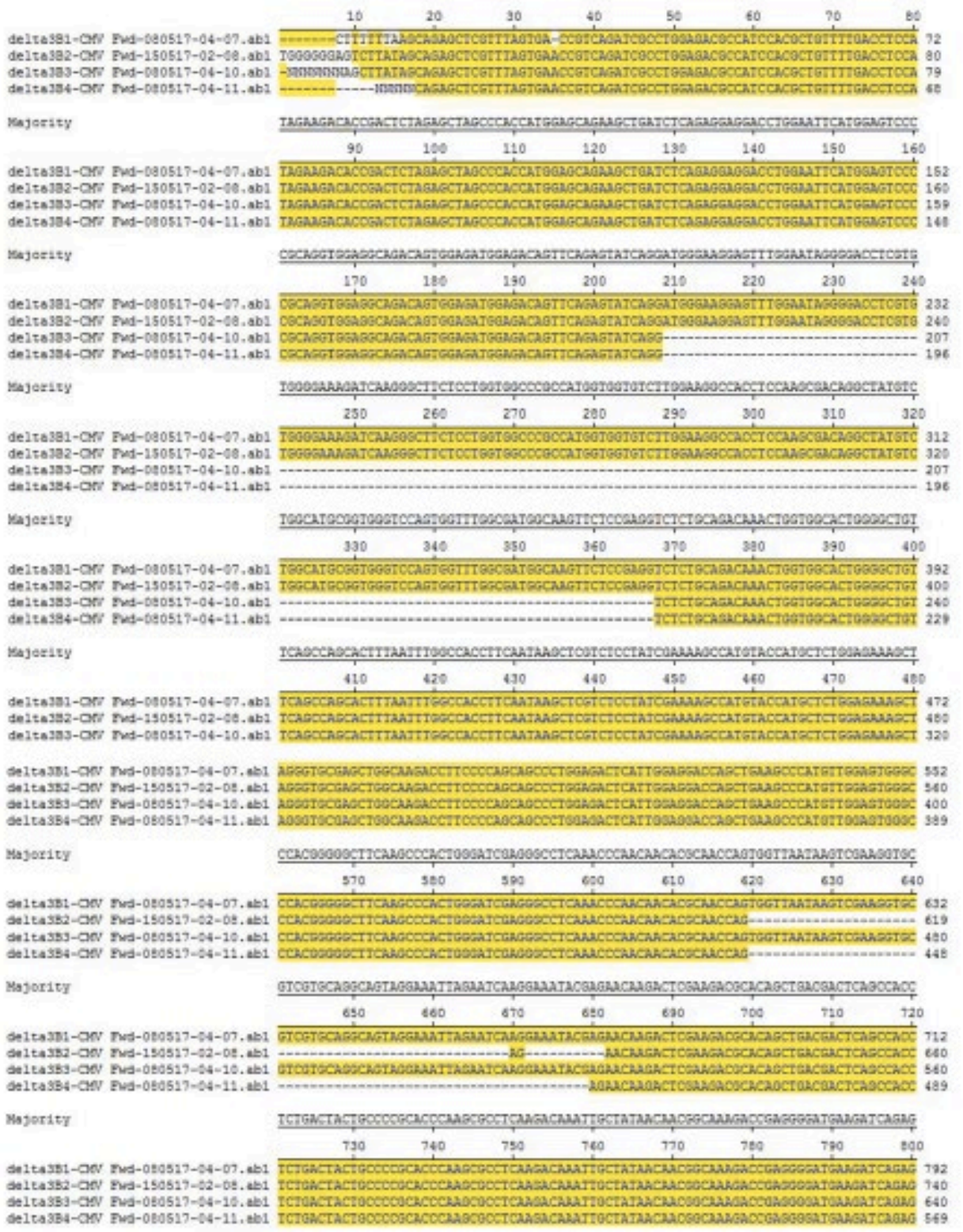


Figure B7 Alignment result among DNMTΔ3B1, DNMTΔ3B2, DNMTΔ3B3, and DNMTΔ3B4. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

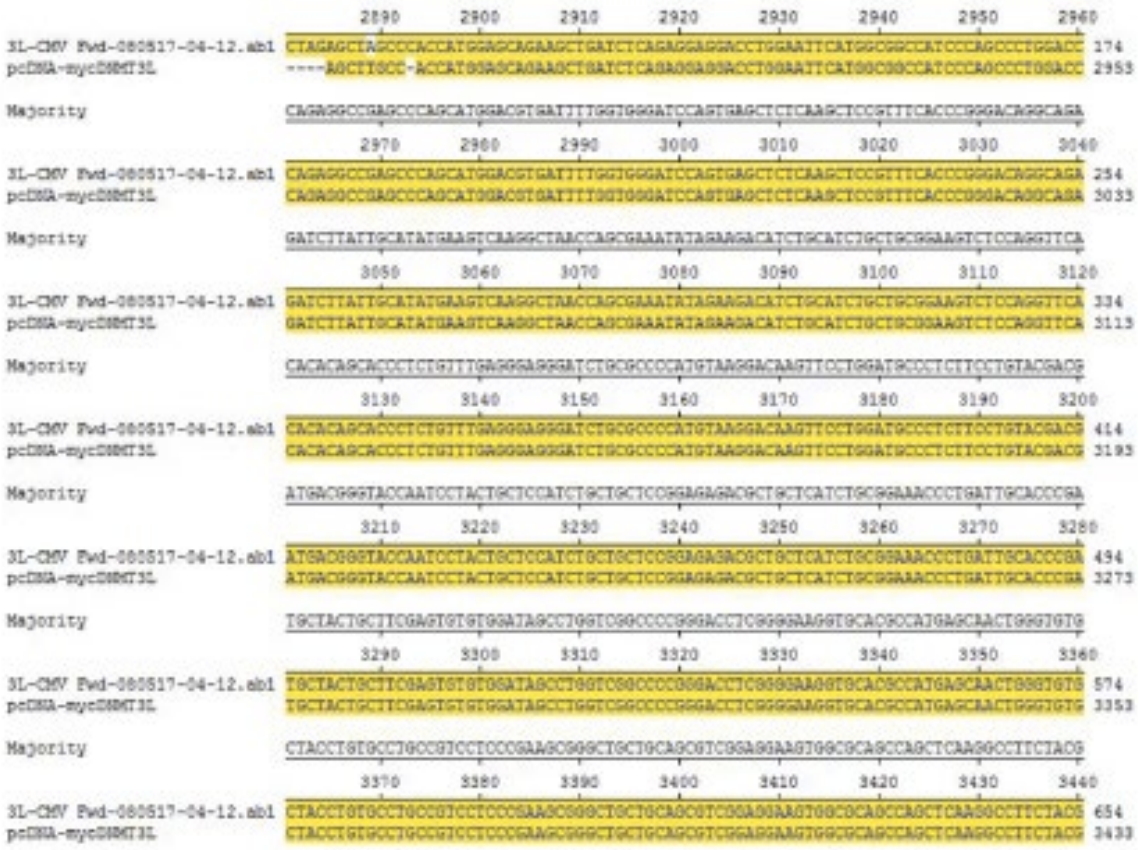


Figure B8 Alignment result between DNMT3L and pcDNA-MycDNMT3L (commercial plasmid). Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

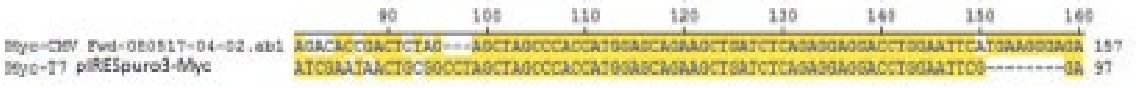


Figure B9 Alignment result between Myc and pIRESpuro3-Myc. Majority is a method to sum the weights of a base in DNA sequences, yellow highlights represent 100% similar sequences with majority, number indicates the number of base pair.

Appendix C: Target genes for each DNMT isoform

Table C1 Target genes of DNMT1

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg19245525	ACACB	**cg08387944**	LINC00964
cg18295758	ADAMTS16	**cg27636676**	LMX1A
cg24985738	ADAMTS2	**cg15671466**	LOC100506457
cg26983228	ARHGAP15	**cg23676348**	
cg05927763	ATXN10	**cg10171489**	LY86
cg24859648	BANF2	**cg23094728**	MAP3K1
cg20943641	CCDC53	**cg23220346**	MIR128-2
cg14085446	CDX1	**cg07284339**	OSBPL5
cg13851989	CER1	**cg25711003**	PHLDB2
cg23013853	CHGA	**cg10207469**	PLX4
cg20348618	CMKLR1	**cg07461239**	PTPRM
cg10985909	CNST	**cg13173924**	SCN1A
cg09185643	CRMP1	**cg16968681**	SHC2
cg26942892	DLGAP2	**cg09868003**	SHC4
cg07611000	EHBP1L1	**cg18886071**	SLAMF1
cg01065960	ERAL1	**cg04914221**	TLN2
cg15865055	FAM55A	**cg15646382**	TMCO6
cg24636477	GDF10	**cg16933440**	VEPH1
cg08053904	GPR139	**cg24885481**	ZNF287
cg20663347	GRIN3A		
cg25751961	KCNJ1		

Table C2 Target genes of DNMT3L

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg08344180	AADACL4	**cg03853861**	EDA2R
cg21211413	ACAN	**cg14435469**	EHHADH-AS1
cg06904000	ACSM5	**cg17313986**	EPHB1
cg24893551	ACVR1C	**cg12665345**	EVC2
cg16754015	ADARB2	**cg06551520**	FAM170B-AS1
cg10533744	ADARB2	**cg09543693**	FAM19A5
cg22069688	ADGRE1	**cg07383415**	FAM49A
cg11425164	AGPAT9	**cg16967830**	GABRA5
cg08624915	AHSP	**cg18172281**	GFRA2
cg03459928	AK7	**cg04096150**	GPR176
cg04188756	AKAP6	**cg25041439**	HYLS1
cg04304338	ANKRD33B	**cg07164224**	IL1R1
cg03501666	ANKS1B	**cg05903710**	IL4R
cg11984673	APBB1IP	**cg02238012**	IRX4
cg20118431	B3GAT1	**cg06828335**	KCNMA1
cg05355436	C10orf107	**cg16324958**	KCNQ1
cg27153327	C16orf73	**cg08947191**	KDR
cg26090020	C17orf102	**cg17031773**	KLK9
cg03343128	C1orf150	**cg10334948**	L3MBTL4
cg01311470	C9	**cg20840054**	LIMS2
cg23550589	CCDC149	**cg18597321**	LINC01091
cg14455169	CCDC162P	**cg02419849**	LINC01107
cg07125635	CHRM5	**cg12958892**	LOC101927285
cg24838240	CLEC17A	**cg05888741**	LOC101928651
cg15965233	COBLL1	**cg19283824**	LOC101928940
cg15784618	CPM	**cg16653838**	LOC101929450
cg02763813	CRTAM	**cg23522965**	LOC101929608
cg26975860	DAB1	**cg16031039**	LOC102477328

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg07855920	MACF1	**cg02399826**	SPOCK1
cg06150437	MEG8	**cg14940332**	SQRDL
cg04251661	MIR4280	**cg09365942**	SUGCT
cg00269670	MIR548F5	**cg15260465**	SUN3
cg14241523	MKLN1	**cg12150401**	TAS2R16
cg00817995	MOV10L1	**cg02606650**	TENM2
cg04434896	NKX2-6	**cg17153727**	TFAP2B
cg07990664	NOS1	**cg09982570**	TRPM1
cg07182163	NRXN3	**cg05150480**	WISP1
cg14079660	PAMR1		
cg12157387	PARD3B		
cg12175689	PDE11A		
cg19077019	PDYN		
cg04993169	PHACTR3		
cg14875564	PLA2G4A		
cg24855857	PLK2;PLK2		
cg03358250	PPP1R1C		
cg20540357	PPP2R2B		
cg22915460	PTK2B		
cg13470291	PTPRE		
cg00533835	PTPRN2		
cg10249637	PTPRN2		
cg08403027	RASL12		
cg11823687	RELN		
cg04444865	RGS6		
cg17674543	RIMS4		
cg17078629	RTP2		
cg01660114	SCD5		
cg20921778	SLC6A13		
cg17264397	SMOC2		
cg25045972	SNORD113-6		
cg11419498	SNX29		

Table C3 Target genes of DNMT3A1

CpG	UCSC_RefGene_Name
cg27039534	C10orf90
cg13395712	C17orf97
cg07613458	EEPD1
cg04726310	FRMD5
cg09203312	GJB6
cg25085070	KIF18A
cg06041732	LINC00907
cg15424570	LINC00907
cg13693143	MIR206
cg27559408	MIR330
cg01307483	NRF1
cg12962308	OR5D16
cg17514068	SYNE2
cg15565110	TMEM67
cg07225782	TRPM8
cg02867216	TSHZ1
cg25743243	ZNF679

Table C4 Target genes of DNMT3A2

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg14426428	AIM1	**cg20593471**	BCL2
cg17095685	ALADL2	**cg15729649**	BICC1
cg16074390	ALDH1A2	**cg18849256**	BICD1
cg17778221	ALG9	**cg01848236**	BIRC6
cg13526659	AMPH	**cg12328330**	BLOC1S4
cg19159985	ANKHD1	**cg02238928**	BMP7
cg09033641	ANKRD11	**cg03291755**	BNC2
cg27079614	ANKRD13A	**cg10649841**	BRD8
cg11672099	ANKRD35	**cg25223634**	C10orf26
cg13355704	ANXA4	**cg04532551**	C19orf47
cg23267831	AOAH	**cg04932658**	C19orf81
cg17727159	ARHGAP31	**cg21268653**	C1orf203
cg12150041	ARHGAP32	**cg18729148**	C21orf62-AS1
cg13376104	ARHGDIB	**cg17903782**	C4orf22
cg06967020	ARHGEF10L	**cg09681278**	C4orf37
cg20092396	ARHGEF37	**cg07204005**	C6orf106
cg04606397	ARID1B	**cg09503608**	C6orf26
cg16677621	ARID1B	**cg19965099**	C6orf48
cg01087456	ARIH1	**cg24612707**	C7orf68
cg20452714	ARTN	**cg03029255**	C8orf31
cg25570306	ASPH	**cg06758746**	C9orf3
cg05342446	ATAD1	**cg21256342**	CACUL1
cg18165237	ATE1	**cg22828602**	CADPS2
cg27629145	ATF6B	**cg09367967**	CALR
cg05433448	ATP2C1	**cg18093691**	CCNI2
cg12097550	ATP5G2	**cg12824456**	CCNY
cg03034900	ATP5I	**cg17862993**	CDH18
cg10075191	ATP6V1A	**cg09781028**	CDK14

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg18090634	CELF1	**cg07587451**	DPYSL2
cg27638662	CEP170	**cg10258107**	DYNC1I1
cg10169382	CGREF1	**cg16893968**	DYNC1I2
cg00239870	CH25H	**cg25329325**	E2F6
cg27018185	CH25H	**cg21978593**	EBF1
cg03542374	CHD2	**cg23029193**	EGFLAM
cg03212674	CLK2	**cg15052505**	EIF2AK1
cg13630493	CLTA	**cg09241638**	ELMO1
cg09006659	CLU	**cg15810415**	ENPEP
cg11662760	COG3	**cg07018561**	ENTPD3-AS1
cg23211065	COPB2	**cg06129726**	EPB41L3
cg10325507	COQ5	**cg16411256**	EPHX2
cg18179075	CPA2	**cg12413918**	ERBB2
cg15095213	CRACR2A	**cg11492274**	ERICH1-AS1
cg07492937	CREM	**cg10426370**	ESM1
cg15029935	CRTC3	**cg08079330**	ESRRB
cg27082285	CTN3	**cg16934235**	EXPH5
cg13180058	CTNNBL1	**cg01297479**	FAM108C1
cg07115822	CXADR	**cg10833066**	FAM109A
cg06817368	DCP2	**cg14140515**	FAM159B
cg10706802	DDX6	**cg00184203**	FAM171B
cg03263272	DENND3	**cg15133477**	FAM205BP
cg27037802	DEPDC5	**cg26062455**	FAM71F1
cg14240608	DESI2	**cg15357165**	FANCC
cg14735313	DH14	**cg15828427**	FBXW9
cg07134608	DHDH	**cg00723017**	FER1L6
cg13880379	DLC1	**cg16794471**	FMNL2
cg21678614	DLGAP1	**cg18174678**	FOXJ3
cg22337620	DLGAP1	**cg10036013**	FOXK1
cg25072592	DLST		

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg12152765	FOXK2	**cg20696451**	HSPA14
cg10551484	GATB	**cg10400423**	HSPA5
cg17430362	GBA2	**cg07310984**	IFIT3
cg13808803	GHR	**cg01396302**	IGF2BP2
cg06245658	GLI2	**cg16181505**	IGF2BP2
cg26875524	GLIS1	**cg19230532**	IGF2BP2
cg17753877	GLIS3	**cg09202145**	IGFBP7
cg07038690	GLTSCR1	**cg09729065**	IKZF3
cg11267359	GMDS	**cg04100971**	IL18
cg02948693	GNB1	**cg08978665**	IL32
cg08038054	GNG11	**cg01806261**	IMMP2L
cg09726355	GPATCH8	**cg23806656**	IMMP2L
cg01987047	GPCPD1	**cg06779086**	INPP5F
cg16204524	GPR113	**cg25238640**	IP
cg04685975	GRB2;GRB2	**cg02835421**	IPO13
cg10020309	GTF2E1	**cg02802590**	ITPRIPL1
cg02749560	GTPBP2	**cg03474430**	KANK1
cg19475988	GULP1	**cg04852685**	KCNMB2
cg24238527	H2AFV	**cg13329032**	KDM2A
cg22277567	HDAC4	**cg26671851**	KDM6A
cg14973347	HDAC9	**cg04458670**	KIAA0947
cg09624120	HEATR6	**cg01885559**	KIDINS220
cg11677533	HHLA3	**cg15819739**	KIFC1
cg21461470	HIST2H2AA4	**cg06326971**	KIFC3
cg24245397	HK1	**cg19823803**	KIT
cg18348836	HOOK1	**cg16167943**	KLB
cg07797397	HOXD1	**cg01260308**	KLHDC2
cg13044985	HOXD1	**cg05241571**	KRTDAP
cg22501143	HP1BP3	**cg20498763**	LAMA2

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg17190362	LAMA4	**cg11346282**	MAP4K3
cg01895696	LATS2	**cg09331409**	MAP6
cg18708502	LATS2	**cg01837497**	MAPK14
cg05527869	LHX1	**cg14455887**	MAST4
cg23966569	LIMS2	**cg12688268**	MBP
cg23345395	LINC00343	**cg25737410**	MCF2L2
cg03647393	LINC01069	**cg21444749**	MDN1
cg25043996	LINC01307	**cg07593915**	MED13L
cg26734132	LINC01330	**cg08581446**	MED13L
cg21817764	LINC01489	**cg21265996**	MED21
cg13572309	LMO3	**cg09221431**	METTL24
cg20017295	LOC100506844	**cg13302154**	MGP
cg21231068	LOC101926975	**cg15930380**	MIER1
cg00524136	LOC101927780	**cg01069675**	MIP
cg02254041	LOC101928103	**cg04933317**	MIR1273H
cg03288924	LOC101929123	**cg25970929**	MIR155HG
cg14132585	LOC101929710	**cg03794238**	MITD1
cg13696490	LOC201651	**cg07113570**	MITF
cg26759486	LOC645434	**cg02296932**	MMD2
cg10445599	LOH12CR1	**cg09637723**	MON1A
cg07263235	LPL	**cg19753469**	MORF4L1
cg22985146	LRRFIP2	**cg06470279**	MRC2
cg06171633	LTK	**cg15709536**	MRPS31
cg09879137	LUC7L3	**cg12935350**	MRPS9
cg08443014	LYPD6	**cg14679803**	MTFMT
cg06473548	MACF1	**cg13628726**	MTMR3
cg09866472	MACF1	**cg23423181**	MTRNR2L4
cg10946938	MACROD1	**cg18121434**	MYO18A
cg26187339	MAD2L1BP	**cg08663213**	MYO18A
cg12396641	MAML3		

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg24581547	MYO5B	**cg25064395**	PDGFRA
cg25551043	MYOF	**cg10206741**	PHC1
cg14125537	NBPF8	**cg01428071**	PHIP
cg00813334	NCOA6	**cg02304767**	PHLPP1
cg20537764	NEK11	**cg05018608**	PIP5K1B
cg06703222	NFAT5	**cg17047829**	PLCE1
cg03449946	NFIA	**cg16222616**	POC1B
cg15484811	NGRN	**cg22956231**	PPP2R2C
cg08546514	NINL	**cg07182343**	PPP6R3
cg26739807	NME2	**cg14217558**	PRDM16
cg26954135	NOCT	**cg06135990**	PRR5-ARHGAP8
cg03446508	NOVA1	**cg27441872**	PSME1
cg01432046	NR5A2	**cg10589408**	PTP4A2
cg17530754	NRIP1	**cg01369817**	PTPN23
cg01654560	NTM	**cg18122501**	PTPRN2
cg12682110	NTM	**cg17698886**	PTPRS
cg17619803	NUMBL	**cg00632811**	PWP1
cg12582028	OAZ3	**cg12130067**	RAB11FIP1
cg03727842	OPHN1	**cg15415783**	RAB40C
cg00516867	OPN1SW	**cg26734848**	RALB
cg13212525	OTOL1	**cg09202322**	RBM18
cg17466151	PACS1	**cg14615491**	RBM33
cg24934408	PADI1	**cg02698507**	RFX2
cg16684992	PAFAH1B1	**cg03191504**	RNF149
cg02187196	PANK2	**cg17796813**	RNF168
cg06903211	PBX1	**cg15672877**	RNF43
cg16155772	PCCA	**cg19083634**	RNFT2
cg15770136	PCDH9	**cg05572461**	RPA3
cg07537734	PDGFC	**cg09401696**	RPS26

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg05651511	RPTOR	**cg16911115**	SRP72
cg14640066	RTN4	**cg05465104**	ST7L
cg17401362	RWDD4	**cg04828068**	STAM2
cg24089111	SARM1	**cg27328824**	STAMBPL1
cg24196318	SCAF4	**cg20716209**	STAT3
cg05463545	SCAPER	**cg05692128**	STIM2
cg04274830	SCARB2	**cg06075946**	STPG1
cg02390614	SCMH1	**cg14994183**	SYN2
cg13001838	SEC61G	**cg21958743**	SYT1
cg03855973	SENP6	**cg13449839**	TAF4B
cg14423617	SERPI5	**cg00274587**	TAPBP
cg00049545	SERPINB13	**cg00555420**	TAPT1
cg14677615	SEZ6	**cg06218079**	TBCD
cg07126559	SGCG	**cg05555876**	TCEA1
cg04297507	SGMS2	**cg07321291**	TCEB1
cg13667782	SH3GLB2	**cg19261940**	THRA1/BTR
cg00961640	SLC16A4	**cg18953822**	TIAM1
cg05487207	SLC39A14	**cg24262469**	TIPARP
cg20392013	SLC39A9	**cg11454957**	TMEFF2
cg00029150	SLC51A	**cg25565793**	TMEM260
cg04799948	SMAD5-AS1	**cg19490180**	TMPRSS6
cg05730460	SMARCA2	**cg26745222**	TOX2
cg19927678	SMG6	**cg13059495**	TPRG1L
cg20188739	SNORD87	**cg12825509**	TRA2B
cg13074203	SNTB1	**cg02854396**	TRHDE
cg05614305	SNX10	**cg20792582**	TRIM24
cg03594550	SOX11	**cg13628022**	TTLL11
cg16920608	SPATA16	**cg03487391**	TTLL4
cg23079866	SPDYC	**cg24829718**	TTLL5
cg08732275	SPSB1		

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg20249071	UBASH3B	**cg18978297**	ZBTB10
cg02020772	UBE2D2	**cg16958493**	ZCCHC7
cg22455082	UBE2K	**cg17692028**	ZDHHC17
cg22212260	UBP1	**cg25766774**	ZDHHC3
cg04839974	UGCG	**cg24857943**	ZFAND3
cg01062424	UHRF1BP1	**cg23661268**	ZFHX3
cg24310913	USP3	**cg16586418**	ZFHX4
cg26561986	USP44	**cg22043381**	ZHX2
cg27257939	VBP1	**cg08769844**	ZNF233
cg09954698	VSX2	**cg06704974**	ZNF438
cg03268306	WDFY4	**cg19866478**	ZNF480
cg01096191	WDR25	**cg04085696**	ZNF559
cg03104358	WDTC1	**cg17712400**	ZNF652
cg21360126	WIPF2	**cg02205936**	ZYG11B
cg19886272	WWP2		
cg23370051	XPR1		

Table C5 Target genes of DNMT3B1

CpG	UCSC_RefGene_Name
cg23967739	A2ML1
cg18502829	ACCN1
cg00584743	APBB2
cg03276883	BDH2
cg23491841	C2orf85
cg13270972	CNPY3
cg07212793	DOCK4
cg06801857	GALNT2
cg19710449	JAKMIP1
cg11235869	KIF26B
cg09385371	LOC153910
cg08485086	MAPK1IP1L
cg26515755	MYH11
cg15368193	MYO5B
cg21610390	MYO5B
cg17274916	POLDIP2
cg08553820	PRKD2
cg06339171	PTPRG
cg01496696	PTPRN2
cg25338587	SYNE1
cg10960152	ZNF713

Table C6 Target genes of DNMT3B2

CpG	UCSC_RefGene_Name
cg23984908	ARMC8
cg06419218	C11orf44
cg01986486	C3orf67
cg06081039	CCDC64
cg20924286	CLDN11
cg21824190	DERA
cg19918866	DIRC3
cg23465427	DLG4
cg20234991	DOC2B
cg24762962	FLJ37543
cg11457308	GALNT16
cg10131699	GALNTL2
cg26310920	GCK
cg17983064	IL3
cg17407969	KNDC1
cg04372796	LOC101929517
cg14351952	PDCD1LG2
cg27007060	RTN4IP1
cg12106308	RUFY4
cg15686094	SNTG2
cg13274534	TEX29
cg10062150	TMEM59
cg23831942	TSHZ2

Table C7 Target genes of DNMT3B3

CpG	UCSC_RefGene_Name
cg12173487	CPNE2
cg12583095	HTR7
cg06580033	ZNF830

Table C8 Target genes of DNMT3B4

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg26975768	ABLIM1	**cg18537730**	IZUMO1
cg03745383	ACCN1	**cg03890215**	LINC00907
cg00354572	ACSS3	**cg00935967**	LOC100133612
cg04977222	ATP13A5	**cg18976321**	LOC102546229
cg27163126	AZIN1-AS1	**cg10081994**	LONRF2
cg03380861	B3GLCT	**cg08643824**	LPXN
cg10687420	C7orf58	**cg26179289**	LY86
cg27167221	CAMTA1	**cg19995387**	MC5R
cg26636590	COX7B2	**cg10350998**	MLLT3
cg07233933	CPEB4	**cg02765998**	MOB3B
cg14776738	DENND1C	**cg06797656**	MYO18B
cg05101674	DOK5	**cg14736911**	MYOZ2
cg12269473	EHF	**cg26657649**	NEDD4L
cg07196212	ESR1	**cg05112298**	NOD2
cg09096933	FGD4	**cg07494218**	OR2C1
cg25980637	FLJ37453	**cg02399892**	OTOA
cg06830784	FRZB	**cg20836372**	OTOA
cg03345391	GCK	**cg17637556**	PACRGL
cg26966698	HECW2	**cg00354484**	PAX8
cg11451043	HLA-DPA1	**cg00586113**	PCSK5
cg00310588	HTR3C	**cg12381317**	PFKM
cg09992309	ISPD	**cg02788195**	PHYHIPL

CpG	UCSC_RefGene_Name
cg11015424	RALGAPA1
cg00754426	RALGAPB
cg03965221	RPGRIP1
cg00554229	RSPO1
cg22063989	RSPO1
cg13440637	RTN1
cg17723122	SLC45A3
cg20413514	SMG6
cg01082554	SMYD3
cg27454528	SMYD3
cg01190484	TRIM40
cg22594238	WIPF2
cg19913551	ZFR2
cg11180789	ZNF615
cg23501467	ZNF93

Table C9 Target genes of DNMT3B5

CpG	UCSC_RefGene_Name
cg15366573	ABHD2
cg17053902	ADAM12
cg02633229	C15orf23
cg11788263	CCSER2
cg15923359	CSGALCT1
cg12640568	DLGAP1
cg15625693	F11
cg20268758	FAM189A1
cg06695675	LOC101927641
cg09307121	LOC101929450
cg21225049	LOC102723828
cg02920604	MYO1E
cg19805562	SEC23A
cg19296884	TSRE1

Table C10 Target genes of DNMTΔ3B1

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg22381317	ACVR1B	**cg01400040**	RORA
cg09302836	ALPK2	**cg09001356**	RPTOR
cg15872575	ASB18	**cg10821341**	SIL1
cg06650659	ATP13A5	**cg19819818**	SLC24A4
cg06717221	ATP8B3	**cg02965511**	SLC51A
cg19378216	C20orf114	**cg24871226**	SSH1
cg19399653	CEACAM8	**cg09044458**	ST5
cg17309085	CNTN5	**cg12603229**	STK4
cg01287361	DJC15	**cg13693826**	SYNE2
cg26763380	EIF3E	**cg21918419**	TAPT1
cg16318688	EPHX4	**cg06490287**	TCTE3
cg11703750	FAM13C	**cg14665901**	TEAD1
cg23752752	FOXK1	**cg16275707**	ZNF10
cg24949632	GABRB1	**cg07790111**	ZNF470
cg19945464	GAS2		
cg11452221	GEFT		
cg26877678	KSR1		
cg12111733	LINC01443		
cg12688894	LINC01564		
cg07680037	LIPC		
cg11304573	LOC101928994		
cg20546365	LOC145845		
cg03894789	MIR874		
cg06970884	MNT		
cg21991396	NLRP3		
cg09259409	OR2W5		
cg07597882	OXT		
cg14930579	PAPD7		
cg19931348	PI3		
cg12406507	POLD3		

Table C11 Target genes of DNMTΔ3B2

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg23891985	A1CF	**cg03191962**	C9orf98
cg15769388	A2ML1	**cg17108476**	CA12
cg05894322	ABCA10	**cg10352688**	CAMK4
cg00747477	ACCN1	**cg24521848**	CCDC108
cg21474880	AFAP1	**cg10106388**	CD244
cg23434264	AKAP13	**cg09781987**	CDYL
cg15930120	AKAP6	**cg10375597**	CEACAM20
cg25533247	AKAP8L	**cg23818780**	CECR2
cg18761400	ALDH8A1	**cg24229206**	CELF2
cg25104512	ALS2CL	**cg00028056**	CHPF
cg22615730	ANXA4	**cg17599620**	CLASP2
cg00341504	ARHGAP28	**cg10223809**	CLEC12A
cg26988221	ARPC2	**cg11937448**	CLEC12B
cg11304664	ATF5	**cg10301072**	CLGN
cg05365685	ATP2B2	**cg05404912**	CLIP3
cg16847719	ATP8A2	**cg13507084**	CNTFR
cg12431207	ATRNL1	**cg03122917**	COL6A5
cg03440454	B3GNTL1	**cg19985724**	CORIN
cg04628938	BCAS3	**cg11270449**	CPB2-AS1
cg09911747	BCL11B	**cg14806927**	CR1
cg14168713	BICC1	**cg05080132**	CTBP2
cg00076072	BPIFB3	**cg14665921**	DCDC2C
cg21488132	C11orf31	**cg27228578**	DH9
cg16216885	C19orf81	**cg03831312**	DIRC3
cg22213391	C1orf106	**cg05511924**	DIRC3
cg01450736	C1orf110	**cg01697794**	DLG4
cg24731625	C1orf229	**cg12017315**	DLST
cg11123493	C5orf13	**cg01708284**	DMRT1

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg22811478	DOCK4	**cg01858205**	GPR108
cg05465131	DRD3	**cg25244238**	GPR108
cg05292310	DSPP	**cg09648243**	GPR39
cg01248810	DYDC1	**cg05458052**	GRHL2
cg17372920	EDDM3A	**cg21007931**	GSG1L
cg10617796	EDIL3	**cg04836362**	GTF2IRD1
cg17877237	EEF1DP3	**cg05309750**	GXYLT2
cg03746008	EIF2B5-AS1	**cg06118312**	GZMA
cg04054100	ENTHD1	**cg22438640**	HCG4
cg17572324	ESYT2	**cg11246805**	HDLBP
cg26404422	ETS1	**cg16638857**	HKDC1
cg05643360	FAM110B	**cg26784412**	HLA-DPB2
cg00205703	FAM131A	**cg05500783**	HLA-DRA
cg16782425	FAM171B	**cg02563789**	HNRNPC
cg02536065	FLI1	**cg02264990**	HOXC4
cg24022651	FNDC3B	**cg00377794**	HUNK
cg19418525	FRAS1	**cg04412506**	ICA1
cg16531955	FRS2	**cg11609995**	IL18R
cg16534289	FSTL1	**cg16319517**	IL1RAP
cg00056280	G12	**cg02349850**	IPO11
cg04769577	G12	**cg16840156**	ITGA11
cg21684809	G12	**cg09934565**	ITPR1
cg00019118	GALNT9	**cg24323597**	KALRN
cg10356397	GBE1	**cg05658173**	KCNK15
cg00416130	GJA4	**cg08856529**	KCNK9
cg19720377	GLP2R	**cg14601981**	KCNMA1
cg13559820	GLT6D1	**cg19584551**	KIAA1217
cg18138010	GMDS-AS1	**cg27587063**	KIAA1751
cg23541617	GNPTAB	**cg14950072**	LAMA1

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg03368758	LDB2	**cg21435394**	NUDT14
cg03730738	LINC00351	**cg18949521**	OR5V1
cg04502126	LINC01258	**cg17473495**	OSBPL3
cg17981966	LMO3	**cg02078558**	OSBPL6
cg00684347	LOC100271832	**cg21191347**	PA2G4
cg03177464	LOC100288798	**cg02644583**	PDCD1LG2
cg14236242	LOC101448202	**cg10989138**	PDE1C
cg11035685	LOC101927769	**cg02202052**	PDE3A
cg23071261	LOC101928162	**cg26062141**	PDE4B
cg09552181	LOC101929596	**cg20133730**	PDK2
cg07152607	LOC105616981	**cg27355501**	PFDN6
cg02088996	LOC285954	**cg25845380**	PID1
cg23275355	LOC392232	**cg26723002**	PLAG1
cg01680573	LOC729991-MEF2B	**cg15034764**	POSTN
cg15790804	LRTM2	**cg25464078**	PPTC7
cg03354707	LTBP4	**cg26493631**	PRDM11
cg04447756	MATR3	**cg09420520**	PRMT8
cg26231094	MBOAT7	**cg14044685**	PRRG4
cg21331791	MCF2L	**cg09267188**	PTDSS1
cg13272357	MIR20B	**cg03398844**	PTK2
cg06671868	MIR548G	**cg20006825**	PTMS
cg26374481	MIR548G	**cg07775266**	PTPRM
cg22877504	MREG	**cg14129439**	PTPRZ1
cg15344442	MYT1L	**cg26654286**	RARA
cg07918168	NCAM2	**cg08384130**	RBFOX3
cg18402101	NDUFC2	**cg09959490**	RBM47
cg02870676	NEDD4L	**cg00430372**	RCAN3
cg01281210	NRG1	**cg19585586**	RPL17
		cg21287940	SCG5

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg06621691	SDK1	**cg10980788**	TMC1
cg27378216	SETBP1	**cg14031220**	TMEM178A
cg05777337	SGK223	**cg13740979**	TNFRSF10B
cg04421348	SH2D4B	**cg21808287**	TNP1
cg15710554	SIL1	**cg20243424**	TNXB
cg04095373	SKAP1	**cg04639174**	TRAF6
cg19734536	SLAMF1	**cg23624321**	TRAF6
cg23642651	SLC13A3	**cg01910804**	TTC6
cg06421013	SLC24A3	**cg11963464**	TTPA
cg04292976	SLC26A9	**cg26161386**	TTPA
cg16923485	SLCO1A2	**cg12142328**	USP34
cg26990023	SMOC2	**cg07337250**	USP43
cg02077920	SMYD3	**cg15405572**	V3
cg16046833	SNX7	**cg13084429**	VWA2
cg18544696	SORL1	**cg17276624**	WBSCR17
cg17284609	SOX6	**cg04372929**	WISP2
cg06728232	SPATA16	**cg17167536**	XKR6
cg02405517	SPC2	**cg20060356**	ZFPM2-AS1
cg14742361	STAU2	**cg11473417**	ZNF251
cg10429523	STL	**cg05859088**	ZNF365
cg04759244	STX16	**cg08370080**	ZNF429
cg10206882	STXBP6	**cg00845602**	ZNF438
cg26535273	SYNGAP1	**cg15604051**	ZNF502
cg08699646	TEPP	**cg19819404**	ZNF718
cg26323602	TEX26	**cg24071762**	ZNF839
cg11265171	TFDP2		
cg12397274	TIG		
cg12625508	TKT		
cg08757862	TLR1		

Table C12 Target genes of DNMTΔ3B3

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg07152070	ADGRG4	**cg18226835**	FSHR
cg19561181	ANKRD60	**cg27154418**	GP6
cg08490364	AP2B1	**cg25762753**	GPR156
cg00950473	APCDD1L	**cg20934765**	GRID2
cg06376392	ARHGAP15	**cg12518734**	HAP1
cg18484278	ARHGAP22	**cg03965949**	HECTD2-AS1
cg26499611	BAIAP3	**cg10759591**	HRNBP3
cg08927738	BCAS1	**cg08067312**	IGFBPL1
cg17512348	C9orf171	**cg16217885**	IL1R1
cg09756599	CALN1	**cg24530225**	IL1R1
cg23388535	CDCP1	**cg26810336**	IL1RAPL1
cg22379697	CHD6	**cg01849284**	KCNMB2
cg04075184	CLASP1	**cg12033029**	KIF5C
cg15501456	CNDP1	**cg10637292**	KRT74
cg04066265	CSGALCT1	**cg14807428**	LARP4B
cg11822265	CTN3	**cg26020743**	LINC00299
cg07089660	DEFA4	**cg05191222**	LINC00703
cg19839614	DGKK	**cg15262365**	LINC00929
cg09055992	DLGAP2	**cg04008237**	LIPE-AS1
cg00529424	DMRT1	**cg22676735**	LOC101929710
cg15665376	DSCR10	**cg02542440**	LOC441601
cg20364776	EEF1AL7	**cg14704531**	LOXHD1
cg27494843	EHF	**cg17838834**	MIATNB
cg14055970	ENG	**cg21876806**	MIR299
cg26381730	ERICH1-AS1	**cg19444866**	MIR6854
cg05949399	ERICH1-AS1	**cg10609662**	MMP24
cg27343917	FAM176A	**cg17964305**	MSRA
cg19218438	FAM19A5	**cg21467050**	MX1

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg08004620	MYOM2	**cg20838631**	TRRAP
cg04717370	MYT1L	**cg12968088**	WDFY4
cg08635101	NBPF25P	**cg01317060**	XYLT1
cg14608275	NCEH1	**cg14289985**	ZNF471
cg12123879	NCR1		
cg11872504	NDN		
cg03466998	NLRP3		
cg18183941	NLRP3		
cg21360918	OR7G1		
cg02462812	PCDHB4		
cg15956339	PHOX2B		
cg17255063	PHOX2B		
cg16323609	PRKAR2A		
cg16490431	REPS2		
cg04751120	RP1L1		
cg09193037	RPEL1		
cg06329972	RYR3		
cg11257864	SERPI5		
cg20254658	SHROOM3		
cg24247040	SKAP1		
cg25221452	SLAMF7		
cg12187305	SLC8A3		
cg03288304	SLC9A4		
cg00975791	SNORD114-18		
cg19200333	SORCS1		
cg18658556	SORCS2		
cg16902385	ST8SIA3		
cg25256067	THSD7A		
cg06196379	TREM1		
cg20936707	TRIM58		

Table C13 Target genes of DNMTΔ3B4

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg17713190	ABLIM3	**cg13770691**	C19orf6
cg23976464	ACCN1	**cg00588517**	C2
cg20240860	ACCS	**cg01383890**	C2CD4A
cg03202077	ADAMTS5	**cg21490417**	CAC1A
cg16559275	ADAMTSL4	**cg25231948**	CAMK4
cg13940693	ADCYAP1	**cg01544580**	CBX1
cg22842879	AFAP1	**cg25051805**	CCDC126
cg14134368	AGFG1	**cg25872744**	CCDC83
cg14667832	ALK	**cg10042846**	CCL13
cg24603490	AMIGO2	**cg13720316**	CCND1
cg16859696	AMMECR1	**cg24729983**	CCNO
cg02548132	ANGPT2	**cg10185424**	CD180
cg15855924	ANKRD10	**cg11874272**	CD86
cg11713658	ARHGAP10	**cg03829839**	CDK4
cg10984236	ARID1B	**cg18094824**	CENPE
cg11090211	ARL5C	**cg20539142**	CHR6
cg14451560	ARRDC1	**cg09449447**	CLDN14
cg07504154	ASB5	**cg25272432**	CLGN
cg21068480	ATOH8	**cg15862165**	CRADD
cg27450074	ATP9B	**cg06794775**	CREB3L4
cg26402735	BASP1	**cg16596604**	CRYM
cg21500300	BCAT1	**cg07752723**	CT45A10
cg04678713	BDH1	**cg21921460**	CTNND2
cg14304249	BEST2	**cg06643622**	CTTNBP2
cg16965449	BRD7	**cg20502337**	CXorf36
cg24671734	BTBD11	**cg03142049**	CYSTM1
cg03359067	BTNL3	**cg02872476**	DBNDD1
cg04358214	C16orf70	**cg17110349**	DIDO1

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg27369423	DKFZP434H168	**cg16415340**	INS-IGF2
cg12954425	DOCK8	**cg21912938**	JAZF1
cg07258372	DT	**cg09244349**	KANSL1
cg05293775	DUSP26	**cg18639125**	KCNF1
cg11806102	EHHADH	**cg04431946**	KCNK12
cg26246486	ELAVL2	**cg07685869**	KIFC3
cg02156680	ENPP2	**cg15199350**	KIRREL3
cg05075705	ENPP3	**cg21807198**	KSR2
cg07482337	EOMES	**cg19909349**	LAMA1
cg24460563	EPAS1	**cg04850395**	LINC01470
cg00157656	ERICH1	**cg09413116**	LOC101928233
cg06752040	FAM107B	**cg23281602**	LOC101928489
cg15904427	FAM122C	**cg17264609**	LOC150568
cg17114257	FCHSD2	**cg03065467**	LOC641518
cg22454769	FHL2	**cg03946923**	LOC641518
cg24561572	FMNL1	**cg04162032**	LOC646627
cg16509658	GABRB2	**cg10222309**	LRRC3C
cg13881405	GADD45G	**cg10316899**	MACF1
cg25435255	GBF1	**cg16896847**	MAFA
cg26295057	GDNF	**cg06852975**	MAGEC1
cg04714954	GLO1	**cg18858249**	MAGI2
cg22702243	GNE	**cg26328687**	MAML2
cg06460174	GON4L	**cg07618155**	MIR124-2HG
cg21351852	GPR84	**cg24603444**	MIR125B1
cg20947470	GT1	**cg10082525**	MIR380
cg12151328	GTPBP8	**cg20751048**	MIR548W
cg05380503	HECW1	**cg02047319**	MIR654
cg13308279	HECW2	**cg24456365**	MMP16
cg26746878	IKZF1	**cg13644221**	MMP20

CpG	UCSC_RefGene_Name	CpG	UCSC_RefGene_Name
cg01157261	MOSPD3	**cg13130398**	RABGAP1L
cg25584313	MROH7	**cg06705767**	RARB
cg14546394	MSC	**cg06237092**	RBM20
cg16751754	MX1	**cg04448201**	RBM28
cg05512413	MYO1D	**cg03273615**	RBM41
cg27523882	NDUFA4L2	**cg08972081**	RBP1
cg18844118	NFE2L3	**cg09566735**	RGS6
cg14260162	OR2S2	**cg12765549**	RGSL1
cg22075486	OR6B1	**cg10827810**	RNF125
cg21515384	OXCT1	**cg23548670**	RNLS
cg20701646	PAX6	**cg25892001**	RTN1
cg13229291	PCGEM1	**cg18426142**	SCN8A
cg11416338	PDE3A	**cg24867550**	SDC2
cg22488256	PDE4B	**cg11772956**	SERPINB8
cg09092093	PDE4D	**cg26840598**	SH2D5
cg21363050	PDGFRA	**cg04489366**	SMPX
cg21799736	PDHA1	**cg09083139**	SMURF2
cg13846114	PEBP4	**cg12167284**	SNORA11B
cg07980164	PHLDB3	**cg15989068**	SOX11
cg16797824	PIK3R1	**cg27463758**	SP110
cg26827893	PLK1S1	**cg02213437**	ST18
cg27022570	POLE2	**cg09147776**	SV2B
cg01195861	POP5	**cg13919174**	SVEP1
cg06649437	PPP4R1L	**cg02015876**	SYNM
cg20557183	PRDM11	**cg06440519**	SYT3
cg15088491	PRLR	**cg02829601**	SYTL3
cg03049691	PROCA1	**cg02833108**	TAF4B
cg00446900	PTPRE	**cg15657686**	TBC1D1
cg02535073	PTRH1	**cg00427635**	TBC1D21

CpG	UCSC_RefGene_Name
cg08448379	TFCP2
cg12430467	TJP1
cg22976313	TMEM179
cg23376554	TMSB4X
cg24098643	TRIM5
cg20467195	TULP3
cg04770550	VAV1
cg13411784	WDR85
cg17133224	WFDC9
cg07657743	WNT7A
cg20769177	WNT9B
cg02171206	WWOX
cg10772868	ZNF235
cg26101904	ZNF385B
cg14393114	ZSWIM2
cg04391391	ZSWIM4

Appendix D: Schematic location of the pyrosequencing assay locations

Yellow highlight represents a CpG site, Green highlight represents a primer, Blue highlight represents a part of EPIC probe. Underline represents a sequence primer.

1. cg01065960: DNMT1

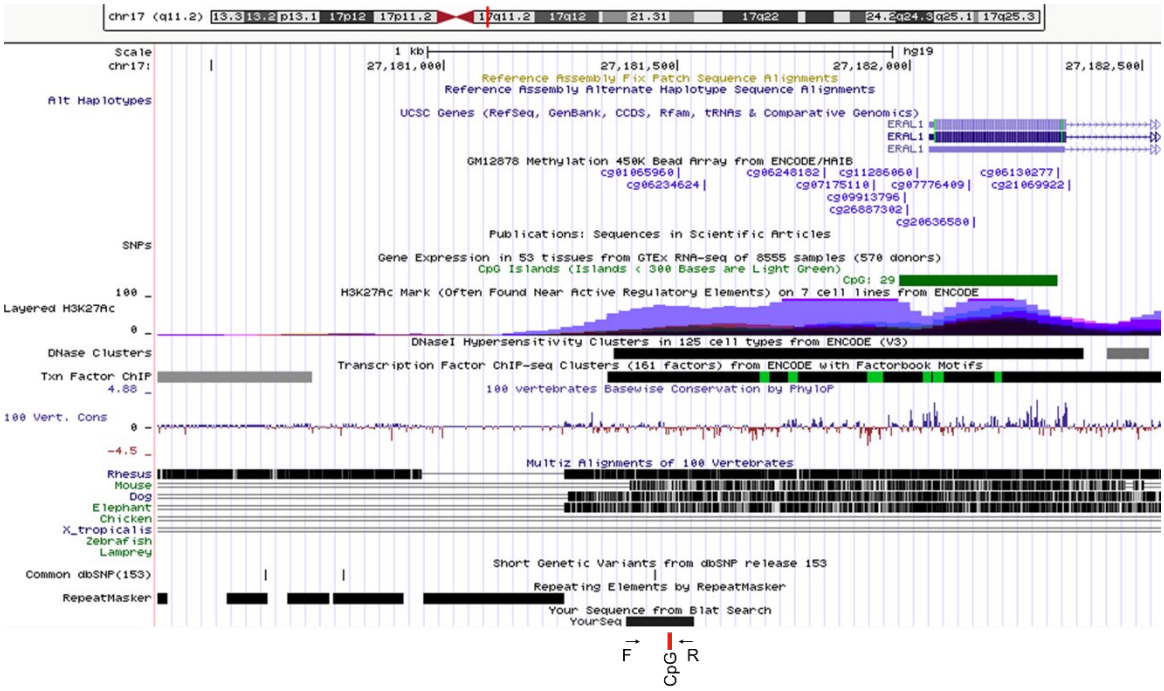
>hg19_dna range=chr17:27181267-27183516 5'pad=0 3'pad=0 strand=- repeatMasking=none

Genomic sequence:

GAAAAAGGGATCACCCGCTCCCTCGCGACATGAGGGCCCACCTGCCAGACTCTTAACACCGATTGAAC
AAGCCTAGCCCCGCGCCAGCTGGGGGCAGCCATTACAGCCGCAAGGCACTGCGGGAACCGACCAGGAG
CGCTCCTACGATTGAACAGAGCGGCTGACGAACCAGCGAGTAGGAGAGAGGTAGGTCTGCCCAAATTG
AAAGCCGAGACACATGGGATGTGGGCGGGAGGAGGAATTGCTCTGGAGACTGGTGGAGGCTAAGGAAA
GGGGAAAATGTGTATTTAGAGTCTGTGCAAGCATGACGGTAGTTTTGGACTCCCATACACAACGTATT
TTTTTCTTCTGCCAAAGTTTTGTAAACGGCCCTTCTGGCCATGTCTCCTAAAAGACTACATTTCCCAT
GAGGCCAAATAGCAGAAATCGTCCTCCCATCGAGGGAGGAATGGTTGTGGGAGTGGGAGGGGGCACTG
TTCCGCCTGTCCCCGTAAGGGTTTTGTTGTTGTTTGTTCCTAGTCCTTTCCATCCTTTGGCCTGCAGG
ATCCTCTCCCAGCCTTAATTTGGCTTGCCCTAGAGGTCAGCGGTGGCCCCGGGCTGTCTGCACCTTTG
CTTTAGGCTAGGTTCCTGGAAGGAGTGGTCCTTGGGTCGGTGTTTGGGATGAAGAGCAGGTCTGGATT
TCCCTCCTGTAAGCCTTGGCTACTGTCTACTATCTGAGGTCATTCCTTTATGGAGAACAACCAGAGGG
TTTGCAAAGGAGGAAGAAGTGGCTTAGCCACCTAAGGTCTCCTAGACCGCTCTAGCAGCCGGAGAAAC
TCATTTACCTTGCTCTGTTTGAGGCCTTGAGCCCTCTTCATCCTTTGTAGCCTCTAAATCAAACACTT
GAA

Bisulfite sequence:

GAAAAAGGGATTATTCGTTTTTTCGCGATATGAGGGTTTATTTGTTAGATTTTTAATATCGATTGAAT
AAGTTTAGTTTCGCGTTAGTTGGGGGTAGTTATTATAGTCGTAAGGTATTGCGGGAATCGATTAGGAG
CGTTTTTACGATTGAATAGAGCGGTTGACGAATTAGCGAGTAGGAGAGAGGTAGGTTTGTTTAAATTG
AAAGTCGAGATATATGGGATGTGGGCGGGAGGAGGAATTGTTTTGGAGATTGGTGGAGGTTAAGGAAA
GGGGAAAATGTGTATTTAGAGTTTGTGTAAGTATGACGGTAGTTTTGGATTTTTATATATAACGTATT
TTTTTTTTTTGTTAAAGTTTTGTAAACGGTTTTTTTGGTTATGTTTTTTAAAAGATTATATTTTTTAT
GAGGTTAAATAGTAGAAATCGTTTTTTTATCGAGGGAGGAATGGTTGTGGGAGTGGGAGGGGGTATTG
TTTCGTTTGTTTTCGTAAGGGTTTTGTTGTTGTTTGTTTTTAGTTTTTTTTATTTTTTGGTTTGTAGG
ATTTTTTTTTAGTTTTAATTTGGTTTGTTTTAGAGGTTAGCGGTGGTTTCGGGTTGTTTGTATTTTTG
TTTTAGGTTAGGTTTTTGGAAGGAGTGGTTTTTTGGGTCGGTGTTTGGGATGAAGAGTAGGTTTGGATT
TTTTTTTTGTAAGTTTTGGTTATTGTTTATTATTTGAGGTTATTTTTTTATGGAGAATAATTAGAGGG
TTTGTAAAGGAGGAAGAAGTGGTTTAGTTATTTAAGGTTTTTTAGATCGTTTTAGTAGTCGGAGAAAT
TTATTTATTTTGTTTTGTTTGAGGTTTTGAGTTTTTTTTATTTTTTGTAGTTTTTAAATTAAATATTT
GAA



cg01065960-BS-F	AGGTTAGGTTTTTGGAAGGAG
cg01065960-BS-R-bio	CCTCCTTTACAAACCCTCTAA
cg01065960-BS-SP	AGGTTAGGTTTTTGGAAGGAG
Sequencing entry	*TGGTTTTTGGGTC/TGGTG*

2. cg20943641: DNMT1

>hg19_dna range=chr12:102455342-102456841 5'pad=0 3'pad=0 strand=- repeatMasking=none

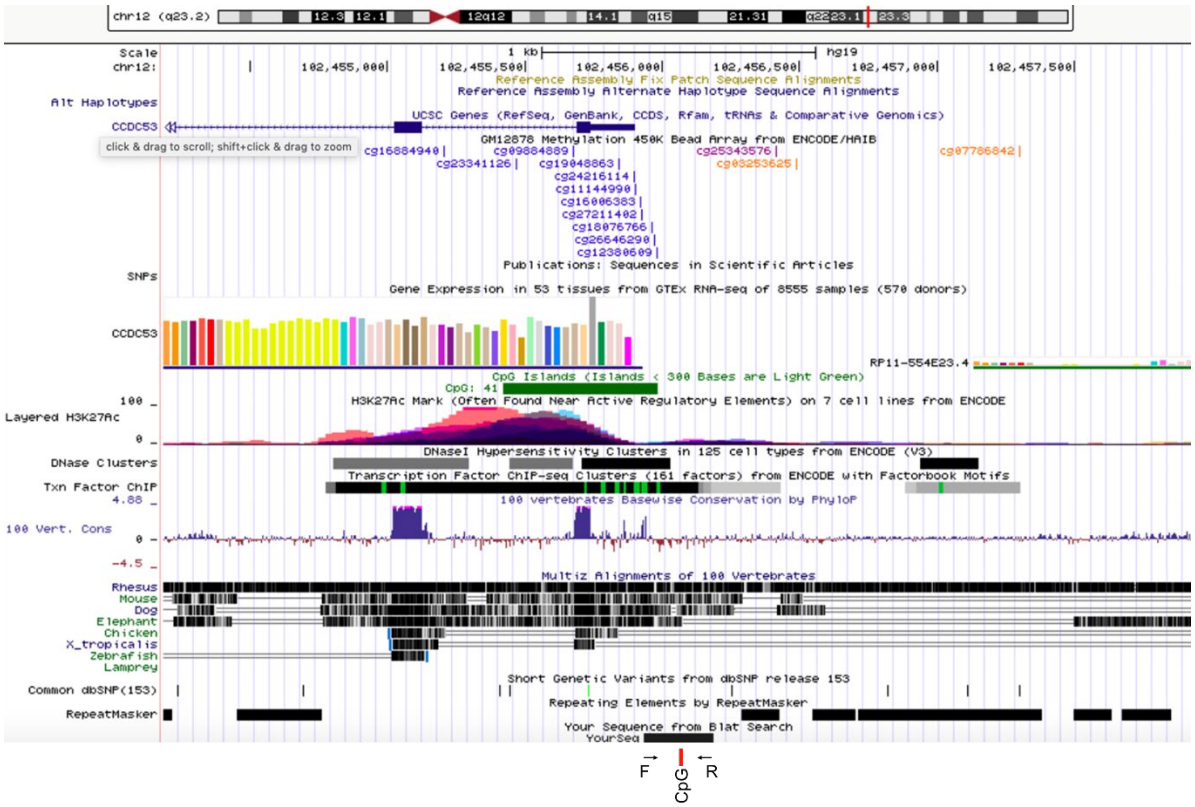
Genomic sequence:

TCATTCTCTCCTTGATTTGTGCGTTTTGTCCAATTCTTTGTTCAAAAGAACCTGGACACCCTCCACCA
GTAACACTCTTACCTCATTTCCTACCATTCTCCATTTATGGGATTATTTGCTTCAACCAAACCAACAG
GCCAGGTGCCAGTGGCTCACACCTATAATCCCAGCACTTAGGGAGGCCAAGGCATGAGGATCATTTGA
GCCAAAGGGTTTGAGACTAAGGTGAGCTATGATCCTCACTGCACGCCAGACTGGGTAACAAAGCGAGA
CCCCCCATCTTTAAAACAAAACTGCCCACTGCAGTCTTGGTCTTCATCTTTTTGATGCTCCCTGTAAG
TGGCAGTCACTTCCGTAGGTTTGTGCCTCACATCTAGTAGTTGACATTTTATTTCATGTTTTAAAAAC
CTTTGGAAGACCAGGCGCGTTGGCTCACGCCTGTAATCCCAGCACTCTGTGAGGCTGAGGCAGGAGAA
TCGTGTGAGGCTGAGGCAGGAGAATCGCTTGAGGCCAGGAGTTCGAGACCAGCCTGGGGAAATAGCGA
GAACTTGTGTGTATACACACACACACGCTTTGGAAGTAGGTCTTATTTCTCATTCTTCAGGGCATAAA
CCTCTCCCGTGCAGCAATCAAACTCAATCTTTGTTTTCCCAGTAAAAGCCTAGCAGAGTATATGGCAC
TTAGGATTAACAACAACAAAAAAAAAATTGTTGCTCAAGGAATGACGGAGTTATTATATAACAGACAG
GACTAAGTAGGGATAACCTCAAAAGGACGCAACTTCTGGCTTTGTTTTTTTGTTTTGTTTTGTTGCCT
CCTTTCTAAGATTTCAGTGCTCTAACTGCATCCTTGCACCACCGCGCGCCACCCCGAAATTGACGTCA
CAATCTGGACTAAAACTACCAATCCCAGAATGCCGAGCGAGGAGGCGGGCTCTCCGGAAGCCGCCGAG
TGATTAGTGAGCGGAGAAGCTTTCTTCCGGCGGGAAGGGCCCCGGAGGCGGGCACTTGGGGGGAAAGT
TGAGACGTGATTACCGGGTTGGGCGGGCCCCATCTGGGAGGGGTTTGTGGGTGAACTCGGGGTCCACC
GCCCGCTGAGGAGATGGATGAGGACGGGCTTCCTCTCATGGGGTCAGGCATAGACCTGACCAAGGTTT
GTAAAAGACGGTGCTAGACTCCCGACGAGCCGGGAGGAGGATGGGGAAGACGCAGCTTTCCGGGGAGA
CACCCACCTTCCGAGTCCCCACCCTGTCACGGCCTCGGGCCCTTGTGACAGGCCTGTGGGGCCTCAGG
AGGGAGAGCCCCACAGTCCACCTTGAAATGGGGTCCCTCCTGAGCTGATTAAGGAGACGGTGTGCATC
TCCCAACTGTCGCTTCCCATCTCGCTTGTCTCTCTGCCGCCGGCGACACACGCCGGTTGCGATAGCTA

GCCCTACCTCTTATTCAGCCCCCAAGATTTTGTGTGTGTTTATGTGTGGTGCTTTTTGCCTGCCAAAA
TTCG

Bisulfite sequence:

TTATTTTTTTTTTGATTTGTGCGTTTTGTTTAATTTTTTGTTTAAAAGAATTTGGATATTTTTTATTA
GTAATATTTTTATTTTATTTTTTATTATTTTTTATTTATGGGATTATTTGTTTTAATTAAATTAATAG
GTTAGGTGTTAGTGGTTTATATTTATAATTTTAGTATTTAGGGAGGTTAAGGTATGAGGATTATTTGA
GTTAAAGGGTTTGAGATTAAGGTGAGTTATGATTTTTATTGTACGTTAGATTGGGTAATAAAGCGAGA
TTTTTTATTTTTAAAATAAAATTGTTTATTGTAGTTTTGGTTTTTATTTTTTTGATGTTTTTTGTAAG
TGGTAGTTATTTTCGTAGGTTTGTGTTTTATATTTAGTAGTTGATATTTTATTTTATGTTTTAAAAAT
TTTTGGAAGATTAGGCGCGTTGGTTTACGTTTGTAATTTTAGTATTTTGTGAGGTTGAGGTAGGAGAA
TCGTGTGAGGTTGAGGTAGGAGAATCGTTTGAGGTTAGGAGTTCGAGATTAGTTTGGGGAAATAGCGA
GAATTTGTGTGTATATATATATATACGTTTTGGAAGTAGGTTTTATTTTTTATTTTTTAGGGTATAAA
TTTTTTTCGTGTAGTAATTAAATTTAATTTTTGTTTTTTTAGTAAAAGTTTAGTAGAGTATATGGTAT
TTAGGATTAATAATAATAAAAAAAAAATTGTTGTTTAAGGAATGACGGAGTTATTATATAATAGATAG
GATTAAGTAGGGATAATTTTAAAAGGACGTAATTTTTGGTTTTGTTTTTTTGTTTTGTTTTGTTGTTT
TTTTTTTAAGATTTTAGTGTTTTAATTGTATTTTTGTATTATCGCGCGTTATTTCGAAATTGACGTTA
TAATTTGGATTAAAATTATTAATTTTAGAATGTCGAGCGAGGAGGCGGGTTTTTCGGAAGTCGTCGAG
TGATTAGTGAGCGGAGAAGTTTTTTTTCGGCGGGAAGGGTTTCGGAGGCGGGTATTTGGGGGGAAAGT
TGAGACGTGATTATCGGGTTGGGCGGGTTTTATTTGGGAGGGGTTTGTGGGTGAATTCGGGGTTTATC
GTTCGTTGAGGAGATGGATGAGGACGGGTTTTTTTTTATGGGGTTAGGTATAGATTTGATTAAGGTTT
GTAAAAGACGGTGTTAGATTTTCGACGAGTCGGGAGGAGGATGGGGAAGACGTAGTTTTTCGGGGAGA
TATTTATTTTTCGAGTTTTTATTTTGTTACGGTTTCGGGTTTTTGTGATAGGTTTGTGGGGTTTTAGG
AGGGAGAGTTTTATAGTTTATTTTGAAATGGGGTTTTTTTTGAGTTGATTAAGGAGACGGTGTGTATT
TTTTAATTGTCGTTTTTTATTTCGTTTGTTTTTTTGTCGTCGGCGATATACGTCGGTTGCGATAGTTA
GTTTTATTTTTTATTTAGTTTTTAAGATTTTGTGTGTGTTTATGTGTGGTGTTTTTTGTTTGTTAAAA
TTCG



cg20943641-BS-F	AAAGTTTAGTAGAGTATATGGTATTTAGGA
cg20943641-BS-R-bio	TTAATAATTTTAATCCAAATTATAAC
cg20943641-BS-SP	TTGTTGTTTAAGGAATGA
Sequencing entry	*C/GGAGTTATTATATA*

3. cg02732111: DNMT3A2

>hg19_dna range=chr4:71728583-71730082 5'pad=0 3'pad=0 strand=+ repeatMasking=none

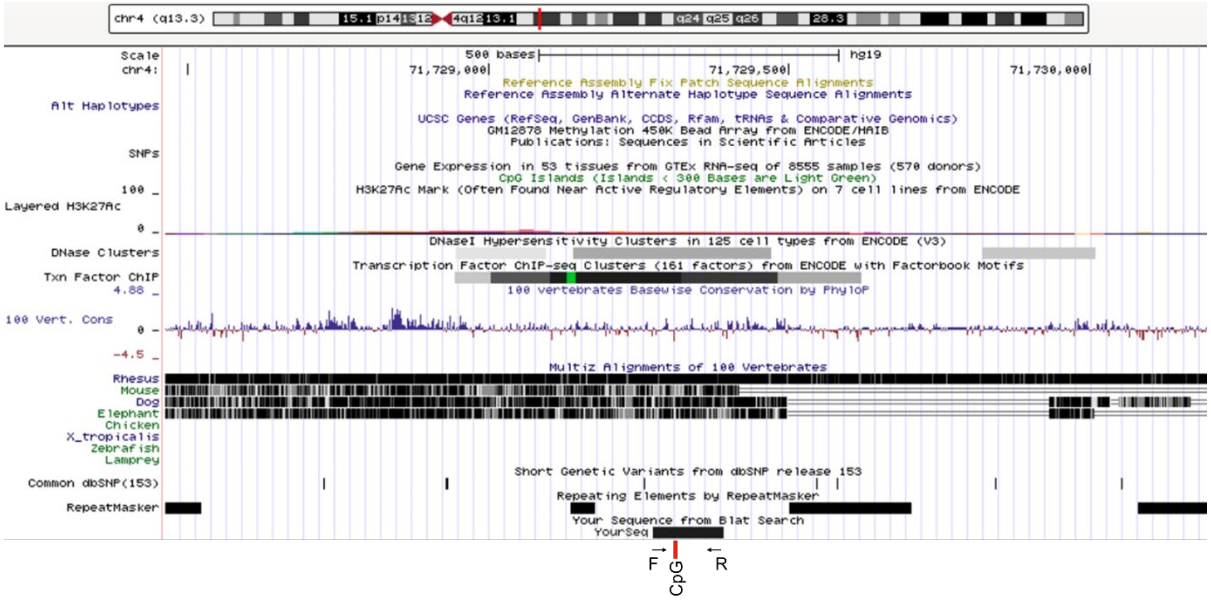
Genomic sequence:

AAAAGAAAACTAGGAAATTTTGTTTAATAGGGACTTGCAAGTTAAGATTCCAGATGCCATCCAGATTT
TCCATTGACTTCCTTCGTGCTGTCTTCCTGAAATGCATTTCTGTATAATGATGCATATGTAGGACCGA
CTTTATGGATAGGTAAAAAGAAGCTATTCCCTAAGTAATGTCATTTGTGGGACCAGGAAGCAAAGCCT
TTCCTCCCTCAACCACCCCCACCCCAAGCTGTCATAGAGTCTCAAATCCCACCATTAGCATTTCTGAA
GAAATAGCACATTAAATGAAGCCACACTAGGGTGTTAACAACATATCAGATAGATTTTGTCCTGACAT
TGGGGACCGTGGGAAATACAGAACCACTAGCAAAGGAGAAAGCACTAAAAATAGAGAAGAGGGAGGCA
ACACTAAGGCAGGGTGAACCAGAGATTCCCAGCAGTCCACAGACAGAGAGCCTTGGTTGATTACATGA
GAAGGACCTGGAATTTGTTCAAAATACTGTTTTTAGGATTCCTCCTTGTATTATTTGTCACAAGAGGT
CTGGGGTGGATTCCTGGGTGATTTCGATGTGCAGCTAAGTTCAGAAACTGTTAACTATGAAGCCACAT
AAAAAGAAACAAACATCAACAACAAAAAAACCCCACATGACTTCCCCACCTCACGTAAACACAGAGTC
AATGTTCCCATCCTGCCTAGGCCCATCACAGTTTGTTCATATGAACGGCTCTTGCCCTAAAGTCTCTA
AGGGAAAGGTTATGGGAAAAGCCAGTGACACATTAGGTACCTGTTGAAATGTGTCACTGACGTTGTGA
AAGTTTTTAAGTCTACTTAAACAACAGTGAGTTTTCAATGAACAGAACTATGATAATGAGTACAGAAC
AACACATAACAGTTTGCCTTTAAAACTTGTCGGCAATGGTGAAACCCCATCTCTACTAAAAATACAAA
AAGTTTAGCCGGGCATGATGGCGCATGCCTGTAATCCCAGCCACTCAGTAGCCTGAGGCAGGAGAATC
GCTTGAACCCAGGCGGCGGAGGTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCGACA
GAGTGAGACTCTGTCTCAAAACAAACAAAAAAACTTCTTGGCAATTGGAATGTCTATGTCTCTTAGGA
TAAGACTTTACTCCTGGCATACCACTCAGAGCAGACCTCTCAAAGTTTATAAAAGAAATTAAGGAGCC
TTGTTAGATGGTAACTATTTACACAAAATTATTTTACCTACTATTATAATTGGTTAATTGATCTTTCA
AAGTTCCACAGTGACTGATGGTAAGTCACTCTGACGCAAATGCCCAGTGCCCTAAATGTTTAAATTTA
TATAAACATTAGGTCATCAGGCCTCGTGATAGCTTTGTGAATTAAGTGGCAAGTTTTAGTTCACTTTA
AACTATAGAAAATAGACAAGGGAGCAAAAGTGCATTAAGGTATAATGCAGTGATTTAAAAAAATATTC
TTGG

Bisulfite sequence:

AAAAGAAAATTAGGAAATTTTGTTTAATAGGGATTTGTAAGTTAAGATTTTAGATGTTATTTAGATTT
TTTATTGATTTTTTTCGTGTTGTTTTTTTGAAATGTATTTTTGTATAATGATGTATATGTAGGATCGA
TTTTATGGATAGGTAAAAAGAAGTTATTTTTTAAGTAATGTTATTTGTGGGATTAGGAAGTAAAGTTT
TTTTTTTTTTAATTATTTTTATTTTAAGTTGTTATAGAGTTTTAAATTTTATTATTAGTATTTTTGAA
GAAATAGTATATTAAATGAAGTTATATTAGGGTGTTAATAATATATTAGATAGATTTTGTTTTGATAT
TGGGGATCGTGGGAAATATAGAATTATTAGTAAAGGAGAAAGTATTAAAAATAGAGAAGAGGGAGGTA
ATATTAAGGTAGGGTGAATTAGAGATTTTTAGTAGTTTATAGATAGAGAGTTTTGGTTGATTATATGA
GAAGGATTTGGAATTTGTTTAAAATATTGTTTTTAGGATTTTTTTTTGTATTATTTGTTATAAGAGGT
TTGGGGTGGATTTTTGGGTGATTTCGATGTGTAGTTAAGTTTAGAAATTGTTAATTATGAAGTTATAT
AAAAAGAAATAAATATTAATAATAAAAAAATTTTATATGATTTTTTTATTTTACGTAAATATAGAGTT

AATGTTTTTATTTTGTTTAGGTTTATTATAGTTTGTTTATATGAACGGTTTTTGTTTTAAAGTTTTTA
AGGGAAAGGTTATGGGAAAAGTTAGTGATATATTAGGTATTTGTTGAAATGTGTTATTGACGTTGTGA
AAGTTTTTAAGTTTATTTAAATAATAGTGAGTTTTTAATGAATAGAATTATGATAATGAGTATAGAAT
AATATATAATAGTTTGTTTTTAAAATTTGTCGGTAATGGTGAAATTTTATTTTTATTAAAAATATAAA
AAGTTTAGTCGGGTATGATGGCGTATGTTTGTAATTTTAGTTATTTAGTAGTTTGAGGTAGGAGAATC
GTTTGAATTTAGGCGGCGGAGGTTGTAGTGAGTCGAGATCGCGTTATTGTATTTTAGTTTGGGCGATA
GAGTGAGATTTTGTTTTAAAATAAATAAAAAAATTTTTTGGTAATTGGAATGTTTATGTTTTTTAGGA
TAAGATTTTATTTTTGGTATATTATTTAGAGTAGATTTTTTAAAGTTTATAAAAGAAATTAAGGAGTT
TTGTTAGATGGTAATTATTTATATAAAATTATTTTATTTATTATTATAATTGGTTAATTGATTTTTTA
AAGTTTTATAGTGATTGATGGTAAGTTATTTTGACGTAAATGTTTAGTGTTTTAAATGTTTAAATTTA
TATAAATATTAGGTTATTAGGTTTCGTGATAGTTTTGTGAATTAAGTGGTAAGTTTTAGTTTATTTTA
AATTATAGAAAATAGATAAGGGAGTAAAAGTGTATTAAGGTATAATGTAGTGATTTAAAAAAATATTT
TTGG



cg02732111-BS-F	TTGTTTAGGTTTATTATAGTTTG
cg02732111-BS-R-bio	TCAATAACACATTTCAACAAATAC
cg02732111-BS-SP	TTGTTTAGGTTTATTATAGTTTG
Sequencing entry	*TTTATATGAAC/TGGTTT*

4. cg16204524: DNMT3A2

>hg19_dna range=chr2:26565034-26566533 5'pad=0 3'pad=0 strand=- repeatMasking=none

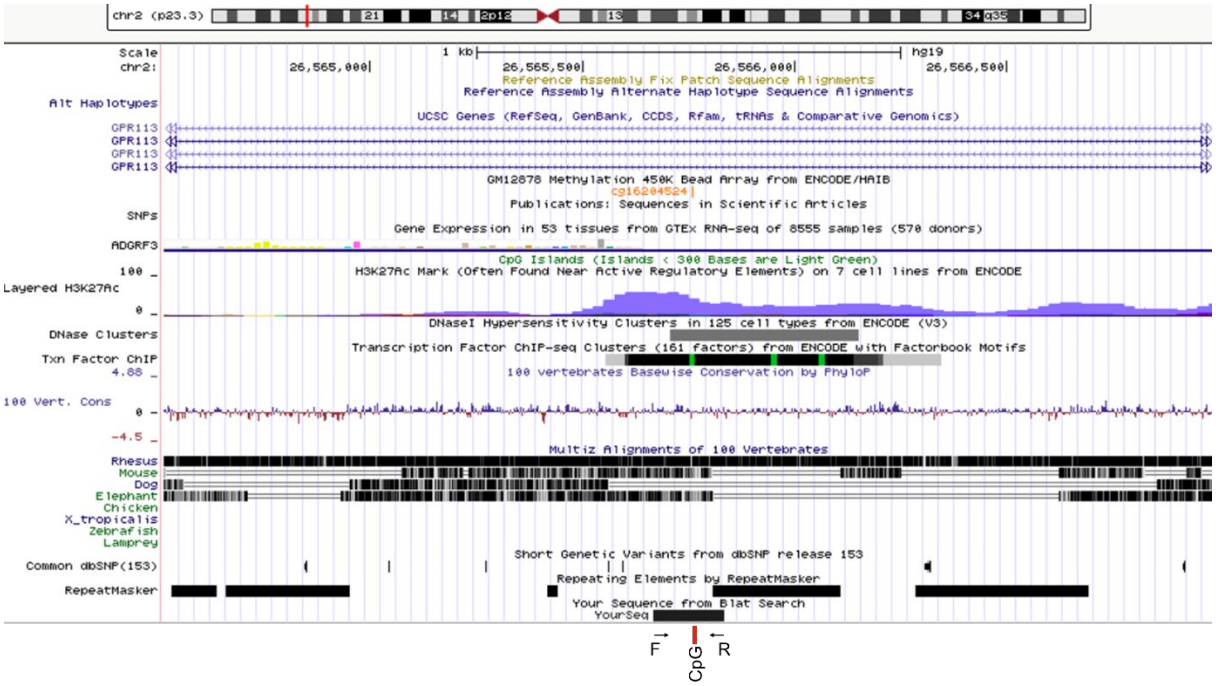
Genomic sequence:

CGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAATATAAAAACAAAATTAGCCGGG
CGTGGTGGCGGGTACCTGTAGTCCCAGCTACTCGGAAGGCTGAGGCAGGAGAATGGTGTGAACCCGGG
AGGCGGAGCTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAGTCTGGGTGACACAGCGAGGCTCTG
TCTCAAAAAAATAAAAAAAAAAAAGATCATTTGATCCTAAGATCTAATTCGATCCATTAGAACCTATT
AGGTACTACGATCCTATTCATCTCTGAAATTCAGTGAAACTGTATATCATATAACTTCTCTGAGATTA
AGTCATCTCATCTGGGAAATCTTGTCTTACCTACAATGATATGGCCCCAGTCTGTTCAATTAAATTCA
TACAAATTTAAGATATTAGTCATGCTCCATATAACCATGTTTCCGTCAATGAGGGACCACATGTGGTC

CCTTAAGATTATAGTGGAGCTGAAAAATTTCTGTCATCTAGTGACATCCTAGCTGTCAAAATGTTGTA
ACACAAGGCATTACTCACGTTCGTGGTAATGCTAGTGTAAACAAACCTACCGCGCTGCCAGTCATCTA
AAAGTATAGCACATATAATTATGTACACTACTGAATACTTGATAATAACTAACTGTCTTACTGGCTTT
TGTATTTACTATCAGAGCATATCCTGTCATTAAGTGATGCATGACTGTATTTTCGCAGAATTCTGTGA
ATCCTAAACCCGGCAGGCATGGTCACGTGGGAAGAGAAAAATGTCATGGAAACAAATTTGCCTTTGGA
ACTAAGTCACCTTTCTGTTTCCTCGTTCCCAAATTTTATTCCAGGAGGCATTTTTCCCTAAGGCAGCC
CAGCTGCCCCGGACTGCTTTAGATTGGCCTTGTTATTCCCATCTGTGCTCTACACGCCACATTCCTGA
GTGTAGTTTATTATTGCTTTGTTTCCCATGTCTGCCTCTTCATCCATTGAAGTTTATTTCTTCTACAT
CTTGTACACTCAAGATCAGGCATTTTGCCTTCCAGAATAATAATCATAGAAAATATCAATTCGTGCTA
GAAACAATATTTTTATTTTTATAATATAACCAAAATAAAGTGTTTTAAGTTCCTGGAGTGCCTGTTTT
ATTTAGTCTGAATAAGAGTGTATCAGATAAATGGAGACAGCCAAAAAATATTGGAGGTTAATTAAGGA
CTAATGTAGTTAATAACTCCAATATAAGTTACTGAAGCCATAAAAATGGGCAGGGTCTTCCAAAGAAG
TTTGTAAAAAGGTATTGATTTAGTTCTCATTTTAAAATAAAAATGAATGCCTTATACACATAAACTCT
GGTTCTCAAAGCAGTTTCACATGACTAACAATCCTGAGCCATGGGCCTGATGGCTGTTCCAATTGTAT
AGAGTGGAACAATAGAATAGTAATATATCTTACCAAAAGTCACCCAGGAAAGAGGAACACAGGTCATC
AGAC

Bisulfite sequence:

CGAGATTATTTTGGTTAATACGGTGAAATTTCGTTTTTATTAAAAATATAAAAATAAAATTAGTCGGG
CGTGGTGGCGGGTATTTGTAGTTTTAGTTATTCGGAAGGTTGAGGTAGGAGAATGGTGTGAATTCGGG
AGGCGGAGTTTGTAGTGAGTCGAGATCGCGTTATTGTATTTTAGTTTGGGTGATATAGCGAGGTTTTG
TTTTAAAAAAATAAAAAAAAAAAAGATTATTTGATTTTAAGATTTAATTCGATTTATTAGAATTTATT
AGGTATTACGATTTTATTTATTTTTGAAATTTAGTGAAATTGTATATTATATAATTTTTTTGAGATTA
AGTTATTTTATTTGGGAAATTTTGTTTTATTTATAATGATATGGTTTTAGTTTGTTTAATTAAATTTA
TATAAATTTAAGATATTAGTTATGTTTTATATAATTATGTTTTCGTTAATGAGGGATTATATGTGGTT
TTTTAAGATTATAGTGGAGTTGAAAAATTTTTGTTATTTAGTGATATTTTAGTTGTTAAAATGTTGTA
ATATAAGGTATTATTTACGTTCGTGGTAATGTTAGTGTAAATAAATTTATCGCGTTGTTAGTTATTTA
AAAGTATAGTATATATAATTATGTATATTATTGAATATTTGATAATAATTAATTGTTTTATTGGTTTT
TGTATTTATTATTAGAGTATATTTTGTTATTAAGTGATGTATGATTGTATTTTCGTAGAATTTTGTGA
ATTTTAAATT*CGGTAGGTATGGTTACGTGGG*AAGAGAAAAATGTTATGGAAATAAATTTGTTTTTGGA
ATTAAGTTATTTTTTTGTTTTTTCGTTTTTTAAATTTTATTTTAGGAGGTATTTTTTTTTTAAGGTAGTT
TAGTTGTTTCGGATTGTTTTAGATTGGTTTTGTTATTTTTATTTGTGTTTTATACGTTATATTTTTGA
GTGTAGTTTATTATTGTTTTGTTTTTTATGTTTGTTTTTTTATTTATTGAAGTTTATTTTTTTTATAT
TTTGTATATTTAAGATTAGGTATTTTGTTTTTTAGAATAATAATTATAGAAAATATTAATTCGTGTTA
GAAATAATATTTTTATTTTTATAATATAATTAAAATAAAGTGTTTTAAGTTTTTGGAGTGTTTGTTTT
ATTTAGTTTGAATAAGAGTGTATTAGATAAATGGAGATAGTTAAAAAATATTGGAGGTTAATTAAGGA
TTAATGTAGTTAATAATTTTAATATAAGTTATTGAAGTTATAAAAATGGGTAGGGTTTTTTAAAGAAG
TTTGTAAAAAGGTATTGATTTAGTTTTTATTTTAAAATAAAAATGAATGTTTTATATATATAAATTTT
GGTTTTTAAAGTAGTTTTATATGATTAATAATTTTGAGTTATGGGTTTGATGGTTGTTTTAATTGTAT
AGAGTGGAATAATAGAATAGTAATATATTTTATTAAAAGTTATTTAGGAAAGAGGAATATAGGTTATT
AGAT



cg16204524-BS-F	ATTTTGTTATTAAGTGATGTATGATTGTAT
cg16204524-BS-R-bio	ACCTCCTAAAATAAAATTTAAAAAC
cg16204524-BS-SP	ATTTTGTGAATTTTAAATT
Sequencing entry	*C/TGGTAGGTATGGTTAC/TGTGGG*

5. cg26286826: DNMT3B4

>hg19_dna range=chr18:65084835-65086334 5'pad=0 3'pad=0 strand=- repeatMasking=none

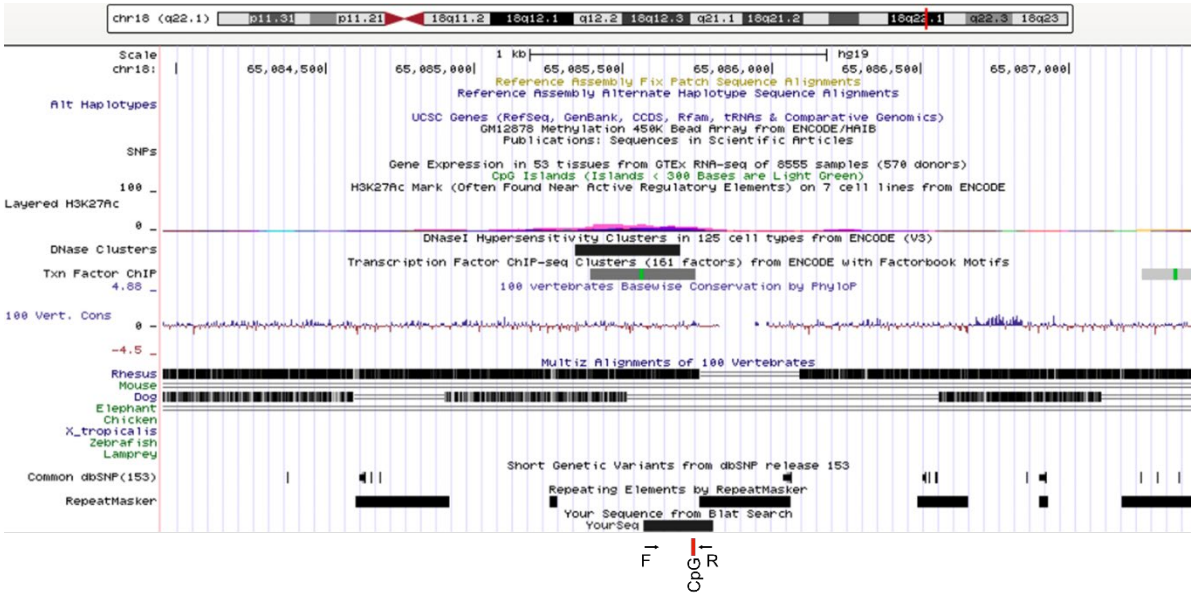
Genomic sequence:

TGAAAGTTTTAAGAATAAAAAGCAAATGCTTGAGAGACATGTCTAAGCCATTAGTTGGTTGATCAGAG
TAATAATTCAACAAATTAATTATTTGACAAGGAAGGACAAGTTTCAAAATCTTCCACAAACTATAATG
AAAATGAGACTGACCACCATCTGTCAATCAGTTCACCAAGTGGCAACAGAACCTAATTAAACTAGTGA
ATCACTAGAAATATGGGAACCTTTGGAAACAGTAAAAGGAAAATACAGGAACTTACGGTCACGCTACT
TTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGC
GCGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGT
AGCTGGGACTACAGGCGCCCGCTACCACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTT
CACCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTG
CTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCCACGCTACTTTTTACCATTCCGAGAAATGTGCGG
TCCATGTATTGCCACCAAACATATATTGATTATTTCAACTATGGGAGTTTACAAGCCTTGAGTGTAAC
CTAGGGTTTGGAATCTGAAATGGAAATGGGTCAGAGCATGAATCA*CGGGCACTAGATCCGAGAGGACT*
CTTTTAGTTGGCTCAGTTGCTTTCACTTTCGTTTGTTGCTGAAGGAGAAACTCACAGCAGAAGGGTGC
ATGAGCTAGAGCTGCTAGTGACTGAATGGAGAGTTCCCTGATTCCTTTCCAGTTCCTTAAGGAAAACT
AATTTGGTGGCTAAAGCAAAGTGAAGCACCTACTTATTTAGAAGAGTCTGGAGGAAATAATTTCCTGT
TTTTTAAAATACCAAGTGAAAGAGGAAACACATCTCCACTCACATGTGTCCTTGCTGTTAGATAAGAC
AGTTTCTTTCAGTATTACAGTCACATATATCTGTGTTCTTTTTTTTTTTTTTTTTTTGCTGTATGAT
CAGACCTACTGCAGTCTTCACTGTGCCCCAACCTATATACACTCTAGCAAACTGATAAGCATTCCAAG
AATAGGTATATGAATCTGTGAACACATTTTGAGCATTTACTCATTCATGTATTAATATTACCTCTAAC
AAGTTTTAAGCTACAATAGCTTAGAAAAAATAGCGATGATCTCACCCAAACTCAAAATTTACCCAAAG
GCTAAACAACATAACCAACCTCATGTAGTAACTCACTGAGAAAGATCAAAAGAAAGAAATGTATTCTT

AGTTCATTGATGAACTGACATCAATATAATTTTTAAACTATAGTACATATTTAGACATGGCCGAGCAC
AGTGGCTCATGCCTGTAATCCCAGCTTTTGGGAGGCCGAGGCAGGCAGATCACCTGAGGTCAGGAGTT
CAAG

Bisulfite sequence:

TGAAAGTTTTAAGAATAAAAAGTAAATGTTTGAGAGATATGTTTAAGTTATTAGTTGGTTGATTAGAG
TAATAATTTAATAAATTAATTATTTGATAAGGAAGGATAAGTTTTAAAATTTTTTATAAATTATAATG
AAAATGAGATTGATTATTATTTGTTAATTAGTTTATTAAGTGGTAATAGAATTTAATTAAATTAGTGA
ATTATTAGAAATATGGGAATTTTTGGAAATAGTAAAAGGAAAATATAGGAATTTACGGTTACGTTATT
TTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTTTCGTTTTGTCGTTTAGGTTGGAGTGTAGTGGC
GCGATTTCGGTTTATTGTAAGTTTCGTTTTTCGGGTTTACGTTATTTTTTTGTTTTAGTTTTTCGAGT
AGTTGGGATTATAGGCGTTCGTTATTACGTTCGGTTAATTTTTTGTATTTTTAGTAGAGACGGGGTTT
TATCGTGTTAGTTAGGATGGTTTCGATTTTTTGATTTCGTGATTCGTTCGTTTCGGTTTTTTAAAGTG
TTGGGATTATAGGCGTGAGTTATCGCGTTCGGTTTACGTTATTTTTTATTATTTCGAGAAATGTGCGG
TTTATGTATTGTTATTAAATATATATTGATTATTTTAATTATGGGAGTTTATAAGTTTTGAGTGTAAT
TTAGGGTTTGGAATTTGAAATGGAAATGGGTTAGAGTATGAATTACGGTATTAGATTCGAGAGGATT
TTTTTAGTTGGTTTAGTTGTTTTTTATTTTCGTTTGTTGTTGAAGGAGAAATTTATAGTAGAAGGGTGT
ATGAGTTAGAGTTGTTAGTGATTGAATGGAGAGTTTTTTGATTTTTTTTTTAGTTTTTTAAGGAAAATT
AATTTGGTGGTTAAAGTAAAGTGAAGTATTTATTTATTTAGAAGAGTTTGGAGGAAATAATTTTTTGT
TTTTTAAAATATTAAGTGAAAGAGGAAATATATTTTTATTTATATGTGTTTTTGTTGTTAGATAAGAT
AGTTTTTTTTTAGTATTATAGTTATATATATTTGTGTTTTTTTTTTTTTTTTTTTTTTTGTTGTATGAT
TAGATTTATTGTAGTTTTTATTGTGTTTTAATTTATATATATTTTAGTAAATTGATAAGTATTTTAAG
AATAGGTATATGAATTTGTGAATATATTTTGAGTATTTATTTATTTATGTATTAATATTATTTTTAAT
AAGTTTTAAGTTATAATAGTTTAGAAAAAATAGCGATGATTTTATTTAAATTTAAAATTTATTTAAAG
GTTAAATAATATAATTAATTTTATGTAGTAATTTATTGAGAAAGATTAAAAGAAAGAAATGTATTTTT
AGTTTATTGATGAATTGATATTAATATAATTTTTAAATTATAGTATATATTTAGATATGGTCGAGTAT
AGTGGTTTATGTTTGTAATTTTAGTTTTTGGGAGGTCGAGGTAGGTAGATTATTTGAGGTTAGGAGTT
TAAG



cg26286826-BS-F	TTTTTTAAAGTGTTGGGATTATAGG
cg26286826-BS-R-bio	CTAAACCAACTAAAAAAATCCTCTC
cg26286826-BS-SP	GGTTAGAGTATGAATTA
Sequencing entry	*C/TGGGTATTAGATTC/TGAG*

6. cg21808287: DNMTΔ3B2

>hg19_dna range=chr2:217724164-217725663 5'pad=0 3'pad=0 strand=+ repeatMasking=none

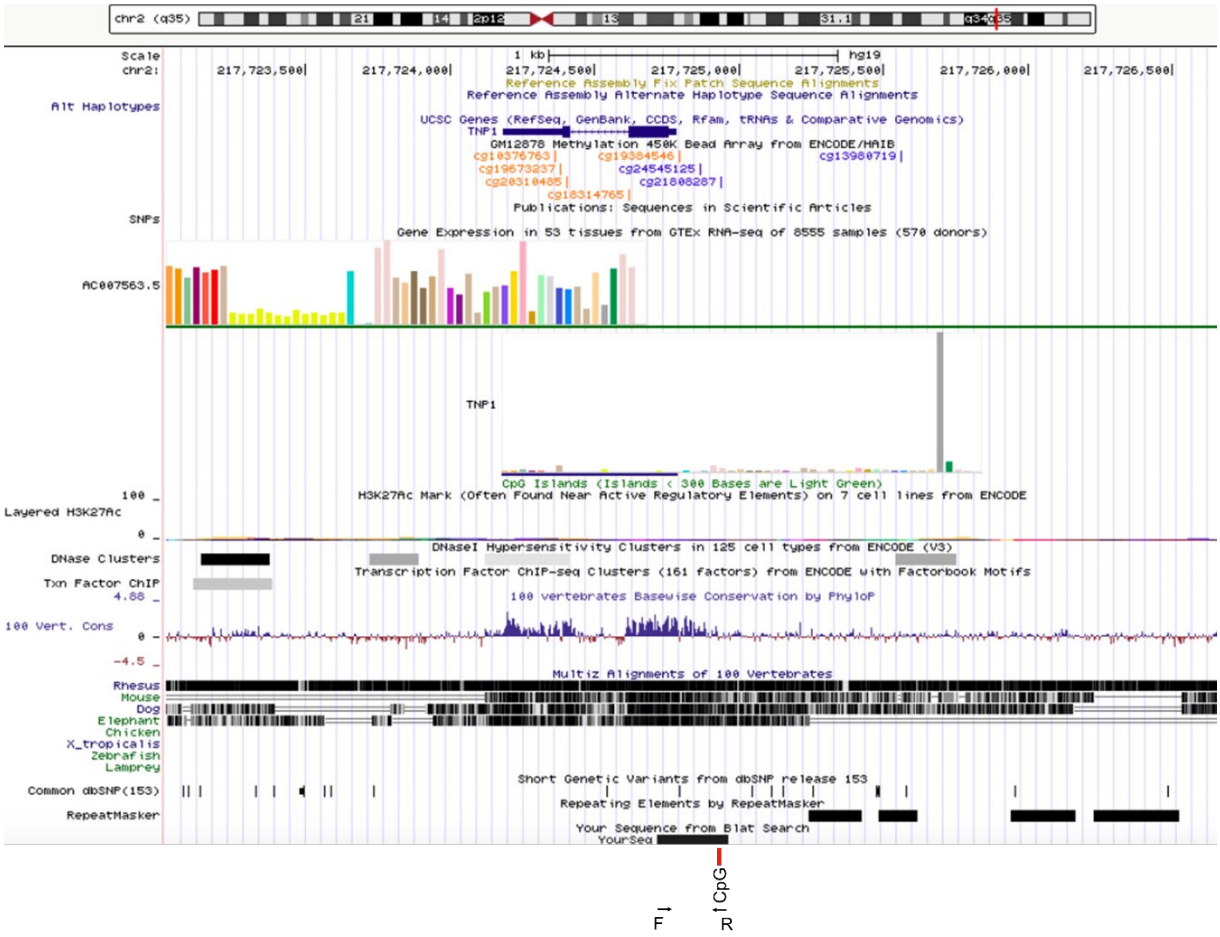
Genomic sequence:

AAAGATCAAAGTAAACATTTCTGCTTTCAAAGTTTTATTGAGATGTTGGGGAAAAACAGCCAACATAT
ACATTCCTCATTTCGTCACAACTGGCATTTGATCCACATTCCATAGGCTCCTCTCTGGCTTTGATCTA
ACATCAGGTCTCCTTGGCAGTTCCCCTTCTGCTGTTCTTGTTGCTGCTTGGTGCTGTGTGAAGCGCAC
CAGGGCAGAGCCCGCTGGGGGCTCACAAGTGGGAGCGGTAATTGCGATTGGCTGTGGGAAAGGACAGG
TTTGCAACAGAGACACGAATTAGAACATTCCCATCAAGACTCTTGTTGCTCAAATCTGCATTTAGTTA
TTGGAAATAATCTCAGGCCAAGATTTCAGGGTCCTGTGAGAGCTGCCTCCAAGAGGGAGGGCCTGGCA
GCAATGTGTGCCTAAGTTTGGCAATCTCCTCCCCATCCCTCACTCACCGTCATCGCCCCGTTTCCTAC
TTTTCAGGTTGCCCTTACGGTATTTTCTTTTGCTGCCACCTCTCTTGACTCCCTTGTGAGGAGATCGG
CTCTTGCTCCTCCTCATGCCATGACTCTTTAATTTGCGGCTGGTCGACATGGTAAGTTCTGCCAAAAT
GAGGGGCTTTGCAAGGCCCGGGGGCCAGGAGTCTGGGTATTTAAGGCCTTGGCCATTGTGACAGGGAA
GGGGCATGGCTCAGCAGGTGGGCGGGGCTCTACTGTGCTGTCACTCACCTTTGCTGCACTTTGGTCAC
ATTTGTGAAATAAGGTGGTCCTTTTCGTGCAGCTCAAGGGCTGCCCTCAGTTTTGAGGTCCTTTTTAT
GTTAAGTGAGGCTTTCCCAGAGAGGTGTTATTGAGAAACTGAATCCCAGCCCTGTGCCAGGCATTGGA
TGAGTACAAGAGACCATTCAGATATGCTTCTCCTGCCTGTAGGAAATGTAACAGTCATTGAGTGTGTG
TGTATGCATGACTGTTTAAGTCTCTCACAGCATCCAGAGATAATAAATTTATTTCTGTAAGGACTTGA
CTCTACAGAACCCAATTCAATATGATCAAAATAGAAATATTCCTTCTTTTTCATCTAGGTCAGGATTT
CTCAGTCTTGGCACTGTTGACACAATGGGCAGGGTAACTTTCAGTGGGCTGTCTTGTGCACCCTAGGA
TGTTTGGCAACATCCCTGGCCATTACCCACTGGACACCAGAGAAGCCTGGCAGCTATGACAACCCAAA
ATATTCCCAAACATTGCCACCTGTTTCTTTGGGAGGCTGGGGAAGTGAGGGAAAAAATGGCTCTGGTT
CAGAACAATTAACCTAATCATTTTTTAAAATTCAGTCACAGATATTGATCACATAGCATATGCCATAC
ATTTAGGTGGTAGAAAATTAGTGGAGGATAAACACACGTAGATACTGTGCTCATGGAACTCTCAGCCC
ACTGGGGAGAAAGCCATTCATCAGATGTCTACATAAATCAGTGAGTAAGTACTTACGGAGTCATTTTG
GCTG

Bisulfite sequence:

AAAGATTAAAGTAAATATTTTTGTTTTTAAAGTTTTATTGAGATGTTGGGGAAAAATAGTTAATATAT
ATATTTTTTATTTCGTTATAATTGGTATTTGATTTATATTTTATAGGTTTTTTTTTGGTTTTGATTTA
ATATTAGGTTTTTTTGGTAGTTTTTTTTTTGTTGTTTTTGTTGTTGTTTGGTGTTGTGTGAAGCGTAT
TAGGGTAGAGTTCGTTGGGGGTTTATAAGTGGGAGCGGTAATTGCGATTGGTTGTGGGAAAGGATAGG
TTTGTAATAGAGATACGAATTAGAATATTTTTATTAAGATTTTTGTTGTTTAAATTTGTATTTAGTTA
TTGGAAATAATTTTAGGTTAAGATTTTAGGGTTTTGTGAGAGTTGTTTTTAAGAGGGAGGGTTTGGTA
GTAATGTGTGTTTAAGTTTGGTAATTTTTTTTTTATTTTTTATTTATCGTTATCGTTTCGTTTTTTAT
TTTTTAGGTTGTTTTTTACGGTATTTTTTTTTTGTTGTTATTTTTTTTTGATTTTTTTTGTGAGGAGATCGG
TTTTTGTTTTTTTTTTATGTTATGATTTTTTAATTTGCGGTTGGTCGATATGGTAAGTTTTGTTAAAAT
GAGGGGTTTTGTAAGGTTCGGGGGTTAGGAGTTTGGGTATTTAAGGTTTTGGTTATTGTGATAGGGAA
GGGGTATGGTTTAGTAGGTGGGCGGGGTTTTATTGTGTTGTTATTTATTTTTGTTGTATTTTGGTTAT
ATTTGTGAAATAAGGTGGTTTTTTTTCGTGTAGTTTAAGGGTTGTTTTTAGTTTTGAGGTTTTTTTTAT
GTTAAGTGAGGTTTTTTTAGAGAGGTGTTATTGAGAAATTGAATTTTAGTTTTGTGTTAGGTATTGGA
TGAGTATAAGAGATTATTTAGATATGTTTTTTTTGTTTGTAGGAAATGTAATAGTTATTGAGTGTGTG
TGTATGTATGATTGTTTAAGTTTTTTATAGTATTTAGAGATAATAAATTTATTTTTGTAAGGATTTGA
TTTTATAGAATTTAATTTAATATGATTAAAATAGAAATATTTTTTTTTTTTTATTTAGGTTAGGATTT
TTTAGTTTTGGTATTGTTGATATAATGGGTAGGGTAATTTTTAGTGGGTTGTTTTGTGTATTTTAGGA
TGTTTGGTAATATTTTTGGTTATTATTTATTGGATATTAGAGAAGTTTGGTAGTTATGATAATTTAAA
ATATTTTTAAATATTGTTATTTGTTTTTTTGGGAGGTTGGGGAAGTGAGGGAAAAAATGGTTTTGGTT
TAGAATAATTAATTTAATTATTTTTTAAAATTTAGTTATAGATATTGATTATATAGTATATGTTATAT

ATTTAGGTGGTAGAAAATTAGTGGAGGATAAATATACGTAGATATTGTGTTTATGGAATTTTTAGTTT
ATTGGGGAGAAAGTTATTTATTAGATGTTTATATAAATTAGTGAGTAAGTATTTACGGAGTTATTTTG
GTTG



cg21808287-BS-F	TTTTATGTTATGATTTTTTAATTTG
cg21808287-BS-R-bio	CTAAAAACAACCCTTAAACTACA
cg21808287-BS-SP	TATATTTGTGAAATAAGGTGG
Sequencing entry	*TTTTTTC/GTGTAGT*

7. cg25533247: DNMTΔ3B2

>hg19_dna range=chr19:15529905-15531404 5'pad=0 3'pad=0 strand=+ repeatMasking=none

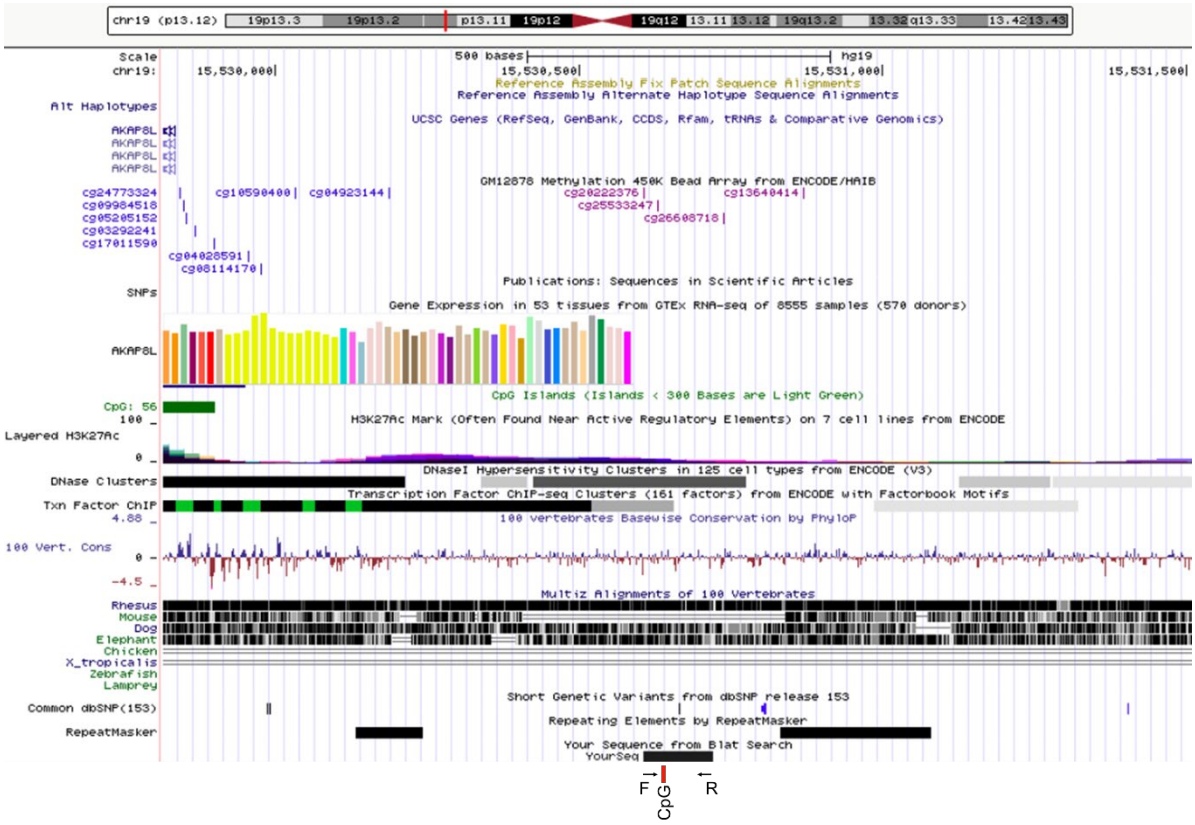
Genomic sequence:

GTCTATCTCTATGCCTCCTCATATCAGTGGCTGGAAGCTGCTGGCAATCTCGTGAGAGCCCCGCCTTG
TCCCCGCCTTCTGTTATTCTAAGGCCTCAGTGATTGGCCTACAGAAGCCGGCAGAGAGGCGGAGTCCT
ATAGAATTTGAGGGGGCGGAGCTGTGCACTACGGAAAGGCCCAAGGCCCAAGCATGGTCGGGCTGGAG
GAGCGCCCGTACCCTGGAGACCCGCCAGAGGGAGTCGGGTCTACGGCTCAGGGCAGGGCTGGGATGCC
AACTCCGCCACCGAGGTGCTGGGTAGGGGACCTAGCTAAGTCTCTCAAATTCTCTGGCCTTCGGCTCC
AAAACAAGCATAGCAAAGCCATGTTGAGAAGAGAAAATTAAAACAGCACCAAACAAGATGCTGGACAC
ACGTTTCAACATTACGTGACAAAGTGTGTGTAAAGAAATCATTCTGGTGCCAGTGCACTGAATAAATG
TGATTAATCGTAACAAAACGATAATATATTTATATGGCATATGCTTATATGTGTAATAATTCCGTTTG
TATTGTGTGTCATTCGAATGCTCCTTCAAACCTGCAAAAAGCTTTTAATCCTCACTTAAAACTGAAGC

TTGGAACTCATGATTTGCTCAAGGTCCTTGACAGTTACGTGTTAAAATGGGAGCTTTGCTTCTAAAAC
CCTCCTGATTCACTCTGGGCTCGAGAGATCCTGCAATAGTGCAGCCGGGCACCAGTCATGTCCTCAGT
CTGTTTCCTGGGCATATGGAAACCCTGCCAGGGACAACTCATAATTTCTAGAGATGACAGAGCTGCAC
CGAAGAGTACTTAACACGTGTGTTTGTGAAAACTCATTAAGCTGCGCAATTACGACTTCTACGCTTCG
TGTGTTTTAAAACTTTTTTTAAGGTATGCTGTTTCTTTTTGATATTGTTGATTTCTTTAGGTGTAATA
GTGGCATTTACAACGTTGAGAGAAACATACTGAAATACTAAGAAGTTAACTGATCTTATGGCCAGGAT
GTGCTTCACACAATTCAGACAGGTAGAGAGAGGAGGAGGAAGCAAGGAGCTATAAGATAAATCTCTGA
CAGCTGAAGACATTGGTCATGGGTTGGTGATGGATAGATGAGAGTCCATTTTACTGTTTGCTCTACTT
TTGTATAGTTGGAATTCTCCATGATAAACTGTTTTAATAAAACTCCCCCCACCCAAATGAGGCTTCTG
CAGAACTTGCTGCCATACGGAACAAAGCCACTTCTGATGTGTTAAGTGAGCAAACTCCAAACATGAAG
CAGAGATCTGACGGATAAGCACTCTGGCTTCGGGGTTCATCTCCCCATCAGCCTCAGGGATGCCTGCC
TGTCTTCCAAACCAAGGCAGGGCCCCAGGACACGTCTTCTGCGCCCTGCAGGTTCGTTCCTTCATCGC
ACTTGCCACAGTCATCATTGTACCGATTTGATTTCGTTCACATCCAACACCCTCAGAGACCACGGCTC
AAAA

Bisulfite sequence:

GTTTATTTTTATGTTTTTTTATATTAGTGGTTGGAAGTTGTTGGTAATTTCGTGAGAGTTTCGTTTTG
TTTTCGTTTTTTGTTATTTTAAGGTTTTAGTGATTGGTTTATAGAAGTCGGTAGAGAGGCGGAGTTTT
ATAGAATTTGAGGGGGCGGAGTTGTGTATTACGGAAAGGTTTAAGGTTTAAGTATGGTCGGGTTGGAG
GAGCGTTCGTATTTTGGAGATTCGTTAGAGGGAGTCGGGTTTACGGTTTAGGGTAGGGTTGGGATGTT
AATTTCGTTATCGAGGTGTTGGGTAGGGGATTTAGTTAAGTTTTTTAAATTTTTTGGTTTTCGGTTTT
AAAATAAGTATAGTAAAGTTATGTTGAGAAGAGAAAATTAAAATAGTATTAAATAAGATGTTGGATAT
ACGTTTTAATATTACGTGATAAAGTGTGTGTAAAGAAATTATTTTGGTGTTAGTGTATTGAATAAATG
TGATTAATCGTAATAAAACGATAATATATTTATATGGTATATGTTTATATGTGTAATAATTTCGTTTG
TATTGTGTGTTATTCGAATGTTTTTTTAAATTTGTAAAAAGTTTTTAATTTTTATTTAAAATTGAAGT
TTGGAATTTATGATTTGTTTAAGGTTTTTGATAGTTACGTGTTAAAATGGGAGTTTTGTTTTTAAAAT
TTTTTTGATTTATTTTGGGTTCGAGAGATTTTGTAATAGTGTAGTCGGGTATTAGTTATGTTTTTAGT
TTGTTTTTTGGGTATATGGAAATTTTGTTAGGGATAATTTATAATTTTTAGAGATGATAGAGTTGTAT
CGAAGAGTATTTAATACGTGTGTTTGTGAAAATTTATTAAGTTGCGTAATTACGATTTTTACGTTTCG
TGTGTTTTAAAATTTTTTTTTAAGGTATGTTGTTTTTTTTTGATATTGTTGATTTTTTTAGGTGTAATA
GTGGTATTTATAACGTTGAGAGAAATATATTGAAATATTAAGAAGTTAATTGATTTTATGGTTAGGAT
GTGTTTTATATAATTTAGATAGGTAGAGAGAGGAGGAGGAAGTAAGGAGTTATAAGATAAATTTTTGA
TAGTTGAAGATATTGGTTATGGGTTGGTGATGGATAGATGAGAGTTTATTTTATTGTTTGTTTTATTT
TTGTATAGTTGGAATTTTTTATGATAAATTGTTTTAATAAAATTTTTTTTATTTAAATGAGGTTTTTG
TAGAATTTGTTGTTATACGGAATAAAGTTATTTTTGATGTGTTAAGTGAGTAAATTTTAAATATGAAG
TAGAGATTTGACGGATAAGTATTTTGGTTTCGGGGTTTATTTTTTTATTAGTTTTAGGGATGTTTGTT
TGTTTTTTAAATTAAGGTAGGGTTTTAGGATACGTTTTTTGCGTTTTGTAGGTTCGTTTTTTTATCGT
ATTTGTTATAGTTATTATTGTATCGATTTGATTTCGTTTATATTTAATATTTTTAGAGATTACGGTTT
AAAA



cg25533247-BS-F	AGAGATTTTGTAATAGTGTAGT
cg25533247-BS-R-bio	ATACAACTCTATCATCTCTAAA
cg25533247-BS-SP	AGAGATTTTGTAATAGTGTAGT
Sequencing entry	*C/TGGGTATTAGTTATGT*

8. cg08927738: DNMTΔ3B3

>hg19_dna range=chr20:52686669-52688168 5'pad=0 3'pad=0 strand=- repeatMasking=none

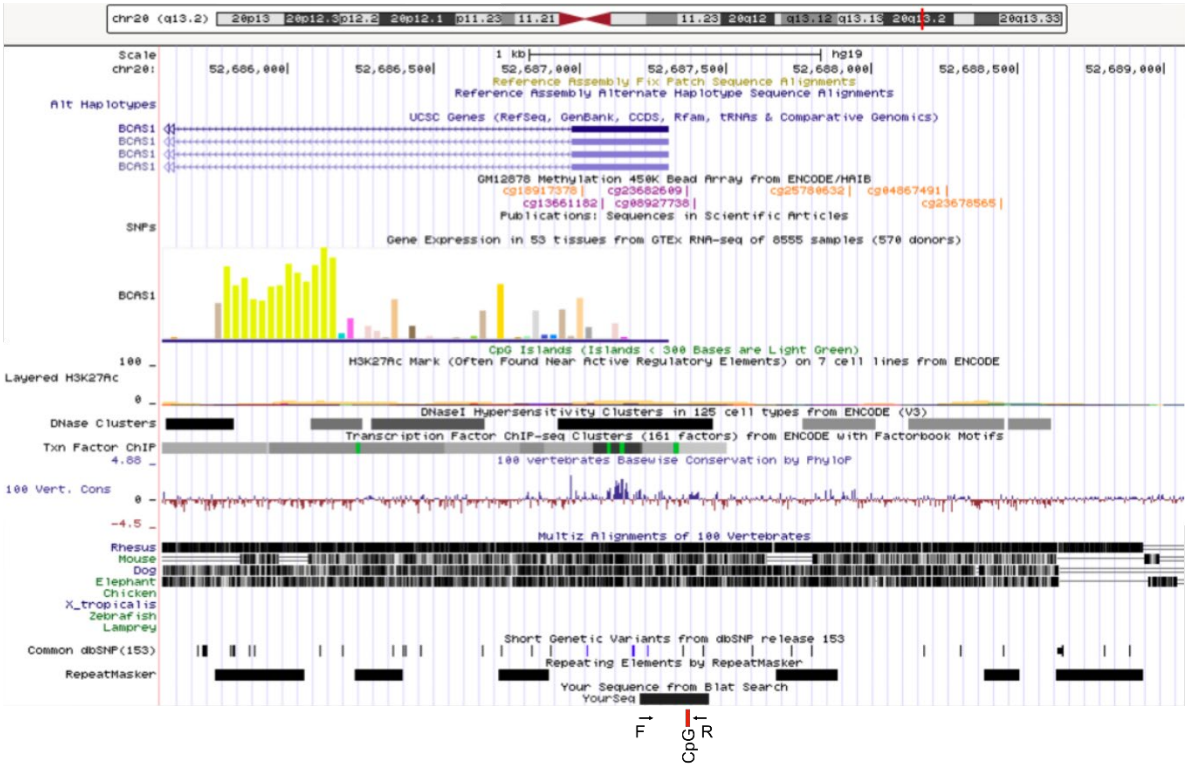
Genomic sequence:

TCTATTTCTCTCCCTTAGCTGCCTTGGCTTTTCTTTTATAATGAGAATCACACCCTGTAATTCTATGT
TTATATTTACAGGATTCTATGTCCAATGACACTGAGTGCAAGAAGGGTAGAGACCCAGCATCTTTTAC
TCACCTCTTTAAACTCTGGCTTATATATATAGCCTGGTGCTTGACAAGTATTCATGATAGGGGTTAAG
CTGAGTATACTGTTGAATAAAACCACAATAATTATTGGTCGCAATTCTCTTGAATGTCTTCTCCAGAG
CACTAAGTAGTGTAACTTATCAAGTTTCCTTTCTTGCATTGGCTGTTAGGAAGCTCAAAGCCATATCT
GCAAACCCTTCATAGCAAGTCTGTAAATCAACAAGAGCATTTATTTTTGATCCTGTTCTCATTTGATG
TTCCAAATTTTGTTCTTCAATCTTCTTTTGTTACATGTTATTTCATGCTCTCATAATCTATTGTATAT
GTCCTTGTAAGACACTTTAAACAACATCCTTGCTTGAATTAAATAAATTGCATAGCCTGTCGAATTTC
TGGCCTGCAAAGAAAAAGAATGCCCTTTCCAACATCATGGTATGTCTCCAGAGCAATAAACCACCCTC
AGGTTACCTGATTTCTGGTGTCCAAACCCCTTTGGGAATAAGGTTTACATGACTAGGGGTAGTTGAGT
GGATGAGAAGCTATTTCCTTCCAATTAGGTGGTTAGGTGGTGGTGGAGGTGGTGTTCTGAAGTCAGCA
GACAGAACTACAAAATCTCACAGAA*CGTGTTTTTCCCTGCTC*TGACCATGACCTACGGTTACTGGGGAA
GGAAACTGCTTCCCCAGGTCAACCCGGCAGCCTCAGTAGGTGAGGGGCACTGGGTAGAATACTTGGGG
TGCCAGGGAGGCATTAATGCGAGAGGAGTCAGGTGCTCAGTTTTTATTGGAGTTGGGAGGGCAGCCCC
ACATCAGGAAGAGAACCTGTTTCTGCAGGATGGTCCGGGGAGAAGGGAGGACTCCACCCAGGCTTGTG
TTTGCCCTGCTCTGTGTATTCAGCCAGCAGGCTCTGCACAAGGAAGCAAAGTGCAGGGAGCCAGGCTC

CACCGACAGCCAGGCACTGGGCAGCACGCACTGGAGACCCAGGACCCTGTGCAGGAGCAGCTCCGGGT
GACACGAGGGGACTGAAGATACTCCCACAGGGGCTCAGCAGGTAACTGCTTTCAGATCCTTTAGGCAT
GGCTGGAGGGTTTGCTGTTTGTGGTAGACGGTGTCATATAGAGGCCAGGGGTCAGACAGGGTCAAGTT
TGATCCCAGCTCTGCCTTCTGTTCACTGTGTGGCCTTGGATATGTGTAAGCCTCAGTTCCACTATCTG
TGAAATGGACCCAAAGATAGCACTCATCTCAAAGTTTTTAGTTTTACAAGGAGCAAATGACAATGTGG
TCAGTGTTAGGCGCATAGCCACATAAATAACTCCTATTTGGTGGAACTATTGAGGGTCATTTCAAGCA
CAGG

Bisulfite sequence:

TTTATTTTTTTTTTTTAGTTGTTTTGGTTTTTTTTTTATAATGAGAATTATATTTTGTAATTTTATGT
TTATATTTATAGGATTTTATGTTTAATGATATTGAGTGTAAGAAGGGTAGAGATTTAGTATTTTTTAT
TTATTTTTTTAAATTTTGGTTTATATATATAGTTTGGTGTTTGATAAGTATTTATGATAGGGGTTAAG
TTGAGTATATTGTTGAATAAAATTATAATAATTATTGGTCGTAATTTTTTTGAATGTTTTTTTTAGAG
TATTAAGTAGTGTAATTTATTAAGTTTTTTTTTTTGTATTGGTTGTTAGGAAGTTTAAAGTTATATTT
GTAAATTTTTTATAGTAAGTTTGTAAATTAATAAGAGTATTTATTTTTGATTTTGTTTTTATTTGATG
TTTTAAATTTTGTTTTTTAATTTTTTTTTTGTTATATGTTATTTTATGTTTTTATAATTTATTGTATAT
GTTTTTGTAAGATATTTTAAATAATATTTTTGTTTGAATTAAATAAATTGTATAGTTTGTCGAATTTT
TGGTTTGTAAAGAAAAAGAATGTTTTTTTTTAATATTATGGTATGTTTTTAGAGTAATAAATTATTTTT
AGGTTATTTGATTTTTGGTGTTTAAATTTTTTTGGGAATAAGGTTTATATGATTAGGGGTAGTTGAGT
GGATGAGAAGTTATTTTTTTTTTAATTAGGTGGTTAGGTGGTGGTGGAGGTGGTGTTTTGAAGTTAGTA
GATAGAATTATAAAATTTTATAGAACGTGTTTTTTTTTGTTTTGATTATGATTTACGGTTATTGGGGAA
GGAAATTGTTTTTTTAGGTTAATTCGGTAGTTTTAGTAGGTGAGGGGTATTGGGTAGAATATTTGGGG
TGTTAGGGAGGTATTAATGCGAGAGGAGTTAGGTGTTTAGTTTTTATTGGAGTTGGGAGGGTAGTTTT
ATATTAGGAAGAGAATTTGTTTTTGTAGGATGGTTCGGGGAGAAGGGAGGATTTTATTTAGGTTTGTG
TTTGTTTTGTTTTGTGTATTTAGTTAGTAGGTTTTGTATAAGGAAGTAAAGTGTAGGGAGTTAGGTTT
TATCGATAGTTAGGTATTGGGTAGTACGTATTGGAGATTTAGGATTTTGTGTAGGAGTAGTTTCGGGT
GATACGAGGGGATTGAAGATATTTTTATAGGGGTTTAGTAGGTAATTGTTTTTAGATTTTTTAGGTAT
GGTTGGAGGGTTTGTTGTTTGTGGTAGACGGTGTTATATAGAGGTTAGGGGTTAGATAGGGTTAAGTT
TGATTTTAGTTTTGTTTTTTGTTTATTGTGTGGTTTTGGATATGTGTAAGTTTTAGTTTTATTATTTG
TGAAATGGATTTAAAGATAGTATTTATTTTAAAGTTTTTAGTTTTATAAGGAGTAAATGATAATGTGG
TTAGTGTTAGGCGTATAGTTATATAAATAATTTTTATTTGGTGGAATTATTGAGGGTTATTTTAAGTA
TAGG



cg08927738-BS-F	AGGTGGTGTTTTGAAGTTAGTAGATAGA
cg08927738-BS-R-bio	CTTCCTAATATAAAACTACCCTCCCA
cg08927738-BS-SP	ATTATAAAATTTTATAGAA
Sequencing entry	*C/GTGTTTTTTTTGTT*

9. cg20364776: DNMTΔ3B3

>hg19_dna range=chr4:106405077-106406576 5'pad=0 3'pad=0 strand=+ repeatMasking=none

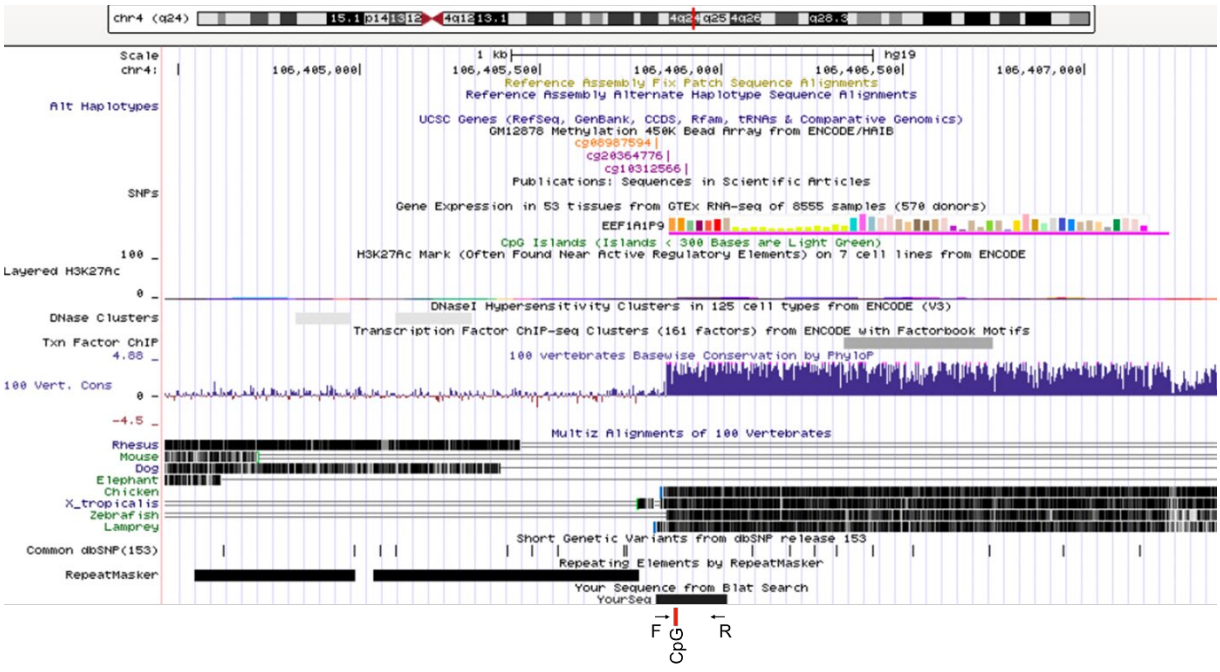
Genomic sequence:

TTTCGAGTGTACTATTTAGTGGTACTTAGCATAGTCACAATATTGTACAACTGCCACTTCCTTCTAGT
TTCAAAACATTTTCATCATCCCAGAAGAATACCTCATATCCTTTAAGTAATCACTTCCTGTTCCCCCC
TGCCACCCGCTGACATCCACTAATCAGCTTGCCTATTTTGGATATTTCACATAAGAGGAATCATATGA
CATGCGGCTTTTTGTGTCTAGTGTCTTTCACTTAGCATAATGCTTTTGGTGTCCATCCACATTCTAGA
ATATATGAAACCTTCATTCCTTTTTATGCCTGAATAATATTTCATTGTATATCATTGTACGTTATGAT
TTGTTTATCACACTCAACAATAAAAAGGCAACCTACAGAATGGGAGAAAATTTTTGCAATCTAGCCAT
CTGACAAAGAGTTAATATCCAGAATCTACAAAGAACTTAAACAAATTTACAAGAAAAAACAGACAACC
CCATCAAAAAGTGGGTGAAGGATATGAACAGACACTTCTCAAGACACTTACGCAGCCAACAAATATAT
GAGAACAGACACTTCTCAAAAGAAGACATTTATGAGGTCAAAAAACATTAAAAAGAGCTCATCATCAC
TGGTCATTGAGAAATGCAAGTCAAAACCACAATGAGATACCATCTCACGCCAGTTAGAATGACGATCA
TTAAAAAGTCAGGAAACTTTTTTGTTTTTTTCTCAATGGGTTTGCCACCAGAACACAGGTATCGTGAA
AACTACCCCTACCTATAAGCCAGAA*CGAGAAAGGAAAAGA*CTCATATCAACACTGTGATCATCGGACA
CGTAGATTCGGGCAAGTCCACCACTACTGGCCATCTGATCTACAAATGCGGTGGCGTCGACAAAAGAA
CCATCGAAAAATTTGAGAAGGAGGCTGCTGAGATGGGAAAGTGCTCCTTCAAGTATGCCTGGGTCTTG
GATAAACTGAAAGCTGAGCGTGAACATGGTATCACCATTGATATCTCTTTGTGGAAATTTGAGACCAG
CAAGTACTATGTGACTATCATTGATGCCCCAGGACACAGAGACCTCATCAAAAACATGATTACAGGGA
CATCTCAGGCTGACTGTGCTGTCTTGATTGTTGCTGCTGGTTTTGGTGAATTTGAAGCTGGTATCTCC
AAGAATGGGCAGACCCGAGAGCACGCCCTTCTGGCTTACACACTGGGTGTGAAACAACTAATTGTTGG

TGTTAACAAAATGGATTCCACTGAGCCACCCTACAGCCATAAGAGATATGAGGAAATTGTTAAGGAAG
TCAGCACTTACATTAAGAAAATTGGCCACAACACCGACACAGTAGCATTTGTGCCAGTTTCTGGTTGG
AATGGTGACAACACGCTGGAGCCAAGTGCTAACATGCCTTGGTTCAAGGGATGGAAAGTCACCCGTAA
GGATGGCAATGCCAGTGGAACCACGCTGCTTGAGGCTCTTGACTGCATCCTACCACCAACTCGTCCAA
CTGA

Bisulfite sequence:

TTTCGAGTGTATTATTTAGTGGTATTTAGTATAGTTATAATATTGTATAATTGTTATTTTTTTTTAGT
TTTAAAATATTTTTATTATTTTAGAAGAATATTTTATATTTTTTAAGTAATTATTTTTTGTTTTTTTT
TGTTATTCGTTGATATTTATTAATTAGTTTGTTTATTTTGGATATTTTATATAAGAGGAATTATATGA
TATGCGGTTTTTTGTGTTTAGTGTTTTTTATTTAGTATAATGTTTTTGGTGTTTATTTATATTTTAGA
ATATATGAAATTTTTATTTTTTTTTTATGTTTGAATAATATTTTATTGTATATTATTGTACGTTATGAT
TTGTTTATTATATTTAATAATAAAAAGGTAATTTATAGAATGGGAGAAAATTTTTGTAATTTAGTTAT
TTGATAAAGAGTTAATATTTAGAATTTATAAAGAATTTAAATAAATTTATAAGAAAAAATAGATAATT
TTATTAAAAAGTGGGTGAAGGATATGAATAGATATTTTTTAAGATATTTACGTAGTTAATAAATATAT
GAGAATAGATATTTTTTAAAAGAAGATATTTATGAGGTTAAAAAATATTAAAAAGAGTTTATTATTAT
TGGTTATTGAGAAATGTAAGTTAAAATTATAATGAGATATTATTTTACGTTAGTTAGAATGACGATTA
TTAAAAAGTTAGGAAATTTTTTTGTTTTTTTTTTTAATGGGTTTGTTATTAGAATATAGGTATCGTGAA
AATTATTTTTATTTATAAGTTAGAACGAGAAAGGAAAAAGATTTATATTAATATTGTGATTATCGGATA
CGTAGATTCGGGTAAGTTTATTATTATTGGTTATTTGATTTATAAATGCGGTGGCGTCGATAAAAGAA
TTATCGAAAAATTTGAGAAGGAGGTTGTTGAGATGGGAAAGTGTTTTTTTAAGTATGTTTGGGTTTTG
GATAAATTGAAAGTTGAGCGTGAATATGGTATTATTATTGATATTTTTTTGTGGAAATTTGAGATTAG
TAAGTATTATGTGATTATTATTGATGTTTTAGGATATAGAGATTTTATTAAAAATATGATTATAGGGA
TATTTTAGGTTGATTGTGTTGTTTTGATTGTTGTTGTTGGTTTTGGTGAATTTGAAGTTGGTATTTTT
AAGAATGGGTAGATTCGAGAGTACGTTTTTTTGGTTTATATATTGGGTGTGAAATAATTAATTGTTGG
TGTTAATAAAATGGATTTTATTGAGTTATTTTATAGTTATAAGAGATATGAGGAAATTGTTAAGGAAG
TTAGTATTTATATTAAGAAAATTGGTTATAATATCGATATAGTAGTATTTGTGTTAGTTTTTGGTTGG
AATGGTGATAATACGTTGGAGTTAAGTGTTAATATGTTTTGGTTTAAGGGATGGAAAGTTATTCGTAA
GGATGGTAATGTTAGTGGAATTACGTTGTTTGAGGTTTTTGATTGTATTTTATTATTAATTCGTTTAA
TTGA



cg20364776-BS-F	TGAAAATTATTTTTATTTATAAGTTAGAA
cg20364776-BS-R-bio	ACTTAAAAAAACACTTTCCCATCTC
cg20364776-BS-SP	TGAAAATTATTTTTATTTATAAGTTAGAA
Sequencing entry	*C/GAGAAAGGAAAAG*

10. cg07504154: DNMTΔ3B4

>hg19_dna range=chr4:177158662-177160161 5'pad=0 3'pad=0 strand=- repeatMasking=none

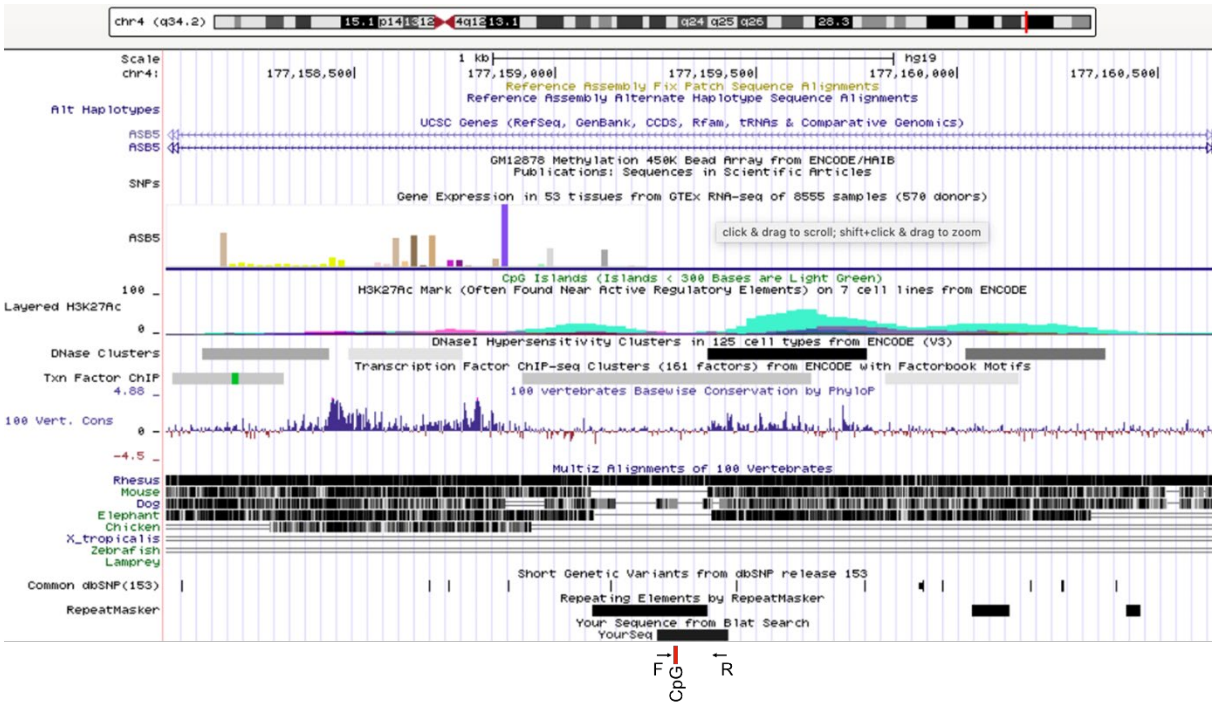
Genomic sequence:

TACCTTTTGTGTGTTAGTCATAGAATCTGGTCTTATCCAAGCTCCTCATTTTATAGAGAAGGAAACTA
TGGTGTCATTAAGCTACTCGCTCCAGGTCACACAAACAGCTGGTAGGATTAGAACTTAATTGAAGATA
AGGAAAAGATATGTGTGGTATTCATTAGCGATAACTTGAGGTAAACACTGGCAGTGTAATATGGCAGC
AAATGTCCATAGGGTAGTGCATAGGACAGAATTATAGTCATAATTTTTTTTTCTCTTACTTGGCTTGT
TACATGAAGACAATAACCCCAGTAATACTTTTTTCCTTTTATATACTAGCAGAAGGTAGAGCCTTAAA
AATACTGATGAAAAAAAGATAGATAACAGAATCTGTCACCATATTTCAGCAAGACAGATTTCAGTATG
TTAATACTCTGTAACGTGATCACCAGCTCAGTTAATCCTTGGAGTCCTATCATAGTTCGGTCGCTGCT
CTTGATTCCATCATAGACTCTGTGCCTTTTTCAACTTTATAGGCCAGTAAAAAATACCAGGCACATGT
TTGGTTTTTCCTCTCATCCTGCCAACTACCCTTTCCTCTGCCTTTTAGGGTGAAAAATCTCAGCTGTA
CTTAAATTGACTGCTGACGTCTTTTGATTTAGAGATGACCTCTAGAAAAATCACAATAGTGAGCTGTG
CCTATATAAATAAGAGGACTGTGATTTTGTATTGAAGCCAATTCTAAACTTAGATCCTTGCAAAATGA
AAACATGAGCTG*GCTACCAGCTGAC*CG*TAATG*CACTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCT
CCTTCTGTCGCCCAGGCTGGAGTGCAGTGGTACAATCTCTGCCTCCCGGGTTCAAGTGATTCTCCCGCC
CTCAGCCTTCTGAGTAGCTGGGACTACAGGCGCCCCGTCCCTATGTCCAGCTAATTTTTGTATTTTTA
GTAAAGACGGGGTTTCACCACGTTGGCCAGGCTGGTCTCTAACCCCTGACCTCAGGTGATCTGCCTGC
CTCGGCCTCCCAAGGTGCTGGGATTACAGGCGTGAGCCACCACGCCCAGCCCTAAGGCACATCTCATA
CTTACTCCTGAGGATGAAGGAGAAAGGCTAACCTAGTCACTGATCACTACTGAGCTTTCTGGCTGCAA
TATTCTGTCTGCAGTCATTATCAACAGGCAATTTTACATTTGATTGATTTGTAGGTTCTTTAAATAGG
CTTTTATCTTTTAACGATAGAGCTCATTTCTATTCTGTTTCGACTAGGAAAAGCAATTATACCCACTT
TTTACTTTCTATATCACACAAATAGTTAAAGACTTCCGGCTGCAGCATGTCTCTGCTTTAAAAATAGA
AAATGCTCCAGCATTTCTGGCCCATCCTCCACATTGCACCTACCAGCCCTGATCCAAGTGTTTCTGAA
TAGAGTTCAATGTTGAAGACATCTCAAGTTAGCACACAGGAACCTTTGTGAAAGTCTGAAGGGAAGGA
AAAA

Bisulfite sequence:

TATTTTTTGTGTGTTAGTTATAGAATTTGGTTTTATTTAAGTTTTTTATTTTATAGAGAAGGAAATTA
TGGTGTTATTAAGTTATTCGTTTTAGGTTATATAAATAGTTGGTAGGATTAGAATTTAATTGAAGATA
AGGAAAAGATATGTGTGGTATTTATTAGCGATAATTTGAGGTAAATATTGGTAGTGTAATATGGTAGT
AAATGTTTATAGGGTAGTGTATAGGATAGAATTATAGTTATAATTTTTTTTTTTTTTATTTGGTTTGT
TATATGAAGATAATAATTTTAGTAATATTTTTTTTTTTTATATATTAGTAGAAGGTAGAGTTTTAAA
AATATTGATGAAAAAAAGATAGATAATAGAATTTGTTATTATATTTTAGTAAGATAGATTTTAGTATG
TTAATATTTTGTAACGTGATTATTAGTTTAGTTAATTTTTGGAGTTTTATTATAGTTCGGTCGTTGTT
TTTGATTTTATTATAGATTTTGTGTTTTTTTTTAATTTTATAGGTTAGTAAAAAATATTAGGTATATGT
TTGGTTTTTTTTTTTATTTTGTTAATTATTTTTTTTTTTGTTTTTTAGGGTGAAAAATTTTAGTTGTA
TTTAAATTGATTGTTGACGTTTTTTGATTTAGAGATGATTTTTAGAAAAATTATAATAGTGAGTTGTG
TTTATATAAATAAGAGGATTGTGATTTTGTATTGAAGTTAATTTTAAATTTAGATTTTTGTAAAATGA

AAATATGAGTTG*GTTATTAGTTGAT*CG*TAATG*TATTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTTT
TTTTTTGTCGTTTAGGTTGGAGTGTAGTGGTATAATTTTTGTTTTTCGGGTTTAAGTGATTTTTTCGT
TTTAGTTTTTTGAGTAGTTGGGATTATAGGCGTTTCGTTTTTATGTTTAGTTAATTTTTGTATTTTTA
GTAAAGACGGGGTTTTATTACGTTGGTTAGGTTGGTTTTTAATTTTTGATTTTAGGTGATTTGTTTGT
TTCGGTTTTTTAAGGTGTTGGGATTATAGGCGTGAGTTATTACGTTTAGTTTTAAGGTATATTTTATA
TTTATTTTTGAGGATGAAGGAGAAAGGTTAATTTAGTTATTGATTATTATTGAGTTTTTTGGTTGTAA
TATTTTGTTTGTAGTTATTATTAATAGGTAATTTTATATTTGATTGATTTGTAGGTTTTTTAAATAGG
TTTTTATTTTTTAACGATAGAGTTTATTTTTATTTTGTTTCGATTAGGAAAAGTAATTATATTTATTT
TTTATTTTTTATATTATATAAATAGTTAAAGATTTTCGGTTGTAGTATGTTTTTGTTTTAAAAATAGA
AAATGTTTTAGTATTTTTGGTTTATTTTTTATATTGTATTTATTAGTTTTGATTTAAGTGTTTTTGAA
TAGAGTTTAATGTTGAAGATATTTTAAGTTAGTATATAGGAATTTTTGTGAAAGTTTGAAGGGAAGGA
AAAA



cg07504154-BS-F	GATTTTTGTAAAATGAAAATATGAGTTG
cg07504154-BS-R-bio	TCCCAACTACTCAAAAAACTAAAAC
cg07504154-BS-SP	GATTTTTGTAAAATGAAAATATGAGTTG
Sequencing entry	*GTTATTAGTTGATC/GTAATG*

11. cg22976313: DNMTΔ3B4

>hg19_dna range=chr14:105069615-105071114 5'pad=0 3'pad=0 strand=- repeatMasking=none

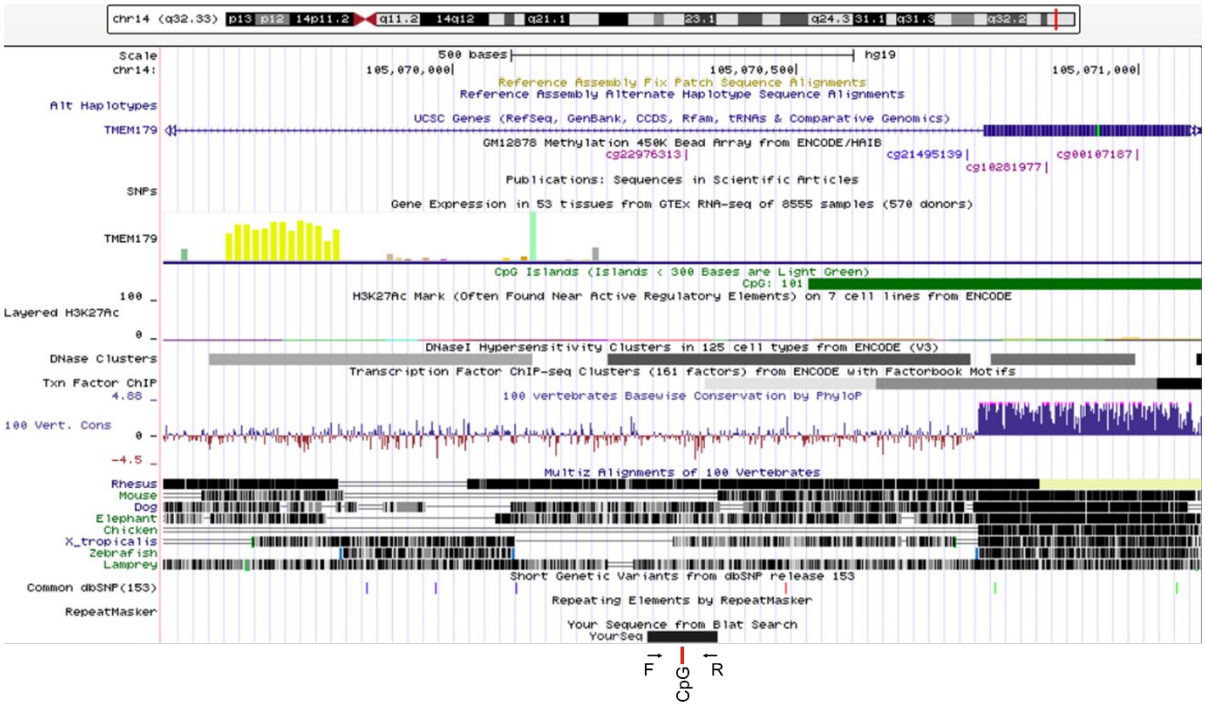
Genomic sequence:

GCGGCCTTCCCGCCGCCGCTCTCGCCCGGGCCGGCCATGGCGCTCAACAATTTCCTTTTCGCTCAGTG
CGCCTGCTACTTCTTGGCCTTCCTGTTCAGCTTCGTGGTGGTGGTCCCGCTGTCCGAGAACGGCCACG
ACTTCCGCGGCCGCTGCCTGCTCTTCACCGAGGGCATGTGGCTGAGCGCCAACCTCACGGTGCAGGAG
CGCGAGCGCTTCACGGTGCAGGAGTGGGGCCCGCCGGCCGCCTGCCGCTTCAGCCTGCTCGCCAGCCT
CCTGTCTCTGCTGCTGGCCGCCGCGCACGCCTGGCGCACGCTCTTCTTCCTCTGCAAGGGACACGAGG
GGTAAGTGGGGGCCGCTCCCGGCGCGGCGCCCCTTCCATGCCCCCAGCGCCCCCAGACCCACCTCCCC
GGGGAGGCGCGCGAGTCCCTCCGTCTCTCAGGCAGCTCCGAGGCCCGCAAATGTCAACAATCCGCCCT

CACCCCCTCCACTCAGACCGCCCTTTGTCTACACACACATCCAGGGACCATGACGACCACCGGTCCTG
CGTACAAGGCCAGTGCCTCCCCTCCCCAGCTGCGCATCACCTCCCGCTTCCGGGCATCTGCTCCCTTG
GGTTGCAGATGTTGCACTGTCATCACCCAGCAGGCCTGGACCTTCCTGGGCGGGCTCAGCCGGGCCAC
ACTCGCCAGGACCGGCGACAACTCCCAGCCCGGTCTAGGGCTTGCACTGAGCTCTCCAGGTCCCAGTG
ACAAGGCGAACTGGGGAGCCCCACCCGGCACCATCGCCCCGGGAAAGGGACGGCAGATGGCTGTCTGG
AGGGGAGCCACCCAGGCCCACCTCGAGCTCAGGCCTTGGATGGGGAGCTGGAGAGAGACTGCAGGGCG
AGGGCTGTGCCTTGTGGGCTCAGCGCCTTCCTCCTCCTGGCCTGCATTTTCCTCTTTCTCTTTCCTGC
TCCGGCTAGATACTGGAGCCCATACAACCTGGTTCACCTGGGGTGCGGTGACACCCTTAGCCAGTGGC
ACCCTGGGGGCCCTCTGGGGGCCCACCCACTCCCTCTGCAGGTGCTCTCTGTGAACCCTCCCACTCCC
TCTGTGAGCCCACCCACTCCCTCTGTGAACCCACCCTCTCCCTCTGTGGGTGCCCTCTGTGAGCCCAT
CTACCACCTCTGTGAATCCACCCACCCCCTCTGTGAACCCACCCACCCCCTCTGTGAGCCCACCCACC
CCCTCTGTGAGCCCACCCACTCCCTCTGTGAACCCACCCACTCCCTCTGTGGGTGCCCTCTGTGGGCC
CTCCCACTCCCCTGGCTGAGGCCTGCTGTGCCCAGGAGCAGCCCGAGATCCTGTAGAGGTTCCGGGAG
AGCACTCATGAAACCCAGGCTCCTCCCAAACTCTGCCCTCACCCCCATCCTTTGTTGTAGGTTTTCAC
TTCCTGCACGCAGGTGGGAGATGAGAAGGGCTCAGAGGGGTGTGGGCGGCAGGAGTGGGGAAGCCCTC
CTGT

Bisulfite sequence:

GCGGTTTTTTCGTCGTCGTTTTCGTTCGGGTCGGTTATGGCGTTTAATAATTTTTTTTTCGTTTAGTG
CGTTTGTTATTTTTTGGTTTTTTTGTTTAGTTTCGTGGTGGTGGTTTCGTTGTTCGAGAACGGTTACG
ATTTTCGCGGTCGTTGTTTGTTTTTTATCGAGGGTATGTGGTTGAGCGTTAATTTTACGGTGTAGGAG
CGCGAGCGTTTTACGGTGTAGGAGTGGGGTTCGTCGGTCGTTTGTCGTTTTAGTTTGTTCGTTAGTTT
TTTGTTTTTGTTGTTGGTCGTCGCGTACGTTTGGCGTACGTTTTTTTTTTTTTGTAAGGGATACGAGG
GGTAAGTGGGGGTCGTTTTCGGCGCGGCGTTTTTTTTATGTTTTTAGCGTTTTTAGATTTATTTTTTC
GGGGAGGCGCGCGAGTTTTTTCGTTTTTTAGGTAGTTTCGAGGTTCGTAAATGTTAATAATTCGTTTT
TATTTTTTTTATTTAGATCGTTTTTTGTTTATATATATATTTAGGGATTATGACGATTATCGGTTTTG
CGTATAAGGTTAGTGTTTTTTTTTTTTAGTTGCGTATTATTTTTCGTTTTCGGGTATTTGTTTTTTTG
GGTTGTAGATGTTGTATTGTTATTATTTAGTAGGTTTGGATTTTTTTGGGCGGGTTTAGTCGGGTTAT
ATTCGTTAGGATCGGCGATAATTTTTAGTTCGGTTTAGGGTTTGTATTGAGTTTTTTAGGTTTTAGTG
ATAAGGCGAATTGGGGAGTTTTATTCGGTATTATCGTTTCGGGAAAGGGACGGTAGATGGTTGTTTGG
AGGGGAGTTATTTAGGTTTATTTCGAGTTTAGGTTTTGGATGGGGAGTTGGAGAGAGATTGTAGGGCG
AGGGTTGTGTTTTGTGGGTTTAGCGTTTTTTTTTTTTTGGTTTGTATTTTTTTTTTTTTTTTTTTGT
TTCGGTTAGATATTGGAGTTTATATAATTTGGTTTATTTGGGGTGCGGTGATATTTTTAGTTAGTGGT
ATTTTGGGGGTTTTTTGGGGGTTTATTTATTTTTTTTGTAGGTGTTTTTTGTGAATTTTTTTATTTTT
TTTGTGAGTTTATTTATTTTTTTTGTGAATTTATTTTTTTTTTTTGTGGGTGTTTTTTGTGAGTTTAT
TTATTATTTTTGTGAATTTATTTATTTTTTTTGTGAATTTATTTATTTTTTTTGTGAGTTTATTTATT
TTTTTTGTGAGTTTATTTATTTTTTTTGTGAATTTATTTATTTTTTTTGTGGGTGTTTTTTGTGGGTT
TTTTTATTTTTTTGGTTGAGGTTTGTTGTGTTTAGGAGTAGTTCGAGATTTTGTAGAGGTTTCGGGAG
AGTATTTATGAAATTTAGGTTTTTTTTAAATTTTGTTTTTATTTTTATTTTTTGTTGTAGGTTTTTAT
TTTTTGTACGTAGGTGGGAGATGAGAAGGGTTTAGAGGGGTGTGGGCGGTAGGAGTGGGGAAGTTTTT
TTGT



cg22976313-BS-F	AGTTTTTTAGGTTTTAGTGATAAGG
cg22976313-BS-R-bio	TAAATAACTCCCCTCCAAACAAC
cg22976313-BS-SP	AGTTTTTTAGGTTTTAGTGATAAGG
Sequencing entry	*C/GAATTGGGGAGTTTTATTC/GGTATTATC/GTT*

12. cg12150401: DNMT3L

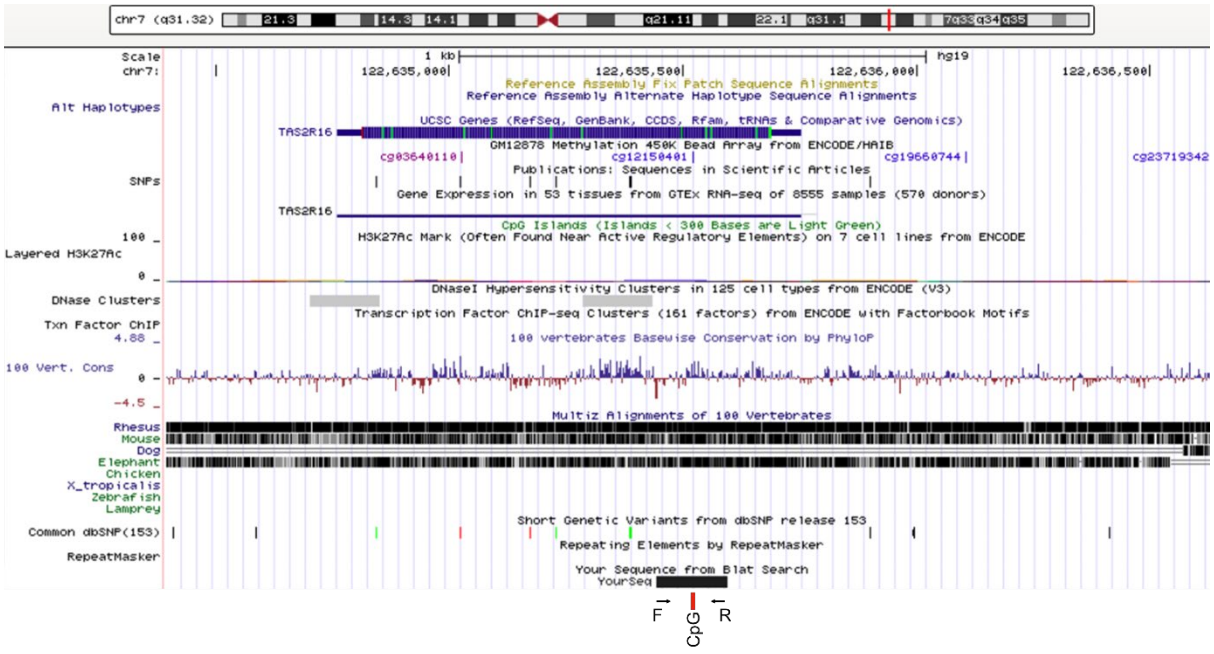
Genomic sequence:

GAAATACTTGAAGGTCTGGGGTTCAGCAATTTTGTGGAATTTCGGAGGAAAGACTTTTCATT
ATCATTTTTATTTACCCTCACACTTTGGTTGTAGCATCCACCACTAGCTGCCATGAGAACTG
GGAAGATAGGGCTTCTGAGGTTCGCTTTGGTACTGCACTCAACAGAGCAGAAACATCATGCA
GGACGTGAATGAACAGGGAAGGATTCATCTTTCTTGTTGCTCAAATTTTAGGTTAAAAGAGT
GAAGATGTCATTTAAAGAAGATAAGGAAATAAGCTAAGTGCACCTCAGAAAAAGAGGAAATA
AACAACATGGTTATGAAATACATATGCCTTTTTGGAATAATCAAATCACCTCAGGAGCAGGA
ATCCACATTTAAACATTTCCCTACATTCAGATGTGATAGCAAACCAGCAAGCAAAGCATAAA
CAAATTCTATAGATAGGAAAATGACAAAATCAAAAATGAGCTTGGGAAAATTTGCTTCCAGA
GAGAGGGGTTTTCTTGTACCAGACAGTGGTGGAATCAGGAACCCATTCTCTACCTGGAAATT
GGTACAGAGATTCAGAGTCTTTGTCCAGGAAGACACTTTGGAGTAGAAGAATGATACCCATC
CAACTCACTGTCTTCTTCATGATCATCTATGTGCTTGAGTCCTTGACAATTATTGTGCAGAG
CAGCCTAATTGTTGCAGTGCTGGGCAGAGAATGGCTGCAAGTCAGAAGGCTGATGCCTGTGG
ACATGATTCTCATCAGCCTGGGCATCTCTCGCTTCTGTCTACAGTGGGCATCAATGCTGAAC
AATTTTTGCTCCTATTTTAATTTGAATTATGTACTTTGCAACTTAACAATCACCTGGGAATT
TTTTAATATCCTTACATTCTGGTTAAACAGCTTGCTTACCGTGTTCTACTGCATCAAGGTCT
CTTCTTTCACCCATCACATCTTTCTCTGGCTGAGGTGGAGAATTTTGAGGTTGTTTCCCTGG
ATATTACTGGGTTCTCTGATGATTACTTGTGTAACAATCATCCCTTCAGCTATTGGGAATTA
CATTCAAATTCAGTTACTCACCATGGAGCATCTACCAAGAAACAGCACTGTAACTGACAAAC
TTGAAAATTTTCATCAGTATCAGTTCCAGGCTCATACAGTTGCATTGGTTATTCCTTTCATC
CTGTTCCTGGCCTCCACCATCTTTCTCATGGCATCACTGACCAAGCAGATACAACATCATAG
CACTGGTCACTGCAATCCAAGCATGAAAGCGCGCTTCACTGCCCTGAGGTCCCTTGCCGTCT
TATTTATTGTGTTTACCTCTTACTTTCTAACCATACTCATCACCATTATAGGTACTCTATTT

GATAAGAGATGTTGGTTATGGGTCTGGGAAGCTTTTGTCTATGCTTTCATCTTAATGCATTC
CACTTCACTGATGCTGAGCAGCCCTACGTTGAAAAGGATTCTAAAGGGAAAGTGCTAGGCCT
AGAGGTTGCCTG

Bisulfite sequence:

GAAATATTTGAAGGTTTGGGGTTTAGTAATTTTGTGGAATTTCGGAGGAAAGATTTTTTATTATTATT
TTTATTTATTTTTATATTTTGGTTGTAGTATTTATTATTAGTTGTTATGAGAATTGGGAAGATAGGGT
TTTTGAGGTTCGTTTTGGTATTGTATTTAATAGAGTAGAAATATTATGTAGGACGTGAATGAATAGGG
AAGGATTTATTTTTTTTGTTGTTTAAATTTTAGGTTAAAAGAGTGAAGATGTTATTTAAAGAAGATAA
GGAAATAAGTTAAGTGTATTTTAGAAAAAGAGGAAATAAATAATATGGTTATGAAATATATATGTTTT
TTTGGAATAATTAAATTATTTTAGGAGTAGGAATTTATATTTAAATATTTTTTTATATTTAGATGTGA
TAGTAAATTAGTAAGTAAAGTATAAATAAATTTTATAGATAGGAAAATGATAAAATTAAAAATGAGTT
TGGGAAAATTTGTTTTTAGAGAGAGGGGTTTTTTTGTATTAGATAGTGGTGGAATTAGGAATTTATTT
TTTATTTGGAAATTGGTATAGAGATTTAGAGTTTTTGTTTAGGAAGATATTTTGGAGTAGAAGAATGA
TATTTATTTAATTTATTGTTTTTTTTATGATTATTTATGTGTTTGAGTTTTTGATAATTATTGTGTAG
AGTAGTTTAATTGTTGTAGTGTTGGGTAGAGAATGGTTGTAAGTTAGAAGGTTGATGTTTGTGGATAT
GATTTTTATTAGTTTGGGTATTTTTCGTTTTTGTTTATAGTGGGTATTAATGTTGAATAATTTTTGTT
TTTATTTTAATTTGAATTATGTATTTTGTAATTTAATAATTATTTGGGAATTTTTTAATATTTTTATA
TTTTGGTTAAATAGTTTGTTTATCGTGTTTTATTGTATTAAGGTTTTTTTTTTTATTTATTATATTTT
TTTTTGGTTGAGGTGGAGAATTTTGAGGTTGTTTTTTTGGATATTATTGGGTTTTTTGATGATTATTT
GTGTAATAATTATTTTTTTAGTTATTGGGAATTATATTTAAATTTAGTTATTTATTATGGAGTATTTA
TTAAGAAATAGTATTGTAATTGATAAATTTGAAAATTTTTATTAGTATTAGTTTTAGGTTTATATAGT
TGTATTGGTTATTTTTTTTTATTTTGTTTTTGGTTTTTATTATTTTTTTTTATGGTATTATTGATTAAGT
AGATATAATATTATAGTATTGGTTATTGTAATTTAAGTATGAAAGCGCGTTTTATTGTTTTGAGGTTT
TTTGTCGTTTTATTTATTGTGTTTATTTTTTATTTTTTAATTATATTTATTATTATTATAGGTATTTT
ATTTGATAAGAGATGTTGGTTATGGGTTTGGGAAGTTTTTGTTTATGTTTTTATTTTAATGTATTTTA
TTTTATTGATGTTGAGTAGTTTTACGTTGAAAAGGATTTTAAAGGGAAAGTGTTAGGTTTAGAGGTTG
TTTG



cg12150401-BS-F	TGGGTAGAGAATGGTTGTAAG
cg12150401-BS-R-bio	CCCAAATAATTATTAAATTACAAAAT
cg12150401-BS-SP	TTATTAGTTTGGGTATTT
Sequencing entry	*TTC/GTTTTTGTTTAT*

13. cg20540357: DNMT3L

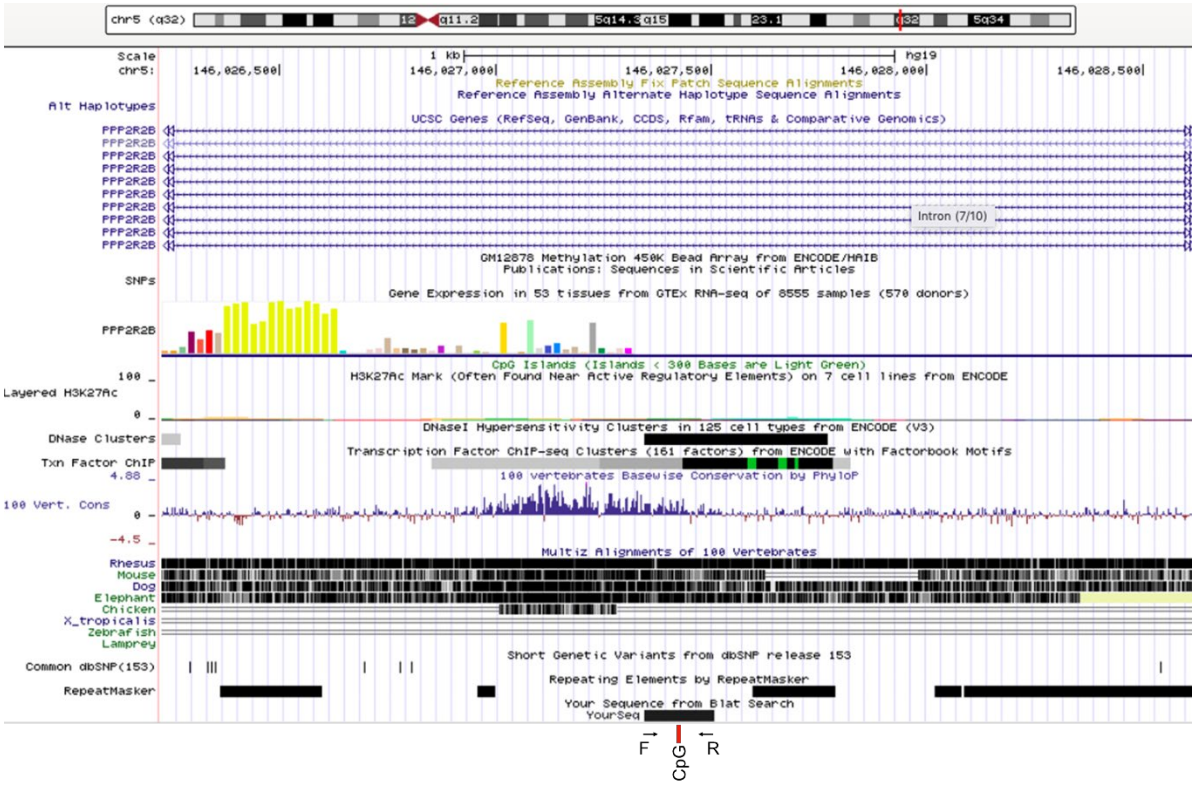
>hg19_dna range=chr5:146026941-146027940 5'pad=0 3'pad=0 strand=- repeatMasking=none

Genomic sequence:

AATGGCTCAAACAAATAAAAGATTATTTTCGTATGAAGTCTAATATTGATGATCCCCATCAGCAGGCA
GCCTCCTCCAATCAGTAATTCAGAAACTCGGGCCTCTTTCATTTGCAGCTCCCCAATGTTACCATGTT
TATAGCAGGACTGTAAAAGGGGAAGGTCATGGAAGATCACAAGTGGGAGATTCCTCTGAGCCATGCCC
TAAAAAAGTATGCATTCCATTGGCTAGAATTCAGCCACATGGCTAACTCTTAACTGCAAGGCAGCCTA
GGAAATGTAGTCTGGCTGTGTGCCCAGGATAAAAAGACCCGGTTTTGGTACACAGCTAGCTAGTCTCC
ACCATGGAGAGCACTCAACCGTCTCCAAGGGCTTTCAGGGAAGGCAGCTTTAATCAGATAGTGGCAAT
TATGAGATATATTTGGAGACAATGAGAACTGACCATCAGGACCATGCCTGGCA*TATTTTGTTGATCAC*
*GGAA*CAGACTCCACATGCTGTCTGGGCCTGCCTACAGATTTTGCCTGGCAGTGCAGCTGCTAATAGAA
TGAGTGTGGGTGGTGGTTCCCTTTTTCTCTCTCCATCTATTCCAAGAGCCAGAAATCTGCTCAGGCTC
AGCTGGCTGTCACTCTGAGCTTCTGCTTCCCCAGTCTAGGTGACATGTTTTCTGAATTTTGCCTCTAG
GCCTTCAAGGGGGAGGGGAGTTTCTTTAGGCATTTTGGGAGACAGATGCCTTGGGGAGAATGCAAAGG
GGAAATATTAACAGCGCAGTTTAGGTTATATTCTGCAGATGTGTGACAGACTGCCAGATGCTGGCATA
CAAATGAGACCGTGCCATTCACCCTTAAAGCAGCCATATGTAAAATAATCATTGTCATCATGAGAACG
ATGATAAGTGATATCTATATAGTTCTTTTCAACCAGACCGCCCTAAGCACTGAGCCTCTCATCACTGA
AGTATAGCCACCTCTGGGGTAGGAAGCGGCCATCTTTCAAAACAGTCT

Bisulfite sequence:

AATGGTTTAAATAAATAAAAGATTATTTTCGTATGAAGTTTAATATTGATGATTTTTATTAGTAGGTA
GTTTTTTTTAATTAGTAATTTAGAAATTCGGGTTTTTTTTATTTGTAGTTTTTTAATGTTATTATGTT
TATAGTAGGATTGTAAAAGGGGAAGGTTATGGAAGATTATAAGTGGGAGATTTTTTTGAGTTATGTTT
TAAAAAAGTATGTATTTTATTGGTTAGAATTTAGTTATATGGTTAATTTTTAATTGTAAGGTAGTTTA
GGAAATGTAGTTTGGTTGTGTGTTTAGGATAAAAAGATTCGGTTTTGGTATATAGTTAGTTAGTTTTT
ATTATGGAGAGTATTTAATCGTTTTTAAGGGTTTTTAGGGAAGGTAGTTTTAATTAGATAGTGGTAAT
TATGAGATATATTTGGAGATAATGAGA*ATTGATTATTAGGATTATGTTTGG**TATATTTTGTTGATTAC*
*GGAA*TAGATTTTATATGTTGTTTGGGTTTGTTTATAGATTTTGTTTGGTAGTGTAGTTGTTAATAGAA
TGAGTGTGGGTGGTGGTTTTTTTTTTTTTTTTTTATTTATTTTAAGAGTTAGAAATTTGTTTAGGTTT
AGTTGGTTGTTATTTTGAGTTTTTGTTTTTTTAGTTTAGGTGATATGTTTTTTGAATTTTGTTTTTAG
GTTTTTAAGGGGGAGGGGAGTTTTTTTAGGTATTTTGGGAGATAGATGTTTTGGGGAGAATGTAAAGG
GGAAATATTAATAGCGTAGTTTAGGTTATATTTTGTAGATGTGTGATAGATTGTTAGATGTTGGTATA
TAAATGAGATCGTGTTATTTATTTTTAAAGTAGTTATATGTAAAATAATTATTGTTATTATGAGAACG
ATGATAAGTGATATTTATATAGTTTTTTTTAATTAGATCGTTTTAAGTATTGAGTTTTTTATTATTGA
AGTATAGTTATTTTTGGGGTAGGAAGCGGTTATTTTTTAAAATAGTTT



cg20540357-BS-F	ATTGATTATTAGGATTATGTTTGG
cg20540357-BS-R-bio	aaaccaccacccacactcat
cg20540357-BS-SP	ATTGATTATTAGGATTATGTTTGG
Sequencing entry	*TATATTTTGTTGATTAC/GGAA*

14. cg04458645: DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1

>hg19_dna range=chr11:43588792-43590291 5'pad=0 3'pad=0 strand=+ repeatMasking=none

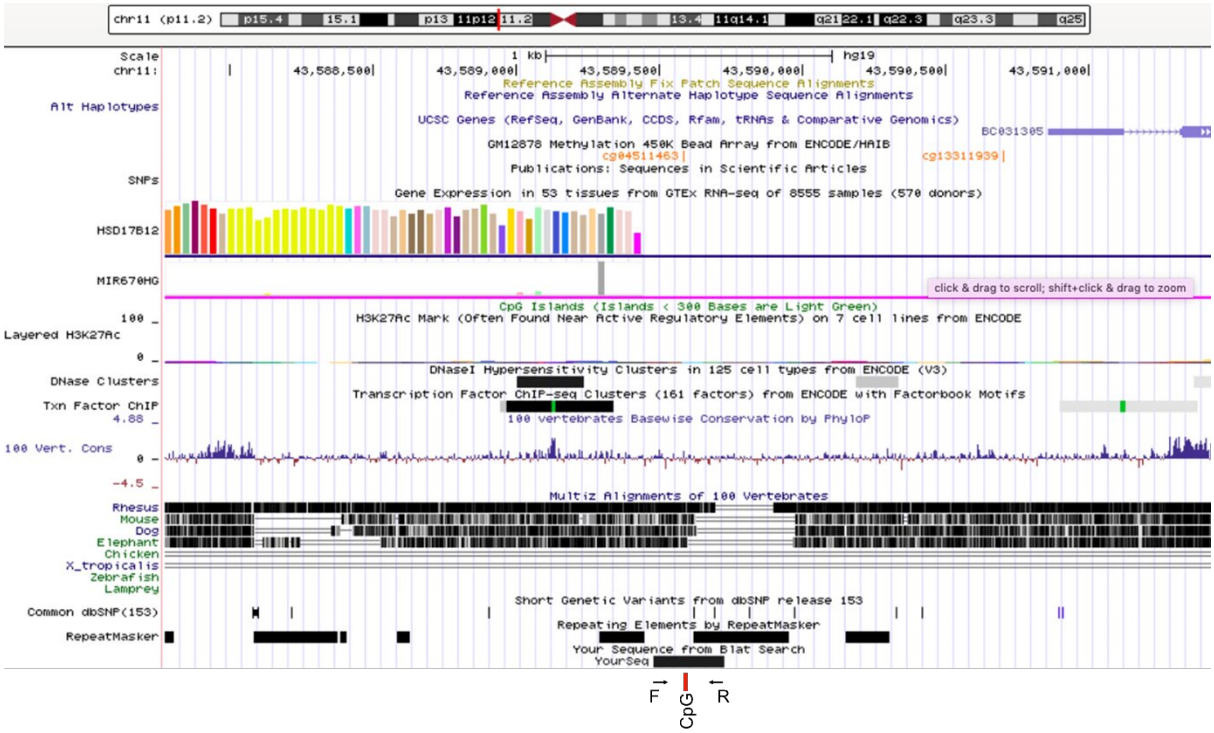
Genomic sequence:

GAACACATTCAAAATCCAAGGGCCAGAAAAATTTTAAGCTCCAGCCTCAGCCTCCGATCTCCTGGGAC
CAGCCCCATAAGAACCATCCATTCTAGCCCAACTTGCAAATGGAGCCCAAGAGTTGAACCTCGGCATG
GGACTCAAGATGTCTCTAAAGTTCAAAAACACCTGGATTACCTGAAAGCAGAACCGGGTCCAGGTCCT
GTGCTGAGGTCTCCAGTTTTCTCTAAGTAAAATATTTCTAGTGGAGAAAAAGCCTGTGGCCTCTGGGC
AGAGATTTGTTCTATAGAGAAATGAAGGAAATGTGCAAGAGAGAACAGTTTTACTCTGGAGCTCCCTC
TACTGGCTCTTTGATAGAAATCCATCGCAGTTATTCCGCCCTGCCTTGTGGGTATCAAGTTAAAGCAT
CCGAGCATCCTGGCACCTTGAAGCAAGGAGAAAAGGTGTCCTGAGGTGCCAAGGAAGATGAAACTGCA
GATACTGGTTTCTTCTGCCTCTTTTACAATTTCATCTAATGTTTTGGGGGTGCCAAGATAGTGCTAAG
CATTTTGTGTGCATTATTTCATTTAATACTCACCATAAACCAGCATTTTACAAATGAGGCAGCTGAGA
CCTGAAGGGATTAAATAAGTTCCTGGCATCACCATAGACTAATAGACCCCCTTAGTGACCCTCATAAT
TCTTATGATTCTCCTGAGTGACAACTCAAGAGAAGCAAGAGTCAACGGACAATGTGAGGTTGGAAGTG
GCACTTGATCTCATCACCTAATATAAAGAGGCTCTGAGATTGCGCCATAAACTTCCTAGGTCCTATAG
AATCTCAGCCCTGTTTGTTTTTTTATTCTTTTAATTTTATTTATTTATTTATTTAAGACGAACT
CTCACTCTGTTGCCCGGGACGGAGTGTGCAGTGGTGTGATCTTGGCTCACTGCAACCTCCACCTCCTG
GTTTCAAGCAATCTTCCCACCTCAGCCTCCCGAGTAGCTAGAATTACAGGCATGCACCATCACACCTC
GTTAATTTTTTCTGTATTTTTAGTAGAGAAAGGGTTTCACCATGTTGGCCAGGCTGGTCTTGATCTTC
TGACCTCAAGTGATCCACCCGCCTCAGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACTGCACC

CAGCCTCCTGTTTGTTTGTTTTTTTTTTAATTCAGCCCTCGAGGCAATTAATTACATGGCAAGGCTTT
GTCCAATCTGCACTCTCACTGTGATGCCCAAGTTGAAGAAAGCAGGGAACTACCACAGTGACTAAGAA
GAGTTTCCATGACAGAAGTGAGGCCTTTTAAAATTTGTGGAACATGTTCACGCCTACTTTTAAAATAA
AATTCCATGAGATAGGTAGGGCAGATAGTGCCCACACATTTTATAGGTTGGGAACTTAAGATACTGAG
TGCCTGTTTCCAGGTTTTGTTTAGGGGCAAATGGCAGAATCAAGACCAGAACTGTCACTCTCAGCTTA
AGTT

Bisulfite sequence:

GAATATATTTAAAATTTAAGGGTTAGAAAAATTTTAAGTTTTAGTTTTAGTTTTCGATTTTTTGGGAT
TAGTTTTATAAGAATTATTTATTTTAGTTTAATTTGTAAATGGAGTTTAAGAGTTGAATTTCGGTATG
GGATTTAAGATGTTTTTAAAGTTTAAAAATATTTGGATTATTTGAAAGTAGAATCGGGTTTAGGTTTT
GTGTTGAGGTTTTTAGTTTTTTTTAAGTAAAATATTTTTAGTGGAGAAAAAGTTTGTGGTTTTTGGGT
AGAGATTTGTTTTATAGAGAAATGAAGGAAATGTGTAAGAGAGAATAGTTTTATTTTGGAGTTTTTTT
TATTGGTTTTTTGATAGAAATTTATCGTAGTTATTTCGTTTTGTTTTGTGGGTATTAAGTTAAAGTAT
TCGAGTATTTTGGTATTTTGAAGTAAGGAGAAAAGGTGTTTTGAGGTGTTAAGGAAGATGAAATTGTA
GATATTGGTTTTTTTTGTTTTTTTTATAATTTTATTTAATGTTTTGGGGGTGTTAAGATAGTGTTAAG
TATTTTGTGTGTATTATTTTATTTAATATTTATTATAAATTAGTATTTTATAAATGAGGTAGTTGAGA
TTTGAAGGGATTAAATAAGTTTTTGGTATTATTATAGATTAATAGATTTTTTTAGTGATTTTTATAAT
TTTTATGATTTTTTTGAGTGATAATTTAAGAGAAGTAAGAGTTAACGGATAATGTGAGGTTGGAAGTG
GTATTTGATTTTATTATTTAATATAAAGAGGTTTTGAGATTGCGTTATAAATTTTTTAGGTTTTATAG
AATTTTAGTTTTGTTTGTTTTTTTATTTTTTTAATTTTATTTATTTATTTATTTATTTAAGACGAATT
TTTATTTTGTTGTTCGGGACGGAGTGTGTAGTGGTGTGATTTTGGTTTATTGTAATTTTTATTTTTTG
GTTTTAAGTAATTTTTTTATTTTAGTTTTTCGAGTAGTTAGAATTATAGGTATGTATTATTATATTTC
GTTAATTTTTTTTGTATTTTTAGTAGAGAAAGGGTTTTATTATGTTGGTTAGGTTGGTTTTGATTTTT
TGATTTTAAGTGATTTATTCGTTTTAGTTTTTTAAAGTGTTAGGATTATAGGCGTGAGTTATTGTATT
TAGTTTTTTGTTTGTTTGTTTTTTTTTTAATTTAGTTTTCGAGGTAATTAATTATATGGTAAGGTTTT
GTTTAATTTGTATTTTTATTGTGATGTTTAAGTTGAAGAAAGTAGGGAATTATTATAGTGATTAAGAA
GAGTTTTTATGATAGAAGTGAGGTTTTTTAAAATTTGTGGAATATGTTTACGTTTATTTTTAAAATAA
AATTTTATGAGATAGGTAGGGTAGATAGTGTTTATATATTTTATAGGTTGGGAATTTAAGATATTGAG
TGTTTGTTTTTAGGTTTTGTTTAGGGGTAAATGGTAGAATTAAGATTAGAATTGTTATTTTTAGTTTA
AGTT



cg04458645-BS-F	TTTTTGAGTGATAATTTAAGAGAAGTAAGA
cg04458645-BS-R-bio	ATAAACCAAAATCACACCACTACAC
cg04458645-BS-SP	TTTTTGAGTGATAATTTAAGAGAAGTAAGA
Sequencing entry	GTTAAC/TGGATAAT

15. cg25843713: DNMT3A2, DNMT3B4, DNMTΔ3B2, DNMTΔ3B3, DNMTΔ3B4, DNMT1

>hg19_dna range=chr18:57347696-57349195 5'pad=0 3'pad=0 strand=- repeatMasking=none

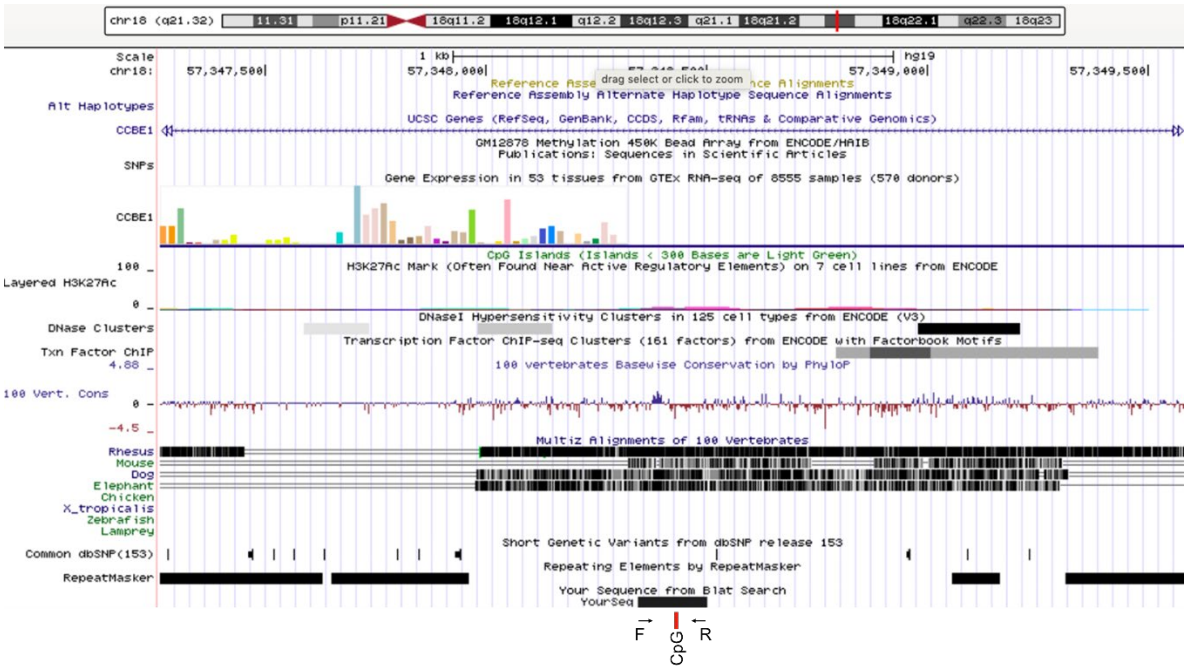
Genomic sequence:

TCCAGGTCTGACCGTCTGACAGGGACCACTCCCACCTTCACAGAGAGCAGAGTAAACAGAAGCGCATT
TAGTTTACAGCCCAGGGGATTTGCTTAACACAAGGGTGATCATCTCGACTGAGGGAGACTCTGCAGTA
CTTATCTCTAATCAGCCTTACAAACATGATAGCCATCTTTGCACTGTGGATTTAATCTCTTGAATGGC
TTTAATGGTTTCTTCCCCAATTATTAAAAAGGAAAAAGTCTTTTTACTCCCCAGGTCTATGCTTTAAA
GTGGGCTGCTATCTGCTGCAAGCTCATGCTGGGTAAGTGCAGGAGCCCATCAAAGTGTTAAGGGCCCT
GGCTGGCCACTGGCTGGCTCGCATTTCACTGCTTGCTTTGAACATGTGCTGTGCCATAGGATAGCAGA
ACACAGGGTGAAGTATCAGCTTCCTCGCTTCATGGAGAGCTGTAGTCCCACGGCACCCACTCCCTTCC
TAAAAAGGAGAAATTGAGTTGCCTTTGGGCAGACTCCTGCTAATAAAGTGGGAAGAAAAGGGGACATC
AGATCAGCTTTCTCATAAAATTTCTTACACATCTTTTATCTTAACATACTGGACTTAGTATGTTTCTT
ATATTTACTAAGGTTTCTATTACAGACTTATTTTTCAACACGTGTCTCCTTCTGAATGTACTTATTTT
ATGGACCAAAGAGCCTCTGCTCAGTGTCTCTAAGGGCCCTCGTAAAGACTTCATGTGCAAATTTGCTG
GCTGTGGCCTTGCTGCCATTTTCATCGTTCTGCTGCTCCGGGTTCTGGTTCCATATTTGGCTCTGTTT
AAAGATTCCGAGGGAAATGGTGACACAAGACAGCTATGCAACTTAGCTATATATGTTTAAAAAGATAC
CAAACAGCTTTCCTTTTTTTTTTTTTTTTTCTCAGCCAAATTACCTTCCTAAAAGGATGAGGCTTCTC
TTGAATAACCCAGATAATTATCAGGTTTTTGTATTTATTAACATCCTTTCCATATTTAGGCAGTCTCA
GCTTAATAAAAAGGTATCACTTGAGTTTAGCATGAGATCAACAGGCCGCTGGTTTGAAGTTTGATTCC
TTCTTTCTCCTGCTTGCAAGTTAATCTGTCTAATCTGCATTTTAACCTACGCTCCTTTGGAAGGAATG

TAGTTAAGGCTGGGTTTGTAACCTGATGGATGAAAGGGAATCTCCACTGCAGCTGTTGTTTCTTCATT
TTTGTTTTGTTTTGTTTTGTTTTGTTTTTTCTTTTTTTTTTGAGATGGAGTCTCGCTCTGTCTCCCAG
GCTGGAGTGCAGTGGCAGCAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCACGCCATTCTCCT
GCCTCAGCCTCCCCAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCCGGCTAATTTTTTGTGTTTT
TAGTAGAGACAGGGTTTCACCATGTTAGCCAGCATGGTCTTGATCTCCTGACCTCGTGATCCTCCCGC
CTCA

Bisulfite sequence:

TTTAGGTTTGATCGTTTGATAGGGATTATTTTTATTTTTATAGAGAGTAGAGTAAATAGAAGCGTATT
TAGTTTATAGTTTAGGGGATTTGTTTAATATAAGGGTGATTATTTCGATTGAGGGAGATTTTGTAGTA
TTTATTTTTAATTAGTTTTATAAATATGATAGTTATTTTTGTATTGTGGATTTAATTTTTTGAATGGT
TTTAATGGTTTTTTTTTTAATTATTAAAAAGGAAAAAGTTTTTTTATTTTTTAGGTTTATGTTTTAAA
GTGGGTTGTTATTTGTTGTAAGTTTATGTTGGGTAAGTGTAGGAGTTTATTAAAGTGTTAAGGGTTTT
GGTTGGTTATTGGTTGGTTCGTATTTTATTGTTTGTTTTGAATATGTGTTGTGTTATAGGATAGTAGA
ATATAGGGTGAAGTATTAGTTTTTTCGTTTTATGGAGAGTTGTAGTTTTACGGTATTTATTTTTTTTT
TAAAAAGGAGAAATTGAGTTGTTTTTGGGTAGATTTTTGTTAATAAAGTGGGAAGAAAAGGGGATATT
AGATTAGTTTTTTTATAAAATTTTTTATATATTTTTTATTTTAATATATTGGATTTAGTATGTTTTTT
ATATTTATTAAGGTTTTTATTATAGATTTATTTTTTAATACGTGTTTTTTTTTGAATGTATTTATTTT
ATGGATTAAAGAGTTTTTGTTTAGTGTTTTTAAGGGTTTTCGTAAAGATTTTATGTGTAAATTTGTTG
GTTGTGGTTTTGTTGTTATTTTTATCGTTTTGTTGTTTCGGGTTTTGGTTTTATATTTGGTTTTGTTT
AAAGATTTCGAGGGAAATGGTGATATAAGATAGTTATGTAATTTAGTTATATATGTTTAAAAAGATAT
TAAATAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGTTAAATTATTTTTTTAAAAGGATGAGGTTTTTT
TTGAATAATTTAGATAATTATTAGGTTTTTGTATTTATTAATATTTTTTTTATATTTAGGTAGTTTTA
GTTTAATAAAAAGGTATTATTTGAGTTTAGTATGAGATTAATAGGTCGTTGGTTTGAAGTTTGATTTT
TTTTTTTTTTTGTTTGTAAGTTAATTTGTTTAATTTGTATTTTAATTTACGTTTTTTTGGAAGGAATG
TAGTTAAGGTTGGGTTTGTAATTTGATGGATGAAAGGGAATTTTTATTGTAGTTGTTGTTTTTTTATT
TTTGTTTTGTTTTGTTTTGTTTTGTTTTTTTTTTTTTTTTTTGAGATGGAGTTTCGTTTTGTTTTTTAG
GTTGGAGTGTAGTGGTAGTAATTTCGGTTTATTGTAATTTTCGTTTTTCGGGTTTACGTTATTTTTTT
GTTTTAGTTTTTTTAGTAGTTGGGATTATAGGCGTTCGTTATTACGTTCGGTTAATTTTTTGTGTTTT
TAGTAGAGATAGGGTTTTATTATGTTAGTTAGTATGGTTTTGATTTTTTGATTTCGTGATTTTTTCGT
TTTA



cg25843713-BS-F	TTTGTTTAGTGTTTTTAAGGGTTTT
cg25843713-BS-R-bio	ACTATCTTATATCACCATTTCCCTC
cg25843713-BS-SP	GGTTTTGTTGTTATTTTTAT
Sequencing entry	C/TGTTTTGTTGTTTC/TGGGT