The role of age-related DNA methylation in the development of age-related diseases

A thesis submitted in part requirement for the degree of Doctor of Philosophy

July 2013

Sanne Dorien van Otterdijk

Northern Institute for Cancer Research,

Faculty of Medical Sciences,

Newcastle University

Abstract

Alterations in DNA methylation can have dramatic effects on gene transcription, and in particular, hypermethylation of promoter associated CpG islands is known to lead to gene inactivation. Altered patterns of DNA methylation play a key role in the development of cancer and may also play important roles in many other diseases. However, the mechanisms which lead to these changes in DNA methylation are unknown. DNA methylation patterns have also been found to change during normal ageing and these changes have similarities to those that occur during the development of cancer. This suggests that for some age-related diseases, most notably cancer, altered patterns of methylation may be an early initiating event and that disease may develop in cells which already possess changes in their DNA methylation landscape. Therefore, this study was designed to examine how methylation levels at a group of genes alters over the life-course and how these relate to methylation changes observed in major age-related diseases (cancer, specifically acute lymphoblastic leukaemia (ALL) and Hereditary Nonpolyposis Colorectal Cancer (HNPCC) patients, and atherosclerosis). DNA was collected from healthy volunteers from different ages and from ALL, HNPCC and atherosclerosis patients. Methylation was quantified using pyrosequencing. The study produced a number of findings: 1) Genes exhibiting variable methylation in PBL samples from healthy volunteers are also highly methylated in leukaemia, suggesting a common underlying mechanism. 2) Increased methylation levels were observed in lymphoid compared to myeloid cells, in healthy individuals, mirroring the patterns seen in leukaemia. 3) A subset of genes exhibiting variable methylation in PBL samples from healthy volunteers and that are highly methylated in leukaemia are aberrantly methylated in HNPCC patients and atherosclerosis patients, suggesting shared risk factors. 4) While methylation levels increase during ageing, a substantial proportion of methylation is already present at birth and may thus alter disease susceptibility throughout life. 5) Blood samples from ALL patients in remission exhibit increased methylation levels (versus controls), not directly related to their leukaemic clone, and the extent of methylation correlates with overall survival.

The studies to date are compatible with a hypothesis in which altered methylation of disease-related genes pre-exists in a subset of haematopoietic cells and that these cells may be at a significantly increased risk of progression to age-related diseases. Furthermore, monitoring DNA methylation may be a valuable tool for early diagnosis of these diseases, as well as for monitoring disease progression in patients.

Acknowledgements

I would like to thank my supervisors Gordon Strathdee, John Mathers and Ioakim Spyridopoulos for their support, guidance and inputs throughout the project, as well as my group members.

Thanks for the numerous members of staff within the University who have provided assistance during this project, under which my assessors, Caroline Relton and Jim Allen, everyone at the flow cytometry facility and the pyrosequencer who helped with technical guidance and the statisticians who were a great help for finding the correct statistical tests.

A special thanks to everyone from the old crucible lab for giving me such a warm welcome in Newcastle. It was a privilege to share a lab with all of you! Thanks for those in the SJS who friendly welcomed us in the new lab after the move and for keeping me company during the, sometimes so hard needed, work breaks.

Thanks to all of my friends who supported me through the past years by either a nice chat or a couple of pints, both here in Newcastle and overseas. Especially for those involved in Monday dinner and Friday drinks; the Newcastle experience wouldn't have been the same without the always good start and end of the week with your great company and sometimes some alcohol ☺. So everyone involved in these activities, but also all the friends that I made beside this, I will miss all of you when I leave Newcastle.

I will finish my acknowledgements with thanking the most important people, without whose support I would not have come this far: my family and Maurice; spreading your wings is easy, when you have someone there to catch you when you fall. So thanks, for always giving me that safety net, to support me with whichever crazy idea I came up next, and most importantly, for always giving me a home. Maurice, I know the past three years apart wasn't easy on both of us, but I couldn't have done it without knowing that you were always there for me. A new chapter of our lives will begin after this, and I'm looking forward to it with you by my side.

Table of contents

List of figures

List of tables

Abbreviations

5hmC	5-hydroxymethylcytosine
5hmU	5-hydroxyluracil
5mC	5-methylcytosine
AID	Activation-induced deaminase
ALL	Acute Lymphoblastic Leukaemia
AMI	Acute Myocardial Infarction
AML	Acute Myeloid Leukaemia
APOBEC	Apolipoprotein B mRNA-editing enzyme complex
APS	Adenosine phosphosulfate
ATP	Adenosine triphosphate
Azacytide	5-azacytosine
BER	Base excision repair
CAD	Coronary atherosclerotic disease
CAPP study	Concerted Action Polyp Prevention study
CCD	Charge coupled device
CE	Clonal evolution
C-DMR	Cancer-specific differentially methylated region
cDNA	Complementary DNA
CHF	Congestive Heart Failure
CIMP	CpG island methylator phenotype
CLL	Chronic Lymphoblastic Leukaemia
CML	Chronic Myeloid Leukaemia
CpG	CpG dinucleotides
CSC	Cancer stem cell
CVD	Cardiovascular disease
Decitabine	5-Aza-2’-deoxycytidine
DMR	Differentially methylation region
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
DNMT	DNA methyltransferase gene
EDTA	Ethylenediaminetetraacetic acid
ES	Embryonic stem cells
FoxO	Forkhead box protein O

G-CSF	Granulocyte-colony stimulating factor
GPCR	G protein-coupled receptor
HAT	Histone acetyltransferase
HDAC	Histone deacetylases
HMT	Histone methyltransferases
HNPCC	Hereditary Nonpolyposis Colorectal Cancer
HTS	High-throughput sequencing
HSC	Haematopoietic stem cell
H3K4me3	Histone H3 Lys trimethylation
ICR	Imprint control region
iPSC	Induced pluripotent stem cells
M	Molar
MBP	Methyl-CpG binding protein
miRNA	MicroRNA
MMR	Mismatch repair
MNC	Mononuclear cells
MSI	Microsatellite instability
mTOR	Mammalian target of rapamycin
MRD	Minimal residual disease
mRNA	Messenger RNA
NCCGP	North Cumbria Community Genetics Project
NER	Nucleus excision repair
NHEJ	Non-homologous end-joining
PBL	Peripheral blood leukocytes
PBS	Phosphate-buffered saline
PcG	Polycomb group
PCI	Primary percutaneous coronary intervention
PCR	Polymerase chain reaction
PI3K	Phosphoinositide 3-kinase
PPi	Pyrophosphate
PRC	Polycomb repressive complex
PRMT	Protein arginine methyltransferase
QC	Quality control
qPCR	Quantitative polymerase chain reaction

ROS	Reactive oxygen species
Rpm	Rounds per minute
RNA	Ribonucleic acid
RNAse	Ribonuclease
RT	Room temperature
SAM	S-adenosylmethionine
SES	Socio-economic status
STEMI	ST elevation myocardial infarction
TDG	Thymine-DNA glycosylase
T-DMR	Tissue-specific differentially methylation region
C-DMR	Cancer-specific differentially methylated region
TET	Ten eleven translocation
TSS	Transcriptional start site
WHO	World Health Organization
µg	Microgram
µl	Microliter
µM	Micromolar

Chapter 1. Background

1.1 Epigenetics

Epigenetic changes are heritable changes in gene expression that do not involve changes in the deoxyribonucleic acid (DNA) sequence. Epigenetic mechanisms mediate diversified gene expression profiles, to allow the generation of the variety of cells and tissues required in multicellular organisms. All cells in an organism contain essentially the same genetic information, but not all genes are expressed simultaneously by all cell types (Moore, 2012). The primary mechanisms that are involved in epigenetic regulation of gene expression include; DNA methylation, which is a biological process by which a methyl group is added to a DNA nucleotide and which is a major epigenetic factor influencing gene activities (Moore, 2012); histone modification, which describes the post translation modification that histones undergo, such as methylation, acetylation, ubiquitylation and phosphorylation, which alters their interaction with DNA and nuclear proteins and affects chromatin structure, and which is involved in controlling transcription (Imhof, 2003); chromatin remodelling, which is the dynamic modification of chromatin architecture that enables access to nucleosomal DNA, and thereby controls gene expression (Saha, 2006); and Ribonucleic acid (RNA) associated gene silencing, where non coding microRNAs (miRNAs), single stranded, evolutionary conserved, small, noncoding RNA molecules, are negative regulators of gene expression that mediate post transcriptional gene repression (Avraham, 2012).

Epigenetic events are necessary for the life of organisms and are important in orchestrating key biological processes, such as imprinting, which allows certain genes to be expressed in a parent-of-origin matter (Fedoriw, 2012); silencing of large chromosomal domains, such as inactivation of the X-chromosome in females (Li, 2012); and differentiation, which allows stem cells and multipotent progenitor cells to differentiate into multiple different cell lineages (Wu, 2006) (Figure 1.1).

This project concentrated primarily on one epigenetic modification, DNA methylation, and therefore the rest of this section will primarily focus on DNA methylation.

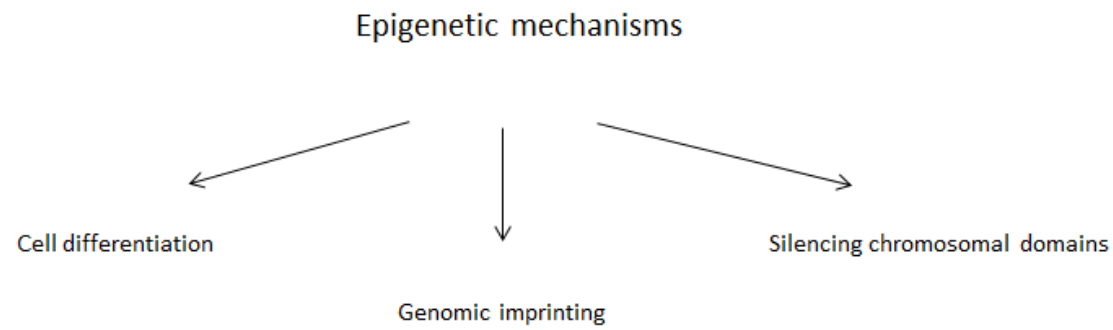


Figure 1.1: The functional roles of epigenetics

Epigenetic mechanisms play important roles in key biological processes, such as; differentiation, which allows stem cell and multipotent progenitor cells to differentiate into multiple different cell lineages; imprinting, which allows certain genes to be expressed in a parent-of-origin specific matter; and silencing of large chromosomal domains, such as inactivation of the X-chromosome in females.

1.1.1 DNA methylation

DNA methylation is a biological process by which a methyl group is added to a DNA nucleotide. In humans, only cytosine bases are susceptible to normal biological methylation and this occurs almost exclusively at cytosine residues which are immediately followed by a guanine, forming the so-called CpG dinucleotide. A small percentage of methylation occurs in non-CG sites, however, this non-CG methylation has almost exclusively been observed in stem cells (Lister, 2009). The methyl group is added to the fifth carbon within the cytosine residue. This does not change the coding properties of the codon in which the base participates.

CpG dinucleotides (CpGs) only comprise around 1% of the genome (Swartz, 1962). Around 85% of these CpGs are spread out in the genome and these sporadic CpG sites found throughout most of the genome are usually methylated, such that 70-80% of all CpGs in the genome are methylated. The remaining 15% of CpG sites are clustered in short regions of genomic DNA (Ehrlich, 1982), so called CpG islands (Cross, 1995). CpG islands comprise about 1-2% of the genome and they encompass the transcriptional start site (TSS) of approximately 60-70% of human protein encoding genes (Bird, 2002; Saxonov, 2006). CpG islands are generally defined as a region of DNA greater than 500 bp, with a GC content of over 55% and an observed CpG content to expected CpG content ratio of over 0.65, as described by Takai *et all* (Takai, 2002), although other slightly different criteria have also been suggested (Gardiner-Garden, 1987). However, it must be noted that all the proposed definitions of CpG islands are based on sequence analysis and there is currently no functional definition of a CpG island. Jones *et al* showed that in the human genome, the frequency of CpGs is about 10 times more frequent in CpG islands than in the surrounding DNA (Jones, 2012). These dense CpG clusters are usually devoid of CpG methylation. Clustering of unmethylated CpGs has allowed the CpG islands to be biochemically isolated as a relatively homogeneous fraction of DNA or chromatin (Jones, 2012). Outside the CpG islands, but within 2kb distance from the TSS is a stretch of CpG sites, named CpG shores (Irizarry, 2009). The region up to 4 kb distance from the TSS is described as CpG shelves (Sandoval, 2011).

The CpG sites in the rest of the genome, outside the CpG islands, shores and shelves, are predominantly present in repetitive elements. These repetitive elements represent around 45% of the human genome and they contain much of the CpG methylation found in normal somatic tissues. The repetitive elements consist of interspersed repeats derived from non-autonomous or autonomous transposable elements and tandem repeats

of simple sequences or complex sequences; such as retroviruses, *LINE1* elements, *Alu* elements and others (Jones, 2012; Weisenberger, 2005).

1.1.2 DNA methyltransferase genes

DNA methylation is carried out by the DNA methyltransferase genes (DNMTs) (Hermann, 2004a), with S-adenosylmethionine (SAM) as the methyl donor (Detich, 2003) (Figure 1.2).

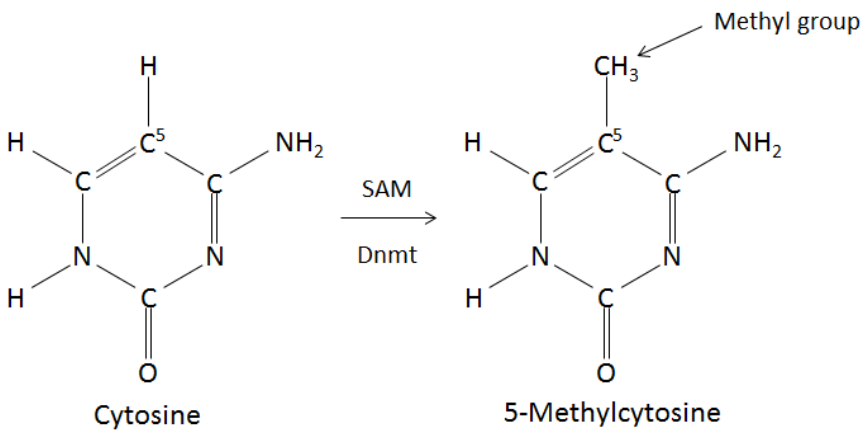


Figure 1.2: The mechanisms of DNA methylation

The mechanism of DNA methylation consists in the production of an altered 5-methylcytosine molecule by the action of the DNMTs. They catalyse the transfer of a methyl group from SAM to the carbon-5 position of a cytosine molecule.

The DNMT protein family in mammals has been shown to have five members, however only three have been shown to have catalytic DNA methyltransferase activity; DNMT1, DNMT3a and DNMT3b (Hermann, 2004a).

DNMT1 had been shown to function primarily in the maintenance of methylation patterns. It was thought to be responsible for the propagation of the DNA methylation information by maintenance of DNA methylation patterns after DNA replication (Hermann, 2004a) and it is able to do this with high fidelity (Hermann, 2004b). DNMT1 was shown to have a preference for hemimethylated sites (Arand, 2012) and whereas unmethylated DNA was a relatively poor substrate for DNMT1, hemi-methylated DNA was efficiently utilised to produce fully methylated sites. This is important to maintain the original methylation patterns of the DNA: during every S-phase, DNA replication converts ~40 million methylated CG sequences into the hemimethylated state and as a result these hemimethylated CG sequences have to be efficiently methylated without methylation of unmethylated CG sites, thus maintaining the original methylation patterns (Hermann, 2004b).

DNMT3a and DNMT3b have been shown to be essential for *de novo* methylation during early mammalian development. Inactivation of both DNMT3a and DNMT3b caused lethality in mice, where knockout of DNMT3b caused early embryonic lethality and inactivation of DNMT3a resulted in death shortly after birth (Okano, 1999) and DNMT3a and DNMT3b deficient induced pluripotent stem cells (iPSCs) were shown to be restricted in their developmental potential (Pawlak, 2011). DNMT3a is required for the establishment of both maternal and paternal imprints (Kaneda, 2004), and has been shown to be important in HSC differentiation (Challen, 2012). DNMT3a is thought to methylate a set of genes or repetitive sequences that are critical during later development or after birth, whereas DNMT3b may play more important roles during early development and is methylating a broader spectrum of target sequences (Okano, 1999).

Even though DNMT1s function is described as the maintenance of methylation patterns, and Dntm3a and Dntm3b being described as being important in the *de novo* methylation, in vivo, neither *de novo* methylation can be exclusively assigned to DNMT3a or DNMT3b nor maintenance methylation exclusively to DNMT1 (Arand, 2012), suggesting a closer relationship between the different DNMTs. A different model has been suggested by Jones *et al* in 2009. They proposed that the bulk of DNA methylation in dividing cells is maintained by DNMT1, whereas DNMT3a and DNMT3b methylate sites that are missed by DNMT1. In this model, the main role of

DNMT1 is in reading the DNA sequence and applying methyl groups opposite to newly replicated hemimethylated CpG sites, without regard for the chromatin configuration in which the particular piece of DNA is located, while DNMT3a and DNMT3b are associated with specific regions of DNA that need to be maintained as highly methylated. If methylated regions remain associated with DNMT3a and DNMT3b, this would ensure that the methylation state of a region is maintained, rather than a specific methylation pattern (Jones, 2009).

1.1.3 DNA de-methylation

Even though DNA methylation is generally thought of as contributing to stable, long-term and heritable transcriptional repression, it is still possible to reverse DNA methylation. Demethylation can either be achieved passively, by the failure of the maintenance of DNA methylation after DNA replication, or actively, by replication-independent processes (Chen, 2011). While until recently there were no known mechanisms for DNA demethylation the existence of such mechanisms had been inferred by observations that DNA methylation undergoes dramatic global changes at specific stages of mammalian development. Shortly after fertilization, the sperm-derived pronucleus undergoes a rapid, global loss of DNA methylation that occurs independently of DNA replication, while the maternal pronucleus undergoes demethylation at a later stage, namely during the first few cell cycles of development and this demethylation is replication-dependent (Sanz, 2010; Wu, 2010).

However, recently a family of proteins, known as the ten eleven translocation (TET) proteins, have been identified that appear to function as the primary DNA demethylases in mammalian cells. These TET proteins have shown to promote active DNA demethylation in mammalian cells *in vitro* and in the mouse brain *in vivo* (Guo, 2011; Tahiliani, 2009). The suggested model of DNA demethylation by the TET proteins starts by the oxidation of 5-methylcytosines (5mCs) into 5-hydroxymethylcytosines (5hmCs). 5hmCs are then deaminated into 5-hydroxyluracil (5hmU) by the deaminases activation-induced deaminase (AID) and apolipoprotein B mRNA-editing enzyme complex (APOBEC). Finally, 5hmU can be excised by 5hmU glycosylases and repaired by the base excision repair (BER) pathway with unmethylated cytosines (Guo, 2011; Cimmino).

Another proposed mechanism of DNA de-methylation is the glycolase-dependent DNA demethylation model. In this model, 5mC is deamiated directly to thymine by AID and

APOBEC, which is then followed by excision of the T/G mismatch by MBD4 and the thymine-DNA glycosylase (TDG) (He, 2011).

Other than these biological mechanisms of demethylation, methylation can be pharmacologically reversed by several drugs, such as the DNMT inhibitors 5-azacytosine (azacytidine) and 5-Aza-2'-deoxycytidine (decitabine) (Yan, 2012; Mund, 2005). These drugs cause an inhibition of the DNMTs, resulting in overall de-methylation. Compounds from both drugs are imported by cells. In the cell it is converted to the activated triphosphate 5-Aza-dCTP and incorporated into the DNA. The presence of a nitrogen at the 5 position of the azacytidine nucleobase results in an irreversible covalent complex with DNMT1 and this triggers proteasome-mediated DNMT1 degradation. Azacytidine can also be converted to the ribonucleoside triphosphate, 5-Aza-CTP, into RNA.

However, these DNMT inhibitors are not selective towards the different DNMTs, are chemically instable and have strong secondary effects, *e.g.* renal toxicity and myelotoxicity (Foulks, 2012; Gros, 2012). Nevertheless, these drugs are clinically used in myelodysplastic syndrome and acute myelogenous leukemia, where dose and exposure time are critical; *in vitro* experiments showed a significant effect on cell viability in higher doses or long exposure times (Kwon, 2012) and an *in vivo* study in mice has shown that high exposure of these drugs can have toxic side-effects and cause lethality's (Chuang).

1.1.4 The role of DNA methylation in cellular processes

DNA methylation plays an important role in numerous cellular processes. It is classically assigned instructional modifications for germ line imprint control regions (ICRs), which are required to establish parental identity of imprinted loci and to maintain differential epigenetic pattern throughout development. Genomic regions of differential DNA methylation, called differentially methylated regions (DMRs), function as ICRs and they have been identified at almost all imprinted gene clusters (Fedoriw, 2012; Smith, 2013).

Another role of DNA methylation is X-chromosome inactivation. Mammalian female cells contain two X chromosomes and as a result, a problem of X-linked gene dosage must be solved. As such, X-chromosome inactivation serves to balance sex-chromosome-linked gene dosage in female somatic cells, by rendering one of the two X chromosomes largely transcriptionally silent by epigenetic mechanisms. DNA

methylation plays an important role in this X-chromosome silencing. This inactivation is mediated upstream of DNA methylation by the non-coding RNA X-inactivation-specific transcript (*Xist*) and Polycomb-group-mediated H3K27 methylation (Fedoriw, 2012; Smith, 2013).

De novo methylation also plays a role in the transcriptional circuitry associated with pluripotency, which is rapidly silenced on differentiation as embryonic programmes are resolved towards specific lineages. *In vivo* experiments showed that differentiation of adult stem cells was associated with small but informative changes in the genomic distribution of DNA methylation (Bock, 2012; Smith, 2013), emphasizing the role of DNA methylation in cell differentiation processes.

1.1.5 The role of DNA methylation in controlling gene expression

Early experiments had provided evidence that DNA methylation was also likely to play a crucial role in repressing gene expression. The state of methylation of CpG islands near gene promoters was found to be associated with the transcriptional activity of a gene (Bird, 1986), and strong negative correlations have been observed between DNA methylation levels and the presence of histone marks that target active genes (Bell, 2011). CpG islands at promoters are hypomethylated corresponding to a permissive chromatin structure in order to poise genes for transcriptional activity (Antequera, 1993) and hypermethylation of gene promoter is related to an inactive state, preventing genes from being expressed. Besides CpG islands near gene promoters, CpG shores have been related to gene expression. In these CpG shores, tissue-specific differentially methylation region (T-DMR) methylation and cancer-specific differentially methylated region (C-MDR) methylation showed a strong inverse relationship with differential gene expression (Irizarry, 2009).

These observations provide evidence that DNA methylation is associated with the transcriptional activity of genes and it can affect this transcription by two mechanisms. Firstly, methylation of cytosine bases can directly inhibit the binding of some transcription factors when present at their target site (Bird, 2002). Transcription factors that have CpGs in their recognition sequences and are thus methylation-sensitive include the Ah receptor, CREB/ATF, E2F, ETS factors, c-Myc, NF-κB and HiNF-P. Other transcription factors are not sensitive to methylation, for example Sp1, CTF and YY1 and, as a result methylation does not hinder binding of these gene-specific transcription factors, but rather prevents the binding of ubiquitous factors, and subsequently transcription (Baron, 2012).

In the second mechanism in which DNA methylation can affect transcriptional activity, both histone modification and DNA methylation act together to influence chromatin structure and as a result gene expression. To fully understand the interaction between DNA methylation and histone modifications, the next section will first go into more detail about the different histone modifications and then the next section will discuss the interplay between these two epigenetic modifications.

1.1.6 Histone modifications

DNA and histones are components of a chromosome, in which the DNA helix is wrapped around core histones to form the DNA structure. Histones are positively charged and can so from a tight interaction with the negative charged DNA phosphate groups, producing dense chromatin. The first step in the formation of this higher-order chromatin is the formation of the nucleosome core particle. This structure comprises an octamer of core histones; two copies of each histone; H2A, H2B, H3 and H4, around which are wrapped 146 base pairs of DNA in 1 ¾ superhelical turns (Luger, 1997). The N-terminal tails of these histone proteins are exposed in the area surrounding the nucleosomes and are subjected to a variety of post translational modifications, including phosphorylation, methylation, acetylation and ubiquitination (Molina-Serrano, 2013). These modifications may, singly or in combination, influence numerous biological processes. The possible complex interplay between the different modifications occurring on the histone tails is being described in the histone code hypothesis. This hypothesis predicts that a pre-existing modification affects subsequent modifications on histone tails and that these modifications serve as marks for the recruitment of different proteins or protein complexes to regulate diverse chromatin functions, such as transcription and gene expression, DNA replication and chromosome segregation (Strahl, 2000).

Acetylation of Lys residues, predominantly in histones H3 and H4 is mediated by enzymes with histone acetyltransferase (HAT) activity, using acetyl-CoA as the acetyl donor. Acetylation removes the positive charge of lysine and alters interactions with DNA and other chromatin-associated proteins, resulting in decreased binding ability of the histones. In general, acetylation of core histone tails correlates with opening of chromatin structure to allow transcription. HATs can be categorized into three major families: the MYST family containing MOZ, Ybf2/Sas3, Sas2, Tip60, the Gcn5 related N-acetyltransferase or GNAT family, and the more recently characterized p300 family (Berndsen, 2008).

Conversely, acetyl groups are removed from the ε-N-acetyl-lysine by histone deacetylases (HDACs), allowing the histones to wrap the DNA more tightly to condense the DNA structure and this is preventing transcription. Mammalian HDACs can be divided into three subclasses. Class I enzymes are homologous to the yeast Rpd3 protein and include the mammalian HDAC1, HDAC2, HDAC3 and HDAC8 enzymes. Class II HDACs (HDAC4, HDAC5, HDAC6 and HDAC7) are larger proteins that are related to the yeast Hda1 protein. Recently, a third class of HDACs, with homology to the yeast Sir2 protein and several putative mammalian members, has been identified. HDACs bind to many different proteins and usually exist in large complexes within the cell (Kramer, 2001).

Both HAT and HDAC activities can be recruited to target genes in complexes with sequence-specific transcription factors and their cofactors. The dynamic interplay between HATs and HDACs dictates the ultimate state of acetylation. The fact that the amino-acid sequence of histones are some of the most highly conserved among eukaryotes and that there is high sequence conservation of HATs, underscores the critical nature of histone acetylation in genome regulation (Berndsen, 2008).

Selected lysines and arginines can be methylated in the N-terminal tail domains of H3 and H4. The enzymes responsible, the histone methyl transferases (HMT), are either lysine or arginine specific and several have now been identified and characterized. Two families of HMTs have been described; protein arginine methyltransferases (PRMTs) catalyze the transfer of methyl groups from SAM to the guanidino nitrogens of arginine residues; and the SETdomain family is a second important signature motif for protein methyltransferases. The known functions of these two family of proteins predict that histone methylation is likely to play important roles in multiple biological processes including transcription, signal transduction, development, and cellular proliferation and differentiation (Zhang, 2001).

Many promoters in embryonic stem (ES) cells harbor a distinctive histone modification signature that combines the activating histone H3 Lys 4 trimethylation (H3K4me3) mark, imparted by the Set/MLL histone transferase and the repressive H3K27me3 mark, imposed by the polycomb group (PcG) proteins. These bivalent domains represent a preprogrammed epigenetic signature that is considered to poise expression of developmental genes, allowing timely activation while maintaining repression in the absence of differentiation signals (De Gobbi, 2011; Han, 2012).

PcGs are histone modifiers that reside in two multi-protein complexes: Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2). PRC2 catalyses the methylation of

H3K27, which is then accompanied by the recruitment of PRC1 members, leading to the formation of repressive chromatin (Radulovic, 2013).

1.1.7 The interplay between DNA methylation and histone modifications

Even though DNA methylation and histone modification are thought to be independent processes, they might influence each other in different ways. Firstly, DNA methylation might direct histone modifications. Methyl-CpG binding proteins (MBP), such as MBD1, MBD2, MBD3 and MeCP2, and HDAC complexes are recruited to the methylated region to induce histone deacetylation and silencing. The chromatin then HMTs which methylate the lysine 9 residues on histone H3 (H3K9) and this stabilises the inactive state of chromatin.

Secondly, histone methylation might direct DNA methylation. Methyl H3-K9 acts as a signal for inactive chromatin by recruiting HP1 to methylated histones, which might in turn recruit DNA methyltransferases directly or indirectly to the silent chromatin to maintain DNA methylation and stabilize the inactive chromatin.

The third way is a model of chromatin remodelling driving DNA methylation. The ATP-dependent chromatin remodelling and DNA helicase activities of proteins might facilitate DNA methylation and histone modification by unwinding nucleosomal DNA to increase its accessibility to DNMTs, HDACs and HMTs. The disruption of these proteins affects both DNA methylation and histone methylation (Li, 2002; Parry, 2011; Bronner, 2010).

All these observations together suggest that DNA methylation does not necessarily play a general role as a transcriptional regulator (Antequera, 2003), but that it rather plays a role in maintaining the silenced state which has been induced by the associated epigenetic marks on histones and other epigenetic molecules (Feng, 2006).

1.2 Ageing and age-related diseases

Important demographic changes occurring across most of the world are leading to an increasing elderly population (Christensen, 2009). Life expectancy has increased steadily over the past 2 centuries (Vaupel, 2010) and according to the World Health Organization (WHO), there will be 2 billion people aged 60 and older in 2050 (WHO, 2013). However, a significant proportion of the extra life years are associated with morbidity, and as a result, aging and age-related diseases are becoming major health priorities.

The fundamental defining manifestation of ageing is an overall decline in the functional capacity of various organs to maintain tissue homeostasis (Kirkwood, 2005) and to respond adequately to physiological needs under stress (Sahin, 2010). Ageing can result in both physical and psychological changes. During ageing, the capacity to adapt declines, which might affect personality, mood, attitude and self-concept of an individual. Changes may occur in perception and memory, leading to learning and problem solving difficulties. During ageing, degenerative changes can occur, such as the inability to see, hear, feel and react to stimulus. Spatial orientation, mobility and motor coordination decline and the work rate of body systems diminish (WHO, 2012).

The stochastic theory of ageing suggest that organisms limit their investment in long-term maintenance of somatic tissues and that, as a consequence, ageing develops through cumulative damage induced at various levels through a gradual accumulation of unrepair cellular defaults (Sozou, 2001).

With the increasing number of aged individuals, age-related diseases are now the dominant health problems in most countries. The incidence of these diseases increases rapidly with age and examples of these age-related diseases include cardiovascular disease, cancer, arthritis, cataracts, osteoporosis, type 2 diabetes mellitus, hypertension, Alzheimer and dementia (WHO, 2012). Even though the incidence of all these diseases increases with age, the most common causes of death worldwide are cancer and cardio-vascular disease (Figure 1.3) and the fraction of deaths attributable to these two disease is expected to increase in the next 10-20 years. Therefore, the main focus will be on these two diseases from here on.

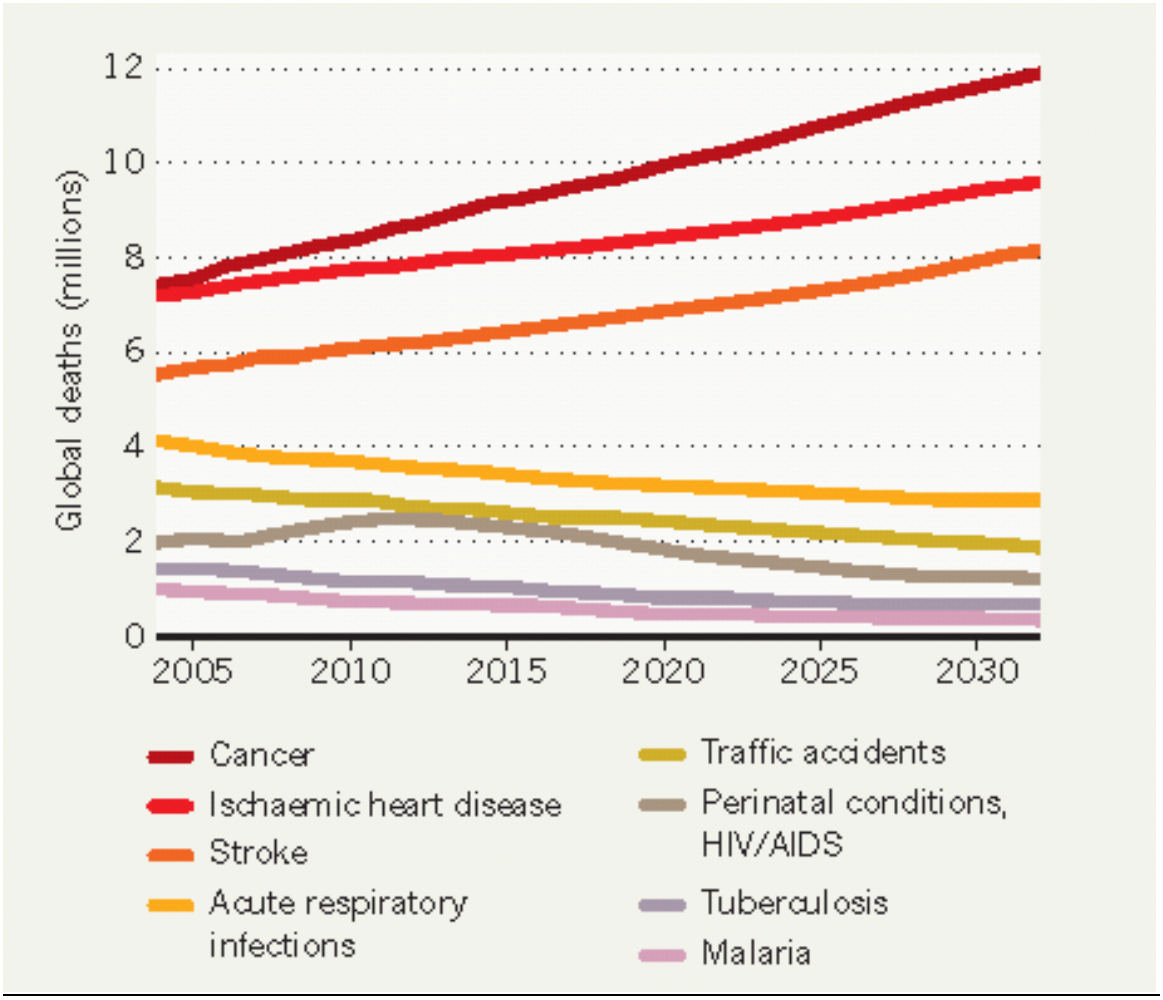


Figure 1.3: Causes of death worldwide

Nowadays, cancer and atherosclerosis are the most common cause of death in the world and this number is expected to increase in the next 10-20 years. This figure is adapted from Cannon *et al* (Cannon, 2013).

1.2.1 Cancer

Cancer is a process in which cells divide and grow uncontrollably and so forming malignant tumours. Cancer arises from a stepwise accumulation of genetic and epigenetic changes that allow neoplastic cells to escape from the homeostatic mechanisms that govern normal cell proliferation. In humans, it has been suggested that at least four to six mutations are required to reach this state (Hahn, 2002).

Acquisition of several types of molecular alterations is necessary to drive a population of normal cells to become a cancer. These alterations include resisting cell death, sustaining proliferative signalling, evading growth suppressors, activating invasion and metastasis, enabling replicative immortality, inducing angiogenesis, obtaining the capability to reprogramming cellular metabolism and evading immunological destruction (Hanahan, 2011). When a single cell accumulates a number of these alterations, usually over a time span of several years, the cells can escape from most restraints on proliferation. The cell and its descendant can than develop additional alterations and accumulate in increasingly large numbers, forming a population that consists mostly of these abnormal cells, which are leading to a tumour formation, which consists of cells with diverse phenotypes.

1.2.1.1 Haematopoietic malignancies

Blood cancers can occur at all ages, but the majority of cases are observed in adults. In fact, 94% of cases diagnosed in the UK every year are adults (LLR, 2013).

There are a number of different types of malignancy that occur in the haematopoietic system, although the majority can be defined as either types of leukaemia or lymphoma. Leukaemia, is a cancer of the tissue in the bone marrow and lymph nodes that manufacture blood cells. It is characterized by the excessive accumulation of leukocytes, or white blood cells. Leukaemias are mainly divided into four subgroups: acute lymphoblastic leukaemia (ALL), acute myeloid leukaemia (AML), chronic lymphoblastic leukaemia (CLL) and chronic myeloid leukaemia (CML) and these types differ substantially in their cellular origin and clinical behaviour. The acute leukaemias are characterized by a rapid increase in the numbers of immature cells (known as blasts), while the chronic leukaemias are characterized by the excessive build-up of relatively mature cells (Sawyers, 1999).

The leukaemic transformation can occur at different points during haematopoietic differentiation. All the blood cells derive from the same pool of haematopoietic stem cells produced in the bone marrow and are then differentiated into

either the lymphoid stem cell pool or the myeloid stem cell pool. The lymphoid stem cells then give rise to the NK cells, T-lymphocytes and B-lymphocytes, while the myeloid stem cells give rise to the monocytes, erythrocytes, and granulocytes. The leukaemic transformation that leads to lymphoid leukaemias occur in the lymphoid precursors and as a result, can occur in both the B-lymphocyte and the T-lymphocyte lineages. Even though the lymphoid leukaemias have been related to age, CLL is largely found in adults over the age of 50 and the incidence of ALL rises after the age of 60 (Frei, 1964), 60% of all ALL cases diagnosed are children under the age of 5, making it the most common form of cancer in children (LLR, 2013).

AML is a disease that is caused by hyperproliferation of myeloid progenitor cells and the resulting leukaemia can have features of granulocytes, monocytes and/or the erythroid cells and megakaryocytes or more closely resemble immature progenitor cells. This disease can occur in children, although more rarely than ALL, and is much more common in adulthood, showing an exponential rise in incidence after the age of 40 (Frei, 1964).

The median age of CML onset is 45 to 55 years. CML is a clonal myeloproliferative disorder that is characterized by the presence of a BRC-ABL fusion, which causes a deregulated activity in the haematopoietic stem cell population, resulting in the over-production of mature myeloid cells. In the blast crisis phase, CML is characterised by a major clonal expansion of immature progenitors, which have either myeloid or lymphoid features (Skorski, 2012; Sawyers, 1999).

1.2.2 Cardiovascular diseases

Cardiovascular diseases (CVDs) are pathologies of the heart, blood vessels and the vascular system of the brain. According to the WHO, CVDs are one of the leading causes of death and disability in the world population (WHO, 2011). Atherosclerosis is the progressive process underlying most cardiovascular diseases. Atherosclerosis involves the formation in the arteries of atherosclerotic plaques, which are characterized by inflammation, lipid accumulation, cell death and fibrosis. These lesions are a heterogeneous mixture of cells, predominantly vascular smooth muscle cells, lymphocytes, macrophages and extracellular lipids, collagen and matrix. The release of growth factors and inflammatory cytokines from these various cell types promotes further accumulation of inflammatory cells and deposition of extracellular matrix components causing the lesion to develop into an advanced plaque and causing the inner layer of the artery wall to thicken (Ross, 1993). This results in a reduction of the artery's

diameter and a decrease in blood flow and oxygen delivery. When this plaque raptures or cracks open, the sudden formation of a blood clot can cause thrombosis. Atherosclerosis can lead to myocardial infarction and heart failure if the blood flow to the heart is blocked and in the arteries that perfuse the brain it can cause ischaemic stroke and transient ischaemic attacks. In other arterial branches, atherosclerosis can results in renal impairment, hypertension, abdominal aortic aneurysms and critical limb ischaemia (Oxford University Press., 2010). Results from Olivetti *et al* suggested that ageing of the human heart is characterized by myocyte loss, reactive myocyte cellular hypertrophy and reduced ventricular mass and that these processes that may underlie the cause of myocardial dysfunction with age (Olivetti, 1991). Several other studies have observed that with age, left ventricular stiffness and vascular and ventricle wall thickness increases. At the same time arterial compliance, endothelial function and ventricular contractility decline with age (Arbab-Zadeh, 2004; Lakatta, 2003). These observations indicate that cardiac functions decrease with age, leading to an increased risk of cardiac dysfunction with age.

1.2.3 Molecular pathways involved in ageing and age-related diseases

Identifying the underlying molecular changes that contribute to ageing and to the development of these diseases related to age will be critical for improving health outcome for elderly patients and also in under-pinning potential preventative strategies. Different molecular pathways of haematopoietic cell ageing have been linked to age-associated changes on the molecular level and these same pathways have been shown to play a role in the development of the age-related diseases cancer and cardiovascular disease.

1.2.3.1 Cell signalling pathways associated with ageing

In most cases, the longevity response is under active control by specific regulatory proteins and with age this protein homeostasis declines and damage accumulates (Kenyon, 2010). These changes lead to a functional decline of the haematopoietic stem cells (HSCs) with an increasing age. Several molecular pathways have been linked to these age-related changes in the haematopoietic system.

The first molecular pathway is the phosphoinositide 3-kinase (PI3K) pathway, which is one of the major pathways modulating cell survival, cell growth, proliferation, metabolism and angiogenesis (Figure 1.4). One of the components of this pathway is the mammalian target of rapamycin (mTOR). Recently, it has been shown that differential

expression of *in vivo* mTOR signalling is associated with human ageing (Harries, 2012). The expression patterns as observed in human studies are similar to the expression patterns as observed in mice studies, where was shown that an increase in mTOR signalling is sufficient to cause premature ageing of HSCs in young mice (Chen, 2009). Age-associated changes have been observed in other components of the PI3K pathway in rat studies; the Forkhead box protein O (FoxO) showed age-induced decreases in the nuclear binding activity and also sirtuin 1 deacetylase decreased with age (Kim, 2012).

The second molecular pathway of age-related changes on the molecular levels is the cellular mortality pathway. Several studies have shown a link between senescence and accelerated ageing (Figure 1.5). For example, a study in mice showed that the induction of senescence, triggered by p53 deficiency, leads to an accelerated aging phenotype (Keyes, 2005). Expression of $p16^{INK4a}$ and *Ink4a/Arf* were shown to correlate *in vivo* with impaired proliferation and failure of HSCs (Krishnamurthy, 2004). Both $p16^{Ink4a}$ and $p19^{Arf}$ (one of the parts of *Ink4a/Arf*) have been implicated as downstream effectors of BMI-1 expression, which was shown to be involved in the control of proliferation and efficient self-renewing cell divisions of adult HSCs (Park, 2003). Indeed, BMI-1 was found to be much more expressed in young individuals compared to older subjects in keratinocytes and the age-related increase of $p16^{INK4a}$ expression in primary human keratinocytes could be principally attributed to the modulation of BMI-1 levels (Cordisco, 2010).

The third molecular pathway that is linked to age-associated changes is the DNA repair pathway (Figure 1.6). Rossi *et al* showed that genetic deficiencies in telomere maintenance, nucleotide excision repair (NER) and non-homologous end-joining (NHEJ) intrinsically diminish long-term reconstituting-HSC function in an age-dependent manner under conditions of stress. Rudolph *et al* showed a diminished capacity of the aged telomerase-deficient mice to respond to a stress. These results indicate that DNA damage may underlie the reduced capacity of stem cells to mediate a return to homeostasis after exposure to injury or stress (Rossi, 2007; Rudolph, 1999).

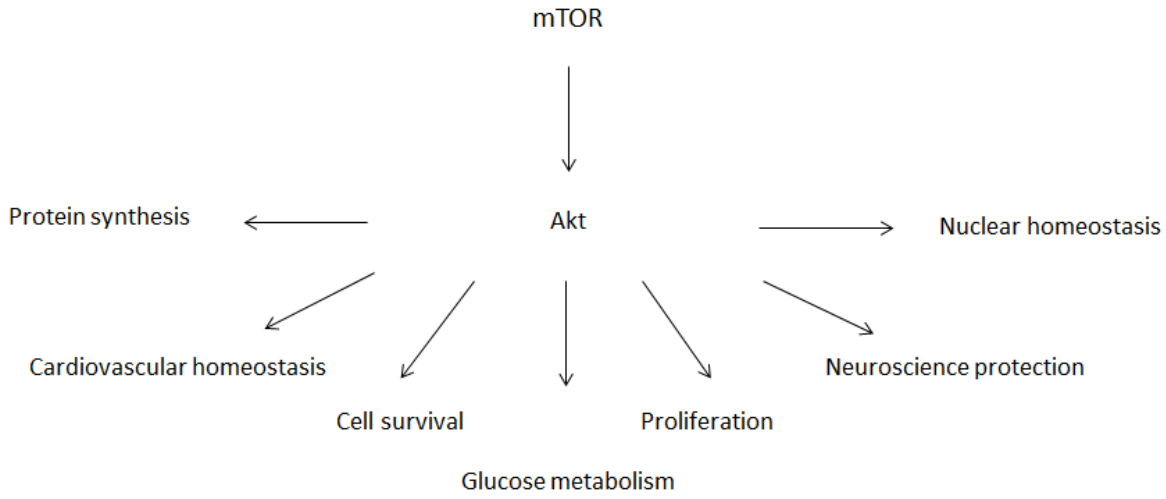


Figure 1.4: The PI3K pathway

A simplified version of the PI3K pathway shows the of this pathway in protein synthesis, cell survival, cell proliferation, glucose metabolism, etcetera (Cellsignal, 2011).

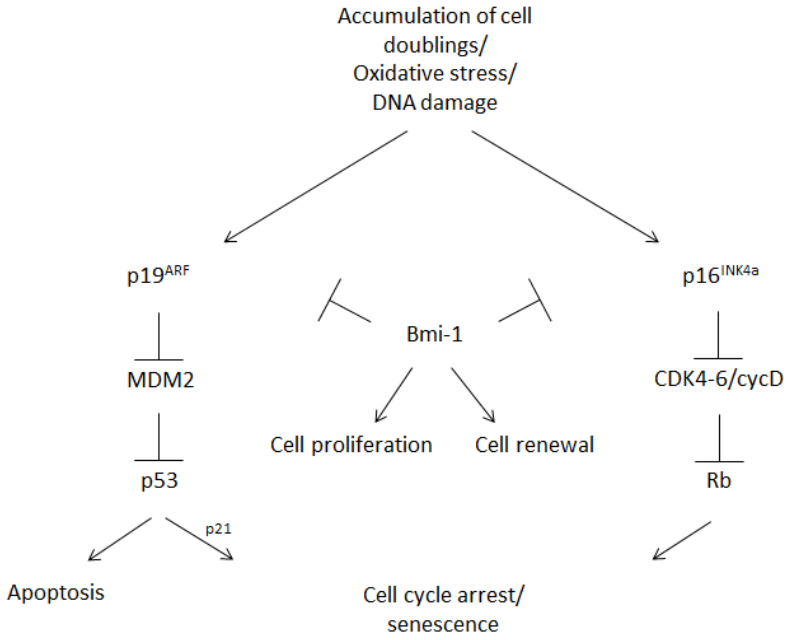


Figure 1.5: The cellular mortality pathway

Several proteins are involved in the cellular mortality pathway, including p19, p16, p53, p21 and Rb.

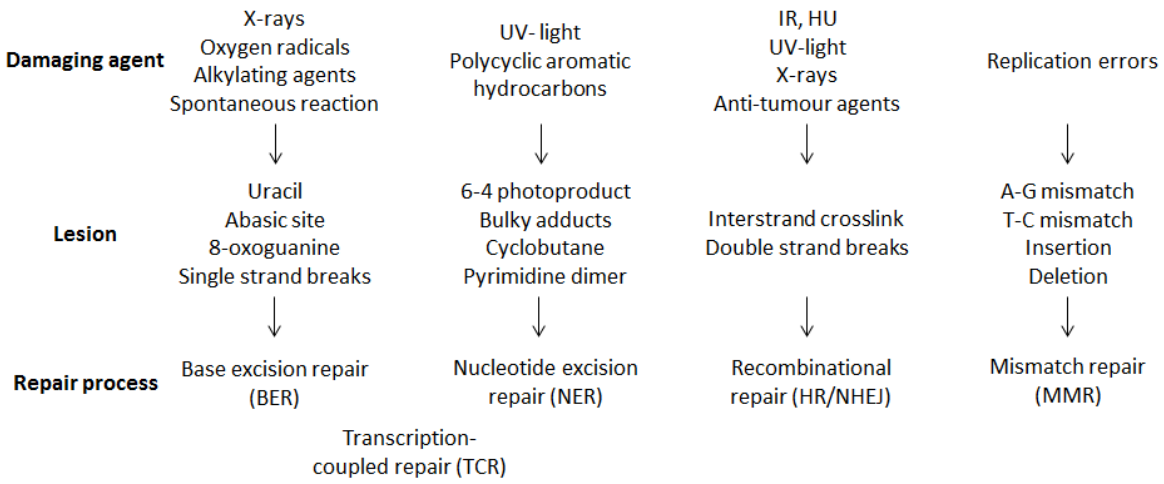


Figure 1.6: The DNA repair pathway

Several DNA repair mechanisms are known. These include base excision repair (BER), Nucleotide excision repair (NER), recombinational repair, which includes both homologous recombination (HR) and non-homologous end-joining (NHEJ) and mismatch repair (MMR) (Jeppesen, 2011).

1.2.3.2 Cell signalling pathways associated with age-related diseases

The pathways that are involved in ageing have also been implicated in the development of both cancer and cardiovascular diseases. With an increasing age, DNA damage accumulates and therefore a properly established DNA damage response network is necessary. If the damage is too extensive to be repaired, the network can trigger the onset of senescence or cell death to eliminate cells. When the DNA remains unrepaired this can lead to the development of cancer (Signer, 2013). Unsuccessful DNA repair can also result in an accumulation of both nuclear and mitochondrial DNA damage. These affected cells undergo senescence or apoptosis and secrete proinflammatory cytokines, causing a build-up of atherosclerotic plaque, leading to cardiovascular diseases (Wang, 2012).

Aging-associated telomere shortening can contribute to the evolution of genome instability and cancer formation by inducing chromosome end resection, fusions, and breakage. Transient telomere dysfunction has been shown to increase cancer initiation and to promote the development of malignant tumours (Begus-Nahrmann, 2012) and deregulation of downstream transcription factors in the PI3K/Akt signalling pathway, such as RUNX2, mTORC1 and mTORC2, have also been indicated to contribute to tumorigenesis (Pande, 2013; Gulhati, 2011). This telomere shortening is also associated with increased levels of cellular senescence, which can facilitate heart failure. However, it is not clear whether short telomere length directly contributes to the development and progression of heart failure, or if the underlying cause of heart failure lead to accelerated telomere shortening (Wong, 2010). And finally, proteins involved in the cellular mortality pathway of ageing, such as p53 and $p16^{Ink4a}$, are known to function as tumour suppressor genes (Nobori, 1994).

1.2.4 Epigenetics changes associated with ageing and age-related diseases

Epigenetic changes have been observed during normal ageing and during the development of several age-related diseases. It may therefore be one of the key players in contributing to the development of these diseases related to ageing.

1.2.4.1 Epigenetic changes associated with ageing

Alterations in the patterns of DNA methylation are one of the hallmarks of ageing (Mathers, 2006). These changes include reduced levels of global DNA methylation across the genome, in conjunction with local areas of hypermethylation, often centring on promoter associated CpG islands. Furthermore, DNA methylation levels have been

observed to change with age in multiple tissue types in both mice and human studies (Wilson, 1983; Maegawa; Teschendorff; Bell, 2012; Issa, 1994), implying that altered patterns of methylation are an inevitable consequence of ageing in mammalian cells. The exact timing of these changes in methylation, and the biological driving forces behind them, remains unclear, but several human and animal studies have suggested that it is not a process solely associated with old age, but that it is an on-going process across the life-course (Bell, 2012; Takasugi). A recent study by Teschendorff *et al* indicated that increased levels of DNA methylation in promoter regions are steadily acquired during the life-course (Teschendorff) and it has been observed by Talens *et al* that epigenetic variation in the population also increases gradually with age. The rate at which these changes occurred differed between different loci and these changes can be substantial at loci regulating transcription of nearby genes (Talens, 2012). While several studies reported that regions near gene promoters become hypermethylated with age, several repetitive sequences lose methylation with increasing age, such as the *Alu elements* (Bollati, 2009), *Line-1* (Zhu, 2012) and *HERV-K* (Jintaridth). Even though age-related methylation changes are observed in many tissues, patterns of DNA methylation are tissue specific (Davies, 2012), and during ageing, individual genes acquire differential methylation patterns in different tissue-types (Maegawa; Thompson, 2010). Inter-individual changes in DNA methylation was shown to have some degree in familial clustering, indicative of a genetic component (Bjornsson, 2008), while another study in monozygotic twins, genetically identical individuals, showed that an increase in epigenetic differences was observed between twin pairs with age in different tissues (Fraga, 2005). Talens *et al* showed that an increasing discordance in epigenetic differences of monozygotic twins occurred during ageing (Talens, 2012) and in the study of Gordon *et al*, monozygotic twins were shown to exhibit a range of within-pair methylation differences at birth, although they showed discordance levels that were generally lower than those seen in dizygotic twins (Gordon, 2012). These observations suggest that external and/or internal factors can have an impact by altering the epigenetic modifications, and that this is not solely relying on genetic information (this section was written as part of a review (van Otterdijk, 2013)). The increased global epigenetic differences with age that was observed in the study of Fraga *et al* was not observed in another monozygotic twin study. In this study, a general change with age in genome-wide methylation was not observed. However, they did identify a subset of loci to be highly correlated with age, which was enriched for genes known to influence age-related diseases, mainly cardiovascular and neurological diseases (Bocklandt).

Besides epigenetic changes, age-related changes in histone modifications have also been observed. In mammals, H3K27me3 regulators play important roles in stem cell aging. BMI1, a member of the mammalian H3K27me3 PRC1, is necessary for the self-renewal of adult stem cells, in large part through repression of the p16INK4a/p19ARF locus (Figure 1.5). Perturbations in H3K27me3 and H3K4me3 regulators both result in defects in stem cell proliferation, suggesting that a delicate balance between activation and repression of distinct sets of target genes is required during the stem cell ageing process (Han, 2012). Genes that harbor both the H3Kme3 and H3K27me3 histone marks, the so called bivalent chromatin domains, were shown to be highly enriched targets for preferential methylation in both ageing and cancer (Rakyan).

1.2.4.2 Epigenetic changes associated with cancer

During cancer development, reduced levels of global DNA methylation is observed, together with hypermethylation of some CpG islands in gene promoters (Gama-Sosa, 1983). These changes in methylation of promoters may mimic the effect of mutations of various tumour suppressor genes or proto-oncogenes and promoter hypermethylation leads to the transcriptional gene silencing. Hypomethylation of regulatory DNA sequences in the genome can activate the transcription of proto-oncogenes, retrotransposons and genes encoding proteins involved in genomic instability and malignant cell metastasis. In this way, both hypermethylation and hypomethylation can play an instrumental role in cancer development (Luczak, 2006).

Hypermethylation is frequently observed in the same genes that are often mutated in familial cancers, emphasising their causal importance in tumorigenesis. For example, loss of function of DNA MMRs, including *MLH1*, are causal for hereditary nonpolyposis colorectal cancer (HNPCC) (Crepin, 2012), while in sporadic colorectal cancer, loss of microsatellite instability (MSI)(an effect of MMR) is due to hypermethylation of *MLH1* (Wheeler, 2000). Another example is the congenital *BRCA1* mutation, which is a cause of breast and ovarian cancer. It is the same gene however, that becomes frequently hypermethylated in sporadic breast and ovarian cancers (Bal, 2012; Press, 2008). In some cases, frequent hypermethylation of genes can be used to identify novel tumour suppressor genes, as was the case for *RASSF1* (Dammann, 2001), *HACE1* (Kucuk, 2012), and *TWIST2* (Thathia, 2012). Hypermethylation of tumour suppressor genes is a feature of several cancers, including the *p15* gene in leukaemia (Wong, 2000), the *Rb* gene in retinoblastoma (Sakai, 1991), and the *VHL* gene in renal

tumours (Prowse, 1997). On the other hand, hypomethylation was oberved in several cancers in tumour promoting or metastasis promoting genes, such as the *uPA* gene in breast cancer (Pakneshan, 2004) and the *MAGEB2* gene in head and neck squamous cell carcinomas (Pattani, 2012) (this section was written as part of a review (van Otterdijk, 2013)).

1.2.4.3 Epigenetic changes associated with cardiovascular disease

Changes in methylation patterns have also been observed during cardiovascular diseases and the genes affected by changes in their methylation patterns are thought to be involved in lipid oxidation, inhibition of endothelial cell migration and formation, the control of cell proliferation and angiogenesis (Friso, 2012; Hiltunen, 2002; Movassagh; Post, 1999). Gene expression is dysregulated in heart failure, and it is hypothesized that a subset of this connecting end-stage disease with different etiologies may be explained by differential DNA methylation, together with other epigenetic mechanisms. The interaction between histone deacetylation and DNA methylation in myocardial disease will be particularly important to understand, since the role of HDAC in cardiac hypertrophy and heart failure has already been established (Gallo, 2008). A global DNA hypo- and hypermethylation has been observed in atherosclerotic lesions and lesion development (Hiltunen, 2003). Hypomethylation of satellite repetitive sequences was observed in diseased hearts compared to healthy hearts (Haider, 2012) and genomic hypomethylation was observed in human atherosclerotic lesions compared to healthy arteries and in proliferating smooth muscle cells *in vitro* and in rabbits *in vivo* (Hiltunen, 2002). Several genes were reported to become hypomethylated in cardiovascular diseases. For example, hypomethylation of *ALOX15* and *F7*, genes that are involved in lipid oxidation, were observed in atherosclerotic plaques and in the venous blood of patients suffering from a coronary atherosclerotic disease (CAD), in the absence of a -323del/ins polymorphism (Hiltunen, 2002; Friso, 2012). Hypomethylation of *AMOTL2*, a gene involved in the inhibition of endothelial cell migration and formation, was observed in the left ventricular of end-stage heart failure patients compared to healthy controls (Movassagh). Besides hypomethylation, several genes were described to become hypermethylated in cardiovascular diseases. For example, as *ERα*, a gene involved in the control of cell proliferation, was observed to become hypermethylated in coronary atheroma's (Post, 1999). In the left ventricular of end-stage heart failure patients, hypermethylation of *PECAM1* and *ARHGAP24* was observed, genes involved in angiogenesis and tube formation (Movassagh).

The research of Chang *et al* showed that homocysteine might promote the methylation of the *FGF2* promoter and in this way, repress the *FGF2* transcriptional activity in arterial endothelial cell proliferation, migration and tube formation (Chang, 2008). Elevated plasma homocysteine levels are a known potential risk factor for vascular diseases. Homocysteine might also change DNA methylation levels in other functional relevant genes, such as *iNOS*, repressing its regulation of blood vessel function (Yideng, 2008) and *APOE*, a gene that is crucial for the hepatic clearance of remnant lipoprotein (Yi-Deng, 2007).

Chapter 2. Project hypothesis

2.1 Project hypothesis

In healthy individuals, CpG islands remain mostly methylation free, whereas most of the non-island associated CpG sites in the bulk of the genome are methylated (Cooper, 1989). Even though patterns of methylation are inherited through cell divisions, the copying of methylation patterns from parental to daughter strand is not 100% efficient and methylation errors accumulate (Ushijima, 2003). The rate at which these errors are accumulated increases with age and during disease development. The changes in DNA methylation patterns observed during ageing are reminiscent of the methylation alterations seen during the development of cancer and potentially other age-related diseases. This includes loss of DNA methylation at the genome wide level in combination with gains in methylation levels in CpG islands in or near gene promoters. This raises the possibility that the acquisition of methylation changes during healthy ageing and during the development of age related diseases might be linked to each other. The methylation patterns of functionally important promoter regions are reported to be more stable as opposed to non-promoter regions and the corresponding genes might be involved in longevity (Bell, 2012; Kaminsky, 2009). Global DNA methylation levels were reported to correlate negatively with frailty measurements in individuals over 65 years of age (Bellizzi). Indeed, there is some evidence that genes that are differentially methylated during normal ageing and those abnormally methylated in diseased tissue show a degree of overlap, suggesting that age related changes in DNA methylation may underlie the methylation defects reported in some age-related diseases. Teschendorff *et al* showed that there was an overlap between the genes that showed age-related methylation changes and genes that were reported previously to be methylated in cancers, such as the genes *TP73* and *SFRP1*, two genes that were shown to become methylated in different types of cancers (Teschendorff). Rakyan *et al* showed a significant correlation specifically between bivalent chromatin domain DMRs that showed age-related hypermethylation and aberrantly methylated promoters in primary AML (Rakyan).

Several hypothesis can be put forward to explain the high frequency with which methylation abnormalities are seen during cancer development. The first hypothesis describes the changes in methylation levels as events that occur during the process of clonal expansion, perhaps because the diseased cells have a reduced capacity for stably maintaining their epigenome. This suggests that methylation changes are not the driving force behind the development of the disease, but rather a side effect of disease development. Many of these methylation changes would likely represent “passengers”,

with no direct role in the pathology of the disease, although some would likely provide a growth or survival advantage and thus be selected for during disease progression. A second hypothesis states that diseases develop in cells that are pre-primed by aberrant patterns of DNA methylation, similar to the epigenetic progenitor origin of human cancer, proposed by Feinberg *et al* (Feinberg, 2006). Thus altered patterns of methylation may pre-exist in a subset of apparently normal cells. If cancer driving mutations occur in a cell with a pre-existing methylation pattern, this cell can rapidly proliferate and produce diseases (Figure 2.1). This hypothesis is supported by the observation that methylation patterns are already present in apparently normal tissues from those at higher cancer risk (Belshaw, 2008) and in pre-cancerous tissues (Sproul, 2012) and by the relatively ubiquitous nature of methylation changes in tumours, suggesting that alterations in DNA methylation must occur early in disease development.

These two hypotheses are not mutually exclusive and both may have a role in cancer development.

Furthermore, as discussed above, the altered methylation found in cancer shows clear similarities to the ageing related methylation that is seen in apparently healthy cells. One potential explanation for this would be that both age-related and cancer-related DNA methylation changes might be driven by similar mechanisms, which is causing the similar patterns of altered methylation during cancer and ageing. If true, this hypothesis would suggest that the altered methylation identified in healthy individuals as they age would potentially have little or no role in the tumourigenic process. For example, the same lifestyle factors which influence the ageing process influence the risk of cancer and of other age-related diseases and modulate patterns of DNA methylation (Tapp, 2012; Mathers, 2010), but this does not have to happen in a causative manner. Alternatively, a second hypothesis would be that age-related DNA methylation may underlie the development of cancers, and possibly other age-related diseases. In this second hypothesis methylation is already present in apparently healthy individuals before the development of the disease and the disease then develops specifically or preferentially in cells with altered methylation patterns (Figure 2.1). Thus increasing accumulation of age related methylation changes would lead to increased risk of development of age related diseases. This hypothesis is supported by the observation that in CLL, global DNA methylation was shown to be relatively stable over time and similar within different CLL compartments (Cahill, 2012), suggesting that cancer cells are not necessarily highly methylation unstable and that continued methylation

instability is not necessary an intrinsic feature of cancer cells. Other evidence to support this hypothesis is that alterations in the DNA methylation levels are already observed in early stages of the disease (van Hoesel, 2013; Heuck, 2013). These observations suggest that methylation patterns must occur very early in cancer development and that cancer cells can stably maintain their methylation patterns (this section was written as part of a review (van Otterdijk, 2013)).

Therefore this project was designed to study the potential for age-related changes in DNA methylation as a causative agent in the development of age-related diseases.

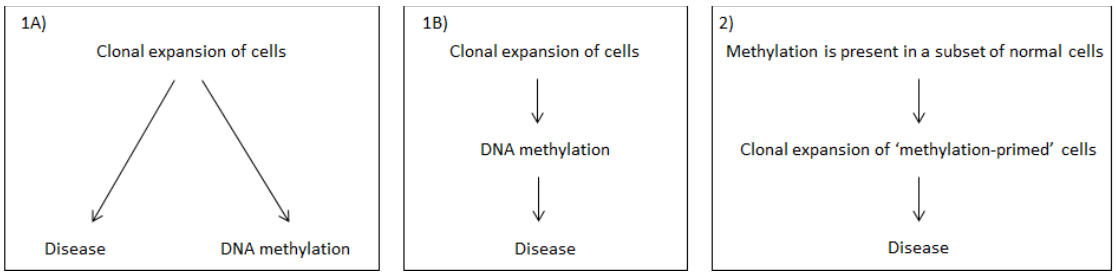


Figure 2.1: The possible role of DNA methylation during disease development

Three possible hypotheses can explain the role of DNA methylation in disease development;

1A) Methylation changes are an effect of the clonal expansion of cells prior to the disease development, but it does not play a role in the development of the disease.

1B) Methylation changes occur during the clonally expansion of cells that occur prior to disease development. The methylation events will then produce further growth advantages, promoting disease development.

2) Methylation levels are already present in a subset of normal cells. Disease develops when these 'methylation-primed' cells clonally expand.

2.2 Project aims

This project was designed to examine the link between methylation changes during normal ageing and during the development of age-related diseases. The primary focus of this study was to understand how abnormal methylation develops in and contributes to age-related diseases and to investigate the potential of altered DNA methylation in apparently normal cells as a significant factor in the development of the two most important age-related diseases, cancer (especially haematological malignancies) and heart disease. Specifically:

- To investigate whether variations in DNA methylation levels at susceptible loci are age related both in terms of overall methylation levels and in terms of variability across the population and if genes exhibiting age related methylation changes are related to genes that exhibit aberrant methylation levels in acute lymphoblastic leukaemia.
- To profile how methylation patterns are distributed among the different alleles in healthy individuals and how these patterns relate to methylation patterns observed in leukaemia patients.
- To investigate whether genes which display cell type specific patterns of leukaemia methylation exhibit similar cell type specific methylation in normal cells and when during the differentiation process altered DNA methylation occurs.
- To investigate whether methylation patterns in non-diseased cells from leukaemia patients exhibit aberrant patterns of DNA methylation.
- To profile DNA methylation from atherosclerosis and HNPCC patients to determine if altered DNA methylation is a common feature of the disease, and to determine whether these genes are related to those targeted by methylation in ageing or in leukaemia.

Chapter 3. Methods

3.1 Samples

3.1.1 Sample collection

DNA was extracted from peripheral blood leukocytes (PBL) samples of 197 healthy volunteers in 4 different age groups. The samples from the neonates (n=49) and young adults (17-42 years of age, with an average of 27.7 years, n=48) were collected as part of the North Cumbria Community Genetics Project (NCCGP) (Chase, 1998). The samples from the adults (50 years of age, n=50) were collected as part of the Newcastle Thousand families study (Pearce, 2009) and the samples from the older adults (85 years of age, n=50) were collected as part of the Newcastle 85^{+} study (Collerton, 2007). The DNA samples from the neonates and the young adults were child-mother pairs, and selected in such a way that all the samples included in this study were paired samples (Table 3.1).

Peripheral blood samples were taken from healthy young adults (24-31 years of age, with an average of 27.6, n=10) and adults (43-60 years of age, with an average of 51.6, n=8) to isolate out specific cell types. Cell types were isolated straight after the blood samples were taken. To isolate out stem cells and early progenitor cells, granulocyte-colony stimulating factor (G-CSF) mobilized CD34 peripheral blood cells were obtained from healthy young adults, via the Newcastle Haematology Biobank. These cells were stored in liquid nitrogen until further implication in medium containing 10% DMSO to prevent cell lysis and to allow cells to remain intact (Table 3.2).

PBL DNA was extracted from patients with ALL, in the age-range of 1-65 years. The samples were taken at diagnosis and, where available, remission samples (between 1 and 4 per patient taken between 4 weeks and 2 years post diagnosis) and relapse samples were also obtained. The leukaemia samples were obtained via the Newcastle Haematology Biobank. From the patients with childhood ALL (1-17 years of age, with an average age of 3.14 years, n=24), 3 subsequently relapsed but they survived the disease. In the adult ALL patients group (21-65 years of age, with an average age of 45.46 years, n=30), 14 subsequently relapsed. Six patients were long time survivors (>5 years). Two of these long term survivors first went into relapse (Table 3.3).

PBL DNA was extracted from patients with HNPCC, also known as Lynch syndrome (age 28-65 years, with an average of 49 years, n=50). They were collected as part of the Concerted Action Polyp Prevention (CAPP) study (Table 3.3).

DNA was extracted from mononuclear cells (MNC), which were obtained from patients who suffered from a heart attack and who participated in the 'Role of T-cells in Coronary Artery Disease' study (age 42-74 years, with an average age of 58.26 years, n=55) (Table 3.3).

All samples were selected at random and further patient details were not provided until after the analyses were performed

	Neonates	**Young adults**	**Adults**	**Older adults**
Age (years)	0	17 – 42 (average 27.7)	50	85
Gestational age (weeks)	35-42 (average 39.6)			
Number of participants	49	48	50	50
Extra information	Related to young adults (child).	Related to the neonates (mother). Pregnant at time of blood collection		
Cohort	NCCG Project	NCCG Project	Newcastle Thousand Families Study	Newcastle 85+ Study

Table 3.1: An overview of the age-related DNA methylation study population

	Young adults	Adults	G-CSF samples
Age (years)	24-31 (average 27.6)	43-60 (51.6 average)	24-33 (average 26.2)
DNA samples	PBL MNC Lymphocytes B-lymphocytes Monocytes	PBL MNC Lymphocytes B-lymphocytes T-lymphocytes Monocytes Granulocytes	Stem- and early progenitor cells Early B-cell progenitors Stem- and early progenitor cells minus B-cell progenitors B-lymphocytes T-lymphocytes Monocytes
Number of participants	10	8	6

Table 3.2: An overview of the cell type specific DNA methylation study population

	Childhood ALL patients	Adult ALL patients	HNPCC patients	Heart disease patients
Age (years)	1-17 (average 3.14)	21-65 (average 45.46)	28-65 (average 49)	42-74 (average 58.26)
Number of patients	24	30	50	55
Source of DNA	PBL	PBL	PBL	MNC
Cohort	Newcastle Haematology Biobank	Newcastle Haematology Biobank	CAPP study	Role of T-cells in Coronary Artery Disease

Table 3.3: An overview of the age-related diseases study population

3.1.2 DNA extraction

DNA was either extracted using a phenol/chloroform protocol, or using the Purelink^{tm} genomic DNA mini kit (Invitrogen).

DNA was extracted from the healthy individuals and the different patient groups using the phenol/chloroform protocol. 340µl of Reagent B (Appendix A, Table A.1) was added to the cell pellet and resuspended. 100µl of 5M sodium perchlorate (Appendix A, Table A.1) was added and the sample was incubated at 37 ºC for 20 minutes, followed by 65 ºC for 20 minutes. 440µl of phenol/440µl of chloroform was added to the tube. The tube was centrifuged for 10 minutes at full speed. Afterwards, the supernatant was transferred to a new tube, 880µl of chloroform was added to this tube and the tube was centrifuged for 10 minutes at full speed. Again, the supernatant was transferred to a new tube and the sample was precipitated by adding 2 volumes of cold 100% ethanol. The tube was centrifuged for 10 minutes at full speed. After washing with 70% ethanol and a 5 minute spin in the centrifuge at full speed, the pellet was dried and resuspended in the appropriate volume of dH_2O. All samples were stored at 4 ºC until further implication.

DNA was extracted from the sorted Lymphocytes, B-lymphocytes, T-lymphocytes, Monocytes, stem and progenitor cells, B-lymphocyte progenitors, and stem and progenitor cells – B-lymphocyte progenitors and the collected PBL and MNC cells from the same individuals, using the Purelink^{tm} genomic DNA mini kit (Invitrogen), according to manufactures protocol. The samples were stored at 4 ºC until further implication.

Quantitation of DNA was performed using a NanoDrop® ND-1000 spectrophotometer which measures the absorbance of UV light at 260 nm passed through a 1µl aliquot of extracted DNA, and performs the necessary calculations according to the Beer Lambert Law to provide DNA concentration (in ng/µl).

3.1.3 Cell type specific cell sorting

B-lymphocytes, T-lymphocytes, monocytes and the stem and early progenitor cells were sorted using the dynabeads method. The cells were sorted from either MNCs, which were obtained after density gradient centrifugation using Ficoll-Hypaque (Biochrom^{AG}); or from G-CSF mobilized CD34 peripheral blood cells. Neutrophils and granulocytes were isolated using density gradient centrifugation from PBL.

3.1.3.1 Isolation of mononuclear cells by gradient centrifugation

A PBL sample was taken and collected in a tube treated with ethylenediaminetetraacetic acid (EDTA tube) (Invitrogen). The blood was diluted in phosphate-buffered saline (PBS) in a 1:1 ratio. Half the amount of the PBL-PBS solution of Bicoll separating solution (Biochrom^{AG}) was added to a 50 ml tube and the PBL-PBS solution was carefully layered on top of the Bicoll. The tube was centrifuged at 700xG for 30 minutes at 4 ºC, after which five layers occurred in the falcon tube (Figure 3.1), including a layer of MNCs. The MNC interphase was carefully aspirated and transferred to a new tube. 1 ml of MNC was pelleted down and this pellet was stored at -80 ºC, to allow DNA isolation from MNCs. The rest of the MNCs were used for further sorting of specific cell types.

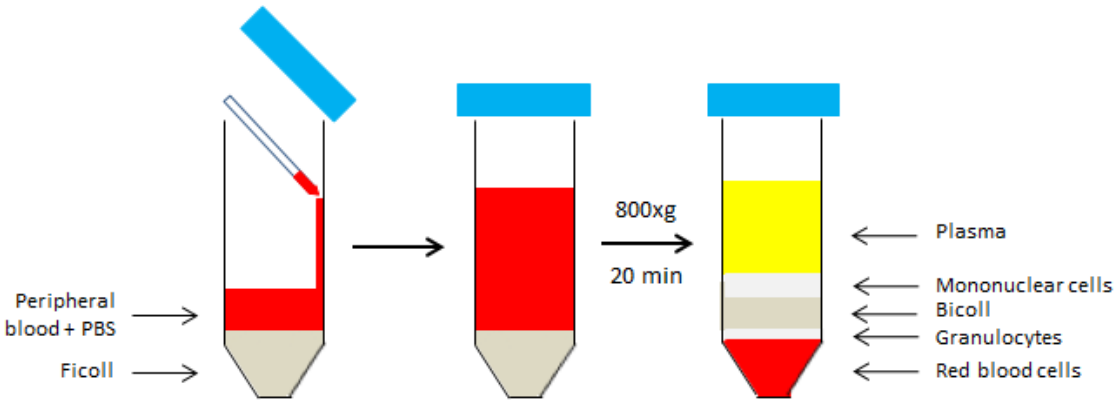


Figure 3.1: MNC cell isolation with density gradient centrifugation using Bicoll separating solution

Treating peripheral blood with a density gradient centrifugation using Bicoll separating solution resulted in different layers of cell types; plasma with thrombocytes and PBS, a layer of MNC, a layer of Bicoll, a layer of granulocytes and cell-debris and a pellet of erythrocytes.

3.1.3.2 Cell-type specific cell isolation

Different cell types were sorted from MNCs or G-CSF mobilized cells, using the Invitrogen dynabeads methods. B-lymphocytes ($CD19^+$ cells), T-lymphocytes ($CD3^+$ cells), Monocytes ($CD14^+$ cells) and stem- and progenitor cells ($CD34^+$ cells) were isolated with this magnetic bead separation methods, which depends largely on the use of antibodies directed against cell surface ligands found specifically on the cell types of interest. This procedure works in 3 steps; first, the beads will bind to the desired target cell, relying on the specific affinity of the ligand on the surface of the beads; secondly, the bead-cell complex will respond to a magnetic field, efficiently separating the bead-cell complex from the rest of the sample and lastly; the bead-cell complex is released in a suitable volume for use in downstream applications (Invitrogen, 2010) (Figure 3.2:).

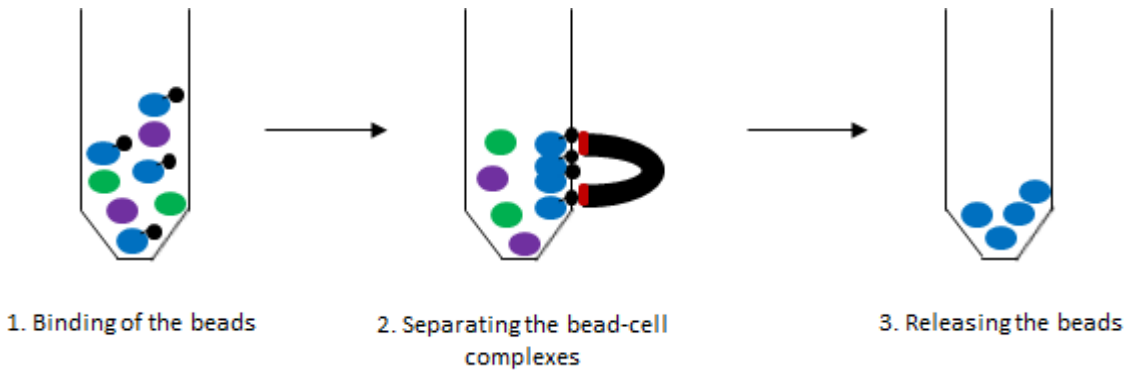


Figure 3.2: Cell sorting using the dynabeads method

The magnetic bead cell separation contains of 3 steps; the beads will bind to the desired target cells; the bead cell complex will respond to a magnetic field to separate the bead-cell complex from the rest of the cells; and afterwards, the cells can be released and used for downstream applications.

3.1.3.2.1 CD19, CD3 and CD14 isolation

The desired volume of CD19, CD14 and CD3 beads (Dynabeads CD19 pan B, Dynabeads CD14 or Dynabeads CD3, Invitrogen) was transferred into a tube (in our study 50µl of beads were used per sample). The beads were pre-washed by adding 1 ml of the cell type specific separation buffer (Appendix A, Table A.1) to the beads. The CD3 and CD14 beads required separation buffer 2, while for the CD19 isolation, separation buffer 1 was required. The tube was placed in a magnetic rack to pellet the beads, the supernatant was discarded and the washed beads were resuspended in the same volume of buffer as the initial volume of the beads.

50µl of beads were added to the MNC sample and incubated for 20 min at 4 ºC on a rotator, rotating 20 rounds per minute (rpm). Afterwards the tube was placed in a magnetic rack. The supernatant was either transferred to another tube (the supernatant from the CD14 isolation contained mainly lymphocytes and was stored) or discarded. The beads-bound cells were washed 4 times by adding 1 ml of the cell type specific separation buffer. After each wash step, the tube was placed in a magnetic rack to pellet the beads and the supernatant was discarded.

The supernatant from the CD14 isolation was pelleted down by centrifuging the sample at 500 x g for 8 minutes. All samples were stored, with the beads attached to the isolated cells, at -80 ºC until further implication.

3.1.3.2.2 CD34 isolation

The stem- and early progenitor cell were isolated during the CD34 dynabeads (Dynal CD34 progenitor cell selection system, Invitrogen). These beads could be detached after cell isolation, to allow the samples for further isolation into lineage specific progenitor cells.

The desired amount of CD34 dynabeads was transferred to a tube (in this study 200µl of beads were used per sample). The beads were pre-washed by adding 1 ml of separation buffer 1 (Appendix A, Table A.1) to the beads. The tube was placed in a magnetic rack to pellet the beads, the supernatant was discarded and the washed beads were resuspended in the same volume of separation buffer 1 as the initial volume of the beads.

The G-CSF mobilized peripheral blood cells sample was defrosted, pelleted down to discard the freezing down media and resuspended carefully in the cell type specific separation buffer to prevent cell clumps. 200µl of beads were added to the sample and vortexed for 2-3 seconds. The sample was incubated for 30 minutes at 4 ºC on a rotator,

rotating 20 rpm. Afterwards, the tube was filled with 1 ml of cold separation buffer 1 and mixed well by pipetting. The tube was placed in a magnetic rack and the supernatant discarded. The bead-bound cells were washed 3 times by adding 1.5 ml of separation buffer 1. After the last wash step, the bead-bound cells were resuspended in 150µl of separation buffer 1. 50µl of this sample was collected and stored at -80 ºC until further implication, with the beads still attached to the cells. The other 100µl was used to detach the beads from the cells, allowing further isolations into lineage specific progenitor cells. To detach the beads, the bead-bound cells were resuspended in 100µl of DETACHaBEAD (Dynal CD34 progenitor cell selection system, Invitrogen) and incubated at room temperature for 45 minutes on a rotator, rotating 20 rpm. 1.5 ml of separation buffer 1 was added to the tube and the tube was vortexed for 2 seconds to enhance detachment of the beads from the cells. The tube was placed in a magnetic rack and the supernatant, containing the detached cells, was transferred into a fresh tube. To obtain the residual cells, the beads were washed 3 times by resuspending them in 500µl of separation buffer 1. All the collected supernatants were pooled together and washed with 10 ml of separation buffer 1 and the cells were centrifuged for 10 minutes at 4 ºC at 400xG to remove excess DETACHaBEAD.

The cell pellet was resuspended in 1ml of separation buffer 1. To isolate out early B-cell progenitors, 50µl of CD19 beads (Dynabeads CD19 pan B) were added to the CD34+ cells and the sample was treated in the same fashion as described earlier in '3.1.3.2.1 CD19, CD3 and CD14 isolation'. After this CD19 isolation, the supernatant was collected and stored, as this sample contains the stem-and early progenitor cells minus the early B-cell progenitors (CD34+CD19-). All the samples were stored at -80 ºC until further implication

3.1.3.3 Isolation of granulocytes

A PBL sample was taken from a healthy individual and collected in an EDTA tube (Invitrogen). Straight after the blood samples were taken, a gradient centrifugation was performed to isolate out the granulocytes.

A red blood cell lysis was performed prior to the gradient centrifugation. The PBL sample was diluted 2:1 with a 6% Dextran solution (Dextran solution, Sigma) and mixed by inverting the tube 15 times. The PBL-Dextran mix was incubated at room temperature for 45 minutes, to allow the red blood cells to sediment. The supernatant with the leukocyte and platelet cell suspension was collected and transferred into a new falcon tube. This tube was centrifuged for 10 minutes at 300xG at 4 ºC to pellet the

leukocytes, but leaving the platelets in suspension. The leukocyte pellet was collected and resuspended in 1 ml of 1x PBS and 9 ml of dH_2O. This mixture was incubated for 15 seconds and then 1 ml of 10x PBS was added. The tube was centrifuged at 300xG for 10 minutes at 4 ºC, which resulted in a milky pellet formation at the bottom of the tube containing the leukocytes. This pellet was resuspended in 3 ml of 1x PBS.
3 ml of Bicoll separating solution (Biochrom^{AG}) was added to a new tube, and the PBS-cell mixture was carefully layered on top of the Bicoll. The tube was centrifuged at 700xG for 30 minutes at 4 ºC. The centrifugation step resulted in the formation of a pellet containing the granulocytes. This pellet was washed and stored at -80 ºC until further implication.

3.1.3.4 Validation of cell isolations

The cell isolations were validated by assessing the expression of cell surface markers specific for the individual cell types in the various isolated cell populations, using either a quantitative Polymerase Chain Reaction (qPCR) to assess messenger RNA (mRNA) expression levels or by flow cytometry using antibodies directed against the cell surface markers.

3.1.3.4.1 Validation of cell isolations by qPCR

To validate the cell isolation for the T-lymphocytes, B-lymphocytes and monocytes, RNA was extracted, from the different cell types. Complementary DNA (cDNA) was synthesised and expression of cell surface markers was measured using a qPCR. This method was used to validate the cell isolations, because the CD19, CD3 and CD14 beads used in this study could not be detached from the cells after isolation. These beads are able to respond to the different antibodies used for flow cytometry, which are supposed to be directed against specific cell surface makers. This makes it not possible to perform the validation of the isolation of these specific cell types by flow cytometry and as a result, the validations of these cell isolations were performed using the qPCR technique.

3.1.3.4.1.1 RNA isolation

The beads-bound cell pellet was resuspended in 1 ml of Trizol (Invitrogen) and incubated for 5 minutes at 30 ºC. 0.2 ml of chloroform was added to the sample, shaken by inverting the tube by hand for 15 seconds and incubated for 3 minutes at 30 ºC. After incubation, the tube centrifuged at 12,000xG for 15 minutes at 4 ºC**.** The aqueous phase

was transferred to a fresh tube and 0.5 ml of isopropanol was added to the sample. The sample was incubated for 10 minutes at 30 ºC and centrifuged at 4 ºC, at 12,000xG for 20 minutes. The isopropanol was removed and the pellet was washed by adding 1 ml of 75% ethanol to the cell pellet. The sample was vortexed and centrifuged at 7,500xG for 5 minutes at 4 ºC. The ethanol was removed and the pellet air dried for 10 minutes. 20µl of dH_2O, free of desoxyribonuclease (DNAse) and ribonuclease (RNAse), was added to the pellet followed by 10 minutes incubation at 57 ºC. After this incubation, the RNA sample was put on ice directly and stored at -80 ºC.

3.1.3.4.1.2 cDNA extraction

cDNA extraction was performed using the Superscript® III Cells Direct cDNA Synthesis System kit (Invitrogen), following the manufacturing protocol: between 0.62 and 2.33µg of total RNA, 1µl OLIGODT (500 µg/µl) and 1µl dNTP (10 mM) were mixed in a PCR tube, and filled up to a total of 12µl with dH_2O. This mixture was incubated for 5 minutes at 65 ºC. 4µl 5x first strand buffer (5xRT), 2µl DTT (0.1M) and 1µl of RNAse out (40 units/µl) was added to the tube and mixed well by pipetting. This mixture was incubated for 2 minutes at 42 ºC and 1µl of superscript III RT (200 units/µl) was added to mixture and again incubated for 50 minutes at 42 ºC. The reaction was inactivated by incubating the sample for 15 minutes at 70 ºC. 1µl of RNAh (40 units/µl) was added and again incubated for 20 minutes at 37 ºC. After this incubation step, the cDNA sample was stored at -20 ºC until further implication.

3.1.3.4.1.3 qPCR

During a qPCR, the progress of the Polymerase Chain Reaction (PCR) can be monitored as it occurs and data is collected throughout the PCR process. The qPCR technique is similar to traditional PCR, with the major difference being that with qPCR the amount of PCR product is measured after each round of amplification while with traditional PCR, the amount of PCR product is only measured at the end point of amplification. The concept of the qPCR is that amplification products are measured as they are produced using a fluorescent label. The fluorescence signal is directly proportional to DNA concentration over a broad range and the linear correlations between PCR product and fluorescence intensity is used to calculate the amount of template present at the beginning of the reaction (Sigma-Aldrich, 2008).

This technique allowed us to test the expression of the different cell surface markers in our sorted cell types. To validate the cell sort, the isolated cell types were

tested for expression of the cell-surface markers; CD19 (B lymphocytes); CD3 (T lymphocytes); CD14 (Monocytes); and GAPDH (control) (Table 3.4).

All qPCR amplifications were carried out in a volume of 10µl, using 0.5µl of cDNA, 150 ng of both the forward and the reverse primer (|Eurofins MWG Operon)(Table 3.4), 5µl SYBRgreen (Sigma-Aldrich) and 4.25µl dH_2O. All samples were done in triplicates.

PCR was performed with thermocycling conditions of 15 min on 95 ºC, followed by 40 cycles of 30 sec on 95 ºC, 30 seconds on the gene specific annealing temperature (Table 3.4), 30 seconds on 72 ºC and 30 seconds on the gene specific reading temperature (Table 3.4), followed by 30 seconds on 95 ºC, 15 seconds on 76 ºC and 15 seconds on 95 ºC (Taqman 7900, Applied Biosystems). Specific cell lines, each known to express one of the individual cell surface markers, were used as controls and one of the positive cell lines was used to produce a standard curve. These cell lines were; HL60 (a promyelocytic cell line), Molt4 (a T lymphoblastic cell line), Lama84 (a chronic myelocytic cell line), Ky01 (a chronic myelocytic cell line) and Nalm6 (a pre B lymphocyte cell line).

	Forward primer	**Reverse primer**	**Annealing temperature (◦C)**	**Reading temperature (◦C)**
CD3	5'-CTTCTCTCGCAAG TGAGC-3'	5'-CAGACATTACAAG ACTGGACC-3'	61	78
CD14	5'-GAAGACTTATCG ACCATGGAG-3'	5'-GCAACTTCTCCGA ACCTCA-3'	55	82
CD19	5'-ACTGCTCGGCCAG TACTATG-3'	5-'GAAAGCGAATGA CTGACCCC-3'	63	80
GAPDH	5′-GTCAAGCTCATT TCCTGGTATG-3′	5′-GTCTACATGGCAA CTGTGAG-3′	57	80

Table 3.4: Primers used for qPCR

3.1.3.4.2 Validation of cell isolation by flow cytometry

The stem and progenitor cell isolation was validated by flow cytometry (FACSCanto, BD Bioscience). Flow cytometry allows us to simultaneously measure and analyse multiple physical characteristics of single cells, as they flow through a beam of light. An antibody, specific for the specific cell type and with a known fluorescence marker, is added to the cells before the cells run through the flow cytometer. This will distinguish the cell surface marker positive cells, from the negative cells.

To validate our cell sort for CD34+ cells, a CD34 antibody with the PE fluorescence marker (BD Biosciences) was used. 10,000 events of live cells were analysed for the marker specificity. When the sorted samples consisted of a minimum of 90 % positive cells, the sort was marked as successful and DNA was isolated from the samples in that specific sort.

3.2 Methylation analysis

Several methods were used to analyse the methylation levels in our sample set; pyosequencing, a method to accurately measure methylation levels in a small number of CpG sites per individual; and bisulphite sequencing, to investigate the distribution of methylation within an individual.

3.2.1 Preparing for methylation analysis

Prior to the methylation analyses, bisulphite modification had to be performed to distinguish between the methylated and unmethylated cytosines, gene specific primers had to be designed and PCRs had to be performed.

3.2.1.1 Sodium bisulphite modification

The bisulphite-based DNA modification is used to discriminate between cytosine and methylated cytosine. DNA is treated with bisulphite salt to convert cytosine residues to uracil, while methylated cytosines are protected from modification and thus remain as cytosine (Figure 3.3). This converts the methylation difference into a sequence difference, which can then be assessed by various PCR based methods or directly assessed by sequencing.

200ng of DNA was modified using the Methylation^{TM} One-Step DNA Modification Kit (Epigentek), following the manufacturing protocol; 200ng of DNA was added to a PCR tube and mixed with 110µl DNA modification solution. The tubes

were placed in a PCR machine with thermocycling conditions of 99°C for 6 minutes, followed by 65°C for 90 minutes. 300µl of modified DNA capture solution was added to a spin column together with the sample from the previous step. The column was centrifuged for 20 seconds at 12,000 rpm and the flowthrough was discarded. The columns were washed 3 times; the first washing step was conducted by adding 200µl of modified cleaning solution to the column, followed by two washing round with 200µl of 90% ethanol. After each round, the tube was centrifuged for 20 seconds at 12,000 rpm and the flow through was discarded. After these wash steps, all samples were eluted in 15µl of elution buffer and centrifuged for 20 seconds at 12,000 rpm. After the modification, the samples were stored at -20 ºC.

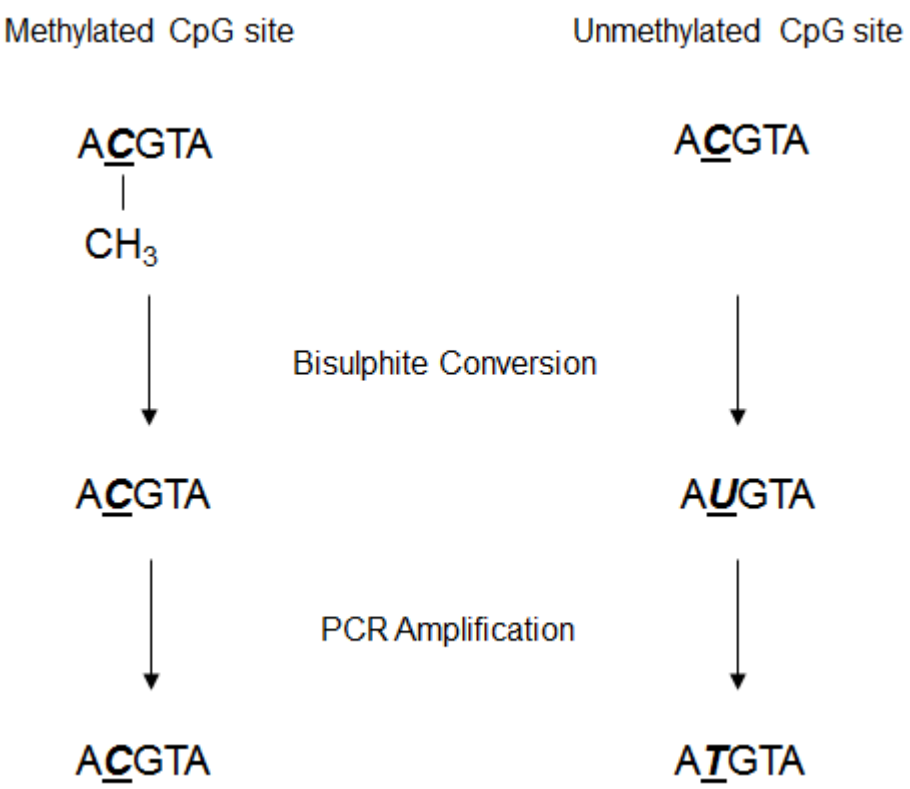


Figure 3.3: Bisulphite modification of DNA

DNA was treated with bisulphite salt. The methylated cytosines remain cytosines, while the unmethylated cytosine residues are converted to uracil in a single-stranded DNA. After PCR amplification, uracil is converted in tymine.

3.2.1.2 Primer design

Gene sequences were obtained from GenBank entry on NCBI and the region surrounding the TSS was examined for the presence of CpG islands, by searching the sequence for high numbers of CpG sites. Primers, specific for sodium bisulphite modified DNA were designed, using the primer designing programme 'PSQ assay' (Biotage). This designed forward and reverse primers for initial amplification and a sequencing primer for use in the final pyrosequencing reaction (Table 3.5). The sequencing primer was designed to be placed between the forward and the reverse primers. One of these primers was designed with a biotin label to allow strand separation during pyrosequencing. All primers were designed as such that all primers had penalty levels lower than 50, indicated by the PSQ assay software, for potential mispriming sites for all the different primers, potential homopolymer sequences adjacent to the target sites, the presence of possible template loops at the amplicon, etc. (for a full list of genes examined in this study, see Appendix A, Table A.2).

	Forward primer	Reverse primer	Sequencing primer	CpG sites	Distance from TSS (bp)
TWIST2	5‘-AACAACTATTTAACAACCCAA CCCAAC-3’	5’-GGGYGAGTTGGAGTTTTTTTT TATGG-3’	5’-CAAAACCTTTCCAACAAC-3’	4	-26 → +208
TUSC3	5’-GAATAGGATGTTTTGTTAGTT-3’	5’-TAACTAAACCCTCCCAAATA-3’	5’-CCAAATAACAACCACTTTAC-3’	6	+25 → +116
HOXD4	5’-GAAATTAATGGTTATGAGT-3’	5’-CCCRAAACTAAAAATCTA-3’	5-‘ATTTGTAGGGYGGTTATTTAGG YGAG-3’	5	+242 → +398
EphA10	5’-ATTATTTATGGGAATTAATTAT TAG-3’	5’-CCCACCAATTAAATATTC-3’	5’-TATTTATGGGAATTAATTATTAG GTA-3’	3	+28 → +143
HAND2	5’-CCAAATTTTAATTATCTTATA-3’	5’-GGATTTTTTAGTAAGATTT-3’	5’-ATTTTATTACCCAAAACTC-3’	4	-29 → +185
HOXA5	5’-CCACAAATCAAACACACATATC-3’	5’-GGTTGGTTGTATTTGGGT-3’	5’-CTACRACTACAATAACATAAATCT-3’	5	+30 → +327
HOXA4	5’-TACACTTCACAAATTAATAACCAT AAACTC-3’	5’-GTTGTTGTAGYGGTAGGTGTTG-3’	5’-AACCCAAATTCCCTCCCTT-3’	4	+31 → +219
IGF2	5’-GGGTTAAGGTAGTTTTTTTGGGA ATG-3’	5’-AAATATAAAAACCTCCTCCACC TCC-3’	5’-AGGGGGTTTATTTTTTTAGGAA-3’	3	DMR0
MLH1	5’-AGTTTTTTTTTTAGGAGTGAAGG-3’	5’-ATAAAACCCTATACCTAATCT ATC-3’	5’-GTAGTATTYGTGTTTAGTTT-3’	6	-569 → -375

Table 3.5: The primers used for the pyrosequencing assays

All primers were designed around the TSS of the gene. The primers for *IGF2* were designed for the DMR0, upstream of exon 2.

3.2.1.3 PCR

All PCR amplifications were carried out in a volume of 25µl, using 1µl of modified DNA (~13.3µg), 2.5µl of manufacturers' 10x buffer (Roche), 0.5µl of dNTPs (20 mM, Roche), 0.2µl of FastStart taq DNA polymerase (5 U/µl, Roche), 1-4mM of $MgCl_2$ Roche) and 150ng of both the forward and the reverse primer (Eurofins MWG Operon) (Table 3.5 and Table 3.6). PCR was performed with thermocycling conditions of one cycle at 95°C for 5 minutes, followed by 40-50 cycles of 95°C for 30 seconds, primer set specific annealing temperature for 30 seconds and 72°C for 30 seconds, followed by a single cycle of 72°C for 5 minutes.

	$MgCl_2$ concentration (mM)	**Annealing temperature (°C)**	**Cycle numbers**
TWIST2	2	63	40
TUSC3	4	55	40
HOXD4	3.5	51	40
EphA10	4	53	45
HAND2	3.5	51	45
HOXA5	2.5	58	50
HOXA4	3	63	40
IGF2	3	63	50
MLH1	4	58	40

Table 3.6: PCR conditions for the pyrosequencing assays

3.2.1.4 Agarose gel electophoresis

PCR amplification was checked prior to pyrosequencing analysis by agarose electrophoresis. An agarose gel was prepared, containing 1.5 mg of agarose (Alpha laboratories) in 150 ml of 1x TAE buffer (Appendix A, Table A.1). 5µl of GelRed™ (Biotium) staining was added for visualisation. 4µl of PCR product was added to each of the slots in the gel, and ran on the gel using electrophoresis on 150 volt for 45 minutes. Gels were analysed using an the G:box imager (Syngene).

3.2.2 Pyrosequencing

The sodium bisulphite modification converted the methylation difference in a sequencing difference. The amount of methylation can now be detected by DNA sequencing.

Pyrosequencing is a DNA sequencing technique that is based on the sequencing by synthesis approach. Quantification of the amount of a specific base can by achieved by an enzymic cascade to measure the release of pyrophosphate (PPi) during DNA synthesis. As a result this system can accurately quantify the amount of specific bases added at a polymorphic site, such as that produced at a CpG site following bisulfite modification.

4 enzymes are included in the pyrosequencing system to allow quantification of the released pyrophosphate; DNA polymerase 1, Adenosine triphosphate (ATP) sulfurylase, luciferase and apyrase (illustrated in Figure 3.4A). The reaction mixture also contains the enzyme substrates adenosine phosphosulfate (APS), d-luciferin and the sequencing template with an annealed sequencing primer to be used as a starting material for the DNA polymerase. During the cascade of enzymatic reactions, visible light is generated that is proportional to the number of incorporated nucleotides. The cascade starts with a nucleic acid polymerisation reaction in which PPi is released from a nucleotide triphosphate, following incorporation into the DNA strand by polymerase. The released PPi is converted to ATP by ATP sulfurylase, which provides the energy to luciferase to oxidise luciferin and generate light. The four different nucleotides are added stepwise, in a cyclic manner, and incorporation is followed using the enzyme ATP sulfurylase and luciferase. The light that is produced will be detected by a charge coupled device (CCD) camera. Apyrase removes unincorporated nucleotides and ATP between the additions of different bases.

This degradation between base additions is crucial for synchronized DNA synthesis asserting that the light signal detected when adding a certain nucleotide only arises from

incorporation of that specific nucleotide (Figure 3.4A). The results are displayed as a pyrogram (Figure 3.4B). The height of the curve is determined by the signal, caused by the activity of luciferase (Ahmadian, 2006; Ronaghi, 2001).

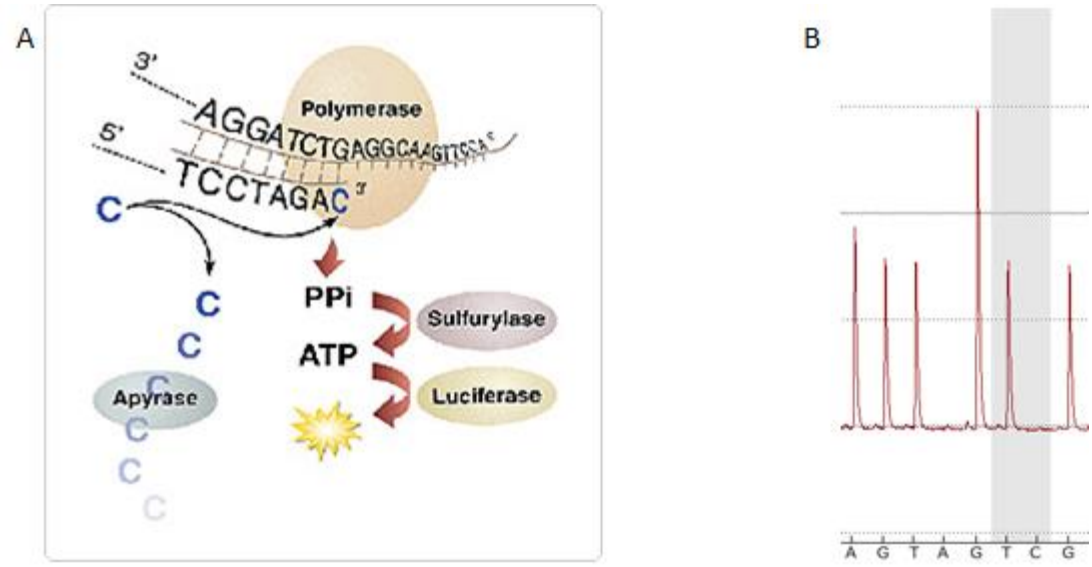


Figure 3.4: Pyrosequencing

A) The nucleotide composition of a DNA-strand is determined by an enzyme cascade system. The single stranded DNA template is first hybridised with the sequencing primer and mixed with the enzymes along with the two substrates APS and luciferin. One of the four nucleotides is then added to the reaction. If the nucleotide is complementary to the base in the template strand, DNA polymerase incorporates it into the growing strand. PPi is released and converted to ATP by sulfurylase in the presence of APS.

Detection is based on the amount of visible light produced during the following ATP-mediated conversion of luciferin to oxyluciferin by the luciferase. Any unincorporated nucleotides are degraded by apyrase, before the next nucleotide is added to the tube and the cascade of reactions repeated (Invitrogen).

B) The light signals are detected and converted into a pyrogram, where the height of the peak is related to one of more identical bases.

Each individual pyrosequencing primer set was validated, by testing CpGenome™ Universal Methylated Control DNA (Millipore) (100% methylated) diluted to different levels of DNA methylation to produce a range of approximately 0-100% methylated DNA. A peripheral blood sample from a young volunteer was used as a negative control of around 2-5 % methylated DNA. This was done to ensure that the level that the pyrosequencer **(**Pyromark Q96 MD pyrosequencer) is measuring reflects the amount of methylated DNA in the sample (5).

For assessing methylation levels in individual samples, 10µl of PCR product was used for each pyrosequencing reaction and 5pmol of sequencing annealing primer (Table 3.5). The pyrosequencer has an internal quality control (QC) and identifies the samples for which the methylation levels were unreliable. All the samples with warnings, or the samples that passed the internal QC but showed low pyrosequencing peak heights, were repeated. All PCRs were performed in duplicate, in such a way that each sample had repeats that both passed the internal QC. If the difference in methylation levels in a sample between the 2 different runs was less than 2%, the average was calculated and used for further analysis. If the difference in methylation levels was greater than 2%, the sample was tested another time. For samples and genes with high levels of methylation (above 40%), a cut off of 5% difference in methylation levels was used.

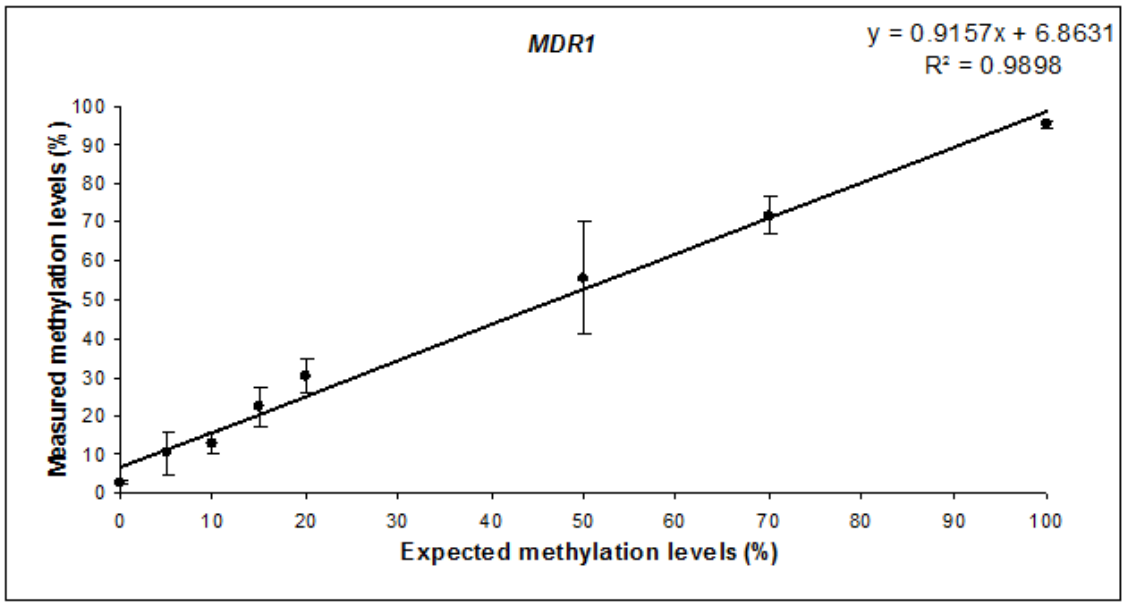


Figure 3.5: Validation of the pyrosequencing primers

To validate the pyrosequencing assay, samples with known levels of methylation (~0-100%) were measured. The average of the CpG sites measured during pyrosequencing was calculated and plotted against the expected methylation values. For the unmethylated control (~0%), a PBL samples was used with methylation levels around 2.7%. The test was performed in duplicates and resulted in a line with an R^2 value, indicating the degree of overlap between the measured and expected methylation levels.

In this graph, *MDR1* is shown as example of one of the genes for which the pyrosequencing assay is validated.

Validation of all the different genes under study was done in collaboration with Hannah Gautrey (Gautrey, 2012).

3.2.3 Bisulphite sequencing

To examine the allelic distribution of identified DNA methylation in greater detail, bisulphite sequencing was performed for the *TUSC3* gene on two ALL remission samples. The sequencing allowed us to measure the methylation levels of 7 CpG sites per allele. 11 alleles were sequenced for both samples.

3.2.3.1 Preparing clones for bisuphite sequencing

To prepare the clones for bisulphite sequencing, the *TUSC3* PCR was performed (as described in 3.2.1.3). This PCR product was cloned into a pCR2.1 vector by mixing 1µl of the PCR product with 1µl of 10x ligation buffer, 2µl of the pCR 2.1 vector (25ng/µl), 1µl T4 DNA ligase (5.0 Weiss units/µl) and 5.5µl of dH_2O (TA Cloning Kit, Invitrogen). This ligation mixture was incubated overnight at 14 °C.

One vial of 50µl frozen One Shot competent cells (Invitrogen) per transformation were thawed on ice. 2µl of ligation product was added into the vial of competent cells and mixed by gently stirring with a pipette tip. The vial was incubated on ice for 30 minutes and afterwards the cells were put in a 42 °C water bath to perform a heat shock. Immediately after the heat shock, the vial was put back on ice. 250µl of S.O.C medium (Invitrogen) was added to the vial of cells. The vials were shaken at 225 rpm in a shaking incubator at 37 °C for 1 hour. Afterwards, the cells were spreaded over an Agar plate (Appendix A, Table A.1) and incubated overnight at 37 °C.

3.2.3.2 Performing minipreps

15 individual colonies were picked for analyses per plate. These colonies were added to a tube containing 5 ml of LB media, containing 50µg/ml kanamycin (Appendix A, Table A.1) and incubated overnight at 37 °C.

The tube was removed from the incubator and centrifuged at 5000 rpm for 10 minutes at 4 °C. A miniprep was performed using the Qiaprep Spin Miniprep kit (Qiagen), according to manufactures protocol; the media was removed and the pelleted cells were resuspended in 250µl of buffer P1 and transferred into an eppendorf tube. 250µl of buffer P2 was added and mixed thoroughly by inverting the tube 4-6 times. 350µl buffer N3 was added and mixed immediately by inverting the tube. The tube was centrifuged for 10 minutes at 13,0000 rpm. The supernatant was transferred into the QIAprep spin column and centrifuged for 1 minute at 13,000 rpm. The spin column was washed by adding 500µl of buffer PB and centrifuged for 1 minute at 13,000 rpm. Afterwards, the column was washed another time by adding 750µl of buffer PE and centrifuged for 1

minute at 13,000 rpm. The flow trough was discarded and the column was centrifuged another time for 1 minute at 13,000 rpm to remove any residual wash buffer. Afterwards, the column was placed in a eppendorf tube, 50µl of elution buffer ED was added and the tube was centrifuged for 1 minute at 13,000 rpm.

3.2.3.3 Analysing the transformants

The product was checked for the insertion by performing an EcoR1 restriction digest. 2µl of Buffer H, 2µl of 10x BSA, 0.5µl of EcoR1 enzyme, 5.5µl of dH_2O and 10µl of the Miniprep DNA was mixed and incubated at 37 °C for 2 hours. The digestion product was ran on a 2% agarose gel (as described in 3.2.1.4). The sample was quantified using the NanoDrop® ND-1000 spectrophotometer (in ng/µl). For the samples that had the PCR product inserted, 100µl per sample was send away for sequencing.

3.3 Statistical analysis

Methylation levels were measured for a set of genes by pyrosequencing. For each of these genes, between 3 and 6 CpG sites in or near the TSS were examined. The statistical analyses were performed on the average levels of methylation of these CpG sites per gene. The methylation levels varied between the different CpG sites, but this variation was relatively consistent between individuals (Figure 3.6).

Our data set consisted of different cohorts of patients and controls. Each cohort was examined individually for normal distribution using the Robvar regression, as well as the equality in variance using the Levene's test for equal variances (STATA). Depending on those results, either parametric- or non-parametric statistical tests (SPSS) were performed. For the unrelated samples, the parametric tests performed were the one way repeated ANOVA, to examine overall methylation differences, and the Independent samples T-test, to examine differences between groups. The non-parametric equivalents for these tests for the unrelated samples were the Kruskal-Wallis test, to examine overall methylation differences, and the Wilcoxon rank sum tests, to examine differences between groups. For the related samples, the parametric tests performed were the one-way repeated ANOVA, to examine overall methylation differences, and a Paired samples T-test, to examine differences between groups. The non-parametric equivalent for these related samples were the Friedman test, to examine

for overall methylation differences, and a Wilcoxon signed rank test, to examine differences between groups.

Correlations between genes were investigated using the Pearson's correlation (SPSS). Differences were called significant when a p value smaller than 0.05 was measured.

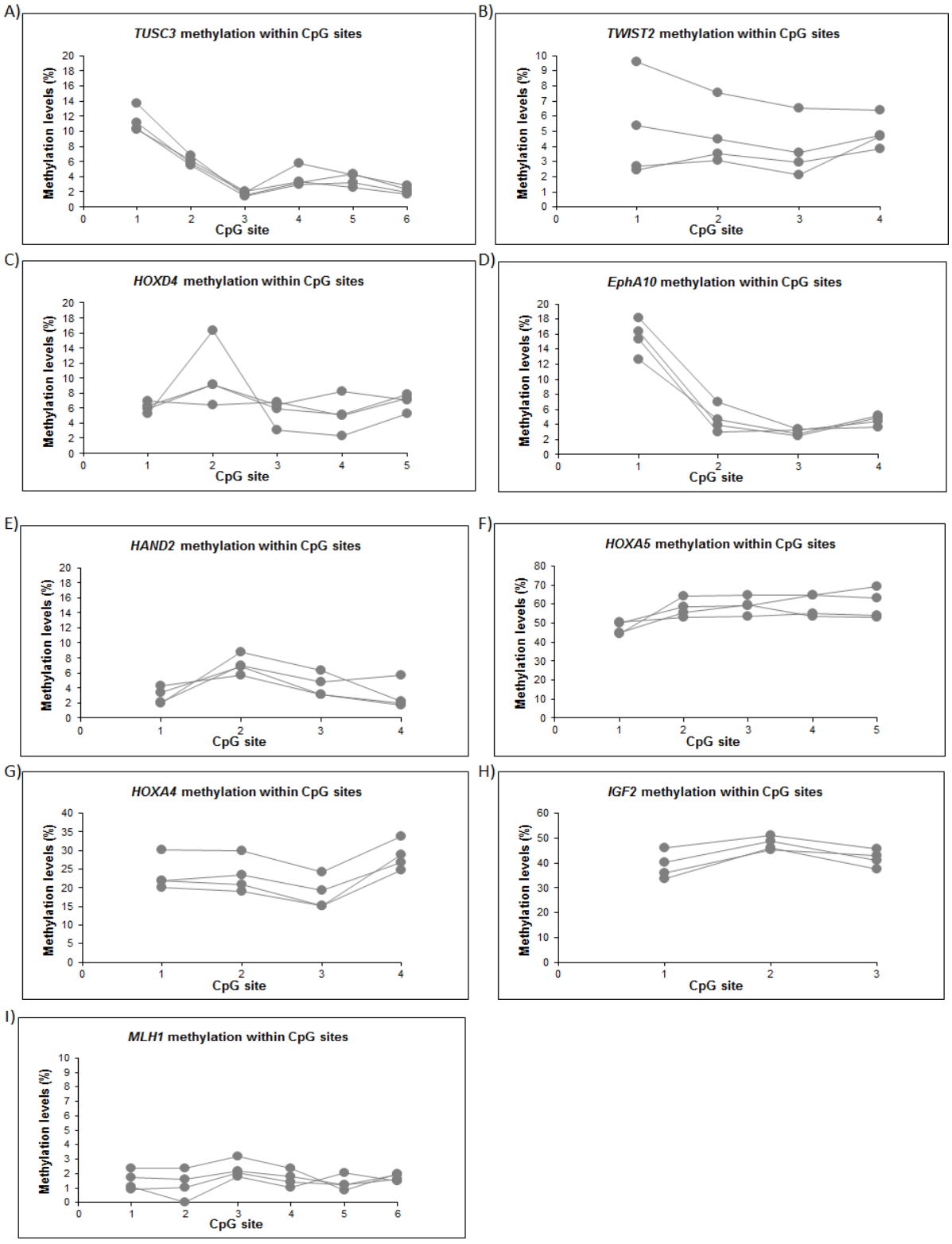


Figure 3.6: Methylation levels across the different CpG sites

Methylation levels per CpG site are shown for four neonate samples, to show overall methylation patterns across the different CpG sites around the TSS from each individual gene. Variation existed across the different CpG sites, however the methylation patterns were very similar within individuals.

Chapter 4. Alterations of DNA methylation patterns with age

4.1 Introduction

Alterations in the patterns of DNA methylation are one of the hallmarks of ageing (Mathers, 2006; Lopez-Otin, 2013). These changes include reduced levels of global DNA methylation across the genome (Bollati, 2009), in conjunction with local areas of hypermethylation, often centring on promoter associated CpG islands (Fraga, 2007). The acquisition of aberrant patterns of methylation with increasing age could play a role in differential susceptibility to develop several age-related diseases. However, the extent and the variance in CpG island methylation across the population is not yet well understood.

Therefore, this study was designed to examine the patterns of DNA methylation in peripheral blood during ageing. Both the change in extent of methylation and the variance in methylation were examined. The genes of interest included seven genes previously described to become methylated during leukaemia development; *TUSC3*, *TWIST2*, *HOXD4*, *EphA10*, *HAND2*, *HOXA5* and *HOXA4* from which the last two genes showed age-related methylation changes in other tissues than blood, and two genes that were not reported to be involved in either leukaemia development or to be methylated during ageing in other tissue types, namely the imprinted gene *IGF2* and *MLH1*, a gene involved in the DNA mismatch repair pathway (for more information about the genes under study, see Appendix A, Table A.2).

4.2 Samples

Peripheral blood leukocyte, or cord blood, DNA was collected from healthy volunteers in different age-groups; neonates (n=49), young adults (n=48), adults (n=50) and older adults (n=50) (Table 4.1). All the samples were collected as part of the NCCGP, The Newcastle Thousand Families Study and the Newcastle 85^{+} study. The analysis of the methylation levels in the Newcastle 85+ population was performed by Hannah Gautrey (Gautrey, 2012).

	Neonates	Young adults	Adults	Older adults
Age (years)	0	17 – 42 (average 27.7)	50	85
Gestational age (weeks)	35-42 (average 39.6)			
Number of participants	49	48	50	50
Gender	33 male 15 female 1 unknown	48 female	18 male 32 female	23 male 27 female
Smoking habits		19 never smoked 2 ex-smokers 4 smokers 23 unknown	19 never smoked 14 ex-smokers 16 smokers 1 unknown	16 never smoked 31 ex-smokers 3 smokers
Extra information	Related to young adults (child)	Related to the neonates (mother) Pregnant at time of blood collection		
Cohort	North Cumbria Community Genetics Project	North Cumbria Community Genetics Project	Newcastle Thousand Families Study	Newcastle 85+ Study

Table 4.1: The age-related methylation study population

The study population consisted of 197 participants in 4 different age groups; neonates, young adults, adults and the older adults.

4.3 Results

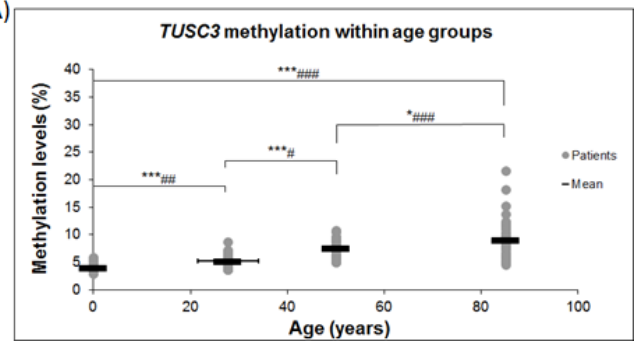
4.3.1 Inter-individual differences in methylation levels within age groups and differences in methylation between age groups

To examine changes in methylation patterns during normal ageing, methylation levels were analysed from 197 individuals in four age groups, between 0 and 85 years of age. Methylation levels for seven of the genes under study increased significantly with an increasing age; *TUSC3*, *TWIST2*, *HOXD4*, *EphA10*, *HAND2*, *HOXA4* and *HOXA5*. For five of the genes under study; *TUSC3*, *TWIST2*, *HOXD4*, *EphA10* and *HAND2* the extent of methylation increased steadily with increasing age. However, although the age-related changes in methylation levels for these five genes were relatively similar, not all between age group differences were statistically significant. *TUSC3*, *HOXD4* and *EphA10* showed significant increases in methylation levels between all 4 age groups ($p<0.05$ for *TUSC3*, $p\leq0.006$ for *HOXD4* and $p\leq0.004$ for *EphA10*). For *TWIST2* no significant difference was observed between the neonates and the young adults, but methylation levels were significantly increasing between the three older age groups ($p<0.001$). *HAND2* showed a significant increase in methylation between the 3 youngest age groups ($p<0.001$), but there were no significant differences observed in methylation levels between the adults and the older adults (Figure 4.1A-E). The age-related methylation differences were observed in the individual CpG sites from the genes analysed, showing that these patterns of age-related methylation are not relying on a subset of CpG sites per gene, but that is was a more gene-specific phenomenon (Figure 4.2 and Appendix B, Table B.1).

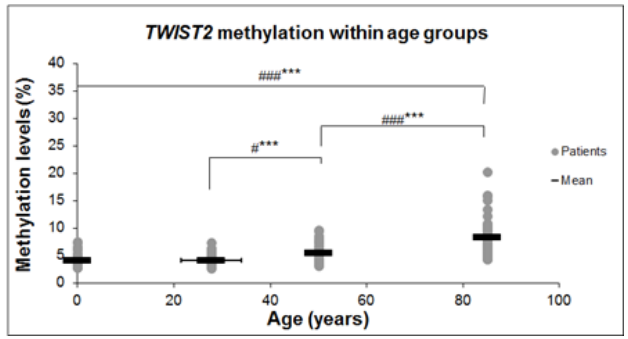
In contrast, *HOXA4*, *HOXA5* and *IGF2* do not show a steady increase in methylation levels across the different age groups and the methylation levels for these three genes were generally more stable than for the other five genes. Compared to these other five genes, *HOXA4*, *HOXA5* and *IGF2* had much higher levels of methylation at all ages. Methylation of *HOXA5* and *IGF2* were significantly higher for the neonates than for the young adults ($p=0.026$ for *HOXA5* and $p=0.003$ for *IGF2*), while methylation levels did not change significantly at older ages. Methylation levels for *HOXA4* were significantly lower in the neonates compared to the young adults ($p<0.001$), while methylation levels were stable among the other age groups (Figure 4.1F-H).

MLH1 was the only gene under study that did not show any significant differences in methylation levels between the different age groups (Figure 4.1I)

A)



B)



C)

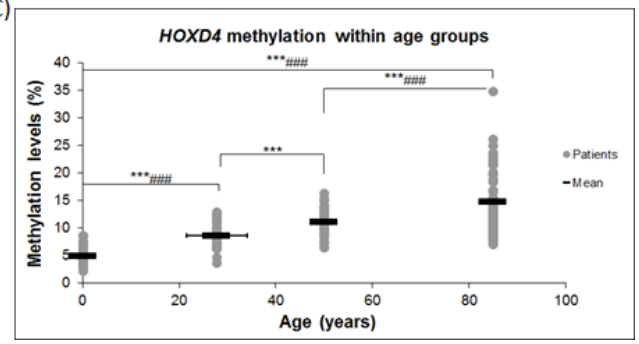


Figure 4.1A-C: Patterns of DNA methylation in different age groups

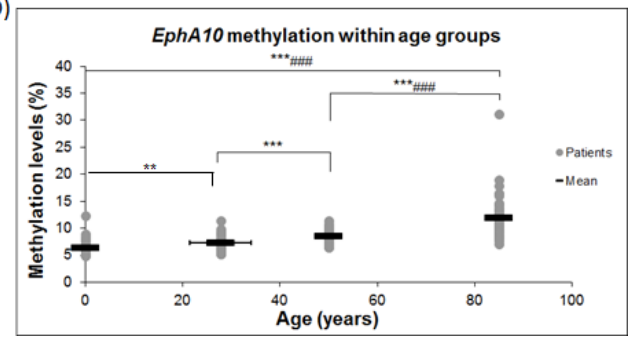
A) For *TUSC3*, the extent of methylation was increasing significantly with an increasing age ($p<0.001$, Kruskal Wallis, SPSS) and within the different age groups ($p<0.001$ between the younger three age groups, $p=0.046$ between the adults and the older adults, Wilcoxon signed rank test and Wilcoxon rank sum test (SPSS)). The variance of methylation was increasing significantly with an increasing age ($p<0.001$) and between the different age groups ($p=0.008$, $p=0.027$ and $p<0.001$, respectively between the four different age groups, Levene's test for equal variances (STATA)).

B) For *TWIST2*, both the extent of methylation and the variance of methylation were increasing significantly with an increasing age ($p<0.001$, Kruskal Wallis (SPSS) and Levene's test for equal variances (STATA)) and between the three older age groups ($p<0.001$ for the extent of methylation between the three older age groups, Wilcoxon rank sum test (SPSS) and $p=0.015$ and $p<0.001$ between the young adults and the adults and the adult and the elderly respectively, for the variance of methylation, Levene's test for equal variances (STATA)).

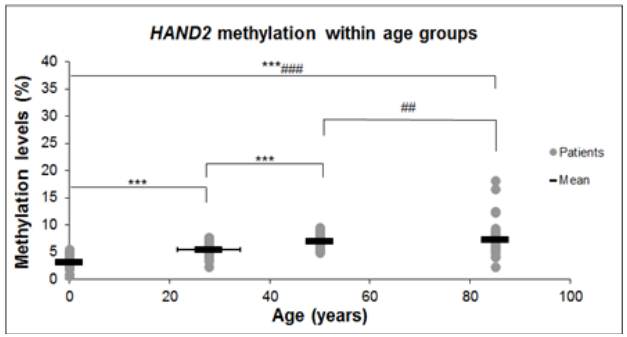
C) For *HOXD4*, both the extent of methylation and the variance of methylation were increasing significantly with an increasing age ($p<0.001$, Kruskal Wallis (SPSS) and Levene's test for equal variances (STATA)). The extent of methylation was increasing significantly between the four different age groups ($p<0.001$ between the three youngest age groups, $p=0.006$ between the adults and the older adults, Wilcoxon signed rank test and Wilcoxon rank sum test (SPSS)). The variance of methylation increased significantly between the neonates and the young adults and between the adults and the older adults ($p<0.001$, Levene's test for equal variances (STATA)).

The analysis of the methylation levels in the Newcastle 85+ population was performed by Hannah Gautrey

D)



E)



F)

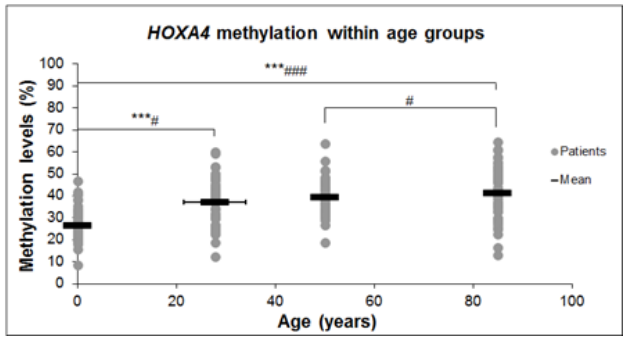


Figure 4.1D-F: Patterns of DNA methylation in different age groups

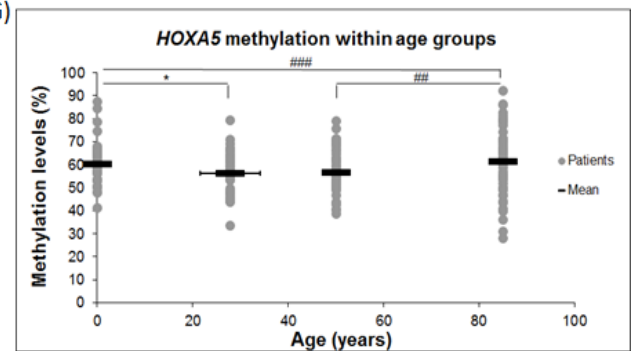
D) For *Epha10*, both the extent of methylation and the variance of methylation were increasing significantly with an increasing age ($p<0.001$, Kruskal Wallis (SPSS) and Levene's test for equal variances (STATA)). The extent of methylation was increasing significantly between the four age groups for *EphA10* ($p=0.004$ between the neonates and the young adults, $p<0.001$ between the older three age groups, Wilcoxon signed rank test and Wilcoxon rank sum test (SPSS)), while the variance of methylation only increased significantly between the adults and the older adults ($p<0.001$, Levene's test for equal variances (STATA)).

E) For *HAND2*, both the extent of methylation and the variance of methylation were increasing significantly with an increasing age ($p<0.001$, Kruskal Wallis (SPSS) and Levene's test for equal variances (STATA)). The extent of methylation increased significantly between the three younger age groups ($p<0.001$, Wilcoxon signed rank test and Wilcoxon rank sum test (SPSS)) and the variance of methylation increased significantly between the adults and the older adults ($p=0.003$, Levene's test for equal variances (STATA)).

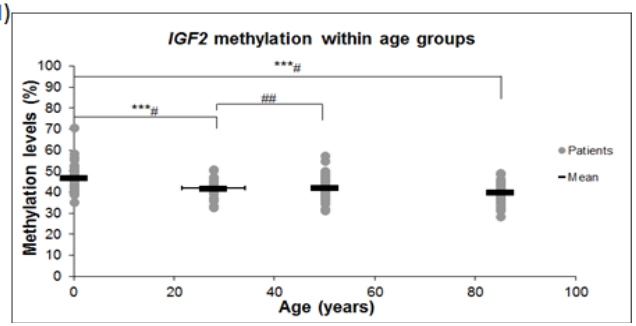
F) For *HOXA4*, both the extent and the variance of methylation were increasing significantly with an increasing age ($p<0.001$, Kruskal Wallis (SPSS and Levene's test for equal variances (STATA)). A significant increase in methylation levels was observed between the neonates and the young adults ($p<0.001$, Wilcoxon signed rank test (SPSS)), but no significant differences in methylation levels were observed between the other age groups. The variance of methylation was increasing significantly with an increasing age and between the neonates and the young adults and the adult s and the older adults ($p=0.015$ and $p=0.011$ respectively, Levene's test for equal variances (STATA).

The analysis of the methylation levels in the Newcastle 85+ population was performed by Hannah Gautrey

G)



H)



I)

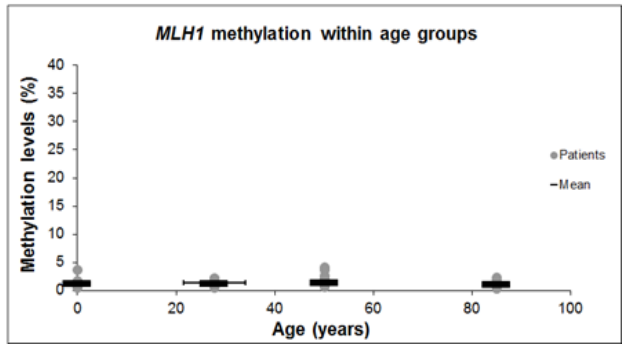


Figure 4.1G-I: Patterns of DNA methylation in different age groups

G) For *HOXA5*, the variance of methylation was increasing significantly with an increasing age ($p<0.001$, Levene's test for equal variance (SPSS)), but no differences were observed in the extent of methylation levels with age. A significantly decrease in methylation levels was observed between the neonates and the young adults ($p=0.026$, Wilcoxon signed rank test (SPSS)). Variance of methylation was increased significantly between the adults and the older adults ($p=0.002$, Levene's test for equal variances (STATA)).

H) For *IGF2*, both the extent of methylation and the variance of methylation were increasing significantly with an increasing age ($p<0.001$ for the extent of methylation, Kruskal Wallis (SPSS), $p=0.05$ for the variance of methylation, Levene's test for equal variance (SPSS)). A significant decrease in the extent of methylation was observed between the neonates and the young adults ($p=0.003$, Wilcoxon signed rank test (SPSS)). Significant differences in variance of methylation were observed between the three younger age groups ($p=0.030$ and $p=0.018$ respectively, Levene's test for equal variances (STATA)).

I) For *MLH1*, no significant difference in both the extent and the variance of methylation was observed with an increasing age, or between individual age groups.

The analysis of the methylation levels in the Newcastle 85+ population was performed by Hannah Gautrey

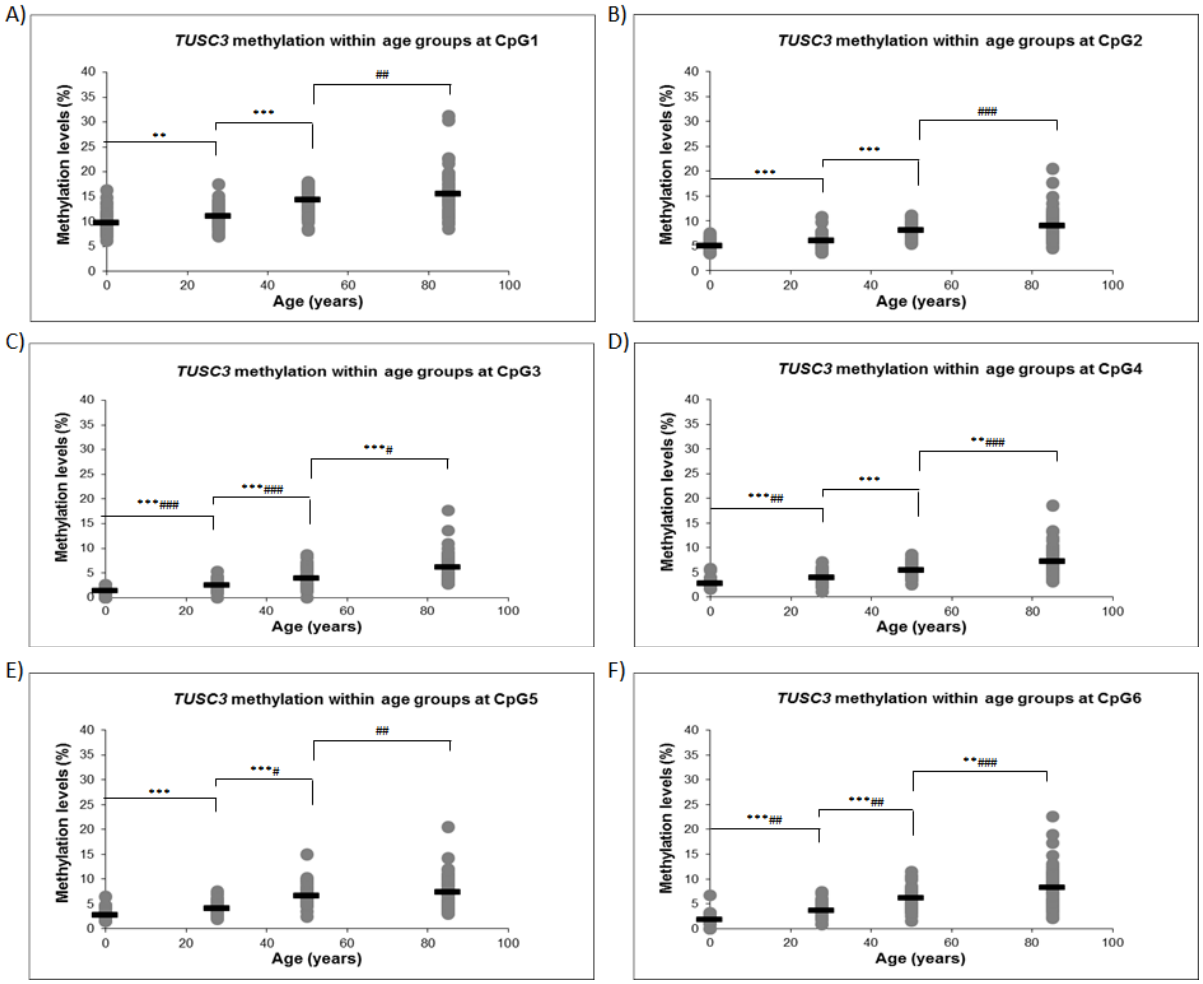


Figure 4.2: Patterns of methylation within age-groups in each individual CpG site for the *TUSC3* gene

The patterns of methylation were relatively similar between the different CpG sites for the *TUSC3* gene, showing that the age-related methylation levels for this gene was not relying on one specific CpG site.

An overview of the CpG site specific methylation changes within age groups for the other genes under study can be found in Appendix B, Table B.1.

4.3.2 Difference in variance of methylation within age groups

To examine differences in the variance of methylation with an increasing age, the differences in methylation patterns in the different age groups were examined for these differences. Seven of the nine genes under study showed an overall significant increase in variance of methylation between the four age groups.

The five genes that showed clear increases in the extent of age-related methylation levels; *TUSC3*, *TWIST2*, *HOXD4*, *EphA10* and *HAND2*, also showed significantly increased levels of variance of methylation between the different age groups. Only for *TUSC3* this significant increase in variance of methylation existed between all four age groups ($p=0.008$, $p=0.027$ and $p<0.001$ respectively).

HOXD4 showed an increase in variance of methylation between the neonates and the young adults and between the adults and the older adults ($p<0.001$), but not between the young adults and the adults. For *TWIST2* an increase in variance of methylation between the young adults and the adults and between the adults and the older adults was observed ($p=0.015$ and $p<0.001$ respectively), but not between the neonates and the young adults and for *EphA10* and *HAND2* a significant increase in variance of methylation was observed between the adults and the older adults ($p<0.001$ for *EphA10* and $p=0.003$ for *HAND2*), but not between the other age groups (Figure 4.1A-E).

The three genes that were relatively stable in the extent of methylation between the different age groups; *HOXA4*, *HOXA5* and *IGF2* did show a significant increase in variance of methylation between the different age-groups. For *IGF2* a significant difference in variance of methylation between the neonates and the young adults and between the young adults and the adults was observed ($p=0.030$ and $p=0.018$ respectively), but not between the adults and the older adults. *HOXA5* showed a significant increase in variance of methylation between the adults and the older adults ($p=0.002$), but not between the other age groups. *HOXA4* showed an increase in variance of methylation between the neonates and the young adults and between the adults and the older adults ($p=0.015$ and $p=0.011$ respectively), but not between the young adults and the adults (Figure 4.1F-H).

MLH1 was the only gene under study that did not show any difference in variance of methylation between the different age groups (Figure 4.1I).

4.3.3 Correlations between genes in the different age groups

For five of the genes under study, *TWIST2, TUSC3, EphA10, HOXD4* and *HAND2*, there was a very clear similarity in the age related changes in the patterns of DNA methylation seen in the four age groups, implying that the factors controlling age related changes at these five loci may be the same or at least very similar. To examine these similarities in greater detail, the methylation patterns were examined for correlations.

In the older adult population, there was a correlation observed between the levels of methylation of all five of these genes ($p<0.05$) (Table 4.2A), such that an individual with high levels of methylation at one of the five loci tends to have high methylation at the other four loci as well. *HOXA5, HOXA4*, *IGF2* and *MLH1* did not show a correlation with these genes, except between *HOXA5* and *EphA10* and *EphA10* and *IGF2* ($p<0.05$). However, despite the strong similarity in their average methylation patterns, the correlation within individuals was not as evident in the adults. Comparisons between the different genes showed similarities in this age group, from which some, but not all, reached statistically significant levels (Table 4.2B). In the young adult age group, fewer correlations were observed between the different genes (Table 4.2C). In the neonates these correlations were even less visible, the only observed correlations were between *HOXD4* and *HOXA5* ($p<0.05$) (Table 4.2D). It should be noted however, that due to the low levels of methylation generally observed in this population, and the lower levels of variation (and to some extent also in the young adult and adult populations), the sensitivity of the pyrosequencing method used might not be sufficient to identify any co-methylation of the genes that may exist in our sample set. It might be that our sample set was too small to detect significant correlations; therefore increased numbers of samples is required to draw more definite conclusions.

These results do suggest that co-methylation occurs in healthy individuals in a pattern that is similar to the CpG island methylator phenotype (CIMP) described in cancer.

A.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
TWIST2		**R=0.471** **P=0.001*****	**R=0.494** **P<0.001*****	**R=0.531** **P=0.006****	**R=0.371** **P=0.008****	R=0.195 P=0.175	R=0.115 P=0.426	R=-0.149 P=0.302
TUSC3			**R=0.612** **P<0.001*****	**R=0.306** **P=0.03***	**R=0.614** **P<0.001*****	R=0.124 P=0.392	R=0.059 P=0.683	R=-0.242 P=0.090
HOXD4				**R=0.312** **P=0.027***	**R=0.527** **P<0.001*****	R=0.112 P=0.437	R=0.090 P=0.532	R=-0.183 P=0.203
EphA10					**R=0.383** **P=0.006****	**R=0.398** **P=0.004****	R=0.094 P=0.517	**R=-0.297** **P=0.036***
HAND2						R=0.270 P=0.058	R=0.077 P=0.594	R=-0.270 P=0.058
HOXA5							R=-0.029 P=0.839	R=-0.020 P=0.890
HOXA4								R=0.024 P=0.867

Table 4.2A: Correlations between the examined genes in the older adult age-group

In the older adult samples (n=50) there were correlations observed between the genes *TWIST2*, *TUSC3, HOXD4, EphA10* and *HAND2* and between *EphA10* with *HOXA5* and *IGF2*. The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

The analysis on the Newcastle 85+ population was performed by Hannah Gautrey.

B.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
TWIST2		**R=0.424** **P=0.002****	R=0.170 P=0.244	**R=0.432** **P=0.002****	**R=0.293** **P=0.041***	R=0.056 P=0.701	R=0.020 P=0.893	**R=-0.419** **P=0.007****
TUSC3			R=0.270 P=0.061	**R=0.291** **P=0.041***	**R=0.444** **P=0.001*****	R=0.097 P=0.501	R=-0.009 P=0.948	R=-0.241 P=0.133
HOXD4				R=0.157 P=0.282	R=0.116 P=0.431	R=0.251 P=0.082	R=-0.017 P=0.908	R=-0.159 P=0.327
EphA10					R=0.130 P=0.373	R=-0.080 P=0.579	R=0.265 P=0.063	R=-0.179 P=0.270
HAND2						R=0.047 P=0.751	R=0.193 P=0.183	R=-0.100 P=0.540
HOXA5							**R=-0.419** **P=0.002****	R=-0.058 P=0.724
HOXA4								R=0.020 P=0.904

Table 4.2B: Correlations between the examined genes in the adult age-group

In the adult age group (n=50), there were correlations observed between *TWIST2* with *TUSC3*, *EphA10, HAND2* and *IGF2*, between *TUSC3* with *EphA10* and *HAND2*, between *HAND2* and *MLH1* and between *HOXA5* and *HOXA4*.

The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

C.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
TWIST2		**R= 0.317** **P=0.028***	R= 0.212 P=0.148	R= 0.138 P=0.349	**R= 0.318** **P=0.028***	**R= 0.338** **P=0.019***	R= 0.008 P=0.959	R= -0.250 P=0.090
TUSC3			R= 0.118 P=0.424	**R= 0.528** **P<0.001****	R= 0.029 P=0.845	R= 0.205 P=0.163	R= 0.041 P=0.783	R= 0.007 P=0.963
HOXD4				R= -0.016 P=0.912	R= 0.203 P=0.166	R= -0.032 P=0.827	R= 0.136 P=0.357	R= -0.011 P=0.939
EphA10					R= 0.076 P=0.608	R= 0.235 P=0.108	R= -0.201 P=0.170	R= -0.073 P=0.627
HAND2						R= 0.192 P=0.191	R= 0.025 P=0.864	R= -0.055 P=0.715
HOXA5							R= -0.271 P=0.063	R= -0.122 P=0.415
HOXA4								R= 0.173 P=0.245

Table 4.2C: Correlations between the examined genes in the young adult age-group

In the young adult samples (n=48) there were correlations observed between *TWIST2* with *TUSC3*, *HAND2* and *HOXA5* and between *TUSC3* and *EphA10*.

The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

D.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
TWIST2		R= 0.194 P=0.182	R= 0.133 P=0.363	R= 0.197 P=0.175	R= -0.012 P=0.937	R= 0.016 P=0.915	R= -0.087 P=0.553	R= -0.154 P=0.317
TUSC3			R= -0.039 P=0.788	R= 0.172 P=0.238	R= 0.068 P=0.642	R= 0.065 P=0.658	R= -0.102 P=0.488	R= 0.068 P=0.661
HOXD4				R= 0.010 P=0.947	R= -0.025 P=0.866	**R= -0.306** **P=0.033***	R= -0.012 P=0.936	R= 0.026 P=0.865
EphA10					R= -0.095 P=0.514	R= 0.075 P=0.609	R= 0.271 P=0.059	R= 0.219 P=0.152
HAND2						R= -0.063 P=0.670	R= -0.066 P=0.651	R= 0.008 P=0.957
HOXA5							R= -0.028 P=0.848	R= -0.052 P=0.735
HOXA4								R= -0.062 P=0.688

Table 4.2D: Correlations between the examined genes in the neonate age-group

In the neonate samples (n=49), the only observed correlation was between the genes *HOXD4* and *HOXA5*.

The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

4.3.4 Methylation levels of correlated genes combined

Patterns of five genes under study; *TWIST2, TUSC3, HOXD4, EphA10* and *HAND2*, showed strong similarities in their patterns of methylation changes with an increasing age and strong correlations were observed between these five genes in the older adults, which suggests that these genes may be influenced by similar mechanisms. Combining the methylation levels at all five genes might give a more accurate measurement of the impact of age on DNA methylation at these (and other similar) genes. To do this, the average methylation levels of these five genes were calculated per individual. Using this combined methylation measure, a significant increase in the extent of methylation with an increasing age is observed, as well as a significant increase between the four individual age groups ($p<0.001$) and a significant increase in the variability of methylation with age was observed and within three of the four age groups ($p\leq0.009$) (Figure 4.3). Interestingly the figure showing these five genes combined produces an almost perfect straight line for the increase in methylation with age, suggesting that methylation is steadily acquired throughout the ageing process and is not specifically a feature of old age.

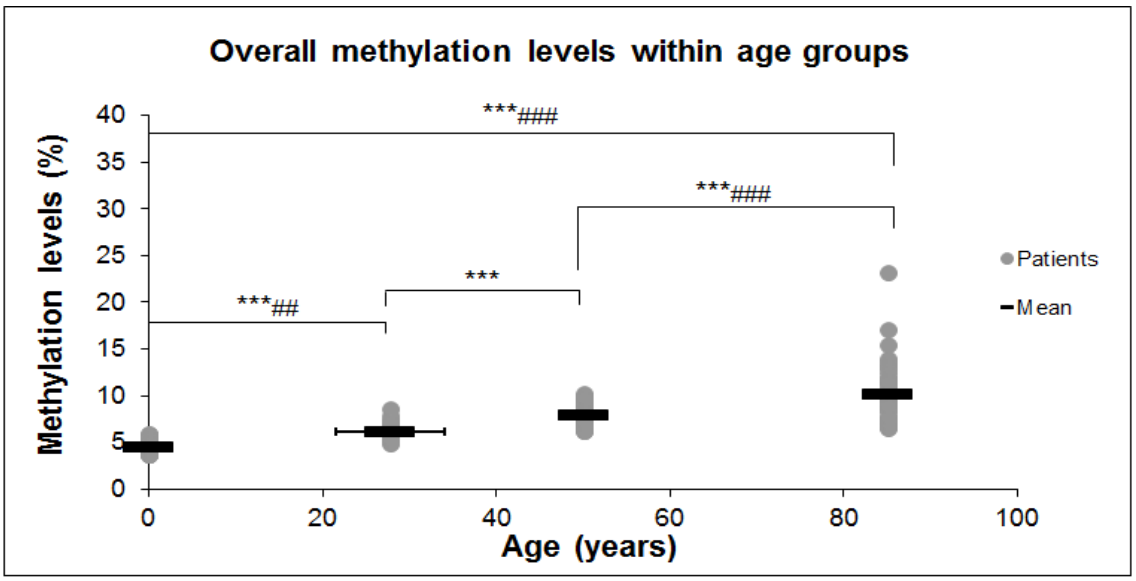


Figure 4.3: The overall levels of methylation increase steadily across the life course

Methylation levels of five genes combined, for which all individual genes showed similar patterns of DNA methylation changes between the different age-groups; *TWIST2*, *TUSC3*, *HOXD4*, *EpHA10* and *HAND2*, showed a significant increase in both the variance and the extent of methylation with an increasing age. The extent of methylation increased significantly between all four age groups ($p<0.001$ between all the age groups, Wilcoxon signed rank test and Wilcoxon rank sum test (SPSS)), while the variance of methylation increased significantly between the neonates and the young adults and between the adults and the older adults ($p=0.009$ and $p<0.001$ respectively, Levene's test for equal variances (STATA). These results suggest that methylation is steadily acquired throughout the ageing process.

The analysis of the methylation levels in the Newcastle 85+ population was performed by Hannah Gautrey.

4.3.5 Methylation levels of the neonates and their mothers

The neonates and the young adults were related, as they were child-mother pairs. To examine the possible effect on methylation of these related groups in greater detail, Pearson's correlations were performed to examine for a potential correlation between methylation levels of the mothers and their new-borns, to examine an effect of the mother's age on methylation levels in the neonates and to examine the effect of the neonates gestational age on their levels of methylation at birth. No correlation was observed between the methylation levels measured in the neonates and the methylation levels of their mothers (the young adults). The mother's age had no significant influence on the methylation levels in the neonates and no correlation was observed between the methylation levels measured in the neonates and their gestational age (Table 4.3).

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation mothers- Methylation neonates	R= 0.220 P= 0.137	R= -0.046 P= 0.759	R= -0.101 P= 0.501	R= 0.050 P= 0.724	R= 0.097 P= 0.517	R= -0.006 P= 0.970	R= 0.075 P= 0.616	R= 0.047 P= 0.793
Mothers age– Methylation neonates	R= 0.052 P= 0.731	R= 0.202 P= 0.173	R= 0.001 P= 0.994	R= -0.048 P= 0.749	R= 0.212 P= 0.152	R= -0.094 P= 0.531	R= 0.029 P= 0.846	R= -0.127 P= 0.394
Gestational age- Methylation neonates	R= -0.267 P= 0.063	R= -0.264 P= 0.067	R= -0.101 P= 0.501	R= 0.087 P= 0.552	R= -0.153 P= 0.293	R= 0.042 P= 0.775	R= -0.020 P= 0.893	R= 0.116 P= 0.426

Table 4.3: Correlations between the methylation levels of the neonates and the methylation levels of their mothers, their gestational age and their mother's age

No correlations were observed between the methylation levels of the neonates and the methylation levels of their mothers, no correlations were observed between the methylation levels of the neonates and their gestational age and no correlations were observed between the methylation levels of the neonates and their mother's age for any of the genes under study (n=47).

The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

4.3.6 Analyses stratification

To ensure that the age-related methylation changes that we observed between the different age groups were not significantly influenced by the different male-female ratios in the different age groups, stratification analyses were performed. First, the differences between the methylation levels of the males and females were examined in the individual age groups that consisted of both male and female participants. A difference in methylation levels was observed in the adult age group for the *TUSC3* gene, where the male participants showed higher average levels of methylation than the female participants, but this difference was not observed for the *TUSC3* gene in the neonate or older adult age groups (Table 4.4). None of the other genes showed significant differences in methylation levels between the male and the female participants (Table 4.4).

The young adult age group consisted of only females, so to perform further stratification analyses on the potential influence of the participant's genders on the overall patterns of methylation changes with age, the male participants in this study were excluded from the analyses. After the exclusion of the male population, the same statistical analyses were performed as before, namely the Kruskal-Willis test and the Wilcoxon signed rank test (SPSS). The general pattern of age-related methylation did not differ between the analyses performed on the whole population or on the female participants only (Table 4.5). No significant differences in the results were observed between the female population alone and the whole study population in the three older age groups. The significant differences between the neonates and the young adults that were observed for *EphA10*, *HOXA5* and *IGF2* when the whole population was included were not observed when only the female participants were included (Table 4.5). However, there were significantly less participants included in the analysis, namely 15 participants per age group, so it is possible that the number of samples was too small to detect significant differences for these genes in the small number of patients.

The smoking habits of the participants differed between the different age groups. To ensure that the methylation differences we observed in the different age groups were not influenced by the number of participant that smoked, stratification analyses were performed. First, the potential differences in methylation levels between the participants that had never smoked and the participants that were current- or previous smokers were examined. In these analyses, the young adult, adult and older adult age groups were included. A difference in methylation levels was observed in the older adult age group, between the participants that had never smoked and the current- and previous smokers

for the *IGF2* gene, but none of the other genes showed a significant difference in methylation levels between the participants that had never smoked and the current- and previous smokers in any of the age groups (Table 4.6).

To perform further stratification analyses on the potential influence of the participants smoking habits on the overall methylation patterns as observed with age, all the current smokers and previous smokers were excluded from these analyses and only the participants that had never smoked were included. After the exclusion of the previous- and current smokers, the same statistical analyses were performed as before, namely the Kruskal-Willis test and the Wilcoxon signed rank test (SPSS). The general pattern of age-related methylation did not differ between the analyses performed on the whole population or on the population only including the participants that had never smoked (Table 4.7). No significant differences were observed in the age-related methylation changes between the analyses performed on the whole study population and on the analyses performed on only the participants that had never smoked for any of the genes under study in the older three age groups. However, the significant differences between the neonates and the young adults that were observed for *EphA10*, *HOXA5* and *IGF2* when the whole population was included, were not observed anymore when only the participants that never smoked were included (Table 4.7). However, there were significantly less participants included in this analysis (namely 19 participants per age group). Therefore, it is possible that the number of samples was too small to detect significant differences for these genes in these small numbers of patients.

A.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels whole population (%)	4.17	3.88	4.98	6.53	3.28	60.52	26.69	46.64
Methylation levels males (%)	4.22	4.00	5.16	4.48	3.39	61.33	26.63	46.20
Methylation levels females (%)	4.08	3.67	4.58	6.68	3.01	58.89	27.01	47.62
Difference between males and females	P=0.885	P=0.155	P=0.225	P=0.824	P=0.417	P=0.081	P=0.747	P=0.973

Table 4.4A: Differences in methylation levels between the male and the female participants in the neonate age-group

In the neonate age group, no significant differences were observed between the methylation levels of the male (n=33) and female (n=15) participants for any of the genes under study (Independent samples T-test for *HAND2*, Wilcoxon rank sum tests for the other genes under study (SPSS)).

B.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels whole population (%)	5.54	**7.51**	11.17	8.59	7.14	56.69	39.58	42.34
Methylation levels males (%)	5.89	**8.27**	11.69	8.65	7.43	59.46	40.62	40.68
Methylation levels females (%)	5.41	**7.11**	10.96	8.59	7.00	55.59	39.17	43.25
Difference between males and females	P=0.443	**P=0.004****	P=0.258	P=0.686	P=0.195	P=0.110	P=0.534	P=0.144

Table 4.4B: Differences in methylation levels between the male and the female participants in the adult age group

In the adult age group, a significant difference was observed between the methylation levels of the male (n=18) and the female (n=32) participants for the *TUSC3* gene (p=0.004), but not for any of the other genes under study (Independent samples T-test for *TUSC3* and *HOXD4*, Wilcoxon rank sum test for the other genes under study (SPSS)).

C.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels whole population (%)	8.41	8.92	14.81	11.98	7.34	61.49	41.51	40.16
Methylation levels males (%)	7.96	8.20	14.06	12.14	7.43	64.76	42.29	39.68
Methylation levels females (%)	8.79	9.26	15.43	11.85	7.27	58.59	40.86	40.57
Difference between males and females	P=0.345	P=0.992	P=0.477	P=0.953	P=0.938	P=0.089	P=0.674	P=0.540

Table 4.4C: Differences in methylation levels between the male and the female participants in the older adult age-group

In the older adult age group, no significant differences were observed between the methylation levels of the male (n=23) and the female (n=27) participants for any of the genes under study (Independent samples T-test for *HOXA4* and *HOXA5*, Wilcoxon rank sum test for the other genes under study (SPSS)).

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Whole population (Neonates-Young adults)	Z=-0.215 P= 0.829	**Z=-5.298 P<0.001*****	**Z=-5.857 P<0.001*****	**Z=-2.841 P=0.004****	**Z=-5.898 P<0.001*****	**Z=-2.231 P=0.026***	**Z=-4.359 P<0.001*****	**Z=-3.003 P=0.003****
Females only (Neonates-Young adults)	Z=-0.454 P= 0.650	**Z=-3.096 P=0.002****	**Z=-3.351 P=0.001*****	Z=-1.477 P= 0.140	**Z=-3.408 P=0.001*****	Z= 0.000 P=1.000	**Z=-2.727 P=0.023***	Z=-1.591 P= 0.209
Whole population (Young adults-Older adults)	**C=72.831 P<0.001*****	**C=62.707 P<0.001*****	**C=44.253 P<0.001*****	**C=77.293 P<0.001*****	**C=36.144 P<0.001*****	C= 3.791 P=0.150	C =3.680 P=0.159	C=3.428 P=0.180
Females only (Young adults-Older adults)	**C=52.596 P<0.001*****	**C=48.548 P<0.001*****	**C=34.826 P<0.001*****	**C=50.710 P<0.001*****	**C=28.073 P<0.001*****	C= 1.195 P=0.550	C= 2.057 P=0.358	C= 4.636 P=0.098

Table 4.5: Gender stratification analyses

When only the female population was included, there were significant differences observed between the analyses done on the whole population and the analyses done on only the female population when comparing the methylation levels of the neonates and the young adults (n=15) for *EphA10*, *HOXA5* and *IGF2*. For the older three age groups, no significant differences were observed between the analyses performed on the whole study population and the female population alone.

The Z is indicating the Z value, the C is indicating the Chi-square value and the P the observed P values (Wilcoxon signed rank test for the analyses performed on the Neonates and Young adults, Kruskal-Willis test for the analyses performed on the other age groups (SPSS)).

A.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels whole population (%)	4.14	5.20	8.65	7.32	5.54	56.36	37.32	42.05
Methylation levels never smoked (%)	4.21	4.28	8.43	7.30	5.51	55.53	40.29	41.23
Methylation previous smokers (%)	3.80	5.05	8.26	8.07	5.94	52.64	28.11	42.30
Methylation levels smokers (%)	3.99	4.93	7.46	6.70	5.96	52.24	36.05	42.60
Difference between never smoked and previous- and current smokers	P=0.783	P=0.860	P=0.597	P=0.916	P=0.317	P=0.549	P=0.113	P=0.930

Table 4.6A: Differences in methylation levels between smokers and non-smokers in the young adult age-group

In the young adult age group, no significant differences were observed between the participants that had never smoked (n=19) and the current- and previous smokers (n=6) for any of the genes under study (Independent samples T-test for *HAND2*, *HOXA5*, *HOXA4* and *IGF2*, Wilcoxon rank sum test for the other genes under study (SPSS)).

B.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels whole population (%)	5.54	7.51	11.17	8.59	7.14	56.69	39.58	42.34
Methylation levels never smoked (%)	5.31	7.56	10.73	8.58	7.23	58.93	38.02	42.85
Methylation previous smokers (%)	5.84	7.85	11.24	8.75	7.15	55.97	40.39	40.37
Methylation levels smokers (%)	5.70	7.22	11.81	8.54	7.06	55.65	41.11	43.35
Difference between never smoked and previous- and current smokers	P=0.068	P=0.894	P=0.234	P=0.894	P=0.697	P=0.182	P=0.251	P=0.335

Table 4.6B: Differences in methylation levels between smokers and non-smokers in the adult age-group

In the adult age group, no significant differences were observed between the participants that had never smoked (n=19) and the current- and previous smokers (n=30) for any of the genes under study (Independent samples T-test for *TUSC3* and *HOXD4*, Wilcoxon rank sum test for the other genes under study (SPSS)).

C.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels whole population (%)	8.41	8.92	14.81	11.98	7.34	61.49	41.51	**40.16**
Methylation levels never smoked (%)	8.87	8.85	16.06	11.50	7.64	64.43	41.71	**43.06**
Methylation previous smokers (%)	7.59	8.49	13.51	11.91	6.82	58.78	41.41	**39.34**
Methylation levels smokers (%)	14.41	13.66	22.13	15.31	11.15	72.92	41.55	**33.20**
Difference between never smoked and previous- and current smokers	P=0.448	P=0.803	P=0.405	P=0.795	P=0.236	P=0.261	P=1.000	**P=0.003****

Table 4.6C: Differences in methylation levels between smokers and non-smokers in the older adult age-group

In the older adult age group, a significant difference was observed between the methylation levels of the participants that had never smoked (n=16) and the current- and previous smokers (n=34) for the *IGF2* gene, but not for any of the other genes under study (Independent samples T-test for *HOXA4* and *HOXA5*, Wilcoxon rank sum test for the other genes under study (SPSS)).

The analysis of the methylation levels in the Newcastle 85+ population was performed by Hannah Gautrey.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	HOXA5	HOXA4	IGF2
Whole population (Neonates-Young adults)	Z= -0.215 P= 0.829	**Z= -5.298 P< 0.001*****	**Z= -5.857 P< 0.001*****	**Z= -2.841 P= 0.004****	**Z= -5.898 P< 0.001*****	**Z= -2.231 P= 0.026***	**Z= -4.359 P< 0.001*****	**Z= -3.003 P= 0.003****
Participants that never smoked only (Neonates-Young adults)	Z= -0.540 P= 0.598	**Z= -2.897 P= 0.004****	**Z= -3.124 P= 0.002***	Z= -1.533 P= 0.125	**Z= -3.408 P= 0.001*****	Z= -1.420 P= 0.156	**Z= -2.953 P= 0.003****	Z= -1.682 P= 0.093
Whole population (Young adults-Older adults)	**C=72.831 P< 0.001*****	**C=62.707 P< 0.001*****	**C=44.253 P< 0.001*****	**C=77.293 P< 0.001*****	**C=36.144 P< 0.001*****	C= 3.791 P= 0.150	C = 3.680 P= 0.159	C=3.428 P= 0.180
Participants that never smoked only (Young adults-Older adults)	**C=22.110 P< 0.001*****	**C=20.217 P< 0.001*****	**C=17.253 P< 0.001*****	**C=26.750 P< 0.001*****	**C=17.664 P< 0.001*****	C= 5.612 P= 0.060	C= 1.120 P= 0.571	C= 2.208 P= 0.332

Table 4.7: Smoking stratification analyses

When only the participants that had never smoked were included, there were significant differences observed between the analyses done on the whole population and the analyses done on only the participants that have never smoked when comparing the methylation levels of the neonates and the young adults (n=19) for *EphA10*, *HOXA5* and *IGF2*. For the older age groups, no significant differences were observed between the analyses performed on the whole study population and the participants that had never smoked.

The Z is indicating the Z value, the C is indicating the Chi-square value and the P the observed P values (Wilcoxon signed rank test for the analyses performed on the Neonates and Young adults, Kruskal-Willis test for the analyses performed on the other age groups (SPSS)).

4.4 Discussion

The results of our study indicate that, for at least five of the genes, clear age-related methylation changes were observed, where both the extent of methylation as well as the variance of methylation increases significantly with age. Our results also suggest that these age-related methylation changes were steadily acquired during the ageing process and that they were not just a feature of an old age. What is causing these methylation changes during aging is currently unclear, but it might be caused by an unstable maintenance of the replication of DNA methylation patterns during cell divisions during aging. In addition to an intrinsic biological effect of aging on DNA methylation levels, the methylation levels might be influenced by several environmental factors. Altered intake of several nutrients can affect DNA methylation. These nutrients include folate, selenium, phytochemicals and polyphenolics and food contaminants, such as arsenic (Mathers, 2006). Air pollution, such as diesel exhaust and PM_{10} exposure, were shown to alter both global and promoter specific DNA methylation levels (Christensen, 2011) and smoking was shown to alter levels of promoter methylation in a dose-dependent matter, where the degree of methylation increases with the intensity of smoking (Talikka, 2012). Older individuals might have been more exposed to all of these environmental modulators of methylation, and the levels of which they are exposed might vary more than in the younger age groups, which might explain the increase in variance of methylation in the older age groups. A study carried out by Fraga *et al* in monozygotic twins showed that an increase in epigenetic differences was observed between twin pairs with an increasing age (Fraga, 2005). These are genetically identical individuals, and the observed increase in the differences in epigenetic patterns in these individuals with an increasing age can possibly be explained by both external and internal factors.

Five of the genes in our study seem to follow similar patterns of age related changes in DNA methylation. Strong correlations between the methylation levels of these five genes were observed in the older age groups, suggesting that changes in the DNA methylation of these genes are closely related to each other in healthy individuals and that there might be similar mechanisms underlying methylation of these genes, and potentially other genes in the genome. The genes under study were identified from the literature and selected due to their role in leukaemia development. It is highly unlikely that these five genes are the only genes in the genome showing such age-related methylation patterns, suggesting that many other genes might undergo similar patterns of methylation. Interestingly, these methylation patterns were observed in leukaemia-

related genes, whereas the genes that were unrelated to leukaemia development, such as *MLH1* and *IGF2*, did not show the same extent of age related methylation changes. This suggests that the genes that become methylation in peripheral blood during normal ageing may be closely related to the genes that become methylated in the cancer type developed from this tissue type. The observation that methylation changes during normal ageing are following similar patterns to the methylation changes during leukaemia development, leads to the hypothesis that increased levels of methylation, acquired during the ageing process, might be a precursor for the development of leukaemia, and potentially other disease related to age. This hypothesis is strengthened by the observations that co-methylation of several genes were also already present in healthy individuals, similar to the CIMP described in cancer, describing the presence of widespread CpG island methylation in tumours.

Three of the other genes under study: *HOXA4*, *HOXA5* and the imprinted gene *IGF2* showed different patterns in methylation levels during aging compared to the earlier discussed five genes. The methylation levels for these genes are relatively stable during aging, although the methylation levels for these genes are changing significantly between the two youngest age groups. Both *HOXA5* and *IGF2* showed significantly higher levels of methylation in the neonates compared to the young adults, while *HOXA4* showed significantly lower levels of methylation in the neonates compared to the other, older, age groups. As there were no samples taken between 0 and 17 years of age it is not clear if the difference in methylation for these genes after birth occurs rapidly or if it is more gradual over the first two decades of life.

DNA methylation levels of all of these three genes have been associated with early development. Both *HOXA4* and *HOXA5* play a role in development and differentiation, while placental DNA methylation changes at the *IGF2*/*H19* imprinted locus are associated with fetal development and birth weight in a normal pediatric population (St-Pierre, 2012). Due to the importance of these genes in early development, it is well possible that the methylation levels of these genes change rapidly during the first years of life.

Our results for *IGF2* also suggests that methylation patterns for imprinted genes are better preserved than for other genes, indicating the importance of the imprinted genes.

In this study, there was no correlation observed between the methylation levels of the neonates and their mother's age. This is contrary to the observations made in the study of Adkins *et al*, where they found a correlation with maternal age and methylation levels in the neonates for 144 CpG sites belonging to 142 genes. The genes that were

associated with oncogenesis and cancer progression were significant over-represented among the genes that correlated with maternal age (Adkins). Unfortunately, our sample set did not allow us to examine the possible link between parental age and the fathers' methylation levels and the methylation levels in the neonates, since in the study of Adkins *et al* a trend was observed between the paternal age and methylation levels in the neonates. The only overlapping gene between our study and the study of Adkins *et al* was the imprinted gene *IGF2*, a gene that showed a correlation between the methylation levels in the neonates and maternal age in their study. In our study, no evidence of a possible correlation between *IGF2* methylation levels in the neonates and the mother's age was observed. However, it is possible that a correlation between the maternal age and methylation levels in the neonates in our study was not detected due to the low number of samples in our study; we examined 49 samples, compared to 96 samples in the study of Adkins *et al*. It is also possible that the different techniques used in these studies can explain the differences in correlations observed. In our study, methylation levels were assessed using pyrosequencing, which is a highly accurate measurement of CpG methylation. In the study of Adkins *et al* they performed a Humanmethylation27 BeadChip method. After the Humanmethylation27 BeadChip, they confirmed some genes on the pyrosequencer. Unfortunately, *IGF2* was not among those genes. However, the methylation results gathered by both methods in their study indicated that the correlation between maternal age and newborn DNA methylation levels is replicable, but that the absolute methylation values derived from the Humanmethylation27 array for some sites may be biased indicators of the true values and may overestimate the true level of inter-individual variation in methylation levels (Adkins). As a result, it is possible that the correlation observed between maternal age and the neonates methylation levels as observed in their study is an overestimation due to the technique used in their study.

The results of our study might suffer from confounding factors related to the study set-up. First of all, the DNA samples of the 4 different age groups were collected as part of different study cohorts. Although the method of DNA isolation is unlikely to influence the methylation results, especially since all DNA samples were stored in a similar fashion after blood collection, it is still a possibility. Our different study populations were not matched for several other factors, such as racial background and socio-economic status (SES), factors that might have influenced the results. Differences in CpG site methylation were reported between neonates with an African-American and Caucasian racial background (Adkins, 2011) and in adults from Caucasian and non-

Caucasian racial backgrounds (Lam, 2012). Correlations were observed between early life SES and methylation levels in adults, where differences in methylation levels were observed in adults with low early-life SES and high early-life SES (Lam, 2012; Borghol, 2012). However, it should be noted that the cohorts used in this study all collected their samples in similar regions and as a result it is likely that most of these factors are relatively similar between the study populations.

The DNA from the neonate samples were isolated from cord blood, while all the other DNA samples in this study were isolated from PBL. The different source of DNA might have influenced methylation levels. The young adults participating in our study were pregnant at time of blood collection. What a pregnancy does to individuals methylation levels is, by our knowledge, not examined into detail, but it might be possible that the methylation levels have been changed due to the pregnancy. Although our results suggest that it is unlikely that these methylation patterns are changed significantly, since the overall methylation levels produced an almost perfect straight line with an increasing age.Overall, our results showed clear age-related methylation changes in several genes under study. Five genes under study showed clear similarities in the change in methylation patterns within different age groups. Three of the other genes under study showed more stable levels of methylation during ageing although increases in variance of methylation were observed in the older age group and one of the genes under study did not show any changes in both the extent and variance of methylation with an increasing age. The genes showing these variable methylation levels during ageing were closely related to genes that were known to be altered in leukaemia, suggesting a potential link between increased levels of methylation and the susceptibility of leukaemia development.

Chapter 5. Cell type specific DNA methylation

5.1 Introduction

Tissues are made up of a number of different cell types. For example, a blood sample contains of plasma and formed elements, such as erythrocytes (red blood cells), thrombocytes (platelets) and leukocytes (white blood cells). The leukocytes are again a mixture of different cell types, including neutrophils, eosinophils, basophils, lymphocytes (T- and B-lymphocytes), monocytes, macrophages and dendritic cells. Unlike the genetic make-up, where generally all the cells are virtually identical, different cell types have different epigenetic make-ups and express genes differentially (Ehrlich, 1982; Byun, 2009). These tissue and cell type specific differences in DNA methylation patterns can cause a number of difficulties in clearly understanding age-related alterations in DNA methylation. Almost all studies are carried out using tissue samples, which are a mixture of different cell types. This means interpretation of the results obtained for methylation patterns in the tissues may be confounded by age-related changes in the cellular make-up of tissues. In addition, the extent of age-related changes may be underestimated if the changes occur only in one specific cell type in the tissue being assessed. The cellular composition of blood changes with age, with a decrease in lymphocyte counts observed at older ages (Huenecke, 2008), possibly as a result of a reduced efficiency of bone marrow or HSCs to generate lymphoid progeny, as was described in several studies using mouse models (Linton, 2004). A study by Teschendorff *et al* showed that these changes in cell composition during ageing could potentially explain hypomethylation that was observed with an increasing age, but this was not the case for age-related hypermethylation, since they were able to correlate the age-associated hypomethylation signature in blood with changes in blood cell-type composition in a set of cancer-related genes, but not so for the age-hypermethylated signature (Teschendorff).

It has been suggested that DNA methylation plays a role in lineage restriction in blood cells and in regulating commitment to lymphoid versus myeloid cell fates (Cedar, 2011). Many genes that are initially methylated in stem cells were found to undergo selective demethylation in a tissue- or lineage-specific manner, which was shown not to be necessarily linked to gene expression. Besides alterations in methylation levels between cell lineages, alterations were also observed during haematopoietic differentiation of specific cell types (Calvanese). Genomic DNA methylation profiles were found to define distinct cell types and profiles of haematopoietic cells of the same type are more closely related to each other than to other cell types (Bocker). It has been proposed that these lineage-specific DNA methylation patterns are established prior to

terminal differentiation of adult progenitor cells (Hupkes). The study of Lee *et al* showed that in progenitor cells before complete B-cell lineage commitment, the promoters of the majority of the CpG islands in gene promoters are hypomethylated. During progression through B-cell stages, these promoter regions maintained their hypomethylation status, whereas alterations in methylation occurred specifically to the gene body regions of genes (Lee, 2012).

This study was designed to examine methylation patterns in different cell types in the promoters of nine genes and to examine how these methylation patterns are maintained during ageing, to assess if the age-related DNA methylation changes measured in PBL could be related to altered cellular make-up of the blood with an increasing age or if these changes in methylation are genuine increased levels of methylation in the different cell types. Another aim of this study was to understand the impact of differences in cellular make up between individuals to assess the potential links between methylation levels and disease development at an individual level.

Methylation differences between cell types and within different age groups were assessed on specific isolated cell types from PBL samples from a group of young adults and a group of older adults. Isolating stem-and early progenitor cells, as well as cells later during differentiation, from a set of G-CSF mobilized samples, allowed us to examine methylation patterns of this same set of genes during cell differentiation. Seven of the genes under study had been previously identified to show age-related methylation changes (as described in Chapter 4), one gene is a known imprinted gene and one was a control gene not affected by age related DNA methylation in PBL.

5.2 Samples

Blood samples were collected from 10 young adults, between 24 and 31 years of age, with an average age of 27.6, and from 8 adults, between 43 and 60 years of age, with an average age of 51.6 (Table 5.1). MNCs were isolated using Biocoll gradient centrifugation. B-lymphocytes, T-lymphocytes and monocytes were isolated from the MNCs using the dynabead separation protocol for these specific cell types, as described in Chapter 3. The monocyte depleted MNC samples were also stored as this sample represents a total lymphocyte population. The granulocytes were isolated by gradient centrifugation. DNA was extracted from the different cell types and from both the PBL and MNC samples.

G-CSF mobilized CD34 peripheral blood cells were collected from six young adults, between 24 and 33 years of age, with an average age of 26.2 and stored in liquid

nitrogen in medium containing DMSO to prevent cell lysis. The cells were defrosted and $CD34^+$ cells were isolated from these samples using the dynabead separation protocol, as described in Chapter 3. Further isolation into early B-cell progenitors led to a collection of $CD34^+CD19^+$ and $CD34^+CD19^-$ cells. B-lymphocytes, T-lymphocytes and monocyte were also isolated from these samples (Table 5.1).

	Young adults	**Adults**	**G-CSF samples**
Age (years)	24-31 (average 27.6)	43-60 (51.6 average)	24-33 (average 26.2)
Gender	4 female 6 male	5 female 3 male	Unknown
DNA samples	PBL MNC Lymphocytes B-lymphocytes Monocytes	PBL MNC Lymphocytes B-lymphocytes T-lymphocytes Monocytes Granulocytes	Stem- and early progenitor cells Early B-cell progenitors Progenitor cells- B-cell progenitors B-lymphocytes T-lymphocytes Monocytes
Number of participants	10	8	6

Table 5.1: The cell type specific methylation study population

5.2.1 Validation of cell isolation

The cell isolation was validated using two different methods. After the CD34 isolation by the use of dynabeads, these beads could be detached, which allowed validation by flow cytometry. After isolation of the $CD34^+$ cells, the cells were stained with a fluorescence antibody specific for CD34 to distinguish the CD34 positive cells, from the negative cells. The beads for the CD19, CD14 and CD3 isolations could not be detached after isolation. These beads were also able to respond to the different antibodies used for flow cytometry, which are supposed to be directed against specific cell surface makers. This made it impossible to perform the validation of the isolation of these specific cell types by flow cytometry and as a result, the validations of these cell isolations were performed using the qPCR technique (as described in Chapter 3).

5.3 Results

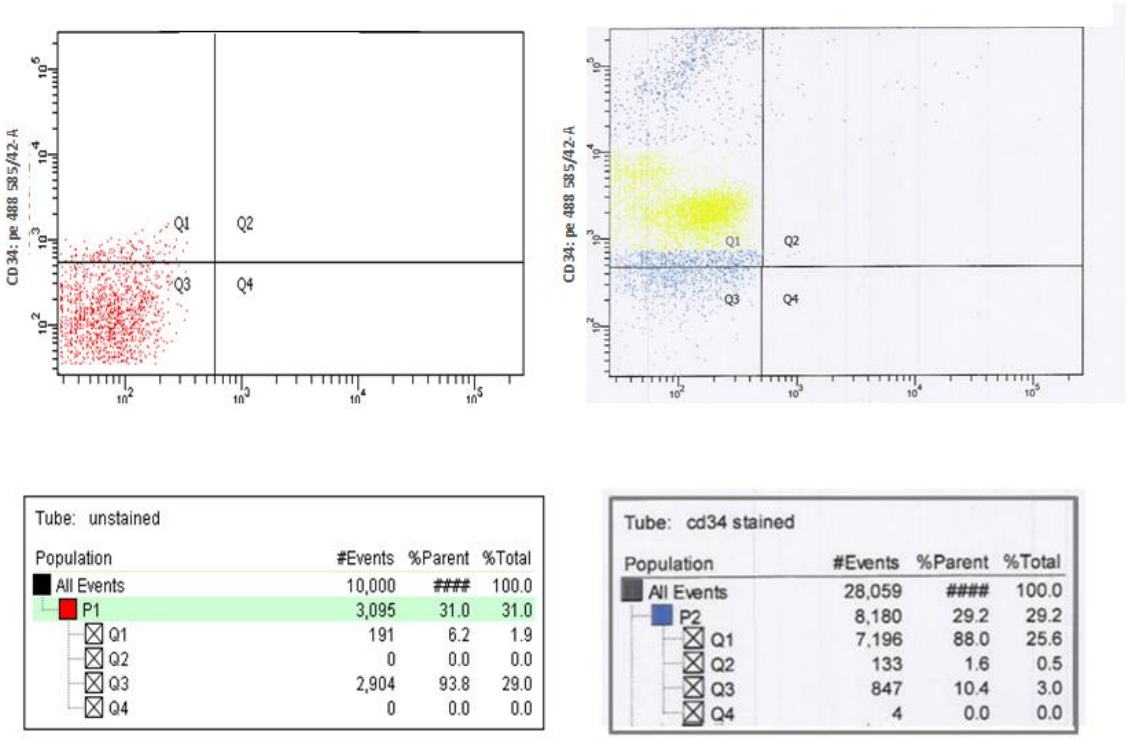
5.3.1 Validation of cell isolation

CD34 cells were isolated from G-CSF mobilized CD34 peripheral blood samples that were stored in liquid nitrogen in medium containing 10% DMSO, by the dynabead isolation method, as described in Chapter 3.

After isolation of the $CD34^+$ cells, the cells were stained with a fluorescence antibody specific for CD34. This showed that after isolation, 88% of the stained isolated cells were positive for the CD34 antibody, compared to 4.4% in for the unstained sample, indicative of the positive isolation (Figure 5.1).

To validate the other cell isolations, CD19, CD14 and CD3 cells were isolated using the dynabead isolation method, as described in Chapter 3, from PBL samples. Expression of cell type specific markers CD19 (a B-lymphocyte marker), CD14 (a monocyte marker), CD3 (a T-lymphocyte marker) and GAPDH (control) was quantified in extracted RNA using qPCR. Cell lines, known to be expressing each marker, were used as positive controls. All of the cell lines expressed the markers they were supposed to be positive for, ensuring us that the qPCR worked properly. As expected, the PBL and MNC samples expressed all the cell type specific markers. The sorted B-lymphocytes expressed CD19, but were negative for CD14 and CD3. The sorted T-lymphocytes expressed CD3, but were negative for CD19 and CD14 expression. The sorted monocytes expressed CD14, but were negative for CD19 and CD3 and the

expression for the MNC sample minus the B-lymphocytes was negative for CD14 and positive for CD19 and CD3 (Table 5.2).



Tube: unstained

Population	#Events	%Parent	%Total
All Events	10,000	###	100.0
P1	3,095	31.0	31.0
Q1	191	6.2	1.9
Q2	0	0.0	0.0
Q3	2,904	93.8	29.0
Q4	0	0.0	0.0

Tube: cd34 stained

Population	#Events	%Parent	%Total
All Events	28,059	####	100.0
P2	8,180	29.2	29.2
Q1	7,196	88.0	25.6
Q2	133	1.6	0.5
Q3	847	10.4	3.0
Q4	4	0.0	0.0

Figure 5.1: Validation of the CD34 isolation by flow cytometry

In the unsorted cells, 93.8% of the cells were placed in the Q3 quarter, indicative of a CD34 negative population, while only low numbers of cells were placed in the upper half of the Q1 quarter.

After sorting CD34 cells, a large number of cells were placed in the upper half of the Q1 quarter, indicative of a CD34 positive population.

Sample	CD14 expression (fold enrichment compared to PBL)	CD19 expression (fold enrichment compared to PBL)	CD3 expression (fold enrichment compared to PBL)
MNC	0.60	0.47	0.45
T-lymphocytes	0.08	0	3.98
B-lymphocytes	0.02	3.55	0
Monocytes	0.20	0.07	0.01
MNC- monocytes	0.02	0.21	1.04

Table 5.2: Validation of the cell isolation using qPCR

Isolation of the T-lymphocytes, B-lymphocytes and the monocytes were validated using qPCR. This table shows the fold enrichment of the individual markers in the different cell types relative to the PBL sample. All the cell types were positively expressing their cell type specific cell surface markers and negative (or close to negative) for the cell surface markers specific for the other cell types.

5.3.2 Cell type specific DNA methylation patterns were observed in the young adult population

Methylation levels were assessed by pyrosequencing in the isolated cell types from the young adults in a set of nine genes, the examine cell type specific methylation patterns.

The five genes under study that were previously described to show similar patterns of age-related methylation changes (as described in Chapter 4); *TUSC3*, *TWIST2*, *HOXD4*, *EphA10* and *HAND2*, exhibited similar patterns of cell type specific methylation differences in the young adults. The lymphoid cells showed significantly higher levels of methylation compared to the myeloid cell type (the monocytes) (Table 5.2). However, it were the B-lymphocytes that showed the highest levels of methylation for all these cell types analysed and the methylation levels in these B-lymphocytes were even significantly increased compared to the total amount of lymphocytes (Table 5.3). For example for the *TUSC3* gene, methylation levels of 6.62% were observed in the PBL sample. The lymphoid cells showed higher levels of methylation, namely 8.11%. However, *TUSC3* methylation levels in the B-lymphocytes were 16.50%, which was significantly higher than any of the other cell types assessed. This was significantly higher than in any of the other cell types (Table 5.3 and Figure 5.2A-E).

HOXA4 and *HOXA5* were previously described to show changes in the variance of methylation with an increasing age, while the extent of methylation was relatively stable with an increasing age. These genes exhibited significantly higher levels of methylation in the normal PBL samples compared to the earlier described five genes in these young adults (38.50% for *HOXA4* and 53.57% for *HOXA5*). Cell type specific methylation changes were observed even in these genes. The methylation levels of the lymphoid cells were significantly higher than the methylation levels in the myeloid cell type (the monocytes) (Table 5.3 and Figure 5.2A F-G).

Even though the imprinted gene *IGF2* was previously described to show similar patterns of age-related methylation as the *HOXA4* and *HOXA5* genes, it did not show any differences in methylation levels within the different cell types in the young adults (Table 5.3 and Figure 5.2AH), nor did the *MLH1* gene. However, methylation levels at this *MLH1* locus were found to be very low in all cell types assessed (Table 5.3 and Figure 5.2AI).

These results show that cell-type specific methylation patterns are present in peripheral blood. The existence of both this cell type specific methylation, which is primarily present in the B-lymphocytes, and age-related DNA methylation, as described

in Chapter 4, in this same set of genes, suggests that the observed methylation changes during ageing are not accumulated in the stem cells (all these cell types derive from the same pool of stem cells), but that they rather occur during the process of differentiation.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*	*MLH1*
Methylation levels PBL (%)	4.86	6.62	10.07	8.27	8.80	53.57	38.50	43.93	1.97
Methylation levels MNC (%)	5.61	8.48	12.50	9.80	8.45	52.62	39.62	43.98	1.54
Methylation levels total lymphocytes (%)	6.51	8.48	13.48	10.73	8.26	54.41	41.23	45.19	1.69
Methylation levels B-lymphocytes (%)	13.74	16.50	22.02	15.80	10.87	56.31	44.58	47.43	2.03
Methylation levels monocytes (%)	3.14	5.07	6.79	6.89	6.34	51.29	36.82	43.41	1.54
PBL – MNC	P=0.169	**P=0.005****	**P=0.028***	**P=0.013***	P=0.878	P=0.199	P=0.169	P=0.575	P=0.056
Lymphocytes- monocytes	**P=0.003****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.022***	**P=0.032***	**P=0.013***	P=0.214	P=0.382
B-lymphocytes- Lymphocytes	**P=0.006****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	P=0.241	**P=0.037***	P=0.374	P=0.183
B-lymphocytes- monocytes	**P=0.003****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	P=0.086	P=0.066

Table 5.3: Methylation levels differ between cell types in the young adult population

Significant differences in methylation levels were observed between the different cell types in the young adults (n=10) for seven of the genes under study (Paired samples T-test for the genes *HOXA5* and *MLH1*, Wilcoxon signed rank test for all the other genes under study (SPSS)).

A complete overview of the statistical analyses performed on the methylation levels between all cell types can be found in Appendix C, Table C.1.

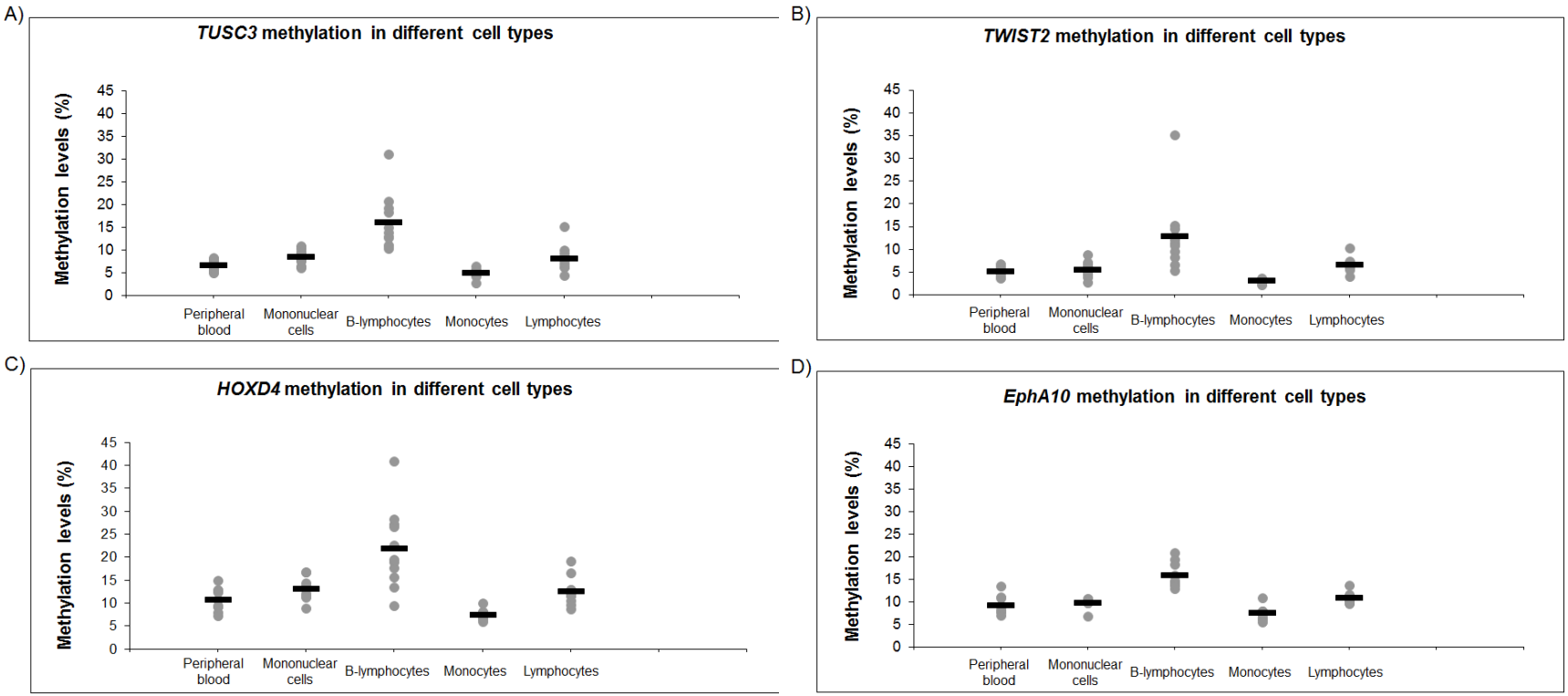


Figure 5.2A-D: Cell type specific methylation patterns in the young adult population for the *TUSC3, TWIST2, HOXD4* and *EphA10*

A-D) Clear cell type specific methylation differences are observed in the young adults for *TUSC3*, *TWIST2*, *HOXD4* and *EphA10*. Significantly higher levels of methylation were observed in the lymphoid cells compared to the monocytes (Table 5.3). The B-lymphocytes exhibited highest levels of compared to the other cell types analysed, namely the monocytes and the total lymphocytes.

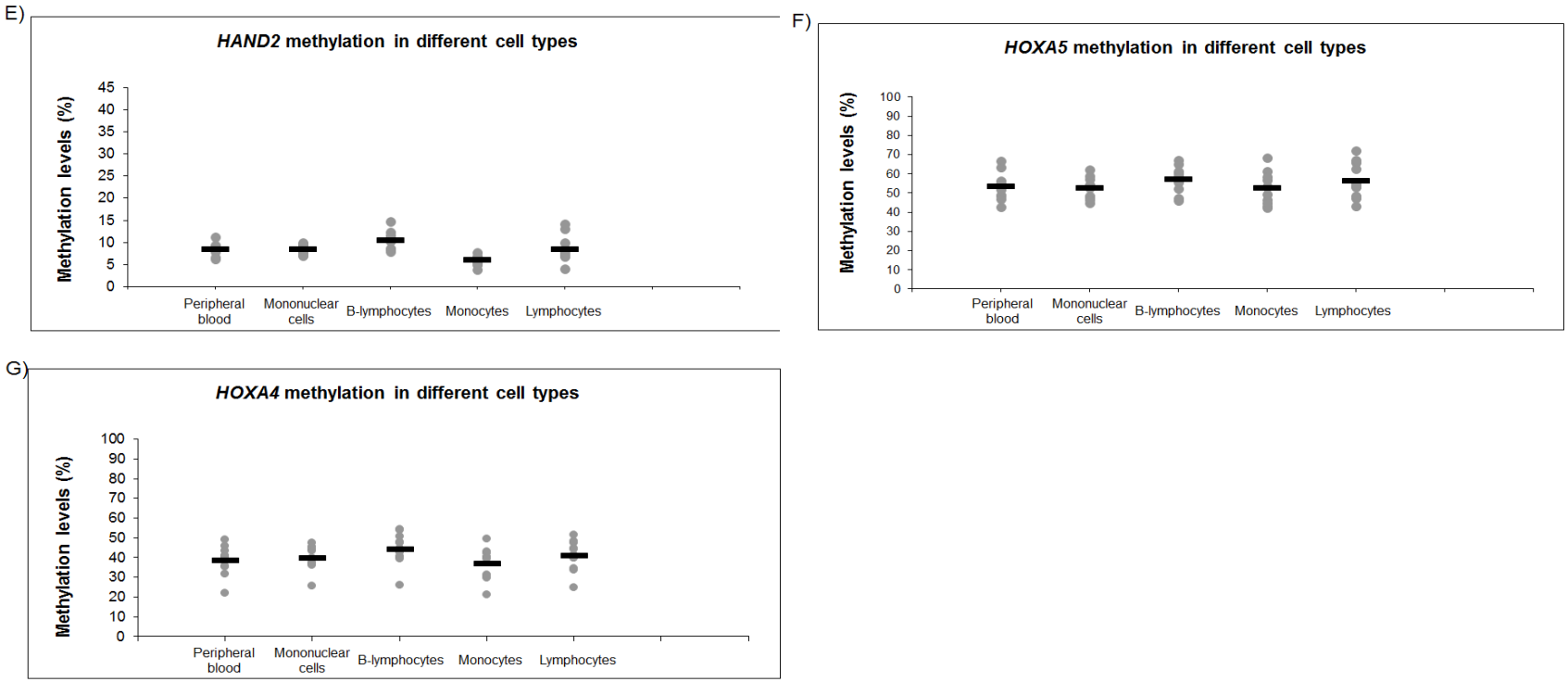


Figure 5.2E-G: Cell type specific methylation patterns in the young adult population for the *HAND2, HOXA5* and *HOXA4*

E-G) Cell type specific methylation differences are observed in the young adults for *HAND2*, *HOXA5* and *HOXA4*. Significantly higher levels of methylation were observed in the lymphoid cells compared to the monocytes (Table 5.3). The B-lymphocytes exhibited highest levels of DNA methylation compared to the methylation levels of the other cell types, namely the monocytes and the total lymphocytes.

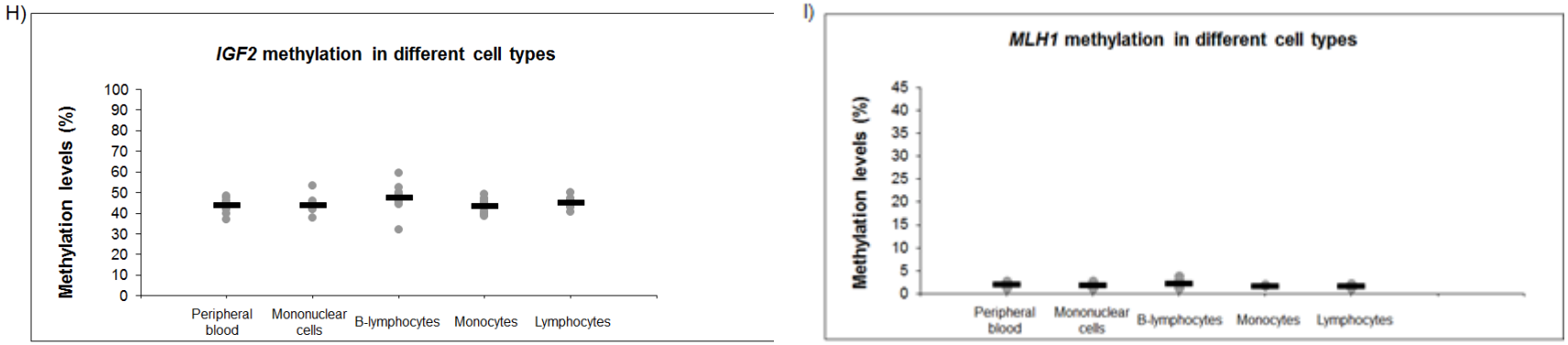


Figure 5.2H and I: Cell type specific methylation patterns in the young adult population for the *IGF2* and *MLH1*

H and I) No significant difference in methylation levels were observed between the different cell types for *IGF2* and *MLH1*.

5.3.3 Cell type specific DNA methylation patterns were observed in the adult population

Cell type specific methylation patterns were observed in the cell types isolated from the young adults. The observations that methylation differences were present between cell-lineages in peripheral blood, suggests that alterations in these cells have to occur during haematopoietic differentiation, since all the cell types in blood are derived from the same pool of stem cells in the bone marrow. An alternative hypothesis for altered epigenetic states to develop in different cell types is that different epigenetic states in stem cells could bias differentiation down one lineage or another, such that highly-methylated stem cells differentiate in one lineage, while the stem cells with low levels of methylation differentiate into the other lineage. It is possible that ageing may result in further differences in methylation levels between cell types, either just by random drift or because different cell types may be differentially sensitive to age-related influences on the epigenome.

To examine cell-type specific methylation patterns with an increasing age, cell types were isolated from a group of adults. In these adults, T-lymphocytes and granulocytes were isolated at the same time as the B-lymphocytes and the monocytes to get a more complete set of cell types in blood.

The five genes that showed both age-related methylation patterns and differential methylation in the specific cell types isolated from the young adults; *TUSC3*, *TWIST2*, *HOXD4*, *EphA10* and *HAND2*, exhibited similar patterns of cell type specific methylation in this group of adults. For these five genes, the lymphoid cells (T-lymphocytes and B-lymphocytes) exhibited more methylation than the myeloid cells (monocytes and granulocytes). Differences in methylation levels were also observed between cell types from the same lineage; the B-lymphocytes exhibited higher levels of methylation than the T-lymphocytes and the granulocytes exhibited higher levels of methylation than the monocytes.

Even though all the genes followed similar patterns of methylation changes in the different cell types, some of the methylation differences between the cell types were smaller for some of the genes, such that some of the differences in those genes fell short of statistical significance (Table 5.4 and Figure 5.3A-E).

As observed in the young adult population, both *HOXA5* and *HOXA4* also showed cell type specific methylation patterns. No significant differences were observed between all the cells from the lymphoid and the myeloid lineage, however

differences in methylation levels were observed between some cell types; for *HOXA5* significant differences were observed between the T-lymphocytes and the myeloid cells and *HOXA4* exhibited differences in methylation levels between the B-lymphocytes and the myeloid cells. Differences in methylation levels were observed between cell types from the lymphoid lineage for *HOXA5*, but neither of these genes showed differences in methylation levels between the cell from the myeloid lineage (Table 5.4 and Figure5.3F-G).

IGF2 did not show any significant differences in methylation levels in the cell types isolated from the young adults. However, in the adults, differences were observed between the B-lymphocytes and the myeloid cells and between the lymphoid cell types (Table 5.4 and Figure5.3H).

These results show that different cell types in peripheral blood still exhibit cell type specific methylation patterns in adults. Differences were even observed in different cell types from the same lineage, showing that these methylation patterns are not just lineage-specific.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels PBL (%)	5.00	6.39	11.32	10.78	9.58	61.22	46.09	42.97
Methylation levels MNC (%)	6.61	8.29	13.88	12.49	10.07	62.03	48.33	44.97
Methylation levels lymphocytes (%)	7.09	8.79	16.27	13.34	10.67	62.60	48.60	41.60
Methylation levels T-lymphocytes (%)	8.28	9.29	16.55	14.67	10.41	64.77	48.24	43.18
Methylation levels B- lymphocytes (%)	12.31	13.28	19.79	16.08	14.13	58.54	53.44	48.03
Methylation levels monocytes (%)	2.75	3.42	6.69	8.07	7.60	56.98	44.08	43.49
Methylation levels granulocytes (%)	4.58	5.71	9.89	9.37	9.18	59.51	44.92	44.33
PBL- MNC	**P=0.036***	**P=0.012***	**P=0.012***	**P=0.036***	P=0.262	P=0.398	P=0.149	P=0.225
Lymphocytes- monocytes	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	P=0.237	P=0.185	P=0.090
Lymphocytes- granulocytes	**P=0.012***	**P=0.017***	**P=0.012***	**P=0.012***	P=0.093	P=0.398	P=0.194	P=0.064
T-lymphocytes- B-lymphocytes	**P=0.012***	**P=0.012***	**P=0.036***	**P=0.327**	**P=0.012****	**P=0.043***	P=0.243	**P=0.017***
Monocytes- granulocytes	**P=0.017***	**P=0.017***	**P=0.012***	**P=0.05***	P=0.208	P=1.000	P=0.541	P=0.445

Table 5.4: Methylation levels differ between cell types in the adult population

Significant differences in methylation levels were observed between the different cell types in the adult population (n=8) for seven of the genes under study (Paired samples T-test for the genes *HOXA4* and *IGF2*, Wilcoxon signed rank test for all the other genes under study (SPSS)).

A complete overview of the statistical analyses performed on the methylation levels between all cell types can be found in Appendix C, Table C.2.

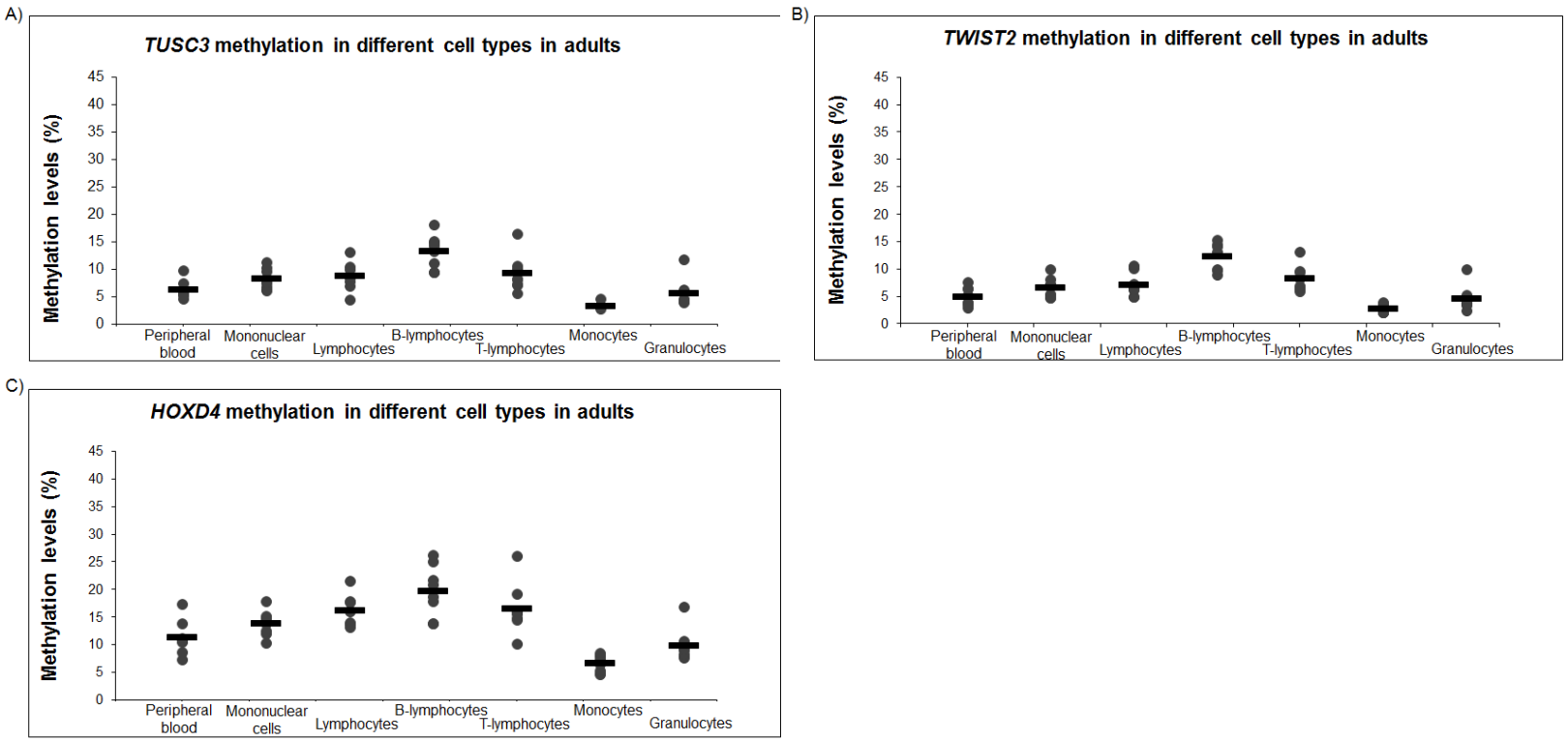


Figure 5.3A-D: Cell type specific methylation patterns in adult population for *TUSC3, TWIST2* and *HOXD4*

A-C) For *TUSC3*, *TWIST2* and *HOXD4*, the cells from the lymphoid lineage showed significantly higher levels of methylation than the cells from the myeloid lineage. Significant differences were observed between the isolated cell types per lineage; the B-lymphocytes showed significantly higher levels of methylation than the T-lymphocytes; and the granulocytes showed significantly higher levels of methylation than the monocytes (Table 5.4).

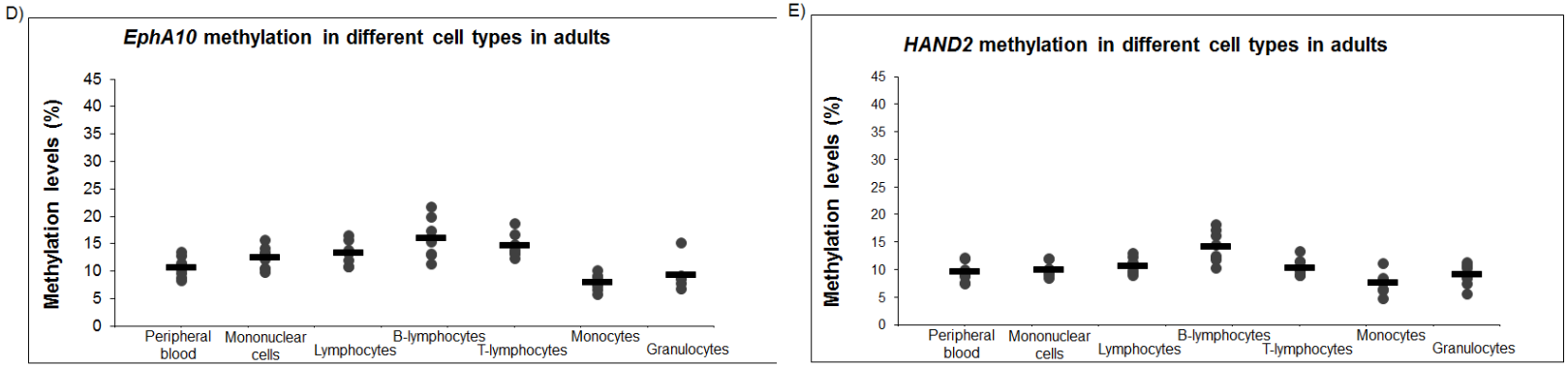


Figure 5.3D and E: Cell type specific methylation patterns in adults for *EpHA10* and *HAND2*

D) For the *EphA10* gene, the cells from the lymphoid lineage showed significantly higher levels of methylation than the cell from the myeloid lineage. No significant differences were observed between the cell types from the lymphoid lineage, whereas significant differences were observed between the isolated cell types from the myeloid lineage; the granulocytes showed significantly higher levels of methylation than the monocytes (Table 5.4).

E) For the *HAND2* gene, significantly higher levels were observed in the isolated B-lymphocytes compared to the isolated cell types from the myeloid lineage and the T-lymphocytes showed significantly higher levels of methylation than the monocytes (Table 5.4).

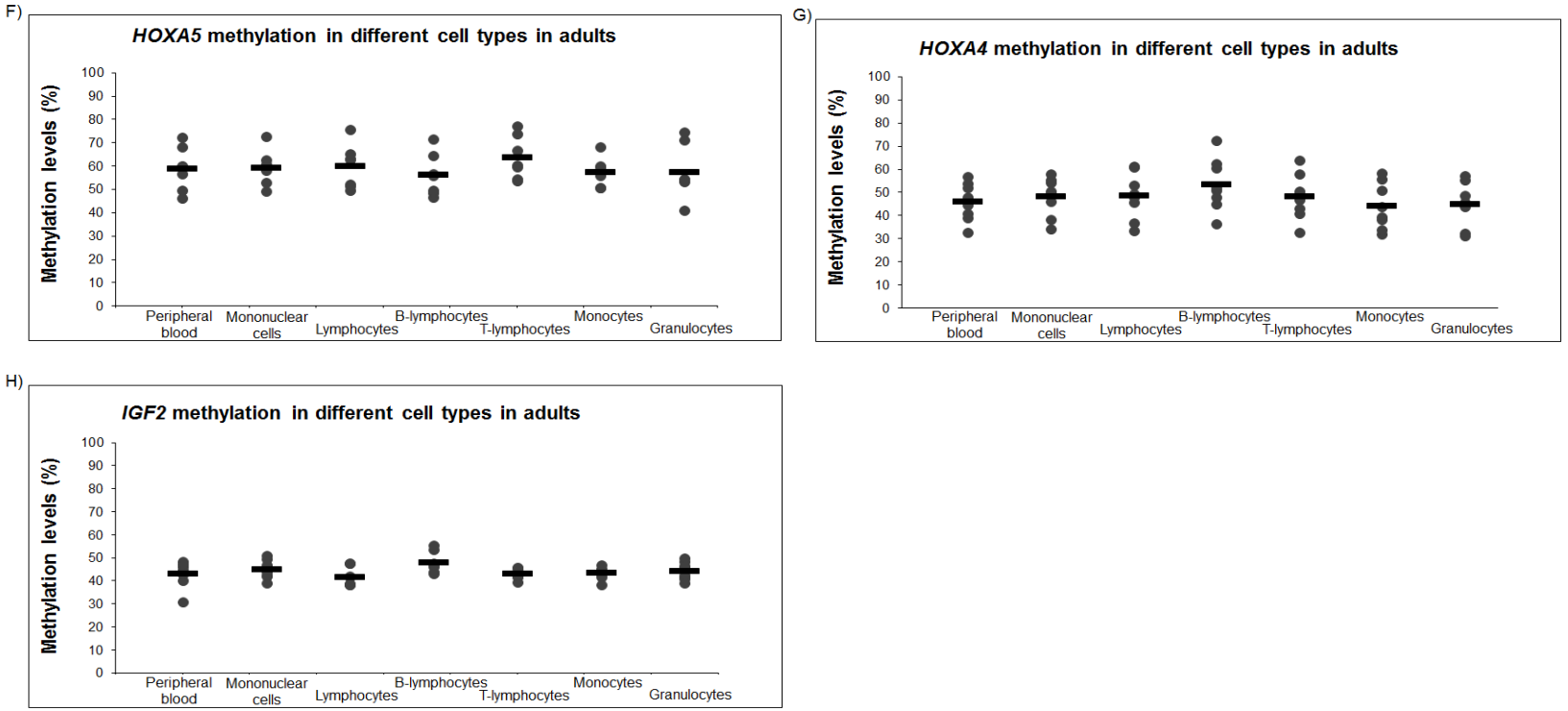


Figure 5.3F-H: Cell type specific methylation patterns in adults for *HOXA4*, *HOXA5* and *IGF2*

F) For the *HOXA5* gene, significantly higher levels of methylation were observed in the T-lymphocytes compared to the other isolated cell types: the B-lymphocytes, monocytes and granulocytes (Table 5.4).

G) For the *HOXA4* gene, there were only significant differences observed between the B-lymphocytes with the lymphocytes, T-lymphocytes and the monocytes (Table 5.4).

H) For the *IGF2* gene, significantly higher levels of methylation were observed in the isolated B-lymphocytes compared to the other isolated cell types: T-lymphocytes, monocytes and granulocytes (Table 5.4).

5.3.4 Age-related methylation changes in the different cell types

To examine the age-related differences in methylation levels in individual cell types, the methylation levels for each individual cell type was compared between the young adults and the adults. The main aim of this comparison was to examine whether the changes that we observed in PBL during ageing occurred in all the different cell types, or in just a subset of the cell types and as a result, whether age-related methylation must occur in the stem cells or if age-related methylation mainly occurs during the process of differentiation.

Five of the genes under study; *TUSC3*, *TWIST2*, *HOXD4*, *EphA10* and *HAND2*, showed similar patterns of both age-related and cell types specific methylation differences, and similar patterns of methylation in the specific cell types was observed in both the young adults and the adults. A general increase in methylation levels were observed for these genes during ageing in PBL and in the lymphocytes. In contrast, the isolated B-lymphocytes did not show consistent increases in methylation at the five loci. In fact, for three of the five loci, *TWIST2*, *TUSC3* and *HOXD4*, methylation levels were reduced in the older age group for this cell type (Table 5.5 and Figure 5.4A-E). The apparent drop in methylation of the B-lymphocytes may be a little misleading as the young adult group contained a single outlier sample with an especially high level of B-lymphocyte methylation. If the comparison is done using the median average (as opposed to the mean) then the apparent reduction in B-lymphocyte methylation is largely lost (Appendix C, Table C.3). However there is still a lack of a clear increase in the levels of methylation in the B-lymphocytes across the 5 genes.

In the monocytes, no differences in methylation levels between the two age groups were observed for *TWIST2* and *HOXD4*, a slight increase was observed for *EphA10* and *HAND2*, while significantly lower levels of methylation was observed in the monocytes for the *TUSC3* gene (Table 5.5 and Figure 5.4A-E).

The *HOXA4* and *HOXA5* genes both showed increased levels of methylation in the adults compared to the young adults in most of the different cell types analysed, with the exception of methylation levels of the B-lymphocytes for *HOXA5*. For this gene, the methylation levels for the B-lymphocytes did not change between the two age groups (Table 5.5 and Figure 5.4F-G).

The *IGF2* gene did not show any sign of differences in methylation levels of the specific cell types between the two age groups (Table 5.5 Figure 5.4H)

In this study, we examined patterns of DNA methylation in different cell types during ageing to examine whether the age-related methylation patterns are restricted to a specific cell type or if there is a more general increase in methylation in all the different cell types. All the blood cells develop from the same pool of stem cells, so to explain the differences in methylation levels between the cell types, the alterations in methylation levels have to occur during haematopoietic differentiation. Ageing may result in further differences in methylation levels between cell types by loss of fidelity of replication of methylation patterns during the differentiation process. However, in our study, there was no evidence to support the hypothesis that an increase in methylation occurred mainly in a specific cell type, arguing against this potential explanation. Instead, our results are more compatible with the hypothesis that altered methylation during ageing accumulates in the stem cells, although our sample numbers were limited, and a greater number of samples have to be examined to draw more definite conclusions.

		TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	1st 5 genes combined	*HOXA5*	*HOXA4*	*IGF2*
PBL	Methylation levels young adults (%)	4.86	6.62	10.07	8.27	8.80	7.72	53.57	38.50	43.93
	Methylation levels adults (%)	5.00	6.39	11.32	10.78	9.58	8.61	61.22	46.09	42.97
	Young adults- Adults	P=0.929	P=0.594	P=0.328	**P=0.004****	P=0.424	P=0.328	P=0.172	P=0.076	1.00
MNC	Methylation levels young adults (%)	5.61	8.48	12.50	9.80	8.45	8.97	52.62	39.62	43.98
	Methylation levels adults (%)	6.61	8.29	13.88	12.49	10.07	10.27	62.03	48.33	44.97
	Young adults- Adults	P=0.310	P=0.790	P=0.248	**P=0.010****	**P=0.026***	P=0.131	P=0.051	**P=0.033***	P=0.534
Lymphocytes	Methylation levels young adults (%)	6.51	8.48	13.48	10.73	8.26	9.49	54.41	41.23	45.19
	Methylation levels adults (%)	7.09	8.79	16.27	13.34	10.67	11.23	62.60	48.60	41.60
	Young adults- Adults	P=0.790	p=0.424	**P=0.026***	**P=0.006****	**P=0.010****	P=0.041*	P=0.430	P=0.155	P=0.083
B-lymphocytes	Methylation levels young adults (%)	13.74	16.50	22.02	15.80	10.87	15.79	56.31	44.58	47.43
	Methylation levels adults (%)	12.31	13.28	19.79	16.08	14.13	15.12	58.54	53.44	48.03
	Young adults- Adults	P=0.722	P=0.286	P=0.594	P=0.929	**P=0.026***	P=0.534	P=0.845	P=0.091	P=0.773
Monocytes	Methylation levels young adults (%)	3.14	5.07	6.79	6.89	6.34	5.65	51.29	36.82	43.41
	Methylation levels adults (%)	2.75	3.42	6.69	8.07	7.60	5.70	56.98	44.08	43.49
	Young adults- Adults	P=0.182	**P=0.001*****	P=1.000	P=0.091	P=0.131	P=0.897	P=0.064	P=0.183	P=0.929

Table 5.5: Methylation levels of the specific cell types in the young adult population compared to the adult population

Statistical differences were observed between the young adults (n=10) and the adults (n=8) in a small number of genes (Wilcoxon rank sum tests (SPSS)).

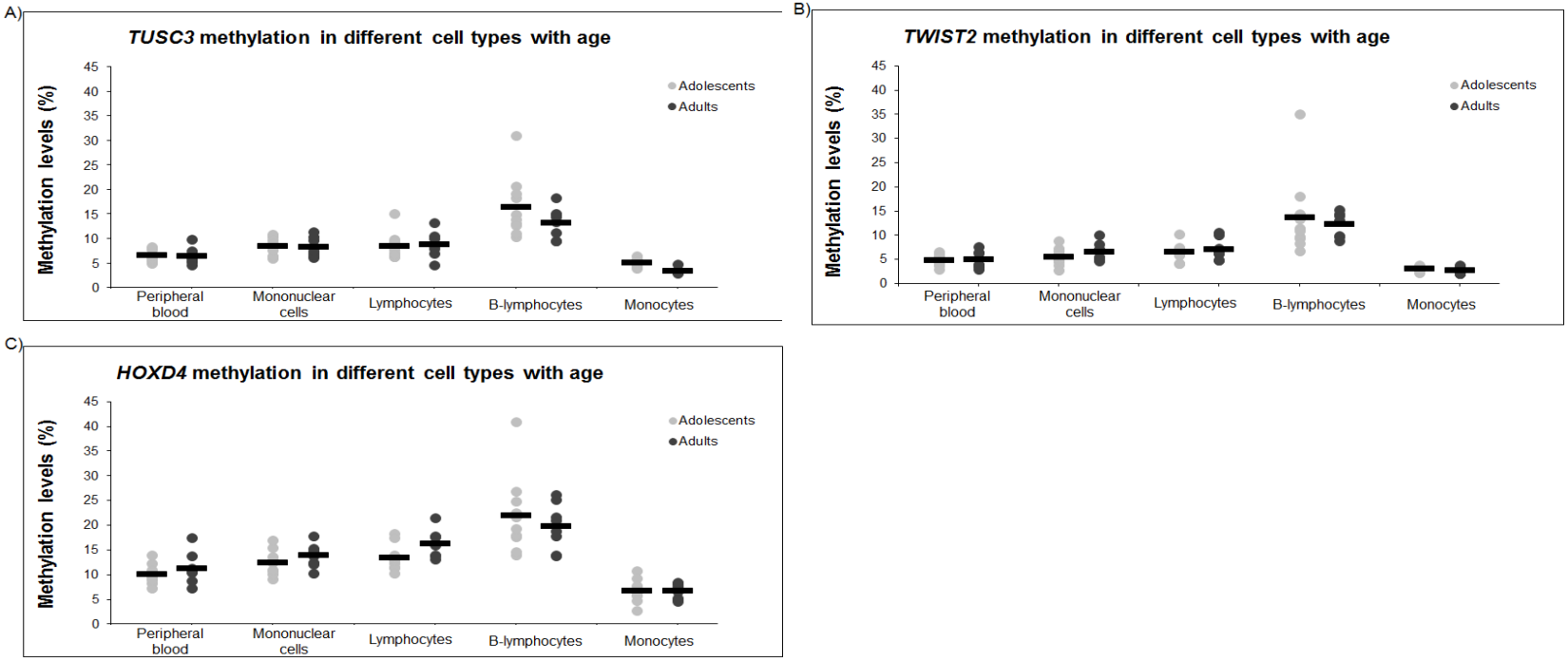


Figure 5.4A-C: Cell type specific methylation patterns with an increasing age for *TWIST2*, *TUSC3* and *HOXD4*

A) No differences in methylation levels were observed between the two age groups for the *TUSC3* gene for the PBL, MNC and lymphocyte samples. The B-lymphocytes and the monocytes showed lower levels of methylation in the older age group, although this difference was only significantly lower in the monocytes (Table 5.5).

B) No significant differences were observed in the methylation of the isolated cell types within the different age-groups for *TWIST2*.

C) An increase in methylation levels with an increasing age were observed in PBL, MNC and the lymphocytes for the *HOXD4* gene, although this difference was only statistically significant in the lymphocytes (Table 5.5).

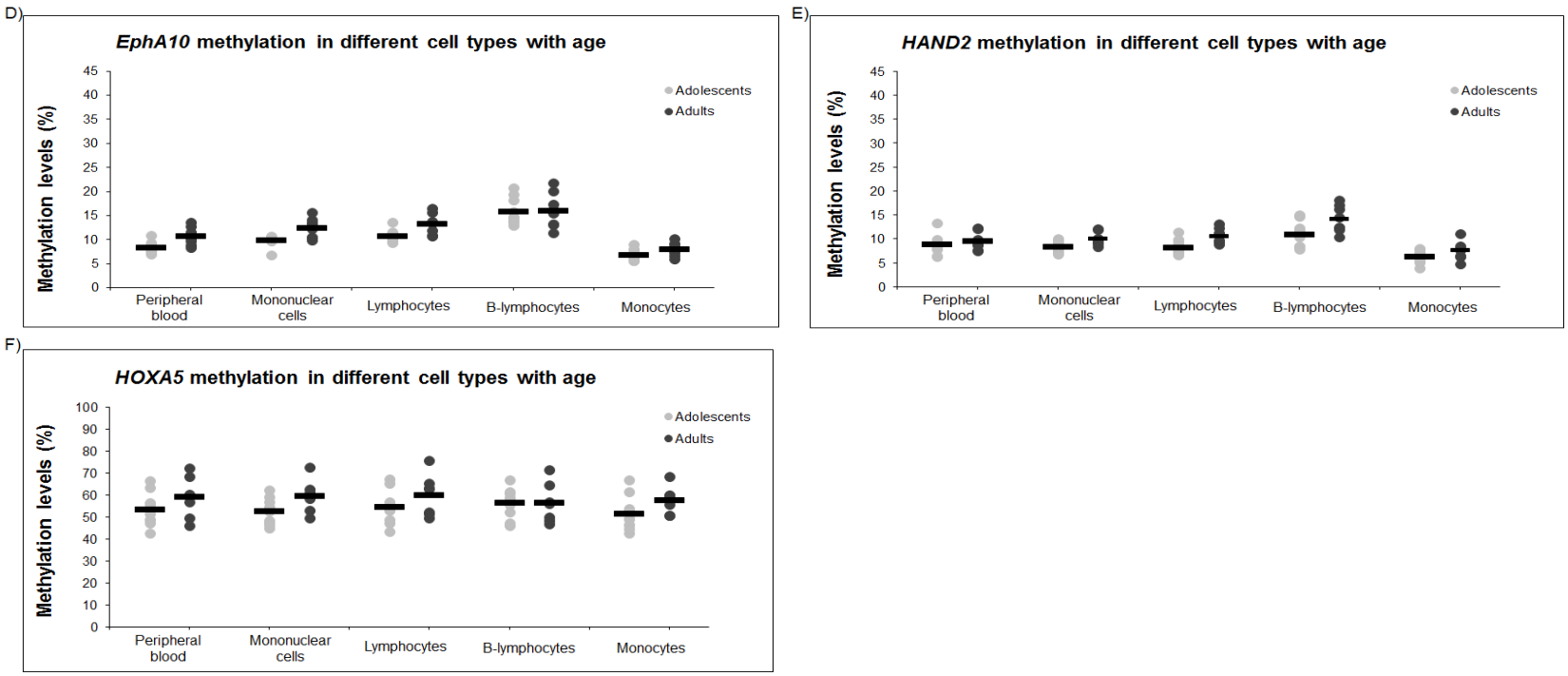


Figure 5.4D-F: Cell type specific methylation patterns with an increasing age for *EphA10*, *HAND2* and *HOXA5*

D) For the *EphA10* gene, significant increased levels of methylation were observed in PBL, MNC and the lymphocytes in the adults compared to the young adults (Table 5.5).

E) Increases in methylation levels were observed in all the different cell types under study with an increasing age for the *HAND2* gene, although this increase was only statistically significant for the MNC, Lymphocytes and the B-lymphocytes samples (Table 5.5).

F) No significant differences in methylation levels in the different cell types were observed within the different age groups for *HOXA5*.

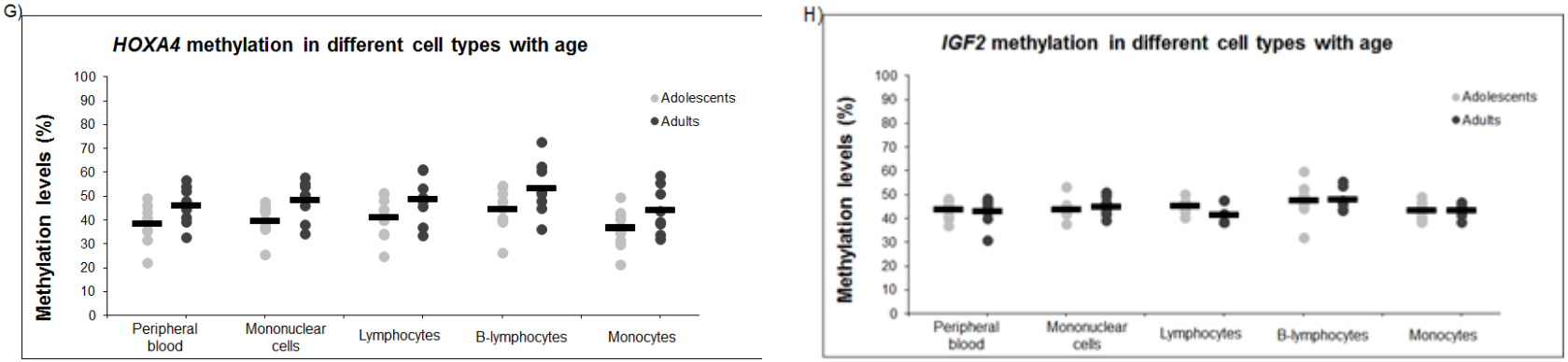


Figure 5.4G and H: Cell type specific methylation patterns with an increasing age for *HOXA4* and *IGF2*

G) *HOXA4* showed higher levels of methylation in all the different cell types under study with an increasing age, although this difference was only statistically significant for the MNCs (Table 5.5).

H) No significant differences in methylation levels in the different cell types were observed within the different age groups for *HOXA5* and *IGF2* (Table 5.5).

5.3.5 Methylation levels in stem- and early progenitor cells

Methylation levels were assessed in isolated stem/progenitor cells ($CD34^+$), early B-cell progenitors ($CD34^+/CD19^+$) and in differentiated cells from a number of G-CSF mobilized samples (isolation as described in section 5.2). This was done to examine whether the differences in methylation patterns in the specific cell types, as observed in the cells isolated from peripheral blood, arise during the process of differentiation, or if these cell type specific methylation patterns are already present in the $CD34^+$ cells. If the differences observed between the different isolated differentiated cells were due to increased lymphoid differentiation of stem cells with increased levels of CpG island methylation, then such differences would be present at all stages of lymphoid differentiation.

Analyses of methylation levels of the isolated $CD34^+$ cells of the five genes that showed clearest differences in methylation patterns within different cell types; *TUSC3*, *TWIST2*, *HOXD4*, *EphA10* and *HAND2*, revealed that $CD34^+$ cells exhibited methylation levels comparable to the methylation levels as observed in the monocytes. No significant differences were observed between the methylation levels of the $CD34^+$ population as a whole and the isolated $CD34^+/CD19^+$ B-cell progenitors for any of the genes under study.

Unexpectedly though, the cell-type specific methylation patterns of the differentiated cells in the G-CSF mobilized samples was clearly different from that seen in isolated cells from peripheral blood samples, most obviously for the differentiated B-lymphocytes. For example, the average methylation across the five age-related genes in the B-lymphocytes from the sorted PBL samples was 15.8% and 15.12% for the younger adults and adults respectively. However, for the sorted G-CSF mobilised samples the average B-cell methylation was much lower (8.0%). The overall pattern of methylation was somewhat conserved, as lymphoid cells in general and B-lymphocytes in particular retained methylation levels higher than in myeloid (monocyte) cells, however the differences were greatly reduced.

No differences in methylation levels were observed between the B-lymphocytes and the T-lymphocytes in the G-CSF mobilized samples. However, both the B-lymphocytes and the T-lymphocytes showed higher levels of methylation than the monocytes for four of these genes (Table 5.6 and Figure 5.5A-E).

For the other two genes under study: *HOXA4* and *HOXA5*, there were no significant differences in methylation levels between all the different cell types isolated from these G-CSF mobilized samples. However, for these two genes, increased levels of

methylation were observed in the $CD34^{+}CD19^{+}$ B-cell progenitor cells, compared to any of the other cell types, including the $CD34^{+}$ cells and the B-lymphocytes. For example for *HOXA5*, average methylation for the $CD34^{+}CD19^{+}$ B-cell progenitor cells were 86.14%, compared to 63.19% and 67.66% for the $CD34^{+}$ cells and the B-lymphocytes respectively (Table 5.6 and Figure 5.5A-E).

These results show that methylation levels do not differ between the stem-and early progenitors as a whole, and the early B-cell progenitors, suggesting that the high levels of methylation observed in the B-lymphocytes are not already present at the first stage of lineage-commitment. However, these results are a bit preliminary due to the low number of samples and due to the unexpected low methylation of the differentiated lymphocytes.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*
Methylation levels stem-and early progenitor cells ($CD34^+$) (%)	4.04	3.83	8.93	7.06	7.32	63.19	35.00
Methylation levels early B-cell progenitors ($CD34^+CD19^+$)	6.71	2.68	8.44	7.26	7.14	86.14	44.47
Methylation levels Stem- and early progenitor cells minus B-cell progenitors ($CD34^+CD19^-$) (%)	4.59	3.00	8.67	6.32	7.82	77.47	32.11
Methylation levels B-lymphocytes (%)	6.70	5.59	10.80	9.13	7.88	67.66	37.03
Methylation levels T-lymphocytes(%)	5.87	5.28	10.26	8.95	7.75	65.04	38.33
Methylation levels monocytes (%)	4.04	3.83	12.4	7.06	7.32	63.19	35.00
Stem-and early progenitor cells- B-cell progenitors	P=0.600	P=0.232	P=0.893	P=0.204	P=0.893	P=0.109	P=0.651
B-cell progenitors- B-lymphocytes	P=0.752	**P=0.038***	P=0.345	P=0.190	P=0.686	P=0.109	P=0.903
B-lymphocytes- T-lymphocytes	P=0.173	P=0.689	P=0.463	P=0.800	P=0.463	P=0.225	P=0.432
B-lymphocytes- Monocytes	**P=0.028***	**P=0.020***	P=0.075	**P=0.025***	P=0.463	P=0.500	P=0.108
T-lymphocytes- Monocytes	**P=0.028***	**P=0.013***	**P=0.028***	**P=0.001*****	P=0.345	P=0.225	P=0.183

Table 5.6: Methylation in the different cell types isolated from the G-CSF mobilized samples

Statistical differences were observed between different cell types in a small number of genes in the cells isolated from the G-CSF mobilized samples (n=6, Paired samples T-test for *TUSC3*, *EpHA10* and *HOXA4*, Wilcoxon signed rank test for the other genes under study (SPSS)).

A complete overview of the statistical analyses performed on the methylation levels between all cell types can be found in in Appendix C, Table C.4.

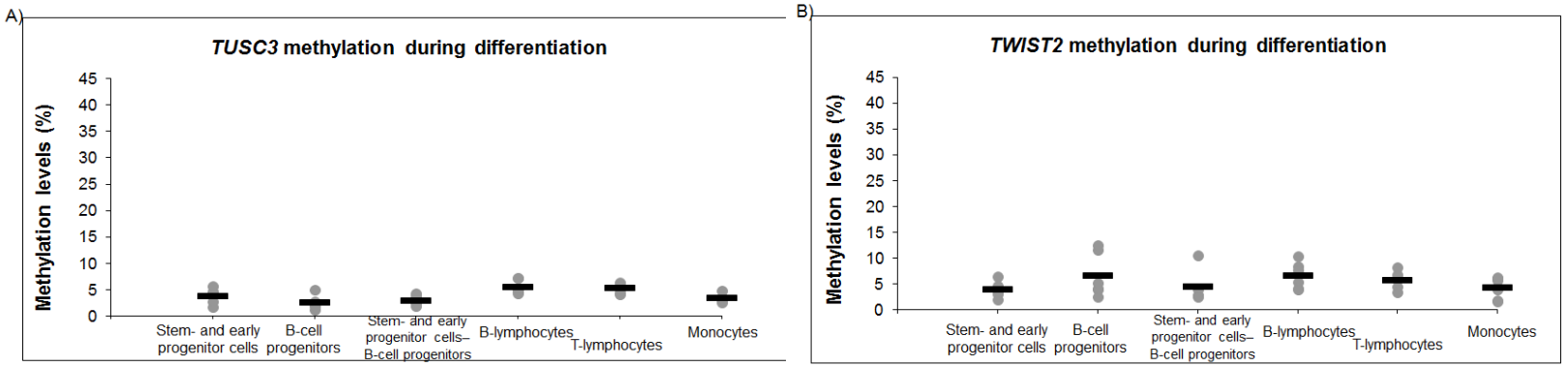


Figure 5.5A and B: Methylation patterns during cell differentiation for *TUSC3* and *TWIST2*

A) Analyses of *TUSC3* methylation patterns in the G-CSF mobilized samples revealed that significant differences were observed between methylation patterns of the B-lymphocytes and $CD34^+$ the stem- and early progenitor cells. Methylation patterns in the B-lymphocyte population as a whole was significantly higher than the methylation patterns in the $CD34^+CD19^+$ B-cell progenitors and both the B-lymphocytes and T-lymphocytes methylation levels were significantly higher than the methylation levels of the monocytes (Table 5.6).

B) *TWIST2* methylation levels were significantly different between the $CD34^+$ stem- and early progenitor cells and the T-lymphocytes and between both the B- and T-lymphocytes and the monocytes in the G-CSF mobilized samples (Table 5.6).

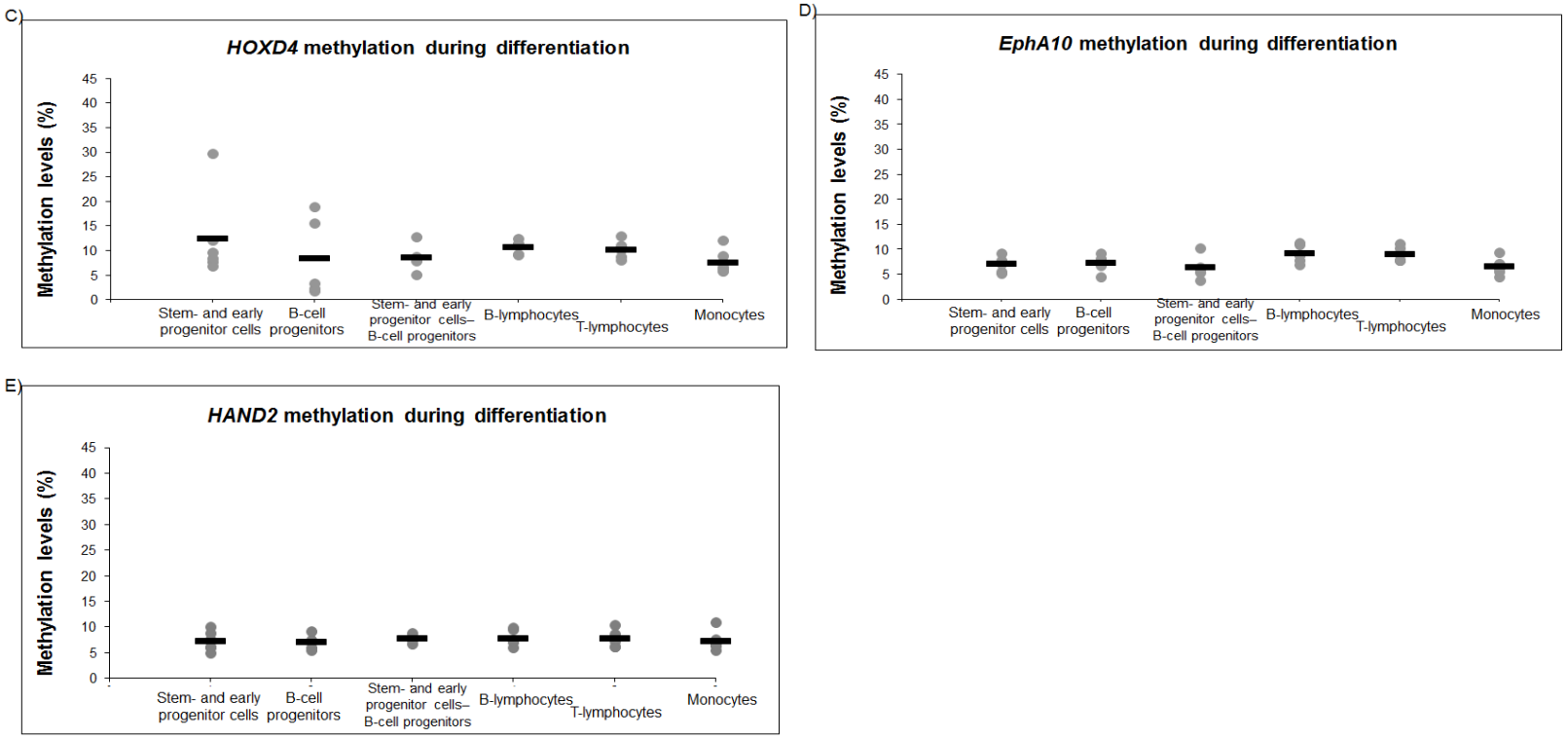


Figure 5.5C-E: Methylation patterns during cell differentiation for *HOXD4, EphA10* and *HAND2*

C) No differences in methylation levels were observed between the $CD34^+$ stem- and early progenitor cells and the cells later during differentiation for the *HOXD4* gene. A significant difference was observed between the differentiated T-lymphocytes and monocytes (Table 5.6).

D) *EphA10* methylation levels were significantly different between the $CD34^+$ stem- and early progenitor cells and the T-lymphocytes and between both the B-lymphocytes and the T-lymphocytes and the monocytes in the G-CSF mobilized samples (Table 5.6).

E) No differences in methylation levels were observed between the stem-and early progenitor cells and the cell later during differentiation or between the different differentiated cell types for *HAND2*.

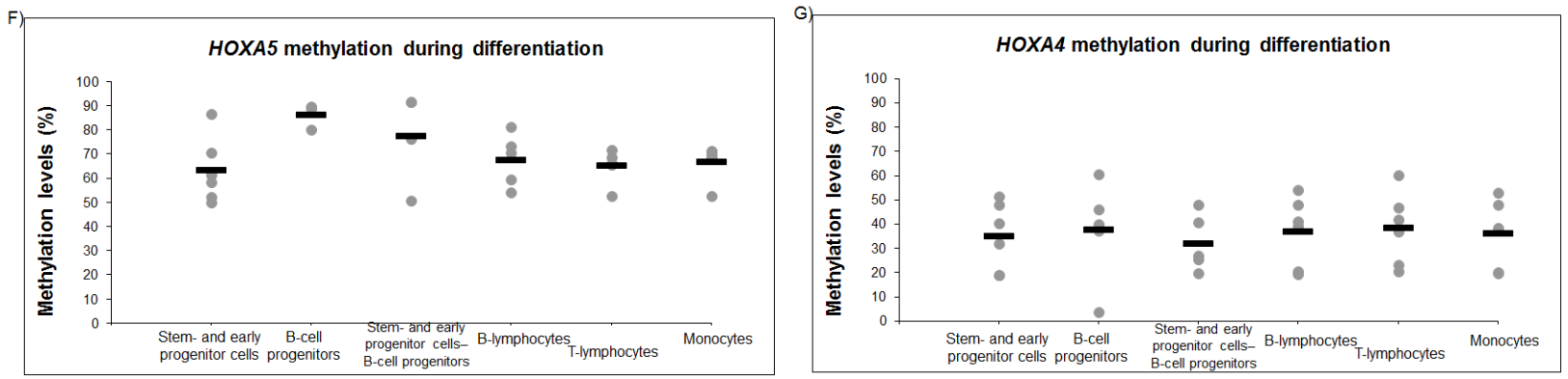


Figure 5.5F and G: Methylation patterns during cell differentiation for *HOXA5* and *HOXA4*

F and G) No differences in methylation levels were observed between any of the isolated cells types for the *HOXA5* and the *HOXA4* genes.

5.3.6 Correlations between genes in the specific cell types

For five of the genes under study, *TWIST2, TUSC3, EphA10, HOXD4* and *HAND2*, correlations between the methylation levels were observed in PBL from healthy individuals. In the specific cell types, again very clear similarities in the patterns of DNA methylation were seen for these five genes. Therefore, the isolated cell types were examined for potential correlations between these five genes, to examine whether co-methylation occurred within the same populations of cells.

In the young adults, strong correlations were observed in the methylation levels of the B-lymphocytes between these five genes, with the exception of *HAND2* methylation with the methylation levels of *TUSC3* and *HOXD4* (Table 5.7).

These results suggest that methylation patterns are mainly driven by co-methylation of genes within the same population of cells and not by inter individual differences in the cellular make-up of peripheral blood. This is also consistent with results obtained in a related study in the Newcastle 85+ study population which demonstrated that after correcting for differences in the % lymphocytes in individual samples the strong co-methylation between the genes still remained (Gautrey, 2012).

	TWIST2	TUSC3	HOXD4	EphA10	HAND2	HOXA5	HOXA4
TWIST2		**R= 0.876** **P= 0.001*****	**R= 0.920** **P<0.001*****	**R= 0.845** **P= 0.002****	**R= 0.691** **P= 0.027***	R= -0.098 P= 0.775	R= 0.200 P= 0.555
TUSC3			**R= 0.842** **P= 0.002****	**R= 0.904** **P< 0.001*****	R= 0.568 P= 0.087	R= -0.343 P= 0.301	R= 0.290 P= 0.387
HOXD4				**R= 0.837** **P= 0.003****	R= 0.6121 P= 0.060	R= -0.471 P= 0.143	R= 0.374 P= 0.257
EphA10					**R= 0.683** **P= 0.029***	R= -0.138 P= 0.685	R= 0.295 P= 0.378
HAND2						R= 0.013 P= 0.970	R= 0.415 P= 0.204
HOXA5							R= -0.419 P= 0.200
HOXA4							

Table 5.7: Correlations between the methylation levels of genes in the isolated B-lymphocytes from the young adults

In the B-lymphocytes, correlations were observed between *TWIST2, TUSC3, HOXD4* and *EphA10* and between *HAND2* with *TWIST2* and *EphA10* in the young adults (n=10)

The test used for these correlations is a Pearson's correlation (SPSS). The R is indicating the correlation coefficient and the observed P the p-values.

The correlations between these genes in the isolated lymphocyte and monocyte samples from the young adult participants can be found in Appendix C (Table C.5).

5.4 Discussion

Our study showed that clear cell-type specific methylation patterns were present in peripheral blood in a subset of genes. It revealed that for seven of the genes under investigation methylation levels were significantly higher in cells from the lymphoid lineage (T-lymphocytes and B-lymphocytes) compared to the cells from the myeloid lineage (monocytes and granulocytes). Differences in methylation levels were not only observed between the cell types from the myeloid and lymphoid lineages, but even in cell types from the same lineage. All these blood cells derive from the same pool of stem cells, resulting in three possible hypotheses that might explain how these cell-type specific methylation differences occur; first of all, methylation differences might accumulate in stem cell populations and stem cells with higher levels of CpG island methylation might preferentially differentiate into lymphocytes. Secondly, accumulation of 'seeds' of methylation might occur in stem cells which are expanded during lymphoid differentiation only, and lastly cell-type specific methylation differences might arise during cell differentiation.

In case of the first hypothesis, which states that stem cells with higher levels of methylation preferentially differentiated into lymphocyte lineages, increased levels of methylation should already be present in early B-cell progenitors. To test this hypothesis, $CD34^+$ stem- and early progenitor cells were isolated from a set of G-CSF mobilized samples, as well as $C34^+CD19^+$ early- B-cell progenitors. Analyses of these two cell populations did not show differences in methylation levels between the $C34^+CD19^+$ B-cell progenitors and the $CD34^+$ population as a whole, arguing against a mechanism involving preferential lymphoid differentiation of stem cells with increased CpG island methylation. However, the unexpectedly low levels of methylation in these samples in the differentiated B-lymphocytes suggest that further confirmation of these results might be useful.

Interestingly, the cell type that showed highest levels of methylation for the genes under study, the B-lymphocytes, is the cell type that is longest lived. This suggests that the observed differences between the cell types might occur during the 'ageing' of the cell during differentiation. Analyses of both the $C34^+CD19^+$ early B-cell progenitors and B-lymphocytes did not show differences in methylation levels between these two cell populations. However, the methylation levels observed for the B-lymphocytes in these G-CSF mobilized samples were significantly lower than the methylation levels of B-lymphocytes isolated from peripheral blood in individuals in a similar age-range. It is possible that, since these G-CSF mobilized samples are enriched

for early progenitor cells, they therefore contain more early B-lymphocytes than a normal blood sample. This might result in lower overall methylation levels in these samples if the methylation levels observed in B-lymphocytes occur during the process of differentiation. However, this means that the levels of early B-cell progenitors must be high to explain the large differences in methylation levels. For example, overall methylation levels in the B-lymphocytes for the *TWIST2* gene was 13.75% in the cells isolated from blood, while this was 6.70% for this cell type in the G-CSF mobilized sample. Another possible explanation is that the B-lymphocytes isolated from the G-CSF mobilized samples underwent less rounds of cell division than the B-lymphocytes isolated from blood and that those cells are therefore biologically younger. These observations are in line with the third hypothesis, which states that cell type specific methylation patterns might occur during the process of differentiation. Differentiated cells differ in their life span, where the B-lymphocytes are the cell type that is longest lived. It will be interesting to further investigate the effect of this expanded life span in the differentiated cells in subsequent studies, to study the full effect of ageing on different cell types during differentiation and to address to question whether the observed differences observed in our study could be explained by the longevity of the B-lymphocytes.

It is possible that ageing may result in further differences in methylation levels between cell types, either just by random drift or because different cell types may be differentially sensitive to age-related influences on the epigenome. To investigate how ageing affects these cell type specific methylation patterns and how the cell-type specific methylation patterns related to the accumulation of age-related DNA methylation in PBL, monocytes, lymphocytes and B-lymphocytes were isolated from adults from different age groups; young adults (with an average age of 27.6 years) and older adults (with an average age of 51.6 year). Small differences in methylation levels were observed between these two age groups for the cell types investigated even in our small group of individuals. The number of samples was low and the methylation differences small, such that some of the differences in methylation fell short of statistical significance. However, these results do suggest that age-related methylation differences occur in several cell types and that they are not occurring in one cell type specifically. This suggest that age-related differences in methylation levels observed for these genes, must occur in a pool of stem cells and that these do not solely occur during the process of differentiation.

The observed differences in methylation levels are of interest in understanding methylation differences during disease development. The inability to maintain methylation levels in stem cells during ageing might be an underlying factor of the increase of incidence of cancer, and potentially other disease, with an increasing age. The CpG island methylation patterns observed here in the isolated lymphoid and myeloid cell types from healthy individuals were mirroring patterns observed during leukaemia development. The genes under study all played a role in lymphoid leukaemia with the exception of both *HOXA5* and *HOXA4*, which are hypermethylated in both myeloid and lymphoid leukaemias, and our results showed that even in healthy individuals, the lymphoid cells contained higher levels of methylation than the myeloid cells for this set of genes. This study also showed that the B-lymphocytes exhibited the highest levels of methylation for the genes investigation compared to any of the other cell types, including the T-lymphocytes. Interestingly, most haematological malignancies are B-lymphocyte derived, again suggesting a link between high levels of methylation and the development of a haematological malignancy. The expression levels of these genes under study were not investigated in these samples. However, even though B-lymphocytes exhibited the highest levels of DNA methylation, a study previously performed in our group showed that, at least three genes under study; *TWIST2*, *TUSC3* and *HOXA4*, were expressed in B-lymphocytes (Thathia, 2012; Strathdee, 2006). This suggests that the high accumulation of DNA methylation observed for these genes in the B-lymphocytes does not only occur in genes that are not expressed in that cell type, in which case it would have had little or no significant biological impact. Instead, the accumulation of DNA methylation in these genes that are expressed in this cell type, suggests that the methylation is not restricted to non-expressed genes and thus may well affect gene expression patterns.

Overall, the results from this study showed that cell-type specific methylation patterns are present in healthy individuals and that these patterns mirror patterns observed during leukaemia development. These patterns are maintained during ageing, and age-related methylation changes were observed in the cell types for the set of genes under investigation. The results shown here are consistent with the hypothesis that low levels of DNA methylation accumulate in stem cells during ageing and that these low levels of methylation are then expanded in lymphoid, but not myeloid, cells. At least for the genes investigated here, the accumulation of age related CpG island methylation is predominantly present in cells of lymphoid origin. This supports the hypothesis that age-related DNA methylation may be a key factor in underlying disease susceptibility,

by the inability of stem cells to stably maintain DNA methylation patterns during ageing.

Chapter 6. DNA methylation patterns during the development of age-related diseases

6.1 Introduction

DNA methylation is most extensively studied in cancers, although changes in DNA methylation have been described in multiple other diseases, such as cardiovascular, neurological and metabolic disorders and autoimmune diseases (Udali, 2012; Jakovcevski, 2012; Bruce, 2011; Grolleau-Julius, 2010). However, while for cancers, a large body of evidence supports a key role for altered DNA methylation in cancer development, for other diseases, the functional significance of the observed changes in DNA methylation patterns is not clear.

The main aim of this study was to test whether the acquisition of altered DNA methylation during normal ageing may have a significant effect on the likelihood of disease development during the life-course, such that individuals who acquire higher levels of altered DNA methylation may be at increased risk of disease development. In order to test this hypothesis, methylation patterns associated with cancer were examined in two of the main age related diseases, cancer and heart disease. This chapter will be divided into two parts. In the first part, the changes in methylation patterns in cancer patients, more specifically leukaemia and HNPCC patients, will be discussed and in the second part the focus will be on changes in methylation patterns in heart disease patients.

6.2 Cancer

During cancer development, reduced levels of global DNA methylation are observed, together with local areas of hypermethylation, particularly at CpG islands in gene promoters (Gama-Sosa, 1983; Bird, 1996). Many of the genes targeted by hypermethylation play roles in tumour suppression. Hypermethylation during cancer development within those promoters serves to turn off these critical genes that otherwise suppress tumourigenesis (Baylin, 2005). Whilst many tumour suppressor genes targeted by hypermethylation during cancer development are tumour-type specific, other tumour suppressor genes are shared by multiple tumour types (Hansen, 2011), such as the *p16/CDKN2* gene (Esteller, 2005), the *RASSF1* gene and *HIC1* gene (Teng). These observations suggest a key role of DNA methylation in the development of cancer and this leads to the hypothesis that those individuals with increased levels of DNA methylation abnormalities may be at a higher risk of cancer development.

6.2.1 Leukaemia

ALL is the most common form of childhood cancer. ALL also occurs in the adult population, although it is a rarer disease in adults, compared with other tumour types. The survival rate between the childhood and adult ALL patients differs significantly. 80-90% of the childhood ALL patients survive the disease. However, ALL is still the 2nd leading cause of cancer mortalities in children (Cancer-Research, 2013). In the adult ALL patients, hematologic remission is obtained in over 90% of the patients, but the long-term survival rate (<5 years) is low (around 40%) (Goldstone, 2008). Therefore, it is important to improve the survival rate of ALL patients. Prediction of patients at high risk of developing leukaemia, or patients that are at high risk of relapse, will help detecting those patients who can benefit from additional treatment.

Hypermethylation of gene promoters is a feature of essentially all cancers, including leukaemia. This methylation can have dramatic consequences since hypermethylation of gene promoters leads to transcriptional silencing. Importantly, many genes with an established anti-tumourigenic role have been shown to be frequently methylated in cancer and promoter associated hypermethylation is believed to be one of the primary mechanisms for inactivating such genes.

Our previous results, as discussed in Chapter 4, suggested a link between genes showing methylation changes with age and genes becoming methylated during leukaemia development. This would be compatible with the hypothesis that the development of age-related changes in DNA methylation may be an important underlying cause of cancer development. Therefore in this part of the study, we investigated whether apparently normal cells from leukaemia patients in remission, exhibited increased levels of methylation and we examined the potential of aberrant patterns of DNA methylation in identifying patients that are at high risk of developing the diseases. This was done by investigating methylation levels during remission in both childhood and adult leukaemia patient in several genes related to leukaemia development.

6.2.1.1 Samples

DNA was purified from peripheral blood collected from 24 childhood ALL patients and 21 adult ALL patients. Samples taken at diagnosis, remission and where applicable, relapse samples were obtained. All samples were obtained from the

Newcastle Haematology Biobank and had been collected with the appropriate ethical approval.

Childhood ALL samples were taken from patients between 1 and 12 years of age, with an average age of 4.2. The adult ALL samples were taken from patients between 17 and 65 years of age, with an average age of 38.8. 14 of the adult ALL patients did not survive the disease, and 6 of them were classified as long-term survivors (> 5 years)(Table 6.1).

The control samples used for the childhood ALL patients in this study were the neonate participants as described in Chapter 4. The controls for the adult ALL patients were the young adults and the adult participants as described in Chapter 4 (Table 6.1).

	Childhood ALL patients		Adult ALL patients	
	Patients	**Controls**	**Patients**	**Controls**
Age (years)	<1-12 (average 4.2)	0 (average 0)	18-65 (average 37.0)	17-50 (average 39)
Number of patients	24	50	28	98
Gender	10 male 8 female 6 unknown	33 male 15 female 1 unknown	14 male 7 female 7 unknown	18 male 80 female
Leukaemia	18 Common 2 B-lymphocyte 3 T-lymphocyte 1 non T-lymphocyte		1 Common 4 B-lymphocyte 8 T-lymphocyte 5 Calla 1 Null 1 Unkown	
Patient outcome	22 non-relapse patients 2 relapse patients 24 long-term survivors		4 non-relapse patients 24 relapse patients 6 long-term survivors	
Available matched DNA samples	22 diagnosis-remission 2 diagnosis-remission-relapse		4 diagnosis-remission 17 diagnosis-remission-relapse 6 diagnosis-relapse	

Table 6.1: An overview of the ALL study population

6.2.1.2 Results

6.2.1.2.1 Relationship of remission methylation patterns with those seen at diagnosis and relapse in ALL patients.

Methylation levels were examined in matched diagnostic, remission, and where applicable relapse, samples from our set of both childhood and adult ALL patients (as detailed in Table 6.1). As expected, methylation levels of the seven loci known to exhibit very frequent methylation changes in ALL were very high in diagnostic samples as well as in the relapse samples. Whereas, in contrast, the non-leukaemia associated gene *IGF2* exhibited relatively stable methylation during both diagnosis and relapse. The methylation levels observed in the diagnostic and the relapse samples showed no significant differences for any of the genes examined. During remission, the methylation levels were significantly reduced compared to both the diagnostic and relapse samples for all the genes under study, with the exception of *IGF2* (Table 6.2 and Figure 6.1)

As detailed above, methylation levels in the remission samples were far lower than those seen in either the diagnostic or relapse sample sets. To examine if there was evidence between a link of the altered methylation patterns in the leukaemia samples and the methylation levels in the matched remission samples, the data was examined for possible correlations. No significant correlations were observed between methylation levels during diagnosis and remission (available for 21 adult samples and 24 childhood samples), with the exception of a limited correlation for *HAND2* in adult ALL. Similarly no significant correlations were observed between methylation levels in the remission samples and the methylation detected in the matched relapse samples (available for 17 adult samples). In contrast, correlations between the diagnostic and relapse samples were observed for 5 of the genes in the adult age group (available for 23 adult samples), namely *TWIST2, TUSC3, EphA10, HAND2*, *HOXA5* and *IGF2* (Table 6.3). In the childhood patients the number of relapsed patients was too small to perform the same statistical analyses.

A.

		TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Adult ALL	Methylation levels at diagnosis (%)	63.03	46.99	70.71	58.93	44.11	78.46	78.33	50.21
	Methylation levels during remission (%)	11.65	12.26	16.22	12.00	7.73	64.44	44.35	41.00
	Methylation levels at relapse (%)	72.38	47.83	74.59	64.15	53.13	81.96	77.62	49.22

B.

		TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Childhood ALL	Methylation levels at diagnosis (%)	51.79	29.60	62.13	45.38	37.26	67.47	65.88	44.82
	Methylation levels during remission (%)	5.77	5.80	9.15	6.72	4.53	59.22	39.10	40.33
	Methylation levels at relapse (%)	28.81	6.49	31.99	20.43	19.43	64.20	51.31	

Table 6.2: Methylation levels at diagnosis, remission and relapse

A) Average methylation levels at diagnosis, remission and relapse in the adult ALL patients (n=28) as assessed by pyrosequencing.

B) Average methylation levels at diagnosis, remission and relapse in the childhood ALL patients (n=24) as assessed by pyrosequencing.

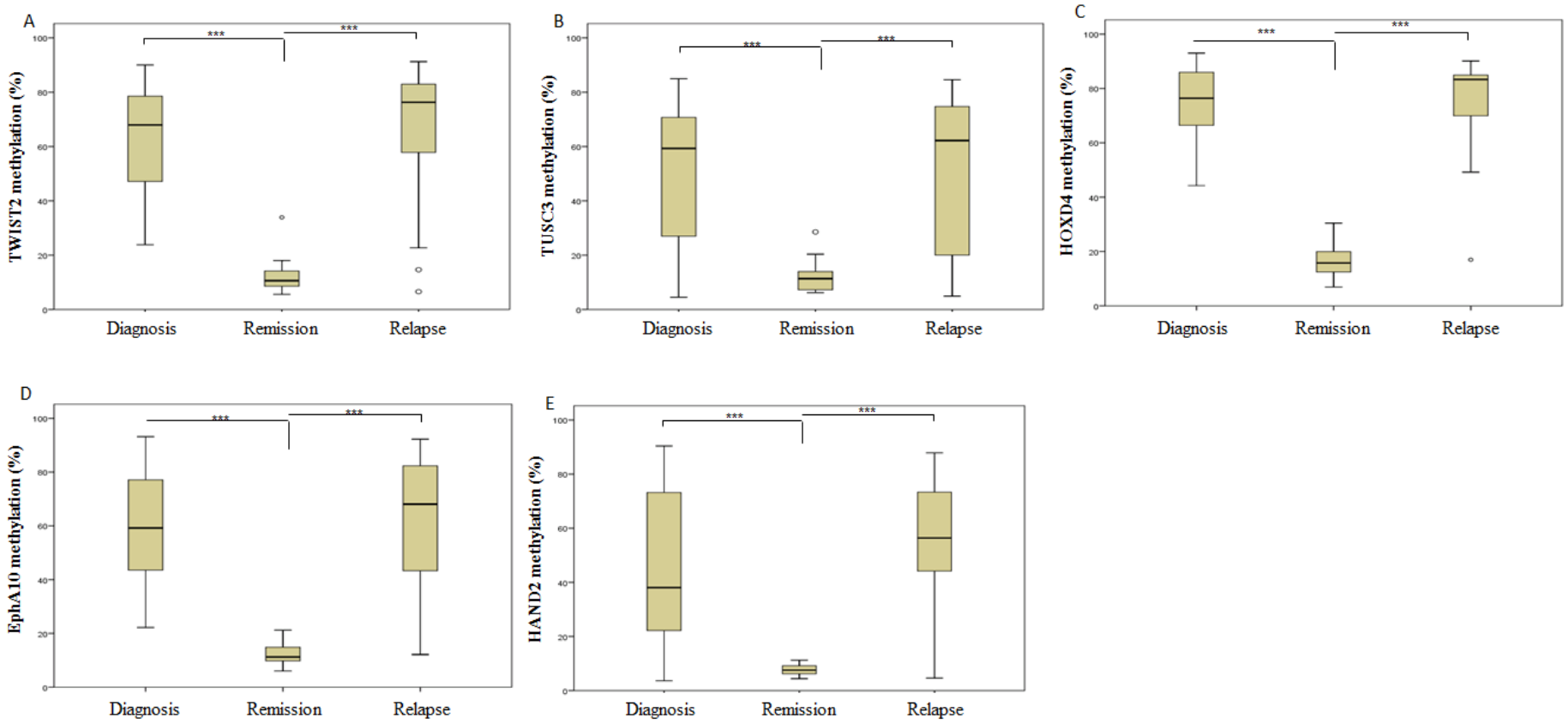


Figure 6.1A-E: Methylation levels during the different phases of adult ALL for *TWIST2, TUSC3, HOXD4, EphA10* and *HAND2*

A-E) The methylation levels in the paired adult ALL samples were significantly reduced during remission compared to both the diagnostic and relapse phase for *TWIST2*, *TUSC3* , *HOXD4*, *EphA10* and *HAND2* ($p<0.001$ for all genes). No significant differences were observed in methylation levels during diagnosis and relapse for any of these genes (Wilcoxon signed rank test (SPSS)).

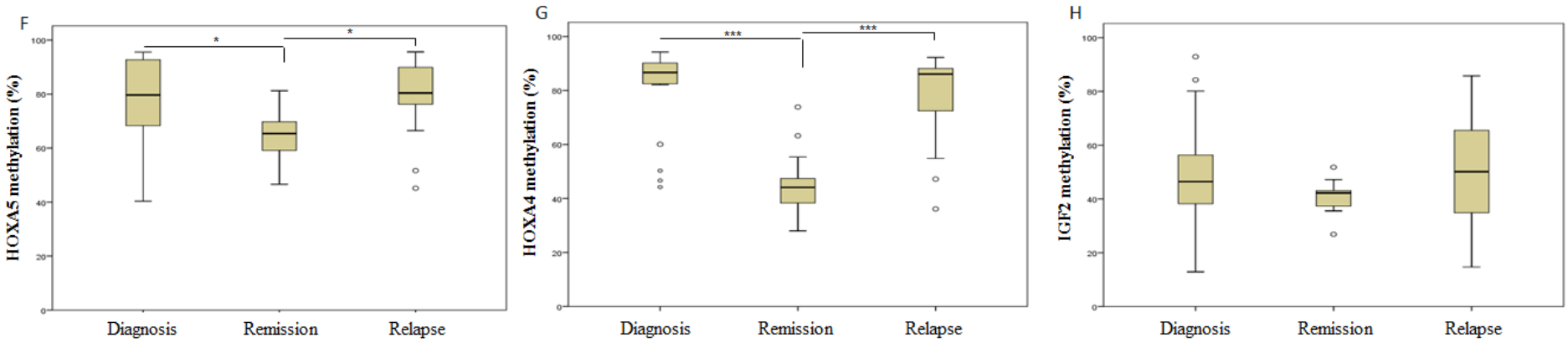


Figure 6.1F-H: Methylation levels during the different phases of adult ALL for *HOXA5* and *IGF2*

F and G) The methylation levels in the paired adult ALL samples were significantly reduced during remission compared to both the diagnostic and relapse phase for *HOXA5* and *HOXA4* ($p=0.013$ and $p<0.001$ respectively). No significant differences were observed in methylation levels during diagnosis and relapse for any of these genes (Wilcoxon signed rank test (SPSS)).

H) No significant differences were observed between the different phases of the disease for the imprinted loci *IGF2* (Wilcoxon signed rank test (SPSS)).

		TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Children	Diagnosis- Remission	R=-0.039 P=0.861	R=0.285 P=0.188	R=0.128 P=0.560	R=0.068 P=0.757	R=0.226 P=0.300	R=0.121 P=0.583	R=0.126 P=0.566	R=0.175 P=0.424
Adults	Diagnosis- Remission	R=-0.185 P=0.421	R=0.285 P=0.211	R=-0.066 P=0.777	R=0.126 P=0.587	**R=0.546** **P=0.010****	R=0.216 P=0.347	R=-0.014 P=0.952	R= 0.348 P=0.122
	Remission-Relapse	R=-0.179 P=0.491	R=-0.151 P=0.563	R=-0.333 P=0.191	R=-0.185 P=0.478	R=0.345 P=0.176	R=0.050 P=0.850	R=0.294 P=0.252	R=-0.007 P=0.979
	Diagnosis-Relapse	**R=0.842** **P<0.001*****	**R=0.813** **P<0.001*****	R=0.224 P=0.294	**R=0.607** **P=0.010****	**R=0.677** **P=0.003****	**R=0.580** **R=0.015***	R=0.454 P=0.067	**R=0.565** **P=0.018***

Table 6.3: Correlations between the ALL diagnostic, remission and relapse samples

No correlations were observed between the diagnostic and remission samples for both the childhood (n=24) and the adult (n=21) patients, except for the *HAND2* gene in the adult population. No correlation was observed between the methylation levels during remission and relapse. Correlations were observed between methylation levels during diagnosis and relapse for 6 of the genes under study; *TWIST2*, *TUSC3*, *EphA10*, *HAND2*, *HOXA5* and *IGF2*.

The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

6.2.1.2.2 Methylation levels during ALL remission are significantly increased compared to healthy controls

Methylation levels of both adult and childhood ALL remission patients were compared to the methylation levels of healthy control patients in a similar age-range. Although methylation levels were reduced in remission samples, compared to the diagnostic and relapse samples, the levels did not fall back to that seen in healthy controls. In childhood ALL a small, but statistically significant increase in methylation was still detectable for six loci at remission (Table 6.4 and Figure 6.2). The only genes not showing a significant increase in methylation in these childhood patients were *EphA10* and *HOXA5*. This increase in methylation was even more pronounced in adult ALL, where remission samples exhibited an increase of about 70% relative to age-matched controls for *TWIST2, EphA10, HAND2, HOXD4* and *TUSC3*. In addition, methylation was also significantly increased in the adult ALL remission patients for both *HOXA4* and *HOXA5* (Table 6.4 and Figure 6.2). These results demonstrate that even during remission significant levels of abnormal methylation are retained.

While this result could potentially be explained by the persistence of a low level of leukaemia cells in the remission samples (minimal residual disease (MRD)) several lines of evidence suggest that this is unlikely. First of all, there was no correlation observed between the methylation levels at diagnosis or relapse with the methylation levels during remission (Table 6.3). Also, for some individuals, methylation levels of the *TUSC3* gene was increased during remission compared to the methylation levels during diagnosis and relapse (and was higher than that seen in any control sample), indicating that the increased methylation cannot be derived from the leukaemic clone (Table 6.5).

MRD data was available for only a subset of patients, however the remission patient with the second highest levels of methylation was demonstrated to be MRD negative (and this sample therefore contained of less than 0.1% leukaemia cells) (Figure 6.3). This demonstrates that increased levels of methylation could be detected even in MRD negative samples, suggesting that, in adult ALL, the high levels of DNA methylation observed in remission cannot be accounted for by residual leukaemic cells.

Another possible explanation for this increase in methylation during remission is that the treatment these patients received altered their methylation levels. Childhood ALL patients receive extensive treatment over a period of 2 years (for girls) or 3 years (for boys). If the increased methylation observed at initial remission was related to

leukaemia treatment, then it is possible that methylation levels may continue to increase while the patients are still receiving treatment. For nine of the childhood ALL samples, two remission samples were available: the first remission sample available was taken when the patients had been under treatment for 3 months, and a second remission sample was available after the patient had been under treatment for 2 years. Assessment of methylation levels in these samples demonstrated that for the five genes associated with ageing related methylation the level of methylation was essentially stable between the initial and long term remission samples, suggesting that additional treatment did not result in any further increase in methylation. In contrast though, the *HOXA4* gene was found to exhibit significant increases in DNA methylation at the second remission time point, suggesting that methylation levels at this gene may be directly influenced by exposure to treatment (Table 6.6, Figure 6.4 and Appendix D, Figure D.).

Another common feature of several cancers, including leukaemia, is CIMP, a phenotype in which multiple genes become co-ordinately methylated in individual patients. Indeed, in our group of ALL patients CIMP was observed for both the childhood and adult ALL patients during diagnosis and relapse. Interestingly, a CIMP like phenotype was also observed in the remission samples in the adult ALL group, where positive correlations were observed between methylation at *TWIST2*, *TUSC3*, *HOXD4*, *EphA10* and *HAND2*, except for *TWIST2* with *HAND2*. In the childhood samples, this CIMP was not observed during remission (Table 6.7). These results showed that, at least in the adult patients, the CIMP is not just a feature of the disease, but it is already present in the non-leukaemic cells of these individuals.

These results show that altered levels of methylation are present in non-leukaemia cells in ALL patients during remission compared to age-matched controls.

A.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels adult ALL remission patients (%)	11.65	12.26	16.22	12.00	7.73	64.44	44.35	41.00
Methylation levels adult controls (%)	4.85	6.38	9.94	7.97	6.35	56.53	38.47	42.40
ALL remission-controls	**P<0.001*****	**P<0.001*****	**P<0.001*****	**P<0.001*****	**P=0.004****	**P=0.001*****	**P=0.029***	P=0.281

B.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels childhood ALL remission patients (%)	5.77	5.80	9.15	6.72	4.53	59.22	39.10	40.33
Methylation levels childhood controls (%)	4.17	3.88	4.98	6.53	3.28	60.52	26.69	46.64
ALL remission-controls	**P<0.001*****	**P<0.001*****	**P<0.001*****	P=0.283	**P<0.001*****	P=0.391	**P<0.001*****	**P<0.001*****

Table 6.4: Methylation levels in the ALL remission patients differ from the methylation levels observed in their age-matched controls

A) Six of the genes under study showed significantly higher levels of methylation in the adult ALL patients during remission (n=21) compared to the healthy controls (n=98) (Wilcoxon rank sum test (SPSS)).

B) In the childhood ALL remission patients (n=24), increased levels of methylation were observed for five of the genes under study compared to the healthy controls (n=50), while one gene, *IGF2*, showed significantly lower levels in these remission patients compared to the controls (Wilcoxon rank sum test (SPSS)).

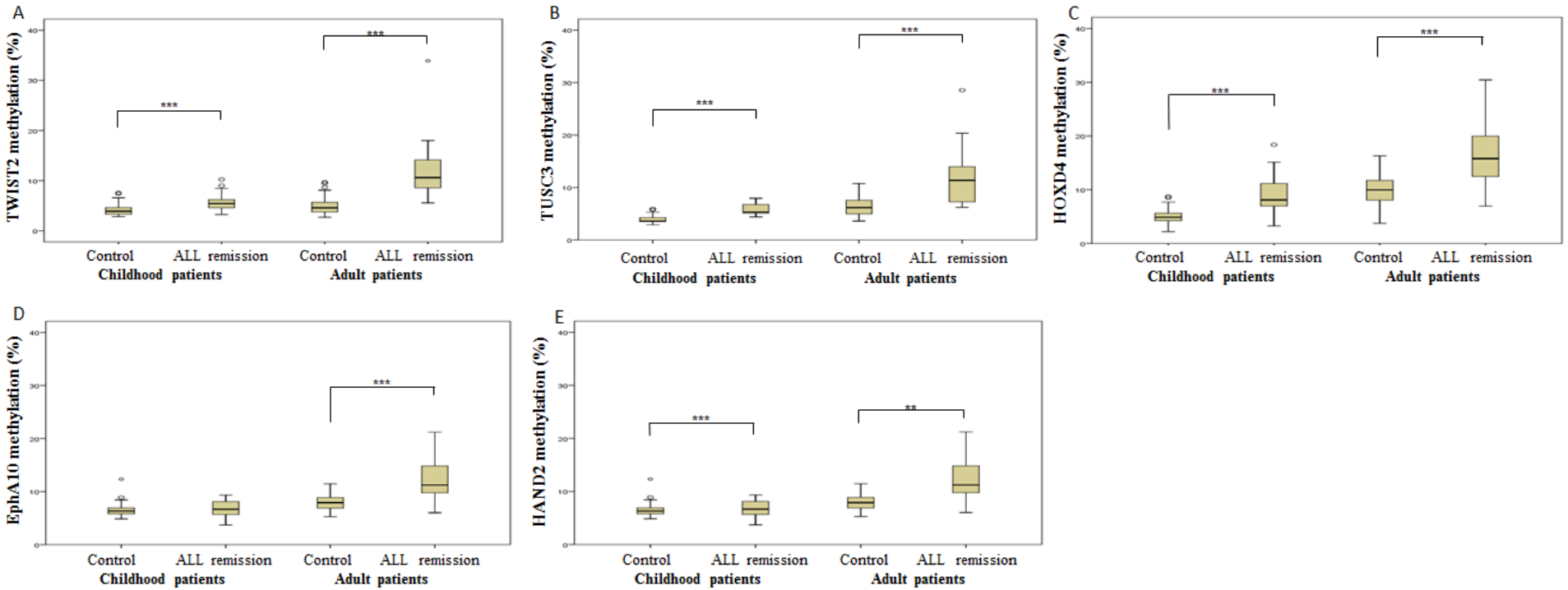


Figure 6.2A-E: Methylation levels in the ALL remission patients differ from the methylation levels observed in their age-matched controls for *TWIST2*, *TUSC3* and *HOXD4*

A-C) A significant increase in methylation levels was observed in both the childhood and adult ALL remission patients compared to the controls for *TWIST2*, *TUSC3* and *HOXD4* (Table 6.4).

D) A significant increase in methylation levels was observed in the adult ALL remission patients compared to the controls for *EphA10* (Table 6.4).

E) A significant increase in methylation levels was observed in both the childhood and adult ALL remission patients compared to the controls for *HAND2* (Table 6.4).

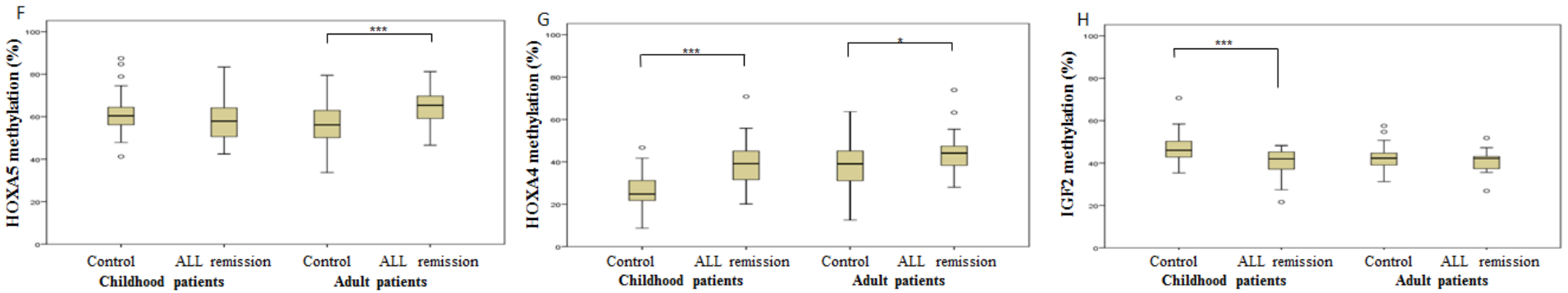


Figure 6.2F-H: Methylation levels in the ALL remission patients differ from the methylation levels observed in their age-matched controls

F) A significant increase in methylation levels was observed in the adult ALL remission patients compared to the controls for *HOXA5* (Table 6.4).

G) A significant decrease in methylation levels was observed in the childhood ALL remission patients compared to the controls for *HOXA4* and a significant increase in methylation levels were observed for this gene in the adult ALL remission patients compared to the controls (Table 6.4).

H) A significant increase in methylation levels was observed in the childhood ALL remission patients compared to the controls for *IGF2* (Table 6.4).

Sample	Methylation at diagnosis (%)	Methylation during remission (%)	Methylation at relapse (%)
Patient 1	**7.61**	**12.70**	**4.91**
Patient 2	**5.21**	**6.90**	**4.99**
Patient 3	**5.31**	**11.21**	**8.49**

Table 6.5: TUSC3 methylation for two patients was increased in the remission sample compared to the diagnostic and relapse samples

For three of the adult ALL patients in this study, methylation levels during diagnosis and relapse were shown to exhibit levels comparable as observed in the healthy controls, while during remission, methylation levels were increased for *TUSC3*.

Patient 1 and 3 showed methylation levels during remission that were higher than seen in any of the control samples.

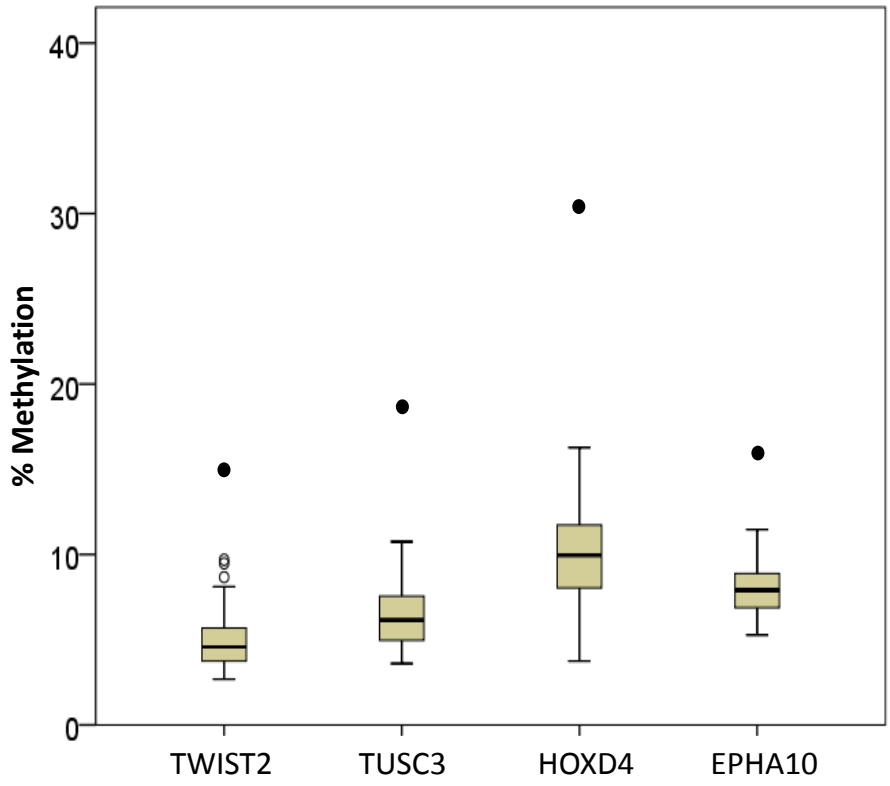


Figure 6.3: Methylation levels in one of the MRD negative ALL patients was higher than that seen in any of the healthy control samples

Methylation levels for four of the genes assessed (as indicated) are represented by box and whisker plots to illustrate the range of methylation levels in healthy controls and a black dot for the remission sample from the MRD negative remission patient. Even though this sample was determined to be MRD negative, it nevertheless still exhibits methylation levels well in excess of that seen in any healthy control sample for multiple genes

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*
Methylation levels at 3 months into remission (%)	5.90	**5.62**	9.95	6.90	4.33	64.46	**37.72**
Methylation levels at 2 years into remission (%)	4.79	**3.55**	8.10	7.28	4.80	55.92	**52.93**
3 Months - 2 Years	P=0.356	**P<0.001*****	P=0.278	P=0.499	P=0.999	P=0.354	**P=0.006****

Table 6.6: Methylation levels are relatively stable during remission

Methylation levels were relatively stable during remission for most of the genes under study for the ten childhood patients for which several remission samples were available. Only the genes *TUSC3* and *HOXA4* showed a significant increase in methylation levels during remission (Paired samples T-test (SPSS)).

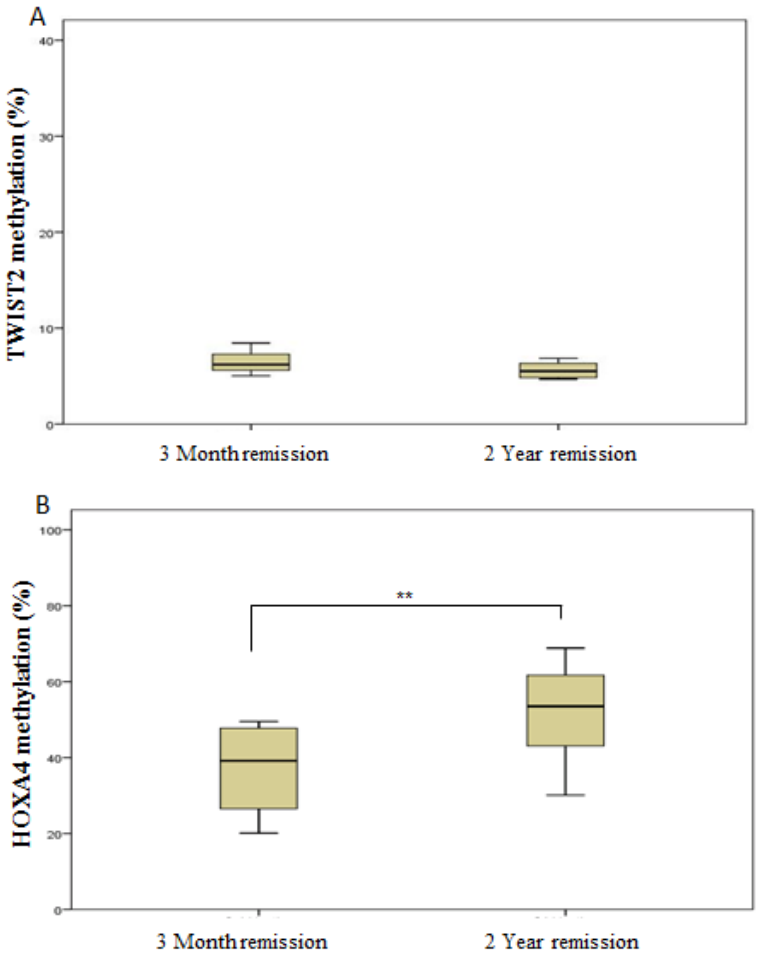


Figure 6.4: Methylation levels are relatively stable during ALL remission

A) Overall methylation levels during remission are relatively stable in the childhood patients. Some patients exhibited a further increase in methylation, while other patients exhibited a reduction later into remission (n=10). The results as shown here for *TWIST2*, was similar for other different genes (Appendix D, Figure D.1), except for *TUSC3* and *HOXA4* (Paired samples T-test, SPSS).

B) *HOXA4* methylation levels are increased after 24 months remission compared to the first remission sample, taken at 3 months into remission (n=10, Paired samples T-test, SPSS).

A.

		TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Adult patients	*TWIST2*		**R= 0.906** **P< 0.001*****	**R= 0.799** **P< 0.001*****	**R= 0.796** **P< 0.001*****	R= 0.287 P= 0.207	R=0.212 P=0.207	**R=0.561** **P=0.008****	R=0.282 P=0.215
	TUSC3			**R= 0.870** **P< 0.001*****	**R= 0.814** **P< 0.001*****	**R= 0.438** **P= 0.047***	R=0.312 P=0.169	**R=0.438** **P=0.047***	R=0.186 P=0.420
	HOXD4				**R= 0.854** **P< 0.001*****	**R= 0.555** **P= 0.009****	R=0.307 P=0.176	R=0.347 P=0.123	R=-0.044 P=0.849
	EphA10					**R= 0.507** **P= 0.019***	R=0.404 P=0.055	**R=0.549** **P=0.010****	R=0.146 P=0.527
	HAND2						R=0.361 P=0.108	R=-0.018 P=0.938	R=-0.115 P=0.619
	HOXA5							R=0.110 P=0.635	R=0.045 P=0.846
	HOXA4								R=0.139 P=0.547
	IGF2								

Table 6.7A: Correlation between genes during remission in the adult ALL patients

The adults ALL patients (n=21) showed a strong correlation between genes during remission, showing a phenotype similar to the CIMP that is observed in cancer patients. The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

B.

		TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Childhood patients	*TWIST2*		R= 0.295 P= 0.162	R= 0.281 P= 0.183	R= 0.324 P= 0.122	R= -0.343 P= 0.101	**R=0.451** **P=0.027***	R=-0.149 P=0.363	R=-0.361 P=0.083
	TUSC3			R= 0.084 P= 0.697	R= 0.396 P= 0.058	R= 0.393 P= 0.058	R=0.108 P=0.615	R=-0.135 P=0.528	R=0.307 P=0.144
	HOXD4				R= 0.272 P= 0.199	R= -0.216 P= 0.310	R=0.382 P=0.066	R=-0.027 P=0.900	R=-0.136 P=0.525
	EphA10					R= 0.302 P= 0.152	**R=0.433** **P=0.034***	R=0.238 P=0.262	R=0.292 P=0.166
	HAND2						R=-0.050 P=0.816	R=0.249 P=0.241	R=0.743 P<0.001
	HOXA5							R=-0.339 P=0.105	R=-0.031 P=0.887
	HOXA4								R=0.270 P=0.202
	IGF2								

Table 6.7B: Correlations between genes during remission in the childhood ALL patients

The correlation between genes as observed in the adult ALL remission patients was not observed for the childhood patients during remission (n=24).

The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

6.2.1.2.3 High levels of CpG island methylation in remission samples is associated with poor outcome in adult ALL

Outcome data was available for the adult ALL samples in this study. Outcome in adult ALL remains poor and indeed, from the 21 adult patients for whom a remission sample was available, six patients exhibited long term survival (>5 years), whereas 15 patients died of the disease. Methylation levels were compared between long term survivors and those that died of their disease. Methylation levels were significantly higher in patients who died of their disease for seven of the genes, while methylation levels in long term survivors was comparable to that see in healthy controls for six of these examined genes. Despite the small sample numbers this difference in methylation levels between the survivors and the non-survivors was statistically significant (Table 6.8 and Figure 6.5) and suggests further study of remission methylation levels as a potential prognostic marker would be merited.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels controls (%)	4.85	6.38	9.94	7.97	6.35	56.53	38.47	42.40
Methylation levels ALL remission-survivors (%)	8.17	8.64	12.95	9.46	6.01	61.18	41.57	39.14
Methylation levels ALL remission- non-survivors (%)	13.04	13.70	17.53	13.02	8.41	65.74	45.46	41.75
Survivor- Non survivor	**P=0.018***	**P=0.016***	**P=0.036***	**P=0.043***	**P=0.020***	P=0.312	P=0.697	P=0.350
Survivor- Control	**P<0.001*****	P=0.157	P=0.759	P=0.503	P=0.387	P=0.354	P=0.329	P=0.568
Non survivor- control	**P<0.001*****	**P<0.001*****	**P<0.001*****	**P<0.001*****	**P<0.001*****	**P<0.001*****	**P=0.039***	P=0.801

Table 6.8: A difference in methylation levels is observed between adult ALL survivors and non-survivors

A significant difference was observed in the methylation levels between the adult ALL survivors (n=6) versus the non-survivors (n=15) for 5 of the genes examined; *TWIST2*, *TUSC3*, *HOXD4*, *Epha10* and *HAND2*. All the genes examined, with the exception of *IGF2*, showed a significant difference in methylation levels between the non-survivors and the controls, while the methylation levels of the survivors versus the controls were only significantly different for *TWIST2* (Wilcoxon rank sum test, SPSS).

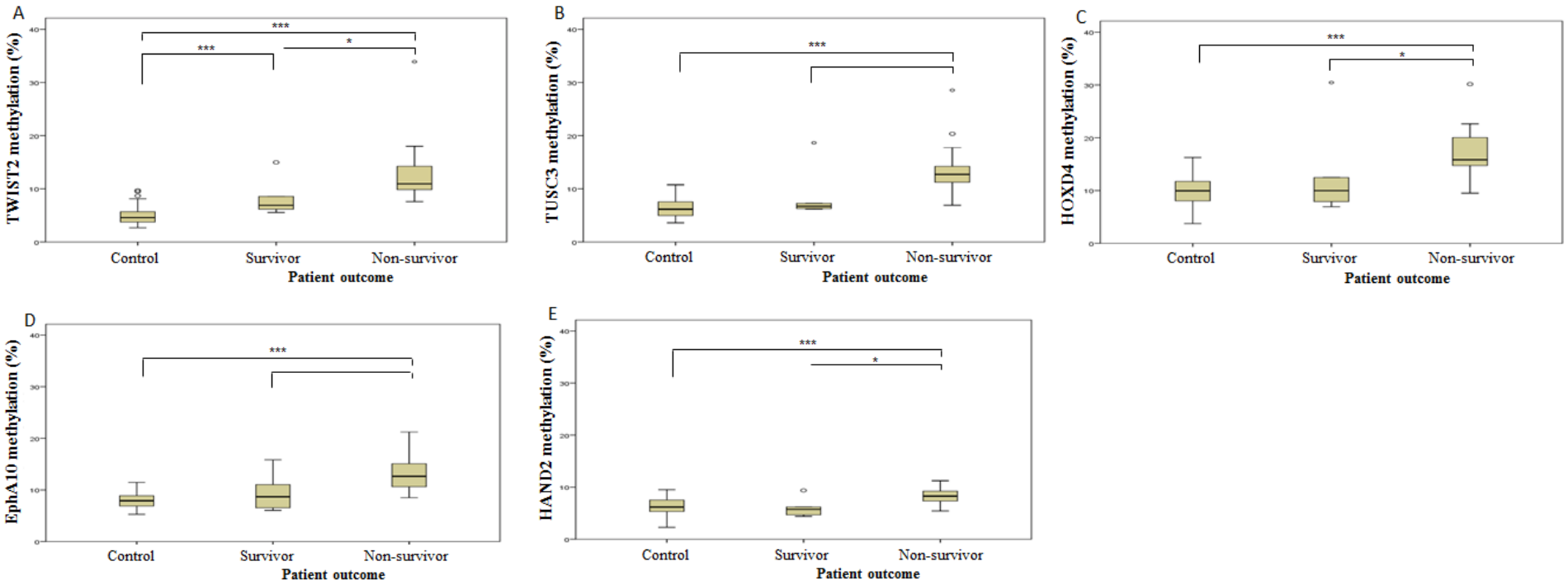


Figure 6.5A-E: Methylation levels during ALL remission in ALL survivors and non-survivors for *TWIST2, TUSC3, HOXD4, EphA10* and *HAND2*

A-E) A significant difference between the methylation levels of the adult ALL survivors (n=6) compared to the non-survivors (n=15) was observed for *TWIST2*, *TUSC3*, *HOXD4*, *EphA10* and *HAND2*. The non-survivors showed methylation levels significantly higher than the controls for these five genes, while only *TWIST2* showed a significant difference in methylation levels between the ALL survivors and the controls (Table 6.8).

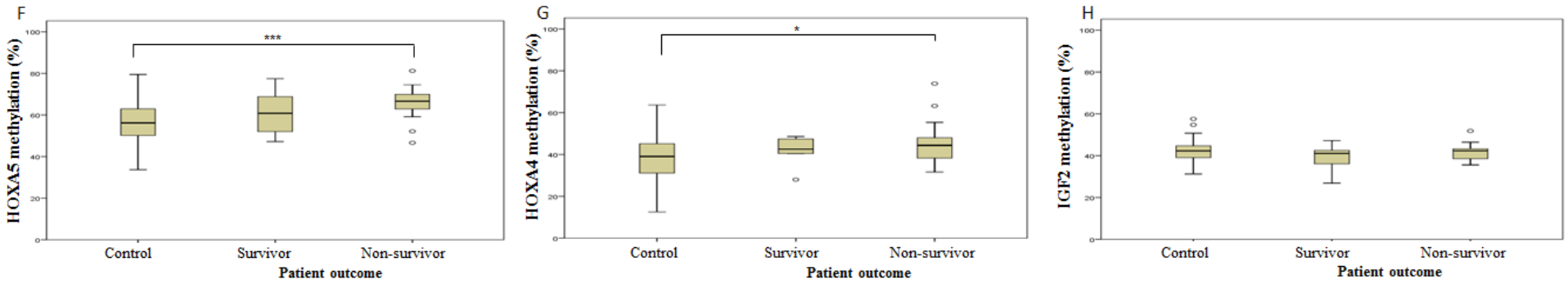


Figure 6.5 F-H: Methylation levels during ALL remission in ALL survivors and non-survivors for *TWIST2, TUSC3, HOXD4, EphA10* and *HAND2*

F and G) No significant differences were observed between the methylation levels during remission of the adult ALL survivors (n=6) and the non-survivors (n=15) for *HOXA5* and *HOXA4*. Significant differences were observed between the methylation levels of the non-survivors and the controls for these genes, whereas no significant differences were observed between the methylation levels of the controls and the ALL survivors (Table 6.8).

H) No differences were observed between the methylation levels during remission of the adult ALL survivors (n=6) and the non-survivors (n=15) for *HOXA5*. A significant difference was observed between the methylation levels of the non-survivors and the controls, whereas no significant difference was observed between the methylation levels of the controls and the ALL survivors (Table 6.8).

6.2.1.3 Discussion

Alterations in DNA methylation, particularly increased methylation of promoter associated CpG islands, is a central feature of cancer development. However, the underlying mechanisms that causes these changes and how this changes across the course of the disease is not well understood. Here we have studied methylation changes during the different phases of ALL. Of particular note, we found that although aberrant methylation was much lower during remission than at either diagnosis or relapse, the level of methylation still remained significantly higher than age matched healthy controls.

The remission methylation levels could not be attributed to a low level of surviving leukaemic blasts and studies of serial remission samples suggests that the increase in methylation is relatively stable. This implies that the increased methylation is either an intrinsic feature of individuals who develop ALL, i.e. the increased methylation existed prior to disease development and may have played an important role in disease development, or alternatively, that the increased methylation is a product of exposure to anti-leukaemic therapy. This could either be due to direct induction of altered methylation by one or more of the specific drugs used or it could be a product of rapid proliferation to re-populate the haematopoietic niche following induction therapy. For at least one locus, *HOXA4*, the increased methylation observed seems likely to be treatment related, as methylation levels were clearly increased when patients had been in remission for 2 years versus the initial remission samples at 2 months, consistent with the idea that more prolonged therapy leads to increased accumulation of altered methylation. This effect of anti-leukaemic treatment would be a particular concern for treatment of childhood ALL patients. With the increasing success of treatment of childhood ALL leading to larger numbers of survivors, the focus for further modifications of treatment will increasingly be on reducing the long term health effects. It is now clear that long term survivors of childhood leukaemia suffer from dramatically increased levels of multiple physical and mental health problems (Kanellopoulos, 2013) and stable alterations in the epigenome could be an important factor contributing to these long term health effects. Furthermore, in adult ALL the patient population is also relatively young, the average age of diagnosis is 55, and so long term health impacts of treatment will also be a concern for this patient population.

This study also suggests that methylation levels measured in remission samples may have potential as a novel prognostic marker, as the levels of methylation at remission significantly correlated with long term survival. Patients who survived long

term generally exhibited methylation levels at or near the level of healthy controls. In contrast, patients that died of their disease exhibited methylation levels significantly above those seen in healthy controls. It has been previously suggested that methylation levels at diagnosis may predict outcome in some ALL patients (Milani, 2010), but this is the first study to find evidence that methylation differences in remission samples may be predictive of subsequent patient outcome, at least in adult ALL. Further studies will be required to confirm this observation and more clearly define the strength of remission methylation levels as a prognostic marker. MRD analysis is already a standard prognostic marker in ALL. The data presented here demonstrate that the increased DNA methylation in adult ALL remission does not appear to be MRD derived, suggesting that remission methylation may provide information that could augment prognosis based on MRD alone, although further studies will be required.

While we did not observe a correlation in these patients between methylation levels during remission and diagnosis or relapse, we did observe a strong correlation between the methylation levels in the diagnostic and relapse samples. This suggests that both the original leukaemia and the relapse are derived from cells with similar methylation patterns. Previous studies have shown that the clonal architecture of relapsed ALL is often very different from that at diagnosis, and that relapses are often derived from clones which were minor at diagnosis (Anderson, 2011; Ding, 2012). The stability of methylation between diagnosis and relapse suggests that most or all clones at diagnosis had very similar patterns of aberrant methylation. This is similar to recent observations in CLL (Cahill, 2012), showing that methylation in this disease is very stable during disease progression. These studies provide further evidence that altered methylation is likely to be one of the earliest events in leukaemia development.

6.2.2 Hereditary nonpolyposis colorectal cancer

HNPCC, or Lynch syndrome, is a genetic disease of autosomal dominant inheritance, caused by a germ-line mutation in one of a set of DNA mismatch repair genes; *MLH1*, *MSH2*, *MSH6*, *MLH3* and/or *PMS2*. This mutation prevents accurate repair of base mismatches and small insertion/deletions produced during DNA replication and this leads to a characteristic mutator phenotype, namely MSI and loss of MMR protein expression (Peltomaki, 1997; Bonnet, 2012). Patients with this germ-line mutation confer an increased risk for various types of cancers, particularly of the colon and the endometrium. It's prevalence in the general population is about 1 in 500, and it causes about 2% to 3% of all colorectal cancers. Lifetime cancer risk for descendant of an affected family member is around 50%, while for the identified gene carriers this risk is much higher (Aaltonen, 1998).

The Amsterdam criteria established diagnostic criteria for HNPCC in 1990. These criteria only described colorectal cancer, but were reviewed in 1999 when endometrial cancer, small intestine cancer, urethral cancer and kidney cancer were also included as HNPCC spectrum cancers. A HNPCC family is clinically diagnosed if; three or more patients with histologically confirmed HNPCC spectrum cancers are present in a family line and one is a first relative to the other two; this cancer develops over two generations; and at least one case is diagnosed at younger than 50 years old (Vasen, 1999).

A new, rarer, type of HNPCC has been found with no pathologic mutation in MMR genes, but epimutation in the promoter region of *hMLH1* or *hMSH2*. These findings suggested that epimutation in germline *hMLH1* or *hMSH2* can be a cause of HNPCC (Chan, 2006; Crepin, 2012). This also suggests that DNA methylation might play a role in HNPCC and that it might be used as a diagnostic marker of identifying HNPCC patients.

Not all HNPCC mutation carriers develop HNPCC related cancers and those that do develop these cancers do that at highly variable ages, with most of these patients being well into adulthood. This suggests that multiple secondary events are likely necessary to develop cancer in these mutation carriers. Several studies reported alterations in methylation levels in the cancers related to HPCC. For example, *MGMT* and *CDKN2* hypermethylation is frequently observed in colorectal cancers (Farzanehfar, 2013; Xing, 2013) as well as hypomethylation of *LINE-1* (Antelo, 2012). Alterations in the

methylation levels of these genes constitute a potentially important feature of colorectal cancer.

A study performed by O'Hagan *et al* showed that cellular reactive oxygen species (ROS), and the accompanied oxidative DNA damage, can potentially lead to aberrant patterns of DNA methylation and transcriptional silencing (O'Hagan). This suggests that deficient DNA repair may be linked to aberrant patterns of DNA methylation. This is of particular interest in these HNPCC patients, since they carry a germ-line mutation affecting key DNA repair pathways. Analogous to what was reported for oxidative damage, this defect in DNA repair may lead to increased levels of DNA methylation in HNPCC patients, which may in turn affect likelihood of developing cancer.

Therefore, in this study we investigated methylation patterns in a set of eight genes in a group of HNPCC patients to investigate whether aberrant patterns of DNA methylation are present in these patients and to study the potential link between DNA mismatch repair defects and promoter methylation. This study was designed to look for an event that would potentially increase the general level of aberrant methylation. As a result, genes that were known to be susceptible to age-related methylation and that showed aberrant pattern of methylation in leukaemia patients were included in this study, as well as the mismatch repair gene *MLH1*.

6.2.2.1 Samples

Peripheral blood DNA samples were obtained from 50 HNPCC patients, between 29 and 75 years of age (with an average age of 53.34), originally collected as part of the CAPP2 study. All patients were proven carriers of a pathologic mismatch-repair mutation or members of a family that met the Amsterdam diagnostic criteria and had a personal history of a cured HNPCC neoplasm. All the samples were taken before the start point of the intervention treatment. Of these 50 patients, 13 carried the *MSH2* mutation and 20 carried the *MLH1* mutation (Table 6.9).

The control samples in this study were the young adults, adults and the older adult participants as described in Chapter 4 (Table 6.9)

	HNPCC patients	Control patients
Age (years)	29-75 (average 53.34)	17-85 (average 54.26)
Number of patients	50	148
	15 male 7 female 28 unknown	41 male 107 female
Mutation	13 MSH2 carrier 20 MLH1 carrier 2 Normal 15 Unknown	

Table 6.9: An overview of the HNPCC study population

6.2.2.2 Results

6.2.2.2.1 Changes in methylation levels in HNPCC patients

Methylation levels, as assessed by pyrosequencing, were examined for differences between the HNPCC patients and the healthy controls to examine if altered methylation levels are a feature of these HNPCC patients. Three of the genes under study, *HAND2*, *HOXA5* and *MLH1*, showed significant differences in methylation patterns between the HNPCC patients and the controls. Methylation levels for *HAND2* and *HOXA5* were significantly increased in the HNPCC patient group. For *HAND2* the methylation levels of the HNPCC patients showed values that were on average 1.46% higher than the controls, namely 8.15% in the HNPCC patient group and 6.69% in the control group (Table 6.10 and Figure 6.6A). The HNPCC patient group showed average levels of 63.46% for the *HOXA5* gene, compared to 58.18% in the control group, resulting in a 5.28% increase in the HNPCC patient group (Table 6.10 and Figure 6.6B). The methylation levels of the *MLH1* gene was significantly reduced in the HNPCC patient group. Overall methylation levels for this gene was low (1.34% in the control group, compared to 1.07% in the HNPCC patient group), resulting in only a decrease of 0.27% in the HNPCC patient group (Table 6.10 and Figure 6.6C). For the other five genes under investigation, *TWIST2*, *TUSC3*, *HOXD4*, *EphA10* and *HOXA4*, no differences in methylation levels were observed between the HNPCC patients and the controls (Table 6.10 and Figure 6.7).

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*MLH1*
Methylation levels HNPCC patients (%)	5.77	7.32	11.21	8.60	8.15	63.46	40.72	1.07
Methylation levels controls (%)	6.05	7.23	11.56	9.32	6.69	58.18	39.50	1.34
HNPCC- Controls	P=0.122	P=0.475	P=0.465	P=0.659	**P<0.001*****	**P=0.002****	P=0.883	**P<0.001*****

Table 6.10: Methylation levels in the HNPCC patients were significantly different compared to the controls for *HAND2, HOXA5* and *MLH1*.

Three of the genes under study showed significantly different levels of methylation in the HNPCC patients (n=50) compared to the controls (n=148) Wilcoxon Signed Ranks test (SPSS)).

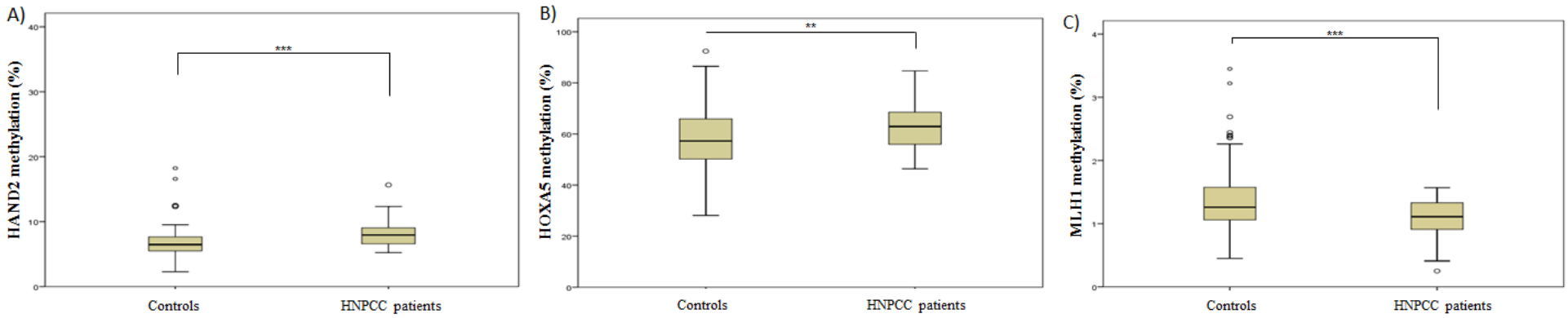


Figure 6.6: Significant differences in methylation patterns between HNPCC patients and controls was observed for *HAND2*, *HOXA5* and *MLH1*

A) Methylation levels for *HAND2* are significantly increased in the HNPCC patient group compared to the controls (Table 6.10).

B) Methylation levels for *HOXA5* are significantly increased in the HNPCC patient group compared to the controls (Table 6.10).

C) Methylation levels are significantly reduced for the *MLH1* gene in the HNPCC patients compared to the controls (Table 6.10).

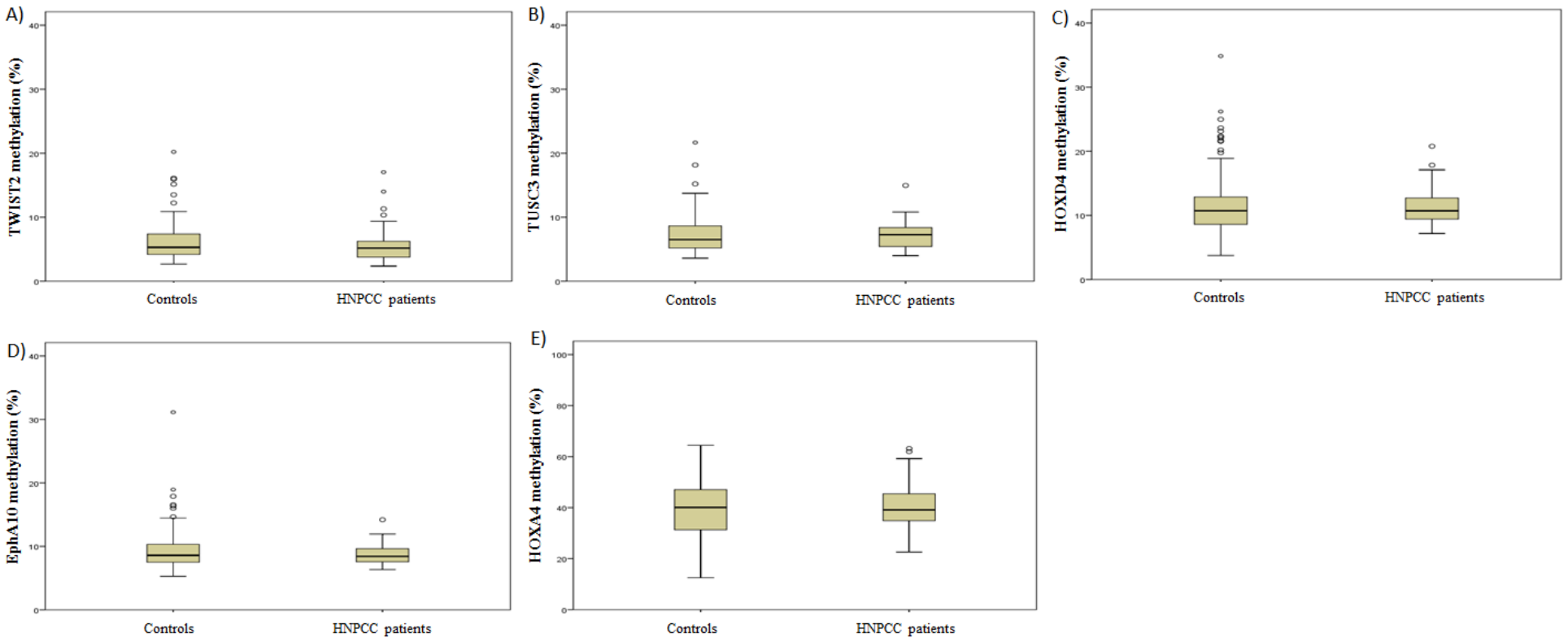


Figure 6.7: No difference in methylation patterns were observed between HNPCC patients and controls for *TWIST2, TUSC3, HOXD4, EphA10* and *HOXA4*

No differences in methylation levels between the HNPCC patients and their age-matched controls were observed for *TWIST2*, *TUSC3*, *HOXD4*, *EphA10* and *HOXA4*.

6.2.2.2.2 Differences in methylation levels between patient mutation groups

Mutations in *MLH1* and *MSH2* are by far the most commonly seen in HNPCC. As they play different roles in mismatch repair function it is possible they may have differential effects on the accumulation of methylation changes. To investigate this, differences in methylation patterns between HNPCC patients from different mutational subgroups (*MLH1* or *MSH2* mutations), were assessed. This analysis found significant differences between these mutational subgroups for both *HOXA5* and *MLH1* (Table 6.11). For *HOXA5*, the patients carrying the *MLH1* mutation showed methylation levels that were on average 6.19% higher than the methylation levels of patients carrying the *MSH2* mutation; the patients carrying the *MSH2* mutation showed average methylation levels of 36.72%, compared to 42.48% for the *MLH1*mutational subgroup. Due to the low levels of methylation observed for the *MLH1* gene, the difference in methylation between the two mutational subgroups for this gene was only 0.40%; the patients carrying the *MSH2* mutation showed average methylation levels of 0.85%, compared to 1.25% for the *MLH1* mutational subgroup (Table 6.11 and Figure 6.8).

These results suggest that the different patterns of methylation observed in the HNPCC patient group can be partly explained by the aberrant patterns of methylation in the patient group carrying the *MSH2* mutation. Indeed, when we examine the differences in methylation of each of the mutational subgroups compared to the controls, a significant difference was observed for *HOXA5* and *MLH1* between the patients carrying the *MSH2* mutation and the controls, but this difference was not observed between the patients carrying the *MLH1* mutation and the controls (Table 6.11 and Figure 6.8). For *HAND2*, an increase in methylation levels was observed in patients carrying both mutations compared to the controls, while this gene did not show a significant difference in methylation levels between the two HNPCC mutational subgroups (Table 6.11 and Figure 6.8). These results suggest that the increase in *HOXA5* methylation and the decrease in *MLH1* methylation are greater in the *MSH2* mutational subgroup compared to the *MLH1* mutational subgroup.

The differences observed between *MSH2* and *MLH1* mutation carriers raises the possibility that while increased methylation at the other 5 genes studies wasn't seen in the whole cohort, that there may be *MSH2* mutation carrier specific effects. However, the other genes under study did not show any significant differences in methylation levels between either the different HNPCC mutation subgroups or between the control patients and patients from either one of the mutational subgroups (Table 6.11 and Figure 6.9).

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*MLH1*
Methylation levels controls (%)	6.05	7.23	11.56	9.32	6.69	58.18	39.50	1.34
Methylation levels *MSH2* mutation carriers (%)	4.66	6.94	11.25	8.60	7.62	67.21	36.72	0.85
Methylation levels *MLH1* mutation carriers (%)	6.30	7.46	10.93	8.40	8.02	61.01	42.48	1.25
MLH1 mutation carriers-*MSH2* mutation carriers	P=0.548	P=0.123	P=0.689	P=0.826	P=0.396	**P=0.014***	**P=0.023***	**P=0.007****
MLH1 mutation carriers-controls	P=0.656	P=0.334	P=0.792	P=0.868	**P=0.029***	P=0.152	p=0.368	P=0.062
MSH2 mutation carriers-controls	**P=0.039***	P=0.539	P=0.856	P=0.682	**P=0.019***	**P=0.001*****	P=0.254	**P<0.001*****

Table 6.11: Methylation levels of HNPCC patients carrying the *MLH1* or the *MSH2* mutation and the controls

Methylation levels differed significantly between the HNPC patients carrying either the *MLH1* (n= 20) or *MSH2* (n=13) mutation for *HOXA5*, *HOXA4* and *MLH1*. The methylation levels of the *MLH1* mutation carriers differed significantly from those of the controls for *HAND2*, while a significant difference between the methylation levels of the *MSH2* mutation carriers and the controls was observed for *TWIST2*, *HAND2*, *HOXA5* and *MLH1* (Independent samples T-test for *HAND2* and *HOXA5*, Wilcoxon Signed Ranks test for the other genes under study (SPSS)).

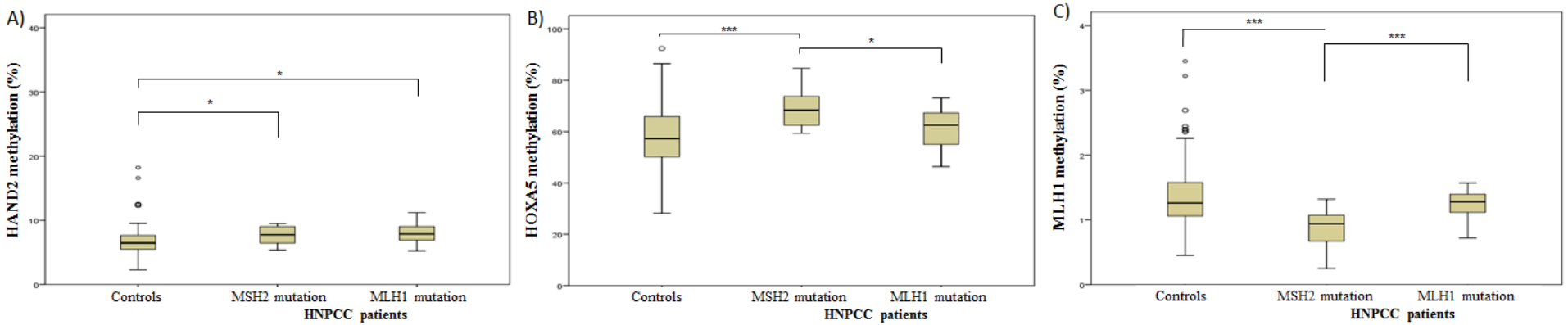


Figure 6.8: Methylation patterns are significantly different between HNPCC patients carrying the *MLH1* and patients carrying the *MSH2* mutation for *HOXA5* and *MLH1*

A) For the *HAND2* gene, methylation levels of both HNPCC mutational subgroups were significantly increased compared to the control patients, whereas there was no significant difference observed in methylation levels between patients carrying either the *MLH1* or *MSH2* mutation (Table 6.11).

B) A significant difference in methylation levels was observed between the HNPCC patients carrying the *MSH2* mutation and the patients carrying the *MLH1* mutation for *HOXA5*. Patients carrying the *MSH2* mutation showed significantly different levels of methylation than the control samples, while this difference in methylation levels was not observed for the patients carrying the *MLH1* mutation compared to the controls (Table 6.11).

C) A significant difference in methylation levels was observed between the HNPCC patients carrying the *MSH2* mutation and the patients carrying the *MLH1* mutation for *MLH1*. Patients carrying the *MSH2* mutation showed significantly different levels of methylation than the control samples, while this difference in methylation levels was not observed for the patients carrying the *MLH1* mutation compared to the controls (Table 6.11).

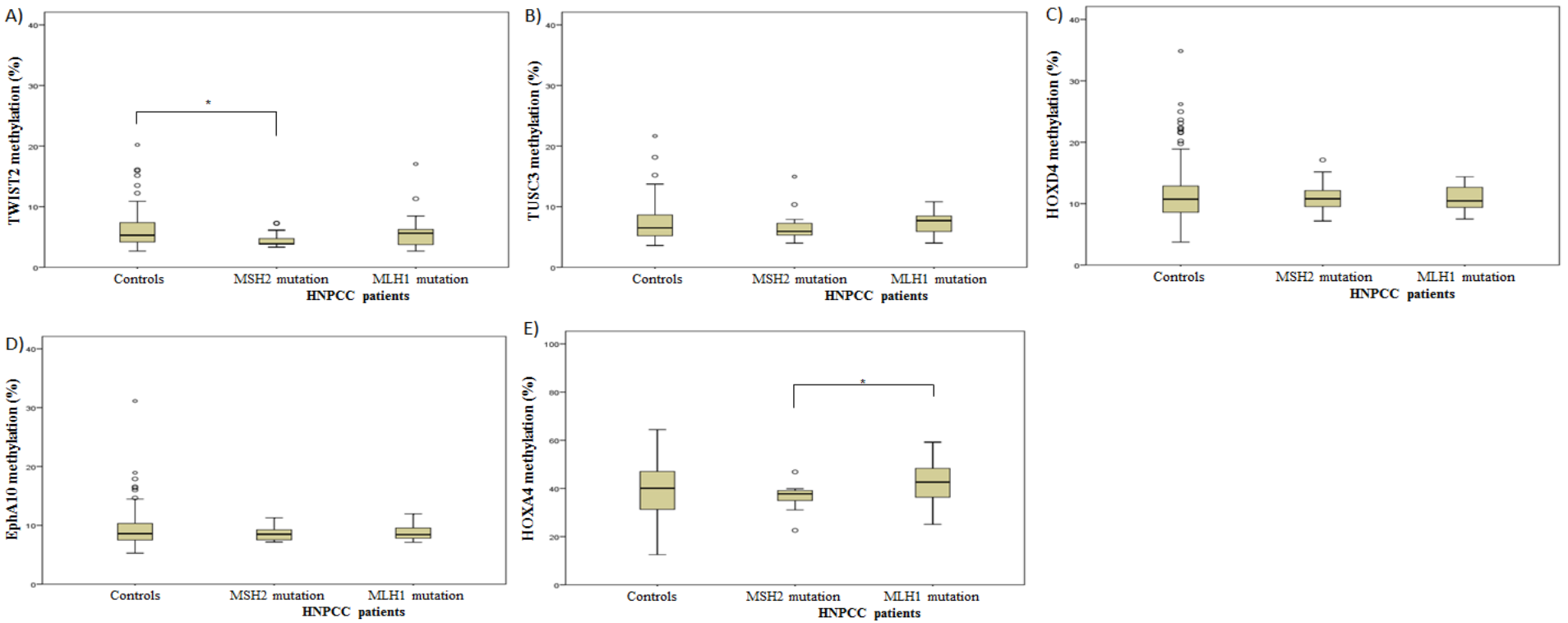


Figure 6.9: Methylation patterns between the HNPCC patients carrying the *MLH1* and patients carrying the *MSH2* mutation for *TWIST2*, *TUSC3*, *HOXD4*, *EphA10* and *HOXA4*

No differences in methylation levels were observed between the HNPCC patients carrying either the *MLH1* or the *MSH2* mutation for *TWIST2*, *TUSC3*, *HOXD4*, *EphA10* and *HAND2*. From these genes, only *TWIST2* showed a significant difference in methylation levels between patients carrying the *MSH2* mutation compared to the controls and only *HOXA4* showed a significant difference in methylation levels between patients carrying the *MLH1* mutation and the controls . These differences were not observed in any of the other five genes (Table 6.11).

6.2.2.3 Discussion

In our study we examined methylation patterns in PBL of HNPCC patients in a set of eight genes. These methylation levels were compared to methylation levels measured in healthy individuals in a similar age-range to examine a potential genome wide effect of methylation changes due to the mutation in the DNA mismatch repair genes these patients are carrying. Our results indicate that such a genome wide effect does not occur in these patients, since aberrant patterns of methylation was only found in three of the eight genes under study; *HAND2, HOXA5* and *MLH1*. However, the presence of altered methylation in this subset of genes does suggest that HNPCC mutation carriers may have frequent methylation changes across the genome and that altered methylation might be a common feature of the condition.

Why it were the genes *HAND2*, *HOXA5* and *MLH1* that showed altered levels of in the HNPCC patients remains currently uncertain. The genes under study were not chosen on the basis of potential functional roles in HNPCC. Also, the tissue analysed in this tissue, PBL, is not the tissue type known to be affected by the condition and it is therefore unlikely that the methylation changes that we detected in these HNPCC patients play a direct role in this condition.

Our results suggest that methylation levels in HNPCC patients are significantly different from patients not carrying this condition, and that substantial methylation defects might be present in these HNCC patients. It is well possible that these methylation defects may be replicated, or are even increased, in tissues and genes that are more directly linked to HNPCC associated conditions. As a result, methylation levels have to be examined in a wider range of tissues and at more loci to get a clearer understanding of the prevalence of methylation differences and their potential impact of the disease.

Interestingly *MLH1*, one of the DNA mismatch repair genes and one of the genes targeted by mutations in a subgroup of HNPCC patients, showed methylation levels that were significantly lower in the patients carrying the *MSH2* compared to the controls, but no apparent decrease in carriers of *MLH1* mutations was observed. While this may at first seem counter-intuitive, one potential explanation could be that mutation of *MSH2* leads to a compensatory up-regulation of transcriptional activity of the MMR genes (or *MLH1* specifically), although a wider study of methylation and expression at all MMR genes would be required to assess this possibility more fully. Also we did not

assess methylation at the *MSH2* promoter and therefore cannot tell if this would have shown a corresponding mutation specific change in methylation status.

Our patient cohort did not allow us to examine whether the observed methylation changes in these HNPCC patients were associated with differential susceptibility to develop one of the HNPCC related diseases, due to the low number of patients and the lack of patients who developed a HNPCC related disease in the follow up period. It could be of interest for further study to examine the potential of DNA methylation in the prediction of disease susceptibility, in both in the affected tissues by HNPCC as well as in PBL as this might mirror the methylation patterns in the tissues.

Overall, the results of this study showed aberrant patterns of DNA methylation in HNPCC patients for three genes under study.

6.3 Heart disease

Heart diseases are currently one of the leading causes of death and disability in the world and the incidence of heart disease increases with age. What, if any, role alterations in DNA methylation have in the development of a heart disease remains currently uncertain, however several studies have identified altered patterns of DNA in a subset of genes which cover functions in lipid oxidation, inhibition of endothelial cell migration and formation, the control of cell proliferation and angiogenesis (Friso, 2012; Hiltunen, 2002; Movassagh; Post, 1999). Our previous study showed that promoter methylation patterns change significantly with an increasing age (as discussed in Chapter 4). This led us to the hypothesis that the rate of accumulation of altered DNA methylation during ageing might have an important impact on the development of heart disease.

Therefore, this study was designed to investigate methylation patterns in patients who had suffered a heart attack, to examine if the heart disease patients exhibited different patterns of methylation as compared to healthy controls. The genes under study included a set of five genes previously identified to show age-related methylation changes; *TWIST2*, *TUSC3*, *HOXD4*, *EphA10* and *HAND2*; and two genes, *HOXA4* and *HOXA5*, that have been implicated to act as positive regulators of the *RGS* genes, especially *RGS2* (unpublished results). The *RGS* genes are involved in the regulation of G-protein coupled signalling and multiple studies have found that reduced expression of *RGS2* is associated with the development of heart disease (Tsang, 2010).

6.3.1 Samples

MNC DNA was extracted from 55 patients, between 42 and 74 years of age (with an average age of 58.3), who had previously suffered a heart attack as part of the 'Role of T-cells in Coronary Artery Disease' study. Clinical characteristics of this cohort are detailed in Table 6.12. The samples from the patients with an Acute Myocardial Infarction (AMI) were taken 24 hours following primary percutaneous coronary intervention (PCI) for ST elevation myocardial infarction (STEMI) and they all had stents 24 hours earlier to unblock the occluded coronary artery. The samples of the chronic patients were all taken at least three months after STEMI. They all had a previous infarct, but were stable at the time of sample collection. The patients with the Congestive Heart Failure (CHF) had larger infarcts than the patients without CHF; these patients had smaller or moderate size infarcts.

The methylation levels from the young adults, adults and older adult participants, as described in Chapter 4, were used as a first control group. The DNA samples from these participants were extracted from PBL. Therefore, as a second control group, the methylation levels from DNA samples extracted from both PBL and MNC were used from 8 adult participants that participated in the cell-type specific methylation study, as described in Chapter 5 (Table 6.12).

	Heart disease patients	Controls 1	Controls 2
Age (years)	42-74 (average 58.3)	17-85 (average 54.26)	43-60 (average 51.6)
Number of patients	55	148	8
Gender	2 female 44 male 9 unknown	106 female 41 male 1 unknown	5 female 3 male
Source of DNA sample	MNCs	PBL	PBL and MNC
Patient group	33 Acute Myocardial Infarction (AMI) 10 Congestive Heart Failure (CHF) 7 Stable post AMI and no CHF 5 unknown		
Amount of vessel affected	23 1 vessel 12 2 vessels 7 3 vessels 13 unknown		
Family history of heart disease	15 yes 13 no 27 unknown		
Previous heart surgery	13 yes 16 no 26 unknown		

Table 6.12: An overview of the cardiovascular disease study population

6.3.2 Results

6.3.2.1 Methylation levels in heart disease patients compared to controls

Variable levels of methylation were observed in the heart disease patients in seven of the genes under study; *TWIST2*, *TUSC3*, *HOXD4*, *EphA10*, *HAND2*, *HOXA5* and *HOXA4*. To examine if these levels of methylation were different than the methylation observed in healthy controls, the methylation levels of the two groups were compared to each other. The methylation levels of these six genes showed increased levels of methylation in these patients compared to the methylation levels observed in their age matched controls (Table 6.13), however the cell composition was different between these two groups (MNC versus PBL). Previous results, as described in Chapter 5, showed significant differences in the methylation levels between the matched MCN and PBL samples in healthy individuals. Overall methylation levels were higher in DNA extracted from MNC compared to the DNA extracted from PBL for all the genes under study, although this difference was relatively small for the genes *HOXA5* and *HAND2* (Table 6.13).

When the increase of methylation levels in the heart disease patients compared to the healthy controls was compared to the increase of methylation levels in the MNC versus PBL, it showed that the increase in methylation levels in the heart disease patients was likely too large to be solely explained by the difference in cell composition for several genes under study, namely *HOXD4*, *HAND2*, *HOXA5* and *HOXA4*. For these genes, an increase in methylation of respectively 4.72%, 2.7%, 7.8% and 7.8% was observed in the heart disease patients compared to the controls, whereas the methylation levels of the MNC samples compared to the paired PBL samples only showed an increase of respectively 2.56%, 0.45%, 0.37% and 2.24%. For the other genes under study, the differences between the heart disease patients versus controls and the MNC versus PBL was less than 2%, suggesting that these differences might be explained by the difference in cell composition (Table 6.13).

To perform some statistical analyses to examine if the methylation levels of these heart disease patients are significantly different from the controls, an Independent samples T-test was performed on the relative differences between the methylation levels of the heart disease patients and the healthy controls 1 and the relative difference between the methylation levels of the paired PBL and MNC controls 2. The relative differences between the patients and controls were calculated by subtracting the average

levels of methylation measured in the control samples from the methylation levels observed in each individual heart disease patient. The relative difference between the paired MNC and PBL were calculated by subtracting the methylation levels observed in the PBL samples from the methylation levels of the paired MNC samples. This analysis revealed statistically significant differences in the methylation of the relative differences of these two groups for seven of the genes under study, namely *TWIST2*, *TUSC3*, *HOXD4*, *EphA10*, *HAND2*, *HOXA4* and *HOXA5* (Table 6.14).

These results suggest that aberrant patterns of methylation are present in heart disease patients, however to confirm these results, methylation analyses will have to be performed on both heart disease patient and control samples with the same sample composition.

	TWIST2	*TUSC3*	***HOXD4***	*EphA10*	***HAND2***	***HOXA5***	***HOXA4***	*IGF2*
Methylation levels heart disease patients (%)	8.74	10.03	**16.28**	12.26	**9.39**	**65.36**	**46.68**	42.66
Methylation levels controls 1 (%)	6.05	7.23	**11.56**	9.32	**6.69**	**58.18**	**39.50**	41.47
Methylation difference heart disease patients- controls 1	2.69	2.8	**4.72**	2.94	**2.7**	**7.18**	**7.18**	1.19
Methylation difference PBL- MNC controls 2	1.61	1.9	**2.56**	1.71	**0.49**	**0.37**	**2.24**	1.99

Table 6.13: Methylation levels of the heart disease patients compared to the controls

For four genes under study; *HOXD4*, *HAND2*, *HOXA5* and *HOXA4*, the differences between the differences in methylation levels between the heart disease patients and the controls 1 and the difference in methylation levels between PBL and MNC was greater than 2% and are therefore unlikely to be solely explained by the differences in cell composition of the DNA.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Relative difference between heart disease patients and controls 1 versus relative difference between PBL and MNC controls 2	**P<0.001***	**P<0.001***	**P<0.001***	**P<0.001***	**P<0.001***	**P=0.029***	**P=0.005***	0.187

Table 6.14: Significant differences were observed between the relative difference in methylation levels between the heart disease patients and the controls

Statistical analysis of the relative differences between the methylation levels of the heart disease patients compared to their age-matched controls 1 versus the relative difference between methylation levels between the paired PBL and MNC from the second control group, showed significant differences in methylation levels for 7 genes under study: namely *TWIST2*, *TUSC3*, *HOXD4*, *EphA10*, *HAND2*, *HOXA5* and *HOXA4* (Independent samples T-test (SPSS)).

6.3.2.2 Differences in methylation levels between patient groups

The heart disease patients under investigation could be divided into different groups of disease characteristics. To investigate for differences in methylation patterns between heart disease patients with a subset of the different disease characteristics, statistical analyses were performed.

No consistent significant differences were seen between the methylation levels of any of the available clinical characteristics of the patients. The only observed differences were between the patients with a family history of heart disease and the patients without such a family history for the *HOXA4* gene, and between the patients with one or three vessels affected by the disease for *HOXA5* (Table 6.15-Table 6.18 and Appendix D, Figure D.2-Figure D.5)

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels AMI group (%)	8.80	9.62	15.85	12.44	9.07	66.91	46.15	43.00
Methylation levels CHF group (%)	8.96	11.23	17.13	11.80	9.39	46.39	47.98	42.81
Methylation levels stable post MI, no CHF group (%)	8.71	9.68	17.84	12.72	10.16	63.20	46.50	42.35
AMI-CHF	P=0.701	P=0.535	P=0.301	P=0.443	P=0.169	P=0.785	P=0.626	P=0.768
AMI-stable post MI, no CHF	P=0.761	P=0.987	P=0.211	P=0.612	P=0.607	P=0.492	P=0.839	P=0.361
CHF-stable post MI, no CHF	P=0.722	P=0.591	P=0.594	P=0.657	P=0.195	P=0.717	P=0.594	P=0.657

Table 6.15: No differences in methylation levels in heart disease patients within different patient groups

No differences were observed between the AMI, CHF or stable post MI, no CHF patient groups for any of the genes under study (Independent samples T-test for *HAND2* and *HOXA5*, Wilcoxon Signed Ranks test for the other genes under study (SPSS))

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels previous surgery (%)	9.31	11.51	18.82	12.57	9.26	65.64	46.75	42.28
Methylation levels no previous surgery (%)	8.80	9.29	15.93	12.15	9.91	65.73	45.05	43.84
Previous surgery-no previous surgery	P=0.539	P=0.105	P=0.087	P=0.614	P=0.957	P=0.672	P=0.219	P=0.511

Table 6.16: Methylation levels did not differ between heart disease patients who underwent a previous surgery and the patients who did not

No differences were observed between the heart disease patients who underwent a previous surgery and the patients who did not for any of the genes under study (Independent samples T-test for *HAND2* and *HOXA5*, Wilcoxon Signed Ranks test for the other genes under study (SPSS)).

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
Methylation levels 1 affected vessel (%)	8.50	9.39	16.15	12.69	8.97	**66.11**	46.22	42.93
Methylation levels 2 affected vessels (%)	9.79	11.39	17.43	12.32	9.98	64.58	48.14	41.41
Methylation levels 3 affected vessels (%)	8.80	9.17	14.88	11.93	9.39	**68.16**	45.71	46.09
1 vessel- 2 vessel	0.065	0.076	0.118	0.972	0.264	0.123	0.889	0.862
1 vessel- 3 vessels	0.418	0.914	0.787	0.391	0.522	**0.034***	0.573	0.315
2 vessels- 3 vessels	0.554	0.375	0.151	0.499	0.086	0.083	0.299	0.482

Table 6.17: Methylation levels between patients with 1, 2 or 3 affected vessels by their heart disease

No differences were observed between the patients with 1, 2 or 3 vessels affected for any of the genes under study, with an exception for *HOXA5* (Independent samples T-test for *HAND2* and *HOXA5*, Wilcoxon Signed Ranks test for the other genes under study (SPSS)).

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	***HOXA4***	*IGF2*
Methylation levels family history (%)	9.06	9.82	15.83	11.82	9.45	66.43	**42.86**	42.51
Methylation levels no family history (%)	8.67	9.88	17.89	12.89	9.85	66.86	**50.52**	43.94
Family history-no family history	0.807	0.558	0.118	0.143	0.270	0.717	**0.036***	0.770

Table 6.18: Methylation levels in heart disease patients with a family history of heart disease and in the patients without such a family history

No differences were observed between the heart disease patients with and without a family history of heart disease for any of the genes under study, with the exception of *HOXA4* (Independent samples T-test for *HAND2* and *HOXA5*, Wilcoxon Signed Ranks test for the other genes under study (SPSS)).

6.3.3 Discussion

An increase in methylation levels was observed in the heart disease patients compared to their healthy age-matched controls for seven of the genes under study. Statistical analyses revealed that these differences in methylation levels are likely not caused by the difference in cell composition between these groups. These analyses are not perfectly validated: the difference in cell composition between these two groups made it not possible to perform direct statistical analyses on the observed methylation levels. Unfortunately, there were no PBL DNA samples available for the heart disease patients, or MNC DNA from a larger group of age-matched controls. As a result, the statistical analyses were performed on the relative differences between the patients and the controls and the previously measured differences in MNC and PBL. These analyses revealed altered levels of methylation in the heart disease patients compared to healthy controls, although the transformation of the data into relative differences might have influenced the statistical analyses. However, for some of the genes under study, for example *HAND2* and *HOXA5*, the differences between the methylation levels of the paired MNC and PBL samples was small. Also, for some of the other genes under study, the methylation differences between the heart disease patients and the controls were over 2% higher than the methylation differences between the paired MNC and PBL samples. These observations together suggest that the increase observed in the heart disease patients in these genes cannot be solely explained by the difference in cell composition and that these patients carry aberrant patterns of DNA methylation. Nevertheless, to confirm these results, methylation levels should be analysed in a greater number of DNA samples with the same cell composition in both the patient and the control group.

The exact role of the methylation changes observed in the heart disease patients in this study remains largely unknown. Interestingly, an increase of both *HOXA4* and *HOXA5* methylation was observed in these heart disease patients. Previous expression profiling results performed in our group showed that multiple *RGS* genes were positively regulated by these genes, with the most extensive upregulation seen for the RGS2 gene (unpublished results). Deregulation of RGS2 is thought to play a crucial role in the pathogenesis of cardiovascular diseases. It does that by deregulation of the G protein-coupled receptor (GPCR) signalling pathways, involved in blood pressure homeostasis, by promoting chronic constriction of the peripheral vasculature, leading to

hypertension and downregulation of *RGS2* is observed in cardiac hypertrophy (Tsang, 2010).

None of the other genes under investigation have been previously described to play a role in the development of heart diseases. However, these genes under study did show clear age-related methylation changes (as described in Chapter 4). Thus one hypothesis that could explain the increased methylation observed would be that heart disease patients have higher levels of age-related DNA methylation in MNCs than healthy individuals. In this study we cannot examine what the role of these methylation changes are in the development of a heart disease, due to a number of limitations in our study set up. First of all, the methylation levels were measured in patients that had already developed a heart disease, making it unknown whether these methylation differences were already present before the development of the disease or if they occurred afterwards. Secondly, we have only examined methylation levels in MNCs in these heart disease patients. Even though these cells are important in the development atherosclerosis, it would be interesting to test whether these methylation changes could be confirmed in different cell types. In our study, we did not examine expression levels, so it is unknown what the effect of the observed increase in methylation levels does to the expression of the genes and if these aberrant patterns of methylation observed in the heart disease patients are having an important biological effect.

Overall, our results show that altered patterns of methylation are present in blood samples of heart disease patients. What the role is of these altered methylation patterns remains currently unknown and further research is necessary to examine this role in greater detail.

6.4 Chapter discussion

The results discussed in this chapter show that the CpG island methylation levels of the genes that were previously identified as being age-related are overlapping with the genes showing aberrant patterns of DNA methylation in the age-related diseases examined; cancer, more specifically leukaemia and HNPCC, and heart disease. Both the leukaemia patients and the heart disease patients showed aberrant patterns of methylation in most of the genes under study, whereas for the HNPCC patients, only a subset of the genes under study showed aberrant patters of methylation. This suggests that, at least in case of patients with a disease related to age, leukaemia and heart

disease, higher levels of age-related DNA methylation was observed in these patients compared to healthy controls.

Even though the methylation levels in all of these patients were significantly different than the methylation observed in the age-matched controls, the extent of methylation measured in these patients varied significantly; while hypermethylation was observed during ALL diagnosis and relapse, which is almost certainly leading to transcriptional silencing, this increase in methylation was not as extreme in the non-diseased samples of these ALL patients, nor in the heart disease and HNPCC patients. However, even though the extent of methylation changes observed in these patients was generally not very large, these methylation changes might be significant at the cellular level and thus biological and expression changes may be present in a subset of cells. As is suggested by the cancer clonal expansion theory, cancer likely develops from clonal expansion of a single mutant cell. This suggests that altered levels of methylation which leads to inactivation of a gene in even a small fraction of cells could have important roles in cancer development. How this relates to the development of other age-related diseases is less clear. It is possible that the inability to maintain DNA methylation during ageing is a result of an overall decline of the body and the inability to maintain important biological processes with an increasing age. This suggests that patients who are biologically older, contain higher levels of DNA methylation and are more likely to develop a disease related to age. In this case changes in DNA methylation could function as a marker of increased risk of developing disease, but perhaps have no direct role in the development of the disease. Alternatively accumulation of altered methylation patterns in a subset of cells may contribute to loss of tissue homeostasis and thus potentially have a more direct impact on disease development.

At the moment, it is unknown whether these increases in methylation levels observed in these patients were already present in these patients before the development of the disease, or if these methylation patterns occurred after disease progression. Therefore, more research is necessary to explore the exact role of these methylation patterns in the development of these diseases. However, increased levels of methylation was also observed for a subset of the genes analysed in the HNPCC patients, patients with a condition leading to an increased risk of developing a disease, although all these patients analysed were disease free at the time of sample collection. This suggests that, at least in the HNPCC patients, aberrant patterns of methylation are already presents in these patients before they develop the disease. It is therefore likely that similar

mechanisms occur in other diseases as well, suggesting that the aberrant patterns of methylation in both the leukaemia remission patients and the heart disease patients might have been pre-existing before the development of the disease. This hypothesis suggests that aberrant patterns of DNA methylation in apparently healthy individuals might be used as a predictive marker of patients that are more likely to develop one of these age-related diseases. This hypothesis is strengthened by the observation that aberrant patterns of DNA methylation during ALL remission might predict future relapse, suggesting a link between underlying methylation patterns and patient outcome.

Overall, we showed that aberrant patterns of DNA methylation is a feature of several diseases related to age. This suggests that patients with these age-related diseases have higher levels of age-related DNA methylation in blood compared to healthy individuals.

Chapter 7. Discussion

7.1 Aberrant patterns of methylation during ageing

The results from our study demonstrated that age-related methylation changes occur in a subset of genes in PBL samples. For at least five of the genes, *TWIST2, TUSC3, HOXD4, EphA10* and *HAND2*, it was shown that with an increasing age, there is also an increase in the extent of methylation as well as an increase in the variance of methylation, which is steadily acquired during the ageing process. Three of the genes under study *HOXA4, HOXA5* and *IGF2* showed different patterns of methylation during aging compared to the other five genes. For these genes, the patterns of methylation were relatively stable during aging, although a difference in variance was observed within the different age groups. For one of the genes under study, *MLH1*, no sign of changes in methylation patterns with an increasing age was observed. In this study, only nine genes were investigated.

The consistency with which multiple genes in our study showed strong similarities in the patterns of age-related methylation, which is steadily acquired over the life span, makes it highly likely that there are other genes following similar patterns of age-related methylation. Further studies are necessary to identify the amount of genes following similar patterns as observed in our study. What is causing these methylation changes during aging remains currently unclear, but it might be caused by an unstable maintenance of the replication of DNA methylation patterns during cell divisions during aging.

Interestingly, the genes that show variable methylation during ageing in PBL are closely related to the genes that become methylated during cancer development from the same tissue, whereas the *MLH1* gene, which is frequently methylated in solid tumours but not in haematological malignancies, did not exhibit increasing methylation with age. This is supported by the analyses of the methylation levels in our leukaemia patients, were a great overlap was observed between the genes showing variable methylation during ageing and those that became hypermethylated during leukaemia development. The methylation patterns of these genes were found to be stable during disease progression; strong correlations were found in our paired ALL diagnostic and relapse samples. This suggests that altered methylation levels are consistent during disease progression, which is supported by the observation that during CLL development, global methylation patterns were shown to be relatively stable with time and within different CLL compartments (Cahill, 2012). These results imply that although methylation is highly abnormal in these patients, it is not necessarily unstable in cancer cells. Other studies have shown that altered levels of DNA methylation were observed

in early stages of the disease (van Hoesel, 2013; Heuck, 2013), suggesting that altered levels of methylation might be an early event in tumour development.

All these observations together, leads to the hypothesis that the susceptibility of developing leukaemia, and potentially other diseases which increases with age, may be caused by the inability of cells to stably maintain the DNA methylation levels during aging and/or cell division. This hypothesis is strengthened by the observation that the observed correlation of the methylation levels of these genes in the older adult population, and in a lower extent also in the younger age groups, is following a similar pattern as the CIMP described in cancer, which is a phenotype that describes the presence of widespread CpG island methylation in tumours. In the older adult population the presence of such a widespread CpG island methylation is also observed in the five genes we tested and *TWIST2, TUSC3, HOXD4, EphA10* and *HAND2* exhibit similar patterns of age related changes in DNA methylation. This again suggests that DNA methylation levels at these genes are closely related to each other in healthy individuals. This CIMP was not driven by changes in the cellular composition of PBL samples during ageing, as it was also seen in specific cell types after cell sorting.

7.2 Methylation patterns during cell differentiation

The results from our study into methylation patterns in specific cell types, showed that, for the genes under investigation in this study, the specific isolated cell types were mirroring patterns as observed during leukaemia development, where the lymphocytes, and most specifically the B-lymphocytes, showed higher levels of methylation than the myeloid cell types. Further investigation into early B-cell progenitors showed that these B-cell lineage specific progenitor cells did not mirror the increase in methylation levels as observed in the isolated differentiated B-lymphocytes, suggesting that cell-type specific methylation differences in CpG island methylation examined in this study must occur during the process of differentiation. Even though several other studies have shown that methylation patterns can change during cell differentiation (Bocker; Calvanese), their results suggested that these methylation changes were related to the differentiation program of that specific cell type. In contrast, our results imply a more general phenomenon, in which it are the B-lymphocytes that acquire increased CpG island methylation of a large group of genes that are also under influence of age-related methylation changes. This acquisition of promoter methylation does not appear to be involved in gene silencing. For example; *TWIST2* accumulated high levels of promoter methylation in B-lymphocytes, even though it was shown that

this gene is also expressed in differentiated B-lymphocytes (Thathia, 2012). At the same time, the methylation levels that were observed in the B-lymphocytes our study were not high enough to suggest a controlled mechanism for promoter methylation in gene activation in all cells.

We also investigated methylation changes in specific cell types related to age. These results showed that age-related methylation changes occur in several cell types and that they are not just occurring in one cell type specifically. Most of the genes under study showed a slight, but not always significant, difference in the methylation levels of the isolated cells between the two age groups. This suggests that age-related differences in methylation levels observed for these genes do not solely occur during the process of differentiation but that they must occur in a pool of stem cells. This inability to stably maintain methylation levels in stem cells during ageing are of interest in understanding methylation changes during disease development. Age-related DNA methylation may be a key factor in underlying disease susceptibility by this inability of stem cells to stably maintain DNA methylation.

The genes included in this study were all lymphoid leukaemia specific, with the exception of *HOXA4* and *HOXA5*, which are both methylated in all types of leukaemia, and we did not investigate methylation patterns during ageing and in specific cell types in blood in myeloid leukaemia related genes. It would be interesting to examine the patterns in this set of genes to investigate if the methylation patterns are different between the different cell types for those genes. In order to identify other genes of interest, a 450K-methylation array was performed on both isolated cell types from the myeloid and lymphoid lineage, as well as for AML and ALL patients. This array was performed to identify more myeloid leukaemia related genes, as well as to identify similarities and differences in the patterns of the lineage-specific cell types from healthy individuals and the leukaemia derived from these cell types.
Unfortunately, the analysis of this array was not completed yet before thesis writing.

7.3 Allelic distribution

To examine the allelic distribution of the methylation levels, as measured by pyrosequencing, bisulphite sequencing was performed on two ALL remission samples. These results showed that the methylation patterns might not be randomly distributed among the different alleles, but that they are more restricted to a subset of alleles. As shown in two ALL remission samples for the *TUSC3* gene (with respectively 28.54% and 20.34% of methylation as assessed by pyrosequencing), most of the alleles remain

methylation free, where only a subset of alleles show methylated CpG sites. For one of the samples there were even no methylated CpG sites detected at all during bisulphite sequencing (Figure 7.1). This suggests that most methylation in this sample is likely to be found in a small percentage of alleles as opposed to being evenly spread across all alleles (such that all alleles had 1 or 2 CpG sites methylated). It is well possible that this pattern as observed in the non-leukaemic cells from two leukaemia patients, also occurs in healthy individuals. Interestingly, the presence of highly methylated alleles is one of the known features of leukaemia, suggesting that even in the non-leukaemic samples, similar patterns of allelic distribution is observed compared to the leukaemia phase. However, the number of alleles analysed so far is small and the lack of methylation identified by this technique suggests some bias towards unmethylated alleles.

Further investigations by HTS will have to give us a more accurate and in-dept analysis of the epi-alleles and will give us more insight on what the methylation patterns, as measured by pyrosequencing, mean in the context of individual alleles or cells. This HTS technique allows us to examine around 100,000 of alleles per individual per gene instead of the small numbers as analysed by bisulphite sequencing. Healthy individuals in different age groups, as well as paired diagnostic and remission samples from six ALL patients, were included in this analysis.

Unfortunately, these analyses were not completed before thesis writing.

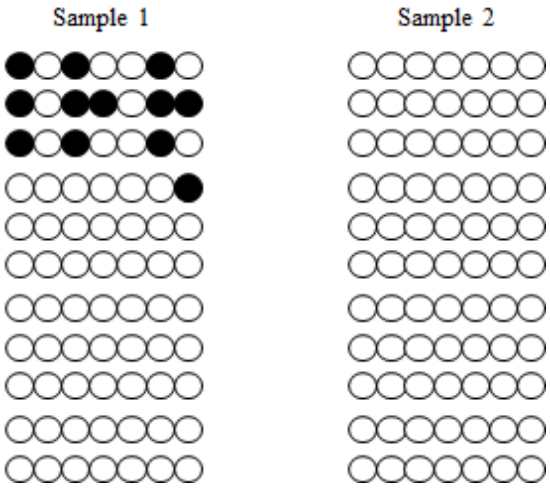


Figure 7.1: Bisulphite sequencing results of two ALL remission patients

Bisulphite sequencing results on two ALL patients that were into remission showed, with methylation levels of respectively 28.54% and 20.34% as assessed by pyrosequencing, showed that the CpG methylation is not randomly distributed among the different alleles. A high number of alleles remain completely methylation free, while other show methylation patterns. No methylated CpGs were identified in sample 2, suggesting that the methylation patterns of the other, not identified alleles, must be hypermethylated to reach a total amount of methylation of 20.34%

7.4 Aberrant patterns of methylation in age-related diseases

Our investigation into aberrant patterns of methylation in diseases related to an increasing age, such as cancer, more specifically leukaemia, and heart disease, revealed that aberrant patterns of methylation was observed both during the development of leukaemia and in patients who had suffered a heart attack. These genes were identified due to their role in leukaemia development, and indeed, seven of the genes under study showed levels of hypermethylation during both the diagnostic and relapsed phase of the disease. Aberrant patterns of methylation were even observed in patients who had suffered a heart attack, suggesting that genes that are under influence of age-related methylation changes also show aberrant patterns of methylation in at least two diseases related to age.

DNA methylation is most extensively studied in cancer and increasing numbers of evidence suggest that DNA methylation plays an important role in this development. Earlier studies already showed that aberrant patterns of DNA methylation do occur in other pathologies in addition to cancer, however, whereas for cancers the causal role of DNA methylation is well established, for other diseases, the direct functional role of these altered methylation patterns is not as clear. During cancer development clonal selection occurs, and thus an epigenetic or genetic change which occurs even in a small subset of cells can have dramatic consequences if those cells clonally expand. Most other diseases do not develop from clonal expansion, and therefore, the effect of changes even in a small number of cells remain less clear. Our results however, showed that aberrant patterns of methylation increases over the life course and is likely to affect a reasonable high number of alleles, which might be sufficient to affect tissue homeostasis. In several diseases, such as atherosclerosis, only local areas of altered cell behaviour are required to develop the disease. The levels of methylation changes observed during ageing and in the patients who suffered from a heart disease might underlie such changes in a fraction of cells. However, this is currently speculation and more research is necessary to investigate this mechanism.

This study was an observational study and therefore cannot tell if altered methylation could be a cause or a consequence of the disease. However, while studying a relatively small number of genes, this study was able to identify methylation differences in multiple diseases, which does lend further credence to the idea that altered methylation may have a more widespread clinical consequences outside of the oncology field.

Several observations together lead to the hypothesis that the observed alterations in the methylation levels of our leukaemia patients might have been pre-existing before the development of the disease, and as a result, might underlie disease susceptibility. First of all, aberrant patterns of methylation were observed even in the non-diseased cells from our set of leukaemia patients. These analyses were performed on the ALL remission samples, and as a result, we can't rule out the possibility that the treatment these patients received altered their methylation levels, such as the chemotherapeutic drugs and/or the rapid proliferation that was required to re-populate the haematopoietic niche. However, it is possible that these methylation patterns were pre-existing before the patients developed the disease. This suggests that patients with higher levels of methylation are more likely to develop leukaemia. This is supported by the observation that the ALL patients that have poor survival showed higher levels of methylation in these non-diseased cells compared to the patients who survived the disease.

Secondly, when we examined methylation patterns in a set of HPCC patients, patients with a condition leading to an increased risk of the disease, although all the patients were disease free at the time of sample collection, three of the genes under study, previously identified of showing both age-related methylation changes and altered methylation levels in our leukaemia and heart disease patients, also showed altered levels of methylation in these HPCC patients. This means that in these patients, altered levels of methylation were present before the development of the cancers related to HNPCC. However, these patients do have an inherited mutation that might cause alterations in their methylation levels, which are not present in the other patients groups.

To further examine the hypothesis of pre-existing alterations in the DNA methylation patterns of gene promoters prior to disease development, DNA samples were collected from the EPIC-Norfolk study (Day, 1999). These samples were collected from patients two to five years before they were diagnosed with either leukaemia or lymphoma and, as a result, all these participants were disease-free at time of sample collection. DNA samples were also collected from age-, gender- and SES-matched controls that stayed disease free in the follow-up years. Analyses of the methylation levels in these individuals will allow us to test whether altered levels of methylation, in patients diagnosed with haematological cancers, are already present before disease development and to test the hypothesis that apparently healthy individuals with higher levels of methylation are more likely to develop the disease.

Unfortunately, these analyses were not completed yet at the moment of thesis writing.

7.5 Cross-talk between methylation changes during ageing and disease susceptibility: the pre-existing methylation hypothesis

All of our results combined led us to the hypothesis that pre-existing methylation may underlie the susceptibility of haematological malignancies, and potentially other cancers. When we go 10-20 years back in time, the main focus was on promoter methylation levels during cancer development. However, what we know now, mainly due to more sensitive techniques, such as pyrosequencing, is that methylation levels are already present in healthy individuals, and in fact, are already present at birth, although at much lower levels than during cancer development. These methylation patterns might not be randomly distributed (as indicated in Figure 7.2A), since our preliminary bisulphite sequencing results showed that methylation patterns are likely more restricted to specific alleles, such that an allele is either densely methylated or remains mostly methylation free. This will result in a pattern in which each cell has one or two densely methylated genes, whereas the other genes in this cell remain mostly methylation free (Figure 7.2 B). This hypothesis might even be expanded, since analyses of our healthy individuals showed strong correlations between genes within individuals, and even within specific cell types. This suggest a co-methylation, where several genes are methylated simultaneously. This might result in a pattern in which there is a small subset of cells which have several genes hypermethylated, which are likely transcriptionally inactive, while these genes remain mostly methylation free in other cells (Figure 7.2C).

This hypothesis suggests that a subset of cells is pre-primed by aberrant patterns of DNA methylation, mirroring patterns as observed during leukaemia progression. This co-ordinated methylation of genes is of particular interest in the development of leukaemia, and potentially other cancers. If a subset of cells is pre-primed by aberrant patterns of methylation, clonal expansion of one of these cells is sufficient to provide the leukaemia-like methylation phenotype. Our results suggest that the amount of these pre-primed cells are increased with an increasing age, providing a potential explanation of why older individuals are more likely to develop such a disease. Subsets of these cells however, are already present at birth and these pre-primed cells may thus alter disease susceptibility throughout life.

The hypothesis that pre-primed methylation patterns might lead to an increased risk of developing diseases is supported by the observation that during ALL remission,

patients with aberrant patterns of methylation subsequently relapsed, while the patients with methylation levels similar to the methylation levels observed in the healthy controls, survived the disease.

All these results together suggest that cancer, and potentially other age-related diseases, develops in cells which are pre-primed by the presence of aberrant patterns of DNA methylation.

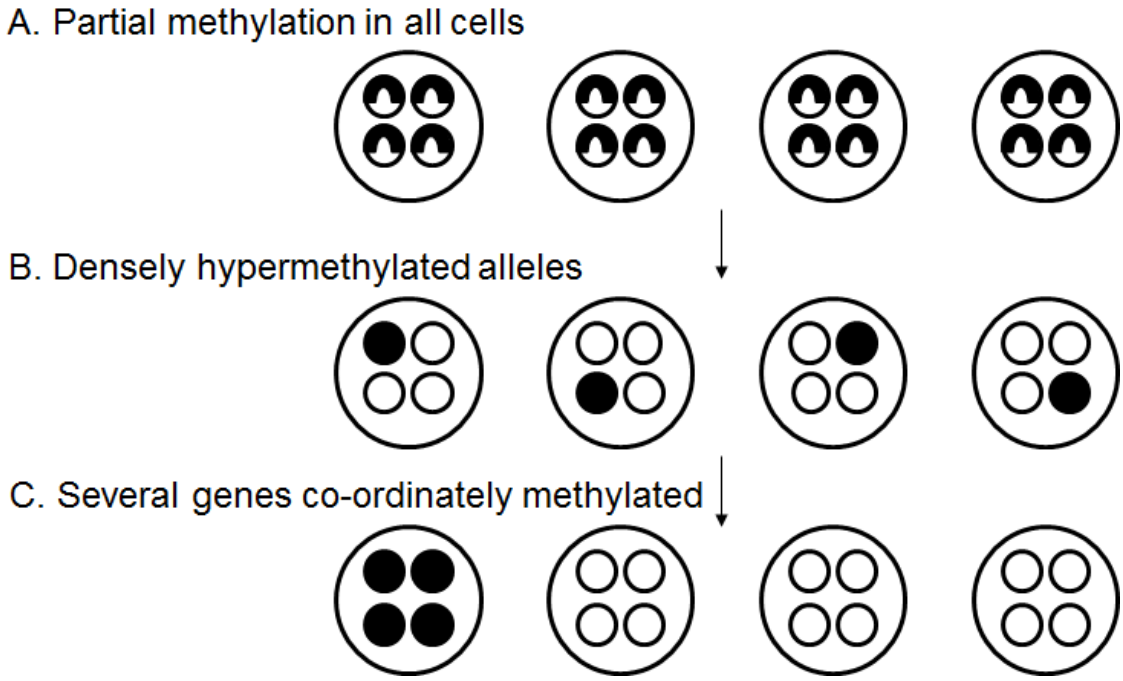


Figure 7.2: Pre-existing methylation hypothesis

All the results of the methylation patterns in healthy individuals combined led us to the hypothesis that pre-existing methylation patterns may underlie the susceptibility of cancer, and potentially other age-related diseases.

Methylation levels of leukaemia-related genes were observed in healthy individuals, and they were even already present at birth. These methylation levels might not be as randomly distributed as indicated by A, since preliminary results suggest that methylation might only be present in a small subset of alleles, while other alleles remain completely methylation free, as indicated by B. However, strong correlations were observed in the methylation levels of a subset of genes, suggesting co-ordinated methylation levels of these genes simultaneously. This might results in patterns of methylation as indicated by C, where a small number of cells show several genes hypermethylated, while other cell remain methylation free for the same set of genes.

Appendix A. Methods

DNA isolation	
Reagent B (Nucleon) 400 mM Tris/Cl 60 mM EDTA 150 mM NaCl 1% SDS Adjust pH to 8.0	***5M sodium perchlorate*** 720 gr Chlorate Adjust to 1 litre with dH_2O
Cell isolations	
Cell seperation buffer 1 PBS + 0.1% BSA + 2 mM EDTA For a total amount of 10ml: 10 ml PBS + 10µl BSA + 4µl 0.5M EDTA	***Cell separation buffer 2*** PBS + 0.1% BSA For a total amount of 10ml: 10 ml PBS + 10µl BSA, ph 7.4
0.5M EDTA Dilute 186.1 gram $Na_2EDTA.2H_2O$ in 800 ml dH_2O Adjust pH to 8.0 with NaOH Adjust volume to 1 litre with dH_2O	
Gel electrophoresis	
50x TAE (1 litre) 242 gram Tris 57.1 ml Acetic acid 18.6 gram EDTA/ 100 ml 0.5M EDTA (pH 8.0) Adjust volume to 1 litre with dH_2O	***1x TAE (1 litre)*** 20 ml 50x TAE 1980 ml dH_2O
PCR loading buffer 30 ml glycerol 70 ml dH_2O Add Orange G until the buffer is coloured orange	***1Kb+ Ladder*** 250µl 1µg/µl 1 Kb+ ladder from invitrogen Add 250µl loading buffer Add 1000µl dH_2O

Table A.1: An overview of the basic buffers used in this study

Bisulphite sequencing	
LB media	***Agar plates***
10g tryptone 5g yeast extract 10g NaCl Ajust volume to 1 litre with dH_2O Adjust pH to 7 After autoclaving, add 50 µg/ml Kanamycin	10g tryptone 5g yeast extract 10g NaCl 15g Agar Ajust volume to 1 litre with dH_2O Adjust pH to 7 After autoclaving, add 50 µg/ml Kanamycin Pour agar solution into the plates, allow to cool down and store 4 ºC
High-throughput sequencing	
TE 1 ml of 1M Tris-Hcl (pH 8.0) 0.2 ml of 0.5M EDTA Adjust volume to 100 ml with dH_2O	

Table A.1: An overview of the basic buffers used in this study

Gene	Description
TWIST2	*TWIST2* (*Dermo-1*) is a basic helix-loop-helix protein, which regulates transcription factors that are known to regulate proliferation and differentiation of myeloid lineage progenitors (Sharabi, 2008). *TWIST2* is one of the main regulators of CD7 in mature T-cells, and the failure of *TWIST2* regulation may results in the generation of abnormal T-cell populations (Koh, 2009). *TWIST2* is shown to act as a negative regulator of osteogenic differentiation, while promoting mesenchymal stromal/stem cell growth and adipogenesis (Isenmann, 2009). In ALL, epigenetic inactivation of *TWIST2* is shown to play a dual role; it alters cells growth and survival properties, and at the same time, it increases the resistance to chemotherapeutic treatment (Thathia, 2012).
TUSC3	*TUSC3,* also known as *N33* or putative prostate cancer tumour suppressor gene, is a magnesium transporter (science, 2012) and it is suggested to act as a potential tumour suppressor gene in several cancers, like breast cancer and pancreatic cancer (Cooke, 2008; Levy, 1999). *TUSC3* is methylated during ALL development, but not during AML development (Scholz, 2005).
EphA10	*EphA10* is a member of the EPH family, which are receptors for ephrin family ligands (Aasheim, 2005), and they are important in the cell signalling pathway (Kandouz, 2012). *EphA10* was found to be hypermethylated in various leukaemia cell lines and ALL patients (Kuang) and a high frequency of *EphA10* expression was observed in CLL cases (Alonso, 2009).
HAND2	*HAND2* is a basic helix-loop-helix transcription factor and it is affecting cell differentiation and cell type-specific gene expression in neural crest-derived noradrenergic sympathetic ganglion neurons. (Hendershot, 2008). It is one of the transcription factors that control cardiac gene expression (Bondue, 2010) and heart development and in mice it has been shown to be able to reprogram cardiac fibroblasts into functional cardiac-like myocytes (Song, 2012).
IGF2	Insulin-like growth factor 2 (*IGF2*), also known as the somatomedin A, is one of the key players being involved in fetal growth and development (St-Pierre, 2012), and is monoallelic expressed on the parental chromosome. In patients with the Beckwith-Wiedemann syndrome it was shown that some of the patients carry a constitutional mutation that results in biallelic expression of *IGF2* (Weksberg, 1993).

Table A.2: Overview of the genes examined in this project

Gene	Description
MLH1	MutL homolog 1 (*MLH1*) is a DNA mismatch repair gene, that is associated with hereditary nonpolyposis colorectal cancer (HNPCC) (Ward, 2012). It was shown that the ATPase domain of *MLH1* is important to transmit the class switch recombination signaling cascade, that lead to the affinity maturation and isotype switching, required to generate protective antibodies, both upstream and downstream of the generation of double-strand breaks (Chahwan, 2012).
HOX genes	The HOX genes encode homeodomain-containing transcription factors which are known to be key regulators of embryonic development. They are also expressed in adult cells, in which they play important roles in the control of cellular differentiation. The human genome contains 4 *HOX* clusters (A-D) (Cillo, 2001).
HOXD4	*HOXD4* is a sequence-specific transcription factor which is part of a developmental regulatory system that provides cells with specific positional identities (science, 2012). *HOXD4* is able to up-regulate the neuronal marker NEFM, but it cannot induce cell cycle arrest and morphologic differentiation (Zha, 2012). It is proposed that a *HOXD4* mutation is involved in the occurrence of childhood ALL (van Scherpenzeel Thim, 2005).
HOXA5	*HOXA5* is an important regulator of myeloid cell proliferation and differentiation. It functions as an important regulator of hematopoietic lineage determination and maturation (Fuller, 1999). *HOXA5* exhibit cell type specific methylation within normal cells, which correlates with gene expression. Indistinguishable levels of methylation was observed in both myeloid and lymphoid cells (Strathdee, 2007b). This gene is hypermethylated in all types of leukaemias. In AML, hypermethylation of *HOXA5* was associated with good prognosis (Strathdee, 2007a).
HOXA4	*HOXA4* might be important for maintaining spatial identity even in the adult aorta and downregulation of *HOXA4* expression may increase susceptibility to aortic disease such as human abdominal aortic aneurysms (Lillvis). As *HOXA5*, *HOXA4* is hypermethylated in all types of leukaemia and hypermethylation of *HOXA4* was associated with good prognosis (Strathdee, 2007a).

Table A.2: Overview of the genes examined in this project

Appendix B. Alterations of DNA methylation patterns with age

		Methylation levels neonates (%)	Methylation levels young adults (%)	Methylation levels adults (%)	Methylation levels older adults (%)		Neonates - Young adults	Young adults-Adults	Adults-Older adults
TUSC3	CpG1	9.72	11.09	14.37	15.49		**P=0.005**	**P<0.001**	P=0.148
	CpG2	4.91	5.98	8.17	8.93		**P<0.001**	**P<0.001**	P=0.121
	CpG3	1.43	2.49	3.92	6.22		**P<0.001**	**P<0.001**	**P<0.001**
	CpG4	2.76	3.86	5.41	7.23		**P<0.001**	**P<0.001**	**P=0.001**
	CpG5	2.76	4.03	6.68	7.40		**P<0.001**	**P<0.001**	P=0.186
	CpG6	1.76	3.61	6.23	8.23		**P<0.001**	**P<0.001**	**P=0.017**
TWIST2	CpG1	3.76	3.36	4.68	8.27		P=0.345	**P=0.001**	**P<0.001**
	CpG2	4.35	4.54	5.78	9.10		P=0.720	**P<0.001**	**P<0.001**
	CpG3	3.33	3.31	4.65	8.72		P=0.821	**P<0.001**	**P<0.001**
	CpG4	5.23	5.08	6.82	8.73		P=0.689	**P<0.001**	**P=0.006**
HOXD4	CpG1	6.58	8.88	10.75	13.84		**P=0.001**	**P=0.002**	**P=0.015**
	CpG2	6.58	11.43	14.57	19.11		**P<0.001**	**P=0.002**	**P=0.001**
	CpG3	3.42	7.64	10.70	14.93		**P<0.001**	**P<0.001**	**P<0.001**
	CpG4	3.32	7.28	9.09	11.90		**P<0.001**	**P=0.002**	**P=0.019**
	CpG5	5.00	8.61	11.13	14.30		**P<0.001**	**P<0.001**	**P=0.002**
EphA10	CpG1	14.97	16.39	18.90	22.91		**P=0.025**	**P<0.001**	**P<0.001**
	CpG2	4.58	5.07	6.55	8.03		P=0.054	**P<0.001**	**P=0.047**
	CpG3	2.00	2.33	3.22	5.00		P=0.137	**P<0.001**	**P=0.001**
	CpG4	4.57	5.52	5.68	8.56		**P=0.001**	P=0.363	**P<0.001**
HAND2	CpG1	2.82	3.84	4.59	5.05		**P<0.001**	**P=0.001**	P=0.715
	CpG2	5.02	8.89	10.48	10.70		**P<0.001**	**P=0.001**	P=0.842
	CpG3	3.54	5.35	6.99	7.13		**P<0.001**	**P<0.001**	P=0.359
	CpG4	1.99	4.13	6.50	6.50		**P<0.001**	**P<0.001**	P=0.308

Table B.1: Difference in levels methylation levels within age groups per individual CpG sites

HOXA5	CpG1	52.65	50.55	51.76	57.24		P=0.347	P=0.414	**P=0.026**
	CpG2	60.91	56.42	58.00	63.11		**P=0.028**	P=0.393	**P=0.048**
	CpG3	61.20	55.57	57.17	61.79		**P=0.002**	P=0.303	P=0.054
	CpG4	63.10	58.88	58.91	63.41		**P=0.046**	P=0.972	P=0.099
	CpG5	64.52	49.45	58.58	63.65		**P=0.003**	P=0.966	P=0.073
HOXA4	CpG1	26.45	34.26	35.49	40.34		**P<0.001**	P=0.584	**P=0.018**
	CpG2	25.98	36.48	38.57	39.94		**P<0.001**	P=0.316	P=0.376
	CpG3	23.62	39.99	43.67	44.46		**P<0.001**	P=0.147	P=0.647
	CpG5	30.64	38.54	40.32	41.55		**P=0.001**	P=0.371	P=0.513

Table B.1: Difference in levels methylation levels within age groups per individual CpG sites

Appendix C. Cell type specific DNA methylation

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*	*MLH1*
PBL – MNC	P=0.169	**P=0.005****	**P=0.028***	**P=0.013***	P=0.878	P=0.199	P=0.169	P=0.575	P=0.056
PBL-Lymphocytes	**P=0.005****	**P=0.005****	**P=0.013***	**P=0.007****	P=0.359	P=0.392	**P=0.013***	P=0.314	P=0.176
PBL- B-lymphocytes	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.028***	**P=0.025***	**P=0.005****	P=0.139	P=0.928
PBL – Monocytes	**P=0.005****	**P=0.009****	**P=0.009****	P=0.203	**P=0.013***	**P=0.032***	**P=0.022***	P=0.959	P=0.057
MNC- Lymphocytes	P=0.074	P=0.721	P=0.241	P=0.037	P=0.799	P=0.143	P=0.143	P=0.515	P=0.547
MNC- B-lymphocytes	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.013***	**P=0.002****	**P=0.005****	P=0.260	P=0.113
MNC- Monocytes	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.007****	**P=0.022***	P=0.319	P=0.037	P=0.760	P=0.979
B-lymphocytes- Lymphocytes	**P=0.006****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	P=0.241	**P=0.037***	P=0.374	P=0.183
Lymphocytes- Monocytes	**P=0.003****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.022***	**P=0.032***	**P=0.013***	P=0.214	P=0.382
B-lymphocytes- Monocytes	**P=0.003****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	**P=0.005****	P=0.086	P=0.066

Table C.1: Differences in methylation levels between cell types isolated from the young adult population

Significant differences in methylation levels were observed between the different cell types in the young adults for seven of the genes under study (Paired samples T-test for the genes *HOXA5* and *MLH1*, Wilcoxon signed rank test for all the other genes under study (SPSS)).

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*
PBL- MNC	**P=0.036***	**P=0.012***	**P=0.012***	**P=0.036***	P=0.262	P=0.398	P=0.149	P=0.225
PBL- Lymphocytes	**P=0.017***	**P=0.017***	**P=0.012***	**P=0.012***	**P=0.012***	P=0.612	P=0.330	P=0.442
PBL- B-lymphocytes	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.018**	P=0.237	P=0.054	**P=0.026***
PBL- T-lymphocytes	**P=0.017***	**P=0.012***	**P=0.035***	**P=0.012***	P=0.575	**P=0.043***	P=0.286	**P=0.919**
PBL- Monocytes	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	P=0.612	P=0.140	P=0.697
PBL- Granulocytes	P=0.572	P=0.263	P=0.484	P=0.123	P=0.750	P=0.499	P=0.386	P=0.521
MNC- Lymphocytes	P=0.575	P=0.327	**P=0.036***	**P=0.012***	P=0.093	P=0.310	P=0.830	**P=0.019***
MNC- B-lymphocytes	**P=0.012***	**P=0.012***	**P=0.017***	**P=0.012***	**P=0.012***	P=0.091	P=0.054	**P=0.003****
MNC- T-lymphocytes	P=0.093	P=0.327	P=0.093	P=0.050	P=0.674	P=0.063	P=0.974	P=0.293
MNC- Monocytes	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	P=0.176	P=0.085	P=0.217
MNC- Granulocytes	P=0.093	**P=0.025***	**P=0.036***	**P=0.025***	P=0.327	P=0.389	P=0.053	P=0.678
Lymphocytes- B-lymphocytes	**P=0.012***	**P=0.012***	**P=0.017***	**P=0.017***	**P=0.025**	P=0.063	**P=0.031***	**P<0.001*****
Lymphocytes- T-lymphocytes	P=0.161	P=0.674	P=0.889	P=0.123	P=0.889	P=0.398	P=0.905	P=0.308
Lymphocytes- Monocytes	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	P=0.237	P=0.185	P=0.090
Lymphocytes- Granulocytes	**P=0.012***	**P=0.017***	**P=0.012***	**P=0.012***	P=0.093	P=0.398	P=0.194	P=0.064
B-lymphocytes- T-lymphocytes	**P=0.012***	**P=0.012***	**P=0.036***	P=0.327	**P=0.012****	**P=0.043***	P=0.243	**P=0.017***
B-lymphocytes- Monocytes	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	P=0.612	**P=0.049***	**P=0.009***
B-lymphocytes- Granulocytes	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.017***	P=0.735	**P=0.047***	**P=0.033***
T-lymphocytes- Monocytes	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.012***	**P=0.025***	**P=0.028***	P=0.131	P=0.713
T-lymphocytes- Granulocytes	**P=0.012***	**P=0.017***	**P=0.012***	**P=0.012***	P=0.123	**P=0.018***	P=0.221	P=0.281
Monocytes- Granulocytes	**P=0.017***	**P=0.017***	**P=0.012***	**P=0.05***	P=0.208	P=1.000	P=0.541	P=0.445

Table C.2: Differences in methylation levels between cell types isolated from the adult population

Significant differences in methylation levels were observed between the different cell types in the adult population for seven of the genes under study (Paired samples T-test for the genes *HOXA4* and *IGF2*, Wilcoxon signed rank test for all the other genes under study (SPSS)).

		TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*	*IGF2*	1st 5 genes combined
B-lymphocytes	Methylation levels young adults (%)	11.14	14.44	20.44	14.64	10.80	57.09	46.18	49.20	14.29
	Methylation levels adults (%)	13.03	13.88	19.85	15.87	13.48	56.51	51.43	47.13	15.22
	Young adults- Adults	P=0.481	P=0.481	P=0.888	P=0.743	**P=0.011***	P=0.743	P=0.114	P=0.773	P=0.290

Table C.3: Median methylation levels in the B-lymphocytes in the young adult population compared to the adult population

One of the participants of the young adult age group exhibited high levels of methylation. To examine if the differences observed between the age groups was driven by that one outlier, the median was determined rather than the mean (Wilcoxon signed rank test (SPSS)).

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*
Stem-and early progenitor cells- B-cell progenitors	P=0.600	P=0.232	P=0.893	P=0.204	P=0.893	P=0.109	P=0.651
Stem-and early progenitor cells- Stem- and early progenitor cells minus B-cell progenitors	P=0.833	P=0.860	P=0.715	P=0.712	P=0.500	P=0.273	0.968
Stem-and early progenitor cells- B-lymphocytes	P=0.460	**P=0.011****	P=0.463	P=0.097	P=0.345	P=0.893	P=0.128
Stem-and early progenitor cells- T lymphocytes	P=0.046	P=0.058	P=0.345	**P=0.021***	P=0.345	P=0.463	P=0.195
Stem-and early progenitor cells- Monocytes	P=0.735	P=0.521	P=0.116	P=0.516	P=0.917	P=0.345	P=0.317
B-cell progenitors - Stem- and early progenitor cells minus B-cell progenitors	0=0.345	P=0.685	P=0.715	P=0.372	P=0.500	P=0.655	P=0.479
B-cell progenitors- B-lymphocytes	P=0.752	**P=0.038***	P=0.345	P=0.190	P=0.686	P=0.109	P=0.903
B-cell progenitors- T-lymphocytes	P=0.917	**P=0.003****	P=0.686	P=0.231	P=0.500	P=0.109	P=0.898
B-cell progenitors- Monocytes	P=0.463	P=0.566	P=0.686	P=0.183	P=0.893	P=0.109	P=0.770
B-lymphocytes- T-lymphocytes	P=0.173	P=0.689	P=0.463	P=0.800	P=0.463	P=0.225	P=0.432
B-lymphocytes- Monocytes	**P=0.028***	**P=0.020***	P=0.075	**P=0.025***	P=0.463	P=0.500	P=0.108
T-lymphocytes- Monocytes	**P=0.028***	**P=0.013***	**P=0.028***	**P=0.001*****	P=0.345	P=0.225	P=0.183

Table C.4: Comparisons of methylation levels in the different cell types isolated from the G-CSF mobilized samples

Statistical differences were observed between different cell types in a small number of genes (Paired samples T-test for *TUSC3, EpHA10* and *HOXA4*, Wilcoxon signed rank test for the other genes under study, SPSS).

A.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*
TWIST2		**R= 0.843** **P= 0.001**	**R= 0.864** **P= 0.001**	**R= 0.800** **P= 0.003**	R= 0.308 P= 0.361	R= -0.531 P= 0.093	**R= 0.701** **P= 0.016**
TUSC3			**R= 0.844** **P= 0.001**	**R= 0.711** **P= 0.014**	R= 0.182 P= 0.593	R= -0.481 P= 0.134	R= 0.573 P= 0.065
HOXD4				**R= 0.628** **P= 0.038**	R= 0.269 P= 0.424	R= -0.514 P= 0.105	R= 0.531 P= 0.093
EphA10					R= 0.490 P= 0.126	R= -0.137 P= 0.689	R= 0.531 P= 0.092
HAND2						R= 0.260 P= 0.440	R= 0.432 P= 0.185
HOXA5							R= -0.226 P= 0.504
HOXA4							

Table C.5A: Correlations between genes in the isolated cell types in the young adults; the lymphocytes

In the lymphocyte samples, correlations were observed between *TWIST2*, *TUSC3*, *HOXD4* and *EphA10* ($p<0.05$). A correlation was observed between *TWIST2* and *HOXA4*. The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

B.

	TWIST2	*TUSC3*	*HOXD4*	*EphA10*	*HAND2*	*HOXA5*	*HOXA4*
TWIST2		R= 0.070 P= 0.839	R= -0.198 P= 0.559	R= 0.313 P= 0.349	**R= -0.258** **P= 0.044**	**R= -0.648** **P= 0.031**	R= 0.532 P= 0.092
TUSC3			R= -0.112 P= 0.743	R= -0.140 P= 0.682	R= -0.312 P= 0.349	R= -0.306 P= 0.359	R= 0.083 P= 0.809
HOXD4				R= 0.546 P= 0.082	**R= 0.651** **P= 0.030**	R= 0.232 P= 0.493	R= 0.515 P= 0.105
EphA10					**R= 0.645** **P= 0.032**	R= 0.046 P= 0.892	R**= 0.730** **P= 0.011**
HAND2						**R= 0.617** **P= 0.043**	R= 0.341 P= 0.305
HOXA5							R= -0.266 P= 0.430
HOXA4							

Table C.5B: Correlations between genes in the isolated cell types in the young adults; the monocytes

In the monocyte samples, there are correlation observed between *HAND2* with *HOXD4* and *EphA10*, between *EphA10* and *HOXA4* and between *HOXA5* with *TWIST2* and *HAND2*. The R is indicating the correlation coefficient and the P the observed p-values (Pearson correlation (SPSS)).

Appendix D. DNA methylation patterns during the development of age-related diseases

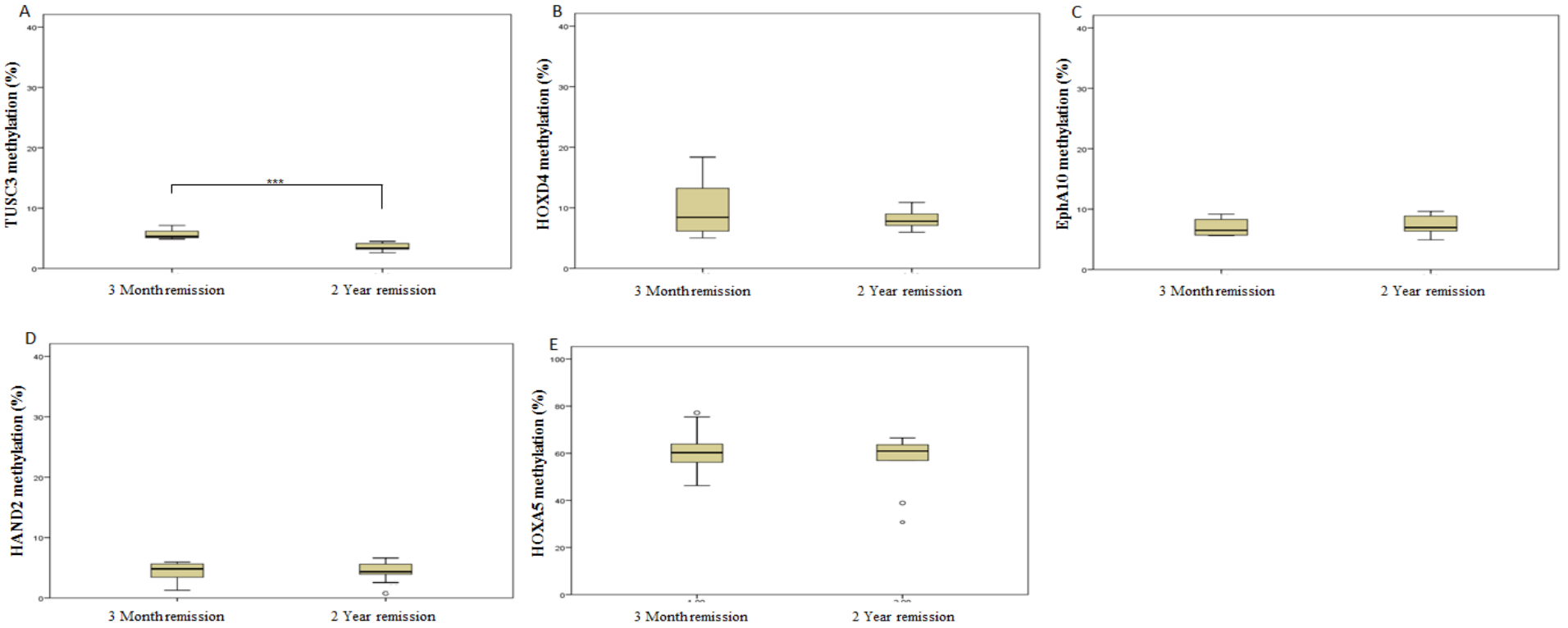


Figure D.1A-E: Methylation levels are relatively stable during ALL remission

A) *TUSC3* methylation showed lower levels after 2 years remission compared to the first remission sample, taken at 3 months into remission.

B-E) Overall methylation levels during remission are relatively stable in the childhood patients for *HOXD4*, *EphA10*, *HAND2* and *HOXA5*. Some patients exhibited a further increase in methylation, while other patients exhibited a reduction later into remission.

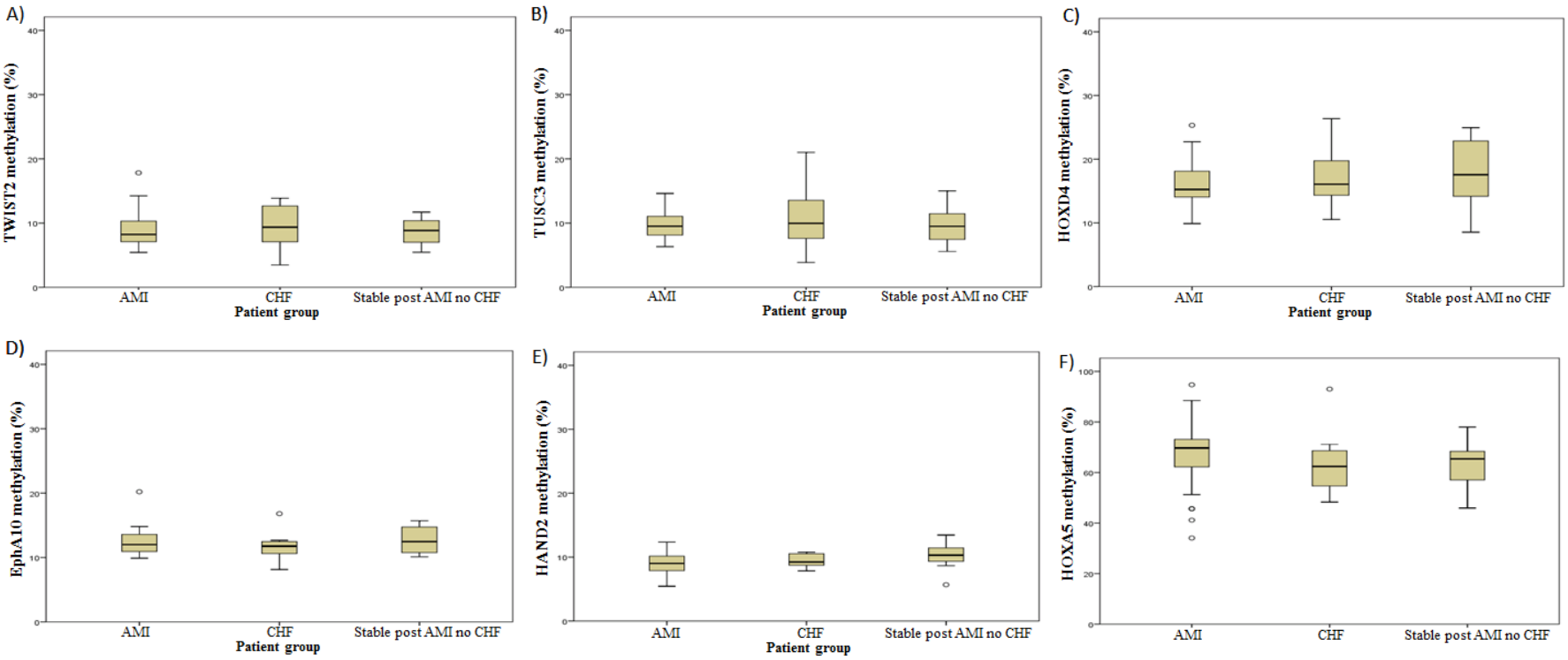


Figure D.2A-F: No differences in methylation levels in heart disease patients within different patient groups

A-F) No differences were observed between the heart disease patients within different patient groups for any of the genes under study.

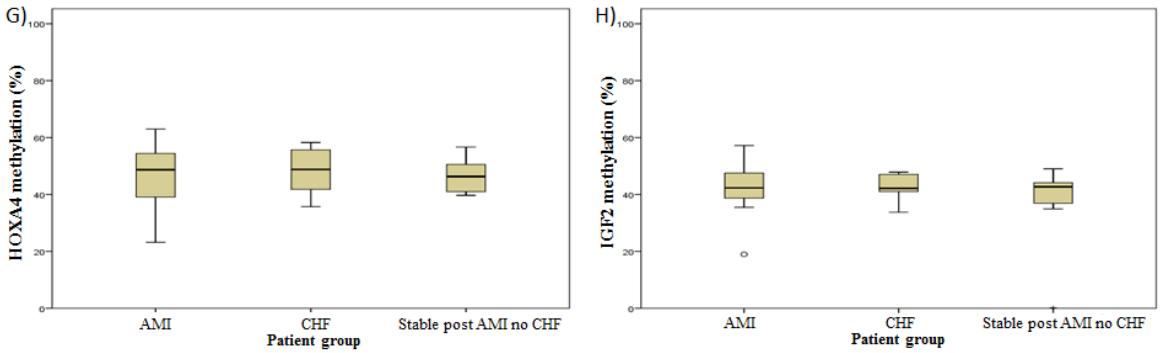


Figure D.2G and H: No differences in methylation levels in heart disease patients within different patient groups

G and H) No differences were observed between the heart disease patients within different patient groups for any of the genes under study.

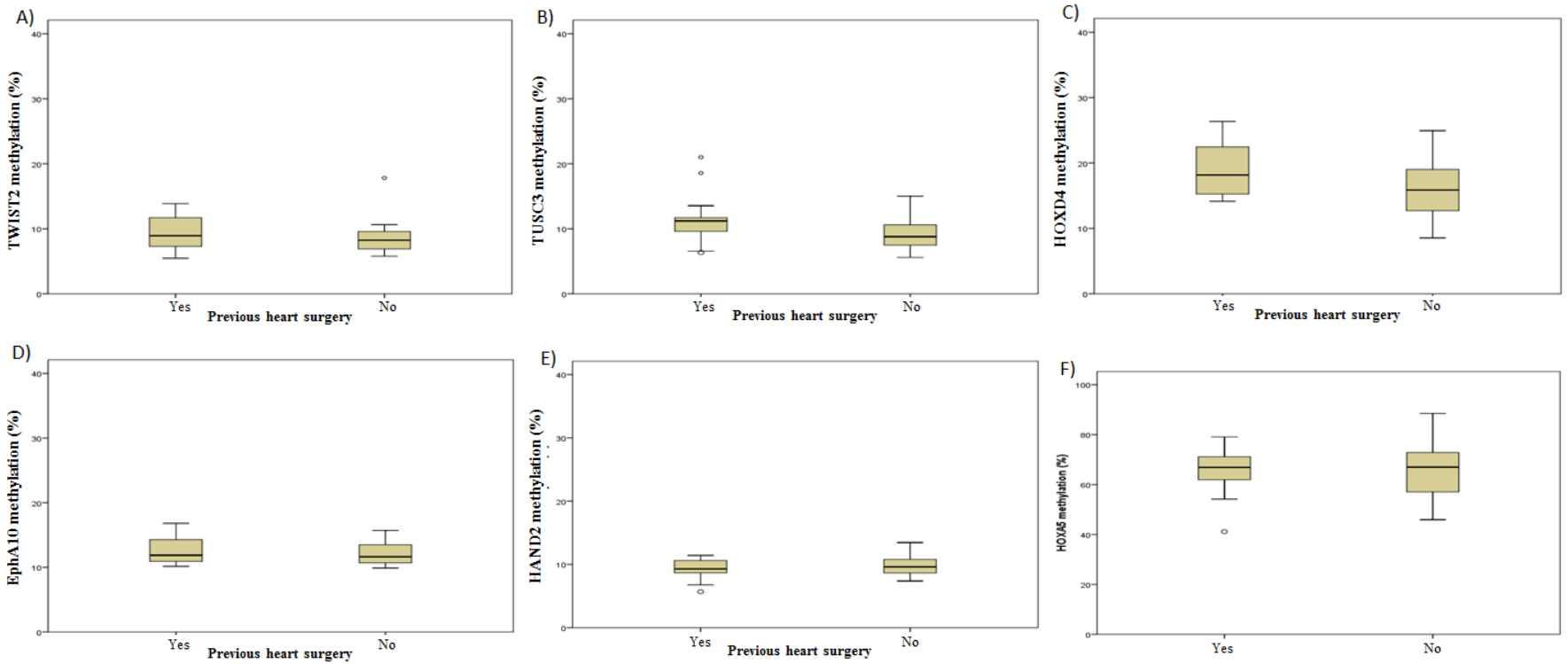


Figure D.3A-F: Methylation patterns between disease patient groups: previous heart surgery

A-F) No significant differences in methylation levels were observed between the heart disease patients who underwent previous heart surgery and patients who did not, for all genes under study.

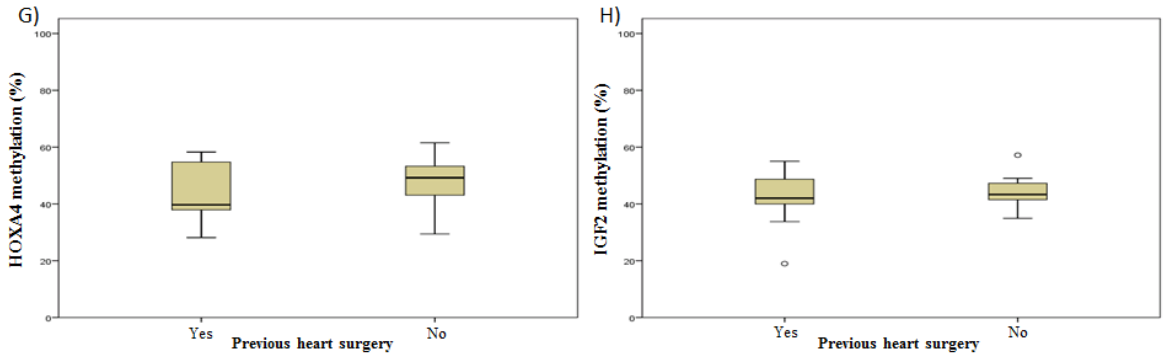


Figure D.3G and H: Methylation patterns between disease patient groups: previous heart surgery

G and H)| No significant differences in methylation levels were observed between the heart disease patients who underwent previous heart surgery and patients who did not, for all genes under study.

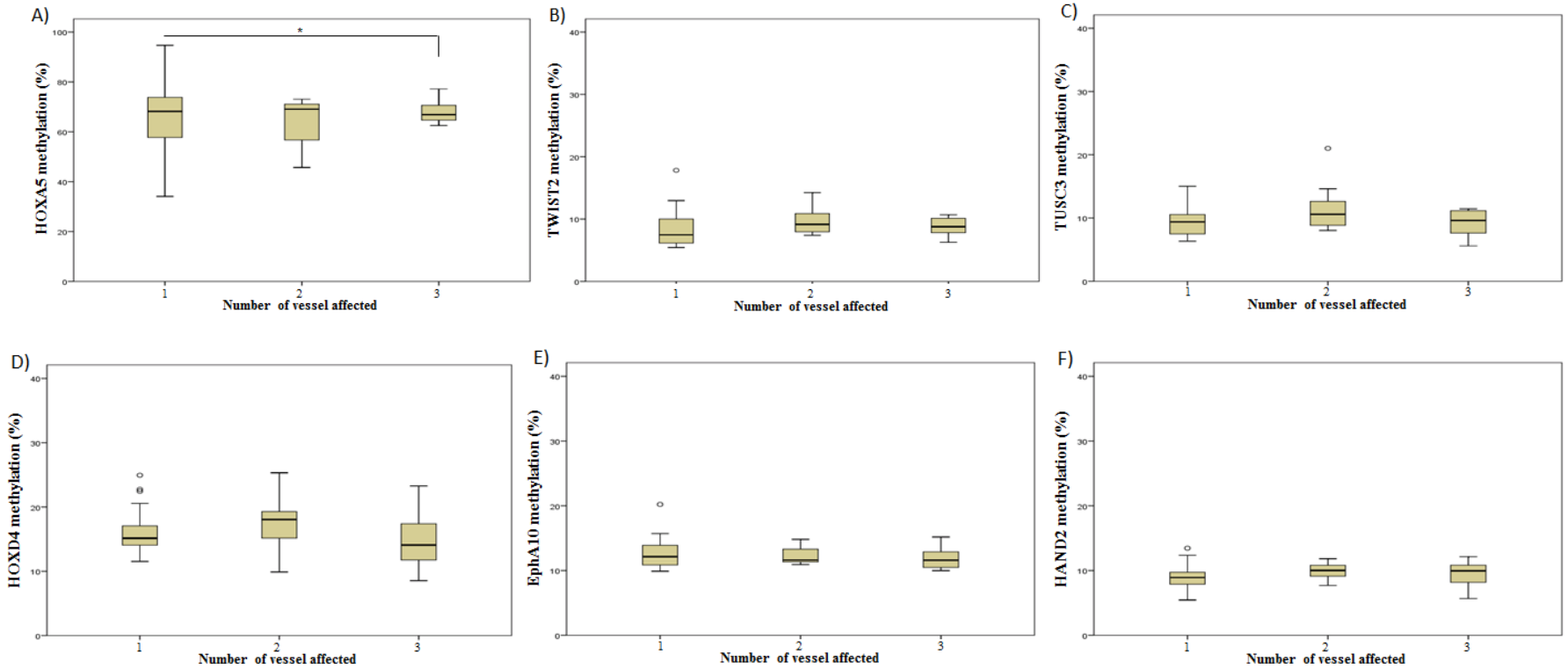


Figure D.4A-F: Methylation patterns between heart disease patient groups: number of vessels affected

A) A significant difference in methylation levels is observed between the patients with one vessel affected and the patients with three vessels affected for *HOXA5* (p=0.034, Independent samples T-test, SPSS).

B-F) No differences in methylation levels were observed between heart disease patients with one, two or three vessels affected by the heart disease for *TWIST2*, *TUSC3*, *HOXD4*, *EphA10* and *HAND2*.

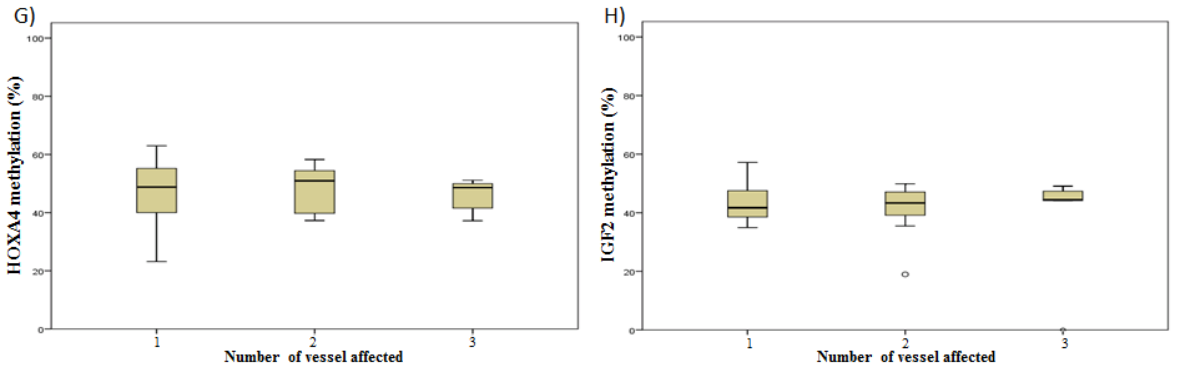


Figure D.4G and H: Methylation patterns between heart disease patient groups: number of vessels affected

G and H) No differences in methylation levels were observed between heart disease patients with one, two or three vessels affected by the heart disease for *HOXA4* and *IGF2*.

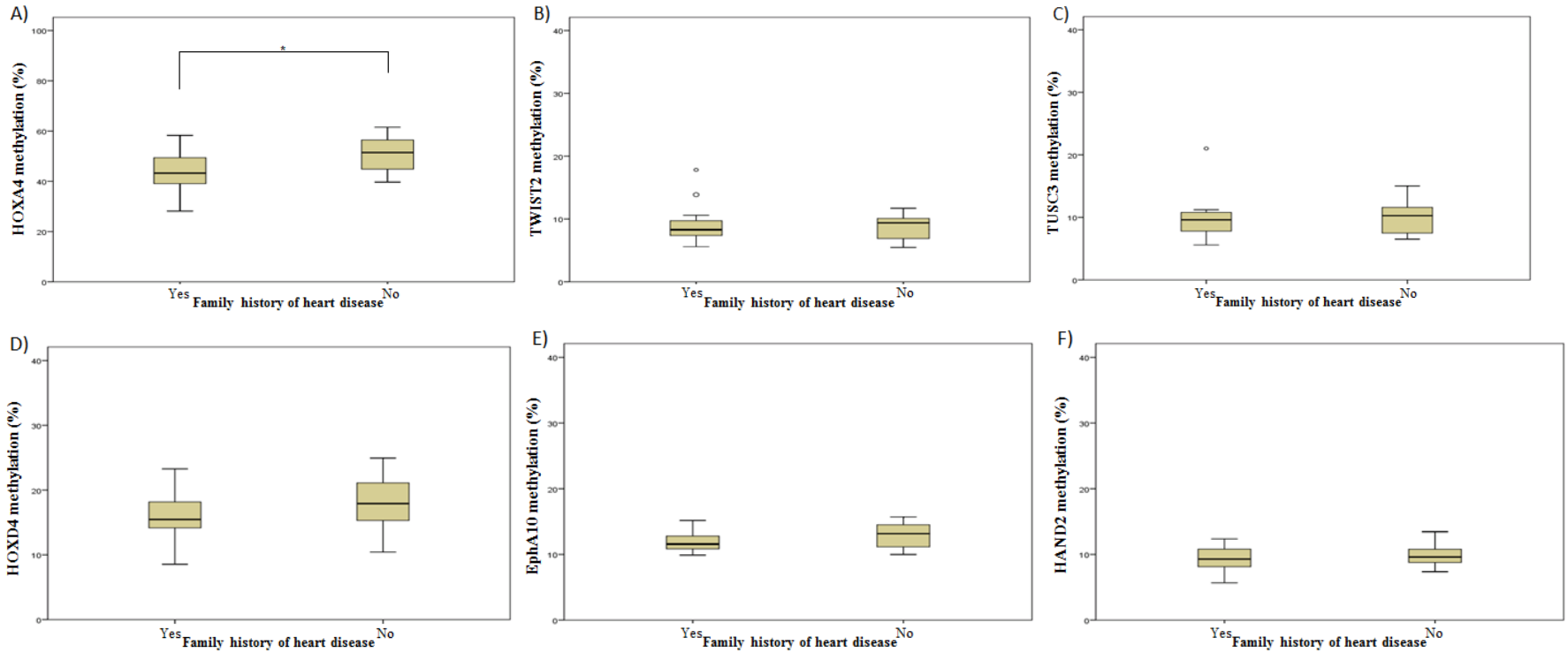


Figure D.5A-F: Methylation patterns between disease patient groups: family history of heart disease

A) A significant differences was observed in the methylation levels of the heart disease patients with a family history of heart disease and the patients without such a family history for *HOXA4* (p=0.036, Mann-Whitney U and Wilcoxon Signed Ranks test, SPSS)

B-F) No significant differences were observed in the methylation levels of heart disease patients with a family history of heart disease and patients without such a family history for *TWIST2*, *TUSC3*, *HOXD4*, *EphA10* and *HAND2*.

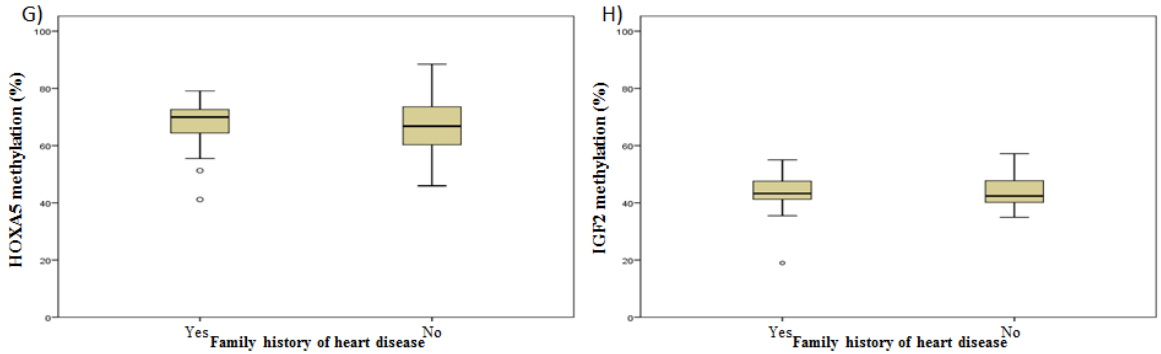


Figure D.5G and H: Methylation patterns between disease patient groups: family history of heart disease

G and H) No significant differences were observed in the methylation levels of heart disease patients with a family history of heart disease and patients without such a family history for *HOXA5* and *IGF2*.

References

AALTONEN, L. A., SALOVAARA, R., KRISTO, P., CANZIAN, F., HEMMINKI, A., PELTOMAKI, P., CHADWICK, R. B., KAARIAINEN, H., ESKELINEN, M., JARVINEN, H., MECKLIN, J. P. & DE LA CHAPELLE, A. 1998. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med,* 338**,** 1481-7.

AASHEIM, H. C., PATZKE, S., HJORTHAUG, H. S. & FINNE, E. F. 2005. Characterization of a novel Eph receptor tyrosine kinase, EphA10, expressed in testis. *Biochim Biophys Acta,* 1723**,** 1-7.

ADKINS, R. M., KRUSHKAL, J., TYLAVSKY, F. A. & THOMAS, F. 2011. Racial differences in gene-specific DNA methylation levels are present at birth. *Birth Defects Res A Clin Mol Teratol,* 91**,** 728-36.

ADKINS, R. M., THOMAS, F., TYLAVSKY, F. A. & KRUSHKAL, J. Parental ages and levels of DNA methylation in the newborn are correlated. *BMC Med Genet,* 12**,** 47.

AHMADIAN, A., EHN, M. & HOBER, S. 2006. Pyrosequencing: history, biochemistry and future. *Clin Chim Acta,* 363**,** 83-94.

ALONSO, C. L., TRINIDAD, E. M., DE GARCILLAN, B., BALLESTEROS, M., CASTELLANOS, M., COTILLO, I., MUNOZ, J. J. & ZAPATA, A. G. 2009. Expression profile of Eph receptors and ephrin ligands in healthy human B lymphocytes and chronic lymphocytic leukemia B-cells. *Leuk Res,* 33**,** 395-406.

ANDERSON, K., LUTZ, C., VAN DELFT, F. W., BATEMAN, C. M., GUO, Y., COLMAN, S. M., KEMPSKI, H., MOORMAN, A. V., TITLEY, I., SWANSBURY, J., KEARNEY, L., ENVER, T. & GREAVES, M. 2011. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature,* 469**,** 356-61.

ANTELO, M., BALAGUER, F., SHIA, J., SHEN, Y., HUR, K., MOREIRA, L., CUATRECASAS, M., BUJANDA, L., GIRALDEZ, M. D., TAKAHASHI, M., CABANNE, A., BARUGEL, M. E., ARNOLD, M., ROCA, E. L., ANDREU, M., CASTELLVI-BEL, S., LLOR, X., JOVER, R., CASTELLS, A., BOLAND, C. R. & GOEL, A. 2012. A high degree of LINE-1 hypomethylation is a unique feature of early-onset colorectal cancer. *PLoS One,* 7**,** e45357.

ANTEQUERA, F. 2003. Structure, function and evolution of CpG island promoters. *Cell Mol Life Sci,* 60**,** 1647-58.

ANTEQUERA, F. & BIRD, A. 1993. Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci U S A,* 90**,** 11995-9.

ARAND, J., SPIELER, D., KARIUS, T., BRANCO, M. R., MEILINGER, D., MEISSNER, A., JENUWEIN, T., XU, G., LEONHARDT, H., WOLF, V. & WALTER, J. 2012. In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLoS Genet,* 8**,** e1002750.

ARBAB-ZADEH, A., DIJK, E., PRASAD, A., FU, Q., TORRES, P., ZHANG, R., THOMAS, J. D., PALMER, D. & LEVINE, B. D. 2004. Effect of aging and physical activity on left ventricular compliance. *Circulation,* 110**,** 1799-805.

AVRAHAM, R. & YARDEN, Y. 2012. Regulation of signalling by microRNAs. *Biochem Soc Trans,* 40**,** 26-30.

BAL, A., VERMA, S., JOSHI, K., SINGLA, A., THAKUR, R., ARORA, S. & SINGH, G. 2012. BRCA1-methylated sporadic breast cancers are BRCA-like in showing a basal phenotype and absence of ER expression. *Virchows Arch,* 461**,** 305-12.

BARON, B. 2012. *Breaking the Silence: The Interplay Between Transcription Factors and DNA Methylation,* Methylation- From DNA, RNA and Histones to Diseases and Treatment.

BAYLIN, S. B. 2005. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol,* 2 Suppl 1**,** S4-11.

BEGUS-NAHRMANN, Y., HARTMANN, D., KRAUS, J., ESHRAGHI, P., SCHEFFOLD, A., GRIEB, M., RASCHE, V., SCHIRMACHER, P., LEE, H. W., KESTLER, H. A., LECHEL, A. & RUDOLPH, K. L. 2012. Transient telomere dysfunction induces chromosomal instability and promotes carcinogenesis. *J Clin Invest,* 122**,** 2283-8.

BELL, J. T., PAI, A. A., PICKRELL, J. K., GAFFNEY, D. J., PIQUE-REGI, R., DEGNER, J. F., GILAD, Y. & PRITCHARD, J. K. 2011. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol,* 12**,** R10.

BELL, J. T., TSAI, P. C., YANG, T. P., PIDSLEY, R., NISBET, J., GLASS, D., MANGINO, M., ZHAI, G., ZHANG, F., VALDES, A., SHIN, S. Y., DEMPSTER, E. L., MURRAY, R. M., GRUNDBERG, E., HEDMAN, A. K., NICA, A., SMALL, K. S., DERMITZAKIS, E. T., MCCARTHY, M. I., MILL, J., SPECTOR, T. D. & DELOUKAS, P. 2012. Epigenome-Wide Scans Identify Differentially Methylated Regions for Age and Age-Related Phenotypes in a Healthy Ageing Population. *PLoS Genet,* 8**,** e1002629.

BELLIZZI, D., D'AQUILA, P., MONTESANTO, A., CORSONELLO, A., MARI, V., MAZZEI, B., LATTANZIO, F. & PASSARINO, G. Global DNA methylation in old subjects is correlated with frailty. *Age (Dordr)*.

BELSHAW, N. J., ELLIOTT, G. O., FOXALL, R. J., DAINTY, J. R., PAL, N., COUPE, A., GARG, D., BRADBURN, D. M., MATHERS, J. C. & JOHNSON, I. T. 2008. Profiling CpG island field methylation in both morphologically normal and neoplastic human colonic mucosa. *Br J Cancer,* 99**,** 136-42.

BERNDSEN, C. E. & DENU, J. M. 2008. Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Current Opinion in Structural Biology,* 18**,** 682-689.

BIRD, A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev,* 16**,** 6-21.

BIRD, A. P. 1986. CpG-rich islands and the function of DNA methylation. *Nature,* 321**,** 209-13.

BIRD, A. P. 1996. The relationship of DNA methylation to cancer. *Cancer Surv,* 28**,** 87-101.

BJORNSSON, H. T., SIGURDSSON, M. I., FALLIN, M. D., IRIZARRY, R. A., ASPELUND, T., CUI, H., YU, W., RONGIONE, M. A., EKSTROM, T. J., HARRIS, T. B., LAUNER, L. J., EIRIKSDOTTIR, G., LEPPERT, M. F., SAPIENZA, C., GUDNASON, V. & FEINBERG, A. P. 2008. Intra-individual change over time in DNA methylation with familial clustering. *Jama,* 299**,** 2877-83.

BOCK, C., BEERMAN, I., LIEN, W. H., SMITH, Z. D., GU, H., BOYLE, P., GNIRKE, A., FUCHS, E., ROSSI, D. J. & MEISSNER, A. 2012. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Mol Cell,* 47**,** 633-47.

BOCKER, M. T., HELLWIG, I., BREILING, A., ECKSTEIN, V., HO, A. D. & LYKO, F. Genome-wide promoter DNA methylation dynamics of human hematopoietic progenitor cells during differentiation and aging. *Blood*.

BOCKLANDT, S., LIN, W., SEHL, M. E., SANCHEZ, F. J., SINSHEIMER, J. S., HORVATH, S. & VILAIN, E. Epigenetic predictor of age. *PLoS One,* 6**,** e14821.

BOLLATI, V., SCHWARTZ, J., WRIGHT, R., LITONJUA, A., TARANTINI, L., SUH, H., SPARROW, D., VOKONAS, P. & BACCARELLI, A. 2009. Decline

in genomic DNA methylation through aging in a cohort of elderly subjects. *Mech Ageing Dev,* 130**,** 234-9.

BONDUE, A. & BLANPAIN, C. 2010. Mesp1: a key regulator of cardiovascular lineage commitment. *Circ Res,* 107**,** 1414-27.

BONNET, D., SELVES, J., TOULAS, C., DANJOUX, M., DUFFAS, J. P., PORTIER, G., KIRZIN, S., GHOUTI, L., CARRERE, N., SUC, B., ALRIC, L., BARANGE, K., BUSCAIL, L., CHAUBARD, T., IMANI, K. & GUIMBAUD, R. 2012. Simplified identification of Lynch syndrome: a prospective, multicenter study. *Dig Liver Dis,* 44**,** 515-22.

BORGHOL, N., SUDERMAN, M., MCARDLE, W., RACINE, A., HALLETT, M., PEMBREY, M., HERTZMAN, C., POWER, C. & SZYF, M. 2012. Associations with early-life socio-economic position in adult DNA methylation. *Int J Epidemiol,* 41**,** 62-74.

BRONNER, C., FUHRMANN, G., CHEDIN, F. L., MACALUSO, M. & DHE-PAGANON, S. 2010. UHRF1 Links the Histone code and DNA Methylation to ensure Faithful Epigenetic Memory Inheritance. *Genet Epigenet,* 2009**,** 29-36.

BRUCE, K. D. & CAGAMPANG, F. R. 2011. Epigenetic priming of the metabolic syndrome. *Toxicol Mech Methods,* 21**,** 353-61.

BYUN, H. M., SIEGMUND, K. D., PAN, F., WEISENBERGER, D. J., KANEL, G., LAIRD, P. W. & YANG, A. S. 2009. Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. *Hum Mol Genet,* 18**,** 4808-17.

CAHILL, N., BERGH, A. C., KANDURI, M., GORANSSON-KULTIMA, H., MANSOURI, L., ISAKSSON, A., RYAN, F., SMEDBY, K. E., JULIUSSON, G., SUNDSTROM, C., ROSEN, A. & ROSENQUIST, R. 2012. 450K-array analysis of chronic lymphocytic leukemia cells reveals global DNA methylation to be relatively stable over time and similar in resting and proliferative compartments. *Leukemia*.

CALVANESE, V., FERNANDEZ, A. F., URDINGUIO, R. G., SUAREZ-ALVAREZ, B., MANGAS, C., PEREZ-GARCIA, V., BUENO, C., MONTES, R., RAMOS-MEJIA, V., MARTINEZ-CAMBLOR, P., FERRERO, C., ASSENOV, Y., BOCK, C., MENENDEZ, P., CARRERA, A. C., LOPEZ-LARREA, C. & FRAGA, M. F. A promoter DNA demethylation landscape of human hematopoietic differentiation. *Nucleic Acids Res*.

CANCER-RESEARCH, U. 2013. *www.cancerresearchuk.org* [Online].

CANNON, B. 2013. Cardiovascular disease: Biochemistry to behaviour. *Nature,* 493**,** S2-3.

CEDAR, H. & BERGMAN, Y. 2011. Epigenetics of haematopoietic cell development. *Nat Rev Immunol,* 11**,** 478-88.

CELLSIGNAL. 2011. *PI3K/Akt signaling* [Online]. www.cellsignal.com.

CHAHWAN, R., VAN OERS, J. M., AVDIEVICH, E., ZHAO, C., EDELMANN, W., SCHARFF, M. D. & ROA, S. 2012. The ATPase activity of MLH1 is required to orchestrate DNA double-strand breaks and end processing during class switch recombination. *J Exp Med,* 209**,** 671-8.

CHALLEN, G. A., SUN, D., JEONG, M., LUO, M., JELINEK, J., BERG, J. S., BOCK, C., VASANTHAKUMAR, A., GU, H., XI, Y., LIANG, S., LU, Y., DARLINGTON, G. J., MEISSNER, A., ISSA, J. P., GODLEY, L. A., LI, W. & GOODELL, M. A. 2012. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet,* 44**,** 23-31.

CHAN, T. L., YUEN, S. T., KONG, C. K., CHAN, Y. W., CHAN, A. S., NG, W. F., TSUI, W. Y., LO, M. W., TAM, W. Y., LI, V. S. & LEUNG, S. Y. 2006. Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nat Genet,* 38**,** 1178-83.

CHANG, P. Y., LU, S. C., LEE, C. M., CHEN, Y. J., DUGAN, T. A., HUANG, W. H., CHANG, S. F., LIAO, W. S., CHEN, C. H. & LEE, Y. T. 2008. Homocysteine inhibits arterial endothelial cell growth through transcriptional downregulation of fibroblast growth factor-2 involving G protein and DNA methylation. *Circ Res,* 102**,** 933-41.

CHASE, D. S., TAWN, E. J., PARKER, L., JONAS, P., PARKER, C. O. & BURN, J. 1998. The North Cumbria Community Genetics Project. *J Med Genet,* 35**,** 413-6.

CHEN, C., LIU, Y. & ZHENG, P. 2009. mTOR regulation and therapeutic rejuvenation of aging hematopoietic stem cells. *Sci Signal,* 2**,** ra75.

CHEN, Z. X. & RIGGS, A. D. 2011. DNA methylation and demethylation in mammals. *J Biol Chem,* 286**,** 18347-53.

CHRISTENSEN, B. C. & MARSIT, C. J. 2011. Epigenomics in environmental health. *Front Genet,* 2**,** 84.

CHRISTENSEN, K., DOBLHAMMER, G., RAU, R. & VAUPEL, J. W. 2009. Ageing populations: the challenges ahead. *Lancet,* 374**,** 1196-208.

CHUANG, J. C., WARNER, S. L., VOLLMER, D., VANKAYALAPATI, H., REDKAR, S., BEARSS, D. J., QIU, X., YOO, C. B. & JONES, P. A. S110, a 5-Aza-2'-deoxycytidine-containing dinucleotide, is an effective DNA methylation inhibitor in vivo and can reduce tumor growth. *Mol Cancer Ther,* 9**,** 1443-50.

CILLO, C., CANTILE, M., FAIELLA, A. & BONCINELLI, E. 2001. Homeobox genes in normal and malignant cells. *J Cell Physiol,* 188**,** 161-9.

CIMMINO, L., ABDEL-WAHAB, O., LEVINE, R. L. & AIFANTIS, I. TET Family Proteins and Their Role in Stem Cell Differentiation and Transformation. *Cell Stem Cell,* 9**,** 193-204.

COLLERTON, J., BARRASS, K., BOND, J., ECCLES, M., JAGGER, C., JAMES, O., MARTIN-RUIZ, C., ROBINSON, L., VON ZGLINICKI, T. & KIRKWOOD, T. 2007. The Newcastle 85+ study: biological, clinical and psychosocial factors associated with healthy ageing: study protocol. *BMC Geriatr,* 7**,** 14.

COOKE, S. L., POLE, J. C., CHIN, S. F., ELLIS, I. O., CALDAS, C. & EDWARDS, P. A. 2008. High-resolution array CGH clarifies events occurring on 8p in carcinogenesis. *BMC Cancer,* 8**,** 288.

COOPER, D. N. & KRAWCZAK, M. 1989. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Hum Genet,* 83**,** 181-8.

CORDISCO, S., MAURELLI, R., BONDANZA, S., STEFANINI, M., ZAMBRUNO, G., GUERRA, L. & DELLAMBRA, E. 2010. Bmi-1 reduction plays a key role in physiological and premature aging of primary human keratinocytes. *J Invest Dermatol,* 130**,** 1048-62.

CREPIN, M., DIEU, M. C., LEJEUNE, S., ESCANDE, F., BOIDIN, D., PORCHET, N., MORIN, G., MANOUVRIER, S., MATHIEU, M. & BUISINE, M. P. 2012. Evidence of constitutional MLH1 epimutation associated to transgenerational inheritance of cancer susceptibility. *Hum Mutat,* 33**,** 180-8.

CROSS, S. H. & BIRD, A. P. 1995. CpG islands and genes. *Curr Opin Genet Dev,* 5**,** 309-14.

DAMMANN, R., YANG, G. & PFEIFER, G. P. 2001. Hypermethylation of the cpG island of Ras association domain family 1A (RASSF1A), a putative tumor suppressor gene from the 3p21.3 locus, occurs in a large percentage of human breast cancers. *Cancer Res,* 61**,** 3105-9.

DAVIES, M. N., VOLTA, M., PIDSLEY, R., LUNNON, K., DIXIT, A., LOVESTONE, S., COARFA, C., HARRIS, R. A., MILOSAVLJEVIC, A., TROAKES, C., AL-SARRAJ, S., DOBSON, R., SCHALKWYK, L. C. &

MILL, J. 2012. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol,* 13**,** R43.

DAY, N., OAKES, S., LUBEN, R., KHAW, K. T., BINGHAM, S., WELCH, A. & WAREHAM, N. 1999. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer,* 80 Suppl 1**,** 95-103.

DE GOBBI, M., GARRICK, D., LYNCH, M., VERNIMMEN, D., HUGHES, J. R., GOARDON, N., LUC, S., LOWER, K. M., SLOANE-STANLEY, J. A., PINA, C., SONEJI, S., RENELLA, R., ENVER, T., TAYLOR, S., JACOBSEN, S. E., VYAS, P., GIBBONS, R. J. & HIGGS, D. R. 2011. Generation of bivalent chromatin domains during cell fate decisions. *Epigenetics Chromatin,* 4**,** 9.

DETICH, N., HAMM, S., JUST, G., KNOX, J. D. & SZYF, M. 2003. The methyl donor S-Adenosylmethionine inhibits active demethylation of DNA: a candidate novel mechanism for the pharmacological effects of S-Adenosylmethionine. *J Biol Chem,* 278**,** 20812-20.

DING, L., LEY, T. J., LARSON, D. E., MILLER, C. A., KOBOLDT, D. C., WELCH, J. S., RITCHEY, J. K., YOUNG, M. A., LAMPRECHT, T., MCLELLAN, M. D., MCMICHAEL, J. F., WALLIS, J. W., LU, C., SHEN, D., HARRIS, C. C., DOOLING, D. J., FULTON, R. S., FULTON, L. L., CHEN, K., SCHMIDT, H., KALICKI-VEIZER, J., MAGRINI, V. J., COOK, L., MCGRATH, S. D., VICKERY, T. L., WENDL, M. C., HEATH, S., WATSON, M. A., LINK, D. C., TOMASSON, M. H., SHANNON, W. D., PAYTON, J. E., KULKARNI, S., WESTERVELT, P., WALTER, M. J., GRAUBERT, T. A., MARDIS, E. R., WILSON, R. K. & DIPERSIO, J. F. 2012. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature,* 481**,** 506-10.

EHRLICH, M., GAMA-SOSA, M. A., HUANG, L. H., MIDGETT, R. M., KUO, K. C., MCCUNE, R. A. & GEHRKE, C. 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res,* 10**,** 2709-21.

ESTELLER, M. 2005. Aberrant DNA methylation as a cancer-inducing mechanism. *Annu Rev Pharmacol Toxicol,* 45**,** 629-56.

FARZANEHFAR, M., VOSSOUGHINIA, H., JABINI, R., TAVASSOLI, A., SAADATNIA, H., KHORASHAD, A. K., AHADI, M., AGHAEI, M. A., KARIMIANI, E. G., MIRZAEI, F. & AYATOLLAHI, H. 2013. Evaluation of

Methylation of MGMT (O-Methyl Guanine Methyltransferase) Gene Promoter in Sporadic Colorectal Cancer. *DNA Cell Biol*.

FEDORIW, A., MUGFORD, J. & MAGNUSON, T. 2012. Genomic imprinting and epigenetic control of development. *Cold Spring Harb Perspect Biol,* 4**,** a008136.

FEINBERG, A. P., OHLSSON, R. & HENIKOFF, S. 2006. The epigenetic progenitor origin of human cancer. *Nat Rev Genet,* 7**,** 21-33.

FENG, Y. Q., DESPRAT, R., FU, H., OLIVIER, E., LIN, C. M., LOBELL, A., GOWDA, S. N., ALADJEM, M. I. & BOUHASSIRA, E. E. 2006. DNA methylation supports intrinsic epigenetic memory in mammalian cells. *PLoS Genet,* 2**,** e65.

FOULKS, J. M., PARNELL, K. M., NIX, R. N., CHAU, S., SWIERCZEK, K., SAUNDERS, M., WRIGHT, K., HENDRICKSON, T. F., HO, K. K., MCCULLAR, M. V. & KANNER, S. B. 2012. Epigenetic drug discovery: targeting DNA methyltransferases. *J Biomol Screen,* 17**,** 2-17.

FRAGA, M. F., BALLESTAR, E., PAZ, M. F., ROPERO, S., SETIEN, F., BALLESTAR, M. L., HEINE-SUNER, D., CIGUDOSA, J. C., URIOSTE, M., BENITEZ, J., BOIX-CHORNET, M., SANCHEZ-AGUILERA, A., LING, C., CARLSSON, E., POULSEN, P., VAAG, A., STEPHAN, Z., SPECTOR, T. D., WU, Y. Z., PLASS, C. & ESTELLER, M. 2005. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A,* 102**,** 10604-9.

FRAGA, M. F. & ESTELLER, M. 2007. Epigenetics and aging: the targets and the marks. *Trends Genet,* 23**,** 413-8.

FREI, E., 3RD & FREIREICH, E. J. 1964. Leukemia. *Sci Am,* 210**,** 88-96.

FRISO, S., LOTTO, V., CHOI, S. W., GIRELLI, D., PINOTTI, M., GUARINI, P., UDALI, S., PATTINI, P., PIZZOLO, F., MARTINELLI, N., CORROCHER, R., BERNARDI, F. & OLIVIERI, O. 2012. Promoter methylation in coagulation F7 gene influences plasma FVII concentrations and relates to coronary artery disease. *J Med Genet,* 49**,** 192-9.

FULLER, J. F., MCADARA, J., YARON, Y., SAKAGUCHI, M., FRASER, J. K. & GASSON, J. C. 1999. Characterization of HOX gene expression during myelopoiesis: role of HOX A5 in lineage commitment and maturation. *Blood,* 93**,** 3391-400.

GALLO, P., LATRONICO, M. V., GALLO, P., GRIMALDI, S., BORGIA, F., TODARO, M., JONES, P., GALLINARI, P., DE FRANCESCO, R.,

CILIBERTO, G., STEINKUHLER, C., ESPOSITO, G. & CONDORELLI, G. 2008. Inhibition of class I histone deacetylase with an apicidin derivative prevents cardiac hypertrophy and failure. *Cardiovasc Res,* 80**,** 416-24.

GAMA-SOSA, M. A., SLAGEL, V. A., TREWYN, R. W., OXENHANDLER, R., KUO, K. C., GEHRKE, C. W. & EHRLICH, M. 1983. The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res,* 11**,** 6883-94.

GARDINER-GARDEN, M. & FROMMER, M. 1987. CpG islands in vertebrate genomes. *J Mol Biol,* 196**,** 261-82.

GAUTREY, H. 2012. PhD thesis: Age-associated changes in promoter CpG island methylation and their potential role in cancer development.

GOLDSTONE, A. H., RICHARDS, S. M., LAZARUS, H. M., TALLMAN, M. S., BUCK, G., FIELDING, A. K., BURNETT, A. K., CHOPRA, R., WIERNIK, P. H., FORONI, L., PAIETTA, E., LITZOW, M. R., MARKS, D. I., DURRANT, J., MCMILLAN, A., FRANKLIN, I. M., LUGER, S., CIOBANU, N. & ROWE, J. M. 2008. In adults with standard-risk acute lymphoblastic leukemia, the greatest benefit is achieved from a matched sibling allogeneic transplantation in first complete remission, and an autologous transplantation is less effective than conventional consolidation/maintenance chemotherapy in all patients: final results of the International ALL Trial (MRC UKALL XII/ECOG E2993). *Blood,* 111**,** 1827-33.

GORDON, L., JOO, J. E., POWELL, J. E., OLLIKAINEN, M., NOVAKOVIC, B., LI, X., ANDRONIKOS, R., CRUICKSHANK, M. N., CONNEELY, K. N., SMITH, A. K., ALISCH, R. S., MORLEY, R., VISSCHER, P. M., CRAIG, J. M. & SAFFERY, R. 2012. Neonatal DNA methylation profile in human twins is specified by a complex interplay between intrauterine environmental and genetic factors, subject to tissue-specific influence. *Genome Res,* 22**,** 1395-406.

GROLLEAU-JULIUS, A., RAY, D. & YUNG, R. L. 2010. The role of epigenetics in aging and autoimmunity. *Clin Rev Allergy Immunol,* 39**,** 42-50.

GROS, C., FAHY, J., HALBY, L., DUFAU, I., ERDMANN, A., GREGOIRE, J. M., AUSSEIL, F., VISPE, S. & ARIMONDO, P. B. 2012. DNA methylation inhibitors in cancer: recent and future approaches. *Biochimie,* 94**,** 2280-96.

GULHATI, P., BOWEN, K. A., LIU, J., STEVENS, P. D., RYCHAHOU, P. G., CHEN, M., LEE, E. Y., WEISS, H. L., O'CONNOR, K. L., GAO, T. & EVERS, B. M. 2011. mTORC1 and mTORC2 regulate EMT, motility, and metastasis of

colorectal cancer via RhoA and Rac1 signaling pathways. *Cancer Res,* 71**,** 3246-56.

GUO, J. U., SU, Y., ZHONG, C., MING, G. L. & SONG, H. 2011. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell,* 145**,** 423-34.

HAHN, W. C. & WEINBERG, R. A. 2002. Modelling the molecular circuitry of cancer. *Nat Rev Cancer,* 2**,** 331-41.

HAIDER, S., CORDEDDU, L., ROBINSON, E., MOVASSAGH, M., SIGGENS, L., VUJIC, A., CHOY, M. K., GODDARD, M., LIO, P. & FOO, R. 2012. The landscape of DNA repeat elements in human heart failure. *Genome Biol,* 13**,** R90.

HAN, S. & BRUNET, A. 2012. Histone methylation makes its mark on longevity. *Trends Cell Biol,* 22**,** 42-9.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

HANSEN, K. D., TIMP, W., BRAVO, H. C., SABUNCIYAN, S., LANGMEAD, B., MCDONALD, O. G., WEN, B., WU, H., LIU, Y., DIEP, D., BRIEM, E., ZHANG, K., IRIZARRY, R. A. & FEINBERG, A. P. 2011. Increased methylation variation in epigenetic domains across cancer types. *Nat Genet,* 43**,** 768-75.

HARRIES, L. W., FELLOWS, A. D., PILLING, L. C., HERNANDEZ, D., SINGLETON, A., BANDINELLI, S., GURALNIK, J., POWELL, J., FERRUCCI, L. & MELZER, D. 2012. Advancing age is associated with gene expression changes resembling mTOR inhibition: evidence from two human populations. *Mech Ageing Dev,* 133**,** 556-62.

HE, X. J., CHEN, T. & ZHU, J. K. 2011. Regulation and function of DNA methylation in plants and animals. *Cell Res,* 21**,** 442-65.

HENDERSHOT, T. J., LIU, H., CLOUTHIER, D. E., SHEPHERD, I. T., COPPOLA, E., STUDER, M., FIRULLI, A. B., PITTMAN, D. L. & HOWARD, M. J. 2008. Conditional deletion of Hand2 reveals critical functions in neurogenesis and cell type-specific gene expression for development of neural crest-derived noradrenergic sympathetic ganglion neurons. *Dev Biol,* 319**,** 179-91.

HERMANN, A., GOWHER, H. & JELTSCH, A. 2004a. Biochemistry and biology of mammalian DNA methyltransferases. *Cell Mol Life Sci,* 61**,** 2571-87.

HERMANN, A., GOYAL, R. & JELTSCH, A. 2004b. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem,* 279**,** 48350-9.

HEUCK, C. J., MEHTA, J., BHAGAT, T., GUNDABOLU, K., YU, Y., KHAN, S., CHRYSOFAKIS, G., SCHINKE, C., TARIMAN, J., VICKREY, E., PULLIAM, N., NISCHAL, S., ZHOU, L., BHATTACHARYYA, S., MEAGHER, R., HU, C., MAQBOOL, S., SUZUKI, M., PAREKH, S., REU, F., STEIDL, U., GREALLY, J., VERMA, A. & SINGHAL, S. B. 2013. Myeloma is characterized by stage-specific alterations in DNA methylation that occur early during myelomagenesis. *J Immunol,* 190**,** 2966-75.

HILTUNEN, M. O., TURUNEN, M. P., HAKKINEN, T. P., RUTANEN, J., HEDMAN, M., MAKINEN, K., TURUNEN, A. M., AALTO-SETALA, K. & YLA-HERTTUALA, S. 2002. DNA hypomethylation and methyltransferase expression in atherosclerotic lesions. *Vasc Med,* 7**,** 5-11.

HILTUNEN, M. O. & YLA-HERTTUALA, S. 2003. DNA methylation, smooth muscle cells, and atherogenesis. *Arterioscler Thromb Vasc Biol,* 23**,** 1750-3.

HUENECKE, S., BEHL, M., FADLER, C., ZIMMERMANN, S. Y., BOCHENNEK, K., TRAMSEN, L., ESSER, R., KLARMANN, D., KAMPER, M., SATTLER, A., VON LAER, D., KLINGEBIEL, T., LEHRNBECHER, T. & KOEHL, U. 2008. Age-matched lymphocyte subpopulation reference values in childhood and adolescence: application of exponential regression analysis. *Eur J Haematol,* 80**,** 532-9.

HUPKES, M., VAN SOMEREN, E. P., MIDDELKAMP, S. H., PIEK, E., VAN ZOELEN, E. J. & DECHERING, K. J. DNA methylation restricts spontaneous multi-lineage differentiation of mesenchymal progenitor cells, but is stable during growth factor-induced terminal differentiation. *Biochim Biophys Acta*.

ILLUMINA 2011. An introduction to Next-Generation Sequencing Technology. *Illumina sequencing technology*.

IMHOF, A. 2003. Histone modifications: an assembly line for active chromatin? *Curr Biol,* 13**,** R22-4.

INVITROGEN 2010. Cell Isolation and Activation. *Dynabeads cell separation guide*.

IRIZARRY, R. A., LADD-ACOSTA, C., WEN, B., WU, Z., MONTANO, C., ONYANGO, P., CUI, H., GABO, K., RONGIONE, M., WEBSTER, M., JI, H., POTASH, J. B., SABUNCIYAN, S. & FEINBERG, A. P. 2009. The human

colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet,* 41**,** 178-86.

ISENMANN, S., ARTHUR, A., ZANNETTINO, A. C., TURNER, J. L., SHI, S., GLACKIN, C. A. & GRONTHOS, S. 2009. TWIST family of basic helix-loop-helix transcription factors mediate human mesenchymal stem cell growth and commitment. *Stem Cells,* 27**,** 2457-68.

ISSA, J. P., OTTAVIANO, Y. L., CELANO, P., HAMILTON, S. R., DAVIDSON, N. E. & BAYLIN, S. B. 1994. Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. *Nat Genet,* 7**,** 536-40.

JAKOVCEVSKI, M. & AKBARIAN, S. 2012. Epigenetic mechanisms in neurological disease. *Nat Med,* 18**,** 1194-204.

JEPPESEN, D. K., BOHR, V. A. & STEVNSNER, T. 2011. DNA repair deficiency in neurodegeneration. *Prog Neurobiol,* 94**,** 166-200.

JINTARIDTH, P. & MUTIRANGURA, A. Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences. *Physiol Genomics*.

JONES, P. A. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet,* 13**,** 484-92.

JONES, P. A. & LIANG, G. 2009. Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet,* 10**,** 805-11.

KAMINSKY, Z. A., TANG, T., WANG, S. C., PTAK, C., OH, G. H., WONG, A. H., FELDCAMP, L. A., VIRTANEN, C., HALFVARSON, J., TYSK, C., MCRAE, A. F., VISSCHER, P. M., MONTGOMERY, G. W., GOTTESMAN, II, MARTIN, N. G. & PETRONIS, A. 2009. DNA methylation profiles in monozygotic and dizygotic twins. *Nat Genet,* 41**,** 240-5.

KANDOUZ, M. 2012. The Eph/Ephrin family in cancer metastasis: communication at the service of invasion. *Cancer Metastasis Rev,* 31**,** 353-73.

KANEDA, M., OKANO, M., HATA, K., SADO, T., TSUJIMOTO, N., LI, E. & SASAKI, H. 2004. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature,* 429**,** 900-3.

KANELLOPOULOS, A., HAMRE, H. M., DAHL, A. A., FOSSA, S. D. & RUUD, E. 2013. Factors associated with poor quality of life in survivors of childhood acute lymphoblastic leukemia and lymphoma. *Pediatr Blood Cancer*.

KENYON, C. J. 2010. The genetics of ageing. *Nature,* 464**,** 504-12.

KEYES, W. M., WU, Y., VOGEL, H., GUO, X., LOWE, S. W. & MILLS, A. A. 2005. p63 deficiency activates a program of cellular senescence and leads to accelerated aging. *Genes Dev,* 19**,** 1986-99.

KIM, D. H., KIM, J. M., LEE, E. K., CHOI, Y. J., KIM, C. H., CHOI, J. S., KIM, N. D., YU, B. P. & CHUNG, H. Y. 2012. Modulation of FoxO1 phosphorylation/acetylation by baicalin during aging. *J Nutr Biochem,* 23**,** 1277-84.

KIRKWOOD, T. B. 2005. Understanding the odd science of aging. *Cell,* 120**,** 437-47.

KOH, H. S., LEE, C., LEE, K. S., PARK, E. J., SEONG, R. H., HONG, S. & JEON, S. H. 2009. Twist2 regulates CD7 expression and galectin-1-induced apoptosis in mature T-cells. *Mol Cells,* 28**,** 553-8.

KRAMER, O. H., GOTTLICHER, M. & HEINZEL, T. 2001. Histone deacetylase as a therapeutic target. *Trends Endocrinol Metab,* 12**,** 294-300.

KRISHNAMURTHY, J., TORRICE, C., RAMSEY, M. R., KOVALEV, G. I., AL-REGAIEY, K., SU, L. & SHARPLESS, N. E. 2004. Ink4a/Arf expression is a biomarker of aging. *J Clin Invest,* 114**,** 1299-307.

KUANG, S. Q., BAI, H., FANG, Z. H., LOPEZ, G., YANG, H., TONG, W., WANG, Z. Z. & GARCIA-MANERO, G. Aberrant DNA methylation and epigenetic inactivation of Eph receptor tyrosine kinases and ephrin ligands in acute lymphoblastic leukemia. *Blood,* 115**,** 2412-9.

KUCUK, C., HU, X., IQBAL, J., GAULARD, P., KLINKEBIEL, D., CORNISH, A., DAVE, B. J. & CHAN, W. C. 2012. HACE1 Is a Tumor Suppressor Gene Candidate in Natural Killer Cell Neoplasms. *Am J Pathol*.

KWON, Y. R., SON, M. J., KIM, H. J. & KIM, Y. J. 2012. Reactivation of Silenced WT1 Transgene by Hypomethylating Agents - Implications for in vitro Modeling of Chemoimmunotherapy. *Immune Netw,* 12**,** 58-65.

LAKATTA, E. G. & LEVY, D. 2003. Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: Part II: the aging heart in health: links to heart disease. *Circulation,* 107**,** 346-54.

LAM, L. L., EMBERLY, E., FRASER, H. B., NEUMANN, S. M., CHEN, E., MILLER, G. E. & KOBOR, M. S. 2012. Factors underlying variable DNA methylation in a human community cohort. *Proc Natl Acad Sci U S A,* 109 Suppl 2**,** 17253-60.

LEE, S. T., XIAO, Y., MUENCH, M. O., XIAO, J., FOMIN, M. E., WIENCKE, J. K., ZHENG, S., DOU, X., DE SMITH, A., CHOKKALINGAM, A., BUFFLER, P.,

MA, X. & WIEMELS, J. L. 2012. A global DNA methylation and gene expression analysis of early human B-cell development reveals a demethylation signature and transcription factor network. *Nucleic Acids Res,* 40**,** 11339-51.

LEVY, A., DANG, U. C. & BOOKSTEIN, R. 1999. High-density screen of human tumor cell lines for homozygous deletions of loci on chromosome arm 8p. *Genes Chromosomes Cancer,* 24**,** 42-7.

LI, E. 2002. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet,* 3**,** 662-73.

LI, Y., TAN, T., ZONG, L., HE, D., TAO, W. & LIANG, Q. 2012. Study of methylation of histone H3 lysine 9 and H3 lysine 27 during X chromosome inactivation in three types of cells. *Chromosome Res*.

LILLVIS, J. H., ERDMAN, R., SCHWORER, C. M., GOLDEN, A., DERR, K., GATALICA, Z., COX, L. A., SHEN, J., VANDER HEIDE, R. S., LENK, G. M., HLAVATY, L., LI, L., ELMORE, J. R., FRANKLIN, D. P., GRAY, J. L., GARVIN, R. P., CAREY, D. J., LANCASTER, W. D., TROMP, G. & KUIVANIEMI, H. Regional expression of HOXA4 along the aorta and its potential role in human abdominal aortic aneurysms. *BMC Physiol,* 11**,** 9.

LINTON, P. J. & DORSHKIND, K. 2004. Age-related changes in lymphocyte development and function. *Nat Immunol,* 5**,** 133-9.

LISTER, R., PELIZZOLA, M., DOWEN, R. H., HAWKINS, R. D., HON, G., TONTI-FILIPPINI, J., NERY, J. R., LEE, L., YE, Z., NGO, Q. M., EDSALL, L., ANTOSIEWICZ-BOURGET, J., STEWART, R., RUOTTI, V., MILLAR, A. H., THOMSON, J. A., REN, B. & ECKER, J. R. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature,* 462**,** 315-22.

LLR. 2013. *leukaemialymphomaresearch.org.uk* [Online]. Available: http://leukaemialymphomaresearch.org.uk/news/resources-press/facts-about-blood-cancers.

LOPEZ-OTIN, C., BLASCO, M. A., PARTRIDGE, L., SERRANO, M. & KROEMER, G. 2013. The hallmarks of aging. *Cell,* 153**,** 1194-217.

LUCZAK, M. W. & JAGODZINSKI, P. P. 2006. The role of DNA methylation in cancer development. *Folia Histochem Cytobiol,* 44**,** 143-54.

LUGER, K., RECHSTEINER, T. J., FLAUS, A. J., WAYE, M. M. & RICHMOND, T. J. 1997. Characterization of nucleosome core particles containing histone proteins made in bacteria. *J Mol Biol,* 272**,** 301-11.

MAEGAWA, S., HINKAL, G., KIM, H. S., SHEN, L., ZHANG, L., ZHANG, J., ZHANG, N., LIANG, S., DONEHOWER, L. A. & ISSA, J. P. Widespread and tissue specific age-related DNA methylation changes in mice. *Genome Res,* 20**,** 332-40.

MATHERS, J. C. 2006. Nutritional modulation of ageing: genomic and epigenetic approaches. *Mech Ageing Dev,* 127**,** 584-9.

MATHERS, J. C., STRATHDEE, G. & RELTON, C. L. 2010. Induction of epigenetic alterations by dietary and other environmental factors. *Adv Genet,* 71**,** 3-39.

MILANI, L., LUNDMARK, A., KIIALAINEN, A., NORDLUND, J., FLAEGSTAD, T., FORESTIER, E., HEYMAN, M., JONMUNDSSON, G., KANERVA, J., SCHMIEGELOW, K., SODERHALL, S., GUSTAFSSON, M. G., LONNERHOLM, G. & SYVANEN, A. C. 2010. DNA methylation for subtype classification and prediction of treatment outcome in patients with childhood acute lymphoblastic leukemia. *Blood,* 115**,** 1214-25.

MOLINA-SERRANO, D., SCHIZA, V. & KIRMIZIS, A. 2013. Cross-talk among epigenetic modifications: lessons from histone arginine methylation. *Biochem Soc Trans,* 41**,** 751-9.

MOORE, L. D., LE, T. & FAN, G. 2012. DNA Methylation and Its Basic Function. *Neuropsychopharmacology*.

MOVASSAGH, M., CHOY, M. K., GODDARD, M., BENNETT, M. R., DOWN, T. A. & FOO, R. S. Differential DNA methylation correlates with differential expression of angiogenic factors in human heart failure. *PLoS One,* 5**,** e8564.

MUND, C., HACKANSON, B., STRESEMANN, C., LUBBERT, M. & LYKO, F. 2005. Characterization of DNA demethylation effects induced by 5-Aza-2'-deoxycytidine in patients with myelodysplastic syndrome. *Cancer Res,* 65**,** 7086-90.

NOBORI, T., MIURA, K., WU, D. J., LOIS, A., TAKABAYASHI, K. & CARSON, D. A. 1994. Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. *Nature,* 368**,** 753-6.

O'HAGAN, H. M., WANG, W., SEN, S., DESTEFANO SHIELDS, C., LEE, S. S., ZHANG, Y. W., CLEMENTS, E. G., CAI, Y., VAN NESTE, L., EASWARAN, H., CASERO, R. A., SEARS, C. L. & BAYLIN, S. B. Oxidative Damage Targets Complexes Containing DNA Methyltransferases, SIRT1, and Polycomb Members to Promoter CpG Islands. *Cancer Cell,* 20**,** 606-19.

OKANO, M., BELL, D. W., HABER, D. A. & LI, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell,* 99**,** 247-57.

OLIVETTI, G., MELISSARI, M., CAPASSO, J. M. & ANVERSA, P. 1991. Cardiomyopathy of the aging human heart. Myocyte loss and reactive cellular hypertrophy. *Circ Res,* 68**,** 1560-8.

OXFORD UNIVERSITY PRESS. 2010. *Concise colour medical dictionary,* Oxford ; New York, Oxford University Press.

PAKNESHAN, P., SZYF, M., FARIAS-EISNER, R. & RABBANI, S. A. 2004. Reversal of the hypomethylation status of urokinase (uPA) promoter blocks breast cancer growth and metastasis. *J Biol Chem,* 279**,** 31735-44.

PANDE, S., BROWNE, G., PADMANABHAN, S., ZAIDI, S. K., LIAN, J. B., VAN WIJNEN, A. J., STEIN, J. L. & STEIN, G. S. 2013. Oncogenic cooperation between PI3K/Akt signaling and transcription factor Runx2 promotes the invasive properties of metastatic breast cancer cells. *J Cell Physiol*.

PARK, I. K., QIAN, D., KIEL, M., BECKER, M. W., PIHALJA, M., WEISSMAN, I. L., MORRISON, S. J. & CLARKE, M. F. 2003. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature,* 423**,** 302-5.

PARRY, L. & CLARKE, A. R. 2011. The Roles of the Methyl-CpG Binding Proteins in Cancer. *Genes Cancer,* 2**,** 618-30.

PATTANI, K. M., SOUDRY, E., GLAZER, C. A., OCHS, M. F., WANG, H., SCHUSSEL, J., SUN, W., HENNESSEY, P., MYDLARZ, W., LOYO, M., DEMOKAN, S., SMITH, I. M. & CALIFANO, J. A. 2012. MAGEB2 is Activated by Promoter Demethylation in Head and Neck Squamous Cell Carcinoma. *PLoS One,* 7**,** e45534.

PAWLAK, M. & JAENISCH, R. 2011. De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. *Genes Dev,* 25**,** 1035-40.

PEARCE, M. S., UNWIN, N. C., PARKER, L. & CRAFT, A. W. 2009. Cohort profile: the Newcastle Thousand Families 1947 birth cohort. *Int J Epidemiol,* 38**,** 932-7.

PELTOMAKI, P. & VASEN, H. F. 1997. Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. *Gastroenterology,* 113**,** 1146-58.

POST, W. S., GOLDSCHMIDT-CLERMONT, P. J., WILHIDE, C. C., HELDMAN, A. W., SUSSMAN, M. S., OUYANG, P., MILLIKEN, E. E. & ISSA, J. P. 1999. Methylation of the estrogen receptor gene is associated with aging and atherosclerosis in the cardiovascular system. *Cardiovasc Res,* 43**,** 985-91.

PRESS, J. Z., DE LUCA, A., BOYD, N., YOUNG, S., TROUSSARD, A., RIDGE, Y., KAURAH, P., KALLOGER, S. E., BLOOD, K. A., SMITH, M., SPELLMAN, P. T., WANG, Y., MILLER, D. M., HORSMAN, D., FAHAM, M., GILKS, C. B., GRAY, J. & HUNTSMAN, D. G. 2008. Ovarian carcinomas with genetic and epigenetic BRCA1 loss have distinct molecular abnormalities. *BMC Cancer,* 8**,** 17.

PROWSE, A. H., WEBSTER, A. R., RICHARDS, F. M., RICHARD, S., OLSCHWANG, S., RESCHE, F., AFFARA, N. A. & MAHER, E. R. 1997. Somatic inactivation of the VHL gene in Von Hippel-Lindau disease tumors. *Am J Hum Genet,* 60**,** 765-71.

RADULOVIC, V., DE HAAN, G. & KLAUKE, K. 2013. Polycomb-group proteins in hematopoietic stem cell regulation and hematopoietic neoplasms. *Leukemia,* 27**,** 523-33.

RAKYAN, V. K., DOWN, T. A., MASLAU, S., ANDREW, T., YANG, T. P., BEYAN, H., WHITTAKER, P., MCCANN, O. T., FINER, S., VALDES, A. M., LESLIE, R. D., DELOUKAS, P. & SPECTOR, T. D. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res,* 20**,** 434-9.

RONAGHI, M. 2001. Pyrosequencing sheds light on DNA sequencing. *Genome Res,* 11**,** 3-11.

ROSS, R. 1993. Rous-Whipple Award Lecture. Atherosclerosis: a defense mechanism gone awry. *Am J Pathol,* 143**,** 987-1002.

ROSSI, D. J., BRYDER, D., SEITA, J., NUSSENZWEIG, A., HOEIJMAKERS, J. & WEISSMAN, I. L. 2007. Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature,* 447**,** 725-9.

RUDOLPH, K. L., CHANG, S., LEE, H. W., BLASCO, M., GOTTLIEB, G. J., GREIDER, C. & DEPINHO, R. A. 1999. Longevity, stress response, and cancer in aging telomerase-deficient mice. *Cell,* 96**,** 701-12.

SAHA, A., WITTMEYER, J. & CAIRNS, B. R. 2006. Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol,* 7**,** 437-47.

SAHIN, E. & DEPINHO, R. A. 2010. Linking functional decline of telomeres, mitochondria and stem cells during ageing. *Nature,* 464**,** 520-8.

SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. & ERLICH, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science,* 239**,** 487-91.

SAKAI, T., TOGUCHIDA, J., OHTANI, N., YANDELL, D. W., RAPAPORT, J. M. & DRYJA, T. P. 1991. Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am J Hum Genet,* 48**,** 880-8.

SANDOVAL, J., HEYN, H., MORAN, S., SERRA-MUSACH, J., PUJANA, M. A., BIBIKOVA, M. & ESTELLER, M. 2011. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics,* 6**,** 692-702.

SANZ, L. A., KOTA, S. K. & FEIL, R. 2010. Genome-wide DNA demethylation in mammals. *Genome Biol,* 11**,** 110.

SAWYERS, C. L. 1999. Chronic myeloid leukemia. *N Engl J Med,* 340**,** 1330-40.

SAXONOV, S., BERG, P. & BRUTLAG, D. L. 2006. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A,* 103**,** 1412-7.

SCHOLZ, C., NIMMRICH, I., BURGER, M., BECKER, E., DORKEN, B., LUDWIG, W. D. & MAIER, S. 2005. Distinction of acute lymphoblastic leukemia from acute myeloid leukemia through microarray-based DNA methylation analysis. *Ann Hematol,* 84**,** 236-44.

SCIENCE, W. I. O. 2012. *The GeneCards Human Gene Database* [Online]. Available: http://www.genecards.org/cgi-bin/carddisp.pl?gene=HOXD4.

SHARABI, A. B., ALDRICH, M., SOSIC, D., OLSON, E. N., FRIEDMAN, A. D., LEE, S. H. & CHEN, S. Y. 2008. Twist-2 controls myeloid lineage development and function. *PLoS Biol,* 6**,** e316.

SIGMA-ALDRICH 2008. qPCR Technical Guide.

SIGNER, R. A. & MORRISON, S. J. 2013. Mechanisms that Regulate Stem Cell Aging and Life Span. *Cell Stem Cell,* 12**,** 152-65.

SKORSKI, T. 2012. Genetic mechanisms of chronic myeloid leukemia blastic transformation. *Curr Hematol Malig Rep,* 7**,** 87-93.

SMITH, Z. D. & MEISSNER, A. 2013. DNA methylation: roles in mammalian development. *Nat Rev Genet,* 14**,** 204-20.

SONG, K., NAM, Y. J., LUO, X., QI, X., TAN, W., HUANG, G. N., ACHARYA, A., SMITH, C. L., TALLQUIST, M. D., NEILSON, E. G., HILL, J. A., BASSEL-DUBY, R. & OLSON, E. N. 2012. Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature,* 485**,** 599-604.

SOZOU, P. D. & KIRKWOOD, T. B. 2001. A stochastic model of cell replicative senescence based on telomere shortening, oxidative stress, and somatic mutations in nuclear and mitochondrial DNA. *J Theor Biol,* 213**,** 573-86.

SPROUL, D., KITCHEN, R. R., NESTOR, C. E., DIXON, J. M., SIMS, A. H., HARRISON, D. J., RAMSAHOYE, B. H. & MEEHAN, R. R. 2012. Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. *Genome Biol,* 13**,** R84.

ST-PIERRE, J., HIVERT, M. F., PERRON, P., POIRIER, P., GUAY, S. P., BRISSON, D. & BOUCHARD, L. 2012. IGF2 DNA methylation is a modulator of newborn's fetal growth and development. *Epigenetics,* 7.

STRAHL, B. D. & ALLIS, C. D. 2000. The language of covalent histone modifications. *Nature,* 403**,** 41-5.

STRATHDEE, G., HOLYOAKE, T. L., SIM, A., PARKER, A., OSCIER, D. G., MELO, J. V., MEYER, S., EDEN, T., DICKINSON, A. M., MOUNTFORD, J. C., JORGENSEN, H. G., SOUTAR, R. & BROWN, R. 2007a. Inactivation of HOXA genes by hypermethylation in myeloid and lymphoid malignancy is frequent and associated with poor prognosis. *Clin Cancer Res,* 13**,** 5048-55.

STRATHDEE, G., SIM, A., PARKER, A., OSCIER, D. & BROWN, R. 2006. Promoter hypermethylation silences expression of the HoxA4 gene and correlates with IgVh mutational status in CLL. *Leukemia,* 20**,** 1326-9.

STRATHDEE, G., SIM, A., SOUTAR, R., HOLYOAKE, T. L. & BROWN, R. 2007b. HOXA5 is targeted by cell-type-specific CpG island methylation in normal cells and during the development of acute myeloid leukaemia. *Carcinogenesis,* 28**,** 299-309.

SWARTZ, M. N., TRAUTNER, T. A. & KORNBERG, A. 1962. Enzymatic synthesis of deoxyribonucleic acid. XI. Further studies on nearest neighbor base sequences in deoxyribonucleic acids. *J Biol Chem,* 237**,** 1961-7.

TAHILIANI, M., KOH, K. P., SHEN, Y., PASTOR, W. A., BANDUKWALA, H., BRUDNO, Y., AGARWAL, S., IYER, L. M., LIU, D. R., ARAVIND, L. & RAO, A. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* 324**,** 930-5.

TAKAI, D. & JONES, P. A. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A,* 99**,** 3740-5.

TAKASUGI, M. Progressive age-dependent DNA methylation changes start before adulthood in mouse tissues. *Mech Ageing Dev,* 132**,** 65-71.

TALENS, R. P., CHRISTENSEN, K., PUTTER, H., WILLEMSEN, G., CHRISTIANSEN, L., KREMER, D., SUCHIMAN, H. E., SLAGBOOM, P. E., BOOMSMA, D. I. & HEIJMANS, B. T. 2012. Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell,* 11**,** 694-703.

TALIKKA, M., SIERRO, N., IVANOV, N. V., CHAUDHARY, N., PECK, M. J., HOENG, J., COGGINS, C. R. & PEITSCH, M. C. 2012. Genomic impact of cigarette smoke, with application to three smoking-related diseases. *Crit Rev Toxicol,* 42**,** 877-89.

TAPP, H. S., COMMANE, D. M., BRADBURN, D. M., ARASARADNAM, R., MATHERS, J. C., JOHNSON, I. T. & BELSHAW, N. J. 2012. Nutritional factors and gender influence age-related DNA methylation in the human rectal mucosa. *Aging Cell*.

TENG, I. W., HOU, P. C., LEE, K. D., CHU, P. Y., YEH, K. T., JIN, V. X., TSENG, M. J., TSAI, S. J., CHANG, Y. S., WU, C. S., SUN, H. S., TSAI, K. D., JENG, L. B., NEPHEW, K. P., HUANG, T. H., HSIAO, S. H. & LEU, Y. W. Targeted methylation of two tumor suppressor genes is sufficient to transform mesenchymal stem cells into cancer stem/initiating cells. *Cancer Res,* 71**,** 4653-63.

TESCHENDORFF, A. E., MENON, U., GENTRY-MAHARAJ, A., RAMUS, S. J., WEISENBERGER, D. J., SHEN, H., CAMPAN, M., NOUSHMEHR, H., BELL, C. G., MAXWELL, A. P., SAVAGE, D. A., MUELLER-HOLZNER, E., MARTH, C., KOCJAN, G., GAYTHER, S. A., JONES, A., BECK, S., WAGNER, W., LAIRD, P. W., JACOBS, I. J. & WIDSCHWENDTER, M. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res,* 20**,** 440-6.

THATHIA, S. H., FERGUSON, S., GAUTREY, H. E., VAN OTTERDIJK, S. D., HILI, M., RAND, V., MOORMAN, A. V., MEYER, S., BROWN, R. & STRATHDEE, G. 2012. Epigenetic inactivation of TWIST2 in acute lymphoblastic leukemia modulates proliferation, cell survival and chemosensitivity. *Haematologica,* 97**,** 371-8.

THOMPSON, R. F., ATZMON, G., GHEORGHE, C., LIANG, H. Q., LOWES, C., GREALLY, J. M. & BARZILAI, N. 2010. Tissue-specific dysregulation of DNA methylation in aging. *Aging Cell,* 9**,** 506-18.

TSANG, S., WOO, A. Y., ZHU, W. & XIAO, R. P. 2010. Deregulation of RGS2 in cardiovascular diseases. *Front Biosci (Schol Ed),* 2**,** 547-57.

UDALI, S., GUARINI, P., MORUZZI, S., CHOI, S. W. & FRISO, S. 2012. Cardiovascular epigenetics: From DNA methylation to microRNAs. *Mol Aspects Med*.

USHIJIMA, T., WATANABE, N., OKOCHI, E., KANEDA, A., SUGIMURA, T. & MIYAMOTO, K. 2003. Fidelity of the methylation pattern and its variation in the genome. *Genome Res,* 13**,** 868-74.

VAN HOESEL, A. Q., SATO, Y., ELASHOFF, D. A., TURNER, R. R., GIULIANO, A. E., SHAMONKI, J. M., KUPPEN, P. J., VAN DE VELDE, C. J. & HOON, D. S. 2013. Assessment of DNA methylation status in early stages of breast cancer development. *Br J Cancer,* 108**,** 2033-8.

VAN OTTERDIJK, S. D., MATHERS, J. C. & STRATHDEE, G. 2013. Do age-related changes in DNA methylation play a role in the development of age-related diseases? *Biochem Soc Trans,* 41**,** 803-7.

VAN SCHERPENZEEL THIM, V., REMACLE, S., PICARD, J., CORNU, G., GOFFLOT, F., REZSOHAZY, R. & VERELLEN-DUMOULIN, C. 2005. Mutation analysis of the HOX paralogous 4-13 genes in children with acute lymphoid malignancies: identification of a novel germline mutation of HOXD4 leading to a partial loss-of-function. *Hum Mutat,* 25**,** 384-95.

VASEN, H. F., WATSON, P., MECKLIN, J. P. & LYNCH, H. T. 1999. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology,* 116**,** 1453-6.

VAUPEL, J. W. 2010. Biodemography of human ageing. *Nature,* 464**,** 536-42.

WANG, J. C. & BENNETT, M. 2012. Aging and atherosclerosis: mechanisms, functional consequences, and potential therapeutics for cellular senescence. *Circ Res,* 111**,** 245-59.

WARD, R. L., DOBBINS, T., LINDOR, N. M., RAPKINS, R. W. & HITCHINS, M. P. 2012. Identification of constitutional MLH1 epimutations and promoter variants in colorectal cancer patients from the Colon Cancer Family Registry. *Genet Med*.

WEISENBERGER, D. J., CAMPAN, M., LONG, T. I., KIM, M., WOODS, C., FIALA, E., EHRLICH, M. & LAIRD, P. W. 2005. Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Res,* 33**,** 6823-36.

WEKSBERG, R., SHEN, D. R., FEI, Y. L., SONG, Q. L. & SQUIRE, J. 1993. Disruption of insulin-like growth factor 2 imprinting in Beckwith-Wiedemann syndrome. *Nat Genet,* 5**,** 143-50.

WHEELER, J. M., LOUKOLA, A., AALTONEN, L. A., MORTENSEN, N. J. & BODMER, W. F. 2000. The role of hypermethylation of the hMLH1 promoter region in HNPCC versus MSI+ sporadic colorectal cancers. *J Med Genet,* 37**,** 588-92.

WHO 2011. Global atlas on cardiovascular disease prevention and control.

WHO 2012. Good healthy adds life to years: Global brief for World Health Day.

WHO 2013. Ageing and Life Course *http://www.who.int/ageing/en/index.html*.

WILSON, V. L. & JONES, P. A. 1983. DNA methylation decreases in aging but not in immortal cells. *Science,* 220**,** 1055-7.

WONG, I. H., NG, M. H., HUANG, D. P. & LEE, J. C. 2000. Aberrant p15 promoter methylation in adult and childhood acute leukemias of nearly all morphologic subtypes: potential prognostic implications. *Blood,* 95**,** 1942-9.

WONG, L. S., VAN DER HARST, P., DE BOER, R. A., HUZEN, J., VAN GILST, W. H. & VAN VELDHUISEN, D. J. 2010. Aging, telomeres and heart failure. *Heart Fail Rev,* 15**,** 479-86.

WU, H. & SUN, Y. E. 2006. Epigenetic regulation of stem cell differentiation. *Pediatr Res,* 59**,** 21R-5R.

WU, S. C. & ZHANG, Y. 2010. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol,* 11**,** 607-20.

XING, X., CAI, W., SHI, H., WANG, Y., LI, M., JIAO, J. & CHEN, M. 2013. The prognostic value of CDKN2A hypermethylation in colorectal cancer: a meta-analysis. *Br J Cancer*.

YAN, P., FRANKHOUSER, D., MURPHY, M., TAM, H. H., RODRIGUEZ, B., CURFMAN, J., TRIMARCHI, M., GEYER, S., WU, Y. Z., WHITMAN, S. P., METZELER, K., WALKER, A., KLISOVIC, R., JACOB, S., GREVER, M. R., BYRD, J. C., BLOOMFIELD, C. D., GARZON, R., BLUM, W., CALIGIURI, M. A., BUNDSCHUH, R. & MARCUCCI, G. 2012. Genome-wide methylation profiling in decitabine-treated patients with acute myeloid leukemia. *Blood,* 120**,** 2466-74.

YI-DENG, J., TAO, S., HUI-PING, Z., JIAN-TUAN, X., JUN, C., GUI-ZHONG, L. & SHU-REN, W. 2007. Folate and ApoE DNA methylation induced by homocysteine in human monocytes. *DNA Cell Biol,* 26**,** 737-44.

YIDENG, J., ZHIHONG, L., JIANTUAN, X., JUN, C., GUIZHONG, L. & SHUREN, W. 2008. Homocysteine-mediated PPARalpha,gamma DNA methylation and its potential pathogenic mechanism in monocytes. *DNA Cell Biol,* 27**,** 143-50.

ZHA, Y., DING, E., YANG, L., MAO, L., WANG, X., MCCARTHY, B. A., HUANG, S. & DING, H. F. 2012. Functional Dissection of HOXD Cluster Genes in Regulation of Neuroblastoma Cell Proliferation and Differentiation. *PLoS One,* 7**,** e40728.

ZHANG, Y. & REINBERG, D. 2001. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev,* 15**,** 2343-60.

ZHU, Z. Z., HOU, L., BOLLATI, V., TARANTINI, L., MARINELLI, B., CANTONE, L., YANG, A. S., VOKONAS, P., LISSOWSKA, J., FUSTINONI, S., PESATORI, A. C., BONZINI, M., APOSTOLI, P., COSTA, G., BERTAZZI, P. A., CHOW, W. H., SCHWARTZ, J. & BACCARELLI, A. 2012. Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis. *Int J Epidemiol,* 41**,** 126-39.